(HOW) DO SECONDARY SCHOOL EFL WRITERS BENEFIT FROM CORPUS-BASED COLLOCATION LEARNING? A STUDY OF USAGE, OUTCOMES, AND PERCEPTIONS

By

Duong Thi Thuy Nguyen

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Teaching English to Speakers of Other Languages – Master of Arts

2023

ABSTRACT

This study investigates the effects of hands-on corpus consultation in a writing class for high school students on the accuracy of collocations used in their academic writing production. 24 Vietnamese high school students who participated in the study engaged in five ninety-minute corpus-training sessions. Corpus of Contemporary American English (COCA) and SKELL were chosen as the main corpus tools. A mixed research method was employed, combining quantitative data with pre-test-and-post-test design and evaluation survey with qualitative data obtained from semi-structured interviews and videotapes recording students' corpus consultation process. The results showed a mixed picture regarding the participants' improvement in the quality of the verb-noun (V-N) collocations used in their essays, measured by the average MI scores and collocation error rates per 100 words, after five corpus-based training sessions. The study also found that there were differences between students who improved and those who did not improve in the way they utilized the corpus tools at different stages of their corpus consultation. The participants in general showed a positive attitude toward direct corpus consultation and the corpus-based lessons. The implications of the findings for academic writing pedagogy and the design of corpus-based lessons in EFL teaching context are discussed.

Keywords: corpus-based instruction, collocations, academic writing, frequency and accuracy

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Charlene Polio, my academic advisor, for her expertise, valuable feedback, and unwavering support throughout my thesis journey. Her guidance enabled me to develop a well-designed mixed-method study, enhance my understanding of data analysis methods, and present my findings effectively. Without her guidance, I would not have made it this far.

I am also deeply thankful to Dr. Sandra Deshors, my second reader, for introducing me to Corpus Linguistics and providing invaluable insights into literature reviews. Her constant support and encouragement have been instrumental in motivating me on my academic path.

I would like to acknowledge Dr. Paula Winke for recommending the use of the Likert scale in coding collocational acceptability during her assessment class. Her suggestions and support have significantly contributed to the progress of my thesis.

Furthermore, I extend my heartfelt appreciation to Dr. Mark Davies for his generosity in offering a discounted COCA subscription for my students, which facilitated access to the corpus tool for my study. I am also grateful to the Fulbright program for sponsoring my study at Michigan State University, enabling my personal and professional development.

I would like to extend my thanks to my friends in the SLS Ph.D. and MA TESOL program for their unwavering support and assistance whenever I needed it. Their advice, feedback, and encouragement have provided me with the strength to persevere and achieve my goals. I am also grateful for the enthusiastic participation of my colleagues and students in Vietnam, Lastly, I would like to express my deepest gratitude to my family and friends for always believing in me.

TABLE OF CONTENTS

1 INTRODUCTION

Learners who want to be more fluent and natural in using a second language need to develop their collocational knowledge. However, research showed EFL learners are limited in collocational knowledge (Nesselhauf, 2004; Durrant & Schmitt, 2009; Laufer and Waldman, 2011) and use insufficient and unnatural collocations in their written productions compared to native speakers (Nesselhauf, 2004). This may be attributed to a lack of exposure to the second language as well as inadequate attention to collocations and ineffective collocation instruction in EFL classrooms. Unlike ESL students who have substantial exposure to English on a regular basis, most EFL learners' exposure to collocations is through textbooks and input from teachers, which as shown in Tsai (2015), accounts for a minority of the collocation repertoire.

Although incidental learning of collocations through extensive reading, listening, and viewing has potential learning benefits (Green, 2020), it takes learners a huge amount of time for such approaches to have visible effects. In addition, it depends largely on the students' motivation to learn the language, their interest in the reading, listening, or viewing materials as well as their ability of self-study.

Corpus-based instruction is a potential approach. Many studies focused on the receptive and productive use of collocations after students are introduced and interact with corpus-driven materials. However, little has been done to measure the frequency distribution and accuracy of collocations used in academic writing after students are instructed to directly use corpus tools to interact with collocations. One reason could be attributed to the difficulty in defining the frequency and accuracy of collocation use in L2 writing.

The current study aims to examine whether hands-on corpus consultation can enable EFL learners to use more collocations in their essays and develop the association strength of

collocations used in their writing productions, how students use corpus tools to assess and revise their own writing, and whether corpus-based instruction is worth the instructional time in EFL classrooms.

2 LITERATURE REVIEW

2.1 Collocations

There has been an inconsistency in the way researchers have defined collocations in the literature. Firth (1957) defined collocations as lexical items that co-occur more frequently than just by chance. More sophisticated definitions were proposed by Nesselhauf (2004) using two different approaches, namely the *frequency-based approach* and the *phraseological approach.* The former refers to the collocations whose individual words occur together more frequently than expected, whereas the latter is limited to collocations constructed by relatively but not completely fixed combinations of words. Furthermore, the words in collocations are not necessarily adjacent but can appear around the node, the main word in the combinations with the window span of 4:4, which means four words before and four words after the node (Goulart, 2019, p.4). This search span has been commonly employed in research on collocations (Yoon, 2016, p.46). Benson et al. (1986) approached collocations from theoretical and pedagogical perspective and categorized collocations into two groups: grammatical collocations consisting of "a dominant word - a noun, an adjective, or a verb and a particle including a preposition or other grammatical constructions" and lexical collocations which do not comprise "a dominant word, including such combinations as verb-noun, adjective-noun, noun-noun, noun-verb, adverb-adjective, adverb-verb" (see Wei, 1999). For the sake of variable control, this study focuses only on one type of collocation, which is verb-noun collocation as many studies suggested this type of collocation poses the most challenge to L2 learners (Nesselhauf, 2003; Chan & Liou, 2005)

Collocations are widely acknowledged to play an integral part in developing L2 learners; proficiency, especially in the productive skills, speaking and writing. A great deal of research claimed collocations are an important factor for distinguishing native and non-native speakers

(Nation, 2001; Koya, 2006). Despite this fundamental role in second language acquisition, collocational knowledge is often neglected in EFL lexical syllabuses, compared to the knowledge of individual words, especially in the high school context. Even ELT materials for high school students covered a small proportion of collocations compared to the amount that should be acquired. Tsai (2015) conducted a study to investigate the profile of collocations in the three most common ELT textbooks used for senior high school students in Taiwan. He created a verb-noun collocation list from the Reference Wordlist prescribed by Taiwanese Ministry of Education and compared it with the list of verb-noun collocations extracted from every text in those textbooks. One of the results revealed that the number of collocations used in those textbooks constituted only a tiny minority in the repertoire of 43,875 collocations from the verb-noun collocation list (Tsai, 2015, p.735).

Another reason for L2 learners' limited knowledge of collocations, notably among advanced learners, is that these types of learners tend to overestimate their knowledge of collocations (Laufer & Waldman, 2011), hindering them from looking up collocations due to the ignorance of their limitations, which makes them susceptible to errors and other problems such as underuse, overuse, and a limited collocation repertoire. This was in tune with other bodies of research stating that EFL learners, even at an advanced level, have a tendency to commit errors with word combinations (Nesselhauf, 2003). Among many types of collocations, verb-noun collocations seem to cause EFL learners the most challenges and they often misuse the verb collocates. This was backed up by previous studies of Nesselhauf (2003). Nesselhauf's investigation into the writing of Germany advanced learners of English revealed that the incorrect selection of verbs was the most common error. She then explained the restricted sense of a collocation made it more difficult for L2 learners if they were unable to differentiate the nuances

of different verbs.

Chan and Liou (2005) proposed three main reasons for L2 learners' failure to acquire verb-noun collocations, including L1 direct translation, inappropriate use of delexicalized verbs such as make/do or add/increase, and insufficient collocational knowledge in terms of semantic differences among such lexemes as synonyms, hypernyms and troponyms (p.232). Therefore, collocations should be explicitly taught in a way that can raise students' awareness about well-formed verb-noun collocations with different verb collocates, especially with the verbs that are syntactically related as mentioned above.

2.2 Teaching collocations in Vietnamese context

Over the last two decades, English education has received a great deal of attention from different stakeholders in Vietnam. With the aim of developing English proficiency for every Vietnamese student, the Ministry of Education has made English a compulsory subject in the national curriculum and the medium of instruction at higher education level. In their review of vocabulary learning, teaching, and testing in Vietnam, Vu and Peters (2021) reported that Vietnamese students' performance on the national high school examination tended to be low, despite huge investment in English development programs. They briefly mentioned the underlying reason for this issue, suggesting the current pedagogical methods in most educational settings below tertiary level are heavily focused on grammar and individual lexical item acquisition, with less emphasis on formulaic language generally and collocations specifically. Concerned teachers and researchers have conducted substantial studies to examine the ways to better English teaching of collocations in Vietnam.

Nguyen and Webb (2016) examined five factors that influence Vietnamese EFL learners' receptive knowledge of collocations, including node word frequency, collocation frequency,

mutual information score (MI score), congruency, and part of speech. The study's results showed that learners' performance on the collocation tests with 2,000- and 3,000-word frequency levels were lower than 50%. Their findings mirrored the results of previous studies, such as Laufer and Waldman (2011), Nesselhauf (2003), and Chan and Liou (2005), about EFL learners' lack of collocation knowledge. In addition, the study confirmed the positive relationship between single-word items and collocations as well as the significant impacts of node word frequency on receptive knowledge of collocations.

Hoang (2018) conducted another study that probed the receptive collocational knowledge of Vietnamese students, which focused on one type of collocation, verb-noun collocations. A semi-structured interview on vocabulary recognition and frequency-based approach to collocations were employed to measure the understanding of verb-noun collocations with academic verbs. The results showed the students had difficulty distinguishing non-academic verb-noun collocations, phrasal verbs, and academic verb-noun collocations, which suggested they were unaware of the concept of collocations and needed more extensive practice exercises. Finally, he called for the development of a framework for teaching verb-noun collocations associated with academic verbs. From understanding how students process collocations and their profile of collocations, teachers and researchers have started investigating different approaches to facilitate learners' collocation use in language production.

In terms of productive knowledge, Cao (2018) examined the effects of using an online collocation dictionary on Vietnamese advanced learners' collocation use in academic writing as well as their perceptions of the use of the dictionary as a supportive tool to learn collocations. Both quantitative and qualitative data were collected, including analyzed written texts, questionnaires, observation, and interviews. The findings demonstrated the benefits of the online collocation

dictionary to students, but with regard to attitudes rather than practical effects as their performance of collocation use in the writing did not show improvement in general. Another finding was that learners made the most mistakes with verb-noun collocations, which is aligned with the results of previous studies.

A recent study on the perceptions of Vietnamese students and lecturers about the use of collocations in enhancing academic writing was implemented by Duong and Nguyen (2021). They used online questionnaires for students and email interviews with the lecturers to obtain both descriptive statistics and qualitative data through thematic analysis. The study found students' demotivation derived from extensive reading and collocational practice preempted them from acquiring collocations. The results also reflected an issue in Vietnamese teaching context, which was the absence of collocations in language instruction and ineffective learning strategies. In short, there is a need for a more interactive and effective way of collocation teaching and learning. This brings up the questions of using corpus tools and corpus consultation as an alternative collocation pedagogical approach in academic writing class.

2.3 Corpus-based instruction in L2 writing

Along with the development of corpus linguistics, the advent of advanced technology in creating corpus tools and the better awareness of L2 instructors about the potential of corpora and corpus tools in analyzing learners' language production as well as promoting learner autonomy, there have been many studies investigating the effects of different corpus-based instruction on L2 learners' performance. Because this study focuses on the use of verb-noun collocations in academic writing, only studies on corpus-based collocation instruction in L2 writing are discussed.

Conway (2011) examined the difference in the frequency and accuracy of vocabulary use between two groups: the comparison group learning with a vocabulary list selected by human

instructors and the treatment group with a list of terms created through corpus analysis. He created a corpus from selected chapters of Harriet Tubman, which was used in a content-based instruction class for fourth level learners, using AntConc to analyze the most frequent words and comparing those with the wordlist composed by a teacher to create a list of vocabulary with the highest frequency and multiple forms. The instruction for the comparison group was varied but excluded the use of collocation analysis or concordancing software, whereas the treatment group was guided to identify vocabulary collocates through text analysis and the use of a concordancer "WordandPhrase.info" as a reference tool to confirm the analysis of the collocations used in the text. The results showed the treatment group outperformed the comparison group in terms of frequency and accuracy of collocation use in their essays. Despite producing positive results, the study focused only on measuring the distribution and accuracy of the collocations in the controlled wordlists created in the beginning.

Another study using corpus-informed teaching materials was conducted by Jafarpour et al. (2013). They compared the effects of the corpus-based approach with a traditional approach to learning collocations of near-synonymous pairs. The same research design was employed with a comparison group and a treatment group. While the former received explicit instruction of collocations, the latter was taught with the materials developed from the concordances in British National Corpus (BNC). The findings confirmed that the benefits of a corpus-based approach to learners' collocational knowledge and writing production outweighed those of a traditional approach. This was instantiated by the lower number of collocation errors in students' post-writing. Nevertheless, how collocation errors were identified was not clearly described in the paper, which seemed to be the limitation of this study.

From these studies, hands-off corpus-based instruction some potential in developing L2 learners' receptive and productive knowledge of collocations. The next question would be whether hands-on corpus consultation would lead to similar promising results. Park (2010) investigated the effects of hands-on corpus-based instruction on L2 learners' lexicogrammatical performance in academic writing. The results showed the efficiency and effectiveness of corpus-based instruction in improving students' lexicogrammatical performance in their writing productions. The study also suggested screen videotapes coupled with collocation query log and written reflection can supplement retrospective interviews as a more reliable data collection method. The screen recordings were used in writing conferences in which participants gave comments on their corpus consultations while they were watching the videos.

It is noteworthy that most studies on corpus-based instruction of collocation use in writing were conducted in tertiary educational settings; few studies focused on lower levels of education. This is understandable given the limited time teachers at secondary levels have in order to cover the rigid syllabus and the rules that prohibit students from using electronic devices such as smartphones or laptops in class. Fang et al. (2021) was one of a few studies investigating the effects of corpus tools and hands-on corpus-based instruction on senior secondary school students' collocation use in IELTS writing. Twenty-two eleventh grade Chinese students participated in the study, but when it comes to data analysis, only the data from seven of them could be used because not all participants completed all tests. This can also be seen as an obstacle when it comes to research on secondary school students. The research instruments included one pre-writing test and one post-writing test, one questionnaire and retrospective interviews. Students did the pre-writing test without any reference tools. Then the instructor used the error data in students' pre-writing texts to design teaching materials for three corpus training sessions. After the treatment, students

did the post-writing test with the reference to corpora such as COCA or Word and Phrase Concordancer and dictionaries if they want. Perhaps because of the small sample size (n=7) in the end, the study did not find any significant difference in the frequency of collocation errors in students' writing between pre-test and post-test. However, the qualitative data showed positive signs regarding students' attitudes toward their experience with corpora and their perception of the efficacy of using corpora as a reference tool in writing.

Acknowledging the scarcity of corpus-based research in pre-tertiary educational settings, Crosthwaite and Steeples (2022) conducted a mixed-method study to explore the effectiveness of hands-on corpus-based instruction in a secondary school. The study involved thirty-one secondary students in Australia, including twenty-four monolingual English speakers and seven students who spoke English as an additional language. The study aimed to investigate the students' receptive and productive knowledge of passive voice structure after a half-year intervention of corpus consultation. The design included a pre-test, a post-test, and a delayed post-test, supported by data from questionnaires and interviews. The results demonstrated that the use of corpora and hands-on corpus-based training was effective in enhancing the learners' acquisition of passive voice, as well as their positive attitudes towards the corpora and corpus consultation.

Inspired by above-mentioned bodies of research, the current study also probed into high school participants with a mixed-method research design including a writing pre-test and a writing post-test combined with questionnaires, screen recordings tracking learners' activities of corpus consultation, and semi-structured interviews. This combination of quantitative and qualitative data as well as retrospective and real-time data can hopefully help us better understand the effectiveness or lack thereof of hands-on corpus-based instruction of collocations to improve students' performance of this lexical feature in their writing productions.

In short, even though research on the effects of different approaches of teaching collocations on the acquisition of this language feature has abounded in literature, most of them focused on the receptive knowledge of collocations and different stakeholders' perceptions of collocations. Empirical studies on hands-on corpus-based instruction are still scarce. In addition, few studies on these issues focused on participants other than university students.

In light of all the above, the present study was conducted to address the following research questions:

1. To what extent does corpus consultation impact the association strength (MI scores) of V-N collocations used in the writing process of advanced EFL learners?
2. How do advanced EFL learners use hands-on corpus consultation in relation to the use of V-N collocations in writing?
3. What are advanced EFL learners' attitudes towards the hands-on use of corpus tools as a resource for learning V-N collocations in writing?

3 METHODOLOGY

3.1 Participants and context

The participants were 16-to-17-year-old English specialized students from a high school for the gifted in a small city in central Vietnam. All of them have studied English for nearly nine years, with two 45-minute English lessons per week. In addition, they have six 45-minute intensive lessons on six aspects of English: Reading, Listening, Writing, Speaking, Grammar, and Vocabulary. The textbook used by all students is *English 11* prescribed by Vietnamese Ministry of Education and other advanced textbooks assigned by the instructor. Thirty participants were recruited from two English specialized classes through flyer distributions; five of them did not participate in all five training sessions, and one of them did not take the writing pretest, so they were excluded from the study, leaving 24 participants, with five males and nineteen females. All participants were given a consent form along with the background survey for their voluntary participation in the study. Each of them received $5 for their full participation at the end of the course. The detailed demographic data is presented in the Table 6 in the Appendix G.

Regarding the participants' English backgrounds, all of them reported to be interested in trying new things. They write in English relatively often with nearly half of the participants writing in English once a week and no one writing less than twice a month. The resources these participants often use when they write in English include online collocation dictionaries (Ozdic), online general dictionaries (Oxford, Cambridge, Longman, Collins, etc.), and relevant websites (online newspapers and academic writing materials). When it comes to collocations, approximately 30% of the total participants strongly agreed that collocations caused them difficulties writing in English, more than 50% partly agreed, and about 8% of them did not think collocations are a challenge. However, all participants showed their interest in learning collocations in English, with

more than 70% of them being extremely enthusiastic about learning English collocations. About 50% of the participants had not heard about corpus or corpora before, and among the remaining who already heard about corpus or corpora, around 70% of them had never used any corpus such as COCA or BNC before, leaving about 30% who already had some experience with COCA.

3.2 Procedures

3.2.1 Writing course procedure

I developed an online course using the Canvas Learning Management System (LMS) and distributed the invitations to all the participants via email. The course was structured around seven modules, with the first two modules being delivered asynchronously due to time constraints. Given the online delivery format, the initial module focused on providing instructions to the participants in key computer literacy skills, including registration for a COCA account, navigation of the Canvas LMS platform, and utilization of various tools such as screen recording software, Google Docs, and Google Drive.

In the second asynchronous module, the participants were provided with a corpus introduction video featuring the following contents: collocation definitions, statistical information related to collocations (MI score and frequency), basic functions in COCA and SKELL for lexical searches, and *collocate* function in COCA as well as *word sketch* in SKELL for collocation searches. The corpus introductory lesson is an interactive video created on PlayPosit so that students could answer the pop-up questions while they were watching the video. As a means of facilitating comprehension of more advanced course materials, a segment of the video demonstrating the utilization of COCA and SKELL was delivered in the participants' first language, Vietnamese.

Subsequent course modules were designed to train the participants to identify good verb-

noun collocations using COCA and SKELL for completing exercises such as gap filling, error correction, and sentence completion, as well as in essay writing process. For further reference, a detailed syllabus and lesson plans are included in Appendix E and F.

The online course was delivered over a period of five weeks, commencing on November 13, 2022, and concluding on December 18, 2022, with a one-week hiatus in observance of the national holiday in Vietnam. Course sessions were held every Sunday, with each class lasting for 90 minutes, from 9:00 to 10:30 PM (GMT), and conducted via Zoom. Prior to each session, I checked attendance as a means of monitoring participant engagement.

Table 1 shows the summary of the asynchronous and synchronous lessons.

Table 1.

Corpus-based lesson summary

Lesson	**Time**	**Focus**	**Activity**
1st asynchronous lesson	Before the online course	Course introduction and Technology help	Watch videos introducing the course and some tools such as Canvas LMS, Google Docs, and the screen recording application
2nd asynchronous lesson	Before the online course	Essay review and Corpus introduction	Watch the interactive videos to review different types of essays and learn about two corpus tools: COCA and SKELL Answer the pop-up questions in the interactive videos

Table 1. (cont'd)

First online lesson	Nov. 13, 2022	Collocation training 1	Use corpus tools as a reference source to complete the following activities: Activity 1: Gap-filling Activity 2: Error identification and correction Activity 3: Sentence building Homework: Creating the virtual corpus (VC) for topic Technology and Environment
Second online lesson	Nov. 27, 2022	Essay writing 1	Activity 1: Brainstorming and Outlining (Group work on Jamboard) Activity 2: Writing the essay (individual) with corpus consultation Activity 3: Peer-feedback Homework: Self-correction
Third online lesson	Dec. 4, 2022	Collocation training 2	Use corpus tools as a reference source to complete the following activities: Activity 1: Gap-filling Activity 2: Error identification and correction Activity 3: Sentence building Homework: Creating the virtual corpus (VC) for topics Crime and Health

Table 1. (cont'd)

Fourth online lesson	Dec. 11, 2022	Essay writing 2	Activity 1: Brainstorming and Outlining (Group work on Jamboard) Activity 2: Writing the essay (individual) with corpus consultation Activity 3: Peer-feedback Homework: Self-correction
Fifth online lesson	Dec. 18, 2022	Collocation training 3	Use corpus tools as a reference source to complete the following activities: Activity 1: Gap-filling Activity 2: Error identification and correction Activity 3: Sentence building Homework: Creating the virtual corpus (VC) for topics Education and Culture

Figure 1 shows some activities used in the collocation training session.

Figure 1.

A sample of lesson activities

Handout #5

- Use COCA to find some common synonyms of the word "viewpoint". Type this query in the search box '=viewpoint'.
    - How many words can you find?
    - Which word has the highest frequency? Click on that word to examine the concordance lines. What do you notice? (word form, patterns)
    - Choose another word you are not familiar with and explore the concordance lines containing that word. What can you learn from that word?
- Now use the collocate function to find some common collocations with these synonyms. Type '=viewpoint' in the first box, and choose 'verb' for the POS in the second box.
    - Which verb has the highest frequency? Can I find other forms of this verb in the search results?
    - Examine the concordance lines of that verb and its other forms. Randomize the concordance lines by choosing a different number from ***"Find sample"***. What patterns of verb-noun collocations can you learn from that? (E.g. taking a long-term view).
    - What do you notice about the noun "position" in the collocations you found?
    - Take notes of five V-N collocations that you are interested in? (E.g. hold the view (that))
    - Share your notes with your partners

Activity 1: Error identification

Some students are writing about whether university education should be free or not. Look at the following sentences from these students' essays and identify collocation errors. Use the collocate function in COCA to check your intuition and correct those errors. (Note: Pay attention to the meaning of the whole sentence.)

1. The ministry can raise the taxes on employed people to free the college education which can raise the knowledge of the society.
2. Many individuals believe that in order to ensure that all students may attend education regardless of their financial situation, the government should make tuition free for all students.
3. To begin with, authorities should advocate financial support to all students to reduce social disparity.
4. In Vietnam, a number of scholarships are given to poor students yearly and help them reach the university.
5. Because if all students have access to university education but don't have to spend money, they will disregard this opportunity.

3.2.2 Data collection procedure

3.2.2.1 Collocation knowledge test

All participants took a collocation knowledge test to determine whether the participants are homogeneous in terms of their collocation knowledge. This test was adapted from the pre-test in Jafarpour et al. (2013) with twenty items of two-choice questions, and the reliability estimate calculated by KR-21 formula was .89 (p.54). However, when I piloted the test with several students with similar levels of the participants in this study, most of them achieved 100% correct answers. To avoid this ceiling effect, I modified the test by changing 20 two-choice questions into 20 gap-filling test items (Appendix C). Due to participant constraints, I was unable to conduct a second pilot test to calculate the reliability of the assessment tool. Instead, I calculated the descriptive statistics of the test items and included them in Appendix J.

3.2.2.2 Writing pre-test and post-test

A writing pre-test was taken by all participants. The task was to write an argumentative essay on the statement that attending a college is better than online learning, under timed condition (60 minutes) without any reference tools. I chose this topic because it is familiar to the participants, and they would be more likely to have ideas to write about. Normally in an exam condition, an essay assignment will be written within about 40 minutes (IELTS) or 30 minutes (TOEFL), but in this task, I set 60 minutes to hopefully eliminate the factor of anxiety which may influence students' writing performance.

The same procedure was applied to the writing post-test, with the same question. The rationale for choosing the same question for the pre-test and post-test is to control the possibility that any changes in the participants' use of verb-noun collocations would not be influenced by the topic and the complexity of the writing prompt. In the writing post-test, however, the participants

were allowed to use the corpus tools as their reference during the writing process. Given the six-week interval between the pre-test and post-test, it is likely that a practice effect was minimal.

The writing prompt used for the course was derived from a genuine question included in the 2008 IELTS writing test, with a topic focus on Education. This topic was selected in light of its relevance to the age and context of the course participants, thereby facilitating the participants' brainstorming process.

3.2.2.3 Background and evaluation questionnaires

Two questionnaires from Fang et al. (2021) were slightly modified and used in this study to obtain background information, students' prior knowledge of corpora, their English learning habits in general, as well as their perceptions about the effectiveness of the hands-on corpus-based instruction to their improvement in V-N collocation use and writing in general. The background questionnaire was created using Qualtrics and translated into Vietnamese before being distributed to the participants at the beginning of the writing course (Appendix A). The same procedure was applied to the evaluation questionnaire, but it was sent to the participants after the writing course had finished (Appendix C).

3.2.2.4 Semi-structured interviews

After the participants completed the evaluation questionnaire with the seven-point Likert scale with 1 for completely agree and 7 for completely disagree, I calculated the mean score for each participant. Six participants were then chosen to participate in a semi-structured interview. These were the participants who had the most positive (n = 3) and most negative evaluations (n = 3) based on the results of the evaluation questionnaire. These participants were chosen based on their mean scores in the evaluation questionnaire. To be more specific, Participant 004, Participant 0013, Participant 0022 had the highest mean scores at 6.89, 6.78, and 6.56, respectively. Participant

0017, 0019, 0025 had the lowest mean scores at 5.67, 5.67, and 1, respectively. All six participants were asked for their permission to voluntarily participate in the interview which lasted 30 minutes. The interviews were conducted in the participants' L1 (Vietnamese) so that it would be easier for them to express their opinions. It is noticeable that Participant 0025 had an extremely low mean score, at 1, which means he had a very negative attitude toward the intervention. However, during the interview, he clarified that it was a mistake. In fact, he highly appreciated the hands-on use of corpus tools.

3.2.2.5 Screen recordings

The participants recorded their writing process via screen capture during various stages of the study, including the writing pre-test, the two intervention essays, the writing post-test, and the peer review and self-correction process. To facilitate the submission of participant videos, a course was created on the Canvas platform. However, due to some participants' preference for Google Drive, a folder was also created on this platform to allow for submission of screen recordings. The videos were collected at various stages of the study, including the writing pre-test, the first and second essays during the course, the peer-feedback stage, the self-revision stage, and the writing post-test. A free screen recorder software was used for this purpose: https://www.apowersoft.com/free-online-screen-recorder.

3.2.3 Data analysis procedure

3.2.3.1 Collocation knowledge test

The scores of the collocation pre-test were imported into Microsoft Excel (ME) for data analysis to check the homogeneity of collocation proficiency of the participants. Given the limited number of participants in the present study, no outliers were removed from the data set. Nevertheless, in the interest of informing future iterations of the V-N collocation test, metrics such

as item facility (IF), item discrimination (ID), standard deviation (SD), KR-21 score, and standard error of measurement (SEM) will be reported.

3.2.3.2 Writing pre-test and post-test

3.2.3.2.1 Identifying the infelicitous collocations

As Yoon (2016) pointed out in his study, it is a daunting task to identify correct and incorrect collocations. Most of other studies referred to the judgement of human raters and then conducted an inter-rater reliability test with the disagreement being resolved through later discussions such as Crossley et al. (2014). Other studies such as Liou (2019) measured the accuracy of collocation use in students' writing by reporting the frequency of collocation errors per 100 words but did not give detailed description of how the errors were identified.

Another way researchers used for identification of collocation errors was checking extracted word combinations from students' writing in different dictionaries. For example, Laufer and Waldman (2011) compared the extracted verb-noun combinations with two dictionaries: *The BBI Dictionary of English Word Combinations* (Benson, 1986) and *The LTP Dictionary of Selected Collocations* (Hill & Lewis, 1997). If the combinations in students' writing matched the collocations in either one of the dictionaries, they were considered collocations.

A similar approach with some modification was adopted by Gao et al. (2019) which suggested that collocations were marked as correct collocations if they could be found in any two of these dictionaries: *Longman Dictionary of Contemporary English (LDCE)* (2015), *Oxford Advanced Learners' Dictionary (OALD)* (2000), *The BBI Dictionary of English Word Combinations* (1995) or in three different texts in the BNC (British National Corpus).

The problems with these approaches were explained in Yoon (2016), including the low inter-rater reliability in case of human judges and inappropriateness of collocations due to different

language context when it comes to dictionary comparison. To solve those issues, in his study, he suggested a way to measure the accuracy of collocations, using the association strength of a combination or in other words, the mutual information (MI score), which "indicates the degree to which two lexical items in a combination occur more often than would be expected by chance" (Yoon, 2016, p.43). An MI score of 3 has been widely accepted by researchers as a significant collocation benchmark (Hunston, 2002; Stubbs, 1995, cited in Yoon, 2016). However, one question that can be taken into consideration is whether a verb-noun combination which has a very high MI score but low frequency in the reference corpus such as *distill insight (MI score: 4.4; frequency: 3)* or *destress holiday (MI score: 5.62; frequency: 2)* can be regarded as an acceptable collocation.

Considering aforementioned factors, in this study, I will extract all verb-noun combinations from students' essays in the pre-writing test and post-writing test, and then check not only MI scores but also the frequency of these combinations in COCA. Yoon (2016) only considered the verb-noun combinations which appeared at least five times in COCA (p.47), whereas Wolter and Gyllstad (2013) set a higher number for the frequency of no fewer than ten occurrences in COCA (p.458). The default setting of frequency for collocation searches in COCA is twenty, so I think it is reasonable to raise a higher bar for the frequency. Therefore, those verb-noun combinations that do not meet the MI score threshold of 3.0 and appear less than 10 times in COCA will be considered as collocation errors. However, another issue arose when the verb-noun combination has an MI score that is lower than 3.0 but extremely high frequency and intuitively completely natural. In such cases, it is unjustified to call it a collocation error although it does not satisfy the above-mentioned criteria set for MI score and frequency. For example, the V-N combination *have opportunity* has a MI score that is lower than 3 (at 1.05) but its frequency is 10,357 in COCA. In

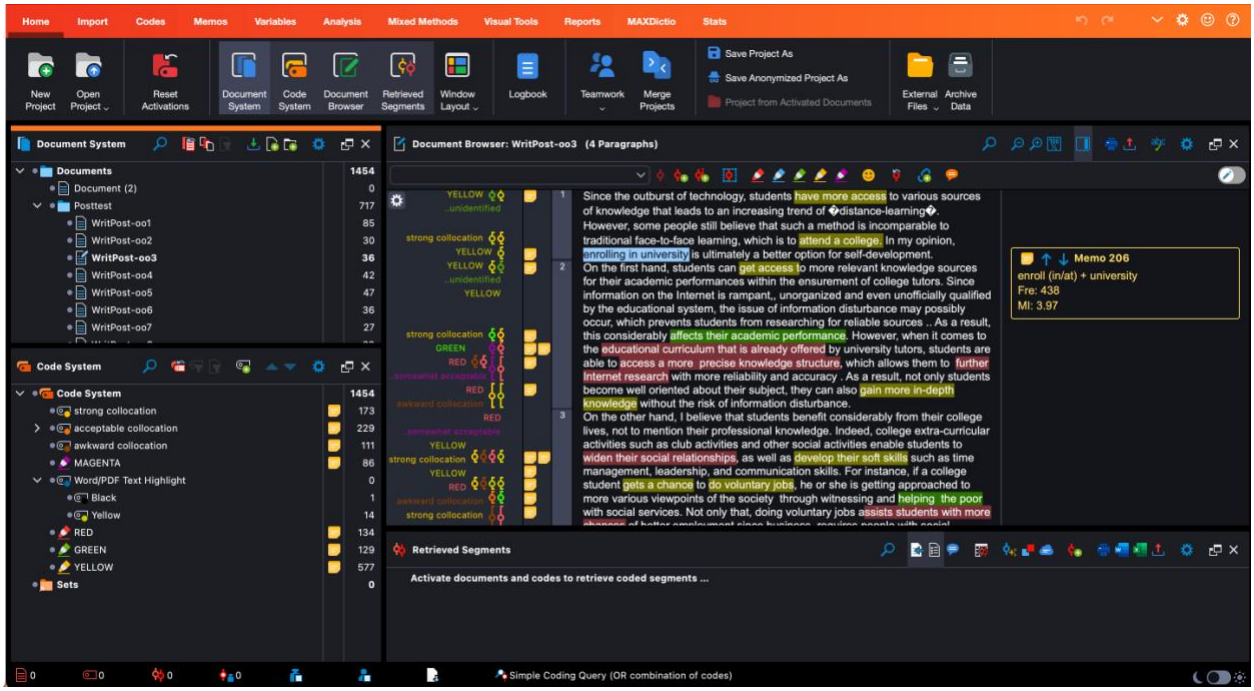
addition, the two L1 raters in this study also rated this combination as a *strong collocation*. If the cut-off points mentioned earlier were adopted, is it justifiable to label the V-N combination *have opportunity* as a collocation error? To solve this dilemma, I developed a 3-point Likert scale of rating, adapted from the three-point Likert scale in Akinci and Yildiz (2017): *awkward collocation, acceptable collocation,* and *strong collocation* to measure the collocation acceptability. The awkward collocations would be considered as collocation errors. The coding procedure and calculating collocation errors would be presented in detail in the methodology section.

3.2.3.2.2 Coding collocational acceptability

After downloading all the essays from Canvas or a Google Drive folder, I read through all the essays and corrected all spelling errors before manually identifying all verb-noun combinations. These combinations were highlighted in the original writing texts before they were imported to MAXQDA for coding collocation acceptability. The highlighted verb-noun combinations were subsequently checked with COCA in terms of their MI scores and frequency for their collocational acceptability to be coded, using the three-point Likert scale mentioned in the previous section. Using MI score and frequency indices in COCA, I set up criteria to categorize the verb-noun combinations extracted from the essays of the participants into three groups: 1 - *awkward collocation* (MI < 1 and/ or 10 < Fre < 100; MI > = 3 and Fre < 10), 2 - *acceptable collocation* (1 < = MI < 3 and Fre > = 10; MI < 1 and Fre > 100), and 3 - *strong collocation* (MI > = 3 and Fre > 10). The awkward collocations would be considered as collocation errors. I used this set of criteria to code the essays of the first half of the participants. Figure 2 illustrates the collocational acceptability coding process in MAXQDA.

Figure 2.

Collocational acceptability coding process in MAXQDA



In addition, fifty percent of the essays were double-coded by two L1 speaker raters (R1 & R2) using their intuitions. Rater 1 has had five years of teaching experience, both in EFL and ESL contexts. Rater 2 has had four years of teaching experience, both in EFL and ESL contexts.

I then calculated the inter-rater reliability between these two raters. The inter-rater reliability obtained from the Spearman's Rank correlation test between these two L1 raters was 0.85. The disagreements were discussed between them to reach a consensus on the final ratings. It was hypothesized that a low inter-rater reliability would be observed between the two L1 raters given that they would be coding the V-N collocations based on their own intuitions, which was suggested in Yoon (2016).

On the other hand, it was expected that the inter-rater reliability between intuition-based raters (R1 & R2) and the corpus-based rating (CBR) would be relatively strong, given that the data utilized by the latter is derived from authentic language as employed by L1 speakers in both written

and spoken contexts. In light of this, it was my intention to compare CBR with the ratings of the two L1 raters to assess the degree to which our results align, with the expectation of observing high inter-rater reliability. If it was the case, I would use CBR for the rest of the data, and only report the inter-rater reliability. However, if the inter-rater reliability was low, I would present two sets of results: one for the intuition-based rating and one for CBR. After that, I would use CBR for the rest of the data to identify the collocation errors.

These errors were tallied, and because the length of the essays is different, I used the standard measure of length, calculating the error rate per 100 words as suggested in the literature to guarantee that the result is comparable across essays of different length.

The Spearman's Rank correlation test was used to calculate the correlation between two L1 raters, rater 1 and rater 2 (using their intuition), between rater 1 and CBR (based on the MI score and frequency information in COCA), and rater 2 and CBR. While Spearman's Rank correlation coefficient can be a useful tool for assessing interrater reliability, it may not be sufficient on its own to fully capture the complexity of the data and the agreement between raters. Adjacent agreement provides a complementary measure of agreement, so I also calculated the adjacent agreement between rater 1 and rater 2, rater 1 and CBR, and rater 2 and CBR.

3.2.3.3 Average MI scores

To obtain MI scores of the verb-noun combinations in the participants' essays, I set certain criteria for the search string. I used the *collocates* function, with a specific verb put in the *word/phrase* box and a specific noun in the *collocates* box on COCA. The window span is four to the right. I also chose *relevance (MI score)* as the sorting method and lemma as the grouping method. These criteria were followed consistently across verb-noun combinations, apart from very common verbs such as *have, do, get* because of COCA default setting. In these cases, the search

string would be opposite, with the noun in the *word/phrase* box and the verb *have/do/ get* in the *collocates* box. The window span now is four to the left. In some contexts, especially in the case of *have* or *do*, the window span needs to be modified to two or three to the left to clean up the data when *have* or *do* plays the role of an auxiliary verb. The decision was made by considering the concordance lines. The average Mutual Information (MI) scores of each participant in the writing pre-test and post-test were computed and juxtaposed to determine whether or not the intervention had led to an improvement in their scores.

3.2.3.4 Background and evaluation questionnaires

Statistical data obtained from Qualtrics was subsequently utilized to inform a general report pertaining to participant backgrounds. Further details have been provided in Table 6, available within Appendix G.

Evaluation questionnaire results were similarly extracted from Qualtrics and analyzed using Microsoft Excel, with participant evaluations for each survey question being computed accordingly. Findings were subsequently presented in the form of a bar chart.

3.2.3.5 Semi-structured interviews

Due to the unavailability of software capable of transcribing Vietnamese language, all interviews were recorded and subsequently transcribed manually. The transcriptions were then translated into English with the utmost care taken to preserve the participants' opinions as accurately as possible. Thereafter, the translated material was imported into MAXQDA for thematic analysis, aimed at identifying recurring patterns and themes within the data. Because my research question was generally about the participants' attitude toward the hands-on corpus-based instruction in learning verb-noun collocations and writing in English, I did not establish predetermined themes. Therefore, to analyze the data, an inductive coding approach was utilized,

in line with the methodology outlined by Kostere and Kostere (2022, p.57-59). First, I read through the interview transcriptions, highlighting meaningful phrases or sentences, and then compared them with the research questions to decide which ones were relevant. I then looked for the connection of these meaningful and relevant phrases or sentences to develop patterns and themes.

3.2.3.6 Screen recordings

The videos recorded by the screen capture software introduced earlier were in *.webm* format, so they were imported to a video converter to be converted into *.mp4* format before being imported into MAXQDA for data analysis.

The analysis of the video recordings followed a procedure similar to that employed in the examination of the interview data. Initially, I viewed the first video, meticulously scrutinizing the patterns that emerged from the participants' utilization of the corpus tools at various stages of the intervention, particularly during the writing post-test. This facilitated the generalization of themes. Subsequently, I replicated this process for the remaining videos, incorporating modifications to the coding, as deemed necessary.

3.2.3.7 Corpus tools

The corpus tools used in this study were the Corpus of Contemporary American English https://www.english-corpora.org/coca/ and SKELL https://skell.sketchengine.eu/#home?lang=en for students' consultation. The Corpus of Contemporary American English (COCA) (Davies, 2008) was employed because of the large size of the corpus and its convenience associated with *collocate* function while SKELL was chosen to supplement COCA because of its user-friendly interface, not to mention the word *sketch function* to search for collocations.

4 RESULTS

In this section, I presented the descriptive stats and statistical test of the participants' average MI scores, the rate of infelicitous V-N collocations, the profiles of corpus consultation from the screen recordings, and the participants' attitude toward the hands-on corpus-based instruction from the evaluation questionnaire and the semi-structured interviews. The result of the collocation knowledge test was presented in Appendix J.

4.1 Average MI scores

Table 2 describes the average MI scores of 24 participants in the writing pre-test and post-test.

Table 2.

Average MI scores

Participant	Ave MI Pre	Ave MI Post
001	**1.24**	**2.15**
002	3.01	2.7
003	**2.05**	**2.82**
004	2.61	2.25
005	3.06	2.14
006	**3.13**	**3.24**
007	**3.16**	**3.47**
008	**1.43**	**2.56**
009	2.38	1.29
0010	**1.3**	**2.35**
0011	3.56	2.91

Table 2. (cont'd)

0012	2.59	2.45
0013	**1.56**	**2.35**
0014	**0.81**	**2.14**
0016	2.06	1.66
0017	2.65	2.5
0018	**2.21**	**3.25**
0019	**1.59**	**2.6**
0020	**2.1**	**3.34**
0021	0.45	0.38
0022	**2.75**	**3.44**
0023	1.96	1.93
0024	2.16	1.64
0025	**2.26**	**3.74**

As can be seen from Table 2, thirteen participants increased the average association strength (MI score) of the verb-noun collocations used in their essays, accounting for about 55% of the participants. The remaining participants had their average MI scores decrease. It is difficult to conclude at this point whether the direct corpus consultation was effective or ineffective in helping these participants improve the quality of verb-noun collocations in their written production. Further analysis from screen recordings and questionnaires were needed to provide more qualitative information for this result.

A paired sample t-test was conducted to analyze the data. Given that I hypothesized that the participants would perform better in the post-test, compared to the pre-test, I chose a one-sided paired sample t-test. The results are shown in Table 3.

Table 3.

Comparison of Pre- and Post-Intervention Scores

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	**Pre-test**	2.17	24	0.78	0.16
	Post-test	2.47	24	0.77	0.16

Paired Samples Correlations

		N	Correlation	p
Pair 1	**Pre-test & Post-test**	24	0.51	0.01

Paired Samples Test

		Paired Differences					t	df	Sig. p
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Difference Lower	Upper			
Pair 1	**Pre-test - Post-test**	-0.30	0.77	0.16	-0.63	0.02	-1.92	23	0.03

Paired Samples Effect Sizes

					95% Confidence Interval	
			Standardized	Point Estimate	Lower	Upper
Pair 1	**Pre-test - Post-test**	**Cohen's d**	0.77	-0.39	-0.80	0.03
		Hedges' correction	0.80	-0.38	-0.78	0.03

The participants had a larger mean score on the average MI scores of the V-N collocations in the post-test (M = 2.47, SD = 0.77) compared to the pre-test (M = 2.17, SD = 0.78). A one-tailed paired samples t-test revealed a significant difference in the test scores between the pre-test and post-test, $t(23) = -1.92$, $p = .03$. However, given the confidence intervals (-.80; .03) and the effect size cohen's d = -.39, the improvement in the average MI scores between the pre-test and the post-test, although statistically significant, needs to be interpreted with caution.

4.2 Infelicitous collocation rates

4.2.1 Identifying the infelicitous collocations

The infelicitous collocations were identified by using a Likert three-point scale with two L1 raters using their intuitions and the third rater (the researcher) using the combined information of MI score and frequency in COCA. Twenty-four essays from the first half of the participants were rated by the three raters, and two types of correlation tests were run to investigate whether there is an agreement on the ratings of collocation acceptability between the two L1 raters and whether two L1 speakers' intuitions and the statistical information (MI score and frequency) in COCA correlate with each other.

It was expected that there would be a low correlation between the two native raters, and in fact, the inter-rater reliability was relatively high with Spearman's Rank correlation coefficient ρ (R1, R2) = .85. Notably, the inter-rater reliability between rater 1 and CBR, as well as between rater 2 and CBR, was not satisfactory, contrary to the anticipated outcomes. Specifically, the correlation coefficients indicated that there was only a moderate level of agreement, with ρ (R1, CBR) = .48 and ρ (R2, R3) = .55. Table 4 presents how different the intuition-based ratings and corpus-based rating was.

Table 4.

Discrepancies in ratings

		R1-R2	
dif-2	dif-1	dif-0	Total
0	66	411	477
		R1-CBR	
dif-2	dif-1	dif-0	Total
33	202	242	477
		R2-CBR	
dif-2	dif-1	dif-0	Total
27	193	257	477

(R1 and R2: L1 speakers, ratings were based on the raters' intuitions; CBR: the rating was based on the combination of MI score and frequency information in COCA; dif-2: the number of collocations that have two-point difference between the two raters; dif-1: the number of collocations that have one-point difference between the two raters; dif-0: the number of collocations that have no difference between the two raters.)

4.2.2 Adjacent Agreement

Adjacent agreement is a measure used in interrater reliability analysis to assess the level of agreement between two or more raters who are evaluating the same set of data. It specifically refers to the degree to which raters agree on whether two adjacent items or observations belong to the same category or have the same attribute. In other words, "percent adjacent measures the number of times the scores were exactly the same plus the number of times the scores were only one level different" (University of North Carolina Wilmington, n.d., Assessment in UNCW, para. 2). Table 5 presents the adjacent agreement between raters.

Table 5.

Inter-rater reliability from adjacent agreement

R1-R2	1
R1-CBR	.93
R2-CBR	.94

4.2.3 Infelicitous collocation rates per 100 words

To determine the frequency of infelicitous collocations per 100 words in 48 essays written by 24 participants during the writing pre-test and post-test, I tallied the occurrences of V-N combinations labeled as 1 (infelicitous collocations). I then divided this count by the total number of words in each essay. The findings are reported in Table 6.

Table 6.

Infelicitous collocation rates per 100 words

Participant	**Pre Miscollocation Rate**	**Post Miscollocation Rate**	**dif**
003	2.94	0.89	-2.05
0018	1.89	0.35	-1.54
0020	2.07	0.72	-1.35
0012	1.43	0.32	-1.11
008	2.33	1.34	-0.98
0010	2.86	2.04	-0.82
0013	2.51	1.89	-0.61
002	1.82	1.26	-0.56

Table 6. (cont'd)

007	0.61	0.29	-0.32
0017	1.35	1.04	-0.31
0019	0.86	0.57	-0.28
0014	2.17	2.02	-0.16
0022	1.57	1.64	0.07
0023	0.90	1.03	0.13
0025	0.40	0.82	0.41
0021	2.08	2.72	0.64
006	1.20	1.98	0.78
0011	1.01	1.82	0.81
0024	1.61	2.46	0.85
009	1.35	2.41	1.05
005	0.67	1.86	1.20
001	2.41	3.79	1.38
004	0.93	2.37	1.44
0016	0.64	2.41	1.77

As can be seen from the table, half of the participants reduced the proportion of infelicitous V-N collocations in their essays between the writing pre-test and post-test after the intervention. The other half, on the other hand, used more infelicitous V-N collocations per 100 words in the post-test, compared to the pre-test. Again, this result needs to be interpreted with caution and triangulated with the qualitative data to fully understand the nature of the changes observed.

4.3 Screen video recordings

In order to understand why some participants improved the accuracy of the V-N collocations used in their essays while some did not, it is necessary to examine how these participants utilized the corpus tools in the writing post-test as well as during the intervention sessions. Considering the potential influence of the regression to the mean phenomenon, which posits that exceptionally high or low scores on a test may tend to converge towards the mean score on subsequent tests, regardless of any external factors, a selective analysis of participants' screen recordings was conducted.

Specifically, four participants were chosen for in-depth analysis based on their initial and subsequent scores on the average MI scores in their pre-test and post-test. These included individuals who exhibited the lowest average MI score but demonstrated significant improvement, those who began with the lowest average MI score but showed no improvement, those who started with the highest average MI score but improved, and those who had the highest MI score initially but showed no improvement. This approach aimed to shed light on the processes and factors contributing to variations in participants' performance on the average MI scores, while also taking into account the potential influence of regression to the mean. The detailed profiles of these four participants were presented as follows.

4.3.1 The profile of Participant 007 (high MI pre – high MI post)

Participant 007 is a 16-year-old male student. According to the background survey, he writes in English three times a month, and his main sources of learning English are Cambridge and Longman dictionaries. He is absolutely interested in trying new things and English collocations though he partly agrees that collocations cause him difficulty in writing English. He has never heard of or used any corpora before. He is often exposed to other sources of English outside

classroom more than three hours a day through news channels such as VOX and BBC. He sometimes pays attention to collocations from those sources but never tries to use them in his writing. He got 14/20 correct answers in the collocation knowledge test. The participant showed a general positive attitude toward the hands-on consultation of corpus tools in learning V-N collocations and writing although regarding the benefit of retaining more collocations after the corpus-based instruction, he did not give an opinion. However, he partly agreed that he could always find the information he needed when he consulted the corpus tools. Table 7 describes the process of corpus consultation of Participant 007 in detail.

Table 7.

Participant 007's profile of corpus consultation

Stage	**Timestamp**	**Action**	**Detail**
Writing post-test	N/A	N/A	not used the corpus tools or other reference sources
Essay 1 - Writing process	N/A	N/A	not used the corpus tools or other reference sources
Essay 2 - Writing process	N/A	N/A	N/A (technical problems)
Essay 1 - Peer feedback	N/A	N/A	N/A (technical problems)
Essay 1 - Self correction	N/A	N/A	N/A (technical problems)

Table 7. (cont'd)

Essay 2 - Peer feedback	02:02-04:06	used COCA to check the accuracy of a verb-noun collocation	wanted to check the V-N collocation *display tendency* with the search pattern *tendency* -> verb collocates -> set the correct part of speech (POS) for the noun, the window span is 4 to the left and 4 to the right -> used MI score (*relevance*) to sort the collocations -> found the verb *display* in the result -> recorded the MI score and frequency -> examined the concordance lines of the verb *display* and noticed the pattern following this collocation
	04:07-08:06		checked the V-N collocation *impact functioning* with search pattern was *functioning* -> verb collocates -> found the results in COCA and examined the concordance lines for the verb *impact* -> recorded the MI score and frequency, but it took him a while to decide whether it is a good collocation or not because of the low frequency

Table 7. (cont'd)

	10:05-11:36		continued checking the V-N collocation *shape mindset* with the search pattern was similar to the above-mentioned in COCA: *mindset* -> verb collocates -> Scanning the verb list, he did not find the verb *shape*. However, he found another verb *develop* with a suitable meaning, a better MI score and frequency -> replaced the original verb with *develop*. He didn't record the MI score and frequency as well as examining the concordance lines of the verb *develop*
	12:21-13:12		checked the V-N combination *achieve outcome* in COCA. The similar search pattern was used -> found the verb *achieve* in the result -> recorded the MI score and frequency. He didn't examine the concordance lines this time
Essay 2 - Self-correction	06:56-11:16	searched for the strong V-N collocations	used the *collocates* function in COCA to search for the verb that collocates with *limit*. After scanning the verb list, he examined the concordance lines of the verb *impose*. It took him a while to think before he continued searching for the verb that collocates with *possibility*, using a similar search pattern *possibility* -> verb collocates -> examined the concordance lines of the verb *eliminate* -> modified his writing into *eliminate the possibility*. He did not record the MI score and frequency

Table 7. (cont'd)

11:22-12:50	searched for the verb that collocates with the noun *likelihood,* using a similar search pattern *likelihood* -> verb collocates in COCA -> found a suitable verb to replace the original one which is marked "not a strong collocation" by his peer -> examined the concordance lines of the verb *predict* and decided to use that verb. He did not record the MI score and frequency

In summary, Participant 007 used the *collocates* function in COCA during his corpus consultation to check the accuracy of the V-N combinations and look for strong collocations. He always established correct part of speech (POS) setting as well as the sorting criteria. He mostly examined the concordance lines for almost every search that he conducted. He also recorded the MI scores and frequency though he sometimes failed to do that.

4.3.2 The profile of Participant 0011 (high MI pre – low MI post)

Participant 0011 is a 17-year-old female student. She writes in English once a week and often uses Cambridge and Oxford Learner's dictionary as well as the online collocation dictionary Ozdic. She is absolutely interested in trying new things as well as English collocations and totally agreed that collocations cause her difficulty in writing in English. She has heard about corpora but had not used any of them before. Outside classroom, she is often exposed to English two hours a day through YouTube videos, podcasts, books, music, movies, and news. She sometimes pays attention to the collocations she is exposed through those sources and sometimes applies them to her writing. She got 18/20 correct answers in the collocation knowledge test. This participant showed an overall positive attitude toward the corpus tools and the hands-on corpus-based

instruction with all “agree” and “totally agree” responses. Table 8 describes the process of corpus consultation of this participant in detail.

Table 8.

Participant 0011’s profile of corpus consultation

Stage	**Timestamp**	**Action**	**Detail**
Writing post-test	24:48-26:08	used other sources rather than the corpus tools – ineffective	She was probably uncertain about the meaning of the verb *emulate* -> looked it up in a bilingual online dictionary -> opened COCA -> did not use it -> instead, went to an online dictionary to check the meaning of the verb *emulate* but did not obtain any information about the noun that collocates with it -> wrote *emulate each other’s positive qualities* in her essay without checking the accuracy of the V-N collocation *emulate quality* with COCA or SKELL
	31:30-32:46	checked the accuracy of a V-N combination	She used the *collocates* function in COCA with the search pattern *experience* -> verb collocates -> established a correct POS setting, sorted the result by MI score (*relevance)* with the window span 4:4 (4 words to the left and 4 words to the right) -> did not find the verb *acquire,* but found another suitable verb with a high MI score and frequency: *enhance* -> replaced *acquire* with *enhance.* She neither recorded the MI score and frequency nor examine the concordance lines of the verb *enhance*
Essay 1 – Writing process	N/A	N/A	not used the corpus tools or any other sources

Table 8. (cont'd)

Essay 2 – Writing process	06:01-06:30	used other sources rather than the corpus tools	looked up the word in a bilingual online dictionary and then checked it with the Oxford Learner's dictionary -> did not use any sources after that for the rest of the writing process.
Essay 1 – Peer feedback	06:16-09:38	checked the accuracy of a V-N combination	check the accuracy of the V-N collocation *boost quality* with the search patter *quality* -> verb collocates -> established a correct POS setting with the window span 4:4 -> found the verb *boost,* but the MI score and frequency are not as high as expected -> did not record the MI score and frequency of this verb -> scanned the verb list to find the verbs that are likely to collocate with *ability* and have a suitable meaning -> made some suggestions for her peers to improve their collocation -> did not examine the concordance lines of the verbs she found

Table 8. (cont'd)

09:46-13:30	checked the accuracy of the V-N collocation *enhance personality* with the similar search pattern and settings -> did not find the verb *enhance* in the result, concluding that it is not a good collocation -> looked at several verbs in the list provided by COCA and checked the meaning of the verb *correlate* in a bilingual online dictionary -> suggested a better collocation for her peer -> did not record the MI score and frequency as well as examining the concordance lines
13:44-17:46	checked whether *boost interaction* was a strong collocation -> used the *collocates* function in COCA with the same settings -> did not find the verb *boost*, concluding that it is not a good collocation -> continued scanning the verb list and paid more attention to those that have a high MI score and frequency -> examined the concordance lines of the verb *promote* ->looked at some other verbs and choose two other verbs which have the similar meaning *enhance* and *foster* to make suggestions for her peer. She did not record the MI score and frequency of these new V-N collocations as well as examining their concordance lines

Table 8. (cont'd)

17:46-20:34	checked the accuracy of *boost (memory) storage* -> used the similar search pattern *storage* -> verb collocates -> did not find the verb *boost,* stating that it is not a good collocation -> looked at other verbs in the result, choosing the verb *expand* to recommend to her peer to maintain the original meaning of the sentence -> did not examine the concordance lines of the verb as well as recording the MI score and frequency -> used SKELL to look for appropriate V-N collocations with *storage,* but it did not work well
20:48-26:16	checked the V-N combination *concrete knowledge* -> first checked the meaning of *concrete* in a bilingual dictionary to understand it in her L1 and looked it up in the Oxford Learner's dictionary -> came back to COCA, using the search pattern *knowledge* -> verb collocates -> did not find the verb *concrete,* concluding that it is not an acceptable collocation -> looked at other verbs and made several suggestions of collocations with high MI scores and frequency for her peer -> considered the meaning appropriateness -> did not examine the concordance lines of any verbs or record the MI score and frequency -> checked the collocation with SKELL as well but did not actually use the result from SKELL

Table 8. (cont'd)

26:18-27:02	checked the verb-noun combination *satisfy craving* -> used the *collocates* function and found the verb *satisfy* immediately with the high MI score and frequency -> did not make comments on that collocation -> did not record the MI score and frequency as well as examining the concordance lines -> did not look at other verbs either
27:06-30:00	checked the V-N combination *enrich potential* -> used the *collocates* function and did not find *enrich* in the result table -> scanned through the results and made suggestions for their peer with the verbs that have a high MI score and frequency -> did not record the information as well as examine the concordance lines
30:02-33:38	checked the accuracy of the V-N combination *enhance interest* -> used the *collocates* function and did not find *enhance* in the result, concluding it is not a good collocation -> looked at other verbs, considering the meaning of the sentence and making suggestions for her peer with the verb whose MI score and frequency are high -> examined the concordance lines of *spark interest* and *generate interest* -> used a bilingual dictionary to look up the meaning of *spark*, and Oxford Learner's Dictionary to check the meaning of *spark* and *generate*

Table 8. (cont'd)

	33:42-37:02	checked the accuracy of the V-N combination *increase curiosity* -> used SKELL first but only looked at the result briefly -> moved to COCA and used the *collocates* function -> did not find the verb *increase* in the result. -> looked at other verbs and found some verbs with a high MI score and frequency to make suggestions for her peer -> quickly looked at the concordance lines for the verb *excite*, but did not record the MI score and frequency of the verb -> used SKELL but did not stop there for long
Essay 2 – Peer feedback	00:59-02:04	checked the accuracy of the V-N combination *determine manner* with a similar search pattern *manner* -> verb collocates -> established a correct POS setting -> scanned through the verb list but did not find *determine,* concluding that it is not a good collocation -> did not take any other actions after that
	02:22-04:01	checked the accuracy of the V-N combination *embark path* with the same search pattern -> scanned the verb list and found the verb *embark* -> recorded the MI score and frequency as well as examining the concordance lines
	04:22-05:10	checked the V-N combination *expect child.* After scanning the verb list, she did not find the verb *expect* -> ignored this combination and moved on with another one -> did not take any other actions regarding the combination

Table 8. (cont'd)

05:12-06:04	checked the V-N combination *discharge hormone* -> scanned the verb list but did not find the verb -> did not make suggestions for her peer -> failed to record the MI score and frequency and failed to examine the concordance lines
06:14-09:29	checked the accuracy of the V-N combination *increase stress/ aggressiveness* with COCA -> found the verb *increase* in the result -> recorded the MI score and frequency -> did not look at other verbs and examine the concordance lines
09:44-10:34	checked the V-N combination *have disorder* using the same procedure -> did not find the verb *have* in the result ->did not look at other verbs or take any other actions
10:38-11:16	replicated the above-mentioned behavior with the V-N combination *carry disorder*
11:23-12:34	checked the accuracy of the V-N combination *put halt* with COCA using the same search pattern -> found the verb *put* in the result and recorded the MI score as well as frequency -> did not examine the concordance lines
12:44-13:45	check the accuracy of the V-N combination *show caring* -> did not find the verb *show* -> looked at the concordance lines of the verb *foster* -> did not make suggestions for her peer or take other actions

Table 8. (cont'd)

13:50-14:25	checked the accuracy of the V-N combination *share advice,* using the *collocates* function -> did not find the verb *share* in the result -> took no other actions afterwards
14:31-15:16	checked the accuracy of the V-N combination *reinforce bond,* using the *collocates* function -> did not find the verb *reinforce* in the result -> took no other actions afterwards
15:18-15:55	checked the V-N combination *tackle issue* in COCA -> quickly found the verb *tackle* -> recorded the MI score and frequency but did not examine the concordance lines
16:04-17:28	looked for *rule out probability, have bond,* and *carry out method* using the *collocates* function in COCA -> scanned the verb list but did not find the verb *rule (out), have,* and *carry (out)* -> took no other actions after that
17:50-18:40	
18:43-19:22	
19:28-20:20	checked the V-N combination *put brake* -> found the verb *put* in COCA and examined the concordance lines -> continued searching for the collocation on Google. However, for some reason, she concluded this collocation was not found.

Table 8. (cont'd)

Essay 1 – Self-correction	00:26-01:51	searched for strong V-N collocations	looked for the verb that collocates with *performance* -> used the noun *performance* as the node and found the verb collocates with the correct setting of POS -> did not look at the result, but replaced the noun *performance* with *result*, using her intuitions without checking with corpora or other sources
	01:58-05:29		searched for the verb that collocates with *interaction* with SKELL and found several verbs that can be suitable -> checked the meaning of *foster* in an online dictionary and then checked the verb with COCA -> examined the concordance lines of *foster* and *mediate* in COCA -> decided to replace the original verb with *foster* -> did not record the MI score and frequency
	05:49-07:32		looked up the meaning of the verbs *validate* and *value* in an online dictionary -> opened COCA but did not conduct any search -> gave up and ignored the feedback of her peer and the instructor

Table 8. (cont'd)

07:50-09:42	searched for the verb that collocates with the noun *enthusiasm* with the search pattern *enthusiasm* -> verb collocates -> scanned the verb list and considered some verbs -> used SKELL to look for suitable verbs -> looked up the verb *wane* in an online dictionary to check the meaning -> replaced the original verb with *generate/ inspire* found in COCA -> did not examine the concordance lines or record the MI score and frequency
10:34-11:24	conducted the search *self-esteem* -> verb collocates in COCA -> scanned the verb list and found some verbs to replace the original inappropriate verb -> did not examine the concordance lines or record the MI score and frequency
11:42-14:24	used SKELL to look for the verb that collocates with the noun *grasp* -> scanned the result in SKELL but did not find any useful outcomes due to the wrong setting of POS -> realized the problem and changed the POS setting but still did not find the needed verb -> came back to COCA and examine the concordance lines of the verb *concentrate* -> chose *gain* without examining the concordance lines or recording the MI score and frequency

Table 8. (cont'd)

	19:02-20:48		continued searching for the verb that collocates with the noun *proposition* in COCA -> examined the concordance lines of the verb *bolster* -> used SKELL and considered some verbs in the result, but it seemed unhelpful -> tried to look up the verb *bolster* in an online dictionary -> did not find any suitable verbs -> replaced the original noun with a different one using her intuition, but it is not suitable in meaning -> did not check the new combination in the corpus tools
Essay 2 – Self-correction	00:34-00:42 01:20-01:30	used intuition instead of consulting the corpus tools	did not use any corpus tools as a reference for her self-correction -> only used her intuition

In summary, Participant 0011 used both corpus tools and other sources in different stages of the writing processes. However, when she consulted COCA, she rarely examined the concordance lines of the collocations she found. She often used her intuition during the process of self-correction and sometimes ignored the feedback provided by her peers and the instructor.

4.3.3 The profile of Participant 0014 (low MI pre – high MI post)

Participant 0014 is a 17-year-old female student. She writes in English once a week, and the main source that she often uses to learn English is online dictionaries. She showed a reluctance to her preference for trying new things but an absolute interest in learning English collocations. She partly agreed that collocations cause her difficulty writing in English. She has heard of corpora but never used them before. Other sources that she uses to expose herself to English outside

classroom two hours a day include YouTube videos, Netflix, online materials. She sometimes pays attention to collocations used in these sources but seldom applies them to her writing. She got 10/20 correct answers in the collocation knowledge test. This participant also showed a general positive attitude toward the corpus tools and the hands-on corpus-based lessons, but she only partly agreed that she could always find the information she needed after consulting the corpora. Table 9 describes the process of corpus consultation of Participant 0014 in detail.

Table 9.

Participant 0014's profile of corpus consultation

Stage	**Timestamp**	**Action**	**Detail**
Writing post-test	38:56-42-30	checked the accuracy of the V-N combination	used COCA to check the V-N combination *perform method* -> used the verb *perform* as the node and looked for the noun collocate -> did not find the noun *method* as well as another suitable noun -> changed their search pattern into *method* -> adjective collocates, which is relatively strange -> did not find the desirable outcome -> used SKELL to check the collocations of the word *educational* -> examined the list of the nouns that can go with *educational* -> did not work -> did not find a new verb or new noun to replace the original ones, thus changing her writing by only keeping *education* and deleting other words in the combination

Table 9. (cont'd)

48:08-48-58	searched for other types of collocations	used the *collocates* function in COCA to search for the adjective that collocates with the noun *cost* -> did not set the sorting by MI score -> only saw the frequency of the collocations -> did not find the adjective *astronomical* in the result -> replaced that adjective with *high* which was seen in the result in COCA with high frequency -> did not examine the concordance lines or record the MI score and frequency
52:12-57:02	searched for other types of collocations	used SKELL to search for the modifier of the noun *networking* -> did not find what she needed -> used the *word* function in COCA to look up *networking* -> examined some concordance lines that contain *networking*, but there was no satisfactory result -> went back to her writing and add *widespread* before *social networking* -> did not use the tools to double-check the new combination -> came back to SKELL and looked for some examples of verbs that collocate with *networking* -> looked at the result for *networking* as subjects (which is not the pattern she was supposed to look for) -> did not find any desired result

Table 9. (cont'd)

Essay 1 - Writing process	00:49-02:26	searched for strong V-N collocations	used the *word* function in COCA to examine the noun *ability* -> changed her mind and used Google to look for the English equivalent of a Vietnamese word *understanding* -> put this noun into COCA, examining it using the *word* function -> did not find what she wanted -> went back to their writing and replaced the noun *ability* with *harmony* -> did not double check the new combination with the corpus tools.
	04:20-05:22	searched for synonyms	used the *word* function in COCA to look for the synonym of the noun *excellence* -> chose the noun *superiority*
	08:49-10:34	checked the accuracy of the V-N combination	used the *word* function in COCA to look up the verb *grant* -> chose the *collocates* information to lead her to the collocation table -> scanned the noun list and examined the concordance lines of the noun *privilege* -> used the collocation that she just found in her writing -> did not record the MI score and frequency
	43:40-44:32	checked the accuracy of the V-N combination	used the *word* function to examine the noun *knowledge* -> might not have found what she needed on the result page -> did not use the result from COCA

Table 9. (cont'd)

Essay 1 - Peer feedback	01:58-02:11	searched for strong V-N collocations	used SKELL to check the accuracy of *bring change* -> found the verb in the result page -> did not double-check the collocation with COCA
	02:13-05:00	checked the accuracy of the V-N combination and looked for strong V-N collocations	searched for the verb collocating with the noun *view* on SKELL -> did not see the verb in the result -> quickly opened COCA to use the *collocates* function to check the previous V-N combination *change* -> verb collocates -> scanned the verb list but did not take any action -> looked at SKELL again to see the result with *view* -> did not find the verb in the original text -> changed her search to look for the noun that collocates with the verb *raise* -> did not find the noun *view* either -> concluded the V-N combination *raise view* is not a good collocation -> used the result in SKELL to make suggestions for her peer: *raise concern/ question*, which are suitable in the context
	05:02-05:36		checked the accuracy of the V-N combination *offer opportunity* -> found the verb *offer* in the result, concluding it a good collocation -> did not double-check it with COCA

Table 9. (cont'd)

05:52-08:06	used the *word sketch* function in SKELL to search for the verb that collocates with the noun *identity* -> did not find the verb *develop* -> moved to COCA with the search pattern *identity* -> verb collocates -> found the verb *develop* and recorded the MI score as well as frequency -> did not examine the concordance lines -> looked at the result on SKELL again to make a suggestion for her peer to improve: *shape identity*
08:10-08:38	checked the accuracy of the V-N combination *exert influence* -> found this in SKELL, concluding it is a good collocation -> did not look at the examples or double-check with COCA
08:38-11:46	searched for the verb that collocates with the noun *preview* -> did not find the verb in the text. She then utilized the *collocates* function in COCA with the search pattern *preview* -> verb collocates -> did not find the verb either -> concluded that *provide preview* is not a good collocation -> went back to the result page in COCA and began looking through some verbs in the result -> switched to SKELL to examine some examples but could not find a suitable verb -> put the verb *provide* in the search box without any better results -> searched with the noun *preview* again in SKELL and examined the examples of the verb *release* -> chose that verb to recommend to her peer

Table 9. (cont'd)

	12:02-12:18		checked the V-N combination *deny request* -> found the verb in the result, concluding it is a good collocation -> did not look at the examples or double-check the collocation with COCA
	12:20-13:10		used word sketch function in SKELL to check the V-N combination *pique interest* with the search pattern *pique* -> noun collocates -> found the noun *interest* in the result -> moved to COCA to double-check this collocation with the search pattern *pique* -> noun collocates -> found the noun *interest* immediately and recorded the MI score as well as frequency -> did not examine the concordance lines
Essay 2 - Writing process	N/A	N/A	N/A (reported technical problems)
Essay 2 - Peer feedback	N/A	N/A	
Essay 1 - Self-correction	N/A	N/A	
Essay 2 - Self-correction	N/A	N/A	

In summary, Participant 0014 utilized both corpus tools, COCA and SKELL, to double-check or search for needed collocations when one of them did not work. She tended to vary the

search patterns and flexibly switched between verb (node) -> noun (collocate) and noun (node) -> verb (collocate). In addition, she used the *word sketch* function which is intended for collocations in SKELL.

4.3.4 The profile of Participant 0021 (low MI pre – low MI post)

Participant 0021 is a 17-year-old female student. She writes in English three times a month and mainly uses online dictionaries such as Cambridge dictionary or newspapers related to the writing topics to learn English. She indicates a potential interest in trying new things and English collocation. She partly agreed that collocations cause here difficulty writing in English. She has not heard of or use any corpora before. Outside classroom, she is often exposed to English one hour a day through inspiring short videos. She sometimes pays attention to the collocations used in these sources and applies them to her writing. She got 12/20 correct answers in the collocation knowledge test. Similar to the previous participants, the participant 0021 showed a general positive attitude toward the corpus tools and the hands-on corpus-based instruction, but she only partly agreed that she could always find the information she needed when she consulted the corpora. This participant's details of corpus consultation are illustrated in Table 10.

Table 10.

Participant 0021's profile of corpus consultation

Stage	**Timestamp**	**Action**	**Detail**
Writing post-test	03:34-04:18	checked the accuracy of collocations with corpus tools	used the *example* function in SKELL to check the examples of the word *home-based* and then *learning trends* -> did not take any specific action after the corpus consultation

Table 10. (cont'd)

05:28-05:58 07:04-07:18		repeated this behavior with the adjective-noun combination *indispensable tool* and the verb-noun combination *acquire knowledge*
11:24-11:38 15:32-15:50 17:40-17:52	checked the accuracy of other types of collocations with corpus tools	used the *example* function of SKELL to check the accuracy of the compound noun *family background,* the adjective-preposition combination *good for,* and the adjective-noun combination *educational institute* -> examined the examples of these combinations
19:18-19:37	checked the accuracy of V-N collocations with corpus tools	continued using the *example* function in SKELL to check the accuracy of the V-N combination *gain horizons* -> did not find the result -> used a different search string *gain knowledge* and found the examples of this combination -> replaced *horizons* with *knowledge* -> did not double-check the result with COCA
28:28-28:53	searched for the synonyms with corpus tools	used the *similar words* function in SKELL to look for the synonym of the word *aforementioned* but did not use the result generated by SKELL -> used Google to check the meaning of this word but no other actions were taken

Table 10. (cont'd)

Essay 1 - Writing process	N/A	N/A	N/A (reported technical problems)
Essay 2 - Writing process	06:48-08:18	searched for V-N collocations with corpus tools - failed attempts	used the *collocates* function to search for the verb that collocates with the word *innate* -> identified this word as a noun, which is incorrect, and established the wrong POS setting in COCA -> COCA did not respond -> did not realize the problem and gave up the search
	19:09-19:40		searched for the verb collocating with the noun *crime* with the search pattern *crime* -> verb collocates -> established inaccurate POS setting -> encountered a technical error -> did not log in again -> COCA did not allow her to continue her search because it said she exceeded the limit number of searches for a day
Essay 1 - Peer feedback	01:43-04:42	checked the accuracy of other types of collocations	used the *collocates* function in COCA with the search pattern *development* -> adjective collocates -> did not establish the correct POS setting and proper sorting criteria, thus failing to see the MI score -> did not vary the search pattern to find the answer

Table 10. (cont'd)

10:06-10:58	checked the accuracy of V-N collocations	continued the search with the pattern *knowledge* -> verb collocates in COCA -> found the verb *possess* in the result and recorded the MI score as well as frequency -> did not examine the concordance lines
11:05-13:42		conducted the next search *difficulty* -> verb collocates in COCA -> did not find the verb *apprehend* -> chose the verb *comprehend* from the result in COCA to recommend to her peer -> did not record the MI score and frequency or examine the concordance lines
13:44-14:30		continued searching with the pattern *tension* -> verb collocates -> scanned the verb list to find the verb *evaluate* and found it -> recorded the MI score and frequency but did not examine the concordance lines
14:32-14:48		checked the accuracy of the V-N combination *achieve result* using the *collocates* function in COCA -> found the verb in the result and recorded the MI score as well as frequency -> did not examine the concordance lines

Table 10. (cont'd)

15:00-15:32	checked the V-N combination *encourage development* -> found the verb in COCA and recorded the MI score as well as frequency but did not examine the concordance lines
15:36-16:20	searched for the verb that collocates with the noun *flaws* -> encountered a technical error -> logged in again -> forgot to establish proper sorting criteria -> did not see the MI score -> did not examine the concordance lines
16:24-17:30	continued her search with the pattern *insight* -> verb collocates in COCA -> did not obtain the MI score due to the inappropriate sorting criteria -> did not find the verb *acquire* after scanning the verb list several times -> did not take other actions
17:32-17:54	used the *collocates* function in COCA with the search pattern *opinion* -> verb collocates -> found the verb *voice* in the result and recorded the frequency information -> did not set proper sorting criteria -> did not see the MI score of the collocation -> did not examine the concordance lines

Table 10. (cont'd)

	17:56-18:18	continued using COCA with the search pattern *confidence* -> verb collocates -> found the verb *boost* in the result and recorded the MI score as well as frequency -> did not examine the concordance lines
	18:20-18:50	continued with the search pattern *need* -> verb collocates -> did not establish the correct POS setting for the noun *need* -> did not receive the desirable result from COCA -> did not double-check with SKELL
Essay 2 - Peer feedback	00:58-03:36	used the *collocates* function to check the verb that collocates with the noun *potential* with the search pattern *potential* -> verb collocates. However, due to the wrong setting of POS, the result generated by COCA was for *potential* as an adjective. The participant did not realize the problem -> scanned the verb list and did not find the verb *recognize* -> made a suggestion of using the verb *identify* instead -> recorded the MI score and frequency of this verb but did not examine the concordance lines
	04:03:04:36	continued using COCA to check the accuracy of the V-N combination *commit crime, break law,* and *trigger response* -> found the verbs in the result and recorded the MI scores as well as frequency -> did not examine the concordance lines

Table 10. (cont'd)

	00:39-05:21		
	05:42-06:48		
	07:40-09:12		used the *collocates* function in COCA to check the accuracy of the V-N combination *needs* -> verb collocates -> did not set proper sorting criteria, so she did not see the MI score -> did not find the verb *provide* in the result, so she chose the verb *met* from the verb list produced by COCA to make suggestions for her peer -> recorded the frequency information but did not examine the concordance lines
	09:16-10:54		used the search pattern *crime* -> verb collocates to check whether the verb *prevent* collocates with *crime* -> found the verb in the result and recorded the MI score as well as frequency -> suggested another verb with a higher MI score *deter* and recorded the MI score and frequency of this collocation -> did not examine the concordance lines
Essay 1 - Self-correction	00:14-00:26	used intuition instead of consulting the corpus tools	corrected their own writing based her peer's feedback without consulting any corpus tools such as COCA or SKELL, using her intuition instead
	01:22-02:14		

Table 10. (cont'd)

05:14-06:09	searched for V-N collocations with corpus tools	used the *collocates* function in COCA to look for the verb that collocates with the noun *connection.* She found the verb *establish* in the result with a high MI score and frequency. She replaced the original verb *have* with this verb. She did not record the MI score and frequency or examine the concordance lines
06:12-06:47		continued with the search pattern *period* -> verb collocates -> chose the verb *undergo* in the result to replace the verb *access* in the original text. Though the verb appeared in the result with an acceptable MI score and frequency, the meaning is not suitable in the context of the sentence. She did not record the MI score and frequency or examine the concordance lines
07:10-07:34		continued with the search pattern *approach* -> verb collocates -> found the verb *adopt* with a high MI score and frequency -> used it to replace the original verb *plan* -> did not record MI score and frequency or examine the concordance lines. The meaning of the new collocation is not completely appropriate but still better than the original V-N combination

Table 10. (cont'd)

Essay 2 - Self-correction	03:24-03:54	searched for V-N collocations with corpus tools	used the *collocates* function in COCA to search for the verb that collocates with the noun *idea* with the search pattern *idea* -> verb collocates -> established an inaccurate POS setting -> encountered a technical problem, possibly because of the wrong POS setting -> did not find the result she wanted
	03:56-05:15		looked for this pattern *behavior* -> verb collocates -> found the verb *observe* and recorded the MI score and frequency -> did not examine the concordance lines
	05:26-05:54		repeated the above-mentioned behavior with the search pattern *likelihood* -> verb collocates -> found the verb *predict* and recorded the MI score as well as frequency -> did not examine the concordance lines
	06:02-07:21		looked for the verb that collocates with the noun *personality* with the search pattern *personality* -> verb collocates in COCA -> scanned the verb list before choosing *develop* to replace the original verb *advocate.* However, the noun *personality* that she looked for was not the correct node, but *disorder* -> did not record the MI score and frequency or examine the concordance lines

Table 10. (cont'd)

07:24-08:12	used COCA to look for the verb that collocates with the noun *care* with the search pattern *care* -> verb collocates -> scanned the verb list and did not find the verb *desire* -> chose *foster* to replace the original verb, but the new collocation was not suitable in the context -> did not record the MI score and frequency or examine the concordance lines

In summary, Participant 0021 encountered many failed attempts of using COCA to search for V-N collocations because she did not establish the accurate POS setting. More importantly, she never examined the concordance lines for the searches she conducted. In addition, she frequently gave up her search when either COCA or SKELL failed to provide her with the results she needed. She also used her intuition during the process of self-correction.

Table 11 summarizes the main differences in the behaviors during the corpus consultation of these four participants.

Table 11.

Main differences in the behaviors of corpus consultation

Participants who improved (high MI pre - high MI post)		**Participants who did not improve (high MI pre - low MI post)**	
Participant oo7	**Participant oo14**	**Participant oo11**	**Participant oo21**

Table 11. (cont'd)

examined the concordance lines (in COCA) for almost every search of V-N collocations	often examined the concordance lines (in COCA) and examples (in SKELL)	rarely examined the concordance lines	never examined the concordance lines
frequently recorded MI scores and frequency of the V-N collocations found	varied her search patterns and flexibly switched between verb (node) -> noun (collocate) and noun (node) -> verb (collocate) for the desirable outcomes	often used her intuition during self-correction process	frequently failed to recorded MI scores and frequency
establish the accurate POS setting and sorting criteria	used both COCA and SKELL for double-checking or searching needed collocations when one of the tools did not work	sometimes ignored the feedback provided by her peers and the instructor	used her intuition during self-correction

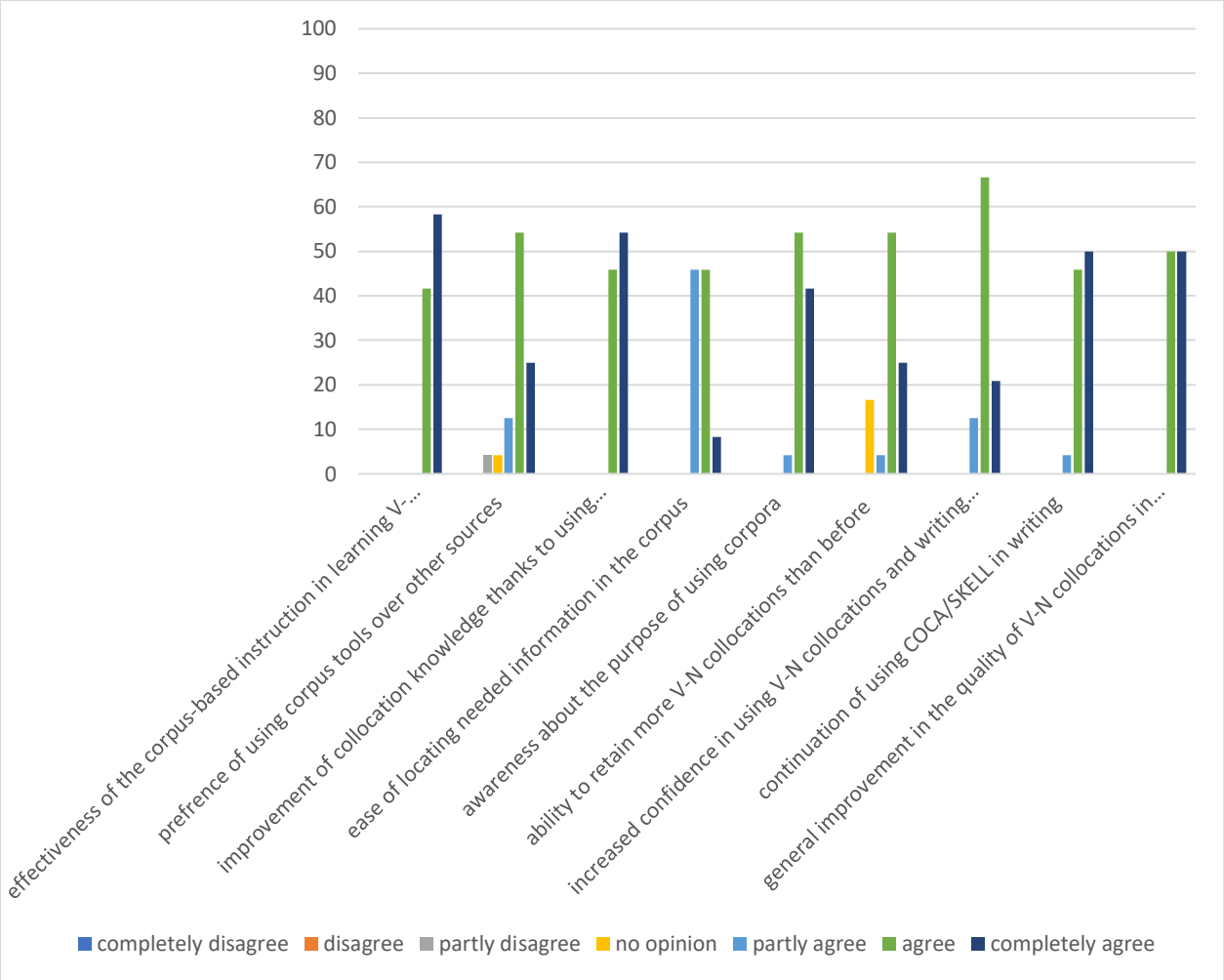
4.4 Students' perceptions of the hands-on consultation of the corpus tools (Evaluation survey)

Figure 3 describes the participants' overall evaluations on the utilization of the corpus tools in learning verb-noun collocations and writing. As can be seen from the chart, the majority of the participants showed a positive attitude toward their experience with the hands-on corpus-based instruction. Typically, 100% of the participants agreed and completely agreed that the corpus-based writing course has improved the quality of the V-N collocations used in their essays.

More than 95% expressed interest in continuing to use the introduced corpus tools for English writing purposes in the future. A mere 5% of participants conveyed some reluctance towards this prospect. More than 95% of the participants indicated the interest in the continuation of using COCA and SKELL in writing in English in the future, with only about 5% of them showing reluctance in doing that.

Figure 3.

Learners' attitude toward using corpora in learning V-N collocations and writing



Another notable finding from the survey is that all the participants agreed that the hands-on corpus-based instruction was significantly effective in improving their repertoire of verb-noun

collocations and that the utilization of the corpus tools helped increase their collocation knowledge. Although a minority of participants expressed no opinion or reluctance regarding the effectiveness of the hands-on corpus-based instruction, none conveyed a negative attitude towards the use of corpus tools or the corpus-based writing course.

4.5 Students' perception of the hands-on corpus-based training (Interview data)

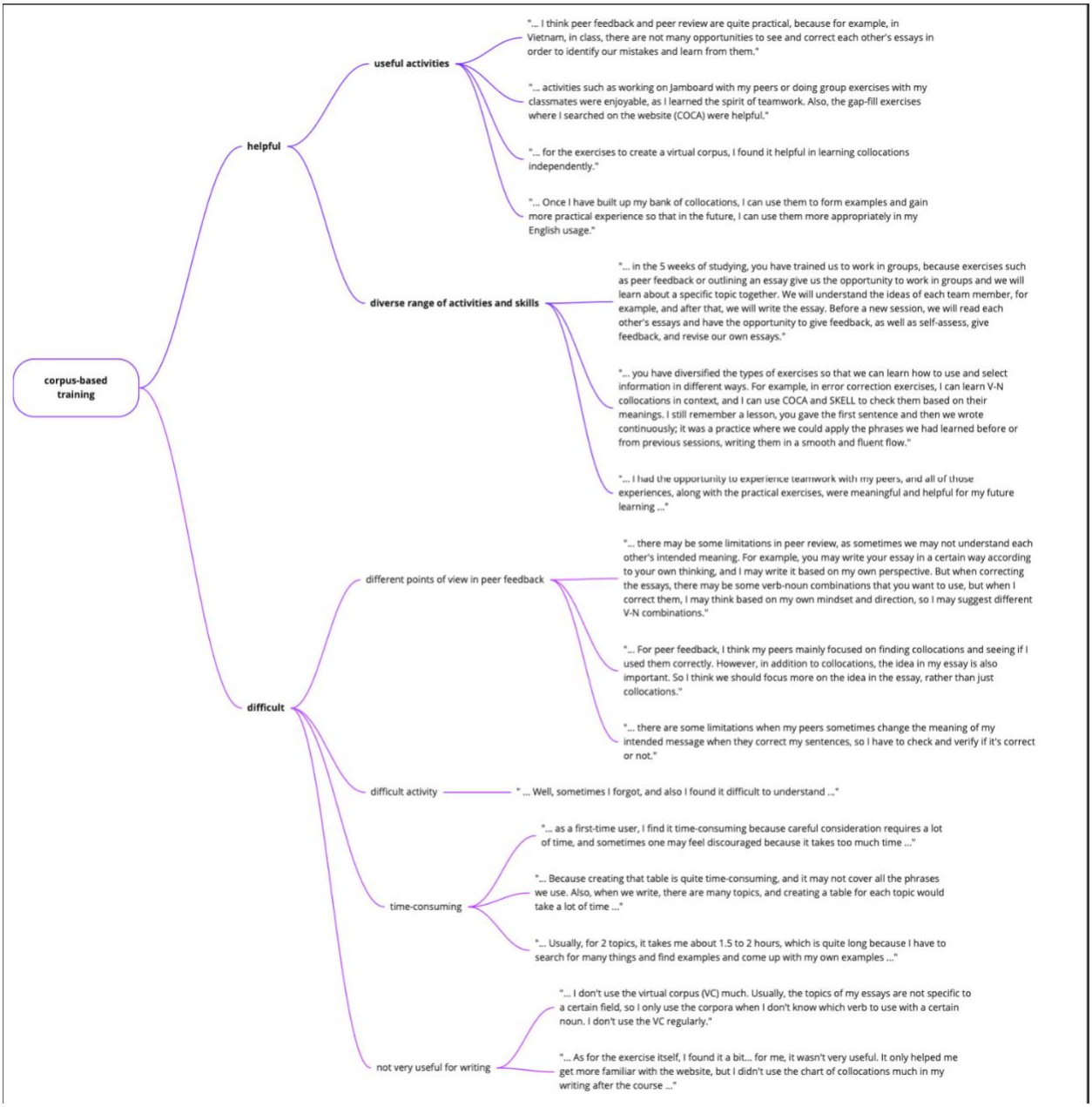
Three participants with the highest mean scores in the evaluation questionnaire, Participant 004 (6.89), Participant 0013 (6.78), and Participant 0022 (6.56) and the other three with the lowest mean scores in the evaluation questionnaire, Participant 0017 (5.67), Participant 0019 (5.67), and Participant 0025 (1), were chosen to participate in the semi-structured interview.

The interview data from Figure 4 gave insights into the participants' perceptions about the hands-on corpus-based course in terms of benefits and difficulties. Students appreciated the diversity of activities and additional skills that they learned from the experience. Participant 004 said in her interview, *"... I had the opportunity to* ***experience teamwork*** *with my peers, and all of those experiences, along with the* ***practical exercises****, were* ***meaningful*** *and* ***helpful*** *for my future learning ...*". This was supported by Participant 0022 when he said, *"... in my opinion, if we were to evaluate it on a scale of 10, I would rate it at 9.5. I will explain why I gave it a 9.5. I gave it a 9.5 because it helps me* ***improve my skills****, such as* ***researching****,* ***writing essays****, and* ***developing critical thinking***." Peer feedback, self-correction, and group brainstorming on Jamboard were highly valued by the participants. Participant 0022 commented, *"... I think* ***peer feedback*** *and peer review are quite* ***practical****, because for example, in Vietnam, in class, there are not many opportunities to see and correct each other's essays in order to* ***identify our mistakes*** *and* ***learn from them*****.**" He added, *"... the essence of these two activities is* ***crucial in my writing*** *and in my* ***acquisition of verb-noun collocations****, because as I mentioned before, I can* ***truly absorb those***

V-N collocations *when I have gone through and* ***received feedback on my mistakes***." Using the **virtual corpus** (VC) in COCA to create their own V-N collocation bank, however, showed a mixed picture. Participants 0013 was very positive about this activity, *"... once I have* ***built up my bank of collocations****, I can use them* ***to form examples*** *and gain more* ***practical experience*** *so that in the future, I can* ***use them more appropriately*** *in my English usage ...*". This was in line with Participant 0022, *"... I remember that I* ***highly value*** *the* ***virtual corpus*** *because it helps me with the* ***topic I want to study****. For example, if I want to learn vocabulary related to health, I have a* ***better vocabulary repertoire*** *to write an essay ...*". Participant 004 complemented, *"... for the exercises to* ***create a virtual corpus****, I found it* ***helpful*** *in* ***learning collocations independently***." However, participant 0019 had different opinion toward the VC activity. *"... I* ***don't use the virtual corpus (VC) much****. Usually, the topics of my essays* ***are not specific to a certain topic****, so I only use the corpora when I don't know which verb to use with a certain noun. I don't use the VC regularly*," he said. Participant 004 commented on another drawback of the VC activity, *"... creating the* ***VC does take a lot of time*** *... it takes me about 1.5 to 2 hours, which is* ***quite long*** *because I have to* ***search for many things*** *and* ***find examples*** *and* ***come up with my own examples*** *...*". This might also be the reason why participants 0025 and 0017 did not complete this activity during the course. However, Participant 004 also expressed that, *"... but* ***what I receive*** *afterwards, in general, it* ***pays off****, as I find many new words, learn many new things that I can* ***absorb and apply more for my future writing***."

Figure 4.

Students' perception of the hands-on corpus-based training



4.6 Students' perceptions of the corpus tools (Interview data)

The data from the semi-structured interview are presented in the form of mind-maps, which was inspired by Crosthwaite and Steeples (2022).

Figure 5 summarizes the interview data showing students' general positive attitudes toward the corpus tools, as indicated in the result of the evaluation questionnaire. Some participants expressed they felt fortunate to have access to the corpus tools. For example, participant 0013 said in her interview, "*... this **corpus is quite new**, especially for countries like Vietnam, so I haven't had much access to it. It's not widely used to help people access English learning resources, and I **feel very fortunate** to have the **opportunity to access it** and find new words to serve my future use of English.*" The participants mentioned a wide range of benefits that the corpus tools provided.

Apart from the main purposes of using the corpus tools to improve the accuracy of their collocation use and learn strong V-N collocations, some other outstanding advantages of using the corpus tools include enhancing confidence in using collocations, retaining collocations better, and learning more independently. Some notable quotes from the interview were presented as follows.

- **Enhancing confidence in using collocations**

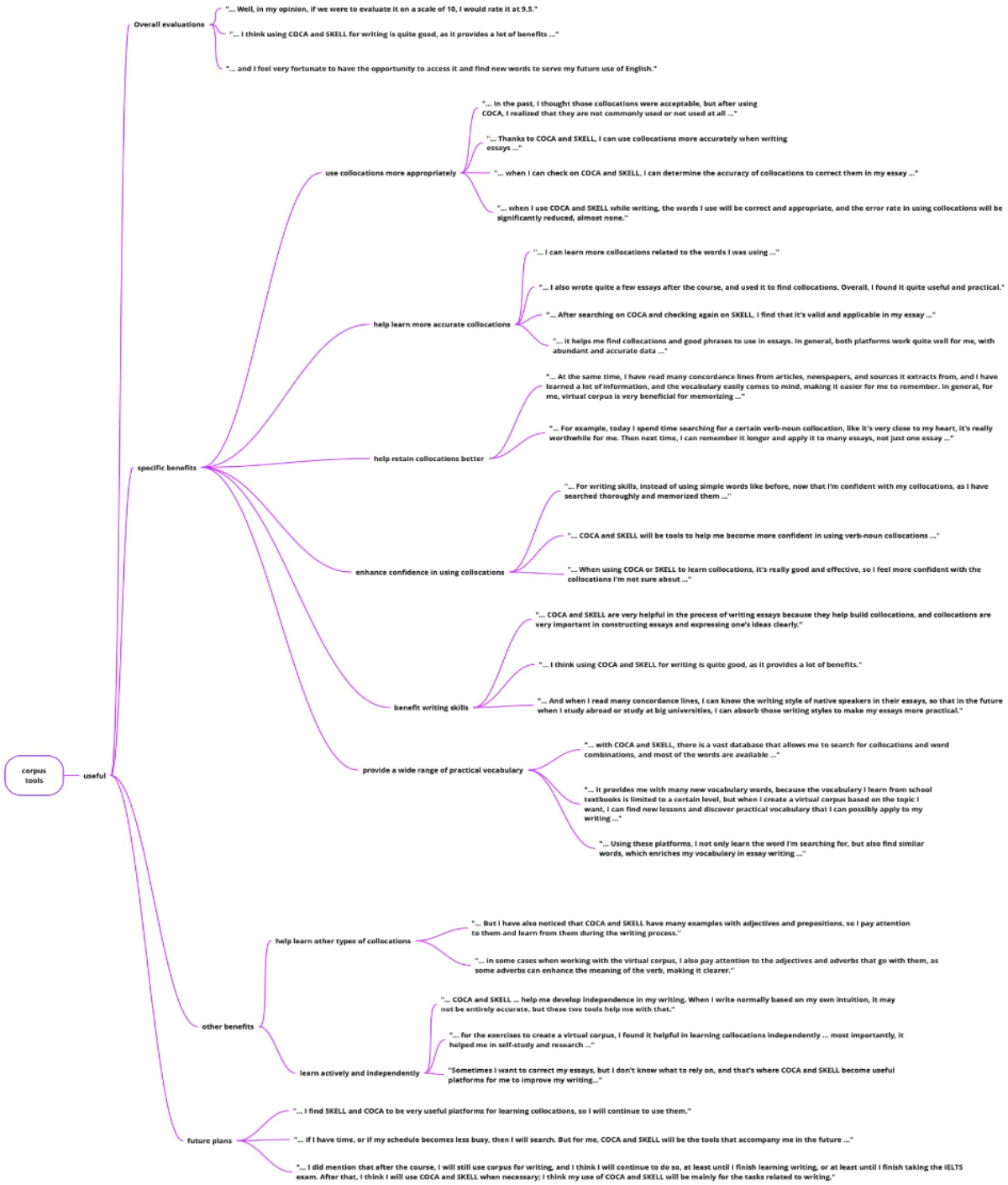
"… COCA and SKELL will be tools to help me **become more confident in using verb-noun collocations**…" (Participant 0022); "… for writing skills, instead of using simple words like before, now that **I'm confident with my collocations**, as I have searched thoroughly and memorized them…" (Participant 0013); "… when using COCA or SKELL to learn collocations, it's really good and effective, so **I feel more confident with the collocations** I'm not sure about …" (Participant 0025).

- **Retaining collocations better**

"…At the same time, I have read many concordance lines from articles, newspapers, and sources it extracts from, and I have learned a lot of information, and the **vocabulary easily comes to mind**, making it **easier for me to remember**." (Participant 0013); "… today I spend time searching for a certain verb-noun collocation, like it's very **close to my heart**,

Figure 5.

Students' perceptions of the corpus tools



it's really worthwhile for me. Then next time, I can **remember it longer** and apply it to many essays, not just one essay. So, although it may take a little time, I think it will be **beneficial for me in the long run**." (Participant 0022); "… when I was doing the virtual corpus, I considered words or phrases that I have encountered before or looked at the context in which they were used in essays and topics related to that context. In general, it is very practical, and I will choose those words to include in my virtual corpus to **make it easier to remember**." (Participant 004).

- **Learning more independently**

"… for the exercises to create a virtual corpus, I found it helpful in **learning collocations independently** … and most importantly, it **helped me in self-study** and research in the future." (Participant 004); "… COCA and SKELL are tools that **help me develop independence in my writing**. When I write normally based on my own intuition, it may not be entirely accurate, but these two tools help me with that." (Participant 0022); "… With the corpus, **I was able to correct some common mistakes** I made with certain verb-noun phrases, and it helped me change my writing style a bit." (Participant 0019).

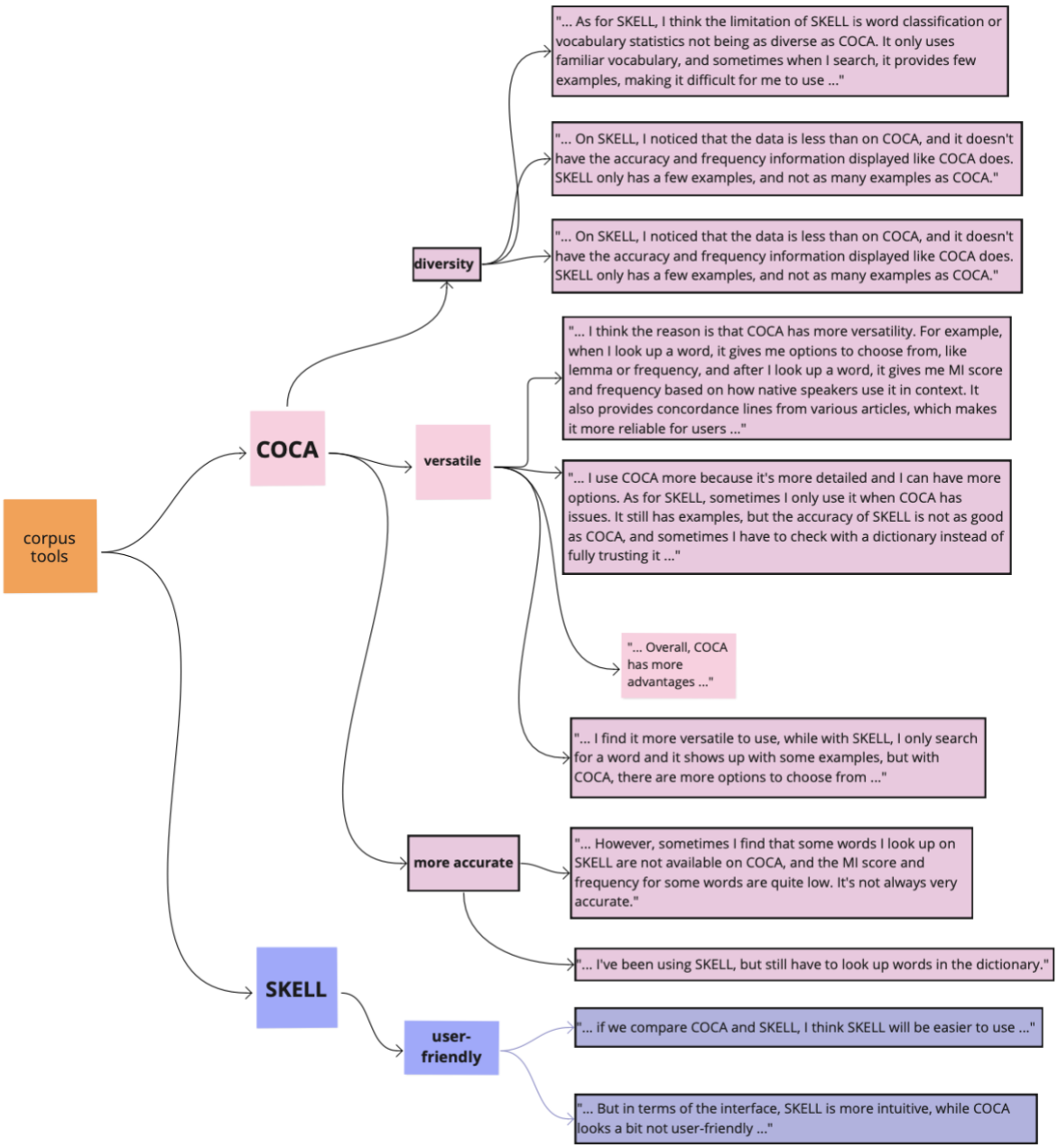
4.7 Students' preferences of the corpus tools (interview data)

The interview data also revealed the participants' preferences for using COCA or SKELL in their consultation (Figure 6). More participants indicated their preferences for COCA over SKELL because COCA provided them with more diverse data and information about V-N collocations. For example, Participant 0013 commented *"... **COCA has more versatility**. For example, when I look up a word, it **gives me options to choose from**, like lemma or frequency, and after I look up a word, it gives me **MI score and frequency** based on how L1 speakers use it in context. It also **provides concordance lines** from various articles, which makes it more reliable for*

users." Participant 004 corroborated this idea by saying "… *I find it* ***more versatile to use****, while with SKELL, I only search for a word and it shows up with some examples, but with COCA,* ***there***

Figure 6.

Students' preferences of the corpus tools



are more options to choose from." They also explained the reasons why they did not use SKELL as frequently as COCA. Participant 0022 shared *"... it* (SKELL) *doesn't provide me with as many skills as COCA because COCA allows me to research better, for example,* ***skills like data analysis****, such as analyzing based on MI scores, frequency, and concordance lines, so that I can find the best option for myself*." Participant 004 added *"... on SKELL, I noticed that the* ***data is less than on COCA****, and it doesn't have the accuracy and frequency information displayed like COCA does. SKELL only has a few examples, and not as many examples as COCA*." Giving a different explanation, Participant 0019 shared, *"... because most of the V-N collocations that I find on COCA, I think I don't need to use the other one. Only when I'm editing my writing and I can't find certain words using COCA, then* ***I use SKELL as a last resort****. So, I don't use SKELL much*."

On the other hand, SKELL has its own strengths. "… *if we compare COCA and SKELL, I think* ***SKELL will be easier to use****,"* Participant 0025 said. This was backed up by Participant 004 when she commented, "… *in terms of appearance,* ***SKELL is more intuitive****, while COCA looks a bit not user-friendly*."

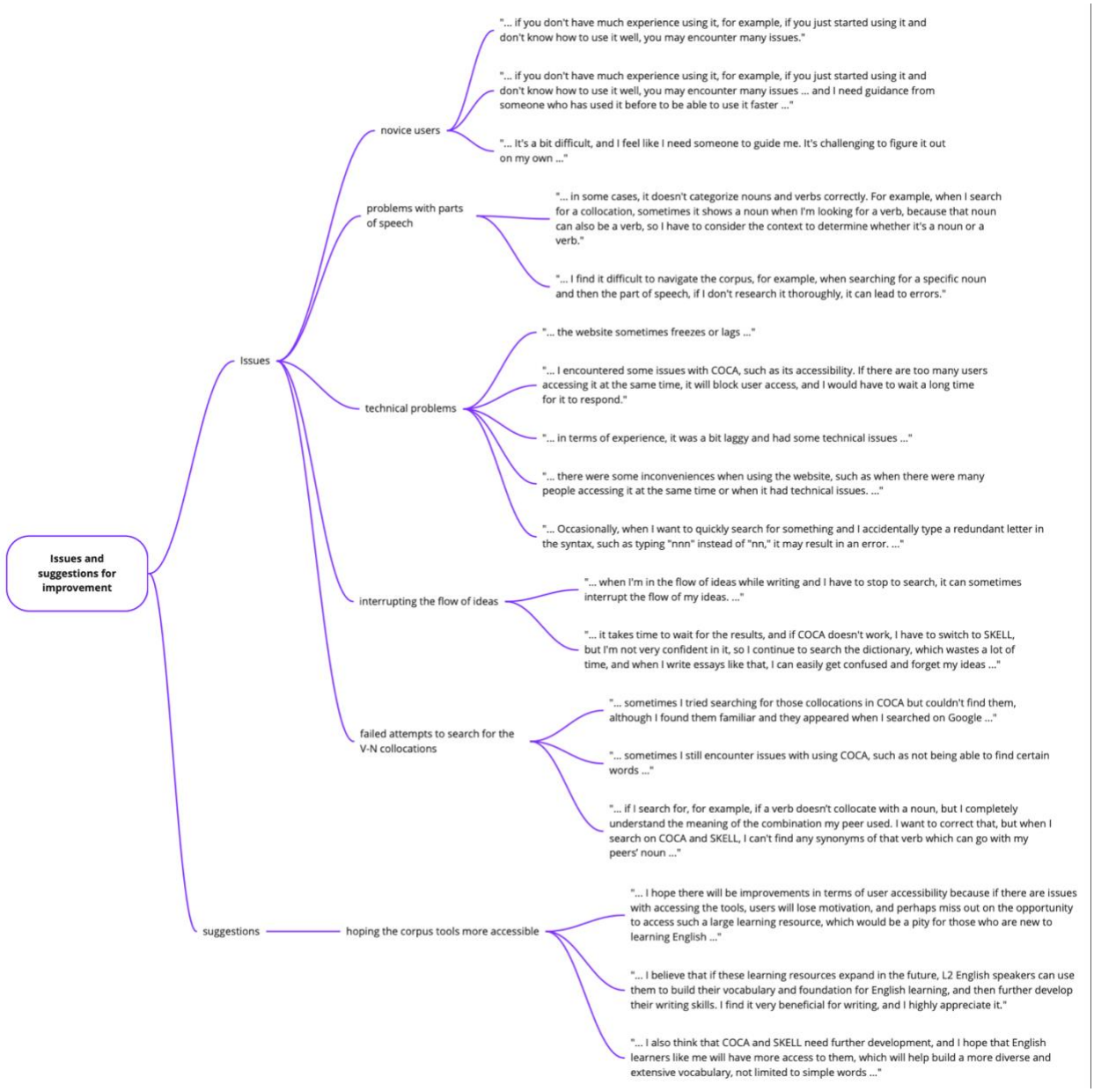
4.8 Challenges and suggestions for improvement (Interview data)

Figure 7 showed students' perceptions related to the issues with corpora and the hands-on corpus-based instruction as well as their desire for improvement. Most participants were concerned about the technical problems in COCA. *"... in my experience,* ***I encountered some issues with COCA, such as its accessibility****. If there are too many users accessing it at the same time, it will block user access, and I would have to* ***wait a long time for it to respond***," Participant 0013 said in her interview. Participant 0022 validated this opinion with his statement *"... sometimes due to network connection or when many people are using it,* ***there may be some glitches*** *in my search.*" Participant 0019 also expressed a similar thought "… *it may be due to my internet connection as I*

sit quite far from the Wi-Fi router, but I think there are many others who ***experience lagging issues*** *like me, although it's less frequent, but it still happens.*"

Figure 7.

Challenges and suggestions for improvement



Another typical issue was related to novice users. Participant 0017 raised this problem in her interview "… if you **don't have much experience** using it, for example, if you just started using it and don't know how to use it well, you may **encounter many issues**." This opinion was validated by Participant 0025. He said it is *"... **difficult to use if you're not familiar with it** (the corpus tools) ... and I **need guidance** from someone who has used it before to be able to use it faster*." This was consistent with what Participant 004 shared, *"... it's a bit **difficult**, and I feel like **I need someone to guide me**. It's challenging to figure it out on my own*."

The participants also expressed their desire to see improvements in the corpus tools. For example, Participant 0013 shared, *"... I **hope there will be improvements in terms of user accessibility** because if there are issues with accessing the tools, users will lose motivation, and perhaps miss out on the opportunity to access such a large learning resource, which would be a pity for those who are new to learning English ...*"

5 DISCUSSION

The study's findings revealed a nuanced understanding of the effects of corpus consultation on the association strength of V-N collocations used in the writing of advanced EFL students. While the results were mixed in terms of the students' improvement in their average MI scores and collocation error rates, the study's participants reported positive perceptions of the hands-on corpus-based training and corpus tools. These findings suggest the potential for integrating corpus tools into English classes in EFL contexts. However, careful consideration of issues related to corpus-based instruction and corpus tools is necessary to address any challenges that may arise. In the subsequent sections, each research question will be discussed in detail.

5.1 RQ1. To what extent does corpus consultation impact the association strength (MI scores) of V-N collocations used in the writing process of advanced EFL learners?

Comparing the average MI scores of V-N collocations in the pre-test and post-test, 13 students had an increase in their average MI scores, while 11 students had a decrease. Furthermore, the results indicated that those with increased MI scores had significant gains (more than 1) from the pre-test to the post-test (Participant 0025, 0014, 0020, 008, 0010, 0018, 0019) while those with decreased scores did not experience significant losses (less than 1, with Participant 009 being the only exception). This combined with the results from the collocation error rates sheds light into the effect of the corpus consultation in advanced EFL students' use of V-N collocations in writing. In general, 16 out of 24 students, accounting for 66.67%, either improved their MI average scores of the V-N collocations or reduced the percentage of V-N collocation errors in their essays between the pre-test and post-test. This partly suggests that the hands-on corpus consultation in the writing process does have a positive impact on the quality of collocations used in advanced EFL students' writing. However, taking into account the regression-to-the-mean phenomenon, which seems to

be the case in this study where most students improving started with low scores and most students whose scores decreased started with higher scores, a closer investigation into what these students did during the whole intervention process is necessary.

Regarding two participants who started with the lowest average MI scores, Participant 0014 who improved and Participant 0021 who did not improve, both had no prior experience using corpus tools, and they do not regularly pay attention to collocations used in sources outside the classroom when writing. However, Participant 0014 writes in English and has slightly more exposure to English outside the classroom than Participant 0021. Additionally, Participant 0014 was more interested in learning English collocations than Participant 0021. Despite some similarities, there are significant differences in their use of corpus consultation, which are evident from their profiles in the result section. First, Participant 0014 used both COCA and SKELL for double-checking or searching needed collocations when one of them did not work. In contrast, Participant 0021 frequently gave up her search when either COCA or SKELL failed to provide her with desirable outcomes. In other words, Participant 0021 rarely compared the results between COCA and SKELL while Participant 0014 did that many times during her corpus consultation. It is also apparent from the data that Participant 0021 encountered numerous failed attempts of using COCA to search for V-N collocations because she did not establish the accurate POS setting. Another significant difference was that Participant 0014 varied her search patterns and flexibly switched between verb (node) -> noun (collocate) and noun (node) -> verb (collocate). In the meantime, Participant 0021 only used one rigid search patter noun (node) -> verb (collocate) although it did not work in many cases. Regarding the use of SKELL, Participant 0014 utilized the *word sketch* function, which is directly designed for collocations, while Participant 0021 used *examples* function instead.

The other two participants, 007 and 0011, began the writing pre-test with the highest average MI scores (one improved and one did not improve, respectively). They both displayed a strong interest in learning English collocations despite having no prior experience using corpora before the course. Although 007 had slightly more exposure to English sources outside the classroom than 0011, both showed no significant application of the collocations used in these sources to their writing. Based on the data, it is evident that both participants primarily utilized COCA during their corpus consultation, as opposed to SKELL, and that they established correct POS setting as well as the sorting criteria by MI scores. However, there were significant differences in the way these two participants utilized the corpus tools at various stages of the intervention procedure.

The most obvious difference was the frequency of examining the concordance lines in COCA during the corpus consultation. While Participant 007 examined the concordance lines of the collocations he found in COCA most of the time as well as recording their MI scores and frequency, Participant 0011 rarely examined the concordance lines where the collocations are used in specific contexts. In addition, during the processes of self-correction, Participant 0011 often neglected the feedback provided by her peer and the instruction or used her intuition to correct her collocation errors instead of consulting the corpus tools to double-check her intuition, which did not happen in the data of Participant 007.

The qualitative data obtained from the screen recording videos has yielded significant insights into the participants' usage of the corpus tools. These insights could potentially explain the mixed picture observed in the average MI scores of the students between the pre-test and post-test. The participants who gained significantly in their average MI scores between the pre-test and post-test demonstrated the following behaviors during their corpus consultation:

- used both COCA and SKELL to double-check the results
- flexibly modify their search pattern to efficiently locate the relevant information
- examined the concordance lines (in COCA) carefully

On the other hand, the participants who experienced losses in their average MI scores between the pre-test and post-test indicated the following characteristics during the corpus consultation:

- used the functions that were not designed for searching collocations such as *word* function in COCA or *examples* function in SKELL
- failed to consult the other corpus tool when the search attempts in either COCA or SKELL were unsuccessful
- established inaccurate POS setting and inappropriate sorting criteria for collocation search (in COCA)
- failed to examine the concordance lines (in COCA) during the corpus consultation
- ignored the feedback provided by their peers or instructor
- used their intuition in self-correction instead of consulting the corpora

Therefore, the findings suggest that corpus consultation potentially has a positive impact on the association strength of V-N collocations that advanced EFL learners use in their writing. However, learner differences and their learning styles, including the way they employ the corpus tools, may influence the quality of their search and their proficiency in using this type of collocations in their essays. The students' attention to both the form and meaning of V-N collocations in concordance lines, which illustrate the usage of collocations in particular contexts, can play a crucial role in facilitating the acquisition of these collocations. This is consistent with what Jafarpour et al. (2013) concluded in their study about the benefits of L2 learners' attention to concordances in their collocation learning (p.57). This also aligned with what students shared in their interviews about

the benefits of using the corpus tools in learning V-N collocations and academic writing. Furthermore, the failure of certain participants to locate the required V-N collocations during corpus consultation due to inappropriate sorting criteria or incorrect POS settings highlights the significance of adhering to the instructor's guidelines when utilizing the corpora to avoid any potential errors.

5.2 RQ2. How do advanced EFL learners use hands-on corpus consultation in relation to the use of V-N collocation in writing?

The findings obtained through the screen recordings of students' corpus consultation at various stages of the writing process, in conjunction with interview data, indicated a higher frequency of using the Corpus of Contemporary American English (COCA) rather than Sketch Engine (SKELL) as a resource for learning verb-noun (V-N) collocations and enhancing their academic writing skills. This tendency may be attributed to several factors, including the sophisticated statistical information, such as MI score and frequency built in COCA, as well as its larger and more diverse corpus size, which offers a wider range of authentic examples of academic writing. An additional reason could lie in the versatility of the *collocates* function itself in COCA as opposed to the *word sketch* function in SKELL. The former allows learners to flexibly modify their search pattern from noun (node) – verb (collocate), or verb (node) – noun (collocate) to direct search (e.g. *insight (node) – give (collocates)*) as well as playing around with the window span. However, some other students still prefer SKELL for its user-friendly interface, which is unfortunately a drawback of COCA.

The findings also suggests that corpus consultation, especially COCA, occurred most frequently during the process of peer feedback when learners needed evidence such as MI scores and frequency of the collocations to provide feedback for their peers, whereas they tended to not

use the corpus tools (or other resources) while they were writing the essays (both in the post-test and two essays during the online training sessions). This can be interpreted based on the data obtained from the interviews when some participants raised their concern about how corpus consultation could take too much time or interrupt their flow of ideas. Another contributing factor to this behavior may be derived from technical errors. Many students reported errors during the online training sessions when they could not load the COCA website or had to wait a long time for COCA to respond after conducting a collocation search. This might have discouraged learners from using the corpus tools to consult their collocation use while they are writing their essays.

5.3 RQ3. What are advanced EFL learners' attitudes toward the hands-on use of corpus tools as a resource for learning V-N collocations in writing?

Regarding learners' perceptions of corpus consultation for learning verb-noun collocations in academic writing, the majority of participants in this study expressed positive attitudes towards corpus tools and the hands-on corpus-based course. The results from Fang et al. (2021) and Crosthwaite and Steeples (2022) confirmed this finding regarding learners' high appreciation for corpora and hands-on corpus-based instruction. Analysis of interview and video data indicated that learners showed a stronger preference for COCA over SKELL despite the former's non-intuitive interface. Crosthwaite and Steeples (2022) also found in their study that their participants showed a primary preference for user-friendly corpora such as Linggle but also appreciated the versatility of the corpus tools (p.15). In addition to the reasons discussed earlier, this preference may be attributed to learner-related factors. As advanced English-major EFL learners, these participants likely value a corpus that offers a comprehensive database of collocations, making COCA more appealing than SKELL in terms of the number of collocations and authentic examples. Moreover, interview data revealed that several participants considered research and critical thinking skills as

additional benefits of COCA consultation, suggesting that advanced learners tend to employ higher-level cognitive skills in their studies. This may be motivated by the importance of learning strategies such as problem-solving skills and independent learning to high school students to prepare for tertiary education (Crosthwaite, 2019, cited in Fang et al., 2021).

Although the findings from the evaluation survey painted a very positive picture about learners' attitude toward the hands-on corpus-based instruction, learners' difficulties in locating the needed collocations in the corpora cannot be overlooked. In other words, half of the learners think that searching for needed V-N collocations is not always easy and successful. This was supported by the findings from the interviews and screen recording videos. Fang et al. (2021) also reported this in their study. According to them, learners faced difficulties in finding the information they intended to use in their writing due to being overwhelmed with abundant concordance lines (p.95). In this study, this challenge could be caused by the novelty of the corpus tools when 50% of participants never heard of corpora and 70% of the participants did not have prior experience of using any corpus tools before the course. The non-intuitive interface of the corpus tools, especially the part of speech (POS) setting in COCA, apparently caused the participants a great deal of difficulty during their corpus consultation.

In addition to the corpus tools per se, the findings also offered insights into learners' perceptions on the learning activities that they experienced throughout the course. In general, students highly appreciated the interactions they had with one another through group work on Jamboard and the peer feedback activity. That opportunity of interaction coupled with the diversity of corpus-based training exercises, corpus consultation, and self-correction activity, according to them, helped them familiarize themselves with the corpus tools, realize their chronic collocation errors, learn more independently, and retain V-N collocations better.

Although students generally reported high levels of satisfaction with various learning activities, there were divergent views regarding the task of creating a personalized verb-noun (V-N) collocation bank using the virtual corpus in COCA. To account for these differences in learning experiences, it is necessary to consider individual learning styles. Within the limited scope of this study, participants expressed highly contrasting opinions about this activity, with some finding it time-efficient and valuable, while others perceived it as unhelpful.

5.4 Limitations

The study, however, has some limitations. First, the number of participants for such a mixed-method study is not insignificant but not large enough. The participants are all English-major advanced students who are strongly motivated in learning English and English collocations as they are aware of the importance of collocations in writing, so they may not be representative for the whole EFL population. These limitations coupled with the exploratory nature of this study limit the generalizability of the findings to a different population or context. In addition, the study just focused on one type of lexical collocation: verb-noun collocations. The intervention period may not have been long enough to allow for a more detailed examination of the participants' progress over time.

5.5 Future directions

Given the limitations of the current study, future studies can extend the duration of the intervention to gain deeper insights into the improvement of students in association strength of collocations used in their writing over time.

Within the scope of another study which is not constrained by time, location, and the number of participants, a quasi-experimental study can be conducted with the design of a pre-test,

a post-test, and a delayed post-test to ensure the ecological validity as well as investigate the long-term impacts of the intervention.

This study represents one of the few efforts to investigate the effectiveness of hands-on corpus-based instruction for pre-tertiary students. However, it has limitations in terms of its focus on a single level of proficiency (i.e., advanced EFL learners) and a specific type of collocations (i.e., V-N collocations). Future studies should explore the impacts of corpus consultation on the acquisition of other types of lexical collocations and among learners of different proficiency levels to provide a more comprehensive understanding of the effectiveness of corpus-based instruction for EFL learners. Additionally, investigating secondary teachers' perceptions of hands-on corpus-based instruction would be a valuable avenue for future research. This would enable a deeper understanding of teachers' attitudes and perceptions towards this type of instruction, which can inform the development of corpus-based teaching materials and teacher training programs.

Given the time constraints of this study, only a selection of representative participants' screen recording videos were analyzed. In future research, it would be beneficial to analyze all of the screen recording videos and compare these results with data collected from simulated interviews conducted immediately after participants consult corpus tools during their writing process. This approach would provide a more detailed understanding of the students' behavior and thinking processes during corpus consultation, shedding light on the underlying reasons for their actions.

5.6 Implications

5.6.1 Methodological implications

As mentioned in the method section, identifying collocation errors is not an easy task, which was echoed in Yoon (2016). The literature suggested that a combination should have a MI

score of 3 or above and a frequency of 10 or above in a reference corpus to be considered a collocation. Additionally, a combination with a negative MI score is deemed unacceptable. However, using this binary scale can lead to V-N combinations with negative MI scores, such as *understand subject, have insight,* or *hold party*, being considered collocation errors. However, L1 raters (in this study) considered *understand subject* and *have insight* strong V-N collocations and *hold* party as an acceptable V-N collocation, despite their negative MI scores. To address this dilemma, in this study, I employed a three-point Likert scale, *1 - awkward collocation, 2 - acceptable collocation, and 3 - strong collocation,* to code the collocation acceptability with two L1 speakers rating the V-N combinations, using their intuitions and me rating the same V-N combinations, using MI score and frequency information in COCA. After the rating process, I calculated inter-rater reliability (IRR) using Spearman's Rank correlation test and adjacent agreement to compare the two types of rating.

The inter-rater reliability (IRR) obtained from the Spearman's Rank correlation test between L1 intuition-based rating and COCA-based rating was moderate, indicating a considerable difference between these two types of rating. However, the IRR from the Adjacent Agreement was high with 0.93 (rater 1 and CBR) and 0.94 (rater 2 and CBR), suggesting that there was a high level of agreement between the L1 raters and corpus-based ratings on the acceptability of the V-N collocations. This suggests that although the L1 raters' intuitions may not completely align with the MI score and frequency information in COCA, two types of ratings were still consistent with each other to a large extent, demonstrating the reliability of using the combination of MI score and frequency information in COCA in the evaluation of collocation acceptability. In addition, the data from this study suggest that the binary scale used to determine collocation acceptability may not be appropriate, and a more nuanced approach, such as the three-point Likert

scale used in this study, could be employed. Further research is needed to determine the optimal methodology for identifying and evaluating collocations in different contexts and populations.

Regarding the rating using the information of MI score and frequency in COCA, the threshold for a strong collocation suggested in the collocation literature was adopted. Therefore, V-N combinations with MI score of three or higher and ten or more occurrences in COCA would be labeled 3 – *strong collocation.* However, no criteria for 1 - *awkward collocation* and 2 – *acceptable collocation* were found in the literature. Using the baseline of the *strong collocations*, I developed the criteria to identify awkward and acceptable based on the V-N combinations' MI scores and frequency as follows:

- 3 - strong collocation: MI > = 3 and Fre > 10
- 2 - acceptable collocation: 1 < = MI < 3 and Fre > = 10 or MI < 1 and Fre > 100
- 1 – awkward collocation: MI < 1 and/ or 10 < Fre < 100 or MI > = 3 and Fre < 10

This study collected data of participants' corpus consultation at different stages of the intervention through screen recording videos, which was not commonly implemented in collocation research. MAXQDA was used to code relevant segments of the videos which help explain what researchers cannot using only quantitative data.

Incorporating screen recording videos into the data collection process to capture participants' corpus consultation at various stages of the intervention has not been widely utilized in previous collocation research. By using MAXQDA to code relevant segments of the videos, this study was able to gain insights beyond what could be obtained through solely quantitative data analysis. As such, this methodological approach has important implications for future research in the field of collocation acquisition.

5.6.2 Pedagogical Implications

The findings of this study could bear implications for EFL teachers because firstly, it suggested that hands-on corpus-based instruction has a great potential in improving the quality of V-N collocation use in academic writing for advanced EFL learners and that it is feasible to be implemented with pre-tertiary students. This was in line with Fang et al. (2021) and Crosthwaite and Steeples (2022). The study provides EFL high school teachers with specific procedures and activities to implement corpus-based instruction in an academic writing class along with the corpus tools that they can use with their students. The insights into the way students consult the corpus tools during the writing processes, peer feedback, and self-correction can enable teachers and instructors to understand what challenges students may face during their corpus consultation and possible reasons for that. With such information, teachers can modify their lesson plans, training activities, and syllabus to facilitate their students' use of the corpus tools as well as the acquisition of collocational knowledge.

Many students in this study faced technical errors while using corpus tools for consultation, resulting from incorrect settings of word parts of speech and sorting criteria in COCA. Despite being provided with lesson interaction videos on how to use corpus tools in detail and being required to watch them before the course started, some students reported not having watched them during the semi-structured interview after the intervention. Therefore, synchronous lessons on corpus introduction and consultation are recommended, whether delivered in online or in-person classes, to ensure that all students are equipped with the necessary knowledge and skills for effective corpus consultation. This approach could reduce the risk of conducting inaccurate searches during consultation by avoiding situations where some students do not watch the lesson videos beforehand.

All five training sessions in this study were conducted online due to the geographical constraint, so when designing the activities, I tried to take advantage of technologies such as interactive videos, Jamboard, PowerPoint slides, or live worksheets to provide learners with as many opportunities to interact with each other and with the learning materials as possible. The diversity and interactive nature of the course received a positive appraisal from students, as evidenced by interview data. Accordingly, this study may serve as a useful reference for EFL instructors considering the integration of corpus consultation in their teaching.

However, online classes have certain limitations. One significant concern is that teachers are unable to observe their students while they are using corpus tools to provide timely assistance in case of technical errors. Therefore, it is recommended that in-person classes may help reduce the likelihood of students encountering technical issues such as incorrect POS settings in COCA, thereby enhancing their experience of corpus consultation. Future research could explore ways to mitigate this challenge in online classes, such as through real-time screen sharing or remote desktop access.

This study suggested that the variation in the impact of corpus consultation on the association strength of V-N collocations is dependent on how the participants utilized the corpus tools. Therefore, to ensure the effectiveness of corpus consultation in improving the quality of V-N collocations in academic writing, it is crucial that students are attentive to the MI score and frequency of collocations in reference corpora such as COCA, and meticulously examine the concordance lines.

6 CONCLUSION

In conclusion, this study investigated the impact of integrating corpus consultation in teaching V-N collocations in academic writing for secondary advanced Vietnamese learners. The results suggest that the use of corpus consultation can significantly improve the association strength of V-N collocations used in learners' essays and reduce the V-N collocation error rates, depending on how they utilized the corpus tools in different stages of the writing process: while-writing, peer feedback, and self-correction/ revision. Additionally, learners' perceptions of the usefulness and effectiveness of corpus consultation as well as the corpus-based lessons were largely positive. However, technical issues with the corpus tools and individual differences in the way learners utilized the tools highlighted the need for careful attention to training and support for learners. Overall, this study provides deeper insights into the potential benefits and challenges of integrating corpus consultation into EFL pre-tertiary educational contexts and offers suggestions for further studies, research methodology, and pedagogical implications.

REFERENCES

Akıncı, M., & Yıldız, S. (2017). Effectiveness of corpus consultation in teaching Verb+Noun collocations to advanced ELT students. *Eurasian Journal of Applied Linguistics*, *3*(1), 93-111. Retrieved February 12, 2023, from https://dergipark.org.tr/en/download/article-file/537431

Assessment in UNCW. University of North Carolina Wilmington. Retrieved April 10, 2023, from https://uncw.edu/assessment/general/interraternotes.html

Benson, M. (1986). *The BBI Combinatory Dictionary of English.* Amsterdam: John Benjamins.

Cao, D., & Badger, R. (2021). Cross-linguistic influence on the use of L2 collocations: The case of Vietnamese learners. *Applied Linguistics Review*. https://doi.org/10.1515/applirev-2020-0035

Chan, T., & Liou, H. C. (2005). Effects of Web-based Concordancing Instruction on EFL Students' Learning of Verb – Noun Collocations. *Computer Assisted Language Learning*, *18*(3), 231–251. https://doi.org/10.1080/09588220500185769

Conway, J. (2011). A Corpus-based Approach to Vocabulary Instruction in the CBI Classroom. Northern Arizona University.

Crossley, S. A., Salsbury, T., & Mcnamara, D. S. (2014). Assessing Lexical Proficiency Using Analytic Ratings: A Case for Collocation Accuracy. *Applied Linguistics*. https://doi.org/10.1093/applin/amt056

Crosthwaite, P., & Steeples, B. (2022). Data-driven learning with younger learners: exploring corpus-assisted development of the passive voice for science writing with female secondary school students. *Computer Assisted Language Learning*, 1–32. https://doi.org/10.1080/09588221.2022.2068615

Davies, M. (2008-) The Corpus of Contemporary American English (COCA). Available online at https://www.english-corpora.org/coca/.

Duong, D. T. H., & Nguyen, N. D. T. (2021). Using Collocations to Enhance Academic Writing: A Survey Study at Van Lang University*:* 17th International Conference of the Asia Association of Computer-Assisted Language Learning (AsiaCALL 2021), Ho Chi Minh, Vietnam. https://doi.org/10.2991/assehr.k.210226.035

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL - International Review of Applied Linguistics in Language Teaching*, *47*(2). https://doi.org/10.1515/iral.2009.007

Fang, L., Ma, Q., & Yan, J. (2021). The effectiveness of corpus-based training on collocation use in L2 writing for Chinese senior secondary school students. *Journal of China Computer-Assisted Language Learning*, *1*(1), 80–109. https://doi.org/10.1515/jccall-2021-2004

Firth, J. R. (1957). Modes of meaning. In J.R. Firth (Ed.), Papers in linguistics 1934-1951 (pp.190-215). Oxford: Oxford University Press.

Gao, B., Zhu, L., Zhu, L., Luo, Z., & Luo, Z. (2019). A Corpus-based Study of Collocation Production of Chinese EFL Learners. *International Journal of Education*, *11*(4), 66. https://doi.org/10.5296/ije.v11i4.15895

Goulart, L. (2019). The use of collocations across proficiency levels: A literature review. *BELT - Brazilian English Language Teaching Journal*, *10*(2), 34129. https://doi.org/10.15448/2178-3640.2019.2.34129

Green, C. (2020). Extensive reading and viewing as input for academic vocabulary: A large-scale vocabulary profile coverage study of students' reading and writing across multiple secondary school subjects. *Lingua*, *239*, 102838. https://doi.org/10.1016/j.lingua.2020.102838

Hill, J., & Lewis, M. (1997). *LTP Dictionary of selected collocations*. Language Teaching Publication.

Hoang, T. M. (2018). Investigating Vietnamese ESL learners' understandings of verb-noun collocations associated with academic verbs. https://hs.kku.ac.th/ichuso/2018/ICHUSO-122.pdf

Jafarpour, A. A., Hashemian, M., & Alipour, S. (2013). A Corpus-based Approach toward Teaching Collocation of Synonyms. *Theory and Practice in Language Studies*, *3*(1), 51–60. https://doi.org/10.4304/tpls.3.1.51-60

Kostere, S., & Kostere, K. (2022). *The Generic Qualitative Approach to a Dissertation in the Social Sciences: A step by Step Guide.* Routledge.

Koya, T. (2006). What is the reality of collocation use by native speakers of English? *Dialogue*, *5*, 1- 18.

Laufer, B., & Waldman, T. (2011). Verb-Noun Collocations in Second Language Writing: A Corpus Analysis of Learners' English: Verb-Noun Collocations in L2 Writing. *Language Learning*, *61*(2), 647–672. https://doi.org/10.1111/j.1467-9922.2010.00621.x

Liou, H. C. (2019). Learner Concordancing for EFL College Writing Accuracy. *English Teaching & Learning*, *43*(2), 165–188. https://doi.org/10.1007/s42321-019-00022-5

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

Nesselhauf, N. (2003). The Use of Collocations by Advanced Learners of English and Some Implications for Teaching. *Applied Linguistics*, *24*(2), 223–242. https://doi.org/10.1093/applin/24.2.223

Nesselhauf, N. (2004). Collocations in a Learner Corpus. Amsterdam: John Benjamins.

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, *21*(3), 298–320. https://doi.org/10.1177/1362168816639619

Park, K., & Kinginger, C. (2010). Writing/thinking in real time: Digital video and corpus query analysis. *Language Learning and Technology*, *14*(3), 31–50.

Staples, S., & Dilger, B. (2018-). Corpus and repository of writing [Learner corpus articulated with repository]. Available at https://crow.corporaproject.org

Tsai, K. J. (2015). Profiling the collocation use in ELT textbooks and learner writing. *Language Teaching Research*, *19*(6), 723–740. https://doi.org/10.1177/1362168814559801

Vu, D. V., & Peters, E. (2021). Vocabulary in English Language Learning, Teaching, and Testing in Vietnam: A Review. *Education Sciences*, *11*(9), 563. https://doi.org/10.3390/educsci11090563

Wei, Y. (1999). Teaching collocations for productive vocabulary development. *Paper Presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages*, 1–31. http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED457690

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition*, *35*(3), 451–482. https://doi.org/10.1017/S0272263113000107

Yoon, H.-J. (2016). Association strength of verb-noun combinations in experienced NS and less experienced NNS writing: Longitudinal and cross-sectional findings. *Journal of Second Language Writing*, *34*, 42–57. https://doi.org/10.1016/j.jslw.2016.11.001

APPENDIX A: BACKGROUND QUESTIONNAIRE

1. What is your gender?
    a) Male
    b) Female
    c) Non-binary/ Third gender
    d) Prefer not to say
2. In general, do you like to try new things?
    a) Definitely not
    b) Probably not
    c) Might or might not
    d) Probably yes
    e) Definitely yes
3. How often do you write in English?
    a. Once a week
    b. Three times a month
    c. Twice a month
    d. Once a month
    e. Less than once a month
4. What resources do you use as reference when you write in English? (e.g. online dictionaries, paper dictionaries, ...)

    _____

5. How familiar are you with English collocations? (Collocations are the words that tend to co-occur with each other. For example: make mistakes, take responsibility)
    a) Not familiar at all
    b) Slightly familiar
    c) Moderately familiar
    d) Very familiar
    e) Extremely familiar
6. Are you interested in learning English collocations?
    a) Definitely not
    b) Probably not
    c) Might or might not
    d) Probably yes
    e) Definitely yes
7. If the answer to Question 6 is "Probably not" or "Definitely not", why are you not interested in learning English collocations?

    _____

8. Have you heard about English corpus/ corpora before?
    a) No
    b) Yes
9. If the answer to Question 8 is Yes, have you used an English corpus before? (e.g. BNC, COCA)
    a) No
    b) Yes
10. If the answer to Question 9 is *Yes*, which English corpus did you use before?

_____

11. How often do you get exposure to English outside classroom? (not including English extra classes)
    a) Less than one hour a day
    b) One hour a day
    c) Two hours a day
    d) Three hours a day
    e) More than three hours a day

12. What sources of English do you often get exposure to? (E.g. movies, TV, news, books, magazines, etc.)

_____

13. Do you take notice of the collocations used in those sources? (frequency)
    a) Extremely unlikely
    b) Somewhat unlikely
    c) Neither likely nor unlikely
    d) Somewhat likely
    e) Extremely likely

14. Do you apply the collocations you noticed from those sources to your writing? (frequency)
    a) Extremely unlikely
    b) Somewhat unlikely
    c) Neither likely nor unlikely
    d) Somewhat likely
    e) Extremely likely

(Adapted from Fang et al. (2021))

APPENDIX B: EVALUATION QUESTIONNAIRE

The following questionnaire is asking for your opinions on using COCA and SKELL as reference tools in your writing process. Please use the scale below to circle the response that most represents your perspectives.

1 - Strongly disagree
2 - Disagree
3 - Somewhat disagree
4 – Neither agree nor disagree
5 - Somewhat agree
6 - Agree
7 - Strongly agree

The five corpus training sessions were helpful in developing my collocation repertoire.	1	2	3	4	5	6	7
I prefer corpus tools as a resource to search and learn collocations when I write an English essay to online collocation dictionaries.	1	2	3	4	5	6	7
Using corpus tools improved my knowledge of collocations.	1	2	3	4	5	6	7
When I search for information in a corpus, I usually get the information that I need.	1	2	3	4	5	6	7
I understand the purpose of using the corpus after the five training sessions	1	2	3	4	5	6	7
I can remember more collocations after working with corpus than before the corpus training.	1	2	3	4	5	6	7
My confidence in using verb-noun collocations and in writing in general has increased by learning about the corpus.	1	2	3	4	5	6	7
I will continue to use COCA and/or SKELL in English academic writing.	1	2	3	4	5	6	7
Overall, the corpus training has improved the quality of verb-noun collocation use in my English essays.	1	2	3	4	5	6	7

(Adapted from Fang et al. (2021))

APPENDIX C: THE TEST OF COLLOCATIONAL KNOWLEDGE

Instruction: Complete the following sentences by filling the blanks with suitable verbs. (You can give multiple answers)

1. Single parents _____ children without a partner's support are entitled to financial help from the government.
2. I've never been very successful at _____ plants.
3. Jack has already _____ a very good reputation as a talented lawyer.
4. He has to stay at home and take care of his wife. She is _____ a baby.
5. Can you _____ the difference between these two pictures?
6. Facing a problem, many decide to _____ suicide.
7. They _____ their first public speech about the abortion in front of hundreds of people in near-freezing temperatures.
8. He _____ a bad accident when he fell asleep while driving.
9. I wish you would stop _____ such a mess in the kitchen.
10. Excuse me, would you mind _____ a photo of me and my friend?
11. You should always check your work carefully in case you have _____ some mistakes.
12. _____ a decision on whether to go or not to go is difficult for him.
13. Nobody in my class has _____ a lesson so far this year.
14. I've been studying French for 6 months now, but I don't feel like I'm _____ progress.
15. She is striving to _____ her ambition to be a fashion designer.
16. The government is keen to _____ consumer confidence and spending.
17. Can you _____ some music, please?
18. He claimed the government has _____ the trust of British people.
19. These tablets should _____ his suffering for a couple of hours.
20. I _____ issue with some of the points made in the speech.

(Adapted from Jafarpour et al. (2013))

APPENDIX D: WRITING PROMPTS

Writing pre-test question

Many people can now study at home through distance-learning programs via Internet or television. But some people say this kind of education cannot be compared to attending a college. To what extent do you agree or disagree?
Provide relevant examples if necessary.
(IELTS Writing Task 2 in Vietnam - 10/23/2008)

Writing post-test question

Many people can now study at home through distance-learning programs via Internet or television. But some people say this kind of education cannot be compared to attending a college. To what extent do you agree or disagree?
Provide relevant examples if necessary.
(IELTS Writing Task 2 in Vietnam - 10/23/2008)

APPENDIX E: THE COURSE SYLLABUS

Essay Writing Course with Corpus Tools

Instructor Information:
Duong Nguyen
Email: nguye863@msu.edu

Course Description:

This Essay Writing Course is designed to provide advanced low high school students with fundamental knowledge and skills to utilize corpus tools such as the Corpus of Contemporary American English (COCA) and SKELL to improve their use of collocations in their essays, to expand their collocation repertoire, and to self-correct the lexical errors in their essays. The course will be 100% online. All materials and information about the course can be found on Canvas. The course will meet one day a week for 90 minutes each class. There are two asynchronous lessons which will be recorded by the instructor and uploaded on Canvas. The primary aim of this course is to raise students' awareness of good English collocations and support them to write more naturally in English with the help of the corpus tools. The course also focuses on developing learner autonomy through discovery learning and collaborative learning.

Course Objectives:

At the end of this course, students will be able to:

- use corpus tools independently to improve their collocation use in their essays
- use corpus tools to self-check and self-correct the use of collocations and lexical items in their own essays
- use corpus tools to enrich their collocation repertoire
- write more naturally in English

Course Structure:

The course includes two asynchronous lessons and five online training lessons. All materials will be stored and accessed through the Learning Management System (LMS) Canvas. All assignments will be submitted to the Google Drive Folder. The course will be divided into seven modules, and there are three tests: the collocation pre-test, the writing pre-test, and the writing post-test.

Module 1: Introduction and Technology Help
Module 2: Essay Review and Corpus Introduction
Module 3: Collocation practice 1
Module 4: Essay 1
Module 5: Collocation practice 2
Module 6: Essay 2
Module 7: Collocation practice 3

Materials:

Supplemental materials, including readings and videos, will be provided by the instructor and will be available online through the Canvas platform.

Student Responsibilities / Expectations:

Students are required to attend all classes, participate in class discussions, and complete all assignments before deadlines. Deadline extension should be requested and granted before the due day.

Essay Writing Course: Assignment List

Week 0 - Module 1 & 2 Topics: • Course Introduction and Technology Help • Essay Review and	**<u>Assignments:</u>** • Watching the videos • Answering the pop-up questions while watching the videos/ doing the quiz after watching the videos • Homework: Making essay outline **Due day: by Sunday, Nov. 13**
Week 1 - Module 3 (Nov. 13) Topics: Collocation Practice 1	**<u>Assignments:</u>** • Upload the recording of using COCA/SKELL to do the collocation practice onto the class Google Drive Folder • Create the virtual corpus with COCA (Topic: Technology & Environment) • Submit the link to the Google Docs/ Notions where you did your homework to Canvas • Upload the recording of creating the **virtual corpus** and using the corpus to search for collocations onto the class Google Drive Folder **Due day: by Saturday, Nov. 19**
Nov. 20	Students are off (Vietnamese Teachers' Day) Remind students to review Week 0 and Week 1 in case they forget what they learned.
Week 2 - Module 4 (Nov. 27) Topics: Essay 1 - Submit your first draft of the essay	**<u>Assignments:</u>** • Upload the recording of using COCA/SKELL as reference when you write the essay and give feedback to your peers • Upload the recording of using COCA/SKELL to self-correct the collocation errors marked by the instructor. • Update your first draft of essay

	Due day: by Saturday, Nov. 26
Week 3 - Module 5 (Dec. 4) Topics: Collocation Practice 2	**<u>Assignments:</u>** • Upload the recording of using COCA/SKELL to do the collocation practice onto the class Google Drive Folder • Create the virtual corpus with COCA (Topic: Crime & Health) • Submit the link to the Google Docs/ Notions where you did your homework to Canvas • Upload the recording of creating the **virtual corpus** and using the corpus to search for collocations onto the class Google Drive Folder **Due day: by Saturday, Dec. 3**
Week 4 - Module 6 (Dec. 11) Topics: Essay 1 - Submit your first draft of the essay	**<u>Assignments:</u>** • Upload the recording of using COCA/SKELL as reference when you write the essay and give feedback to your peers • Upload the recording of using COCA/SKELL to self-correct the collocation errors marked by the instructor. • Update your first draft of essay **Due day: by Saturday, Dec. 10**
Week 5 - Module 7 (Dec. 18) Topics: Collocation Practice 3	**<u>Assignments:</u>** • Upload the recording of using COCA/SKELL to do the collocation practice onto the class Google Drive Folder • Create the virtual corpus with COCA (Topic: Education & Culture) • Submit the link to the Google Docs/ Notions where you did your homework to Canvas • Upload the recording of creating the **virtual corpus** and using the corpus to search for collocations onto the class Google Drive Folder **Due day: by Saturday, Dec. 17**

APPENDIX F: LESSON PLANS

Asynchronous Lesson Plan: Introduction to Corpora (COCA and SKELL)

Target Audience and Context: 11th grade students

Learner Level: Intermediate High/ Advanced Low

Class size: ~30

Class length: 90 mins

Topic: Essay Review

Objectives:

After the lesson, SWBAT:

- define collocations and know some statistical information related to collocations
- use some basic functions in COCA and SKELL to conduct lexical searches
- use collocates function in COCA and word sketch in SKELL to search needed collocations and expand their repertoire of collocations

Materials and resources:

- Teacher will record a video of the lesson and turn it into an interactive video in which some quizzes will be integrated to check students' understanding of what they watch (using PlayPosit)
- Slides: https://tinyurl.com/27v245th

Time & Materials	**Procedures**
10 mins	Warm-up: • Students look at some examples of collocations. • Students think of the definition of collocations. "What do you think a collocation means?" • Students look up the definition of collocations in dictionary and compare it with their own definitions • Teacher shares the definition of collocations from the perspective of corpus linguistics

10 mins	Activity 1: Definition of collocations • Teacher gives an example of collocation (E.g. *make mistakes*, not *do mistakes*) • Teacher introduces the *node* and the *collocate*
35 mins	Activity 2: Introduction of COCA • Teacher opens COCA and walk students through all basic functions on the interface of COCA (List, Chart, Word, Collocates, Compare) • Teacher focuses on *collocate* function, showing students how to search a collocation with a given noun or a verb • Teacher explains the results appearing on the screen, especially the statistics: frequency and MI score (Mutual Information score) and their importance • Teacher shows students how to see the collocation is used in context by examining concordance lines
30 mins	Activity 3: Introduction of virtual corpus • Teacher shows students how to create a virtual corpus • Teacher shows students how to choose the key nouns from the virtual corpus created and find the collocations with those nouns • Teacher shows students a way to keep records of the collocations they found from the corpus in an organized way, using Notion/ Google Docs
5 mins	Activity 4: Introduce SKELL • Teacher shows students how to use SKELL to double check collocations with the results in COCA.

Asynchronous Lesson Plan: Review of Essays

Target Audience and Context: 11th grade students

Learner Level: High-intermediate/ Low-advanced

Class size: 30

Class length: 90 mins

Topic: Essay Review
Objectives:
After the lesson, SWBAT:

- identify an appropriate structure of an essay
- differentiate between different types of essays
- identify typical features of each type of essay

Materials and resources:

- Teacher will record a video of the lesson and turn it into an interactive video in which some quizzes will be integrated to check students' understanding of what they watch (using PlayPosit)
- Handouts: https://tinyurl.com/3jphdtf2
- Slides: https://tinyurl.com/3pujutxn

Time & Materials	**Procedures**

25 mins	**Activity 1:** Model essay deconstruction
	• Ss work in groups of 4 and deconstruct the model essay.
	  • How many parts can you identify in the essay?
	  • What is the purpose of each part?
	  • Look more specifically at each paragraph and answer the following questions:
	  o What is the focus (main point) of each paragraph?
	  o How do the paragraphs relate to each other? (E.g. ideas, linking devices, etc.)
	  o How do the sentences in a paragraph relate to each other?

25 mins	**Activity 2:** Analyze the question
	• Underline keywords in the questions
	• Pay attention to the question requirement
	  • Fact
	  • Opinion
	  • Task
	  • Topic
	  • Restrictions
	• Match the essay types with the appropriate essays

40 mins	**Activity 3:** Model essay analysis - Text annotation (Google Docs)
	• Identify typical language used in each type of essay
	  • Modality
	  • Hedges and boosters
	  • Cause-and-effect relationships (e.g. caused)
	  • Relationship of addition (e.g. in addition)
	• Pay attention to specific structure of each essay
	  Strong answer
	  Balanced
	  Reasons - solutions (block/ point-by-point)

	Homework: Make outlines for 4 essays in Activity 3

Corpus-based Lesson Plan #1

Target Audience and Context: 11th grade students

Learner Level: High-intermediate/ Low-advanced

Class size: ~30

Class length: 90 mins

Topic: Verb-Noun collocations (Part 1)
Objectives:
After the lesson, SWBAT:

- use COCA and/ or SKELL to identify strong verb-noun collocations in a sentence
- use COCA and/or SKELL to identify verb-noun collocation errors and correct them
- develop a sense of collocation use when constructing a sentence
- develop an individual collection of verb-noun collocations in topic "Technology" and "Environment"

Materials and resources:

- Slides: https://tinyurl.com/3pujutxn
- Handout: https://tinyurl.com/2unuf3tc

Time & Materials	**Procedures**
15 mins	**Warm-up:** • Which verbs do you think can go with the noun "information"? • Use COCA to check your answers. • Open COCA, choose the collocate function -> Type 'information' in the Word/Phrase box (choose the POS as _nn). Choose the POS in the *collocate* position as Verb. Click *Find collocates.* • What is the verb with the highest frequency and MI score? • Open the concordance lines and examine the pattern of the verb-noun collocation identified. What do you see?
20 mins	**Activity 1: Sentence completion** Choose appropriate verbs to complete the following sentences. Use the collocate function feature and concordance lines in COCA to check whether the chosen verbs **strongly** collocate with the nouns that follow.

	1. In a preseason drill last fall, Vegas Golden Knights goalie Marc-Andre Fleury, who had joined the expansion team from the Pittsburgh Penguins, a perennial power, _____ his anger after allowing a series of goals. (vented/ released) 2. But since assuming the general manager role in 2011, he said he has _____ his anger into "positive energy", building a formidable team capable of toppling countries who recruit from populations several thousand times larger than the roughly 70,000 strong Iroquois Nation. (channeled/ fueled) 3. Additionally, our energy audit can reveal where simple enhancements and maintenance activities can _____ dividends to the family budget and the environment. (pay/ give) 4. According to Palmer, who _____ the committee, there is little doubt that the group will call for full participation by lawyers in the mediation process, including serving as neutrals, and, on behalf of parties, attending sessions and serving as coaches. (chairs/ heads) 5. Their terms of usage should _____ regulations that preserve the privacy of non-offenders as well as comply with legislative safeguards for security. (conform to/ ensure)
20 mins	**Activity 2: Error identification** These sentences are extracted from students' paragraphs. Identify verb-noun combinations in the following sentences and then use COCA as reference to decide which V-N combination is not a good collocation. 1. To cope with such prevalent health problems as obesity, those who are and are not incurring this situation could also implement the following methods. 2. Lastly, training willpower to be on a diet is also essential. 3. If people continue to keep this harmful lifestyle, they will fall ill more frequently and lose their health to alarm level. 4. These are a few methods that are helpful in decreasing obesity. 5. However, there are also various proposed ways to settle this growing problem.

35 mins	**Activity 3: Sentence building** Imagine you are writing an essay about the Advantages and Disadvantages of Automobiles on Society and the World. Continue writing to develop the following points (3-5 sentences). Identify verb-noun collocations you may use in your sentences. Use the collocate function in COCA to check whether your use of verb-noun collocations is correct. 1. Many people desire to own cars because this considerably enhances their standards of living. 2. Notwithstanding cars are convenient, the damage they do to the environment is a problem that must be solved. *(Source: Crow Corpus)*
10 mins	Homework: • Create an academic virtual corpus about the topic "technology" and "environment" and make a list of 5 key nouns for each topic from that virtual corpus (choose the nouns that are the most relevant to the topics). Then search for 5 V-N collocations for those selected nouns and keep a record with frequency and MI score. Check concordance lines and take notes of 2 examples from the corpus, and then write 1 example of your own, using searched collocations. Provide a simple definition/ meaning/ Vietnamese equivalence of the collocations found. • Record screen while creating the virtual corpus and keeping records of verb-noun collocations. Then submit the recording to the Google Drive folder using the link on Canvas

Corpus-based Lesson plan #2

Target Audience and Context: 11th grade students

Learner Level: High-intermediate/ Low-advanced

Class size: ~30

Class length: 90 mins

Topic: Verb-Noun collocations (Part 2)

Objectives:

After the lesson, SWBAT:

- use COCA and/ or SKELL to self-evaluate their use of verb-noun collocations in their essays
- use COCA and/or SKELL to give peer-feedback to their peers on their use of verb-noun collocations in their essays
- use COCA and/or SKELL to self-correct their verb-noun collocation errors

Materials and resources:

- Slides: https://tinyurl.com/3ysyxjh2
- Jamboard: https://tinyurl.com/4cya2b6a
- Homework instruction: https://tinyurl.com/3hchebs2

Time **Materials**	**Procedures**
5 mins	Icebreaker • Poll: https://PollEv.com/duongnguyen376

20 mins	• Read the essay prompt and
	  ○ analyze the essay question (3 mins)
	  ○ brainstorm ideas (5 mins) on Jamboard: https://jamboard.google.com/d/1W6I_njgsHvw9jbryd6jB_r-IrW6t5--1DDc28iz_MT0/edit?usp=sharing
	  ○ make the outline (12 mins)
	  ○ https://docs.google.com/presentation/d/1tBcVYdz88J2ZmVasYOkpC8Gc_38id0AY4PvFtqerX-U/edit?usp=sharing
	
	Question:
	
	• **Option 1**
	
	A teacher's ability to relate well with students is more important than excellent knowledge of the subject being taught. To what extent do you agree or disagree?
	
	Write at least 250 words. Use specific reasons and examples to support your answer.
	
	• **Option 2**
	
	Television advertising directed toward young children (aged two to five) should not be allowed. To what extent do you agree or disagree?
	
	Write at least 250 words. Use specific reasons and examples to support your answer.

40 mins	**Activity 2: Essay writing** **Question:** • **Option 1**

	A teacher's ability to relate well with students is more important than excellent knowledge of the subject being taught. To what extent do you agree or disagree? **Write at least 250 words. Use specific reasons and examples to support your answer.** • **Option 2** Television advertising directed toward young children (aged two to five) should not be allowed. To what extent do you agree or disagree? **Write at least 250 words. Use specific reasons and examples to support your answer.** Students write the essay on their own. During this process, students can use COCA or SKELL as a reference, but students are encouraged to use the corpus after they finish the essay to avoid disrupting their flow of ideas.
25 mins	**Activity 3: Peer-feedback** • Students give feedback to their peers about the use of verb-noun collocations in the essays **Steps:** 1. Identify and underline verb-noun combinations in your peer's essay 2. Use the collocate function to check whether those word combinations are acceptable collocations. Record the frequency and MI score. 3. Remember to check the concordance lines to see how the collocations are used in specific contexts

	4. Think critically, based on what you did in step 2 and 3, and decide whether your partner's use of collocations is acceptable or not. If it is not, make a comment on that. 5. Remember to record the screen when you use COCA/SKELL to give feedback on the use of collocations in your peers' essays.
	Homework: • Use COCA/SKELL to self-correct the collocation errors marked by your peer

Corpus-based Lesson Plan #3

Target Audience and Context: 11th grade students
Learner Level: High-intermediate/ Low-advanced
Class size: 30-35
Class length: 90 mins

Topic: Verb-Noun collocations (Part 3)

Objectives:

After the lesson, SWBAT:

- use COCA and/ or SKELL to identify strong verb-noun collocations in a sentence
- use COCA and/or SKELL to identify verb-noun collocation errors and correct them
- develop a sense of collocation use when constructing a sentence
- develop a larger repertoire of verb-noun collocations in topic "crime" and "health"

Materials and resources:

- Slides: https://tinyurl.com/2jx3y3t8
- Handout: https://tinyurl.com/hvfh9ehc

Time & Materials	**Procedures**

15 mins	**Warm-up:**
	• Use COCA to find some common synonyms of the word "crime". Type this query in the search box '=crime'.
	  • How many words can you find?
	  • Which word has the highest frequency? Click on that word to examine the concordance lines. What do you notice?
	• Now use the collocate function to find some common collocations with these synonyms. Type '=crime' in the first box, and choose 'verb' for the POS in the second box.

	• What nouns can go with the verbs "commit", "charge", "accuse"? • Take notes of five V-N collocations that you are interested in? (E.g. commit a really heinous crime") • Share your notes with your partners
20 mins	**Activity 1: Error identification** Some students are writing about the reasons why ex-inmates reoffend after their release and the solutions to this problem. Look at the following sentences from these students' essays and identify collocation errors. Use the collocate function in COCA to check your intuition and correct those errors. (Note: Pay attention to the meaning of the whole sentence.) 1. As a result, many former inmates are forced to commit crimes again in order to continue their life or just to alleviate the aversion they face from the general population, such as robbery for money or murder for vengeance. 2. The ideal solutions should be enacted by the authorities. 3. There are a variety of reasons why this is the case and effective measures need to be taken urgently to tackle the phenomenon. 4. As a result, many ex-criminals have to commit crimes again to sustain their lives or to just relieve aversion they receive from the general public. 5. Therefore, the administrators as well as the society should cooperate to lessen the aforementioned reasons.
20 mins	**Activity 2: Gap-filling** Read the following sentences carefully to understand the meaning. Then try to fill in the blanks with the suitable verb, and use the collocate function in COCA to check your intuition and modify your answers if necessary. (There may be more than one answer for each sentence.) 1. Young and first-time offenders should not be mingled with seasoned criminals while in jail to _____ the likelihood of being reintroduced to crimes. 2. In addition, the government should _____ initiatives to provide financial assistance as well as vocational training. 3. Small funds may be provided to assist inmates on their release in adjusting, but the emphasis should be _____ on education. 4. When they apply for a job, they may be looked down on and they cannot _____ a chance of landing a decent job.

	5. While there are some reasons for this problem, a number of feasible solutions should be adopted to _____ the consequences. —---------- 6. After a frantic series of late night meetings among environment ministers, the conference's decisive ministerial segment went smoothly for the report's supporters. First the G-77 and China endorsed it, then the EU ministers, and finally the United States, which _____ scorn on those who opposed the findings. 7. However, the government and society can _____ certain methods to deter them from reoffending.
35 mins	**Activity 3: Sentence building** Imagine you are writing an essay about "crime punishment" . Continue writing to develop the following points (3-5 sentences). Identify verb-noun collocations you may use in your sentences. Use the collocate function in COCA to check whether your use of verb-noun collocations is correct. 1. Some may argue that rehabilitation is a "soft punishment" and is not fair to the victims of the crime. 2. The purpose of death penalty is to serve as a crime deterrent. *(Source: Crow Corpus)*
10 mins	**Homework:** • Create an academic virtual corpus about the topic "crime" and "health" and make a list of 5 key nouns for each topic from that virtual corpus (choose the nouns that are the most relevant to the topics). Then search for 5 V-N collocations for those selected nouns and keep a record with frequency and MI score. Check concordance lines and take notes of 2 examples from the corpus, and then write 1 example of your own, using searched collocations. Provide a simple definition/ meaning/ Vietnamese equivalence of the collocations found. • Record screen while creating the virtual corpus and keeping records of verb-noun collocations. Then submit the recording to the following link:

Corpus-based Lesson Plan #4

Target Audience and Context: 11th grade students
Learner Level: High-intermediate/ Low-advanced
Class size: ~30
Class length: 90 mins

Topic: Verb-Noun collocations (Part 4)

Objectives:

After the lesson, SWBAT:

- use COCA and/ or SKELL to self-evaluate their use of verb-noun collocations in their essays
- use COCA and/or SKELL to give peer-feedback to their peers on their use of verb-noun collocations in their essays
- use COCA and/or SKELL to self-correct their verb-noun collocation errors

Materials and resources:

- Slides: https://tinyurl.com/yckw5z2x
- Jamboard for brainstorming and outlining: https://tinyurl.com/ycynbtyb
- Handout: https://tinyurl.com/3e52tczv

Time & Materials	**Procedures**
5 mins	Icebreaker: • Poll: https://PollEv.com/duongnguyen376

20 mins	Read the essay prompt and • analyze the essay question (3 mins) • brainstorm ideas (5 mins) on Jamboard: • make the outline (12 mins) • Duration: 20 mins **Question:** Science finds that by studying the behavior of three-year-old children, they can predict if they will grow up to be criminals. To what extent is criminality related to human nature? Is it possible to stop children from being a criminal when they grow up? Write at least 250 words. Use specific reasons and examples to support your answer. (Ielts Writing Task 2 - Vietnam - 06/27/2019)
40 mins	**Activity 2: Essay writing** Question: **Science finds that by studying the behavior of three-year-old children, they can predict if they will grow up to be criminals. To what extent is criminality related to human nature? Is it possible to stop children from being a criminal when they grow up?** (06/27/2019) Students analyze the question and write the essay on their own. During this process, students can use COCA or SKELL as a reference, but students are encouraged to use the corpus after they finish the essay to avoid disrupting their flow of ideas.
25 mins	**Activity 3: Peer-feedback** • Students give feedback to their peers about the use of verb-noun collocations in the essays **Steps:** • Identify and underline verb-noun combinations in your peer's essay • Use the collocate function to check whether those word combinations are acceptable collocations. Record the frequency and MI score. • Remember to check the concordance lines to see how the collocations are used in specific contexts • Think critically, based on what you did in step 2 and 3, and decide whether your partner's use of collocations is acceptable or not. If it is not, make a comment on that. • Remember to record the screen when you use COCA/SKELL to give feedback on the use of collocations in your peers' essays.

	Homework: Use COCA/SKELL to self-correct the collocation errors marked by your peer

Corpus-based Lesson Plan #5

Target Audience and Context: 11th grade students

Learner Level: High-intermediate/ Low-advanced

Class size: ~30

Class length: 90 mins

Topic: Verb-Noun collocations (Part 5)

Objectives:

After the lesson, SWBAT:

- use COCA and/ or SKELL to identify strong verb-noun collocations in a sentence
- use COCA and/or SKELL to identify verb-noun collocation errors and correct them
- develop a sense of collocation use when constructing a sentence
- develop a larger repertoire of verb-noun collocations in topic "education" and "culture"

Materials and resources:

- Slides: https://tinyurl.com/5n7jk6j6
- Handout: https://tinyurl.com/48w2f5dw

Time & Materials	**Procedures**

15 mins	**Warm-up:**
	• Use COCA to find some common synonyms of the word "viewpoint". Type this query in the search box '=viewpoint'.
	  • How many words can you find?
	  • Which word has the highest frequency? Click on that word to examine the concordance lines. What do you notice? (word form, patterns)

	• **Choose another word you are not familiar with and explore the concordance lines containing that word. What can you learn from that word?** • Now use the collocate function to find some common collocations with these synonyms. Type '=viewpoint' in the first box, and choose 'verb' for the POS in the second box. • Which verb has the highest frequency? Can I find other forms of this verb in the search results? • Examine the concordance lines of that verb and its other forms. Randomize the concordance lines by choosing a different number from ***"Find sample"***. What patterns of verb-noun collocations can you learn from that? (E.g. taking a long-term view). • What do you notice about the noun "position" in the collocations you found? • Take notes of five V-N collocations that you are interested in? (E.g. hold the view (that)) • Share your notes with your partners
20 mins	**Activity 1: Error identification** Some students are writing about whether university education should be free or not. Look at the following sentences from these students' essays and identify collocation errors. Use the collocate function in COCA to check your intuition and correct those errors. (Note: Pay attention to the meaning of the whole sentence.) 1. The ministry can raise the taxes on employed people to free the college education which can raise the knowledge of the society. (suggested answer: raise the knowledge -> enhance the knowledge) 2. Many individuals believe that in order to ensure that all students may attend education regardless of their financial situation, the government should make tuition free for all students. (suggested answer: attend education -> pursue higher education) 3. To begin with, authorities should advocate financial support to all students to reduce social disparity. (suggested answer: advocate financial support -> offer/lend financial support) 4. In Vietnam, a number of scholarships are given to poor students yearly and help them reach the university. (suggested answer: reach the university -> attend the university) 5. Because if all students have access to university education but don't have to spend money, they will disregard this opportunity. (suggested answer: disregard this opportunity -> squander/ waste this opportunity)

20 mins	**Activity 2: Gap-filling** Read the following sentences carefully to understand the meaning. Then try to fill in the blanks with the suitable given verbs, and use the collocate function in COCA to check your intuition and explain your choice. 1. You can also _____ an **online** class now or take a full time 4 week class over the summer abroad and get some experience. (take/ get) (take) 2. In 2016 more than 6.3 million students at over 4,700 colleges and universities choose to _____ an online course (enroll in/ apply for) (enroll in) 3. Unpaid interns, who often work only part time, agree to a trade-off: volunteering their time to learn and _____ **hands-on** experience, and perhaps a future paid staff position, at a company in the field of their choice. (gain/ take) (gain) 4. When they are included into an established daily routine, they provide a built-in authentic purpose for writing, increase writing time, and _____ **communication** between school and home. (facilitate/ support) (facilitate) 5. Primary examples of technical motives include: desire to integrate systems cross-functionally (internal and external system integration), reduce software maintenance, and _____ technology capacity **constraints.** (ease/ liberate) (ease)
35 mins	**Activity 3: Sentence building** Imagine you are writing an essay about the ways to cope with culture shock. Continue writing to develop the following points (3-5 sentences). Identify verb-noun collocations you use in your sentences. Use the collocate function in COCA to check whether your use of verb-noun collocations is correct. 1. I believe that it is important for international students to keep their own identity while dealing with culture shock. 2. International students should make friends with domestic people. *(Source: Crow Corpus)*
10 mins	**Homework:** • Create an academic virtual corpus about the topic "education" and "culture" and make a list of 5 key nouns for each topic from that virtual corpus (choose the nouns that are the most relevant to the topics). Then search for 5 V-N collocations for those selected nouns and keep a record with frequency and MI score. Check concordance lines and take notes of 2 examples from the corpus, and then write 1 example of your

	own, using searched collocations. Provide a simple definition/ meaning/ Vietnamese equivalence of the collocations found. • Record screen while creating the virtual corpus and keeping records of verb-noun collocations. Then submit the recording to the assigned link on Canvas

APPENDIX G: DEMOGRAPHIC AND BACKGROUND INFORMATION OF THE PARTICIPANTS

Table 12.

Demographic and background information of the participants (part 1)

Student	Age	Preference for novelty	Frequency of writing in English	Sources used for writing	Collocations cause difficulties in writing English	Interest in learning collocations
001	17	absolutely yes	three times a month	online dictionaries such as Tflat, Cambridge, Longman	totally disagree	very interested
002	17	probably yes	twice a month	online dictionaries	partly agree	very interested
003	17	probably yes	once a week	online dictionaries	partly agree	very interested
004	17	absolutely yes	once a week	online dictionaries	partly agree	very interested
005	17	absolutely yes	three times a month	online dictionaries such as Oxford, Cambridge	partly agree	very interested
006	17	probably yes	once a week	Ozdic, Oxford dictionary	totally agree	very interested
007	16	absolutely yes	three times a month	Cambridge and Longman dictionaries	partly disagree	very interested
008	17	probably yes	three times a month	Oxford dictionary, Ozdic, relevant newspapers, samples	no opinion	quite interested

Table 12. (cont'd)

009	17	probably yes	three times a month	online dictionaries, academic writing materials from self-collections and courses	totally agree	very interested
0010	17	absolutely yes	once a week	Cambridge Dictionary	partly agree	very interested
0011	17	absolutely yes	once a week	Ozdic, Cambridge dictionary, Oxford learner's dictionary	totally disagree	very interested
0012	17	absolutely yes	once a week	newspapers, online dictionaries	totally agree	very interested
0013	17	absolutely yes	once a week	Cambridge dictionary, Longman dictionary, Ozdic	partly agree	very interested
0014	17	probably yes	once a week	online dictionaries	partly agree	very interested
0016	17	absolutely yes	three times a month	online dictionaries	partly agree	quite interested
0017	17	probably yes	three times a month	online dictionaries such as Oxford, Cambridge, Collins	partly agree	quite interested
0018	17	probably yes	twice a month	online dictionaries and relevant websites	totally agree	very interested
0019	17	absolutely yes	twice a month	Thesaurus and Cambridge dictionary	partly agree	no opinion

Table 12. (cont'd)

0020	17	probably yes	twice a month	online dictionaries	partly agree	very interested
0021	17	probably yes	three times a month	Cambridge dictionary, newspapers relevant to topics for writing	partly agree	quite interested
0022	17	absolutely yes	once a week	Oxford and Cambridge dictionaries	totally agree	very interested
0023	17	probably yes	twice a month	Cambridge dictionary, Lexico, other online dictionaries, websites with contents about collocations and synonyms	totally agree	very interested
0024	17	probably yes	once a week	online dictionaries such as Longman, Oxford, Cambridge	partly agree	very interested
0025	17	absolutely yes	three times a month	Cambridge dictionary	totally agree	very interested

Table 13.

Demographic and background information of the participants (part 2)

Student	Awareness about the existence of corpora	Prior use of corpus	The corpus used	Exposure to English outside English classroom	Sources of English exposure outside school	Notice of collocations in those sources	Use of those colloca-tions in writing English
001	no	no	N/A	two hours/ day	YouTube, Ted Talks	quite often	some-times
002	yes	yes	COCA	two hours/ day	movies, news-papers, and news	sometimes	some-times
003	no	no	N/A	three hours/ day	YouTube, TV	sometimes	some-times
004	yes	yes	COCA	two hours/ day	blog, movies	quite often	often
005	yes	no	N/A	less than one hour/day	books, news-papers	sometimes	seldom
006	no	no	N/A	one hour/ day	BBC Learning English	sometimes	seldom
007	no	no	N/A	more than three hours/day	VOX, BBC	sometimes	never
008	yes	no	N/A	two hours/ day	Internet, online news-papers, YouTube	sometimes	some-times
009	yes	yes	COCA	more than three hours/day	YouTube, TV series, news-papers, books, podcasts	quite often	some-times
0010	no	no	N/A	two hours/ day	music, movies, YouTube videos	sometimes	some-times

Table 13. (cont'd)

0011	yes	no	N/A	two hours/ day	YouTube videos, podcasts, books, music, movies, news	sometimes	some-times
0012	yes	no	N/A	more than three hours/day	movies, music, news-papers, books	very often	quite often
0013	yes	yes	COCA	more than three hours/day	films, podcasts, daily news from Gmail	quite often	quite often
0014	yes	no	N/A	two hours/ day	YouTube videos, Netflix, online materials	sometimes	seldom
0016	no	no	N/A	two hours/ day	movies, texts, music	sometimes	some-times
0017	no	no	N/A	two hours/ day	news, movies, books	sometimes	some-times
0018	no	no	N/A	three hours/ day	movies, music, social media platforms, news channels	sometimes	seldom
0019	no	no	N/A	one hour/ day	no answers	sometimes	some-times
0020	yes	yes	COCA	three hours/ day	movies, music, news, news-papers	sometimes	some-times

Table 13. (cont'd)

0021	no	no	N/A	one hour/ day	inspiring short videos	sometimes	some-times
0022	yes	yes	COCA	two hours/ day	podcasts	sometimes	very often
0023	no	no	N/A	more than three hours/day	news channels, movies, novels, news-papers	sometimes	some-times
0024	no	no	N/A	less than one hour/day	movies, news-papers	sometimes	seldom
0025	yes	yes	COCA	three hours/ day	movies, books	quite often	quite often

APPENDIX H: V-N COLLOCATION RATINGS

Table 14.

V-N collocation ratings (L1 speakers' ratings vs. COCA-based ratings)

Student	Essay	Collocation	Frequency	MI score	R1	R2	CBR
001	1	utilize foundation	0	0	1	1	1
		attend program	152	1.65	3	3	2
		partake college	0	0	1	1	1
		hold merit	17	0.7	3	3	1
		have contact	2627	0.2	3	3	2
		understand subject	104	-0.64	3	3	2
		contact professor	14	0.27	3	3	1
		exchange opinion	23	3.57	2	2	3
		have talk	2566	-0.44	2	2	2
		strengthen relationship	117	3.83	3	3	3
		enhance state	38	-0.83	1	1	1
		absorb knowledge	44	3.58	3	3	3
		obtain knowledge	86	3.13	2	2	3
		supply learner	0	0	2	3	1
		allocate time	122	2.87	3	3	2
		acquire knowledge	483	6.46	3	3	3
		participate (in) course	32	-0.6	2	2	1
		host course	11	-1.37	2	2	1
	2	join course	26	-1.98	3	3	1

Table 14. (cont'd)

have merit	561	0.59	3	3	2
have benefit	1253	-0.57	3	3	2
attend course	69	0.51	2	2	1
offer contact	29	-0.34	1	1	1
resolve question	172	2.62	1	1	2
raise question	1891	3.89	3	3	3
have insight	636	-0.13	3	3	2
crave knowledge	5	2.05	3	2	1
have opportunity	10357	1.05	3	3	2
enrich life	267	3.88	3	3	3
enhance skill	215	4.83	3	3	3
strengthen bond	65	4.65	3	3	3
hold party	163	-0.23	2	2	2
invite friend	495	4.2	3	3	3
invite teacher	26	0.81	2	2	1
deepen relationship	64	5.4	3	3	3
give advantage	562	2.12	2	3	2
widen horizon	13	5.76	3	2	3
utilize foundation	0	0	1	1	1
act encyclopedia	0	0	1	1	1
give access	1282	2.49	3	3	2
have problem	21797	0.49	3	3	2

Table 14. (cont'd)

		gain insight	6.72	7.38	3	3	3
		surf Internet	117	7.36	3	3	3
		investigate aspect	18	1.8	2	2	2
		hold knowledge	37	-0.95	1	1	1
		give instruction	332	1.72	3	3	2
		act lifeline	1	1.17	1	1	1
		provide assistance	1426	5.89	3	3	3
		promote understanding	306	5	3	3	3
		hold debate	66	-0.06	2	2	1
		assist student	561	3.89	3	3	3
		advance horizons	0	0	2	2	1
		alter education	6	-0.86	1	1	1
		discuss matter	427	2.97	3	3	2
002	1	support idea	770	2.21	3	2	2
		attend college	1430	5.68	3	3	3
		gain insight	672	7.38	3	3	3
		follow presentation	14	0.28	2	2	1
		keep track	4547	4.67	3	3	3
		keep up with lecture	0	0	2	2	1
		reduce possibility	129	2.97	3	3	2
		attend class	311	3.51	3	3	3
		receive education	550	3.27	2	2	3

Table 14. (cont'd)

	acquire curriculum	3	1.02	1	1	1
	take part in club	2	0.44	2	2	1
	enhance health	153	2.94	1	2	2
	sharpen skill	153	7.64	3	3	3
	offer opportunity	769	3.85	2	3	3
	apply theory	151	2.86	2	3	2
	apply experience	42	-0.54	2	2	1
2	support education	466	2.11	2	2	2
	receive feedback	292	5.67	3	3	3
	complete work	497	1.86	3	3	2
	keep track	4547	4.67	3	3	3
	raise question	1891	3.89	3	3	3
	ask (for) explanation	53	3.91	3	2	3
	attend college	1430	5.68	3	3	3
	ensure principle	12	-0.01	1	1	1
	have attitude	1518	-0.35	3	2	2
	follow lecture	10	0.08	3	2	1
	comprehend lecture	0	0	2	2	1
	utilize device	11	1.73	1	1	1
	have difficulty	3722	1.59	3	3	2
	gain access	1935	6.92	3	3	3
	support idea	770	2.21	3	2	2

Table 14. (cont'd)

		receive education	550	3.27	3	2	3
003	1	provide support	3761	4.02	3	3	3
		keep pace	1617	4.8	2	2	3
		utilize approach	25	2.68	1	1	2
		bring effect	45	-1.99	2	1	1
		provide flexibility	204	4.67	3	2	3
		arrange timetable	0	0	2	2	1
		bring joy	526	4.14	3	3	3
		bring passion	59	1.09	2	2	2
		come across issue	8	-0.39	3	2	1
		adjust timetable	4	4.37	2	2	1
		deal with problem	1029	2.81	3	3	2
		give opinion	594	1.62	3	3	2
		give question	177	-2.4	1	1	1
		undergo pressure	1	-1.3	2	2	1
		absorb lesson	14	2.72	1	1	2
		exceed level	63	2.73	2	2	2
		take exam	450	3.03	3	3	3
		reach region	21	-0.57	2	1	1
		solve issue	345	2.8	3	3	2
		take class	1180	0.63	3	3	2
		provide opportunity	1940	4.31	3	3	3

Table 14. (cont'd)

	keep contact	397	1.25	3	3	2
	improve skill	871	5.02	3	3	3
	stand chance	1780	4.51	3	3	3
	expand relationship	25	0.4	1	1	1
	enhance skill	215	4.83	3	3	3
	stand complementarity	0	0	1	1	1
	attend class	311	3.51	3	3	3
	decrease opportunity	8	0.69	1	2	1
	inhale knowledge	0	0	1	1	1
	distract concentration	2	1.88	1	1	1
	lessen possibility	13	4.07	2	1	3
	diminish motivation	5	3.34	2	2	1
	diminish passion	1	0.5	1	2	1
2	have access	14510	1.62	3	3	2
	attend college	1430	5.68	3	3	3
	enroll (in) university	438	3.97	3	3	3
	get access	1821	0.5	3	3	2
	affect academic performance	34	5.28	2	2	3
	offer curriculum	48	2.13	2	2	2
	access structure	6	-1	1	1	1
	further research	20	2.68	3	3	2

Table 14. (cont'd)

		gain knowledge	551	5.2	3	3	3
		widen relationship	2	0.06	1	1	1
		develop skill	2045	6.06	3	3	3
		get chance	7167	1.88	3	3	2
		do job	21689	1.07	3	3	2
		help the poor	662	3.03	3	3	3
		assist student	561	3.89	3	3	3
		gain advantage	527	5.87	3	3	3
		build (relationship)					
		circle	0	0	1	1	1
004	1	join class	131	1.18	3	3	2
		offer merit	16	1.43	2	1	2
		save time	2298	1.56	3	3	2
		attend lesson	6	-0.45	3	3	1
		gain knowledge	551	5.2	3	3	3
		obtain information	1129	5.4	2	2	3
		download content	74	4.17	2	3	3
		deal with advantage	0	0	2	1	1
		create atmosphere	570	5.01	2	3	3
	2	take class	1180	0.63	3	3	2
		attend session	166	4.53	2	2	3
		possess privilege	6	2.47	1	1	1

Table 14. (cont'd)

		save time	2298	1.56	3	3	2
		get access	2783	1.85	3	3	2
		download material	71	2.05	3	3	2
		miss lesson	14	-0.28	2	2	1
		provide experience	504	1.31	2	2	2
		provide environment	483	2.88	2	2	2
		comprehend lesson	2	0.41	2	2	1
		address concern	216	3.19	3	3	3
		enrich development	8	1.86	2	2	1
		exchange thought	8	1.63	2	2	1
		exchange opinion	23	3.57	1	1	3
		offer opportunity	1442	4.76	3	3	3
		enhance skill	215	4.83	3	3	3
		produce performance	58	0.91	2	2	1
		impose limitation	366	5.83	2	2	3
		gain knowledge	551	5.2	2	2	3
		allow performance	27	-0.97	1	1	1
		allow education	49	-0.97	1	1	1
005	1	enroll (in) education	37	3.46	1	1	3
		bring advantage(s)	23	-1.26	2	2	1
		follow curriculum	56	2.09	3	3	2
		have benefit	1872	-1.08	3	3	2

Table 14. (cont'd)

	reduce burden	283	5.27	2	2	3
	lessen risk	78	5.36	2	3	3
	access knowledge	50	2.76	3	3	2
	keep distance	1005	3.22	3	3	3
	attend class	311	3.51	3	3	3
	solve problem	7030	6.85	3	3	3
	acquire knowledge	483	6.46	3	3	3
	have opportunity	10357	1.05	3	3	2
	learn subject	105	0.24	2	3	2
	conduct class	36	1.23	3	3	2
	provide information	6754	5.05	3	3	3
	follow progress	161	2.49	1	2	2
	shape character	49	3.16	3	3	3
	stand chance	1780	4.51	3	3	3
	mold personality	1	2	2	2	1
	pursue education	341	4.09	3	3	3
	bring benefit	234	0.94	1	2	2
	attend college	1430	5.68	3	3	3
	attend university	594	4.1	3	3	3
	boost development	40	2.37	2	2	2
2	enroll (in) education	37	3.46	1	1	3
	bring advantage(s)	23	-1.26	2	2	1

Table 14. (cont'd)

follow curriculum	56	2.09	2	2	2
have benefit	1872	-1.08	3	3	2
reduce burden	283	5.27	3	3	3
lessen risk	78	5.36	3	3	3
save time	2298	1.56	3	3	2
access knowledge	50	2.76	3	3	2
cite example	231	4.89	2	2	3
keep distance	1005	3.22	3	3	3
attend period	6	-1.54	2	2	1
tackle trouble	2	-0.81	1	1	1
gain information	354	3.12	2	3	3
participate (in) college	43	0.62	3	3	1
participate (in) university	16	-1.12	3	3	1
stand chance	1780	4.51	3	3	3
conduct class	47	1.62	3	3	2
provide data	1252	3.04	2	3	3
follow progress	161	2.49	2	2	2
shape character	49	3.16	2	2	3
offer opportunity	1442	4.76	3	3	3
mold personality	1	2	3	3	1
bring benefit	234	0.94	2	2	2

Table 14. (cont'd)

		boost development	40	2.37	2	2	2
006	1	take advantage	12983	5.76	3	3	3
		attend college	1430	5.68	3	3	3
		reap benefit	869	8.62	3	3	3
		take part in program	46	3.04	3	3	3
		engage interest	48	0.89	1	1	1
		assign homework	56	7.59	3	3	3
		make announcement	823	3.24	3	3	3
		have attention span	130	1.46	3	3	2
		gain skill	281	4.25	3	3	3
		attend class	311	3.51	3	3	3
		allow progress	43	0.61	3	3	1
		have insight	762	-0.28	2	2	2
		give suggestion	175	1.52	3	3	2
		participate (in) class	128	2.23	3	3	2
		join lesson	6	-1.54	2	2	1
		expand (social) circle	46	2.93	3	3	2
		establish relationship	468	4.34	3	3	3
		expand relationship	25	0.4	1	1	1
		do favor	5452	2.82	3	3	2
		suffer (from) depression	1101	5.46	3	3	3

Table 14. (cont'd)

	take (into) consideration	1569	3.84	3	3	3
	make choice	3947	2.58	3	3	2
2	take advantage	12983	5.76	3	3	3
	attend course	69	0.51	3	3	1
	engage interest	48	0.89	3	1	1
	make use (of)	5623	0.02	3	3	2
	extend (attention) span	0	0	2	1	1
	provide schedule	27	-0.67	3	3	1
	gain appreciation	114	6.13	3	3	3
	advance career	221	5.13	3	3	3
	reap benefit	965	8.77	3	3	3
	enroll (in) lesson	0	0	3	3	1
	provide insight	1230	5.93	3	3	3
	achieve goal	1739	6.33	3	3	3
	evaluate strength	28	2.63	3	3	2
	evaluate weakness	10	2.94	3	3	2
	expand (social) circle	55	3.18	2	3	3
	establish relationship	468	4.34	3	3	3
	befriend people	22	2.14	3	3	2
	do favor	5452	2.82	3	3	2

Table 14. (cont'd)

		suffer (from)					
		depression	1101	5.46	3	3	3
		make choice	3947	2.58	3	3	2
007	1	play role	13534	6.5	3	3	3
		pursue education	341	4.09	3	3	3
		bring merit	7	-0.94	2	1	1
		access resource	178	4.54	2	2	3
		facilitate process	277	4.54	1	2	3
		solve task	147	2.13	3	3	2
		gain knowledge	551	5.2	3	3	3
		save cost	140	0.59	2	3	2
		prevent spread	444	4.76	3	3	3
		do wonders	412	1.17	3	3	2
		attend education	78	1.45	1	1	2
		encourage student	1151	3.96	2	2	3
		motivate student	936	3.47	2	2	3
		gain insight	672	7.38	3	3	3
		create condition	610	3.01	2	3	3
		accumulate source	0	0	1	1	1
		land (a) job	443	4.2	3	3	3
		expand circle (of friends)	46	2.93	3	3	2

Table 14. (cont'd)

		have opportunity	10357	1.05	3	3	2
	2	get access	2783	1.85	2	3	2
		pursue education	341	4.09	3	3	3
		establish habit	32	2.96	2	2	2
		fulfill duty	170	6.51	3	2	3
		stimulate growth	455	7.36	3	3	3
		popularize form	3	3.28	2	2	1
		attend education	78	1.45	1	1	2
		have interaction	387	-1.7	2	3	2
		facilitate progress	24	2.98	2	2	2
		provide material	351	2.08	3	3	2
		gain insight	672	7.38	3	3	3
		realize dream	280	2.74	3	3	2
		attend university	594	4.1	3	3	3
008	1	facilitate use	141	1.22	1	1	2
		have edge	1104	-1.55	3	3	2
		boost attention	4	-0.61	1	1	1
		foster concentration					
		span	0	0	1	1	1
		attend college	1430	5.68	3	3	3
		have influence	3546	0.51	3	3	2
		lift spirit	57	3.25	3	3	3

Table 14. (cont'd)

	create environment	1431	4.77	3	3	3
	save time	2298	1.56	3	3	2
	prepare uniform	1	-2.06	2	1	1
	guarantee result	75	2.01	3	2	2
	boost quality	45	3.16	1	2	3
	maximize power	82	3.08	2	2	3
	optimize study	0	0	2	1	1
	access source(s)	27	1.39	3	2	2
	* make advantage	7	-4.35	1	1	1
	retain material	11	0.85	3	3	1
	find pile (of papers)	30	-1.74	2	2	1
	play role	13534	6.5	3	3	3
	overlook significance	11	4.83	1	1	3
2	have perk	75	-1.09	3	2	1
	monitor performance	154	4.44	2	2	3
	supervise class	5	1.43	3	3	1
	draw attention	2940	6.93	3	3	3
	improve (concentration) span	0	0	1	1	1
	do wonders	412	1.17	3	3	2
	run risk	1574	3.52	3	3	3
	bear fruit	615	6.89	3	3	3

Table 14. (cont'd)

		save time	2298	1.56	3	3	2
		choose time	433	-0.59	3	3	2
		attend class	311	3.51	3	3	3
		accelerate progress	75	5.84	2	2	3
		comprise source	4	0.91	1	1	1
		search (for) information	190	2.99	3	3	2
		bring benefit	234	0.94	1	1	2
009	1	gain knowledge	551	5.2	3	3	3
		visit lecture	0	0	1	1	1
		attend lecture	46	4.06	2	2	3
		continue information	0	0	1	1	1
		give opportunity	2377	3.08	1	3	3
		complete study	139	1.55	3	3	2
		gain learning	11	0.34	2	1	1
		manage balance	48	2.29	2	2	2
		enroll (in) program	100	4.13	3	3	3
		take admission	26	-1.55	1	1	1
		save time	2298	1.56	3	3	2
		manage finance(s)	155	4.7	3	3	3
		obtain education	112	2.51	3	2	2
		save money	7197	5.43	3	3	3

Table 14. (cont'd)

	2	achieve potential	176	3.29	3	2	3
		choose term	29	-0.87	1	1	1
		choose condition	18	-1.37	2	2	1
		reach horizons	2	-1.58	2	2	1
		provide resource(s)	757	3.29	3	3	3
		give exam	43	0.53	3	3	1
		provide material	351	2.08	3	3	2
		manage lifestyle	4	1.31	2	2	1
		take on job	132	1.89	2	2	2
		use technology	1966	3.03	3	3	3
		save time	2298	1.56	3	3	2
		save money	7197	5.43	3	3	3
		have hour	3349	-1.6	1	1	2
		follow routine	84	2.74	3	3	2
		adapt (to) schedule	6	1.18	2	2	1
		gift freedom	0	0	1	1	1
		bring benefit	234	0.94	1	1	2
0010	1	pursue education	341	4.09	3	3	3
		have benefit	630	-0.56	3	3	2
		afford course	15	-0.44	3	2	1
		offer course	2000	3.01	3	3	3
		impede path	3	2.33	2	2	1

Table 14. (cont'd)

	impede choice	2	0.6	1	2	1
	provide lecture	6	-1.26	2	2	1
	establish class	21	-0.34	2	2	1
	attend class	311	3.51	3	3	3
	take example	2239	1.21	1	1	2
	supply material	33	2.59	2	2	2
	have propensity	184	0.35	2	3	2
	follow form (of study)	49	-1.62	2	2	1
	access form (of study)	14	-0.28	2	2	1
	achieve contact	7	-1.21	1	1	1
	save time	2298	1.56	3	3	2
	spend time	24350	4.9	3	3	3
	take advantage	12983	5.76	3	3	3
	attend classroom	10	0.57	1	1	1
	adopt method	32	1.23	2	2	2
2	reap benefit	965	8.77	3	3	3
	overtake class	1	-0.21	1	1	1
	prepare video	12	-0.86	2	2	1
	offer opportunity	1442	4.76	3	3	3
	pursue education	341	4.09	3	3	3
	prevent opportunity	4	-3.07	1	1	1
	eradicate obstacle	0	0	2	2	1

Table 14. (cont'd)

		reduce time	636	0.38	3	3	2
		waste time	4883	5.28	3	3	3
		enroll (in) course	61	3.42	3	3	3
		require equipment	125	3.26	2	2	3
		comprehend lesson	2	0.41	1	2	1
		facilitate environment	20	2.13	1	1	2
		consult teacher	12	0.17	2	2	1
		promote teamwork	10	4.9	3	3	2
		promote collaboration	48	4.4	3	3	3
		create opportunity	321	2.04	3	3	2
0011	1	gain (in) popularity	103	5.83	3	3	3
		offer selection	116	3.12	3	2	3
		provide access	1695	4.41	3	3	3
		register (for) course	14	-0.27	3	3	1
		equalize opportunity	17	5.62	2	3	3
		have job	10724	-0.53	3	3	2
		necessitate approach	4	2.78	1	1	1
		attend lecture	46	4.06	3	3	3
		create balance	185	2.3	3	3	2
		attend college	1430	5.68	3	3	3
		play role	13534	6.5	3	3	3
		act (as) incentive	13	1.18	2	2	2

Table 14. (cont'd)

	emulate quality	2	0.6	1	1	1
	partake (in) activity	16	4.18	2	3	3
	stand chance	1780	4.51	3	3	3
	guarantee (job) prospect	0	0	3	3	1
	gain appreciation	144	6.13	3	3	3
	gain (competitive) edge	156	3.82	2	3	3
	facilitate learning	256	6.41	3	3	3
	gain skill	281	4.25	3	3	3
	gain experience	709	4.12	3	3	3
2	gain (in) popularity	103	5.83	3	3	3
	offer selection	116	3.12	2	2	3
	apply (for) college	159	1.86	3	3	2
	provide access	1695	4.41	3	3	3
	download content	77	4.17	2	3	3
	register (for) course	14	-0.27	3	3	1
	equalize opportunity	17	5.62	1	1	3
	have job	10724	-0.53	3	3	2
	necessitate method	0	0	2	1	1
	take class	1180	0.63	3	3	2
	access lecture	1	-0.52	2	2	1
	desire (career) prospect	0	0	1	1	1

Table 14. (cont'd)

		have tendency	1976	1.35	3	3	2
		pursue education	341	4.09	3	3	3
		attend college	1430	5.68	3	3	3
		act (as) incentive	22	1.94	3	3	2
		emulate quality	2	0.6	1	1	1
		enhance performance	441	5.71	3	3	3
		lack skill	396	6.08	3	3	3
		lack ability	282	5.48	3	3	3
		partake (in) activity	21	4.57	3	3	3
		join (college) community	354	2.21	2	2	2
		enhance experience	349	4.06	1	1	3
		acquire (job) prospect	4	1.44	2	1	1
		gain appreciation	114	6.13	3	3	3
		gain (an) edge	156	3.82	3	3	3
		facilitate program	30	0.76	2	2	1
		attend classroom	10	0.57	1	1	1
		gain skill	281	4.25	3	3	3
		gain experience	709	4.12	3	3	3
0012	1	gain popularity	103	5.83	3	3	3
		attend college	1430	5.68	3	3	3
		have impact	13042	1.5	3	3	2

Table 14. (cont'd)

	develop environment	44	0.69	1	2	1
	discover diversity	5	0.76	1	1	1
	confront challenge	59	3.22	1	2	3
	clarify lesson	2	-0.16	2	1	1
	interrupt concentration	4	2.47	2	2	1
	lower (academic) performance	23	1.49	3	2	2
	join club	961	5.14	3	3	3
	give opportunity	2377	3.08	2	3	3
	make friend(s)	3132	0.84	3	3	2
	take class	1180	0.63	3	3	2
	use device	515	2.09	3	3	2
	pursue education	341	4.09	3	3	3
	attend university	594	4.1	3	3	3
2	use technology	1966	3.03	3	3	3
	attend university	594	4.1	3	3	3
	have impact	13042	1.5	3	3	2
	encourage student	1151	3.96	3	3	3
	have chance	18180	1.35	3	3	2
	study subject	75	2.01	3	3	2
	offer class	304	1.97	3	3	2
	have understanding	3946	0.19	3	3	2

Table 14. (cont'd)

		evaluate learning	49	3.47	3	3	3
		evaluate development	26	1.09	2	2	2
		shape personality	13	3.26	3	3	3
		attend college	1430	5.68	3	3	3
		make friend	3132	0.84	3	3	2
		hinder opportunity	0	0	2	3	1
		take class	1180	0.63	3	3	2
		attend class	311	3.51	3	3	3
		develop habit	105	3.52	3	3	3
		mold character	17	4.07	3	3	3
0013	1	have impact	13042	1.5			2
		bring advantage(s)	23	-1.26			1
		control pace (of learning)	55	2.84			2
		have chance	18180	1.35			2
		access content	209	4.95			3
		understand aspect	74	0.14			1
		utilize distance learning	0	0			1
		consume (energy) source	2	-1.49			1
		produce emissions	64	3.47			3

Table 14. (cont'd)

	reach material	7	-2.44	1
	attend class	311	3.51	3
	provide interaction	35	0.31	1
	involve experiment	2	-1.22	1
	have skill(s)	3200	-0.42	2
	do experiment	280	-1.61	2
	explain principle	23	-0.08	1
	boost improvement	0	0	1
	activate (critical) thinking	9	3.33	1
	express viewpoint	13	3.87	3
	enrich knowledge	23	4.07	3
	have ability	7094	0.42	2
	revise lesson	4	2.41	1
	pass exam	364	6.32	3
	choose (distance) learning	13	-0.71	1
	attend college	1430	5.68	3
	foster development	215	5.5	3
2	attend class	311	3.51	3
	bring benefit	234	0.94	2
	have impact	13042	1.5	2

Table 14. (cont'd)

*have freedom	1963	-0.05	2
create environment	1431	4.77	3
stand chance	1780	4.51	3
affect (academic) performance	519	4.65	3
affect productivity	51	4.41	3
have requirement	301	-1.44	2
deal with challenge(s)	191	2.16	2
access program	42	0.74	1
attend college	1430	5.68	3
offer opportunity	769	3.85	3
formulate foundation	3	2.06	1
stimulate critical thinking	77	5.82	3
pass exam	364	6.32	3
gain chance	14	-0.71	1
comprehend aspect (of lectures)	0	0	1
handle thought	16	-0.73	1
offer clarification	13	3.41	3
shape personality	13	3.26	3
skip class	65	2.84	2

Table 14. (cont'd)

		impose regulation	28	3.43	3
		impose supervision	4	3.39	1
		create level (of discipline)	134	-0.22	2
		mold characteristic	0	0	1
		bring merit	7	-0.94	1
		advance development	63	2.72	2
0014	1	have class	750	-2.48	2
		do the former	32	-4.41	1
		have access	14510	1.62	2
		travel (a long) distance	186	4.29	3
		manage syllabus	0	0	1
		miss (a) lesson	14	-0.28	1
		obtain knowledge	86	3.13	3
		pay money	2701	2.99	2
		pay cost	823	2.11	2
		find course(s)	213	-1.92	2
		teach lectures	22	2.08	2
		suffer (from) debt	5	-0.7	1
		attend course	69	0.51	1
		benefit people	429	1.09	2
		learn information	311	0.78	2

Table 14. (cont'd)

	have (an) environment	64	-5.1	1
	spend money	8959	5.69	3
	get internship	67	0.55	1
	build networking	7	1.52	1
	prepare (for) career	34	2.98	2
	provide environment	483	2.88	2
	reassess ability	3	2.72	1
	give advantage(s)	562	2.12	2
	have education	798	-2.84	2
2	hold (the) view	337	1.12	2
	access knowledge	50	2.76	2
	pay cost	823	2.11	2
	benefit students	256	2.11	2
	do homeschooling	3	-0.03	1
	decrease concentration	7	3.02	1
	decrease productivity	10	4.37	3
	provide curriculum	76	1.91	2
	censor content	35	6.13	3
	have access	14510	1.62	2
	give opportunity	35	6.13	3
	possess networking	0	0	1
	attend college	1430	5.68	3

Table 14. (cont'd)

		employ students	0	0	1
		hold (work) fair	40	-0.7	1
		choose approach	53	-0.06	1
		join education	16	-1.93	1
0016	1	provide opportunity	1940	4.31	3
		provide facilities	154	1.91	2
		do program(s)	117	-4.19	2
		gain education	72	1.26	2
		gain knowledge	551	5.2	3
		provide (a) structure	210	1.8	2
		enable individual	43	1.64	2
		provide intelligence	161	1.92	2
		create environment	1431	4.77	3
		create checks	18	-2.56	1
		reduce number	1632	3.94	3
		take advantage	12983	5.76	3
		provide atmosphere	75	1.76	2
		attain wisdom	5	3.36	1
		enhance skill	215	4.83	3
		participate (in) activity	191	3.24	3
		do courses	121	-3.01	2
		bring honesty	16	1.05	2

Table 14. (cont'd)

	2	achieve (full) potential	176	3.29	***	3
		attend college	1430	5.68		3
		attend university	594	4.1		3
		choose (learning) term	29	-0.87		1
		choose condition	18	-1.37		1
		reach horizons	2	-1.58		1
		provide resource(s)	757	3.29		3
		give exam	43	0.53		1
		provide material	351	2.08		2
		manage lifestyle	4	1.31		1
		take on (a) job	132	1.89		2
		use technology	1966	3.03		3
		save time	2298	1.56		2
		save money	7197	5.43		3
		have (studying) hour	3349	-1.6		2
		follow routine	84	2.74		2
		adapt (to) schedule	6	1.18		1
		gift freedom	0	0		1
		bring benefit(s)	234	0.94		2
0017	1	take course	1722	0.35		2
		support viewpoint	31	2.92		2
		meet people	3279	1.55		2

Table 14. (cont'd)

	provide advantage(s)	109	1.28	2
	appreciate discussion	27	0.41	1
	attend classroom	311	3.51	3
	obtain skill(s)	24	1.3	2
	cultivate ability	12	2.82	2
	gain skill(s)	281	4.25	3
	provide resource(s)	757	3.29	3
	participate (in) course	32	-0.6	1
	provide comprehension	11	0.84	1
	answer question	11376	6.42	3
	clear (up) misunderstanding	38	5.96	3
	assist learners	30	5.09	3
	offer encouragement	145	5.77	3
	offer motivation	8	-0.1	1
2	take class	1180	0.63	2
	provide benefits	996	3.32	3
	make use	5623	0.02	2
	obtain knowledge	86	3.13	3
	obtain skills	24	1.3	2
	save time	2298	1.56	2
	use time	2384	-0.37	2

Table 14. (cont'd)

		foster growth	114	5.34	3
		rule out possibility	494	9.05	3
		provide books	57	-2.33	1
		provide resource(s)	757	3.29	3
		attend course	69	0.51	1
		provide comprehension	11	0.84	1
		receive instruction	406	4.92	3
		answer queries	55	4.8	3
		clear (up) ambiguity(ies)	11	3.66	3
		provide support	3761	4.02	3
		provide motivation	130	3.05	3
		provide advantage(s)	109	1.28	2
		overlook (the) value	12	1.68	2
		provide environment	483	2.88	2
0018	1	enroll (in) learning	2	1.04	1
		oppose idea	98	2.49	2
		attend class	311	3.51	3
		have influence	3546	0.51	2
		engage (in) discussion	264	4.43	3
		encourage people	1844	2.87	2
		boost motivation	13	3.68	3

Table 14. (cont'd)

	perform (academic)			
	achievement	1	-2.5	1
	neglect lesson	1	0.39	1
	alleviate problem	136	4.22	3
	provide opportunity	1940	4.31	3
	participate (in) lesson	15	0.87	1
	conduct experiment	141	5.22	3
	understand knowledge	105	-0.21	2
	impose regulation	28	3.43	3
	join college	21	-1.5	1
	have chance	18180	1.35	2
	broaden view	21	2.84	2
	improve skill(s)	871	5.02	3
	concur (with)			
2	viewpoint	2	4.16	1
	attend college	1430	5.68	3
	have impact	13042	1.5	2
	encourage child	478	2.32	2
	have chance	18180	1.35	2
	offer class(es)	304	1.97	2
	give understanding	242	0.83	2
	use Internet	1246	3.17	3

Table 14. (cont'd)

		use TV	151	-0.2	2
		evaluate progress	129	4.85	3
		attend university	594	4.1	3
		achieve goal	1739	6.33	3
		shape character	49	3.16	3
		get chance	7167	1.88	2
		hinder growth	65	5.85	3
		shape personality	13	3.26	3
		develop habit	105	3.52	3
		develop personality	60	2.6	2
		attend college	1430	5.68	3
		promote development	703	4.85	3
		seek education	103	1.32	2
0019	1	attend college	1430	5.68	3
		discuss view	28	-0.76	1
		give opinion	594	1.62	2
		enhance flexibility	16	4.28	3
		have connection	2250	0.03	2
		arrange timetable	0	0	1
		attain skill(s)	24	3.54	3
		have skills	3200	-0.42	2
		navigate material	3	-0.24	1

Table 14. (cont'd)

	use software	953	3.41	3
	have experience	8746	-0.02	2
	attend class	311	3.51	3
	have interaction	387	-1.7	2
	receive response	200	2.39	2
	bring experience	257	0.04	2
	gain experience	709	4.12	3
2	conduct experiment	141	5.22	3
	attain skills	24	3.54	3
	have skills	3200	-0.42	2
	navigate material	3	-0.24	1
	use website	229	1.23	2
	use app	430	2.74	2
	use software	953	3.41	3
	have experience	8746	-0.02	2
	enhance experience	349	4.06	3
	provide background	297	2.9	2
	conduct experiment	141	5.22	3
	watch video	2186	4.68	3
	gain understanding	1011	6.74	3
	have interaction	387	-1.7	2
	receive reply	79	2.67	2

Table 14. (cont'd)

		develop technology	468	3.44	3
		provide experience	504	1.31	2
		enroll (in) college	197	5.91	3
		bring advantage(s)	23	-1.26	1
0020	1	attend college	1430	5.68	3
		reap popularity	0	0	1
		attend classroom	311	3.51	3
		have opportunity	10357	1.05	2
		share doubt(s)	22	-0.64	1
		share misunderstanding	2	-0.2	1
		stimulate creativity	35	6.58	3
		boost confidence	213	6.85	3
		set up contact	15	-0.38	1
		provide flexibility	204	4.67	3
		provide cost	147	-0.11	2
		allow (a number of) user	243	2.7	2
		reduce fee(s)	68	2.61	2
		have freedom	1963	-0.05	2
		interfere (with) schedule	13	2.6	2
		join university	29	-1.34	1

Table 14. (cont'd)

	2	join college	21	-1.5	1
		gain popularity	103	5.83	3
		attend class	311	3.51	3
		bring benefit(s)	234	0.94	2
		provide flexibility	204	4.67	3
		provide cost	147	-0.11	2
		allow (a number of) user	243	2.7	2
		reduce expenses	162	4.15	3
		reduce fee	68	2.61	2
		have freedom	1963	-0.05	2
		disrupt schedule	10	2.96	2
		provide opportunity	1940	4.31	3
		stand chance	1780	4.51	3
		boost productivity	136	7.79	3
		boost confidence	213	6.85	3
		maintain contact	379	4.31	3
0021	1	attend university	594	4.1	3
		have program	1445	-2.62	2
		have course	122	-4.38	2
		have education	1704	-2.33	2
		acquire knowledge	483	6.46	3

Table 14. (cont'd)

interact (with)			
teacher(s)	38	1.79	2
have background	1573	-0.75	2
have ability	6237	0.65	2
earn money	2354	5.71	3
enroll lesson	0	0	1
take course	1722	0.35	2
have opportunity	10357	1.05	2
make skill(s)	46	-3.55	1
support (educational)			
institute	6	-2.38	1
gain knowledge	551	5.2	3
demand			
qualification(s)	3	1.7	1
have prospect	153	-2.01	2
create (a) generation	103	1.2	2
give environment	40	-2.24	1
have activity(ies)	145	-3.82	2
face challenge(s)	1284	4.97	3
learn (novel) thing	1736	1.33	2
alternate (educational)			
institute	0	0	1

Table 14. (cont'd)

2	have program	1445	-2.62	2
	join college	21	-1.5	1
	support learner	43	2.9	2
	acquire knowledge	483	6.46	3
	interact (with)			
	teacher(s)	38	1.79	2
	have background	1573	-0.75	2
	have ability	6237	0.65	2
	do job	21689	1.07	2
	enroll lesson	0	0	1
	take course	1722	0.35	2
	have opportunity	10357	1.05	2
	make skill(s)	46	-3.55	1
	support (educational)			
	institute	6	-2.38	1
	put (into) practice	952	1.28	2
	demand			
	qualification(s)	3	1.7	1
	demand skill(s)	24	1.41	2
	have prospect	153	-2.01	2
	create generation	103	1.2	2
	give environment	40	-2.24	1

Table 14. (cont'd)

		have activity(ies)	145	-3.82	1
		face challenge	1284	4.97	3
		learn (novel) thing	1736	1.33	2
		alternate (educational) system	0	0	1
		meet friend(s)	883	1.74	2
0022	1	gain traction	393	9.48	3
		reap benefit	869	8.62	3
		get access	1821	0.5	2
		spend time	24350	4.9	3
		place reliance (on)	44	5.13	3
		establish learning	33	2.03	2
		take advantage	12983	5.76	3
		allocate expenditure(s)	4	5.11	1
		endure impact(s)	1	-1.92	1
		exert effect(s)	119	5.22	3
		evaluate connection(s)	6	0.32	1
		build up relationship(s)	859	3.59	3
		accumulate experience	13	1.52	2
		replace people	131	-0.83	2
		retreat (into) shell(s)	10	4.43	3

Table 14. (cont'd)

	immerse (into) virtual world	35	3.41	3
	stand chance	1780	4.51	3
	boost (critical) thinking	2	-0.18	1
	have (good) grasp	669	1.25	2
	nurture student	0	0	1
	boost convenience	1	1.14	1
	cut down time	70	0.89	2
	cut down cost(s)	4713.386	1.82	3
	spend time	24350	4.9	3
	spend cost(s)	21	-2.21	1
	restate view	6	3.56	1
	give lesson	291	1.24	2
2	gain traction	869	8.62	3
	attract interest	142	3.38	3
	enroll (in) college	197	5.91	3
	hold view	416	1.44	2
	apply education	57	0.35	1
	allocate budget	16	4.43	3
	save money	7197	5.43	3
	support children	711	1.15	2
	utilize Internet	23	3.47	3

Table 14. (cont'd)

		take advantage	12983	5.76	3
		remove barrier(s)	278	5.8	3
		follow schedule	103	1.88	2
		engage (in) class	23	0.63	1
		understand lesson	43	-0.7	1
		lose opportunity	170	1.5	2
		increase risk	1555	5.22	3
		develop performance	68	0.99	1
		raise hand	1496	3.49	3
		express opinion	528	5.64	3
		have question	6837	0.04	2
		stand chance	1780	4.51	3
		apply knowledge	465	4.38	3
		gain insight	6.72	7.38	3
		nurture value(s)	7	1.56	1
		reap benefit	869	8.62	3
		restate view	6	3.56	1
		intensify education	4	1.23	1
		guarantee development	16	0.62	1
0023	1	have effect	12851	1.03	2
		pursue education	341	4.09	3
		complete training	181	4.05	3

Table 14. (cont'd)

		facilitate learner(s)	10	3.84	3
		improve skill(s)	871	5.02	3
		approach skill(s)	6	-0.25	1
		educate people	906	3.54	3
		educate skill	0	0	1
		get degree	1331	0.43	2
		bring impact	14	-2.71	1
		attend university	594	4.1	3
		have advantage	3509	0.35	2
	2	pursue education	341	4.09	3
		participate (in)			
		institution(s)	15	-0.01	1
		facilitate individual	27	1.48	2
		arrange time	105	0.35	2
		enhance productivity	96	6.62	3
		improve skill(s)	871	5.02	3
		approach skill(s)	6	-0.25	1
		educate skills()	0	0	1
		bring benefit(s)	234	0.94	2
		have effect(s)	12851	1.03	2
0024	1	stand chance	1780	4.51	3
		save money	7197	5.43	3

Table 14. (cont'd)

	spend money	8959	5.69	3
	improve infrastructure	123	4.37	3
	open (up) opportunity	170	1.35	2
	access education	82	2.47	2
	allocate money	117	5.04	3
	improve facility(ies)	66	2.15	2
	reward student(s)	40	2.23	2
	have access	14510	1.62	2
	attend education	78	1.45	2
	receive education	550	3.27	3
	provide experience	504	1.31	2
	provide interaction	35	0.31	1
	acquire experience	47	1.65	2
	acquire interaction	0	0	1
	involve experiment(s)	4	-0.22	1
	apply theory	151	2.86	2
	play role	13534	6.5	3
	put into function	0	0	1
	have (a) view	4352	-1.6	2
	offer skill(s)	48	0.26	1
	have place	5631	-0.89	2
2	receive education	550	3.27	3

Table 14. (cont'd)

		enroll (in) university	438	3.97	3
		bring benefit	234	0.94	2
		substitute method	4	1.27	1
		pick up pace	346	6.01	3
		eliminate cost(s)	47	1.29	2
		require cost(s)	47	-0.42	1
		offer course	2000	3.01	3
		operate course	13	-1.5	1
		gain knowledge	551	5.2	3
		teach theory	57	1.19	2
		master field	2	-1.05	1
		eliminate air (of interaction)	0	0	1
		reduce efficiency	83	3.93	3
		reduce skill(s)	5	-2.44	1
0025	1	enroll (in) education	37	3.46	3
		provide benefit	996	3.32	3
		offer advantage(s)	226	3.21	3
		achieve aim	91	3.3	3
		have freedom	1963	-0.05	2
		pick schedule	16	-0.63	1
		obtain degree	126	3.71	3

Table 14. (cont'd)

	find (a) job	3680	2.05	2
	use tool(s)	1607	3.57	3
	offer program(s)	498	1.85	2
	access program(s)	54	1.1	2
2	enroll (in) education	37	3.46	3
	attend college	1430	5.68	3
	attend university	594	4.1	3
	take course	1722	0.35	2
	provide opportunity	1940	4.31	3
	advance knowledge	138	4.53	3
	acquire skill(s)	561	6.69	3
	finish course	97	0.63	1
	place emphasis	600	6.71	3
	promote learning	239	4.77	3
	lack (practical) aspect(s)	3	-0.13	1

APPENDIX I: SEMI-STRUCTURED INTERVIEW QUESTIONS

Semi-structured Interview after the intervention

These questions are intended to ask for participants' opinions about their experience after five corpus training sessions. The interview may last about 30 mins.

1. What do you think about the five online corpus-based lessons?
2. What do you think about the assignment of creating your own bank of verb-noun collocations using virtual corpus in COCA?
3. What do you think are the benefits and drawbacks of using COCA and SKELL in your writing process?
4. What do you think are the benefits and drawbacks of using COCA and SKELL in giving feedback to your peers' essays?
5. What do you think are the benefits and drawbacks of using COCA and SKELL in self-correcting verb-noun collocations in your own essays?
6. What do you think about the benefits and drawbacks of doing peer-feedback and self-correction to your collocation acquisition and writing skills?
7. In general, what do you think about using COCA and SKELL to learn verb-noun collocations and write in English?
8. What are some advantages and disadvantages you had during five weeks of intervention?
9. Would you continue using COCA and SKELL in the future? Why?/ Why not?

The researcher will ask the participants for their permission to conduct the interview and record their answers. All audio recordings will be coded and deleted after the study is finished.

APPENDIX J: COLLOCATION KNOWLEDGE TEST RESULTS

Table 15.

Collocation knowledge test descriptives

Item	IF	ID
1	0.71	0.6
2	0.67	0.6
3	0.71	0.7
4	0.21	0.1
5	0.67	0.6
6	0.75	0.6
7	0.71	0.1
8	0.67	0
9	0.96	0
10	1	0
11	0.96	0.1
12	0.92	-0.2
13	0.67	0.7
14	0.67	0.4
15	0.63	0.4
16	0.58	0.2
17	0.67	0.5
18	0.46	0.5
19	0.67	0.7

Table 15. (cont'd)

20	1	0
Test average IF	0.71 (min 0.21, max 1.0)	
Test average ID	0.33 (min -0.2, max 0.7)	
KR-21 score	0.58	
Test score SD	3.6	
Test score average	14.25 (min 5, max 19)	
Test score SEM	2.16	