Journal of English for Academic Purposes 39 (2019) 97–108

Contents lists available at ScienceDirect

Journal of English for Academic Purposes

journal homepage: www.elsevier.com/locate/jeap

How useful are corpus tools for error correction? Insights from learner data

Natalia Dolgova [a, *], Charles Mueller [b]

[a] English for Academic Purposes (EAP) Program, George Washington University, 801 22nd St NW #610, Washington, DC, 20052, USA
[b] Fuji Women's University, Sapporo, Hokkaido, Japan

ARTICLE INFO

Article history:
Received 12 May 2018
Received in revised form 22 March 2019
Accepted 25 March 2019
Available online 28 March 2019

ABSTRACT

This study contributes to the growing literature on data-driven learning in ESL/EFL instruction, examining how advanced English learners choose to use corpus tools for error correction on their own following classroom training. Controlled L2 production and survey data from U.S.-based EAP students were collected over the course of four consecutive semesters and then analyzed quantitatively and qualitatively. The research examined how learners approach the correction of different type of errors using corpus tools and which types of errors could be more successfully corrected with these tools. The results of a Chi Square Test indicated that participants' success at error correction depended on the type of error being addressed ($p < .001$). While local lexicogrammatical errors were chosen by learners to be corrected in over half of all participant revisions, these corrections were inappropriate more often than corrections of register errors. Survey data and qualitative analysis of errors further revealed that participants often struggle with identifying the context of the concordance lines and sorting through multiple lines to identify patterns. The paper concludes with a discussion of implications and pedagogical recommendations.

1. Introduction

1.1. Research on corpora in L2 classrooms

Over the last two decades, successful use of corpus tools has been well-documented in research literature focusing on language use and applications to classroom activities. This research reflects a greater appreciation of corpora utilization in both materials design and as a novel means of promoting learner agency. Leech (1997) has thus proposed the distinction between *indirect* and *direct* types of pedagogical corpus applications: while *indirect* applications inform syllabus and material design for teachers, *direct* applications include hands-on activities and tasks that ideally promote learner independence and autonomy.

In direct approaches, DDL, or data-driven learning (Cresswell, 2007; Geluso & Yamaguchi, 2014; Johns, 1991), has been widely used in language pedagogy over the course of the last few decades. This approach places students at the center of a learning process in which they gain agency as they pose a research question, observe corpus output, and draw generalizations based on observed trends. While DDL is often generally understood as "using the tools and techniques of corpus linguistics for

* Corresponding author.
E-mail addresses: natalia@gwu.edu (N. Dolgova), mueller@fujijoshi.ac.jp (C. Mueller).

https://doi.org/10.1016/j.jeap.2019.03.007
1475-1585/

pedagogical purposes" (Gilquin & Granger, 2010, p. 359), previous studies have also explored the use of DDL for teaching and learning of specific target language structures, such as linking adverbials (Boulton, 2009), verb-adverb (Daskalovska, 2015) and verb-noun (Chan & Liou, 2005) collocations as well as collocations at large (Chon & Shin, 2013; Hyland, 2008), abstract nouns (Huang, 2014), light verbs and prepositions (Mueller & Jacobsen, 2016), and passive voice (Smart, 2014). When discussing DDL, it is useful to differentiate between deductive and inductive approaches (Flowerdew, 2012), which are discussed in greater detail in the next section.

1.2. Deductive vs. inductive approaches to L2 corpus use

As can be inferred from the term itself, *deductive*, or top-down, approaches entail learners search for data available in corpora that allow them to confirm previously known, explicit hypotheses about a given language structure. It has been suggested in recent studies (e.g., Charles, 2007; Liu & Jiang, 2009) that deductive corpus searches should serve as a foundation for students learning to work with and navigate a corpus interface. If students are given hypotheses they need to confirm or refute upfront, they are likely to feel better supported and more confident engaging in the subsequent corpus search process.

On the other hand, *inductive*, or bottom-up, approaches place the focus on learners themselves under the assumption that learners do not need a pre-formulated hypothesis and can observe language usage in context, based on which they can generate rules on their own and subsequently modify their own production in accordance with such rules (Flowerdew, 2009; Gavioli, 2005; Hyland, 2008; Meunier, 2002). To capture the essence of a more extreme version of this approach, Gavioli (2005) used the metaphor of "spies" to refer to ESP students learning about practices of a discourse community in order to infiltrate it as they formed research questions and searched for answers in a corpus. The main challenge of the inductive approach is the significant learning required to navigate corpus output and identify relevant information.

Some would argue that deductive and inductive approaches, instead of being viewed as opposing each other, should be regarded as complementary. For instance, Charles (2007) advocates reconciling top-down and bottom-up instructional approaches to provide more balanced instruction. When combined, these two approaches enable students to focus both on function (while analyzing sentences in context) and on form (while analyzing the semantics of morpho-syntactic patterns occurring within sentences). Like Charles (2007), Mizumoto and Chujo (2016) adopt an inclusive view, suggesting that the choice of approach depends on the type of learning task.

The Fig. 1 below summarizes how both inductive and deductive approaches to corpus investigations can be used in an integrated manner.

1.3. Remaining issues

Despite many documented benefits, a number of pedagogical challenges to employing corpus tools for language learning and teaching remain. First, many teachers working in EFL and ESL settings are unfamiliar with corpus interfaces and tools. Also, a considerable investment of time is often required for learners to gain familiarity with corpora for purposes of language learning (Frankenberg-Garcia, 2012), especially when learning to use the corpus interface, as documented by the vast majority of researchers who have written on the topic (see Chang, 2014, for a recent and comprehensive overview). To provide an example, Liu and Jiang (2009) point out that students may struggle with sorting through copious output and irrelevant examples and with processing unknown words in search results. Chan and Liou (2005) indicated that learners who may be comfortable with or successful at most corpus search tasks may have an underlying preference for inductive learning. These problems pose challenges to educators attempting to incorporate corpora into instructional contexts based on rigid or pre-established syllabi that do not allocate adequate time for instruction of corpus search techniques.

Echoing these concerns, Schenk and Cho (2012) found that the use of corpus methodology is not always effective and that even advanced learners of English did not "substantively apply observations from the corpus to the editing process, despite

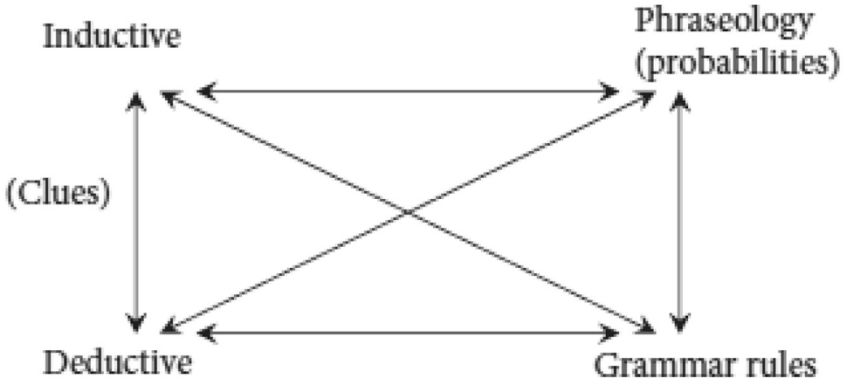

Fig. 1. Flowerdew's (2009) dynamic paradigm for corpus investigations (p. 407).

systematic guidance and extensive prior experience in the English field" (p. 174). In other words, even though the learners in their study acknowledged the benefits of corpus use, they were unable to appropriately utilize corpus findings to improve their own final written product. Additionally, Pérez-Paredes, Sánchez-Tornel, Alcaraz Calero, and Jiménez (2011) state that insufficient research has been done to date to offer adequate insight into corpus user experience, and it is thus difficult to foresee and account for all possible challenges in advance. Until additional research proves otherwise, Boulton's (2009) recommendation that corpora should be used in combination with other methods, materials, and strategies remains valid.

1.4. How do learners engage with corpora for correcting errors?

Notwithstanding the above-mentioned challenges, some research suggests that corpus tools can be useful for error correction in the L2 writing process. In fact, according to Gilquin and Granger (2010), one of the main functions of data-driven learning is to facilitate error correction processes. Flowerdew (2010) likewise suggests that "corpus consultation seems to be most effective for the revising process" (p. 452). Lastly, Quinn (2015) points out that the focus on correction of individual learner errors helps learners easily see the benefit of corpus use as they are able to see direct applicability of the corpus to their own language learning.

In an empirical investigation of error correction processes, Gaskell and Cobb (2004) examined how lower-intermediate English learners used concordances for correcting errors and found that learners achieved much greater success following instructions for using online concordance links for error correction. As would be expected, without instructions, success rates were significantly lower. As for error types, the language features most easily corrected with corpus tools were word order, capitals/punctuation, and pronouns, while errors involving articles, subject-verb agreement, and noun plurals were not affected by corpus tools and, in fact, increased.

Also, Tono, Sataki, and Miura (2014) focused on just one aspect of corpus use, corpus-assisted error correction accuracy in controlled settings. They found that omission and addition errors were easily corrected with corpus tools but misformation errors were not. As there exist many intervening variables affecting potential success, such as learner proficiency, type and length of corpus instruction, and so on, the authors call for more research on applications of data-driven learning to L2 writing.

Most recently, Bridle (2019) investigated, among other foci, how error types and learner personality may contribute to learners' choice to rely on corpus tools for correction. The results demonstrated that learners chose to use corpus tools for the "Wrong Word" and "Formal/Informal" errors (Bridle, 2019, p. 63). Other error types, designated by the author as "Articles", "Grammar", "Incomprehensible", and "Missing Words" accounted for a minor portion of correction attempts using corpus tools; instead, learners relied on their own knowledge or on dictionaries/reference books to correct those error types. Grammatical errors also resulted in a greater number of incorrect corpus-based corrections.

As can be seen from these studies, most researchers have limited the number or range of errors learners could consider for correction in order to provide necessary limits to the process of searching the corpus. Written corrective feedback was typically used as the main method for directly (pointing out type) or indirectly (pointing out error but not type) identifying errors for the students.

However, despite the fact that errors are typically made salient to students, successful processing of the corpus data and subsequent internalization of corpus-derived rules may be dependent on a number of factors. Park (2012) analyzed the microgenesis process in learner-corpus interactions, which focused on specific episodes in the three student participants of the study that led to qualitative changes in learning. He found that "learner's achievements depend both on their ability to interpret and exploit the search results" (p. 361) and thus concluded that corpus use is beneficial to learners only when they take proper time and effort to "evaluate search results based on careful analysis" (p. 381). Without such analysis, positive outcomes often fail to appear.

It remains essential to continue exploring learners' experiences and the processes they rely on when sorting through corpus findings. While Park's (2012) study provided significant qualitative insights into the process, it is important to gain greater awareness of how learners go about correcting errors on their own outside of class and over instructional time intervals exceeding a few corpus training sessions. Indeed, engaging learners in out-of-class corpus activities has been put forward as a needed area of research related to the applications of corpus tools to writing instruction (Flowerdew, 2010).

The present study thus aims to investigate the effectiveness of advanced learners' independent use of corpus tools for error correction to gain insight into their long-term uptake of appropriate corpus use strategies. It is hoped that the study will shed light on the usefulness of corpus tools for developing learners' agency in correcting their errors appropriately. The details pertaining to the study's design and procedure are outlined in the next section.

2. Methods

Combining qualitative and quantitative research approaches, the present study aims to investigate how learners choose to use corpus tools on their own for error correction purposes. Qualitatively, this study explores how students approach and experience the use of corpus tools, as conveyed through survey responses. The quantitative research focuses on the analysis of L2 production data and aims to determine the types of errors that appear to be most conducive to corpus-based correction. The specific research question (RQ) is as follows:

RQ: For which error types is the use of corpus tools most effective as measured by the successful learner uptake of correct forms?

In this context, uptake is operationalized as providing an appropriate written correction for a previously identified error.

2.1. Study context and procedure

The study took place in the context of English for Academic Purposes (EAP) courses. The participants were graduate students in the sciences (computer science, engineering, etc.) enrolled in a required upper-level EAP course at a large private U.S. university. All participants had a TOEFL iBT score between 80 and 104, signifying that their proficiency fell into the upper intermediate to advanced range. The choice of population was based on the assumption that their advanced language proficiency would increase the likelihood that they possessed sufficient metalinguistic awareness needed for the successful use of corpus tools. Furthermore, it was expected that background in STEM disciplines and familiarity with analyzing data patterns would allow the participants to more readily appreciate the potential benefits of applying data-driven language techniques to their own language learning. Moreover, because these students would not receive any systematic English language support beyond their EAP classes, it was felt that they would be motivated to explore the potential behind corpus tools in order to further their independent language learning in the future.

2.2. Procedure

The study took place over the course of four consecutive semesters with students from intact EAP classes. Initially, 175 participants took part in the study. All classes were taught by the same instructor (one of the authors). Within each semester, the data collection process lasted six weeks and included the following three parts:

1) A quasi-experimental procedure tested the participants' prior knowledge and ability to use corpus tools. The data were collected using pre- and post-tests preceding and following a corpus training session.
2) An online survey, focused on the participants' impressions of the corpus technology, was distributed the week following the post-test.
3) A mixed-method follow-up phase consisted of qualitative and quantitative analysis of L2 writing errors of self-selected groups of participants produced during the three weeks following the completion of corpus treatment. The participants were instructed to use corpus tools to correct some of the previously identified errors from their final paper drafts.

The details for each of the parts are described in separate sections below.

2.3. Materials

Part 1: Corpus use training. Pre-test and post-test forms during this stage focused on the participants' ability to use corpus tools for language learning and error correction. The tests consisted of fill-in-the-blank questions within naturally-occurring discourse contexts; a few representative test questions are available in Appendix A. There were 25 items on each test; all of them were adapted from the Corpus of Contemporary American English (COCA) discourse excerpts and focused on typically problematic grammatical structures. Specifically, the following structure types – preposition choices, phrasal verbs, common verb-noun collocations (including register-specific ones), subject verb agreement, and verb tense/form choice – were represented through five test items each. All excerpts came from highly specialized contexts such as academic journals to minimize the possibility of the participants' prior familiarity with the vocabulary. Participants were asked to provide answers using any available tools, such as dictionaries and online corpora. All tests were timed and had to be completed within 30 minutes.

Corpus use training. The pre-test was followed by basic training on using corpus tools that consisted of three 60-minute workshops. The workshops combined elements of deductive (teacher-guided practice) and inductive teaching methodologies. The more detailed schedule was as follows:

- Day 1 entailed a basic introduction to the corpus (COCA and wordandphrase.info) interface, the concept of lexical register (academic vs. spoken), and a number of straightforward corpus searches for appropriate collocation and preposition options for common academic words.
- Day 2 training modeled collocations, grammar and prepositions corpus searches in greater depth.
- Day 3 included inductive activities with participants exploring specific queries related to previous topics. Participants watched a demonstration and had the opportunity to practice many other query types (e.g., queries targeting local and global lexicogrammatical issues), focusing on frequency using the LIST function of the COCA interface.

During the class meeting following the completion of the treatment, the participants spent the first 15 minutes reviewing the basics of corpus queries and then completed a post-test. Pre-tests and post-tests were coded based on correctness of choices supplied in the blanks.

To ensure that the training was effective for familiarizing participants ($n = 175$) with the corpus tools, a paired t-test demonstrated a significant, $t(174) = 50.87$, $p < .001$, $d = 3.9$, difference between the pre- ($M = 17$, $SD = 5.8$) and post-test ($M = 22$, $SD = 4.5$) scores, which represented a 30% improvement in average scores. The t-tests and the post-test performance suggest that participants could indeed carry out independent corpus searches based on the skills they acquired during the training.

Part 2: Online survey. Immediately following the post-test, participants were offered the option of participating in a short voluntary online survey on their perceptions regarding their experience with corpus training and tools. Less than a half of all participants, 76 out of the 175 (43%) who took part in the training chose to contribute their responses to the survey. The survey included six questions (see Appendix B), which focused on the participants' experience with and impression of the corpus tools.

Part 3: Using corpus tools for editing. In the third phase of data collection, which took place approximately four weeks after the completion of post-tests, participants had the option of using corpus tools for correcting a minimum of five previously marked errors in the final drafts of their research papers. These errors had been earlier identified by the instructor as part of her written corrective feedback on first drafts of students' research papers. The students' errors were originally pointed out through a color-coding system that included four general categories: vocabulary/lexis, grammar, punctuation/spelling, and register. In the coding scheme, lexical errors involved vocabulary replacement, while grammatical errors reflected non-targetlike grammatical usage. The number of categories (four) was fairly limited due to students' own level of metalinguistic awareness. In other words, these are the categories that learners are likely to reliably perceive in their writing. The authors felt the errors, identified in this manner, were sufficiently distinct from each other and would be salient to learners.

Extra credit points towards the final paper grade were offered to students in exchange for submitting a form documenting their corpus-assisted corrections and corresponding explanations. It was assumed that the extra credit option would help provide an incentive for learners to explore the use of corpus tools. Participants could choose on their own which errors to correct with the help of the corpus; moreover, they could utilize or rely on any elements of the corpus training to complete the corrections. Students who chose to participate in this phase of data collection completed special forms (following a standardized template), in which they provided the original and the corrected contexts of their errors, the words that were changed using the corpus tools, the procedure they followed for each correction, and a brief explanation of their reasoning behind the completed correction choices. Among the total participants, 63 (slightly over a third) chose to complete this part of the study and apply corpus strategies to error correction in their final papers. One limitation of this study design is that individual differences and per subject analysis of errors were not accounted for.

Participant error correction submissions (a total of 304 tokens after discarding missing and incomplete data) were coded based on the type of correction and on whether or not the correction was target-like or appropriate for the context. Following the analysis of collected data, the correction types were divided into three categories: (1) local lexicogrammar, comprising the categories of word choice, collocation, word inflection, preposition collocation, and part of speech; (2) global lexicogrammar, which includes phrasal/clausal level chunks, connective choices, and word order; and (3) register.

The coding options for the target-like/non-target-like corrections were 2 (two) for target-like/appropriate, 1 (one) for partially appropriate/not ideal, and 0 (zero) for inappropriate or unnecessary corrections. To ensure intercoder reliability, 15 percent of the data was coded by another coder; the intercoder reliability coefficient was 91%. The results of data analysis are reported in the following section.

3. Results

This section will begin with the presentation of quantitative results of this study, transitioning into the discussion of qualitative findings from the survey and from error analysis. In response to the quantitative research question concerning the effectiveness of corpus tools for boosting short-term noticing of specific linguistic challenges, errors were coded for correction type and the degree of appropriateness for the local context, as described in the Methods section. On average, each participant submitted evidence of correction for approximately five errors (some participants had one or two, while others had up to seven correction examples that could be used for the analysis). Descriptive statistics pertaining to participants' corrections are provided in Table 1.

Table 1
Frequency and categories of error types.

Error Category	Number of Tokens (Percentage)
Local lexicogrammar	166 (55%)
Global lexicogrammar	41 (13%)
Register	97 (32%)
Total	304 (100%)

Table 2
Types of local and global lexicogrammatical errors: Descriptive statistics.

Local lexicogrammar		Global lexicogrammar	
Part of speech choice	9 (5%)	Word order	2 (5%)
Preposition collocation	46 (28%)	Phrasal/clausal level chunks	36 (88%)
Word choice (meaning-related)	67 (40%)	Linking adverbials/connectors	3 (7%)
Word collocation	3 (2%)		
Word inflection	42 (25%)		
Total:	166 (100%)	Total:	41 (100%)

As can be seen from Table 1, slightly over half of participants' corrections involved local lexicogrammatical issues. Register was the second highest correction category with 32% of all errors, while global lexicogrammatical issues accounted for 13%. Furthermore, local and global lexicogrammatical errors were divided into subtypes, as indicated in Table 2 below.

In the local lexicogrammar category, meaning-related word choice corrections were the leading type at 40%. To clarify, the *word choice* category involved adding replacements for words that were not originally inappropriate in the context (i.e., students replaced the original expression because they wanted to convey their intended meaning more precisely). They were followed by preposition collocations at 28% and word inflection corrections at 25%. Part of speech choices and non-preposition word collocations were the least "popular" correction categories at 5% and 2% respectively. In the global lexicogrammar category, the category of clausal and phrasal-level chunks occupied the majority of corrections at 88%, with changes in word order and linking adverbials being fairly negligible at 5% and 7% respectively. Comparing the local and the global lexicogrammar categories, it is clear that local lexicogrammatical error corrections seem to be a more intuitive corpus correction choice for learners, with their total numbers being four times that of global lexicogrammar.

To allow for statistical analysis and meet requirements for independent observations, a pared dataset was created by taking the first correction from each participant. A chi-square test of independence was performed to examine the relationship between error types and participants' successful correction of these errors in this dataset. The assumptions for using chi-squared formulas were met. Namely (1) there were an independent variable (i.e., error type), and a dependent variable (i.e., +/− nativelike), and both variables were nominal; (2) the data consisted of frequency counts; (3) each of the data points were independent and were counted in only one cell; and (4) all expected frequencies were greater than five. The test indicated that the relationship between error type and targetlike corrections was significant, χ^2 (2, $N = 304$) = 23.79, $p < .001$), $V = 0.28$, with effect size approaching the medium range. The results are shown in the contingency table (Table 3). As can be seen, the relationship can be primarily attributed to participants' under performance on local errors and over performance on register errors.

To explore possible explanations for these findings, it may be helpful to examine the corpus-assisted corrections in greater detail through a qualitative lens.

3.1. How exactly did participants explore vocabulary through corpus?

This section will address how participants approached correcting prevalent error types and the problems that arose in the process.

Register. To correct register errors, the majority of the participants reported (in the completed forms tracking their changes) first finding register information (i.e., frequency for one of the five registers: spoken, academic, fiction, magazine, newspaper) in the COCA and depending on the results, proceeding with a correction. For situations where this was helpful or necessary, they also reported doing a search for the base/root vocabulary item on wordandphrase.info, accompanied by checking the highest percentage for academic usage, and picking a more appropriate synonym from the list of synonyms there. As was clear from the quantitative findings, most register errors were indeed given an appropriate target correction. A number of representative examples are shown in Table 4. Inappropriate register corrections occurred rarely.

Word choice: Appropriate substitutions. Vocabulary errors were often marked as such because the word was either inappropriate or redundant in the immediate context. In cases of redundancy in the original context, participants relied on the synonym features in the wordandphrase.info results interface to find acceptable substitutions (during the corpus training

Table 3
Chi-square contingency table for success of corrections for each error type.

	Targetlike (Expected Count) [Chi-Square for Cell]	Nontargetlike (Expected Count) [Chi-Square for Cell]
Local	97 (115.8) [3.0]	69 (50.24) [7.0]
Global	31 (28.6) [0.2]	10 (12.4) [0.5]
Register	84 (67.7) [4.0]	13 (29.4) [9.1]

Table 4
Examples of register corrections.

Original version	Corrected version	Was the change felicitous? (Y/N)
widely	extensively	Y (more academic register)
there are usually	it usually includes	Y (more precise/academic verb choice)
talked about	covered	Y (more academic register)
tough to overturn	difficult to overturn	Y (more academic register)
spent	allocated	Y (more academic register)
seems not so necessary	seems not necessary	Y (removal of a boosting particle 'so')
it is not necessary	it is not critical	Y (more academic register)

participants reported feeling "empowered" by this feature). To illustrate cases of successful lexical substitution in participant data, *interesting* became *noteworthy, stating* became *declaring, facilitate* became *expedite* (while the meaning slightly changed, the new meaning still made sense in local context), and *usage* was replaced with *application.* Before and after quotes below illustrate successful substitutions.

> Original: Before conducting analysis, the dataset was *improved*. (The feedback suggested that the original choice of verb was too generic.)
>
> Correction: Before conducting analysis, the dataset was *refined*. (*Refined* worked well as a more focused verb in this context.)

Word choice: Inappropriate substitutions. Unfortunately, some substitutions were either awkward or inappropriate for the context. To illustrate, a participant tried rephrasing *a lot of earlier works* as *many foregoing explorations*, which sounds unnatural and artificial in an academic paper (in fact, *foregoing explorations* does not appear even once in COCA).

In another example, the original feedback question asked whether in the phrase *a secret guideline* (referring to unspoken rules of academic writing and discourse organization), it was necessary to think of these guidelines as "secrets," that is, something that was intentionally guarded as hidden information. The participant did not fully understand the question in the feedback regarding the word *secret* and, concluding that a synonym would likely solve the problem, replaced *secret* with *clandestine.* While the new word does mean 'hidden', a *clandestine* guideline is not a natural means of expressing this notion. As in the previous example, *clandestine guideline* also does not appear as a collocation in COCA. These infelicities suggest that students do not always take advantage of corpora as a means of confirming whether synonym substitutions result in a collocation that appears within a corpus.

In another instance, a participant, referring to a cumulative set of concepts/activities in the context of graduate school, revised a sentence by replacing *stuff* with a supposedly more formal synonym *paraphernalia*.

> Original: *it is a student guide and discuss* ***stuffs***
>
> Correction: *it is a student guide and discuss* ***paraphernalia*** *…*

The participants' explanation accompanying the corrected sentence was that in academic writing, "paraphernalia" was "a more formal word that means the equipment consisting of miscellaneous articles needed for a particular operation or sport." The problem is that sports were not discussed in this context; in fact, the focus was on general topics one must consider when applying to graduate school, rather than on tangible articles needed for a sport activity. One can infer that the participant assumed that the word belonged to an academic register based on a simple heuristic (i.e., academic words tend to be longer). Unfortunately, the participant failed to consider the detailed semantic profile of the word, which could have been more fully explored through a careful analysis of the word's collocation patterns in the corpus.

Word inflection and part of speech errors. These errors often concerned issues like subject-verb agreement, verb forms, appropriate adjective and adverb forms, and appropriate part of speech for the context. Examples of successful corrections of such errors are provided in Table 5.

Participants often reported doing only basic frequency searches for a given phrase and possible alternatives, replacing their original erroneous forms with the more commonly occurring versions. In that sense, the missing element (i.e., the source of error) was the absence of a correct ending/suffix; accordingly, when the corpus results showed a variety of ending/suffix

Table 5
Examples of felicitous word inflection and part of speech corrections.

Original Version	Corrected Version	Why Was the Change Felicitous?
we are *expect*	we are *expected*	Added passive marking
many researches	*much research*	Changed count to noun-count marking
little statistics papers	*few* statistics papers	Changed count to non-noun marking
there *have* been a great interest	there *has* been a great interest	Change in verb form
similar *phenomenon*	similar *phenomena*	Plural noun form
we analysis	we *analyze*	Changed part of speech

options attached to the original root, participants could readily fix the problem by noting the addition and plugging it into the target context.

It seems that the relative ease of word inflection error corrections was associated with the fact that appropriate forms were immediately accessible in the corpus. In situations when this was not the case (i.e. when participants actually had to consider local sentence structure and not just plug in the corpus version), error correction was not as straightforward. For instance, substitutions sometimes involved inappropriate forms: e.g., an original phrase *present an analyze* (color-coded by instructor as 'grammar' error) was replaced with *present an investigate,* which suggests that this participant did not correctly understand the nature of the error, thinking that s/he just needed to replace the lexical choice instead of correcting the word form, and subsequently did not check for form accuracy and appropriateness of substitution in this context.

In another instance, instead of supplying the missing comparative marker (as indicated by instructor feedback) in *are generally* ***knowledgeable*** *than students*, the participant chose to ignore the grammatical error and opted for replacing the adjective instead, rephrasing the original as *are generally* ***professional*** *than students*. The participant thus used the corpus where there was no need for correction in the first place.

Finally, syntactic ambiguity of the word *urge* in the definition "*Autonomy is urge to lead own life"* propelled one learner to change the noun *urge* into a passive verb form: *Autonomy is urged to lead own life*, where there was actually no underlying need to make this change. While the new sentence was grammatically correct, it did not make sense.

Preposition collocation errors. It is likely that preposition collocations were heavily featured as popular correction choices because learners (1) had many errors with preposition selection, and (2) were taught specific patterns to address preposition problems (including those in phrasal verbs) during the corpus training. Participants used the LIST function to find multiple preposition options following/preceding a given noun or verb and then compared contexts of the corpus results with the erroneous contexts in their own writing. Common successfully corrected errors included examples such as: **in** his agenda → ***on*** *his agenda*; *section* ***in*** *a paper* → *section* ***of*** *a paper* (the meaning of partition made more sense than location in this context); *classify four sections* (missing preposition) → *classify* ***into*** *four sections*; ***in*** *average* → ***on*** *average*, and so on.

The example below shows the original and the corrected versions of an appropriately used preposition in a wider sentence context.

Original: … and used for further research ***on*** the same field.

Correction: … and used for further research ***in*** the same field.

This participant included the following explanation for the replacement: "Because it is the research 'in' the particular study/field, but not 'on' the subject." It is obvious that searching the corpus allowed this participant to notice semantic patterns governing the use of prepositions and rely on this newly acquired knowledge to select a more appropriate preposition for the target context.

Despite the fact that many errors could be readily corrected through an examination of corpus lines, the replacement preposition was, in some cases, still wrong or was not the most appropriate choice based on the context. In one example, "… has been studied ***about*** the introduction" was replaced with "… has been studied ***on*** the introduction". Here, the context clearly suggested location, and thus the correct choice should have been "in." A different participant replaced the preposition in the phrase "successful **for** obtaining tests" with the preposition *of*, rephrasing the statement as "*successful* ***of*** *obtaining tests"*; the correct preposition should have been *in*. These observations imply that in cases of inappropriate lexicogrammatical corrections, participants did not consult the full local context so as to use context information to inform the selection of an appropriate replacement. These examples highlight a particularly thorny problem with L2 learners' use of corpora. Many corpus searches rely on intuitions (which L2 learners often appear to lack) regarding the strength of association between left- and right-hand collocates of a to-be-revised word or phrase (an association expressed in terms of mutual information measures in corpus research).

Global lexicogrammatical corrections. These corrections were usually concerned with changing/replacing phrasal/clausal level chunks, changing word order, and adding or removing words, including connective phrases. A few examples of successful clausal/phrasal-level replacements are provided in Table 6 below.

As was the case with other error categories, a number of corrections were unsuccessful because they did not consider how the new version fit the meaning of local context. Consider the sentence examples below.

Original: *Different study fields have* ***different*** *frequency of appearance of individual moves*

Correction: *In different study fields the frequency of appearance of individual moves is* ***altered****.*

Table 6
Examples of felicitous phrasal/clausal corrections.

Original version	Corrected version
previously, few researches focus	researchers were interested in
they are faced with different groups	they serve different groups
were found by the evidence in data collection	features were found by the evidence obtained from the data collection

This participant meant to replace the word and the corresponding sentence structure around the word 'different' in the meaning of *representing a variety of something*; however, the new context conveyed the original meaning as *changed in comparison to the original appearance/characteristics*. As can be seen from the local context, the participant misunderstood that the meaning of *change* should in fact not be included here.

In the following example, the original error was color-coded as a grammatical error, with an added comment that necessary elements of sentence structure were potentially missing. The participant chose to rely on a corpus search to find a preposition that could be used after *evaluated*, explaining that this addition would likely make the sentence more acceptable; however, the new version was not grammatically correct either.

> Original: *10 RAs were evaluated* ***whether*** *and to what extent the proposed models for introduction part were suitable for the academic articles in computer science field.*
>
> Correction: *In this paper, 10 RAs were evaluated* ***about whether*** *and to what extent the proposed models for introduction part were suitable for the academic articles in computer science field.*

3.2. Does a clear error correction pattern emerge in participant data?

The qualitative analysis of error corrections suggests that the errors that were easiest to correct were those that appeared verbatim in the corpus and for which a more careful consideration of context was unnecessary. Whenever participants needed to rely on the analysis of local context either in the original error context or in the corresponding corpus excerpt, they often failed to proceed with these more complicated steps. Viewed negatively, the use of an online tool for error correction in some cases may give participants an expectation that the corpus can provide ready answers without the need for further judgment and discernment. In this sense, the corpus tools can, at times, have a negative effect on learners' agency and involvement in the error correction process.

3.3. Can survey findings provide an explanation?

As mentioned in the Methods section, participants were invited to provide feedback using a brief survey that was designed to gain further insights into their experience using a corpus. A total of 76 participants responded to the online survey on effectiveness of corpus tools for language learning. Key findings of the survey are provided below with the numbers indicating the percentage of participants who chose each corresponding option:

- Most (97%) participants had no prior exposure to corpus tools.
- Navigating the unfamiliar interface (43%) and choosing correct search criteria, i.e. elements of the pre-search phase (42%), were deemed to be more challenging than sorting through the output and corpus lines (35%).
- Learners thought that corpus tools were most useful for learning vocabulary (65%), usage patterns associated with phrasal verbs (77%), and collocations (69%), and less useful for grammar (50%) and register (31%).
- Among participants, 80% agreed or strongly agreed with the statement that corpus tools were more useful than dictionaries for English learning, and 93% of participants agreed or strongly agreed that corpus tools should be introduced in classroom instruction.
- Most (97%) participants planned to use COCA in the future in their own writing.

In sum, it seems that the participants clearly recognized the benefits of using corpus tools for improving their language use in academic writing. In particular, they signaled that employing corpus tools was very helpful for learning vocabulary and usage patterns, a perception that is, in part, supported by some of the error analysis findings discussed in the previous section. As demonstrated by the descriptive statistics of error types chosen for correction, local lexicogrammatical errors, especially word choice and preposition collocations, were fairly amenable to corpus-based correction. Also, contrary to participants' expectations, they often did rely on the corpus to address word inflection errors.

However, as statistical analyses showed, even though local lexicogrammatical errors were the most natural candidates for correction with corpus tools, wrong or inappropriate corrections were also more common in those cases. In other words, these results show that the error type itself was only part of the issue. As demonstrated by the analysis of participants' submissions, successful error correction was usually associated with understanding the local context correctly.

Furthermore, combining the findings from the analysis of errors and the survey, it can be concluded that successful use of corpus tools for error correction is viewed by participants as an activity requiring a significant degree of confidence as well as a certain inclination for consulting and manipulating longer stretches of text. As a rule, participants appeared to accept the results of corpora searches too readily without critical analysis. They thus became less interested in examining implications of the context of the target language on their own. They often chose what looked to be the easiest option instead of checking and comparing several versions and choosing the most contextually-appropriate choice. This suggests that complications arising during the corpus search process may deter participants from seeking answers from corpus tools for language error correction. The more general implications of these findings are provided in the discussion section.

4. Discussion

The current study examined the effectiveness of independent corpus use to repair errors based on typical instructor feedback. The data collection process and the rate of student participation during different phases of data collection imply that using corpus tools in the classroom is inherently challenging, and students are not always willing to exert extra effort to take full advantage of the tools. For the participants who chose to rely on corpus tools for error corrections, local lexicogrammatical and register errors were the predominant categories; however, despite being the most popular error type, local lexicogrammatical errors were more likely to result in an inappropriate correction. Register errors, on the other hand, were associated with fewer inappropriate corrections. More generally, these findings support previous work by Chang (2014) who concluded that general corpora are best used for finding collocations, synonyms and exact expressions and identifying distinctions in vocabulary use across genres. In an effort to make sense of these results, it is worth mentioning that learners' perceptions (as demonstrated through survey data) did not always coincide with their actual performance.

The analysis of qualitative data revealed that learners did not always understand the nature of their errors despite the fact that error types were made explicit by the instructor. While pointing out errors is common in related research, perhaps additional measures are necessary to ensure that students adequately understand the nature of their errors before engaging in corpus-assisted correction.

Along related lines, our findings also demonstrated that many of the erroneous corrections were inappropriate due to participants' failure to consider the local (original or target) context in detail, as they often failed to fully consider the contextual information in concordance lines. This failure may reflect (1) a lack of awareness of the importance of context, (2) an inability to fully utilize such information due to limitations in their linguistic proficiency, or (3) difficulty in fully exploiting such information due to time constraints. Similarly, Schenk and Cho (2012) also found that learners had problems making comparisons between linguistic features and appropriateness of given forms in local discourse, as well as noticing differences in grammar use across contexts.

Future researchers may wish to explore the source of such failures via alternative data-collection procedures. For example, an additional follow-up phase could be included in which the researcher informs participants about instances in which they failed to give adequate consideration to context. Participants' subsequent error corrections could then be examined.

As the analysis of errors demonstrated, independent use of corpus tools seems to be best suited for contexts that do not require extensive grammaticality judgment and/or advanced metalinguistic awareness (cf. the discussion of types of learners most suited to corpus use in Römer, 2011). More generally, it appears that the use of corpus tools cannot replace or compensate for a deeper and more nuanced understanding of language, which indirectly supports the conclusion of Geluso and Yamaguchi (2014), who found that corpus-assisted learning was not promoting target-like use of pragmatics, an area that required a deeper awareness of and reliance on contextual information.

Successful error correction is tightly intertwined with overall language proficiency. Given that this study's participants were placed in the upper-level EAP course, the authors did not focus on the nuances of individual differences and proficiency, but there may have been an interaction between the learners' specific proficiency levels and types of errors they could correct that was not accounted for by methods used in this study. Future research should work to clarify potential interactions of this type. Additionally, future research could consider exploring possible interactions between successful utilization of corpus tools and learner factors such as metalinguistic awareness, language learning motivation, or learner type as explored by Bridle (2019), which could be measured using standard assessments.

More generally, the findings also indicate that instructors and their feedback play a vital role for students' understanding of how corpus tools can be used. First, as students are introduced to the corpus tools for the first time, effective and extensive guidance from the instructor is of paramount importance. Second, the typical color-coding approach to indicating errors utilized in this study might not be sufficient to facilitate successful use of corpus resources. Instructors should strive to help students correctly understand what the errors are and how those should be corrected. Future research should investigate the need for more precise feedback and its interaction with or effect on successful use of corpus tools.

From a pedagogical standpoint and taking into consideration typical limitations in classroom time, this implies that increasing one's reliance on teacher-created corpus-informed materials may provide greater benefits to learners than pedagogical materials or approaches that rely solely on autonomous data-driven learning. Along similar lines, Römer (2011) suggested that "corpus researchers would do well to help create more DDL exercises and corpus-derived teaching materials in general" (p. 216). Considering other possible uses of corpora in the classroom, this technology may be particularly useful for non-native teachers of English, who, as advanced or near-native level learners themselves, are highly motivated and metalinguistically aware and thus more likely to master the corpus tools that would help them undertake autonomous linguistic investigations.

In addition to lack of data on individual differences and proficiency, the current study has several other limitations. The open-ended nature of the experimental task, while making the study more realistic from an ecological perspective, made the quantitative results more difficult to interpret. Moreover, the limited response rate to survey questions may have failed to reveal a range of attitudes and experiences by participants who were less enthusiastic about using an online corpus. For this reason, participants' reported enthusiasm for using corpus tools in the future must be interpreted with some caution.

Finally, future investigations of corpus tools could benefit from more theory-driven approaches that rely on specific linguistic theories and theories of language acquisition (e.g., Systemic Functional Linguistics, Cognitive Linguistics, Socio-Cultural Theory, discourse theories, and so on, cf. Charles, 2007). In other words, it would be worthwhile to examine the

effect of complementing technology-enhanced instruction with a systematic and theory-driven approach to language instruction.

Appendix A

Sample pre- and post-test questions (one example per question category).

1. Preposition choice: We believe that advances _____ social psychology have important implications for the measurement, as well as conceptualization, of job satisfaction.
2. Phrasal verbs: Drawing _____ the social cognition, judgment, and life satisfaction literature, we illustrate how satisfaction measures are heavily dependent on the context.
3. Verb-noun collocations: We _____ an in-depth analysis of the problem.
4. Subject verb agreement: A wide range of computer applications and programs _____ now therefore available in all these areas for individual and classroom use.
5. Verb tense/form choice: However, even though many studies have reported an increased use of technology in education, there _____ very little research reported on the effectiveness of such use.

Appendix B

Questions of Online Survey on Learners' Perceptions of Corpus Use.

1) Had you heard about corpus/corpora before taking this class?
2) If you answered *yes* to the previous question, have you used corpora before you took this class? (If you did, please indicate the name of corpus/corpora in the 'Comment' box.)
3) If this was your first time using the corpus, please indicate the types of difficulties you may have experienced (mark all that apply).
   - Difficulty navigating the unfamiliar interface
   - Difficulty choosing correct search criteria
   - Difficulty sorting through and/or dealing with too many sentences in the output
   - Difficulty creating and following search sequences
   - Other
4) What do you think the use of corpus resources is good for? Mark all that apply.
   - Vocabulary
   - Grammar
   - Usage patterns (in particular, phrasal verbs, etc.)
   - Learning collocations
   - Learning contextual meaning of synonyms
   - Academic register
   - Other (please specify)
5) Please indicate your opinion in regard to the following statements.
   - I would like to use COCA or other corpus for my English writing in the future
   - The corpus is more helpful than a dictionary for my English writing
   - I believe corpus use is more helpful for writing than for reading in English
   - I believe that instruction on corpus-searching strategies should be used in all/many English courses
   - I would recommend the use of corpora to my English-learning friends
6) Please summarize your general impression about using corpus resources for language learning purposes. Both positive and negative feedback is welcome.

References

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL, 21*(1), 37–54.
Bridle, M. (2019). Learner use of a corpus as a reference tool in error correction: Factors influencing consultation and success. *Journal of English for Academic Purposes, 37*, 52–69.
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.
Chan, T., & Liou, H.-C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning, 18*, 231–251.
Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes, 6*(4), 289–302.
Chon, Y. V., & Shin, D. (2013). A corpus-driven analysis of spoken and written academic collocations. *Multimedia-Assisted Language Learning, 16*(3), 11–38.
Cresswell, A. (2007). Getting to 'know' connectors? Evaluating data-driven learning in a writing skills course. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 267–288). Amsterdam, The Netherlands: Rodopi.

Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning, 28*(2), 130–144.
Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics, 14*(3), 393–417.
Flowerdew, L. (2010). Using corpora for writing instruction. In A. O'Keeffe, & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 444–457). London, UK: Routledge.
Flowerdew, L. (2012). *Corpora and language education*. Basingstoke, UK: Palgrave Macmillan.
Frankenberg-Garcia, A. (2012). Raising teachers' awareness of corpora. *Language Teaching, 45*(4), 475–489.
Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System, 32*(3), 301–319.
Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam/Philadelphia: John Benjamins.
Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL, 26*(2), 225–242.
Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching. In A. O'Keeffe, & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359–370). London: Routledge.
Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL, 26*(2), 163–183.
Hyland, K. (2008). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18*(1), 41–62.
Johns, T. (1991). Should you be persuaded? Two samples of data-driven learning materials. *English Language Research Journal, 4*, 1–16.
Leech, G. N. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1–23). London, UK: Longman.
Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal, 93*(1), 61–78.
Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 119–141). Amsterdam, The Netherlands: John Benjamins.
Mizumoto, A., & Chujo, K. (2016). Who is data-driven learning for? Challenging the monolithic view of its relationship with learning styles. *System, 61*, 55–64.
Mueller, C., & Jacobsen, N. (2016). A comparison of the effectiveness of EFL students' use of dictionaries and an online corpus for the enhancement of revision skills. *ReCALL, 28*(1), 3–21.
Park, K. (2012). Learner–corpus interaction: A locus of microgenesis in corpus-assisted L2 writing. *Applied Linguistics, 33*(4), 361–385.
Pérez-Paredes, P., Sánchez-Tornel, M., Alcaraz Calero, J. M., & Jiménez, P. A. (2011). Tracking learners' actual uses of corpora: Guided vs non-guided corpus consultation. *Computer Assisted Language Learning, 24*(3), 233–253.
Quinn, C. (2015). Training L2 writers to reference corpora as a self-correction tool. *ELT Journal, 69*(2), 165–177.
Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics, 31*, 205–225.
Schenk, A., & Cho, W. C. (2012). The efficacy of corpus-based pedagogical techniques for academic writing. *Multimedia-Assisted Language Learning, 15*(2), 167–186.
Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL, 26*(2), 184–201.
Tono, Y., Sataki, Y., & Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL, 26*(2), 147–162.

Natalia Dolgova is Teaching Assistant Professor of English for Academic Purposes at the George Washington University, Washington, DC, USA.

Charles Mueller is Professor of English at Fuji Women's University, Sapporo, Hokkaido, Japan