LEARN Journal: Language Education and Acquisition Research Network
ISSN: 2630-0672 (Print) | ISSN: 2672-9431 (Online)
Volume: 18, No: 1, January – June 2025

Language Institute, Thammasat University
https://so04.tci-thaijo.org/index.php/LEARN/index

The Effects of the ILH-based DDL on the Ability of Thai EFL Learners to Use Academic Collocations

Kietnawin Sridhanyarat[a], Supong Tangkiengsirisin[b,*]

[a] kietnawins@gmail.com, Language Institute, Thammasat University, Thailand
[b] supong.t@litu.tu.ac.th, Language Institute, Thammasat University, Thailand
* Corresponding author, supong.t@litu.tu.ac.th

APA Citation:
Sridhanyarat, K., & Tangkiengsirisin, S. (2025). The effects of the ILH-based DDL on the ability of Thai EFL learners to use academic collocations. *LEARN Journal: Language Education and Acquisition Research Network. 18*(1), 190-211. https://doi.org/10.70730/AAUA5255

Received 02/05/2024

Received in revised form 29/07/2024

Accepted 17/08/2024

ABSTRACT

The purpose of this study is two-fold: 1) to investigate the effects of Data-Driven Learning (DDL) framed within the Involvement Load Hypothesis (ILH) on Thai learners' use of academic collocations and 2) to examine how Thai learners utilized the involvement load (IL) components (need, search, and evaluation) to master academic collocations. It is hypothesized that the more learners actively engage with the IL components in a DDL task, the greater their development will be in acquiring new collocations. Participants *(n* = 31) were provided with repeated exposure to authentic L2 input (COCA) through DDL. They explored and analyzed three collocation patterns: adjective-noun, verb-noun, and adverb-past participle. The learners also verbalized their thoughts while completing the DDL tasks. A multiple-choice test and a collocation judgement test were constructed to assess the students' academic collocation competence. The results demonstrated that the ILH-based DDL generated a significant effect on Thai learners' performance of academic collocations. Thai learners also utilized all of the IL components in acquiring

academic collocations through DDL. The search component was used most frequently, followed by the evaluation and need components. These findings provide useful insights into both the predictions of the ILH and L2 pedagogy regarding academic collocation learning.

Keywords: academic collocations, COCA, Data-Driven Learning (DDL), Involvement Load Hypothesis (ILH)

Introduction

The use of appropriate L2 collocations is regarded as the most significant indicator of overall lexical proficiency (Crossley et al., 2015). Conceptualized as a word combination that co-occurs frequently in spoken and written discourses, a collocation also serves as a crucial aspect of nativelike English proficiency (Mallikamas & Pongpairoj, 2005; Vu & Peter, 2022). Collocations are thus essential for L2 learners who wish to achieve language output fluency. However, using appropriate academic collocations poses a great challenge to L2 learners regardless of their proficiency level (Khantiwong & Thiengthong, 2022). Many L2 writers struggle with using accurate collocations in their academic writing (Nesselhauf, 2003; Siyanova-Chanturia, 2015). Furthermore, L2 learners are likely to use only a small range of collocations, albeit correctly (Laufer & Waldman, 2011).

Regarding L2 pedagogy, several previous studies have investigated effective methods that facilitate L2 collocation learning (e.g., Kulsitthiboon & Pongpairoj, 2018; Zhang, 2017). In the digital era, Data-Driven Learning (DDL) is one of the most common approaches to teaching and learning L2 collocations (Saeedakhtar et al., 2020). DDL has become popular among L2 researchers and teachers due to its emphasis on authentic language use and the use of technological tools to facilitate language learning (Alanazi, 2023; Johns, 1990; Rezaee et al., 2015). As far as DDL is concerned, learners are motivated to independently explore and analyze target features using authentic language data, which ultimately allows them to discover certain rules inductively (Hunston, 2002; Johns, 1990). In the Thai EFL context, however, little attention has been paid to an investigation of the effects of DDL conceptualized within the Involvement Load Hypothesis (ILH) on Thai learners' performance of academic collocations. This study aimed to examine the influence of the ILH-based DDL on the ability of Thai learners to use academic collocations. The investigation also considered to what extent Thai learners employed the three cognitive dimensions of Involvement Load (IL) components (need, search, and evaluation) to master academic collocations through DDL. Essentially, the inclusion of the ILH in the realm

of DDL enabled the researcher to clearly see how and whether the participants had strong need and evaluation. This serves as a mechanism through which the participants could become more self-reliant in learning academic collocations.

Given the research objectives mentioned above, this study attempted to address the following questions:

1) What are the effects of the ILH-based DDL on Thai EFL learners' ability to use academic collocations?

2) How do Thai EFL learners employ the IL components to improve their ability to use academic collocations through the ILH-based DDL?

Arguably, the findings obtained would offer valuable insights into both the effectiveness of the ILH as a tool for evaluating tasks for academic collocation learning and L2 pedagogy regarding academic collocation learning.

Moreover, it should be noted that the current study made use of the ILH because it enabled the researcher to determine the extent to which DDL tasks assigned could be more effectively structured which would eventually contribute to academic collocation learning.

Literature Review

Data-Driven Learning (DDL)

In the digital age, many scholars have shown interest in using DDL to facilitate L2 collocation learning. For example, Kartal and Yangineksi (2018) investigated the influence of using corpus tools on student teachers' knowledge of verb-noun collocations over a period of four weeks. Huang and Tsao (2019) also examined the impact of NetCollo, a web-based exploration tool, on L2 learners' use of adjective-noun and verb-noun collocations.

As initially established by Johns (1990), DDL is viewed as inductive learning where students are motivated to take responsibility for their own learning. Under a DDL circumstance, learners notice particular linguistic items within concordance lines and analyze their use in context (Boulton, 2012). Three steps are involved in conventional DDL: identification, classification, and generalization. Identification is the first stage where learners explore and investigate specific linguistic features in concordance lines. This step is crucial for L2 learning as it directs learners' attention to searching for the meaning and form of an unknown linguistic feature. Word searches can be conducted through software tools or paper-based materials (Hunston, 2002). Classification is the second step where learners use various DDL strategies, such as identifying, grouping, comparing, and inferencing to acquire target structures (Kulsitthiboon & Pongpairoj, 2018). The last step is

generalization that allows learners to utilize their DDL skills to practice using target features. Given the aforementioned steps, DDL has emerged as an important approach to L2 collocation learning and teaching as it provides learners with repeated exposure to linguistic features through enriched authentic L2 input (Luo, 2016; Saeedakhtar et al., 2020). Once learners notice new linguistic features present in authentic L2 data, they are more likely to succeed in language learning (Robinson, 1995).

The salient characteristics of DDL are as follows. DDL is based on the use of corpora or a collection of authentic L2 input such as texts or even learner-generated L2 data (Yoon, 2011). DDL is a student-centered approach that empowers students to take responsibility for their own learning, ultimately increasing learner autonomy (Luo, 2016). Rather than relying solely on their teacher's assistance, learners can independently explore target language features through L2 data, potentially leading to success in L2 learning (Boulton, 2017). This is because DDL is characterized by discovery learning where learners are motivated to autonomously discover linguistic patterns and rules by themselves rather than relying on explicit instruction and explanations (Bernardini, 2002). DDL fosters cooperative learning where students work together to analyze L2 data, exchange ideas, and share discoveries (Kulsitthiboon & Pongpairoj, 2018).

In the extant literature, there has been no exploration of how learners use the IL components to master academic collocations through DDL. Therefore, in the present study, DDL is incorporated into the ILH to elucidate the extent to which the ILH-based DDL enhances the ability of Thai learners to use academic collocations.

The Involvement Load Hypothesis (ILH)

The ILH has been used by numerous researchers (e.g., Hulstijn & Laufer, 2001; Kim, 2008) as a tool to predict the effectiveness of a task for incidental vocabulary learning. According to the ILH, it is hypothesized that the retention of new words depends on the IL of a task in which three dimensions of IL (need, search, and evaluation) are involved. Need, a motivational dimension, can be defined as the necessity of completing a specific task to learn unfamiliar words. Need is manifested in three levels: absent, moderate, and strong. In the absence of need, learners do not feel compelled to perform a task. Moderate need arises when learners' motivation to complete tasks is prompted by either the task itself or by the teacher. Strong need is imposed by learners themselves, such as when they independently search for unknown words to incorporate into their speech or writing. Search, as a cognitive factor, involves the process of searching for word forms and meanings. Search can be manifested in two forms: absence

and presence. In the presence of search, learners are tasked with actively seeking word meanings and forms, whereas its absence entails no such requirement for word search. Evaluation, as a cognitive process, involves the comparison of new word forms and meanings with similar forms and meanings. Evaluation can occur at three levels: absent, moderate, and strong. In the absence of evaluation, learners neither recognize nor produce any word form. Moderate evaluation concerns learners choosing the most appropriate word from a list of options to fill in a blank. Strong evaluation arises from the requirement of students to use target words in authentic contexts, such as sentence writing and essay writing.

It is important to note that the need component aligns with the identification step, where participants are required to hypothesize about word combinations before conducting word searches and applying their DDL strategies to the use of academic English collocations. The search component corresponds to the classification step as students are encouraged to use their observing and identifying strategies to facilitate their use of academic collocations. The evaluation component is integrated into both the classification and generalization steps. In the classification step, learners utilize DDL strategies, such as comparison, differentiation, grouping, and inference, to study academic English collocations. In the generalization step, they are prompted to independently apply their observational skills to the use of academic collocations.

As far as previous studies are concerned, there has been limited systematic investigation on the effects of DDL incorporated into the ILH on the ability of Thai learners to use academic collocations. The current investigation thus examined the effects of the ILH-based DDL on the ability of Thai learners to employ academic collocations. The study also investigated the degree to which Thai learners utilized the IL components of need, search, and evaluation, as provided in the DDL tasks, to enhance their proficiency in academic collocations.

Related Studies Regarding DDL and the ILH

Learners across different L2 proficiency levels frequently struggle with the use of inappropriate collocations (Wolter & Gyllstad, 2011). Inadequate exposure to collocations in authentic L2 input is a crucial factor that characterizes the difficulty of collocational use (Saeedakhtar et al., 2020; Vu & Peters, 2021). Therefore, numerous researchers have explored various methods to enhance L2 collocational competence. Recently, DDL has been adopted as an effective approach to facilitate L2 collocation learning (Rezaee et al., 2015). By conducting concordance searches, learners can repeatedly encounter authentic L2 data, prompting them to explore and analyze

collocations and ultimately achieve nativelike proficiency of English (Boulton, 2010).

For example, Chan and Liou (2005) conducted a one-group pretest-posttest study to examine the influence of DDL on Chinese college students' use of verb-noun collocations. A Chinese-English bilingual concordancer serves as a valuable source for collocation learning. The results demonstrated that the students achieved significant learning achievements in verb-noun collocations following exposure to DDL activities. The learners further expressed positive attitudes toward the use of DDL to enhance the learning of verb-noun collocations.

Similarly, Rezaee et al. (2015) reported on the positive impact of DDL based on two scaffolding types—student-student support and teacher-student support—on the use of collocations among 120 Iranian intermediate learners of English. The learners were divided into four different learning modes: DDL with peer-peer support, DDL with teacher-student support, teacher-student support only, and neither DDL nor support. The findings revealed that the first three experimental groups of the learners achieved learning improvements more significantly than the control group. The participants also expressed positive opinions regarding the implementation of both DDL and scaffolding to enhance their use of L2 collocations.

As noted, there has been little focus on incorporating DDL into the ILH to predict the effectiveness of a DDL task to improve the ability of Thai learners to use academic collocations. However, the predictions of the ILH have been used to investigate single words rather than groups of words like collocations. To address this gap, the current study examined the influence of DDL on the ability of Thai students to use academic collocations. The investigation also explored the IL components that Thai students utilized to enhance their proficiency in English collocations.

Methodology

Participants

31 Thai EFL students were recruited to participate in this study. Their English proficiency varied between B1 and B2 CEFR levels as determined by a STEP test, an internal proficiency English test provided by their university. This research adds to the literature regarding how intermediate learners use the IL components to improve their academic collocation fluency and proficiency (Boulton, 2017). The accurate use of academic collocations is a key indicator of lexical proficiency, which contributes to the enhancement of L2 learners' academic writing skills (Crossley et al., 2005). The selection of the 31 students aligned with the number of the students enrolled on the

researcher's course. Prior to the intervention, the participants were asked to sign a consent form to confirm their voluntary participation. Their participation in this study, or lack thereof, did not impact their grades or scores.

Research Instruments

The research instruments included the multiple-choice test (MCT) and the collocation judgement test (CJT). Think-aloud techniques were also used to understand Thai students' use of the IL components as they were performing three DDL tasks: adjective-noun, verb-noun, and adverb-past participle. Think-alouds were selected for this investigation as they provide primary evidence to support claims concerning learners' use of the IL components during the learning process (Padilla & Leighton, 2017).

Prior to the actual study, both the MCT and CJT as well as the think-aloud tasks were determined to ensure their content validity and reliability. Regarding validity, three Thai instructors of English, whose expertise in Applied Linguistics and ELT, were asked to evaluate the tests using the IOC, resulting in a significant value of 0.99. As determined against the IOC, the target think-aloud tasks demonstrated their acceptable content validity (George & Mallery, 2003), with 0.89 for the adjective-noun task, 0.93 for the verb-noun task, and 0.88 for the adverb-past participle task, respectively.

Subsequently, the reliability of the MCT and CJT was examined by piloting with 15 students who shared a similar L2 background as those participating in the actual research. Based on KR20, the MCT and CJT were of significantly statistical reliability KR20 = 0.82 and KR20 = 0.80, respectively. The inter-rater reliability of the think-aloud data was determined using a percentage agreement technique (Goodwin, 2001), and the results indicated statistical reliability (r = 0.83).

Research Design

Table 1

DDL-based Instructional Design

Sessions	Study Treatment	Time (Minutes)
1	Introduction	30
2	Pretest (MCT and CJT)	40
3	DDL and think-aloud training	60

4	Adjective-noun collocations	60
5	Verb-noun collocations	60
6	Adverb-past participle collocations	60
7	Wrap-up activity	60
8	Posttest (MCT and CJT)	40

As shown in Table 1, this research was conducted over a period of eight weeks. Initially, the participants were first introduced to the objectives of this study and what they had to do during the instructional intervention. In the second week, the participants took the pretests (i.e., MCT and CJT), which lasted approximately 40 minutes. Before being exposed to academic collocations through DDL, the participants received training in DDL and think-aloud tasks. That is, the participants first registered to use the COCA to become acquainted with the corpus. Next, an adjective-noun collocation task was assigned to them. This would ensure that they would be capable of carrying out the task in subsequent weeks. During the study treatment, the students performed DDL tasks involving adjective-noun, verb-noun, and adverb-past participle collocations. In the final two sessions, a wrap-up session was conducted, and the MCT and CJT were administered as posttests to measure the students' ability to use academic collocations after the DDL intervention.

Target Collocations

Lexical collocations pose more challenges to L2 learners in terms of production when compared to grammatical collocations. This is because lexical collocations exhibit higher variability, making them harder to predict (Yamashita & Jiang, 2010). However, grammatical collocations are often treated as fixed expressions, which facilitates easier use. Therefore, only lexical collocations were examined.

Adjective-noun (e.g., *significant influence*, *concerted effort*, and *major implications*), verb-noun (e.g., *convey meaning*, *perform* (*a*) *task*, and *demonstrate competence*), and adverb-past participle (e.g., *clearly related*, *greatly influenced*, and *deeply rooted*) collocations were derived from a large corpus. According to Ackermann and Chen (2013), the academic English collocations (AEC) examined were recruited from the Pearson International Corpus of Academic English (PICAE), which includes over 25 million words. As a result, the AEC comprises pedagogically relevant cross-disciplinary combinations frequently used in academic English. Additionally, these target collocations consist of single words that are likely to co-occur within a span of ±2 words. This

specific span was chosen because it has been observed to be more frequent compared to other spans, indicating a higher likelihood of encountering these collocations in academic texts. Target collocational patterns—adjective-noun, verb-noun, and adverb-past participle—were distributed in the tests in conformity with their frequency in academic texts. Each test contained four lexical items, totaling 24 items across all tests. The distribution of the collocations was carefully planned to minimize cognitive fatigue, which could have impaired participants' ability to perform optimally (Ling et al., 2014). Excessive cognitive strain might have led to results that would distort the participants' ability to use academic English collocations.

These three patterns of collocations were then reexamined against the COCA using two criteria. First, they had to be the most frequently used. A word combination was regarded as a collocation if it appeared in the top 25 frequency list as determined by the COCA. Second, a minimum MI score of 3 was applied as a criterion to determine if such word combinations were regarded as accurate collocations. Finally, distribution was employed to examine the frequency with which target collocations appeared in academic contexts.

Data Collection

As shown in Table 1, the MCT and CJT were administered to the participants before the study intervention. Next, a 60-minute DDL training session was conducted to familiarize the students with DDL worksheets focusing on adjective-noun, verb-noun, and adverb-past participle collocations. Subsequently, six participants were asked to verbalize their thoughts as they were completing each hands-on DDL task for 60 minutes through the COCA. They were divided into the high-proficiency learner group and the low-proficiency learner group based on their pretest proficiency in academic collocations, their willingness to verbalize their thoughts, and their talkativeness. Three participants were categorized as the high-proficiency learner group as their pretest scores were higher than the mean ($\bar{x}$ = 1.97). The remaining three, whose pretest scores fell below the mean, were categorized as the low-proficiency learner group. A 60-minute session was also conducted to conclude the implemented DDL activities. After exposure to concordance lines, the same tests were administered as posttests to evaluate the participants' proficiency in academic collocations.

Data Analysis

Both quantitative and qualitative data were gathered. Based on the Paired-Samples T Test, the quantitative data obtained from both the pretest and posttest were first compared. The Paired-Samples T Test was deemed appropriate as it involved two measurements (MCT and CJT) taken at two separate times (pretest and posttest) from the same group of individuals (Ross & Willson, 2017). The think-aloud data were transcribed by the researchers and put into three categories: need, search, and evaluation. The inter-rater reliability of the data was then examined by a Thai research associate who holds a Ph.D. in ELT and had experience doing qualitative research. Using a percentage agreement technique, the think-aloud data were significantly reliable ($r = 0.94$).

Results

Thai Learners' Ability to Use Academic Collocations

The results indicated a notable influence of DDL on the ability of Thai learners to use academic collocations. The analysis of the results was conducted through a Paired-Samples T Test, with further details provided in Table 2.

Table 2

Comparison of Overall Pretest and Posttest Scores

Tests	Pretest	Posttest	T	Sig. (2-tailed)
MCT **CJT**	1.97	10.06	-27.27	0.00

As illustrated in Table 2, the posttest mean score (10.06) obtained from both the MCT and the CJT is statistically higher than the pretest score (1.97) of the same tests at the level of significance ($p < 0.05$). Based on this evidence, it can be concluded that Thai EFL learners demonstrate improved accuracy in the use of academic collocations due to the ILH-based DDL.

The learners in the advanced group were labelled as high proficiency student 1, high proficiency student 2, and high proficiency student 3, while the remaining learners were labelled as low proficiency student 1, low proficiency student 2, and low proficiency student 3. The think-aloud reports of the two learner groups were presented as follows.

Think-aloud Reports of High-Proficiency Learner Group

In each DDL task, there were eight questions designed to measure the participants' use of IL. Examples of think-aloud reports generated by the high-proficiency learner group are shown in the following items. Items 1 and 2 focus on the use of need, item 3 is related to the use of search, and the remaining five items (4-8) pertain to the use of evaluation.

In think-aloud reports 1 and 2, high proficiency student 1 used the need component to learn academic collocations.

> (1) Item 1 asked me to look for the meaning of an academic collocation. I think it means a word combination that is used together in academic contexts. A collocation consists of a node and a collocate. Collocates for a node should share similar meanings and be found in similar contexts. (High proficiency student 1)

> (2) Item 2 required me to find adjectives that occur frequently with the node *influence*. (High proficiency student 1)

In think-aloud report 3, high proficiency student 2 identified adjectives that appeared frequently with the node *influence*. This evidence demonstrates that high proficiency student 2 employed the search component to learn adjective-noun collocations through DDL.

> (3) Upon searching for adjectives in the COCA, I find that *political*, *positive*, *bad*, *significant*, and *great* co-occur frequently with the noun i*nfluence*. (High proficiency student 2)

Think-aloud reports 4-8 demonstrate the use of the evaluation component in learning academic collocations through DDL. As shown in think-aloud report 4, high proficiency student 2 compared adjectives that expressed the same meaning for node *influence*.

> (4) In the frequency list, I can group adjectives that share the same meaning for the noun *influence*. Group 1 consists of

> *powerful* and *strong*; Group 2 consists of *bad* and *negative*; Group 3 consists of *great* and *major*. (High proficiency student 2)

High proficiency student 3 utilized the evaluation components by verbalizing their thoughts as illustrated in think-aloud reports 5-7, which involves using comparing, differentiating, grouping, and inferencing strategies.

> (5) I can find one concordance line *There are the casino operators that hold significant influence at the state level* that contains the collocation *significant influence*. (High proficiency student 3)

> (6) I think the collocations *powerful influence* and *strong influence* convey the meaning *the power to influence people or things*. (High proficiency student 3)

> (7) One concordance line that contains the use of *powerful influence* is *Many generations have been blessed by his powerful influence on the field*. (High proficiency student 3)

High proficiency student 2 also used the evaluation component to improve their proficiency in academic collocations. This evidence is revealed in think-aloud report 8.

> (8) *Powerful* occurs frequently with various nouns, such as *tool*, *man*, *people*, and *weapon*. I've found that *powerful* is likely to collocate with objects, such as *powerful tool* and *powerful weapon* and with people, such as *powerful people* and *powerful men*. (High proficiency student 2)

In summary, the think-aloud reports reveal that high proficiency learners used all of the IL components—need, search, and evaluation—to master academic collocations through DDL. However, it should be noted that these advanced learners skipped specific items related to the need and evaluation components.

Think-aloud Report of Low-Proficiency Learner Groups

The low-proficiency learner group also verbalized their thoughts while completing the same DDL tasks as the high-proficiency learner group. The low-proficiency learners' use of IL to improve their ability to use academic collocations is presented as follows.

In all DDL tasks, low proficiency students were encouraged to employ the need component before using the other two components to master academic collocations. An example of using the need component is

presented in think-aloud report 1, articulated by low proficiency student L1. While it is important to note that not all low proficiency students completed item 1, which required them to use the need component, they consistently completed item 2 in all DDL tasks, which encouraged them to use the need component.

> (1) An academic collocation refers to a group of words that co-occur frequently in academic context. Academic collocations can be found in academic writing and articles. (Low proficiency student 1)

> (2) What adjective collocates are used frequently with the node *influence*? (Low proficiency student 1)

As shown in think-aloud report 3, Low proficiency student 2 attempted to analyze adjectives that co-occurred frequently with the node influence. This item was specifically designed to investigate the extent to which Thai EFL learners used the search component to master academic collocations through DDL.

> (3) In the frequency list, *political*, *positive*, *bad*, *significant*, *strong*, and *powerful* appear frequently with the noun *influence*. (Low proficiency student 2)

In think-aloud reports 4 and 5, low proficiency student 2 used the evaluation component, employing comparing, grouping, and inferencing strategies to learn academic collocations. In think-aloud report 4, low proficiency student 2 compared two verbs *develop* and *demonstrate*, noting that they share the same meaning for the noun *competence*. However, it should be noted that while these verbs do not share the same meaning, they do collocate with *competence*.

> (4) I think *develop* and *demonstrate* are collocates that express the same meaning for the node *competence*. (Low proficiency student 2)

> (5) I find one concordance line for the collocation *demonstrate competence*. The concordance line I found is *They attempt to demonstrate competence relative to others*. (Low proficiency student 2)

Low proficiency student 3 used the evaluation component to master adverb-past participle collocations. Specifically, she analyzed an adverb that appeared frequently with the past participle *influenced* that expresses one

meaning presented in the adverb-past participle collocation task. An example of low proficiency student 3's learning of adverb-past participle collocations is presented in think-aloud report 6.

> (6) This item requires me to provide collocations that express two different meanings: to affect or change how someone or something develops, behaves, or thinks and to cause someone to change a behavior, belief, or opinion. I think *greatly influenced* expresses the first meaning, while *significantly influenced* conveys the second meaning. (Low proficiency student 3)

Low proficiency student 3 further used the evaluation component to learn the adverb-past participle collocation through DDL. She explored and analyzed one concordance line containing the collocation *significantly influenced*. Low proficiency student 3's think-aloud report is presented in item 7 below.

> (7) I think the concordance line *Performance quality was significantly influenced by TQM practices dimensions* demonstrates how the collocation *significantly influenced* is used. (Low proficiency student 3)

However, low proficiency student 3 did not complete item 8, which required her to use the evaluation component to group, compare, and differentiate adverb-past participle collocations through DDL. This item specifically assessed the use of the evaluation component.

In conclusion, it appears that low proficiency students utilized all of the IL components—need, search, and evaluation—to master academic collocations through DDL. Despite the presence of multiple need and evaluation components in each DDL task, low proficiency students did not consistently use all of the need and evaluation components. They tend to skip specific items that require them to employ either need or evaluation components.

Overall IL Components Used by Thai Learners

Table 3

IL Components Employed by Thai EFL Learners

IL Components	Frequency	Percentage	Ranks
Need	28 out of 36	77.78	3
Search	36 out of 36	100	1
Evaluation	72 out of 90	80%	2

Discussion

It is found that DDL enhances Thai learners' proficiency in using academic English collocations. Specifically, the participants demonstrated increased accuracy in employing various collocational patterns such as adjective-noun, verb-noun, and adverb-past participle following the intervention. These findings highlight the positive impact of DDL on enhancing Thai EFL learners' collocational competence.

DDL as implemented in this study has demonstrated a positive impact on the effective acquisition of academic collocations. This is attributed to the participants' exposure to authentic input, enabling them to internalize the studied collocations. The authentic input facilitated the transition from mere exposure to meaningful intake, thereby aiding participants in acquiring the targeted collocations with relative ease. These findings are consistent with Rezaee et al. (2015), who advocate that learners enhance their collocational competence through discovery learning, enabling them to analyze language data independently and identify patterns. Vyatkina (2016) also supports the effectiveness of DDL, highlighting its inductive approach to learning collocations, which enhances learners' engagement and motivation. Additionally, Lee and Lin (2019) suggest that DDL promotes L2 vocabulary acquisition by empowering learners to take ownership of their learning process. Thus, due to its inductive learning methodology and utilization of authentic language data, DDL proves to be an effective approach for mastering academic English collocations.

The COCA serves as a valuable source of authentic data for L2 learning (Johns, 1990). Consequently, engaging in DDL activities can significantly enhance Thai students' acquisition of academic collocations. These findings underscore the role of repeated exposure to collocations within meaningful contexts in fostering proficiency among Thai learners in academic English collocations. The current students were exposed multiple times to concordance lines containing targeted academic collocations. The results demonstrate that the ILH-based DDL enhances Thai learners' proficiency in using academic English collocations. It can be argued that DDL facilitates discovery learning, enabling students to transform input into meaningful intake. Specifically, the COCA's provision of repeated exposure to concordance lines supports learners in acquiring academic English collocations. Nation (2001) and Schmitt (2010) further validate these findings by suggesting that repeated encounters with unfamiliar words in diverse contexts can enhance vocabulary acquisition.

The search component was found to play the most significant role in learning academic collocations, followed by evaluation and need, respectively. These findings strongly corroborate the predictions of the Involvement Load

Hypothesis (ILH) proposed by Laufer and Hulstijn (2001). According to the ILH, the effective retention of unfamiliar words is positively influenced by three task-related factors: need, search, and evaluation. Therefore, it can be inferred that learners' active engagement with these IL components correlates with substantial advancements in acquiring academic collocations.

Additionally, search appeared to contribute significantly to the learning of academic collocations, followed by the evaluation and need components, respectively. These findings contrast with previous research. For instance, Laufer and Hulstijn (2001) suggested that the search component might have the least impact on initial learning, while Kim (2008) proposed that strong evaluation is crucial for vocabulary acquisition. Tang and Treffers-Daller (2016) argued that evaluation plays the primary role in incidental vocabulary learning, followed by the need component, with the search component showing no significant effect. Similarly, Yanagisawa and Webb (2021) supported the idea that evaluation has the greatest influence on vocabulary learning, followed by need, while search has no positive effect.

Interestingly, in this study, the search component emerged as the most influential factor for learning academic collocations through DDL. This divergence from previous findings may be explained by the emphasis of DDL on allowing learners to actively search for unknown words using authentic data sources such as the COCA (Boulton & Cobb, 2017). Therefore, the learners predominantly utilized the search component due to the nature of DDL, which requires them to explore and analyze target items within concordance lines provided by the COCA.

Given the noticeable differences in how each group approached the use of the need, search, and evaluation components, it appears that both groups completed their DDL tasks differentially. The more capable group tended to skip the need stage because of the repetitive nature of the questions. In contrast, the less proficient group consistently completed this stage, presumably believing that repetition would reinforce their understanding of academic collocation concepts through DDL.

Implications and Recommendations

The findings reveal that DDL enhances the ability of Thai EFL learners to use academic collocations. This study suggests that repeated exposure to authentic input enables Thai learners to internalize target academic collocations. In addition, DDL promotes academic collocation learning as it empowers Thai learners to facilitate active engagement and ownership in their acquisition process of academic collocations. Therefore, teachers and researchers should consider these crucial aspects when designing

tasks or activities for collocation learning in general, and academic collocation learning in particular.

The results also demonstrate that the components of need, search, and evaluation play a pivotal role in facilitating the effective acquisition of academic collocations among Thai learners. These insights are valuable not only for researchers aiming to develop a comprehensive model for teaching and learning academic collocations but also for educators seeking to enhance their L2 vocabulary instruction, particularly in the area of collocations. This knowledge prepares teachers to effectively teach L2 vocabulary.

According to the ILH, the retention of new vocabulary depends, in large measure, on the level of attention learners devote to the word (Laufer & Hulstijn, 2001). The findings suggest that the ILH serves as a valuable tool for teachers to assess the efficacy of tasks aimed at collocation learning. Specifically, the investigation emphasizes an equitable distribution of the IL components (need, search, and evaluation), where each component plays an equally vital role in the learning process of academic collocations. Consequently, academic collocations processed with higher levels of IL are more likely to be learned successfully compared to those with lower IL.

Based on the findings reported, several critical issues warrant further investigation. Future research endeavors should focus on gaining a deeper understanding of how the three dimensions of IL and other related factors contribute to academic collocation learning. Notably, there is a gap in the literature concerning systematic investigations into the effects of varying degrees of the need component on academic collocation acquisition. Researchers could explore the impact of strong versus moderate need on learning academic collocations, providing valuable insights into effective teaching strategies.

While this study revealed that the search component positively influenced Thai learners' proficiency in using academic English collocations, it did not explore the potential effects of different levels of the search component (e.g., moderate versus strong search) on collocation learning through DDL. Future studies should incorporate these variables to ascertain their influence on the acquisition of academic English collocations. This inquiry could contribute to task effectiveness for L2 collocation development or acquisition.

Moreover, the use of sentence completion exercises as a form of controlled practice may be justifiable. However, subsequent research could investigate how different types of practice, such as controlled exercises versus naturalistic tasks like essay writing, impact L2 learners' proficiency in using academic collocations through DDL. This comparative approach will shed light on optimal instructional practices for enhancing collocational competence.

Conclusion

The primary objective of this study was to assess the impact of DDL on Thai learners' proficiency in using academic English collocations. It also aimed to examine how Thai learners employed IL components during the acquisition of academic collocations through DDL. The findings indicate that DDL significantly enhanced Thai learners' capability to produce academic collocations. The participants demonstrated utilization of all the IL components in mastering academic collocations through DDL, with the search component being the most frequently employed, followed by the evaluation component, and then the need component.

One limitation is the restricted availability of target academic collocations for specific patterns. Consequently, the same word with different parts of speech (e.g., *influence* and *influenced*) appeared in both the adjective-noun and adverb-past participle collocation tasks. In this study, consistency was not considered a criterion when selecting target academic collocations. Therefore, learners might accurately use certain congruent noun-noun collocations (e.g., *data analysis* and *survey data*) without the influence of the ILH-based DDL. Only six participants were involved in the think-aloud sessions. As a result, the findings from the think-aloud reports may not fully represent the actual IL components utilized by Thai EFL learners. Additionally, the investigation employed a one-group pretest-posttest design exclusively. Therefore, the findings may not sufficiently capture the broader nature of academic collocation learning or the positive effects of DDL on academic collocation acquisition beyond the scope of this investigation.

About the Authors

Kietnawin Sridhanyarat: A Ph.D. Candidate in the Ph.D. in English Language Teaching Program (International Program), the Language Institute of Thammasat University. His research interest centers around Second Language Acquisition and its applications to English language teaching. He is currently the Deputy Director of the G&E Center at Silpakorn University, Thailand.

Supong Tankiengsirisin: An Associate Professor at the Language Institute of Thammasat University, Thailand. He earned his Ph.D. in English Studies from the University of Nottingham, UK. He serves as the Director of Advanced Center for Testings at Language Institute, Thammasat University (LITU-ACTs), and is currently Past President of Thailand TESOL Association. He also serves on several Editorial Committees for international

academic journals. His areas of interest include *Global Englishes*, *discourse analysis*, *ESP*, *ELF*, and *second language writing*.

References

Ackermann, K., & Chen, Y. (2013). Developing the academic collocation list (ACL)—A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes, 12*, 235-247.

Alanazi, Z. (2023). Data-driven learning tasks and involvement load hypothesis. *World Journal of English Language*, *13*(2), 23-32.

Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann, & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165-182). Edwin Mellen Press.

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

Boulton, A. (2012). Beyond concordancing: Multiple affordances of corpora in university language degrees. *Procedia - Social and Behavioral Sciences*, *34*, 33 – 38.

Boulton, A. (2017). Data-Driven learning and language pedagogy. In S. L. Thorne., & M. Stephens (Eds.), *Language, education and technology* (pp. 181-192). Springer.

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

Chan, T. P., & Liou, H. C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, *18*(3), 231–250. https://doi.org/10.1080/09588220500185769

Goodwin, L. D. (2001). Interrater agreement and reliability. *Measurement in Physical Education and Exercise Science*, *5*(1), 13–34.

Crossley, S. A., Salsbury, T., & McNamara, D. A. (2015). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, *36*(5), 570–590.

George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference. 11.0 update*. Boston: Allyn & Bacon.

Huang, P-Y., & Tsao, N-L. (2019). Using collocation clusters to detect and correct English L2 learners' collocation errors. *Computer Assisted Language Learning*, *34*(3), 270-296. https://doi.org/10.1080/09588221.2019.1607880

Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, *51*, 539–558.

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

Johns, T. (1990). From printout to handout. Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, *10*, 14-34.

Kartal, G., & Yangineksi, G. (2018). The effects of using corpus tools on EFL student teachers' learning and production of verb-noun collocations. *PASAA*, *55*, 100-125.

Khantiwong, W., & Thienthong, A. (2022). Thai EFL learners' knowledge of congruent and incongruent academic L2 collocations. *LEARN Journal: Language Education and Acquisition Research Network*, *15*(1), 809-835.

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, *58*, 285–325.

Kulsitthiboon, S., & Pongpairoj, N. (2018). Cooperative corpus consultation for acquisition of adjective + preposition collocations. *GEMA Online® Journal of Language Studies*, *18*(3), 57-72.

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22(*1), 1-26. https://doi.org/10.1093/applin/22.1.1

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, *61*(2), 647–672.

Ling, G., Mollaun, P., & Xi, X. (2014). A study on the impact of fatigue on human raters when scoring speaking responses. *Language Testing*, *31*(4), 479-499.

Lee, P., & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, *81*, 14-25. https://doi.org/10.1016/j.system.2018.12.011

Luo, Q. (2016). The effects of data-driven learning activities on EFL learners' writing development. *SpringerPlus*. https://doi.org/10.1186/s40064-016-2935-5

Mallikamas, P., & Pongpairoj, N. (2005). Thai learners' knowledge of English collocations. *HKBU Papers in Applied Language Studies*, *5*, 1-28.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Pres.

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

Padilla, J. L., & Leighton, J. P. (2017). Cognitive interviewing and think aloud methods. In B. D. Zumbo, & A. M. Hubley (Eds.), *Understanding and investigating response processes in validation research* (pp. 211–228). Springer.

Rezaee, A. A., Marefat, H., & Saeedakhtar, A. (2015) Symmetrical and asymmetrical scaffolding of L2 collocations in the context of concordancing. *Computer Assisted Language Learning*, *28*(6), 532-549.

Robinson, P. (1995). Review article: Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*(2), 283-331.

Ross, A., & Willson, V. L. (2017). Paired Samples T-Test. In *Basic and advanced statistical tests* (pp. 17-19). Springer.

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.

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

Siyanova-Chanturia, A. (2015). Collocation in beginner learner writing: A longitudinal study. *System*, *53*, 148-160.

Tang, C., & Treffers-Daller, J. (2016). Assessing incidental vocabulary learning by Chinese EFL learners: A test of the involvement load hypothesis. In G. Yu & Y. Jin (Eds.), *Assessing Chinese learners of English* (pp. 121-149). Palgrave Macmillan.

Vu, D. V., & Peters, E. (2021). Vocabulary in English language learning, teaching, and testing in Vietnam. *A review. Education Sciences*, *11*(9), 563. https://doi.org/10.3390/educsci11090563

Vu, D. V., & Peter, E. (2022). A longitudinal study on the effect of mode reading on incidental collocation learning and predictors of learning gains. *TESOL Quarterly*. https://doi.org/10.1002/tesq.3111

Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language, Learning and Technology*, *20*(3), 159-179. http://dx.doi.org/10125/44487

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, *32*(4), 430-449. https://doi.org/10.1093/applin/amr011

Yamashita, J., & Jiang, N. (2010). L1 Influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly*, *44*(4), 647–668.

Yoon, C. (2011). Concordancing in L2 writing class: An overview of research and issues. *Journal of English for Academic Purposes*, *10*, 130-139.

Zhang, X. (2017). Effects of receptive-productive integration tasks and prior knowledge of component words on L2 collocation development. *System*, *66*, 156-167. https://doi.org/10.1016/j.system.2017.03.019)