

A CORPUS-BASED INVESTIGATION OF ENGLISH NEAR-SYNONYMS: ASSESS, EVALUATE, AND MEASURE

Kietnawin Sridhanyarat^{1*} and Supakorn Phoocharoensil²

¹ Faculty of Archaeology, Silpakorn University, Thailand

² Language Institute, Thammasat University, Thailand

ABSTRACT

***Corresponding author:**
Kietnawin Sridhanyarat
sridhanyarat_k@su.ac.th

Received: 12 April 2022

Revised: 5 August 2022

Accepted: 9 August 2022

Published: 19 April 2023

Citation:

Sridhanyarat, K. and
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.

Near-synonyms have been the subject of a considerable amount of research within the field of corpus linguistics. Insights gained from research regarding near-synonyms allow teachers to design effective English for Specific Purposes (ESP) courses for students who wish to become more proficient in English. Thus, this corpus-based research is devoted to investigating how three near-synonyms, namely *assess*, *evaluate*, and *measure*, are distributed across eight genres. This study further attempts to examine what noun collocates co-occur frequently with the target lexical items. Determined against the Corpus of Contemporary American English (COCA), the findings illustrate that *assess*, *evaluate*, and *measure* appear most frequently in academically-based contexts. It is also evident that *assess*, *evaluate*, and *measure* are generated least frequently in less formal settings (i.e., fiction, TV and movie subtitles, and spoken). The results suggest that the near-synonyms *assess*, *evaluate*, and *measure* are utilized less frequently in colloquial language. Interestingly, the target synonyms appear frequently in web pages, a genre newly generated in the COCA. This emerging evidence provides support for the notion that *assess*, *evaluate*, and *measure* are frequently found in academic discourses in web pages. Regarding collocation, the target synonyms share the same noun collocates (i.e., *effect*, *effectiveness*, and *impact*). In addition, *assess* co-occurs with *reliability* and *validity* more frequently than *evaluate* and *measure*. Pedagogical implications drawn from the findings are also offered in the study.

Keywords: Academic discourses; COCA; corpus linguistics; collocations; near-synonyms

1. INTRODUCTION

Vocabulary is generally recognized as a language feature that plays a crucial role in both second language (L2) learning and pedagogy (Flowerdew, 2022; Phoocharoensil, 2020a; Walker, 2011; Zhang, 2017). Among various linguistic phenomena of L2 vocabulary (e.g., lexical bundles, idioms, and phrasal verbs), collocations, defined as a group of words that frequently co-occur, have provided facilitative support for learners in several respects (Walker, 2011). In academic writing, the mastery of collocations allows novice writers to generate texts as naturally and proficiently as those produced by professional writers (Flowerdew, 2022; Sridhanyarat, 2018). Collocations are thus considered a salient aspect that facilitates writers to become more successful in generating academic articles for journal publication (Flowerdew, 2022; Vathanalaoha and

Tangkiengsirisin, 2018). In today's world of digital communication, collocations are regarded as a key aspect that contributes substantially to mutual intelligibility among speakers from various cultural backgrounds (Jenkins, 2015; Roshid et al., 2018). For these reasons, L2 collocations have been a popular subject of research among a number of scholars within the domain of corpus linguistics (e.g., Phoocharoensil, 2020a; Phoocharoensil, 2020b; Timmis, 2015; Walker, 2011). By definition, a corpus refers to a collection of authentic language used in a particular genre (Brazil, 1995; Walker, 2011). In corpora, there exist both spoken and written discourses that are naturally occurring (Brazil, 1995; Timmis, 2015; Webb and Nation, 2017). In the relevant literature, genre has been used as a useful tool in describing lexical patterns that occur naturally in a particular context. This allows users to gain exposure to authentic usage of a L2 and ultimately leads to successful communication (e.g., Boontam and Phoocharoensil, 2022; Daskalovska, 2015; Kruawong and Phoocharoensil, 2022; Kartal and Yangineksi, 2018; Pellicer-Sánchez et al., 2022; Webb et al., 2012). Recently, numerous researchers have undertaken studies on synonyms regarded as central components of L2 vocabulary to add pedagogical implications to the relevant literature (Petcharat and Phoocharoensil, 2017; Phoocharoensil, 2020a; Phoocharoensil, 2020b; Walker, 2011).

As noted earlier, corpus-based investigations of L2 synonyms are of paramount importance for researchers and teachers. Insights derived from corpus-based research assists in designing English for Specific Purposes (ESP) courses geared particularly toward students from particular fields of study. More concisely, pedagogical implications drawn from corpus-based research findings assist in facilitating students within a particular discourse community to become proficient and successful in English. As such, the current investigation seeks to address the following research questions:

1. To what extent are three near-synonyms, namely *assess*, *evaluate*, and *measure* distributed across different genres as established by the COCA?
2. What collocational patterning co-occurs frequently with the targets *assess*, *evaluate*, and *measure*?

2. LITERATURE REVIEW

2.1 The notion of near-synonyms

This study reviews theoretical frameworks regarding synonyms. In corpus-based research, synonyms fall into two categories: strict synonyms and near-synonyms. The former category of synonym, also known as perfect synonyms, refers to two words that have the same meaning (Jackson and Amvela, 2000; Phoocharoensil, 2020a; Phoocharoensil, 2020b; Taylor, 2002; Yang, 2016). For example, *dad* and *father* are considered strictly synonymous with each other as their meanings are exactly the same (Cruse, 1986). An outstanding characteristic of the absolute synonyms *dad* and *father* is that they can be replaced interchangeably. The latter form of synonym, near-synonyms, are two words that share similar meanings and cannot be employed interchangeably in all linguistic environments. For example, *repair* cannot be substituted interchangeably for its near-synonym *mend*, as in the sentence *My father used to repair his shoes* when it comes to the context of clothes (Phoocharoensil, 2020b: 3). In this study, only near-synonyms are accounted for. The rationale behind this choice consists of two aspects. The first reason why perfect synonyms are disregarded in this study is that they are rare in number (Cruse, 1986; Palmer, 1997). A language where perfect synonyms are frequently used can be considered redundant. Secondly, in order to give the text produced coherence, it is important for learners to avoid strict synonyms.

Again, this study aims to make a distinction between three synonyms, namely *assess*, *evaluate*, and *measure*, with a particular focus on their distribution across genres and collocations. The three synonyms are defined by the Longman Dictionary of Contemporary English (2014) as follows.

Assess

1. To make a judgment about a person or situation after thinking carefully about it

Evaluate

1. To judge how good, useful, successful something is
2. To carefully consider something to see how useful or valuable it is

Measure

1. To find the size, length, or amount of something, using standard units such as inches, meters, etc.
2. To judge the importance, value, or true nature of something

Overall, *assess*, *evaluate*, and *measure* under study are considered near-synonyms which refer to the ability to judge the importance, value, or true nature of something.

2.2 Identification of near-synonyms

In the relevant literature, near-synonyms can be identified by different means. The current study adapts the method of distinguishing near-synonyms from a seminal study of Phoocharoensil (2020b). In this respect, near-synonyms are thus differentiated based on four aspects: formality, collocations, semantic preference, and semantic prosody.

Considering register, a pair of near-synonyms can occur in different degrees of formality. For example, *assistance* tends to occur in formal contexts, while its near-synonym *help* appears in less formal situations. Therefore, these two near-synonyms can be identifiable in terms of formality. Other near-synonyms that can be distinguished according to levels of formality are presented below.

Formal words	Informal words
require	need
inform	tell
postpone	put off
inquire	ask

(Longman Dictionary of Contemporary English, 2014)

Near-synonyms can also be recognized based on collocations. In the relevant literature, collocations can be defined differently. A collocation can be viewed based on semantic transparency. For example, *play football* is considered a free combination as it is produced in a literal sense. However, *draw a conclusion* serves as a collocation as it demonstrates a low level of semantic transparency (Gyllstad and Wolter, 2015; Nesselhauf, 2005). In this study, collocation is defined as a group of words that co-occur frequently in spoken and written discourses. Under this frequency-based model, *strong coffee* is identified as a collocation over *heavy coffee* because the former combination appears more frequently than the latter (Davies, 2021; Webb and Nation, 2017; Webb et al., 2012).

Occasionally, near-synonyms can be distinguished in accordance with semantic preference, by which collocates for a specific node are classified into themes (Cheng, 2012; Sinclair, 2004). In this study, for instance, *impact*, *effect*, and *effectiveness* have been assigned to the theme EFFECT rather than HUMANS.

The concept of semantic prosody can also be adopted as a criterion to identify near-synonyms. For example, the word *cause* tends to co-occur with negative collocates such as *damage* and *delay*. However, *bring about*, a near-synonym of *cause*, is likely to appear with positive nouns such as *growth* and *happiness* (Phoocharoensil, 2020b: 4).

2.3 Previous studies

In this study, previous corpus-based studies focusing particularly on English synonyms are explicitly reviewed. As elucidated in the introduction, interest in corpus-based research has gained popularity among applied linguistics researchers.

In a corpus-based study, Walker (2011) examined groups of semantically associated nouns (i.e., *issue*, *factor*, and *aspect*) in the domain of business English. The Bank of English (BoE) and the British National Commercial Corpus (BNCs) were used to analyze data. It was discovered that the learning of such lexical items can be explained by means of particular linguistic features that influence their construction. For example, the node *process* is linked more frequently to adjectives that show negative attributes (i.e., *lengthy process* and *slow process*) rather than those referring to positive ones (e.g., *quick*, *fast*, and *painless + process*). This negative semantic prosody is not connected with noun phrases that contain the noun *process* or *processes* (e.g., *learning process* and *manufacturing process*). It was also revealed that the collocates *growth* and *feel-good* are used more frequently with *factor* than *issue* and *aspect*. This is probably due to the fact that *growth factor* is used in a more technical sense, while *feel-good factor* is formed in a sense of subjectivity to something. As discussed, the way some collocations are formed can be explained by linguistic features or the processes that play a part in the formation. This unparalleled insight makes the process of learning L2 collocations more successful.

In addition, Aroonmanakun (2015) investigated the two synonyms *fast* and *quick*. Specifically, this study aimed at exploring possible collocational patterns for the targets *fast* and *quick*. Determined against the COCA, the results showed that *fast* and *quick* share different collocates. To elaborate, *fast* tended to co-occur with noun collocates associated with movements (e.g., *attack*, *boat*, and *car*), whereas *quick* co-occurred with noun collocates related to actions taken within a short period of time (e.g., *breakfast* and *reaction*).

Jirananthiporn (2018) also determined how two words: *problem* and *trouble*, vary across the COCA genres. The results revealed that *problem* is used more frequently than *trouble*, and the noun *trouble* is found more frequently in informal settings. Regarding semantic preference, *problem* is likely to co-occur with formal verbs, such as *acknowledge* and *exacerbate*, while informal verb collocates, such as *like* and *want*, tend to be used frequently with the word *trouble*.

Similarly, Petcharat and Phoocharoensil (2017) analyzed three synonyms, namely *appropriate*, *proper*, and *suitable*. In their study, the three aspects of formality, collocations, and grammatical features were taken into account. The data in this study came from multiple sources, e.g., the Longman Dictionary of Contemporary English and the COCA. The findings showed that the target synonyms varied from context to context. This study provides support for the notion that near-synonyms cannot be used interchangeably in all situations.

Phoocharoensil (2020a) paid particular attention to investigating how *error*, *fault*, and *mistake* vary across genres. He also examined possible verb and adjective collocates of the target near-synonyms. Drawn from the COCA, the results indicate that *error* is typically used in academic texts—formal contexts. However, *fault* and *mistake* are used more frequently in less formal contexts, specifically in TV and movie subtitles. Considering collocations, it is evident that *error* has the greatest number of verb collocates (30 collocates), whereas *mistake* and *fault* have only 12 and 5 verb collocates, respectively. The findings further demonstrate that particular verb collocates (i.e., *correct* and *bias*) appear more frequently with the synonyms *error* and *mistake* than with *fault*. From the findings, it is suggested that *error* and *mistake* serve as synonyms that are more closely associated with each other than *fault*.

In accordance with the above synonym studies, Phoocharoensil (2020b) determined how three synonyms: *consequence*, *outcome*, and *result*, are distributed throughout eight genres according to the COCA. In his investigation, possible collocational patterns appearing with the target words were also examined. Derived from the COCA, the findings illustrate that *consequence*, *outcome*, and *result* co-occur most frequently in formal contexts, appearing frequently in academic texts, with *result* being used frequently in research-based contexts. The findings also demonstrate that the occurrences of the target features are lowest in informal registers; *consequence* appears with verbs and adjectives used in a negative sense; and interestingly, it is found that *outcome* has a wide variety of collocates.

As discussed, English synonyms have been the subject of research among numerous researchers in the field of applied linguistics. This line of corpus-based research aimed at investigating how particular synonyms were distributed across genres and what frequently collocated with them.

3. METHOD

3.1 Data collection and analysis

Under investigation, the Corpus of Contemporary American English (COCA) is recognized as the major source of data. There are two reasons for selecting the COCA. Firstly, the COCA is the only large, genre-balanced corpus of American English. Secondly, the COCA is associated with many other types of English corpora, which provides invaluable insights into English variation. It covers more than one billion words of text (25+ million words each year 1990–2019) from eight different genres (i.e., spoken, fiction, popular magazines, newspapers, academic texts, TV and movie subtitles, blogs, and other web pages) (Davies, 2021). In the terminology of the COCA, genre refers to spoken or written discourses that share the same communicative purposes within a particular community. For example, the synonyms *assess*, *evaluate*, and *measure* are mainly distributed in academic texts. In corpus-based research, the COCA further offers authentic American English data that assists the analyses of genre as well as linguistic and lexical features more practically and systematically (Phoocharoensil, 2020a; Phoocharoensil, 2020b; Schmitt, 2010; Thamratana, 2013). Thus, it is reasonable to claim that the COCA serves as a reliable source of data that facilitates the analysis of the genre and lexical features under study.

In this study, data were analyzed in accordance with two main aspects: frequency and distribution, as well as collocational patterns of the target near-synonyms and semantic preference. As elucidated, the current study aimed at investigating how the target synonyms: *assess*, *evaluate*, and *measure*, varied across genres. Thus, these target lexical items were determined against the COCA to see whether they occurred frequently in formal or informal texts. Furthermore, this corpus-based investigation attempted to explore which noun collocates co-occurred frequently with the verbs *assess*, *evaluate*, and *measure*. Verb-noun collocations were targeted because this pattern is considered particularly challenging for learners when compared to others such as adjective-noun and verb-preposition collocations (e.g., Nesselhauf, 2003; Peters, 2016; Sridhanyarat, 2018). Under study, frequency and Mutual Information (MI) scores or values were adopted as criteria for the data analysis. In corpus-based research, frequency is regarded among researchers to be an important variable when it comes to data analysis. In the relevant literature, collocates for particular nodes in the top-20 frequency list from the COCA are considered significant (Cheng, 2012). This benchmark has also been used effectively to identify collocational association by many researchers in the field of corpus linguistics (e.g., Kruawong and Phoocharoensil, 2022; Phoocharoensil, 2020a). Accordingly, in this study the top 25 most frequently listed noun collocates represented in the COCA were targeted. Only noun collocates which had an MI score of ≥ 3 were selected for analysis, corresponding with the level of significance used in collocational analyses in many

previous studies (e.g., Cheng, 2012; Hunston, 2002; Pavia et al., 2019; Phoocharoensil, 2020b; Vu and Peters, in press).

The extracted noun collocations were then grouped based on semantic preference. Specifically, collocates with similar meanings were categorized within a specific theme. Within corpus-based research, particular collocations may vary greatly from context to context (Phoocharoensil, 2020b). Thus, the grouping of noun collocates into a particular theme allowed the researchers to achieve better insight into how the target synonyms were used with their noun collocates.

4. RESULTS AND DISCUSSION

In this section, findings are presented in response to the two established research questions under investigation. In order to gain a thorough understanding of the current results, the two research questions are repeated here. This study attempted to answer to what extent the target synonyms *assess, evaluate, and measure* vary across genres. It also aimed to discover possible noun collocates of the target near-synonyms *assess, evaluate, and measure*, determined against the COCA. Accordingly, this section first features the results regarding the overall frequency of the near-synonyms *assess, evaluate, and measure*. The section then highlights the findings concerning the noun collocates that co-occur commonly with the target synonyms. The results for the first research question are illustrated under Subheading 4.1 below.

4.1 Frequency and distribution of *assess, evaluate, and measure* across genres

Table 1: Frequency and Distribution of *Assess, Evaluate, and Measure* Across Eight Genres

Genres	<i>Assess</i>		<i>Evaluate</i>		<i>Measure</i>		Total frequency
	Frequency	Per million	Frequency	Per million	Frequency	Per million	
Spoken	1,215	9.63	724	5.74	4,421	35.05	
Fiction	369	3.12	219	1.85	2,807	23.72	
Magazines	1,731	13.73	1,491	11.82	8,999	71.37	
Newspapers	1,284	10.55	1,288	10.58	8,327	68.40	
Academic texts	11,034	92.11	8,206	68.50	22,016	183.79	
TV and movie subtitles	307	2.40	258	2.01	1,288	10.06	
Blogs	1,445	11.24	1,893	14.12	7,274	56.56	
Web pages	2,175	17.50	2,289	18.42	9,024	72.63	
Total	19,560		16,368		64,156		100,084

As illustrated in Table 1, *measure* occurs the most frequently among the three near-synonyms. Specifically, the target verb *measure* (64,156 tokens) occurs in the COCA approximately three times more frequently than the other two verbs *assess* with 19,560 tokens and *evaluate* with 16,368 tokens.

Regarding genre, it is evident that the near-synonyms *assess, evaluate, and measure* are typically generated in formal contexts rather than informal contexts. As demonstrated in Table 1 by the number of occurrences, the target lexical items mainly appear in academic texts. More importantly, *measure* with 22,016 tokens is used the most frequently of all the three targets. The highest frequency in academic texts is accompanied by the occurrences of the target items *assess* and *evaluate* with 11,034 tokens and 8,206, respectively.

In addition, Table 1 reveals that the target synonyms are distributed in a similar fashion. It is evident that *measure, assess, and evaluate* occur with their lowest frequencies in informal contexts. In particular, *measure* appears with its lowest frequency in the three genres *TV and movie subtitles* (1,288 tokens), *fiction*, (2,807 tokens), and *spoken* (4,421 tokens). Regarding informality, the distribution of *assess* is the same as that of *measure*. *Assess* is produced with its lowest frequency throughout TV and movie subtitles (307 tokens), fiction (369 tokens), and spoken (1,215 tokens). Of *evaluate*, its frequency is the lowest in informal contexts (i.e., fiction with 219 tokens, TV and movie subtitles with 258 tokens, and spoken with 724 tokens).

Table 1 further demonstrates that the three near-synonyms appear frequently in web pages, a new genre embedded in the COCA. The frequent occurrences of *assess, evaluate, and measure* indicate that web pages contain both academic and non-academic texts (e.g., Phoocharoensil, 2020a; Phoocharoensil, 2020b). Under observation, the targets *assess, evaluate, and measure* are mainly represented in web pages, and are thus

considered academically specific and formal as nouns that collocate with such items are likely to appear in research-based contexts (e.g., *assess the reliability*, *evaluate the efficacy*, and *measure the effectiveness*).

With respect to the degree of formality, the results of this study correspond to previous studies supporting that some synonyms vary in terms of formality. For example, *error* occurs most frequently in academic texts when compared to the near-synonyms *fault* and *mistake* (Phoocharoensil, 2020a). Jirananthiporn (2018) further supports that some synonyms differ in their degree of formality. In particular, it is evident that *problem* occurs more frequently than *trouble* in academic contexts.

The findings for Research Question 2 are presented in the section below. At this juncture, the findings regarding noun collocates of the near-synonyms *assess*, *evaluate*, and *measure* are shown first. The results are then presented concerning semantic preference of the target near-synonyms.

4.2 Noun collocates and semantic preference of *assess*, *evaluate*, and *measure*

In order to gain further understanding of what nouns frequently collocate with *assess*, *evaluate*, and *measure*, the findings are outlined in two respects. First, noun collocates of the target items are highlighted in Table 2. Next, semantic preference, which allows us to classify collocates for a specific node into themes, is demonstrated in Tables 3, 4, and 5.

Table 2: Noun Collocates with *Assess*, *Evaluate*, and *Measure*

Rank	<i>Assess</i>			<i>Evaluate</i>			<i>Measure</i>		
	Noun collocate	Frequency	MI Score	Noun collocate	Frequency	MI score	Noun collocate	Frequency	MI score
1	Impact	397	6.12	Effectiveness	316	8.51	Success	832	5.38
2	Situation	330	5.35	Performance	300	5.76	Performance	394	4.19
3	Damage	287	6.35	Effect(s) 378	Effects (210) Effect (168)	5.53 4.70	Progress	343	4.89
4	Effect(s) 416	Effects (262) Effect (154)	5.59 4.32	Quality	183	5.12	Effectiveness	284	6.38
5	Risk(s) 388	Risk (257) Risks (131)	5.03 6.19	Impact	175	5.20	Distance(s) 359	Distance (266) Distances (93)	4.56 6.15
6	Effectiveness	218	7.71	Information	160	3.36	Quality	261	3.36
7	Quality	214	5.09	Teachers	133	4.39	Impact	230	3.62
8	Performance	299	4.91	Program(s) 222	Programs (120) Program (102)	4.22 3.06	Levels	187	3.26
9	Extent	197	6.22	Progress	107	5.18	Amount	184	3.15
10	Health	168	3.24	Situation	107	3.98	Achievement	179	4.98
11	Differences	145	5.14	Evidence	102	3.68	Temperature	175	4.46
12	Needs	145	3.63	Results	93	3.59	Size	173	3.14
13	Progress	139	5.30	Efficacy	85	7.73	Satisfaction	159	5.26
14	Value	133	4.16	Options	83	5.00	Effects	154	3.11
15	Degree	131	4.92	Success	80	3.97	Activity	153	3.29
16	Relationship	128	3.98	Differences	76	4.47	Strength	152	3.75
17	Level(s) 222	Level (127) Levels (95)	3.45 3.99	Extent	73	5.05	Extent	141	4.03
18	Ability	117	4.21	Outcomes	65	5.55	Attitudes	129	4.66

Table 2: Noun Collocates with *Assess, Evaluate, and Measure* (Continued)

Rank	<i>Assess</i>			<i>Evaluate</i>			<i>Measure</i>		
	Noun collocate	Frequency	MI Score	Noun collocate	Frequency	MI score	Noun collocate	Frequency	MI score
19	Skills	115	4.49	Risk(s)	Risk	3.28	Speed	124	3.29
	Validity	113	7.17		(64) Risks	5.19			
20	Knowledge	111	4.03	Relationship	59	3.12	Degree	122	3.10
21	Changes	98	3.71	Claims	58	4.18	Intelligence	110	3.13
22	Reliability	90	6.92	Candidates	57	4.44	Length	102	3.77
23	Status	90	4.34	Changes	55	3.13	Inflation	101	4.55
24	Role	84	3.19	Ability	54	3.35	Outcomes	101	4.21
25	Potential	80	3.56	Players	54	3.33	Height	83	3.79

As revealed in Table 2 above, only 25 noun collocates with a score of ≥ 3 MI are recruited for each verb. In this study, the target words are recognized as near-synonyms as they co-occur with different collocates (Phoocharoensil, 2020b; Yang, 2016). For example, *reliability* and *validity* co-occur with *assess* more frequently than with *evaluate* and *measure*. Among the three near-synonyms, there are also shared noun collocates. Specifically, *assess, evaluate, and measure* collocate frequently with the nouns *effect, effectiveness, and impact*. This suggests that the three near-synonyms are associated closely with one another. This proposed idea is supported by several previous studies (e.g., Cruse, 1986; Firth, 1957; Phoocharoensil, 2020a; Phoocharoensil, 2020b; Sinclair, 1991; Stubbs, 1995), in that some near-synonyms can be closely related to one another when they share the same collocates. However, the target near-synonyms may occur with other noun collocates that are not demonstrated in the current study. This phenomenon is known as a weak collocation. For example, the verb *assess* may appear with the collocate *growth* as in *assess growth*, a less frequent collocation in this study. Previous literature (e.g., Hill, 2000; Phoocharoensil, 2020a) has demonstrated that weak collocations are not found frequently in daily life. Thus, this useful insight allows teachers to make decisions about what collocations to introduce to their students in English language classrooms.

Moreover, the target verbs are found to co-occur with different noun collocates. For example, *validity* and *reliability*, which are used in association with testing or research instruments, occur frequently with the verb *assess*. As determined against the COCA, the verb *evaluate* appears frequently with nouns associated with humans (i.e., *candidates* and *players*).

It is also evident that the word *measure* co-occurs frequently with nouns relevant to the act of measuring the size or quality of something (e.g., *distance, height, speed, and temperature*). The target near-synonyms seem to appear frequently with shared and different noun collocates. This phenomenon suggests that English near-synonyms depend heavily upon semantic preference. Scholarly discussion with reference to the concept of semantic preference and its crucial role in English synonyms is featured below in Tables 3–5. Table 3 shows that noun collocates of the verb *assess* can be assigned to eight themes.

Table 3: Semantic Preference of Noun Collocates with *Assess*

Semantic preference	Noun collocates of <i>assess</i>
1. EFFECT	effect, effectiveness, impact
2. QUALITY	performance, potential, progress, quality, reliability, role, situation, status, validity, value
3. POSSIBLE DANGER or HARM	damage, risk
4. EXTENT	degree, extent, level
5. CHANGE	changes, difference
6. ABILITY	ability, knowledge, skills
7. LEARNING	learning
8. MISCELLANEOUS	needs, perceptions, relationship

Table 3 demonstrates that the nouns *effect, effectiveness, and impact* under the theme EFFECT commonly co-occur with the verb *assess*. For example, *assess* is commonly used with *impact* as shown in the sentence in Item 1.

1. *A wide range of models are used to assess the impact of different types of shocks and spillovers through various channels.*

It is also evident that *assess* is used to measure the quality of something. QUALITY appears to have the highest number of collocates (i.e., *performance, potential, progress, quality, reliability, role, situation, status,*

validity, and *value*) that belong to the verb *assess*. Example sentences that demonstrate the use of *assess* with QUALITY are shown in Items 2–3 as follows.

2. *Cohen's coefficient was used to **assess the reliability** of the data.*
3. *Confirmatory Factor Analysis (CFA) was carried out **to assess the validity** of the measuring instrument.*

As can be seen in Tables 3, 4, and 5, the noun collocates of the verb *assess* share four themes (i.e., EFFECT, QUALITY, EXTENT, and ABILITY) with those of the verbs *evaluate* and *measure*. To elaborate, the same three nouns *effect*, *effectiveness*, and *impact* are used under the theme EFFECT. There are three noun collocates: *performance*, *progress*, and *quality*, which are shared by the three verbs *assess*, *evaluate*, and *measure* under the theme QUALITY. However, these target verbs co-occur with different noun collocates within the same theme. Under QUALITY, the verb *assess* collocates frequently with *reliability* and *validity*; the verb *evaluate* is strongly associated with particular noun collocates such as *efficacy*, *results*, and *success*; and the verb *measure* strongly co-occurs with *achievements* and *strength*.

As discussed above, the shared noun collocates of the three synonyms are recognized as near-synonyms (Phoocharoensil, 2020a; Stubbs, 1995). In this study, there are different noun collocates that belong to the theme QUALITY (e.g., *assess reliability* as shown in Item 2, *evaluate efficacy* as revealed in Item 15, and *measure achievement* as in Item 26).

According to the COCA, *assess* is employed to describe situations that are related to the theme POSSIBLE DANGER OR HARM, to which the two words *damage* and *risk* are assigned. An example of this observation is illustrated in Item 4 below.

4. *Park staff will **assess the damage** on Saturday and reopen as soon as possible.*

Evidently, *degree*, *extent*, and *level*, which are organized under EXTENT, co-occur frequently with *assess*. This collocational pattern is illustrated in the following sentences.

5. *The scale was designed to **assess the degree** of difficulty in performing activities during daily living.*
6. *Responders did not **assess the level** of risk inside the theater.*

As demonstrated in the COCA, *changes* and *difference* under the theme CHANGE collocate frequently with the target item *assess*. An example of this collocational pattern is shown in Item 7 below.

7. *This article uses repeat photography to **assess changes** in transportation infrastructure and accessibility in the country of Honduras.*

Within the theme ABILITY, the COCA reveals that *assess* is typically used with *ability*, *knowledge*, and *skills*. This collocational pattern is demonstrated in Items 8 and 9.

8. *Japanese and US researchers developed an advanced model to **assess the ability** of black carbon particles to absorb sunlight.*
9. *We haven't come up with an alternative way to **assess the knowledge** students acquire in the pursuit of their degree.*

In this study, we have assigned the noun collocate *learning*, co-occurring with the verb *assess*, to the theme LEARNING. An example of this verb-noun collocation is revealed in Item 10.

10. *We administered a 28-item multiple-choice test to **assess learning** of science content that was part of the intervention.*

In addition, abstract nouns (i.e., *needs*, *perceptions*, and *relationship*) that are identified under the theme MISCELLANEOUS tend to appear with the verb *assess* as determined against the COCA. The use of *assess* with *needs*, *perceptions*, and *relationship* are demonstrated in the sentences below.

11. *The current GME system also fails to properly **assess the needs** of the physician workforce.*
12. *The purpose of this study was to **assess the perceptions** of high school principals regarding RTI implementation.*
13. *Pearson's correlation was used to **assess the relationship** between the concentrations of adiponectin and periodontal parameters.*

Presented in Table 4 are nouns which collocate with the verb *evaluate*.

Table 4: Semantic Preference of Noun Collocates with *Evaluate*

Semantic preference	Noun collocates of <i>evaluate</i>
1. EFFECT	effect, effectiveness, impact
2. QUALITY	efficacy, outcomes, performance, programs, progress, quality, results, situation, success
3. POSSIBLE DANGER or HARM	risk
4. EXTENT	extent
5. CHANGE	change, difference
6. ABILITY	ability, knowledge, skills
7. LEARNING	learning
8. HUMANS	candidates, players, teachers
9. STATEMENT	claims, evidence, information
10. MISCELLANEOUS	option, relationship

As illustrated in Table 4, the verb *evaluate* appears with nouns assigned to 10 themes, seven of which are the same as those discovered with the verb *assess*. Under the theme EFFECT, there are three words: *effect*, *effectiveness*, and *impact*, which correspond to the verb *evaluate*. An example that represents the use of *evaluate* with EFFECT is shown in the sentence below.

14. *The purpose of this review was to **evaluate the effectiveness** of school-based applications of the FRIENDS program for preventing the onset of anxiety.*

As previously noted, the noun collocates categorized under this theme are the same as those observed with the target *assess*.

Under the theme QUALITY, the words *efficacy*, *outcomes*, *programs*, *results*, and *success* do not appear with the word *assess*, but co-occur with the word *evaluate*. Some examples extracted from the concordance lines are included in the following items.

15. *At present it is too early to **evaluate the efficacy** of these measures.*

16. *Efforts to **evaluate the outcomes** of case management services are underway.*

As checked against the COCA, only the noun *risk* under the theme POSSIBLE DANGER OR HARM is used frequently with the target *evaluate*, as shown in Item 17.

17. *They understand how to **evaluate risk**, allocate capital, and manage their time.*

With respect to the theme EXTENT, only the noun *extent* appears to co-occur with the verb *evaluate*. An example of this collocational feature is revealed in Item 18 below.

18. *The panel is primarily interested in the studies that attempt to **evaluate the extent** to which interventions or approaches have an impact on student outcomes.*

Interestingly, CHANGE and ABILITY share the same collocates (i.e., *change* and *difference* and *ability*, *knowledge*, and *skills*, respectively) when they appear with the verbs *assess* and *evaluate*. In Table 4, it is evident that the target *evaluate* is used with concrete nouns associated with humans—i.e., *candidates*, *players*, and *teachers*. Some sentences extracted from the concordance lines with this collocational pattern are:

19. *The scores on these tests help schools **evaluate candidates** from a range of schools to understand how they compare to each other.*

20. *Consultants were hired to help redirect teaching efforts and **evaluate teachers** and staff.*

As analyzed using the COCA, the verb *evaluate* appears with the collocates *claims*, *evidence*, and *information* as assigned to the theme STATEMENT. An example of this collocational structure includes:

21. *Children would benefit most by learning to gather facts, **evaluate evidence**, think critically, and draw conclusions.*

In this study, the remaining words *option* and *relationship* assigned to the theme MISCELLANEOUS tend to co-occur with the verb *evaluate*. Some uses recruited from the concordance lines are illustrated in the following sentences.

22. *This is typically how the respondent would **evaluate different options** in reality.*

23. *The present preliminary study was designed to **evaluate the relationship** between self-efficacy and to access PMBR information.*

Table 5: Semantic Preference of Noun Collocates with *Measure*

Semantic preference	Noun collocates of <i>measure</i>
1. EFFECT	effect, effectiveness, impact
2. QUALITY	achievement, outcomes, performance, progress, quality, strength, success
3. EXTENT	amount, degree, extent, level
4. MEASUREMENT	distance, height, inflation, length, size, speed, temperature
5. ABILITY	intelligence
6. OPINION	attitudes, satisfaction
7. MISCELLANEOUS	activity

As can be seen in Table 5, semantic preference of the verb *measure* falls under seven themes, one of which is the same as that observed in the near-synonyms *assess* and *evaluate*. More importantly, *effect*, *effectiveness*, and *impact* are used with *measure*. Some instances of EFFECT derived from the concordance lines include:

24. *The purpose of this study was to **measure the effectiveness** of the HELPS program.*

25. *Further, very few rigorous studies have been conducted to **measure the impact** of health IT on nursing practice and patient outcomes.*

Within QUALITY, it is evident that nouns, i.e., *achievement*, *outcomes*, *performance*, *progress*, *quality*, *strength*, and *success*, typically collocate with the verb *measure*. As obtained from the concordance lines, some examples showing possible collocates of the verb *measure* are shown in the following items.

26. *The 92 items in this test are designed to **measure achievement** in the basic skills commonly found in state and district curricula.*

27. *Most people **measure success** in wealth.*

It appears that the theme EXTENT has various nouns (i.e., *amount*, *degree*, *extent*, and *level*) that accord with the verb *measure*. Items 28 and 29 below demonstrate this collocational pattern.

28. *Close study via the new technology allowed the researchers to **measure the amount** of magnesium they contained.*

29. *The Rutgers scientists use brain scanners to **measure the levels** of excitement in women.*

In this study, the COCA also yields interesting results regarding semantic preference of the verb *measure*. As displayed in Table 5, *measure* naturally co-occurs with *distance*, *height*, *inflation*, *length*, *size*, *speed*, and *temperature*, which are highlighted under the theme MEASUREMENT. Some instances extracted from the concordance lines are presented in the following items.

30. *Parallax allows us to directly **measure the distance** to a star without assuming any of its properties.*

31. *I thought the best way to **measure inflation** would be through disposable incomes.*

32. *You can actually **measure the size** of the sun by hand.*

Regarding ABILITY, only *intelligence* occurs frequently with the target *measure*. An example sentence is shown in Item 33.

33. *I **measure my intelligence** by the quality of facts and analysis at my fingertips.*

According to the COCA, *measure* is commonly used to determine people's satisfaction and attitudes, which are categorized under OPINION. An example that demonstrates the use of *measure* with the theme OPINION is included in Item 34.

34. *I don't **measure my satisfaction** or success by the big events.*

Under observation, the word *activity*, which serves under the theme MISCELLANEOUS, is typically used with the verb *measure*. Item 35 demonstrates the *measure* + activity structure.

35. *Volunteers in the study wore a headset to **measure electrical activity** with an electroencephalograph, or EEG.*

Previous studies have shown that there are differences in collocations between synonyms (Cheng, 2012; Phoocharoensil, 2020a; Yang, 2016). In this study, for example, *assess* tends to co-occur frequently with nouns associated with the quality of a research study (i.e., *assess reliability* and *validity*), and *evaluate* is strongly associated with people (i.e., *candidates*, *players*, and *teachers*), and nouns relevant to measurement (i.e., *distance*, *height*, *inflation*, *length*, *size*, *speed*, and *temperature*).

5. CONCLUSION

This corpus-based research aimed to address two research questions. Firstly, it attempted to investigate how three near-synonyms, namely *assess*, *evaluate*, and *measure* were distributed across eight genres as determined against the COCA. Secondly, it intended to explore what noun collocates co-occurred frequently with the target lexical items. The findings demonstrated that *assess*, *evaluate*, and *measure* appeared most frequently in academic texts. Such evidence provides strong support for the formal distribution of the synonyms *assess*, *evaluate*, and *measure*. More concisely, the target near-synonyms *assess*, *evaluate*, and *measure* are likely to represent discourses based in formal contexts. As revealed in the findings, the near-synonyms *assess*, *evaluate*, and *measure* are produced least frequently in less formal settings (i.e., fiction, TV and movie subtitles, and spoken). The results portrayed suggest that the synonyms *assess*, *evaluate*, and *measure* are used less frequently in colloquial language. Interestingly, the target near-synonyms appear to be found frequently in web pages, a genre newly generated in the COCA. This emerging evidence provides support for the notion that *assess*, *evaluate*, and *measure* are frequently used in academic discourse in web pages.

Regarding collocational patterns, the target verbs tend to co-occur with the same semantic preference. For example, the targets *assess*, *evaluate*, and *measure* collocate frequently with *effect*, *effectiveness*, and *impact* as assigned to the theme EFFECT. Under observation, it is likely that the target synonyms share different collocates although they are assigned to the same semantic preference. Within the theme QUALITY, the verb *assess* typically co-occurs with the nouns *reliability* and *validity*, which are associated with research studies. However, the verb *evaluate* is more likely to co-occur with the collocates *outcomes* and *results*. Additionally, *achievement* and *strength* are regarded as common collocates of the word *measure*.

The findings of this study provide useful guidelines pertinent to the field of ELT in several respects. The results obtained allow teachers to design English for Academic Purposes (EAP) courses that can be oriented toward advancing students' writing skills. ELT professionals can also adapt insights gained from the current research to develop test items to measure the English skills of advanced learners. Based on the COCA, the words *assess*, *evaluate*, and *measure* are generally used in academic contexts. This evidence implies that the target words are used less frequently in everyday general English and thus should be introduced to students whose proficiency is of an advanced level. Teachers whose interest centers around ELT or language testing may consider target verbs that are suitable for more proficient learners in EAP courses. This useful insight will help learners become more proficient in English.

Some limitations of this study should be elucidated in several aspects. Only the COCA is utilized as the major tool for the data analysis. Therefore, the research findings may not truly reflect the actual distribution of the synonyms across the eight genres. As Xiao and McEnery (2006) suggest, two comparable corpora, regarded as supplementary data, provide more reliable quantification for researchers. Thus, adopting more corpora tools to conduct further studies could allow for more clarification of how the near-synonyms *assess*, *evaluate*, and *measure* are distributed across genres.

REFERENCES

- Aroonmanakun, V. (2015). Quick or fast: A corpus based study of English synonyms. *LEARN Journal: Language Education and Acquisition Research Network* 8(1): 53–62.
- Boontam, P. and 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.
- Brazil, D. (1995). *A Grammar of Speech*. Oxford: Oxford University Press.
- Cheng, W. (2012). *Exploring Corpus Linguistics: Language in Action*. London: Routledge.
- Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning* 28(2): 130–144.
- Davies, M. (2021). *The New Corpus of Contemporary American English (COCA 2021)*. Corpus of Contemporary American English. [Online URL: <https://www.english-corpora.org/coca/>] accessed on October 1, 2021.
- Firth, J. R. (1957). *Papers in Linguistics, 1934–1951*. London: Oxford University Press.
- Flowerdew, L. (2022). Application of Swalesian genre analysis to academic writing pedagogy: A corpus perspective. *LEARN Journal: Language Education and Acquisition Research Network* 15(1): 1–9.
- Gyllstad, H. and Wolter, B. (2015). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning* 66(2): 296–323.
- Hill, J. (2000). Revisiting Priorities: From Grammatical Failure to Collocational Success. In M. Lewis (Ed.), *Teaching Collocation: Further Developments in the Lexical Approach*, pp. 47–69. London: Commercial Colour Press Plc.

- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Jackson, H. and Amvela, E. Z. (2000). *Words, Meaning and Vocabulary: An Introduction to Modern English Lexicology*. London: Cassell.
- Jenkins, J. (2015). *Global Englishes: A Resource Book for Students*. 3rd ed. London: Routledge.
- Jiranthiporn, S. (2018). Is this problem giving you trouble? A corpus-based examination of the differences between the nouns *problem* and *trouble*. *Thoughts* 2018(2): 1–25.
- Kartal, G. and Yangineksi, G. (2018). The effects of using corpus tools on EFL student teachers' learning and production of verb-noun collocations. *PASAA* 55: 100–125.
- Kruawong, T. and Phoocharoensil, S. (2022). A genre and collocational analysis of the near-synonyms *teach*, *educate* and *instruct*: A corpus-based approach. *TEFLIN Journal* 33(1): 75–96.
- Longman Dictionary of Contemporary English*. (2014). Essex: Pearson Education.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics* 24(2): 223–242.
- Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam: John Benjamins Publishing Company.
- Palmer, F. R. (1997). *Semantics*. Great Britain: Cambridge University Press.
- Pavia, N., Webb, S. and Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition* 41(4): 745–768.
- Pellicer-Sánchez, A., Siyanova-Chanturia, A. and Parente, F. (2022). The effect of frequency of exposure on the processing and learning of collocations: A comparison of first and second language readers' eye movements. *Applied Psycholinguistics* 43(3): 1–30.
- Petcharat, N. and Phoocharoensil, S. (2017). A corpus-based study of English synonyms: Appropriate, proper, and suitable. *LEARN Journal: Language Education and Acquisition Research Network* 10(2): 10–24.
- Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research* 20(1): 113–138.
- Phoocharoensil, S. (2020a). Collocational patterns of the near-synonyms error, fault, and mistake. *The International Journal of Communication and Linguistic Studies* 19(1): 1–17.
- Phoocharoensil, S. (2020b). 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.
- Roshid, M. M., Webb, S. and Chowdhury, R. (2018). English as a business lingua franca: A discursive analysis of business e-mails. *International Journal of Business Communication* 59(1): 83–103.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. London: Palgrave Macmillan.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Sridhanyarat, K. (2018). Thai learners' acquisition of L2 collocations: An interlanguage perspective. *GEMA Online Journal of Language Studies* 18(1): 1–21.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language* 2(1): 23–55.
- Taylor, J. R. (2002). Near synonyms as co-extensive categories: 'high' and 'tall' revisited. *Language Sciences* 25(3): 263–284.
- Thamratana, P. (2013). *A Corpus-Based Analysis of Reduce, Decrease, Diminish, Dwindle and Decline*. Master's Thesis. Thammasat University, Thailand.
- Timmis, I. (2015). *Corpus Linguistics for ELT: Research and Practice*. New York: Routledge.
- Vathanalaotha, K. and Tangkiengsirisin, S. (2018). Genre analysis of experiment-based dental research article abstracts: Thai and international journals. *3L: Language, Linguistics, Literature® The Southeast Asian Journal of English Language Studies* 24(3): 1–14.
- Vu, D. V. and Peters, E. (in press). A longitudinal study on the effect of mode of reading on incidental collocation learning and predictors of learning gains. *TESOL Quarterly*.
- Walker, C. P. (2011). A corpus-based study of the linguistic features and processes which influence the way collocations are formed: Some implications for the learning of collocations. *TESOL Quarterly* 45(2): 291–312.
- Webb, S. and Nation, P. (2017). *How Vocabulary Is Learned*. Oxford: Oxford University Press.
- Webb, S., Newton, J. and Chang, A. (2012). Incidental learning of collocation. *Language Learning* 63(1): 91–120.
- Xiao, R. and McEnery, T. (2006). Collocation, semantic prosody and, near synonymy: A cross-linguistic perspective. *Applied Linguistics* 27(1): 103–129.
- Yang, B. (2016). A corpus-based comparative study of *learn* and *acquire*. *English Language Teaching* 9(1): 209–220.
- Zhang, X. (2017). Effects of receptive-productive integration tasks and prior knowledge of component words on L2 collocation development. *System* 66: 156–167.