

Using Corpus-based Instruction to Explore Writing Variation Across the Disciplines: A Case History in a Graduate-level Technical Editing Course

Ryan K. Boettger
University of North Texas

Abstract: Understanding the linguistic and rhetorical patterns of an academic discipline strengthens students' abilities to write in professional settings. Data-driven learning and corpus-linguistic methods can increase this understanding and should be considered valuable contributors to any writing curriculum. In this paper, I present a case history on integrating corpora in a graduate-level technical editing course to teach students about writing variation. Students applied this knowledge to edit research-based texts for non-native English speakers within the STEM disciplines. Though this case history focuses on corpora in a technical editing course, the approaches I describe transfer to any course with a writing component as well as across grade levels and student proficiencies. I conclude by addressing the barriers associated with integrating corpus-based learning into the classroom.

Linguistic and rhetorical patterns reflect an academic discipline's subject area and its methods for building knowledge (Conrad, 1996, 2001; Hyland, 2004; Stoller, Jones, Costanza-Robinson, & Robinson, 2005). Understanding these patterns strengthens students' abilities to write across academic disciplines and in professional settings. In this paper, I present a case history of how graduate technical editing students identified language patterns with corpus-linguistic methods.

Corpus linguistics is an applied linguistics approach that uses computer-assisted techniques to explore authentic language data and facilitates large-scale analyses of writing patterns, expanding students' depth and breadth of genres and language when compared to traditional instruction. Quantitative findings reveal language patterns that writers use, showing the most typical language choice for certain functions in certain contexts. Subsequently, qualitative analysis provides interpretation of reasons for typical and unusual choices (Conrad, 1996).

The purpose of integrating these instructional approaches was to raise developing editors' genre awareness as well as introduce them to writing conventions used across the disciplines. Students applied this knowledge to edit research-based texts for clients in the STEM fields who spoke English as a second language. Though this case history is focused on using corpora in a technical editing course, the approaches I described transfer to any course with a writing component as well as across

Across the Disciplines

A Journal of Language, Learning and Academic Writing

DOI: <https://doi.org/10.37514/ATD-J.2016.13.1.03>

wac.colostate.edu/atd

ISSN 554-8244

Across the Disciplines is an open-access, peer-review scholarly journal published on the WAC Clearinghouse and supported by [Colorado State University](#) and [Georgia Southern University](#). Articles are published under a [Creative Commons BY-NC-ND license](#) (Attribution-NonCommercial-NoDerivs) ISSN 1554-8244. Copyright © 1997-2017 The WAC Clearinghouse and/or the site's authors, developers, and contributors. Some material is used with permission.

grade levels and student proficiencies. I conclude by addressing the barriers that educators associate with integrating corpora into their classrooms.

Literature Review

In this section, I provide an overview of how corpus-based approaches align with the data-driven learning pedagogical movement. I then discuss how these approaches help raise students' genre awareness and develop their data information literacies. These discussions are connected to teaching writing across the disciplines and preparing future technical editors.

Situating Corpus-Based Learning

Corpora and subsequent teaching approaches contribute to data-driven learning (or DDL). DDL contrasts with traditional deductive, lecture-based methods and promotes students' active engagement with the subject through technology. Instruction is typically designed to (i) foster students' active engagement, (ii) encourage inductive activities that allow students to explore a topic on their own terms, (iii) promote interaction between students, and (iv) provide students with output-focused activities to apply this new knowledge (Hanson & Wolfskill, 2000; Willis & Willis, 2007).

Corpus-linguistic methods reveal language patterns that contextualize writing nuances across disciplines rather than promote general, and potentially prescriptive, writing principles. Hyland's (2004) corpus analyses of published research articles in eight academic disciplines revealed tremendous insight into writing variation; for example, engineers prefer to *report* information while philosophers *argue* and biologists *describe*. Corpora integration and learning also outperforms traditional teaching methods in all areas of language competency, including vocabulary (Chujo, Anthony, & Oghigian, 2009; Cobb, 1997, 1999; Cresswell, 2007), idiomatic expression (Chan & Liou, 2005; Liu, 2010), and grammar (Garner, 2011, unpublished).

Corpus-based learning complements the aims of the technical editing course. Developing editors must learn to analyze a text and then apply their problem-solving abilities to make it more usable and comprehensible, which could include editing for content, organization, style, and visual design and illustrations (Rude & Eaton, 2011). Similarly, editors must learn interpersonal skills, such as listening to their clients' needs, explaining the editing process, and ensuring the rhetorical goals of the document are fulfilled (Rude, 2010). To refine these skills, technical editors must understand writing variation and genre construction across disciplines.

Raising Genre Awareness

Corpora function as databases of authentic language data. Educators have access to a variety of free corpora, which I will address in the final section. Corpora house a variety of genres (or text types), including academic and professional texts as well as student-written and expert (or published) texts. These text types can assist in raising students' genre awareness, which scholars like Johns (2002) maintain is essential in instilling the "the rhetorical flexibility necessary for adapting [students'] socio-cognitive genre knowledge to ever-evolving contexts" (p. 238).

However, pairing linguistics approaches with genre teaching often elicits strong opinions from writing studies scholars. For almost forty years, researchers have explored genre through three different scholarly traditions (see Hyon, 1996; Russell, Lea, Parker, Street, & Donahue, 2009). Many North American scholars promote the New Rhetorical tradition, which emphasizes the activity of genre. An important contribution to this ideology is Miller's (1984) description of genre as "typified

rhetorical actions based in recurrent situations" (p. 159). New Rhetoricians approach genres as complex social practices that evolve from communal needs rather than linguistic templates (Miller, 1984; Schryer, 1993). Resulting WAC pedagogies then emphasize raising students' genre awareness rather than their genre acquisition (Devitt, 2004, 2009).

On the other hand, proponents of English for Specific Purposes (ESP) believe that genre construction correlates with the objectives of the communication situation, influencing lexio-grammatical features, organization, content, and disciplinary conventions (Hyland, 2004; Johns, 2002; Paltridge, 2000; Swales & Feak, 2011). ESP-based research has addressed genre awareness within diverse academic disciplines, including engineering, nursing, law, chemistry, archaeology, history, and literature (Kuteeva, 2013; Maswana, Kanamaru, & Tajino, 2015; Staples, 2015; Stoller & Robinson, 2013; Tessuto, 2015). The large amount of corpus-driven research related to ESP might lead New Rhetoricians to question to value of corpora in their classrooms and its relationship to raising students' genre awareness.

In fact, corpus-linguistic methods complement many of the underpinnings of New Rhetorical Studies. Sinclair (1991) argues that corpus analyses lead to new observations and models of language; the approach examines how specific language patterns are actually used rather than how these patterns should be used. A text's form and function though are inseparable: a form is not just random, but intricately tied to performing a specific function, be it syntactic, pragmatic, or meta-discoursal. Corpus linguistics is fundamentally different from a truly *formal* approach to language, in which forms are often seen as arbitrary, random, and therefore, inconsequential to functions and studied in their right entirety. These ideas challenge previous orthodoxies of language patterns but have been supported via substantive, empirical analyses of authentic language data (for example Hunston & Francis, 2000; Römer, 2009; Sinclair, 2004).

Though the three genre traditions promote diverse pedagogies, its scholars unite around the belief that the awareness of a text's form, function, and situation (albeit in different proportions) facilitates students' development of expertise. To promote genre awareness, Johns (2002) recommends that students analyze, discuss, and reproduce the structures and language elements of the targeted text to better understand how it varies across disciplines: "...if [students] don't study textual variety and the disciplinary ideologies that infuse genres, they can, and do, fall on their faces when they attempt to read and produce texts in their classrooms" (p. 249-50). Genre awareness is particularly important for technical editors because they cannot effectively analyze a text if they do not understand its expected conventions or how the intended readers traditionally engage with the information. Awareness then becomes a critical component to a technical editing course, which scholars like Rude equate with performance classes: "student editors become conscious of their craft by observing and analyzing the writing of others. This awareness of craft increases their competence in writing and as future practitioners" (p. 64).

Developing Data Information Literacies

The technologies and approaches associated with corpus-based learning also promote students' data information literacies. These literacies encompass the ability to think critically about concepts and arguments as well as read, interpret, and evaluate information (Schield, 2004).

The increasing availability of large-scale data coupled with user-friendly analysis and visualization tools have changed research practices across disciplines. Research teams must now submit data management plans in order to obtain federal funding from agencies like the National Science Foundation (NSF). NSF's *Cyberinfrastructure Vision of 21st Century Discovery* recommends enhanced instructional methods for data observation and interpretation that engage students and challenge

traditional discipline-based curricula (p. 38). However, scholars have questioned whether faculty or students are prepared to meet these challenges (Carlson, Fosmire, Miller, & Nelson, 2011).

Technical communicators, for instance, are typically expected to possess data management skills. Pflugfelder (2013) states that technical communicators' relationships with computerized data management professionals have intensified alongside the pressures of communicating with industry experts who work with data warehousing methods and object-oriented and post-relational database systems. The complexities related to analyzing and interpreting large data sets also utilizes one of a technical communicator's greatest strengths—the ability to produce persuasive narratives from data (p. 19).

These expanding employment requirements are important for aspiring technical editors, who are often associated as either only a proofreader or only a quality assurance analyst (Corbin, 2010). However, editing instructors are not necessarily adapting their courses or approaches to the technologies, genres, and processes that were nonexistent a few years ago. Similarly, editing students' proficiencies with current technologies has also been questioned. Rude observed that students are often "surprisingly unaware" of available tools in word processors like page numbering, running headers, and styles, and digital markup and document types were typically "alien concepts" (p. 62). She argues that technology has influenced contemporary technical editing and educators must, in turn, adapt their curricula to these evolving practices.

In the following case history, I offer quantitative and qualitative insights into how 14 technical editing students used corpora to complete a comprehensive edit of a research text and engage with their clients. The corpus work was intended to make these student editors more mindful of the writing nuances in their client's discipline and therefore produce a more appropriate final product.

Approaches and Materials

In this section, I describe the student participants who contributed to this case history and then explain why I used corpus-linguistic methods with this population. I then outline the various ways students were exposed to corpora and the tools they used to explore disciplinary writing conventions.

Student Editors

Fourteen students contributed to this case history. Twelve students were majors in the MA program in professional and technical communication, and two other students were doctoral students in English. The two doctoral students were concurrently enrolled in the certificate program in teaching technical writing. The average age of these students was 29.7 years ($sd = 5.14$, median = 30), and 57.1% ($n = 8$) were female. All student editors spoke English as their first language except Marisol, whose first language was Spanish. (I refer to students by pseudonyms throughout.) Data were collected with IRB approval.

This case history emphasizes how students used corpora to complete their final course project. Each student editor was paired with a client who spoke English as a second language. The pairs worked together on a client's research writing that conformed to the IMRD format (Introduction-Methods-Results-Discussion). Editors met with their clients twice; once to discuss the editorial process and again to discuss the final edits. During this final meeting, editors tutored their clients on the 2-3 issues that would improve their future writing. The clients were graduate students in the STEM disciplines, including materials science, electrical engineering, environmental science, psychology, and

geography. Clients represented a variety of native speaker backgrounds, including Thai, Farsi, Arabic, Korean, and Hausa.

Instructional Approach

I chose corpus-based instruction because it met the needs of my student population. A MA program in professional and technical communication is traditionally classified as a professional degree; students enroll in the program with the end goal of working as a technical communicator rather than transitioning into a doctoral program. Professional degree programs are also common to economics, engineering, and psychology. Students therefore require a skillset that can be applied to their future workplace responsibilities. Corpus-learning offers a methodology for exploring different rhetorical and linguistic variation, which is arguably more useful than teaching variation itself. In particular, technical editors work with subject matter experts from various academic backgrounds, necessitating linguistic awareness that aligns with disciplinary expectations.

I taught corpus lessons using a common four-step structure to DLL (Chujo et al., 2009). Each unit began with students seated at their computers, running a related corpus and the text processing tool: (i) students followed a worksheet of exercises that familiarized them with the text processing tool as well as directed them to examine a corpus for particular information or patterns. (ii) Students shared their findings with the class, and I explained any identified patterns and rules. I provide explicit explanations at this stage, so students could confirm or correct their hypotheses. (iii) Students were given a second worksheet of follow-up exercises for homework that also encouraged them to use a corpus in ways that were specific and meaningful to them. (iv) I provided students feedback on their homework. I followed this cycle during the first eight weeks of the semester. I then transitioned to primarily using the first two steps so students could focus their outside class time on editing and working their client. Students were also given more in-class work time, allowing me to provide feedback that was more focused on individual editing abilities and client experiences.

The Corpora

The overriding learning objectives to my lessons were to help students examine writing variations across academic disciplines, investigate other linguistic and rhetorical features for their own editing purposes, and apply these findings to their editing process or translate them to their clients. I achieved these objectives with three corpora:

- **Corpus of Contemporary American English (COCA).** COCA is the largest freely available corpus of American English with 450 million words that are divided among spoken, texts, fiction, magazines, newspapers, and academic journals (Davies, 2008-present). The academic journal register contains texts from nine disciplines.
- **STEM Student Writing Corpus.** I compiled a corpus of student-written IMRD-format texts from various STEM disciplines. The 50,000-word corpus of developing STEM writing provides an important contrast to the expert academic writing in COCA.
- **Editorial Comments Corpus.** I compiled this 28,000-word corpus of editorial comments from previous students' assignments to help students apply understanding of linguistic politeness strategies in editorial comments.

The corpus compilation process is one of the perceived barriers to using corpora in the classroom. The final section of this paper addresses this barrier with more depth; however, the process mainly involves saving content as a text file and therefore less time-consuming than often believed.

My students explored the course-compiled corpora with the free text processing tool AntConc (Anthony, 2011). AntConc was designed for use in technical writing courses, and it provides the functionality needed to test language hypotheses (Anthony, 2005, 2009). For example, a preliminary step in understanding writing variation across disciplines is to explore vocabulary. AntConc allows users to generate word lists, which contain the words that are most frequent in the corpus. Users can further their understanding by generating a keyword lists. Keywords are derived by comparing words in the targeted corpus to a reference corpus.

To illustrate this function, a technical editor might want to understand vocabulary differences between electrical engineering and mechanical engineering. Results from Hyland's study found that *we* was a significantly unique word to electrical engineering. Students might hypothesize then that electrical engineering is (a) a more collaborative discipline than mechanical engineering, or (b) perhaps the personal pronoun is more standard in electrical than mechanical engineering. The hypothesis could be tested in a number of ways, such as reviewing the specific contexts that contain *we* (called concordances in corpus linguistics and in AntConc) or by generating a plot that graphically represents where *we* occurs within the texts. Student editors might discover that high shares of *we* cluster around the middle, which is typically the materials and methods section of an IMRD-organized text.

Corpus Lessons

In this section, I describe students' introduction to corpora as well as how I facilitated their exploration of writing variation across disciplines. These lessons helped developing editors analyze how writing is socially constructed within disciplinary cultures and thereby encourage them to edit based on this knowledge rather than generalist principles. Though I created these lessons for a specific course, they can be adapted for students across grade levels and academic majors to better reflect how language influences their discipline.

Semantic Prosody and Collocation: Introducing Corpora

I introduced students to corpora during the second class period through the concept of semantic prosody and collocation. Semantic prosody describes ways in which seemingly neutral words take on positive or negative associations based on their collocations (Louw, 1993). Collocations, in turn, are "the company a word keeps," or the words that frequently co-occur with a given word or phrase (Firth, 1957). Though students are typically unfamiliar with the terminology, the concepts actually extend their prior knowledge of a word's denotation and connotation.

To begin this lesson, I asked students to write a sentence that included the noun *teenager*. Students then read their sentences, which included "Teenagers listen to loud music," and "The teenager smoked at the mall." As students read their sentences, I wrote down words they associated with *teenager*, such as *loud* and *smoked*. A pattern soon emerged that *teenager*, a seemingly neutral word, had a negative semantic prosody. I next introduced students to the collocation function in COCA (AntConc also has this function for exploring your compiled corpora). Students identified the collocates of *cause*, another word with perceived neutrality (Stubbs, 1995). The results indicated that *death*, *problems*, *damage*, *concern*, and *disease* were collocates of *cause*, giving this word a negative semantic prosody (High blood pressure is the third leading *cause* of *death* in the U.S.).

This initial lesson demonstrated how corpora contextualized language differently than traditional methods and often in more meaningful ways. Understanding semantic prosody and collocation makes students more critical communicators, and the methods used to understand these features offer an alternative to a thesaurus or other language references. Students soon transferred this

understanding to assist them in their editing assignments. For example, Bryant used COCA to make word choice suggestions to his first editing client. He suggested using the phrase *legislative branch* in a text on US policymaking because it was more associated with the subject than the author's original phrase *legislative party*.

Registry Variation: Distinguishing Between Features of Spoken and Written Communication

To move toward exploring writing across the disciplines, the second series of corpus lessons focused on language variations across registers. As noted earlier, the contents of COCA are equally divided among five registers, including academic writing and speech communication. Linguistic and rhetorical patterns in academic writing typically contrast with oral communication, so exploring the same pattern within both registers is often a valuable pedagogical tool. Developing students often struggle with professional writing because their writing style is similar to their speech patterns. An issue technical editors may encounter in their clients' work is the presence of colloquialisms used in spoken language. In the documentary, "Do You Speak American?," for example, the assistant managing editor of *The Columbus Dispatch* discussed how his print journalists were being affected by the spoken journalism of radio and TV, including the misuse of words like *importantly*, *nonplussed*, and *bemused* (Cran & MacNeil, 2005). The editor affirmed his employees were good journalists but could not always distinguish between features of spoken and written communication.

To illustrate linguistic awareness across registers, I adapted materials from a corpus activity on *must* (Reppen, 2010) and an exercise from my students' technical editing textbook (Rude & Eaton, 2011). Reppen selected *must* for her register awareness activity because of how the intensifier is used differently in spoken communication and academic writing. In speech, *must* often appears in sentences where personal pronouns are in the subject position whereas inanimate nouns or people referred to by their professions take the subject spot in academic texts. Select concordance lines from COCA (see Table 1) illustrate these differences with words like *your* and *I* accompanying *must* in the spoken register and *programmer*, *verbs*, and *module* in the academic register. Likewise, *must* is not always used to convey obligation in spoken texts as also illustrated in the concordance lines. Mixing elements of these registers then can create unnecessarily terse communication that impedes action or compliance—you *must* finish all work today versus all work *must* be completed before 5pm. The earlier described behaviors from *The Columbus Dispatch* journalists suggest that even expert writers struggle with distinguishing certain words or phrases across registers, so developing communicators would likely benefit from concentrated instruction in this area.

Table 1. Different uses of must in spoken and academic registers.

Spoken Register	Academic Register
...she said, "Susanne, call the police." ROBIN-ROBERTS-1-A# (Off-camera) Your heart must have been racing. SUSANNE-ARNOLD-1F# I was frightened for her, very frightened. I	...door The door is locked. To make the game interesting, a programmer must populate the virtual world with places and things that have rich behavior. Most any...
...if he does go out, what a way. GEORGE-STEPHANOPOU# (Off-camera) How special	Things must be defined and registered in much the same way as places. Verbs must be collected into a list to be used by the game's command parser.

it must have been for you to watch this, and I'm just wondering...	
...quick, and we need to develop that kind of military. DAVID-AXELROD-1-0# But I must say, I don't think the President needs to justify his national security bona...	...define-verbs followed by define-everywhere, then allows any number of other declarations. The module must end with a place expression, which is used as the starting location for the...

The editing textbook materials were two versions of a memo on a company's snow policy. The original version was written by the company CEO, who described his new policy in a detached, brusque tone. For example, a passage in the original memo used the word *must* four times, three of which were followed by the personal pronoun *you* (If you work from home you *must* be connected to the office via VPN during your work hours). The revised memo, written by a former employee, presented the same information with more inviting language. The intended purpose of these materials was to show the impact of editing for tone and *you*-attitude, but they also highlight nuances between registers.

Based on previous experiences with the snow policy exercise, I knew the instances of *must* bothered student editors. When the issue arose in the current class, I encouraged them to return to COCA and revisit corpus approaches to addressing their language questions. I demonstrated how to access the five registers and encouraged them to use *must* as their search word. Marisol was the first to determine that *must* occurred in the spoken register the least and in the academic register the most. Students then examined the concordance lines in each register and gradually came to similar conclusions as Reppen, namely that *must* is primarily used in academic texts to convey obligation but typically paired with inanimate nouns rather than personal pronouns. The mixing of these conventions in the original snow policy memo—using personal pronouns and *must* to convey obligation—likely impeded employee compliance and showed the author's faulty awareness to audience.

During the lesson, I also observed more students internalize the functionality of the COCA interface, the corpus terminology, and the relevance of their findings. For example, Courtney questioned the use of *docked* in the snow policy memos (you will be docked pay) and hypothesized the word had a negative semantic prosody. In fact, her COCA research found that *docked* had a more neutral semantic prosody and collocated with words like *boat*, *yacht*, *yard*, and *battleship*. Further, she identified *cut*, *reduced*, and *deducted* as close synonyms of *docked* and concluded that any would have been better word choices.

Citation Signals and Metadiscourse: Exploring Writing across the Disciplines

Once students were familiar with corpora and the related approaches, they explored language features within the academic register of COCA as well as the class-compiled corpus of student STEM writing. I created a series of lessons around the following concepts:

- **Hedges and boosters.** These are communication strategies used for increasing or decreasing the force of statements. I initially introduced students to these strategies in relation to the style and tone of their editorial comments. Both concepts were revisited in relation to students' final project where they analyzed the force of statements in research writing, particularly in the literature review and results section. (See Hyland, 2004, Appendix 3 for a list of common hedges and boosters). Understanding these language patterns are important

for developing writers who either overstate the significance of their results or understate how their results contribute to a broader research conversation.

- **Citation signals.** These broadly classify how writers mark their ideas (Swales, 1990). Integral citations include the name of the cited author within the body of the sentence (Smith (2015) recommended corpus lessons in all writing courses). Non-integral citations reference the author within parenthesis, brackets, or by superscript (Corpus lessons have been recommended for all writing courses (Smith 2015)). Hyland (2004) found that expert academic writers in all examined disciplines but philosophy used significantly more non-integral than integral citations. As a result, these citations emphasized the idea or subject over the author who conducted the research. Understanding why authors choose integral and non-integral citations is important for developing writers, who oftentimes overuse integral citations to compensate for their uncomfortableness in accurately summarizing or paraphrasing someone else's research.
- **Metadiscourse.** This encompasses how writers project themselves into their writing and also manage their communicative intentions (Hyland, 2004, p. 109). Hyland divides metadiscourse into textual and interpersonal. *First, next, and finally* are frame makers, which writers use to signal a change in topic. This is textual metadiscourse. The phrases *I find it interesting that* and *it is alarming to note* are attitude markers, which writers use to express surprise, importance, or obligation. This is an example of interpersonal metadiscourse. As illustrated by both, textual metadiscourse simply organizes content for the readers while interpersonal metadiscourse establishes a connection between the writer and readers (See Hyland 2004, Appendix 3). Understanding metadiscourse can help developing writers build confidence in communicating and organizing their ideas as well as persuading readers to accept these ideas.

Results from Final Editing Project

In this section, I report how my students applied corpus learning to their final editing project. This project included a comprehensive edit of a 25-30 page research text for a client who spoke English as a second language. Students were also required to tutor on the 2-3 issues that would improve their client's future writing. Corpus-based materials and approaches were not required for this assignment, but nine students (64% of the class) created client materials from their own corpus investigations or based on in-class materials. The following discusses the four themes that emerged from students' reflections on the editing process and working with their clients.

Validated work of editors

One major theme that I noticed from the final project materials was that students used their corpus training to validate their editing approaches. For example, Marisol read four journal articles in information systems to familiarize herself with her client's research topic and the common writing style. Though she did not perform a corpus analysis on these pieces, she found her independent review connected to the class work on metadiscourse and citation signals and thus validated her instincts when editing the client's paper.

Tony also applied his instruction on integral and non-integral citations to his client's paper, a doctoral student in psychology. Unlike Marisol, Tony analyzed several journal articles in this discipline with AntConc but appeared confused when the results confirmed how his client was already communicating ideas: "I'm not entirely clear on what these results mean, and how I can use them to help edit my client's paper. It appears that my client is falling right in line with the way that published

authors are using citations in their own papers." Though the client used citation signals that aligned with the standards in her discipline, Tony seemed to question the value of having this knowledge confirmed. I later discussed this issue with Tony, who acknowledged that his editing training primed him to recognize what was wrong with his client's writing rather than what clients intuitively did well. After further reflection, he admitted that sharing the citation findings might have helped his client's confidence with her writing ability and explicated a language pattern she likely did not know was standard in her discipline.

Built confidence in editorial decisions and established authority with clients

Another theme I observed was that students used corpus materials to build confidence in their editing abilities, thereby establishing authority with their client. Marie recounted that she removed what she considered unnecessary metadiscourse from her client's thesis chapters, such as relational markers *in regard to* and *it is crucial to note*. When the client questioned these suggestions, Marie offered a handout on metadiscourse types. She supplemented this information with a variety of original and revised sentences from the client's work that further justified her decision: "I explained [to the client] that he has a message he wants to convey to readers, and when he adds too much metadiscourse, his message gets buried under all the words. He seemed to understand it better after that." Though Marie did not cite results from a corpus analysis, she used related materials and concepts to establish her editorial authority.

Other students established authority with clients by referencing their own corpus research. Courtney said her client's overuse of hedges detracted from her sense of authority. When preparing to discuss this issue with her client, a logistics systems doctoral student, Courtney created a handout that defined hedging and explained how it could weaken academic writing. She also provided Hyland's (2004) list of hedge words (p. 188) and three excerpts from the *Journal of Business Logistics* and asked her client to analyze their effectiveness compared to the client's original paper:

Together we identified the words that mitigated their impact, and discussed strategies to make these statements stronger. [The client] was surprised that there was a term for the excessive diplomacy in her writing; she said this was something she had noticed, but had never been able to articulate. She attributed her unwillingness to be fully assertive to the fact that she's a relatively new PhD student—sometimes it's difficult to believe you're the authority.

Encouraged editor-client communication

Another prominent theme was how corpus learning motivated constructive editor-client communication. While the previous example demonstrated how students like Courtney used corpus approaches to engage their client, others actually demonstrated these methods. Kimbra loaded AntConc onto her client's computer to justify some of her editorial suggestions. Similarly, Leeann showed her client how to compile a corpus in an effort to improve his future academic writing:

I explained to [the client] that since he was already saving articles related to his field of study, he could easily create a reference corpus that would benefit him beyond his dissertation. I told him if he writes any more research papers that a corpus would come in handy, and that the more articles he loads to it the stronger the results get.

In addition to building this specialized corpus for her client, Leeann analyzed it for metadiscourse and created a handout with the results. These types of exchanges provided clients with content to improve their future writing, extending the value of technical editors beyond a single project.

Motivated self-reflection and improved professionalization

The final theme suggested that corpus lessons encouraged students to consider how writing varied across disciplines, and how this knowledge might impact future editing experiences. For example, Courtney's educational background was in the humanities, but she also observed that certain academic disciplines commonly use language devices like passive voice construction that she was taught to avoid in her own writing. This type of instruction reflects a humanistic-influenced approach to teaching style, especially in technical writing (Wolfe, 2009). But while technical writing is a people-oriented field, writing for end users, customers, and clients, engineering and biology are object-oriented fields. Corpus analyses of research texts within both disciplines conclude that researchers used passive constructions with clear purposes like emphasizing and organizing information to read more effectively (Conrad, 1996; Ding, 2001). Courtney reflected on and eventually reconciled these language nuances through her editing process: "I've spent my academic career entrenched in the humanities, so it was initially odd to 'fix' a paper by applying passive voice."

Results from Post-Treatment Survey

After final projects were graded and the semester's grades were posted, I invited students to respond to a survey that assessed their opinions on corpus-based learning, including enjoyment and choice, perceived value, transference, and how often they consulted corpora. Survey questions were inspired by instruments used in previous studies (Farr, 2008; Yoon & Hirvela, 2004) as well as my own interests related to using corpora in professional contexts. Students registered their agreement level on 15 statements on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). Table 2 (below) includes the raw data. All students participated in this survey, and the following reports the quantitative and qualitative results.

Enjoyment and choice

The first three survey questions related to students' enjoyment of the corpus activities and their choice in using these approaches. Overall, 85.7% of students enjoyed the course's corpus-based approaches ($M = 5.64$), and strongly agreed that they understood the purpose of using these approaches in their technical editing course ($M = 6.14$). Understanding the purpose of corpora in a technical editing classroom is arguably an accomplishment in itself, and the responses to this question yielded the highest statistical mean in the survey. Similarly, the majority of students also expressed a desire to work with corpora in future classes ($M = 5.36$).

Perceived value

The next five questions related to students' perceived value of corpora to their professional careers. Students typically agreed their general technical communication education benefited from their exposure to corpora ($M = 6.07$) and somewhat agreed that this exposure made them stronger technical editors ($M = 5.64$) and benefited their understanding of language systems ($M = 5.64$). Other students like Erik questioned the relevance of corpora: "I wish I were more confident that my exposure to corpora had made me a better editor. I'm still too unsure how to effectively apply it. And how can a practitioner go about creating a corpus? Ehhhh, there's a lot of *so what?* I still need addressed." Finally, students somewhat agreed that their understanding of large, complex datasets

benefited from their exposure to corpora ($M = 5.93$). However, students neither agreed nor disagreed that this exposure made them more comfortable with applying for jobs that required experience with large, complex datasets ($M = 4.86$). Courtney acknowledged that phrases like *big data* and *text mining* still seemed "intimidating and cryptic," but she was also more confident in her abilities to discuss and navigate the related concepts in professional settings.

Transference

Another five questions gauged how students might use corpora in their future professional careers. Students somewhat agreed they would use corpora for future technical editing projects ($M = 5.57$) and for future technical communication projects ($M = 5.79$). Tony expanded on how he could use corpora: "I can see a use for corpus-based approaches when sifting through old documents—I could enter a search term and every document that included that term could be pulled up. This would help in identifying all the different ways the term was used and inform future documents."

Students slightly agreed they would use the results from corpus-based inquiries to justify editorial suggestions to a primary decision maker ($M = 5.86$). Along these lines, students agreed they would use corpus-generated results to justify how communicators from different disciplines wrote ($M = 6.14$), and slightly agreed they would use these results to propose changes to writing practices in their workplace ($M = 5.79$). "From what I gathered," Erik wrote, "corpus studies are best for determining best practices. I think that's useful if you are leading a team or establishing standards within your company." Additionally, before the semester had concluded, students like Courtney, Kimbra, and Leeann had already integrated corpus approaches into their professional careers. After the semester's end, Leeann emailed the following:

I plan to meet with my team leader to suggest a special project. I want to analyze all of our current documentation and firm up the style guide we've been creating. I think using the AntConc tool will help us check our documents for things we want to be consistent with across all of our documents (e.g. writing "type" or "enter" in certain contexts), as well as things we want to stop doing. It can also help us make sure we are using the terms consistently.

Actual use

Students' desire to work with corpora in the future professional settings, however, was somewhat disconnected to how often they actually consulted them. Students stated they voluntarily consulted corpora for their technical editing projects about 30% of the time ($M = 3.29$) and less than 10% for projects in other courses ($M = 2.50$).

Desire to use corpora compared to actual usage might correlate to students' comfort level with both the corpus terminology and related tools. Five students commented that it took them longer than expected to learn AntConc. Marie, for example, said she understood how to use this tool in class, but then felt lost when she tried to use it at home. As reported earlier, Marie used concepts from the corpus lessons into her final ESL editing project, but did not actually conduct her own analysis. Additionally, Nicholas wrote that the results he generated with the corpora made some projects difficult to finish:

Instead of judging something holistically, I looked at the results, and then reflected on what I thought about those results. Sometimes the results didn't match my cursory

holistic impression. While it was a more informed decision, I also spent time reconciling my initial impression with the actual results.

It is perhaps easy to dismiss Nicholas' complaint that corpora promoted his critical thinking; however, the comment also captures how this type of engagement can refocus language analysis and understanding. Nicholas' experiences might also reflect the process of learning to manage and digest the large amounts of data that corpus analyses can produce. As noted earlier, Rude observed that many editing-related technologies were often "alien concepts" to students (p. 62). Other researchers have also discovered that faculty expect their graduate students to carry out data management and handling activities; however, they also expressed some hesitation in teaching these skills to their students themselves (Carlson et al., 2011). In other words, the increasing availability of all these data sets has not simultaneously increased our own critical data analysis skills. In Nicholas' case, an oversaturation of information might have further complicated his performance on the assigned tasks.

Table 2 - Post-treatment Survey Data

Question	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree	Mean (SD)
<i>Enjoyment and Choice</i>								
I enjoyed using corpus-based approaches in this course.	0	0	0	2	4	5	3	5.64 (1.01)
I understood the purpose of using corpus-based approaches in this course.	0	0	0	0	2	8	4	6.14 (0.66)
I want to use corpus-based approaches in future courses.	0	0	1	3	4	2	4	5.36 (1.34)
<i>Perceived Value</i>								
My general technical communication education benefited from my exposure to corpora.	0	0	0	0	2	9	3	6.07 (0.62)

My understanding of language systems benefited from my exposure to corpora.	0	1	0	1	3	5	4	5.64 (1.39)
My understanding of large, complex data sets (aka "big data") benefited from my exposure to corpora.	0	0	0	1	4	4	5	5.93 (1.00)
I am more comfortable applying for a job that emphasizes large, complex data sets because of my exposure to corpora.	0	1	1	3	4	4	1	4.86 (1.35)
I am a stronger technical editor because of my exposure to corpora.	0	0	0	2	3	7	2	5.64 (0.93)
<i>Transference</i>								
I would use corpus-based approaches for future technical editing projects.	0	1	0	2	3	3	5	5.57 (1.28)
I would use corpus-based approaches for future technical communication projects.	1	0	0	1	2	6	4	5.79 (1.53)
I would use the results from corpus-based	0	0	0	1	3	7	3	5.86 (0.86)

inquires to justify editorial suggestions to a primary decision maker.								
I would use the results from corpus-based inquires to justify how communicators from different disciplines write.	0	0	0	1	1	7	5	6.14 (0.86)
I would use the results from corpus-based inquires to propose changes to writing practices in my place of work.	0	1	0	1	3	3	6	5.79 (1.48)
Question	Rarely (10%)	Occasionally (30%)	Sometimes (50%)	Frequently (70%)	Usually (90%)	All of the time	Never	Mean (SD)
<i>Actual Use</i>								
I consulted corpora for this course outside the in-class activities and homework.	0	5	4	3	1	0	1	3.29 (1.44)
I consulted corpora for projects outside this course.	5	4	2	1	1	0	1	2.50 (1.79)

Integrating Corpus-Based Learning across The Disciplines

This case history focused on integrating corpus-based learning in a graduate-level technical editing course; however, the approaches I described can transfer to any course with a writing component as well as across grade levels and student proficiencies. Despite the acknowledged value of corpus-based learning, educators still perceive a series of barriers to its implementation. In this final section,

I address the three primary barriers to corpus learning, including a lack of discipline-specific corpora, a lack of knowledge in using text processing tools, and a lack of knowledge of appropriate teaching materials.

Lack of discipline-specific corpora

One barrier to integrating corpus-based learning is the lack of publically available corpora that contain relevant professional and technical texts. Even fewer corpora contain student-written as opposed to expert-written texts. The lack of these corpora become more pressing in view of the observed differences in the writing conventions and text types across disciplines, such as chemistry (Stoller et al., 2005; Stoller & Robinson, 2013), biology (Conrad, 1996), and computer science (Laurence Anthony, 2001). Appendix A lists six corpora that contain relevant student-written texts (proposals, reports, research papers) and expert-written texts (research papers, correspondence, blog postings) from a variety of academic disciplines.

Additionally, corpus compilation is particularly useful for teachers who want to tailor instruction to their students' needs. For example, Lee and Swales (2006) helped students compile a corpus of their own writing (term papers, dissertation drafts, unedited versions of published papers) that they explored alongside a corpus of published research articles. Corpus compilation projects can help students better understand their specific writing strengths and deficiencies as well as provide teachers with a resource they can continue to develop for future classes. Compiling corpora for classroom use involves converting original texts into text-only formats (ASCII). For teachers, this task becomes easier if they ask students to save their own writing as text files. As an example, if 15 students contributed four pieces of academic writing (500-1,000 words each), they could compile a 30,000-60,000-word classroom corpus with minimal time investment. See Reppen (2010), Chapter 4 for instructions on creating corpora for classroom use.

Involving students in the corpus compilation process engages them in their writing development and introduces them to data management skills. The latter was particularly important for students in this case history because they were preparing to become practicing technical communicators. By the semester's end, three of my students had transferred these classroom skills and compiled their own corpora at work. They used these corpora for a variety of purposes, such as maintaining style guide consistency, justifying editorial decisions to clients, and revising text types commonly used within the organizations.

Lack of knowledge in using text processing tools

Another barrier to integrating corpus-based learning is the lack of text processing tools and interfaces that appeal to users not trained in corpus linguistics. This accessibility gap discourages even foreign language teachers from using corpora and available tools (Meunier, 2010). Acquiring knowledge of relevant text-processing tools involves a learning process synonymous with learning any new technology, but the availability of user-friendly tools should not deter educators from integrating corpus-based learning.

Throughout this case history, I described COCA and the text processing tool AntConc. Again, COCA is the largest freely available corpus of American English, and the web-based interface allows students the flexibility to explore authentic language data from any computer. Students in this case history used COCA to explore vocabulary and semantic prosody; however, COCA contains a series of other functions that facilitate language learning through parts of speech within academic writing, creative writing, and speech communication. Teachers interested in integrating COCA into their classrooms will benefit from the tutorials offered via TheGrammarLab . AntConc was the free text processing tool

that my students used to explore the class-compiled corpora. AntConc includes a series of linguistic analysis features that were described earlier, such as concordances, collocates, word lists, and keyword lists. The video tutorials offered by creator Laurence Anthony (2015) are assessable to novice users and demonstrate the variety of ways students can use to explore writing.

Lack of knowledge of appropriate teaching materials

A final barrier to integrating corpus-based learning is a lack of knowledge in creating appropriate teaching materials. Corpora can be used to teach disciplinary variation in grammar and style, genre, and registry, but teachers should first identify their overriding learning objective. Corpus-based lessons should be staggered throughout the unit to help comprehension and also because not all students will respond to the approach. As an instructional starting point, Appendix B outlines seven manageable writing concepts that can be taught with corpora. In addition to defining each concept, I provide a rationale for why the concept should be taught; recommended methods for exploring the concepts, and references to key corpus-based publications. A variety of other resources are also available to teachers (see Conrad, 2008; O'Keeffe, McCarthy, & Carter, 2007; Reppen, 2010).

As teachers become more comfortable with corpus approaches and begin reading corpus-driven research, they will naturally begin to create their own lessons. Conrad (2008) offers three findings that have informed her corpus-based teaching: (i) the language used to persuade and argue in academic writing differs from the language used in everyday arguments like newspaper editorials or conversations, (ii) The patterns of language used across academic disciplines reflect subject area and a discipline's methods for building knowledge, and (iii) formulaic chunks of language (or lexical bundles) aid in reader comprehension because they are familiar and commonly used (p. 117). In academic writing, some of the most frequently occurring lexical bundles (3-, 4-, and 5-word phrases) include *in order to*, *one the other hand*, and *at the end of the* (see Hyland, 2008 for more information on this topic).

Corpus approaches can be powerful instructional components to teaching writing across the disciplines because they reveal language patterns that would be otherwise unobservable. Results of this pedagogical case history suggest that students typically enjoyed engaging with corpora and found these experiences valuable in working with clients from various academic backgrounds and validating their editorial decisions. Educators across the disciplines, particularly within the STEM areas, also acknowledge that their students' writing skills would likely benefit from corpus-based learning (Mudraya, 2004). STEM students' preferred learning styles are typically more active than reflective, more visual than verbal, and more sensing than intuitive (Felder & Henriques, 1995; Felder & Silverman, 1988). Allowing students to form and explore their own language hypotheses via corpora then activates these preferred learning styles and reflects the tenants of DDL. Integrating corpus-based approaches become more salient when considered alongside the idea that STEM students' preferred learning styles often contrast with STEM educators' preferred instructional approaches (Felder & Silverman, 1988; Felder & Spurlin, 2005). Corpora will not always provide students with definitive answers to their language questions, but they can encourage their discovery of preferred usage patterns and phenomena worth investigating.

Appendix A - Corpora Relevant to Teaching Writing

In addition to compiling corpora for classroom purposes, instructors have access to variety of corpora relevant to professional and technical writing. Below is a list of available corpora that are focused on American English. See Reppen (2010), Appendix B for additional resources on available corpora and corpus tools.

- Air Traffic Control (ATC) Corpus—houses 70 hours of recorded conversations between controllers and aircrafts in three majors US airports <https://catalog.ldc.upenn.edu/LDC94S14A>.
- Blog Authorship Corpus—houses the collected posts of 19,320 bloggers gathered from blogger.com. The corpus incorporates a total of 681,288 posts and over 140 million words, approximately 35 posts and 7250 words per person <http://u.cs.biu.ac.il/~koppel/BlogCorpus.htm>.
- Business Letters Corpus—houses 1 million words of US and UK business letters, most of which are model letters from textbooks <http://www.someya-net.com/concordancer/>.
- Michigan Corpus of Academic Spoken English (MICASE)—houses 1.8 million words of transcribed academic speech <http://quod.lib.umich.edu/m/micase/>.
- Michigan Corpus of Upper-Level Student Papers (MICUSP)—houses 2.6 million word of A-graded, student-written texts from 16 different disciplines. Text types include proposals, reports, and research papers. The MISCUSP interface allows for complex searches by discipline, writer level, and writer characteristics (e.g., native-speaker status, major, gender) <http://micusp.elicorpora.info/>.
- Professional English Research Consortium Corpus (PERC) – houses 17 million words of copyright-cleared English academic journal texts in science, technology, engineering, and mathematics. <https://scn.jkn21.com/~percinfo/>.

Appendix B - Writing-Related Corpora Lessons

Concept	Definition	Examples	Rationale	Methods	Resources
Demonstrative pronouns	Pronouns that are either free standing or function as determiners	<i>this</i> (plural <i>these</i>), <i>that</i> (plural <i>those</i>)	Not attending demonstratives like <i>this</i> with a noun phrase if typically considered a mark of poor writing. However, students are often deliberate in their choice to leave <i>this</i> unattended (e.g., delivering bad news, summarizing)	frequencies, concordances, collocations	Boettger and Wulff (2014); Wulff, Römer, and Swales (2012)
Hedges	Words or phrases used to avoid making definitive statements or to lessen the impact of those statements	<i>can, would, might, could, should, kind of, sort of, typically</i>	Hedging impacts the message being conveyed and readers' perceptions of the writer. Writers hedge intentionally and unintentionally, so the results can buffer a message or make the writer appear passive and noncommittal.	frequencies, collocations, n-grams	Hyland (1998, 2004)

Lexical bundles	Units of more than two words that occur repeatedly over many texts and many authors	<i>in order to, on the other hand, it should be noted that</i>	Frequently occurring word bundles shape meaning and contribute to a text's coherence. Bundles vary across genres, registers, disciplines, and native speaker status	concordances, collocations, n-grams	Biber et al. (1999); Cortes (2004); Hyland (2008)
Linking adverbials	An type of adverbial that explicitly states the relationship between two units of discourse	<i>however, though, for example, nevertheless, therefore, i.e., e.g.</i>	Linking adverbials are necessary features to technical writing because they create organization and aid in comprehension	concordances, collocations, n-grams	Biber et al. (1999); Friginal (2013)
Near synonyms	Words or phrases that a typically similar in meaning, but different in distribution	<i>continually v. continuously, mistake v. error, for instance v. for example, somewhat v. fairly</i>	Word choice is important in technical writing, and poor word choices can convey unwanted connotations or attitude. Certain near-synonyms often correlate with specific verbs and are linked to the concept of semantic prosody	frequencies, concordances, collocations	Lee & Swales (2006)
Passive constructions	Features of transitive sentences that occur when the object of an action occupies the subject position	This conclusion <i>was supported by</i> the common garden experiments of Gaudet and Keddy (1989)	Communicators in object-oriented fields like engineering and biology typically use more passive constructions than communicators in people-oriented fields like technical communication	frequencies, concordances	Conrad (1996); Ding (2001); Wolfe (2009)

Reporting verbs	A verb class that conveys the action of speaking and used with both direct and reported speech	<i>argue, illustrate, observe, think</i>	Reporting verbs correlate with how writers synthesize information from secondary sources. Students typically overuse <i>show</i> and <i>find</i> , and the verbs often correlate to specific disciplines	frequencies, concordances, collocations	Friginal (2013); Hyland (2004); Thomas and Hawes (1994)
-----------------	--	--	--	---	---

References

- Anthony, Laurence. (2001). Characteristic features of research article titles in computer science. *Professional Communication, IEEE Transactions on*, 44(3), 187-194.
- Anthony, Laurence. (2005). AntConc: design and development of a freeware corpus analysis toolkit for the technical writing classroom. Paper presented at the Professional Communication Conference Proceedings.
- Anthony, Laurence. (2009). Issues in the design and development of software tools for corpus studies: The case for collaboration. In P Baker (Ed.), *Contemporary Corpus Linguistics* (pp. 87-104). London, UK: Continuum Press.
- Anthony, Laurence. (2015). AntLab, from <https://www.youtube.com/channel/UC5bgZtdgKyj5un66ZOGGN2A>
- Anthony, Laurence. (2011). AntConc (Version 3.2.4w). Tokyo, Japan: Waseda University. Retrieved from <http://www.antlab.sci.waseda.ac.jp/>
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan, & Finegan, Edward. (1999). *Longman grammar of spoken and written English*. London/New York: Pearson Education Limited.
- Boettger, Ryan K, & Wulff, Stefanie. (2014). The naked truth about the naked this: Investigating grammatical prescriptivism in technical communication. *Technical Communication Quarterly*, 23(2), 115-140.
- Carlson, Jacob, Fosmire, Michael, Miller, CC, & Nelson, Megan Sapp. (2011). Determining data information literacy needs: A study of students and research faculty. *portal: Libraries and the Academy*, 11(2), 629-657.
- Chan, Tun-pei, & Liou, Hsien-Chin. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning* 18(3), 231-251.
- Chujo, Kiyomo, Anthony, Laurence, & Oghigian, Kathryn. (July, 2009). DDL for the EFL classroom: Effective uses of a Japanese-English parallel corpus and the development of a learner-friendly, online parallel concordancer. Paper presented at the American Association for Corpus Linguistics. Liverpool, UK.
- Cobb, Tom. (1997). Is there any measurable learning from hand-on concordancing? *System*, 25(3), 301-315.
- Cobb, Tom. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning*, 12(4), 345-360.
- Conrad, Susan M. (2008). Myth 6: Corpus-based research is too complicated to be useful for writing teachers. In Joy Reid (Ed.), *Writing myths: Applying second language research to classroom teaching* (pp. 115-139). Ann Arbor: University of Michigan Press.
- Conrad, Susan M. (1996). Investigating academic texts with corpus-based techniques: An example from biology. *Linguistics and Education*, 8(3), 299-326.
- Conrad, Susan M. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. In Susan M. Conrad and Douglas Biber (Eds.), *Variation in English: Multidimensional Studies* (pp. 94-107). Harlow: Pearson Education/Longman.

- Corbin, Michelle. (2010). The editor within the modern organization. In Avon J. Murphy (Ed.), *New perspectives on technical editing*. Amityville, NY: Baywood Publishing Company, Inc.
- Cortes, Viviana. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397-423.
- Cran, William, Buchanan, Christopher, MacNeil, Robert, Cassidy, Orlagh, Palmer, Allan, Frost, Joe, & Foss, Paul. (2005). *Do you speak American?: Episode 1*. Princeton, NJ: Films for the Humanities & Sciences.
- Cresswell, Andy. (2007). Getting to 'know' connectors? Evaluating data-driven learning in a writing skills course. In Encarnacion Hildalgo, Luis Quereda & Juan Santana (Eds.), *Corpora in the foreign language classroom* (pp. 267-287). Amsterdam: Rodopi.
- Davies, Mark. (2008-present). *The Corpus of Contemporary American English: 450 million words, 1990-present*. Retrieved from <http://corpus.byu.edu/coca/>
- Devitt, Amy J. (2004). *Writing genres*. Carbondale, IL: Southern Illinois University Press.
- Devitt, Amy J. (2009). Teaching critical genre awareness. In Charles Bazerman, Adair Bonini & Debora Figueiredo (Eds.), *Genre in a changing world* (pp. 337-351). Fort Collins, CO: The WAC Clearing House.
- Ding, Daniel. (2001). Object-centered—How engineering writing embodies objects: A study of four engineering documents. *Technical Communication*, 48(3), 297-308.
- Farr, Fiona. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, 17(1), 25-43.
- Felder, Richard M, & Henriques, Eunice R. (1995). Learning and teaching styles in foreign and second language education. *Foreign Language Annals*, 28(1), 21-31.
- Felder, Richard M, & Silverman, Linda K. (1988). Learning and teaching styles in engineering education. *Engineering Education*, 78(7), 674-681.
- Felder, Richard M, & Spurlin, Joni. (2005). Applications, reliability and validity of the Index of Learning Styles. *International Journal of Engineering Education*, 21(1), 103-112.
- Firth, John R. (1957). *Papers in linguistics 1934-1951*. London: Oxford University Press.
- Friginal, Eric. (2013). Developing research report writing skills using corpora. *English for Specific Purposes*, 32(4), 208-220.
- Garner, James R. (2011, unpublished). Does data-driven learning lead to better academic writing? University of Alabama.
- Hanson, David, & Wolfskill, Troy. (2000). Process workshops - a new model for instruction. *Journal of Chemical Education* 77(1), 120-130.
- Hunston, Susan, & Francis, Gill. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Cambridge: MIT Press.
- Hyland, Ken. (1998). *Hedging in scientific research articles* (Vol. 54). Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Hyland, Ken. (2004). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor, MI: University of Michigan Press.
- Hyland, Ken. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21.
- Hyon, Sunny. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly*, 30(4), 693-722.
- Johns, Ann M. (Ed.). (2002). *Genre in the classroom: Multiple perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Kuteeva, Maria. (2013). Graduate learners' approaches to genre-analysis tasks: Variations across and within four disciplines. *English for Specific Purposes*, 32(2), 84-96.
- Lee, David, & Swales, John. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25(1), 56-75.

- Liu, Dilin. (2010). Using corpora in treating lexico-grammatical errors in ESL writing. Paper presented at the Proceedings of the 2010 International Conference on ELT Technological Industry and Book, Neipu, Taiwan.
- Louw, Bill. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Mona Baker, Gill Francis & Elena Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 240-251). Philadelphia: John Benjamins Publishing.
- Maswana, Sayako, Kanamaru, Toshiyuki, & Tajino, Akira. (2015). Move analysis of research articles across five engineering fields: What they share and what they don't. *Ampersand*, 2, 1-11..
- Meunier, Fanny. (2010). Learner corpora and English language teaching: Checkup time. *Anglistik: International Journal of English Studies*, 21(1), 209-220.
- Miller, Carolyn R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70(2), 151-167.
- Mudraya, Olga V. (2004). Need for data-driven instruction of engineering English. *IEEE Transactions on Professional Communication*, 47(1), 65-70.
- O'Keeffe, Anne, McCarthy, Michael, & Carter, Ronald. (2007). *From corpus to classroom: Language use and language teaching*. New York, NY: Cambridge University Press.
- Paltridge, Brian. (2000). Genre knowledge and teaching professional communication. *IEEE Transactions on Professional Communication*, 43(4), 397-401.
- Pflugfelder, Ehren Helmut. (2013). Big data, big questions. *Communication Design Quarterly Review*, 1(4), 18-21.
- Reppen, Randi. (2010). *Using corpora in the language classroom*. New York, NY: Cambridge University Press.
- Römer, Ute. (2009). The inseparability of lexis and grammar: Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics*, 7(1), 141-162.
- Rude, Carolyn D. (2010). The teaching of technical editing. In Avon J Murphy (Ed.), *New perspectives on technical editing*. Amityville, NY: Baywood Publishing Company, Inc.
- Rude, Carolyn D., & Eaton, Angela. (2011). *Technical editing* (5th ed.). Boston: Longman.
- Russell, David R., Lea, Mary, Parker, Jan, Street, Brian, & Donahue, Tiane. (2009). Exploring notions of genre in "academic literacies" and "writing across the curriculum": Approaches across countries and contexts. In Charles Bazerman, Adair Bonini, & Debora Figueiredo (Eds.), *Genre in a changing world* (pp. 459-491). Fort Collins, CO: The WAC Clearinghouse.
- Schild, Milo. (2004). Information literacy, statistical literacy and data literacy. *IASSIST Quarterly*, 28(2/3), 6-11.
- Schryer, Catherine F. (1993). Records as genre. *Written Communication*, 10(2), 200-234.
- Sinclair, John. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, John. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge.
- Staples, Shelley. (2015). Examining the linguistic needs of internationally educated nurses: A corpus-based study of lexico-grammatical features in nurse-patient interactions. *English for Specific Purposes*, 37, 122-136.
- Stoller, Fredricka L., Jones, James K., Costanza-Robinson, Molly S., & Robinson, Marin S. (2005). Demystifying disciplinary writing: A case study in the writing of chemistry. [Special issue on the linguistically diverse student]. *Across the Disciplines*, 2. Retrieved from <https://wac.colostate.edu/atd/lds/stoller.cfm>
- Stoller, Fredrika L., Costanza-Robinson, Molly S, & Robinson, Marin S. (2013). Chemistry journal articles: An interdisciplinary approach to move analysis with pedagogical aims. *English for Specific Purposes*, 32(1), 45-57.
- Stubbs, Michael. (1995). Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of Language*, 2(1), 23-55.
- Swales, John. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, John M, & Feak, Christine B. (2011). *Navigating academia: Writing supporting genres*. Ann Arbor, MI: University of Michigan Press.

- Tessuto, Girolamo. (2015). Generic structure and rhetorical moves in English-language empirical law research articles: Sites of interdisciplinary and interdiscursive cross-over. *English for Specific Purposes*, 37, 13-26.
- TheGrammarLab. (2015). YouTube, from <https://www.youtube.com/user/TheGrammarLab>
- Thomas, Sarah, & Hawes, Thomas P. (1994). Reporting verbs in medical journal articles. *English for Specific Purposes*, 13(2), 129-148.
- Willis, Dave, & Willis, Jane. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.
- Wolfe, Joanna. (2009). How technical communication textbooks fail engineering students. *Technical Communication Quarterly*, 18(4), 351-375.
- Wulff, Stefanie, Römer, Ute, & Swales, John M. (2012). Attended/unattended this in academic student writing: Quantitative and qualitative perspectives. *Corpus Linguistics and Linguistic Theory*, 8(1), 129-157.
- Yoon, Hyunsook, & Hirvela, Alan. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13(4), 257-283.

Contact Information

Ryan K. Boettger
Department of Technical Communication
University of North Texas
1155 Union Circle #305298
Denton, TX 76203-5017
Email: ryan.boettger@unt.edu

Complete APA Citation

Boettger, Ryan. (2016, February 28). Using corpus-based instruction to explore writing variation across the disciplines: A case history in a graduate-level technical editing course. *Across the Disciplines*, 13(1). Retrieved from <https://wac.colostate.edu/docs/atd/articles/boettger2016.pdf>