วารสารอิเล็กทรอนิกส์
ทางการศึกษา

OJED, Vol. 13, No. 2, 2018, pp. 536-550

O J E D

An Online Journal
of Education
http://www.edu.chula.ac.th/ojed

ผลของการเรียนรู้ด้วยวิธีข้อมูลขับเคลื่อนโดยใช้คลังข้อมูลภาษาบนอินเทอร์เน็ตที่มีต่อความรู้ทางไวยากรณ์ของนักศึกษาที่เรียนรายวิชาภาษาอังกฤษสำหรับวิทยาศาสตร์ในมหาวิทยาลัยขอนแก่น

EFFECTS OF DATA-DRIVEN LEARNING USING ONLINE CORPUS TO ENHANCE GRAMMATICAL KNOWLEDGE OF STUDENTS IN ENGLISH FOR SCIENCES COURSE AT KHON KAEN UNIVERSITY

นายอรรถวิทย์ บัวหงษ์ *
Athawit Buahong
ผู้ช่วยศาสตราจารย์ ดร. ปราณภา โหมดหิรัญ **
Asst. Prof. Prannapha Modehiran, Ph.D.

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาผลของการเรียนรู้ และ 2) ประเมินทัศนคติของผู้เรียนต่อการเรียนรู้ไวยากรณ์ด้วยวิธีข้อมูลขับเคลื่อนโดยใช้คลังข้อมูลภาษาบนอินเทอร์เน็ต การวิจัยนี้เป็นการวิจัยแบบทดลองกลุ่มตัวอย่างเป็นนักศึกษาสาขาวิทยาการคอมพิวเตอร์ เทคโนโลยีสารสนเทศ และ ภูมิสารสนเทศศาสตร์ มหาวิทยาลัยขอนแก่น ได้มาจากการเลือกแบบเจาะจง จำนวน 59 คน ซึ่งลงทะเบียนเรียนรายวิชาภาษาอังกฤษสำหรับวิทยาศาสตร์ ภาคฤดูร้อน ปีการศึกษา 2559 แบ่งเป็นกลุ่มทดลอง 29 คนและกลุ่มตัวอย่าง 30 คน โดยกลุ่มทดลองได้รับการสอนโดยใช้การเรียนรู้ด้วยวิธีข้อมูลขับเคลื่อนโดยใช้คลังข้อมูลภาษาบนอินเทอร์เน็ต ในขณะที่กลุ่มควบคุมได้รับการสอนแบบปกติ ทั้งสองกลุ่มได้รับการทดสอบ ไวยากรณ์ก่อนและหลังเรียน เครื่องมือที่ใช้ในการวิจัยได้แก่ แผนการสอนและสื่อการเรียนรู้ด้วยวิธีข้อมูลขับเคลื่อนแบบทดสอบก่อนและหลังการเรียน แบบสอบถาม สถิติที่ใช้ในการวิจัยคือ T-test, Mean และ Standard Deviation ผลการวิจัยพบว่า

1) ค่าคะแนนเฉลี่ยของการทดสอบหลังการใช้วิธีการศึกษาคลังข้อมูลภาษาบนอินเทอร์เน็ตของกลุ่มทดลองสูงกว่าค่าคะแนนของกลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 2) ทัศนคติต่อการใช้วิธีศึกษาคลังข้อมูลภาษาบนอินเทอร์เน็ตต่อความรู้ทางไวยากรณ์ มีค่าเฉลี่ย 3.53 ซึ่งแสดงให้เห็นว่านักศึกษาส่วนใหญ่มีทัศนคติในระดับเห็นด้วยเชิงบวกกับกิจกรรมและสื่อการสอนของการเรียนรู้ด้วยวิธีข้อมูลขับเคลื่อนโดยใช้คลังข้อมูลภาษาบนอินเทอร์เน็ต

* Master's Degree Student, Department of Curriculum and Instruction
Faculty of Education, Chulalongkorn University, Bangkok, Thailand
E-mail Address: athawit@hotmail.com
** Advisor and Assistant Professor, Department of Curriculum and Instruction
Faculty of Education, Chulalongkorn University, Bangkok, Thailand
E-mail Address: Praneemod@gmail.com
ISSN 1905-4491

Abstract

The purposes of this study were to: 1) investigate the effects of data-driven learning using online corpus to enhance students' grammatical knowledge, 2) explore students' opinions towards using data-driven learning using online corpus to enhance grammatical knowledge. This research was of an experimental design. The participants were selected by the purposive sampling method; 59 of them were majoring in Computer Science, Geo Informatics, and Information and Communication Technology and enrolled in English for Sciences, in the summer semester of academic year 2016 at Khon Kaen University. The experimental group, consisting of 29 students, experienced data-driven learning using online corpus while the control group of 30 students experienced conventional teaching. Both groups were tested with the pre-and posttest. The research instruments were lesson plans, worksheets, a pre-test and posttest, and a 5-rating Likert-scale questionnaire with open-ended questions. The data analysis used the statistics of t–test, means and standard deviation. The results showed that: 1) the post-test mean score of the experimental group was higher than the control group at a 0.05 level of significance, and 2) the result of the questionnaire indicated the average of 3.53 which means most students expressed positive agreement towards the data-driven learning activities and materials using online corpus.

คำสำคัญ: คลังข้อมูลภาษาบนอินเทอร์เน็ต / การเรียนรู้ด้วยวิธีข้อมูลขับเคลื่อน
KEYWORDS: ONLINE CORPUS / DATA–DRIVEN LEARNING

Introduction

Of the four strands, namely meaning-focused input, meaning-focused output, language-focus learning, and fluency development in any language instruction, each strand should be treated equally in term of amount of time spent (Nation and Newton, 2009). Regarding the language-focus learning, linguistic elements such as grammar is an aspect of the focus. Nation (2009) noted that grammar should be taught deliberately because communicative and other focus activities are not sufficient for language acquisition.

Correspondingly, grammar plays a substantial role in leading the use and application of language; making a basic structure of complete and meaningful sentences. The role of grammar can take various dimensions and varies by the situation and context. A major difficulty is that the examples of language use are limited to the sentence level and completely context-free which is of aspects of the circumstance of actual language use taken as relevant to meaning (H.G. Widdowson, 2000). Grammar explanation without discourse context fails to provide students with sufficient information about grammar meaning and use (Celce-Murcia, 2000). The alternative way of dealing the mentioned problems of teaching grammar is to use the samples from naturally-occurring written or spoken text which is collected and stored in the corpus database.

From time to time, English language learners are bored when learning grammar in a conventional classroom setting where teachers primarily deliver the instruction by

presenting grammatical rules and the students have to drill the exercise; probably most teachers tend to teach like this as the way they were taught.

With these concerns and the trends of teaching, it has brought about the implication of using corpus into the researcher's class to study the effects of using online corpus-based learning or so-called data-driven learning to enhance the grammatical knowledge. Hence, trying the opposite way by involving students as an active learner would be an alternative: summarizing grammar rules from their own observation, possibly under teachers 'proper guidance (O'Keeffe, McCarthy & Carter, 2007).

However, some scholars have suggested that data-driven learning (DDL) has been conducted and reported to work effectively with advanced students (Conrad, 2005; Hunston, 2002). Besides, a number of previous research conducted with English language major or high proficient students. Therefore, this study aimed to investigate and observe the effects and responses from participants of other disciplines when using online corpus to study grammar, and recommend ways for using it more effectively in the classroom.

Research Questions

1) To what extent does the data-driven learning using online corpus help enhance students' grammatical knowledge?
2) What are students' opinions towards the data-driven learning using online corpus to enhance their grammatical knowledge?

Literature Review

1) Online Corpus

Thank to corpus linguistics and the development of instructional technology, there are many of corpora where real utterances were collected and stored on online database. The corpus shows frequency occurrence for words, expressions, and various grammar structures in the naturally occurring discourse (Nation, 2001). For example, it has been discovered that there are more than 30 possible patterns of conditionals which is not limited to the three classic types known as first, second, and third conditionals (O'Keeffe, McCarthy & Carter, 2007).

In the corpora, concordancer is built-in searching tool for the language instances in a corpus, or a list of KWIC (Key Word in Context) examples, here it is called concordance. Each may come from a different part of the corpus and is not related to the previous or the next lines; however, each concordance line has the option to view the full context.

Corpus has been incorporated in abundant research and teaching material development and the results appeared positive. For example, Aston (2001) said that when students involve themselves in the corpus-based tasks, the learners feel in change of their

own learning, and the motivation is increased. Also, the students who learned the collocation with the help of online concordancer gained more knowledge and had better results in all parts of the test. Daskalovska (2013) as well as Cheng, Warren, and Xun-feng (2003) studied the effects of using corpus with the English major students, it showed that it is possible and worthwhile to add corpus linguistics to the curriculum because it enhances the other subjects learning when students are confronting with the problem of a grammatical point while writing work for other subjects; the students used the corpus as a reference. Interestingly, in Cheng's, few students pointed out that dictionaries may not provide much information for grammar or word reference as a corpus because the corpus reveals the additional functions of words which are not described in a traditional dictionary.

2) Data-driven Learning (DDL)

By integrating the features of corpus and its benefits into the language pedagogy, it can play key roles in learning. Corpus-based approach is interrelated with the principle of Data-Driven Learning. The concept is about the way learners develop an ability to see the patterns in the target language and to generalize about language form and use. Put it in another way, learners explore the language data which is in the form of concordance lines and then they are expected to formulate the rules of the pattern. Here, students are going to learn as a language researcher. The teacher is no longer a resource of knowledge, but a facilitator who provides the guidance to complete the task. It is also believed that what learners can find out for themselves is better remembered than what they are simply told (Ellis, 2003, p.163).

DDL demonstrates similar concept of second language acquisition theory that emphasizes the role of input, noticing, and output in the language learning process (Schmidt, 1990), which allowing learners to test their hypotheses about language use against the data they observe, in turn, noticing is likely to help learners modify their output so that it more closely approximates target language norms, or the conventions of particular disciplines or genres (Aston, 2001; Gavioli, 2001; Meunier, 2002; Zanettin, 2001 cited in Keck & Kim, 2014).

Generally, presentation-practice-production (PPP), a deductive method, has been practiced in schools where the knowledge is presented directly. By adapting and moving to the new idea of involving students in greater degree in raising language awareness; becoming more inductive. The "three Is" (Illustration, Interaction, Induction) has been proposed; illustration means students are searching the input terms examining the data by looking at the chunk of language feature on the corpus, interaction means raising awareness through observation and discussion about the target language structure, and induction means conclusions are drawn and rules are generated by students for target language features. (McCarthy & Carter, 1995, P.217)

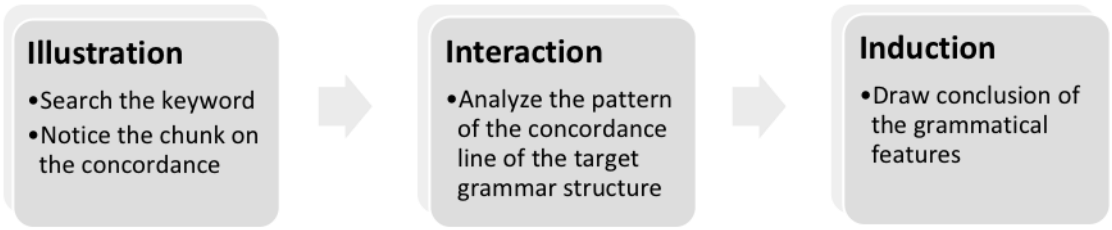

Figure 1: Flow chart displaying three-Is steps proposed by McCarthy & Carter (1995)

Research Methodology

Participants

The participants were 59 students who enrolled in English for Science course at Khon Kaen University, Thailand in the academic year of 2016. This group of students was purposively selected as participants in this study because they have to pass the course to fulfill their degree as required by the curriculum of faculty of Science. In addition, prior to taking this course, they must pass three pre-requisite fundamental English courses. These students were majoring in Computer Science, Information and Communication Technology, and Geo Informatics. Their level of English proficiency was the equivalent regarding the class sectioning placement. These students were divided into two groups; the experiment group for 29 students and the control group for 30 students according to their sectioning enrollment. All of them were informed that they were going to be involved in the study and they were fully aware that the participation had no impact on their course grade.

Research Instruments

There were 3 research instruments in this study; lesson plans with worksheets, pre- and post-test, and questionnaire.

1) *Lesson plans and the worksheets*

Lesson plans and worksheets were designed based on the topics in the book entitled Technical English 3 by David Botany as required in the course (See Table1). Since different treatments were given to the experiment group and control group. The lesson plans and worksheets for experiment group have incorporated the corpus-based or data-driven learning concept and the three Is, Illustration, Interaction, and Induction learning and teaching steps. Together with the online BYU-BNC corpus (http://corpus.byu.edu/bnc/), which is free for educational purpose, has been selected to use in this study. The adapted worksheets were revised by native-English speaking teacher.

Concerning the worksheet of each lesson, beginning with 'illustration', the students worked in group of 3-4 and used the assigned keyword to search the concordance lines in

the corpus on the selection of the academic register to notice the language structure. After that, in the step of 'interaction', responding to the guided questions, students viewed their concordance lines and answered the guided questions with their peers. In this step, they compare the form of the occurrence grammatical patterns and dealing with the task. Finally, in the step 'induction', students shared what the findings are in the group to make conclusion about the form and use of a grammar topic with the facilitation of teacher when in need. On the other hand, the control group will be taught with conventional PPP teaching method on the same topics of grammar.

2) *Pre-test and post-test*

The pre-test and post-test were in the completion format consisted of 30 questions and the both tests were identical. Each lesson contributed to 10 items which marked out of 30. One point was allocated for each question. The test was designed to evaluate the student' grammatical knowledge learned from the lessons. The test aimed at measuring participants' outcomes in the following objectives in Table 1. Regarding the instructions of the test, In lesson 1, two sentences were provided without the relative pronoun and then students had to join them correctly using the correct relative clause concept. In lesson 2, students were asked to write the third conditional sentences with opposite meaning from the situation provided and in lesson 3, situations were given, student had to analyze the incidents and then write the future situation that already completed by using future perfect concept. The vocabulary used in the item questions were from the lesson in the coursebook.

Learning objectives of each lesson
Lesson 1: Relative clause
1. Student will be able to generate the rules or pattern in relative clauses
2. Student will be able to use the correct relative pronoun
3. Student will be able to link two clauses with relative clause.
Lesson 2: Third conditional
1. Student will be able recognize the form of the third conditional.
2. Student will be able to write third conditional sentence for unreal situation in the past
Lesson 3: Future Perfect
1. Student will be able to generate the form of future perfect
2. Student will be able to write future perfect from the provided situations

Table 1: The table displaying objectives of each grammar lesson.

3) Questionnaire

In this current study, the questionnaire was to used explore the participants' opinions toward the data-driven learning using corpus which is employed to enhance their grammatical knowledge. It was adapted from the questionnaire in their research of investigating ESL students' opinions toward language corpus use in L2 writing by Yoon & Hirvela (2004). However, some questions were modified to suit this study. The items of questionnaire divided into 3 aspects regarding students themselves, teachers, and the data-driven learning activities. These questions relating to the quantitative data consisted of Likert-scale questions asking the respondents to rate the degree of agreement on a scale of 1–5 (1= strongly disagree, 2= disagree, 3= neutral, 4= agree, 5= strongly agree). Another part was to inquire the qualitative data; the respondents could voluntarily answer the open-ended questions. Furthermore, in order to avoid misunderstanding of the questions, all questions were translated into Thai language.

These research instruments were validated using the index of item – objective congruence (IOC) by the experts on the degree to which they do or do not measure specific objectives listed by the research instrument designer. The experts were from Khon Kaen University Language Institute, Faculty of Education at Khon Kaen University, and Faculty of Arts at Ubon Ratchathani University. The IOC results for each instrument were: lesson plan and worksheets =0.86, Pre- and posttest = 1, and the Questionnaire = 0.95. which is above 0.5 showing that all research instruments were acceptable and reliable to use for collecting the data. Comments and suggestions were taken into revision such as giving more examples, more concise and clearer instructions as well as changing some unclear test questions. The instruments were piloted with a few students and they commented on the problems about the unclear test instructions and typo errors. In addition, they suggested to provide instructions in Thai which is the native language of the students.

Data Collection Procedure

A week before the treatment, both groups of the participant were tested on their grammatical knowledge on the topics of relative clause, third conditional, and future perfect which is pre-test and later on the post-test. The following week, the participants of the experimental group were told that they would learn the grammar topics in their course book by data-driven learning using the online corpus. In order to answer the first research question regarding to what extent the data-driven learning using online corpus helps enhance students' grammatical knowledge, the post-test was used after the treatment to compare the results.

Both groups of the participants spent an equal amount of time in learning the grammar. The post-test was administered one week after the treatment was completed. The format and the test questions were the same as pre-test but in different order.

To answer the second research questions regarding students' opinions towards data-driven learning; the data was collected using the questionnaire; it consisted of two main parts. The first one was the series of questions regarding DDL learning and the second section was open-ended questions. The administration of questionnaire data collection was after all lessons have been completed.

Data Analysis

Firstly, the Normality test was conducted to find out the normality of distribution of the pre-test and post-test scores of the experimental group and control group. The gathered data from the pre-test and post-test were statistically compared using Paired-sample t-test and the Independent sample t-test. Regarding the Paired-sample t-test, it was employed to investigate whether the effect was shown in each group after learning grammar. Independent sample t-test was used to test whether there was the significant difference between the post-test means score of the experimental group and control group at the 0.05 level. The statistical data from the questionnaires were also calculated and demonstrated in descriptive statistics; means and standard deviation. In addition, the qualitative data from the open-ended questions of the questionnaires were identified and used to support or refute the statistical results from the pre-test and post-test and questionnaires.

Results

At the beginning, the test of Normality was conducted to see the distribution of the scores of the participants before and after learning grammar. Shapiro-Wilk's was selected to report the result because it was more appropriate for small sample sizes (less than 50 samples).

Table 2 : The Test of Normality of the Pre-test and Post-test in both Experimental Group and Control Group Using Shapiro-Wilk

Tests of Normality

Score	Group	Shapiro-Wilk		
		Statistic	df	Sig.
Pre	Experimental	.973	29	.631
	Control	.950	30	.171
Post	Experimental	.945	29	.134
	Control	.938	30	.081

The above table showed that the p-values were greater than 0.05; no significant difference in both groups regarding the pre-test and post-test scores. This means that scores

were normally distributed and indicated that the participants made the same distribution of the test scores. Therefore, the parametric test can be used in the next analysis process.

Regarding the first research question, it was to investigate to what extent the data-driven learning using online corpus helps enhance students' grammatical knowledge. The collected data was analyzed using the following statistics.

Paired-sample t-test was employed to calculate the difference between pre-test and post-test in each group. And then independent samples t-test was conducted to find out which group had more effect after learning the grammar lessons.

Table 3: The Comparison of Pre-and Post-test of the Experimental Group analyzed by a Paired Samples T-test

Test	n	Mean ( )	S.D.	t	df	Sig
Pre	29	17.86	4.04	8.404	28	.000
Post	29	24.59	2.90			

$p < .05$

From the table 3, the result showed that there was significant difference between pre-test and post-test scores in experimental group. The p-value was less than 0.05 and t=8.404.

Table 4: The Comparison of Pre-and Post-test of the Control Group analyzed by a Paired Samples T-test

Test	n	Mean ( )	S.D.	t	df	Sig
Pre	30	15.47	5.19	9.507	29	.000
Post	30	19.47	5.54			

$p < .05$

From the table 4 above, the result showed that there was significant difference of mean scores in control group before and after learning grammar. The p-value was less than 0.05 and t=9.507.

Even though both groups had significant effect after learning grammar as shown in table 3 and table 4, then the independent samples t-test was conducted to calculate the difference of the mean score from both groups in order to find out that which group had more effect after learning the grammar lessons. The result revealed that there was a significant difference between the experimental group and the control group at the 0.05 level; $p < .05$, t=4.470. See Table 5.

Table 5: The Comparison of Mean Scores of the Post-test between Experimental group and Control group analyzed by an Independent Samples T-test

Group	n	Mean ( )	S.D.	t	df	Sig

Experimental	29	24.59	2.90	4.470	44.085	0.002
Control	30	19.47	5.54			

$p < .05$

According to table 5, it can be interpreted that the data-driven learning using online corpus significantly outperformed the conventional method in learning grammar. In other words, the data-driven learning method using online corpus can enhance the grammatical knowledge of the students. The mean score of the experimental group ($\bar{x}$ = 24.59) was higher than the control group (= 19.47).

Speaking of the second research question about students' opinions towards the data-driven learning using online corpus to enhance students' grammatical knowledge, the 5-Likert scale questionnaire was used to gathered data from the respondents. The items of question in the questionnaires are shown in table 4.

Table 5: The students' opinions toward the data-driven learning using online corpus

Statements	Degree of Agreement		Interpretation
	Mean	S.D.	
1. I think that learning of English grammar by data-driven learning using the online corpus improved my English grammatical knowledge	3.45	0.95	Agree
2. I think that data-driven learning using the online corpus increased my confidence in using grammar.	3.64	0.79	Agree
3. I think I actively participate in the lesson in data-driven learning using the online corpus.	3.66	0.81	Agree
4.. I think I can discover the grammar rules by myself or in group.	3.64	0.96	Agree
5. I think it is important that the teacher integrates collaborative learning activity into corpus-based analysis	3.42	0.93	Agree
6. I think the concordances are easy to read and understand	3.45	0.95	Agree
7. I think the data-driven activities learning using the online corpus are more challenging than the activities in the course book.	3.30	1.01	Agree
8. I think the corpus is very useful grammar learning resource.	3.64	0.79	Agree
9. I think data-driven learning motivates me to do more grammar exercise	3.51	0.83	Agree
10. I think the corpus requires basic knowledge of grammar for the analysis	3.66	0.81	Agree
Total	**3.53**	**0.88**	**Agree**

(1= strongly disagree, 2= disagree, 3= neutral, 4= agree, 5= strongly agree)

In overall, the ratings obviously demonstrated the opinions on "agree" with the total grand mean of 3.53 which indicated that most students acknowledged either totally or to a certain extent of agreement to the contribution of data-driven learning using online corpus in enhancing students' grammatical knowledge. In short, the students had positive opinion towards the treatment.

In this study, the data-driven learning method was designed to support the students as the language researchers consulted the online corpus database to find out about the target language structure by themselves. The questionnaire item no.3 revealed similarly at the average of 3.66 that the students had active role to learn grammar when they did so through the data-driven learning method. Correspondingly, item no. 4, at mean of 3.64 showing that students could discover the grammar structure by themselves or in group. Moreover, item no. 2 showed that the confidence was increased when using the grammar with the mean of 3.64. Concerning the research question whether the treatment was effective, in the questionnaire item no.1, students noted that the data-driven learning using online corpus improve their English grammatical knowledge at the mean of 3.45. Students became active learners because the data-driven learning method because it gave them more opportunities to work independently.

Interestingly, some respondents viewed item no.10 that the corpus requires basic knowledge of grammar for the analysis task at the mean of 3.66 . This was also supported by the rating result of the item no. 7 which of online corpus-based activities is more challenging than the activities in the course book. Speaking of the nature of data-driven learning task, students had to deal with the concordance lines and noticed it. After that without directly taught, they had to analyze the structure and the meaning by themselves. This explains why it is more challenging and requires some basic knowledge of grammar such as part of speech and tense. In data-driven learning activities, students viewed that it is important that the teacher integrates collaborative learning activity into corpus-based analysis at the average of 3.42 because some students may find that the task was beyond their ability.

Regarding the online corpus, students agreed that the corpus is very useful grammar learning resource at mean of 3.64 and at this level of students; they found that the concordances are easy to read and understand

The results from the open-ended questions indicated that the students had expressed both positive and negative opinions about implementing data-driven learning using online corpus in grammar class. With respect to the learning and instruction, they expressed followings:

1) The problem when dealing with the online corpus was the technical problems such as internet speed, IP address, and the specification of the computer. These are not big deal since the students were from the computer science area.

2) Learning English grammar by data-driven learning using the online corpus was time-consuming for some students because preferred informed grammar in the book because the rules are already presented and they can read and start drilling. However, there were a few of students like the corpus-based activity which it is not boring like sitting in the traditional grammar lecture class.

3) The students believed that the data-driven learning requires English syntactic knowledge to analyze the concordance and sometimes consumed a lot of time. Furthermore, they noted that it is a fun way to learn grammar because they were encouraged a lot of thinking. And teacher must lead the corpus analysis task: interaction and induction parts because some grammar topics were too difficult for them such as relative clause when the relative pronoun was omitted.

To sum up, the answers to the research questions were positive because Pre-test and Post-test scores indicated that the data-driven learning using online corpus contributed more effective results in enhancing grammatical knowledge than the conventional method. Implementing online corpus into the instruction has provided the learners a useful and effective method for facilitating grammar learning. Also, they enjoyed learning the treatment. For some students with lower levels of English proficiency, corpus analysis is problematic especially during the interaction and induction steps. Facilitation of the instructor is necessary for completing the task.

Discussion

The findings of this study showed the effective results of data-driven learning tasks using online corpus in learning grammar which it was demonstrated outperformance to the conventional method. It suggested that using data-driven learning in learning English grammar have an effect on students' achievement.

The concept of data-driven learning is believed to promote the self-discovery learning which the learner's role become more independent when learning, the teacher acts as the facilitator and the concordance on the corpus is used as the knowledge resource. It is also in compliant with the four Cs of the framework for 21st century learning that Creativity and innovation, Critical thinking and problem solving, Communication, and Collaboration are to be promoted in learning and instruction (Partnership for 21st Century Learning, 2015). According to the qualitative data from the open-ended questions in the questionnaire, some of learners had difficulty when exercising the data-driven leaning tasks because of the sentence structure and vocabulary shown in the output of the concordance. Concerning to this point, the data-driven learning performed well in promoting the critical thinking and problem-solving skill of a learners when they were working on three-I steps in order to generate the target language structure. Besides, the tasks were designed in group work because some students may have insufficient syntactic knowledge, these encouraged students in collaboration and communication skills in order to meet the objective of the particular tasks. Students were successful in task completion because they were allowed to use Thai while working on their assigned exercises; they were not needed to prepared for the explanation in English.

This corpus-based experience was in such a limited time frame of 3 grammar lessons with the total of 9 hours and both pre-test and post-test were the same, it is not expected that the learners will change the way they learn grammar. However, it was observed that students started to feel for the data-driven learning method. Even though they do not become fully self-directed learners yet, at least they observed the target language structure on the corpus and tried to understand it. This demonstrated important gains and valuable consequence of how independency has been developed.

In contrast to the conventional method, the role of the teacher in Data-driven learning became less active. The teacher was allowed to give guidance and observe the learners closely in order to help them when the learners faced problems. As noted in the feedback of students from the questionnaire, the concordance results from the corpus should be simplified by controlling the amount of language input because it was time-consuming to complete the task.

In addition, speaking of the students' ability to discover the language target structure either by themselves or with the assistance from peers, some students addressed that the teacher was needed to lead the activities. The reasons behind this feedback could be that some of these university students possess an extent of English proficiency. Still, it is suggested that the teacher should inform the advantages of this self-learning as well as the responsibility for their own learning. Plus, teacher should keep in mind not to give them the direct answer and try to promote the peer assistance learning.

Even though the lessons were well-planned; the worksheets were equipped with guided questions that would have generated the expected answers. There were a number of points came up during the class such as the grammar questions about the basic structure. Training for fundamental grammar structure and corpus searching skill is one way that could help students succeed in exploring concordances which was absent in this study. According to Sun (2003), it is more difficult for students who had experience in learning grammar deductively to change to the inductive method. Therefore, a training of how to learn through the data-driven learning must be conducted to deal with a large amount of information on the online corpus.

Recommendation for Future Research

This study lasted 3 weeks and the lesson time was total of 9 hours. The time difference between pre-test and post-test was only three weeks. It is obvious that the study time was such a very short period and the pre-test and post-test were the same. Students might be able to remember the test items. The larger scale of study time such as a

semester or more lessons of grammar topics would show the different results and feedback of student. Conversely, administering the post-test at the end of each grammar topic instead of testing it until the end of the treatment would showed different effects and demonstrated other aspects of how scores have been developed as well.

This study presents the results of learning when exploring the features of concordance data. The learning strategies of the students used in learning could be further studied. From the class observation, students have different learning strategies when dealing with the corpus which could reflect and suggest the development of the data-driven teaching materials in the future.

In addition, this study was on the learning of grammar, turning theory of data-driven learning into practice of other notions of language learning would have been another research area and contribute to the development of teaching and learning materials. For example, the feature of the concordancer can be used to search about the vocabulary regarding the frequency or collocation.

References

Aston, G. (2001). *Learning with corpora*. Houston: Athelstan.

Celce-Murcia, M. (2000). *Discourse and context in language teaching: a guide for language teachers.* New York: Cambridge University Press.

Cheng, W., Warren, M., & Xun-feng, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, *31*(2), 173–186. doi:10.1016/ S0346-251X(03)00019-8

Conrad, S. (2005). Corpus linguistics and L2 teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 393-409). Mahwah, NJ: Lawrence Erlbaum Associates.

Daskalovska, N. (2013). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, *28*(2), 130–144. doi:10.1080/09588221.2013.803982

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

H.G. Widdowson. (2000). *Linguistics*. Shanghai: Shanghai Foreign Language Education Press

Keck, C. M., & Kim, Y. (2014). *Pedagogical grammar*. Amsterdam: John Benjamins Publishing.

McCarthy, M. & Carter R., (1995) Grammar and the spoken language. *Applied Linguistics. 16 (2),* 141-158

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge

Nation, I. S. P. (2001). *Learning Vocabulary in another language*. Cambridge: Cambridge University Press.

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

Partnership for 21st Century Learning. (2015). *The Partnership for 21st Century Skills*. Retrieved from http://www.p21.org/about-us/p21-framework

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158. doi:10.1093/applin/11.2.129

Sun, Y. C. (2003). Learning process, strategies and web-based concordancers: A case study. *British Journal of Educational Technology*, *34*(5), 601-613.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing, 13*(4), 257-283. doi:10.1016/j.jslw.2004.06.002