

OPEN ACCESS

 Manuscript ID:
 EDU-2022-10024458

Volume: 10

Issue: 2

Month: March

Year: 2022

P-ISSN: 2320-2653

E-ISSN: 2582-1334

Received: 03.11.2021

Accepted: 15.02.2022

Published: 01.03.2022

Citation:

Boontam, Punyapa. "The Effect of Teaching English Synonyms through Data-Driven Learning (DDL) on Thai EFL Students' Vocabulary Learning." *Shanlax International Journal of Education*, vol. 10, no. 2, 2022, pp. 80–91.


DOI:

<https://doi.org/10.34293/education.v10i2.4458>



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

The Effect of Teaching English Synonyms through Data-Driven Learning (DDL) on Thai EFL Students' Vocabulary Learning

Punyapa Boontam
Nakhon Sawan Rajabhat University, Thailand
 <https://orcid.org/0000-0002-3390-1150>
Abstract

This paper examines to what extent paper-based data-driven learning (DDL) activities can improve Thai EFL students' vocabulary learning concerning three English synonymous adjectives, which are naughty, disobedient and rebellious, and explores the attitudes of the participants towards learning through the DDL activities. This study was a one-group pre-test post-test design conducted over a four-week period. The convenience-sampled participants were 30 Thai EFL students studying in the Faculty of Education, majoring in English. There were three research instruments in this study including pre/post-test, Thai version of attitudes questionnaire and semi-structured interviews. The instructional treatment, the paper-based DDL materials, was designed based on the criteria of meanings, collocations, formality and grammatical patterns. The corpus data representing each synonymous adjective was drawn from The British National Corpus (BNC) with the assistance of the online tool Sketch Engine (SkE). Using paired-sample t-test, the results indicate a statistically significant difference in the mean scores ($p < 0.001$) suggesting that their vocabulary knowledge had been developed after learning through DDL activities. Furthermore, the participants had positive attitudes towards learning through DDL as they found it enjoyable and effective in learning English. Thus, the findings of the study have a profound impact in the fields of English language teaching as it is clearly seen that corpus data serves as a useful tool to explore the authentic language patterns and help enhance learners' metalinguistic awareness.

Keywords: Data-Driven Learning, Corpus Linguistics, English Synonyms, Vocabulary Learning

Introduction

In a digital technology-driven age, it is undeniable to say that learning English is far more convenient and accessible for ESL/EFL learners around the world. Owing to the digital technology, which is both available online and offline, it could enable students to have more exposure to a number of English learning resources. Precisely, English language skills play a vital role in the 21st century that could help students achieve their goals and lead to career opportunities as English serves as a tool that enables them to communicate effectively with people from other countries (Kirkpatrick, 2007). Nevertheless, among four main skills of English, vocabulary is considered as the most essential skill because lexical competence and lexical knowledge are regarded as the foundation for learners to comprehend both written and spoken languages and lead to successful communication (Nation, 1994; Ling, 2005; Alquahtani, 2015).

In the same vein, the lack of vocabulary knowledge can limit the number of sentences that they would like to produce and consequently lead to learners' frustration (Caro & Mendinueta, 2017). Considering this problem, ESL/EFL learners tend to experience the problem on lexical variety when they are assigned to do the writing task in English (Ferris, 2009).

Hence, the vocabulary knowledge can enhance learner's communicative skills and self-confidence in speaking English. Nonetheless, one of the effective methods for teaching vocabulary is using synonyms. Investigating synonyms can help enhance students' metalinguistic knowledge, for example, part of speech, register, and collocations (Shaw, 2011). Liu (2010) also explained that learners with the ability to use various synonyms would be able to convey their message effectively and rarely fail to interpret the meaning. However, there is a concern when teaching synonyms to students known as thesaurus errors (Ferris, 2009). This problem tends to occur when learners apply synonyms in the wrong context. Therefore, teachers need to develop students' metalinguistic awareness by teaching them register and collocational differences of synonyms. Considering synonyms usage, knowing register of the words is helpful for students to distinguish the different usage of the synonyms, such as the differences between spoken and written language (Shaw, 2011). Therefore, ESL/EFL teachers and learners should be aware that some synonyms cannot be used interchangeably in some particular contexts. In order to differentiate the target synonyms, the corpus-based analysis could be applied when teaching and learning synonyms as it is regarded as a useful and reliable approach. Currently, the incorporation of corpus data or corpus consultation in designing and writing language teaching materials like grammar books or dictionaries has gained its popularity because these materials represent authentic use of the language (Varley, 2008).

Despite the importance of English, it is implicitly considered as one of the most struggling subjects in Thailand as they tend to encounter difficulties when learning English, especially when it comes to speaking in real-life communication due to their limited vocabulary knowledge (Sasum & Weeks, 2018). There are many possible causes affecting students' limited vocabulary knowledge such as the lack of the implementation of authentic teaching material and the conventional vocabulary teaching method (e.g. memorization and one-to-one translation method). Nation (2001) stated that vocabulary knowledge and learning cannot be acquired by only relying on remembering the

definition of a word. As such, data-driven learning (DDL), the inductive approach driven by the corpus information (Johns, 1986), should be taken into consideration as it is considered as a useful method in developing EFL learners' vocabulary knowledge. Unlike other traditional approaches, DDL learners are required to generalise the language patterns by themselves through analysing the concordance lines which reflects the authentic language use. The incorporation of corpora could also enhance the metalinguistic awareness of learners, for example, parts of speech, collocations, alternative meanings, spelling, specific uses, and register related to the context (Liu, 2013; Shaw, 2011; Nation, 2001). There have been some studies on using corpus data in vocabulary teaching. For example, Varley (2008) indicated that corpus consultation plays a positive role in vocabulary acquisition and increases the students' awareness of lexico-grammatical usage. Kartal and Yangineksi (2018) found that concordance lines served as an effective tool for Turkish EFL student in learning collocations, especially on the production of verb-noun collocations. Interestingly, Yaemtui and Phoocharoensil (2019) also found that DDL can enhance the collocational knowledge of both low-proficiency and high-proficiency Thai EFL students. Despite the benefits of corpus-driven approach, little is known about teaching synonyms through data-driven learning. Due to the research gap and problems mentioned above, the researcher was motivated to carry out the present study in order to examine the effect of DDL approach on teaching English synonyms for Thai EFL students.

The present study aims to answer the following research questions:

- To what extent does data-driven learning improve Thai EFL learners' vocabulary knowledge?
- What attitudes towards English vocabulary learning do learners in this DDL class hold?

Literature Review

Corpus Linguistics

O'Keeffe, McCarthy and Carter (2007) have defined a corpus as "a collection of texts, written or spoken, which is stored on a computer" (p.1). Therefore, corpus linguistics is the study of language through investigating the authentic data

in naturally-occurring context which is beneficial for language learners in terms of authenticity, speed and reliability (Sinclair, 1991; Lindquist, 2009). In order to explore the corpus data, concordance lines, a list of sentences showing an occurrence of the target word in a particular context (Lindquist, 2009), need to be analysed systematically. Additionally, the presentation of corpus data through Key-Word-In-Context (KWIC) concordance lines in which the search word or phrase called the ‘node’ is presented in the centre of the line can help enhance learners’ awareness through investigating the lexicogrammatical patterns (Sripicharn, 2002). As such, the learners can look at words or phrases which precede or follow the node word. Likewise, learners can investigate the most frequent collocations of the target words which occur in a specific context by looking at the frequency figures; as shown in Table 1.

Table 1: First 10 Most Frequent Collocates of Naughty in BNC

	Frequency	All	%	MI
Boy	69	12560	0.55	6.64
Being	42	86636	0.05	3.14
Naughty	36	662	5.44	9.95
Children	27	45334	0.06	3.44
Child	23	23486	0.10	4.15
Bit	21	26204	0.08	3.86
Girl	20	14383	0.14	4.66
Boys	18	7667	0.23	5.42
Nice	11	12588	0.09	3.99
School Boy	8	496	1.61	8.20

As can be seen in Table 1, it is likely that the word naughty tends to co-occur with the noun collocates that are equivalent to children, such as boy, children, child, girl, boys and schoolboy. These findings could help EFL/ESL learners to understand the contextual usage of the word effectively and lead them to the further investigation of language patterns as this could shed the light on other new words or grammar rules that they have never experienced before (Jones & Waller, 2015).

Data-Driven Learning (DDL)

Tim Johns initially introduced Data-driven learning (DDL) in the 1980s. This inductive approach allows learners to formulate their own language patterns and rules by investigating the authentic language input through concordance lines (Johns, 1991). In other words, DDL could be described as the corpus-designed activities (Bennett, 2010) as the use of corpora is applied into teaching directly and learners are provided with hand-on experience in analysing the actual data. Moreover, corpus-based/ driven learning is regarded as the excellent discovery learning opportunities for learners (Liu, 2013) as it can promote learner autonomy (Sripicharn, 2002), and enhance their vocabulary learning retention (Cobb, 1997). It is noted that learners become the “language detectives” during DDL classroom as Johns (1997) propounded the idea that “every student a Sherlock Holmes”. To elaborate, the roles of DDL learners are both a language learner and a language researcher.

Teaching Vocabulary through DDL

Read (2010) proposes that corpus analysis and vocabulary knowledge are naturally associated. In other words, corpus linguistics has been a great influence on vocabulary learning (Nation, 2001; Schmitt, 2000). However, using corpus-based approach in teaching can be very challenging (Liu, 2013). The teacher can have students work in groups, as group work has been found to be effective in corpus-based learning (Liu & Jiang, 2009). Similarly, Hunston (2002) suggests that low-proficiency learners should experience DDL through *teacher-led* activities, on the other hand, high-proficiency learners should be taught through discovery learning or learner-led learning. Moreover, DDL teachers should perform as a facilitator throughout the learning process to help learners overcome the difficulties in manipulating actual data and draw a proper conclusion from it (Chambers, 2010). It is also suggested that signposting and guiding questions should be frequently used by the teachers in order to raise the students’ awareness (Flowerdew, 2012). Interestingly, instead of providing learners with the direct access to immediate hand-on concordance lines, paper-based DDL materials would be more

beneficial to beginning or low-proficient learners who have less experiences and limited background knowledge in discovering the language patterns effectively (Kirschner, Sweller & Clark, 2006). In order to help raise learners' consciousness of the inter-lingual differences in the words being learned, translation task is very helpful for learning lexical items whose meaning and usage in L2 differ noticeably from those in learners' L1 (Laufer & Girsai, 2008; Nation, 2001, as cited in Liu, 2013).

Corpus-based Study on English Synonyms

Oxford Advanced Learner's Dictionary (2005) defines synonym as "a word or expression that has the same or nearly the same meaning as another in the same language" (p. 1557), such as big and large. Synonyms can be categorised as strict synonym and loose synonym. Regarding strict synonyms, two words can be used interchangeably, meaning that one could substitute another with no effects on meanings and language patterns. In contrast, loose synonyms may share some overlapping meanings and they cannot replace each other in some contexts (Jackson & Amvela, 2000). Hence, loose synonyms are commonly found in English whereas strict synonyms are almost non-exist (Palmer, 1997).

However, this definition given above is ambiguous since some synonymous words cannot be used interchangeably and some convey different connotations (Gu, 2017). For instance, Chung (2011) found that synonyms create and produce shares some senses of meaning, but the corpus data drawn from Brown Corpus and Frown Corpus reveals that create seems to co-occur with things required a high level of creativity whereas produce tends to co-exist with objects required a low level of creativity. Thus, corpus-based analysis serves as a tool to analyse the synonyms systematically according to register difference, frequency and collocations (Biber, Conrad & Reppen, 2006; Moon, 2010). These criteria can be used to distinguish synonyms effectively. Dealing with genre difference, degree of formality could be revealed here. For example, if a word is frequently used in an academic context, it can be indicated that that word mostly occurs in formal context. Obtained from COCA corpus, Cai (2012) found that great is used more frequently than other six synonymous

adjectives: *awesome, excellent, fabulous, fantastic, terrific, and wonderful*. Interestingly, it is evident that *fabulous, fantastic, great, terrific and wonderful* mostly occur in spoken genre, whereas *awesome and excellent* are commonly used in magazines. The findings suggest that these synonyms are common in informal context. In terms of collocations, different collocates that co-exist with the target synonyms can implicitly identify whether a set of synonyms can substitute each other or not. Sumintana and Tawilapakul (2019) investigated the synonyms cure and heal, using the corpus data obtained from COCA. It was found that cure is frequently followed by disease and problems; e.g. *cancer and disease*, while *heal* frequently co-occurs with emotional injuries and something abstract; e.g. *wounds and people*. Hence, they are loose synonyms.

Lastly, grammatical patterns can unveil the similar or different syntactic pattern of the synonyms which can help to distinguish the strict and loose synonyms. Gu (2017) examined the synonyms obtain and gain through BNC corpus. They found that the syntactic patterns of these two words are similar in the pattern of "verb + object noun" and "verb + adverb". Nonetheless, the result shows that obtain is more frequently used in the passive voice.

In conclusion, to truly understand the semantic differences between two synonyms, we will need to read and analyse the concordance lines of the result tokens (Liu, 2013). Although there are a lot of corpus-based analysis on English synonyms, still not many research studies conducted on the application of such findings in developing an innovation in English Language Teaching. Thus, this is considered as a research gap that the researcher aimed to bridge.

Previous Studies

Jafarpour, Hashemian and Alipour (2013) investigated the effectiveness of corpus-based approach toward teaching collocation of near-synonymous pairs. They found the participants who were taught through corpus-based approach performed significantly better than those who were taught by the traditional approach. The results suggested that the traditional approach could not improve the collocation knowledge of L2 learners. Similarly, Soruç and Tekin (2017) found that DDL

method can develop the vocabulary knowledge of the students as the results shows that those who were taught through DDL gained significantly higher scores than those who were taught through traditional instruction on both the immediate and the delayed posttest. The students also had positive attitudes towards DDL as they described DDL as “freedom,” “comfortable,” “relaxing,” “technological,” and “feeling myself responsible”. In the same vein, Choksuansup and Tangpijaikul (2017) found that Thai EFL students who studied through concordancing-based teaching outperformed those who were given traditional vocabulary exercises in the vocabulary retention test and had more positive effects on their vocabulary learning. According to the reviewed literature, it can be concluded that the implementation of corpus-based analysis of English synonyms to teaching is still limited.

Methodology

Research Design

A one-group pre-test post-test design was employed in the study. The instructional period lasted four weeks in which the participants were taught through DDL activities once a week for two hours. A pre-test related to the target synonyms was given to the participants before the instructional period. A post-test with the same test tasks as the pre-test was administered immediately after the intervention. All of the participants were asked to complete the Thai version of 5-point Likert scale attitudes questionnaire adapted from Phoocharoensil (2012) to examine their attitude towards learning English vocabulary through DDL. Lastly, in order to triangulate the data, five randomly-selected participants were interviewed in their L1 Thai language to yield exact insightful information.

Participants

Before conducting the research, the consent forms were given to the participants in order to gain their permission and the researcher was allowed to conduct this study with them. The convenience sampling was used to recruit the participants from the second-year EFL students in the Faculty of Education, majoring in English at a state university in Thailand. Firstly, there were 55 students whose L1 language background was Thai in the class. Even

though they are English major students, they still possess low levels of English proficiency. In order to control the participants’ English proficiency, a free online vocabulary level test was used to select the participants (<https://www.oxfordonlineenglish.com/english-level-test/vocabulary>). The scores of the vocabulary placement test are corresponding to the levels of the Common European Framework of Reference for Languages (CEFR). Only A2 learners were included in the study to reassure that elementary or beginning learners are capable of learning English through paper-based DDL. The proficiency test showed that only 30 participants were rated as A2; thus, they were recruited to participant in the study.

Data Collection

Instruments. There were three research instruments in this study including pre/post-test, Thai version of attitudes questionnaire and semi-structured interviews. The pre/post-test was divided into three parts, which were 15 items of gap-filling, five items of grammaticality judgement test with English-Thai translation and three items of sentence building. The total score was 40. The questionnaire (adapted from Phoocharoensil, 2012) consists of 12 statements of a 5-point Likert scale and four open-ended questions.

Instructional Treatment-Innovation. The instructional treatment was paper-based DDL materials of three synonymous adjectives which were naughty, rebellious, and disobedient. The tasks included in these corpus-informed materials were designed based on the criteria of meanings, collocations, formality and grammatical patterns which were considered as the criteria for synonym distinction (Biber, Conrad & Reppen, 2006; Moon, 2010). These synonyms are appeared to share some similar meanings in Thai; as a result, there might be times that Thai EFL learners use these English synonymous adjectives interchangeably and end up with producing ungrammatical sentences due to thesaurus errors (Ferris, 2009). Therefore, analysing the language patterns of the target synonyms occurred in different contexts through concordance lines can enlighten learners to notice the different usages of the given words. The corpus data representing each synonymous adjective was drawn from The British

National Corpus (BNC) with the assistance of the online tool Sketch Engine (SkE) to construct the innovation in this study. The paper-based corpus-informed materials were developed by the researcher with the careful selection of concordance lines and activities according to the reviewed literature (see Appendix). Sixty concordance lines per each synonym were given for learners in order to increase the exposure to language patterns and provide them with sufficient data to generalise the rules effectively.

During the learning process, the students had to analyse the provided concordance lines and complete the tasks in groups with teacher’s assistance. The researcher also performed as a teacher in this study. The researcher explained how to analyse the concordance lines and complete the tasks step-by-step to the students in their L1, Thai, in order to help them cope with difficulties that might cause during learning. After their discovery, the students in each group had to share and present their ideas to the class. The teacher concluded their findings to prevent the overgeneralisation of the discovered rules. The frequency of collocates and register of the synonyms found in the corpus-based analysis were shown to them at the end of their investigation to confirm their discovered language patterns.

Data Analysis

In order to examine the effectiveness of using DDL in teaching synonyms, the paired samples t-test was used to compare the means of pre-test and post-test. Moreover, the results from the 5-point Likert scales questionnaire were analysed and triangulated with the interview results to elicit the deeper concerning the attitudes of the learners towards learning English through paper-based DDL activities.

Results

The present study yields instructive results divided into three parts as follows:

Effectiveness of Data-Driven Learning (DDL) on Vocabulary Learning

The effectiveness of the innovation, paper-based DDL, on vocabulary learning was evaluated by the difference of the pre-and post-test mean scores. The mean scores of the participants increased from 12.07 (SD = 1.95) in the pretest to 16.07 (SD = 1.39) in the posttest. To illustrate, the higher mean score reveals that the students’ vocabulary knowledge concerning the three synonyms has been developed after learning through DDL.

Table 2: Paired Sample T-Test of Overall Mean Scores

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Vocabulary Knowledge Scores	Pretest - Posttest	-4.000	1.983	.362	-4.74	-3.260	-11.050	29	.000

Note. * $p < .05$.

As can be seen in Table 2, using paired sample t-test, the p value was $<.001$. Thus, the result implies that there was a statistically significant difference between the pre-test and the post-test at a significant level of 0.5. Thus, it is evident that the lexical knowledge of the students has been significantly enhanced.

Considering the production of the students on vocabulary knowledge, the sentences produced by the participants in the sentence building task from pre- and post-test were compared in order to determine


the students’ development regarding the use of the target synonyms including *naughty*, *rebellious* and *disobedient*. Interestingly, it was found that their writing performance regarding the use of learnt synonyms in the correct context had been improved after the intervention because most of them could describe the pictures by using the target synonyms more appropriately in the posttest. Notably, the majority of the participants could produce more grammatical and meaningful sentences, suggesting a variety of vocabulary use and sentence structures.

This task reveals that the concordance lines provided in the paper-based DDL materials play a crucial role in reinforcing learners to produce authentic language.

Table 3 shows some produced sentences from the pretest and the posttest by the participants. For example, S2 used *rebellious* as a verb which was ungrammatical in the pretest, but she/he could use *rebellious* as an adjective to modify the noun, teens, correctly in the posttest. It is interesting to point out

that the variety of vocabulary used by the students was also improved since S2 used phrasal verb, *call out*, which was found in the concordance line in the DDL worksheet. Therefore, it is a strong evidence to show that the students can develop their lexicogrammatical awareness and metalinguistic awareness in the sense that their produced sentences tend to be more accurate and natural.

Table 3: Examples of Sentences Produced by the Participants in the Tests

Participant	Test Item No.2	Pretest	Posttest
S1		Look at them! Their dressing are very <i>rebellious</i> . They have to respect the host.	Look at those <i>rebellious couple</i> . No one tells them about manner.
S2		John and Jane walk in the street and they see one guy and one girl who <i>rebellious</i> they look so weird.	Two old generation look at two <i>rebellious</i> teens that want to call out for their punk style.

Attitudes of Students towards Data-Driven Learning (DDL)

In order to explore the students' attitudes towards

learning through DDL, the questionnaire and interview results were analysed.

Table 4: Students' Attitudes towards DDL Method

Category	Statements	Mean	SD
Satisfaction	I think that learning though DDL method is enjoyable and not dull.	4.33	0.62
	I find DDL method more challenging than other traditional methods that I have experienced before.	4.83	0.37
	Learning though DDL method is difficult for me.	3.33	1.03
	I enjoy discovering the language patterns or rules by myself.	4.08	1.04
	In the future, I want to learn through DDL method in other English lessons.	4.75	0.43
Content	I understand the lesson better in DDL class.	4.58	0.64
	After studying through DDL method, I could retain the learnt vocabulary in long-term memory.	4.67	0.47
	Throughout the DDL activities, I was encouraged to actively think and express my idea.	4.33	0.62
	Investigating the concordance lines helps me obtain new vocabulary, grammatical patterns and new sentence structures.	4.50	0.50
	Scaffolding techniques used by the teacher can enhance my understanding.	4.83	0.37
	Learning through DDL can foster the collaborative skill.	4.50	0.76
	I think that DDL can promote autonomy in learning and encourage me to self-study outside the class.	4.58	0.64
OVERALL		4.44	0.63

Note. n = 30.

Table 4 reveals that most of the participants have a positive attitude towards the DDL method. Particularly, the majority of the participants strongly agreed that DDL was fun and challenging (x=4.33)

servicing as a useful tool to learn English as it could better their understanding (x=4.58), enhance vocabulary retention (x=4.67) and promote active learning (x=4.33), collaborative skill (x=4.50) and

autonomy ($x=4.58$). It is also vital to be noted that the almost all of the participants tend to feel eager to learn other English lessons through DDL method ($x=4.75$). On the other hand, some of the students found DDL method difficult to learn. Regarding their answers from the open-ended questions in the questionnaire, they stated that DDL might be a little bit difficult for low proficiency learners due to the limited vocabulary knowledge. However, they could overcome the difficulties by consulting dictionary, discussing with their friends and asking the teacher. Thus, the answers from the open-ended questions were corresponding to the questionnaire result that the use of scaffolding techniques by the teacher during the class enhanced their understanding ($x=4.83$).

Furthermore, the findings above were consistent with the interview results, as the interviewees found analysing the concordance lines beneficial for them.

- S1: *“DDL provides a variety of sentences that could help me understand how to use the given synonyms in different contexts”*
- S2: *“I think that it is like killing two birds with one stone. You can learn vocabulary and grammar at the same time.”*
- S3: *“It could create a positive atmosphere for learning that we can share and discuss the findings with our friends”*
- S4: *“It is better than waiting for the teacher to spoon-feed me”*
- S5: *“The teacher is important in guiding and helping us throughout the investigation to make sure that we are on the right track”*

In conclusion, the interview and questionnaire results reflect that the students had positive attitudes towards learning through DDL.

Discussion

RQ1: To What Extent does Data-Driven Learning improve Thai EFL Learners' Vocabulary Knowledge?

Paper-based DDL is considered as a useful tool in developing the students' vocabulary knowledge which is consistent with other previous studies. To some extent, results such as this align with those found in Jafarpour, Hashemian and Alipour (2013) and Choksuansup and Tangpijaiikul (2017). The results

show that DDL activities can promote collaboration in the class through working in groups (Liu & Jiang, 2009), enhance vocabulary retention (Cobb, 1997), foster active learning and critical thinking (Liu, 2013), and increase learners' autonomy (Sripicharn, 2002). With regard to the effectiveness of DDL on developing learners' vocabulary knowledge, it is noteworthy that concordance lines drawn from the BNC play a vital role in increasing learners' lexical variety as they could reflect authentic language patterns (Johns, 1991) and enhance learners' lexicogrammatical awareness (Sripicharn, 2002; Varley, 2008). It is evident that the participants can produce more meaningful and grammatical sentences after learning through DDL because corpus-influenced materials lead them to the investigation of new vocabulary and grammatical patterns that they have never experienced before (Jones & Waller, 2015). Even though there are a number of studies conducted to confirm the effectiveness of data-driven learning method, little is known about English synonyms teaching through DDL. Thus, this study reveals a salient result that could bridge the research gap by advocating the incorporation of corpus-based analysis of English synonyms in vocabulary teaching. Particularly, in terms of paper-based corpus-informed material on the target synonyms, the students had a chance to work on the language by themselves as a detective. Owing to the benefits of the innovation, the students were exposed to the register difference, frequency and collocations which helped them analyse the synonyms systematically (Biber, Conrad & Reppen, 2006; Moon, 2010) and raise their metalinguistic awareness (Liu, 2013; Shaw, 2011; Nation, 2001). By following the criteria, students will be able to distinguish synonyms; as a result, their lexical knowledge and lexical competence will be enhanced (Shahzadi, Asghar & Javed, 2019; Liu, 2010; Phoocharoensil, 2010) and thesaurus errors will be decreased (Ferris, 2009).

Moreover, the corpus-based data-driven learning conducted in the present study was evaluated as an effective innovation in teaching new English vocabulary for EFL/ESL students for several reasons. First, in terms of novelty, this paper-based DDL activity is considered as a new method for Thai EFL learners and it is also not well-known in Thai EFL

classroom since the participants have experienced learning through DDL for the first time. Secondly, this method could change the teaching style and the way teachers develop their teaching material as the study could shed the light on the effectiveness of using DDL in the classroom and raise the awareness of Thai EFL teachers to apply corpus-based data into their teaching and designing materials. Lastly, regarding development, the implementation of DDL in the English classroom could explicitly enhance learners' understanding, improve their English skills and create more positive atmosphere in the classroom since almost all of the participants strongly agree that DDL activities should be more advocated in Thai EFL classroom.

RQ2: What are Thai EFL Students' Attitudes towards DDL Approach?

The results from the questionnaire and the interview indicate that the students had positive attitude towards the DDL method which are concur with other previous studies (Soruç & Tekin, 2017). It is noticeable that the teacher plays such a crucial role in precipitating the success of DDL (Hunston, 2002; Chambers, 2010; Flowerdew, 2012) since the majority of the students stated that the use of scaffolding techniques by the teacher during the class enhanced their understanding. Similar words, the role of change agents should be taken into consideration when implementing the DDL activities because it is one of the crucial factors that plays a critical role on the success of the innovation in the present study. To elaborate, the researcher taught the students through the teacher-led activities to facilitate them who were considered as beginning DDL learners (Hunston, 2002) and assigned them to work in groups (Liu & Jiang, 2009). In the same vein, signposting and guiding questions were frequently used throughout the discovery to raise students' awareness by the teacher (Flowerdew, 2012). Hence, a change agent should be trained and fully understand the nature of DDL and the benefits of corpus data in order to help learners overcome the difficulties and draw a conclusion from the given corpus data as most effectively as possible (Chambers, 2010).

Conclusion

Corpus-based data-driven learning in the present study is considered as a useful innovation in teaching and learning new vocabulary for EFL students as it might help shed light across L2 education. It should be noticed that the findings of this study can have a profound impact in the fields of English language teaching as it is clearly seen that corpus data serves as a useful tool to explore the authentic language patterns and help enhance learners' metalinguistic awareness. As English teachers, we should apply this knowledge to vocabulary teaching and use corpus evidence, i.e. concordance lines and frequency, to design corpus-informed teaching materials as it has been discovered that corpora can provide some additional data that is non-existed in the dictionary. In addition, corpus-based teaching approaches should be fostered in the 21st century English classroom where we are driven by data in the digital age.

Implications and Recommendations

To some extent, the generalisability of the results such as these might be limited due to the insufficiency period of instruction and the small number of participants. However, this study could serve as a preliminary research that helps researchers and practitioners gain better understanding in corpus-based DDL method on vocabulary teaching. Therefore, these limitations should be taken into consideration for the further studies in order to strengthen the validity and reliability of the research findings. In terms of practical implications, a teacher should serve as a facilitator and lead the students, a language detective, to the successful discovery. Prior to the instruction, a teacher should be well-prepared in terms of designing the paper-based DDL materials and planning the activities. Especially for Thai EFL context, the teachers should always be aware that the role of students will be totally changed as DDL learners are no longer passive. Generally speaking, this could be another challenge for Thai EFL teachers to transform the traditional classroom to be the active discovery learning classroom. Likewise, in terms of designing corpus-influenced material or paper-based DDL worksheet, teachers need to carefully select the concordance lines from the online corpus sources (e.g. BNC and COCA) that are suitable for the English

proficiency or background knowledge of the students in order to minimise the difficulties that they might encounter during learning. So far, a lot of researchers have conducted studies on the effectiveness of DDL on students' writing performance or tests, little is known about oral performance. Hence, for future studies, it is recommended that speaking ability or oral production of learners should also be evaluated in order to examine the effect of DDL in other perspectives. As a final note, it should be added that a pilot teaching should be conducted prior to the instruction in real setting as the teachers can change, adapt or develop their teaching strategies and DDL material to benefit the learners as much as possible.

References

- Alqahtani, Mofareh. "The Importance of Vocabulary in Language Learning and how to be Taught." *International Journal of Teaching and Education*, vol. 3, no. 3, 2015, pp. 21-34.
- Bennett, Gena R. *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*. University of Michigan Press, 2010.
- Biber, Douglas, et al. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press, 2006.
- Cai, Jiutian. *Is it "Great" Enough? A Corpus-based Study of "Great" and its Near Synonyms*. Ball State University, 2012.
- Caro, Keiby, and Nayibe Rosado Mendinueta. "Lexis, Lexical Competence and Lexical Knowledge: A Review." *Journal of Language Teaching and Research*, vol. 8, no. 2, 2017, pp. 205-13.
- Chambers, Angela. "What is Data-driven Learning?." *The Routledge Handbook of Corpus Linguistics*, edited by Anne O'Keefe and Michael McCarthy, Routledge, 2010.
- Choksuanup, Nuchara, and Montri Tangpijaikul. "The Effect of Concordancing-based Teaching on Vocabulary Learning and Retention on Thai EFL Learners." *Chophayom Journal*, vol. 28, no. 3, 2017, pp. 11-19.
- Chung, Siaw-Fong. "A Corpus-based Analysis of "Create" and "Produce"." *Chang Gung Journal of Humanities and Social Sciences*, vol. 4, no. 2, 2011, pp. 399-425.
- Cobb, Tom. "Is there any Measurable Learning from Hands-on Concordancing?" *System*, vol. 25, no. 3, 1997, pp. 301-15.
- Ferris, Dana. *Teaching College Writing to Diverse Students Populations*. University of Michigan Press, 2009.
- Flowerdew, Lynne. *Corpora and Language Education*. Palgrave Macmillan, 2012.
- Gu, Bing-jie. "Corpus-based Study of Two Synonyms-Obtain and Gain." *Sino-US English Teaching*, vol. 14, no. 8, 2017, pp. 511-22.
- Jackson, Howard, and Etienne Amvela. *Words, Meaning and Vocabulary: An Introduction to Modern English Lexicology*. Athenaum Press, 2000.
- Jafarpour, Ali Akbar, et al. "A Corpus-based Approach toward Teaching Collocation of Synonyms." *Theory and Practice in Language Studies*, vol. 3, no. 1, 2013, pp. 51-60.
- Johns, Tim. "Micro-concord: A Language Learner's Research Tool." *System*, vol. 14, no. 2, 1986, pp. 151-62.
- Johns, T. "Should you be Persuaded: Two Examples of Data-driven Learning." *ELR Journal*, vol. 4, 1991.
- Johns, T. *Contexts: The background, Development and Trialling of a Concordance-based CALL Program*. Routledge, 1997.
- Jones, Christian, and Daniel Waller. *Corpus Linguistics for Grammar: A Guide for Research*. Routledge, 2015.
- Kartal, Galip, and Gonca Yangineksi. "The Effects of Using Corpus Tools on EFL Student Teachers' Learning and Production of Verb-Noun Collocations." *PASAA: Journal of Language Teaching and Learning in Thailand*, vol. 55, 2018, pp. 100-25.
- Kirkpatrick, Andy. *World Englishes Hardback with Audio CD: Implications for International Communication and English Language Teaching*. Cambridge University Press, 2007.
- Kirschner, Paul A., et al. "Why Minimal Guidance during Instruction does not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-based, Experiential, and Inquiry-

- based Teaching.” *Educational Psychologist*, vol. 41, no. 2, 2006, pp. 75-86.
- Laufer, Batia, and Nany Girsai. “Form-focused Instruction in Second Language Vocabulary Learning: A Case for Contrastive Analysis and Translation.” *Applied Linguistics*, vol. 29, no. 4, 2008, pp. 694-716.
- Lindquist, Hans. *Corpus Linguistics and the Description of English*. Edinburgh University Press, 2009.
- Ling, Lai Yu. *Teaching Vocabulary Learning Strategies: Awareness, Beliefs, and Practices. A Survey of Taiwanese EFL Senior High School Teachers*. University of Essex, 2005.
- Liu, Dilin. “Is it a Chief, Main, Major, Primary, or Principal Concern? A Corpus-based Behavioral Profile Study of the Near-Synonyms.” *International Journal of Corpus Linguistics*, vol. 15, no. 1, 2010, pp. 56-87.
- Liu, Dilin. “Using Corpora to Help Teach Difficult-to-distinguish English Words.” *English Teaching*, vol. 68, no. 3, 2013, pp. 27-50.
- Liu, Dilin, and Ping Jiang. “Using a Corpus-Based Lexicogrammatical Approach to Grammar Instruction in EFL and ESL Contexts.” *The Modern Language Journal*, vol. 93, no. 1, 2009, pp. 61-78.
- Moon, Rosamund. “What can a Corpus Tell us about lexis?” *The Routledge Handbook of Corpus Linguistics*, edited by Anne O’Keeffe and Michael McCarthy, Routledge, 2010, pp. 345-58.
- Nation, I.S.P. *New Ways in Teaching Vocabulary*. TESOL, 1994.
- Nation, I.S.P. *Learning Vocabulary in Another Language*. Cambridge University Press, 2001.
- O’Keeffe, Anne, et al. *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge University Press, 2007.
- Oxford Advanced Learner’s Dictionary*. Oxford University Press, 2005.
- Palmer, F.R. *Semantics*. Cambridge University Press, 1997.
- Phoocharoensil, Supakorn. “Language Corpora for EFL Teachers: An Exploration of English Grammar through Concordance Lines.” *Procedia - Social and Behavioral Sciences*, vol. 64, 2012, pp. 507-14.
- Read, J. “Methodological Issues in Research on Vocabulary Learning.” *19th International Symposium on English Teaching*, 2010, pp. 63-73.
- Sasum, Sunisa, and Bruce Weeks. “Why Some Thai Students cannot Speak English Fluently?” *RSU International Research Conference*, 2018, pp. 361-67.
- Schmitt, Norbert. *Vocabulary in Language Teaching*. Cambridge University Press, 2000.
- Shahzadi, Amina, et al. “Effectiveness of Corpus in Teaching English Synonyms.” *Corporum: Journal of Corpus Linguistics*, vol. 2, no. 1, 2019, pp. 51-65.
- Shaw, Erin Margaret. *Teaching Vocabulary through Data-Driven Learning*. Brigham Young University, 2011.
- Sinclair, John. *Corpus, Concordance, Collocation*. Oxford University Press, 1991.
- Soruç, Adem, and Bilal Tekin. “Vocabulary Learning through Data-driven Learning in an English as a Second Language Setting.” *Educational Sciences: Theory & Practice*, vol. 17, 2017, pp. 1811-32.
- Sripicharn, Passapong. “Introducing Data-Driven Learning: Theory, Practice, and Evaluation.” *Journal of Liberal Arts*, vol. 2, no. 2, 2002, pp. 100-11.
- Sumintana, Natapaun, and U. Tawilapakul. “A Corpus-based Study of the English Synonyms Cure and Heal.” *8th LITU International Graduate Conference*, 2019.
- Varley, Steve. “I’ll Just Look that up in the Concordance: Integrating Corpus Consultation into the Language Learning Environment.” *Computer Assisted Language Learning*, vol. 22, no. 2, 2008, pp. 133-52.
- Yaemtui, Wachirapong, and Supakorn Phoocharoensil. “Effectiveness of Data-driven Learning (DDL) on Enhancing High-proficiency and Low-proficiency Thai EFL Undergraduate Students’ Collocational Knowledge.” *Asian EFL Journal*, vol. 23, no. 3.2, 2019.

Appendix

Examples of Paper-Based DDL Teaching Material on the Synonym Rebellious

- a symbol of ‘growing up’. Teenagers like to experiment and to feel rebellious , and problems can arise if they are not given guidance in the use of alcoho
- ular species. Some of his birds look mischievous, others furious or rebellious , some even forlorn. They twist on their perches in positions never
- the little man who is enslaved and craves authority and is at the same time rebellious . ? It is at this point that his analysis starts to become much more sp
- ... One can speculate that this boy might some day join one of the rebellious students who fight the establishment in an astonishingly infantile way, expe
- d the castle, tightlygarrisoned and well-supplied as a vital base against the rebellious Welsh, had little comfort to offer the Lady Percy and her women. F
- other in formidable acres of widow’s black. Isabella was becoming rebellious , said the head teacher, and it took an hour to modify this to fanciful.
- rent. He knows that the West would still not send its troops to help rebellious East Europeans, because Russia is still a superpower. His idea of
- nment did not apply it. In a new and nervous democracy, beset by rebellious soldiers, the ruling Radical party always gave in when pushed by public ser
- h the pressures of competition, rather that she was merely going through a rebellious streak.

“ I was on teams with older girls and if they drank I would f

- f the government, which admitted that it had failed to exert its authority over rebellious troops in the Agadez region who had taken the law into their own hands.
- ommitted to it now. And from somewhere inside she felt a spurt of rebellious determination. ‘I’ll make it work. I will ...’ And th
- ave been a perilous situation, but, clasped in his arms like this, she felt both rebellious and faintly ridiculous. Knowing he was in the right, that she had ta
- /s> It was a classic holiday picture. The two parents with their rebellious adolescent son. Anyone could have guessed the relationship with
- s no problems, but estranged parents using the child to fight their bat ties or rebellious teenagers seeking recognition of their maturity and autonomy against over

Direction: Investigate the given concordance lines and answer the following questions.

1. What noun phrases are followed by the adjective, disobedient?

2. Identify the grammatical patterns.
3. From your analysis, in what context can this word be used? Formal VS Informal?
4. How will you translate this word in Thai language?
5. Your sentence examples using this word.

Author Details

Punyapa Boontam, Nakhon Sawan Rajabhat University, Thailand, **Email ID:** punyapa.b@nsru.ac.th