USING CORPORATA TO AID IN LEARNING COLLOCATIONS

BY

FEI CHEN WEN

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Arts in the Teaching of English as a Second Language in the Graduate College of the University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Master’s Committee

Associate Professor Tania Ionin, Chair
Associate Professor Randall Sadler
Abstract

The potential of corpora, language databases comprised of authentic language materials from a variety of sources, has gradually trickled down to ESL and EFL classrooms (McCarthy, O’Keeffe, & Walsh, 2010) and has been associated with data-driven learning (DDL) where learners observe language patterns and improve language accuracy. This study examined whether the familiarity with corpora would improve learners’ preference to use corpora and the ability to solve V-N collocation problems, learners’ knowledge of V-N collocations and learners’ perception of the usefulness of corpus tools for solving collocation problems. Three groups of participants were recruited for three treatments: a one-time workshop combined with long-term in-class practices (TG), a one-time workshop only (CG), and no treatments (BG), and a pre- and post-test comparison was made to measure collocation knowledge gain of the three groups. Besides the post-test, the post-test with reference tools were given immediately after the post-test where the use of reference tools was allowed. The reference tools used on the post-test with reference tools and the performance of those who used corpora and those who did not was analyzed between the post- and post-test with reference tools. The results showed that those receiving treatments (TG and CG) that also chose corpora as reference tools were more successful than those who did not. Also, although collocation knowledge gain was not reflected in the score improvement from the pre- and post-test in all three groups, a correlation between the frequency of out-of-class use of corpora and pre- and post-test score improvement was found in TG, but not in CG nor BG. More positive perception of corpora as reference tools was found in TG than CG, but both TG and CG rated favorably for corpora’s utility to solve lexical collocation problems. This study contributes to the current literature by applying DDL to lexical collocations, specifically V-N collocations and by examining the factors influencing learners’ preference of reference tools, the ability to solve V-N collocation problems, potential knowledge gain of collocations and the perception of learners.
Table of Contents

I. Introduction .................................................................................................................................................. 1

II. Literature Review ....................................................................................................................................... 3

1. Types of Collocation ................................................................................................................................ 4

   1.1 Lexical Collocations: Verb-Noun collocation ...................................................................................... 6

   1.2 Source of collocation errors ................................................................................................................ 7

2. Corpora in The Writing Classroom: Self-correction and Data-Driven Learning ..................................... 8

   2.1 Corpora as a Reference Tool .................................................................................................................. 8

   2.2 Corpora and Data-driven Learning (DDL) ............................................................................................ 10

3. Research Questions ..................................................................................................................................... 13

III. Methods ..................................................................................................................................................... 14

1. Participants .................................................................................................................................................. 14

2. Procedure .................................................................................................................................................. 15

3. Treatment .................................................................................................................................................. 17

   3.1 Treatment A: Corpora for Collocations Workshop .............................................................................. 17

   3.2 Treatment B: Verb-Noun Class Exercises using Corpora .................................................................. 21

4. Materials .................................................................................................................................................. 22

   4.1 Pre- and Post-test and Pre- and Post-questionnaire ............................................................................ 22

   4.2 Pre- and Post-Test Item Selection ....................................................................................................... 25

      4.2.1 Academic Word List (ACL) ............................................................................................................. 25

      4.2.2 Content Validity ............................................................................................................................... 25

IV. Results ....................................................................................................................................................... 29

1. Predictions .................................................................................................................................................. 29

2. Analysis of Results .................................................................................................................................... 31

V. Discussion ................................................................................................................................................... 38

VI. Limitations ................................................................................................................................................. 51

VI. Conclusion and Implications ..................................................................................................................... 53

References ....................................................................................................................................................... 55

Appendix .......................................................................................................................................................... 60
A. pre-test questionnaire, pre-test .............................. 60

B. Post-test questionnaire, post-test, post-test with reference tools .......................... 66
I. Introduction

For second language learners, writing has long been considered a difficult and tedious task (Jafarpour, Hashemian, & Alipour, 2013). Nowalk (2010)’s study explored writing needs of college level students with intermediate or advanced proficiency and found that students struggle with accurate expressions of what they truly mean, and students themselves consider grammar along with vocabulary instructions the most beneficial for such problems in writing. Despite learner perception of the need of grammar instruction at college-level academic ESL courses, it is found that advanced and intermediate learners make very few errors in grammar; they, however, make plenty of collocation mistakes, i.e. word combination errors (e.g. Laufer and Waldman, 2011). Conventionally defined as word combinations in a continuum of varying degrees of arbitrariness (Kathpalia & Ling, 2008), collocations play a significant role in coherence and cohesion of language, which are essential for overall L2 mastery (Brashi, 2009), and perhaps even more so for writing.

Gledhill (2000, p.1) considers it “impossible for a writer to be fluent without a thorough knowledge of the phraseology of the particular field he or she is writing in.” While the importance of collocational competence for academic writing has been widely acknowledged, most ESL classes fail to go beyond arousing learner awareness of collocations (Nesselhauf, 2003). The current study aiming to teach collocations, specifically verb-noun collocations, adopts Data-driven Learning (DDL) as a part of graduate level ESL course in the United States. DDL uses corpora as a teaching/learning tool, also characterized as inductive and discovery learning through concordance lines queried from corpora, i.e. excerpted lines of authentic language data containing the
queried word (Daskalovska, 2013) Through DDL focusing on V-N collocations, the current study attempts to not only arouse learner awareness of lexical collocations, but also train learners’ skills to use corpora and enable self-learning. To examine the effect of the long-term (10 weeks) collocation learning through DDL, three groups were recruited: the treatment group (TG) received one workshop (class) of corpora use and weekly exercises for 10 weeks, the control group (CG) received only one workshop (class) of corpora use, and the baseline group (BG) received no treatment. The study aims to address the three broad research questions: Does familiarity with corpora improve learners’ preference to use corpora and the ability to solve V-N collocation problems? Does exposure to corpus tools improve learners’ knowledge of V-N collocations? Does exposure to corpus tools influence learners’ perception of the usefulness of corpus tools for solving collocation problems?
II. Literature Review

Collocations are restricted combinations that stand out to native speakers of English, but not so much by non-native speakers of English (Ackermann & Chen, 2013). They contain “some element of grammatical or lexical unpredictability or inflexibility” (Nation, 2001, p.324) that are central to proficient language use (Siyanova & Schmitt, 2008). As Siyanova and Schmitt (2008) pointed out, collocations are a recognized problem in L2 learners’ spoken and written language (Nesselhauf, 2003). Bahns and Eldaw (1993) indicate that collocations are a main source of advanced learner errors, and in Granger and Bestgen (2014) intermediate students are also found to overuse high-frequency collocations but underuse low-frequency but strongly associated collocations. Errors found in learner corpus analyses by Laufer and Waldman (2011) show that L2 learners experience difficulties with collocations regardless of their levels of proficiency, their native language, and types of tasks, further supporting Bahns and Eldaw’s (1993) view that learners’ collocational competence does not evolve along with vocabulary knowledge.

Collocation problems constitute one strand of writing difficulties of word choice (Leech, 1994). They are “complementary” word choice problems in contrast with “superficial” problems that are “single co-textual variable problems” (p.86). While ESL writing instructors may introduce aspects of vocabulary such as definitions, connotations and denotations, both teachers and students may not notice these more or less fixed combinations, as collocations are composed of frequent individual words that are usually transparent in meaning (Laufer and Waldman, 2011). Academic vocabulary specific to different disciplines, coupled with difficulties of collocations, could lead to poor quality
of academic writing, as vocabulary plays a tremendous role in writing (Nation, 2001). Below, the literature on the definition of collocations and common source and types of errors associated will first be discussed, and then the literature on the use of corpora for collocations in writing classrooms will follow.

1. Types of Collocation

Different definitions of collocations have been construed since Firth (1957) first proposed that the meaning of words should be derived from their co-occurrence with other lexemes in contexts (Hill, 2000). While there have been debates about the number of words collocations should consist of, ranging from two words to much longer lengths such as idioms, the current study adopts Benson et al. (1997)’s definition of lexical and grammatical collocations given that this definition has often been used as the basis of collocation-related research (Farrokh, 2012). The two types of collocations are mainly composed of two chunks. Grammatical collocations are composed of an open class word-nouns, adjectives, and verbs—followed or preceded by a preposition or grammatical structure such as an infinitive or clause. Lexical collocations are comprised mainly of combinations of open class words such as nouns, adjectives, verbs, and adverbs, although closed class words such as articles “the” and “a” can exist in between of the two open class words as syntactic rules require. Table 1 lists subcategories of grammatical collocations defined by Benson et al. (1997), and Table 2 lists subcategories of lexical collocations and select examples in Kuo (2009).
Type | Pattern | Example
--- | --- | ---
G1 | Noun + preposition | blockade against
G2 | Noun + to-infinitive | He was a fool to do it.
G3 | Noun + that clause | He took an oath that he would do his duty.
G4 | Preposition + noun | by accident; in advance
G5 | Adjective + preposition | fond of children
G6 | Adjective + to-infinitive | It is important to speak up.
G7 | Adjective + that clause | He was surprised that his dog ate all its dog food.
G8 | Verb of 19 types + a grammatical structure | eg. pattern A: the dative movement transformation verbs. eg. He sent the book to his brother (p.24). See Benson et al (1997) for detailed examples of various such category.

Table 1: Types of Grammatical Collocations in Benson et al. (1997) modified based on Alsulayyi (2015, p.33)

<table>
<thead>
<tr>
<th>Type</th>
<th>Pattern</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Verb + noun/pronoun</td>
<td>set a record</td>
</tr>
<tr>
<td>L2</td>
<td>Verb + noun</td>
<td>crush resistance</td>
</tr>
<tr>
<td>L3</td>
<td>Adjective + noun</td>
<td>strong tea</td>
</tr>
<tr>
<td>L4</td>
<td>Noun + verb</td>
<td>bombs explode</td>
</tr>
<tr>
<td>L5</td>
<td>Noun 1 + noun 2</td>
<td>a piece of advice</td>
</tr>
<tr>
<td>L6</td>
<td>Adverb + adjective</td>
<td>closely acquainted</td>
</tr>
<tr>
<td>L7</td>
<td>Verb + adverb</td>
<td>apologize humbly</td>
</tr>
</tbody>
</table>

Table 2: Types of Lexical Collocations in Benson et al. (1997) modified based on Kuo (2009, p.145).

Jafarpour, Hashemian, and Alipour (2013) pointed out that most recent research has endeavored to explore teaching of grammatical collocations, while much less is done on lexical collocations. However, Bahns and Eldaw (1993) indicated that lexical collocations are especially problematic for L2 learners. The study will focus on one type of lexical collocations listed above—L2: Verb-Noun collocations.
1.1 Lexical Collocations: Verb-Noun collocation

In the translation items in Bahns and Eldaw (1993), verb collocates (in contrast with nouns as bases) take up 23% of all the lexical words but they constitute 48% of all errors. Such observation reveals verb-noun combinations as a major source of difficulty for learners. Nesselhauf (2003) in an attempt to uncover learners’ verb-noun collocation difficulties through learner corpus analysis found verb-noun collocations as an area of highest rate of errors. For instance, an example of such can be “exert influence,” where exert can take quite a few other nouns such as control, pressure, authority, attention and power. This is because the verb in a collocation has a restricted sense, which makes its correct use more difficult if learners cannot fully distinguish subtle differences among verb collocates (Nesselhuaf, 2003). Another reason verb-noun collocations are difficult to acquire is that nouns in verb-noun collocations often carry salient semantic meanings, whereas verbs are often de-lexicalized, i.e. less semantically salient; examples that illustrate such phenomenon are make a mistake, where make is often used with different purposes in English and is semantically vague, and near-equivalence pairs such as lie and tell a lie, where tell does not reveal as much meaning as the noun lie or the verb lie itself (Boers et al., 2014). With semantic insignificance, Boers et al. (2014) noted that learners may not notice the verbs when exposed to verb-noun combinations, even for advanced learners. Nesselhuaf (2003), observing the same phenomenon, suggested the teaching of verbs in verb-noun collocations.
1.2 Source of collocation errors

The use of synonyms has been identified as a common source of collocation errors (Biskup, 1992 in Liu 1999). According to Chan and Liou (2005), Liu (2002)’s examination of lexical miscollocations found that 87% of miscollocation errors were of the V-N type and 56% of the lexical collocation errors were synonym errors. Phoocharoensil (2014) pointed out that learners’ misuse of synonyms in collocations show learners’ analogy of similar words. For example, when trying to say achieve an aim, learners may put reach in place of achieve. English synonyms are different from each other in subtle ways, one being different collocates with synonymous words. The example given earlier achieve an aim, could take reach as its collocate if aim is replaced with goal, a synonym of aim. McCarthy, O’Keeffe, and Walsh (2010) commented on such idiosyncrasy of synonyms: “it is often how the words collocate with other words that can show up differences” (p. 32). Inappropriate use of synonyms is used interchangeably in all contexts by learners (Biber et al., 1998 in Boers, Demecheleer, Coxhead, & Webb, 2014). This serves to corroborate Sinclair (1991)’s hypothesis that non-native speakers adopt open choice principle for collocations, rather than acquire them in chunks or as wholes. Boers et al. (2014) further explains difficulties as such can be due to formal similarities of synonyms, besides semantic ones. Given that V-N collocations are the most problematic area for learners, and synonym errors account for much of the errors, the current study designed treatments that target synonym mistakes frequently made for V-N collocations.
2. Corpora in The Writing Classroom: Self-correction and Data-Driven Learning

2.1 Corpora as a Reference Tool

While it is a conventional practice for language teachers to correct errors in writing tasks, the effects of feedback in error correction is not undisputed. According to Geiller (2014), some researchers argue that error correction is useless while others hold that it leads to more grammatical accuracy. This debate is central to the literature on effects of corrective feedback; however, effects of lexical feedback are relatively underdeveloped. The diversity of lexical errors, in contrast with rule-governed grammar errors, is likely the cause of such underdevelopment. While the debates remain unsettled, one thing is widely agreed—learners’ ability to self-correct and even self-identify errors, demonstrating the highest level of autonomy, is an ultimate goal of language teachers. As Todd (2001) pointed out, the ability to self-correct is a “global goal of language learning” (p.94). Advocates of autonomy in language learning hold that “the capacity for autonomy is innate but suppressed by institutional learning” (Benson, Grabe, and Stoller 2001, p.25). With an aim to cultivate learner’s autonomy, the current study trains learners to self-identity and correct V-N collocation errors during their writing process using corpora as reference tools, ie. language databases composed of authentic language data from native speakers of a language.

Reference tools such as the most traditional monolingual or bilingual dictionaries, have been the most common to assist learners in lexical errors. Nurmukhamedov (2015) examined the effectiveness of a paper dictionary, an online dictionary, and a corpus tool in helping learners with collocation productions and found online tools to be more suitable for learners’ needs. More recently, Google and corpora have become popular
reference tools for second language learning. Geiller (2014)’s study of how Google can be used to correct their “untreatable” errors, ie. lexical errors, revealed that although Google output can be “messy,” it is useful for learners that would like to “emulate the language of the press” (p.38). Comprised of authentic language data from a wide variety of genres, corpora facilitate the writing process as both “a composing and a revision tool” with their ability to allow learners to appropriately express their ideas (Quinn, 2015, p.165). Boulton (2009) pointed out the wide range of corpora used in his examination of 27 corpora-related studies. Two publicly available corpora: Corpus of Contemporary American English (COCA) and Word and Phrase Info would be introduced to learners of the current study as reference tools. COCA is the largest corpus of American English that has a nice balance of the following genres: spoken, fiction, magazine, newspaper, and academic. The data in the 5 genres are evenly distributed, with each genre including around 90 million words. Word and Phrase Info is an alternate COCA interface that provides information of the most frequent 20-30 collocates of the queried word and, lists out the synonyms of the query, and displays 200 concordance lines.

Corpora are especially useful for distinguishing between collocation environments of synonyms with its customizable commands. The concordance format of corpora enables users to notice “regularities in use that tend to remain unobserved when the same words or phrases are met in their normal contexts” (p.9.), thus an important tool for linguistic exploration of collocations (Hunston, 2002). While other reference tools such as the dictionaries and web tools other than corpora can be helpful, the customizable functions of corpora as well as their presentation of the output that allows for the generalization of patterns are valuable for learners as reference tools for lexical
collocations. The current study will examine whether corpora are more effective tools compared to other reference tools for V-N collocation problems. Besides, the learners’ perception of the effectiveness of corpora and whether there has been a change in the habit of the use of reference tools across time will also be discussed.

### 2.2 Corpora and Data-driven Learning (DDL)

Data-driven learning (DDL), which involves the use of corpora to train skills of analyzing a large amount of text, observing recurring patterns, and falsifying and confirming the patterns to form conclusions, is considered a form of inductive and discovery learning (Boulton, 2009). Todd (2001) investigated whether students can induce rules from corpora and self-correct errors, and the result confirmed learners’ ability. He coded student errors in their assignment drafts, instructed students to search in a self-selected corpus, had students compile language data they gather from corpus search, and required students to induce patterns based on the data compiled (see table 3 for sample of the exercise). Based on the inductive error-correction process in Todd (2001), Sun and Wang (2003) adopted the same cognitive process for error correction tasks as treatment for their inductive group and presented rules for the same tasks to their deductive group. The immediate post-test showed that the inductive group improved significantly over the deductive group. Chan and Liou (2005) designed 5 units of collocations and used concordancing in three units and non-concordancing in the other two. Their pre- and post-test included items that were taught in both concordancing and non-concordancing units, and they found that learners showed significant improvement on those taught in the concordancing units, whereas those in non-concordancing units did
not reflect learner improvement. While Boulton (2010) points out that relationship of ability to self-correct using corpora may not necessarily translate directly into effective learning “outcome,” research adopting similar cognitive process such as Sun and Wang (2003) and Chan and Liou (2005) demonstrate promising effects of learning of collocations through inductive learning with the help of concordance lines.

<table>
<thead>
<tr>
<th>For the body, the best scheme is the following linear organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based upon these findings to SESC following ISESS ’97 SESC determined the second survey was this second workshop</td>
</tr>
<tr>
<td>In the following top of text, page 2 and following we have summarized the major ideas to be a candidate for the following obstacle avoidance and corridor a control for mobile robots, The following procedures: representation of deals with methods for improvement of</td>
</tr>
</tbody>
</table>

Rules for *following*:

*Following* is used at start of sentences

*Following* is expanded noun word.

Work to be corrected: “*Following*, in this section we will explain some detail about GR101…” Following is followed by ‘comma’.

Correction: “*Following* we will explain some detail about GR101…”

Table 3: Todd (2001, p.97)’s sample of students’ inductive process

In a survey of 27 DDL focused empirical studies in Boulton (2010), he pointed out that empirical studies using DDL method often fail to mention of DDL as a constructive learning approach that promotes use of a wide array of cognitive skills and allows for practice of learning from authentic language data (Boulton, 2010). DDL as a discovery-based and inductive learning approach, is a “more natural approach than the
intellectually rigorous rule-based approach characteristic of much traditional teaching” and can “increase motivation where learners are allowed to pursue their own queries, leading to greater autonomisation and life-long learning” (Boulton, 2010, p.16). In the book edited by Leńko-Szymańska and Boulton (2015), Flowerdew highlights the lack of existing second language acquisition theories underpinning DDL. She examines existing DDL studies and highlights their focus on the noticing hypothesis proposed by Schmidt, constructivism theories (also associated with process-oriented approach), or sociocultural theory first proposed by Vygotsky based on their designs. Her examination reveals that several studies highlight the different “features” or “classroom practices” of how DDL can be utilized to increase learner noticing, adapt to different learning styles, or to achieve learning through interactions in class and with corpora. Although corpora sometimes intimidate teachers as well as students due to the excerpted output and the technical command inputs (Boulton, 2009), with the optimism of the capability of corpora-enabled DDL triggering noticing, allowing for individualization, and learning through interaction, the current study explores the effects of different familiarity with corpora with different treatments on learners’ preference to choose corpora and the ability to solve V-N collocation problems. Further, whether learners are able to internalize collocation knowledge and to distinguish between collocations of synonyms after different treatments will be explored.
3. Research Questions

With research findings showing positive effects of collocation learning through data-driven learning using concordance lines, the current research aims to examine effects of such practice on verb-noun collocations. Because learners are prone to producing synonyms of collocates to go with the base words as defined by formal linguists (Hausmann, 1989 as cited in Jaén (2007)) in which case the verb is the collocate and the noun is the base (Boers et al. 2014), specific process using corpora was taught to tackle such problem as treatments. The research questions are as detailed below:

RQ1. Does familiarity with corpora affect learners’ preferences for choosing corpora over other reference tools and/or learners’ ability in solving V-N collocation problems?

RQ2. Do learners gain knowledge of V-N collocations when they receive exposure to corpus tools?

RQ3. Does exposure to corpora influence learners’ perception about using corpora to solve collocation problems?
III. Methods

1. Participants

Participants were graduate students at the higher level, ESL501, of the two-level ESL course sequence at University of Illinois at Urbana Champaign. At the time of the study, students admitted to UIUC graduate programs were required to take ESL courses if their TOEFL IBT score ranges from 80-102 out of 120 during the time of admission, unless they are proficienced out of the course through a diagnostic test given in the second class.

Three groups of participants were recruited from ESL501. The treatment group (TG) consisted of 14 students from one section of ESL501 courses. The control group (CG) comprised 8 students from another section of ESL501 courses. 12 students were recruited across ESL501 sections in the baseline group (BG), but only 7 completed the post-tests. Therefore, the data of the 5 participants from BG who failed to complete the whole procedure were excluded. Participants native language background and the years of English learning and education is shown in table 4 and 5 below.
<table>
<thead>
<tr>
<th>Participant Gender and Native Language/ Groups</th>
<th>TG (N=14)</th>
<th>CG (N=8)</th>
<th>BG (N=7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/Female</td>
<td>8/6</td>
<td>5/3</td>
<td>6/1</td>
</tr>
<tr>
<td>Mandarin Native Speaker</td>
<td>10</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>European Languages (Spanish, Italian)</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Southeast Asian Languages (Thai, Indonesian, Konkani, Bengali)</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Other Languages (Turkish, Korean, Kazakh)</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 4. Participants’ Gender and Native Languages

<table>
<thead>
<tr>
<th>Participant Background/ Groups</th>
<th>TG (N=14)</th>
<th>CG (N=8)</th>
<th>BG (N=7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/Female</td>
<td>8/6</td>
<td>5/3</td>
<td>6/1</td>
</tr>
<tr>
<td>Average years of English learning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (years)/ Std</td>
<td>13.32/5.52</td>
<td>11.34/6.10</td>
<td>6.86/2.80</td>
</tr>
<tr>
<td>Average years of stay in an English-speaking country</td>
<td>1.17/0.62</td>
<td>3.96/3.82</td>
<td>0.73/0.12</td>
</tr>
<tr>
<td>Avg. years of English Learning in higher education in the United States</td>
<td>1.24/0.65</td>
<td>2.60/2.90</td>
<td>0.60/0.20</td>
</tr>
</tbody>
</table>

Table 5. Participants’ English Learning and Education Background

2. Procedure

The study took place in the spring of 2015 semester, which lasted for 15 weeks.

In week 3, all participants were given the pre-test questionnaire and the pre-test. In week 13-14, after the third major assignment on research paper writing was due, all participants were given the post-test questionnaire, the post-test, and the post-test with reference tools (see appendix for the questionnaires and tests).
In week 4, a 50-minute workshop on Corpora for Collocations (Treatment A) was given to TG and CG. Between weeks 5-12, participants in TG were asked to turn in 2 self-selected verb-noun combination problems for class exercises (Treatment B). All participants in TG were given an average of 6 questions in the beginning of class for class exercises. Class exercises lasted approximately 30 minutes on a weekly basis. The distinction of the treatments TG and CG received and the zero treatment of BG is designed so the effects of the different amounts of exposure of corpora can be examined in terms of the learners’ preference to choose corpora and the ability to solve V-N collocation problems (RQ1), the knowledge gain of collocations (RQ2), and the learners’ perception of corpora (RQ3). A flow chart of treatments given to different groups can be found in Table 6 below.

<table>
<thead>
<tr>
<th>Week 3</th>
<th>Treatment Group (TG)</th>
<th>Control Group (CG)</th>
<th>Baseline Group (BG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-test questionnaire and pre-test</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Week 4</td>
<td>Treatment A: 50-minute workshop</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Week 5-12</td>
<td>Treatment B: in-class practice</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Week 13-14</td>
<td>post-test questionnaire and post-test, post-test with reference tools</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Procedure and Treatments given to TG, CG, and BG
3. Treatment

3.1 Treatment A: Corpora for Collocations Workshop

A fifty-minute workshop on using corpora to solve collocation problems was given to TG and CG. This treatment provided both groups the one-time exposure to corpora of the same quality and was the only treatment CG received, whereas TG further received treatment B. The first 5-10 minutes were spent on awareness raising and introduction of collocation types. Then two 20 minute activities were designed for hands-on use of COCA and Word and Phrase Info as reference tools. The activities started with fill-in-the-blanks passages that included comprehensive lexical collocation types (adjective+noun, adverb+verb, etc.) as departure points for the following collocation queries with COCA and Word and Phrase Info. Time allotted for the workshop is tabulated in Table 7.

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 mins</td>
<td>Collocation Awareness-Raising</td>
</tr>
<tr>
<td>5 mins</td>
<td>Collocation Types</td>
</tr>
<tr>
<td>20 mins</td>
<td>Fill-in-the-blanks; Word and Phrase Info Intro</td>
</tr>
<tr>
<td>20 mins</td>
<td>Fill-in-the-blanks; COCA Intro</td>
</tr>
<tr>
<td>2 mins</td>
<td>Q&amp;A</td>
</tr>
</tbody>
</table>

Table 7. Collocation Workshop Activities

In the 20-minute activities introducing COCA and Word and Phrase Info, participants read example passage 1 with collocates left as blanks in the cloze format and
were instructed to come up with their own answers. An excerpt of passage 1 is shown below:

___(1)___(adj. +) evidence has shown that, despite its ___(2)___(adj. +) environmental impacts

Then they were introduced to using Word and Phrase Info, where information of frequency of the query used in different genres, 20-30 collocates categorized by parts of speech, a list of synonyms, and 200 concordance lines from COCA were presented, to check if their answers were listed in Word and Phrase Info. See Figure 1 for a sample of Word and Phrase Info Interface.

![Figure 1. Sample of Word and Phrase Info Interface](image)

Participants were told to use Word and Phrase Info to check if the answers they formulated were “appropriate” based on Word and Phrase Info. Those that participants came up with as answers and were listed on Word and Phrase Info were elicited from participants and reinforced as appropriate combinations. Then participants were told to
ignore the answers not listed on *Word and Phrase Info* after checking Word and Phrase Info at the stage, and the class continued with reading of passage 2 and filling in the blanks for the second passage. This is for participants to notice insufficiency of *Word and Phrase Info* as the only tool to solve specific collocation problems they encounter and to understand different functionalities of *Word and Phrase Info* and *COCA*.

After participants finished passage 2, two functions in *COCA* were then introduced to participants. First, they were shown to use symbols for different parts of speech that collocate with the base. This is the Collocates Function on *COCA*, where users specify parts of speech using the drop down menu from the POS list, and the range of the occurrence of collocates can also be narrowed down. The maximum range of the part of speech of collocates near the query is 9 words before the queried word and/or 9 words after. See Figure 2 for a sample of *COCA* interface where the Collocates function can be used.

![Figure 2. Function of Collocates on COCA](image.png)

The Collocates function enables users to see the environment where the queries exist and the words of specific parts of speech filtered. However, the meanings of the collocations may be varied and not the ones users are looking for. Participants were taught a more refined search skill--to use the synonym function--if they had a word in
mind but were not sure if it actually collocates with the base, or if they knew it collocates with the base but wanted to use a more specific, precise, or even more natural word as the collocate. Using the synonym function, users are able to narrow down to collocations of specific meanings. See Figure 3 for a sample of command of the synonym function and returned results.

In summary, participants receiving the workshop were instructed to refer to Word and Phrase Info for information of common collocates, list of synonyms, and concordance lines of the query, as well as COCA for collocates of a specific part of speech they are looking for their base words, and using synonym function on COCA to find specific meanings they aim for.
3.2 Treatment B: Verb-Noun Class Exercises using Corpora

Treatment B provides more exposure to corpora to TG and is where DDL takes place on a regular basis in class. It is hypothesized to be the variable that may 1.) influence learners’ preference of reference tools and performance to use corpora in contrast with CG and BG (RQ1), 2.) lead to knowledge of V-N collocations given the DDL experiences (RQ2), and 3.) foster a positive perception of corpora use (RQ3). A week after the Corpora for Collocations workshop, participants in TG were asked to submit 2 verb-noun combinations they had doubts about while doing assignments. They were asked to first indicate what verb and what noun it is that they had doubts about, and then they were asked to put down the sentence they were trying to phrase. The verbs of the submitted items were then turned into blanks for in-class practices to elicit answers from all participants on Google form. However, in the first and second times of class exercises, participants pointed out that some sentences lacked enough contextual clues for them to fill in the answers as sometimes ideas may have flowed from previous sentences in the paragraph. After the first two exercises, the questions were given along with the verbs and nouns that the submitter intended to use, but a question mark was put after the verb, serving as a cue for a trouble source that needed working on. For example:

E.g. Satoshi ________ a study on analyzing whether birth order has genuine causal effect on general intelligence.

conduct? study

Participants were told that they could use the cue as the answer if they believed that is the most appropriate verb for the combination. In the above example, they may use “conduct”
as the answers they render most appropriate and keep it as is, or they may come up with
other verbs if they think “conduct” is not appropriate for the V-N combination.

In each exercise, participants answered all 6 questions. After they write down
individual answers to the questions, all responses were collected and published to the
participants. In pairs or in groups depending on how the class is structured for the lesson
of the day, participants were assigned one question to work with. They followed the step
by step instructions of using Word and Phrase Info to see if any answers submitted were
acceptable. If the answers submitted were not listed, they then consulted COCA to solve
the problems. Participants marked acceptable or unacceptable for the submitted answers,
but the final step of the exercise required them to list three most frequent verbs used in
the same sense for the verb-noun combination. They were asked to decide the best verb(s)
for the combination and provide their reasoning. Then all participants were asked to
check each group’s selection and justification of the best verb collocates against the
answers they originally put down for themselves in the beginning of the in-class practice.
This step allowed individualized feedback for all 6 exercises even though participants
only worked with one of the six exercises. If time allowed, the instructor went through
the answers with the class and prompted them to explain in more detail if the
explanations were not clear enough.

4. Materials

4.1 Pre- and Post-test and Pre- and Post-questionnaire

The current study utilized a pre- and post- collocation test based on Academic
Word List formulated by Ackermann and Chen (2013), along with pre- and post-
questionnaires. The pre- and post- tests are in the multiple choice format and contain the
same test items with shuffled item orders and item options. The multiple choice questions
require participants to choose the best 2 out of 5 as the answers to the questions. That is,
there are two answers to each question, and the remaining three options are considered
incorrect (see below for an example). This is because although from frequency of
occurrences in corpus, one collocate can be more frequently used than others with a base,
the "most frequent" combination does not necessarily render the lesser frequent
combinations wrong or inappropriate; synonymy can be viewed as a cline of sameness of
synonym words which take similar and the overlapping augments and those that are
much less the same that do not allow the same augments to be taken. Behavioral-profiling,
a method to observe distributions of words in relation to other words and in its
constructions using corpora, has proven that both methodologies are used to shed light on
fine-grained semantic similarities of words (D. Liu, 2013). Therefore, synonym verbs
preceding same base nouns listed on ACL are used as correct answers. For other items
where ACL has only one V-N combination, corpora are consulted for a second best (or in
some cases, more frequent used) options.

Eg. Examiners typically _______ data from multiple sources,

including records, attorneys, caretakers, and the youth.

(A) collect    (B) receive    (C) obtain    (D) gain    (E) attain

Ans: (A)(C)

The post-test is followed by the post-test with reference tools, which contained the same
test items as the pre- and post-test and had the same order of the post-test, but in the post-
test with reference tools, participants were allowed to use reference tools of any kind,
such as online dictionaries, Google or corpora. Although it would be ideal to require learners in TG and CG to use corpora on post-test with reference tools to examine learners’ performance to use corpora for V-N collocation problems given different exposure of corpora, the login process and corpus skills required may cause delays for the tests. The time designed for the post-test with reference tools may end up being spent on the researcher helping learners with technical problems. The selection of reference tools was thus not restricted and left to the learners’ choice, but learners were asked to indicate in the end of the post-test with reference tools what reference tool they used for the test.

The pre-test questionnaire contained the same items for all three groups; the post-test questionnaire contained overlapping items with the pre-test questionnaire to gauge change of the habit of reference tool use, and slightly different questions were included depending on treatments participants underwent. The pre-test questionnaire included questions about a.) personal information, b.) vocabulary help during writing-dictionary, c.) vocabulary help during writing-other tools The post-test questionnaire basically included the same questions for all TG, CG and BG, but an extra section, section d). experiences with using corpus for word choice, asked about participants’ habit of use of corpora, the frequency of corpora use, the perception of effectiveness of corpora for V-N problems, and the intention of future use was included for TG and CG. Finally, questions 25-26 in section d on the effectiveness of class practices were specifically included for TG (See appendix for the questionnaires and tests.)
4.2 Pre- and Post-Test Item Selection

4.2.1 Academic Word List (ACL)

Academic Collocation List (ACL) is appropriate as the basis of the collocation test for the current study, as it includes lexical collocations and is formulated with the general academic population in mind. It used a mixed method of corpus examination and expert consultation in selecting items to be included on Academic Collocation List. What distinguishes ACL from previous academic lists such as Durrant’s (2009) is that it consists mainly of lexical collocations instead of grammatical ones. Lists containing frequent “formulas” instead of collocations such as Durrant’s (2009), are mainly composed of one closed-class word and one open-class word, which did not yield much pedagogical value (Ackermann & Chen, 2013). Ackermann and Chen (2013) thus identified the need of a collocation list that provides lexical collocations, i.e. two open-class components, and contended that such list be conceived based on frequency of word occurrences in corpora as well as expert judgment for the most pedagogical value.

4.2.2 Content Validity

Basing its test items on the 311 V-N collocations on ACL, the current study does not select academic collocation items from scratch. However, given the fact that academic collocation learning is just a component of the whole academic writing course, time allotted for both the pre- post- tests and questionnaires had to be restricted to at most half a class period (30 minutes) in order to control reliability through the monitored environment of the test in class whenever possible. The problem for the test design then, came down to which items out of the 311 should be selected. In other words, which items
and which distracters should be chosen to construct a valid test for the purpose of the current study. In his attempt to craft a valid and reliable English adjective-noun collocation test for L1 Spanish learners, Jaén (2007) first filtered through the most frequent nouns (the base) from corpora for item selection. Then to reduce items to a reasonable number, he sorted frequency of the nouns, followed by the sorting of the most frequent adjective collocates. However, the current study did not follow the same process as Jaén (2007); as the most frequent and pedagogically valuable academic collocations have been pre-selected in the Academic Collocation List, the current study focused on accommodating as many verbs in the 311 verb-noun entries on ACL as possible for its pre- and post-tests. Thus more weight was given to the verbs (the collocates) in the selection process, then the nouns. That is, the most frequent collocates (the verbs) on ACL were firstly included, then the most frequent bases (the nouns).

Among the 311 V-N collocation entries, there were repeating verbs and nouns. They verbs were firstly sorted based on the frequency of occurrences (see Table 8 for an example of the sorting.) A pilot test on the items was given to three native-speakers of English. The items whose two target answers were not both chosen were ruled out. The final item selection included all verbs that appeared three or more times either as the target answer or the distracters with exceptions of “publish”, “convey”, and “become”\textsuperscript{1}. Nouns that appeared five or more times were included as base words with exceptions of

\textsuperscript{1}These exceptions, although appearing three or more times, are rather semantically salient or hard to be replaced with other verbs, making the selection of the second answer that is also a synonym difficult, if not impossible.
“evidence.” The selection process yielded 21 items for the verb-noun collocation test, with 30 different verbs including 9 pairs of synonymous verbs both listed as collocates with their respective nouns, and 21 nouns from ACL.3

<table>
<thead>
<tr>
<th>The Academic Collocation List</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>Addition</td>
<td>Component</td>
<td># of occurrences</td>
<td>Component II</td>
<td># of occurrences</td>
</tr>
<tr>
<td>875</td>
<td>gain</td>
<td>access</td>
<td>3</td>
<td>6 (to)</td>
<td></td>
</tr>
<tr>
<td>876</td>
<td>gain</td>
<td>information</td>
<td>3</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>877</td>
<td>gain</td>
<td>insight</td>
<td>3</td>
<td>4 (into)</td>
<td></td>
</tr>
<tr>
<td>878</td>
<td>gather</td>
<td>data</td>
<td>2</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>879</td>
<td>gather</td>
<td>information</td>
<td>2</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>909</td>
<td>give</td>
<td>access</td>
<td>10</td>
<td>6 (to)</td>
<td></td>
</tr>
<tr>
<td>910</td>
<td>give</td>
<td>consent</td>
<td>10</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>911</td>
<td>give</td>
<td>consideration</td>
<td>10</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>912</td>
<td>give</td>
<td>emphasis</td>
<td>10</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>913</td>
<td>give</td>
<td>evidence</td>
<td>10</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>914</td>
<td>give</td>
<td>feedback</td>
<td>10</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Table 8. An Example Sorting by occurrences of verbs and nouns of V-N combinations in ACL

As already mentioned, ACL itself included collocations that are “synonym collocations,” i.e. collocations that are likely interchangeable in many cases. Given that each question requires learners to choose two answers, such synonym collocations (9 out of 21 items) were used in the same test item as target answers. For example, item 1 in Table 9 is an example of “collect” and “obtain” as the correct collocates with the base “data.” Both verbs were listed as frequent collocates with “data” on ACL and can be

2 The exception of the noun is due to the fact that another selected noun “insight” has the same collocates of “give” and “provide” with “evidence” but “insight” (freq=9893) is a lower frequency word based on COCA than “evidence” (freq=87692) and considered more worthy for testing given the students are at the graduate level.

3 Only 37 entries among the 311 verb-noun collocations on ACL contained verbs and nouns that were both unselected for the test, which means that the other 274 verb-noun collocations had either the verb selected as collocate options or the noun included in the test question as the base word or had both the verbs and nouns selected. Note that the verbs in the 37 entries unselected appeared on the ACL one or two times, and the nouns one to four times, thus counted as separate entries. 3 out of the 37 entries contained phrasal verbs or verb followed by prepositions: “carry out,” “contribute to,” and “engage in,” and they were intentionally excluded from the test as the treatment of the test did not address such verb types.
considered synonym pairs and thus selected for the pre- and post-tests as target answers (see table 9 for an example of the test items where blue highlights target answers derived from ACL). For the other 12 items of the test where only one collocation is listed on ACL, the second target answer was firstly derived from collocates listed on *Word and Phrase Info*. If no collocates was considered to match the meaning of the context, *COCA* was then referenced. As for choice of synonyms as distracters for the multiple choice questions, since the current study aims to investigate learner ability to distinguish between synonyms that actually collocate with the bases, synonyms provided in *Word and Phrase Info* interface were firstly selected if considered distracting enough, i.e. the combination was not too obviously impossible and if the word is not too difficult; then if there were still not enough distracters, functions on *COCA* were performed to derive more synonyms that could be used as distracter following the above two criteria, namely, that the combinations were not considered obviously impossible and the difficulty of the words selected was not beyond learners’ capacity.

Table 9. Sample test items with target answers and distracters (blue means indicates verbs are on ACL, green indicates the verbs are derived from *Word and Phrase Info* and *COCA*, red highlights distracters.)
IV. Results

1. Predictions

In order to answer RQ1, “Does exposure to corpora affect learners’ preferences for choosing corpora over other reference tools and/or learners’ ability in solving V-N collocation problems?” two analyses were performed. Firstly, the reference tools indicated to have been used by all three groups at the end of the post-test with reference tools will be tabulated. It is expected that TG would have the most participants choosing corpora than CG, and none in BG. For those using corpora in CG, it is predicted that they had been using corpora outside of class. In short the number of participants preferring to use corpora for V-N collocations is expected to be TG>CG>BG. Secondly, based on the selection of reference tools, the performance of score improvements between the post-test and post-test with reference tools will be based on those using corpora (the corpora-using group) and those who did not (the non-corpora-using group). It is predicted that the corpora-using group would be more successful than the non-corpora-using group given that corpora are taught specifically for lexical /V-N collocations and corpora are useful for observations of language patterns. In short, the comparison of the performance is corpora-using group> non-corpora-using group.

To answer RQ2 “Do learners gain knowledge of V-N collocations when they receive exposure to corpus tools?” two analyses will be conducted. Firstly, the score improvements from the pre- and post-tests for all three groups will be performed. It is expected that TG would show signs of knowledge gain because they may have incidentally practiced collocation combinations that overlap with the test items, or that in
the process of treatments, their collocation knowledge was expanded because of the
treatment requiring them to self-correct. CG may exhibit knowledge gain if they had been
using corpora on their own after the one-time treatment. BG not receiving any treatments,
is not expected to show knowledge growth of V-N collocations. In short, TG is expected
to show more knowledge gain than CG and BG, with CG’s outcome of knowledge
growth undetermined. Therefore, the comparison of knowledge gain is predicted to be
TG>CG>BG or TG>CG=BG. Secondly, a correlation analysis between the frequency of
out-of-class use of corpora and the score improvement between the pre- and post-tests for
all three groups will be performed. This is because the out-of-class use of corpora is an
intermediary factor derived from the two treatments that could potentially lead to
knowledge gain. If out-of-class use of corpora proves to be an important factor for
knowledge gain, it is predicted that TG and CG’s out-of-class frequency use of corpora
may demonstrate a correlation with score improvements between the pre- and post-test as
more exposure can potentially strengthen learners’ knowledge growth, but given TG and
CG’s different treatments, the strength of the correlation is predicted to be stronger in TG
than CG. This is because TG is believed to be the most familiar with corpora and the
more they use corpora outside of class, the more knowledge gain there will be, whereas
CG not being as familiar with corpora may only have limited knowledge gain. For BG,
no prediction can be made since they did not receive exposure to corpora and thus no out-
of-class use existed. In short, the correlation of out-of-class use of corpora and score
improvement between the pre- and post-test is predicted to be TG >CG>BG.

For RQ3 “Does exposure to corpora influence learners’ perception about using
corpora to solve collocation problems?” two analyses will be performed. Firstly, the data
in section \( d \) (see appendix B) of the post-test questionnaire will be tabulated, along with the question (Q14) in both the pre-test questionnaire and the post-test questionnaire asking about the reference tool learners use the most frequently other than dictionaries. It is predicted that from the above, TG would feel more positively than CG about the use of corpora by indicating that they had been using corpora since the treatments, and that they would continue using corpora in the future. CG is expected to feel a little less positively than TG given the shorter treatment (treatment A only). Therefore, the perception about using corpora to solve collocation problems is predicted to be TG>CG. Secondly, an analysis of Q25 and Q26 included only for TG on the experiences and perception of the effectiveness of the aspects of the in-class treatment will be conducted. It is predicted that TG would feel positively about both by stating positive experiences and giving higher ratings (4 or 5 out of 5).

2. Analysis of Results

To answer RQ1, those who indicated the use of corpora as a major tool at the end of the post-test with reference tools and those who did not, were grouped into corpora-using and non-corpora-using groups respectively. The composition of the two groups from TG, CG, and BG are shown in table 10, and the reference tools indicated to have been used in post-test with reference tools is tabulated in table 11. As predicted, TG has the most users of corpora, CG the middle, and BG none. A paired t-test was conducted to compare the score improvements of post and post-test with reference tools of the two groups. The result confirms that prediction that the corpora-using group demonstrated a statistically
significant result $t(15) = 4.89, p = 0.0001$, whereas the non-corpora-using group did not $t(12) = -.57, p = 0.14$.

<table>
<thead>
<tr>
<th>tool/group</th>
<th>corpora-using group</th>
<th>non-corpora-using group</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG (N=14)</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>CG (N=8)</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>BG (N=7)</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 10. The Composition of the Corpora-using Group and Non-corpora-using Group

<table>
<thead>
<tr>
<th>tool/group</th>
<th>non-corpora-using group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>5</td>
</tr>
<tr>
<td>non-use of tools</td>
<td>0</td>
</tr>
<tr>
<td>other tools</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 11. The Reference Tools Used by the Non-corpora-using Group

To examine RQ2, a mixed 3X3 ANOVA was conducted, with one within-subjects factor (test, 3 levels) and one between-subjects factor (group, 3 levels). Both the results of the main effect of the groups $F(2, 81) = 1.69, p = 0.19$ and tests $F(2, 81) = 1.97, p = 0.15$ were statistically non-significant. The interaction of the two factors were also non-significant, $F(4, 81) = 1.813, p = 0.13$. However, as the sample sizes are small, it is possible that effects are not detected in the analysis. A repeated-measures ANOVA on the effect of tests is thus performed within each individual group to compare the effect of the
treatments on pre-test, post-test, and post-test with tools. The analysis of the treatment group (TG) revealed that there was a significant effect between tests, whereas for the control group (CG) and baseline group (BG), there was no significant effect (see table 12 below).

<table>
<thead>
<tr>
<th></th>
<th>TG</th>
<th></th>
<th>CG</th>
<th></th>
<th>BG</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DF</td>
<td>F</td>
<td>p-value</td>
<td>DF</td>
<td>F</td>
<td>p-value</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>-----</td>
<td>---------</td>
<td>-----</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>2, 26</td>
<td>6.10</td>
<td>0.01*</td>
<td>2, 14</td>
<td>0.14</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>2, 12</td>
<td>0.20</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 12. Results of Repeated-Measures ANOVA of TG, CG, and BG

As significant difference is only found in the Repeated-Measures ANOVA in TG, a post-hoc analysis for this group is further conducted in order to pin down on which test performances actually differed in statistical terms. The post-hoc Tukey test showed that TG’s performance on pre-test did not differ from post-test (p=0.99), whereas that on the post-test with reference tools differed significantly from post-test at p = 0.04, and the performance on the pre-test also differed significantly from post-test with reference tools at p = 0.05. The prediction on TG having more knowledge gain than CG is thus rejected.

Although the treatments TG and CG received controlled for the variable of long-term in-class exposure of corpora, the out-of-class exposure has to be considered as an intermediary factor. The frequency of the out-of-class use of corpora self-reported in the post-test questionnaire is tabulated as table 13 below.
<table>
<thead>
<tr>
<th>Groups</th>
<th>TG</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every Week</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Every Two Weeks</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Every Three Weeks</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Every Four Weeks</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>None</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 13. learners’ self-reported (out-of-class) frequency of corpora use

A Pearson product-moment correlation coefficient was computed to evaluate the relationship between the frequency of out-of-class corpora use and the score improvement between the pre- and post-test. The result showed that TG had a moderate-strong positive correlation between the frequency of corpora use and score improvement between pre- and post-test with reference tools \( r=0.65, p=0.01 \) and between the frequency of out-of-class corpora used and pre- and post-test score improvement \( r = 0.51, \ p = 0.05 \), but no correlation was found between the frequency of out-of-class corpora use and post-test and post-test with reference tools. The computations of CG’s reported frequency of out-of-class use of corpora and the improvement of test score improvements all showed non-significant correlations (see table 14 below). The result confirmed the prediction that TG’s out-of-class use of corpora is correlated with the score improvement between pre- and post-test but rejected the prediction that there is a correlation in CG.
To answer RQ3, whether there is change in their preference to use corpora as reference tools throughout the semester was examined. The answers to Q14 in pre- and post-test questionnaire asking learners what reference tools other than dictionaries they used most often was analyzed and presented in Table 15. Contrary to the prediction, TG did not choose to use corpora as the most frequently used reference tools after treatments; CG also did not take up using corpora after the treatment.

<table>
<thead>
<tr>
<th>Group/Tool</th>
<th>Google</th>
<th>Corpora</th>
<th>Others (Thesaurus, etc.)</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG (N=14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 3</td>
<td>12</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(86%)</td>
<td>(14%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 13</td>
<td>11</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(79%)</td>
<td>(14%)</td>
<td>(7%)</td>
<td></td>
</tr>
<tr>
<td>CG (N=8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 3</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(75%)</td>
<td>(25%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 13</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(75%)</td>
<td>(12.5%)</td>
<td>(12.5)</td>
<td></td>
</tr>
<tr>
<td>BG (N=7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 3</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(72%)</td>
<td>(14%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 13</td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(86%)</td>
<td>(14%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 15. Most Frequently Used Reference Tool Other Than Dictionaries (Q14)
Further analysis of learners’ perception of corpora effectiveness was based on section d “Experiences with Using Corpus for Word Choice” of the post-test questionnaire.

Questions 25-26 were included only for TG to find out TG’s general perception of effectiveness Q25: “Describe your experience using Corpus in ESL501D. Did you like the exercise in class? Why or why not?” and the different aspects’ effectiveness of the in-class practices Q26: “Rate effectiveness of in-class word exercises using corpus in the following aspects- submit combination difficulties; step-by-step instructions; instructor help; peer help; group presentation of best choices.” Based on a scale of 5, Q26’s result is presented in table 16. As predicted, TG gave high ratings to the aspects of the in-class practices (treatment B) with only7% giving scores lower than 3 out of 5 for “step-by-step instructions,” “peer help,” and “group presentations of best choices.”

<table>
<thead>
<tr>
<th>Aspect/Rating</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>submit a combination</td>
<td>9</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(64%)</td>
<td>(36%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>step-by-step instructions</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(64%)</td>
<td>(29%)</td>
<td>(7%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>instructor help</td>
<td>9</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(64%)</td>
<td>(36%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>peer help</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(36%)</td>
<td>(57%)</td>
<td>(7%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>group presentations of best choices</td>
<td>9</td>
<td>4</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>(64%)</td>
<td>(29%)</td>
<td></td>
<td></td>
<td></td>
<td>(7%)</td>
</tr>
</tbody>
</table>

Table 16. TG’s perception of effectiveness of aspects of in-class practices

Other questions in the section were included in both TG and CG’s questionnaire, asking whether they had been using corpora since treatment A (Q17), what they mainly used corpora for (Q19), the perception of effectiveness of corpora for lexical versus other problems (Q21-22), and the intention of future use (Q28) were analyzed and tabulated in tables 17-19. Confirming the prediction, more TG indicated the use of corpora and the
intention of future use than CG. Also, in terms the perception of corpora’s functionality to help with lexical collocation problems, although TG and CG both felt positively about the effectiveness of corpora, TG seemed to feel more positively (93% giving rating higher than 3) than CG (75% giving rating higher than 3). Results of the questions in the same section with an open-ended question format will be discussed in the discussion section.

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG (N=14)</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(92%)</td>
<td>(8%)</td>
</tr>
<tr>
<td>CG (N=8)</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>(50%)</td>
<td>(50%)</td>
</tr>
</tbody>
</table>

Table 17. The Use of Corpora Since The 50-minute Workshop (Q17)

<table>
<thead>
<tr>
<th>Groups/Rating</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG (N=14)</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(50%)</td>
<td>(36%)</td>
<td>(14%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG (N=8)</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(12.5%)</td>
<td>(12.5%)</td>
<td>(75%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 18. The Likelihood to Use Corpora in the Future (Q28)

<table>
<thead>
<tr>
<th>Groups/Rating</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG (N=14) lexical collocations</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(43%)</td>
<td>(50%)</td>
<td>(7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TG (N=14) others</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7%)</td>
<td>(50%)</td>
<td>(36%)</td>
<td>(7%)</td>
<td></td>
</tr>
<tr>
<td>CG (N=8) lexical collocation</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(12.5%)</td>
<td>(62.5%)</td>
<td>(25%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG (N=8) others</td>
<td></td>
<td>3</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(37.5%)</td>
<td>(50%)</td>
<td>(12.5%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 19. The Perception of Effectiveness of Corpora for Lexical Collocation Versus Other Problems (Q21-22).
V. Discussion

RQ1. Does familiarity corpora affect learners’ preference to choose corpora as reference tools and the ability in solving V-N collocation problems?

The result showed that the corpora-using group consisting of TG and CG was able to solve V-N collocation problems, whereas the non-corpora group consisting of CG and BG did not show a statistical significant result. Different exposure to corpora seems to have influenced learners’ preference in choosing corpora as the reference tools. The success of the corpora-using group also reflects that the exposure to corpora is decisive for learners’ ability to solve V-N collocation problems.

Although other reference tools than corpora may be effective for solving V-N collocation problems, the result shows that the corpora-using group was successful whereas the non-corpora-using group was not. Given that those in the corpora-using group all received exposure to corpora (either treatment A or treatment A and B), the exposure proves beneficial to learners in using corpora to solve collocation problems. At the same time, it is worth examining what tools the non-corpora-using group used and discussing the potential reasons they were unhelpful for V-N collocation problems.

Nearly half of the learners in BG and CG groups chose Google as the reference tool for V-N collocation problems. Given their lack of score improvement, Google as a reference tool for V-N collocations (or even lexical collocations in general) thus, does not seem effective. Some of the most commonly seen search results Google returns are dictionary results. Geiller (2014) pointed out the insufficient information provided in dictionaries as well as online dictionaries, which could be the source of difficulty to solve V-N collocation problems. Other search results may be part of a paragraph or discourse with
the lexical chunks students put in the search box. In examining untreatable errors, ie. lexical errors, Guiller (2014) utilized Google to help learners with lexical errors and suggested that search results enable students to replace inaccurate expressions by taking the language chunks found in the searches, especially for writings that involve news language, namely the expressions that “catch up with world events” (p.40). Although the news genre is considered academic English, the specific genre commonly found through Google may be insufficient for a wider-range of academic V-N collocations. Also, since the current study target synonym errors, although learners may have found collocation combinations through search results, the match of meaning may not be the same as the context of the items on the test, which led to learners’ failure to accurately solve the problems. That is, Google search results may expose learners’ to news-related collocation combinations containing their target keywords; however, the customizability of Google is much less useful to help learners pin down on an expression they have in mind.

However, the sharp contrast of the percentage of participants choosing to use corpora in TG (92%) and CG (40%), makes the variable differing the two groups--treatment B--a likely reason for the different preferences of learners’ choice of reference tools. The semester-long in-class practice (treatment B) TG received may have led to learners’ confidence of both the effectiveness of corpora as a reference tool for V-N collocation problems and their skills in using corpora for such tasks. Besides, having received the most exposure to corpora working with V-N collocation problems among the three groups, TG may have been highly aware of the complexity of the target, i.e. V-N collocations due to treatment B. According to Nurmukhamedov (2015), task-induced involvement proposed by Laufer and Hulstijn (2001) serves as the theoretical
underpinning for noticing and awareness for the use of corpora for language learning and teaching. The in-class practices TG received were the designed tasks that involved a problem-solving process of the stages of \textit{need, search,} and \textit{evaluation}, which prompted learners to assess their needs, search through corpora, and evaluate the results. TG’s awareness of the high error rates they themselves, as the literature suggests, likely frequently experienced in their own writing process that is triggered by the weekly exercises, motivated them to take advantage of the reference tools (most using corpora) to find the correct word choice when allowed. Receiving only a one-time workshop only 3 out of 8 (40\%) in CG chose to use corpora and were able to solve collocation problems, but the rest 5 that belonged in the non-corpora-using group were unsuccessful in their attempts. Not only does the exposure to corpora (treatment A only) affect CG’s preference in choosing to use corpora, but CG’s lower (than TG) awareness due to the one-time only in-class exposure to the target form may have accounted for the lack of improvement for the non-corpora-using group. This finding complements Nurmukhamedov’s (2015) research in answering whether the group given tasks to engage with corpora would be more successful than the group that were not due to the different degrees of noticing of the target form. TG’s stronger preferences for corpora as well as statistically significant score improvement between post-test and post-test with reference tools suggest that task-induced involvement is essential for a higher rate of both preferences and accuracy with more noticing triggered by the tasks. As BG and CG were either not exposed to the training at all or had only a 50-minute workshop, the non-existent and minimal awareness-raising may also have caused learners’ simply selecting the same answers for post-test with reference tools as those on the post-test when they
were not allowed to use reference tools, or in some cases, random choosing of answers perhaps due to impatience or the lack of time.

Regardless of the amounts of exposure (treatments A or treatments A and B), learners in the corpora-using group all became able to solve V-N collocation problems with corpora. However, the analysis combining participants from both TG and CG ignored 1.) the different treatments of the in-class practice (treatment B) TG and CG received and 2.) the frequency of the out-of-class use of corpora. Although learners from both groups were successful in using corpora to solve V-N collocation problems, it does not mean the different treatments have no effects on learners. Namely, whether learners received long-term in-class practices (treatment B) and the frequency of out-of-class corpora use should not be discounted. While the current study did not require the use of corpora on the post-test with reference tools and could not tease apart the effects of long-term in-class practices (treatment B) and the frequency of out-of-class use of corpora on learners’ ability to solve V-N collocation problems, with 3 out of 8 in CG still choosing to use corpora on post-test with reference tools, the variable of long-term in-class practices alone does not suffice as the sole reason that explains why learners chose corpora and were able to improve significantly from the post-test to post-test with reference tools. An examination of self-reported frequency of out-of-class corpora use of those who used corpora and those who did not in CG, revealed that the 3 participants using corpora for post-test with reference tools had not been using corpora, and had been using corpora once per week and one per three weeks respectively. 4 in CG indicated that they had been using corpora on their own, and 2 of them (using corpora one per week and per three weeks) chose to use corpora to solve V-N problems on the post-test with reference tools.
Given the statistically significant score improvement of the corpora-using group, it is possible that the out-of-class corpora use after the 50-minute workshop also plays a role in learner’s preference and performance to use corpora to solve V-N collocation problems. Therefore, long-term in-class practices (treatment B) does not appear to be the only deciding factor for learners to choose and become able to use corpora as reference tools for V-N collocation problems, but rather, both the long-term in-class exposure to corpora (treatment B) and the frequency of out-of-class use of corpora influenced learners’ preference for corpora tools as well as their performance. That is, the training and the semester long in-class practice tasks TG received that were voluntarily taken up by CG after receiving the one-time workshop on their own, contributed to their preferences and ability to solve V-N collocation problems using corpora although TG was the only group demonstrating a statistically significant improvement between the post- and post-test with reference tools.

RQ2. Do learners gain knowledge of V+N collocations when they receive exposure to corpus tools?

RQ2 is rejected as the result shows that all three groups did not gain knowledge of V-N collocations regardless of their different amounts of exposure to corpus tools. This question, however, asks the bold question of whether practicing solving self-selected V-N collocation problems enable learners to be better at solving V-N problems in general. The assumption is that the more they practice self-selected V-N collocation, the more overlaps with the frequent collocations they will encounter and possibly internalize. Research such as Sun and Wang (2003) and Chan and Liou (2005) examined the
acquisition of specific collocations by controlling for inductive and deductive teaching of groups and learning units. Their finding showed that 1.) the inductive learning group were able to learn specific collocations and 2.) the collocations taught inductively were better acquired than the ones taught deductively. As Boulton (2010) pointed out, the examination of the outcome of learning involved in the use of corpora has been varied with different target forms and objectives. Although the current research question has a broader scope in asking about the outcome of inductive learning through DDL, it predicted positive results as found in Sun and Wang (2003) and Chan and Liou (2005). With the assumption made for this research question, the non-significant result of all three groups in showing a general knowledge growth of V-N collocations calls for a more thorough and better-designed methodology to answer this question.

Despite the non-significant score improvement between the pre- and post-test, TG’s result of the correlation between the frequency of out-of-class use of corpora and the pre- and post-test score difference demonstrated a moderate-strong positive correlation, whereas CG’s did not. The explanation for such finding is two-fold: 1.) Although TG as a group have all received in-class practice, the higher frequency they used corpora outside of class, the more they improve between the pre- and post-test. 2.) That CG’s reported time of use of corpora did not correlate with the score improvement between pre- and post-test whereas TG’s did, may have to do with their different treatments: the long-term in-class practice (treatment B). Such discovery is inspiring in that it points to the possibility of qualitative growth of collocation knowledge of TG with the frequency of out-of-class use of corpora on top of the treatments they received. However, a few situations have to be ruled out for such interpretation of the test results.
Firstly, it is important to distinguish between whether students rely on rote memory of the pre-test questions to answer the post-test questions, and whether the same items were encountered during the treatment and thus acquired by the students. As the pre-test and post-test were administered 10 weeks apart, the former scenario does not seem to explain the improvement of scores. As for the latter scenario, to cultivate learner autonomy, the current study had asked students to self-identify sources of V-N difficulties and record problematic combinations they encountered in their writing process. Of the self-identified problematic 92 entries of V-N combinations that TG students work with in class together, 45 verbs and 21 nouns appear also in the V-N collocation test. However, the exact same V-N collocations submitted by students that are also included as test items are only three. This is not surprising, as other studies such as Todd (2001) also found that only 3 of the 50 lexical items indicated as errors in his study were common to more than one student.

Given that the analysis of correlation between reported time of the out-of-class use of corpora and score improvement between pre- and post-test yields contrasting results in TG versus CG, it is worth discussing the potential reasons why TG, with long-term in-class practices, demonstrated a more favorable result for collocation learning than CG considering their frequency of out-of-class use of corpora. The potential of collocation knowledge growth may be explained by the overlaps of frequent academic combinations. ACL itself contained several collocations whose verbs are interchangeable with another entry with the same base word (the noun). In fact, because of the overlap of the frequent academic collocations in meaning, ie. synonymous collocations, the test used in the current study was able to include 9 (out of 21) synonym verbs listed in ACL as target answers to the base nouns. As the synonymous collocations, selected in ACL based
on frequency and expert judgment suggest, these combinations appear frequently in academic writing. Therefore, it can be assumed that learners in TG encountered at least some, if not all, of the collocations, in their own writings. Also, encountering some collocations which learners may have worked with using corpora either in class or out of class may have more significance than it appears. Being aware of collocation difficulties and having possibly been using wrong or less frequent synonyms for appropriate collocates, learners in TG may have, through DDL to find out appropriate collocations with synonym errors, not only learned the “appropriate” collocates to go with the base words but also made distinctions of how the synonyms differ. That is, with DDL specifically designed to solve V-N collocation problems, learners’ attention may have been drawn to which verbs, being synonymous verbs, collocate with which nouns more frequently and thus more appropriate. Building on existing knowledge of the words they were familiar with, learners may have made jumps of acquisition of collocations of synonymous words.

Judging from CG’s lack of score improvement between pre- and post-test and the lack of correlation between the frequency of out-of-class use of corpora and pre- and post-test score improvement, however, it appears that for qualitative knowledge growth to take place, simply teaching students, in the case of the current study, one class period to use the tool (treatment A) and allowing them to use it to their liking, was not enough in promoting collocation knowledge growth. This could be because of not only the technical problems as an obvious difficulty that requires intentional trainings to overcome, but also the awareness strengthened by being required to turn in 2 self-identified collocation problems, and the data-driven learning composed of the induction process as well as the
final group presentations TG underwent in class. RQ3 will discuss these aspects in the perspectives of the students in TG.

RQ3. Does exposure to corpora influence learners’ perception about using corpora to solve collocation problems?

The result of Q26 included only for TG in the post-test questionnaire showed that more than half (64%) of TG gave a rating of the effectiveness of the following aspects 5 out of 5: submitting a problematic combination, step-by-step instructions, instructor help and group presentations of best collocation choices. Peer help appears to be the less effective to TG in helping them with corpora use with the lower percentage of learners assigning a score of 5 in comparison with all the other aspect. It came as a surprise that learners viewed submitting a combination of collocations per week as highly favorable, but it confirms the fact that learners’ awareness was raised through self-identifying V-N collocation problems and learners found the noticing useful. The positive perception (64% rating 5 on a scale of 5) also confirms that DDL’s induction process is considered useful and beneficial among learners. Through the step-by-step instructions, learners searched different corpora Word and Phrase Info and COCA, marked the responses everybody contributed acceptable or not acceptable, and presented with their partners the best choices of the V-N collocations they found to the class. Although the result of RQ2 did not show a statistical significance of TG’s score improvement between pre- and post-test and thus indicated no knowledge growth, with all above aspects given a 64% rating of 5, it could be said that learners felt that they benefited from the DDL process to work with collocations. An adaptation of treatments A and B should be considered for future
research to examine the effects of DDL on overall collocation knowledge growth. As for the relatively lower rating TG assigned to peer help, it may have been considered less effective in the process of the exercises because corpora interface indeed does not appear user-friendly (Boulton, 2009). Learners may find it easier to ask their questions directly to the instructor, who is more knowledgeable about the search functions and can help expand or narrow the searches to the users’ needs.

The result of the other parts of section d in the post-test questionnaire asking both TG and CG if they had been using corpora since the 50-minute workshop shows that most learners in TG continued using corpora (also due to the incorporation of treatment B as part of the curriculum), whereas only half of the learners in CG did. Besides, when asked to rate the likelihood to continue/start to us(ing) corpora in the future, half of TG gave the highest rating of 5, whereas half of CG gave the rating of 3. The long-term in-class practice (treatment B) differing TG and CG seems to have affected TG’s preference to use and continue using corpora, perhaps due to their higher familiarity and thus confidence and more positive perception of the effectiveness of corpora. Compared to TG, CG having received only a one-time workshop (treatment A), did not seem to have a high percentage of corpora use, nor a strong willingness to continue using or start to use corpora.

In another light, when asked to rate the corpora’s effectiveness to solve lexical collocations versus other problems, regardless of the different treatments, both TG and CG considered corpora as useful in solving collocation problems. 93% in TG and 75% in CG gave a rating of 4 and 5. In contrast, for the effectiveness of corpora to solve other problems, 57% in TG and 37.5% in CG gave a rating of 4 and 5. More specifically,
learners in both TG and CG pointed out that corpora are useful in finding ‘V-N combinations’ and “the most frequent word for your sentence,” as well as “providing similar and alternative words” to allow for word variations. The results thus indicate that giving learners training of a one-time workshop of corpora for (lexical) collocations is enough in helping learners see and identify with the effectiveness of corpora in solving such problems. Although one learner in CG pointed out that corpora helped him/her to find out the “most common prepositions with vocabulary,” only one instance in all the participants surveyed in TG and CG mentioned corpora’s strength other than its capacity to solve V-N collocation and lexical collocation problems that is highlighted in the treatments both groups received. It reveals that the role of training of use of corpora for specific aspects of writing problems, whether short (as treatment A) or long (treatment B) is important as a starting point for learners, from which they were taught to explore the possibility and potential of reference tools that could potentially help them with autonomous learning.

However, despite the fact that TG demonstrated a high tendency to continue using corpora and both TG and CG’s positive perception of corpora as reference tools to solve lexical collocation problems, when asked what tools other than dictionaries learners most frequently use, the top preference, or at least what first came into learners’ mind, did not switch to corpora from the widely popular Google search after the treatments. Whether it was before and after the treatments, all three groups had more than or about three quarters of learners choosing Google as the preferred tool other than dictionaries as reference tools, and the distribution of the most commonly consulted reference tools—Google, Corpora, other tools, non-use of tools—in the three groups barely changed throughout
the semester. The reasons that learners chose Google search as the most frequently referenced tool are indicated to be “convenience,” “translations,” “synonyms,” both before and after treatments, but after treatments, features such as “words in context” started to appear in CG’s responses, and “combinations,” “word matches,” and “frequency” started to appear in TG’s responses. These features of Google that coincidentally are also considered to be strengths of corpora, were mentioned by TG and CG to be present in Google search. While learners may perceive Google a handy tool that could potentially solve a wide range of problems including lexical/ V-N collocation problems, the result of RQ1’s non-corpora-using group formed by BG and CG with BG: 3/7 and CG: 5/8 choosing Google, showed that they lacked improvements when allowed reference tools. For writing teachers, the importance to encourage the use of corpora, to transform the positive perception of corpora effectiveness as reference tools into actions, should thus be highlighted.

A look into the reasons underlying the above phenomenon can thus provide writing teachers some input into how to tackle the problem of learners’ lack of action of corpora use, given that the use of corpora can potentially lead to more accuracy in V-N collocations and knowledge growth. When asked about the least helpful aspect of corpora, learners in TG and CG complained about the complicated “login process,” “complicated interface,” problems of restricting the search to words of a specific parts of speech and “too much information” provided, other than corpora being “time-consuming” and “not user-friendly.” These weaknesses, have unfortunately, influenced learners’ preference for Google search over corpora besides dictionaries after the treatments provided in the current study, despite the positive perception of the effectiveness of corpora as useful
collocation reference tools. Since the presentation of corpora using concordance lines is almost universally adopted by different kinds of corpora, which at the same time could look intimidating to learners but also beneficial and helpful for rule-finding, writing teachers can design activities to get learners used to the output of corpora. The requirement to log in to the corpora may be the case for the corpora selected for the current study, i.e. COCA and Word and Phrase Info. Therefore, when choosing corpora for a writing class, writing teachers should take the technical aspect of corpora into account and if possible, choose corpora that are more user-friendly to increase learners’ willingness for corpora use.
VI. Limitations

Despite the positive findings of the study, it is not without flaws. Future researchers wishing to examine the effects of corpora and DDL could consider the limitations to further expand on the findings of the current research. First, the tests (pre-test, post-test, and post-test with reference tools) containing the same 21 items, was not validated and recalibrated among learners of different language backgrounds, except for the pilot-test. Also, the design of choosing 2 out of 5 collocations was due to the difficulties of test design with at least four synonym verbs and the complex nature of collocations some of which appear to be much stronger while others less strong. The requirement for learners to choose 2 answers suggests to the learners that two of the combinations are “correct” and perhaps “equally correct,” while one may be much stronger or frequently used than the other. However, due to the limited time and resources the current study was allowed, the above two technicalities of test-design were adopted. Also, as briefly mentioned, examining learners’ ability to use corpora for V-N problems based on their indication of the reference tool instead of the differences of the treatments they received, was unable to help the study pin down on the exact effects of long-term practices and one-time workshop’s respective effects on learners’ ability to solve V-N collocation problems. While it is true that most of the corpora-using group consisted of TG, future research design could restrict the use of reference tools to corpora in order to examine different the effects of treatments on using corpora as a reference tool. Finally, the bold assumption on the general qualitative growth in general calls for a more restricted range of collocations targeted, especially given that the study failed to detect any significant improvement between the pre- and post-test given at the beginning and the end of the semester. Future researchers aiming to examine qualitative collocation
knowledge growth could perhaps, other than letting learners self-select problems they encounter in their own writings, give a list of base words in the form of pairs of synonyms and require some of them (for example, choose 3 out of 10) pairs) depending on learners’ needs in writing tasks to trigger the task-induced process. For example, teachers could give goal/aim or study/research as pairs and require learners to incorporate them into the writing tasks. In doing so, there is a more limited pool of collocation combinations that help narrow down on the collocations learners encounter. Research that takes on such a design can compose a collocation test consisting of the target collocations on top of general academic collocations which the current study designed based on ACL. In incorporating the two sections of items, practiced and general collocations, the degree to which learners transfer knowledge to other base words or collocations can be better pinpointed.
VI. Conclusion and Implications

Recognizing the need of future research, the current study proved similar positive results of using corpora to help with the learning of collocations as the studies of Sun and Wang, (2003) and Chan and Liou (2005). Corpora as reference tools have been confirmed to be useful in helping learners solve V-N collocation problems, as not only demonstrated by the difference of means test between the corpora and non-corpora using group, but also learners’ positive perception of corpora as reference tools for V-N collocation problems, especially by TG. Based on the comparison of accuracy of answers between the post-test and the post-test with reference tools, corpora are speculated to be a more effective tool than Google search given training. ESL/EFL teachers, thus, should consider incorporating corpora into their curriculum as reference tools for its effectiveness and potential to promote learning. At the same time, several other factors could also account for learners’ preference to choose corpora as reference tools and performance, including learner motivation and awareness of target forms, the availability and knowledge of effective tools, skills to use reference tools, and also convenience and difficulties to use reference tools.

The finding of RQ2 sheds light on the insufficiency of frequency of use of corpora to promote collocation knowledge growth. Taking into account the differences of treatments, findings of RQ2 pointed to the important role of the long-term in-class practices that involve a guided induction process potentially leading to collocation knowledge growth. The reasons may be that the frequent academic collocations are highly repetitive, and the learning of synonymous words is easier at the later stage of lexical development of learners, when they have large vocabulary. Learners in the current study undoubtedly falls under the category of advanced learners, given that academic
English is known to be of lower frequency compared to daily English. Since the more TG used corpora outside of class, the more they improved between the pre- and post-test, ESL/EFL teachers should not only introduce and incorporate corpora but also encourage out-of-class use in students’ own time.

The overall findings are encouraging for ESL educators who wish to teach learners to use corpora tools for lexical problems given their obvious strength for the purpose. Incorporating corpora to the regular curriculum also makes learners more autonomous through increased awareness, and inductive learning with DDL also has the potential to lead to qualitative knowledge growth in the long run. However, the positive perception of the effectiveness of corpora in helping with lexical collocation problems, specifically in helping with distinguishing synonyms, as found in the study, did not translate into the change of habit of using corpora as the most preferred reference tool and stuck with Google for its convenience. As pointed out in the discussion, while Google was able to help with some lexical errors as the result of Geiller (2014) suggested, the finding of the current study reflects the insufficiency of Google to help learners with V-N collocation problems. Learners seem to perceive the process to be time-consuming and likely to encounter technical problems such as being required to log in. ESL/EFL teachers should thus try to minimize technical difficulties as such by choosing different corpora or simplified interface that is more user-friendly. In sum, the advantages of corpora, when given proper training, can be maximized to help with solving collocation problems and potentially promote knowledge growth.
References


Nesselhauf, N. (n.d.). Collocations in a Learner Corpus -. Retrieved from https://books.google.hu/books?hl=en&lr=&id=RasDwVvBp00C&oi=fnd&pg=PR9&ots=joEWBgDFnF&sig=mD4-pKFEHb3fAsmHPR7qFV1Luy4&redir_esc=y#v=onepage&q&f=false


Appendix

A. pre-test questionnaire, pre-test

Components
I. Short Questionnaire (16 questions)
II. Main Questions (21 questions)

***This task is NOT graded on the basis of accuracy. It is only graded for completion.***

I. Short Questionnaire
a. Personal Information
1. Your Subject Number *
   Please find the email containing your subject number. Failure to use the correct subject number so will lead to loss of compensation.

2. What is/are your native language(s)? *

3. Do you speak other languages? What level are you in these languages (advanced, intermediate, beginner)? *

4. When did you begin your education in the U.S.? *

5. List all your visits/stays and purposes of visits/stays in the U.S. to the best of your memory. *

6. In general, how confident are you in using English vocabulary in writing without using any tools? *

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremely unconfident</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Extremely confident</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. Vocabulary help during writing-Dictionary
7. When writing, how often do you use the dictionary when you need information about English vocabulary? *

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Always</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. In general, how confident are you in using English vocabulary in writing after consulting a dictionary? *

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-Not Applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-Extremely Unconfident</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-Extremely Confident</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

9. What in dictionaries specifically help you with word choice? *
   Put N/A if you never use dictionaries when writing.

10. What functions or features do you wish are included in dictionaries? *
    If you never use dictionaries when writing, think about features you’d like dictionaries to have.
c. Vocabulary help during writing-Other tools

11. When writing, how often do you use other tools when you need information about English vocabulary?*

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Never</th>
<th>Always</th>
</tr>
</thead>
</table>

12. Check tools you use to find words to use when writing. Check all that you use regardless of frequency.*
Put other tools or N/A in "Other" if no options fit your situation.

- Thesaurus
- Google
- Corpus

Other: 

13. If you checked "Corpus" in the question above, please indicate the name(s) of the corpus you use in the box below.*
Put N/A if you didn't check "Corpus" in #12 

14. Which of the tools above do you use most frequently? *
Put other tools or N/A in "Other" if no options fit your situation.

- Thesaurus
- Google
- Corpus

Other: 

15. What features of the tools were the main reasons you used the tools (instead of dictionaries) for proper vocabulary?*
Put N/A if you never use other tools.

16. In general, how confident are you in using English vocabulary in writing using tools other than a dictionary?*

- 0-Not Applicable
- 1-Extremely Unconfident
- 2
- 3-Neutral
- 4
- 5-Extremely Confident

II. Main Questions (21 questions)
Check boxes below to indicate you've read and understood guidelines for this section.

* Please DO NOT use any tool to find answers to the questions. This helps with more accurate interpretation of the research.
* There are a total of 21 questions for you to answer. Choose TWO answers that best match the context of the sentences.
The sentences are mainly academic sentences. However, you do not need to have advanced knowledge of the academic fields to answer the questions.

Some references in the questions are not clear, but they do not hinder a general understanding of the sentences. Words such as "it," "this," "table 1,"... may refer to concepts/objects outside of the sentences.

Your careful responses are highly appreciated. Please spend no more than 30 minutes for this section.

***This task is NOT graded on the basis of accuracy. It is only graded for completion.*** For accurate interpretation of the research, please refrain from using any tools for the test.

II. Main Questions (21 questions)
Choose TWO answers that best match the context of the sentences. Please do not refer to any reference tools.

Please try finishing the test uninterrupted within 30 minutes. Record the current time on your computer system. *

Example: 11:00 AM

1. Examiners typically _______ data from multiple sources, including records, attorneys, caretakers, and the youth. *
   - obtain
   - attain
   - gain
   - receive
   - collect

2. But if you really believe that you need to _______ the goal, you will find your way through the process to achieve it. *
   - obtain
   - achieve
   - reach
   - complete
   - gather

3. Only a few stories _______ insight into the way Islam fits into converts' life histories. *
   - present
   - deliver
   - provide
   - supply
   - give

4. I have talked to them about how we could _______ the likelihood that they will get an overpayment. *
   - reduce
   - lessen
   - drop
   - decrease
   - lower

5. Receiving approval to _______ research has become more complicated because of the increasing complexity of nursing studies, difficulty accessing subjects, and rising concern over legal and ethical issues. *
   - run
   - assume
   - undertake
6. It is necessary to support this process, to _______ the development of relations of various public groups and youth. *

- increase
- promote
- boost
- encourage
- elevate

7. Deaf children _______ difficulties in their perception and production of speech, and their own speech may be hard for others to understand. *

- undergo
- experience
- endure
- bear
- face

8. The small child in the beginning cannot _______ the distinction between what is dream and what is real. *

- build
- compose
- yield
- make
- draw

9. Symptoms are indicators of problems that reveal something is not as it should be and help you to _______ the problem. *

- identify
- comprehend
- grasp
- distinguish
- recognize

10. They do not _______ the issue of data emerging from other sources while the trial is being conducted. *

- address
- create
- pose
- begin
- tackle

11. First, all three models _______ a tendency to underpredict larger flows. *

- illustrate
- demonstrate
- present
- describe
- show
12. In order to _______ the argument I refer to some recent research of mine on trust in the corporate sector in the country, in the context of economic globalization. *
- develop
- expand
- progress
- grow
- advance

13. Teachers may _______ the requirement for the unscheduled staff development day by attending professional development sessions for a total of six hours *
- reach
- achieve
- meet
- cover
- fulfill

14. You can _______ parameters in the development environment and then test your application by running it from Visual Studio. *
- establish
- distinguish
- characterize
- set
- impose

15. You can _______ this method to process JavaScript code when the page is loaded in the browser. *
- consume
- take
- apply
- exercise
- use

16. Table 1 could _______ information on the population density and demographics of all the cities where you are thinking of locating your new enterprise. *
- contain
- surround
- include
- involve
- enclose

17. For this reason, I decided to _______ the role of language and gender behaviour in the family. *
- consider
- weigh
- inspect
- regard
- examine

18. Introduced wildlife species _______ the potential for several, if not numerous, adverse ecological consequences (Table 6). *
- keep
19. Web search engines ______ access to free Internet resources, so some question the need to catalog them. *
- own
- have
- hold
- enjoy

20. However, there are a number of secondary factors that also ______ consideration. *
- allocate
- allow
- provide
- lend
- distribute

21. Again, there is no suggestion that the child's right to ______ contact with his or her parents also needs protecting. *
- undertake
- hold
- maintain
- preserve
- make

Please record the current time on your computer system. *
Example: 11:00 AM
Submit
B. Post-test questionnaire, post-test , post-test with reference tools

Components
I. Short Questionnaire (28 questions)
II. Main Questions without Reference Tools (21 questions)
III. Main Questions with Reference Tools (21 questions)

***This assignment is NOT graded on the basis of accuracy. It is only graded for completion.***

I. Short Questionnaire

a. Personal Information

1. Your Subject Number *
Please find the email containing your subject number.

2. Gender *
- Male
- Female

3. Including time spent learning English at your home country, how long have you been learning English? *
Put # years, # months

4. How long have you stayed in an English-speaking country (America, U.K., New Zealand, Canada, Australia)? Count all the time periods you've stayed in such countries, including studying in UIUC. *
Put # years, # months

5. How much time in question #4 is spent on studying for a degree or for English learning? *
Put # years, # months

6. At this point of the semester of your study at UIUC, how confident are you in using English vocabulary in writing without using any tools? *

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Extremely unconfident</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

b. Vocabulary help during writing-Dictionary

The rest of the questionnaire should be answered based on your perception as of this point of the semester.

7. When writing, how often do you use the dictionary when you need information about English vocabulary? *

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

8. In general, how confident are you in using English vocabulary in writing after consulting a dictionary? *

- 0-Not Applicable
- 1-Extremely Unconfident
- 2
- 3-Neutral
- 4
- 5-Extremely Confident
9. What in dictionaries specifically help you with word choice? *
Put N/A if you never use dictionaries when writing.

10. What functions or features do you wish are included in dictionaries? *
If you never use dictionaries when writing, think about features you'd like dictionaries to have.

11. When writing, how often do you use other tools when you need information about English vocabulary?*

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

12. Check tools you use to find words to use when writing. Check all that you use regardless of frequency. *
Put other tools or N/A in "Other" if no options fit your situation.

- Thesaurus
- Google
- Corpus
- Other: _____________________

13. If you checked "Corpus" in the question above, please indicate the name(s) of the corpus you use in the box below. *
Put N/A if you didn't check "Corpus" in #12

14. Which of the tools above do you use most frequently? *
Put other tools or N/A in "Other" if no options fit your situation.

- Thesaurus
- Google
- Corpus
- Other: _____________________

15. What features of the tools were the main reasons you used the tools (instead of dictionaries) for proper vocabulary? *
Put N/A if you never use other tools.

16. In general, how confident are you in using English vocabulary in writing using tools other than a dictionary? *

- 0-Not Applicable
- 1-Extremely Unconfident
- 2
- 3-Neutral
- 4
- 5-Extremely Confident

d. experiences with using corpus for word choice [ONLY FOR CG and TG; Q25-26 ONLY TG]
17. In the beginning of the semester, you were given a workshop on using corpus to solve collocation problems. Have you been using corpus? *

Corpus are databases such as COCA and Word and Phrase Info

☐ Yes, I already were using it before the workshop
☐ Yes, I started using it after the workshop.
☐ No, I have not been using it since the workshop.
☐ Other: 

18. If you answered yes in question#17, how often have you been using corpus outside of class to solve ANY problems? *

☐ Not applicable, I have not been using it.
☐ Roughly once a week
☐ Roughly once every two weeks
☐ Roughly once every three weeks
☐ Roughly once every four weeks
☐ Other: 

19. If you answered yes in question#17, what have you been using corpus MAINLY for? *

☐ Not applicable, I have not been using it.
☐ to check grammar
☐ to check meaning of words in context
☐ to look for appropriate word that goes together with another word
☐ Other: 

20. If you answered no in question#17, why have you not been using corpus? *

☐ Not applicable, I have not been using it.
☐ don't know how to use it
☐ don't think it meets my needs
☐ don't have time to use it.
☐ don't like the appearance of corpus
☐ don't think it is useful
☐ Other: 

21. How helpful do you think corpus is in helping you solve LEXICAL COLLOCATION PROBLEMS? *

example of lexical collocation problems are: save money, heavy rain, criticize sharply, highly correlated.

   1  2  3  4  5

   Not helpful  |  |  |  |  | Extremely helpful

22. How helpful do you think corpus is in helping you solve OTHER PROBLEMS that are not collocation problems? *

example of other problems are: grammatical collocations such as what prepositions to use after verb, whether a word is being too formal or informal.

   1  2  3  4  5

   Not helpful  |  |  |  |  | Extremely helpful

23. Describe what aspect of corpus you find/think is MOST helpful. Give an example of such search experiences on corpus in detail. If you have not been using corpus, think of an instance you experienced in the workshop. *
24. Describe what aspect of corpus you find/think is LEAST helpful. Give an example of such your search experiences on corpus in detail. If you have not been using corpus, think of an instance you experienced in the workshop. *

25. Describe your experience using Corpus in ESL501D. Did you like the exercise in class? Why or why not? *

26. Rate effectiveness of in-class word exercises using corpus in the following aspects. *

<table>
<thead>
<tr>
<th></th>
<th>Submit Combination Difficulties</th>
<th>Step-by-Step Instructions</th>
<th>Instructor Help</th>
<th>Peer Help</th>
<th>Group Presentation of Best Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-Extremely Helpful</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>3-Neutral</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>1-Not helpful</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

27. What do you think is the most challenging aspect of using corpus? *

28. Do you think you will continue/start using corpus in the future? *

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

II. Main Questions (21 questions)
Check boxes below to indicate you've read and understood guidelines for this section.
*  
- Please DO NOT use any tool to find answers to the questions. This helps with more accurate interpretation of the research.
- There are a total of 21 questions for you to answer. Choose TWO answers that best match the context of the sentences.
- The sentences are mainly academic sentences. However, you do not need to have advanced knowledge of the academic fields to answer the questions.
- Some references in the questions are not clear, but they do not hinder a general understanding of the sentences. Words such as "it," "this," "table 1,"... may refer to concepts/objects outside of the sentences.
- Your careful responses are highly appreciated. Please spend no more than 30 minutes for this section.

For accurate interpretation of the research, please DO NOT USE any tools for the test.

***This assignment is NOT graded on the basis of accuracy. It is only graded for completion.***

II. Main Questions WITHOUT Reference Tools (21 questions)
Choose TWO answers that best match the context of the sentences. Please do not refer to any reference tools.

Please try finishing the test uninterrupted within 30 minutes. Record the current time on your computer system. *

Example: 11:00 AM

1. Only a few stories _______ insight into the way Islam fits into converts' life histories. *

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>give</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td></td>
</tr>
<tr>
<td>supply</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td></td>
</tr>
<tr>
<td>present</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td></td>
</tr>
</tbody>
</table>
2. But if you really believe that you need to _______ the goal, you will find your way through the process to achieve it. *
   - obtain
   - achieve
   - reach
   - gather
   - complete

3. Examiners typically _______ data from multiple sources, including records, attorneys, caretakers, and the youth. *
   - receive
   - attain
   - obtain
   - gain
   - collect

4. It is necessary to support this process, to _______ the development of relations of various public groups and youth. *
   - increase
   - elevate
   - boost
   - promote
   - encourage

5. Receiving approval to _______ research has become more complicated because of the increasing complexity of nursing studies, difficulty accessing subjects, and rising concern over legal and ethical issues. *
   - run
   - undertake
   - perform
   - conduct
   - assume

6. I have talked to them about how we could _______ the likelihood that they will get an overpayment. *
   - drop
   - lower
   - lessen
   - decrease
   - reduce

7. Symptoms are indicators of problems that reveal something is not as it should be and help you to _______ the problem. *
   - grasp
   - distinguish
   - identify
   - recognize
8. The small child in the beginning cannot _______ the distinction between what is dream and what is real.*
- comprehend
- compose
- yield
- make
- draw
- build

9. Deaf children _______ difficulties in their perception and production of speech, and their own speech may be hard for others to understand. *
- undergo
- bear
- face
- experience
- endure

10. In order to _______ the argument I refer to some recent research of mine on trust in the corporate sector in the country, in the context of economic globalization. *
- expand
- advance
- grow
- progress
- develop

11. First, all three models _______ a tendency to underpredict larger flows. *
- show
- describe
- illustrate
- demonstrate
- present

12. They do not _______ the issue of data emerging from other sources while the trial is being conducted. *
- begin
- address
- tackle
- pose
- create

13. You can _______ this method to process JavaScript code when the page is loaded in the browser. *
- use
- exercise
- take
- apply
- consume

14. You can _______ parameters in the development environment and then test your application by running it from Visual Studio. *
- set
15. Teachers may ______ the requirement for the unscheduled staff development day by attending professional development sessions for a total of six hours.*
- impose
- establish
- distinguish
- characterize

16. Introduced wildlife species ______ the potential for several, if not numerous, adverse ecological consequences (Table 6). *
- own
- have
- hold
- enjoy
- keep

17. For this reason, I decided to ______ the role of language and gender behaviour in the family. *
- inspect
- weigh
- examine
- consider
- regard

18. Table 1 could ______ information on the population density and demographics of all the cities where you are thinking of locating your new enterprise. *
- involve
- enclose
- include
- surround
- contain

19. Again, there is no suggestion that the child’s right to ______ contact with his or her parents also needs protecting. *
- maintain
- preserve
- hold
- undertake
- make

20. However, there are a number of secondary factors that also ______ consideration. *
- earn
- command
- request
- deserve
21. Web search engines _______ access to free Internet resources, so some question the need to catalog them. *
  □ distribute
  □ provide
  □ lend
  □ allow
  □ allocate

Please record the current time on your computer system. *
Example: 11:00 AM

III. Main Questions with Reference Tools (21 questions)

Check boxes below to indicate you've read and understood guidelines for this section.
*

□ Please DO USE any tool to find answers to the questions. This helps with more accurate interpretation of the research.
□ There are a total of 21 questions for you to answer. Choose TWO answers that best match the context of the sentences.
□ The sentences are mainly academic sentences. However, you do not need to have advanced knowledge of the academic fields to answer the questions.
□ Some references in the questions are not clear, but they do not hinder a general understanding of the sentences. Words such as "it," "this," "table 1,"... may refer to concepts/objects outside of the sentences.
□ Your careful responses are highly appreciated. Please spend no more than 30 minutes for this section.
For accurate interpretation of the research, please DO USE tools your normally would use to find answers to this section.

***This assignment is NOT graded on the basis of accuracy. It is only graded for completion.***

III. Main Questions WITH Reference Tools (21 questions)
Choose TWO answers that best match the context of the sentences. Please DO refer to reference tools to find the best answers.
Please try finishing the test un uninterrupted within 30 minutes. Record the current time on your computer system. *

Example: 11:00 AM

1. Only a few stories _______ insight into the way Islam fits into converts' life histories. *
  □ supply
  □ deliver
  □ provide
  □ present
  □ give

2. But if you really believe that you need to _______ the goal, you will find your way through the process to achieve it. *
  □ complete
  □ obtain
  □ achieve
  □ reach
  □ gather

3. Examiners typically _______ data from multiple sources, including records, attorneys, caretakers, and the youth. *
  □ collect
obtain
attain
receive
gain
4. It is necessary to support this process, to _______ the development of relations of various public groups and youth. *

promote
elevate
boost
courage
increase
5. Receiving approval to _______ research has become more complicated because of the increasing complexity of nursing studies, difficulty accessing subjects, and rising concern over legal and ethical issues. *

assume
undertake
perform
run
conduct
6. I have talked to them about how we could _______ the likelihood that they will get an overpayment. *

lessen
decrease
reduce
drop
lower
7. Symptoms are indicators of problems that reveal something is not as it should be and help you to _______ the problem. *

distinguish
identify
grasp
recognize
comprehend
8. The small child in the beginning cannot _______ the distinction between what is dream and what is real.*

yield
make
compose
build
draw
9. Deaf children _______ difficulties in their perception and production of speech, and their own speech may be hard for others to understand. *

experience
undergo
10. In order to _______ the argument I refer to some recent research of mine on trust in the corporate sector in the country, in the context of economic globalization. *
- endure
- face
- bear

11. First, all three models _______ a tendency to underpredict larger flows. *
- demonstrate
- present
- illustrate
- show
- describe

12. They do not _______ the issue of data emerging from other sources while the trial is being conducted. *
- address
- begin
- pose
- tackle
- create

13. You can _______ this method to process JavaScript code when the page is loaded in the browser. *
- consume
- apply
- take
- exercise
- use

14. You can _______ parameters in the development environment and then test your application by running it from Visual Studio. *
- distinguish
- characterize
- establish
- impose
- set

15. Teachers may _______ the requirement for the unscheduled staff development day by attending professional development sessions for a total of six hours *
- fulfill
- achieve
- cover
- reach
- meet
16. Introduced wildlife species _______ the potential for several, if not numerous, adverse ecological consequences (Table 6). *
   - own
   - hold
   - have
   - enjoy
   - keep

17. For this reason, I decided to _______ the role of language and gender behaviour in the family. *
   - examine
   - consider
   - weigh
   - inspect
   - regard

18. Table 1 could _______ information on the population density and demographics of all the cities where you are thinking of locating your new enterprise. *
   - contain
   - involve
   - enclose
   - surround
   - include

19. Again, there is no suggestion that the child's right to _______ contact with his or her parents also needs protecting. *
   - make
   - hold
   - maintain
   - undertake
   - preserve

20. However, there are a number of secondary factors that also _______ consideration. *
   - deserve
   - request
   - earn
   - require
   - command

21. Web search engines _______ access to free Internet resources, so some question the need to catalog them. *
   - distribute
   - allow
   - lend
   - allocate
   - provide

Please record the current time on your computer system. *
Example: 11:00 AM
Please check the box of the tool you used the most frequently when answering part III of the test.
   - Google
☐ Online Dictionaries
☐ Corpora
☐ Other: [Box for input]

Submit