Data-Driven Learning (DDL) and Its Effectiveness in Teaching Preposition Collocations to Thai High School Students

Pimnada Khemkullanat[1*], Anchalee Watcharajinda[2] and Arintar Pongsapang[3]

[1,2]Lecturer, Department of English, Didyasarin International College, Hatyai University.
[3]Undergraduate student, Department of English, Didyasarin International College, Hatyai University.
[*]**Corresponding author, E-mail:** pimnada@hu.ac.th

Abstract

This study explores the effectiveness of Data-Driven Learning (DDL) in enhancing preposition collocational knowledge of EFL learners implementing a corpus-based DDL approach, utilizing the Corpus of Contemporary American English (COCA) as a resource for instructional materials. A group of 37 Grade 11 Thai EFL students participated in the study, engaging in a structured DDL instruction over multiple sessions. A pre-test and post-test were administered to assess students' accuracy in collocation use, while a questionnaire was used to examine their attitudes toward learning through DDL. The results indicated a statistically significant improvement in students' collocational accuracy, suggesting that DDL facilitated an understanding and more precise usage of prepositional collocations. Furthermore, most participants expressed positive attitudes toward DDL, finding it engaging and beneficial for learning collocations in context. A few students nonetheless found independent pattern identification through DDL challenging. This study suggests that while DDL is effective for collocation learning, prompting students' prior knowledge of collocation and incorporating guided instruction alongside DDL may optimize learning outcomes and support diverse learning preferences.

Keywords: Data-Driven Learning, Preposition Collocations, Thai EFL Students

Introduction

Effective communication is crucial for English language learners (ELLs) as it enables them to express their messages and share their ideas. Moreover, it helps them avoid misunderstandings while enhancing their ability to articulate their thoughts clearly in both spoken and written forms. To develop effective communication skills, learners must first grasp the core principles of sentence construction. That is, it is essential for them to acquire a functional knowledge of grammar tailored to their individual needs for their accurate and effective communication (Mehmonova, 2022).

Even though teaching grammar components does not necessarily aim to make learners grammarians, it helps them communicate effectively (Weaver, 2018 cited in

Chiekezie, 2021). Understanding grammar indeed means understanding the language (Chiekezie, 2021; Mehmonova, 2022). Even so, in the language classroom, it was still discovered that learners encountered difficulties related to the structures of the language. One of the most recurring issues is the proper use of collocational patterns, specifically those involving preposition combinations. Recently, several studies (e.g., Chuenchaichon, 2022; Qamariah & Wahyuni, 2020; Suraprajit, 2021; Yusuf et al., 2021) have shown that frequent errors include the addition of unnecessary preposition and the omission of the necessary preposition. However, the causes of these errors are multifaceted. A major contributing factor is the influence of learners' first language (L1), where structural disparities between languages can result in incorrect assumptions about English collocations (Hashemi et al., 2012). Another significant source is the insufficient knowledge of collocations (Boonraksa & Naisena, 2022; Thongvitit & Thumawongsa, 2017), often resulting from limited exposure to authentic language use or a lack of explicit teaching (Dokchandra, 2019).

As extensively documented, collocational knowledge is essential for English as a foreign language (EFL) students, as it enhances their communicative competence. Effective use of collocations contributes to more natural and accurate language production, allowing learners to express their ideas with precision and fluency. Since collocations are a core aspect of native-like speech patterns, their appropriate use helps learners sound more natural and intelligible in both spoken and written communication (Tian & Della Amirulloh, 2024). Without sufficient collocational knowledge, EFL learners may struggle with unnatural or incorrect word combinations. Moreover, a strong grasp of collocations allows students to understand phraseological units when their meanings are not from individual components. These fixed expressions often pose a challenge for learners due to their non-literal meanings. In fact, a word in isolation may lack clarity, as words do not occur independently in natural language (Stubbs, 2001). Also, collocations account for a significant portion of daily communication, appearing in both spoken and written discourse, with a large percentage of what people say and hear daily (Hill, 2000). Therefore, students should receive appropriate instruction to support their learning and acquisition of collocational knowledge.

As stated earlier, students continue to face difficulties in learning preposition collocation patterns. To address this challenge, effective teaching methods such as explicit teaching and Data-Driven Learning (DDL) utilizing language corpora have been suggested (He & Xie, 2024; Liu & Gablasova, 2023). Evidently, DDL allows students to analyze actual language data, helping them understand language patterns through exposure to real-life language usage (Muftah, 2023). Unlike traditional methods (e.g., rote memorization, grammar translation, dictionary-based learning) as they do not promote natural language

acquisition or meaningful use, DDL is believed to provide students with the opportunity to actively engage with language input, fostering a deeper and more meaningful learning experience (Chambers, 2010). By analyzing multiple instances of collocations in their authentic usage, learners can develop an automatic use of the collocations they are learning (Yuvayapan & Yükselir, 2021). Research further suggests that presenting collocations in diverse and meaningful contexts helps students identify recurring patterns and apply them accurately in their speech and writing (Lindquist, 2018; Muftah, 2023).

Given the limitations of traditional methods and the growing recognition of DDL's benefits, this study adopts a DDL approach to assist Thai EFL students improve their preposition collocation knowledge. While DDL has been widely implemented in university-level English instruction due to its effectiveness for learners with advanced L2 proficiency, there remains a noticeable gap in research on its application in high school language learning. This gap suggests the need for further exploration of how DDL can be tailored to younger learners or those at lower proficiency levels. Additionally, existing research on data-driven vocabulary learning has primarily focused on word meanings rather than contextual usage, including collocations. While collocations are sometimes mentioned in vocabulary instruction, they are often not the primary focus, highlighting the necessity for further studies in this area (Li et al., 2024). Building on the identified challenges and pedagogical recommendations, this study addresses the following research questions:

1. For Thai high school students, is there a statistically significant difference in pre- and post-test accuracy scores after using DDL to learn preposition collocations?
2. What are the Thai high school students' attitudes and feedback on the use of DDL in preposition collocation learning?

Hypotheses

1. H1: Students who undergo DDL instruction will show a statistically significant improvement in preposition collocation accuracy.
2. H2: Thai high school students will demonstrate a generally cheerful outlook toward the use of Data-Driven Learning (DDL) in learning preposition collocations.

Purposes of the Study

This study aims to investigate the effectiveness of Data-Driven Learning (DDL) in enhancing Thai high school students' knowledge and usage of English preposition collocations, with a focus on verb–preposition and adjective–preposition combinations. Specifically, the purposes of this study are as follows:

1. To examine the effectiveness of DDL in improving students' accuracy in using prepositional collocations in English. This will be assessed by comparing pre-test and post-test scores following a DDL instruction.
2. To explore students' attitudes toward learning preposition collocations through the DDL approach, including their engagement with paper-based concordance lines and their perceived benefits or challenges.

Literature Review

1. Preposition Collocation and its Learning.

Collocations consist of words that tend to appear together in language with regularity and predictability (Lewis et al., 2000). However, Sinclair (1991) presents a different view, suggesting that collocations can also follow unpredictable patterns, making them more context-dependent but less formulaic. While Lewis et al. (2000) emphasizes the structured and systematic nature of collocations, Sinclair (1991) highlights their dynamic and sometimes unexpected combinations. This divergence in perspectives underscores the complexity of collocations, as they can exhibit both fixed and semi fixed patterns. Despite this, grammatical collocations encompass several types. As a subcategory of collocations, prepositional collocations involve the combination of a content word with a function word—most commonly a preposition—or other grammar structures, namely clauses or infinitives (See Benson et al., 1997).

In contemporary English language teaching, there is an increasing emphasis on the importance of using collocations accurately and appropriately (Demir, 2017), as enhances both comprehension and production skills, enabling more precise and natural expression (Tian & Della Amirulloh, 2024). Despite their massive significance, collocations still pose challenges for EFL learners. Research has highlighted persistent issues such as incorrect word pairings: omission of necessary prepositions and addition of unnecessary prepositions (Chuenchaichon, 2022; Qamariah & Wahyuni, 2020; Suraprajit, 2021). Considering this, one of the most prominent sources of these errors is first language interference, where structural differences between L1 and English create incorrect assumptions about collocation patterns (Hashemi et al., 2012). For instance, learners whose native languages do not require prepositions in certain verb collocations may mistakenly omit them in English (e.g., I am waiting my mother instead of I am waiting for my mother). Conversely, learners may also add prepositions unnecessarily due to direct translation from their L1, leading to incorrect assumptions. For example, while English uses good at something, some learners might incorrectly use good in something if their native language follows a different structure. Moreover, limited exposure to authentic language input further exacerbates these issues, as many learners rely on textbooks rather than encountering collocations in

natural spoken or written discourse. It has been found that some teachers predominantly use English textbooks as their primary resource (Boontam & Phoocharoensil, 2016), even though these materials occasionally contain language that is not actually used in real communication (Mindt, 1996, cited in McEnery & Xiao, 2011). Taking these issues into account, these teaching and learning methods are not entirely applicable or appropriate for the students. Contextualized learning and repeated exposure to authentic collocations through DDL therefore are seen as efficient strategies for enhancing learners' ability to use collocations accurately (Lindquist, 2018; Muftah, 2023).

2. Data-Driven Learning (DDL) on Collocation Learning

Collocation learning is a critical aspect of language acquisition, and Data-Driven Learning (DDL) has emerged as a classroom approach for helping learners identify and internalize collocational patterns. The application of DDL, commonly referred to as applying corpora in language learning, is prevalently associated with Tim Johns' pioneering work. Johns (1991) proposed that learners should actively engage as researchers of the target language, while teachers serve as facilitators and coordinators of their linguistic inquiry. DDL emphasizes human interaction with authentic machine-generated data, enabling students to recognize language patterns in real-world language use (O'Keeffe & McCarthy, 2022). To facilitate collocation learning, DDL can be implemented through either computer-based or paper-based approaches, depending on classroom context and learner proficiency and/or experience with their independent searches (Chujo et al., 2012; Huang, 2014). Computer-based DDL provides learners with direct access to large corpora and interactive tools that facilitate real-time language analysis. This approach is particularly beneficial for learners who are comfortable navigating digital resources. On the other hand, paper-based DDL, which involves the use of printed concordance lines and structured worksheets, can be more suitable for learners in environments with limited access to digital tools or for those at lower proficiency levels who may find corpus software overwhelming, but it may may be limited in interactivity and data range.

In this study, paper-based DDL was implemented to enhance students' collocation learning, allowing them to engage with the chosen corpus data under the teacher's guidance based on the student's learning objectives. To ensure that the material is suitable for students' proficiency levels, teachers can modify concordance lines and adjust collocational input to prevent cognitive overload (Römer, 2011). Both computer-based and paper-based DDL provide valuable opportunities for learners to observe and acquire collocations in authentic language use (Saeedakhtar et al., 2020). By being exposed to corpus data, students develop greater awareness of natural word combinations, which contributes to improved accuracy, fluency, and overall communicative competence in English (Chambers, 2010; Lindquist, 2018; Satake, 2022). This exposure to real-life language

use allows learners to observe recurrent collocational patterns and gain a deeper understanding of language in context (He & Xie, 2024).

3. Recent Research on DDL for Collocation Learning

DDL has been recognized as an effective method for teaching collocations, particularly in addressing challenges associated with preposition usage. Studies have shown that learners who engage with corpus-based materials demonstrate improved accuracy in collocation use, as they develop an automatic recognition of word pairings through repeated exposure (Yuvayapan & Yükselir, 2021). Furthermore, DDL promotes learner autonomy by allowing students to actively explore and verify language usage rather than passively receiving information. By encouraging learners to discover language patterns inductively, this approach has been acknowledged for its role in boosting linguistic awareness and fostering independence, as noted by Johns (1991) (Kwarikunda et al., 2022). Indeed, several studies have examined the impact of DDL on collocation learning in EFL settings, consistently highlighting its effectiveness in improving learners' collocational competence. For instance, Saeedakhtar et al. (2020) found that both computer-based and paper-based DDL were equally effective regarding the use of verb-preposition collocation. Similarly, Satake (2022) demonstrated that students learning collocations through DDL produced a greater number of the collocations than those relying on dictionaries, reflecting DDL's role in fostering collocational awareness. Additionally, Çekiç (2022) compared DDL to deductive learning and found that while both methods improved collocation knowledge, DDL users retained more collocations and engaged in deeper processing. However, He and Xie (2024) noted that despite DDL's positive impact on noun collocation retention, some learners expressed uncertain or negative perceptions, suggesting a need for deeper investigation into participants' perspectives.

Considering the above studies, though many existing studies focus on word meanings, fewer have examined collocations as a primary learning target. Among those that do, most emphasize noun-verb and adjective-noun lexical combinations, with prepositional collocations receiving comparatively less attention, despite their vital role in accurate language use. Besides, data-driven vocabulary learning has primarily been implemented at the university level, as DDL is believed to be more efficient for learners with higher L2 proficiency. Nevertheless, research on its practical application in high school settings remains limited, particularly regarding collocation learning (Li et al., 2024). Building on the findings from recent existing research, this present study seeks to contribute to the literature by examining the effectiveness of DDL for teaching prepositional collocations to Thai EFL learners by addressing the research gaps related to high school students and preposition-focused collocation learning focusing on verb and adjective preposition collocations, and students' attitudes toward learning through DDL.

Methodology

1. Participants of the Study

The participants of this study were Grade 11 students from Hatyai Wittayalai 2 School, Hatyai District, Songkhla Province. A purposive sampling method was employed to select students with low proficiency in English, as determined by the school's standardized English assessment. A total of 37 students participated in the study. The selection process ensured that the participants were suitable for addressing the study's objectives by focusing on learners who were expected to benefit most from explicit instruction in preposition collocations.

2. Instruments

2.1 The selected corpus

The Corpus of Contemporary American English (COCA) was utilized as the primary resource for creating exercises aimed at teaching preposition collocations. It provides real-world language usage across diverse contexts, including fiction, academic texts, spoken language, newspapers, and magazines, making it an ideal tool for exposing learners to authentic collocational patterns. The corpus allowed students to observe verb-preposition and adjective-preposition combinations in naturally occurring language.

2.2 Teaching materials

Materials on prepositional collocations were developed to align with the students' learning objectives. While the targeted collocations were drawn from the students' coursebook, additional authentic examples from COCA were incorporated to enhance the accuracy and contextual relevance of the exercises. The materials included four exercises, each consisting of ten items. The first part required students to fill in each gap with the correct preposition, while the second part asked them to write their own sentences using the collocations from the first part. These activities aimed to reinforce both recognition and productive use of collocations.

2.3 pre-and post-tests

A pre-test and post-test were administered to measure students' knowledge of verb- and adjective-preposition collocations before and after DDL. The tests were 50 minutes long and comprised three sections with a total of 15 items. The first part required students to fill in each gap with the correct preposition. The second part involved selecting and inserting the correct verb or adjective + preposition collocation. The third part required students to correct errors in prepositional collocations. The tests were validated by three experts, and their Index of Item-Objective Congruence (IOC) score was 1.00, confirming the high validity of the tests.

2.4 Attitude Questionnaire

A 10-item questionnaire was adapted from previous research (Pan, 2024; Tosun & Sofu, 2023) to assess participants' attitudes and feedback on the learning experience. The questionnaire was administered immediately after the post-test to capture students' reflections on their experience using DDL. The questionnaire consisted of three sections. The first section collected general information about the participants. The second section contained five-point Likert scale items (Strongly Agree = 5 to Strongly Disagree = 1) to examine students' attitudes regarding the usefulness of DDL and the challenges faced while learning collocations through DDL. The last section included an open-ended question to allow students to provide additional comments and suggestions about their learning experience.

3. Data Collection Procedure

The data was collected during the second semester of the academic year 2024, between December and February, in a fundamental English class. Each week, a 50-minute session was allocated for DDL instruction. The research procedure was divided into three main stages: planning and preparation, teaching and experimentation, and post-experimentation.

3.1 Planning and Preparation

COCA was selected as the reference corpus for the study. Learning activities were designed to focus on verb- and adjective-preposition collocations. The materials were carefully developed to incorporate both corpus-based examples and structured practice exercises to facilitate students' gradual understanding of collocational patterns. Prior to implementation, the materials underwent a review by language experts to confirm their suitability for the learners' proficiency level and ensure alignment with the research objectives.

3.2 Teaching and Experimentation

A pre-test was administered to measure students' initial understanding of verb- and adjective-preposition collocations. This was followed by an introduction to the topic, where one of the researchers, as the teacher, explained the key principles of collocations and provided examples to illustrate their use in different contexts. The instruction emphasized how prepositions function in collocational patterns and how meaning can change depending on the prepositional choice.

The DDL approach was implemented by guiding students through the analysis of language data from the designed materials. Students were encouraged to identify common prepositions that frequently appear in verb- and adjective-preposition collocations, recognizing patterns through repeated exposure. To reinforce understanding, students engaged in structured exercises that required them to match verbs and adjectives with

their appropriate prepositions and construct sentences using the collocations they identified.

Throughout the activities, the teacher observed the students' interactions with the exercises and their ability to recognize and apply collocations correctly. Interventions were provided as needed, particularly when students raised their hands to seek clarification or assistance. The teacher facilitated discussions to help students reflect on their learning, encouraging them to articulate their reasoning behind specific prepositional choices.

3.3 post-experimentation

Following the completion of the DDL instruction, the post-test was administered to measure students' learning gains in the collocations after learning. The results of the post-test were analyzed to determine the effectiveness of the DDL approach in improving students' collocational competence. Afterwords, the students were asked to provide their attitudes and feedback on their experiences through questionnaire implementation.

4. Data Analysis

The pre-test and post-test results were analyzed using statistical software to determine the improvement in students' ability to use preposition collocations correctly. The data was analyzed using a paired sample t-test and was conducted to identify whether the differences between the pre-test and post-test scores were statistically significant. Following that, the effect size was calculated to determine the magnitude of the DDL approach's effect on student learning.

Responses from the Likert-scale questionnaire were analyzed using descriptive statistics, including percentages, to determine students' attitudes of the usefulness and challenges of learning prepositional collocations through DDL. Open-ended responses were qualitatively analyzed to identify recurring themes and suggestions for improvement.

Results

This section presents the findings based on data collected from pre- and post-writing tests, as well as the attitude questionnaire.

1. Pre-test and post-test results

Table 1 demonstrates a substantial improvement in students' collocational knowledge after engaging in DDL.

Table 1: Pre-test and post-test results

Test	N	$\bar{x}$	S. D.	Paired difference means	Minimum	Maximum	Sig. (2-tailed)	Cohen's d
Pre-test	37	4.84	2.75	9.48	0	10	.00	0.87
Post-test	37	14.32	0.74		13	15		

* Significant at 0.05

As illustrated in Table 1, a statistically significant difference was found when comparing the pre-test ($\bar{x}$ = 4.84, SD = 2.75) and the post-test ($\bar{x}$ = 14.32, SD = 0.74), with 9.48 of score difference (p < 0.05). The large effect size (d = 0.87) was also found, indicating a substantial impact of the DDL treatment on students' learning outcomes (Cohen, 1988).

The results from the paired-sample t-test, as shown in this table, indicated a practically and statistically significant difference between the pre-test and post-test scores, suggesting a notable improvement in students' performance after learning the collocations through DDL.

2. Questionnaire results

The data collected from the questionnaire revealed the students' attitudes toward the usefulness and the challenges of DDL, as can be seen in figure 1 and 2.
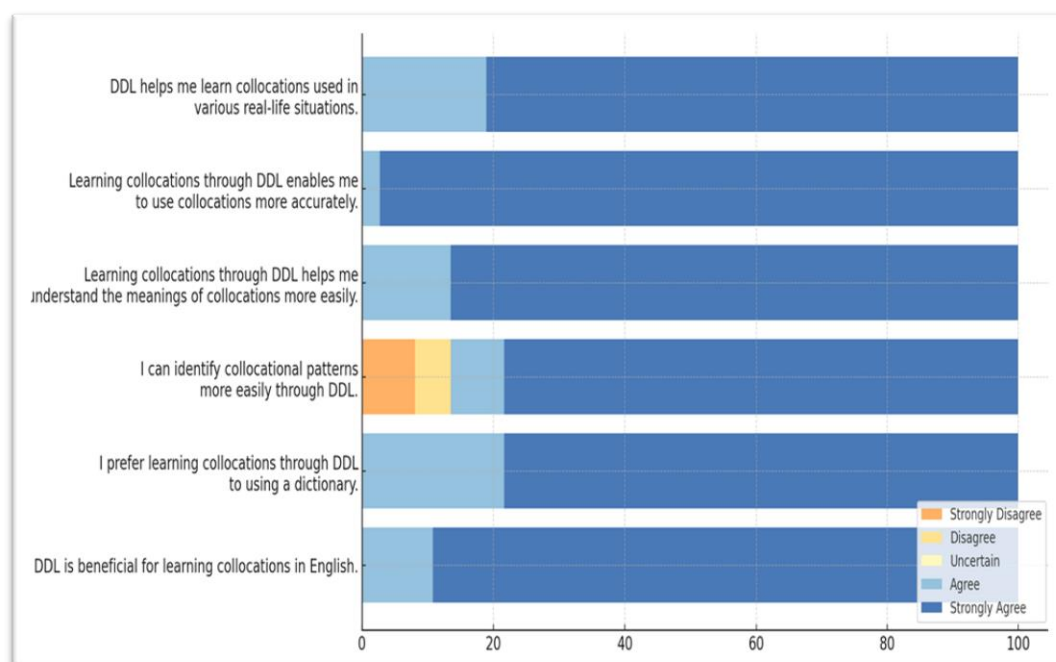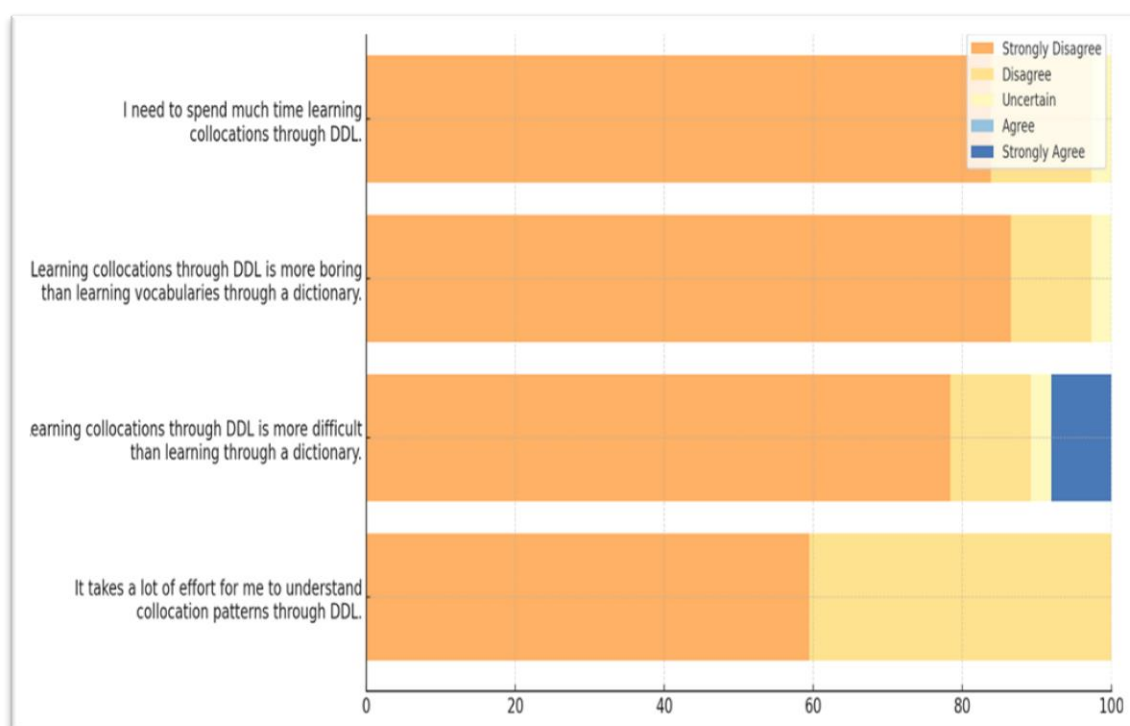

Figure 1: Students' Attitudes Toward the Usefulness of DDL

In response to the usefulness of DDL, the findings indicate that most students held positive attitudes toward its effectiveness. That is, 97.30% of students strongly agreed that learning collocations through DDL helped them to use collocations more accurately. In consistent with this, 89.19% of the students believed that DDL was beneficial for learning collocations in English. Additionally, a substantial percentage of students acknowledged that DDL helped them understand the meanings of collocations more easily (86.49%) and learn collocations used in various real-life situations (81.08%). Furthermore, 78.38% of students stated that they preferred learning collocations through DDL to using a dictionary. Although most students responded positively to the usefulness of DDL, a small number (8.10%) Strongly expressed disagreement with the statement that they could identify collocational patterns more easily through DDL.

Figure 2: Students' Attitudes Toward the Challenges of DDL

Regarding the attitudes toward the challenges of DDL, the results indicate that a substantial majority of students (86.49%) strongly disagreed that learning collocations through DDL was more boring than learning through a dictionary. Likewise, 83.78% of students strongly disagreed that learning collocations through DDL required excessive time. In terms of difficulty, 78.38% of students strongly disagreed that learning collocations through DDL was more difficult than learning through a dictionary; however, 8.11% strongly agreed with this statement. When asked about the effort required to understand

collocational patterns, 59.46% of students strongly disagreed, further suggesting that most students did not find DDL particularly demanding.

Conclusions and Discussion

The results obtained from both the pre- and post-tests, as well as the questionnaire responses, highlighted the effectiveness of DDL in improving students' collocation knowledge in prepositional collocations. The statistically significant improvement in students' performance suggests that the implementation of paper-based DDL contributed to a notable improvement in their ability to use the collocations. These findings confirm that learners who engaged in DDL developed a practical grasp of the collocations, allowing them to apply these patterns more accurately in their language production. This aligns with Schmidt's (1990) Noticing Hypothesis, which emphasizes that language learners must consciously attend to linguistic patterns for acquisition to occur. Through DDL, students were repeatedly exposed to prepositional collocations in meaningful contexts, which may have facilitated their recognition of these structures. This finding also resonates with Chambers and O'Sullivan's (2004) perspective on autonomous learning, which highlights the importance of fostering self-directed learning in language acquisition. Through DDL, students could engage in autonomous examining of data, leading to self-discovery of collocational patterns. It was stated that learners are more likely to retain linguistic information when they discover it independently rather than through passive learning (Ellis, 2003). Unlike traditional deductive teaching methods, which present rules and rely on memorization, DDL allows learners to engage in inductive learning by discovering linguistic patterns on their own.

The students' positive attitudes toward DDL further validate its effectiveness as an instructional approach. A substantial majority of students found DDL useful, believing that it helped them to use collocations more accurately and was beneficial for learning collocations. These findings align with previous research (Yuvayapan & Yükselir, 2021; Saeedakhtar et al., 2020), which has demonstrated that DDL approach improves accuracy in collocation use. The ability to repeatedly observe collocations in different contexts can allow learners to internalize the patterns and use them more effectively in their own language output. These findings are also consistent with Satake (2022), who found that DDL leads to greater collocational awareness. The finding that a significant number of students preferred DDL to traditional dictionary-based learning methods suggests that DDL instruction is not only effective but also motivating for learners (He & Xie, 2024).

Although the majority favored DDL, a small number of students expressed a preference for dictionary to learn collocations. This might be attributed to differences in individual learning styles, where some learners feel more comfortable with structured and

defined explanations rather than the inductive approach of DDL. Additionally, some students may have faced difficulties due to insufficient knowledge of collocations (Boonraksa & Naisena, 2022). For these learners, dictionaries have possibly offered a sense of clarity and immediacy in understanding collocations, as opposed to the exploratory and discovery nature of DDL (Johns, 1991). However, it should be noted that traditional dictionaries often do not provide adequate support for collocation learning, as they typically focus on individual word definitions rather than word combinations. Many dictionaries allow users to search for isolated words rather than entire collocational parings, limiting their usefulness for learners who need to understand how words function together in natural discourse. This limitation suggests that while dictionaries can be helpful for vocabulary learning, they may not be the most effective tool for acquiring collocational knowledge. Besides, the results also indicate that the challenges typically associated with DDL, such as excessive time consumption and difficulty in understanding collocational patterns, were not significant concerns for most students. In this study, the students nonetheless worked with selected concordance lines from a massive corpus, simplifying the learning process and reducing the risk of overwhelming them. In any case, the teacher-guided implementation of DDL provided necessary scaffolding, ensuring students could navigate and interpret corpus data independently. By consulting the target corpus, teachers can also make pedagogical decisions to determine what to teach while utilizing the target corpus to develop authentic teaching materials (Römer, 2011).

Overall, with respect to the key findings, it was suggested that implementing DDL into classroom instruction can enhance EFL learners' collocational competence while fostering greater autonomy in language learning. However, given the challenges some students faced, further research is needed to explore ways to optimize DDL, particularly in terms of scaffolding, teacher facilitation, and blended approach. By refining the implementation of DDL, educators may maximize its potential in supporting students' ability to use collocations more naturally and accurately in real-world communication.

Recommendations

Based on the findings of this study, several practical recommendations can be made for effectively implementing DDL in English language classrooms to improve students' collocational competence:

1. For high school EFL learners with limited access to technology or less familiarity with corpus tools, paper-based DDL is a practical and effective option. Teachers can pre-select concordance lines from accessible corpora like COCA and present them in worksheet form, allowing students to focus on noticing patterns without the cognitive load of navigating digital interfaces.

2. To prevent confusion or cognitive overload, it is recommended that teachers carefully select examples of collocations, adjusting certain lines with overly complex grammar or vocabulary.

3. Before presenting corpus data, activate students' background knowledge related to the target collocations. This helps learners make meaningful connections and better anticipate patterns, increasing the effectiveness of DDL tasks.

This study highlights the potential of DDL as an effective approach for teaching preposition collocations in EFL contexts. The findings demonstrate that engaging students in DDL could benefit from emphasizing natural and context-rich usage, leading to improved accuracy and more natural language use. The study also highlights the positive effect of DDL, as most students found it to be useful for collocation learning, with many preferring it over traditional dictionary. Nevertheless, despite its benefits, this study acknowledges several limitations. Firstly, the sample size was small and limited to a specific group of EFL students, which may restrict the generalizability of the findings. Future research could examine how DDL affects learners across a wider range of language proficiencies and educational backgrounds. Secondly, some students experienced difficulties in independently identifying collocational patterns, suggesting that DDL may not be equally effective for all learners. Lower-proficiency students may require additional scaffolding and teacher guidance to fully benefit from DDL. Another limitation of this study is the possibility of researcher-teacher bias, as the researcher also served as the instructor during the implementation of the DDL. This dual role may have influenced students' responses, attitudes, or performance, either consciously or unconsciously. To minimize this bias, the researcher adhered strictly to a standardized lesson plan, used pre-developed instructional materials, and ensured that the same level of support was given to all participants. In addition, students were encouraged to provide honest and anonymous feedback through self-administered questionnaires without the presence of the researcher. Lastly, while this study focused on paper-based DDL, the integration of digital tools and interactive corpus software could further enhance students' ability to explore collocations effectively. Given these limitations, further studies should investigate ways to optimize DDL implementation, ensuring that it is adaptable to diverse classroom contexts and learner needs. By addressing these challenges, educators can maximize the potential of DDL in enhancing collocational competence and overall language proficiency among EFL learners.

References

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., & Phoocharoensil, S. (2016). Effectiveness of English preposition learning through data-driven learning (DDL) [Unpublished master's thesis]. Thammasat University.

Chambers, A. (2010). What is data-driven learning? In A. O'Keeffe & M. McCarthy (Eds.), The Routledge handbook of corpus linguistics (pp. 345–358). Routledge.

Chambers, A., & O'Sullivan, I. (2004). Corpus consultation and advanced learners' writing skills in French. ReCALL, 16(1), 158–172.

Chiekezie, P. N. (2021). Lexical collocations in the English sentences: An overview. Pinisi Journal of Art, Humanity and Social Studies, 1(1), 1–5.

Chuenchaichon, Y. (2022). An error analysis of written English paragraphs at lexical, syntactic, and paragraph levels made by Thai EFL non-English major students. 3L: Southeast Asian Journal of English Language Studies, 28(2), 96–108.

Chujo, K., Anthony, L., Oghigian, K., & Uchibori, A. (2012). Paper-based, computer-based, and combined data-driven learning using a web-based concordancer. Language Education in Asia, 3(2), 132–145.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Demir, C. (2017). Lexical collocations in English: A comparative study of native and non-native scholars of English. Journal of Language and Linguistic Studies, 13(1), 75-87.

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

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

Firth, J. R. (1957). Papers in linguistics. Oxford University Press.

Hashemi, M., Azizinezhad, M., & Dravishi, S. (2012). The investigation of collocational errors in university students' writing majoring in English. Procedia-Social and Behavioral Sciences, 31, 555-558.

He, M., & Xie, Q. (2024). Empowering autonomy in language learning: The sustainable impact of data-driven learning on noun collocation acquisition. Humanities and Social Sciences Communications, 11(1), 1-14.

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). Thomson Heinle.

Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. ReCALL, 26(2), 163-183.

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. English Language Research Journal, 4, 1–16.

Kwarikunda, D., Schiefele, U., Muwonge, C. M., & Ssenyonga, J. (2022). Profiles of learners based on their cognitive and metacognitive learning strategy use: Occurrence and relations with gender, intrinsic motivation, and perceived autonomy support. Humanities and Social Sciences Communications, 9(1), 1–12.

Lewis, M. (Ed.). (2000). Teaching collocation: Further developments in the lexical approach. Thomson Heinle.

Li, L., Qi, X., Tang, X., & Wang, J. (2024). Effectiveness of high school students' DDL of verb–noun collocations in an EFL context. Frontiers in Educational Research, 7(3), 106–114.

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

Liu, T., & Gablasova, D. (2023). Data-driven learning of collocations by Chinese learners of English: a longitudinal perspective. Computer Assisted Language Learning, 38(3), 612-637.

McEnery, T., & Xiao, R. (2011). What can a company offer in language teaching and learning? In E. Hinkel (Ed.), Handbook of research in second language teaching and learning (Vol. 2, pp. 364–380). Routledge.

Mehmonova, Y. C. (2022). Lexical and grammatical collocations in the English sentences. Eurasian Research Bulletin, 8, 43-47.

Muftah, M. (2023). Data-driven learning (DDL) activities: do they truly promote EFL students' writing skills development? Education and Information Technologies, 1–27.

Pan, Z. (2024). EFL learners' use of data-driven learning and their attitudes in the improvement of English-speaking skills. International Journal of Learning, Teaching and Educational Research, 23(8), 98–115.

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

Saeedakhtar, A., Bagerin, M., & Abdi, R. (2020). The effect of direct and hands-off data-driven learning on low-intermediate learners' verb-preposition collocations. System, 91, 102268.

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

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

Suraprajit, P. (2021). An analysis of errors in English essays written by Thai non-English major students. Theory and Practice in Language Studies, 11(1), 12-19.

Thongvitit, S., & Thumawongsa, N. (2017). A corpus-based study of English collocations found in the abstracts of research articles written by Thai EFL writers. International Journal of Social Science and Humanity, 7(12), 751–755.

Tian, L., & Della Amirulloh, R. (2024). L1 influence on EFL learners' use of English collocations. The Journal for Research and Practice in College Teaching, 9(1), 1–18.

Tosun, S., & Sofu, H. (2023). The effectiveness of data-driven vocabulary learning: Direct concordancing through a pedagogical corpus. Journal of Language and Education, 9(3), 176–190.

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.

Yuvayapan, F., & Yükselir, C. (2021). Understanding Turkish EFL students' perceptions about collocations and investigating their collocational errors in descriptive and argumentative essays. International Journal of Curriculum and Instruction, 13(3), 2179–2194.