Correcting Language Errors in EFL Writing by the Use of COCA

Chi-An Tung*
Shu-Ying Chang
Fang-May Peng

Feng Chia University
Taiwan

ABSTRACT

This study explores the use of the Corpus of Contemporary American English (COCA) by a cohort of intermediate EFL writers to see if, and to what extent they used corpus consultation to mitigate language errors, the problems they encountered in the process, and how they reacted to the use of COCA. This study draws some conclusions based on the data collected from twelve EFL writers' written drafts before and after corpus consultation, their records of corpus inquiries, interview transcripts, and a survey on the use of the corpus. First, frequent users of COCA seemed to correct lexicogrammatical errors via corpus consultation while the less frequent users of COCA tended to rely on other resources for such error correction. Second, most students had problems with identifying the error types, analyzing, and interpreting the concordance output. However, with more training on the use of the corpus and more time allowed to correct errors, these problems seemed to be ameliorated. Finally, despite the negative responses from the less frequent users, the majority of the students were positive toward the use of the corpus. In light of these findings, some pedagogical implications and suggestions are provided.

KEYWORDS: COCA, error correction, EFL writers

*Corresponding Author

Introduction

For decades, sentence-level errors in L2 writing has been a problematic area to student writers and teachers. Though some researchers like Krashen (1981) and Truscott (1996) believed that grammar was unteachable, grammar correction has been proved effective and beneficial to L2 writers (e.g. Chandler, 2003; Ferris, 1995, 1997; Ferris & Roberts, 2001). However, in the process-based L2 writing, language error correction was ignored, for L2 writers engaged themselves more in improving content and organization. In the 1990s, the rise of corpus data-driven learning (DDL) (Johns, 1991) shed some light on enhancing the quality and sustainability of sentence-level error correction in L2 writing, for its nature and functions were found beneficial to L2 writers. Some empirical studies have corroborated the following benefits. It could boost the student writers' language awareness, decrease error frequency (Biber, Conrad & Reppen, 1998), improve their command of lexicogrammatical rules and patterns (Boulton, 2009; Conroy, 2010), cultivate their critical thinking and problem-solving skills (O'Sullivan, 2007), promote discovery learning, develop learner autonomy (Chambers, 2005), and improve their writing products (e.g., Hegelhemier, 2006; Liu & Jiang, 2009; Yoon, 2008; Yoon & Hirvela, 2004). In addition, the teacher can gain more time to address the global and discourse aspects of L2 writing as the responsibility of learning sentence skills is transferred to the students (Johns, 1991).

Due to these benefits, corpus consultation has been adopted as a part of L2 writing curriculum design to tackle the problem of language errors. Some studies used concordance feedback in the domain of L2 writing for a general purpose (Chambers, 2005; Liu & Jiang, 2009; Yoon & Hirvela, 2004), while some practiced corpus consultation in the domain of L2 for an academic purpose (Gaskell & Cobb, 2004; Lee & Swales, 2006; Sun, 2007; Yoon, 2008). However, to date, most of these studies were conducted in an ESL setting, and the ones set in an EFL environment were mostly EAP-oriented and targeted at more proficient L2 writers. There is a paucity of research investigating how intermediate EFL writers use corpus consultation to mitigate lexicogrammatical errors, what problems they encountered in the process, and how they reacted to such a process. Therefore, this study aims to address these concerns by comparing the drafts produced before and after corpus consultation, and analysing the questionnaires on the use of COCA and interview transcripts.

Literature review

DDL, corpus consultation and L2 writing

In the late 1980s, corpus development with a new interest in exploiting authentic data induced DDL, a term coined by Tim Johns (1991), aiming at promoting language learners' acquisition of lexicogrammatical knowledge through corpus consultation. As Johns (1991, 2002) explained, in the process of corpus DDL, learners were exposed to actual language use as language items or as patterns shown in multiple real-life contexts. Then they actively explored the available samples of authentic language and detected the patterns themselves. In the following two decades, some studies on corpus consultation, grounded in DDL, found that language learners not only acquired lexical and syntactical knowledge more effectively (Boulton, 2009; Conroy, 2010) but also gained a variety of benefits as stated earlier.

Despite the benefits, a few concerns were also noted in the literature of corpus use and L2 writing. They were: a) Who will benefit more from corpus consultation, the more proficient learners or the comparatively low ones? b) Can learners be trained to use corpus technology, and if yes, to what extent? c) What are learners' attitudes toward corpus use in language learning? In the past, it has been found that corpus consultation may not be for all, as those with less language proficiency were slow

in finding appropriate words or patterns to execute their intended meanings, and confined their activities to error correction and text revision (Aston, 2001). For those who had difficulties or were less familiar with inductive learning, analysing the results was "tedious, time-consuming and laborious" (Chambers, 2005, p. 120). Corpus consultation seemed to be more for those who "prefer unstructured, discovery-oriented learning" (Bloch, 2007, p. 187). Some studies (see Yoon, 2011) corroborated this inference that consultation benefited advanced L2 learners more, and in line with this belief, quite a few studies were targeted at advanced learners (Lee & Swales, 2006; Sun 2007; Yoon, 2008). However, Yoon and Hirvela (2004) found that given longer training of corpus use, intermediate ESL learners could benefit from corpus use as much as their advanced peers. Gaskell and Cobb (2004) noted that although low-intermediate ESL learners did not make overall writing improvement, there was some significant reduction in some of the error types, and the learners seemed to enjoy using concordancing as a learning tool. In addition, Chang and Sun (2009) found that given scaffolding prompts—procedural, elaborative and reflective prompts, Taiwanese senior high students improved their proofreading performance on verb + preposition collocation.

These studies seem to suggest that it is not merely an issue of who will benefit more from corpus consultation. What matters is how the training in relation to corpus consultation is conducted and for how long. A few researchers (Boulton, 2010; Gaskell & Cobb, 2004; Stapleton & Radia, 2010; Yoon, 2008; Yoon & Hirvela, 2004) have endorsed this proposition.

As to the third concern, to date, most studies on corpus consultation, and language teaching and learning have found that most ESL learners, given guided and gradual training, gave positive feedback to such an approach (e.g., Conroy, 2010; Geluso, 2011; Lee and Swales, 2006; O'Sullivan & Chambers, 2006). However, in EFL classrooms set in Asia, voices for corpus use were mixed with complaints. It was probably due to the conventional model of Asian pedagogy which was much teacher-centered and discouraged inductive learning (Smith, 2011). In addition, it was probably the gaps between the students' expectations of the task and the performance that exhausted the students (Smith, 2011). It was likely that their limited L2 proficiency or lack of some vocabulary knowledge inhibited them from generating lexicogrammatical rules and patterns (Liu & Jiang, 2009). The factors involved are plenty and complicated, so further exploration is needed to see whether and to what extent EFL writers in Asia use corpus consultation to tackle their language problems in EFL writing.

New trends of grammar instruction, and corpus consultation in EFL writing

In the 2000s, the debate of grammar teachability waned gradually as grammar instruction was marching into a new direction. From the perspective of corpus linguistics, Conrad (2000) noted that with the advancement of computer technology and a rejuvenated interest in grammar instruction, "three changes prompted by corpus-based studies of grammar have the potential to revolutionize the teaching of grammar" (p. 549). The three changes are register-specific description of English grammar, the integration of vocabulary teaching and grammar teaching, and a shift from structural accuracy to the appropriate conditions of language use. The three changes have their roots in focus-on-form grammar instruction which stresses the importance of raising learners' grammatical consciousness and directing their attention to the forms and meanings in the contexts (Conrad, 2000; Liu & Jiang, 2009). Conrad's observation was later endorsed by Liu and Jiang (2009) who proposed a unified approach, an integration of the three trends, to grammar instruction in an L2 writing classroom. They found the approach viable in enhancing learners' language awareness and better command of lexicogrammatical rules and patterns. In addition, the learners appreciated the importance of contexts in lexicogrammatical choices.

It seems that the suggested unified approach would be a highly viable means to correct language errors in L2 writing. However, Smith (2011) voiced a few caveats when constructing a learner corpus

for general English learning in Taiwan: a) Inductive learning might be a problem for Taiwanese teachers and EFL learners, and b) the forms of gap-fill error correction exercises might demotivate these learners for serendipitous learning because they have been trained to complete similar exercises in their prior learning about vocabulary and grammar. To some extent, Smith's caveats are true. Inductive learning and independent learning are less promoted in an EFL classroom in Taiwan. Some studies have recorded complaints from both students (Sun, 2007; Smith 2011) and teachers (Sun & Wang 2003) that corpus consultation was too time-consuming. In addition, most of the EFL learners in Taiwan receive little critical thinking training (Tung, Chang & Peng, 2013; You, 2001), so they need to develop cognitive skills to process induction. Furthermore, in an EFL writing classroom, the student writers usually depend on the teacher as a more reliable source to correct their local errors (Liou & Peng, 2009; Yang, Badger & Yu, 2006). Therefore, it is a great challenge for these EFL writers to take charge of the following actions all by themselves: identify the errors, use the concordancer, analyse the concordance output and generate grammatical patterns or usages of words.

To sum up, these problems seem to mitigate the viability of using corpus consultation to solve language problems in L2 writing. However, these problems should serve as caveats, not deterrents, because some of these problems were presented from a perspective of integrating corpus use into a general English class (Smith, 2011) and some were based on the studies conducted in the EAP domain (Sun, 2007) where the subjects were usually advanced learners. There is a dearth of research investigating the viability of tertiary students' use of corpus consultation in an EFL writing class. Therefore, the research questions posed are as follows:

1. Can corpus DDL help intermediate EFL writers correct the local errors (i.e. sentence-level errors) in writing? If yes, to what extent?
2. How do intermediate EFL writers conduct the corpus consultation process when revising texts?
3. What are the writers' attitudes toward corpus consultation for L2 writing?

Method

Participants and settings

The subjects were 12 English majors taking English Composition IV at a private university in central Taiwan. The students were English majors from an intact class and they were all female, aged from 19 to 21. All of them, except a transfer student and a repeater, have received training in peer review and critical thinking in English Composition III in the previous semester. The twelve students had been learning L2 writing skills for at least three semesters. In general, most of them were considered intermediate L2 learners as evidenced from their written products and their reported TOEIC or General English Proficiency Test (GEPT) scores. The class met once a week for two hours. The students had fair computer proficiency, but they had no knowledge of language corpora or their pertinent use or functions.

Corpus use training and procedure

The corpus used for this study was COCA, a large database consisting of 450 million entries derived from spoken and written sources developed by Brigham Young University. It was chosen partly because of its large size and availability as a free corpus and partly because of the vast written sources which seemed appropriate for a L2 writing course. The introduction of COCA started after the students were familiar with the conventional requirements of argumentative essay writing. The students were trained to get familiar with the basic functions of COCA. They were given instructions on, and demonstrations of, how to conduct a concordance search and interpret the concordance output.

Then, in the following two weeks, they were advised to use COCA to revise the lexical or grammatical errors underlined by the teacher in their first drafts. Meanwhile, they were welcome to use the tutorial hours to consult the teacher on how to correct these errors. Later, before they submitted their second drafts, a list of frequent lexical and grammatical error types was introduced and distributed to students for their reference to identify error types by themselves. After their submission of the second drafts, the teacher conducted a general class discussion, focusing on the lexical and grammatical errors commonly shared by most of the students in their second drafts, and the teacher demonstrated how to use corpus consultation to correct the errors. The discussion continued for another week. Finally, students submitted their final drafts along with the records of their corpus search. The period of the introduction of COCA and practices lasted for five weeks, with ten in-class hours and four tutorial hours. To get a general feeling of the students' responses to the use of COCA, a survey on the use of COCA was conducted, and seven students consented to be interviewed.

Instruments and data collection

The instruments used for this study consisted of the following: a) students' work, including corpus search results and their written drafts of the argumentative essay; b) transcripts of the semi-structured interviews; and c) a survey on the use of COCA. A total of 24 written drafts, including the second and final drafts, were collected for analysis. The first drafts were not collected because they were usually produced in free writing style and they were usually greatly revised after the writers elicited comments on content and organization. The second and final drafts were shuffled for blind review. In these drafts, the underlined lexical and grammatical errors were coded by error types. The first two researchers of this study coded three drafts randomly chosen and then had a discussion on how to build up the coding system. After the system was built, both coded half of the drafts (12) and the inter-rater reliability was satisfactorily high (r=.95). Then the first researcher proceeded with coding the remaining 12 drafts alone.

Each interview lasted about 15 to 20 minutes. Interviewees were first shown their second and final drafts, and they recalled the process of their corpus use, the problems they encountered, and the solutions, if any, they found, and stated their general feelings toward the corpus. The interviews were conducted in Chinese, audio-recorded, transcribed, and translated. The survey consisted of two parts, and it was based on the one used by Yoon and Hirvela (2004). However, in the first part on biographical information, nine out of the 23 items were removed for they did not match the setting of this study. Age was added as an item, so the items in this section totalled 15. In the second part, all the 42 items were retained, but seven were rephrased to reflect the setting of this study. The participants were asked to indicate their degree of agreement on a scale of 1-7 instead of a 6-point scale with an option of N (no opinion) as used in Yoon and Hirvela's (2004) questionnaire. The questionnaire was translated into L1 and proofread by the researchers. The Cronbach alpha value of item-total correlation of the questionnaire was .94.

Results and discussion

Students' written drafts and corpus search results

The second and final drafts of the argumentative essays and the search results were compared to gain some information on how the use of COCA contributed to local error correction. A system of classification of errors was established with reference to the previous taxonomies established by James (1998), Chan (2010), O'Sullivan and Chambers (2006), Murrow (2005), and Darus and Subramaniam (2009). Because of the wide range of error types and the complexity of some errors, the error types were classified into 19 types to best reflect the common types of errors found in this EFL writing class. These types were, in nature, lexical and grammatical errors. Discourse errors as

Chan (2010) found among her Hong Kong ESL writers were not considered in data analysis of this study.

Nineteen error types and numbers of the errors pertinent to these types are summarized in Table 1. Most of the error types were identified in the previous taxonomies, except the error types –*L1 expressions*, and *fragments*. In this study, L1 expressions were defined as a clause or phrase that was phrased in Chinese syntax, and the multiple errors in it inhibited understanding the meaning without the help of L1. For example, this error type included sentences like "Besides, city kids are easy to move house," (EM-2nd draft). Fragments included such errors as "Then won't have enough time to play with their friends" (AN-2nd draft).

In this system, the corpus search was also classified into three types in terms of their contribution to error correction. If the errors shown in the second drafts were corrected in the final drafts due to corpus search, they were marked as "positive changes." If the errors still existed, they were marked as "negative changes." "Unidentifiable changes" was marked if the original errors were deleted due to idea rephrasing or sentence reconstruction in the final draft.

Table 1. Error types and use of coca by twelve student writers

Error types	Errors					COCA		
		Final drafts[1]				Positive changes	Negative Changes	U[2]
	2nd draft	New	Er-ctd	Old	U			
Articles missing or wrong use (a, the)	24	4	18	6	0	7	0	0
Verbs (tense/missing/serial verbs)	23	5	16	3	4	5	2	2
Nouns (plural or singular)	16	8	6	6	4	5	2	0
Wrong word choices	13	9	3	8	2	2	0	0
Prepositions (missing or wrong use)	9	2	6	1	2	3	2	2
L1 expressions (Chinese English)	6	3	3	1	2	1	0	0
Adjectives	5	3	4	1	0	3	1	0
Infinitives, gerunds, participles	5	3	3	1	1	1	0	0
Misuse of Vt/Vi	5	1	1	3	1	0	0	0
Parts of speech	5	3	2	2	1	1	1	0
Fragments	4	5	0	3	1	0	0	0
Subject-verb agreement	4	5	3	1	0	0	0	0
Conjunctions (missing or wrong use)	3	3	0	3	0	0	0	0
Mechanical errors –misspelling	3	2	0	3	0	0	0	0
Pronoun	3	1	3	0	0	0	0	0
Run-on	3	1	0	3	0	0	0	0
Non-parallel structure	2	0	1	0	1	1	0	0
Dangling modifiers	1	0	1	0	0	0	0	0
Mechanical errors	1	0	1	0	0	0	0	0
Total	135	58	71	45	19	29[3]	8	4

Notes:
1. In the final draft, old errors refer to the errors found in the second draft and still existent in the final draft. New errors mean the errors that were not found in the second draft but emerged in the final draft due to global or local revisions. Er-ctd refers to the errors that have been corrected and disappeared in the final drafts. U refers to those errors found in the second draft but it is hard to identify whether they have been corrected due to content revision or sentence reconstruction.
2. COCA U means the changes are unidentifiable due to sentence reconstruction. The original errors cannot be found for a comparative study.
3. The 41 corpus inquiries include 29 inquiries that led to positive changes and eight that resulted in negative changes. Four unidentifiable inquiries cannot be classified into either category because the writers decided to revise the idea and change the sentence structure.

By examining the numbers of errors by error type, it was found that the errors made in relation to articles, verbs, nouns, word choices/inappropriate words, and prepositions ranked the top five, and they accounted for 63% and 50% of the total errors made in the second and final drafts. It is important to note that these error types were not confined to a few individual students, but shared by most of the students. Furthermore, it was found that among these five error types, numbers of three error types--articles, verbs and prepositions—were greatly reduced in the final drafts. Out of 24 article errors made in the second drafts, 18 were corrected; 16 out of 23 verb errors and six out of nine preposition errors were corrected

To see whether the corpus enquiries were related to the errors corrected, students' search results were examined. It was found that 22 out of a total of 41 (54%) corpus enquiries were made in the top five error types—articles, verbs, nouns, wrong word choices and prepositions. In addition, 22 out of 29 (76%) of the corpus enquiries induced positive changes in these error types. These might indicate that

corpus consultation helped EFL writers correct the errors in these five types. To verify this inference, the number of the corpus enquiries made by the students was thoroughly checked. It was found that the corpus consultation seemed to be limited to half of the students. Among the 12 students, six used the concordancer more frequently than the other six. Therefore, this study further examined the written drafts and the search results of these six frequent users. Out of the 69 errors made by the frequent users in the second draft, 38 errors were corrected. In addition, the search results showed that 21 corpus inquiries conducted by these six users contributed to positive changes (Table 2). Therefore, the accountability of corpus consultation for error correction in this group was 55% (21/38). In contrast, in the less frequent users group, out of 66 errors in the second drafts, 33 were corrected. Eight corpus enquires were made, and they all led to positive changes; therefore, the accountability rate was 24% (8/33).

Table 2. Error types and use of COCA by the six frequent users and the six less frequent users

Error types	Frequent Users			Less Frequent Users		
	2nd draft	Er-ctd[1]	COCA	2nd draft	Er-ctd	COCA Pt-changes
			Pt-changes[2]			
Articles missing or wrong use (a, the)	18	15	6	6	3	1
Verbs (tense/missing/serial verbs)	9	4	4	14	1 2	1
Nouns (plural s or singular)	11	3	3	5	3	2
Wrong word choice	7	1	2	6	2	0
Prepositions (missing or wrong use)	5	3	1	4	3	2
L1 expressions (Chinese English)	2	1	0	4	2	1
Adjectives (wrong use of compound, missing, wrong comparative)	3	3	2	2	1	1
Infinitives, gerunds, participles (wrong use of participle as S.)	3	2	1	2	1	0
Misuse of Vt/Vi	1	0	0	4	1	0
Parts of speech (use v. as a n. or a n. as a v.)	2	1	1	3	1	0
Fragments	1	0	0	3	0	0
Subject-verb disagreement	1	1	0	3	2	0
Conjunctions (missing or wrong use)	1	0	0	2	0	0
Mechanical errors –misspelling	0	0	0	3	0	0
Pronoun (disagreement in number and reference, missing)	1	1	0	2	2	0
Run-on	0	0	0	3	0	0
Non-parallel structure	2	1	1	0	0	0
Dangling modifiers	1	1	0	0	0	0
Mechanical errors—lower or upper case	1	1	0	0	0	0
Total	69	38	21	66	33	8

Notes:
1. Er-ctd refers to the errors that have been corrected and disappeared in the final drafts.
2. COCA Pt-changes refers to the positive changes made in the final draft due to the corpus search.

By further examining the corpus enquiries made by the frequent users, it was found that five of the seven negative changes were confined to student ST (pseudonym) who had problems with identifying and interpreting the errors. Therefore, it seems that these problems had a negative impact on corpus consultation and contributed to the low accountability of error correction.

When compared with the less frequent users', COCA accountability rate (24%) for error correction, the frequent users' (55%) was much higher. Although the frequent users also had a slightly higher

error correction rate of 55% (38/69) in contrast to 50% (33/66), it is too early to infer that there is a causal relationship between the use of COCA and the success of error correction. The data only indicated that COCA frequent users depend on the corpus for error correction while the less frequent users probably depend on other resources or tools to help them correct the errors.

Interviews

Seven students were interviewed. Four (JM, ST, RT, SA) from the frequent-user group, and three (CA, KR, AG) from the less-frequent-user group agreed to be interviewed. During the interviews, their second drafts, final drafts, and the corpus search results were presented to help them recall the process of using COCA. In addition, they explained the problems they encountered, stated their likes or dislikes about the use of COCA, and gave some suggestions or comments. In general, most of them agreed to the following: a) There was a need for more time and training in using COCA; b) if this need was fulfilled, COCA would be helpful in mitigating the lexical or grammatical errors in their writing; c) it was not easy to identify the error types themselves; d) COCA was a useful tool to help them learn about the authentic language use.

Somehow, these shared concerns underscored three core problems as evidenced in their interview transcripts: a) error type identification, b) level of familiarity with the syntax of enquiries and the functions of the corpus, and c) level of difficulty in interpreting the concordance output. Though the errors had been underlined by the teacher, most often the interviewees found it hard to make the right corpus enquiries in just one attempt. As SA stated, "If the phrase is underlined, I will use the whole phrase to check. It fails. Then I use some words in the phrase. This or that word and see. If I can see, then I will use COCA" (SA-01). Similar problems were expressed by the other four interviewees. However, for most students, if without the time constraint, they made several attempts until they found possible ways to correct the errors.

To help correct the errors and use COCA in a more efficient and time-saving way, some students like SR first resorted to online dictionaries or electronic dictionaries to find the English word/phrase that matched their Chinese meanings and then keyed in the word/phrase in COCA with the KWIC function. Some like JM consulted her friends to pinpoint the error type and then make the COCA enquiry. She then read some of the sample sentences and the pertinent contexts produced by the concordancer to see if she could find some sample sentences to express her intended meanings as closely as possible. For frequent corpus-users like SR and JM, they tried to use corpus consultations to correct the underlined errors, though it was "time-consuming" (RT-01), and the COCA functions were "complicated" (JM-02).

However, for less frequent corpus-users, their responses were relatively negative. They considered COCA as too inconvenient (AG-02), too complicated to learn (CA-02), and too time-consuming to be useful (KR-01). As AG recalled, "If I feel something is wrong, I just revise it but I won't use COCA to check…Most of the time, the COCA sample sentences are not what I need" (AG-03). AG admitted that she depended heavily on online sources like Google translation software or Yahoo dictionary to tackle the language problems if necessary. KR, a more proficient EFL writer, who used the corpus the least frequently confessed that she did not like COCA. She said that she had no need for it and it was not useful. When asked if she would consider using COCA to help her acquire more authentic language use, she simply declined. It was surprising to know that she, a proficient English learner who scored 785 on TOEIC, had little motivation to try it.

Though KR responded negatively to COCA, the other six interviewees expressed willingness to receive more training in using COCA, and use the corpus to learn more about authentic language use. Some used it for other purposes like drafting up a speech, a partial requirement of a speech class or

checking the lexical or grammatical errors in a mail to a DJ at an English-speaking radio station in Taiwan. More or less, despite their differences in the frequency of using the corpus, most of the interviewees were positive toward the use of COCA in tackling the language problems in L2 writing. As JM said, they believed that COCA was authentic and credible in this respect.

Survey results

Based on the clusters of questions as suggested in Yoon and Hirvela (2004), the survey findings of this study are presented in the order of four sub-sections: a) students' assessment of advantages of corpus use, b) students' problems/difficulties in corpus use, c) students' responses to corpus use in writing instruction, and d) students' overall evaluations of corpus use in L2 writing. With regard to students' assessment of advantages of corpus use, 90.9% of the students reported positively about the corpus use relative to writing and the usage of vocabulary, and 72.7% of the students thought that corpus use helped with understanding vocabulary and grammar. However, 45.5% and 54.5% of the students believed that corpus use was helpful with reading and the standardized English tests. The results were expected as the corpus was primarily used to correct the underlined errors in the second drafts, not for full understanding of vocabulary and grammar.

In terms of students' problems or difficulties in corpus use, 81.8% of the students had a problem with access to the corpus because there was a limit of use set for non-registered and registered users. In addition, 81.8% of them had difficulty with reading unfamiliar words in the output; 90.9% of them felt that analysing the concordance output was time-consuming; 63.6% of them felt it difficult to analyse the output. Regarding students' responses to corpus use in writing instruction, most students responded positively to corpus use in writing instruction, although they were required to use it. Over 80% of them felt that the search techniques and hands-on practices were easy and helpful. In addition, over 70% of them agreed that they learned from the corpus search and it was a helpful means to solve writing problems. What was intriguing was that 81.8% of them believed that constructing prototype was easy, but in reality some used the wrong prototypes for search, and this resulted in negative changes of the errors. Finally, with regard to students' overall evaluation of corpus use, 81.8 % of the students would recommend it to other classmates, and 72.7% of them desired to use it in future writing. Overall, 63.6% of them believed that corpus use was useful for writing.

Based on these findings and discussion, the answers to the three research questions can be concluded as follows. First, corpus DDL helped intermediate EFL writers correct lexical and grammatical errors in writing, but only accounted for half of the error correction if the writers used corpus search frequently. Second, when processing corpus enquiries to correct errors, some students encountered difficulties in identifying the error types, understanding unfamiliar words in the concordance output, and analysing the output. Less frequent corpus users would use other online resources like online dictionaries, Google or Yahoo Translator to solve the problems themselves or ask their peers for assistance. Finally, most of the students gave positive feedback to the use of COCA in the writing class. They understood that the corpus was beneficial and useful to the sentence-level revision of the texts, and they intended to keep using it in the next writing course if they could get more familiar with the corpus and reduce the time for corpus search.

Conclusion

It is impossible to generalize from such a small-scale study whether all intermediate EFL writers will be able to solve the language problems in EFL writing if they consult a corpus to improve a text where the local errors have been underlined in the process. However, it is suggested that corpus consultation may be a viable means for EFL learners to participate more actively in local-error correction, if the following condition is fulfilled, that is, students are given sufficient training, guidance and time to

practice making corpus inquiries. At the time of this study, the students had little experience with corpus consultation. Most of them still had problems with analysing and interpreting the output, and they needed guidance and assistance in this respect. However, they were not overwhelmed by frustration as reported in the previous studies. On the contrary, they believed that COCA was a credible tool and the authentic texts offered them a good opportunity to learn authentic language use.

Some pedagogical implications emerged from the results of this study. First, using corpus consultation to correct local errors in EFL writing can be useful, but students also need other learning resources or tools to tackle the language problems, or even to facilitate the effectiveness of corpus consultation. Second, students should be given more time to use corpus consultation. Finally, at the initial stage of corpus use practices, teachers should provide tutorials on the students' individual errors so that their interests in using the corpus can be enhanced. Finally, EFL writing teachers should be aware that corpus consultation is not a panacea to correct all the errors emerging in the EFL students' writing

This study has some limitations. The sample size was small so the quantitative data should be treated more as a reference, rather than for any summative conclusions. Some of the students' corpus search results were incomplete as some students forgot to record part of their corpus enquires, and they had to recall information to supplement the data. Finally, unified course requirements and the teaching schedule restrained the teacher from devoting enough time to address students' needs for more training in the use of COCA. Despite all these limitations, in general, this study found that the activity of corpus consultation was beneficial to these EFL writers and with more practice of corpus use, it is believed that the students would learn more, use the corpus more often, and write better. Further research could investigate whether with a longer time frame and more time for corpus use, learners can gradually learn inductively and develop autonomy in error corrections and improving writing.

Acknowledgements

References

Aston, G. (2001). Learning with corpora: An overview. In G. Aston (Ed.) *Learning with corpora (*pp. 7-45). Bologna, Italy: Athelstan.

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

Bloch, J. (2007). *Technologies in the second language composition classroom*. Ann Arbor, MI: University of Michigan Press/ESL.

Boulton, A. (2009). Data-driven learning: Reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics, 35*(1), 1-28.

Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*(3), 534-572.

Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology, 9*(2), 111-125.

Chan, A. (2010). Toward a taxonomy of written errors: Investigation into the written errors of Hong Kong Cantonese ESL learners. *TESOL Quarterly, 44*(2), 295-319.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of second language writing*, *12*(3), 267-296.

Chang, W.-L., & Sun, Y.-C. (2009). Scaffolding and web concordancers as support for language learning. *Computer Assisted Language Learning, 22*(4), 283-302.

Conrad, S. (2000). Will corpus linguistics revolutionalize grammar teaching in the 21st century? *TESOL Quarterly, 34*(3), 548-560.

Conroy, M. A. (2010). Internet tools for language learning: University students taking control of their writing. *Australasian Journal of Educational Technology*, *26* (6), 861-882.

Darus, S., & Subramaniam, K. (2009). Error analysis of the written English essays of secondary school students in Malaysia: A case study. *European Journal of Social Sciences*, *8*(3), 483-495.

Ferris, D. R. (1995). Can advanced ESL students be taught to correct their most serious and frequent errors. *CATESOL journal*, *8*(1), 41-62.

Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, *31*(2), 315-339.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of second language writing*, *10*(3), 161-184.

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, *32*(3), 301-319.

Geluso, J. (2011). Phraseology and frequency of occurrence on the web: native speakers' perceptions of Google-informed second language writing. *Computer Assisted Language Learning*, *26(2),* 144-157.

Hegelhemier, V. (2006). Helping ESL writers through a multimodal, corpus-based, online grammar resource. *CALICO Journal, 24*(1), 5-32.

James, C. (1998). *Errors in language learning and use*. London, New York: Longman.

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English language research journal*, *4*, 1-16. Retrieved from http://wordsmithtools.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf

Johns, T. (2002). Data-driven learning: The perpetual challenge. *Language and Computers*, *42*(1), 107-117.

Krashen, S. D. (1981). *Second language acquisition and second language learning. Retrieved from* http://www.sdkrashen.com/SL_Acquisition_and_Learning/

Lee, D., & Swales, J. (2006). A corpus-based EAP course for New NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, *25*(1), 56-75.

Liou, H.-C., & Peng, Z.-Y. (2009). Training effects on computer-mediated peer review. *System*, 37, 514-525.

Liu, D., & Jiang, P. (2009). Using a Corpus‐Based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, *93*(1), 61-78.

Murrow, P. (2005). Analysis of Grammatical Errors in Students' Writing-indicators for curricula enhancement. 松江工業高等専門学校研究紀要, *40*, 19-26. Retrieved from http://www2010.matsue-ct.ac.jp/tosho/kiyou40/pdf/k-report02.pdf

O'Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, *19*(3), 269-286. DOI: http://dx.doi.org/10.1017/S095834400700033X

O'Sullivan, Í., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of second language writing*, *15*(1), 49-68.

Smith, S. (2011). Learner construction of corpora for general English in Taiwan. *Computer Assisted Language Learning, 24*(4), 291-316.

Stapleton, P., & Radia, P. (2010). Tech-era L2 writing: Towards a new kind of process *ELT Journal 64*(2), 175-183.

Sun, Y. C. (2007). Learner perceptions of a concordancing tool for academic writing. *Computer Assisted Language Learning*, *20*(4), 323-343.

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.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning 46*(2), 327-369

Tung, C.-A., Chang, S.-Y., & Peng, F.-M. (2013). A better form of peer review with the integration of critical thinking training in an EFL writing class--CMPR or OPPR? *Feng Chia Journal of Humanities and Social Sciences, 25*, pp. 95-127

Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing, 15*, 179-200.

Yoon, C. (2011). Concordancing in L2 writing class: An overview of research and issues. *Journal of English for Academic Purposes*, *10*(3), 130-139.

Yoon, H. (2008). More than a Linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning and Technology, 12* (2), 31-48.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of second language writing*, *13*(4), 257-283.

You, Y. L. (2001). To write well is to think clearly: Suggestions for teaching EFL writers how to write logically. Proceedings of the Eighteenth Conference of English Teaching and Learning in the Republic of China, Taiwan, 391-406.