
Developing autonomy through awareness of textual features in academic texts

H.M. McGarrell

hmcgarrell@brocku.ca

BROCK UNIVERSITY

Abstract

Research on ESL writing suggests that current pedagogical practices typically train developing writers to submit their texts to their teacher for comments and “correction”, without opportunities to develop independent self-editing skills. These practices focus primarily on sentence-level grammar, vocabulary and mechanics. This study reports on an exploration into introducing university-level writers to online corpus tools to develop their ability to edit their own work. Ten native and non-native speaking participants analysed corpus materials and published journal articles relevant to their discipline to compare selected uses of grammar and vocabulary with their own texts. Insights from triangulated data sources (participant texts, participant and teacher interview data) show how participants edited their work using the corpus tools. Benefits related specifically to awareness of lexical conventions and sentence structure.

Key words: self-editing, autonomy, learner concordancing, corpus analysis

Résumé

La recherche sur la rédaction en anglais langue seconde montre que la pratique pédagogique actuelle consiste généralement à demander aux rédacteurs en formation de soumettre leurs textes au professeur pour commentaires et « correction », sans leur donner l'occasion d'apprendre à s'autoréviser de façon indépendante. Cette pratique est axée principalement sur la grammaire des phrases, le vocabulaire et la mécanique. L'étude rend compte d'un exercice d'initiation de rédacteurs de niveau universitaire à des outils de corpus en ligne pour les aider à développer leurs aptitudes à réviser leur propre travail. On a demandé à dix participants de langue maternelle anglaise et non anglaise d'analyser le contenu d'un corpus et d'articles publiés dans des revues de leur discipline afin de comparer certains usages de la grammaire et du vocabulaire à ceux de leurs propres textes. Les constatations émanant de sources de données triangulées (textes rédigés par les participants et données tirées d'entrevues

avec les participants et les professeurs) montrent comment les participants ont révisé leur travail à l'aide des outils de corpus. Les apprentissages concernaient spécifiquement les conventions lexicales et la structure des phrases.

Mots-clés : autorévision, autonomie, établissement de la concordance par les apprenants, analyse de corpus

Graduate students are typically expected to have advanced academic writing skills with which to reflect their subject matter knowledge in appropriate papers and, eventually, a thesis. Yet, both native and non-native speakers of English (NS and NNS) find that they have had insufficient opportunities to develop appropriate writing skills during their undergraduate years (Addison and McGee, 2010; Ondrusek, 2012; Sallee, Hallett and Tierney, 2011) as testing practices in increasingly large classes at the undergraduate level favour multiple choice or short answer items. Butler and Britt (2011) argue that many graduate students are ill-prepared for the demands of graduate writing, yet graduate programs rarely offer in-class attention to develop and refine students' writing skills (Sallee et al., 2011). One of the skills many graduate students lack is the ability to edit their own work. Their previous experiences led them to rely on their teachers to point out writing problems through corrective feedback, giving students neither opportunity nor motivation to develop autonomous editing skills. In process approaches to writing (Osterholm, 1986; Zamel, 1982), editing is viewed as the last task before writers submit or “publish” their work. The development of such skills in combination with corpus materials — the focus of the study presented in the following paper — is increasingly of interest, as discussed in the next section.

Background

Written corrective feedback (CFW) has been widely researched and debated by researchers and language professionals in first and subsequent (L1 and L2) writing, especially since Truscott's (1996) controversial claim that grammar correction was ineffective for developing writers. Gaskill and Cobb (2004) observed that many learner errors appear to be immune to CFW and researchers continue to explore how best to provide CFW on L2 writers' written work. Swain and Lapkin (1985) hypothesized that learners need opportunities to notice how aspects of their output differ from native speaker language. Learners' active engagement in the analysis of their own output is thus seen as a means of increasing uptake through activities that require cognitive tasks, which in turn is anticipated to result in language learning gain. One way to engage learners in such noticing activities is through data-driven learning (DDL), as proposed and exemplified in Johns (1986, 1991). DDL activities lead learners to inves-

tigate authentic examples of the target language on their own to discover how selected aspects of language function in actual language use. Similarly, Allwright and Bailey (1991, p. 107) stressed the need for learners to “be able to make self-initiated self-repairs” and Ferris (1995) recommended that writing teachers use activities that develop learners’ ability to self-correct their work to promote language development as well as learner autonomy. Despite such endorsements of self-correction, relatively few studies are available that document teaching activities designed to encourage self-correction or learners’ development in self-correcting their writing output. Although greater emphasis on self-correction in writing courses is unlikely to solve the controversy around CFW, it offers opportunities for language development as well as learner autonomy (Lai, Shum and Tian, 2014).

DDL is often linked to corpus materials to develop reading, grammar, speaking and writing skills. The large data banks available through corpora are useful as they offer naturally occurring texts that can be queried for specific language occurrences including words, phrases and structures in oral or written use. Boulton (2010) observed that the majority of studies focus on advocating the benefits of corpus inquiries for learners, with relatively few studies offering actual data to support claims made. However, some recent qualitative and/or quantitative studies have explored learner attitudes toward DDL and corpus investigation (Gaskell and Cobb, 2004; Lee and Swales, 2006; Razeq, 2014; Yoon and Hirvela, 2004), vocabulary learning (Liu, 2013), the development of learner training modules (Feng, 2014; Lee and Swales, 2006; Liu, 2013; Quinn, 2015) and, increasingly, the use of corpus material to improve aspects of L2 writing. The latter category includes Gilmore (2008), who asked 45 Japanese undergraduate students to revise their first drafts, on which the instructor had underlined non-standard language, through corpus work. After a 30-minute introduction to corpus work and a 60-minute practice session, the students revised their texts and raters judged 61% of the revisions as resulting in more appropriate language compared to the drafts. Participant feedback on the experience with corpus material was generally positive but it is noted that some participants had not submitted revised texts. Yoon (2008) reports that in a qualitative study of six ESL graduate students in the USA, the participants became more independent and confident writers. The six case studies reflect considerable individual variation in how students used corpus materials in their writing assignments, but all six participants used the materials to improve grammar and word usage questions, with searches on verbs being the most frequent. Feng (2014) describes a corpus-based self-editing training program implemented with three Chinese intermediate level graduate students who had no previous experience with corpus work. The study offers detailed descriptions of the training the participants received and the outcomes of the

training. Analysis of revised versions of three consecutive essays from each participant shows a decrease in errors, but the students commented on feeling overwhelmed with data and the time commitment required to carry out the self-editing. Finally, Yoon and Jo's (2014) small study compares direct and indirect use of corpus materials and shows that participants who had direct access to corpora frequently corrected errors in their writing.

The above studies indicate that developing writers benefit from the use of corpus materials for self-correction but that students find the technique challenging and time-consuming. The participants were at intermediate or lower levels and from disciplines that did not introduce them to corpus linguistics as part of their studies; their texts were typically on different topics and genres. None of the studies reviewed include data that reflect developing NS writers' editing abilities. The purpose of the present study was to explore how high intermediate to advanced level graduate students in Applied Linguistics experience self-editing through the use of corpus materials. More specifically, the study explores how NS and NNS graduate students with little or no experience with self-editing their academic writing draw on corpus tools to self-edit their work, a text with consistent topic and genre for all participants. The tools used were Lextutor (Cobb, 2015; for a review see Sevier, 2004) and the Corpus of Contemporary American English (COCA) (Davies, 2008), both available wherever an Internet connection is available.

Methodology

The study took place at a medium-sized North American university with participants from a graduate program in applied linguistics. As part of their course work, participants had been exposed to Lextutor and COCA through three journal articles that illustrated the use of such tools for investigations related to second language acquisition. A hands-on one-hour workshop allowed them to learn to construct basic searches and queries in both Lextutor and COCA. The course instructor encouraged students to use Lextutor and COCA independently, but no additional course work was devoted to the tools. The students were, however, encouraged to focus on careful analysis of vocabulary and constructions in the journal articles on which their assignment was based. The assignment consisted of a 700- to 800-word argumentative paper; thus, topic and genre were consistent for all the students. They were asked to submit their work electronically by a given date, take an additional week to edit their work with corpus tools and then resubmit. The participants produced these texts as homework and thus had access to any resources they wished to consult. Students who wished to release their work for the study were asked to submit both their final draft text and the subsequent edited version to the researcher.

The ten participants in the present study included four NS of English (NS),

three Arabic and three Mandarin speakers (NNS). The four male and six female participants were between 23 and 28 years old, with a mean age of 26.2 years. They had indicated that their experience with academic writing was between 0 and 3 years: the Mandarin speakers had had no academic writing opportunities prior to their arrival in Canada, but a required two-month intensive English for Academic Purposes course preceded their start to graduate studies. Three of the four NS had no formal instruction in academic writing and their undergraduate degree program required only informal reflective essays, short answer questions and multiple-choice questions. One of the four NS participants had completed one course in which an academic paper was expected, but had received no writing instruction or guidance. The NS and NNS participants thus had similar academic writing backgrounds.

The texts submitted by the ten participants are of similar length and range from 710 to 814 words, averaging 784 words. The participants' responses to a questionnaire and notes from short oral followup meetings provide additional data. The questionnaire was completed during a meeting about the study, after the participants had submitted their work and given consent to participate in the study. In addition to biographical data, the questions elicited details about participants' language background, language learning and, especially, academic writing experience and the process they followed to produce the paper they had submitted to their instructor. The short followup interviews of about ten minutes served to elicit information about specific changes participants made to their drafts as well as to probe their motivation for making these changes rather than potential others. These qualitative data serve to provide insight into the strategies or thought processes that led writers to make the changes they made.

Data analysis consisted of comparing the two text versions from each participant, then categorizing each change made into a change in mechanics, vocabulary, construction or other. Mechanics included changes to spelling, punctuation, replacement of a contraction with its non-contracted equivalent; vocabulary changes included all changes to words or phrases; construction changes related to structural changes at the sentence level such as changes in tense. Notes from the oral interviews were organized into themes to provide insight into individual participants' editing experience as well as similarities and differences among the experiences of the different participants.

Findings

The first section below presents the number and types of changes participants made to their texts, followed in the second section by selected comments from oral interviews that illustrate some of the participants' thinking around a particular change. The last section offers some observational comments from the

instructor about issues that arose during the editing activity.

Participants' edits in numbers

An analysis of the changes in each participant's text shows that changes were primarily at the level of vocabulary, mechanics and sentence construction but also included a small number of changes to organization and content. The latter relate to the revision process more than editing and are not considered further. Both NS and NNS participants edited their work, but NS edits were fewer in number and related more to academic language while the NNS's edits focused on vocabulary and non-standard constructions. Participants made between 3 and 10 changes (average 6.6) to their texts, as shown in Table 1. The largest number of edits involved vocabulary, followed by mechanics, then constructions.

TABLE 1
NS and NNS writers' self-edits

Participant	Mechanics	Vocabulary	Construction	Organization	Content	Totals
NS						
SENSEM14	2	2	—	—	—	4
SENSEF23	1	3	—	—	2	6
SENSEF26	—	4	—	1	—	5
SENSEF29	1	—	—	1	1	3
NNS						
SEMAM31	4	4	1	—	—	9
SEMAM35	5	4	1	—	—	10
SEMAM38	3	5	1	—	—	9
SEARF40	2	4	2	—	—	8
SEARF42	2	3	—	—	—	5
SEARF49	1	2	1	—	3	7
Totals:	21	31	6	2	6	66

Changes in vocabulary included replacement of informal or general expressions with more formal or specific ones. For example, several participants replaced "do" with "study", "determine" or "carry out", or "look at" with "examine", as appropriate to the context. Changes to mechanics addressed primarily misspellings but also missing punctuation marks and elimination of contractions. The changes to constructions varied among participants. One par-

participant reformulated a long sentence into two shorter ones, another changed a direct question into an indirect one and a third addressed a misplaced modifier. Three changes involved the verb phrase. The majority of the changes made resulted in standard English sentences, but numerous other language problems were not addressed during the editing phase.

According to the interviews, intuition or an observed item in the assigned readings motivated participants' selection of items to edit. The main strategy they used in their editing consisted of taking a phrase or single word and checking it against academic databases in Lextutor or COCA to determine whether and how the expression was used. When unable to locate an expression, participants tried to identify the key word and its collocates to find a suitable alternative. When a phrase did not reflect standard academic English, the participant was unable to find the phrase, leaving him or her without resolution to the question. Another type of lexical change occurred when participants replaced a verb more typical of oral English (e.g. "find out", "look at") with one more typical of academic discourse (e.g. "investigate", "determine"). Changes at the level of mechanics involved minor local changes such as changing a contraction to its long form, correcting a misspelled word or changing punctuation. Not all changes were successful, either because the changes resulted in a non-standard construction or in an atypical expression.

Insights from participant interviews

Questionnaire and interview data suggest that the participants found editing their work challenging in terms of knowing how to edit and the time required to work with corpus materials. Several participants used intuition to guide their editing. For example, a NS of Arabic felt that one of the sentences in her text did not "sound right" and determined to revise it. The relevant section from her text was "Peer feedback has fruitful merits for the learners" (SEARF42). The edited version of her text was "Peer feedback has many benefits for the learners." In the follow-up interview, the participant indicated that she was unsure about the word 'fruitful' and explained that

I searched COCA ... found that 'fruitful' [is] usually used in scientific papers ... and scholars that mainly talk about multiplying. I searched also to see which nouns collocate (examples) but could not find 'fruitful merits' and I don't mean 'fruitful approach' or something like that. (SEARF42IN)

Still unsure about how to use the expression, she opted for what she considered a "safe" version, one she had verified and located in COCA but that did not include the word "fruitful" she had intended to use. Similarly, participant SEMAF31 felt that her use of "reflective" in her sentence "... their marks were reflective of ..." was inappropriate. Unable to find the exact phrase in Lextutor,

she searched COCA for “reflect” with a wildcard ending and found confirmation for “. . . their marks reflect this,” the version she chose to use in her edited version. In the followup interview, she explained:

I didn't like the sentence but couldn't figure out what was wrong with it. I looked for 'reflective' in a concordance; reflective is used differently but I don't know exactly how, so I changed it to a different verb form. I looked for a similar expression in the readings but couldn't find it . . . maybe it's not used in English.
(SEARF49)

One of the Mandarin-speaking participants was not quite successful in resolving his problem with the sentence “In order to collect data, the following procedures were made.” He explained that he had a sense that the verb was not appropriate and searched strings as well as collocations. In the end, he settled for “In order to collect data, the following procedures were operated,” explaining that he found “operating procedures” but not “procedures were operating.” He added that “. . . the corpus can't tell you what's wrong, it just shows what's used, so it's not always useful.”

NS participants also had reason to consult corpus tools. One of the NS participants was unsure whether an infinitival or participial construction was required in his sentence:

Despite the arguments against and for the peer feedback role in class, it is viewed to be one of the most beneficial and effective areas, providing students with opportunities to interact, communicate and acquire the language.
(SENSEM14)

He settled for “. . . to provide students with opportunities to interact, communicate and . . .” During the followup interview, when questioned about his use of “against and for,” he was surprised by his phrasing and suggested that he was too preoccupied with the infinitival question that he had missed this problem, adding “sometimes, I miss the obvious.” Another NS participant expressed eloquently a discovery she had made about the use of informal verbs in her academic writing in comparison to published work, an issue several other participants alluded to more obliquely during interviews:

I wish I'd known about how to check whether a word is used in informal spoken or formal written language in my undergrad. One of my profs kept writing 'informal' in the margin of my papers and I didn't know what he meant. I mean, it's how I talk, right? So, what's the problem . . . (laugh). Now I know. If I want to use a short, two-word verb, I look for a longer one — if necessary in the online Thesaurus — and check it in COCA.
(SENSEF26)

However, one of the NNS participants was not quite so successful with a similar issue she was trying to resolve: “I want to know how to say ‘find out’ in

academic writing but how can I find that in a corpus?" (SEARF42). While some of the participants drew on the language used in their assigned reading, others had not yet developed the technique of learning from models that published authors' provided.

One of the NS participants explored the Lextutor website and discovered the VocabProfile section, a part of the site that allows users to submit a text to determine vocabulary complexity, i.e. how much of the vocabulary in the text is part of the first 1,000 or second 1,000 words, the Academic Word List or off-list. She reports that

When I first checked my draft in VocabProfile, 80% of the words I used were from the K1 or K2 list. I couldn't believe it, I'm a native speaker, I'm supposed to have a good vocabulary. It was an eye opener. Now I pay careful attention to vocabulary when I read to express my ideas better, I mean, more like somebody who knows the subject matter. I'm using more verbs like 'suggest', 'argue', 'claim' rather than just 'say'.
(SENSEF23)

Her discovery shows how beneficial the kind of analysis required for self-editing can be and how "noticing" leads to learning and development, as suggested in Swain and Lapkin (1985), for both L1 and L2 writers.

Although the participants experienced successes with their self-editing, not all attempts to improve the texts were successful, as shown above. Some of the participants pointed out that self-editing with corpus linguistics tools can be frustrating and all of them commented on it being time-consuming. They identified a number of hurdles that might prevent them from using these tools regularly. For example, SENSEF23 pointed out that "COCA is difficult to use without teacher guidance. It has too much information and not everything is relevant ... It's very useful but doesn't have answers for all my questions." SEARM40 independently indicated that "Corpus is good but it takes too much time ... lots of reading ... and I don't know how to search, what to write to get the good answers."

Insights from the instructor

The editing activity took place outside of class but the participants were able to consult with the instructor during office hours. Several participants requested that the instructor indicate what should be edited (which the instructor declined to do), others sought advice on the construction of queries as well as on how to identify items for editing. The instructor suggests that the time investment required rather than lack of insights prevented participants from making additional edits to their work. Additional training, even for graduate students exposed to corpus materials as part of their studies, would seem appropriate. The instructor also suggested that more emphasis should be placed on discus-

sions of academic writing to draw developing writers' attention to disciplinary conventions (e.g. jargon or specific expressions, transitions).

Discussion and conclusion

As illustrated in the examples above, data indicate that participants were able to identify several areas in their texts that required editing and that they were able to draw on corpus material to resolve their questions. The findings show that both NS and NNS writers benefited from corpus materials but that NS focused on lexical expressions more than constructions. Despite the participants' more advanced language abilities compared to previous studies, they would likely benefit from additional training. And, similar to previous studies, the participants felt that editing would be more efficient if teachers were to indicate where editing should take place. One of the participants, SENSEF26, suggested that self-editing should be introduced much earlier to allow developing writers to get used to the idea, pointing out that "... once you are used to the teacher doing it, it's too late to be motivated to do it yourself." Similarly, SEARF49 maintained that "... it is easier for the instructor to do this task ... this is the job my instructor is expected to do." Similar comments were reported in earlier studies and seem to reflect the fact that developing writers are socialized into relying on their teachers for CFW, with little opportunity or need to develop self-editing skills. While the participants appreciated the wealth of information available through corpus materials, they were mindful of the time required to edit their work, thus echoing participant feedback from similar studies (Gilmore, 2008; Quinn, 2015). Participant comments confirm that self-editing needs to be introduced early in writing classes to help writers develop the notion that they are able to reflect on their own work and improve it with the help of tools appropriate for their level of ability.

In summary, the findings of this exploratory study with highly motivated participants, i.e. graduate student writers, confirms that these writers need to develop their self-editing skills to become autonomous writers. Findings suggest that corpus linguistics tools are useful for these writers but that they need support to help them identify suitable searches. Past experience has taught them that if they wait for the teacher's comments they can simply correct their mistakes, a strategy that seems to turn teachers into proofreaders and prevents developing writers from paying attention to differences between their own and expert writers' texts, i.e. noticing (Swain and Lapkin, 1985). Whereas the former appears to develop a culture of dependency, the latter is likely to help these writers develop their language and writing abilities. A small exploratory study such as the one described in the preceding paper cannot be used to generalize. It does, however, show that for the ten participants in this study, self-editing with corpus tools is productive. Introduction to self-editing

techniques during regular writing courses at earlier stages of writing development would no doubt be advantageous as it would limit teacher dependency and instead cultivate a sense of autonomy. All the participants pointed out that self-editing draws on techniques that can be practised, but these techniques require time to develop. One broad implication of this study is, therefore, that Ferris' (1995) recommendation that teachers introduce their students to level-appropriate self-editing techniques early on has merit: the practice would likely help instil in developing writers the habit as well as the confidence needed to become autonomous writers.

References

- Addison, J. and S. McGee. 2010. Writing in high school/Writing in college: Research trends and future directions. *College Composition and Communication*, 62, pp. 147–179.
- Allwright, D. and K. Baily. 1991. *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge: Cambridge University Press.
- Boulton, A. 2010. Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60, pp. 534–572.
- Butler, J.A. and M. Britt. 2011. Investigating instruction for improving revision of argumentative essays. *Written Communication*, 28, pp. 70–96.
- Cobb, T. 2015. The Compleat Lexical Tutor, v. 8. Available at: www.lextutor.ca.
- Davies, M. 2008–. *The Corpus of Contemporary American English: 450 million words, 1990–present*. Available at: corpus.byu.edu/coca/.
- Feng, H.-H. 2014. A pilot study: The use of corpus concordancing for ESL learner self error-correction. *Journal of Interactive Learning Research*, 25, pp. 5–25.
- Ferris, D. 1995. Teaching students to self-edit. *TESOL Journal*, 4, pp. 18–22.
- Gaskell, D. and T. Cobb. 2004. Can learners use concordance feedback for writing errors? *System*, 32, pp. 301–319.
- Gilmore, A. 2008. Using online corpora to develop students' writing skills. *ELT Journal*, 63, pp. 363–372.
- Johns, T. 1986. Micro-concord: A language learner's research tool. *System*, 14, pp. 151–162.
- Johns, T. 1991. From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *English Language Research Journal*, 4, pp. 27–45. Thematic issue: *Classroom concordancing*, T. Johns and P. King (eds.).
- Lai, C., M. Shum and Y. Tian. 2014. Enhancing learners' self-directed use of technology for language learning: The effectiveness of an online training platform. *Computer Assisted Language Learning*, Advance Online Publishing. Available at: www.tandfonline.com/doi/abs/10.1080/09588221.2014.889714.
- Lee, D. and J. Swales. 2006. A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25, pp. 56–75.

- Liu, D. 2013. Using corpora to help teach difficult-to-distinguish English words. *English Teaching*, 68, 3, pp. 27–50.
- Ondrusek, A. 2012. What the research reveals about graduate student' writing skills: A literature review. *Journal of Education for Library and Information Science*, 53, pp. 176–188.
- Osterholm, K. 1986. Writing in the native language. In B. Wing (ed.), *Listening, reading and writing analysis and application*. Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages, pp. 17–43.
- Quinn, C. 2015. Training L2 writers to reference corpora as a self-correction tool. *ELT Journal*, 69, pp. 165–177.
- Razeq, A. 2014. University EFL learners' perceptions of their autonomous learning responsibilities and abilities. *RELC Journal*, 45, pp. 321–336.
- Sallee, M., R. Hallett and W. Tierney. 2011. Teaching writing in graduate school. *College Teaching*, 59, pp. 66–72.
- Sevier, M. 2004. The Compleat Lexical Tutor, v. 4. *TESL EJ*, 8,3, pp. 1–10. Available at: tesl-ej.org/ej31/m2.html.
- Swain, M. and S. Lapkin. 1985. Problems in output and cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, pp. 371–391.
- Truscott, J. 1996. The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327–369.
- Yoon, H. 2008. More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning and Technology*, 12,2, pp. 31–48. Thematic issue: *Technology and Learning to Write*, ed. J. Bloch. Available at: llt.msu.edu/vol12num2/yon.pdf.
- Yoon, H. and A. Hirvela. 2004. ESL student attitudes towards corpus use in L2 writing. *Journal of Second Language Writing*, 13, pp. 257–283.
- Yoon, Y. and J. Jo. 2014. Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning and Technology*, 18, pp. 96–117.
- Zamel, V. 1982. Writing: The process of discovering meaning, *TESOL Quarterly*, 16, pp. 195–209.