Liuqin Fang, Qing Ma* and Jiahao Yan

The effectiveness of corpus-based training on collocation use in L2 writing for Chinese senior secondary school students

<https://doi.org/10.1515/jccall-2021-2004>

Published online September 6, 2021

Abstract: Corpus tools are known to be effective in helping L2 learners improve their writing, especially regarding their use of words. Most corpus-based L2 writing research has focused on university students while little attention has been paid to secondary school L2 students. This study investigated whether senior secondary school students in China, upon receiving corpus-based training under the framework of data-driven learning (DDL), could improve their vocabulary use, especially the use of collocations, in their writing for the International English Language Testing System (IELTS) test. Twenty-two students aged 16–18 in a senior secondary school in Nanchang, China who were planning to take the IELTS exam participated in the study. Corpus of Contemporary American English (COCA) and Word and Phrase were the main corpora that the participants used to learn various search functions. Pre-writing and post-writing tests were administered to measure the effect of corpus training. In addition, a questionnaire and interviews were used to collect students' perspectives and attitudes. The results indicate that students made improvement in word selection after three corpus training sessions, and their attitudes towards corpus use were positive even though they were restricted from using computers to access corpora inside their school.

Keywords: collocation; data-driven learning; IELTS academic writing; secondary school students

1 Introduction

In view of rapid globalization, it is not surprising that the number of English language learners continues to increase as English is embraced as the

***Corresponding author: Qing Ma**, The Education University of Hong Kong, Hong Kong, China, E-mail: maqing@eduhk.hk

Liuqin Fang, The Gertrude Simon Lutheran College, Hong Kong, China, E-mail: liuqinfang07@gmail.com

Jiahao Yan, The Education University of Hong Kong, Hong Kong, China, E-mail: yjiahao@eduhk.hk

contemporary lingua franca. Meanwhile, many researchers have noted that vocabulary and grammar learning can be considered as the major focus of second/foreign language acquisition (Pan & Xu, 2011; Song & Chen, 2017; Zhang, 2011). Academic writing in English is an essential skill for students planning to enter university; however, writing in English is a major challenge for students in China because traditional teaching approaches emphasize correct grammar rather than well-structured and well-argued academic writing (Zhang & Luo, 2004). Consequently, many Chinese students are capable of recognizing grammatical errors but are deficient in writing academically (Cai, 2013; Zhang, 2018), particularly regarding word choice. This is because their academic vocabulary is limited and they are unfamiliar with how words in English are habitually juxtaposed or collocated (Nation, 2001). In the past decades, there is a rapidly increasing number of Chinese students taking IELTS. However, the failure to appropriately use the learned vocabulary knowledge in their academic writing is one of the biggest obstacles to achieve higher scores (Yao, 2014).

Johns (1991) promoted data-driven Learning (DDL) where English learners use language data sets, such as corpora, to search for semantic information, which has proved effective in learning collocations. Considerable research has been conducted and found out that corpus tools are effective in helping L2 learners improve their vocabulary use and collocation competence in academic writing (Chan & Liou, 2005; Charles, 2007; Kennedy & Miceli, 2010; Lee & Swales, 2006; Sun & Wang, 2003; Todd, 2001; Wu, 2016). However, most research on corpus-based L2 writing focused on university students, while there has been little research carried out in secondary school settings. Therefore, the present study aims to investigate the effect that corpus-tools' training has on secondary school students' academic writing with the ultimate aim of improving their IELTS performance.

2 English writing and vocabulary learning of Chinese students

Traditionally in China, students are encouraged to produce error-free sentences from elementary school to university, and teachers focus on delivering grammar rules and analyzing students' academic performance based on dictation and quizzes (Duan & Yang, 2016). Under this learning approach, students have limited opportunities to learn content-based knowledge where unknown words are taught through context. Despite the focus on vocabulary knowledge, previous research has revealed that Chinese students have difficulty in converting their receptive vocabulary to productive vocabulary in their writing (Wen, 2018; Zhai, 2016; Zhou,

2010); they also lack skills and strategies to write in English (Sang, 2017). Conveying a clear written message using appropriate words is a major challenge for Chinese learners in their academic writing (Clark & Yu, 2020). Chinese students may have acquired a passive knowledge of many lexical items throughout a long period of English learning, but they still fail to access this vocabulary knowledge to express their thoughts clearly (Li & Schmitt, 2009).

In recent years, a growing number of Chinese students have been studying abroad, most of them taking IELTS as a qualifying standard. It is estimated that 50% of IELTS test takers in Asia are from mainland of China. Among the IELTS's four language skills, Chinese students tend to get lower score in writing compared to scores in the other skills (British Council, 2020b). In the IELTS writing Task 2, test takers are required to write at least 250 words in a formal, academic style, which is difficult for many Chinese students (An, 2013; Hao, 2016; Mayor, 2006). Despite a large vocabulary and a solid foundation in grammar, Chinese students' scores on IELTS academic writing are below average (Yao, 2014). A common error occurred in IELTS writing is inappropriate collocation use (Futagi, Deane, Chodorow, & Tetreault, 2008).

To increase vocabulary competence, research reveals that a greater focus is needed on vocabulary learning, especially collocations (Conzett, 2000; Men, 2015; Webb, Newton, & Chang, 2013). Knowledge of collocations is important for fostering language learning (Nation & Webb, 2011); however, Chinese learner's writing competence has been found to be weak in this area (Hou, 2014; Zou, 2019). Zou (2019) summarized six major collocation types in Chinese student writing, namely, noun-noun, noun-verb, verb-noun, adjective-noun, verb-adjectives, and adjective-adverb, where more than half of the errors belong to verb-noun collocations. Hou (2014) also identified verb-noun as well as adjective-noun collocations as the two most frequent collocation error types. Zou (2019) explained that these collocation errors can be attributed to factors such as first language transfer and misuse of synonyms. Several studies have demonstrated that exposing students to concordance lines from corpora is a useful approach to improve their collocation performance (Li, 2017; Saeedakhtar, Bagerin, & Abdi, 2020; Vyatkina, 2017; Wu, 2016).

3 Data-driven learning

Beginning with the work by Johns (1991) and Sinclair (1991), the application of DDL has been advocated by scholars because corpora provide rich authentic examples regarding how words are used in real contexts. Gilquin and Granger (2010) also promoted DDL as a field with pedagogical value for language learners. DDL allows students to learn and internalize frequency patterns and contextual information

regarding the appropriate use of words. Under this approach, learners are described as “language detectives” (Johns, 1997:101) as they can actively detect how words are used in context. Research has revealed considerable benefits from using corpora as a pedagogical tool in L2 writing (Chang, 2014; Crosthwaite, 2017; Kennedy & Miceli, 2010; Yoon & Hirvela, 2004).

One of the most important roles of corpus use is error correction (Crosthwaite, 2017; Vyatkina, 2017; Yoon & Hirvela, 2004). It is found that corpus-led DDL instruction can help L2 students identify and correct errors in their writing (Crosthwaite, 2017; Li, 2017; Saeedakhtar et al., 2020; Todd, 2001; Vyatkina, 2017; Wu, 2016). Wu (2016) noted two broad types of collocation errors, (1) Adj. + noun and (2) Verb + noun, and found that Chinese learners with both high and low proficiency levels improved their collocation knowledge with the help of corpus tools. Li (2017) focused on verb-preposition collocations and found that Chinese postgraduates who had no previous knowledge of corpora displayed a significant improvement in using this type of collocation in their academic writing.

Furthermore, after being introduced to the basic functions of corpora, learners can use corpora as a reference/editing tool without receiving any comments or error correction from their teachers (Charles, 2012, 2014; Kennedy & Miceli, 2010; Yoon & Hirvela, 2004). Kennedy and Miceli (2010), for example, designed a longitudinal training program and found that their students were able to use various search functions when facing collocation difficulties, and subsequently built their own corpus for later consultation. Studies have also revealed that students hold positive attitudes and perceptions towards corpus use (Charles, 2012, 2014; Crosthwaite, 2017; Yoon & Hirvela, 2004). Charles (2012) investigated 40 English for Academic Purposes (EAP) students’ evaluations of using their self-created corpora to solve collocation errors and found that 70% of the students wanted to use corpora in their EAP studies over the long-term.

However, most research on DDL for improving students’ use of collocations and word choice in their writing was conducted at the university level or above (Charles, 2012, 2014; Crosthwaite, 2017; Kennedy & Miceli, 2010; Wu, 2016; Yoon & Hirvela, 2004). Boulton and Cobb (2017) conclude that “unfortunately, there is little [DDL] research with high school learners” (p. 357). Crosthwaite (2019), therefore, recommends more research on the application of DDL to pre-tertiary students. Because there are many Chinese secondary students who need to perform well on the IELTS for acceptance into overseas universities, the present study investigates whether secondary school students can correct collocation errors and use appropriate vocabulary in their IELTS-oriented writing after being trained to use corpus tools. These secondary students’ attitudes towards the corpus resources were also collected and analyzed. Thus, the following research questions are relevant:

- 1) In what ways does corpus training benefit secondary school students' collocation competence and lexical quality in IELTS academic writing?
- 2) How do senior secondary school students perceive corpus training, and what attitudes do they hold towards corpora use in EAP writing?

4 Methodology

4.1 Participants

Twenty-two grade eleven Chinese students volunteered to participate in the study. They were studying at a senior high school in Nanchang, Jiangxi, and wished to pursue higher education overseas. The participants were aged 16–18, and had been learning English as a foreign language for about 14 years. The average English proficiency level among the participants was 5.0 on the IELTS exam. For confidentiality, pseudonyms are used throughout the paper in alphabetical letters (from A to V). Normal ethical procedures were followed throughout the study.

4.2 Research instruments

In total, three instruments were designed for data collection in the current study: (1) one pre-test and one post-test, (2) one questionnaire (see Appendix II), and (3) interviews (see Appendix III) with one intervention study (corpus training) between the two writing tests. The two tests were compositions written by students before and after the intervention (three training sessions) both of which were assessed to investigate whether students were able to reduce their error frequency (and as a formative exercise).

4.3 Procedure

Students were firstly asked to complete the pre-test. Following Wu (2016), students were not allowed to refer to any tools so the researcher could evaluate students' baseline writing skills on collocation knowledge. Then, by examining errors in the students' pre-test scripts, the researcher designed the teaching materials for the corpus training sessions. Three corpus training sessions were then conducted (see Section 4.3.1) to demonstrate how corpora function as an alternative reference tool for academic writing. Following the training sessions, students then completed the post-test with guidance from the corpora to measure their change of collocation

knowledge. Apart from consulting COCA and Word and Phrases Concordance, students were also allowed to use dictionaries which are their habitual reference tools in writing. Finally, a questionnaire was administered (see Appendix II), and interviews were conducted to determine the students’ attitudes as well as their overall evaluation of corpus use.

The prompts for the two compositions focused on educational matters and were taken from the IELTS writing database on Task 2:

Pre-test prompt: “In the past, the role of teachers was to provide information. Today, students have access to wide sources of information. There is, therefore, no role of teachers in modern education. To what extent do you agree or disagree?”

Post-test prompt: “In the past, lectures were used as a way of teaching large numbers of students, but now with the development of technology for education, many people think there is no justification for attending lectures. To what extent do you agree or disagree?”

4.3.1 Training sessions

Three corpus training sessions (see Appendix IV) were designed to introduce corpus-based reference tools for students to consult as they wrote. Three major types of collocation errors were addressed in this study as shown in the literature review, namely adjective + noun, verb + preposition, and verb + noun. The collocations were addressed in the three training sessions (30 min each), and these errors were collected from students’ scripts in the pre-test. During the three training sessions, COCA and Word and Phrase Concordance were used as the main reference tools. Participants were taught to use various search functions from the two corpora and were asked to examine the correctness of collocations as well as to provide suggested answers to the extension tasks either in class or at home (see Table 1). The aim of the training sessions was to provide alternative tools for enhancing the students’ writing competence, i.e., using the *collocates* function of COCA and the *frequency list & analyze texts* function of Word and Phrase Concordance. Meanwhile, concordance lines from the corpora could provide students with rich authentic examples to discover language patterns (see Tables 1 and 2).

Table 1: Three examples of collocation errors for each corpus training.

	Examples of collocation errors	Types of errors	Suggested answers based on corpora
Q1	*Meet questions	Verb + noun	*Encounter questions
Q2	*Apply *to scholarship	Verb + Preposition	*Apply *for scholarship
Q3	*Suited way	Adj. + noun	*Suitable way

Table 2: The Display of concordances.

Verb + Preposition	Concordances
Apply to	<div>others (1971 : 11) . This also seems to apply to the villages of High Versilia . The " cats " of he extent to which each of 12 statements applied to their lives during the previous two weeks . E Kirby developed a kind of taxonomy that applies to these changing arts (dance , drama , music ,</div>
Apply for	<div>having served as a witness when Palmisano , applied for naturalization in 1899. sup118 Probi nts admit they are deportable , but then they apply for one of more forms of relief , such as s to reach Canada and preventing them from applying for political asylum once in the country.</div>

Each corpus training session was divided into three parts: lead-in activity (to check students’ comprehension and increase their awareness of the usage of collocations); hands-on corpus search (to show students how to consult corpora to check the correctness of collocations and to explore the synonyms of the target words); and extension tasks (to provide students with more opportunities to work individually). Each corpus training session focused on two search functions of COCA: (1) Collocation and (2) Synonyms. Some simple search functions catering to the learners’ current writing competence were designed in the teaching guide as students had no background knowledge of corpora (see Appendix IV for a sample of instructional materials design).

In IELTS writing, students are required to use different expressions to avoid repetition as well as to prevent monotonous-sounding writing. Therefore, synonym learning was chosen in the design of the teaching guides. The *collocates* function from COCA was introduced to help students examine collocation accuracy. The *frequency lists* function from Word and Phrase was presented to help learners generate language patterns, and the *analyze texts* function was an alternate way to search for synonyms. An inductive method was adopted for instruction to activate learners’ schemata to become familiar with corpora along with the searching functions. The three corpus training sessions were delivered by one of the researchers at a computer lab in a large classroom setting.

4.4 Data collection and analysis

4.4.1 Assessing the students’ writing

It is generally accepted that students’ progress can be adopted as a measure of effectiveness of a given teaching innovation. In this case, a pre-writing task (the

pre-test) before corpus training intervention and a post-writing task (the post-test) were conducted to measure learners' performance and progress using IELTS-oriented prompts. Four marking criteria were used in the assessment of the two tests: "task achievement, coherence and cohesion, lexical resource, and grammatical range and accuracy" (British Council, 2020a). In addition, the raters marked the scores on lexical quality separately based on the IELTS rubric with reference to British Council (2020a) (see Appendix I). Students' writing samples were independently rated by the one of the authors and one writing teacher in the school. Disagreements were resolved via discussion. Mean scores and standard deviation measures were generated.

4.4.2 Questionnaire and interviews

The questionnaire and the interviews were designed to measure learners' perceptions of and overall attitudes towards corpus usage after the training sessions. Two parts were included in the questionnaire: (1) Background information and an examination of students' prior knowledge of corpora and (2) An evaluation and investigation of their attitudes towards corpora use. Fifteen items from the evaluation part were designed with a 7-point Likert preference scale, ranging from "strongly agree" (7 points) to "strongly disagree" (1 point) (see Appendix II). Inspired by Yoon and Hirvela (2004) and Crosthwaite (2017), the questionnaire data were categorized into four dimensions: (1) Reflection and perception of corpus use after corpus training; (2) Evaluation of a corpus; (3) Learners' autonomy in choosing a corpus to consult when writing; and (4) Attitudes towards corpus use. Semi-structured interviews (Appendix III) were conducted to gain insight into students' attitudes towards corpora; most questions were taken from Chang (2014).

The interview data were analyzed to answer research question 2 regarding the learners' attitudes and overall evaluation of the corpus training and usage. Eleven of the 22 participants were interviewed individually in 10–15 min. The interview responses were recorded and then transcribed and translated by the authors. Each participant viewed the transcripts and approved of the contents. The interview data were then coded independently by two of the researchers and an inter-rater reliability of 85% was reached. The two sets of codes were then reduced to one set via discussion between the two raters.

5 Results

The results are mainly presented in two parts based on the two research questions: (1) The effectiveness of corpus training and its usage as measured by frequency of

errors and quality of writing and (2) The students’ overall attitudes towards corpora. The effectiveness of corpus training was divided into two categories: (1) Collocation error frequency and (2) Lexical quality scores.

5.1 Collocation error frequency

The effectiveness of corpus training and its usage was measured based on two aspects: (1) The frequency of collocation errors in the two writing tests and (2) The lexical quality scores in the two tests. The results are displayed in Tables 3–6. As the study took place in evenings during students’ self-study period and all students participated in the study on a voluntary basis, some students missed either the pre-test or the post-test for various reasons. There remained seven students who took both the pre-test and the post-test. So we decided to focus on the scripts of the seven students.

Table 3: Frequency of collocation errors in the pre-test and post-test ($n = 7$).

	Number of collocation errors in pre-test	Number of collocation errors in post-test
Mean	2.86	2.14
Standard deviation (SD)	1.57	1.77

Only a certain number of students’ data were used as they had a full data set. Due to practical constraints, some students’ data was missing.

Table 4: T-test of error frequency in collocation.

t-test	Mean	Variance	Observations	df	t-stat	t-critical	p
Pre-test	2.86	2.48	7	6	1.51	1.94	0.09 ^a
Post-test	2.14	3.14	7				

^a $p < 0.05$ means significant.

Table 5: Means and standard deviation of writing scores ($n = 7$).

	Pre-test	Post-test
Mean	6.86	7.50
Standard deviation (SD)	1.07	0.91

Only a certain number of students’ data were used as they had a full data set. Due to practical constraints, some students’ data was missing.

Table 6: T-test of writing scores.

t-test	Mean	Variance	Observations	df	t-stat	t-critical	p
Pre-test	6.86	1.14	7	6	-6.97	1.94	0.00 ^a
Post-test	7.50	0.83	7				

^a $p < 0.05$ means significant.

The mean error frequency of the seven students in their pre-test and post-test indicates that learners made a minor improvement of 0.72 in the post-test. However, the result of the t-test ($t = 1.94$; $p = 0.09 > 0.05$) showed that there was an insignificant effect on the frequency of errors between the pre- and post-writing tasks, possibly due to the relatively small sample size.

5.2 Lexical quality and writing scores

The mean writing score of the post-test (7.5) (Table 5) indicates that the students' quality of writing increased by an average score of 0.64 over the pre-test. In addition, the result of the writing scores of the t-test ($t = 1.94$; $p = 0.00 < 0.05$) suggested that there was a statistically significant difference between the pre- and post-writing tasks, which showed that learners made slight improvements in writing scores after three corpus training sessions.

5.3 Questionnaire

The questionnaire was divided into two categories, (1) The background information of students' prior knowledge in writing and (2) The evaluation of their attitudes towards the corpora. Students indicated in the survey that the e-dictionary was the most popular tool to consult when encountering any vocabulary problems in writing before receiving any corpus training. However, after receiving the corpus instruction, most participants expressed that they would seek help from the corpora to solve any problems they encountered.

The latter part of the questionnaire was divided into four dimensions (see Table 7). The results indicate that learners' overall attitude towards corpus training was positive, with a mean score of 5.87 out of 7. Their perceptions and reflections about corpus use after corpus training were positive with a mean of 5.88. This result indicates that students believed that the three corpus training sessions were useful for learning new words and collocations when writing academic essays. Students highly appreciated using the corpus (mean = 5.93) and their attitude

(mean = 5.94) towards the corpus was the highest. They expressed a strong willingness to continue using corpora, such as COCA and Word and Phrase in their future English learning and academic writing. However, learner autonomy (mean = 5.67), received a relatively low score. This is understandable since full learner autonomy in corpora use is likely to be developed after sustained usage.

5.4 Interview

Coding of the interviews generated two main codes – benefits and difficulties of corpus use. These are elaborated here with excerpts from the students.

5.4.1 Benefits of corpus use

5.4.1.1 Corpus as a beneficial reference tool in writing

The search functions of the corpora (COCA and Word and Phrase Concordance) motivated the students to actively infuse their search results into their academic writing, especially the functions, “collocates” and “analyze texts,” which they could consult to find appropriate synonyms to replace the target words. Most students claimed that the corpora facilitated and guided them to develop their

Table 7: Result of the questionnaire.

Self-regulating dimensions	No. of items	Sample survey items 7-point Likert-scale questions (1 = <i>strongly disagree</i> ; 7 = <i>strongly agree</i>) <i>N</i> = 22, <i>response rate</i> = 100%	Mean (<i>Max</i> = 7; <i>Min</i> = 1)	SD
Reflection and perception of corpus use after corpus training	4	e.g. Using the corpus is helpful for learning the usage of collocations	5.881.632	
Evaluation towards corpus	4	e.g. My confidence in writing in English has increased by learning about the corpus	5.931.567	
Learners’ autonomy	3	e.g. When I search for information in a corpus, I usually get the information that I need	5.671.676	
Attitudes towards corpus usage	4	e.g. After the three training sessions, I wish our teachers can introduce corpus more	5.941.401	
Overall			5.871.561	

writing abilities. They were more willing to use corpora to find suitable synonyms and to examine how to appropriately use words and collocations. Prior to the corpus training, learners tended to translate directly from Chinese when writing English essays as noted in literature (Zheng & Chang, 2014). For example:

The size of my vocabulary is small. If I encounter any questions in choosing which word to use in IELTS writing, I will think in Chinese first, and then I will translate it into English. (A)

I would look up the Chinese meaning in an electronic dictionary first and then choose a word that I had not used before in my writing so that my essay would be more academic. (B)

After being introduced to the concepts of corpora, many students showed interests in using corpora to help them overcome difficulties in collocating words. Students indicated that they appreciated the teaching guides from the corpus training sessions, which were practical and closely matched the needs of the IELTS writing topics for which they were preparing. Moreover, the use of collocation errors drawn from students' previous writing motivated them to find ways to solve their own problems. During the training aimed at collocations and synonyms, students enjoyed being given more opportunities to do individual work by using corpora to identify appropriate words to collocate with target words. Some students looked in their notebooks and used corpora to check whether the phrases they had previously written were correct.

IELTS candidates are also required to use a variety of words and expressions in the writing section of the test according to the marking criteria; this requirement motivated students to turn to corpora searching to decrease the monotony. Student C and E mentioned that:

I have benefited from Word and Phrase Concordance in finding synonyms. Meanwhile, I also used both Word and Phrase and COCA to find the words which can be juxtaposed with these synonyms. Furthermore, the collocation and synonym learning through Word and Phrase and COCA provided me a new learning opportunity for using different expressions to avoid repeating the same word and enhanced language accuracy, which improved the quality of my writing. (C)

Writing is my weak point. I always use one word to express my opinion in my writing, so corpora offered me a chance to look up the collocation and synonym usage which helped me improve my writing ability. (E)

5.4.1.2 Advantages of using concordance lines

Two advantages of using concordance lines were (1) The rich authentic examples provided for generating language patterns and (2) The opportunity to train how to summarize the usages of collocations. Specifically, summarizing requires students to generate different language patterns for two similar collocations, such as “apply

to” and “apply for.” Student B reported that summarizing required him to think critically, while student D stated that concordance lines are more efficient than dictionaries in terms of self-discovery:

Summarizing requires my deep knowledge of vocabulary. For example, when we had the training to find out the difference between “apply for” and “apply to” through the concordance lines, I was not able to figure it out without the help from the teacher. After I tried two to three times, I found out that the concordance lines provided me examples to analyze what kinds of words can be added after the target word. (B)

Concordance lines require me to self-discover the language patterns while dictionaries present all the patterns for learners. In this case, learners will just copy the phrases without knowing the usage. However, I can find the language patterns through concordance lines and use them in my writing which can make me feel successful. Thus, I think I can improve with the help of corpora. (D)

5.4.2 Difficulties students encountered in using corpus tools

5.4.2.1 Practical problems of corpus use in classrooms with short training sessions

The interview responses also indicated that it was difficult for learners to search for information through the corpora due to limited training. During the training session about verb + preposition collocation, the sheer abundance of concordance lines appearing in their second language overwhelmed the students, preventing them from promptly deciphering language patterns. Having only three 30 min training sessions was also insufficient for learners to be fully familiar with the search functions. The excerpts below illustrate:

It takes me a long time to search for one word because I am not familiar with the searching functions. (E)

We had never heard of corpora before, so it was hard for us to remember all the searching steps in such a short time. (C)

When I use these two corpora, it is easy for me to mix up the order of searching for words. For example, I always forget when I should change the order of the words to either the left side or the right side. Thus, I think it will be better if I can receive more corpus-training. (F)

5.4.2.2 Constraints of electronic device usage in classrooms

Due to the school policy, students’ use of electronic devices during class was restricted, which reduced the capacity of students to become familiar with corpora. Although some students were able to use their mobile phones to look up

collocations, the small screen proved to be a challenge for them to continue using corpora to search for words. It was also found that some students did not have access to corpora for their post-writing task. Therefore, the reference tool they could use was dictionary.

I did not use corpora because I cannot use laptops at school. I used my mobile phone to search for words, but the words and screens are too small. We are studying for IELTS, and it required handwriting. Thus, we are not allowed to use computers to write essays at school. (H)

I really appreciated the training sessions because I learned something new that can be applied to my academic writing. However, the sessions were too short, and our writing teachers never mentioned corpora before. So, I do not think I will use corpora at this moment, but I will use it later when I enter university. At that time, I will have my own laptop. Moreover, we are learning IELTS, which requires us to write essays in a written form. If I am learning TOEFL (Test of English as a Foreign Language), then I think it is definitely a good resource, and our teacher will allow us to use computers in school because TOEFL requires candidates to use the computer to write the essay. (G)

6 Discussion and implications

6.1 Benefits of using corpus tools to facilitate writing

The higher scores on IELTS writing task 2 show that the students improved their lexical use and tended to reduce errors in their writing after corpus-based training. According to the questionnaire, the students had barely heard of the concept of collocations and had no prior knowledge of corpora in English vocabulary learning. In this sense, the study achieved two aims, namely, to familiarize students with collocations and provide them with tools to improve their writing. Similar to existing studies on DDL (Crosthwaite, 2017; Li, 2017; Saeedakhtar et al., 2020; Vyatkina, 2017; Wu, 2016; Yoon & Hirvela, 2004), students made noticeable improvement on collocations and vocabulary use. We therefore conclude that Chinese secondary students can benefit from DDL in their language learning, especially their collocation competence.

This finding contributes to the application of DDL for international English learners. According to Futagi et al. (2008), collocation errors are not only common in Chinese IELTS examinees, but also frequently appear in academic writing by other L2 learners, including college students in Saudi Arabia and international students in American universities. Same as Futagi et al. (2008), our study highlights the need for learners to focus on collocation strings and distinguish synonyms based on their collocations. Smirnova (2017) also found that the experience

of using corpora proved effective for EFL learners in a Russian university. She claimed the use of corpora improved students' understanding of usage patterns, helped them correct collocation errors autonomously, and more importantly, increased their scores in the IELTS writing section. In addition, students in the current study expressed their willingness to use corpora when practicing for other tests, such as TOFEL and CET (College English Test). Therefore, it is suggested that non-native English learners refer to corpora to improve their writing competence to increase their scores in English tests.

The inductive thinking involved when searching linguistic items and looking for language use patterns in corpora helps language learners take charge of their own learning beyond the examples provided in textbooks (Crosthwaite, 2019). In the current study, the secondary students also showed the ability to correct errors on their own. Searching and examining language use in corpora provided them with the opportunity to correct writing errors. Crosthwaite (2019) suggests that the more pre-tertiary students interact with corpora data, the more they develop problem solving skills and autonomous learning, which are important learning strategies required after graduation. However, corpus-using strategies take time to develop. Charles (2014), for example, showed that 70% of students practiced for 12 months before getting accustomed to using corpora. Likewise, Yoon and Hirvela (2004) found learners needed a long time to develop their abilities in editing lexicogrammatical errors in the corpora. This explains why some students from the present study claimed that the limited number of training sessions prevented them from being fully familiar with various search functions embedded in the corpus websites. Early studies, Thurstun and Candlin (1998) and Sun (2000), revealed that students may not be confident in reading concordance lines in English. This concern was also evident in our study. Thus, it appears that if corpora training were to become a part of the English curriculum, students would require prolonged training to become more skillful and confident in using search functions and benefit more from corpus use.

6.2 Attitudes and overall evaluations of corpora

6.2.1 Student attitudes and overall evaluations of corpora

The results reveal that students held positive attitudes towards corpora use in academic writing. Even though they were not able to fully familiarize themselves with the search functions, they regarded corpora as beneficial alternative tools to better understand language patterns as well as to enhance their confidence and ability in writing. This finding is consistent with Yoon and Hirvela (2004), who

stated that learners were positive towards corpora as they regarded them as helpful reference tools not only for obtaining word usage but also for developing writing skills. Chang (2014) also showed that learners appreciated the benefits corpora provide, even though they had not yet developed a deep knowledge of corpus use.

There appear to be several reasons for the students' positive attitudes towards corpora usage as indicated by responses to the questionnaire. Apart from arousing students' attention to the lexical mistakes offered by the researcher during the training session, some students used corpora to check whether the previously written phrases in their notebooks were appropriate after the training session. The questionnaire responses implied that dictionaries were commonly used by students before corpus training sessions; however, after consulting corpora, they found that words from dictionaries were not always appropriate for certain contexts. Similar to Chan and Liou (2005), students were able to transfer collocation patterns in examples they learned during the lesson to new sentences they wrote outside of class. In addition, after training, many learners asked the researcher to teach them more search functions, such as "all forms" and "parts of speech" from the Word and Phrase Concordance (see Figure 1).

Figure 1: Search functions of Word and Phrase that students were interested to learn.

Students showed their appreciation for learning about collocations through corpora even though the training sessions were limited. Some students stated that using various expressions can enhance one's ability to use lexical resources. Most of the students were fully engaged in learning about collocations during the three training sessions, especially the hands-on activities, which encouraged them to further practice, leading to a highly positive attitude towards using corpora.

Nevertheless, some negative aspects were evident. Because self-discovery plays an important role in corpora usage, particularly for the pattern "verb + preposition," it was found that some students were unable to develop their own language patterns through concordance lines without any hints from their teacher even after discussions with their groupmates. Also, some students had a difficult time locating information similar to what they wanted when faced with such an abundance of concordance lines, which caused them to give up. These reasons probably explain why learners' autonomy was rated the lowest in the survey among all dimensions.

6.2.2 Technology issues related to corpora use

One salient factor that caused learners to have a negative evaluation of corpus usage was the unstable speed of the Internet. Due to the school policy of banning computers or electronic devices, the training sessions were conducted in the school's computer lab, where the speed of the Internet was relatively slow, thus reducing the students' practice time. A few of the students explained that the slow speed of the Internet was one possible factor inhibiting the effectiveness of the training. It prevented students from having more opportunities to practice search functions. In the early days of corpora usage, Sun (2000) also noted that network speed and stability are two major problems. This challenge thus remains.

Although students were generally positive about corpora usage, they had not yet gained the habit of using them. According to Charles (2014), developing a habit of using corpora means that students can continue to use them after the training is completed. However, because the students could access computers only either during the training sessions or at home on the weekend, the opportunities for students to consult corpora and develop a behavioral pattern of using them were limited.

One solution, noted by students, is to create a record of the collocations they searched from the corpora. Some students made a record of the suggested answers from their corpus search to replace their errors during the training, which they could later refer to for further study or review. Some of them also copied some concordance lines so that they could learn the difference between two phrases in their free time. This notetaking was especially important in the present context of their English learning because students were restricted from using electronic devices in the classroom. Further, teachers need to urge their schools to provide facility for students while encouraging students to be pro-active in their learning by keeping records of the new collocations they have learned.

7 Conclusion

The study explored the effectiveness of corpus training sessions in IELTS academic writing among secondary school students by assessing their ability to learn from a corpus-based intervention; afterwards, their attitudes towards the corpora were gathered. The results of this research shed positive light on the efficacy of corpus training for improving senior secondary school students' lexical usage. Specifically, the findings revealed the students collocated better in their writing after the intervention and had mostly positive views about using corpora. Moreover, the questionnaire responses suggest that students had ample opportunities to use new vocabulary or phrases in their weekly writing practice. Students also received adequate support from their writing teachers for further improvements, such as

grammatical errors and coherence. However, corpus tools are relatively new to Chinese teachers and students, and not all writing teachers are able to point out collocation errors in students' texts or even know about corpora, indicating a need for bringing corpora usage into language teacher training.

The findings also revealed some limitations and difficulties that the students encountered in using corpus tools. One key element that emerged from this study is that it is difficult for Chinese secondary school students to have access to electronic devices to use corpora. Students could only use computers during the training sessions and on weekends. Therefore, we recommend that schools provide students with computers after class to access corpora and create records of concordance lines for certain language items so that they can analyze these corpus data in free time. Although only three corpus training sessions were conducted, the findings revealed the positive effects of DDL for pre-tertiary students. Since learning corpus search functions and the analysis of concordance lines can be a very complicated process (Heather & Helt, 2012; Naismith, 2017), it is suggested that instructors provide adequate corpus training for students so that they can be more autonomous and confident in corpora use over the course of their long-term language learning.

In conclusion, the study showed that in general, secondary L2 learners can improve their writing skill and are positive towards corpora usage. However, due to practical constraints, only seven participants' full data sets were used in the quantitative scoring part of the study, which limits its generalizability; thus, the results are indicative only. The qualitative parts of the study gathered positive views from a larger group of students indicating good potential. Although the learners' writing scores improved, collocations and synonyms learned through corpora are still only one instructional element among a myriad of elements that go into L2 writing.

There are a few limitations in the study. First, this is a small-scale study that only examined limited data on student consultation of corpora as reference tools for academic writing. Second, due to some practical constraints caused by the school policy, some students did not make full use of corpora, which may have undermined the results of the study. Even though, our data is indicative of good potential of corpus use in enhancing secondary school students' collocation learning in academic writing. Future studies researching corpora training on language students can consider having a larger number of participants while including an expanded list of the type of collocation errors L2 students make.

Research funding: The article was funded by the CRAC project (Ref: 03AAB) by the Education University of Hong Kong and it was also supported by Project of Discipline Innovation and Advancement (PODIA)-Foreign Language Education Studies at Beijing Foreign Studies University (Ref: 2020SYLZDXM011).

Appendix I: Criteria for lexical quality dimension (adapted and edited from British Council, 2020a)

Band	Lexical Quality
9	<ul style="list-style-type: none"> – “uses a wide range of vocabulary with very natural and sophisticated control of lexical features; rare minor errors occur only as ‘slips’” – 1 or below collocation error
8	<ul style="list-style-type: none"> – “uses a wide range of vocabulary” – “fluently and flexibly to convey precise meanings” – “skilfully uses uncommon lexical items but there may be occasional inaccuracies in word choice and collocation” – “produces rare errors in spelling and/or word formation” – 2–3 collocation errors
7	<ul style="list-style-type: none"> – “uses a sufficient range of vocabulary to allow some flexibility and precision” – “uses less common lexical items with some awareness of style and collocation” – “may produce occasional errors in word choice, spelling and/or word formation” – 4–5 collocation errors
6	<ul style="list-style-type: none"> – “uses an adequate range of vocabulary for the task” – “attempts to use less common vocabulary but with some inaccuracy” – “makes some errors in spelling and/or word formation, but they do not impede communication” – “produce few collocation errors” – 6–8 collocation errors
5	<ul style="list-style-type: none"> – “uses a limited range of vocabulary, but this is minimally adequate for the task” – “may make noticeable errors in spelling and/or word formation that may cause some difficulty for the reader” – 9–12 collocation errors
4	<ul style="list-style-type: none"> – “uses only basic vocabulary which may be used repetitively, or which may be inappropriate for the task” – “has limited control of word formation and/or spelling; errors may cause strain for the reader” – 13–15 collocation errors
3	<ul style="list-style-type: none"> – “uses only a very limited range of words and expressions with very limited control of word formation and/or spelling errors may severely distort the message” – “errors may severely distort the message” – 16–18 collocation errors
2	<ul style="list-style-type: none"> – “uses an extremely limited range of vocabulary; essentially no control of word formation and/or spelling” – 19–21 collocation errors
1	<ul style="list-style-type: none"> – “can only use a few isolated words” – 21 or more collocation errors
0	<ul style="list-style-type: none"> – “does not attend” – “does not attempt the task in any way” – “writes a totally memorized response”

Appendix II: Questionnaire about using COCA & Word and Phrase Concordance in EAP writing (adapted from Yoon & Hirvela, 2004)

Background information

1. Gender: Male ____ Female ____
2. Country where you were born: _____
3. First language: _____
4. In general, do you like to try new things? Yes _____ No _____
5. How often do you write an English essay?
 - a. One time a week
 - b. Two times a week
 - c. More than three times a week
6. What resources do you use for English writing? (e.g., dictionary)

7. Had you heard about corpus before the corpus training? Yes ____ (see Q. 8) No (see Q. 10) ____
8. Had you used corpus before the corpus training? Yes ____ (see Q. 9) No (see Q. 10) ____
9. Which corpus did you use before? _____
10. Which corpus did you use for your second writing? COCA ____ Word and Phrase Concordance ____ or please provide the name of the corpus you used _____

Evaluations of using COCA & Word and Phrase Concordance

The following questions are regarding your opinions on using COCA & Word and Phrase Concordance. Please use the scale below to *circle* the response that most represent your perspectives.

1. Strongly disagree
2. Disagree
3. Somewhat disagree
4. No opinion
5. Somewhat agree
6. Agree
7. Strongly agree

1. The three corpus training sessions were helpful for my writing	1	2	3	4	5	6	7
2. The corpus is a complementary tool to help my English academic writing apart from using dictionaries.	1	2	3	4	5	6	7
3. Using the corpus is helpful for learning the usage of vocabulary	1	2	3	4	5	6	7
4. Using the corpus is helpful for learning the usage of collocations	1	2	3	4	5	6	7
5. Using the corpus improved my English academic writing skill	1	2	3	4	5	6	7
6. When I had problems in the second writing task, I searched for help in COCA	1	2	3	4	5	6	7
7. When I had problems in the second writing task, I searched for help in Word and Phrase Concordance	1	2	3	4	5	6	7
8. When I search for information in a corpus, I usually get the information that I need	1	2	3	4	5	6	7
9. I understand the purpose of using the corpus after the three training sessions	1	2	3	4	5	6	7
10. I want to continue to use corpus in the future	1	2	3	4	5	6	7
11. After the three training sessions, I wish our teachers can introduce corpus more	1	2	3	4	5	6	7
12. My confidence in writing in English has increased by learning about the corpus	1	2	3	4	5	6	7
13. I will continue to use COCA in English academic writing	1	2	3	4	5	6	7
14. I will continue to use Word and Phrase in English academic writing	1	2	3	4	5	6	7
15. Overall, the corpus is a very useful resource for my English academic writing	1	2	3	4	5	6	7

Appendix III: Semi-structured interview about using COCA in EAP writing (Chang, 2014)

Which corpus search function(s) do you think is(are) useful?

Why do you think COCA or Word and Phrase Concordance is or isn't helpful?

Did you encounter any difficulty in using COCA & Word and Phrase Concordance? If so, how did you overcome it?

After you learned about COCA & Word and Phrase Concordance, do you think it is useful when compared to a dictionary? Why?

In what ways does COCA & Word and Phrase Concordance benefit your academic writing? What difficulties did COCA & Word and Phrase Concordance help you resolve?

How long does it take for you to search one collocation in COCA & Word and Phrase Concordance?

How frequently do you use COCA or Word and Phrase when you were doing the second writing task?

Do you still want to continue to use COCA or Word and Phrase Concordance after the three training sessions? Why or Why not?

How important is it for teachers to introduce corpus in class? Why?

Appendix IV: Teaching guides for the corpus training session (verb + noun)

Step 1: Matching exercises

(Please match the following clauses into two complete sentences)

As soon as students meet	they will search for the information on the Internet.
It is possible to express	questions,
When students encounter	opinions
It is possible to tell	without hurting people.

Purpose: To check students’ comprehension and raise their awareness of the usage of collocations

Teacher Tips: Teachers can use the matching activity to check students’ knowledge of the usage of collocations.

Teacher Actions: Using some open-ended questions to ask students why they match these clauses together.

Step 2: Hands-on corpus search

- Show students how to process the vocabulary ‘meet’ and ‘encounter’ on the *collocates* function of COCA.

Meet

ListChartCollocatesCompare KWIC

meetWord/phrase [POS]

_nn*Collocatesnoun.ALL

+4321001234+

Find collocatesReset

☐ Sections

Texts/Virtual

Sort/Limit

Options

Encounter

ListChartCollocatesCompare KWIC

encounterWord/phrase [POS]

_nn*Collocatesnoun.ALL

+4321001234+

Find collocatesReset

☐ Sections

Texts/Virtual

Sort/Limit

Options

- Look at the *frequency* page to see whether “questions” can be juxtaposed after meet & encounter.

Meet

SEARCH	FREQUENCY	CONTEXT	ACCOUNT
SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP_1]			
1	CONTEXT	FREQ	
1	PEOPLE	632	
2	DEMAND	173	
3	MR	136	
4	STANDARDS	134	
5	STUDENTS	105	
6	CRITERIA	96	
7	STATE	92	
8	RECORD	91	
9	EXPECTATIONS	85	
10	WOMEN	75	
11	FRIENDS	70	
12	PRESIDENT	68	
13	MEN	61	
14	DR	55	
15	FACE	55	
16	REQUIREMENTS	52	
17	STUDENT	52	

Encounter

SEARCH	FREQUENCY	CONTEXT	ACCOUNT
SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP_1]			
16	OPPOSITION	8	
17	TROUBLE	8	
18	DATA	7	
19	HOSTILITY	7	
20	STUDENTS	7	
21	HOMING	6	
22	INFORMATION	6	
23	PATIENTS	6	
24	POLICE	6	
25	ROADBLOCKS	6	
26	SESSION	6	
27	CONDITIONS	5	
28	FAMILY	5	
29	MEMBERS	5	
30	MUSIC	5	
31	QUESTIONS	5	
32	RACISM	5	
33	SESSIONS	5	
34	WORDS	5	
35	ZONES	5	
36	BARRIERS	4	
37	DIFFERENCE	4	
38	PRELUDE	4	
39	WOMEN	4	
40	ZONE	4	
41	CLIENTS	3	

- Show students how to process the vocabulary “question” on the *collocates* function of COCA and to see what kind of verbs are commonly collocated with “questions”.

Questions

List Chart **Collocates** Compare KWIC

questions

 Word/phrase [POS]

_v*

 Collocates

verb.ALL

+ 4 3 2 1 0 0 1 2 3 4 +

Find collocates

Reset

☐ Sections ☐ Texts/Virtual ☐ Sort/Limit ☐ Options

SEARCH	FREQUENCY	CONTEXT	ACCOUNT
SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP_1]			
1	ASK	2039	
2	ANSWER	1923	
3	ASKING	1748	
4	HAVE	917	
5	ASKED	916	
6	RAISES	738	
7	ARE	714	
8	RAISED	674	
9	RAISE	668	
10	RAISING	481	
11	ANSWERED	353	
12	HAD	230	
13	WERE	214	
14	TAKE	191	
15	ANSWERING	171	
16	MICROD	161	
17	ANSWERS	158	
18	ASKS	153	
19	TAKING	123	
20	POSE	107	
21	ADDRESS	102	

Purpose: To show students how to use corpus to search and check collocations.

Teacher Actions: Walking around to check if students know how to search it on corpus. If not, provide individual help.

Step 3: Extension task

- First, students process the vocabulary “tell” and “express” on the *collocates* function of COCA.
- Second, check the frequency page of these two words and see if “opinions” can be collocated together.
- Third, process the vocabulary “opinion” on the *collocates* function of COCA and to see what kind of verbs are commonly collocated with “opinion”.

Tell

[List Chart](#)
[Collocates](#)
[Compare KWIC](#)

Word/phrase

Collocates

Express

The screenshot shows the KWIC interface with the following elements:

- List Chart** button
- Collocates** button (highlighted with a blue border)
- Compare KWIC** button
- express** in the **Word/phrase** field
- Word/phrase** label and **[POS]** label
- _nn*** in the **Collocates** field
- Collocates** label and **noun.ALL** label
- +** button
- 4** button
- 3** button
- 2** button
- 1** button
- 0** button (highlighted with a blue background)
- 0** button (highlighted with a green background)
- 1** button (highlighted with a green background)
- 2** button
- 3** button
- 4** button
- +** button
- Find collocates** button
- Reset** button
- Sections** button
- Texts/Virtual** button
- Sort/Limit** button
- Options** button

Tell

SEARCH	FREQUENCY	CONTEXT	ACCOUNT
SEE CONTEXT: CLICK ON WORD OR SELECT WORDS > (CONTEXT) > (HELP) >			New Help
1	US	2045	
2	US STORES	773	
3	US ME	440	
4	US	234	
5	US MO	180	
6	US TRADE	143	
7	US MAM	142	
8	US OFFICE	135	
9	US JONES	130	
10	US STUDENTS	122	
11	US KIDS	119	
12	US LBS	114	
13	US PRODUCTS	109	
14	US MODERN	98	
15	US REDUCE	90	
16	US PATENTS	87	
17	US SALES	81	
18	US DAD	78	
19	US MRS	78	
20	US CONGRESS	75	

Express

	SEARCH	FREQUENCY	CONTEXT	ACCOUNT
6	INTEREST	77		
7	CRUISER	75		
8	FEELINGS	70		
9	SCHTTE	69		
10	CONCERN	68		
11	ANGER	67		
12	BUS	67		
13	GRATITUDE	66		
14	ADROCKEY	61		
15	IDEAS	61		
16	FRUSTRATION	58		
17	LANES	57		
18	LANE	52		
19	GRINDING	51		
20	SUPPORT	50		
21	LOVE	48		
22	CONFIDENCE	45		
23	EMOTIONS	39		
24	CO	34		
25	SYMBALARY	34		
26	BUSSES	33		
27	DOUBTS	31		
28	PERMISSION	21		

Opinion

[List](#)
[Chart](#)
[Collocates](#)
[Compare KWIC](#)

Word/phrase [POS]

Collocates

☐ Sections
 ☐ Texts/Virtual
 ☐ Sort/Limit

	SEARCH	FREQUENCY	CONTEXT	ACCOUNT
6	☐ CONCURRING	11
7	☐ INQUIRY	10
8	☐ INFORMED	9
9	☐ HAVE	8
10	☐ SHARE	7
11	☐ SUBMIT	7
12	☐ SWAY	7
13	☐ WRITING	6
14	☐ WRITTEN	6
15	☐ RE	5
16	☐ EXPRESSED	5
17	☐ ARE	5
18	☐ BE	4
19	☐ EXPRESS	4
20	☐ POLARIZE	4
21	☐ THINK	4
22	☐ RECEIVED	3
23	☐ TURN	3

Step 4: Synonym learning

- Showing students how to find synonyms of “questiond” that are collocated with “encounter” and the synonyms of “encounter” that are collocated with “question”.

Questions

List Chart **Collocates** Compare KWIC

question

 Word/phrase [POS]

=encounter

 Collocates [POS]

+ 4 3 2 1 0 0 1 2 3 4 +

Find collocates

Reset

☐ Sections Texts/Virtual Sort/Limit Options

Encounter

List Chart **Collocates** Compare KWIC

encounter

 Word/phrase [POS]

=question

 Collocates [POS]

+ 4 3 2 1 0 0 1 2 3 4 +

Find collocates

Reset

☐ Sections Texts/Virtual Sort/Limit Options

Question

SEARCH		FREQUENCY		CONTEXT		ACCOUNT	
SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...] iWeb QUESTION
	CONTEXT		FREQ			
1		COMBAT		1			

Encounter

SEARCH		FREQUENCY		CONTEXT		ACCOUNT	
SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]							IWeb ENCOUNTER
	☐	CONTEXT		FREQ			
1	☐	DIFFICULTY		11	<div></div>		
2	☐	UNCERTAINTY		3	<div></div>		
		TOTAL		14			

Purpose! To show students some other possible ways to say the target collocations instead of keeping using one.

Step 5: Extension task (synonyms)

- Students search individually how to find synonyms of “opinion” that are collocated with “express” and the synonyms of “express” that are collocated with “opinion”.

Opinion

List Chart **Collocates** Compare KWIC

opinion

 Word/phrase [POS]

=express

 Collocates [POS]

+ 4 3 2 1 0 0 1 2 3 4 +

Find collocates

Reset

☐ Sections Texts/Virtual Sort/Limit Options

Express

List Chart **Collocates** Compare KWIC

express

 Word/phrase [POS]

=opinion

 Collocates [POS]

+ 4 3 2 1 0 0 1 2 3 4 +

Find collocates

Reset

☐ Sections Texts/Virtual Sort/Limit Options

Opinion

SEARCH	FREQUENCY	CONTEXT	ACCOUNT
SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...] iWeb OPINION			
	☐	CONTEXT	FREQ
1	☐	DEFINITE	7
2	☐	EXPRESS	4
3	☐	SAY	2
4	☐	SPECIFIC	2
5	☐	STATE	2
6	☐	PRECISE	1
7	☐	EXACT	1
		TOTAL	19

Express

SEARCH	FREQUENCY	CONTEXT	ACCOUNT
SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...] iWeb EXPRESS			
	☐	CONTEXT	FREQ
1	☐	OPINION	4
2	☐	BELIEF	3
		TOTAL	7

Step 6: Homework

Use COCA to check the collocations in the following sentences and to find a synonym(s) to replace them.

Although modern technology can provide lots of information, it will also **bring some disadvantages**.

I **said** some unpleasant **comments** about her work. (**Say comments**)

Did you **have** some **exercise** today? (**Have exercise**)

References

- An, Y. 安亚伦 (2013). [Study of repetition in Chinese IELTS candidates' English written texts with an angel of systemic-functional linguistics] 功能语言学视角下的中国雅思考生写作文本中的重复手法研究. Henan: Henan University.
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393.
- British Council. (2020a). *IELTS TASK 2 Writing band descriptors [public version]*. Retrieved from https://takeielts.britishcouncil.org/sites/default/files/ielts_task_2_writing_band_descriptors.pdf.
- British Council. (2020b). *Test taker performance 2019*. Retrieved from <https://www.ielts.org/research/test-taker-performance>.
- Cai, L. J. (2013). Students' perceptions of academic writing: A need analysis of EAP in China. *Language Education in Asia*, 4(1), 5–22.
- Chan, T. P., & 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.
- Chang, J. Y. (2014). The use of general and specialized corpora as reference sources for academic English writing: A case study. *ReCALL*, 26(2), 243–259.
- Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes*, 6(4), 289–302.
- Charles, M. (2012). 'Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes*, 31(2), 93–102.
- Charles, M. (2014). Getting the corpus habit: EAP students' long-term use of personal corpora. *English for Specific Purposes*, 35, 30–40.
- Clark, T., & Yu, G. (2020). Beyond the IELTS test: Chinese and Japanese postgraduate UK experiences. *International Journal of Bilingual Education and Bilingualism*, 1–19. <https://doi.org/10.1080/13670050.2020.1829538>.
- Conzett, J. (2000). Integrating collocation into a reading and writing course. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 70–87). Hove: Language Teaching Publications.
- Crosthwaite, P. (2017). Retesting the limits of data-driven learning: Feedback and error correction. *Computer Assisted Language Learning*, 30(6), 447–473.
- Crosthwaite, P. (2019). Data-driven learning and younger learners: Introduction to the volume. In P. Crosthwaite (Ed.), *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners* (pp. 1–10). London: Routledge.
- Duan, Y. H., & Yang, X. Y. (2016). Difficulties of Chinese students with their academic English: Evidence from a China-United States university program. *SFERC*, 2016, 45–51.
- Futagi, Y., Deane, P., Chodorow, M., & Tetreault, J. (2008). A computational approach to detecting collocation errors in the writing of non-native speakers of English. *Computer Assisted Language Learning*, 21(4), 353–367.
- Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching. *The Routledge handbook of corpus linguistics*, 359, 370.
- Hao, Y. (2016). *A contrastive study of lexical chunks in high- and low-score IELTS compositions* [Master's thesis]. Shanghai International Studies University, Shanghai.

- Heather, J., & Helt, M. (2012). Evaluating corpus literacy training for pre-service language teachers: Six case studies. *Journal of Technology and Teacher Education*, 20(4), 415–440.
- Hou, J. (2014). Analyzing collocation errors in EFL Chinese learners' writings based on corpus. *Higher Education of Social Science*, 7(1), 87–91.
- Johns, T. (1991). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, 10, 14–34.
- Johns, T. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A. Wichmann, S. Fligelstone, T. McNery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100–115). Harlow: Addison Wesley Longman.
- Kennedy, C., & Miceli, T. (2010). Corpus-assisted creative writing: Introducing intermediate Italian learners to a corpus as a reference resource. *Language, Learning and Technology*, 14(1), 28–44.
- Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25(1), 56–75.
- Li, S. (2017). Using corpora to develop learners' collocational competence. *Language, Learning and Technology*, 21(3), 153–171.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18(2), 85–102.
- Mayor, B. M. (2006). Dialogic and hortatory features in the writing of Chinese candidates for the IELTS test. *Language Culture and Curriculum*, 19(1), 104–121.
- Men, H. (2015). *Vocabulary increase and collocation learning: A corpus-based cross-sectional study of Chinese EFL learners* [Doctoral dissertation]. Birmingham City University, Birmingham.
- Naismith, B. (2017). Integrating corpus tools on intensive CELTA courses. *ELT Journal*, 71(3), 273–283.
- Nation, I. S. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nation, I. S., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle, Cengage Learning.
- Pan, Q., & Xu, R. J. (2011). Vocabulary teaching in English language teaching. *Theory and Practice in Language Studies*, 1(11), 1586–1589.
- Saeedakhtar, A., Bagerin, M., & Abdi, R. (2020). The effect of hands-on and hands-off data-driven learning on low-intermediate learners' verb-preposition collocations. *System*, 91, 102268.
- Sang, Y. (2017). Investigate the “issues” in Chinese students' English writing and their “reasons”: Revisiting the recent evidence in Chinese academia. *International Journal of Higher Education*, 6(3), 1–11.
- Sinclair, J. (1991). *Corpus concordance collocation*. Oxford: Oxford University Press.
- Smirnova, E. A. (2017). Using corpora in EFL classrooms: The case study of IELTS preparation. *RELC Journal*, 48(3), 302–310.
- Song, M., & Chen, L. P. (2017). A review on English vocabulary acquisition and teaching research in recent 30 years in China. *Science Journal of Education*, 5(4), 174–180.
- Sun, Y. C. (2000). *Using on-line corpus to facilitate language learning*. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages. British Columbia: Canada.
- Sun, Y. C., & Wang, L. Y. (2003). Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, 16(1), 83–94.

- Thurstun, J., & Candlin, C. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes*, 17(3), 267–280.
- Todd, R. W. (2001). Induction from self-selected concordances and self-correction. *System*, 29(1), 91–102.
- Vyatkina, N. (2017). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology* 20(3), 159–179.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91–120.
- Wen, Q. (2018). The production-oriented approach to teaching university students English in China. *Language Teaching*, 51(4), 526–540.
- Wu, Y. J. A. (2016). The effects of utilizing corpus resources to correct collocation errors in L2 writing – Students’ performance, corpus use, and perceptions. In S. Papadima-Sophocleous, L. Bradley, & S. Thouesny (Eds.), *CALL communities and culture – Short papers from EUROCALL 2016* (pp. 479–484). Dublin: Research-publishing.net.
- Yao, S. (2014). An analysis of Chinese students’ performance in IELTS academic writing. *The New English Teacher*, 8(2), 104–138.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13(4), 257–283.
- Zhai, L. (2016). A study on Chinese EFL learners’ vocabulary usage in writing. *Journal of Language Teaching and Research*, 7(4), 752–759.
- Zhang, B. (2011). *A study of the vocabulary learning strategies used by Chinese students* [Master’s thesis]. Kristianstad University, Kristianstad.
- Zhang, Z. (2018). Academic writing difficulty of Chinese students: The cultural issue behind Chinese and British academic writing styles. *Studies in Literature and Language*, 17(2), 118–124.
- Zhang, W., & Luo, L. (2004). Some thoughts on the development and current situation of college English teaching in China. *Foreign Language World*, 3, 2–7.
- Zheng, C., & Chang, C. E. (2014). Shi chuandi yiyi haishi xuanyao xingshi: Waiyu xiezuogongcheng zhong de xing-yi chanjie xianxiang [Communication or show-off: A study of the confusion in meaning-structure relationship in undergraduate English writing]. *Modern Foreign Languages*, 37(4), 513–524.
- Zhou, S. (2010). Comparing receptive and productive academic vocabulary knowledge of Chinese EFL learners. *Asian Social Science*, 6(10), 14.
- Zou, Q. (2019). A corpus-based study of verb-noun collocation errors in Chinese non-English majors’ writings. In *4th international conference on contemporary education, social sciences and humanities (ICCESSH 2019)*. Paris: Atlantis Press.

Bionotes

Liuqin Fang

The Gertrude Simon Lutheran College, Hong Kong, China
liuqinfang07@gmail.com

Liuqin Fang is an English teacher at Gertrude Simon Lutheran College. She received her Master’s Degree in TESOL from The Education University of Hong Kong and her Bachelor’s Degree in TESOL

Education from Brigham Young University Hawaii. Her research interests include corpus linguistics, data-driven learning (DDL), and English for academic purposes (EAP).

Qing Ma

The Education University of Hong Kong, Hong Kong, China
maqing@eduhk.hk

Qing Ma is an associate professor at the Department of Linguistics and Modern Language Studies, The Education University of Hong Kong. Her main research interests include second language vocabulary acquisition, corpus linguistics, computer assisted language learning (CALL) and mobile assisted language learning (MALL).

Jiahao Yan

The Education University of Hong Kong, Hong Kong, China
yjiahao@eduhk.hk

Jiahao Yan is a Research Associate and post-graduate student at the Education University of Hong Kong. He has taken part in a number of research projects and come up with many insight research outputs. His research interests include corpus linguistics, data-driven learning (DDL) and English for academic purposes (EAP).