



Synonymy in Context: Analyzing the Usage of *Coming* and *Upcoming* in American English

Supakorn Phoocharoensil

Language Institute

Thammasat University

2 Phra Chan Alley, Phra Borom Maha Ratchawang

Phra Nakhon, Bangkok, Thailand, 10200

Tel: +66089-799-9323

Email: supakorn.p@litu.tu.ac.th

Abstract: This corpus-based study investigates the differences in the usage between the near-synonymous adjectives *coming* and *upcoming*, both meaning “happening soon.” Drawing on the Corpus of Contemporary American English (COCA), the research compares their frequency, genre distribution, and typical noun collocations to reveal distinctions in formality, collocational patterns, and semantic prosody. The findings indicate that *coming* is more prevalent in informal contexts such as spoken English and fiction, whereas *upcoming* tends to occur in more formal, written genres like newspapers and weblogs. Although both adjectives share collocates related to time periods and events, *coming* frequently co-occurs with nouns denoting natural disasters and negative events, suggesting a negative connotation, while *upcoming* is more strongly associated with organized events and entertainment. These distinctions highlight the limited substitutability of near-synonyms and emphasize the importance of collocation and context in vocabulary instruction. Pedagogical implications involve incorporating corpus-based tools into language instruction to foster learners’ awareness of subtle usage differences, with particular attention to the characteristic collocations and connotative meanings linked to each synonym.

Keyword: synonym, corpus, genre, collocation, formality

Received: April 20, 2025 Revised: May 21, 2025 Accepted: June 9, 2025

Introduction

There exists a large number of words whose core meanings appear to be very similar. Such words with closely related meanings, known as synonyms, often pose challenges for English language learners, particularly those with limited exposure to English (Nation, 2022). These learners tend to assume that synonyms, such as *main* and *major*, are always interchangeable in every context of use. While both synonyms can collocate with the same nouns (e.g., *concern*, *challenge*, *goal*, *issue*, *point*, *problem*, *reason*, and *source*), some combinations are more natural than others. For example, while *main road* and *main entrance* sound natural in English, *major road* and *major entrance* may strike native speakers as unusual or awkward (Phoocharoensil, 2022). In other words, substituting one synonym for another can often result in uncommon or awkward word combinations, which may cause various communication issues. Since collocations are essential for achieving fluency (Saito, 2020), using non-standard combinations, such as *do a trip* instead of *make a trip*, can disrupt the natural flow of language. Furthermore, the degree of formality should also be considered, as some

synonyms are more associated with formal contexts, whereas others tend to appear more frequently in informal settings. For instance, although both *do a favor* and *render a favor* are grammatically correct, learners who use the latter in casual conversation may sound unnatural or overly formal.

No near-synonyms are completely identical in all aspects, and substituting one for its synonym may lead to errors for L2 learners of English (Szudarski, 2018). It is widely accepted that most language learners face challenges in distinguishing between the different meanings and uses of synonyms, which often results in confusion (Narkprom, 2024; Selmistraitis, 2020). Gass et al. (2023) assert that learners of English who possess an extensive vocabulary have a considerable advantage in terms of language production. Achieving proficiency in English vocabulary requires learners to develop an understanding of the correct usage of synonyms, as varying word choice can mitigate redundancy in written discourse. Mastery of a broad range of vocabulary not only facilitates effective communication in daily interactions but also contributes to a deeper comprehension of lexical nuances and syntactic structures, thereby offering learners a wider array of options for selecting contextually appropriate words.

It is therefore important for EFL/ESL teachers to raise students' awareness of the limited substitutability of synonyms and then help them identify the main usage differences between synonymous words in the same set. Not only teachers but researchers have been interested in making distinctions between synonyms frequently used. An array of recent studies has used an abundance of data from large corpora, e.g. COCA, to draw some convincing, reliable generalizations on the common usage of synonyms, concentrating on the genres of which the synonyms are characteristic and the typical collocations with which they are combined. These studies rely on different criteria to distinguish synonyms, including degree of formality (Boontam & Phoocharoensil, 2022; Chaengchenkit, 2023; Chaokongjakra, 2023; Imsa-ard, 2021; Narkprom, 2024; Phoocharoensil, 2020b, 2021a, 2021b, 2022; Sridhanyarat & Phoocharoensil, 2023), connotations (e.g., Panrat & Yanasugondha, 2024; Phoocharoensil, 2020a, 2021a, 2021b; Selmistraitis, 2020), collocations (e.g., Boontam & Phoocharoensil, 2022; Chaengchenkit, 2023; Chaokongjakra, 2023; Crawford & Csomay, 2024; Jarunwaraphan & Mallikamas, 2020; Narkprom, 2024; Phoocharoensil, 2020a, 2020b, 2021a, 2021b, 2022; Sridhanyarat & Phoocharoensil, 2023), and colligations (e.g., Phoocharoensil, 2021a, 2021b).

The focus of this study was on the two synonymous adjectives *coming* and *upcoming*, both of which share the same core meaning of 'happening very soon'. They are interchangeable in some contexts, as in (1). It is noteworthy that *coming* is more preferable than *upcoming* in such a context as (2), whereas *upcoming*, as opposed to *coming*, is a more natural collocate of *exam* in (3). The researcher first examined the senses of meaning of both synonyms through three learner dictionaries, namely the Longman Dictionary of Contemporary English (2014), the Oxford Advanced Learner's Dictionary (2020), and the Cambridge Dictionary (2013). The Corpus of Contemporary American English (COCA) also provided comprehensive data for the distribution of the target synonyms across genres and the identification of the common noun collocates frequently accompanying them.

- (1) I'm looking forward to the **coming/ upcoming** holiday season.
- (2) The **coming** years will bring significant changes to the industry.
- (3) I have been preparing for the **upcoming** exam.

Literature Review

This section presents an overview of the main concepts that the present study concerns, i.e. corpus-based vocabulary instruction, near-synonyms, and the criteria for synonym differentiation.

Corpus-based Vocabulary Instruction

Language corpora, i.e. large, organized collections of authentic language data, act as useful sources for linguistic researchers and language teachers (Charles & Frankenberg-Garcia, 2023; Durrant, 2022; Pérez-Paredes, 2020). Studies using corpus data can focus on a variety of linguistic aspects, such as lexis, syntax, pragmatics, discourse, sociolinguistics, historical linguistics, stylistics, translation studies. (e.g., Lange & Leuckert, 2019; Lindquist & Levin, 2018; O'Sullivan et al., 2025; Viana & O'Boyle, 2022). Access to corpus data allows for an examination of language use, language pattern identification, and hypothesis testing (Friginal, 2018). Insights from an enormous corpus enable language learners to make more reliable language observations and generalizations from the concordance lines, or lists of real examples showing how a specific word or phrase is used in context within a corpus, that corpus software offers (O'Keeffe & McCarthy, 2024).

It can be seen that corpus data substantially support language teaching and learning (Cheng & Lam, 2024), and one major area in which corpus-informed research often contributes to concerns vocabulary. Teachers can apply corpus data to English Language Teaching (ELT) in several ways, ranging from determining high-frequency vocabulary, exploring common collocations, creating textbooks based on authentic language, constituting word lists for a course in English for Specific Purposes (ESP), promoting learner autonomy through such inductive approaches as Data-Driven Learning (DDL), exploring the degree of formality of vocabulary characterizing each particular genre (Szudarski, 2018). Another lexical aspect that the current corpus-based study mainly addresses is near-synonyms. This study aimed at analyzing subtle differences in meaning and usage between the synonyms *coming* and *upcoming*. Advanced technology facilitates discriminating near-synonyms systematically by looking at a massive corpus like COCA.

Near-Synonyms

One crucial lexical relation is synonymy, defined as a bilateral or symmetrical sense relation in which two or more language forms take an identical conceptual or propositional meaning (Carter, 2012). Examples of synonyms in English include *big*, *large*, *vast*, *huge*, and *gigantic*. Two categories of synonyms exist: absolute synonyms and near-synonyms, the latter of which is the main focus of this study. Absolute or strict synonyms are referred to as words that can be used in place of one another in every possible context without changing the original meaning, tone, or nuance of the sentence (Cruse, 1986). Nonetheless, it is worth noting that such absolute synonyms are extremely rare, if not non-existent, in language as there will be redundancy in language if such words whose meaning and usage are completely identical are allowed. In the historical development of the English language, instances in which two lexical items possessed highly similar meanings and exhibited the potential to become strict synonyms often resulted in semantic divergence, with one term undergoing a shift in meaning or becoming obsolete over time. For instance, in earlier stages of the English language, speakers exclusively employed the native term *heaven* to denote the physical expanse above the earth, encompassing celestial phenomena such as clouds, the sun, the moon, and stars. Subsequently, the word *sky* was borrowed from Old Norse to represent the same concept. For a period, these two lexical items functioned as synonyms and coexisted in semantic competition. Over time, however, a distinct differentiation emerged. In contemporary English, *sky* refers specifically to the observable firmament, while *heaven* has come to signify the spiritual domain associated with divinity and the afterlife (Jackson & Amvela, 2022).

The second category, i.e. loose synonyms or near-synonyms, refers to a word that has a similar meaning to another but is not identical in all contexts of use. In contrast to strict synonyms, loose synonyms can generally be substituted for one another in some situations, but differences in connotation, usage, register, or stylistic tone may prevent them from being interchangeable in every context. These variations apparently reflect distinctions in formality, affective meaning, or

collocational patterns. Clear examples of near-synonyms are *finish* and *end*. Although both synonymous verbs are interchangeable in (4), only *finished my homework* rather than *ended my homework* is considered a natural collocation, as in (5), while it is very common to use *the movie ended* instead of *the movie finished*, as shown in (6). It is the responsibility of English educators to acquaint students with commonly occurring near-synonyms and to elucidate the distinctions in their usage (Phoocharoensil, 2025).

(4) He *finished/ended* his speech with a joke.

(5) I *finished/*ended* my homework.

(6) The movie *ended/*finished* at 10 p.m.

Synonym Differentiation Criteria

Linguists use several criteria to distinguish near-synonyms, such as connotation, collocations, degree of formality, and dialects (Jackson & Amvela, 2022), which will be discussed in detail below:

Connotations

Words can share some denotative meaning but differ in connotations, or emotional associations that words carry beyond literal meaning (Jackson & Amvela, 2022). *While* *cheap* and *inexpensive* both take the meaning ‘costing little money or less than usual’, *cheap* often carries a negative connotation, implying low quality or poor value, as in (7), and *inexpensive* has a neutral or even positive connotation, suggesting affordability without sacrificing quality, as in (8).

(7) He bought some *cheap* snacks.

(8) He found an *inexpensive* way to travel.

Corpus-based researchers explored connotations associated with near-synonyms (e.g., Panrat & Yanasugondha, 2024; Phoocharoensil, 2020a, 2021a, 2021b; Selmistraitis, 2020). Panrat and Yanasugondha (2024), in their analysis of the four synonymous adjectives *clear*, *obvious*, *apparent*, and *evident*, found that *evident*, in contrast to the other three near-synonyms, tends to convey a more positive connotation due to its frequent collocation with nouns that carry positive meanings, e.g. *enthusiasm*, *popularity*, *delight*, *empathy*, *sincerity*, *charisma*, and *dedication*. In Phoocharoensil’s (2020a) study, a detailed examination of the semantic preferences in verb and adjective collocations in COCA reveals subtle distinctions in how the synonyms *consequence*, *result* and *outcome* are used. The word *consequence* is generally linked to words with negative connotations. Although *result* and *outcome* are near-synonyms of *consequence*, it does not exhibit a strong connection to negative or adversarial contexts. Differences in connotations between synonyms are also clearly presented in the study of Phoocharoensil (2021a), in which *persist* and *persevere* were investigated in COCA. The results indicate that *persist* is frequently used in negative contexts, as revealed by semantic prosody analysis, which highlights its strong association with adversity. The subject noun collocates in the examples provided include terms with negative connotations, such as *the stagnant economy*, *unfortunately these conditions*, *active residues of pharmaceuticals*, *symptoms*, *power outages*, *gender-based preferences for competition*, and *our perturbation to the atmospheric concentrations*. In addition to these noun collocates, several surrounding words and phrases also reflect a negative tone, including expressions like *these threats*, *arguing on and on*, *harder to blame*, *their oppressive structure*, *behave irrationally*, *unfortunately*, *the disastrous rise of misplaced power*, *pain*, and *an act of crass stupidity*. On the other hand, its synonym *persevere* frequently appears in Christian contexts, as evidenced by its association with lexical items or phrases commonly linked to religious themes. These include references such as *God*, *saints*, *prayer*, *disbelieved*, *guard against evil*, and expressions like *in the midst of dryness and spiritual aridity*, *Please give me the strength and faith*, *a Catholic school*, *are more faithful to*, and *I ask that you trust in God*. These collocates reflect *persevere*’s strong

connection to spiritual endurance and religious faith.

Collocations

A very effective criterion to differentiate a synonym from another is to look at the collocational pattern in which synonyms usually appear (Xiao, 2015). Although near-synonyms share similar denotative meanings, they are not consistently interchangeable in terms of collocations, i.e. words that frequently occur together (Xiao & McEnery, 2006). That is, each synonym typically combines with a different set of collocates. English language learners who assume that synonymous words can be used with identical collocational patterns are likely to produce expressions that sound unnatural in the target language. As Thornbury (2002, p. 7) notes, “even the slightest adjustments to the collocation – by substituting one of its components for a near-synonym...turns the text into non-standard English.” Consequently, numerous studies examining subtle distinctions between English near-synonyms highlight collocational behavior as a key criterion for differentiation, looking at the data from COCA (e.g. Boontam & Phoocharoensil, 2022; Chaengchenkit, 2023; Chaokongjakra, 2023; Crawford & Csomay, 2024; Narkprom, 2024; Phoocharoensil, 2020a, 2020b, 2021a, 2021b, 2022; Sridhanyarat & Phoocharoensil, 2023).

Among the studies on synonymous nouns and verbs is Boontam and Phoocharoensil (2022), who analyzed the synonymous verbs *expand*, *widen*, and *broaden*, using collocational behavior as one of the key criteria. The findings from COCA revealed that these verbs are not fully interchangeable in all contexts, primarily due to variations in collocation patterns. Although all three verbs shared similar semantic preferences—such as references to areas, knowledge/experiences, and business contexts—they differed significantly in their noun collocates. Only two nouns, *scope* and *horizons*, were common across all three verbs, indicating distinct lexical usage patterns in real-life contexts.

Phoocharoensil (2020a) also conducted a study on the commonly occurring verb and adjective collocates of the near-synonyms *consequence*, *result*, and *outcome*. The analysis revealed differences in usage based on semantic tendencies. The term *consequence* is predominantly linked to collocates with negative connotations, whereas *result*, though semantically similar, is more frequently used in research-related contexts and does not strongly associate with negative meanings. *Outcome*, in contrast, displays a broader semantic range, often appearing with the collocates related to specific topics or categories, such as *psychological*. In another study, Phoocharoensil (2020b) analyzed the synonyms *error*, *fault*, and *mistake* through their collocational behavior. The findings indicated that *error* has the widest variety of verb collocates, while *fault* has the fewest. Similarly, *fault* also has the least variety of adjective collocates. The study also noted key semantic patterns, particularly that *error* and *mistake* share collocates associated with seriousness (e.g., *fatal*, *grave*, *glaring*, *grievous*) and those related to correction (e.g., *correct*, *rectify*). Phoocharoensil (2021a) further examined the synonymous verbs *persist* and *persevere*, which both denote the act of continuing despite difficulty. However, their usage diverges based on context: *persist* is often linked with negatively connoted nouns, while *persevere* is more commonly found in Christian-related contexts and in phrases expressing determination.

Additionally, Kruawong and Phoocharoensil (2022) investigated the synonymous verbs *teach*, *educate*, and *instruct*, focusing on their object noun collocates. The findings showed that *teach* is primarily associated with academic subjects, *educate* with societal and social work contexts, and *instruct* with legal or formal language, highlighting their distinct domains of use. Narkprom (2024), furthermore, found that the verb *restrict* commonly appears with terms linked to governance and individual rights, such as in contexts where laws and regulations limit liberties. In contrast, *constrain* tends to co-occur with vocabulary related to governance and economic matters, as seen in phrases like *rules constrain spending*. Additionally, these verbs exhibit distinct collocational patterns involving nouns with opposite morphological orientations: *restrict* often combines with

nouns containing morphemes that suggest inward movement or limitation (e.g., *acc-* in *access*, as in *restrict access*), whereas *constrain* is typically paired with nouns marked by morphemes indicating outward movement or expansion (e.g., *ex-* in *expansion*, as in *constrain expansion*).

Chaengchenkit's (2023) analysis of noun and adverb collocates reveals that *cease*, *halt*, and *stop* share certain collocates and exhibit similar semantic preferences, supporting their classification as synonyms. However, notable differences also emerge in their collocational behavior and semantic associations. *Halt* is associated with a broader range of thematic categories than *cease* and *stop*, particularly in contexts related to transportation and law. Each verb also tends to co-occur with distinct sets of nouns: *cease* primarily with terms linked to war, conflict, or damage; *halt* with those denoting movement or change; and *stop* with nouns reflecting emotion or action. Regarding adverb collocates, the three synonymous verbs show further differences. *Cease* is most frequently modified by adverbs relating to speed, while *halt* is commonly paired with adverbs indicating duration. For *stop*, adverbial collocates are evenly distributed across themes. Notably, only *cease* co-occurs with an adverb implying low speed, and only *halt* co-occurs with an adverb suggesting a prolonged or indefinite duration.

Synonymous adjectives have also been examined using corpus-based approaches. Phoocharoensil (2022) conducted an in-depth analysis of the adjectives *primary*, *main*, and *major*, revealing notable distinctions in their collocational patterns. The adjective *major* frequently co-occurs with nouns associated with sports (e.g., *championship*, *league*) and business (e.g., *airline*, *corporation*, *manufacturer*). In contrast, *primary* tends to collocate with terms related to healthcare (e.g., *caregiver*, *physician*, *provider*) and political contexts (e.g., *election*, *voter*). The adjective *main* commonly appears alongside nouns referring to physical locations (e.g., *entrance*, *gate*, *road*, *street*), culinary topics (e.g., *course*, *dish*, *ingredient*, *menu*), and literary elements (e.g., *character*). In a subsequent study, Phoocharoensil (2025) explored adjective–noun collocations involving *poisonous* and *venomous*, highlighting their semantic proximity when referring to toxic animals or attacks involving toxins. Both adjectives share noun collocates such as *creature* (e.g., *animal*, *snake*, *species*) and *attack* (e.g., *bite*). However, some important distinctions emerge. *Poisonous* is more likely to co-occur with harmful substances (e.g., *fume*, *gas*, *vapor*) and chemicals (e.g., *mercury*), whereas *venomous* more often collocates with anatomical features of dangerous animals (e.g., *fangs*, *stings*, *spines*). Furthermore, *venomous* also appears in metaphorical contexts, combining with abstract nouns such as *debate*, *reply*, or *hatred*, a usage not typically observed with *poisonous*.

Chaokongjakra (2023) also examined the synonymous adjectives, focusing on the collocational patterns of the synonyms *important*, *significant*, and *crucial*. All three adjectives are frequently used in academic writing, with *important* being the most prevalent. In terms of formality, no notable distinction is observed between *significant* and *crucial*. While these adjectives share certain collocational patterns, each also displays distinct noun and adverb collocates, reflecting unique semantic associations. For example, *significant* is often linked to concepts involving quantity and downward trends, whereas *crucial* tends to occur in contexts related to the political domain. The analysis of the noun and adverb collocates further reveals both shared and different usage patterns among these near-synonyms. *Important* frequently co-occurs with adverbs of intensity, *significant* is often modified by adverbs associated with academic fields or disciplines, and *crucial* is uniquely associated with the theme of culture. These findings not only affirm the synonymy of the adjectives but also underscore the subtle distinctions in their usage, particularly in terms of semantic and contextual preferences.

Degree of Formality

In addition to connotations and collocations, near-synonyms can also be differentiated in terms of

degree of formality. Jackson and Amvela (2022) demonstrate that certain synonyms within the same lexical set are distinctly associated with varying degrees of formality. For example, *ask* is typically used in informal contexts, whereas *plead* and *appeal* are generally considered more formal (Phoocharoensil, 2010). Notably, the inappropriate use of an informal synonym in a formal context—or vice versa—can lead to discrepancies or inconsistencies in language use. For this reason, learners should be aware of the degree of formality of each synonym they use to ensure the appropriateness and compatibility with the specific context.

A growing body of corpus-based research has recently examined the distribution of near-synonyms across various genres, with a particular emphasis on degrees of formality using the Corpus of Contemporary American English (COCA) (e.g., Boontam & Phoocharoensil, 2022; Chaengchenkit, 2023; Chaokongjakra, 2023; Imsa-ard, 2021; Narkprom, 2024; Phoocharoensil, 2020b, 2021a, 2021b, 2022; Sridhanyarat & Phoocharoensil, 2023).

Imsa-ard (2021) examined the distribution of *focus*, *emphasize*, and *concentrate*. Both *focus* and *emphasize* exhibit a high degree of formality, as reflected in their frequent occurrences within academic texts. In terms of overall formality, these two verbs are considered formal, supported by their high frequency in academic genres. However, when examining frequency per million words, *focus* appears to be more prevalent in academic contexts than *emphasize*. The association of *focus* and *emphasize* with formal English is further reinforced by their notably low frequency in informal genres. Specifically, *focus* appears least frequently in fiction and in TV and movie subtitles, while *emphasize* is least frequent in the same informal genres (Moon, 2010). In contrast, *concentrate* occurs most frequently in the magazine genre, which lies between formal and informal registers (Language Institute Thammasat University, 2020). Furthermore, the second and third most frequent occurrences of *concentrate* are found in fiction and TV subtitles—genres typically associated with informal language use.

Boontam and Phoocharoensil (2022), investigating the synonymous verbs *expand*, *widen*, and *broaden*, found that *expand* and *broaden* occur predominantly in formal registers, especially within academic sub-genres such as science and finance. In contrast, *widen* is more frequently found in fiction, reflecting a tendency toward informal usage. Their collocational patterns further reinforce these distinctions: *expand* often co-occurs with formal nouns like *boundaries*, *capabilities*, and *repertoire*, while *broaden* is frequently paired with nouns such as *agenda*, *discussion*, and *perspectives*. Conversely, *widen* typically appears alongside informal vocabulary such as *road*, *pool*, *mouth*, and *lips*.

Similarly, Chaengchenkit (2023) analyzed the synonymous verbs *stop*, *cease*, and *halt*, revealing clear formality-based distinctions. *Stop* was found to be informal, appearing most frequently in the TV and movie subtitles genre, and least frequently in academic texts. In contrast, *cease* and *halt* demonstrate higher levels of formality, with low frequencies in spoken genres and higher occurrences in formal contexts. Specifically, *cease* is most frequent in web pages, particularly legal and fictional sub-genres, while *halt* predominantly appears in newspapers, especially within international news sections. In a related study, Narkprom (2024) reported that the verbs *restrict* and *constrain* are strongly associated with academic genres, with their highest frequencies found in academic texts. These findings are consistent with their inclusion in Coxhead's (2000) Academic Word List (AWL), further supporting their formal and scholarly work.

A number of studies have focused on synonymous nouns. For instance, Phoocharoensil (2020a) conducted a systematic analysis of the similarities and differences among the noun synonyms *consequence*, *result*, and *outcome* across eight text types in the COCA corpus. The findings indicate that these terms are predominantly used in academic texts, highlighting their strong association with formal registers. This formal tendency is further supported by their low frequencies in informal

genres: *consequence* and *outcome* occur least frequently in TV and movie subtitles, while *result* appears least often in fiction. Similarly, Phoocharoensil (2020b) explored the usage distinctions among the synonyms *error*, *fault*, and *mistake*, offering valuable insights into their stylistic variation. His analysis reveals that *error* is chiefly found in academic contexts, affirming its formal character, whereas *mistake* and *fault* are more typical of informal registers, particularly in media such as television and film subtitles.

In addition to noun synonyms, research has also examined synonymous adjectives. Chaokongjakra (2023) found that the adjectives *important*, *significant*, and *crucial* appear more frequently in formal than in informal contexts. Data on frequency show that these adjectives are most prevalent in academic texts: *important* is least used in fiction, while *significant* and *crucial* register their lowest frequencies in TV and movie subtitles. Supporting this trend, Chaokongjakra further observed that all three adjectives are characteristic of formal academic discourse, with *major* also appearing notably in newspapers and magazines. The relatively low occurrence of these terms in informal genres such as fiction and audiovisual media reinforces their alignment with formal language use.

Although there have been corpus-based studies distinguishing usage differences between synonymous adjectives (Chaokongjakra, 2023; Phoocharoensil, 2022, 2025), it seems that no study so far has investigated the near-synonyms *coming* and *upcoming* using the updated function in COCA known as *Word*, which shows the distribution of the search words across genres and also identifies common collocates with which they are used. The present study employs two major criteria—degree of formality across genres and collocations—to differentiate between *coming* and *upcoming*. The reasons why both adjectives were selected lie in the fact that *upcoming* is classified as a C1 word based on the CEFR, which is worth teaching to English learners, while *coming* is the near-synonym of *upcoming* identified by the Cambridge Advanced Learner's Dictionary. Moreover, the differences between synonyms “...are often subtle and difficult to discern” (Liu, 2013, p. 28). That is, the semantic subtlety of adjectives and their dependence on context can also pose challenges. Adjectives like *coming* and *upcoming* may appear similar in meaning but can have different connotations depending on usage. It is worth noting that, while *coming* and *upcoming* can also function as nouns, their frequencies as nouns are far lower than their frequencies as adjectives, thus justifying the selection of the adjectives rather than the nouns.

This study aims to address the following research questions:

1. What are the core meanings of the synonyms *coming* and *upcoming*?
2. What is the distribution of the synonyms *coming* and *upcoming* across various genres?
3. Which nouns commonly collocate with the synonyms *coming* and *upcoming*?

Methodology

Data Collection Procedure

This study utilized the Corpus of Contemporary American English (COCA) as its primary data source. Recognized as one of the most comprehensive and widely referenced corpora for American English, COCA contains over one billion words, with approximately 25 million words added annually between 1990 and 2023. As of March 2020, the corpus covers eight distinct genres: spoken language, fiction, popular magazines, newspapers, academic writing, TV/movie subtitles, blogs, and webpages (Language Institute Thammasat University, 2020).

An outline of the individual genres in COCA is presented as follows:

Table 1

The Descriptions of Genres in COCA

Genres	Descriptions
TV/Movie Subtitles	Consists of 128 million words from American TV and movie subtitles sourced from OpenSubtitles and showcases informal and highly colloquial language—often even more so than everyday spoken language.
Spoken	Contains 127 million words transcribed from unscripted dialogue on over 150 radio and TV programs, including All Things Considered (NPR), Newshour (PBS), and talk shows like Hannity and Colmes and Jerry Springer.
Fiction	Includes 120 million words taken from short stories, plays, literary and children's magazines, popular media, and film scripts.
Popular Magazines	Contains 127 million words from nearly 100 magazines spanning a wide range of subjects, including news, health, finance, religion, and lifestyle.
Newspapers	Comprises 123 million words sourced from ten U.S. newspapers, including USA Today, The New York Times, The Atlanta Journal-Constitution, and The San Francisco Chronicle, across a range of sections.
Academic Journals	Includes 121 million words from nearly 100 peer-reviewed journals organized according to the Library of Congress classification, representing a broad spectrum of academic fields.
Blogs	Contains 125 million words from U.S.-based blogs featured in the GloWbE corpus, originally collected during the period when Google allowed blog-specific search queries.
Webpages	Comprises 130 million words from general U.S. web content—a subset of the GloWbE corpus—featuring blogs and various other online materials.

As part of Brigham Young University's suite of 'mega-corpora', COCA has been extensively used in linguistic research and English Language Teaching (ELT) (Friginal, 2018). Its strengths lie in the ability to track word frequency across balanced genres, analyze common collocates, and compare lexical patterns across registers. Being constantly updated, it remains a reliable resource for examining contemporary native-speaker English, aiding ELT practitioners in keeping pace with current usage trends. As highlighted by Ma and Mei (2021), COCA enables users to explore subtle distinctions in meaning and usage between similar words, making it particularly useful for systematically differentiating near-synonyms through analysis of typical collocates and genre-based tendencies.

Because of its many advantages, COCA was chosen as the primary tool to explore the usage differences between the near-synonyms *coming* and *upcoming*. In addition to its large size, COCA offers advanced concordance capabilities—comparable to software like AntConc or WordSmith Tools—that allow users to search for specific terms and study their usage patterns, including how they commonly co-occur with other words. The latest version of COCA also includes a feature called "Word," which gives detailed information about the 60,000 most frequent words, such as their distribution across different genres, definitions, related topics, and typical collocates (Davies, 2020). To make sure that only adjectival uses of *coming* and *upcoming* were analyzed, since both can also function as nouns, the researcher applied part-of-speech tagging using the code adj.ALL under the CHART interface. This ensured that the frequency data reflected only their adjective forms across various genres.

Data Analysis

This study aims to explore the research questions previously outlined. As an initial step, the researcher referred to three learner dictionaries—the *Oxford Advanced Learner's Dictionary*, the *Longman Dictionary of Contemporary English*, and the *Cambridge Dictionary*—to examine the definitions of *coming* and *upcoming*. This was done to confirm their synonymy in the sense of 'happening very soon,' along with reviewing example sentences featuring both words. Following this, the Corpus of Contemporary American English (COCA) was employed to analyze the frequency and distribution of the two synonyms across COCA's eight genre categories. Since the sizes of these sub-corpora vary, the comparison of occurrences was made using normalized frequency, i.e. frequency per million words, to ensure accuracy (Phoocharoensil, 2020a).

Next, the study analyzed the noun collocates of these synonymous adjectives, focusing on nouns that commonly co-occur with each term. The selection of collocates was based on both their collocational strength and frequency, with Mutual Information (MI) scores, automatically calculated and provided by COCA, serving as a primary measure. While high MI scores reflect strong associative strength between words, frequency was also considered essential to avoid emphasizing rare or atypical pairings not representative of natural language use (Cheng, 2012). Therefore, the study prioritized collocates with both frequent occurrence and high MI scores (≥ 3), identifying the top 30 high-frequency noun collocates in COCA for each synonym.

Subsequently, the study analyzed the shared noun collocates of *coming* and *upcoming*, classifying them according to semantic preference and grouping nouns with related meanings. Noun collocates that were semantically similar or reflected common conceptual domains were categorized under the same thematic label. For example, nouns such as *concert*, *episode*, *film*, *movie*, and *preview* were grouped under the theme *ENTERTAINMENT*. In addition to the researcher's analysis, all collocates were independently coded by a second coder holding a Ph.D. in English Language Teaching. Thematic categories were determined based on an inter-coder agreement rate of 85% or higher. In cases where discrepancies arose between the two coders, discussion was undertaken to reach a consensus on the final categorization. This analysis aimed to reveal potential semantic overlap between the synonyms by identifying collocates that commonly appear with both. Moreover, the investigation highlighted collocates that are uniquely associated with one of the synonyms, thereby illustrating the distinct lexical environments and usage patterns of each target synonym.

Results

This section presents a comparative analysis of the meanings of *coming* and *upcoming* as defined in three learner dictionaries. Additionally, data from the Corpus of Contemporary American English (COCA) will be examined to illustrate the distribution of these two synonyms across eight distinct genres, thereby providing insights into their relative degrees of formality. Furthermore, common noun collocates with Mutual Information (MI) scores of 3 or higher will be identified and subsequently categorized into thematic groups based on their semantic similarities.

Table 2*The definitions of coming and upcoming from three learner dictionaries*

Synonym	Oxford Advanced Learner's Dictionary	Longman Dictionary of Contemporary English	Cambridge Advanced Learner's Dictionary
<i>coming</i>	1) happening soon; next e.g., in the coming months e.g., This coming Sunday is her birthday.	1) happening soon e.g., The Pilgrims prepared for the coming winter. e.g., In the coming year, many people will have their homes damaged or their property stolen.	1) happening soon: e.g., We look forward to greater success in the coming year. e.g., I'll be back this coming Friday.
<i>upcoming</i>	1) going to happen soon e.g., the upcoming presidential election e.g., a single from the band's upcoming album	1) happening soon e.g., He was preparing the federal budget for the upcoming fiscal year. e.g., The wrinkles around his mouth pulled down sternly as he talked about the upcoming missions.	1) happening soon e.g., Tickets are selling well for the group's upcoming concert tour. e.g., Check the website for a calendar of upcoming events.

In Table 2, the definitions of *coming* and *upcoming* from the three dictionaries namely the Oxford Advanced Learner's Dictionary, the Longman Dictionary of Contemporary English, and the Cambridge Dictionary, are exhibited. It seems to be very clear from the three reliable sources that both adjectives are similar in meaning of 'happening soon', thus being considered synonyms of each other. As can be seen from some of the available examples, both target synonyms can sometimes be combined with the same nouns, such as *year*, which in a way ensures their being synonymous but presents a real challenge to English learners.

Table 3*Distribution of the Synonyms Coming and Upcoming across Genres according to Frequency*

Genre	<i>coming</i>		Genre	<i>upcoming</i>	
	Frequency	Per million		Frequency	Per million
spoken	70,203	556.57	weblogs	3,495	27.17
TV and movies subtitles	66,145	516.46	newspapers	2,946	24.20
fiction	37,869	320.05	magazines	2,888	22.90
weblogs	30,133	234.29	webpage	2,645	21.29
webpage	26,462	212.97	spoken	1,601	12.69
newspapers	25,778	211.74	academic texts	871	7.27
magazines	21,259	168.60	fiction	587	4.96
academic texts	8,304	69.32	TV and movies subtitles	288	2.25
Total	286,153		Total	15,321	

As illustrated in Table 3, the adjective *coming* (286,153 occurrences) appears significantly more frequently than its counterpart *upcoming* (15,321 occurrences), indicating that *coming* is much more commonly used. When examining the genres where these synonyms appear, *coming* is notably linked with informal language, as it is most frequent in spoken English (556.57 per million words), TV and movie subtitles (516.46 per million words), and fiction (320.05 per million words), respectively. It is important to highlight that among these three genres typically reflective of colloquial English, TV and movie subtitles represent especially informal usage. Much of this language data is sourced from American television series and shows, often featuring dialogue that

is as informal as, or even more formal than, that found in the spoken corpus. Moreover, *coming* shows its lowest frequency in academic writing (69.32 per million words), reinforcing its rarity in the more formal contexts.

On the other hand, *upcoming* tends to be more prevalent in written and relatively formal genres. Its highest frequencies are found in weblogs (27.17 per million words), newspapers (24.20 per million words), and magazines (22.90 per million words), all of which typically exhibit a more formal tone. It is of interest that *upcoming* occurs least often in fiction (4.96 per million words) and TV and movie subtitles (2.25 per million words) respectively, i.e. genres considered to be highly informal, which clearly suggests that it is seldom used in casual or conversational English. An analysis of the two synonymous adjectives indicates that, despite having the same core meaning, they vary in formality, with *upcoming* occurring more frequently in formal contexts than *coming*. Simply put, the two near-synonyms differ with respect to their degree of formality.

Table 4

Noun Collocates of Coming and Upcoming in COCA

Rank	Noun Collocate	<i>Coming</i> Frequency	MI Score	Noun Collocate	<i>Upcoming</i> Frequency	MI Score
1	year	2705	3.77	election	1059	5.77
2	week	1884	5.09	season	683	4.66
3	month	1546	5.15	event	558	4.53
4	day	1070	3.13	film	415	4.26
5	decade	721	5.71	movie	361	3.93
6	season	283	3.51	show	260	3.22
7	election	274	3.94	album	238	5.38
8	century	152	3.39	project	219	3.32
9	weekend	151	4.00	release	196	5.05
10	storm	140	4.35	series	195	3.49
11	winter	136	4.28	trip	189	4.06
12	generation	98	3.30	meeting	178	3.65
13	crisis	95	3.50	episode	159	4.67
14	battle	83	3.29	holiday	148	4.74
15	apocalypse	82	7.44	session	145	4.33
16	attraction	74	5.42	conference	141	3.64
17	collapse	73	5.00	trial	134	3.55
18	wave	59	3.13	tour	126	4.23
19	revolution	45	3.30	wedding	121	4.34
20	kingdom	37	4.07	debate	111	3.24
21	millennium	34	4.96	schedule	100	4.41
22	anticipation	23	4.82	concert	90	4.64
23	messiah	21	5.45	weekend	83	3.02
24	dawn	20	3.69	visit	76	3.53
25	flood	20	3.28	winter	75	3.30
26	shortage	20	3.73	summit	74	4.93
27	catastrophe	18	4.51	preview	67	6.25
28	cliff	18	3.14	preparation	62	4.26
29	anarchy	17	5.61	anniversary	60	4.51
30	preview	17	4.38	birthday	58	3.46

The nouns that strongly collocate with the synonyms *coming* and *upcoming*, as identified by a mutual information (MI) score of ≥ 3 , are presented in Table 4. As linguistic evidence supports the status of both adjectives as near-synonyms, it is apparent that several noun collocates are shared by the two words, such as *election*, *season*, and *preview*. It should be noted that some other nouns, although able to combine with both *coming* and *upcoming*, were not extracted for specific reasons. First, some nouns co-occur with both adjectives but occur less frequently than those listed in the

top 30 in Table 4, and were therefore not selected. For example, the noun *anticipation*, identified as a collocate of *coming*, can also be modified by *upcoming*, but it did not meet the selection criteria owing to its low frequency. Similarly, nouns such as *summit* and *anniversary*, which are common collocates of *upcoming*, can also occur with *coming*; however, they were excluded due to low frequency. Another reason why certain nouns that can collocate with both target synonyms appear in only one of the two lists is that their MI scores with the other adjective fell below the threshold of 3, i.e. the statistically significant level. Although *episode* and *summit* were strong collocates of *upcoming*, their MI scores with *coming* were below 3 (2.18 for *episode* and 2.16 for *summit*), which explains why they did not appear on the list of collocates.

Table 5

An Analysis of the Semantic Preference in Noun Collocates of Coming

Themes	Collocates
TIME PERIOD	century, day, decade, millennium, month, season, week, weekend, winter, year
DISASTER	apocalypse, battle, catastrophe, collapse, cliff, crisis, flood, shortage, storm
POLITICAL SYSTEM & GOVERNANCE	anarchy, election, kingdom, revolution
PEOPLE	generation, messiah
ENTERTAINMENT	attraction, preview
NATURE	dawn, wave

After the noun collocates were extracted based on the frequency and MI-score criteria, they were also categorized according to the semantic preference. All the noun collocates being classified into six themes are shown in Table 5. Of all the themes, TIME PERIOD, including *century*, *day*, *decade*, *millennium*, *month*, *season*, *week*, *weekend*, *winter*, and *year*, as exemplified in (7), contain the highest number of collocates, followed by DISASTER, comprising *apocalypse*, *battle*, *catastrophe*, *collapse*, *cliff*, *crisis*, *flood*, *shortage*, and *storm*, as in (8), whereas PEOPLE, ENTERTAINMENT, and NATURE, as illustrated in (9)-(11) respectively, are among the lowest with only two members. Another theme with four collocates is POLITICAL SYSTEM AND GOVERNANCE, containing *anarchy*, *election*, *kingdom*, and *revolution*, as in (12).

(7) If you're aiming to fill something for a weekend for example, you might make the offer on a Thursday and have an editorial that talks about what you could do with your ***coming weekend***.

(8) From the far north they heard a low wail of the wind, the long grass bowed in waves before the ***coming storm***.

(9) And these institutions are the places where this ***coming generation*** is learning what an elephant is.

(10) Kathy and I didn't absorb as much as we might have of the movie we had actually come to see (The Revenge of the Giant Leeches), since right before it they showed a movie trailer for a ***coming attraction*** entitled A Summer Place.

(11) Mercury and Spica are noticeable for being close together in the sky and for their brightness before the ***coming dawn***.

(12) He must and will be severed from the Presidency in this ***coming election***.

Table 6*An Analysis of the Semantic Preference in Noun Collocates of Upcoming*

Themes	Collocates
ENTERTAINMENT	album, concert, episode, film, movie, preview, project, release, show
TIME PERIOD	anniversary, birthday, holiday, schedule, season, session, weekend, winter
EVENT	debate, conference, event, meeting, summit, trial, wedding
TRAVEL	tour, trip
POLITICAL SYSTEM & GOVERNANCE	election

Like the classification of the noun collocates of *coming* according to their semantic preferences, those typically co-occurring with *upcoming* were also grouped into themes based on their similarities in meaning, as demonstrated in Table 6. The theme ENTERTAINMENT was found to have the highest number of collocates, *album, concert, episode, film, movie, preview, project, release, and show*, as illustrated in (13). This was immediately followed by TIME PERIOD, a theme comprising eight collocates, *anniversary, birthday, holiday, schedule, season, session, weekend, and winter*, as exemplified in (14). Ranking third in frequency is the theme EVENT, which includes the collocates *debate, conference, event, meeting, summit, trial, and wedding*. Two noun collocates, *tour* and *trip*, were classified under TRAVEL, while only one, *election*, belongs to POLITICAL SYSTEM & GOVERNANCE.

(13) Here's a James Bond history lesson and how it relates to the ***upcoming film***.

(14) The new stores offer more options for price-conscious consumers during the ***upcoming holidays***.

(15) I'll do some research to get me prepared for an ***upcoming meeting***.

(16) Walton said Rondo was already fighting to travel with the Lakers on this ***upcoming trip*** to Orlando, Miami and Cleveland starting on Saturday.

(17) You don't need to know what your " friends " think about the ***upcoming election***.

It is worth mentioning that the analysis of the noun collocates frequently appearing with *coming* and *upcoming* reveals three shared themes, TIME PERIOD, ENTERTAINMENT, and POLITICAL SYSTEM & GOVERNANCE. Despite these overlapping themes, the two synonyms tend to co-occur with different sets of nouns. More specifically, *coming* frequently collocates with everyday monosyllabic words related to calendar units (i.e. *day, week, month, and year*) and with longer-term time-span expressions (i.e. *decade, century, and millennium*).

Another key observation is that although both synonyms share the same theme ENTERTAINMENT, the noun collocates of *upcoming* cover a wider range of entertainment-related terms than those of *coming*, with *preview* being the only collocate they share. In addition, *coming* is more strongly associated with POLITICAL SYSTEM & GOVERNANCE, having four relevant collocates, whereas *upcoming* has only one (*election*), which both words can co-occur with.

The themes that distinguish the two adjectives include DISASTER, PEOPLE, and NATURE, all of which are linked to *coming* rather than *upcoming*. Conversely, *upcoming* has its own distinct themes, namely EVENT and TRAVEL. Another notable finding concerns the DISASTER theme, which characterizes *coming* but not *upcoming*, suggesting that *coming* may carry a semantic prosody associated with negative connotations (e.g. *apocalypse, catastrophe, collapse, crisis, and flood*), whereas *upcoming* does not seem to collocate with nouns carrying negative meanings.

Discussion

The findings of this study confirm the near-synonymous status of the adjectives *coming* and *upcoming*, as both share the core denotational meaning of ‘happening soon,’ a conclusion supported by definitions in three learner dictionaries, thereby addressing the first research question. A detailed analysis of their distribution across genres in the COCA corpus, along with an examination of their common noun collocates, addresses the second and third research questions. In relation to the second question, differences in genre-specific frequencies reveal distinctions in formality: *coming* is more frequent in informal contexts, such as spoken language, TV and movie subtitles, and fiction, while *upcoming* is more prevalent in formal written genres, including weblogs, newspapers, and magazines. These variations in formality offer a useful criterion for distinguishing between the two synonyms and identifying the types of texts in which each typically appears. This finding is consistent with previous research (e.g., Boontam & Phoocharoensil, 2022; Chaengchenkit, 2023; Chaokongjakra, 2023; Narkprom, 2024; Phoocharoensil, 2020b, 2021a, 2021b, 2022; Sridhanyarat & Phoocharoensil, 2023).

In relation to the third research question, an analysis of the semantic preferences of adjective–noun collocations revealed that *coming* and *upcoming* share three common thematic domains: TIME PERIOD, ENTERTAINMENT, and POLITICAL SYSTEM & GOVERNANCE, supporting their near-synonymous nature. While both adjectives collocate with certain shared nouns, such as *election*, *weekend*, and *preview*, several collocates appear to be exclusive to one or the other. Notably, the data indicate a clear pattern: *coming* is more frequently associated with terms denoting calendar units (e.g., *day*, *week*) and extended time spans (e.g., *decade*, *century*), whereas *upcoming* tends to co-occur with special events or occasions (e.g., *anniversary*, *birthday*). This distinct collocational behavior serves as a useful means of differentiating between near-synonyms (Phoocharoensil, 2020a, 2022). The study underscores how collocational patterns can effectively reveal subtle semantic distinctions between *coming* and *upcoming*, echoing findings from earlier research (e.g., Boontam & Phoocharoensil, 2022; Chaengchenkit, 2023; Chaokongjakra, 2023; Crawford & Csomay, 2024; Imsa-ard, 2021; Narkprom, 2024; Phoocharoensil, 2020a, 2020b, 2021a, 2021b, 2022; Sridhanyarat & Phoocharoensil, 2023).

In addition to the differences in formality and typical collocational patterns, the two synonyms also diverge in terms of connotation. Unlike *upcoming*, the adjective *coming* appears in collocations associated with disasters or crises, as it frequently co-occurs with words that convey negative meanings, such as *apocalypse*, *catastrophe*, *collapse*, *cliff*, *crisis*, *flood*, and *storm*, all of which denote unpleasant or destructive events, as exemplified in (18)–(22). This negative semantic prosody, defined as the consistent evaluative meaning a word acquires through its common collocates (Louw, 1993), is evident in the frequent noun collocates of *coming*, distinguishing it from its synonym *upcoming*, which rarely carries such negative associations. Utilizing connotation as a basis for differentiating near-synonyms proves to be an effective approach for identifying subtle semantic distinctions, in line with findings from previous studies on synonymy (Panrat & Yanasugondha, 2024; Phoocharoensil, 2020a; Phoocharoensil, 2021a).

(18) The ***coming apocalypse*** will be when civil war erupts in America for the 2nd time.

(19) I provide such an explanation, and in the process illustrate how these deep biological and genetic mechanisms for morality interact with the social evolution of humans to result in the particular vector of moral concern which is pushing us towards a ***coming catastrophe***.

(20) The best we can do at this point is prepare and survive the ***coming collapse***.

(21) The ***coming crisis*** could make some of the toughest foreign policy challenges facing the United States – from energy insecurity to the trade deficit to global warming – even more

difficult to resolve.

(22) In this way, they could run more of the **coming flood** through the turbines, generating more electric power, more cash, and waste less of that lucrative water through the flood gates above the spillway of the dam.

Conclusion and Pedagogical Implications

This corpus-based study investigated the synonymous adjectives *coming* and *upcoming*, with a particular attention to their degree of formality and their collocational patterns. The two adjectives differ significantly in terms of formality: *coming* is more prevalent in informal or spoken contexts, such as conversational language, television and film subtitles, and fictional texts, whereas *upcoming* occurs more frequently in written forms, including weblogs, newspapers, and magazines. As noted by Xiao (2015), analyzing the collocational environments in which these adjectives appear offers an effective means of distinguishing them. The noun collocates of *coming* and *upcoming* reveal overlapping thematic domains, namely time periods, entertainment, and political systems, with five shared collocates: *election*, *preview*, *season*, *weekend*, and *winter*. Importantly, *coming* strongly collocates with nouns relating to people, nature, and disasters, whereas *upcoming* is more typically associated with nouns referring to events and travel. Furthermore, an unexpected finding is the negative connotation conveyed by *coming*, as evidenced by its frequent co-occurrence with disaster-related nouns.

The findings of this study highlight the importance of understanding that synonyms have limitations in their substitutability, a key insight for English language learners. Vocabulary instruction should therefore emphasize this point, enabling students to become more aware of the context-specific and collocational appropriateness of near-synonyms. Choosing less suitable synonyms can lead to awkward or unnatural phrasing, a frequent issue for English language learners (Phoocharoensil, 2021a). To address this, the study suggests that teachers introduce data-driven learning (DDL), a corpus-based method that allows learners to observe and discover patterns in language use (Boontam & Phoocharoensil, 2024; Yamtui & Phoocharoensil, 2019). By guiding students in analyzing corpus data from native speakers, teachers can help them independently uncover the subtle differences between *coming* and *upcoming*, while offering support when needed. For learners with lower proficiency or in classrooms with limited time, using paper-based DDL materials is advisable, as it allows for more efficient time management. In contrast, for advanced learners who have access to online corpora like COCA, teachers can encourage them to independently investigate authentic language data to complete tasks designed to be more challenging and complex. A sample DDL lesson plan on the synonyms *coming* and *upcoming* is available in the appendix.

The current study has some limitations that should be considered. First, the selection of the top 30 noun collocates, based on a Mutual Information (MI) score of 3 or higher, may have excluded other relevant nouns present in the COCA database. These nouns were not included because they either did not appear in the top 30 frequency list or had an MI score below 3. Consequently, the results should be interpreted cautiously, as a broader selection could yield additional noun collocates. Second, the study's focus on American English using COCA data limits its scope. Future research could explore synonyms in other first-language English varieties, such as British English, or even in learner English. Furthermore, future studies might employ other criteria, like colligations, in the analysis of these synonyms or other synonym sets. Finally, while this study used MI scores due to their automatic availability in COCA, alternative statistical measures such as t-scores, z-scores, and log-likelihood could potentially produce different results (Gablasova et al., 2017).

References

- Boontam, P., & Phoocharoensil, S. (2022). Broaden your horizons: Distribution and collocational patterns of the English synonyms *expand*, *widen*, and *broaden*. *The International Journal of Communication and Linguistic Studies*, 20(1), 107-123. <https://doi.org/10.18848/2327-7882/CGP/v20i01/107-123>
- Boontam, P., & Phoocharoensil, S. (2024). Effects of data-driven learning on writing complexity, accuracy, and fluency (CAF) of Thai EFL Learners. *PASAA*, 68(1), 229-271.
- Cambridge advanced learner's dictionary (2013). Cambridge University Press.
- Carter, R. (2012). *Vocabulary: Applied linguistic perspective*. Routledge.
- Chaengchenkit, R. (2023). A corpus-based study of the synonyms "cease," "halt," and "stop." *LEARN Journal: Language Education and Acquisition Research Network*, 16(1), 473-494.
- Chaokongjakra, W. (2023). Is it 'important,' 'significant,' or 'crucial'? A corpus-based study of English synonyms. *LEARN Journal: Language Education and Acquisition Research Network*, 16(2), 512-532
- Charles, M., & Frankenberg-Garcia, A. (Eds.) (2023). *Corpora in ESP/EAP writing instruction: Preparation, exploitation, analysis*. Routledge.
- Cheng, W. (2012). *Exploring corpus linguistics. Language in action*. Routledge.
- Cheng, W., & Lam, P. (2024). What can a corpus tell us about language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 299-312). Routledge.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Crawford, W. J., & Csomay, E. (2024). *Doing corpus linguistics*. Routledge.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press.
- Durrant, P. (2022). *Corpus linguistics for writing development: A guide for research*. Routledge.
- Friginal, E. (2018). *Corpus linguistics for English teachers*. Routledge.
- Gablasova, D, Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(1), 155-179 <https://doi.org/10.1111/lang.12225>
- Gass, S. M., Behney, J., Plonsky, L., & Huntley, E. (2023). *Second language acquisition: An introductory course* (6th ed.). Routledge.
- Imsa-ard, P. (2021). A corpus-based study on the meanings, distribution, collocations, and formality of *concentrate*, *emphasize*, and *focus*. *NIDA Journal of Language and Communication*, 26(36), 76-98.
- Jackson, H., & Amvela, E. (2022). *Words, meaning, and vocabulary: An introduction to modern English lexicology*. Bloomsbury Academic.

- Jarunwaraphan, B., & Mallikamas, P. (2020). A corpus-based study of English synonyms: Chance and opportunity. *rEFlections*, 27(2), 218-245. <https://so05.tci-thaijo.org/index.php/reflections/article/view/248710>
- Kruawong, T., & Phoocharoensil, S. (2022). A genre and collocational analysis of the near-synonyms *teach*, *educate*, and *instruct*: A corpus-based approach. *TEFLIN Journal*, 33(1), 75-97. <http://dx.doi.org/10.15639/teflinjournal.v33i1/75-97>
- Lange, C., & Leuckert, S. (2019). *Corpus linguistics for World Englishes: A guide for research*. Routledge.
- Language Institute Thammasat University (2020, June 25). *The New Corpus of Contemporary American English (COCA 2020)* [Webinar]. <https://www.youtube.com/watch?v=zW9I9C9rPD4>
- Lindquist, H., & Levin, M. (2018). *Corpus linguistics and the description of English*. Edinburgh University Press.
- Liu, D. (2013). Using corpora to help teach difficult-to-distinguish English words. *English Teaching*, 68(3), 22-39.
- Longman dictionary of contemporary English (2014). Pearson Education.
- Ma, Q., & Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *Journal of China Computer-Assisted Language Learning*, 1(1), 177–190. <https://doi.org/10.1515/jccall-2021-2008>
- Moon, R. (2010). What can a corpus tell us about lexis? In A. O’Keefe & M. McCarthy, (Eds.), *The Routledge handbook of corpus linguistics* (pp. 197-211). Routledge.
- Narkprom, N. (2024). A corpus-based investigation into English synonymous verbs ‘restrict’ and ‘constrain’. *LEARN Journal: Language Education and Acquisition Research Network*, 17(1), 688–714.
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.
- O’Keeffe, A., & McCarthy, M. (Eds.). (2024). *The Routledge handbook of corpus linguistics*. Routledge.
- O’Sullivan, J., Amador-Moreno, C. P., & Barron, A. (2025). *Corpus linguistics for sociolinguistics: A guide for research*. Routledge.
- Oxford advanced learner’s dictionary (2020). Oxford University Press.
- 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.
- Pérez-Paredes, P. (2020). *Corpus linguistics for education: A guide for research*. Routledge.
- Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences*, 3(10), 227–245.

- Phoocharoensil, S. (2020a). A genre and collocational analysis of *consequence*, *result*, and *outcome*. *3L: Language, Linguistics, Literature. The Southeast Asian Journal of English Language Studies*, 26(3), 1–16. <http://doi.org/10.17576/3L-2020-2603-01>
- Phoocharoensil, S. (2020b). Collocational patterns of the near-synonyms *error*, *fault*, and *mistake*. *The International Journal of Communication and Linguistic Studies*, 19(1), 1–17. <https://doi.org/10.18848/2327-7882/CGP/v19i01/1-17>
- Phoocharoensil, S. (2021a). Semantic prosody and collocation: A corpus study of the near-synonyms *persist* and *persevere*. *Eurasian Journal of Applied Linguistics*, 7(1), 240–258. <http://dx.doi.org/10.32601/ejal.911269>
- Phoocharoensil, S. (2021b). Multiword units and synonymy: Interface between collocations, colligations, and semantic prosody. *GEMA Online® Journal of Language Studies*, 21(2), 28–45. <http://doi.org/10.17576/gema-2021-2102-02>
- Phoocharoensil, S. (2022). *Primary*, *main*, and *major*: Learning the synonyms through corpus data. *GEMA Online® Journal of Language Studies*, 22(4), 76–89.
- Phoocharoensil, S. (2025). Exploring collocational patterns and genres: An analysis of the use of *poisonous* and *venomous* in American English. *World Journal of English Language*, 15(2), 202–212.
- Saito, K. (2020). Multi- or single-word units? The role of collocation use in comprehensible and contextually appropriate second language speech. *Language Learning*, 70(2), 548–588. <https://doi.org/10.1111/lang.12387>
- Selmistraitis, L. (2020). Semantic preference, prosody and distribution of synonymous adjectives in COCA. *GEMA Online® Journal of Language Studies*, 20(3), 1–18. <http://dx.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*, 23(1), 208–219.
- Szudarski, P. (2018). *Corpus linguistics for vocabulary: A guide for research*. Routledge.
- Thornbury, S. (2002). *How to teach vocabulary*. Longman.
- Viana, V., & O'Boyle, A. (2022). *Corpus linguistics for English for academic purposes*. Routledge.
- Xiao, R. (2015). Collocation. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 106–124). Routledge.
- Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics*, 27(1), 103–129. <https://doi.org/10.1093/applin/ami054>
- Yaemtui, W., & Phoocharoensil, S. (2019). Effectiveness of data-driven learning on enhancing high-proficiency and low-proficiency Thai EFL undergraduate students' collocational knowledge. *Asian EFL Journal*, 23(3.2), 290–314.

Appendix

Lesson Title: Understanding the Synonyms *Coming* and *Upcoming*

Level: B2 (Upper-Intermediate)

Time: 45–60 minutes

Focus: Vocabulary – Synonyms (*coming* vs *upcoming*)

Skills: Reading, Inductive Reasoning, Vocabulary Use

1. Warm-up (5 minutes)

Ask students:

- What's the difference between *coming* and *upcoming*?
- Can you give an example using either word?

Write a few examples from students on the board and highlight different contexts (e.g., events, deadlines, seasons).

2. Concordance Exploration (15 minutes)

Task A: Look at the concordance lines below.

Concordance: “coming” (as an adjective)

1. We are looking forward to the coming holiday season.
2. She plans to move to New York in the coming months.
3. Inflation is expected to rise in the coming year.
4. The coming weeks will be critical for the negotiations.
5. In the coming days, the weather will improve.

Concordance: “upcoming”

1. The company announced its upcoming product launch.
2. Tickets are now available for the upcoming concert.
3. He's preparing for his upcoming exams.
4. The upcoming election has sparked heated debates.
5. Don't miss our upcoming webinar on digital marketing.

3. Discovery Questions (Inductive Task) (10–15 minutes)

Ask learners to **work in pairs or small groups** and answer the following:

1. What types of nouns follow *coming*?
2. What types of nouns follow *upcoming*?
3. Are there any differences in **formality** or **tone**?
4. Is one more common in spoken English than the other?
5. Which one sounds more **planned** or **scheduled**?

Let them share their findings with the class. Guide the discussion with these possible answers:

- *Coming* often appears with **general time words**: days, weeks, months, year, season.
- *Upcoming* usually appears before **specific events**: exam, meeting, concert, webinar.

- *Upcoming* implies something that has already been **scheduled or arranged**.

4. Practice Exercises (15–20 minutes)

Task B: Fill in the blanks with either *coming* or *upcoming*.

1. We will discuss this again in the _____ weeks.
2. Are you attending the _____ conference in Berlin?
3. The _____ year looks promising for the economy.
4. He's working hard to prepare for the _____ exam.
5. The _____ months are likely to bring changes.

Answer Key:

1. coming
2. upcoming
3. coming
4. upcoming
5. coming

5. Wrap-up (5 minutes)

Ask:

- Can you now explain the difference between *coming* and *upcoming*?
- When might you choose one over the other?