

Online Corpus Consultation in L2 Writing for In-service Teachers of English

Moonbok Lee (Korea Institute for Curriculum and Evaluation)
Dongkwang Shin (Korea Institute for Curriculum and Evaluation)
Yuah Vicky Chon (Korea Institute for Curriculum and Evaluation)

Lee, Moonbok, Shin, Dongkwang, & Chon, Yah Vicky. (2009). Online corpus consultation in L2 writing for in-service teachers of English. *English Teaching*, 64(2), 233-254.

Use of corpus has traditionally been of interest for syllabus designers and lexicographers. However, the recent paradigm shift in research on corpus has shifted to how corpus consultation may have an affect on the writing performance of L2 writers. In the study, using a repeated measures design, two groups of in-service English teachers were asked to write two compositions on two writing prompts where subjects were provided with online referencing at *Concord Writer 2* to help writers solve their lexical and grammatical problems. When the writing products were rated, compositions from the post-training writing tasks showed significant improvements in vocabulary, but not for grammar. Comments made by raters in the written feedback for the compositions imply that L2 writers may notice their problems more immediately if the lexical and grammatical errors can be pointed out to them for independent corpus consultation. Results of the study indicate importance of training in online corpus consultation skills for the different types of errors that may occur during the process of writing. The study is expected to provide implications for training in online corpus consultation skills in L2 writing.

I. INTRODUCTION

Within the Korean context among both EFL teachers and students, there has been less concern with L2 writing compared to the other skills of English due to the current exam-oriented system of Korean secondary schools (Yuah Vicky Chon & Hae-Dong Kim, 2005). In the same vein, there is lack of interest in consultation skills that may be needed for referencing online corpora during the L2 writing process. Also, despite the more recent recognition of the advantages of corpus consultation for L2 writing, there is scarcity of

meaningful empirical research investigating learners' use of corpora. Explanation on what effects such exercises might have on the students' writing performance is even more lacking in the field (Gaskell & Cobb, 2004; O'Sullivan & Chambers, 2006; Stevens, 1991; Yoon & Hirvela, 2004). However, since it is the EFL/ESL teachers that are expected to teach corpus referencing skills of L2 writing, we then need greater understanding of what actually happens when the teacher uses corpora in L2 writing. The purpose of the study, therefore, was to examine what effect corpus referencing has on the writing performance of EFL teachers who in our study were participants of an in-service teacher training program. Based on the writing products, this study also considered how corpus referencing skills for L2 writing could be improved in teacher-training workshops so as to help teacher-trainees implement corpus referencing skills within their own writing and ultimately guide EFL students to benefit from classroom activities where corpus can be used as a pedagogical tool for L2 writing.

II. BACKGROUND

1. Use of Corpora for L2 Learning

The growing interest in corpora and concordancing in recent years has provided invaluable resources in the context of language learning and teaching. A corpus, in the context of modern linguistics, is understood to imply "a finite body of text, sampled to be maximally representative of a particular variety of a language, and which can be stored and manipulated using a computer" (McEnery & Wilson, 2001, p. 73). A concordancer is the tool most often used in corpus linguistics to examine corpora. Johns (1988, p. 9) defines a concordancer as follows: "A concordancer is a computer program that is able to search rapidly through large quantities of text for a target item (morpheme, word, or phrase) and print out all the examples it finds in the contexts in which they appear." The word form under examination appears in the center of each line with an extra space at each side of it and the surrounding context to its left and right, facilitating the study of the immediate context of the keyword and allowing patterns to become clearly visible.

While corpora in the former days were used for developing syllabi, instructional materials and reference works, typically by focusing on the most frequent items found in the corpora (i.e., '*behind-the-scenes*'), the more recent approach (i.e., '*on stage*') focuses on directly utilizing raw concordances to exemplify real-life usage of a target item (Sorell & Shin, 2007). This approach has been promoted as 'data-driven learning' by Johns (1986, 1988, 1991). Data-driven learning focuses on the discovery procedure by which learners inductively derive and deductively apply generalizations by categorizing data from corpora.

This procedure is based on the rationale that inductive learning can occur from multiple examples (Ellis, 1996; Skehan, 1998), and that data-driven learning using corpora can provide a concentrated source of English input in classrooms where the amount of real English input available for learners is usually very limited. However, the distinction between the two strands has diminished and the recent development of corpus analysis software and computer networking makes it easier for learners to access free online concordancers (e.g., Tom Cobb's LEXTUTOR, <http://www.lex Tutor.ca>) which combine common features of both strands.

As a motivation to our study, we found that one of the practical uses of corpora and concordancers for pedagogical purposes is in L2 writing where learners can check the corpus for authentic uses of target lexical or grammatical items (Thurstun & Candlin, 1998). Exposure to these examples of genuine language has been known to (a) enrich learners' understanding of specific uses of target words in a wide variety of contexts and (b) expand their L2 linguistic repertoire. Students' encounters with these multiple samples of discourse combinations should then contribute to growth as L2 writers. In light of such benefits of corpora use in L2 writing, we investigated how inclusion of corpus consultation via concordancers had an effect on the writing performance of teacher-trainees who would possibly transfer these skills to their students in corpus-based L2 writing.

2. Use of Corpora as a Pedagogical Tool in L2 Writing

Current research and development on corpus linguistics at the present lacks empirical studies on how concordances can help L2 writers improve their writing performance. The use of corpora and concordancing in the language-learning environment began as early as 1969 (McEnery & Wilson, 1997, p. 12), but it was the work of Tim Johns in the 1980s (1986) and others which brought it to public attention. Important developments occurred in the 1990s, beginning with publications advocating the use of corpora and concordancing in language teaching (Tribble & Jones, 1990). The first empirical study of learners' consultation of corpus printouts (Stevens, 1991) was followed by a number of studies of the effects of direct corpus consultation (Bernardini, 2000, 2002; Bowker, 1998; Cobb, 1997; Kennedy & Miceli, 2001, 2002; Sun, 2003; Turnbull & Burston, 1998). Recently, there has been increasing interest in the use of corpora to improve language learners' writing skills, and researchers have begun to investigate how learners use corpora in L2 writing and error correction. Those few studies have been conducted within the area of investigating L2 writers' attitudes toward corpus use in L2 writing (Yoon & Hirvela, 2004), and how corpus consultation may help learners correct their writing errors (Gaskell & Cobb, 2004; O'Sullivan & Chambers, 2006).

Yoon and Hirvela (2004) examined students' corpus use behavior and their perceptions

of the strengths and weaknesses of corpora as a second language writing tool in two ESL academic writing courses. The study involved a survey of student attitudes and follow-up interviews with selected students from the two ESL academic writing courses at a large Midwestern American university. A majority of the students were from East Asia, predominantly Chinese (32%) and Korean (23%). The classes worked with the Collins COBUILD Corpus sampler (100 million words), which was attractive in part because of its availability to students as a free web sampler, and also due to the availability of the database from written sources (e.g., letters, newspaper articles, memos, advertisements). The instructor emphasized its value in enhancing students' knowledge of vocabulary and grammar, and thus its positive effects on the development of writing ability.

The results of the survey indicated that the majority of students in both groups had positive feelings about corpus use relative to writing. There was a particularly strong belief in the idea that the corpus use improved their general writing skills, and that the corpus was particularly useful for acquiring usage patterns of words and enhancing their writing skills.

In connection to our study, the relatively low scores regarding grammar are most noticeable. Even though corpora as pedagogical tools have drawn attention with respect to their grammatical value, the mean scores reported in the surveys were not especially high relative to the emphasis placed on the grammatical usefulness of corpora in the classroom instructions. The reason for this was not clear to the researchers, but they speculated that this may be due to the students having their own perceptions about grammar, perceptions that did not match those of advocates of the corpus approach. Given that a large majority (82%, as reported earlier) had come from Asian countries, where traditional structural and grammar-translation approaches to language teaching have long dominated, it was likely that this was how they had learned about English grammar, in which case they may not have related the lexical approach of corpora and the emphasis on collocations to grammatical value.

Gaskell and Cobb (2004) reported on some preliminary results gained from attempts to make concordance information available to 20 adult Chinese lower-intermediate EFL learners enrolled in a 15-week writing course at a university in Montreal, Canada. Most of the participants had an undergraduate degree from China, and all had received at least three years of English instruction via the grammar-translation method. The research question of interest to the present study was to seek if learners could use concordances to reduce their errors in writing. Here the concordancer (Virtual Language Center's concordancer: <http://vlc.polyu.edu.hk/concordance/> Consulted in April 2004) was used to mainly provide feedback for sentence-level writing errors, in particular with regard to grammatical errors. As a part of the training procedure, one of the researchers gave feedback to each student's assignment, including online concordance links for five typical

errors. The students were required to revise the text for final submission, and for each of the concordance-linked errors to submit a form explaining specifically what correction had been made based on concordance information.

The results of the study indicated that 50% of the students had found improvements in their ability to use many of the grammar points targeted in the course. However, only 8 of the 20 learners attributed this to concordance work where some students explained that their exposure to multiple examples of English structures had helped them understand how to use constructions they had been having trouble with in the beginning of the course. As one of the main research questions, when the researchers also investigated which errors appeared to be affected by using the concordance, the post-testing writing samples, which were written on the same topic and under identical conditions as the pre-test samples, showed no decrease at all in terms of overall errors (i.e., 396 errors before, 405 after) where the researchers attribute this to the slightly longer texts in the post-tests.

As a whole, Gaskell and Cobb's study showed that lower level learners are willing to use concordances for feedback in correcting writing errors at the word and sentence levels, and that some learners can be led to conduct independent concordancing. However, as the researchers admit, the study leaves many questions about grammar concordancing that future studies should look at, and points out that adapting concordances for lower level learners' grammar development is less straightforward than for lexical development.

Another study that involves corpus consultation in L2 writing was conducted by O'Sullivan and Chambers (2006). The study presented the second phase of a research project at the University of Limerick involving native speakers of English at both the masters and undergraduate levels who were given the opportunity to engage in corpus consultation in order to improve their writing skills in French (see Chambers & O'Sullivan, 2004 for the first phase of this study). For the researchers, the study was motivated by the need to investigate the potential of corpora in the promotion of L2 writing skills in general and the role of concordance data as a means of assisting error correction.

To obtain results of the study, the researchers compared the essays that had been produced using traditional resources, such as dictionaries and grammar resources, with those corrected with the aid of a corpus. The researchers also analyzed the feedback and evaluation forms completed by the students in order to gain information on the types of errors corrected by the students. From the changes which resulted from consulting the corpus, the researchers established a system of classification of errors based on previous taxonomies (i.e., Corder, 1974; Ferris, 2002; James, 1998; Richards, 1994). To the researchers, most notably grammatical errors (i.e., use of prepositions) followed by lexical errors (i.e., word choice or inappropriate vocabulary) appeared to improve the most as a result of students engaging in corpus consultation. This differs from the results of Yoon and Hirvela where students reported that they did not benefit from corpus use for

grammatical errors. O'Sullivan and Chambers interpreted that namely prepositions and word choice appeared to be one of the most common sources of errors within the two categories due to native language interference between English and French.

Having recognized that there is lack of research on the use of corpora as a pedagogical tool for L2 writing, in particular with use of recent freeware online corpus and concordancer, the present study tries to add findings with regard to how non-native teacher-trainees performed with the concordancer as a writing aid, and the research questions are presented as follows. However, in trying to answer one of the research questions, we were aware that with short-term training in corpus use, it would be difficult to see immediate improvements in grammar (Gaskell & Cobb, 2004; Yoon & Hirvela, 2004) which is an aspect of 'language use' in our study.

- 1) When rated by analytic scoring, to what extent did the L2 learners' writing performance improve with use of corpus consultation (i.e., *Concord Writer 2*) as a pedagogical tool for L2 writing?

In the present study, we are not interested in how corpus consultation can potentially improve the writing proficiency of teacher-trainees (which would require a study of a long-term scale), but in the effectiveness of the *Concord Writer 2* as a pedagogical tool for resolving the writers' lexical and grammatical problems.

- 2) What are the writers' problems found according to rater comments? Were they problems that could have been resolved with corpus consultation?

III. METHOD

1. Participants and Context

The study took place in the summer of 2008 during an in-service teacher training program held at a teacher college of Education in Incheon, Korea. Participants of the study were forty-one in-service English high school teachers who were from the Kyongi and Incheon areas. The subjects, who were teaching English at public high schools, could be regarded as advanced learners of English.¹ In order to investigate the effects of corpus consultation on the quality of the writers' compositions as indicated by the overall mean

¹ One of the researchers of the study acted as an instructor for the training sessions, but the length of time did not allow him to conduct a formal proficiency test.

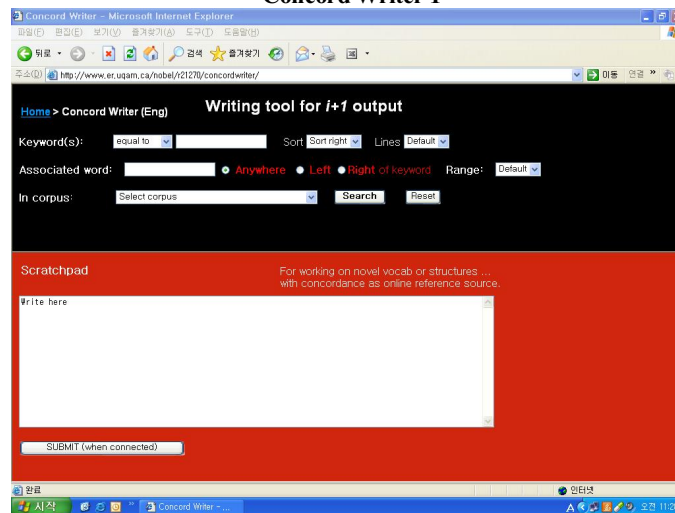
scores of the compositions and written feedback from the raters, we used two intact classes, each referred to as Group 1 ($n=20$) and Group 2 ($n=21$) respectively hereafter. The two groups were used as a means of obtaining subjects rather than as a means of comparing the two groups. In total, there were 16 male and 25 female teachers, and their teaching experiences were approximately four years.

2. Tasks and Procedures

1) Instruments

Since the purpose of the present study is to investigate how online corpus consultation can improve the L2 learners' writing performance, *Concord Writer (CW)*, the online program developed at the University of Quebec at Montreal under supervision of Tom Cobb was utilized. *Concord Writer, Version 1 (CWI)* which is the first version has four major components. The first component of the program is the 'scratchpad' where students can directly write their compositions. See Figure 1 for the details of the program.

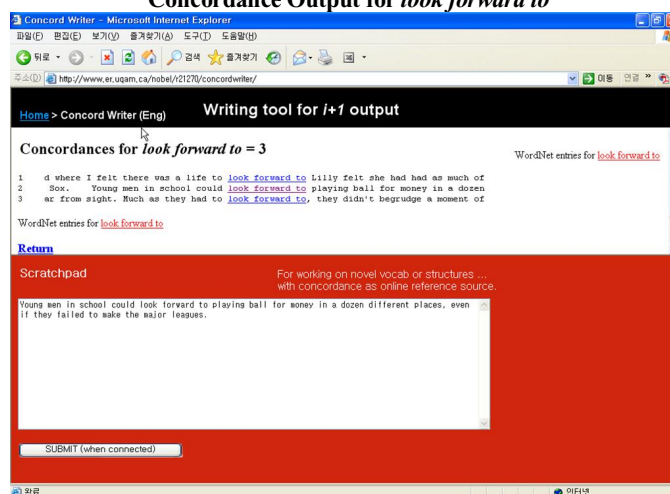
FIGURE 1
Concord Writer 1



The second component is the 'concordancer' which can be selected at 'Select corpus.' *CWI* offers 12 types of different corpora, and can offer up to thousands of example sentences for a target word. One of the corpora available is the Brown Corpus consisting of one million words of written American English and is the first kind of an electronic corpus. There are also other smaller scale written and spoken corpora available in *CWI*.

In order to initiate a search for a target word, the user has to select a corpus, type in the target word or phrase at 'keyword(s)' and click 'search' to obtain portions of sentences, with the keyword positioned in the middle of them which is known collectively by the more technical term, 'concordance output.' Students can study the output to see how a keyword operates in context with closely related words. For instance, with 'look forward to,' the corpus will provide output as seen in Figure 2 as to what can follow after the target phrase. From the corpus, the learner can check that it is a gerund or a noun, and not a basic verb, that can follow after 'look forward to' so that corpus consultation can be seen as useful for the area of lexico-grammatical patterning (O'Sullivan & Chambers, 2006). Also, clicking on one of the sentences from the output will provide context of the target search word item. This all demonstrates an example of inductive learning which is expected to provide effective means for resolving lexico-grammatical problems.

FIGURE 2
Concordance Output for *look forward to*

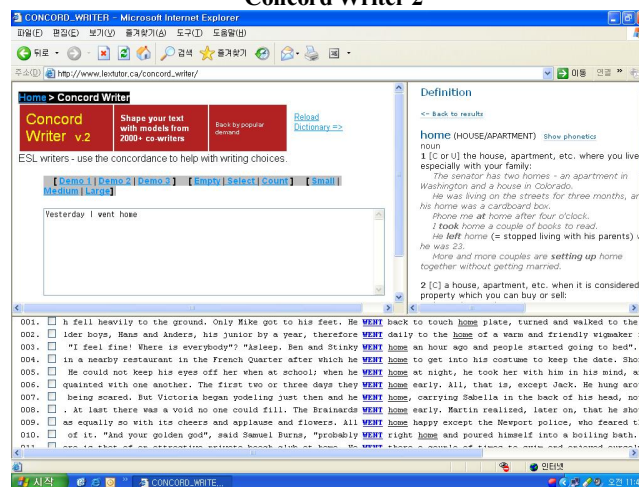


The program also allows the users to look up collocations with regard to the target word. For instance when the word 'go' is typed in for the 'keyword(s)', and 'home' for the 'associated word', this will produce corpus output where the user can find that the preposition 'to' cannot fit between 'go' and 'home'. With this information at hand, students can then compose their own sentences using the keyword with increased knowledge of how to link the word with appropriate collocates. Also, for lexical gap problems, *WordNet (Version 2.1)* (<http://wordnet.princeton.edu/perl/webwn>) allows the user to consult a target word at any time during the writing process. The forte of *WordNet* is that synonyms are well-organized and polysemes are provided with information on frequencies according to the respective meanings so that learners can ascertain which

words to learn first and maximize learning gains.

The most recent version of *CW* is *Concord Writer, Version 2 (CW2)* (http://www.lextutor.ca/concord_writer/) where the most noticeable feature is that interfacing² (Nesi, 1999) is possible with the concordancer, monolingual dictionary and the writing pad so that students can directly write their compositions within a single screen and also call up the dictionary or concordancer while writing (See Figure 3).

FIGURE 3
Concord Writer 2



Based on the premise that an overwhelming amount of information to process may be rather detrimental to the learners' process of inductive learning (Thurstun & Candlin, 1998), *CW2* displays a maximum of 100 example sentences. Other functions that are available at *CW2* are word count functions, and shrink or enlarge functions offered for the writing pad. *The Cambridge Advanced Learner's Dictionary Online* is also provided at *CW2*. In the actual task of the present study, only the concordancer at *CW2* was allowed for use so that we could examine the effects of corpus consultation on L2 learners' writing performance.

2) Training with *Concord Writer*

As a procedure of the study and also as a part of the in-service teacher training program, training with *CW2* was conducted before the learners were asked to write their second

² 'Interfacing' refers to the facility to call up a reference source when working with another application.

compositions in L2 (see later for the design of the study). Training with *CW2* was conducted for two hours for explaining the theoretical justifications and functions of the program. At this stage, the students were shown various applications of the online program for aiding them with their writing problems, in particular how to employ strategies to help them find what they are looking for and ways to fine-tune their searches which would be useful in helping them find target structures within the corpus. Here it was explained that the key to success in using the program lies in the selection of the key word in order to obtain the concordance output. For instance, when meeting a word choice problem with expressions such as ‘all the people’ and ‘the all people’ where the writer is not sure about word order, *CW2* will show that ‘all the people’ is correct. To find this target, the user has to type in ‘people’ as the key word, and ‘all’ as an associated word. Another example with the key word ‘avoid’ would show that nouns and gerunds, but not a basic verb can come after the word. Also, when the query is to find words that can come between ‘pick’ and ‘up’, ‘pick’ may be entered, so that only words that collocate after ‘pick’ appear, and enter ‘up’ to find the target. Attention was also paid to guiding students in the use of corpora and concordancing for finding not just lexical items but also a combination of lexis and grammar (i.e., lexico-grammatical patterns).

3) Writing Tasks and Procedures

Each teacher-trainee was asked to write two writing tasks: pre-training and post-training writing tasks. There was training on the use of *CW2* before the second writing task (i.e., post-training task) which was to be analyzed in comparison to the first writing task (i.e., pre-training task) for investigating the effectiveness of corpus consultation in improving the L2 learners’ writing performance. For the two writing tasks, writing prompts were counterbalanced so that subjects would not all be writing for identical writing prompts in each of the sessions (see Table 1). By doing this, any order effect could be eliminated. Writing prompts which were considered interesting to the in-service teachers were selected from the topics of the Test of Written English of TOEFL. One writing prompt asked the subjects to write on the reasons for attending college or university, and the other asked the subjects to explain how movies or television influence people’s behavior (See Appendix for the complete writing prompts).

TABLE 1
Design of the Writing Tasks

	Group 1	Group 2
Pre-training Task (Task 1)	Writing prompt 1	Writing prompt 2
Training	<i>Concord Writer 2</i>	
Post-training Task (Task 2) with use of <i>Concord Writer 2</i>	Writing prompt 2	Writing prompt 1

For Groups 1 and 2 respectively, the study tasks including training took 4 hours as a whole to complete and they were all trained by the same instructor. Since the aim of the study was to investigate the effectiveness of corpus consultation for solving the L2 learners' writing problems through the comparison of the two writing tasks, the subjects were allowed to use corpus and concordancers only in the second writing task. For each task, the subjects were given 50 minutes to write, but considering the extra time needed for searching *CW2*, the subjects were not pressured to finish on time during the second writing session. After the first writing task, training was conducted after which the second writing task was conducted. To eliminate any intervening variables (e.g., practice with using *CW2* that may have influenced the second writing task), the post-training writing task was conducted after a 15-minute break after the training session. If more time had been provided (e.g., a few days to a few weeks), it is likely that some teacher-trainees, being highly motivated learners, would have become skillful at using the program while others would have abandoned using it. However, the major reason was that there were situational constraints which did not allow the teacher-trainees to come back for a separate session.

3. Data Analysis

The writing tasks yielded 82 writing products from the 41 subjects. In order to calculate the differences of writing scores between the essays collected in Task 1 and Task 2, the first step of the analysis was to score the essays. The writing products were scored by two experienced instructors who were native speakers of English having worked at the tertiary levels for 5 and 8 years respectively. The writing products were rated with Jacobs, Zinkgraf, Wormuth, Hartfiel and Hughey's (1981) ESL Composition Profile, the analytical writing scale which produces a score for content (30 points), organization (20), vocabulary (20), language use (25), and mechanics (5), producing a total of 100 points. As such, according to descriptions of the criteria provided in the writing scale, the raters read the writing products to assign scores for each aspect of writing (i.e., content, organization, vocabulary, language use, mechanics) where problems with grammar and sentence structure are included in 'language use.' For the writing scores, there was an inter-judge reliability of 0.87 between the two raters with Cronbach α at a significant level ($p < 0.05$).

As a second stage of the analysis, the raters were also asked to write comments on the type of errors found in the compositions and to provide suggestions for how they could be corrected.

IV. RESULTS AND DISCUSSION

1. Corpus Referencing and L2 Learner Compositions

This section answers research question 1: When rated by analytic scoring, to what extent did the L2 learners' writing performance improve with use of corpus consultation as a pedagogical tool for L2 writing? After the two raters scored the compositions, the writing scores produced by Group 1 and Group 2 were calculated for the overall mean scores of pre-training and post-training writing tasks. The difference in the overall mean scores between the pre-training and post-training writing tasks indicated an increase of 2.90 which indicate a smaller difference than we expected. However, the large standard deviations (Pre-training task: SD =9.07, Post-training task: SD =9.21) indicate that there were large variations between the subjects, and this is not surprising due to the statistically small size of the trainee group ($n=41$) and that some learners had avidly consulted the corpus while others had not during their writing of the study tasks. Although computer literacy skill was not of interest in the present study, it was apparent during the session that some trainees were not comfortable with using the computer so that working with the online corpus and concordancer would have interfered with searching and finding target lexico-grammatical patterns. In fact, the negative numbers in Table 2 that appeared for the difference between pre-training and post-training mean scores may suggest that some of the subjects' writing scores fell when corpora and concordancer were in use. It is not clear as to how this happened³ but this probably occurred when the writers needed extra time to consult corpora while writing.

³ To investigate how some L2 writers' writing scores fell in the post-training session, studies with the think-aloud method would be needed to provide an explanation. In this study, we were primarily interested in looking at the compositions from a product-oriented view.

TABLE 2
Pre-training and Post-training Writing Tasks

(n=41)	Pre-training Task		Post-training Task		Difference between Pre-training and Post-training Tasks			
	Mean	SD	Mean	SD	Mean	SD	<i>t</i>	<i>p</i>
Content	20.83	3.41	22.63	3.37	1.81	4.41	2.62	0.01*
Organization	15.05	2.17	15.68	2.27	0.63	3.00	1.35	0.18
Vocabulary	13.68	2.20	15.00	2.05	1.32	2.37	3.56	0.01*
Language use	16.85	3.54	16.07	4.10	-0.78	3.88	-1.29	0.21
Mechanics	3.39	0.70	3.32	0.72	-0.07	0.88	-0.53	0.60

Note: SD = standard deviation, * = statistically significant ($p < .05$)

With regard to the difference of overall total mean scores between pre-training and post-training tasks, the paired-sample *t*-test indicated a borderline significance level between the two conditions (overall mean difference= 2.90, $t = 2.03$, $p = 0.05$) which suggest that corpus consultation did assist writers improve their performance to some extent in L2 writing. In fact, seeing that the difference was significant between the two conditions for 'content' and 'vocabulary,' we can infer that better use of vocabulary in the corpus consultation condition contributed to the rater's improved evaluation of the compositions in the post-training writing tasks.

Here we pay separate attention to grammar since previous research (O'Sullivan & Chambers, 2006) has pointed out the effectiveness of corpus consultation for improving knowledge in the combination of lexis and grammar used by native speakers. In fact, the subjects in our study did not benefit from corpus consultation for improving grammatical aspects of their compositions. A possible explanation for this is that the teacher-trainees, having been trained with the grammar-translation approach for most of their lives, may have regarded inductive learning via corpus consultation ineffective for helping them solve lexical or grammatical problems during writing. To the teacher-trainees, it seems grammar as presented through a corpus orientation did not conform to the teachers' notion of learning about grammar. In the same vein, Yoon and Hirvela (2004) report that in relation to the grammatical usefulness of corpora, their students who were predominantly Asian did not rate corpus consultation as favorably as they rated its usefulness for vocabulary or writing skill development. The authors attribute this to the unfamiliarity of students with the notion of lexico-grammatical patterning: "grammar as presented through a corpus orientation did not conform to the notion of learning about grammar" (p. 269).

The results also point to the need for further training if the subjects are to benefit from corpus consultation for reducing lexico-grammatical errors that can occur in their

compositions. As a part of training, we realized through qualitative analysis⁴ of the compositions that during the initial training stage for using corpora and concordancers, guidance needs to be provided for teacher-trainees, for instance by marking their errors and demonstrating corpus searches (O'Sullivan & Chambers, 2006) or by including online concordance links for typical errors that may occur with the students (Gaskell & Cobb, 2004). Also, when there are occasions for teacher-training, teachers should be pushed to write at an '*i+1*' level (Swain, 1985; Swain & Lapkin, 1995) so that every opportunity can be provided for them to notice their linguistic problems and use reference sources such as the *CW*. Through this process, teacher-trainees can be encouraged to self-correct their errors and be led to inductive learning through corpus use. However, as to how much training time is sufficient for effective use of corpora is open to question since the semester spent for training in previous studies (i.e., O'Sullivan & Chambers, 2006; Yoon & Hirvela, 2004) indicated that it was only in the former study that we saw an obvious improvement in grammar.

2. Rater Feedback and Implications for Corpus Referencing

This section answers research question 2: What are the writers' problems found according to rater comments? Were they problems that could have been resolved with corpus consultation?

We examined the raters' comments to see how some of the L2 writers' errors could eventually be resolved through corpus consultation. Overall comments for vocabulary were 'good words but improper use of words', 'the writing shows an adequate range of words, but not a sophisticated range' and that 'meanings become unclear due to wrong choice of words' which also includes problems with use of prepositions and collocations. Table 3 and 4 indicate a sample of the different type of errors and the suggested corrections provided by the raters. The system of errors was established based on previous taxonomies established by researchers such as Corder (1974), Ferris (2002), James (1998), and Richards (1994). The categories of errors are lexical errors (word choice, informal usage, idioms); grammatical errors (prepositions, articles, singular/plural, adjectives, tenses); and syntactic errors (sentence structure, word order).

⁴ Here interest in counting the types of errors is beyond the scope of the present study. We leave this for a future study.

TABLE 3
Lexical Errors and Suggested Corrections

Errors	Corrections
- <i>A country missile* another.</i>	- <i>A country bombs another.</i>
- <i>The world is becoming a global village. Therefore we need to know even the reverse* of our own place in the globe</i>	[meaning not known]
- <i>The characters in the movies do* actions with incredible super power. [collocation problem]</i>	- <i>The characters in the movies perform actions with incredible super power.</i>

Note: *= unacceptable

For language use, which in most cases was related to grammatical and syntactic problems, a majority of the comments were on how grammatical errors or awkward sentence structures made the meaning difficult to understand or ambiguous. Examples are listed as follows in Table 4:

TABLE 4
Language Use Errors and Suggested Corrections

	Errors	Corrections
	- <i>Television and movies make the reality exaggerate.*</i>	- <i>Television and movies exaggerate reality.</i>
Syntactic Errors	- <i>Lots of people go to the movies and watch television when they are free.</i>	- <i>Lots of people go to the movies and watch television when they have free time.</i>
	- <i>Movies and television are too much exposed to people.*</i>	- <i>People are exposed too much to televisions and movies.</i>
Grammatical Errors	- <i>committed a suicide*</i>	- <i>committed suicide</i>
	- <i>The description of the room was on* newspapers. (preposition)</i>	- <i>The description of the room was in the newspapers.</i>
	- <i>When the movie is including many violences, it will have bad effect on people's behavior.*</i>	- <i>When the movie includes a lot of violence, it will have a bad effect on people's behavior.</i>

Note: *= unacceptable

In fact, here we note that teacher-trainees would be able to reduce some of their errors in L2 writing through corpus consultation. However, online corpus consultation searching skills are critical to the success of finding target lexico-grammatical patterns within the concordance output. In the following, we provide examples where errors can be avoided with corpus consultation skills. The errors are those that were pointed out by the native English speaker raters. For instance, the sentence 'the world is very wide' did not sound natural to the raters. The covert lexical error may have originated from L1 interference (i.e., *sesang-eun neoldda*) or the L2 writer may have wrongly guessed that 'world' collocates

with ‘wide’ from the compound ‘worldwide.’ If the writer had checked the corpus with ‘world’ as the key word and ‘wide’ as the associated word (Figure 4), the concordance output would have shown that the two words are predominantly used as compounds (i.e., world-wide) rather than as a subject and a complement so that there is no collocational relationship between ‘world’ and ‘wide.’

FIGURE 4
Use of Key Words at Concordancers

The screenshot shows a search interface with the following elements:

- Word(s):
- with properties: equals
- sorted by: word after keyword
- with associate:
- to left/right: either [reset](#)

With regard to grammatical errors, the phrasal verb ‘committed a suicide’ is ungrammatical because the word ‘suicide’ is not countable. By looking up either ‘commit’ or ‘suicide’ in *CW2*, users can check word association usages as demonstrated in Figure 5. In addition, by extended searches with the verb ‘commit,’ the L2 writer will be able to infer that the verb often collocates with words carrying negative connotations, such as ‘crime,’ ‘murder’ and ‘adultery.’

FIGURE 5
Concordance Output with *committed* and *suicide*

his brother, who afterwards **COMMITTED** suicide. Watson told me tha
water beneath her. If I ever **COMMITTED** suicide, she thought, I wou
that because Mr. Foster had **COMMITTED** suicide in July 1993 prosec
nd it is feared that she has **COMMITTED** suicide. The written accoun
er I'm surprised you haven't **COMMITTED** suicide yet after dropping
a, Col. Gen. Vladimir Janko, **COMMITTED** suicide after he was implic
born in Vienna in 1894. She **COMMITTED** suicide in Oxford on 10 Apr
born in Vienna in 1935, who **COMMITTED** suicide in 1952. Waismann d

Most frequent uses of corpus consultation are for checking prepositions (O’Sullivan & Chambers, 2006) in prepositional phrases or prepositional verbs. For example, in the phrase ‘after the description of the room was on newspaper,’ we can check in *CW2* as shown in Figure 6 that preposition ‘in’ is needed with the definite article ‘the’ preceding ‘newspaper’ so that it is ‘in the newspapers’ rather than ‘on newspaper.’ As such, concordancers often prove useful for prepositional errors which can be considered one of the most common types of grammatical errors among language learners. However, for

queries on prepositions, consulting a corpus with the preposition as the keyword should be avoided since there would be many word combinations that can occur with the target preposition. Thus, choosing a proper search word is the key to successful corpus consultation.

FIGURE 6

Concordance Output with *newspapers*

reading the travel ads in the [NEWSPAPERS](#) and magazines will give you t
 ll. His face was always in the [NEWSPAPERS](#), sometimes in cartoons that s
 es, who called himself, in the [NEWSPAPERS](#), Art the Great, or The Great :
 e sake of earning space in the [NEWSPAPERS](#). There were others who climbe
 elating his peccadilloes to the [NEWSPAPERS](#). State keeps the junketeering
 ake sweeping statements to the [NEWSPAPERS](#). Miriam had not yet goaded hir
 l suite in Bury Street. To the [NEWSPAPERS](#) he talked about his unquiet l:

Another example was retrieved from one of the rater's comments. The sentence with the problematic preposition was '*These days Korea has been split in* two parts, one of which is for the import of American beef and the other which is against it.*' This is the case where the writer should have used 'split into' instead of 'split in.' In this instance of a prepositional error, the corpus as shown in Figure 7 demonstrates the difference between the two. Examination of the output would show that the two words 'split' and 'in' are not semantically related whereas the verb 'split' collocates with the preposition 'into.'

FIGURE 7

Concordance Output with *split in* and *split into*

t status would mean a [SPLIT IN](#) the party, said the Vilr
 Both originated in a [SPLIT IN](#) the nationalist movement
 Republic and Slovakia [SPLIT IN](#) 1993, the Czech Governme
 rday showed the usual [SPLIT IN](#) government reaction. The
 rday showed the usual [SPLIT IN](#) government reaction. The

The corporation appears [SPLIT INTO](#) two camps, with strong s
 the Forms The forms are [SPLIT INTO](#) sections. These sections
 . Man-made fibres can be [SPLIT INTO](#) two groups, those that a
 rvices. The forms can be [SPLIT INTO](#) two groups: forms NEWOED
 at metropolitan life had [SPLIT INTO](#) two trends- expanding in

However, corpus consultation hardly contributed to reducing syntactic errors. To resolve syntactic error problems, users need to find sentence-level target expressions with the same structure within a corpus search, but it is difficult to do this when concordance output contains a variety of sentence structures for the user to analyze so that there is a heavy

cognitive load when the writer has to do this during writing. As such, it can be seen that corpus consultation is most effective for resolving most notably grammatical errors (i.e., use of prepositions) followed by lexical errors (i.e., word choice or inappropriate vocabulary) (O'Sullivan & Chambers, 2006). The types of searches also illustrate that teacher-trainees should try to become active in their attempts to consult the corpus in particular for lexico-grammatical problems. This would provide possibilities for teachers to transfer these skills to other L2 writers, who in most cases would be their students.

V. CONCLUSION

1. Main Research Findings

In comparison to corpus use studies that have been conducted primarily with students, this study examined how online corpus consultation affects the writing performance of in-service teachers who can be considered advanced learners of English. As a whole, the writing products indicated that use of corpus consultation in writing did not automatically resolve all the language learners' problems, and this was due to various reasons involving training, different notions on learning grammar, and lack of knowledge on consultation skills. The results implied that teacher-trainees would need to have their errors pointed out to them particularly in the initial stage of training so as to facilitate learners to use concordancers and find target lexico-grammar patterns that are used by native speakers.

For the lexico-grammatical problems that occurred, we saw that the value of the concordancer lies in the fact that it can make correct forms of the language (e.g., prepositions) more salient to the learner, and therefore potentially lead to greater learning benefits. For word choice problems, we saw that the concordancer helped the learners, particularly in comparison with the dictionary, to see how the words should be used in the correct context, while providing learners with examples of words in multiple contexts as demonstrated by O'Sullivan and Chambers (2006). It should also be noted that it is through production (e.g., writing) that learners have the opportunity to explore new forms and match them to communicants' expectations so that an error on a page is an important opportunity for learning (Swain, 1985, 1995). Last but not least, the study illustrates the importance of training in corpus consultation skills as already documented in the dictionary literature that consultation skills for referencing Internet dictionary sources (Yuah Vicky Chon, 2009) is not an automatic skill that learners can bring.

2. Limitations and Recommendations for Further Research

With the time available with the teacher-trainees, we were not able to take control of the subjects' different computer skills that may have interfered with the corpus consultation process. Also, extra time spent on training with corpus consultation would have allowed us to more accurately ascertain why we did not see a significant difference between the overall mean scores of pre-training and post-training tasks. However, considering that the subjects were teachers with a strong background in self-directed learning and good metalinguistic knowledge of L2 (e.g., grammar rules), we were confident that the training sessions offered to them were sufficient to help them improve their corpus consultation skills. The sessions furthermore did not allow us to provide more time for drafting and revising through which we may have been able to see a reduction in the number of errors through corpus consultation. However, within the scope of our study, counting the number of errors was not an area of interest to us.

In this study, use of corpus was evaluated only from the writing products of the L2 writers. However, with use of think-aloud as conducted in previous studies that investigate consultation skills used for referencing online dictionaries (e.g., Yuah Vicky Chon, 2009), we would be able to devise taxonomy of corpus use strategies in L2 writing. We leave this process-oriented study for future research.

REFERENCES

- Bernardini, S. (2000). Systematising serendipity: Proposals for concordancing large corpora with language learners. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 225-234). Frankfurt: Peter Lang.
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165-182). Amsterdam: Rodopi.
- Bowker, L. (1998). Using specialised monolingual native-language corpora as a translation resource: A pilot study. *Meta*, 43(4), 631-651.
- Chon, Y. V. (2009). The electronic dictionary for writing: A solution or a problem? *International Journal of Lexicography*, 22(1), 23-54.
- Chon, Y. V., & Kim, Hae-Dong (2005). University EFL learners' prior experience of writing instruction in L1 and L2. *English Teaching*, 60(1), 97-124.
- Cobb, T. (1997). Is there any measurable learning from hands on concordancing? *System*, 25(3), 301-315.

- Composition: A practical approach*. Rowley, MA: Newbury House.
- Corder, S. P. (1974). Error analysis. In J. P. Allen & S. P. Corder (Eds.), *Techniques in applied linguistics* (pp. 122-154). London: Oxford University Press.
- Ellis, N. C. (1996). Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18, 91-126.
- Ferris, D. (2002). *Treatment of error in second language student writing*. Ann Arbor, MI: University of Michigan Press.
- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, 32, 301-319.
- Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.
- James, C. (1998). *Errors in language learning and use: Exploring error analysis*. London: Longman.
- Johns, T. (1986). Micro-concord: A language learner's research tool. *System*, 14(2), 151-162.
- Johns, T. (1988). Whence and whither classroom concordancing? In T. Bongaerts, P. De Haan, S. Lobbe & H. Wekker (Eds.), *Computer applications in language learning* (pp. 9-27). London: Foris.
- Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. *Classroom concordancing: English Language Research Journal*, 4, 1-13.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology*, 5(3), 77-90.
- Kennedy, C., & Miceli, T. (2002). The CWIC project: Developing and using a corpus for intermediate Italian students. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 183-194). Amsterdam: Rodopi.
- McEnery, T., & Wilson, A. (1997). Teaching and language corpora. *ReCALL*, 9(1), 5-14.
- McEnery, T., & Wilson, A. (2001). *Corpus linguistics* (2nd ed.). Edinburgh: Edinburgh University Press.
- Nesi, H. (1999). A user's guide to electronic dictionaries for language learners. *International Journal of Lexicography*, 12(1), 55-66.
- O'Sullivan, I., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing*, 15, 46-68.
- Richards, J. C. (Ed.). (1994). *Error analysis: Perspectives on second language acquisition*. London: Longman.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Sorell, J., & Shin, D. (2007). The next step in concordance-based language learning:

- Constructing an online language learning resource for high-frequency vocabulary and collocations. *The International Journal of Learning*, 13(12), 217-222.
- Stevens, V. (1991). Concordance-based vocabulary exercises: A viable alternative to gap-fillers. *Classroom concordancing: English Language Research Journal*, 4, 47-63.
- Sun, Y. (2003). Learning processes strategies and web-based concordancers: A case study. *British Journal of Educational Technology*, 34(5), 601-613.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input and second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Sidelhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.
- Thurstun, J., & Candlin, C. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes*, 17(3), 267-280.
- Tribble, C., & Jones, G. (1990). *Concordances in the classroom: A resource book for teachers*. Harlow: Longman.
- Tribble, C., & Jones, G. (1997). *Concordances in the classroom*. Houston, TX: Athelstan.
- Turnbull, J., & Burston, J. (1998). Towards independent concordance work for students: Lessons from a case study. *On-CALL*, 12(2), 10-21. Retrieved August 26, 2008, from the World Wide Web: <http://www.cltr.uq.edu.au/oncall/turnbull122.html>.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13, 257-283.

APPENDIX

Writing prompts used in the study tasks

Writing prompt 1

People attend college or university for many different reasons (for example, new experiences, career preparation, increased knowledge). Why do you think people attend college or university? Use specific reasons and examples to support your answer.

Writing prompt 2

How do movies or television influence people's behavior? Use reasons and specific examples to support your answer.

Applicable levels: teacher education, tertiary education, general education
Key words: teacher training, corpus, concordancing, L2 writing, consultation

Lee, Moonbok
Division of Curriculum, Teaching, and Learning
Korea Institute for Curriculum and Evaluation (KICE)
25-1, Samchung-Dong, Jongno-Gu,
Seoul 110-230, Korea.
Tel: (02) 3704-3699
Fax: (02) 3704-3570
Email: himoon@kice.re.kr

Shin, Dongkwang
Division of English Education Research
Korea Institute for Curriculum and Evaluation (KICE)
148, Anguk-Dong, Jongno-Gu,
Seoul 110-240, Korea.
Tel: (02) 3704-3958
Fax: (02) 3704-3954
Email: sdhera@kice.re.kr

Chon, Yuah V.
Division of the College Scholastic Ability Test
Korea Institute for Curriculum and Evaluation (KICE)
25-1, Samchung-Dong, Jongno-Gu,
Seoul 110-230, Korea.
Tel: (02) 3704-3828
Fax: (02) 3704-3740
Email: vylee52@kice.re.kr

Received in March, 2009
Reviewed in April, 2009
Revised version received in May, 2009.