

A Corpus-based Study of English Near-synonyms: *Careful, cautious, and wary*

Preeyanan Assawawattanasuntorn^a, Vanlee Siriganjanavong^{b,*}

^a p.assawaw@gmail.com, Language Institute, Thammasat University, Thailand

^b vanlee.s@litu.tu.ac.th, Language Institute, Thammasat University, Thailand

* Corresponding author, vanlee.s@litu.tu.ac.th

APA Citation: Assawawattanasuntorn, P., & Siriganjanavong, V. (2005). A corpus-based study of English near-synonyms: careful, cautious, and wary. *LEARN Journal: Language Education and Acquisition Research Network*, 18(2), 869-892. <https://doi.org/10.70730/UAXF5168>

Received 09/04/2025	ABSTRACT
Received in revised form 27/06/2025	This corpus-based study investigates the similarities and differences among the three synonymous adjectives: <i>careful</i> , <i>cautious</i> , and <i>wary</i> in terms of genre distribution, collocational patterns, semantic preference, and semantic prosody. Data were drawn from the Corpus of Contemporary American English (COCA). The results reveal that <i>careful</i> occurs widely across both formal and informal genres, with the highest frequency in TV and movie subtitles. In contrast, <i>cautious</i> and <i>wary</i> are predominantly found in more formal genres, such as newspapers and academic text. Collocational patterns further distinguish that <i>careful</i> frequently co-occurs with nouns such as <i>attention</i> , <i>consideration</i> , <i>analysis</i> , <i>planning</i> , and <i>examination</i> . <i>Cautious</i> commonly collocates with <i>approach</i> , <i>optimism</i> , <i>step</i> , <i>investor</i> , and <i>consumer</i> . <i>Wary</i> tends to appear with nouns such as <i>eye</i> , <i>investor</i> , <i>look</i> , <i>consumer</i> , and <i>glance</i> . While <i>cautious</i> and <i>wary</i> share several collocates and semantic preferences, <i>careful</i> does not overlap significantly with either. The majority of noun collocates with <i>careful</i> fall under the theme reflecting cognitive or analytical processes. Most noun collocates of <i>cautious</i> tend to appear in financial or economic contexts, while those of <i>wary</i> primarily relate to people and social perception. In terms of semantic
Accepted 05/07/2025	

	<p>prosody, <i>careful</i> exhibits neutral tones, whereas <i>cautious</i> exhibits neutral to negative prosody and <i>wary</i> is generally associated with a more negative connotation. These findings demonstrate that, despite their synonymous definitions, <i>careful</i>, <i>cautious</i>, and <i>wary</i> are not fully interchangeable across contexts due to distinct linguistic and semantic patterns.</p> <p>Keywords: synonym, genre distribution, collocation, semantic preference and prosody</p>
--	---

Introduction

English is one of the world's most widely used languages. It has become the primary language in various areas, including computer programming, global business, and higher education (Potter & Crystal, 2020). Mastering English enables learners to communicate effectively and opens doors to many opportunities (Malokani et al., 2023).

To master the English language, vocabulary knowledge is vital. Without vocabulary knowledge, it is impossible to convey anything (Wilkins, 1972). Vocabulary knowledge is widely recognized as a fundamental skill for other skills (Viera, 2018). Learners with rich vocabulary knowledge tend to succeed in language learning, whether in spoken or written forms (Schmitt, 2000). Understanding the meanings of words enables learners to use a language more confidently in various contexts (Walters, 2004).

Nevertheless, synonyms seem to be a barrier for learners wishing to become successful in language learning (Laufer, 1990; Yeh et al., 2007). Not all synonyms are interchangeable in all possible contexts of use, causing difficulties for learners to differentiate similarities and differences among synonyms. In other words, some words cannot be substituted by their synonyms without affecting the original meaning or connotations.

Among the countless synonyms, the words *careful*, *cautious*, and *wary* are one of those synonyms that may cause confusion. Since their meanings are so similar, they can easily be confused and misused by learners. According to Merriam-Webster Dictionary (n.d.), the adjective *careful* is defined as “marked by wary caution or prudence; marked by attentive concern and solicitude; or marked by painstaking effort to avoid errors or omissions”. *Cautious* is defined as “careful about avoiding danger or risk” and *wary* is defined as “marked by keen caution, cunning, and watchfulness, especially in detecting and escaping danger.” These definitions clearly indicate that the three adjectives are closely related in meaning. In addition, these three synonymous adjectives are present in the list of the Longman Communication 9000 (LDOCE6, 2014), a list of the 9000 most frequent used

words in English, compiled based on their relevance and frequency in both spoken and written communication. The adjective *careful* belongs in the high-frequency group, while *cautious* and *wary* are classified in the medium-frequency and low-frequency groups of the Longman Communication 9000, respectively. Without a doubt, these three synonymous adjectives, namely *careful*, *cautious*, and *wary*, are worth investigating to find their similarities and differences.

Consequently, a corpus-based study of synonymous adjectives *careful*, *cautious*, and *wary* was conducted, with an emphasis on their distribution across genres, collocational patterns, semantic preference, and semantic prosody. Such analysis could provide deeper vocabulary insight and support more effective English language teaching through authentic, contextualized usage. As such, the study seeks to address the following research questions:

1. How are the three synonyms, *careful*, *cautious*, and *wary*, distributed across different genres?
2. What are the noun collocations that frequently co-occur with the synonyms *careful*, *cautious*, and *wary*?
3. What are the semantic preferences and prosody of the noun collocations that frequently co-occur with the synonyms *careful*, *cautious*, and *wary*?

Literature Review

A synonym refers to a word or phrase that has similar senses of meaning with another word or phrase (Webb & Nation, 2017). Similarly, the Longman Dictionary of Contemporary English or LDOCE (Online access) defines the term ‘synonym’ as “a word with the same meaning as another word in the same language”. Based on these definitions, words that have similar meanings can be regarded as synonyms. In corpus-based research, synonyms can be divided into two main categories: absolute (or strict) synonyms and loose (or near) synonyms. (Sridhanyarat & Phoocharoensil, 2023). Absolute synonyms imply complete interchangeability in every context, without affecting the original meaning, style, or connotation (Cruse, 1986). On the contrary, near-synonyms refer to words that have similar meanings; still, they cannot be used interchangeably in every context (Phoocharoensil, 2020).

The target synonyms are defined by the Longman Dictionary of Contemporary English (Online access) as follows:

<i>Careful</i>	<ol style="list-style-type: none"> (1) “trying very hard to avoid doing anything wrong or to avoid damaging or losing something.” (2) “paying a lot of attention to details, so that something is done correctly and thoroughly”
----------------	--

<i>Cautious</i>	(1) “careful to avoid danger or risks.”
<i>Wary</i>	(1) “someone who is wary is careful because they think something might be dangerous or harmful.”

A closer examination of their definition(s) reveals some important distinctions. *Careful* is a broader term with two key definitions, covering both general prudence and specific contexts of avoiding harm or mistakes. In contrast, *cautious* and *wary* each have a more limited scope with only one primary definition; both focus on avoiding danger or harm. Overall, *careful*, *cautious*, and *wary* are considered near-synonyms, as they all relate to a sense of attentiveness and the desire to avoid mistakes, harm, or risk.

According to Jackson and Amvela (2007), aspects such as formality, collocation, and semantic prosody help differentiate near-synonyms. This study adopts the approach of Sridhanyarat and Phoocharoensil (2023), demonstrating four aspects to differentiate synonyms: formality, collocations, semantic preference, and semantic prosody.

Regarding formality, synonyms can be distinguished in terms of formality. Although some synonyms may share similar senses of meaning, their differences may lie in their formality. One word may often occur in a formal context, whereas its synonym may commonly appear in informal context (Phoocharoensil, 2020). Formal words are words that typically occur in formal contexts such as journals, academic texts, and textbooks. On the contrary, informal words are words that are commonly found in casual contexts, such as in everyday conversation, television, and movies (Niwasworakarn et al., 2023). The following are examples of English words that vary according to their degree of formality:

Formal	Informal
purchase	buy
request	ask
begin	start

LDOCE (Online version)

With respect to collocations, collocation refers to pairs of words that are frequently used together. According to LDOCE, collocation is “the way in which some words are often used together, or a particular combination of words used in this way”. For instance, the verb ‘assess’ seems to co-occur with word ‘reliability’ more frequently than the verb ‘evaluate’ and ‘measure’ even though they share a core meaning (Sridhanyarat & Phoocharoensil, 2023). Collocation appears in a statistically significant manner and is often identified using measures such as Mutual Information (MI) scores in combination with frequency data, which reveal how strongly and regularly

words co-occur in natural language use. Therefore, considering possible collocates is a significant way to differentiate synonyms (Phoocharoensil, 2020).

Furthermore, the phenomenon of collocation and semantic preference are closely related, as semantic preference refers to a semantic environment in which vocabulary items frequently appear together (Phoocharoensil, 2021). This relationship plays an essential role in grouping collocates, enabling the classification of lexical items according to their semantic similarity. In addition, regarding semantic prosody, semantic prosody is closely connected with the concept of collocation and can also serve as a criterion to identify near-synonyms (Phoocharoensil, 2021). According to Louw (1993), semantic prosody is “a consistent aura of meaning with which a word is imbued by its collocates” (p. 157). In other words, semantic prosody concerns the meaning that emerges from the broader context in which a word appears, rather than from the word in isolation (Hunston, 2002). For instance, the word *cause* has a high tendency to appear with negative collocates such as *anxiety*, *problem*, and *damage*.

In conclusion, synonyms may differ in terms of their level of formality and may appear with different collocates. Additionally, synonyms tend to occur in a particular semantic environment and may carry an implicit positive or negative tone based on how they are typically used. For these reasons, understanding the degree of formality, collocation, semantic preference, and semantic prosody is essential for accurately distinguishing synonyms.

Previous Related Studies

Numerous corpus-based studies have examined various types of English near-synonyms. As for verb synonyms, Phoocharoensil (2021) compared *persist* and *persevere* using the Corpus of Contemporary American English (COCA), showing *persist* is linked to negative contexts in academic texts, while *persevere* appears more often on webpages, associated with determination. Muangmeeka and Kanokpermpon (2022) examined *join*, *attend*, and *participate* using Oxford English Dictionary (OAAD) and COCA, highlighting grammatical and collocational differences. For instance, the noun *session* collocates with *attend* and *participate*, whereas it is not a common collocate of the verb *join*. Additionally, the study also pointed out that dictionaries may not list all potential collocational patterns. Similarly, Sridhanyarat and Phoocharoensil (2023) investigated *assess*, *evaluate*, and *measure* across eight COCA genres, identifying distinct noun collocates and shared semantic preferences.

Regarding synonymous adjectives, Thummasoonthorn (2020) studied *dangerous*, *hazardous*, and *risky*, using data from Longman Dictionary of Contemporary English (LDOCE), Merriam-Webster Dictionary, and COCA, noting differences in meaning, formality, and collocations, with *hazardous* being the most formal. Askrapond (2020) analyzed *flammable*, *inflammable*, *combustible*, and *incendiary*, using the online Longman Dictionary of Contemporary English, online Oxford Advanced Learner's Dictionary, COCA, and British National Corpus (BNC), finding core meaning similarities but differences in collocations, grammatical patterns, and dialects; for instance, *oil* collocates with *flammable* and *combustible*, while *dust* only with *combustible*. Selmistratis (2020) examined *succinct*/*concise*, *coherent*/*cohesive*, and *precise*/*accurate* in academic texts within COCA, identifying variations in semantic preference and genre distribution but similar semantic prosody. More recently, Phoocharoensil (2022) studied *primary*, *main*, and *major* using COCA, finding all linked to formal English but differing in collocational patterns and semantic preferences.

Overall, these studies emphasize the importance of using corpus data to distinguish the similarities and differences between synonyms. Their findings demonstrate that synonyms vary in contextual usage. Some tend to be used in more formal contexts, while others are more common in informal settings. Additionally, certain synonyms frequently occur with different collocates, reflecting differences in typical usage patterns and associated meanings. Furthermore, previous studies emphasize that examining semantic preference and semantic prosody can provide an in-depth understanding of the connotative meanings of synonyms. These aspects together contribute to a more comprehensive understanding of synonym differentiation in authentic usage and provide support that near-synonyms cannot be used interchangeably in all contexts.

However, while numerous studies have examined various synonyms, a corpus-based analysis of *careful*, *cautious*, and *wary* remains unexplored. Therefore, this study will explore the similarities and differences among these adjectives regarding their distribution across genres, collocations, semantic preferences, and semantic prosody.

Methodology

In this study, the target words were three synonymous English adjectives, namely, *careful*, *cautious*, and *wary*. These words were selected due to their frequent occurrence in communication, as confirmed by their listing in the Longman Communication 9000, highlighting the 9000 most common words in spoken and written English. In carrying out this study, the data retrieved from the Corpus of Contemporary American English (COCA) were

analyzed. It is regarded as a contemporary and regularly updated corpus that represents authentic and current usage of the English language. This makes COCA an invaluable resource for English language learning and teaching, as it provides data for teachers to demonstrate, and for learners to directly observe how English is used in real-world contexts. In addition, it contains a diverse range of genres, including spoken, fiction, popular magazines, newspapers, and academic texts, which allows for a more comprehensive analysis of language use.

With regard to data collection, for the first research question, frequency data, frequency per million data, and overall frequency data of each target word across eight different genres were generated and analyzed. For the second research question, to explore which nouns frequently co-occur with the target synonyms, frequency levels and the statistical corpus-based method known as Mutual Information (MI) score, which measures collocational strength, were used for analysis. Schmitt (2010) emphasized the importance of considering MI scores together with frequency for determining collocates. The collocates were determined based on MI scores of at least 3 because of its significance in measuring collocational strength (Cheng, 2012). Additionally, a window span of 4L/4R was used for identifying noun collocates of each word, following the default setting in COCA, which is widely adopted in corpus-based studies (Davies, 2008). As a result, the top 30 most frequently listed noun collocate with an MI score above 3 were selected for data analysis. To address the third research question regarding semantic preference and semantic prosody, a total of 100 concordance lines were randomly selected for each target word, resulting in a dataset of 300 concordance lines for qualitative analysis. The extracted noun collocations were then categorized manually into groups, following their semantic preference and their semantic prosody.

Results and Discussion

In this section, the findings are presented according to the three research questions. The findings related to the first research question are presented in the subheading below.

Frequency and Distribution of *Careful*, *Cautious*, and *Wary* across Genres

The findings regarding the frequency and distribution across eight different genres of *careful*, *cautious*, and *wary* are illustrated in Table 1. In COCA, *careful* has the highest raw frequency, followed by *cautious* and *wary* respectively. *Careful* occurs approximately five times more frequently than

cautious and occurs approximately seven times more frequently when comparing to *wary*. Specifically, *careful* occurs in COCA with 47,899 tokens while *cautious* and *wary* occur with 8,547 and 6,615 tokens, respectively. Regarding the number of frequencies, it can be inferred from the result that *careful* is more common and widely used than *cautious* and *wary*.

Table 1

Frequency and Distribution of Careful, Cautious, and Wary across Eight Genres

Genres	<i>Careful</i>		<i>Cautious</i>		<i>Wary</i>	
	Frequency	Per Million	Frequency	Per Million	Frequency	Per Million
Spoken	5,670	44.95	1,324	10.50	448	1.16
Fiction	7,793	65.86	1,172	9.91	1,318	11.14
TV/movie subtitles	**12,661	98.86	*424	3.31	*149	3.55
Blogs	4,472	34.77	922	7.17	676	5.26
Web pages	4,728	38.05	903	7.27	737	5.93
Magazines	4,978	39.48	1,317	10.44	1,266	10.04
Newspapers	*3,138	25.78	**1,406	11.55	**1,362	11.19
Academic texts	4,459	37.22	1,079	9.01	659	5.50
Total	47,899		8,547		6,615	

Notes: ** refers to the genre with the highest frequency.

* refers to the genre with the lowest frequency.

Regarding genres, as illustrated in Table 1, the target synonym *careful* is most commonly used in TV and movie subtitles and appears the least in newspapers genre. In particular, *careful* mainly appears in TV and movie subtitles, fiction, and spoken genres with 12,661 tokens, 7,793 tokens, and 5,670 tokens, respectively. In addition, *careful* appears in magazines, web pages, blogs, academic texts, and newspapers genres with 4,978 tokens, 4,728 tokens, 4,472 tokens, 4,459 tokens, and 3,138 tokens, respectively. On the other hand, *cautious* has the highest frequency in newspapers and appears the least in TV and movie subtitles. Particularly, *cautious* mainly appears in newspapers, spoken, and magazines genres with 1,406 tokens, 1,324 tokens, and 1,317 tokens, respectively. In addition, *cautious* appears in fiction, academic texts, blogs, web pages, and TV and movie subtitles genres with 1,172 tokens, 1,079 tokens, 922 tokens, 903 tokens, and 424 tokens, respectively. Taking the adjective *wary* into consideration, similarly to *cautious*, *wary* mainly appears in newspapers and occurs the least in TV and movie subtitles. Specifically, *wary* appears in newspapers, fiction, and magazines

genres with 1,362 tokens, 1,318 tokens, and 1,266 tokens, respectively. Additionally, *wary* appears less in web pages, blogs, academic texts, spoken and TV and movie subtitles genres with 737 tokens, 676 tokens, 659 tokens, 448 tokens, and 149 tokens, respectively.

Noticeably, considering the number of frequencies of *cautious* and *wary* distributed across various genres, as shown in Table 1, *cautious*, and *wary* are mainly associated with formal genre, supported by having the highest frequency in newspapers (11.55 per million with 1,406 tokens and 11.19 per million with 1,362 tokens, respectively). This suggests that *cautious* and *wary* are more commonly used in formal context as these genres are typically associated with a higher degree of formality. Furthermore, *cautious* seems to be more formal than *wary*, as indicated by its frequency in the academic genre and the newspaper genre. Specifically, *cautious*, and *wary* appear in newspapers with 11.55 per million and 11.19 per million, respectively and appear in academic texts with 9.01 per million and 5.50 per million, respectively.

Unlike *cautious* and *wary*, which are more prevalent in formal genres, *careful* occurs with the highest frequency in informal genre, as evidenced by their highest frequency in TV and movie subtitles (98.86 per million). However, *careful* also occurs frequently across a broader range of genres including both informal (fiction and spoken genres with 65.86 per million, and 44.95 per million, respectively) and formal (academic texts and newspapers genres with 37.22 per million, and 25.78 per million, respectively). This wide distribution across formal and informal genres suggests that *careful* is not restricted to a specific register but belongs to a more neutral or versatile register.

Overall, *cautious* is the most formal adjective, frequently used in newspapers and academic texts. *Wary* is also formal but slightly less common in academic contexts. In contrast, *careful* is more neutral and versatile, appearing widely across both formal and informal genres; it has a particularly strong presence in informal settings, especially in TV and movie subtitles. With respect to the genre distribution and formality degree, the findings correspond to previous research studies supporting the current findings (Askrapond, 2020; Phoocharoensil, 2020a; Thummasoonthorn, 2020) that synonyms can vary in terms of formality.

Noun Collocates with *Careful*, *Cautious*, and *Wary*

This section presents the findings related to the second research question, with a focus on analyzing the common noun collocates of the three target synonyms.

Table 2*Noun Collocates with Careful, Cautious, and Wary*

<i>Careful</i>				<i>Cautious</i>				<i>Wary</i>			
Noun Collocates	F	M	Noun Collocates	Fre	M	Noun Collocates	Fre	MI	Scor		
	r	I		que	I		que	Scor			
	e	S		nc	Sc		nc	e			
	q	c		or	or		or				
	u	or		e	e		e				
	e	e		n	c		y				
1	attention	617	3.46	Approach	268	5.07	eye	266	3.84		
2	consideration	588	5.59	Optimism	195	8.39	investor	72	4.86		
3	analysis	484	3.29	Step	99	3.25	look	70	3.03		
4	planning	441	5.15	Investor	84	4.72	consumer	48	3.49		
5	examination	306	5.27	Consumer	49	3.16	glance	43	5.80		
6	reading	277	3.46	Spending	30	3.39	stranger	40	4.66		
7	observation	178	3.55	interpretatio	30	3.71	expressio	29	3.22		
8	selection	153	3.39	n	Tone	29	n	politician	24	3.16	
9	scrutiny	131	4.98	outlook	20	4.44	gaze	17	4.11		
10	monitoring	130	4.95	Sip	14	4.86	buck	16	3.07		
11	preparation	91	3.17	stance	14	3.85	outsider	14	4.54		
12	calculation	66	3.60	lender	13	4.36	lawmaker	13	3.86		
13	Aim	63	3.23	prediction	13	3.17	regulator	10	4.01		
14	deliberation	58	5.12	dealings	8	4.44	rival	9	3.67		
15	handling	58	4.14	glance	8	3.01	local	9	3.49		
16	inspection	51	3.17	capitalist	7	4.43	construct	9	3.29		
17	placement	45	3.24	lending	7	4.29	or	9	3.15		
18	listening	40	4.18	borrower	7	4.24	competit	9	4.12		
19	supervision	32	3.37	nod	7	3.43	or	7	4.11		
20	wording	28	4.57	endorsemen	7	3.30	shopper	7	3.90		
21	coordination	24	3.03	t	forecast	7	3.30	trout	7	3.79	

<i>Careful</i>				<i>Cautious</i>				<i>Wary</i>	
Noun	F	M	Noun	Fre	M	Noun	Fre	MI	
Collocates	r	I	Collocates	que	I	Collocates	que	Scor	
	e	S		n	Sc		n	e	
	q	c		cy	or		cy		
	u	or			e				
	e	e							
	n								
	c								
	y								
22	balancing	23	4.75	sniff	6	6.15	predator	7	3.46
23	Editing	23	3.43	pragmatism	6	5.73	stance	7	3.21
24	crossing	23	3.13	realism	6	4.01	ambush	5	5.07
25	cultivation	20	3.96	skepticism	6	3.76	scam	5	3.88
26	Driving	17	4.29	surrounding	6	3.74	environmentalist	5	3.86
27	stewardship	16	4.08	shopper	6	3.53	coyote	5	3.71
28	weighing	14	5.62	optimist	6	5.32	feminist	5	3.45
29	experimentati on	14	3.21	liberalizatio n	5	4.80	prey	5	3.41
30	calibration	13	4.21	demeanor	5	4.26	ruler	5	3.25

Table 2 displays the noun collocates associated with the three synonyms: *careful*, *cautious*, and *wary*. To validate their significance, these nouns were selected based on their frequency and an MI score of ≥ 3 . Only 30 noun collocates with a score of ≥ 3 was recruited for each adjective and ranked by the number of frequencies.

As demonstrated in Table 2, *careful* mostly co-occurs with the nouns *attention*, *consideration*, *analysis*, *planning*, and *examination*, while *cautious* frequently precedes such nouns as *approach*, *optimism*, *step(s)*, *investor(s)*, and *consumer*. Furthermore, the noun collocates that often follow *wary* are *eye(s)*, *investor(s)*, *look*, *consumer(s)*, and *glance*. In addition, five noun collocates, namely *consumer*, *glance*, *investor*, *shopper*, and *stance* are shared by the two synonymous adjectives: *cautious* and *wary*. This indicates that the two adjectives are synonyms because they share the same collocates (Muangmeeka & Kanokpermpoon, 2022; Thummasoonthorn, 2020).

Noticeably, the top 30 most frequent noun collocates list, as demonstrated in Table 2, shows that although cautious and wary share some overlaps, *careful* does not share common collocates with *cautious* and *wary*. This indicates that there are no noun collocates shared by all three target synonyms within this list. Nevertheless, additional shared collocates may exist outside of the top 30. However, it is essential to mention that this is because other noun collocates that appear with target synonyms were not chosen for

presentation in this study due to their low MI score. This phenomenon is known as a weak collocation (Phoocharoensil, 2022; Sridhanyarat & Phoocharoensil, 2023).

Semantic Preference and Semantic Prosody

The findings for the third research question regarding semantic preference and semantic prosody of the target words are presented.

Semantic Preference and Semantic Prosody of Noun Collocates with Careful

In this section, semantic preference and semantic prosody of the common noun collocates with *careful* are described.

Table 3

Semantic Preference of Noun Collocates with Careful

Semantic preference	Noun collocates of Careful
1. ADMINISTRATIVE ACTIVITY	balancing, coordination, handling, inspection, monitoring, placement, planning, preparation, stewardship, supervision
2. COGNITIVE/ANALYTICAL PROCESS	aim, analysis, attention, calculation, consideration, deliberation, editing, examination, experimentation, observation, selection, scrutiny, weighing
3. MOTOR SKILLS/EXECUTION	calibration, crossing, cultivation, driving
4. COMMUNICATION/LANGUAGE	listening, reading, wording

Table 3 illustrates noun collocates of the adjective *careful* that were placed in their semantic categories into four different themes with respect to semantic preference. The first theme ADMINISTRATIVE ACTIVITY includes the nouns *balancing*, *coordination*, *handling*, *inspection*, *monitoring*, *placement*, *planning*, *preparation*, *stewardship*, and *supervision* as exemplified in the sentences below.

(1) "... lead to false security regarding the need for ***careful handling*** of pesticides and proper environmental stewardship."

(2) "...already back to normal, under the ***careful stewardship*** of the Anglo-American invasion force..."

The second theme includes the nouns *aim*, *analysis*, *attention*, *calculation*, *consideration*, *deliberation*, *editing*, *examination*, *experimentation*, *observation*, *selection*, *scrutiny*, and *weighing*. This second theme is labeled as COGNITIVE/ANALYTICAL PROCESS theme, involving mental effort and critical thinking. In addition, the majority of noun collocates commonly associated with *careful* are related to this theme. Examples of this collocational structure are below.

(3) "...she pays ***careful attention*** to language and rhythm, using splendid verbs."

(4) "... Beijing's position on Taiwan is the product of a ***careful weighing*** of China's strategic interests."

Interestingly, taking the noun *weighing* into consideration, according to Longman Dictionary of Contemporary English or LDOCE, *weighing* also means to consider something carefully so that you can make a decision about it. In sentence (4), *weighing* clearly corresponds to this meaning, as it is used metaphorically to indicate a careful evaluation and deliberation of strategic interests, rather than referring to any physical measurement of weight. When *weighing* is combined with the adjective *careful*, it further emphasizes the deliberate and thoughtful nature of the decision-making process.

The third category MOTOR SKILLS/EXECUTION, which refers to the physical ability to carry out tasks or actions effectively and accurately, contains the nouns *calibration*, *crossing*, *cultivation*, and *driving*, as exemplified in the sentences below.

(5) "... she introduced alcohol with such ***careful calibration*** that her blood flowed exactly as she desired..."

(6) "I saw nothing of the ***careful cultivation*** of rice paddies and fields that make so much of agrarian Southeast Asia beautiful."

The fourth theme COMMUNICATION/LANGUAGE includes the nouns *listening*, *reading*, and *wording*. Some sample concordance lines are displayed in the following sentences.

(7) “A ***careful reading*** of the pre-IPCC assessments strongly suggests that the...”

(8) “To engage in problem solving experiences requires ***careful listening*** to contributions made.”

To gain a more profound understanding of the target synonyms, the noun collocates of *careful* were closely examined regarding their semantic prosody. As shown in example sentences (1-8), *careful* consistently conveys a neutral prosody, suggesting thoroughness and thoughtful behavior. However, in situations involving physical risk or safety, such as *careful driving*, the prosody can lean slightly toward caution or risk-awareness but without suggesting negativity.

Semantic Preference and Semantic Prosody of Noun Collocates with Cautious

In this section, semantic preference and prosody of the common noun collocates with *cautious* are presented.

Table 4

Semantic Preference of Noun Collocates with Cautious

Semantic preference	Noun collocates of Cautious
1. COGNITIVE/ANALYTICAL PROCESS	endorsement, interpretation, pragmatism, prediction, realism, step
2. ATTITUDE OR PERSPECTIVE	approach, dealings, demeanor, optimism, optimist, skepticism, stance, tone
3. SENSORY INPUT/ NONVERBAL SIGNAL	glance, nod, sip, sniff
4. FINANCE AND ECONOMICS	borrower, capitalist, consumer, forecast, investor, lender, lending, liberalization, outlook, shopper, spending
5. MISCELLANEOUS	surrounding

Table 4 illustrates noun collocates of *cautious* that were placed into categories. The semantic preference of *cautious* falls into four different themes.

As illustrated in Table 4, the theme COGNITIVE/ANALYTICAL PROCESS includes the nouns *endorsement*, *interpretation*, *pragmatism*, *prediction*, *realism*, and *step*. Examples of this collocational structure are below.

(9) “However, as Copeland (2000) points out, there are other more **cautious predictions** that Turing makes elsewhere.”

(10) “Five years ago, the National Institutes of Health issued a **cautious endorsement** of surgery for people with morbid obesity.”

The nouns *approach*, *dealings*, *demeanor*, *optimism*, *optimist*, *skepticism*, *stance*, and *tone* are assigned under the theme ATTITUDE OR PERSPECTIVE. Some examples extracted from the concordance lines are included in the following sentences.

(11) “Bigger competitors like Walmart Inc. and Target Corp. would typically take a more **cautious approach**.”

(12) “I think right now they're taking a **cautious stance**. They haven't issued a formal statement yet, but the idea is...”

The theme SENSORY INPUT/NONVERBAL SIGNAL contains the nouns *glance*, *nod*, *sip*, and *sniff*. An example is shown in the sentence below.

(13) “Elliana was aware of everyone's **cautious glances**, every gasp of disbelief...”

As illustrated in Table 4, most noun collocates which frequently appear with *cautious* predominantly fall under the theme FINANCE AND ECONOMICS, including the nouns *borrower*, *capitalist*, *consumer*, *forecast*, *investor*, *lender*, *lending*, *liberalization*, *outlook*, *shopper*, and *spending* as exemplified in the sentences below.

(14) “BMO Capital Markets analyst Tim Long said that in light of a **cautious outlook** BMO has reduced its 2013 earnings per share ...”

(15) “**Cautious consumers** might want to wait for the inevitable subsequent patches, but those who do...”

(16) “... growth in global goods trade in 2011 to 5.8 per cent from an already **cautious forecast** of 6.5 per cent...”

The remaining noun *surrounding* is assigned to the theme MISCELLANEOUS since it cannot be clearly categorized under any of the other semantic themes.

In terms of semantic prosody, *cautious* often conveys a sense of concern, hesitation, or protective awareness. An analysis of randomly selected concordance lines (sentences 11-16) reveals that the noun collocates of *cautious* range from neutral (e.g., *cautious approach*, *cautious dealings*) to slightly negative prosody (e.g., *cautious forecast*, *cautious glance*), depending on the context.

Semantic Preference and Semantic Prosody of Noun Collocates with Wary

In this section, semantic preference and semantic prosody of the common noun collocates with the adjectives *wary* are shown.

Table 5

Semantic Preference of Noun Collocates with Wary

Semantic preference	Noun collocates of Wary
1. PEOPLE OR SOCIAL ENTITIES	competitor, constructor, environmentalist, feminist, foreigner, lawmaker, local, newcomer, outsider, politician, rival, ruler, stranger
2. FINANCE AND ECONOMICS	investor, consumer, shopper, regulator
3. SENSORY INPUT/ NONVERBAL SIGNAL	expression, eye, gaze, glance, look
4. THREATS AND DANGERS	ambush, buck, coyote, predator, prey, scam
5. ATTITUDE OR PERSPECTIVE	stance
6. MISCELLANEOUS	trout

Table 5 displays the noun collocates *wary* assigned to five different themes. Table 5 demonstrates that the nouns *competitor*, *constructor*, *environmentalist*, *feminist*, *foreigner*, *lawmaker*, *local*, *newcomer*, *outsider*, *politician*, *rival*, *ruler*, and *stranger* typically co-occur with the adjective *wary*. These collocates are assigned according to the theme PEOPLE OR SOCIAL ENTITIES, referring to individuals, groups, or social entities, as exemplified in the sentence below.

(17) "...gradual compromise can proceed without gambling on a single grand bargain among **wary rivals** in Baghdad."

Under the theme FINANCE AND ECONOMICS, the nouns *investor*, *consumer*, *shopper*, and *regulator* are commonly used with *wary* as shown in the sentences below.

(18) “**Wary regulators** and environmentalists have introduced detailed rules and oversight ...”

(19) “**Wary investors** dumped stock and ran to gold.”

The theme SENSORY INPUT/NONVERBAL SIGNAL includes the nouns *expression*, *eye*, *gaze*, *glance*, and *look*. Some examples extracted from the concordance lines are shown below.

(20) “I exchanged a **wary look** with Adam and withdrew the gun from my waistband.”

(21) “They gave me **wary glances**, probably still impressed by my ferocity earlier.”

In addition, the nouns *ambush*, *buck*, *coyote*, *predator*, *prey*, and *scam* are assigned under the theme THREATS AND DANGERS. The theme refers to physical threats, adversarial encounters, or deceptive dangers in both natural and social contexts as exemplified in the sentences below.

(22) “...a career in social services could be the thing to inspire you; just be **wary of scams** and tricksters.”

(23) “...coyotes isn't just the price of their pelts or the heart-thumping approach of a **wary predator**.”

Under the theme ATTITUDE OR PERSPECTIVE, the nouns *stance* commonly appears together with *wary*. An example extracted from the concordance lines is shown below.

(24) “...the United States, South Korea's sunshine policy, and Japan's more **wary stance** toward North Korea.”

Lastly, the remaining noun, which is the noun *trout*, is assigned to the theme MISCELLANEOUS.

Interestingly, the noun *trout* can also possess a figurative meaning, referring to an old or unattractive woman. However, this particular sense did

not appear in the concordance lines selected for analysis, although it may still be present elsewhere within the corpus.

Upon a thorough analysis of semantic prosody, the noun collocates of *wary* exhibit a negative prosody reflecting meanings of suspicion, vigilance, defensiveness, and anticipation of threat or deception. For example, *wary of strangers* or *a wary glance* reflecting a sense of guardedness or suspicion. The prosody here is clearly more negative in comparison with *careful* and *cautious*.

A Summary of Semantic Preference and Semantic Prosody of Noun Collocates with *Careful*, *Cautious*, and *Wary*

In this section, a summary of the analysis of the noun collocates with the three target synonyms regarding their semantic preference and semantic prosody are further examined.

Table 6

Shared Noun Collocates and Semantic Preferences of Careful, Cautious, and Wary

Semantic preference	Noun collocates of Careful	Noun collocates of Cautious	Noun collocates of Wary
FINANCE AND ECONOMICS	-	consumer, investor, shopper	consumer, investor, shopper
SENSORY INPUT /NONVERBAL SIGNAL	-	glance	glance
ATTITUDE OR PERSPECTIVE	-	stance	stance

As discussed earlier in the collocational analysis section, *careful* does not share any noun collocates with *cautious* and *wary*. However, five noun collocates: *consumer*, *glance*, *investor*, *shopper*, and *stance* are shared by *cautious* and *wary*.

Regarding semantic preference, an examination of the shared noun collocates among the adjectives *cautious* and *wary* reveals some overlap. As can be seen in Table 6, interestingly, there are three similar noun collocates, *consumer*, *investor*, and *shopper*, shared under the same theme, FINANCE AND ECONOMICS, for both adjectives *cautious* and *wary*. Additionally, the noun *glance* appears with both *cautious* and *wary* within the SENSORY INPUT/NONVERBAL SIGNAL theme, and the noun *stance* associates with *cautious* and *wary* under the ATTITUDE OR PERSPECTIVE theme.

Noticeably, some noun collocates, such as *consumer*, *investor*, and *shopper*, may overlap with other semantic categories. These nouns can also be classified under the theme of PEOPLE and SOCIAL ENTITIES. However, they were categorized based on their most salient and contextually frequent usage patterns observed in the corpus, which predominantly aligned with financial or economic contexts.

Table 7

Summary of Semantic Preference and Semantic Prosody of Noun Collocates with Careful, Cautious, and Wary

Adjective	Semantic preference	Semantic Prosody
Careful	1. ADMINISTRATIVE ACTIVITY 2. COGNITIVE/ANALYTICAL PROCESS 3. MOTOR SKILLS/ EXECUTION 4. COMMUNICATION/LANGUAGE	Neutral
Cautious	1. COGNITIVE/ANALYTICAL PROCESS 2. ATTITUDE OR PERSPECTIVE 3. SENSORY INPUT/NONVERBAL SIGNAL 4. FINANCE AND ECONOMICS	Neutral to Negative
Wary	1. PEOPLE OR SOCIAL ENTITIES 2. FINANCE AND ECONOMICS 3. SENSORY INPUT/NONVERBAL SIGNAL 4. THREATS AND DANGERS 5. ATTITUDE OR PERSPECTIVE	Negative

Table 7 demonstrates the summary of semantic preference and semantic prosody of noun collocates with *careful*, *cautious*, and *wary*. Regarding semantic preference, the noun collocates of the adjective *careful* share only one theme, COGNITIVE/ANALYTICAL PROCESS, with those of the adjective *cautious*. In addition, the noun collocates of the adjective *cautious* share three themes: SENSORY INPUT/NONVERBAL SIGNAL, FINANCE AND ECONOMICS, and ATTITUDE OR PERSPECTIVE, with those of the adjective *wary*. However, the noun collocates of the adjective *careful* do not share any themes with those of the adjective *wary*.

Furthermore, regarding semantic prosody as illustrated in Table 7, while *careful*, *cautious*, and *wary* all denote attentiveness and a desire to avoid harm or error, they differ significantly in semantic prosody. The overall

semantic prosody of noun collocates of *careful* demonstrate a neutral prosody, whereas noun collocates of *cautious* demonstrate a neutral to negative prosody. *Wary*, on the other hand, tends to exhibit a predominantly negative prosody. In particular, the collocates of *careful* carries a generally neutral tone, suggesting responsibility, precision, and thoughtful action. The collocates of *cautious* demonstrates a neutral to negative prosody, often implying sensible restraint but occasionally hinting at hesitation or indecision. However, the collocates of *wary*, by contrast, tend to have a negative prosody, conveying mistrust, skepticism, or guardedness. Examples of concordance lines, illustrating the prosodic behavior of *careful*, *cautious*, and *wary*, are presented below.

(25) “Rethinking the role of the expert in custody proceedings involves a ***careful balancing*** of the expert's potential roles...”

(26) “The job of a scientist is to establish facts by ***careful experimentation***.”

(27) “I think it is appropriate that the process follow the most ***cautious interpretation*** of the law.”

(28) “Expect customers to take the news with ***cautious optimism***.”

(29) ... cautioned that more ***cautious spending*** by customers and higher costs could hurt its business.”

(30) “Some politicians say the monarch has been forced to take only ***cautious steps*** toward economic reforms...”

(31) “...the fears of a ***wary newcomer*** still armed against the sucker punches...”

(32) “...something that will make it harder to bring the ***wary consumer*** who has already left, back.”

The findings support the dictionary definitions of the target synonyms. *Careful*, defined as “trying very hard to avoid doing anything wrong or to avoid damaging or losing something,” is associated with neutral, responsible behavior and is widely used in contexts requiring precision and attention. This is consistent with its frequent collocations related to administrative activity, motor skills, and communication, and its neutral prosody. *Cautious*, meaning “careful to avoid danger or risks,” emphasizes risk avoidance and aligns with analytical, perceptual, and financial contexts,

showing restraint and occasional hesitation. *Wary*, defined as “being careful because something might be dangerous or harmful,” introduces a distinctly negative emotional layer, highlighted in its frequent collocates related to people, threats, and distrust. The findings show that *wary* is used when caution is motivated by suspicion or fear, which clearly aligns with its negative semantic prosody.

Overall, these findings are consistent with previous studies suggesting that near synonyms, despite sharing core meanings, often differ in their collocational patterns, as well as in their semantic preferences and prosody. The findings of this present study confirm what Phoocharoensil (2021) and Selmistraitis (2020) have demonstrated, that near-synonyms can be distinguished by examining their collocational patterns, semantic preferences and semantic prosody.

Conclusion

Despite being synonyms, *careful*, *cautious*, and *wary* differ in their occurrence across genres and formality, as well as in their common noun collocates, semantic preference, and semantic prosody. Among the three adjectives, *careful* is the most flexible, commonly used in both informal and formal contexts. It appears frequently in everyday language as well as in academic writing. In contrast, *cautious* and *wary* are more appropriate in formal contexts.

In addition, *careful* is most appropriate when referring to administrative tasks, cognitive processes, or physical and communicative action while *cautious* is suited to contexts involving analytical thinking, personal attitudes, sensory input, nonverbal expression, and especially financial or strategic decision-making. *Wary*, however, is best reserved for situations involving a potential threat, mistrust, or skepticism such as those relating to people, institutions, economic uncertainty, or danger. In terms of semantic prosody, *careful* typically conveys a neutral prosody, reflecting precision, responsibility, and attentiveness, making it suitable for tasks or actions requiring thoroughness. *Cautious* has a generally neutral to negative prosody, implying hesitation or deliberate restraint, especially in situations involving risk or uncertainty. In contrast, *wary* carries a negative prosody, often associated with mistrust, suspicion, or fear. It is most appropriate in contexts where emotional alertness or skepticism is intended.

In summary, since *careful*, *cautious*, and *wary* cannot be used interchangeably in every context, it is confirmed that *careful*, *cautious*, and *wary* are near synonyms rather than strict synonyms. Therefore, understanding the level of formality, typical noun collocates, semantic preferences, and prosody

or tone associated with each adjective helps ensure that the word chosen aligns accurately with the intended meaning and context.

Acknowledgements

We extend our sincere gratitude to all the anonymous reviewers of this study for their valuable time and critical insights. Special thanks go to Assistant Professor Dr. Rungsima Jeanjaroonsri for her thoughtful comments and constructive feedback. Her suggestions, along with those of the reviewers, have significantly enhanced the clarity and overall quality of this work.

About the Authors

Preeyanan Assawawattanasuntorn: A Master of Arts student in the English Language Teaching Program, Language Institute, Thammasat University, Thailand.

Vanlee Siriganjanavong: An assistant professor at the Language Institute at Thammasat University in Bangkok, Thailand. She received her PhD in Educational and Applied Linguistics from Newcastle University in the UK. Her research interests include vocabulary learning strategies, reading strategies, discourse analysis, and corpus linguistics.

References

Askrapond, P. (2020). *A corpus-based study of the synonyms flammable, inflammable, combustible, and incendiary* [Master's thesis, Thammasat University]. Thammasat University Digital Repository.
https://ethesisarchive.library.tu.ac.th/thesis/2020/TU_2020_62210_42259_14402_14138.pdf

Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Routledge.

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

Davies, M. (2008–). *The Corpus of Contemporary American English (COCA)* [Corpus]. <https://www.english-corpora.org/coca/>

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.

Jackson, H., & Amvela, Z. E. (2007). *Words, meaning and vocabulary: An introduction to modern English lexicology* (2nd ed.). Continuum.

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.

Longman Dictionary of Contemporary English (6th ed.). (2014). Pearson Education.

Louw, B. (1993). Irony in the text or insincerity in the writer? In M. Bake G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology*. John Benjamins.

Malokani, D. K. A. K., Kazi, S., Chandio, S. P., Mumtaz, S. N., Darazi, M. A., & Farooqui, N. S. (2023). The correlation between English language proficiency and perceived career opportunities: Empirical analysis. *Remittances Review*, 8(4), 4818-4827.

Merriam-Webster. (n.d.). *Careful, cautious, & wary*. In *Merriam-Webster.com dictionary*. Retrieved June 20, 2025, from <https://www.merriam-webster.com/dictionary>

Muangmeeka, P., & Kanokpermpoon, M. (2022). *A corpus-based study of English near-synonyms: join, attend and participate* [Master's independent study, Thammasat University]. https://ethesisarchive.library.tu.ac.th/thesis/2022/TU_2022_64210_42323_17552_25168.pdf

Niwesworakarn, N., Phoocharoensil, S., Petkaew, J., & Intasingh, S. (2023). Collocational patterns of the synonyms *join, participate, and attend*: A corpus-based study. *The International Journal of Communication and Linguistic Studies*, 22(1), 63-68.

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

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.

Phoocharoensil, S. (2022). *Primary, main, and major*: Learning the synonyms through corpus data. *GEMA Online Journal of Language Studies*, 22(4), 76-95. <https://doi.org/10.17576/gema-2022-2204-05>

Potter, S., & Crystal, D. (2020). English language. *Encyclopedia Britannica*. <https://www.britannica.com/topic/English-language/Characteristics-of-Modern-English>

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Selmistraitis, L. (2020). Semantic preference, prosody and distribution of synonymous adjectives in COCA. *GEMA Online Journal of Language Studies*, 20(3), 1-18. <https://doi.org/10.17576/gema-2020-2003-01>

Sridhanyarat, K., & Phoocharoensil, S. (2023). A corpus-based investigation of English near-synonyms: *Assess, evaluate, and measure*. *Humanities, Arts and Social Sciences Studies*, 43(2), 208-219.

Thummasoonthorn, J. (2020). *Dangerous, hazardous, or risky: A corpus-based study of English synonyms* [Master's independent study, Thammasat University]. Thammasat University Digital Repository.
https://ethesisarchive.library.tu.ac.th/thesis/2020/TU_2020_62210_42093_14281_14070.pdf

Viera, R. T. (2018). Vocabulary knowledge in the production of written texts: A case study on EFL language learners. *Revista Tecnológica ESPOL – RTE*, 30(3), 89-105.
<http://www.rte.espol.edu.ec/index.php/technologica/article/view/628/377>

Walters, J. (2004). Teaching the use of context to infer meaning: A longitudinal survey of L1 and L2 vocabulary research. *Language teaching*, 37(4), 243-252.

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

Wilkins, D. (1972). *Linguistics in language teaching*. Edward Arnold.

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