The Effectiveness of Corpus-Assisted Approach in Learning Grammatical Collocations of Thai Undergraduate Students in an EFL Classroom

PIMNADA KHEMKULLANAT*
Faculty of Liberal Arts, Prince of Songkla University, Thailand
SOMRUEDEE KHONGPUT
Faculty of Liberal Arts, Prince of Songkla University, Thailand
Corresponding author email: pimnadakhem15@gmail.com

Article information	Abstract
	The present study implements a corpus-assisted approach with data-driven learning (DDL) in the EFL classroom to investigate its effectiveness in learning target grammatical collocations (verb-, adjective-, and noun-preposition collocations) of Thai undergraduate students and to examine the extent to which the students incorporate the collocational knowledge learned into their writing. Forty students who were inexperienced in DDL in one intact class at a private university in southern Thailand participated in this study. The participants learned through scaffolded paper-based DDL and autonomous computer-based DDL for a total of 10 weeks in an English for Communication course, which aims to develop their communicative abilities. Pre- and post-writing tests, a stimulated recall interview, and a semi-structured interview were employed as the data collection instruments. The writing test results indicate that the participants' collocational knowledge significantly improved in all patterns (p = 0.00), with a large overall effect size (d = 1.26). The interview results uncover that most participants could accurately: 1) classify the types of the target collocational patterns; 2) identify the hidden usage of the content words with varying prepositions; and 3) elucidate some key considerations when using collocations for their written communication. The results also suggest that the participants have acquired several collocations other than those targeted in DDL. The study concludes with pedagogical implications for DDL implementation and limitations in conducting DDL lessons.

INTRODUCTION

In the context of English as a foreign language (EFL), students at all levels, especially in higher education, are expected to be able to effectively communicate in written English (Chintaradeja, 2020; Seensangworn, 2017). Mastering writing nevertheless can be undeniably arduous for EFL students since writers need to be knowledgeable of paragraph unity, cohesion, and coherence as well as grammar. Still, it was discovered that learners greatly struggled with

writing, and the most serious problem was dealing with grammatical components (Boonyarattanasoontorn, 2017). Several recent studies on the analysis of students' writing have revealed that among various types of writing errors that occurred, collocational patterns associated with prepositions were found to be one of the most frequent types of the errors. Precisely, the troublesome areas involved addition, omission, and misuse of prepositions (Kampookaew, 2020; Latupeirissa & Sayd, 2019; Qamariah & Wahyuni, 2020; Suvarnamani, 2017; Yusuf et al., 2021). This leads to the focus on grammatical collocations in this current study.

Despite the fact that the sources of the errors could be from first language interference and/ or a shortage of collocational knowledge (Boonraksa & Naisena, 2022; Wang & Shaw, 2008; Yumanee & Phoocharoensil, 2013), learning about collocations is crucial as it enables the learners to enhance their communicative competence. Mis-collocations in written production can affect the quality of students' writing since they might lead to misinterpretation of the message and increase the difficulty in interpreting the text. In fact, prioritizing grammatical collocations equips learners with essential skills for effective communication, structural accuracy, and linguistic versatility across numerous contexts. The knowledge of collocations minimizes the risk of making grammatical mistakes in language output, thus leading to more accurate and more understandable written language for readers. Additionally, focusing on grammatical collocations could aid in understanding and using idiomatic expressions appropriately. These expressions often have meanings that are not easily deduced from the individual words themselves. Since collocations concerns with the core meaning of words and they are an inexplicable and indispensable part of language learning (Duan & Qin, 2012), teaching collocations requires an effective teaching method.

One efficient method suggested by substantial studies is implementing corpora in language classrooms where learners can be exposed to the authentic language forms and patterns in a variety of contexts (Friginal, 2018; Lindquist, 2018). As Muftah (2023) asserted, the prevalent collocations exhibited in diverse contexts allow the learners to recognize and subsequently remember the patterns of collocations. Moreover, it was uncovered that corpus use is effective for teaching grammatical collocations, particularly the new ones (Vyatkina, 2016a), and it could be used by the learners regardless of their grammar proficiency (Lin, 2021). Several previous studies (e.g., Fang et al., 2021; Jantarabang & Tachom, 2021; Preradovic et al., 2019; Sun & Hu, 2023) have explored students' views and attitudes towards using corpora and revealed that the students have responded favorably to the direct corpus use or so-called data-driven learning (DDL). However, more research is needed in a broader area, including students with different mother tongues, nationalities, and educational backgrounds (Bridle, 2019). Furthermore, research on the specific effects of the direct corpus use, especially regarding collocation learning, is rare. Thus, it is still uncertain how corpus data can aid in students' collocation learning or how the knowledge they have processed and retrieved from DDL can assist them in producing the collocations (Satake, 2022).

The present study therefore was designed to investigate the effectiveness of DDL in learning grammatical collocations of Thai EFL undergraduate students and examine the extent to which they incorporated the collocations learned into their writing. The results have implications for

teaching the collocations in similar EFL contexts, as the teachers would be able to decide whether the adopted approach can be effective and/or the selected corpus is suitable for their classrooms, and what they should consider when planning DDL lessons. To shed light on such matter, the following research questions were asked:

1. Do pre-test and post-test scores differ significantly after learning grammatical collocations through a corpus-assisted approach in the EFL classroom?
2. To what extent do the Thai EFL students incorporate in their writing their knowledge of grammatical collocations learned through concordances?

LITERATURE REVIEW

Collocational knowledge and collocational errors

To most linguists, words that are frequently used together with other words in a certain grammatical structure are referred to as collocations (Benson et al., 1997; Cruse et al., 1986; Nation, 2005). They are important to language learners as they help facilitate the production of the target language in a manner that is both natural and grammatically correct to native speakers (Dokchandra, 2019; Men, 2017). Hence, the learners need to have receptive collocational knowledge to identify the co-occurrence of words as well as productive collocational knowledge to use these combinations accurately in speech and writing (Alsakran, 2011; Kamarudin et al., 2020). Collocations should be learned since "words seldom occur in isolation" (Wallace 1982, as cited in Duan & Qin, 2012, p. 1891) and collocations "are found in up to 70% of everything we say, hear, read, or write" (Hill, 2000, p. 53). Moreover, collocational knowledge contributes to lexical development by diversifying vocabulary resources, leading to lexical variety in writing (Dokchandra, 2019; Duong & Nguyen, 2021).

Although it is commonly agreed that knowing the company a word has is crucial and valuable, learners often have trouble with learning and using collocations, regardless of English proficiency level (Nagy, 2020; Yaemtui & Phoocharoensil, 2018). This could be due to the inexplicable nature of collocations with no logical justification or explanation for the frequent occurrence of a certain word combination (Benson, 1989; Duan & Qin, 2012). Accordingly, errors probably occur when students' mother tongue is directly translated into the target language (Wangsirisombat, 2011). Huang (2001) pinpointed that language interference could occur, but the interference of the native language is exacerbated when learners attempt to translate collocations, especially those with the native language that differs in the forms or patterns of native speakers. When an ill-formed word combination is generated, it might result from limited collocational knowledge/competence in the target language (Boonraksa & Naisena, 2022; Yumanee & Phoocharoensil, 2013). While Zughoul (1991) pointed out that learners use collocations incorrectly since their exposure to the language is insufficient, Duong and Nguyen (2021) believed that collocational errors might arise if learners have inappropriate knowledge of collocations. As collocational errors are concerned, it is necessary for learners to possess knowledge of collocations and for teachers to include collocation instruction in any EFL course (Darvishi, 2011).

Grammatical collocation learning

Grammatical collocations comprise a dominant word, namely a noun, adjective, or verb, accompanied by a preposition or grammatical structure. The typical patterns of the collocations include noun + preposition/to infinitive/that clause, adjective + preposition/to infinitive/that clause, verb + preposition, and preposition + noun (Benson et al., 1997). To learn collocations, students might employ different learning strategies. Traditionally, the open-choice principle was adopted when constructing collocations (Farghal & Obiedat, 1995). Nonetheless, Sinclair (1991) argued that this principle should not be applied because free combinations of every word are not allowed in the case of collocations. In his view, most language production and comprehension stem from pre-assembled linguistic chunks. This concept is recognized as the idiom principle, which is believed that a portion of language can be fixed or semi-fixed expressions observable in a corpus. These expressions carry meanings that are linked to the entirety of the expressions. Sinclair's idiom principle stands in opposition to the open-choice principle, where language is believed to be processed by selecting individual words and combining them together according to grammar rules. In addition, students tend to use literal translation strategy when they are unable to identify the correct collocations to put in use. Some may choose to do the translation work in a way that their thought is transferred word-for-word (Wangsirisombat, 2011), making students adopt an inappropriate way of using and learning collocations regardless of acceptable (e.g., busy with work, interested in your article) or unacceptable (e.g., married with his friend, decision for this matter) collocations. Another common strategy is synonymy. Several students often try to replace a synonym for a certain word in the target language, assuming that the collocations are faultless or useable (see Phoocharoensil, 2011). Based on the aforementioned studies, these strategies are neither suitable nor successful for collocation learning. Therefore, it is imperative for language learners to receive adequate guidance or effective learning approach.

Using corpora in learning grammatical collocations

In recent years, with the accessibility of several language corpora and the rapid development of computational technology, corpora have had a significant impact on education (Pérez-Paredes, 2020). In corpus linguistics, collocation denotes the notion that key aspects of word meaning are not included within the word directly but rather the association it has with other words that it regularly co-occurs with. By thoroughly observing a word usage, it frequently has a preference with other words nearby, before or next to it. Language learners should learn the word by its collocates in an authentic context (McEnery & Hardie, 2011). As stated earlier, the occurrence of collocations can be unexplainable as there is no discernible rationale for the grouping of words. Using corpora can be effective to learn collocations because it enables learners to gain exposure to countless naturally occurring examples of the collocations across contexts.

In language classrooms, corpora have been found to be used either indirectly or directly by students. For indirect application, language teachers use the target corpus to help in making a pedagogical decision of what to teach and/or develop their teaching materials (Römer, 2011), while direct application means that students are trained to use the corpus tool to figure out

meaning, lexical, and grammar rules by themselves. If the concordances do not provide sufficient information for them to interpret, they can obtain extra texts by inputting other keyword(s) to search for additional contextual clues (Yoon & Jo, 2014). According to McEnery and Xiao (2011), the application of corpora was more indirect than direct. Perhaps, it was because the direct use has some limitations, including students' familiarity with corpora (Frankenberg-Garcia, 2012), availability of internet access and computers (Yaemtui & Phoocharoensil, 2018), time constraints due to curriculum pacing (Kaya et al., 2022), and size of target corpus or the access to it (O'Keeffe & McCarthy, 2022).

Though direct corpus use has several limitations that teachers should consider before adopting it, it has various advantages worth employing. The direct application of corpora is commonly known as data-driven learning (DDL), and is commonly referred to as Tim Johns' pioneering work in the 1980s, bringing corpus-based practices into the classroom. Adopting DDL, learners should be approached as language researchers with access to linguistic data driving their own learning, while teachers should act as if they are the directors and coordinators of the learner-initiated linguistic research (Johns, 1986). As remarked by Johns (1991), the answers discovered by learners are better than those simply offered by teachers since they allow learners to learn how to autonomously achieve their language learning goals. Thus, DDL learners are afforded the opportunity to unlimitedly access corpus data, engage in real-life everyday communication, and be exposed to numerous examples of collocation in authentic contexts (Rivera, 2021). Because of its student-centered characteristics (Pérez-Paredes, 2022), this learning approach can foster and maintain learner autonomy during the learning process (Boulton & Cobb, 2017). Furthermore, DDL as an inductive learning method was widely acknowledged to activate learners' higher cognitive skills: hypothesizing and inferencing (Flowerdew, 2015). As Corino and Onesti (2019) noted, what is taught could remain longer in memory if higher-order thinking skills are activated.

Theoretically, the notion of DDL appears to be associated with noticing hypothesis, one of the theories of second language acquisition (SLA) (Flowerdew, 2015). In DDL, the unique feature of corpus tool known as KWIC (Key Word In Context) function can be said to raise learners' attention to a target linguistic item. This is because KWIC concordances feature recurrent patterns, thereby enabling learners to simply recognize such patterns in a specific context (Boontam, 2022; Sriphicharn, 2002). To assist learners to learn the target grammatical collocations, it is vital to provide them the chance to analyze corpus data to notice the hypothesis through the implementation of Johns's DDL procedures (Johns, 1991): identification, classification, and generalization. As asserted by Fauzanz et al. (2022), while learners are engaging in a systematic process of identifying, analyzing prominent patterns, and generating the rules, their analytical skills are also enhanced, enabling them to apply these skills in studying other grammatical rules.

To date, many studies have investigated collocation learning and found the improvement of learners' collocational knowledge. For instance, Li (2017) investigated the use of DDL to develop learners' collocational competence on verb-preposition collocations of sixty Chinese postgraduate students majoring in English. An experiment was carried out to compare two groups of students: one DDL group and one rule-based group. After learning, the DDL group showed greater use

of collocations and a higher rate of the accuracy than the rule-based group. In parallel with these findings, Özbay and Olgun (2017) employed DDL for teaching adjective-preposition collocations to sixty Turkish high school learners. DDL was used in the experimental group, whereas the students in the control group were given a traditional teaching method. The results showed that the DDL group had a high advantage over the control group in learning the collocations.

In addition to the above studies, Saeedakhtar et al. (2020) compared the role of hands-on DDL (students directly access to corpus data) and hands-off DDL (students learn from paper-based concordances) on learning verb-preposition collocations by sixty Iranian pre-intermediate students. The students were allocated into three different groups: hands-on DDL, hands-off DDL, and a control group. The findings from the immediate posttest showed that the hands-on and hands-off DDL groups outperformed the control group. The findings also suggested that both DDL groups improved the ability to use the target collocations. More recently, Satake (2022) conducted a study on the effect of corpus used by fifty-five Japanese intermediate students. The students were divided into the treatment group (corpus users) and the control group (dictionary users). In using verb-noun collocations in essay writing, the results revealed that the corpus users outputted more suitable collocations than dictionary users.

Based on the review of previous studies, the empirical investigations conducted on DDL agree that it is beneficial for collocation learning. It is however important to note that these studies predominantly centered on comparing DDL with traditional or deductive methods. The learning procedures applied by Johns (1991) have been given inadequate attention to be discussed in DDL methodology for assisting learners to notice the target items being learned. It also remains unclear how the knowledge students have processed from DDL can aid them in learning collocations (Satake, 2022). Accordingly, this study is intended to investigate the effectiveness of DDL in learning grammatical collocations of Thai students as well as examine the extent to which they incorporated the collocational knowledge learned through concordances into their writing.

METHODOLOGY

Research design and participants

The research adopted a quasi-experimental design in which both qualitative and quantitative methods were employed. The participants were 18 female and 22 male first-year students aged between 18 and 21 years old in one intact class from a private university in southern Thailand. Through purposive sampling, they were selected since they were non-English major students who studied English as a Foreign Language (EFL) in a Thai program. The chosen course from which the data were collected was an English for Communication course, one objective of which aims at improving and promoting the students' English communicative ability in the written form of communication. Despite students' familiarity with internet usage, none of them experienced language corpora and concordancing prior to the study. As such, the selected corpus used to design the materials and conduct the study was the original British National Corpus (BNC) since it is genre-balanced and contains a large collection of written texts.

Instruments

To collect the data, training handouts, teaching materials, writing tests, a stimulated recall interview, and a semi-structured interview were employed.

1. Training handouts

The two training handouts were designed for two training sessions: paper-based concordance and computer-based concordance. The paper-based concordance handout included a review of English word classes and their functions and a guide for students to learn the target preposition collocations from concordance line printouts, while the computer-based concordance handout was for the students to learn the collocations independently through an online searching concordancer. Each handout contained 18 KWIC-format sentence samples of the collocations retrieved from the corpus along with collocation exercises (multiple-choice questions). All the concordance lines were equally distributed according to the target patterns of the collocations.

2. Teaching materials

Three sets of paper-based DDL and three sets of computer-based DDL were created for target collocation learning. The paper-based DDL was used as a scaffolding source to assist students to learn through DDL before learning independently using computer-based DDL. The target patterns involving content words, namely verb, adjective, and noun plus prepositions were selected since they were discovered to be frequently misused by EFL students (Kampookaew, 2020; Latupeirissa & Sayd, 2019; Qamariah & Wahyuni, 2020; Suvarnamani, 2017; Yusuf et al., 2021). They were also reflected in the researcher's preliminary analysis of collocational errors made by students enrolled in the English for Communication course in the academic year 2021. The list of specific words used as the target patterns was then identified based on the errors found.

Both paper- and computer-based materials had two parts. The first part dealt with an explanation of the target collocations, while the second part contained collocation exercises with different tasks, namely gap-filling, error correction, and sentence building. Nonetheless, it should be noted that the distinction between the two materials was that the paper-based DDL contained 18 concordance lines selected from the corpus and adjusted by the researcher to suit the context of learning and the levels of students in the course. For the computer-based DDL, students were not given the sentence samples; instead, they were asked to search for the concordance lines in the corpus by themselves.

3. Pre- and post-writing tests

The two parallel writing tests designed to measure the students' collocational knowledge of prepositions were administered before and after DDL instruction. All students were assigned to correctly match 15 content words covering three target patterns (see Table 1) with 7 prepositions given (i.e., for, to, of, with, in, at, on) and use them along with other grammatical collocations they had known to compose a 150 to 200-word descriptive paragraph prompted by a set of

pictures adopted from Heaton (1975) (see Appendix A). To evaluate their actual ability to accurately use grammatical collocations learned within the context of written communication, the students were not allowed to consult any dictionaries and corpus resources. The allocated time for each test was 60 minutes. For the validity of the tests, the two tests were verified by three experts in the field, and the Index of Item-Objective Congruence (IOC) score of each test was 1.00. To ensure the suitability of the tests and the allocated time, these tests were piloted with 30 EFL first-year students in another foundation English course. The pilot study indicated that the students were able to complete the tests within the designated time frame.

Table 1
The target content words used in the pre- and post-writing tests

• Pre-writing test					
Pattern	**Content Word**				
Verb	ask	arrive	agree	think	help
Adjective	angry	similar	bad	afraid	crowded
Noun	trouble	solution	visit	decision	information
• Post-writing test					
Pattern	**Content Word**				
Verb	belong	remind	prepare	laugh	point
Adjective	tired	allergic	ready	busy	disappointed
Noun	concern	change	effect	problem	difficulty

4. Stimulated recall interview

Upon the completion of the posttest, 5 students who received the highest gain scores and 5 students with the lowest gain scores were chosen as the interviewees. The stimulated recall interview was carried out face-to-face in Thai to elicit the students' reflections on their collocational knowledge learned through concordances and how they incorporated this knowledge in their writing. The completed posttest was served as a stimulus in the interview where students were asked to clarify, elaborate, or explain the respective employed collocations. Each interview session took approximately 25-30 minutes.

5. Semi-structured interview

The interview was also processed face-to-face with the same 10 students selected for the stimulated recall interview. The purpose of the semi-structured interview was to confirm the data on the effectiveness of learning grammatical collocations through the concordances in DDL instruction. The Thai language was used to allow the students to effectively articulate their thoughts and obtain the detailed information as much as possible.

Data collection procedure

The study was carried out over a period of 10 weeks for one semester. The class which met twice a week for four hours was taught by the researcher. The last one hour was weekly allocated for DDL. In total, the research procedure involved three stages: pre-experimental stage, experimental stage, and post-experimental stage.

In the pre-experimental stage, all participants were asked to write a pre-writing test with picture prompts for an hour without consulting any dictionaries. After taking the test, one-hour training of paper-based concordance was provided to prepare students for DDL by introducing what a concordance is and explaining how it can be used to learn the grammatical collocations. As students were required to pattern the collocations in reaching the generalization stage of DDL, this training was to also review word classes in English to assure that all students had adequate background knowledge to identify each word class and its function.

The experimental stage consisted of three phases: paper-based concordance instruction, computer-based concordance training, and computer-based concordance instruction. During the first three hours, the three target patterns of the collocations were instructed using paper-based DDL. Then the computer-based concordance training was provided for one hour to have all learners practice using a concordancer before autonomously using it to learn the three target patterns in the next three hours through computer-based DDL instruction. Throughout the course, all learners were requested to work in pairs and switch their partners after learning each pattern because changing partners could allow them to collaboratively exchange new ideas and obtain different viewpoints (Dankittikul & Laohawiriyanon, 2018). While students were working with DDL exercises, the teacher frequently intervened during the paper-based phase; however, teacher intervention in the computer-based phase occurred only when students were likely to struggle with the exercise. After each exercise was completed, the answers with explanations were given to ensure that all students perceived the correct generalized patterns.

In the post-experimental stage, the students were tested to measure their collocational knowledge after learning through DDL. Subsequently, 10 selected students were interviewed and their responses were audio-recorded. Principally, as the stimulated recall protocol should occur as soon as possible after the incident so that the students could retrieve their memories promptly (Fox-Turnbull, 2009), the interviews took place within two days after the post-writing test was completed. The interviews were conducted individually with each student to ensure that the interviewees felt comfortable sharing their thoughts without any concern of being compared to others.

Data analysis

The pre- and post-writing tests were analyzed to detect the grammatical collocation errors in students' use of the three target collocation patterns. To evaluate how students correctly used grammatical features of the language, the frequency of the accurate use of each pattern was identified using Pica's obligatory occasion analysis (Pica, 1983), where the number of the student's correct uses of the target pattern is regarded as the number of correct suppliance in context. The obligatory occasion is the number of chances in which the target pattern is necessary to be used to construct a grammatically accurate sentence. The suppliance in non-obligatory contexts is considered the number of overuses of the pattern in the sentence. The following is the formula for the analysis.

$$\frac{\text{n correct suppliance in context}}{\text{n obligatory occasions + n suppliance in non-obligatory contexts}} \quad X\ 100$$

The accuracy scores of the accurate use of the three target patterns were coded and calculated by the researcher and one non-native teacher with 5-year experience of teaching English. Then Cohen's Kappa statistic was used to determine the reliability of the coders. The computed inter-coder reliability for the pre-writing test was 0.86 and the post-writing test was 0.88. After the accuracy scores of all patterns were identified, the statistical differences between the two tests were determined using a paired sample t-test. For the data gathered from the stimulated recall and semi-structured interviews, they were verbatim transcribed and translated into English. Afterwards, the data were qualitatively analyzed by coding and categorizing into themes.

RESULTS

This section reports the results of the study based on the data obtained from the pre- and post-writing tests as well as the stimulated recall and semi-structured interviews.

1. Results from pre- and post-writing tests

The available data indicated a notable improvement in the collocational knowledge of the students who have undergone DDL, as can be seen in Table 2.

Table 2
The accuracy scores of the pre- and post-writing tests

Patterns	Pre-writing test (%)		Post-writing test (%)		t	Sig. (2-tailed)	Cohen's d
	Means	S.D.	Means	S.D.			
Verb-preposition collocation	59.93	29.43	78.31	26.75	3.00*	0.00	0.47
Adjective-preposition collocation	23.59	34.65	80.95	34.41	7.87*	0.00	1.24
Noun-preposition collocation	38.79	39.99	79.58	34.49	4.87*	0.00	0.77
Overall	**40.77**	**34.69**	**79.62**	**31.88**	**7.96***	**0.00**	**1.26**

Note: n = 40, p < .05

Table 2 reveals that students' knowledge of collocations improved significantly in all patterns after learning through DDL. Overall, the accuracy score in the pre-writing test increased from

40.77 (SD = 34.69) to 79.62 (SD = 31.88) in the post-writing test with 38.85 of score difference (t = 7.96, p = 0.00), and a large effect size value (*d* = 1.26) was found. When considering the target collocational patterns individually, the biggest difference was found in adjective-preposition collocation, with a large effect size (*d* = 1.24), followed by noun-preposition collocation with a quite large effect size (*d* = 0.77), and verb-preposition collocation with a quite medium effect size (*d* = 0.47), respectively.

2. Results from the stimulated recall interview

The data collected from the stimulated recall interview with 10 selected students revealed the extent of students' collocational knowledge they acquired from DDL lessons and the ways they incorporated their acquired knowledge in their writing. The key results were interpreted and presented as follows:

2.1 The students' collocational knowledge

2.1.1 The ability to notice collocational patterns

The data obtained showed that the students' focus while writing was the grammatical collocations and the patterns they learned through concordancing. Eight students indicated that they have remembered the content words, namely verbs, adjectives, and nouns that are used with the grammatical words or prepositions as they were the target patterns in the DDL lessons. Most students could simply recall and identify the patterns of target preposition collocations they employed in the post test. They mentioned that multiple practices of identifying collocational patterns helped them induce and notice the collocational patterns, making them certain of the collocations they had to use in the test. Some students, especially those with high gain scores, also observed that all preposition collocations they were asked to write in the test were from the DDL lessons. For instance, one interviewee said:

Excerpt 1:

> *"I mainly focused on the targeted patterns of the collocations, namely verbs, adjectives, and nouns that are used with prepositions. I have remembered that they were from DDL exercises. When I took the test, it was easy for me because I practiced patterning a lot in the class and I have acquired collocational knowledge considerably. So, I know whether or not they are grammatically correct." (Student with high gain scores 2)*

Nonetheless, a few students, particularly those with low gain scores, could not recognize some target words they learned; they reported that they sometimes made a prediction of the collocations based on their prior knowledge of the parts of speech.

2.1.2 The accuracy of the grammatical collocations

Concerning the accuracy of students' knowledge of collocations, seven students were able to explain the correct usage of the collocations they used in the posttest. They could also elaborate on what to consider to use the collocations accurately. For example,

Excerpt 2:

> *"Depending on the object or the meaning I want to convey, the same verb, adjective, or even noun can be followed by different prepositions but not all prepositions. For example, the prepositions 'for' and 'to' collocate with the word 'ready'. So, the collocation 'ready for' indicates the readiness of someone for something that will happen in the future, while the collocation 'ready to' indicates the readiness of their action to do something." (Student with high gain scores 3)*

The student in Excerpt 3 explained the adjective 'ready' that can be used with the prepositions 'for' and 'to'; he could also explain the use of the patterns 'ready for + noun' and 'ready to + verb'.

Excerpt 3:

> *"For the adjective 'ready' that can be used with the prepositions 'for' and 'to', 'ready' with 'for' (ready for + noun) is used to point out that someone is prepared for a certain event that will be happening, and if 'ready' with 'to' (ready to + verb) is used, it implies that someone would like to indicate that he is prepared to perform a certain action." (Student with low gain scores 1)*

2.1.3 The variety of the grammatical collocations

Further analysis of the data showed the variety of grammatical collocations students employed in their post-writing test after learning through DDL. That is, both high and low gain score students reported that they utilized the collocations learned in DDL exercises and used other preposition collocations discovered while working with the concordancer. For example,

Excerpt 4:

> *"Although I have learned only three patterns in the exercises, I believe I have known a lot of collocations to use when I have to write in English, comparing with the pretest which I knew only some basic verbs + prepositions." (Student with high gain scores 4)*

Of the interest here is when closely analyzing students' writing tests, the researcher found that they obviously used more collocations in the posttest than the pretest, as presented in Table 3.

Table 3
The variety of the correct use of collocations in the pre- and post-writing tests

Patterns	Pre-writing test	Post-writing test
Verb + preposition	go to/for/with, think of, arrive in/at, ask for, suggest to, start to, begin to, plan to, seem to, agree with, look at/for, promise to, listen to, agree to, talk to, decide to, walk on, search for, bring to, travel in/to, seek for	go to/for, think of, arrive in/at/on, ask for, wait for, belong to, remind of, prepare for/to, laugh at, point at, succeed in, help with/in, look for/at, deal with, laugh with, discuss with, participate in, apologize to, argue with, lead to, agree with/to, response to, compare for/with, result in, dream of, believe in, *cover in, continue to, step on, assist in*
Adjective + preposition	afraid of, angry with/at, similar to, impressed with, embarrassed to	afraid of, angry with/at, similar to, tired of, allergic to, ready for/to, busy with/to, disappointed in/with, proud of, good at/for, satisfied with, full of, skillful at, bad at, good at, responsible for, suitable for, comfortable in, capable of, *careful with, grateful for, mad at, delighted with*
Noun + preposition	information on, solution to, trouble with	concern in/with, change in, effect on, impact on, problem of, difficulty in/with, benefit of, solution to, argument with, lack of, responsibility for, success in, difference in, intention of, smell of, decision on, *dedication to, point of*

Note: The italicized collocations are not in the target collocations.

Surprisingly, as Table 3 indicates, several students stated that there were some new grammatical collocations they discovered in concordances and they have also used those collocations in their writing. One of the interviewees commented,

Excerpt 5:

> *"There are other new collocations such as continue to, delighted with, dedication to that I discovered in computer-based DDL. As the instruction of the test says that I should write more, I therefore add them in my writing too." (Student with high gain scores 1)*

Another compelling finding is that some students employed the same content word with different prepositions that were not in the target lists (e.g., arrive on, angry with, concern in). To elaborate, one student mentioned that while she was concordancing, she found that 'arrive' with the preposition 'on' was not in the DDL exercises. Since she knows that it is one of the grammatical collocations, she used it in the posttest to indicate the arrival date of someone.

2.2 How the students incorporated collocational knowledge in writing

Almost all students addressed that they incorporated the grammatical collocations in their writing based on the correct usage and the meaning in use. Mainly, students thought of the usage of a certain collocation and its function before and while using it. When asked about

verb-preposition collocation, students were unanimous in the view that a verb in combination with a preposition was used to delineate a particular action of the subject. They further added that their selection of the preposition depends on the preceding verb, and the verb does not function properly without the preposition or with different preposition. Likewise, adjective-preposition collocation was used to describe or modify the subject of the sentence (e.g., noun and pronoun). In the case of noun-preposition collocation, students used a noun in combination with a preposition to refer it to a thing or an idea in each sentence.

Regarding the meaning in use, the student responses indicated that the collocations they employed were appropriate for communicating their intended message. They stated that the meaning they wished to convey plays a crucial part in selecting a certain collocation. Interestingly, some students reported that they also considered the collocates or the surrounding words of the collocations. For instance, one student recalled the use of the adjective-preposition collocations: 'ready to' and 'ready for'. The student knows how to use these two collocations after learning through DDL, and what she remembered is that they are ready for something (ready for + noun) to happen, and ready to do something (ready to + infinitive verb). Also, she noted that 'ready to' is more common than 'ready for', according to her corpus search result.

3. Results from the semi-structured interview

The data retrieved from the semi-structured interview revealed what students thought of learning grammatical collocations through DDL and its influence on their writing improvement.

3.1 Views on learning grammatical collocations through DDL

Evidenced from the interview data, the majority of the students viewed learning grammatical collocations through DDL as an effective method to learn the target collocations and several others that were not in the target lists. The DDL learning steps could lead them to become aware of the collocations' patterns and their uses in different contexts. As the students remarked,

Excerpt 6:

> "Collocations can be confusing because there are no specific rules or explanations, and the list appears to be infinite. So, looking at various examples demonstrating how they are used is the effective way to learn. In DDL, particularly the step of generalization, I eventually noticed and became aware of the target patterns as well as their applications such as in academic and non-academic texts." (Student with low gain scores 2)

Excerpt 7:

> "I believe learning the new collocations from many examples offered by the corpus is effective, especially the collocations used in different situations. And the concordance lines in the KWIC make it simpler for me to analyze, generalize, and notice the correct patterns of collocations I am trying to learn." (Student with high gain scores 4)

Also seen in Excerpt 7, the student reported that concordance lines could help her learn the collocational patterns. This point was also elaborated by another student who explained that they were more helpful than learning patterns through a dictionary. Various concordance lines could help her develop a deeper comprehension of collocational patterns and their uses. As she commented,

Excerpt 8:

> *"Initially, I used an online dictionary to search for some verb-preposition collocations. Despite the fact that the meaning was shown, a few examples were given. In DDL, I was able to confirm what I thought by analyzing the concordance lines until I really noticed the collocational patterns and gain deeper understanding of the collocations during the classification step of DDL. As I carefully read more concordance lines, I become more acquainted with the usage of the collocations." (Student with high gain scores 5)*

3.2 Views on the influence of collocations on students' writing improvement

During the interview, students were asked to explain how they think the grammatical collocations influence their writing improvement. Overall, students' responses revealed their awareness of collocations, which leads to their correct use of collocations in their writing. As explained by a student,

Excerpt 9:

> *"In using 'afraid of', I had used 'I afraid cockroach' without 'am' and 'of' countless times. I didn't know my sentence was totally grammatically wrong. After learning through DDL, I remember that if I write afraid, I have to write 'am' because 'afraid' is an adjective along with 'of' because it is the collocate of afraid." (Student with high gain scores 3)*

Furthermore, students stressed that learning collocations leads to their vocabulary expansion and their writing improvement as they can write more accurate sentences. They commented that,

Excerpt 10:

> *"The collocations enlarge my vocabulary bank. In addition to this, knowing the words in pairs helps me to constitute a sentence more accurately. Sometimes, the content word used with an incorrect preposition doesn't only make the sentence odd, but it also results in misunderstanding of the message. For example, I can never use the adjective-preposition collocation 'good for' if I want to say that I am 'good at' something." (Student with high gain scores 1)*

Excerpt 11:

> "I think the collocations influence my writing improvement since I have observed that
> I have made fewer collocational errors. If I practice more and I have flawless use of the
> grammatical features, I may have adequate collocational competence to achieve higher
> level of language proficiency or near native fluency." (Student with high gain scores 5)

DISCUSSION

The overall result obtained from the writing tests indicates that after utilizing the corpus with data-driven learning method, students' collocational knowledge significantly improved. This result confirms the previous studies which found students' substantial improvement in collocational knowledge in the test scores after they participated in DDL (Vyatkina, 2016b; Yaemtui & Phoocharoensil, 2018). When comparing the effect size of all patterns, it was discovered that the greatest practical improvement occurred with adjective-preposition collocation, followed by noun-preposition collocation, and verb-preposition collocation. This may be explained by the fact that English nouns and verbs have various inflectional markers in their forms. As we inflect many nouns by putting -s or -es after them to make the plural forms, we add -s or -es to various verbs if the subjects are singular in the present simple tense (and other different forms to express English tenses). There are several words that can be both verbs and nouns, such as change, smell, and visit, making the identification of nouns and verbs difficult. Nonetheless, Nesselhauf (2003) asserted that both verb and noun collocations resulted from a strong interference from learners' first language – usually because of the incorrect choice of verb and noun along with the production of an incorrect prepositional combination. For instance, by comparing English and Thai, waiting for someone/something is expressed by the verb 'wait' and the preposition 'for' in English, whereas it does not necessitate the inclusion of any preposition in Thai. Regarding these reasons, Thai students probably were better at adjective-preposition collocation than noun- and verb-preposition collocations since the adjective is more straightforward and it does not inflect in its form when used as an adjective-preposition collocation.

Also revealed from the stimulated recall interview and semi-structured interview, students learned knowledge of collocations through DDL, and to some extent they could also apply it in their writing. It was found that most students could accurately i) classify the types of the target collocational patterns, ii) identify the hidden usage of the content words with different prepositions, and iii) explain what should be considered when employing particular collocations in writing, namely the object of a sentence and the meaning of a collocation, as well as selecting contextually appropriate collocations for their written communication. Surprisingly, their explanations also suggest they have acquired collocations in addition to the targeted ones in DDL lessons. These results seemed to point out that DDL is suitable for learning grammatical collocations and their uses, and the learning procedures could lead to students' acquisition of the collocations. Students' acquiring ability could result from students' noticing ability developed through Johns's (1991) DDL learning procedures. As they had to induce the regularity of linguistic patterns appeared in the concordance lines, they had to notice the salient features of the

target collocations. With this inductive strategy, learners were urged to dynamically engage in complex cognitive processes, namely inferencing and hypothesizing (Flowerdew, 2015) and other cognitive skills necessary for language learning, including noticing, exploring, analyzing, interpreting, differentiating, and verifying (Alshammari, 2019; O'Sullivan, 2007). Taken together, practicing regularly under the learning procedures together with being exposed to multiple contextual examples of the collocations, learners were able to learn the target collocations, particularly the novel ones.

The students' ability to notice can be also explained by Schmidt's noticing hypothesis (1990). According to Vyatkina (2016a), noticing can be facilitated by repeated exposure to large corpus data or the target feature (input enrichment) (Trahey & White, 1993), and emphasized target feature in the corpus (input enhancement) (Sharwood Smith, 1993). As Vyatkina (2016a) addressed that "rich and enhanced input (e.g., concordance lines) serves as material for learners' noticing and analysis of language patterns" (p. 208), the KWIC concordance lines accordingly have the ability to highlight linguistic features that can attract learners' attention to the target language item being studied. When working with concordance lines presented in KWIC format, learners can easily notice the correct collocational patterns, raising their awareness and facilitating their recognition of co-occurring patterns of the language item more easily (Sriphicharn, 2002). Learners can also apply what Ellis (2003) called "discovery learning" principle as they are likely to remember what they have discovered on their own better than those that they have been told. Furthermore, by analyzing multiple concordance lines, students could deepen their understanding of collocational patterns, though they may need more time to understand collocations than using a dictionary (Satake, 2022).

According to the noticing hypothesis viewpoint, student's ability to notice the target language features is an essential starting point towards accurately incorporating the knowledge in writing. The language features should be consciously noticed in the input for them to be acquired. However, the acquired knowledge can be supplied by different factors, including individual processing ability, readiness to notice, perceptual salience of linguistic features, frequent practice, and effective instructional strategies (Schmidt, 1990). From this study, although almost all students demonstrated their ability to notice collocation patterns, it was found that a few students with low gain scores were unable to do so, possibly attributable to their limited knowledge of English or individual learning style. As Hughes (2010) asserted, it is plausible for low-proficiency learners to encounter difficulty in comprehending unfamiliar words in a corpus. In addition, if students are passive learners or are used to deductive learning, grammar translation, or teacher-directed instruction, it might make them feel demotivated or uncomfortable to learn through inductive DDL. Indeed, regardless of the proficiency level, this pedagogical approach can pose challenges for those habituated to traditional classroom settings (Boulton, 2010; Dankittikul & Laohawiriyanon, 2018). As such, it is likely that DDL may be more advantageous for active learners than passive learners (Flowerdew, 2012).

Considering the qualitative results, it also indicated how students incorporated the acquired collocational knowledge in their writing. When students had to incorporate the knowledge in writing, they considered the correct usage and the meaning in use when planning on how to employ particular collocations before starting to write to produce the correct and meaningful

use of the collocations. The findings suggest that students were aware of grammatical accuracy of collocational patterns – particularly the correct forms of the collocations. Additionally, when it comes to writing, the students realized that the accurate application of a collocation is important in communicating their intended message. In fact, Firth (1935, as cited in O'Keeffe et al., 2007) stressed that "the meaning of a word is as much a matter of how it combines with other words in actual use (i.e., its collocations) as it is of the meaning it possesses in itself" (p. 59). Since the meaning stems from the structure of the language, the readers can attain the meaningful language through the correct formation of its segmental structures, e.g., words, collocations, phrases, and sentences (Ibrahim & Yunus, 2018).

The results reported also shed light on how students think the grammatical collocations learned influence their writing improvement. The students' insights showed that they valued collocations as a mechanism assisting them to become native-like writers of English as they could expand their vocabulary bank and could write more straightforward and accurate sentences. This yields an interesting finding that despite being novices of English and non-English major students, they could highlight the necessities of using and mastering collocations in writing. Practically, the incorporation of collocations enriches students' lexical repertoire and facilitates the accurate and clear expression of their thoughts and written ideas (McCarthy & O'Dell, 2017; Saito, 2020). These elements serve the need for achieving a level of language proficiency akin to that of native speakers (Wood, 2002).

CONCLUSION

The present study was set out to investigate the effectiveness of DDL in learning grammatical collocations in the EFL classroom and examine the extent to which students incorporated the knowledge learned through DDL in their writing. With reference to both quantitative and qualitative results of the current study, it could consequently draw a conclusion that DDL implementation is an effective learning method as it improved students' collocational knowledge and supported their acquisition of the collocations. To some extent, students were able to apply the knowledge of collocations in their writing. The students' accurate use of the collocations can increase the language accuracy of their written communication.

DDL, however, also has certain limitations. Some obstructions, such as individual learning style and insufficient knowledge of English, might hinder students' maximally successful learning. It is suggested DDL teachers conduct the lesson based on students' needs and analytical abilities. The teachers may deliberate either to adopt a guided DDL, where the teachers play the role of assistant guiding the learners through each sequential step of learning, or an autonomous DDL, where learners work through the DDL steps on their own without direct assistance from the teachers. A corpus may contain incorrect language information (Jaihow, 2018). Therefore, after each DDL exercise is completed, it is important to provide an explanation to enable students to determine whether the patterns they have generated were correct as well as to ensure that the target collocational patterns were consciously noticed.

Although the effectiveness of DDL in learning grammatical collocations has been confirmed in

this current study, it is imperative to acknowledge a number of limitations. First, the study was carried out in an intact class where the students' proficiency levels varied. Often, classrooms can contain learners with varying English proficiency levels; nevertheless, if possible, they should be classified into their respective proficiency groups so that the teacher can design exercises that best suit each group of students. While using the same exercises for varying proficiency levels may present challenges for some students, with careful planning and attention to fairness, it is recommended to provide clear instructions and expectations. That is, the teacher needs to clearly communicate expectations and instructions for each exercise, so students know what is expected of them. This can also help students approach the exercises with the appropriate mindset. Another limitation is the fact that the participants of this study were limited to non-English major undergraduate students. Accordingly, the findings may not be applicable to other settings or different levels of education.

In addition, to promote DDL learners as much as possible, it would be beneficial for future research to investigate how much influence each different factor (e.g., language proficiency, learning style, processing ability, readiness to notice) has on learners' capacity to notice the target language items. For this study, the interviews were administered to collect data on collocation knowledge reflection; other instruments, such as learner diaries, teacher observation, and think-aloud protocol, could be used to elicit the self- or prompt-reflection of learners. Moreover, the results of the study were constrained as they only exclusively indicated learners' immediate application of the acquired knowledge in their writing after the experimental stage. Accordingly, a longitudinal study should be carried out to examine students' knowledge retention or the long-term effects of DDL in learning grammatical collocations.

THE AUTHORS

Pimnada Khemkullanat is an assistant lecturer at Hatyai University. She has been studying for a master's degree in Applied English Language Studies at the Faculty of Liberal Arts, Prince of Songkla University, Thailand. Her areas of interest include applied corpus linguistics and English language teaching (student engagement).
pimnadakhem15@gmail.com

Somruedee Khongput is an assistant professor at the Faculty of Liberal Arts, Prince of Songkla University, Thailand. She received her PhD in Education from The University of New South Wales, Australia. Her research interests include teacher education, and writing instruction and assessment.
somruedee.kh@psu.ac.th

REFERENCES

Alsakran, R. A. (2011). *The productive and receptive knowledge of collocations by advanced Arabic-speaking ESL/ EFL learners* [Master's thesis, Colorado State University]. Mountain Scholar. https://dspace.library.colostate. edu/handle/10217/48104

Alshammari, S. R. (2019). Data driven learning and teaching of prepositions in ESL: A study of Arab learners. *International Journal of Educational Research and Innovation, 12*, 153-168. https://doi.org/10.46661/ ijeri.3809

Benson, M. (1989). The structure of the collocational dictionary. *International Journal of Lexicography, 2*(1), 1-14.

Benson, M., Benson, E., & Ilson, R. (1997). *The BBI dictionary of English word combinations*. John Benjamins Publishing.

Boonraksa, T., & Naisena, S. (2022). A study on English collocation errors of Thai EFL students. *English Language Teaching, 15*(1), 164-177.

Boontam, P. (2022). The effect of teaching English synonyms through data-driven learning (DDL) on Thai EFL students' vocabulary learning. *Shanlax International Journal of Education, 10*(2), 80-91. https://doi.org/10.34293/education.v10i2.4458

Boonyarattanasoontorn, P. (2017). An investigation of Thai students' English language writing difficulties and their use of writing strategies. *Journal of Advanced Research in Social Sciences and Humanities, 2*(2), 111-118.

Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*(3), 534-572.

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning, 67*(2), 348-393.

Bridle, M. (2019). Learner use of a corpus as a reference tool in error correction: Factors influencing consultation and success. *Journal of English for Academic Purposes, 37*, 52-69.

Chintaradeja, P. (2020). An error analysis in paragraph writing in academic writing class of Thai undergraduate students, Rajamangala University of Technology. *Journal of Education, Prince of Songkla University, Pattani Campus, 31*(2), 64-76.

Corino, E., & Onesti, C. (2019). Data-driven learning: A scaffolding methodology for CLIL and LSP teaching and learning. *Frontiers in Education, 4*, 1-12. https://doi.org/10.3389/feduc.2019.00007

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

Dankittikul, T., & Laohawiriyanon, C. (2018). Effect of paper-based concordance on Thai low proficiency English language learners' logical connector knowledge. *Veridian E-Journal, Silpakorn University, 11*(4), 61-78.

Darvishi, S. (2011). The investigation of collocational errors in university students' writing majoring in English. *International Conference on Education, Research and Innovation, IACSIT Press, Singapore, 18*, 52-56.

Dokchandra, D. (2019). Thai EFL learners' collocational competence and their perceptions of collocational difficulty. *Theory and Practice in Language Studies, 9*(7), 776-784.

Duan, M., & Qin, X. (2012). Collocation in English teaching and learning. *Theory & Practice in Language Studies, 2*(9), 1890-1894.

Duong, D. T. H., & Nguyen, N. D. T. (2021). Using collocations to enhance academic writing: A survey study at Van Lang University. *Proceedings of the 17th International Conference of the Asia Association of Computer-Assisted Language Learning, 533*, 275-287. https://doi.org/10.2991/assehr.k.210226.035

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Fang, L., Ma, Q., & Yan, J. (2021). The effectiveness of corpus-based training on collocation use in L2 writing for Chinese senior secondary school students. *Journal of China Computer-Assisted Language Learning, 1*(1), 80-109.

Farghal, M., & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral.1995.33.4.315

Fauzanz, A., Basthomi, Y., & Ivone, F. M. (2022). Effects of using online corpus and online dictionary as data-driven learning on students' grammar mastery. *LEARN Journal: Language Education and Acquisition Research Network, 15*(2), 679-704.

Firth, J. R. (1935). The technique of semantics. *Transactions of the Philological Society, 34*(1), 36-72.

Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.

Flowerdew, L. (2015). Data-driven learning and language learning theories. *Multiple Affordances of Language Corpora for Data-driven Learning, 69*, 15-36.

Fox-Turnbull, W. (2009). Stimulated recall using autophotography-A method for investigating technology education. In A. Bekker, I. Mottier & M. J. de Vries (Eds.), *Proceedings of the PATT-22 Conference: Strengthening the position of technology education in the curriculum* (pp. 204-217). International Technology and Engineering Educators Association.

Frankenberg-Garcia, A. (2012). Raising teachers' awareness of corpora. *Language Teaching, 45*(4), 475-489.

Friginal, E. (2018). *Corpus linguistics for English teachers: Tools, online resources, and classroom activities*. Routledge.

Heaton, J. B. (1975). *Beginning composition through pictures*. Longman.

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 47-69). Oxford University Press.

Huang, L.-S. (2001). Knowledge of English collocations: An analysis of Taiwanese EFL learners. *Proceedings of the Texas Foreign Language Education Conference, 6*(1), Article ED465288.

Hughes, R. (2010). What a corpus tells us about grammar teaching materials. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 401-412). Routledge.

Ibrahim, B., & Yunus, K. (2018). Perspectives on corpus linguistics: The methodological synergy in second language pedagogy and research. *Arab World English Journal, 9*(3), 84-97.

Jaihow, P. (2018). *Corpus use by student writers: Error correction by Thai learners of English* [Doctoral dissertation, University of Birmingham]. E-theses Repository. https://etheses.bham.ac.uk//id/eprint/8274/2/Jaihow18PhD.pdf

Jantarabang, N., & Tachom, K. (2021). An investigation of students' experiences with corpus-based approach on high school students' English paragraph writing. *Journal of Saeng Khom Kham Buddhist Studies, 6*(1), 1-15.

Johns, T. (1986). Micro-concord: A language learner's research tool. *System, 14*(2), 151-162.

Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning materials. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 1-16). Birmingham.

Kamarudin, R., Abdullah, S., & Aziz, R. A. (2020). Examining ESL learners' knowledge of collocations. *International Journal of Applied Linguistics and English Literature, 9*(1), 1-6.

Kampookaew, P. (2020). An analysis of grammatical errors made by Thai EFL university students in an EAP writing class: Issues and recommendations. *rEFLections, 27*(2), 246-273.

Kaya, Ö. F., Uzun, K., & Cangır, H. (2022). Using corpora for language teaching and assessment in L2 writing: A narrative review. *Focus on ELT Journal, 4*(3), 46-62.

Latupeirissa, D. S., & Sayd, A. I. (2019). Grammatical errors of writing in EFL class. *International Journal of Linguistics, Literature and Culture, 5*(2), 1-12.

Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology, 21*(3), 153–171.

Lin, M. H. (2021). Effects of data-driven learning on college students of different grammar proficiencies: A preliminary empirical assessment in EFL classes. *Sage Open, 11*(3), 1-15.

Lindquist, H. (2018). *Corpus linguistics and the description of English*. Edinburgh University Press.

McCarthy, M., & O'Dell, F. (2017). *English collocations in use: Intermediate* (2nd ed.). Cambridge University Press.

McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

McEnery, T., & Xiao, R. (2011). What corpora can offer in language teaching and learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 364-380). Routledge.

Men, H. (2017). *Vocabulary increase and collocation learning: A corpus-based cross-sectional study of Chinese learners of English*. Springer.

Muftah, M. (2023). Data-driven learning (DDL) activities: Do they truly promote EFL students' writing skills development? *Education and Information Technologies*, 1-27. https://doi.org/10.1007/s10639-023-11620-z

Nagy, T. (2020). On the nature of collocations and ways of teaching them in foreign language classes. *Hungarian Educational Research Journal, 10*(1), 16-31.

Nation, P. (2005). *Handbook of practical second language teaching and learning*. Routledge.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied linguistics, 24*(2), 223-242.

O'Keeffe, A., & McCarthy, M. (2022). *The routledge handbook of corpus linguistics* (2nd ed.). Routledge.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

O'Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL, 19*(3), 269-286.

Özbay, A., & Olgun, O. (2017). The application of DDL for teaching preposition collocations to Turkish EFL learners. *The International Journal of Research in Teacher Education, 8*(3), 1-10.

Pérez-Paredes, P. (2020). *Corpus linguistics for education: A guide for research*. Routledge.

Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning, 35*(1-2), 36-61.

Phoocharoensil, S. (2011). Collocational errors in EFL learners' interlanguage. *Journal of Education and Practice, 2*(3), 103-120.

Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition, 6*(1), 69-78.

Preradovic, N. M., Posavec, K., & Unic, D. (2019). Corpus-supported foreign language teaching of less commonly taught languages. *International Journal of Instruction, 12*(4), 335-352.

Qamariah, H., & Wahyuni, S. (2020). An analysis of students' grammatical errors in writing English text in the second grade students of SMTI Banda Aceh. *Getsempena English Education Journal, 7*(1), 58-71.

Rivera, J. G. G. (2021). *A corpus-based approach to teach verb+ noun collocations to intermediate level English language learners*. University of Tsukuba.

Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics, 31*, 205-225.

Saeedakhtar, A., Bagerin, M., & Abdi, R. (2020). The effect of hands-on and hands-off data-driven learning on low-intermediate learners' verb-preposition collocations. *System, 91*, 1-14.

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.

Satake, Y. (2022). The effects of corpus use on L2 collocation learning. *Japan Association for Language Teaching Computer Assisted Language Learning Journal, 18*(1), 34-53.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Seensangworn, P. (2017). Writing problems and writing strategies of English major and non-English major students in a Thai university. *Manutsat Paritat: Journal of Humanities, 39*(1), 113-136.

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

Smith, M. S. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*(2), 165-179.

Sriphicharn, P. (2002). Introducing data-driven learning: Theory, practice, and evaluation. *Journal of Liberal Arts, Thammasat University, 2*(2), 100-111.

Sun, X., & Hu, G. (2023). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research, 27*(3), 660-688.

Suvarnamani, S. (2017). A study of grammatical and lexical errors in descriptive writing of first year arts students at Silpakorn University. *Humanities, Arts and Social Sciences Studies*, *17*(2), 239-264.

Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition, 15*(2), 181-204.

Vyatkina, N. (2016a). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL, 28*(2), 207-226.

Vyatkina, N. (2016b). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology, 20*(3), 159–179.

Wallace, M. J. (1982). *Teaching vocabulary*. Heinemann Educational Books.

Wang, Y., & Shaw, P. (2008). Transfer and universality: Collocation use in advanced Chinese and Swedish learner English. *ICAME Journal, 32*(1), 201-232.

Wangsirisombat, N. (2011). *A study of collocational errors in Thai EFL learners' writing* [Master's thesis, Thammasat University]. Thammasat University Digital Collections. https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:119896

Wood, D. (2002). Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal, 20*(1), 1-15.

Yaemtui, W., & Phoocharoensil, S. (2018). *The effectiveness of data-driven learning (DDL) on teaching English collocations to Thai EFL students* [Doctoral dissertation, Thammasat University]. Thammasat University Digital Collections. https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:142211

Yoon, H., & Jo, J. (2014). Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning & Technology, 18*(1), 96-117.

Yumanee, C., & Phoocharoensil, S. (2013). Analysis of collocational errors of Thai EFL students. *LEARN Journal: Language Education and Acquisition Research Network, 6*(1), 88-98.

Yusuf, Y. Q., Mustafa, F., & Iqbal, R. M. (2021). An inquiry into grammatical errors in writing committed by high achieving EFL students. *International Journal of Language Studies, 15*(2), 1-22.

Zughoul, M. R. (1991). Lexical choice: Towards writing problematic word lists. *International Review of Applied Linguistics, 29*(1), 45-60.

APPENDIX A

Writing tests

Pre-writing test

Directions: Match all content words in box A with the correct prepositions in box B (Each preposition can be used more than once). Then use them to describe the pictures below. Other grammatical collocations should also be used (150 – 200 words).

Box A

ask	angry	trouble
arrive	similar	solution
agree	bad	visit
think	afraid	decision
help	crowded	information

Box B

for	to	of	with	in	at	on

A trip to an unfamiliar place

Last year, Jennie and James went on a trip to a place where they had never been before.
....
....
....
....
....

Post-writing test

Directions: Match all content words in box A with the correct prepositions in box B (Each preposition can be used more than once). Then use them to describe the pictures below. Other grammatical collocations should also be used (150 – 200 words).

Box A

belong	tired	concern
remind	allergic	change
prepare	ready	effect
laugh	busy	problem
point	disappointed	difficulty

Box B

for	to	of	with	in	at	on

Wet paint in the new house

Yesterday, Peter's mother entered the wet living room while he and his father were painting.
....
....
....
....
....
....
....
....
....