ĐẠI HỌC HUẾ
TRƯỜNG ĐẠI HỌC NGOẠI NGỮ

NHIỀU TÁC GIẢ

KỶ YẾU
HỘI THẢO QUỐC TẾ

NGHIÊN CỨU LIÊN NGÀNH VỀ NGÔN NGỮ
VÀ GIẢNG DẠY NGÔN NGỮ LẦN THỨ VII

PROCEEDINGS THE 7th INTERNATIONAL CONFERENCE

INTERDISCIPLINARY RESEARCH IN LINGUISTICS
AND LANGUAGE EDUCATION

NHÀ XUẤT BẢN ĐẠI HỌC HUẾ

ĐẠI HỌC HUẾ
TRƯỜNG ĐẠI HỌC NGOẠI NGỮ

KỶ YẾU HỘI THẢO QUỐC TẾ
NGHIÊN CỨU LIÊN NGÀNH VỀ NGÔN NGỮ VÀ GIẢNG DẠY NGÔN NGỮ LẦN THỨ VII

PROCEEDINGS THE 7th INTERNATIONAL CONFERENCE
INTERDISCIPLINARY RESEARCH IN LINGUISTICS AND LANGUAGE EDUCATION

NXB ĐẠI HỌC HUẾ

TRƯỜNG ĐẠI HỌC NGOẠI NGỮ, ĐẠI HỌC HUẾ
UNIVERSITY OF FOREIGN LANGUAGES, HUE UNIVERSITY

NHIỀU TÁC GIẢ

KỶ YẾU HỘI THẢO QUỐC TẾ

NGHIÊN CỨU LIÊN NGÀNH
VỀ NGÔN NGỮ VÀ GIẢNG DẠY NGÔN NGỮ
LẦN THỨ VII

PROCEEDINGS OF THE
7TH INTERNATIONAL CONFERENCE

INTERDISCIPLINARY RESEARCH
IN LINGUISTICS AND LANGUAGE EDUCATION

HUE UNIVERSITY PUBLISHING HOUSE

Hue, 2022

.

MỤC LỤC

EFFECTIVENESS OF USING CORPUS AS A SELF-CORRECTION TOOL IN EFL WRITING

Nguyen Hoang Hanh An[1], Nguyen Vu Quynh Nhu[2]

[1,2]University of Foreign Languages, Hue University

Abstract

This study aims to exlore the effectiveness of corpus in suppoting students in their error self-correction in writing as well as their attitudes towards the corpus consultation. Ten participants, including third-year and fourth-year students, were trained to use searchable corpora to advance their products in the writing process. The participants were required to write six essays on specific topics and independently corrected their errors through corpus research. An analysis of error correction was employed as a quantitative approach, while the learning log and interview were used as qualitative data collection instruments. The findings from the study also showed that in spite of some drawbacks, the majority of participants still expressed their positive attitudes towards the use of corpus as a self-guided error-correction instrument to help them improve their writing skills as well as their writing confidence.

Keywords: Corpus, error self-correction, benefits, challenges, attitudes.

1. Introduction

The ability to produce effective English communication has grown essential for academic success in many countries. Of all the English skills, writing is considered the most difficult skill that English as a Foreign Language (EFL) learners are expected to acquire. One of the key reasons is that beyond acquisition of learning strategies, writing requires a high proficiency of the target language. Advanced learners still have great difficulty with native-like collocations and idiomaticity. As Kim (2009) stated, even when advanced students make few or no grammatical errors in their writing, "their work can still sound non-native-like or unnatural, due to the differences in frequency with which a certain linguistic structure is used or to unknown conventions of the particular genre in which they are writing" (p. 1). Hence, students and teachers need comprehensive and authentic language resources to provide information beyond what they find in textbooks or reference books.

The use of authentic language materials, which are texts, audio, videos, etc., that have been produced for native speakers, in foreign language teaching has been considered important for language learners. Authentic language materials have been widely applied in language classrooms because of the richness that authentic texts provide in terms of cultural and linguistic contents or the availability of materials that are relevant, suitable,

and interesting for specific groups of learners. Moreover, with the development of technology and information, language educators have recently taken advantage of technology to enhance the quality of pedagogy through easy access to a greater number of authentic language samples. With the advent of computers and the Internet, there is a growing interest in using corpus for educational purposes. A corpus is defined as a database collection of authentic instances of language use from a wide variety of sources (Yoon & Hirvela, 2004). According to Kim (2009), for EFL students and teachers, the use of a searchable data corpus offers an excellent resource to check whether their words are used appropriately. With the increased availability of technological devices, including computers, laptops and smartphones, in classrooms and at home, students now have more opportunities to access corpus tools freely while writing. This study, therefore, was conducted with the aim of investigating the effectiveness of corpus applied in EFL students' error self-corrections in writing and their attitudes towards it, since corpus is a useful tool but unfamiliar to most of the students at Hue University of Foreign Languages (HUFL). Students could use corpus as a support search tool to revise their essays and correct different types of errors by themselves. In addition, students' attitudes towards the use of corpus as self-correction tool on writing were also examined. Based on the above-mentioned research purpose, the study was carried out to answer the following research questions:

- What types of errors can students correct with the corpus?

- What are the students' experiences with the use of corpus as self-correction tool for their writings?

- What are the students' attitudes towards the use of corpus as self-correction tool for their writings?

2. Literature Review

2.1. Definition of corpus

As Sinclair (2005) defined, a corpus is "a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research." (p. 19). Koo (2006) simply defined corpus in modern linguistics terms as a large collection of language production that can be used to investigate lexis, syntax, text and discourse, regional differences, differences between learners and native speakers, and historical changes. It may consist of texts taken from numerous genres, including newspapers, journals, books or other speech productions. From the aforementioned, the term "corpus" can be described as a large collection of written and spoken texts sourced from various registers, "which can be found in a digital form stored in a vast databases" (Dobrić, 2009, p. 360). A corpus, whether for specialized or general purposes, may comprise from around a few thousand to

hundreds of millions of words of text. The availability of such a vast volume of verified data increases opportunities for empirical observation of linguistic phenomena at the word, clause, phrase, and textual levels of analysis (Francis, 1993).

2.2. Corpus use in error self-correction

Self-correction is necessary for students to push themselves to reach the goal of longitudinal learning and is an effective way to foster students' motivation for self-development, asss well as help to provide them with more opportunities to learn from their own mistakes. Several studies have been conducted to examine the effectiveness of corpus research as a self-correction tool for EFL students. Corpora are regarded to be valuable in L2 pedagogy because they provided learners with target language input, particularly information about the word or phrase frequency with which certain linguistics items are employed and the most popular word pairings, or collocations. The most frequently used linguistic items in authentic examples are often the most beneficial for learners, so that "it does provide a more solid basis than relying only on intuitions and accepted practice" (Biber & Reppen, 2002, p. 201) by allowing students to choose the most suitable word that goes with other words. Another useful application of corpus is to expose students to an enormous number of authentic texts in order to expand their understanding of the word functions in particular genre-specific contexts (Yoon & Hilvera, 2004) and study naturally occurring language for grammatical patterns, word usage, semantic and pragmatic features and textual discourse (Flowerdew, 2009), consequently producing more natural expressions in their essays. Corpus, by providing authentic examples and encouraging "noticing", or "awareness-raising", assists students of all levels in developing skills that are necessary for dealing with the lingusitic problems they encounter. Corpus work can promote learner autonomy in various forms, depending on the learner's level of language proficiency and autonomy. The use of concordances, particularly in the area of error-correction, has demonstrated tremendous potential (O'Sullivan & Chambers, 2006; Yoon, 2011). Increasing students' confidence in writing when using corpus is another effective application of corpus. By understanding numerous authentic examples from various context, students will be more self-assured to their decision of word choice, as well as more confident in lexical and grammatical accuracy in writing (Yoon & Hilvera, 2004; Luo & Liao, 2015).

While there are several benefits that corpora offer to support students writing and checking errors, some learners may have obstacles in acquiring necessary skills to take advantage of it, leading to their difficulties in query formulation and result interpretation. In addition, some concordance lines are particularly sophisticated and provide language data in formats that are difficult to inteprete. Learners may feel overwhelmed when facing such complex looking linguistic input (Yoon & Hilvera, 2004). Because of the difficulties in the result interpretation, learners may find that searching for the answer to correct their

errors in corpus is a time-consuming procedure. Similarly, Cobb (1997) commented on the burden of time necessary for students to become familiar with new technology like corpus.

A large body of research exists investigating students' response to corpus use. Some findings suggest that students have positive attitudes towards the application of corpus in L2 learning (Baghestani, 2011; Kim, 2009; Yoon & Hilvera, 2004). On the other hand, in some emperical studies, respondents expressed mixed attitudes, including negative reactions, towards corpus consultation (Chang, 2010; Xue-hua & Ming-ying, 2013). In Xue-hua and Ming-ying's study (2013), learners that had positive attitudes towards corpus found it useful for word use in context and in word comparison; in contrast, learners who have negative attitudes would find corpus time-consuming, as they were overwhelmed with too many analyses they had to make.Many researchers have been successful in examining the effectiveness of the corpus-based activities for students' improvement in language proficiency and their development in grammatical knowledge as well as vocabulary acquisition. Previous studies have also shown that students' attitudes towards corpus application to writing and fixing errors are generally positive. However, despite of the growth in corpus consultation in various regions around the world, there is a scarcity of study research on EFL students' experience and attitudes towards corpus use in Vietnam. Hence, this study is designed to address this gap by evaluating the effectiveness of corpus as a self-correction tool for EFL students in writing, the benefits and challenges as well as students' attitudes towards the corpus application.

3. Methodology

3.1. Research design

To address the three research questions, this study adopted a mixed-methods approach to investigate the effectiveness of corpus as an instrument supporting students' error self-correction in EFL writing. By integrating quantitative and qualitative data, the researcher would be able to gain a more comprehensive understanding of the research problem than either method alone. In this study, quantitative data consisted of an analysis of error correction, students' ratings of the level of difficulty of error types that can be fixed by using corpus, and a comparison of the difficulty levels reported by participants in writing their essays on week 1 and week 6. The qualitative data collection instruments consisted of learning logs and structured interviews, which played an important role in examining the EFL students' experiences on corpus use in revising essays and their reactions to the use of corpus on their error self-correction in writing.

3.2. Participants

The participants chosen for this research were 10 students at the English Department, HUFL while the study was conducted. There were 3 third-year students and 7 fourth-year students enrolled in different majors, including English Language Teaching, Interpretation,

English for Tourism, and English Language Teaching at Primary School. They have already finished the Writing 5 module in which they learnt how to write an essay; as a result, learning how to use corpus gives them an opportunity to evaluate whether corpus is a useful tool to help them enhance the quality of their essays. The majority of the participants (70%) considered themselves as intermediate-level writers, and others evaluated themselves as advanced-level or proficient-level writers. The computer skills of the participants ranged from basic to advanced level. Meeting these prerequisites about English competence and computer skills might enable the participants to use corpus properly. The students had no prior experience in using corpus to improve their writing skills or language competence before, which means the results of the study could be evaluated under more controlled conditions.

3.3. Procedure of the study

Two corpora that were employed as prefered choice for research tools were Corpus of Contemporary American English (COCA) and British National Corpus (BNC). COCA (www.english-corpora.org/coca/) was developed by Professor Mark Davies from Brigham Young University and initially launched on the Internet on February 20, 2008. It has become one of the largest freely-available corpus of English with a size of more than a billion words derived from spoken and written sources. BNC (www.english-corpora.org/bnc/) was originally created by Oxford University Press in the 1980s - early 1990s, and it contains 100 million words of text texts from a wide range of genres, which can be a good resource for EFL learners as it provides numerous number of examples of different usages. Both corpora possess multiple of useful functions and features that rendered great assistance in supporting students in their writing and error-correction. LIST function offered students a list of results of the searching words' frequency; CHART and SECTION allowed students to compare the language usage in different genres of texts; WORD function worked as a dictionary that provided much more information that students might not have known about the searching word; COLLOCATES and COMPARE enabled students to determine which collocate would co-occur with a given word; SYNONYMS feature was a useful tool for students in paraphrasing sentences to avoid repetition; POS TAG allowed students to specify the part of speech of individual words; and WILDCARD assisted students in fixing spelling mistakes or finding specific word that co-occured with others.

The participants were invited to join two phases of the research data collection process. In the first phase, corpus training was required to ensure the student participants use the corpus efficiently. Before beginning the study, the researcher arranged a training program for all participants. The researcher spent two days in one week introducing two corpora COCA and BNC to the participants so that they would familiarize themselves with the use of corpus and its functions. Each training session lasted for 90 minutes. In the first stage of training, the researcher introduced the basic information of two corpus platforms

that were applied in this study, namely COCA and BNC. The participants received the handouts of the introduction session during the training, and after the first training finished, they received supplementary documents introducing the use of COCA and BNC via email as well as some basic searching exercises for them to become more familiar with those corpora. In the second stage, a sample essay with highlighted error corrections was shown for the participants, and the researcher explained and demonstrated how to fix the errors by using COCA and BNC. By doing this, the participants could understand the necessary procedure to follow during the time the study was conducted. For the rest of the training, the participants were required to consult the corpus to revise another sample essay in which the error corrections were highlighted, and then the researcher would comment on their performance in using corpus on error-correction and advise them if they had any problems.

In the second phase, each participant was asked to write an essay each week for 6 weeks continuously. After receiving the participants' work, the researcher highlighted the errors in their first drafts and then returned them to the participants. After receiving the comments on their errors from the researcher, the participants corrected the highlighted errors by consulting the two corpora, and wrote down the correction or new words/phrases beside the errors in the original essay. After that, they handed in the second version of their essay to the researcher. The researcher kept a record of the errors that participants had committed in the first draft and the changes they had made in the second version. At the end of the experiment, though most of the students completed all of the essays, some of them missed one or two essays; therefore, the researcher collected 56 essays in total.

3.4. Data collection and analysis

To analyze how effectively participants employed the corpus to improve their writing, the researcher adopted the analysing method of the number and types of errors from Kim's research (2009). Each correction was rated with one of three possible symbols, namely + (plus), - (minus) and 0 (zero). + indicates that the error was improved with more natural language after being revised; - indicates that the errors became worse or less natural after revision; and 0 indicates that there was no change from the initial state of the error in question. The corrections were manually numbered and categorized into different ratings

In addition, all participants were required to maintain a learning log during the course of the study. Each week, the students were asked to fill in three columns of a chart after making their revisions, including the errors made in the first draft, type of the errors and the corrections made by consulting the corpus. Some guided questions were given to investigate participants' experience and attitudes towards corpus.

An text analyzer tool was also applied to rate the difficulty levels of language mastery of student texts. Participants' essays collected during the first week, when students had not been exposed to the corpus use, would be compared to the essays obtained in the

last week of the study, when participants had practiced using corpus and became familiar with it.

Lastly, interviews were conducted to allow students to give their opinions about corpus use in detail. The interviews were carried out via Zoom and Google meeting, and each interview lasted about 20-25 minutes depending on each participant. In order to keep track of the data, both audio-recording and note-taking were applied. The interviews were recorded and transcribed for qualitative data analysis. The respondents' answers to the interview questions were summarized and presented in the form of quotations when necessary to further clarify the students' experiences and attitudes towards corpus.

4. Findings

4.1. Types of errors that could be corrected with the use of corpus

Table 4.1: Types of errors that could be easily corrected with corpus

Types of errors	Word choice	Preposition	Spelling	Grammar	Word order	Register
Students' votes	70%	90%	20%	30%	90%	20%

Table 4.2: Types of errors that could be difficult to be corrected with corpus

Types of errors	Word choice	Preposition	Spelling	Grammar	Word order	Register
Students' votes	20%	0%	60%	50%	0%	40%

As was clearly seen in Table 4.1, the three categories that were selected as the most easily corrected errors by the participants were Preposition, Word order and Word choice. Preposition and Word order ranked first with 90%, followed by Word choice with 70%. Grammar, Spelling and Register, with only 30%, 20% and 20% respectively, were three options not many students perceived could be easily fixed.

On the other hand, in terms of types of errors that could be difficult to be corrected with corpus, as depicted in Table 4.2, more than half of participants (60%) believed that among six types of errors, Spelling was the most difficult error type to be fixed with corpus. Grammar and Register errors came after that with 50% for Grammar and 40% for Register. As a large number of students said that Preposition, Word choice and Word order were the most easily fixed errors, it was no surprise that only 20% of participants thought that it was difficult to use corpus to correct Word choice errors, and no one chose Preposition and Word order for this question.

4.2. Students' experience on the use of corpus in essay error self-correction

4.2.1. Improvement in error-correction

Table 4.3: Total number of errors in six categories in essays

Total	Word choice	Preposition	Spelling	Grammar	Word order	Register
264	139	43	17	45	7	13
100%	52.7%	16.3%	6.4%	17%	2.7%	4.9%

It can be seen from Table 4.3 that Word choice, Preposition and Grammar were the most frequent types of errors students committed in their writing. With 52.7%, Word choice was the type of errors that participants made most. Following Word-choice errors were Grammar and Preposition, which accounted for 17% and 16.3% of the total respectively. Spelling, Register and Word order made up less significant parts among the errors with 6.4%, 4.9% and 2.7% respectively.

Table 4.4: Total number of corrections in six categories in essays

Total corrections	**Type of errors**	**+**		**0**		**-**	
264	**Word choice**	116	83.5%	14	10,1%	9	6.4%
	Preposition	38	88,4%	4	9.3%	1	2,3%
	Spelling	17	100%				
	Grammar	41	91.2%	2	0.4%	2	0.4%
	Word order	7	100%				
	Register	12	92.3%	1	7.7%		
Total		231	87.5%	21	8%	12	4.5%

As shown in Table 4.4, 87.5% of the total errors in the collected essays corrected were able to improve the essays with more natural language, whereas only 8% of them showed no improvement and 4.5% became worse. Spelling and Word order were two types of errors that were 100% appropriately corrected by the participants. Word choice, Preposition, Grammar and Register errors were successfully rectified at a relatively high rate, ranging from 83.5% to 92.3%

Table 4.5: Number of errors needing corrections in six categories from Week 1 to Week 6

Total corrections	**Type of errors**	**Week 1**	**Week 2**	**Week 3**	**Week 4**	**Week 5**	**Week 6**
264	**Word choice**	35	28	28	20	16	12

Preposition	12	9	5	6	6	6
Spelling	5	4	3	2	2	1
Grammar	9	9	9	7	6	5
Word order	3	0	1	2	1	0
Register	3	2	3	3	0	2
Total	**67**	**52**	**49**	**40**	**31**	**26**

As illustrated in Table 4.5, the number of errors witnessed a sharp decline from the first week to the last week of the study. Using corpus continuously for 6 weeks significantly reduced the total number of errors in students' essays, from 67 to only 27. Over the same period, the number of errors in choosing the appropriate words and phrases also fell by one third (from 35 to 12), yet still accounting for the highest proportion. Meanwhile, as compared to the beginning, students' incorrect use of Preposition and Grammar was reduced by half, happening only 6 and 5 times respectively in the last week. In addition, a downward trend could be clearly seen in the number of Spelling and Word order errors, while the Register errors remained unchanged over the period of the study. The findings, therefore, reflected that students had nearly no difficulties correcting errors regarding Preposition, Word order and Word choice, which they perceived as the most easily corrected ones.

4.2.2. Benefits of corpus in error self-correction

A majority of participants in the interview attributed benefits of various contexts in corpus, so they could remember the meaning of words from several examples in different genres, rather than learning new vocabulary in isolation. One student said that owing to a variety of real-life examples from corpus, he could acquire more thorough understanding of the meaning of a word as well as its usage in different contexts. Another commented that learners could imitate the patterns of the target language derived from the authentic texts to a certain extent to produce more appropriate and native-like expressions. In addition, frequency in query results was a beneficial aspect of corpus that was mentioned by several interviewees when they were asked about the merits of corpus use in correcting students' essays. A student expressed, *"Frequency helps me realize that some phrases that seemed right to me were not popularly used by the native speakers."*, while another said that due to the frequency reported in query results, she felt more confident in the quality of her essays when using sentences that seemed to be more appropriate and commonly used in natural language.

Another remarkable and substantial benefit that corpus offered as a supportive instrument for students' writing is the various functions and features.

Table 4.6: Functions of corpus that are helpful for students in essay self-revision

Func-tions	Section	POS tag	Wild card	Syno-nyms	List	Chart	Word	Collo-cates	KWIC	Com-pare
Stu-dents' votes	20%	50%	40%	90%	80%	10%	80%	100%	50%	40%

With respect to beneficial functions that participants preferred using, COLLOCATES, SYNONYMS, WORDS and LIST were considered the most useful features for their self-correction in writing, ranging from 80% to 100% in approval. POS TAG and COMPARE were voted by half of the participants, whilst WILD CARD and KWIC were favored by 40% of participants. SECTION and CHART were least preferred with only 20% and 10% respectively.

Finally, corpus made remarkable changes in students' learning English.

Table 4.7: Comparison of CERF levels of essays in Week 1 and Week 6

Subjects	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
Week 1	B2	B1	B1	B2	B2	B1	B1	B1	B1	C1
Week 2	B2	B2	B2	C1	B2	B2	B2	B1	B2	C1

As illustrated in Table 4.7 there was significant improvement in the participants' writing competence. Specifically, 6 out of 10 students improved their writing skills thanks to the regular use of corpus, while the remaining showed no improvement. A respondent said, *"The corpus helps me improve and boost my writing skills a lot with more accurate language features in real certain contexts that I want to express, and that makes me feel more confidence in my writing."* Another participant commented, *"Before using corpus, my writing ability was limited mainly due to inappropriate word usage and low vocabulary level. Sometimes it is difficult to find synonyms to use to suit the style and context of the essay. After using corpus, my writing ability has changed markedly, that is, the overall level of vocabulary has been improved, sentences are produced more naturally by regularly reading contexts when using corpus.".* Secondly, during two months training and practicing using corpus to correct errors, some students reported that they had formed a new habit of using corpus to check their word choices when they were not sure about them or to avoid making mistakes in their writings. Finally, more than half of the participants in the interviews reported that they had become less reliant on teachers' feedback as well as more proactive in revising their errors on their own. One student admitted that she used to rely on teachers' error correction and explanation, and she would feel anxious and confused when correcting the errors in her essays because she was not sure about her correction. However, after being exposed to corpus, she was able to resolve the problems herself without the guidance from teachers and therefore felt more confident with her correction.

4.2.3. Challenges of corpus in error self-correction

One of the main problems that all the participants encountered during the course of the study was technical problems. The main technical problems that discouraged students from utilizing corpus in writing was unexpected blockage and errors for unknown reasons. *"This kind of interruption is really annoying and stressful sometimes."*, one respondent complained. In addition, the restricted number of searches of free accounts was considered an obstacle that prevented students from fully utilizing the corpus. One student commented: *"After 4-5 searches, there would be a message appearing, asking if you want a premium license, which was quite inconvenient and time-consuming to me. And 50 searches per days is really not enough, of course. If I want to make more queries, I have to pay to get the premium account, but that's way too expensive."*

In the learning logs, some students commented that it took them much time to use the corpus as a checking tool. *"The procedure to use the corpus is not so easy. When I looked up the dictionaries or typed a word in Google, it is often much easier and faster to get the results I wanted."*, said one interviewee. On the other hand, others argued that once they familiarized themselves with the corpus tool, they could use it to solve their writing problems in a short span of time. In addition, the overwhelming number of concordance examples could be a factor that affected the amount of time the students spent on corpus search. One student wrote in her learning log that when she clicked on a word she had chosen from the frequency list, she had to read an exhaustive number of example sentences to figure out the word pattern and the word usage, which sometimes distracted her from focusing on the particular language item that she wanted to search and made her frustrated. Another respondent shared her experience in her learning log, *"Some example sentences are too difficult to understand, so I have to find the easier one among the list of concordance sentences to understand the word meaning and its usage. So, I guess it's a little bit time-consuming and took me a lot of effort to understand the meaning of the authentic texts."*

Another major problem that the participants might face when using corpus in essay self-correction is the difficulties in result interpretation. Firstly, the most common problem that students dealt with when searching corpus was the high frequency word did not fit the students' intended ideas. One student said, *"Sometimes I thought that my word choice was right, but when I typed the word/phrase in the query box, there was no answer, if yes, its frequency was very low, so I had to find an alternative word. However, the word with higher frequency was not what I wanted to express."* Secondly, some students reported that it was sometimes difficult for them to understand the example sentences. One participant wrote in his learning log, *"Some authentic texts were really hard to understand, as there were too many new vocabularies as well as grammar structures in there, which made me lose my motivation for analyzing the texts"*. A large quantity of new words in corpus texts

that prevented students from grasping the whole idea of the concordance lines should be taken into consideration.

Lastly, the difficulties in carrying out a search in corpus might affect the students' experience with corpus as a self-correction tool in writing. In the first two weeks, the participants did not know how to search the errors in order to revise them or what they should type to get the results. One student said, *"Sometimes, even when I could recognize the error, but I didn't know how to use corpus to correct it."* However, at the end of the experiment, the number of students facing this problem has significantly decreased. Another problem that students had to deal with was the situation in which the corpus cannot execute a query of more than three words. "*I received no answer when searching some strings that contain more than three or four words, so in my opinion, the corpus is ineffective in this circumstance."*, a participant shared his experience in his learning log.

4.3. Students' attitudes towards corpus using

With respect to the participants' attitudes, most of them did not have positive attitudes towards corpus in the first two weeks, as they had never been exposed to this instrument before. One student even considered using corpus to be "*a complicated and time-consuming procedure*". However, at the end of the experiment, 7 out of 10 students believed that corpus was a supportive tool for revising their essays, as well as expressed greater interests in using corpus for error rectification. *"The more I use corpus, the more I like it."*, said an interviewee. In contrast, 3 participants did not show much fondness for corpus use. In terms of students' confidence in writing, two-thirds of the interviewees reported that they became more confident in L2 writing after experiencing the use of corpus for writing and error-correction. When asked whether they had tried applying corpus in other courses or for other purposes, the majority of students reported that they would use corpus to some degree to polish their academic works in other courses at university. Furthermore, a minority of students suggested that corpus could be used for translation purposes in an academic environment. The students were also enthusiastic about introducing corpus to their friends, as they believed that the rich source of authentic text and the various functions of corpus would help language learners develop their writing competencies. Finally, concerning the application of corpus into teaching and learning at HUFL, they all agreed that corpus should be taught in writing classes in university as this instrument is extremely beneficial for their language learning in tertiary education and students may depend on them for future reference.

5. Discussion and Implications

In relation to the easily-fixed error types, Preposition, Word order and Word choice were three types of errors that the students found easy to be corrected with the help of corpus. This attributed to the provision of a list of various parts of speech tags, by which

students could specify the precise word class of each collocate, and other supporting functions that aided students significantly in their correction. On the other hand, Spelling, Grammar and Register were considered error types that are difficult to fix when using corpus. Several reasons were given, including the spelling errors that corpus itself contained, the difficulties in forming queries and interpreting answers, and inadequate or incorrect prior knowledge to rectify errors independently without the guidance or explanation from instructors. This echoes the findings of Kim (2009) and Baghestani (2011).

Regarding to the improvement in error-correction, the result revealed that in the students' essays, Word choice, Grammar and Preposition were the most frequent types of errors that the students committed, whilst Spelling, Register and Word order errors made up less significant parts. This finding indicated that with the corpus use, the students could identify their errors and were able to correct them independently, which led to significant improvement in the quality of their writings. Additionally, when comparing the errors committed by the students in the first week to those spotted in the last week of the study, it was obviously seen that the number of errors dropped dramatically. The significant reduction in errors provided persuasive evidence that corpus may be applied to writing as an effective and useful instrument to aid students in essay self-revision. Therefore, corpus could be considered a beneficial tool for EFL students to improve their writings themselves as several types of errors can be corrected with the use of corpus. This finding showed correspondence to the findings from the previous studies (Lou & Liao, 2005; O'Sullivan & Chambers, 2006; Yoon & Hilvera, 2004).

The students could gain several benefits from corpus consultation in the process of error correction. Firstly, the authenticity of the language enables them to broaden their understanding of word meaning and its usage from several examples of various genres. Learners could explore varied sources of natural and authentic language from real-life contexts in corpus and could partially imitate the patterns of the target language generated from the authentic texts in order to produce more suitable and natural expressions. This result of the study is consistent with previous research (Flowerdew, 2009; Yoon, 2011; Yoon & Hilvera, 2004). Secondly, the students could check whether a word or phrase was frequently used by native speakers, indicating that language learners would gain a better understanding of how words are actually used and how to choose the most suitable phrases for their writing by using the frequency feature. They believed that seeing if a word or a phrase was low, medium, or high frequency could help them to determine whether it was acceptable and frequently used by native speakers. Thirdly, it was evident that the multiple functions and features that corpus possesses offered great assistance in supporting students in their writing as well as for self-correcting errors. By utilizing those beneficial features, students were able to form queries to determine if they had produced proper sentences, or whether the errors they had corrected were acceptable or not. Furthermore, by using corpus

to check errors, find more academic words and paraphrase terms to avoid repetition as well as organize their ideas more coherently, the students showed improvement in the writing quality, especially their collocational and prepositional knowledge. A number of empirical studies by various researchers also reported similar findings (Kim, 2009; O'Sullivan & Chambers, 2004; Quinn, 2015). Some students formed a new habit of using corpus to check their word choices when they were not sure about them or to avoid their mistakes in their writing. This supports the findings in Yoon's study (2011), which suggested that students used the corpus not only to address linguistics problems, but also to satisfy their curiosity about whether the corpus might provide better alternative expressions.

With respect to the challenges that the students encountered during the experiment, the first disadvantage that all the participants suffered from was related to technical problems. The main obstacle that prevented students from smooth corpus browsing was the unexpected blockages for unknown reasons that made students feel stressed and frustrated. Additionally, the limited search times per day and the interruption of advertisement after 4-5 searches were also the problems students encountered when they signed in the two corpora with free accounts, which might impede them to extensively exploit the corpus consultation. Secondly, a number of students considered the query formulation and the interpretation of concordance data in corpus to be effort- and time-consuming procedures. The respondents remarked that they did not feel so confident to decide what to type in the query box to revise their errors. This might be due to the insufficient time to practice and get familiar with corpus use. Also, the overwhelming number of concordance examples could be a factor that affected the amount of time the students spent on corpus search. A large quantity of new words in corpus texts that prevented students from grasping the whole idea of the concordance lines should also be taken into consideration. The findings echoed the studies of Baghestani (2011) and Bridle (2019). Therefore, more training and practice for the application of corpus into checking errors should be provided to help the students get involved in using corpus. Moreover, the most commonly occurring issue that the students experienced when exploring corpus was that a high frequency word did not fit their intended ideas. In this case, when the students are seeking for a specific search term and the initial inquiry does not yield the desired result, they may try to alter their search to find a similar structure by modifying their search term. Congrad (2001) suggested that writers should not be entirely reliant on the frequency of corpus but should be creative in their writing. Lastly, the difficulties associated with doing a search in corpus might have an effect on the students' experience with corpus as a self-correction tool for writing. It was challenging for students to find a suitable expression from a collection of various authentic texts. Students sometimes did not know how and what to query to get the results, or they could not execute a query of more than three words. In this case, other resources, such as Google, would do a better job in helping students search for the desired information.

Overall, the participants' attitudes towards corpus use were positive with a vast majority of students believing that corpus was a useful and efficient tool to help them revise their essays. There was a considerable change in students' attitudes from skepticism to optimism during 6 weeks practicing using corpus in their revision processes, which appeared to be mainly a result of their greater familiarity with corpus consultation. With the various concordance lines with authentic materials, they could explore various aspects of language and improve their essays independently, which resulted in their boosting confidence in writing competencies. The findings were similar to what had been found by Yoon & Hilvera (2004), Kim (2009), and Bridle (2019). As they had experience the effectiveness of corpus in detecting and rectifying errors in compositions, the students agreed that they would apply it regularly in the future and also introduce this helpful instrument to their friends. Moreover, as they had valuable experiences with corpus consultation, students suggested that corpus should be taught and trained in writing classes to help students gaining numerous benefits that corpus could offer in enhancing their language proficiency in tertiary education.

6. Conclusion

The use of corpus remains a relatively new concept at many academic institutions; however, recent research appears to indicate that corpus can contribute a great deal to improving students' error self-correction in writing. This includes the present study on investigating students' experience on corpus consultation, proving evidence for the benefits of corpus use in error correction in EFL writing. As shown through the students' feedback, corpus was a versatile and effective tool for the students in supporting their error correction as well as L2 acquisition with its multifunctional features and its provision of authentic materials. It also fostered students' language awareness as well as facilitated greater autonomy of language learners. Despite its positives, the disadvantages that the students had to experience when using corpus were inevitable, including technical issues, time availability, and difficulties in result interpretation as well as in query formulation. Those challenges could be overcome with students' practice and development of language proficiency over time.

The present study provides information on the students' experience on the benefits and challenges when applying corpus in their errors self-correction process and how they react to it. This paper, however, involved the participation of a very small number of subjects of a local university in Vietnam. Therefore, similar studies with a larger group of participants should be supported in future research. Other academic institutions can carry out research to constitute nationwide and worldwide initiatives to pursue corpus application for student academic development.

References

Baghestani, S.P. (2011). *Addressing language errors in L2 students' writing: Can corpora help?*. Master thesis. American University of Sharjah.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use.* Cambridge: Cambridge University Press.

Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition, 24*(2), 199–208.

Bridle, M. (2019). Learner use of a corpus as a reference tool in error correction: Factors influencing consultation and success. *Journal of English for Academic Purpose, 37,* 52-69.

Chang, J.Y. (2010). Postsecondary EFL students' evaluations of corpora with regard to English writing. *The SNU Journal of Education Research, 19,* 57-85.

Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System, 25*(3), 301–315.

Congrad, S. (2001). Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly*, *34*, 548–560.

Dobrić. N. (2009). Corpus linguistics - the basic form of linguistic analysis. *Language, Literature and Identity,* 359-363.

Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics, 14*(3), 393–417.

Francis, G. (1993). A corpus driven approach to grammar—principles, methods and examples. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology* (pp. 137-156). Amsterdam/Philadelphia: Benjamins.

Kim, Y.J. (2009). *Effectiveness of on-line corpus research in L2 writing: Investigation of proficiency in english writing through independent error correction.* Master's thesis. University of North Texas, American.

Lou, Q., & Liao, Y. (2015). Using corpora for error correction in EFL learners' writing. *Journal of Language Teaching and Research, 6*, 1333-1342.

O'Sullivan, I., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of second language writing, 15*, 49-68.

Quinn, C. (2015). Training L2 writers to reference corpora as a self-correction tool. *ELT Journal*, *69*(2), 165–177.

Sinclair, J. (2005). Corpus and text - basic principles. *Developing linguistic corpora: A guide to good practice*. Oxford: Oxbow Books.

Xue-hua, A., & Ming-ying, X. (2013). An empirical research on DDL in L2 writing. *US-China Education Review*, *3*(9), 693-701.

Yoon, C. (2011). Concordancing in L2 writing class: An overview of research and issues. *Journal of English for Academic Purposes, 10,* 130-139.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing, 13*, 257-283.

HIỆU QUẢ CỦA VIỆC SỬ DỤNG KHỐI NGỮ LIỆU NHƯ MỘT CÔNG CỤ TỰ SỬA LỖI TRONG VIẾT TIẾNG ANH

Tóm tắt

Nghiên cứu này được thực hiện nhằm mục đích khám phá hiệu quả của việc sử dụng khối ngữ liệu trong việc hỗ trợ sinh viên tự sửa lỗi trong bài viết cũng như thái độ của họ đối với việc tham khảo khối ngữ liệu. 10 sinh viên tham gia vào thực nghiệm, bao gồm sinh viên năm 3 và năm 4, đã được hướng dẫn cách sử dụng khối ngữ liệu nhằm cải thiện bài viết của họ trong quá trình viết. Những sinh viên này được yêu cầu viết 6 bài với các chủ đề khác nhau và sau đó tự sửa lỗi của họ thông qua khối ngữ liệu. Phân tích sửa lỗi được sử dụng như phương pháp định lượng, trong khi nhật ký học tập và phỏng vấn được sử dụng làm công cụ thu dữ liệu định tính. Kết quả nghiên cứu chỉ ra rằng mặc dù gặp phải một số khó khăn khi sử dụng khối ngữ liệu, đa số những sinh viên tham gia vẫn bày tỏ thái độ tích cực đối với việc sử dụng khối ngữ liệu như một công cụ tự sửa lỗi giúp họ cải thiện kỹ năng Viết cũng như tăng thêm sự tự tin khi viết.

Từ khóa: Khối ngữ liệu, tự sửa lỗi, lợi ích, khó khăn, thái độ.