A CORPUS-BASED STUDY OF COLLOCATIONS OF ENGLISH SYNONYMS: STUDENT AND LEARNER
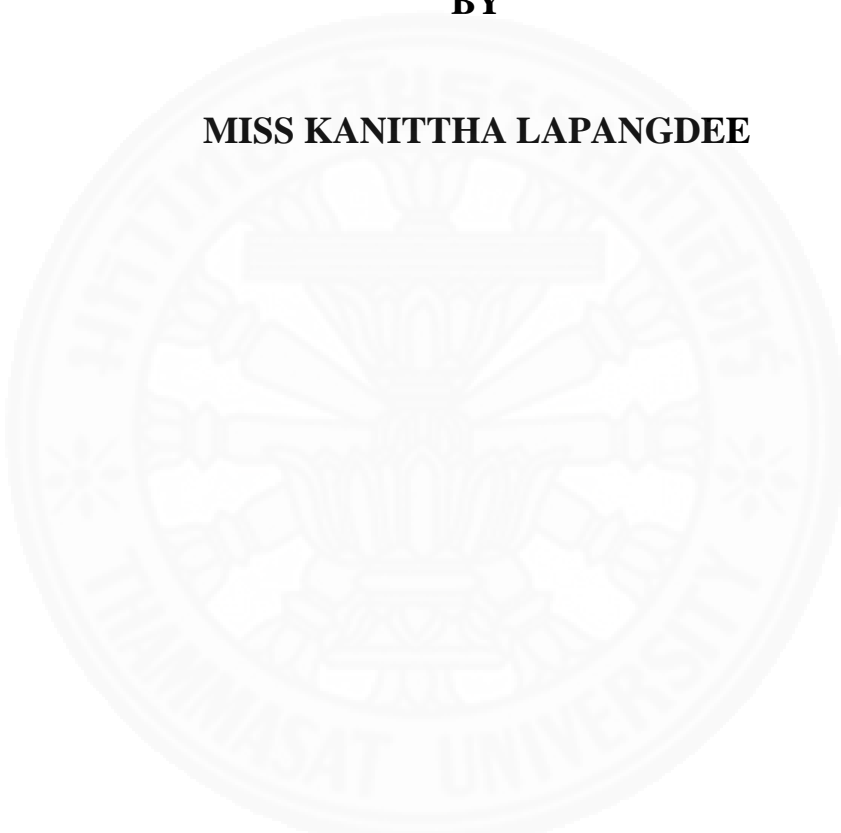
BY

MISS KANITTHA LAPANGDEE

AN INDEPENDENT STUDY PAPER SUBMITTED IN PARTIAL
FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN
ENGLISH LANGUAGE TEACHING
LANGUAGE INSTITUTE
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2017
COPYRIGHT OF THAMMASAT UNIVERSITY

A CORPUS-BASED STUDY OF COLLOCATIONS OF ENGLISH SYNONYMS: STUDENT AND LEARNER

BY

MISS KANITTHA LAPANGDEE

AN INDEPENDENT STUDY PAPER SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF ARTS IN
ENGLISH LANGUAGE TEACHING
LANGUAGE INSTITUTE
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2017
COPYRIGHT OF THAMMASAT UNIVERSITY

THAMMASAT UNIVERSITY

LANGUAGE INSTITUTE

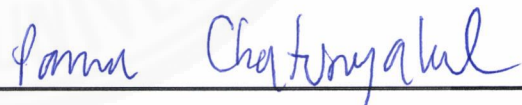INDEPENDENT STUDY PAPER

BY

MISS KANITTHA LAPANGDEE

ENTITLED

A CORPUS-BASED STUDY OF COLLOCATIONS OF ENGLISH SYNONYMS:
STUDENT AND LEARNER

was approved as partial fulfillment of the requirements for
the degree of Master of Arts in English Language Teaching

on June 22, 2018

Chairman _____

(Panna Chaturongakul, Ph.D.)

Member and Advisor _____

(Assistant Professor Supakorn Phoocharoensil, Ph.D.)

Dean _____

(Associate Professor Supong Tangkiengsirisin, Ph.D.)

Independent Study Paper Title	A CORPUS-BASED STUDY OF COLLOCATIONS OF ENGLISH SYNONYMS: STUDENT AND LEARNER
Author	Miss Kanittha Lapangdee
Degree	Master of Arts
Major Field/Faculty/University	English Language Teaching Language Institute Thammasat University
Independent Study Paper Advisor	Assistant Professor Supakorn Phoocharoensil, Ph.D.
Academic Years	2017

ABSTRACT

The present study aimed to (1) investigate the possible noun collocates of two synonyms, i.e., student and learner, and to (2) find differences between student and learner in terms of noun-noun collocations. In this study, student and learner serve as an attributive which is used to precede and modify other nouns. The data were derived from the Corpus of Contemporary American English (COCA). For the task of synonym choice, the corpus-based data provided shared noun collocates of two synonyms. The result showed that student and learner share a similar meaning, but differ in terms of collocations. Furthermore, it was discovered that there are less-preferred collocations of learner as student seems to have more possible collocates. For this reason, the two synonyms cannot substitute another in all contexts. In addition, the corpus-based data are useful as they provide many examples of a large proportion of the vocabulary for language learning.

Keywords: synonyms, collocations, less-preferred collocations, corpora

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

In writing, repeating the same words several times can make writing uninteresting; therefore, synonyms are used to make the writing more vivid and captivating. In some contexts, a word cannot be replaced with another. As Inkpen and Hirst (2002) provided support for this idea that even though they are synonyms; one does not collocate well with other word choices for the sentences. Lin (1998, as cited in Pearce, 2001) explains that "even though *baggage* and *luggage* are synonyms, only *baggage* can be modified by *emotional, historical* or *psychological*". Gitsaki, Daigaku and Taylor (2000, as cited in Pearce, 2001) give several examples of English problematic word combinations, e.g., *several* and *many* as synonyms. Only *many* but not *several* can be used with *thanks*.

Since a number of previous studies were focused on different types of collocations, for instance, noun-noun collocations, adjective-noun collocations, adverb-noun collocations, and verb-noun collocations, *student* and *learner* have not been chosen as a research object in those studies. Generally, *student* and *learner* serve as a 'noun', which is "a word or group of words that represent a person, a place, a thing or activity, or a quality or idea" (Longman Dictionary of Contemporary English, 2005, p.1122). However, it is often found that *student* and *learner* serve as an *'attributive'*, which is used to precede and modify other nouns (Timmis, 2015). According to Longman Dictionary of Contemporary English (LDOCE, 2005), *student* collocates with other nouns, e.g., student *body*, student *government*, student *loan* and student *teaching* (Longman Dictionary of Contemporary English, 2005, p.1650), whereas *learner* is used with other nouns, e.g., learner *driver* (Longman Dictionary of Contemporary English, 2010, p.916). In order to fill the research gaps, collocations and synonyms were examined altogether and corpus data were used to see the possible noun collocates of *student* and *learner*.

1.2 Research Objectives

    (1) To investigate the possible noun collocates of *student* and *learner*

    (2) To find differences between *student* and *learner* in terms of noun-noun collocations

1.3 Research Questions

    (1) What are the possible noun collocates of *student* and *learner* found in the Corpus of Contemporary American English (COCA)?

    (2) To what extent are *student* and *learner* different in noun-noun collocations?

1.4 Scope of the study

In this study, two synonyms, i.e., *students* and *learner* which serve as *"attributive"* are investigated only in terms of collocations. The data were derived from the Corpus of Contemporary American English (COCA).

1.5 Limitations of the Study

The study focuses on the possible noun collocates of *student* and *learner* through one corpus data source, namely the Corpus of Contemporary American English (COCA), and only 600 concordance lines of *student* and *learner* were analyzed in the study. Due to the small corpus size, the findings may not represent all the possible noun collocates of both synonyms that native speakers of English use. Furthermore, since the data were drawn from American English, the findings may not be generalized to other varieties of English, e.g., British English.

1.6 Significance of the study

The study is significant in the following ways:

    (1) The data were gathered from the genuine communications of people through corpus data source which are more authentic than artificial texts.

(2) The results of the study show the noun-noun collocations and less-preferred collocations of *student* and *learner* which are useful in English writing.

1.7 Definition of Terms

(1) *Synonym*: a word which has the same or similar meaning, e.g., *job, task,* and *assignment.*

(2) *Attributive*: a noun which is used to precede and modify other noun, e.g. a *student* assignment.

(3) *Collocation*: a word which collocates with the target word and does not lead to unnatural readings, e.g., *a daunting task.*

(4) *Less-preferred collocation*: a word which should not be used with the target word unless there is no better alternative and does not lead to unnatural readings, e.g., *a daunting job.*

(5) *Anti-collocation*: a word which must not be used in lexical choice as it leads to unnatural readings, or grammatical errors, e.g., *a daunting assignment.*

1.8 Organization of the study

The rest of the study is organized in the following ways:
Chapter 2 "Review of Literature" consists of the definition and types of synonym and collocation. Furthermore, related studies are reviewed. Chapter 3 "Research Methodology" comprises data source, data collection procedure, and data analysis. In Chapter 4, the findings will be presented and discussed. Chapter 5, finally, includes the conclusions, recommendations for further studies, and pedagogical implications.

CHAPTER 2
REVIEW OF LITERATURE

2.1 Synonym

Oxford Advanced Learner's Dictionary (2010, p. 1515) gives a definition of *a synonym* as "a word or expression that has the same or nearly the same meaning as another in the same language: *'big'* and *'large'* are synonyms." In a study of synonyms, two major types, strict and loose synonyms, are often referred to.

Strict and Loose Synonyms

Strict synonyms refer to "the words that can be used interchangeably in all contexts and the substitution must not result in a change in meaning, style and connotation of the context" (Jackson & Amvela, 2000, as cited in Phoocharoensil, 2010). However, Kreidler (1998, as cited in Petcharat & Phoocharoensil, 2017) claimed that no two words share all the same linguistic features and it is considered unnecessary redundancy if there is a word which can fully replace another in all contexts. On the other hand, *loose synonyms* refer to the words that can be used interchangeably where their meaning overlap but one cannot replace another (Jackson & Amvela, 2000). Synonyms can be distinguished by the following criterion: *Collocations.*

2.2 Collocations

Nattinger and DeCarrio (1992, as cited in Moehkardi, 2002, p. 53) give a definition of collocations which are "strings of word that seem to have certain mutual expectancy or a greater-than-chance likelihood that they will co-occur in any text". Sinclair (1991, as cited in Khodadady, 2012, p. 42) also defines a collocation as "the co-occurrence of two or more words within a short space of each other in a text." According to Benson, Benson and Ilson (1986), there are two types of collocations, i.e., grammatical collocations and lexical collocations. Benson et al. (1986) list various combinations of grammatical collocations: a noun, a verb, and an adjective plus a particle e.g. preposition, while lexical collocations consist of noun + noun; adjective + noun; adverb + noun; verb + noun; verb + adverb (Bahns, 1993).

For the task of synonym choice, two other collocations are implemented (less-preferred collocation and anti-collocation). Pearce (2001, p. 3) gives the definition of less-preferred collocation as "the word which tends not to be used with the target word although, if used, it does not lead to unnatural readings", whereas anti-collocation is defined as "the word which must not to be used with the target word since it will lead to unnatural readings" (p. 3). Inkpen and Hirst (2002, p. 67) defines less-preferred collocation as the word which should not be used in lexical choice unless there is no better alternative, while anti-collocation refers to the word which must not be used in lexical choice. For example, three synonyms: *task, job,* and *assignment* with the adjective *'daunting'*, *a daunting task* is a collocation; *a daunting job* is a less-preferred collocation, and *a daunting assignment* is an anti-collocation.

As mentioned above, the frequency plays a role among collocations, less-preferred collocations, and anti-collocations. Pearce (2001, p. 3) provided the example collocates of *baggage* and *luggage* by drawing the data from AltaVista (a Web search engine), the British National Corpus (BNC), and the Cambridge International Dictionary of English (CIDE). The occurrences count with both *baggage* and *luggage* and the data show the frequency of collocates as shown in the Table 1.

Table 1. The collocates of "baggage" and "luggage"

Collocates	AltaVista		BNC		CIDE	
	baggage	*luggage*	*bagage*	*luggage*	*baggage*	*luggage*
allowance	**3279**	**5**	*0*	*502*	**1**	*0*
car	**3324**	**4**	*0*	*357*	**1**	*0*
carts	*806*	**2**	*0*	**1268**	-	-
compartment	*2890*	*1*	**3**	**5144**	*0*	**1**
hall	**197**	**2**	*0*	*41*	**-**	-
handler	**1448**	**5**	*0*	*83*	**1**	*0*
rack	*164*	*0*	**8**	**14773**	*0*	**1**

According to Table 1, each pair of numbers consists of two occurrences of the

collocates in the first column. The first of each pair collocates with *baggage,* whereas the second collocates with *luggage*. The high frequency is presented in bold, while the low frequency is shown in italic, and the dash represents zero co-occurrence. It can be seen that even the occurrences count with two synonyms, the results are different. Among these three corpus data, AltaVista provides more obviously the frequency of two synonyms. However, this web engine is unavailable since 2013 ("AltaVista," n.d.). In addition, it was found that although the data from the Cambridge International Dictionary of English (CIDE) and the British National Corpus (BNC) represent number 0, it does not represent non-existence of occurrences. Furthermore, Inkpen and Hirst (2006) supported that a low number of co-occurrences indicate a less-preferred collocation or even an anti-collocation.

In the study of two synonyms; *student* and *learner* in terms of frequency, *student* is one of the 1000 most common words in spoken and written English. In other words, *student* is the first 1000 words to be taught in both spoken and written, while *learner* is not one of the 3000 most common words in spoken and written English (Longman Dictionary of Contemporary English, 2005).

2.3 Corpus

A Corpus is defined as "an electronically stored collection of texts that is exploited using specialized software" (Hunston, 2006, p. 234). A *Corpus* is "used to test hypotheses about language and to provide quantificational data about language use." In addition, a corpus has influenced on a number of linguistics applications, e.g., translation and language teaching (Hunston, 2006, p. 234). Likewise, Huang and Yao (2015, p. 949) gives the definition of *corpus* as "a collection of examples of language in use that are selected and compiled in a principled way".

Benefits of Corpora

Sinclair (1996, as cited in Granger, 2002, p. 5) describes the value of corpora for quality is that all the data is gathered from the genuine communications of people unlike data gathered in artificial conditions which are rarely fully natural, e.g., textbooks. Chambers (2005, p. 115) also claims that corpora clearly provide superior

resources for the study of language including many examples of a large proportion of the vocabulary. Moreover, Hunston (2006, p. 234) supports that "corpora provide an insight into recurring patterns of language that are difficult to observe in other ways."

Language Corpora in Language Learning

In the foreign or second language learning and teaching, the Corpus of Contemporary American English (COCA) "has significant advantages over other free corpora in terms of vocabulary study as it gives a sufficient patterning of English lexis and grammar" produced by American English native speakers (Davies 2009, as cited in Ghaniabadi, Marjane, & Zareian, 2015). Granger (2002) stated that the foreign language context usually involves some degree of artificiality which is not authenticity, however, corpus-based data provides genuine communications which lead learners to access authentic language.

A number of previous studies were conducted by using corpora. Ghaniabadi et al. (2015) conducted a study by using a collection of written texts by Iranian EFL learners and the Corpus of Contemporary American English (COCA) in order to reveal some facts about the frequency of the use of adjective-noun collocations according to noun abstractness. The data were derived from a written text which required academic vocabulary and grammatical sentences and compared to the frequency in COCA. It was found that the adjective+abstract noun collocations were more efficiently and frequently used by Iranian EFL learners. Moreover, it was found that a corpus provides simultaneously example sentences with frequency searches.

Phoocharoensil (2010) studied five synonyms, i.e. *appeal, request, plead, beg, and ask* . The data were collected from three standard learner's dictionaries, i.e., Oxford Advanced Learner's Dictionary (2005), Cambridge Advanced Learner's Dictionary (2009), and Longman Dictionary of Contemporary English (2009), and the Corpus of Contemporary American English (COCA). The result showed that even these five synonyms share similar meaning, they still differ regarding connotations, styles, dialects, grammatical patterns, and collocations. Therefore, the five synonyms cannot be used interchangeably in all contexts.

Petcharat and Phoocharoensil (2017) investigated three synonyms, *i.e. appropriate*, *proper*, and *suitable*. The data were collected from three dictionaries, namely Macmillan Collocations Dictionary (2010), Longman Advanced American Dictionary 3rd edition (2013), and Longman Dictionary of Contemporary English 6th edition (2014), and the Corpus of Contemporary American English (COCA). The results showed that even though the three synonyms share the same meaning, they differ in terms of collocations, detailed meanings, degrees of formality, and grammatical patterns. Therefore, the three synonyms cannot substitute the others in all contexts. It was also found that the corpora provide some additional data which is non-existent in the three dictionaries.

Apart from the studies that focus on the Corpus of Contemporary American English (COCA), there are studies which used other data sources of corpus-based data, e.g., Chung, 2011; Jafarpour, Hashemian, & Alipour, 2013; Xia, 2013.

Chung (2011) studied two synonyms: *create* and *produce*. The data were collected from the Brown Corpus and the Frown Corpus. The results showed that *create* and *produce* have two similarities. First, they appear most frequently as a bare infinitive and in the *–ed* form. Second, they have two overlapping senses, i.e., *"bring into existence/cause to happen, occur, or exist"* and *"create or manufacture a man-made product"*. On the other hand, *create* and *produce* have certain differences. *Create* are often to be followed by abstract objects that are fewer in quantity and higher possibility in creativity, e.g., *create jobs*, *create world*, and *create problems*, whereas *produce* tends to be followed by objects that are more in quantity and lower in creativity, e.g., *produce machines*, *produce seeds*, and *produce goods*.

Jafarpour, Hashemian, and Alipour (2013) also studied the effects of learning collocations of near-synonyms from the corpus-based approach, and from the traditional approach. There were two groups of Iranian L2 learner participants, i.e., experimental group and control group, who were similar in terms of proficiency and collocational competence. The experimental group was taught through concordance lines from the British National Corpus (BNC), while the control group was taught collocations directly by the teacher. They were given a definition of collocations,

examples of different types of exercises, and follow up exercises. The data were collected from a pre-test, a post-test, a pre-writing and a post-writing. Then they were compared between two groups. The results showed that the corpus-based approach has more effectiveness on collocation learning and in developing writing fluency of L2 learners. Even though Iranian L2 learners have problems with collocations in their writings, the corpus-based approach helps them use the words correctly, and makes collocations more apparent to L2 learners. In addition, teaching collocations to the L2 learners through the concordance lines improves their collocational competence, increases their awareness of combining words naturally, and reduce miscollocations.

Xia (2013) did a case study of *ability*, which has high frequency in the Chinese Learner English Corpus (CLEC). A study was conducted on the production of verb-noun collocations by Chinese EFL learners based on Chinese Learner English Corpus (CLEC), the International Corpus of Learner English ICLE, and the British National Corpus (BNC). The data from Chinese Learner English Corpus were collected and analyzed. Then, they were compared to the data from International Corpus of Learner English and the British National Corpus. The study showed that there is a clear difference between the use of verb-noun collocations between Chinese EFL learners and English native speakers. Moreover, Chinese EFL learners also give less satisfactory performance in verb-noun collocations in terms of correctness and appropriateness.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Data Source

The language data in this study were drawn from the Corpus of Contemporary American English (COCA) created by Mark Davies from Brigham Young University. The corpus contains more than 560 million words of text (20 million words each year 1990-2017) and it is divided among academic texts, newspapers, popular magazines, spoken, and fiction, which provides a sufficient patterning of English vocabulary and grammar (Davies, 2008; Ghaniabadi et al., 2015).

3.2 Data Collection Procedure

The principle source of data was the Corpus of Contemporary American English (COCA), where each of 300 concordance lines of *student* and 300 concordance lines of *learner* were randomly searched for. Then those lines were sorted out to make sure that *student* and *learner* served as attributive and collocated with other nouns. In other words, verb collocates of *student* and *learner* would be excluded, e.g., a certain *student* receives a particular grade (drawn from Davie, 2008). There were 136 tokens of *student* and 43 tokens of *learner* to be analyzed.

3.3 Data Analysis

The study was aimed at investigating possible noun collocates of *attributive*s, i.e., *student* and *learner*. The number of possible noun collocates of two synonyms were counted and ordered in the frequency. Then collocations of *student* and *learner* were analyzed.

CHAPTER 4
RESULTS AND DISCUSSION

4.1 Results and Discussion

Table 2. The noun collocates of *"student"* in COCA

Collocates	Freq.	Collocates	Freq.	Collocates	Freq.
learning	12	satisfaction	2	interviews	1
achievement	11	success	2	journalist	1
teacher	7	variables	2	knowledge	1
teaching	7	aspiration	1	litigants	1
body	5	assessment	1	motivation	1
loan	5	attrition	1	perspective	1
performance	5	autonomy	1	progress	1
population	5	candidate	1	representative	1
behavior	4	characteristics	1	resistance	1
engagement	4	choice	1	responses	1
group	4	choreography	1	scores	1
ideas	3	club	1	services	1
interest	3	coaches	1	shooter	1
participants	3	complaints	1	strategist	1
work	3	council	1	understanding	1
data	2	development	1	union	1
growth	2	effort	1	visitations	1
movement	2	enrollments	1	voices	1
need	2	evaluation	1	writers	1
outcome	2	exchange	1		
participation	2	government	1		

From Table 2, three hundred concordance lines with the attributive *study* were derived from COCA. After manual search, it showed sixty-one noun collocates of *study* in order of frequency. The most frequent noun was *learning* (12 tokens), followed by

achievement (11 tokens), *teacher* and *teaching* (7 tokens), *body, loan, performance,* and *population* (5 tokens), *behavior, engagement*, and *group* (4 tokens), *ideas, interest, participants,* and *work* (3 tokens), *data, growth, movement*, *need, outcome, participation, satisfaction, success,* and *variables* (2 tokens). The least frequent nouns found in COCA were thirty-seven noun collocates of *study*, i.e., *aspiration, assessment, attrition, autonomy, candidate, characteristics, choice, choreography, club, coaches, complaints, council, development, effort, enrollments, evaluation, exchange, government, interviews, journalist, knowledge, litigants, motivation, perspective, progress, representative, resistance, responses, scores, services, shooter, strategist, understanding, union, visitations, voices,* and *writers* (1 token).

Then, three hundred concordance lines with the attributive *learner* were derived from COCA. After manual search, it showed twenty-eight noun collocates of *learner* in order of frequency as shown in Table 3.

Table 3. The noun collocates of *"learner"* in COCA

Collocates	Freq.	Collocates	Freq.	Collocates	Freq.
characteristics	5	behavior	1	preferences	1
outcomes	5	choice	1	profile	1
needs	3	competency	1	progress	1
achievement	2	differences	1	responses	1
identities	2	experience	1	subjects	1
participation	2	groups	1	success	1
population	2	independence	1	variables	1
variability	2	organization	1	workbooks	1
assignments	1	performance	1		
attainment	1	perspectives	1		

The most frequent noun collocates of *learner* were *characteristics* and *outcomes* (5 tokens) followed by *needs* (3 tokens), *achievement, identities, participation,* and *population* (2 tokens). The least frequent nouns found in the corpus were twenty nouns, i.e., *assignments, attainment, behavior, choice, competency, differences,*

experience, groups, independence, organization, performance, perspectives, preferences, profile, progress, responses, subjects, success, variables, and *workbooks* (1 token).

According to Table 2 and Table 3, there were fifteen shared noun collocates of both attributives: *student* and *learner,* i.e., *achievement, behavior, characteristics, choice, groups, needs, outcomes, responses, participation, performance, perspectives, population, progress, success, and variables.* Then each noun which appeared only in Table 1 or Table 2 was manually searched for from COCA in order to find out whether the noun collocates of either *student* or *learner* can be used together with another. For example, *data* in the Table 1 is one of the noun collocates of *student.* After manual tagging, it showed that *learner* can be used together with the noun *data* as well.

Table 4. The shared collocations of *"student"* and *"learner"* in COCA

Collocates	*Student*	*Learner*	Collocates	*Student*	*Learner*
achievement	✓	✓	growth	✓	✓
assessment	✓	✓	ideas	✓	✓
assignments	✓	✓	identities	✓	✓
attainment	✓	✓	independence	✓	✓
autonomy	✓	✓	interest	✓	✓
behavior	✓	✓	knowledge	✓	✓
characteristics	✓	✓	learning	✓	✓
choice	✓	✓	motivation	✓	✓
competency	✓	✓	needs	✓	✓
data	✓	✓	organization	✓	✓
differences	✓	✓	outcomes	✓	✓
engagement	✓	✓	participants	✓	✓
enrollment	✓	✓	participation	✓	✓
evaluation	✓	✓	performance	✓	✓
experience	✓	✓	perspectives	✓	✓
groups	✓	✓	population	✓	✓

Collocates	*Student*	*Learner*	Collocates	*Student*	*Learner*
preferences	✓	✓	success	✓	✓
profile	✓	✓	teacher	✓	✓
progress	✓	✓	understanding	✓	✓
representative	✓	✓	variability	✓	✓
responses	✓	✓	variables	✓	✓
satisfaction	✓	✓	workbooks	✓	✓
subjects	✓	✓			

According to Table 4, there were forty-five noun collocates of *student* and *learner*, i.e., *achievement, assessment, assignments, attainment, autonomy, behavior, characteristics, choice, competency, data, differences, engagement, enrollment, evaluation, experience, groups, growth, ideas, identities, independence, interest, knowledge, learning, motivation, needs, organization, outcomes, participants, participation, performance, perspectives, population, preferences, profile, progress, representative, responses, satisfaction, subjects, success, teacher, understanding, variability, variables,* and *workbooks,* as exemplified in (1) – (90).

(1) ...How have you helped to raise **student achievement**, and how did...

(2) ...arguments for why each construct should influence **learner achievement**.

(3) **Student assessment** is embedded within lessons from multiple perspectives...

(4) ...physical education receives instruction in **learner assessment** techniques...

(5) ...to craft measurable behavior-based objectives for **learner assignments**.

(6) ...a very natural style of writing on **student assignments**, instructors may want...

(7) This information may become objectives for **learner attainment.**

(8) ... intended to directly measure **student attainment** of only a portion...

(9) ...classroom environment that is amenable to openness and **student autonomy**...

(10) ... to promote self-directed learning and to increase **learner autonomy**...

(11) Their observations of **student behavior** may be specific to their group...

(12) ...motivating **learner behavior** in wanting to learn more within the ongoing unit...

(13) ...the matter of **student characteristics** or motives for writing violently…

(14) ...study was designed to describe the **learner characteristics** of the students...

(15) How do you provide opportunities for **student choice** in your classroom?

(16) ...teacher decision making in the curriculum versus **learner choice** of activities...

(17) ...a critical **learner competency** is: "Manage time effectively; "...

(18) ... instruction that emphasized rhythm increased **student competency** in meter...

(19) ...no restriction on the flow of **student data** through education technology...

(20) Learning analytics is the collection and subsequent use of **learner data**....

(21) ...appreciation of **learner differences** required for successful cooperative...

(22) **Student differences** also play a key role in interactions with faculty...

(23) Encouraging **student engagement** with future tie-in texts gives students...

(24) ...to use it to enhance writing instruction and **learner engagement**.

(25) ....have led to a push for high schools to complete for **student enrollment**.

(26) With the increase in adult **learner enrollment** and the realization…

(27) ...promotion purposes, the conduct of **student evaluations** is a primary device...

(28) ... EFCAs on affect ventilation, **learner evaluation**, postdicted acceptance...

(29) The premise was for school librarians to mirror the **learner experience**...

(30) ...the **student experience** is enriched with lyrics in different languages...

(31) The **student group** progressed from an average of approximately 46 WRCM...

(32) ...greater than the scores for expressive language for all three **learner groups**...

(33) Ultimately, **student growth** can be a legitimate factor in judging teacher...

(34) ...to improve the instruction, (2) to promote teacher and **learner growth** and...

(35) The data analysis revealed several **student ideas** for situating writing course...

(36) ... pre-service teachers praised learners, accepted **learner ideas**, and posed...

(37) ... and contexts to inform the development of their **learner identities**...

(38) ...building on the association of **student identities** with global...

(39) ... lessons scaffolded the models as ***learner independence*** was purported...

(40) …interest is that 84% of the teachers report improved ***student independence***...

(41) ... their awareness of ***student interest***, engaging learning experiences...

(42) Projects are determined by students based on ***learner interest*** and value.

(43) This IPM environment enhanced ***student knowledge*** while fostering motivation...

(44) ... benefits of PALS are improved ***learner knowledge***, use, and mastery...

(45) ...my interest has shifted toward ***student learning*** and brain development.

(46) ...In J. W. Keefe (Ed.), ***Learner learning*** styles: Diagnosing and prescribing.

(47) ...which promotes ***student motivation***, reading, and academic achievement...

(48) ...she also engaged and sustained ***learner motivation*** by providing a choice...

(49) ...transition assessment identifies areas of ***student need*** that can be addressed...

(50) ...and agenda were established to evaluate ***learner needs*** and learner...

(51) ...constitute ***learner organization*** and enable learner organization to live...

(52) Any ***student organization*** was free to leave Inter Varsity at any time…

(53) ...an intervention caused an increase in a ***student outcome*** such as…

(54) ...cause and effect, reducing education to teacher input and ***learner outcome***.

(55) Some ***student participants*** were worried about how they would access...

(56) ...to allow the teachers and ***learner participants*** to be desensitized...

(57) ...exploration courses, or during ***student participation*** in a part-time job...

(58) ...to encourage interaction and ***learner participation*** is highlighted...

(59) How will you teach it and how will you evaluate ***student performance***?

(60) ...preceding characteristics may well have an impact on ***learner performance***.

(61) ...what your physical space looks like from a ***student perspective***, give students...

(62) ...some focus primarily on **learner perspectives** while others emphasize...

(63) Taiwan, the ***student population*** in higher education has gone through changes...

(64) ...it is important to recognize that the ***learner population*** and learning...

(65) ...established to evaluate learner needs and **_learner preferences_**.

(66) As it generates an awareness of **_student preferences_** in relation to faculty...

(67) ...literature is clear that differentiation by **_learner profile_** is a best practice...

(68) …identification process includes gathering data to develop a **_student profile_**...

(69) ... pushing to deploy reliable, external data to measure **_student progress_**.

(70) ...dated and kept from one evaluation to the next to notice **_learner progress_**.

(71) ...a **_student representative_** on the Arizona Board of Regents, came to the...

(72) ...governing bodies and all secondary schools have **_learner representative_**...

(73) **_Student responses_** were analyzed by the program Krkwic...

(74) ...documentation of **_learner responses_** to managerial and instructional tasks.

(75) It entails evaluating **_student satisfaction_** and achievement of learning outcomes.

(76) ...can improve a learning environment by increasing **_learner satisfaction_**.

(77) ...obtained from the teacher subjects and parents of the **_learner subjects_**.

(78) ...a reasonable level of complexity for use with the **_student subjects_** who would...

(79) All educators share a mission of working toward **_student success_**.

(80) INDICATORS OF **_LEARNER SUCCESS_** "How do teacher librarians...

(81) The time afforded for each **_student teacher_** to practice their teaching...

(82) ... education teachers, assistants and English Language **_Learner teacher_** go into...

(83) ...questioning is foremost in developing **_student understanding_**.

(84) ...and replicas as they enlarge **_learner understanding_** of daily lessons...

(85) ...content systems for their potential to support **_learner variability_** within...

(86) ...devoted to the three sources of **_student variability_** frequently...

(87) Two of the five **_student variables_** were found significant at the.05 level...

(88) Gender and Ability are **_learner variables_** that have not been widely explored...

(89) They already had the accompanying **_learner workbooks_**...

(90) ...age-appropriate stories, teachers' manuals, **_student workbooks_** and reading...

The sentences (1) – (90) represented the examples of lexical collocations (student and learner + noun) (Benson et al., 1986 as cited in Moehkardi, 2002).

Table 5. The different collocations of *"student"* and *"learner"* in COCA

Collocates	*Student*	*Learner*	Collocates	*Student*	*Learner*
aspiration	✓	✗	litigants	✓	✗
attrition	✓	✗	loan	✓	✗
body	✓	✗	movement	✓	✗
candidate	✓	✗	resistance	✓	✗
choreography	✓	✗	scores	✓	✗
club	✓	✗	services	✓	✗
coaches	✓	✗	shooter	✓	✗
complaints	✓	✗	strategist	✓	✗
council	✓	✗	teaching	✓	✗
development	✓	✗	union	✓	✗
effort	✓	✗	visitation	✓	✗
exchange	✓	✗	voices	✓	✗
government	✓	✗	work	✓	✗
interviews	✓	✗	writers	✓	✗
journalist	✓	✗			

According to Table 5, there were only twenty-nine nouns collocating with *student*, i.e., *aspiration, attrition, body, candidate, choreography, club, coaches, complaints, council, development, effort, exchange, government, interviews, journalist, litigants, loan, movement, resistance, scores, services, shooter, strategist, teaching, union, visitation, voices, work, and writers,* as can be seen in (91) – (119).

(91) ...importance of increased levels of ***student aspiration*** in the larger context.

(92) ...Development and validation of an instrument that predicts ***student attrition***...

(93) ...are in schools at which 90 to 100 percent of the ***student body*** are kids of color.

(94) ... endorsed a Republican ***student candidate***; he won one of the 20 seats in U.C.

(95) **Student choreography** and works-in-progress performed by...

(96) Westside currently does not recognize any **student club** that advocates...

(97) A total of 64 **student coaches** participated in this study.

(98) ...and quickly, with a structured system to address **student complaints**...

(99) He was a kid in high school in **student council** when the revolution began.

(100) ...better understand their role in **student development** as measured by...

(101) One approach to focus **student effort** would be the use of a checklist that...

(102) ...it would be education, helping children, **student exchange** and cultural...

(103) ...two high-school boys active in **student government** and bound for college...

(104) **Student interviews** were focused broadly on the case study…

(105) I was a registered Democrat at the time and I was a **student journalist**.

(106) …rules key anti-bias law applies to **student litigants** only if school...

(107) …out on a financial limb, taking out a **student loan** to pay for my education...

(108) ...research also studies **student movement** between 2-year and 4-year…

(109) The lack of **student resistance** became a guilty pleasure because…

(110) **Student scores** were therefore summed to provide a total composition…

(111) ...serving as vice president of **student services** and had an opportunity to...

(112) ...the shape suggests a person pointing a gun at you, the **student shooter**.

(113) But our **student strategist** does not wish to place even partial or subsidiary…

(114) During his **student teaching**, Wesley taught elementary PE for 6 weeks...

(115) ...Sigma Phi Epsilon set up a big seesaw in front of the **student union** at..

(116) ...and colleges were very willing to participate in **student visitations**...

(117) **Student voices** should be taken into account in the development of such…

(118) They talked about the craft of grading and commenting on **student work**.

(119) …new literacies to create third space classrooms that engage **student writers**...

In terms of word frequency, the number of noun collocates of high frequency words might be more than that of the low frequency words as there are the preponderance of noun collocates of *student* over *learner*. According to Longman Dictionary of Contemporary English (2005), in terms of frequency, *student* is one of the 1000 most common words in spoken and written English. In other words, *student* is the first 1000 words in both spoken and written English that should be taught, while *learner* is not one of the 3000 most common words in spoken and written English. Therefore, the possible noun collocates of *learner* are outnumbered by the possible noun collocates of *student*. According to Table 4 and Table 5, there are forty-five shared noun collocates of *student* and *learne*r, e.g., *achievement* and *behavior*, whereas twenty-nine nouns, e.g., *effort* and *loan*, are used only with *student.* The results are in line with what Chung (2011) proposed in that although some words, e.g., *manufacture*, are shared by both synonyms, i.e., *create* and *produce*, some other words, e.g., *farming*, are found to be used only in *produce*. Moreover, the data is supported by Inkpen and Hirst (2002) in terms of less-preferred collocations which word should not be used in lexical choice unless there is no better alternative. Those twenty-nine nouns are less-preferred collocations of *learner* as *student* is better alternative to be collocated with when automatically generating the text. In addition, the results of this study are similar to what Phoocharoensil (2010), and Petcharat and Phoocharoensil (2017)   found about synonyms that even some words share the same meaning, they still differ in terms of collocations. The results are also in line with what Pearce (2001) proposed in that although some words are synonyms, the occurrences of collocates are different.

CHAPTER 5
CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This study is a corpus-based investigation of collocations of English synonyms *student* and *learner*. The findings of this study suggest the number of possible noun collocates of *student* and *learner* derived from the Corpus of Contemporary American English (COCA). In terms of word frequency, the possible noun collocates of *learner* are outnumbered by the possible noun collocates of *student*. Hence the results show that *student* and *learner* share a similar meaning, but differ in terms of collocations. For the task of synonym choice, the corpus data provide shared noun collocates of *student* and *learner*. Furthermore, they demonstrate that there are less-preferred collocations of *learner* as *student* seems to have more possible collocates. In addition, this study has proposed that corpora have significant advantages in terms of vocabulary study as they provide many examples of a large proportion of the vocabulary for language learning. The value of corpora is that all the data is gathered from the genuine communications of people. Corpora also provide an insight into recurring patterns of language that are difficult to observe in other ways. Students should be taught to be aware of the differences of English synonyms which cannot replace others in all contexts, and students should be instructed to use corpora in order to find further language information apart from dictionaries or textbooks.

5.2 Recommendations for Further Studies

Due to the small corpus size, the findings of this study may not represent all the possible noun collocates of both synonyms that native speakers of English use. Furthermore, since the data were drawn from American English, the findings may not be generalized to other varieties of English. For further studies, increasing the number of concordance lines or corpus size will represent more noun collocates of both synonyms. Furthermore, other sources of corpus-based data may show some other different noun-noun collocations in addition to the Corpus of Contemporary American

English. For these reasons, there seems more likely to be generalized to other varieties of English.

5.3 Pedagogical Implications

The major findings are summarized as follows:

First, the possible noun collocates of *learner* are outnumbered by the possible noun collocates of *student* as a result of word frequency. Second, even *student* and *learner* share a similar meaning, they still differ in terms of collocations. In addition, there are less-preferred collocations of *learner* as *student* seems to have more possible collocates.

These findings bring some pedagogical implications for English language teaching in the following ways.

First, teacher should taught students to be aware of the differences of English synonyms which cannot be used interchangeably in all contexts. Second, teacher should provide students the authentic examples from corpus-based data as corpora offer students opportunity to access authentic language that they may not have in traditional materials, e.g., textbooks or dictionaries. Finally, teacher should instruct students how to access corpus-based data as students may take responsibility for their own language learning and get a boarder view of English language with the teacher's guidance.

REFERENCES

Boulton, A., Carter-Thomas, S., & Rowley-Jolivet, E. (Eds.). (2012). Corpus-informed research and learning in ESP. Amsterdam: John Benjamins.

Chambers, A. (2005). Integrating corpus consultation in language Studies. *Language Learning & Technology*. Retrieved from http://llt.msu.edu/vol9num2/pdf/chambers.pdf

Chung, Siaw-Fong. (2011). A corpus-based analysis of "create" and "produce". *Chang Kung Journal of Humanities and Social Sciences, 4*(2) (October 2011), 399- 425.

Davies, Mark. (2008-) The corpus of contemporary American English (COCA): 520 million words, 1990-present. Available online at https://corpus.byu.edu/coca/.

Ghaniabadi, S., Marjane, E. M., & Zareian, G. (2015). The use of adjective+noun collocations by Iranian EFL learners based on noun abstractness. *International Journal of Linguistics*, *7*(5), 86. https://doi.org/10.5296/ijl.v7i5.8143

Granger, S. (2002). A bird's eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 3–33). Amsterdam/ Philadelphia: Benjamins. https://doi.org/10.1017/CBO978 1107415324.004

Huang, C.-R., & Yao, Y. (2015). Corpus linguistics. In Wright, J. D. (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., pp. 949–953). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.52004-2

Hunston, S. (2006). Corpus linguistics. In Brown, K. (Ed.), *Encyclopedia of Language and Linguistics* (2nd ed., pp. 234–248). https://doi.org/10.1016/B0-08-044854-2/00944-5

Inkpen, D. (2007). A statistical model for near-synonym choice. *ACM Transactions on Speech and Language Processing*, *4*(1), 1–17. https://doi.org/10.1145/1187415.1187417

Inkpen, D., & Hirst, G. (2006). Building a lexical knowledge-base on near-synonym differences. *Computational Linguistics*, *32*(2), 223–262.

Inkpen, D., & Hirst, G. (2002). Acquiring collocations for lexical choice between near-synonyms. *Annual Meeting of the ACL*, *9*(July), 67–76. https://doi.org/10.3115/1118627.1118636

Jafarpour, A. A., Hashemian, M., & Alipour, S. (2013). A Corpus-based Approach toward Teaching Collocation of Synonyms. *Theory and Practice in Language Studies*, *3*(1). https://doi.org/10.4304/tpls.3.1.51-60

Khodadady, E., & Shamsaee, S. (2012). Formulaic sequences and their relationship with speaking and listening abilities. *English Language Teaching*, *5*(2), 39–49. https://doi.org/10.5539/elt.v5n2p39

Longman dictionary of contemporary English (2005). Essex: Pearson Education Ltd.

Moehkardi, R. (2002). *Grammatical and Lexical English Collocations*: Some Possible Problems to Indonesian Learners of English. Humaniora, volume XIV No 1.

Oxford advanced learner's dictionary (2010). Oxford: Oxford University Press

Pearce, Darren: 2001, 'Synonymy in collocation extraction', in *Proc. of the NAACL 2001 Workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations*, CMU.

Petcharat, N., & Phoocharoensil, S. (2017). A corpus-cased study of English synonyms: appropriate, proper, and suitable. *LEARN Journal*, *10*(2), 10-24.

Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences, 3*(10), 227-245.

Timmis, I. (2015). *Corpus linguistics for ELT*. New York: Routledge.

Xia, L. (2013). A corpus-driven investigation of Chinese English learners' performance of verb-noun collocation: A case study of Ability. *English Language Teaching*, *6*(8), 119–124. https://doi.org/10.5539/elt.v6n8p119

BIOGRAPHY

Name	Miss Kanittha Lapangdee
Date of Birth	July 29, 1991
Education Attainment	2014: Bachelor of Education, Silpakorn University
Work Position	English teacher at Nongyasaiwitthaya School, Suphanburi