

# *A comparison of the effectiveness of EFL students' use of dictionaries and an online corpus for the enhancement of revision skills*

CHARLES M. MUELLER

*Fuji Women's University, Japan  
(email: mueller@fujijoshi.ac.jp)*

NATALIA D. JACOBSEN

*George Washington University, United States  
(email: natalia@gwu.edu)*

---

## Abstract

Qualitative research focusing primarily on advanced-proficiency second language (L2) learners suggests that online corpora can function as useful reference tools for language learners, especially when addressing phraseological issues. However, the feasibility and effectiveness of online corpus consultation for learners at a basic level of L2 proficiency have been relatively unexplored. The current study of Japanese-L1 (first language) learners in an EFL (English as a foreign language) context ( $N = 117$ ) addresses these gaps in research. A preliminary investigation (Experiment 1) examined EFL learners ( $n = 78$ ) as they used the Corpus of Contemporary American English (COCA; Davies, 2008–) to revise essays. Experiment 2 ( $n = 39$ ) used a within-subjects comparison to determine whether participants attained greater accuracy in supplying the missing word in a gap-fill test when using an electronic dictionary or COCA. The survey results from the two experiments revealed that participants generally found using an online corpus difficult. In Experiment 2, a paired-samples  $t$ -test showed that participants, at an alpha of  $p = .05$  two-tailed, were marginally better able to answer test questions when using the online corpus than they were when using an electronic dictionary,  $p = 0.030$ . The implications of the study within the context of previous research are discussed along with pedagogical recommendations and possible avenues for future research.

Keywords: corpus, corpus consultation, COCA, error correction, revision, collocations

---

## 1 Introduction

When producing texts, non-native speakers (NNSs) are often unsure how to resolve problems arising from gaps in their linguistic knowledge. NNSs become aware of these gaps when, for example, they are unsure about a particular lexical choice. At other times, gaps are identified by others, as when their instructor circles a phrase or sentence, or puts a question mark in the margin of their paper. Students' typical response to this problem has been to turn

to reference works, especially dictionaries. Such resources are able to resolve some questions, particularly those regarding lexical choices. Unfortunately, they often provide inadequate information about certain facets of language, such as how words typically combine (i.e., collocations). Dictionaries can also provide inadequate insights into which combinations are more common within a specific genre and which expressions NSs (native speakers) actually use, what Bachman (1990: 97) calls “sensitivity to naturalness”.

As a possible alternative to conventional reference works, a number of studies, especially in the last decade, have explored the feasibility of students’ use of computerized corpora for revision of writing and error correction. The sudden interest in this area has been, in part, spurred by the increasing availability of free and well-organized corpora with sophisticated search interfaces, such as the Corpus of Contemporary American English (COCA, Davies, 2008–). There are a number of potential benefits to corpus consultation, as it can: (1) enhance metalinguistic knowledge (Yoon, 2008); (2) draw attention to the link between linguistic expressions and context (Aston, 2001); (3) provide greater exposure to authentic language (Viana, 2010; for a critical discussion, see Widdowson, 1990); (4) promote a more process-oriented view of language learning (O’Sullivan, 2007); (5) promote learner autonomy (Yoon, 2008); (6) increase students’ confidence (Yoon & Hirvela, 2004); (7) develop research skills; (8) enable students to more effectively revise their writing; and (9) promote “noticing the gap” (see Swain, 1995) as learners observe discrepancies between the solutions that occur to them spontaneously and the solutions suggested by a corpus. The present study aims to determine whether online corpus consultation provides L2 learners with an effective alternative to dictionaries when addressing language problems.

## 2 Review

A large portion of studies in the area of corpus consultation for error correction have been qualitative studies of learners engaged in essay revision tasks, with a focus on students’ correction of lexical errors (e.g., Todd, 2001), grammatical errors (e.g., Gaskell & Cobb, 2004), or both (e.g., Chambers & O’Sullivan, 2004; O’Sullivan & Chambers, 2006). Within this area of research, Koo’s (2006) qualitative study is notable as it involved an examination of both corpus-consultation and the use of conventional reference tools. Koo used both screen capture technology and stimulated recall to examine processes and outcomes of ten Korean graduate students with advanced English proficiency. The participants’ task was to paraphrase an English news article using a concordancing program, dictionaries (L1-L2, L2-L1, and L2-L2), and thesauri. The results indicated that among choices of reference tools, the participants used the corpus most frequently (47.4% of the time). They often combined tools, and when they did so, were more successful in finding solutions to their writing problems. Collocations were the most common reason (69.0%) that they consulted the corpus, and in their corpus searches, they most frequently used a verb (42.4% of searches) in their query to find information about a preposition (23.2% of searches). The study suggests that students with training that provides adequate guidance may prefer to consult a corpus to solve writing problems, and that students are aware that particular problems, such as determining collocates, may be more easily solved using a corpus.

In another qualitative study, Chang (2014) observed ten Korean graduate students’ use of COCA and a specialized academic corpus over a period of 22 weeks. Participants received initial workshop training followed by weekly feedback. Data collection involved a diverse

range of methods including search logs, interviews during weekly meetings, and a survey. Participants used the corpora to write a range of academic texts. The results indicated that several participants found that COCA was useful, as it allowed highly specific searches that could be limited to a particular part of speech and genre. On the other hand, only two participants felt that COCA was easy to use.

Boulton (2009) conducted one of the few empirical studies to compare the use of corpus data and conventional reference materials as a reference and learning source. French learners of English were given information on linking adverbials (e.g., *but*, *actually*, *whereas*, etc.). The participants were divided into four groups: (1) the short context (SC) group saw five short concordance lines (around 40 words) for each item; (2) the key word (KW) group saw eight concordance lines (around 8 words) for each item; (3) the bilingual dictionary (BD) group saw entries taken from a large bilingual dictionary; and (4) the general usage (GU) group saw notes on the adverbials taken from a manual on usage. Participants were given a pre-test (Time 1), a test using information sheets which differed based on condition (Time 2), and a post-test (Time 3). All measures consisted of a gap-fill test with two parts: a concordance line with a missing word or a longer context of several sentences with a missing word. The results showed that the two groups with corpus information (i.e., the SC and KW groups) performed better on the Time 2 test, indicating that consultation of corpus lines was more useful than traditional references for this task. The Time 3 scores (which measured learning) showed no significant differences between groups. The study shows that participants who were not at advanced levels of proficiency and who had no prior training in how to use concordance lines were still able to use corpus-based materials to make better lexical choices. An earlier study by Boulton (2008) that had low-proficiency learners without prior training use corpus lines to answer questions about phrasal verbs likewise found that the participants, even at this proficiency level, were able to improve their performance.

In sum, qualitative research suggests that learners who receive adequate training can make use of online corpora to effectively remedy errors when revising their writing. Moreover, some recent quantitative research indicates that corpus data as a reference tool can be more effective than conventional resources such as dictionaries or usage manuals.

The current study is designed to address several gaps in the existing research on corpus consultation for revision and errors. First, while much research has been conducted with high-proficiency learners who often have both research experience and fairly advanced computer skills, little research has examined learners with the lower proficiency levels typical of many undergraduate college students learning English in a non-immersion environment. For this reason, the current experiments examined the performance of low-proficiency EFL undergraduate students in Japan. Second, while there has been a rapid increase in rigorous qualitative studies in this area during the last decade, more quantitative research employing online corpora is needed to confirm the trends and patterns that have emerged in this research.

The preliminary experiment involved qualitative methods and was designed so as to examine participants' revisions within a highly generalizable context, the revision of a first draft of an essay based on instructor comments. Experiment 2 used both qualitative and quantitative methods to compare the effects of corpus consultation and dictionary use. Because the two experiments involved similar participants and survey questions, the general discussion of their results appears in the final discussion section.

### 3 Method (Experiment 1)

#### 3.1 Aims and research questions

Experiment 1 sought to determine EFL undergraduates' subjective impressions of using an online corpus for essay revision. Kennedy and Miceli (2001) list four steps in corpus consultation: (1) formulating a question; (2) devising a search strategy; (3) observing the results and selecting relevant examples; and (4) drawing conclusions. Both of the experiments in this study focused on the final three steps. This was done for practical reasons. It was felt that lower-proficiency students may fail to correctly identify errors and that this would result in an inadequate pool of data for analysis. The interpretation of results should therefore take into account the fact that participants were not required to identify errors so as to formulate a question.

#### 3.2 Participants

The participants were female students at a women's university in Japan. All were in a department focusing on English language and literature. Eighty-six participants were recruited from four intact classes; eight of them failed to attend one of the sessions, so were dropped from the study, leaving 78 participants: 42 freshmen from two first-year English reading courses, 20 sophomores from an English writing course, and 16 juniors from two third-year English essay-writing courses. The participants can be generally described as highly motivated learners who had fairly rudimentary knowledge of and little interest in computers. Based on the TOEFL ITP scores of incoming freshmen and the scores of juniors who take the TOEFL or IELTS for study abroad, the freshmen can generally be described as A2 level within the Common European Framework of Reference (CEFR), whereas the juniors were generally at level B1 with some participants as high as B2. The sophomores ranged from the A2 to B1 level.

#### 3.3 Instruments

To acquaint participants with the use of online corpora, they were provided with a five-page handout that discussed the following: (1) the visual inspection of corpus lines to determine left-hand and right-hand collocates; (2) examination of token frequency of two phrases to determine which one is more common in spoken English; (3) examination of the frequency of words within various genres to determine whether they are primarily associated with academic or informal registers; (4) precise searches for left-hand or right-hand collocates using part-of-speech (POS) settings; and (5) the need to attend to specific features in the corpus results such as semantic prosody (i.e., the association of words with collocates that have either positive or negative meanings) to determine the most appropriate word for a given context. To acquaint learners with different corpora, two short tasks were performed using the BNC (the British National Corpus, 2007) and the Michigan Corpus of Academic Spoken English (Simpson, Briggs, Ovens & Swales, 2002). The remaining tasks were performed using COCA. A room with computers running Windows and with Internet access was used for the corpus training and practice tasks.

Two “student essays” were created (for an example, see Appendix A) based on Japanese student essays from the International Corpus of Learner English (Granger, Dagneaux, Meunier & Paquot, 2009) to ensure that the errors were realistic. Each essay contained fifteen errors or infelicities: five involved inappropriate collocates (e.g., *do an effort* instead of *make an effort*); five involved inappropriate preposition use (e.g., *control on our thought processes* instead of *control over our thought processes*); and five involved inappropriate word choice (e.g., *alcohol can cause body problems* instead of *alcohol can cause physical problems*).

“Instructor” comments on these errors were then created using Microsoft revisions mark-up. The comments, made to resemble typical teacher comments, stated that the target word was inappropriate for the context and asked the participants to find an alternative. These materials were developed with several considerations in mind. First, they loosely resembled comments the participants might receive on papers prior to a revision stage in writing. Second, even when students revise their writing without instructor feedback, they are likely to follow a similar sequence in which they (1) identify a problem in their writing and then (2) search for an alternative. To record participants’ responses, separate answer sheets with the same sentences but with the errors replaced by blanks were created.

In many of the qualitative studies on corpus consultation for error correction, participants have edited their own essays. This was not done in the current study for two reasons. First, the use of the same essays with fixed categories of errors provided a means of eliciting participants’ impression of using the corpus for a range of item types within a short span of time. Second, the use of three categories of item types (i.e., light verbs, preposition choice, and lexical choice) helped determine whether this particular population of learners could effectively use corpus consultation for these specific linguistic targets. The first experiment thus helped determine the choice of item types used in the second experiment.

In addition to the student essays, a survey was created to elicit participants’ perceptions of the training and the use of corpora as a tool to facilitate revision. All but one of the responses were on a seven-point Likert scale. The remaining item asked for a written response, which could be made in either English or Japanese.

### **3.4 Procedure**

All experimental activities were carried out as part of the participants’ regular classes. Participants were given a 90-minute training session on the use of corpora using the five-page handout, which guided participants through the procedures for using the corpus to answer various questions regarding lexical choices. After going over the steps in the handout, the participants conducted more autonomous searches as they attempted to answer problems on the worksheet, while the experimenter walked around the class, providing one-on-one assistance to participants who had questions or were experiencing difficulties. After this training, participants were asked to complete a number of similar tasks as homework (designed to take about one hour), which was to be submitted a week later at the beginning of the next class. One week later, they were given one of the student essays and asked to make revisions. To create a wider range of items, the two student essays were counterbalanced so that half of the participants in each class received one version (which discussed the benefits of exercise) and half received the other (which discussed the dangers of alcohol). At the end of the 90-minute class, the participants filled out the survey.

#### 4 Results (Experiment 1)

Table 1 shows participants' responses to the survey questions. As can be seen from Item 6, the participants were unfamiliar with corpora, 51 of the 78 participants putting '1' for no prior knowledge whatsoever. When asked whether they envisioned themselves using corpora to facilitate revisions of their writing, the participants gave middling responses. From among the categories related to vocabulary, grammar, and pragmatics, they displayed slightly more enthusiasm for using the corpus to solve problems concerning pragmatics (Item 4). This response is understandable: the participants were shown during the training phase that corpora can provide information on actual use of a form within a specific mode (speech or writing) and genre (e.g., academic writing). Conventional reference materials are generally unable to provide this information for all but the most common expressions.

Participants, particularly the juniors, also displayed some interest in using corpora to revise their research (Item 5). All the participants are required to write a long thesis in English for graduation, so juniors, who were soon to begin writing this long paper, were particularly interested in having a new tool to assist with lexical and stylistic choices as they wrote.

The three groups reported some difficulty in using the corpus (Item 1). Based on observations during the exercise, these difficulties involved several areas of corpus use. First, many participants experienced difficulty when using the corpus interface. This was partly due to lack of experience with computers, and the complexity of the task was exacerbated by the fact that the COCA interface was entirely in English. Several participant comments reflected these technical difficulties:

**Freshman:** It seems like this would be very difficult until I get used to it. In the class, it was so difficult that I could only understand a little. Even so, the ability for the user to divide results by decade, situation, and gender was very interesting! ☺

**Junior:** I dislike it. It's too difficult. I'm bad at using computer a little. To speak honestly, it's too complicated for me.

Second, participants often had a difficult time designing an optimal search query. This was especially true when the error to be revised involved collocates that were separated by intervening words that would ideally be omitted from consideration in the query string. One contributing factor may be participants' failure to understand and apply the concept of strength of association (i.e. the relationship gauged by mutual information measures), or it could be related to difficulty in parsing the sentence correctly so as to determine the proper phrase boundaries.

Table 1 *Descriptive results of Experiment 1 survey results (7-point Likert scale)*

Item	Description	Total	Freshmen	Sophomore	Junior
		<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
1	Ease of use	2.7 (1.4)	2.6 (1.3)	2.4 (1.7)	3.0 (1.6)
2	Future use for vocabulary	4.0 (1.5)	3.9 (1.5)	4.0 (1.8)	4.2 (1.4)
3	Future use for grammar	3.9 (1.6)	3.6 (1.4)	3.8 (2.0)	4.6 (1.5)
4	Future use for pragmatics	4.3 (1.6)	4.0 (1.6)	4.3 (2.1)	4.9 (1.5)
5	Future use for research	4.1 (1.6)	3.8 (1.6)	4.1 (1.8)	4.6 (1.3)
6	Prior knowledge of corpora	2.1 (1.8)	2.2 (1.9)	1.7 (1.1)	2.2 (1.9)

Third, the participants, particularly when revising items involving poor word choice, had difficulty choosing between two or more choices that appeared frequently in the search results (i.e., the corpus lines). This same issue has been reported in previous accounts of learners' use of corpora (e.g., Ädel, 2010).

In addition to the Likert scale responses, the final question asked what they liked or disliked about using the corpus. They were allowed to respond in either English or Japanese. Some responses to the final survey question reflected difficulties:

**Freshman:** It is very difficult for me to use the corpus. The corpus of sentence is very difficult.

**Sophomore:** The corpus is very useful if I could determine which word is right.

**Junior:** Like: It's convenient due to abundant information. Dislike: It's hard to use due to abundance of information.

**Junior:** These are too much examples to find the correct word. It take time to search.

A more general problem was that the search results often included low-frequency words or expressions or were difficult to understand without examining more of the source text. Other research on participants' use of corpora (e.g., Geluso & Yamaguchi, 2014) has reported similar difficulties. This issue was reflected in the following response to the final survey question:

**Junior:** I dislike it. I can't do PC well. And I don't have enough vocabularies to understand sentences of corpus.

Finally, a number of participants expressed enthusiasm at having a tool at their disposal that increased their autonomy and provided insights into authentic English usage. In addition to verbal comments during the lessons, a number of responses to the final survey question reflected this:

**Freshman:** I can understand true meaning of any words. I like it!

**Sophomore:** I like it. I can use real English with it.

**Junior:** I've found a corpus very useful. It lets me find more natural English expression, vocabulary, and grammar. An English tool (service?) is hard for me to use perfectly.

The results of Experiment 1 agree with some previous research (e.g., Gilmore, 2008) that has suggested that lower-proficiency learners may have difficulty carrying out certain steps in corpus consultation, particularly when it comes to interpreting search results. At the same time, some of the participants' more enthusiastic appraisals suggest that they saw the corpus as a useful tool with great potential.

## 5 Method (Experiment 2)

### 5.1 Aims and research questions

Some participants' comments in Experiment 1 suggest that corpus consultation is effective; however, these comments are purely subjective impressions. To determine whether EFL

learners can, in fact, benefit from using a corpus to address some specific language problems after these have been correctly identified, Experiment 2 incorporated a quantitative methodology. The key research question was whether electronic dictionaries or an online corpus search after a brief training period would lead to better solutions to language problems related to collocation and register.

## 5.2 Participants

The participants ( $n = 39$ ) were female freshman EFL students from two first-year English reading classes at the same women's university in Japan, none of whom had participated in Experiment 1. From among the 47 participants who began the study, eight were dropped from all analyses due to their absence during one of the test or training sessions. Within the six months prior to the experiment, 35 participants had taken the paper-based TOEFL ITP (Level 1); their scores are shown in Table 2. Within the CEFR, the score range would put most participants at the A2 (Basic User) level and some of the more proficient participants at the B1 (Independent User) level.

Within six months of the experiment, all but one participant had taken Nation and Beglar's (2007) Vocabulary Levels Test (VLT), from the 1,000 to 10,000 word level. Participants were told to refrain from guessing and, to ensure that guessing did not affect the scores, a third of a point (reflecting the four choices per item) was subtracted for each error. Participants had a mean estimated vocabulary size of 2,796 words ( $SD = 712$ ; range = 1,500–4,167). According to most research on the lexical knowledge required by learners to function independently in English, the participants' mean vocabulary size would be regarded as inadequate. Thornbury (2002), in his summary of the research, claims that 3,000 word families are considered by most researchers to cover the basic vocabulary of English. A study by Laufer and Ravenhort-Kalovski (2010), which focused specifically on reading ability, suggests that knowledge of 4,000 to 5,000 word families (corresponding to about 95% of text) allows learners to read with some guidance, whereas a vocabulary of around 8,000 word families (corresponding to around 98% of text coverage) is required for learners to read independently without reliance on dictionaries.

## 5.3 Instruments

A one-page worksheet was created to introduce participants to the concept of collocation and register (i.e., formal and informal English). The corpus training materials were similar to those used in Experiment 1 except that the examples, after one initial example using the BNC, all used the COCA corpus. The first part of the worksheet taught participants how to use the chart

Table 2 TOEFL ITP (Level 1) scores for Experiment 2 participants

	Score scale	<i>M</i> ( <i>SD</i> )	Range
Total score	310–677	419 (35)	360–507
Listening comprehension	31–68	43 (3)	34–50
Structure and written expression	31–68	42 (3)	37–51
Reading comprehension	31–67	40 (7)	31–51

function to determine the genres in which a word occurs most often. The examples focused on distinguishing words associated with informal speech from words that were appropriate for academic writing. The next part of the worksheet concentrated on using the LIST display while searching for a word's collocates by designating the search span (to the left or right) and the part of speech of the collocate, and the participants were shown how to identify optimal spans. Two heuristics were introduced: (1) the strategy of disregarding intervening words if possible so as to search for the minimal feasible span; and (2) the strategy of incrementally considering larger spans if the smaller span yields too many results.

The participants were also taught how to categorize search results in terms of features such as semantic prosody (i.e., positive or negative connotations; for a discussion, see Stubbs, 1995) and patterns of use. For example, searching for the nouns that appear as a subject before the verb *occur*, they were led to see that the nouns are often negative (e.g., *injuries*, *accidents*) and that many of the nouns are associated with health problems and accidents. The final section of the worksheet had participants apply what they had learned to several practice problems. This section was assigned as homework and was designed to take about an hour to complete. For review after the corpus workshop, a one-page worksheet was created with ten gap-fill problems targeting collocations and register (informal vs. academic) issues.

For the in-class corpus training and testing, participants had access to individual computers running Windows with Internet access. All participants also had access to their own stand-alone electronic dictionary. Because participants use electronic dictionaries extensively in their classes, all owned expensive models with large dictionaries. Both L1-L2 and L2-L1 dictionaries were used to complete the experimental tasks.

To assess participants' revision skills, two tests were created. Both Test Form A (Appendix B) and Test Form B had 12 items, with four items corresponding to each of the following three categories: (1) preposition use; (2) use of light verbs (e.g., *do*, *make*, *take*) before nouns; and (3) choice of a lexical item appropriate for academic writing. Based on previous research (e.g., Koo, 2006; Tono, Satake & Miura, 2014), these item types were chosen so as to reflect common issues in EFL students' writing that were likely to be solved through the use of a corpus search.

The four light verb items on the two test forms varied so that the items had zero, one, three, or four intervening content words. For example, the collocations *do laundry*, *make noise*, *have a bite*, and *give a description* occurred as follows:

- *Jane usually \_\_\_\_\_ laundry...* (zero)
- *When I turn on my computer, it \_\_\_\_\_ a strange noise.* (one)
- *... he \_\_\_\_\_ only two small bites of the cake.* (three)
- *Sally \_\_\_\_\_ the local police a very detailed description of the man.* (four)

Unlike Experiment 1, Experiment 2 used a typical test format with blanks. As in Experiment 1, a survey was created, but with slightly revised items.

#### **5.4 Procedure**

In the week prior to the experimental tasks, all participants received a 30-minute training session on collocations and the differences between formal and informal English. One week

later, the Block A participants (who were all in the same English reading class) took the gap-fill test in the Dictionary condition. The choice to provide training only for the Corpus condition was based on the fact that the participants had years of experience using dictionaries and had extensive training in dictionary use, prior to Experiments 1 and 2, in their required English reading course. Additionally, the participants all had access to high-quality electronic dictionaries and were very adept at using them.

Test forms were counterbalanced so that half the participants received Form A, and half, Form B, in the two conditions. In the following week, the participants received a 90-minute training session on use of COCA. As in Experiment 1, they went through exercises in a computer room as a class and then practiced individually. At the end of the session, they were given homework on using the corpus, designed to take approximately an hour to complete. One week later, they returned to the computer room to carry out a short review and then complete the gap-fill test in the Corpus condition. Block B participants (who were all in a different English reading class from those in Block A) did the same experimental tasks, but they completed the corpus training and test prior to taking the test using a dictionary. At the end of these sessions, both blocks were given a survey asking them about their impressions of using the corpus. The sequencing of tasks is shown in Table 3.

## 6 Results (Experiment 2)

The descriptive statistics for the gap-fill tests in the Corpus condition ( $n = 39$ ) are shown in Table 4. There were 12 items on the test with four in each category of prepositions, light verbs, and register.

The descriptive statistics for the Dictionary condition ( $n = 39$ ) are shown in Table 5.

As can be seen, the participants in the Dictionary condition were accurate on just over half the items, while scores in the Corpus condition were slightly higher. Advantages for the Corpus condition can be attributed to performance on light verb items and, to a lesser extent, preposition items. The standard deviation was slightly higher for register items in the Corpus condition, suggesting some variation in participants' ability to use the corpus to distinguish academic from informal registers. Judging from scribbled notes in the margins of their test papers, several participants had used the chart function (as instructed during training) to examine candidate words' appearance in various genres, but they simply compared the frequency at which each word appeared in academic texts. This strategy worked if the words being compared had similar frequencies in the corpus, but it failed when words differed in terms of overall frequency. Participants should have instead examined each word's relative tendency to appear within each genre.

It could also be that dictionaries are quite useful in determining which registers and genres are associated with a word. Generally speaking, lexicographers who create L2-L1 bilingual dictionaries are acutely aware of the importance of register as an aspect of word usage and therefore tend to provide L1 equivalents that have register and genre profiles similar to the L2 word being defined. The participants, who have extensive experience using dictionaries, were evidently able to determine the English words' association with an informal or formal register based on this information.

In the statistical analysis using a repeated measures design, Revision Tool served as a within-subjects independent variable with two levels: Corpus and Dictionary; gap-fill test scores

Table 3 Sequencing of Experiment 2 tasks: Block A

Day 0	Day 7	Day 14	Days 14–21	Day 21
collocations training	gap-fill test (using dictionary)	corpus training	corpus homework	corpus review
discussion of formal vs. informal English				gap-fill test (using COCA) survey
<i>Sequencing of Experiment 2 tasks: Block B</i>				
Day 0	Day 7	Day 7–14	Days 14	Day 21
collocations training	corpus training	corpus homework	corpus review	gap-fill test (using dictionary)
discussion of formal vs. informal English			gap-fill test (using COCA)	survey

Table 4 Descriptive statistics for gap-fill tests in the Corpus condition

	%	<i>M</i>	<i>SD</i>	Range
Total	63.0%	6.9	1.9	2–11
Prepositions	55.1%	2.2	0.8	0–4
Light verbs	65.4%	2.6	0.8	1–4
Register	68.6%	2.7	1.3	0–4

Table 5 Descriptive statistics for gap-fill tests in the Dictionary condition

	%	<i>M</i>	<i>SD</i>	Range
Total	56.4%	6.8	2.1	3–12
Prepositions	47.4%	1.9	1.0	0–4
Light verbs	53.2%	2.1	1.1	0–4
Register	68.6%	2.7	1.1	1–4

served as the dependent variable. All items were scored as worth one point. A paired-samples *t*-test was conducted to compare the scores in the Corpus and Dictionary conditions. At an alpha of  $p = 0.05$ , two-tailed, participants scored significantly higher in the Corpus condition,  $t(38) = 2.25$ ,  $p = 0.030$ , 95% CI[0.1, 1.5],  $r^2 = .118$ . As can be seen from the *r*-squared value, only a small percentage of the variance is explained by the independent variable Revision Tool. Due to the low number of items in the three categories of the test (prepositions, light verbs, and register), no statistical analysis of these categories was conducted.

In the survey, the participants all reported that they had no previous experience using an online corpus. A seven-point Likert scale was used to ask them whether: (1) the corpus was easy to use; (2) they thought they would use a corpus in the future to learn more about specific words; (3) the corpus was useful for making revisions; and (4) they understood the sentences that appeared in the corpus. The descriptive statistics for the Likert-scale items are shown in Table 6.

The final three open response items, which could be answered in either English or Japanese, asked what participants liked or disliked about using the corpus, what they found most difficult about using the corpus, and whether they would like to learn more about using the corpus for revision – and if so, which areas of English they would like to focus on.

One area of interest in the study was whether the participants experienced difficulty using the corpus and what particular aspects of using the corpus they found difficult. Most participants reported considerable difficulty. Among participants who elaborated on the causes of this, six mentioned general problems in using the computer (two participants mentioned that they do not own a computer); one specifically mentioned difficulty negotiating the interface, as the webpage elements all appeared in English. Other difficulties included trouble understanding how to determine the appropriate span when searching for collocates (two responses), difficulty using the chart function, and trouble choosing the best word for a query. This latter response is understandable as searches require some intuitions regarding the strength of association between two words. The middling responses to Item #4

suggest that part of the difficulty may have been a result of working with naturalistic data, which, in the case of corpus query responses, consist of short sentences removed from their original context. Limitations in participants' proficiency undoubtedly made it difficult for them to evaluate some of the corpus search results (cf. Park, 2012) and choose a solution within the time allotted for the task. Participants' difficulties in this regard were undoubtedly exacerbated by their lack of knowledge and experience using a corpus.

In their written comments, five participants mentioned specifically that they liked using a corpus to search for the appropriate preposition. This fits in well with the gap-fill test results showing a 7.7% boost in performance when using the corpus instead of a dictionary for preposition queries. It is odd that no participants mentioned the advantage of using the corpus to determine collocations involving light verbs, but this could simply reflect their lack of metalinguistic terminology to describe this area of language. Three participants mentioned a desire to search for more appropriate academic language using the corpus.

The participants' performance on the light verb items was compared in terms of the number of intervening content words; the results are shown in Figure 1. In both conditions, performance dropped when there were many intervening words between the missing light verb and the collocating noun. The effect of intervening words is similar in both conditions; however, performance on the light verb items with no intervening words was higher in the Corpus condition. Because only two items (one for each test form) represent each of the four categories, interpretations related to this section of the results should be regarded as tentative.

## 7 General discussion

The qualitative results of the two experiments suggest that freshmen EFL students in Japan, after receiving a short training in the use of an online corpus, have fairly positive attitudes toward corpus consultation, although this enthusiasm is offset by reported difficulty in learning to conduct and analyze searches. The quantitative results of Experiment 2 show that in spite of these difficulties, corpus consultation leads to an enhanced ability to solve language issues when these issues have been identified as problems (as they were in the gap-fill test). Among the three areas of language examined, the greater effectiveness of the online corpus appears to be in identifying the light verbs that collocate with a given noun and the prepositions that collocate with a given verb or noun. Dictionaries and corpus consultation appear to be equally useful for determining whether a word is associated with academic register.

The results thus support Boulton's (2009) findings that corpora, compared to conventional reference resources, can improve learners' capacity to solve certain language

Table 6 *Descriptive results of Experiment 2 survey results (7-point Likert scale)*

Item	Description	<i>M</i> ( <i>SD</i> )	Range
1	Ease of use	2.8 (1.4)	1–6
2	Future use for word queries	4.4 (1.7)	1–7
3	Usefulness for making revisions	3.8 (1.5)	1–7
4	Understanding of sentences in query responses	3.4 (1.3)	2–7

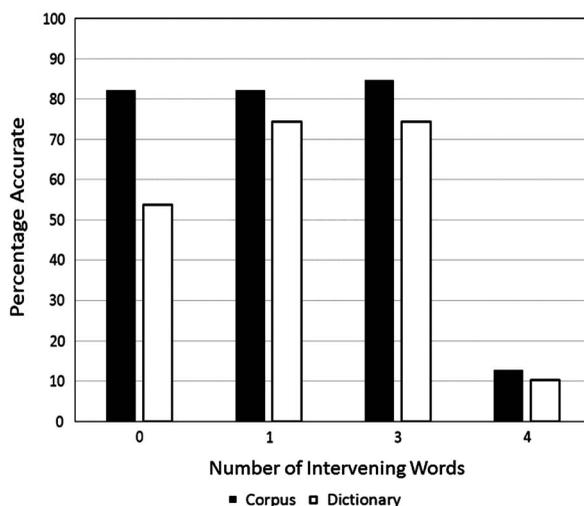


Fig. 1. Graph showing the performance of participants on light verb items.

problems. However, unlike the Boulton study, in which participants used prepared corpus lines, the participants in the current experiments used an online corpus. If we consider Kennedy and Miceli's (2001) analysis of corpus consultation in terms of four steps (i.e., formulating a question, devising a search, observing results, and drawing conclusions), the Boulton study can be regarded as a demonstration that corpus consultation has advantages over conventional resources when learners are required to perform only the last two of the four steps, whereas the current study shows an advantage for corpus consultation even when learners are also required to devise a search (i.e., when they must perform all steps except for formulating a question).

Recent qualitative research on corpus consultation for error correction (Chang, 2014; Koo, 2006) strongly suggests that corpus consultation has some advantages over conventional resources for error correction, especially in the case of learners with higher proficiency and research skills. Their findings suggest that this is so even when the learners must perform all four of the steps mentioned above. One advantage of these studies is their generalizability to typical L2 writing contexts; however, further quantitative research is needed to confirm these findings. Unfortunately, the question of the effectiveness of corpus consultation for all four steps is difficult to address using a quantitative design. Ideally, a study would include all steps of corpus consultation, but this is generally impractical, as individual learners' L2 knowledge and errors are different. More specifically, there is a methodological problem in establishing an appropriate dependent variable; the same list of targets cannot be presented to a group to assess each learner's ability to identify and correct his or her own errors. A more practical solution may be to break the problem into two parts: (1) an assessment of learners' ability to accurately identify their own errors (the first of the four steps outlined by Kennedy and Miceli); and separately, (2) an assessment of their ability to autonomously perform the remaining three steps, as examined in this study.

Future research in this area should continue to explore the use of online corpora, as these tools hold great promise in terms of fostering autonomy and boosting learner agency. While online corpora pose a significant initial learning hurdle, they ultimately provide learners

with another powerful tool to solve language issues that are often imperfectly treated in conventional reference materials. Yet it must be acknowledged that learners' autonomous use of online corpora brings with it a number of issues.

One problem to be resolved is how to make online corpora more accessible to lower-proficiency learners such as those in this study. Some of the participants reported difficulty interpreting corpus results due to limitations in their lexical knowledge. One confusing aspect of searching corpus lines, especially for learners with lower proficiency, is choosing the appropriate collocation pattern when more than one appears in the results. Tono *et al.* (2014: 159) give the example of a participant who mistakenly used *of* when filling in the missing word in "he will graduate \_\_\_ his university"; the participant was apparently confused by the preponderance of "graduate of" in the corpus results, ignoring the fact that *graduate* can occur as both a verb and a noun. More advanced online tools such as COCA are useful as they allow part-of-speech (POS) searches that can, if used correctly, prevent such confusion.

Another issue concerns the development of manuals on best practices for teachers, who sometimes experience difficulties using corpus tools in the classroom (Bunting, 2013). Such manuals need to go beyond mere technical descriptions so as to address the practical heuristics that best solve the types of language problems that learners frequently encounter. These heuristics must take into account the typical abilities of language learners at a wide range of proficiency levels. Pedagogical advice also needs to incorporate findings regarding the amenability of various language problems to corpus consultation (see, for example, Gaskell & Cobb, 2004; Tono *et al.*, 2014). Fortunately, some recent work (e.g., Kennedy & Miceli, 2010) has provided advice on how corpus workshops can be used to foster learners' corpus consultation strategies.

Finally, while the current study examined dictionary and corpus consultation in isolation, more quantitative research is needed to determine optimal methods for combining the two (cf. Kennedy & Miceli, 2010; Koo, 2006). In the survey results of both experiments, participants often reported frustration in choosing between corpus lines. One potentially effective strategy to overcome this problem would be to search for the word in a dictionary to determine whether the meaning of the word in the corpus results was that intended in the text to be revised. Another strategy could involve a follow-up search to verify that the word found in a dictionary search often appeared as a collocate of the other adjacent content words in the sentence, and, if this was the case, whether it had the intended meaning within those contexts. In other words, students may need more alternative strategies that can be employed when their initial search data do not lead them to a clear answer.

## 8 Conclusion

This paper addressed the feasibility of online corpus consultation by low-proficiency L2 learners who are attempting to repair errors or find solutions to questions about appropriate language use. The results provide positive news for those considering the introduction of online corpora to their students: online corpus consultation, even for learners at fairly low proficiency levels, appears to have practical benefit in enhancing learners' ability to solve language issues. This suggests that online corpus consultation should also be of use to learners attempting to correct errors in their writing.

These practical benefits are likely to be accompanied by several additional benefits not explored in this study. According to Swain (1995), one important function of output practice such as writing is that it encourages metalinguistic reflections on language. Corpus data may be particularly conducive to fruitful metalinguistic reflections, particularly when the linguistic targets involve syntagmatic patterns or the association of words with a particular genre or time period. For this reason, future research on corpus consultation should also examine its effects on learners' metalinguistic knowledge.

The current study has a number of limitations. First, the corpus training was of much more limited duration than that of some of the qualitative studies that have investigated the same topic (e.g., Chang, 2014). It therefore is not clear whether the difficulties reported by a considerable number of participants were inevitable given their proficiency level or whether these difficulties would have been largely ameliorated if more training and greater opportunities for practice had been provided. Second, the effectiveness of online corpus consultation relative to the use of conventional references is difficult to characterize without a more detailed understanding of (1) the types of errors learners of different backgrounds make; (2) the proportion of errors they recognize as questionable; and (3) the proportion of these errors that they can repair using online corpora or other tools such as dictionaries, thesauri, or specialized manuals on usage. With these caveats noted, the current study provides evidence that an online corpus can be at least as useful as a dictionary, if not more so in some circumstances. Considered together with previous research, the study suggests that educators should consider introducing online corpora to a wider range of students learning a foreign language.

## References

- Ädel, A. (2010) Using corpora to teach academic writing: Challenges for the direct approach. In M.-C. Campoy, B. Bellés Fortuño and M.-L. Gea-Valor (eds.), *Corpus-based approaches to English language teaching* (pp. 39–55). London: Continuum.
- Aston, G. (2001) Learning with corpora: An overview. In G. Aston (ed.), *Learning with corpora*. Houston, TX: Athelstan, 7–45.
- Bachman, L. F. (1990) *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Boulton, A. (2008) Looking for empirical evidence of data-driven learning at lower levels. In B. Lewandowska-Tomaszyk (ed.), *Corpus linguistics, computer tools, and applications: State of the art*. Frankfurt: Peter Lang, 581–598.
- Boulton, A. (2009) Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, **21**(1): 37–54.
- British National Corpus*, version 3 (BNC XML Edition). (2007) Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Retrieved from <http://www.natcorp.ox.ac.uk/>.
- Bunting, J. D. (2013) *An investigation of language teachers' explorations of the use of corpus tools in the English for Academic Purposes (EAP) class*. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3591056).
- Chambers, A. and O'Sullivan, Í. (2004) Corpus consultation and advanced learners' skills in French. *ReCALL*, **16**(1): 158–172.
- Chang, J.-Y. (2014) The use of general and specialized corpora as reference sources for academic English writing: A case study. *ReCALL*, **26**(2): 243–259.
- Davies, M. (2008–) The Corpus of Contemporary American English: 450 million words, 1990–present. Available online at <http://corpus.byu.edu/coca/>

- Gaskell, D. and Cobb, T. (2004) Can learners use concordance feedback for writing errors? *System*, **32**(3): 301–319.
- Geluso, J. and Yamaguchi, A. (2014) Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, **26**(2): 225–242.
- Gilmore, A. (2008) Using online corpora to develop students' writing skills. *ELT Journal*, **63**(4): 363–372.
- Granger, S., Dagneaux, E., Meunier, F. and Paquot, M. (eds.). (2009) *International corpus of learner English, Version 2*. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Kennedy, C. and Miceli, T. (2001) An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology*, **5**(3): 77–90.
- Kennedy, C. and Miceli, T. (2010) Corpus-assisted creative writing: Introducing intermediate Italian learners to a corpus as a reference resource. *Language Learning & Technology*, **14**(1): 28–44.
- Koo, K. (2006) *Effects of using corpora and online reference tools on foreign language writing: A study of Korean learners of English as a second language*. (Unpublished doctoral dissertation). University of Iowa, Iowa City, IA.
- Laufer, B. and Ravenhorst-Kalovski, G. C. (2010) Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, **22**(1): 15–30.
- Nation, P. and Beglar, D. (2007) A vocabulary size test. *The Language Teacher*, **31**(7): 9–12.
- 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.
- O'Sullivan, Í. and Chambers, A. (2006) Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing*, **15**(1): 49–68.
- Park, K. (2012) Learner-corpus interaction: A locus of microgenesis in corpus-assisted L2 writing. *Applied Linguistics*, **33**(4): 361–385.
- Simpson, R. C., Briggs, S. L., Ovens, J. and Swales, J. M. (2002) *The Michigan Corpus of Academic Spoken English*. Ann Arbor, MI: The Regents of the University of Michigan.
- Stubbs, M. (1995) Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language*, **2**(1): 23–55.
- Swain, M. (1995) Three functions of output in second language learning. In G. Cook and B. Seidlhofer (eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson*. Oxford: Oxford University, 125–144.
- Thornbury, S. (2002) *How to teach vocabulary*. Essex: Pearson Education.
- Todd, R. W. (2001) Induction from self-selected concordances and self-correction. *System*, **29**(1): 91–102.
- Tono, Y., Satake, Y. and Miura, A. (2014) The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL*, **26**(2): 147–162.
- Viana, V. (2010) Authentic English through the computer: corpora in the ESOL writing classroom. In S. Kasten (ed.), *Effective second language writing*. Alexandria, VA: TESOL, 163–168.
- Widdowson, H. G. (1990) *Aspects of language teaching*. Oxford: Oxford University Press.
- Yoon, H. (2008) More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning & Technology*, **12**(2): 31–48.
- Yoon, H. and Hirvela, A. (2004) ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, **13**(4): 257–283.

## Appendix A

### “Student Essay” (form A) used in Experiment 1

As human beings, we all strive to get physical, mental, and spiritual health.<sup>1</sup> These three areas of personal health are an essential component of happiness. Fortunately, exercise

provides us with a cost-effective way to improve our overall state of health and increase our general sense of well-being.

Exercise's promotion of physical health is well known. Exercise, particularly those activities that involve the entire body, has an invigorating effect to the body.<sup>2</sup> From a physiological point of view, exercise helps make a healthy cardiovascular system.<sup>3</sup> This is especially important in our modern lives, in which we move around so little in the day.<sup>4</sup>

An often ignored strong point<sup>5</sup> of exercise involves mental health. Yet this may be the most important benefit. Exercise, after all, provides an important distraction from everyday concerns and, as result, does much to kill off stress.<sup>6</sup> After a quick exercise session, people generally feel that their problems disappear or at least are forgotten for a while. Research also shows that exercise helps us to think more clearly.

Of course, exercise is never easy, and most people often have times when they don't take care<sup>7</sup> their exercise routine. The important thing is that we do an effort.<sup>8</sup> Over time, as we push ourselves to change our former habits and exercise more, we learn how to gain better control on our own thought processes.<sup>9</sup> In a sense, the daily struggle to improve ourselves is directly related to our own spiritual development. As we get discipline<sup>10</sup> we learn to take responsibility and become better people in the process.

In short, exercise makes us healthy in body, mind, and spirit. Of course, we should not expect too much at first. Remember that even a short amount of exercise can have a difference.<sup>11</sup> And as we begin to exercise more, we will gain the satisfaction of accomplishment, the pride in achieving what we originally set out to do. Moreover, our own healthy and wholesome lifestyles can inspire those by us<sup>12</sup> to lead healthier lives. In my case, I remember when I started exercising. The lady I work with said to me a compliment,<sup>13</sup> saying that I looked much younger recently. So the next time you find yourself bored and listless, sitting in front of the TV, get up and go for a long walk or a jog through the park. Or invite a friend to go play tennis! If you take more time<sup>14</sup> doing exercise, you'll be healthier, happier, and more fulfilled. So start exercising daily and most importantly, keep on it!<sup>15</sup>

1. "Get" is slightly awkward here, and it isn't an academic word.
2. "Effect to the body" sounds odd. The preposition "to" isn't right here.
3. "Helps make" is awkward and doesn't sound very academic.
4. "In the day" sounds strange here. The preposition "in" isn't right.
5. "Strong point" sounds awkward here.
6. I understand what you're saying here (more exercise = less stress), but "kill off stress" is awkward. The word "stress" is okay here, but "kill off" sounds odd.
7. "Don't take care" is awkward.
8. I understand what you mean, but "do an effort" sounds strange. The word "effort" is okay, but "do" sounds strange.
9. "Control on thought processes" sounds odd. The preposition "on" isn't correct here.
10. "Get discipline" sounds awkward.
11. "Have a difference" is awkward. The word "difference" is OK, but "have" sounds odd.
12. "Inspire those by us" sounds strange. The preposition "by" doesn't work here.
13. The word "compliment" makes sense here, but "say a compliment" sounds strange.
14. "Take more time" sounds strange. I don't think you want the word "take" here.
15. I can see what you mean to say here, but the phrase "keep on it" is awkward.

## Appendix B

### Experiment 2 gap-fill (test form A)

#### Part 1

Directions: *Fill in the blank with the preposition that best completes the sentence.*

Example:

Questions: The cup is    the table.

Answer: The cup is on the table.

1. Alcohol use    teenagers is rising.
2. Although she got many "A"s, she wasn't satisfied    her grades.
3. The cat was crying, so Tom got a ladder and climbed    the roof to rescue it.
4. The store's much better now that it's    new management.

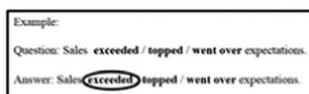
#### Part 2

Directions: *Fill in the sentence with the verb that best completes the sentence.*

5. Sally    the local police a very detailed description of the man who robbed the bank.
6. Jane usually    laundry on the weekend.
7. He was on a diet, so he    only two small bites of the cake.
8. When I turned on my computer, it    a strange noise.

#### Part 3

Directions: *The following sentences appeared in formal writing. Circle the word that best completes the sentence. Keep in mind that the word you choose should sound formal.*



9. The government's new policies involve a **great deal of/lots of/tons of** risk.
10. Young **children/kids/whippersnappers** usually learn foreign languages more slowly than older children and adults.
11. Compared to the poor and middle class, **loaded/wealthy/well-heeled** people generally spend less of their monthly income.
12. Childhood **flab/obesity/paunchiness** has become a serious problem in the U.S. due to the overconsumption of sweets and other high-calorie foods.