Teachers, Researchers, and Communities of Practice: Building a Corpus to Support Graduate Education

Thomas Peele, Vivian Stoll, and Andréa Stella

ABSTRACT: The authors of this essay discuss the impact of corpus collection and analysis on the writing program at The City College of New York, CUNY, the digital literacies encouraged by the corpus collection process, and how corpus studies can be used to support genre awareness and build communities of practice in basic writing classrooms and among graduate students. New graduate student instructors collected the corpus of student essays and later used it in their classrooms to both introduce and reinforce what students already know about rhetorical moves and genre conventions. Since these corpus-based assignments were derived from our own student writers, they showed our graduate students what our undergraduate students already knew about academic writing and helped these first-year writers build upon that knowledge—the assignments showed them, visually, the kind and frequency of rhetorical moves of the argument essay in an academic community of practice. The corpus collection and analysis process provided new graduate student instructors with hands-on experience in one strand of composition research. As they began the process of learning to teach academic discourse to basic writers, English language learners, and other students of composition, they were also actively learning the discursive practices and analytical modes of composition researchers.

KEYWORDS: argument; basic writing; communities of practice; composition; corpus studies; discourse communities; graduate education; rhetorical genre studies

Thomas Peele *is an Associate Professor at The City College of New York, where he serves as the writing program administrator and coordinates the WAC initiative. His current research interests include corpus linguistics and qualitative studies of student knowledge transfer.* ***Viv Stoll*** *holds an MA in Language and Literacy from The City College of New York, where she has taught Freshman Composition and Writing for Engineering. Her experience as a musician and sound designer informs her interest in multimodal composition. She tutors at the writing centers of City College and the Center for Worker Education, CUNY.* ***Andréa Stella*** *holds an MFA in Creative Writing from City College, where she teaches first-year composition and Writing for Engineering. She is enrolled in the doctoral program in English at the Graduate Center, CUNY.*

> I can recall breaking down paragraphs and attempting to understand what made a paragraph a paragraph. I tried to understand what it was about the content in the first sentence that made it an introductory sentence, how it connected to the second sentence and the purpose of the content in the second sentence, how a line of reasoning was threaded throughout a paragraph and how it was concluded. I tried to understand how writing worked on a macro (meaning and content) and micro (punctuation and structure) level. Draft after draft, I would use a newly learned mechanism of writing.—Jamil Shakoor (Schnee and Shakoor 94)

> No matter what we think of these rules, obey is the only option. Every community formed its own language. . . If we are in school, this community of practice, then we have to follow the[ir] rules, because that's how this community works. People who can't follow the rules will be left out of the community, no matter how intelligent they are.—Marian (Carter, "Redefining" 119)

The corpus collection and analysis that we describe in this article introduced co-authors Viv Stoll and Andréa Stella, two new graduate student instructors of composition at The City College of New York (CCNY), The City University of New York (CUNY), to the multiple ways in which computer technology could be engaged in the service of writing pedagogy, and how corpus analysis could be used as a pedagogical tool in the classroom. The above epigraphs, drawn from basic writing students at Kingsborough Community College, CUNY, and Texas A&M Commerce, respectively, illustrate our motivation: to support students in the complex process of understanding one set of typical rhetorical structures of academic writing. At the same time, Tom, as CCNY's Writing Program Administrator, wanted new graduate student instructors to be enrolled in the composition teaching practicum, immersed in the discursive practices of rhetoric-composition researchers at the beginning of their teaching careers. As they introduced their students, comprised of a mix of English language learners, basic writers, high school, honors, and mainstream composition students, to the discursive practices of academic writing, they were simultaneously joining the discourse com-

munity of rhetoric-composition researchers. At our institution, the teaching practicum generally aims to support new instructors as they experience the challenges and rewards of creating assignments, responding to essays, and assigning grades, but they rarely have the opportunity to engage in classroom-based research.

The writing program's eventual ability to develop a corpus is the result of a series of disconnected events. For three years, using the model of revision analysis described by Lester Faigley and Stephen Witte in "Analyzing Revision," Tom had worked with graduate students to code revisions in essays written by first-year composition students. They focused on the categories provided by Faigley and Witte: surface changes and meaning changes (402-405). This kind of analysis, while helping to illuminate for graduate student instructors the kinds of revision that their first-year students were likely to make, was also frustratingly narrow in scope. Similar to Faigley and Witte's study, the graduate students originally involved in this project coded about six essays each, and compared their results. While we could see the revisions made by these six undergraduate students in great detail, we couldn't reliably extrapolate any broader patterns from this narrow sample. Given that this work took place between 2012 and 2014, and that all of the essays were collected digitally, the hand-coding of a very limited set of essays also seemed anachronistic. Tom wondered if it wasn't possible to conduct a larger scale analysis of revision.

In 2015, three things happened that made this large-scale analysis of revision possible. First, Tom attended Duncan Buell and Chris Holcomb's presentation at the Conference on College Composition and Communication, "First-year Composition as Big-Data: Natural Language Processing and Portfolio Assessment," which described the large-scale, detailed study of revision that he wanted to conduct but for which he did not have a corpus of student essays. Then, in the summer of 2015, he received a call from the Provost's office. At the end of the fiscal year, unspent grant monies had been returned, and they were in search of ways to spend it in support of faculty development. With this money, Tom was able to pay thirty-five current and former graduate student composition instructors $500 each to collect the first- and second-drafts of four assignments. This faculty development effort created a 6,311,220 million-word corpus containing first- and second-drafts of 4,280 essays (approximately 2,140 first-drafts and 2,140 second-drafts). He collected the four required essays in the first semester of the composition sequence: a literacy narrative, an expository essay, an exploratory essay, and a research essay. The corpus analysis that we describe was based on 548 final

drafts (1,465,091 words) of the argument-research essay. From the outset, then, the corpus project had graduate student development at its center. In order to collect the essays, instructors had to learn to use the Assignment tool in the Blackboard (Bb) course management system (Peele, "Blackboard"). Later in this essay, Viv describes the impact this process had on her pedagogy.

After the essay collection process was complete, Tom was repeatedly frustrated by his inability to recruit and retain a computer scientist to undertake the revision analysis. Even though he had already secured IRB approval for a study of the corpus, for reasons having to do with his home institution's Byzantine structure for approving non-CUNY employees, he needed to find an internal, CUNY colleague with whom to collaborate. While a few expressed interest, none were able to commit to the project. The 2015 publication of Laura Aull's book length, corpus-driven study of student writing, *First-Year University Writing: A Corpus-Based Study with Implications for Pedagogy*, though, showed him how he could conduct an analysis of rhetorical moves by using simple, free technology—Laurence Anthony's concordance software, AntConc—in order to use the corpus in graduate teaching and for basic writing and composition pedagogy. As Andréa describes later in this essay, incorporating corpus analysis into her teaching impacted her first semester as an instructor of a disciplinary-specific course, Writing for Engineering. True, Tom hasn't yet been able to conduct the study of revision in student essays that he had planned, but the corpus has provided a database of student essays from CCNY students, the study of which offers ample support for discussions of patterns in student writing across all classes.

In this essay, we describe the basis of the corpus study, the impact that the collection and analysis process had on Viv and Andréa, and how the study of rhetorical moves in student writing helped shape the philosophy and structure of the writing program. This study of local corpora provided graduate student instructors a record of the rhetorical moves that their students were making in the same social context in which they were teaching, knowledge that would guide instruction. As composition pedagogy researchers, graduate students were simultaneously joining a community of practice while learning to teach the discursive conventions of academic communities of practice.

COMMUNITIES OF PRACTICE IN CORPUS-DRIVEN BASIC WRITING AND COMPOSITION PEDAGOGY

We use Shannon Carter's description of communities of practice as a framing device for her concept of rhetorical dexterity to ground our discussion of corpus-driven basic writing and composition pedagogy. Our students were unlike Carter's in that they had not been labeled "not ready for college level literacy" as a result of state-mandated tests ("Redefining" 95). The City University of New York, of which City College is a part, does not rely on written entrance exams. However, City College is both a Hispanic and a minority-serving institution; the most current available data estimates that median parent income for our students is low, at $40,200 ("Economic"). Also CCNY students have diverse language backgrounds, and are likely to be first in their families to attend college, immigrants or from immigrant families ("City Facts"), and be less familiar with the expectations of academic writing than mainstream composition students. They also have diverse language backgrounds ("City Facts"). Although our students have not been marginalized in the same way that Carter's students have been, they share many of the demographic characteristics that could lead them to be labeled basic writers.

Before we begin our exploration of communities of practice, we want to emphasize that in our discussion of our study of the corpus, we focus exclusively on the argument essay only for reasons of time and space. We do not make any recommendations, explicit or implied, about the best content for basic writing courses. As our curriculum shows, we assign essays in a variety of genres. Space does not permit us to include discussions of more than one genre of essay. That genre—the loosely defined argument essay—provides us with a reference point for what Carter describes as a "familiar" community of practice ("Redefining" 99). In "Redefining Literacy as a Social Practice," Carter describes a pedagogical approach aimed at improving students' rhetorical dexterity, "that is, the ability to effectively read, understand, manipulate, and negotiate the cultural and linguistic codes of a new community of practice (the academy) based on a relatively accurate assessment of another, more familiar one" (99). In our context, we used the corpus as an example of a familiar community of practice. Admittedly, our use of loosely-defined argument essays is a far cry from the examples that Carter provides, which include practices familiar to her students, "like skateboarding, photography, basketball, *Halo 2*, and cheerleading" ("Redefining" 105). Carter notes that writing assignments can function "within the context of what we know about how literacy functions in the world beyond the artificial 'school' litera-

cies we often celebrate" (103). The specific assignment that Carter describes is, we imagine, significantly more complex and in many ways more satisfying in that it makes connections between non-academic and academic literacies. We contend, however, that the academic literacy represented by the argument essay is no less authentic than other literacies. Since we built our own corpus, these essays were literally written by our students; moreover, the form is ubiquitous in rhetoric-composition classes and, in variations, across the curriculum. Corpus analysis helps students to assess their own discursive practices in academic writing with a high degree of accuracy, and they can build upon that knowledge as they learn more about that genre in the college environment.

Aull's study, which examines various rhetorical moves that students make in response to a range of assignments, provides a useful model for the ways in which we might assess rhetorical dexterity. Based on a corpus of "19,433 essays written by FY students at two four-year institutions, as well as information about the 91-million word reference corpus of published academic writing from 1990-2013," Aull's multiple studies "suggest that students need more guidance about how academic writers use personal evidence as argumentative views (versus opinions) and also that there may be a connection between open-ended prompt questions and FY arguments that are generalized and personalized" (15). While Aull's study does not specifically illuminate our own, it does model multiple ways in which corpus studies might inform discussions of patterns in first-year writing with both graduate student instructors and with composition students.

As another model for our project, Zak Lancaster's study of rhetorical moves in student essays compares a narrow range of rhetorical moves in the same corpora that Aull used. By comparing the ways in which professional writers, advanced college students, and pre-college level high school students introduce objections to their claims (discussed in more detail below), Lancaster shows that writers in all three categories place a high value on interpersonal relationships. As Lancaster notes, writers "in all three groups preferred to acknowledge objections namelessly. . . They likewise preferred to interact with readers' views indirectly, without attributing propositions pointedly to the reader" (451). For example, instead of using the wording "some readers may challenge," student writers in particular are much more likely to use a formula such as "some would/may/might argue that" (451). The nameless objections ("some" instead of "readers") allow the writer to "project a reader-in-the-text with whom the writer can negotiate meanings without impinging on the reader's face. Since, that is, the alternative view is

left unattributed, the face-threatening-act (FTA) of purporting to know what the reader is thinking is mitigated" (448). Such information is potentially valuable to basic writing students as they acquire rhetorical dexterity as it shows how advanced academic writers make use of some of the rhetorical moves that are a common focus of instruction in basic writing classrooms—how to structure written arguments that engage multiple perspectives. The rhetorical information revealed by corpus analysis provides another kind of map for students in academic writing. Changes in digital technology make it possible for instructors to create a variety of corpora, including a corpus made up of their own students' essays, to produce a highly local map of these moves.

All our models echo the aims of rhetorical genre studies and writing across the curriculum pedagogy (Bawarshi and Reiff; Bazerman; Miller; Reiff and Bawarshi; Swales). Among other goals, rhetorical genre studies aims to demystify the elements of genre for students who have not been immersed in them while at the same time maintaining that genres are not fixed and hardened formulas but arise in social contexts and shift as culture shifts. Obvious examples of this are the text message, the social media post, and the comments section of digitally-delivered newspaper articles, none of which existed in their current form twenty years ago but all of which follow recognizable if malleable genre conventions and are seamlessly integrated into the lives of traditional-age college students in 2019. This aim, to introduce students to the elements of various genres as those genres are influenced by students' work, social, and academic lives, resonates strongly within basic writing studies, which has long understood the value of making explicit the conventions of academic writing and the importance of student subjectivity (Bartholmae; Bizzel; Delpit; Elbow; Peele and Antinori; Rose; Shaughnessy).

OUR CORPUS ANALYSIS PROJECT

As Viv and Andréa describe, the corpus collection and study during their graduate educations informed their teaching in multiple ways. It served to generate ideas about how to incorporate other digital practices and helped to demystify and isolate the typical rhetorical moves in academic arguments; the rhetorical moves that were mapped in our corpus analysis helped define, for instructors and students, the discursive practices of this academic community of practice. At the same time, including graduate students in a large-scale research project situated them within the community of practice of rhetoric and composition researchers. The corpus study showed how genre

conventions inform essays written by undergraduate writers and career academics; it demonstrated to them that all of our undergraduates—basic writers, mainstream composition students, and English language learners— are aware of and make use of identifiable rhetorical moves, as it provided a method for making clear to themselves and to their students some of the moves that are expected in argumentative essays.

Thus far, we have grounded the corpus collection and analysis project within basic writing and rhetorical genre studies; we now turn to three salient engagements with the process—collecting the data, analyzing the data, and more widely interpreting the data points—as well as to the specific impact that collection and analysis had on two new graduate student instructors: Viv and Andréa.

Collecting the Data: Competing and Collaborating in Expanded Spaces

In the fall of 2015, Viv was a master's degree student in the Language and Literacy program at City College, and it was her first time teaching composition as an adjunct instructor, so much of the protocol was new to her. Viv was extremely happy to find that new instructors would be paid to participate in a large-scale corpus collection project that included training on how to create a syllabus and assignments, incorporate the required textbook, as well as use Bb for posting assignments, communicating with students, and collecting their work. The collection process had a significant, positive impact on Viv. The requirement to use the Bb Assignment tool to collect the essays prompted her to develop multiple digital literacies, which affected how she archived, graded, and evaluated student essays. Because of the digital collection process, she had not only individual essays from individual students, but, conceptually, a body of digital essays, the study of which might reveal information that would be useful to the class overall. She was led to consider the ways in which her students' essays were not only texts, but specifically digital texts. If essays are mediated on a digital platform, how might she be able to make use of that platform's affordances?

Little did she anticipate that her participation in the corpus collection process would lead her to a significantly expanded sense of place. As she developed her digital processes, she broadened the technological and cultural base upon which her course rested. By using automated processes to archive student essays, she reproduced the archival processes of large-scale, publicly available databases such as YouTube, which automatically collects, organizes,

and makes its content available. Instead of resisting digital processes, Viv followed the logic of the database (Johnson-Eilola).

As a result of the emphasis on the digital collection process, by the time the fall semester began, Viv was prepared to run her class totally in the digital realm. All student writing for the class would be submitted digitally so that she could easily collect the drafts for submission to the corpus project. Because she was working fully in the digital domain, Viv could:

- preserve a clean copy of the student's original work
- type comments on the essay then "save as" to preserve her comments
- upload her responses to Bb for students to review
- submit unmarked first- and second-drafts of all essays to the corpus project

Viv developed a methodology for collecting and responding to students' work. She used the Assignment tool on Bb where students could upload their first drafts. She downloaded these as Microsoft Word documents, read them, and responded with in-text, marginal, and final comments. She would then "Save As" and upload her responses to each student. After considering Viv's comments and those of their peers, students would revise their essays, then upload the final drafts to a new Bb Assignment link. From there, Viv would access and download them for final evaluation and grading. She sent both sets of the essays to the corpus project archive as she received them from the students.

To compare students' first and final drafts, Viv used the "track changes/compare documents" function in Microsoft Word. This digital tool allowed her to combine a student's first and final drafts into one document in order to easily view every change made from simple word choice to major content revisions. Being able to compare the two drafts in this way helped Viv quickly determine what kind of revisions were made and if they were substantive and meaningful—the exact criteria she understood to be most relevant for the corpus project.

Viv was inspired by what she saw as a significant correlation between the larger, program-wide goals of the corpus project and her goals as a teacher. The corpus project might illuminate the global structure of writing classes at City College, supporting the development and evaluation of writing program pedagogies and curricula. As an active composition pedagogy researcher, Viv was brought directly into that community of practice along

with the other thirty-four paid researchers. The corpus collection process directly influenced her to become more finely attuned to ways of looking at and evaluating students' writing using digital technologies. For basic writers specifically, this kind of attention—a detailed map of the changes that they have made in response to instructor and peer review—demonstrated the seriousness with which instructors read and respond to students' work and also situates that work in the social context of peer and instructor review.

By thoroughly analyzing what they had changed, added, or omitted between drafts, Viv felt that she was ostensibly viewing maps of her students' minds, and by interpreting these maps, she gained crucial information about these students, individually and as a class. She was also able to present her findings across the corpus of their texts in visual form or patterns. As every instructor does, Viv made determinations about learner types, language levels, English language skills, and familiarity with U.S. academic culture and writing standards. These insights helped her to adjust her teaching and communication practices to better meet individual students' needs and those of the class as a whole, thus addressing her part of the collective goal of understanding the techniques and processes that students employ in their writing and improve how writing is taught.

In retrospect, Viv notes that being conscious of her class's contribution to this larger endeavor motivated her in a dialectical combination of competition and cooperation. Competitively, she experienced a drive to ensure that the contributions from her students were equally as significant and meaningful as those from other classes, so she pushed harder for substantive results than she might have otherwise. Simultaneously, she was inspired by the collaborative nature of the project and felt supported as a member of a larger cohort of instructors who were also immersed, many for the first time, in teaching groups of students from a wide range of backgrounds. As an active member of this cohort of researchers, the isolation that Viv might have felt as a new instructor was mitigated. In this way, her experience mirrored the experience of basic writing students, who often feel isolated in unfamiliar and intimidating new environments. Viv was empowered through her membership in a group that was potentially leading toward transformative change. The culmination of her experiences elicited a powerful, visceral sense that she was working in a space that extended far beyond the confines of her own classroom.

Analyzing the Data: Troubling They Say / I Say

The data collection process took place in the fall semester of 2015, and Tom prepared the data for analysis in the summer of 2016. That summer and fall, he conducted preliminary analyses of the data for presentation at the 2017 Conference on College Composition and Communication (Peele, "Cultivating"). He first presented the corpus analysis to students in a graduate composition pedagogy class in spring of 2017, a move he hoped would reduce their isolation—the long hours that new instructors spend alone, grading essays, without very much in the way of an external reference point—and provide them with a sense of belonging to a larger, collective body of instructors who were addressing the same issues.

Many of the students in that class, including Andréa, were new composition instructors. He used a part of the corpus—the argument-based research essays—to show these students how the rhetorical moves that CCNY students make compare to students in other colleges and to professional writers, to introduce graduate student instructors to corpus analysis, and to persuade instructors to focus on the role of digitally-mediated collection and transmission of student writing in a contemporary academic setting. Using the CCNY corpus, Tom initially asked graduate students to conduct a form-function analysis (described below) on a few of the argument essays. To conduct this analysis, he provided the graduate students with a list of sentences that had been drawn from the argument-based research essays. Most of the graduate students resisted the idea of student writing as data. They feared that by looking at a massive collection of essays, we were stripping away each author's individual voice—the very aspect of the essay that gave it value. As they discovered, however, and as Andréa describes in the next section, corpus analysis relies heavily on human interpretation.

To frame in-class activity and discussion, we replicated a part of Lancaster's study examining the rhetorical moves drawn from Gerald Graff and Cathy Birkenstein's text, *They Say/I Say (TSIS)*. Lancaster isolates three rhetorical moves that appear in an argument essay—the introduction of an *objection* to the argument that the student is making (that is, the introduction of a point of view that more or less opposes the argument that the student is making), *concessions* to the objections (the moves that writers make to admit that their opposition might have a point), and *counter* wordings (the moves that writers make to disagree with objections to their arguments). Examples of rhetorical moves for introducing objections include "some readers might object that," or "it could be argued that." Using Anthony's

concordance tool to identify these wordings, we created a concordance of all of the occurrences of the search terms. The tool allows the researcher to click between the concordance and the sentence in the context of the whole essay, so instructors were able to judge how the search terms were being used. The software could create, for example, a list of every sentence in which the word "argued" appeared, and then, by clicking to the full essay, help the reader see if the wordings were functioning as objections to the writer's argument.

This in-class form-function activity directed graduate students to look at a particular rhetorical move to decide whether or not the students had used the form *it could be argued that* as a means of *introducing an objection* or for some other reason, such as emphasizing their own argument. Graduate students were able to see the patterns that emerge in students' texts across the writing program and to experience the difficulties of coding texts. In class, we looked both at the wordings provided by Graff and Birkenstein and alternative wordings for introducing objections, making concessions, and offering counter arguments which Lancaster presents in his study. We found that CCNY students, just like the students and professional, academic writers in Lancaster's study, were much more likely to use nameless rather than named objections (for detailed results of the study, see Peele, "Is"). Our study revealed the same general patterns that Lancaster discovered, and new graduate student instructors in the graduate seminar, and their students, developed a clearer understanding of the rhetorical moves that students might imitate as they worked toward rhetorical dexterity in academic contexts.

The introduction of this analysis of a local corpus of student writing also illuminated for graduate students how we interpret academic essays. We noted that we read specific moves differently, debating whether or not students had in fact, for example, introduced an objection to an argument or instead expanded their original idea. As Andréa describes in the next section, it was harder for us to agree on how the wordings were functioning than we had anticipated. As with Viv's experience, Andréa's participation in a collaborative grading process helped her to feel both that she belonged at the institution, as a member of the community, and that she was receiving specialized training for working with students with a wide range of writing proficiencies. Because she was conducting a very close reading of specific moves in hundreds of essays, she was able to fairly quickly develop a list of the multiple ways that students might struggle with this fundamental yet complex set of rhetorical moves—an experience that she would not have had by reading just one class set of essays during her first semester teaching the course.

Figure 1. AntConc-generated Concordance of "Argue"

As members of this community, we also wondered how much of a difference it made that the linguistic formulas in *TSIS* were not representative of the formulas that writers generally use. Should the aim be to use *TSIS* heuristically, as Don Kraemer has recently argued in the pages of *JBW*? To what extent were instructors teaching their own linguistic preferences under the guise of correct usage, and how did the teaching of formulaic genres mesh with other aims of the composition classroom? We don't offer any answers to these questions here, but the use of our corpus and the close examination of *TSIS* proved generative of thoughtful discussions concerning the aims of composition and how best to reach them. Corpus-driven genre studies in this class introduced the rhetorical-move, genre structure concept that Graff and Birkenstein describe in their book, gave graduate students instructors an opportunity to see how CCNY students conform to patterns of rhetorical practice that are evident in other colleges and in professional, academic writing, and to trouble the concept that many of us develop in isolation: that *our* reading of a student essay is, in some essential way, the only possible reading.

Interpreting Rhetorical Moves Beyond Data Points

Andréa's first semester teaching first-year composition, in the fall of 2016, was also during her final year of graduate school at City College for her MFA in Creative Writing; and in spring of 2017, she also enrolled in our graduate course, Composition Pedagogies and Rhetorical Theories, where Tom introduced the class to corpus analysis. Having previously worked as a qualitative researcher, Andréa did not have the same fears as her graduate peers about the potentially dehumanizing effect of turning student essays into data and of stripping away human voice in that process; Andréa was intrigued by the possibility of capturing students' rhetorical moves through a large scale analysis. During the Composition Pedagogies class, Tom hired Andréa as a research assistant, and Viv, Tom, and Andréa began to prepare for a corpus analysis presentation at the 2017 Computers and Writing conference. Building on the in-class activity described above, they analyzed the corpus to see how CCNY student essays would compare to Lancaster's results.

For their particular data set, Tom and Andréa analyzed 548 argument-based research essays to find the rhetorical moves students were using to *entertain an objection*. To conduct the study, they used the search terms that Lancaster provides in his study to create concordances (451, 453, 455). For example, they created a concordance of the word "argue," then coded the sentences according to how the word was used in the sentence. If the student attributed an objection to her own claim to a recognizable person or group (i.e., "opponents to a ban on tobacco argue" or "professors argue") they assigned a "1" to this use. If the student introduced an objection but did not name a specific person or group (i.e. "others argue" or "some argue") they assigned a "2." They did not code uses of "argue" that did not introduce an objection.

If part of the concern surrounding corpus data analysis is the potentially dehumanizing effect of turning student essays into data points, Tom and Andréa can confidently assert that, after scoring all 548 student argument essays, the coding of rhetorical moves in student essays depends heavily on human interpretation. Before starting their initial round of coding, Tom provided Andréa with background on the process and expectations for scoring the essays using Lancaster's research as the model. Andréa had previously encountered coding but never within the context of linguistic analysis. Andréa stepped into the role of mentee in these preliminary discussions, which gave her space to interrogate the required assignments in tandem with the coding process. For the coding itself, the first person who scored

the essays hid the column of scores in the spreadsheet; the other researcher then scored the student essays while remaining unaware of the first scores.

Once they were both finished, they did the big reveal to see where they agreed and disagreed. After the first round of coding, they were in agreement only about seventy percent of the time. The essays on which they disagreed entertained objections in a variety of ways they did not expect. This brought up several questions about how to understand the student's intent, the importance of placement of the rhetorical move within the context of the essay, and the larger question of how to structure assignments. The following sample from a student's essay shows one kind of discussion that Tom and Andréa had about coding rhetorical moves.

> However, the other party believes that enforcing gun control can be harmful because they believe that this would violate the Second Amendment and in addition, they argue that without their guns in their possession, they feel unsafe and not being able to protect themselves.

On the one hand, since "argue" is directly preceded by "they," it should be coded 2, since a specific entity is not identified in this particular construction. On the other hand, the referent for "they" could easily be read as either "the other party," which counts as an unnamed group, or it could be read as a group who "believes that enforcing gun control can be harmful," which earns it the code of 1. In a similar vein, how far back into the essay should the coders go to find the referent for a pronoun that occurs just before the search term? In some cases, students named an objector only once, early in the essay, pages before the use of the pronoun referent. Similarly, Tom and Andréa noted several instances in which the student used this language not as a means to introduce a concession but rather as a way to support their own argument. They also noted the use of this form in an essay that never actually made an argument (even though one was intended), but instead offered a series of perspectives. Yes, the student used the rhetorical moves, but did the form follow the function?

Working with the CCNY corpus also gave Andréa exposure to a large set of student essays, providing her with insight into what CCNY students are being asked to do in the classroom and how they are composing texts. While working on the corpus analysis, Andréa was teaching a section of Writing for Engineering. She wanted to give her students a macro understanding of why they were being asked to write different texts in specific ways. Andréa quickly

harnessed the engineering students' penchant for data to her advantage by showing them the corpus study as a new entry point for understanding essay writing and genre analysis.

After a meeting with Tom, Andréa used the classroom projector to show the rapt engineers what the corpus of essays looked like as data. Andréa explained the rhetorical moves that were occurring in a few of the sentences that she and Tom disagreed on. By using this kind of example, which pin-pointed the occasional opacity found in student writing, the engineering students were able to clearly see the gap between what a reader confronts and what students sometimes assume is being understood. The presentation was followed by a low-stakes, small group discussion in which the students shared different sentences of their own with each other and compared them to the genre expectation being asked of them. Students were able to see, for example, in their research proposals, that writers typically follow a range of rhetorical moves to introduce objections. The visual presentation allowed students to see how what they were doing fit into the overall pattern of what other students were doing and what the expectations were, and it had the possibility to lead to discussions about the values represented by these rhetorical moves. As Amy Devitt writes, "critical genre awareness. . . can help students maintain a critical stance and their own agency in the face of disciplinary discourses, academic writing, and other realms of literacy" (337). At the same time, she notes that unlike "scholars merely studying genre, students wishing to participate in the academy or discipline or profession cannot simply disengage but must follow that distancing with enlightened participation" (338). Visual corpus analysis, then, opened the space for students both to see, literally, the extent to which they were conforming to genre expectations, and to evaluate, in a preliminary fashion, their own roles in reproducing discursive structures.

As a result of the in-class corpus analysis, Andréa collaborated with the students to revise assignments so that they better reflected students' new understanding of genre and rhetorical moves. To the delight of both Andréa and the engineers, the updated assignments were built from evidence-based writing analysis and felt more grounded in a language that they all could understand. The Lab Report and Technical Report essays were previously modeled after the forms in the professional field, focusing on the macro structure of each genre. The revised assignments focused on both the genre construct of the overall essay as well as points where the students could test rhetorical moves.

CURRICULAR CHANGE IN THE WRITING PROGRAM: HEIGHTENED RHETORICAL GENRE AWARENESS AND TRANSFER

At the end of the corpus collection and analysis, Tom and Andréa distributed an evidence-based, rhetorical moves assignment at a faculty development training. In the assignment, they pointed out, among other things, that CCNY students very often used "some/ many/ one might/ could/ may object/ disagree/ argue that. . ." to introduce an objection, but almost never used "At first glance," "On the surface," "It might seem that," or "It could be argued that." To make a concession, students often used "of course." In the assignment sheet, we suggested that they consider other options, such as "Although I," "While I," "Yes," "It is true," "While it is true," "Clearly," "Obviously," or "Certainly" (Peele, "Is").

This low-stakes writing assignment also asks students to use the library's *Opposing Viewpoints in Context* database to find brief articles that offer opposite viewpoints on one issue and to identify the rhetorical moves that the authors make. The assignment mirrors *TSIS*'s template technique and, we hope, helps students make connections between their use of genre conventions in their own writing and how those same moves are used by professional writers. Using the corpus as a frame of reference, and comparing rhetorical moves across corpora, marked a shift in how our program discusses student writing. The corpus analysis and faculty development around it has begun to move us away from the assignment sequence that had been in place—a literacy narrative, an expository essay, an exploratory essay, and a research essay—toward a curriculum that asks students to study genre explicitly in order to support their transfer of writing knowledge from composition to other classes. The new curricular model is based both on corpus analysis, which is now a part of teaching practicum for new, graduate student instructors of composition, and on the study of transfer presented by Kathleen Yancey, Liane Robertson, and Kara Taczak.

WRITING AND LEARNING TO TEACH IN SOCIAL CONTEXTS

As we have noted, some strands of basic writing and rhetorical genre studies in composition aim to help students situate their own literacy practices within broader social contexts. The more students are able to see how their literate lives fit within a continuum of literate practices, the more equipped they will be to draw on those practices as they begin the process of learning how to write academic essays in college. As Carter notes in her

description of the basic writing program that she developed at Texas A&M Commerce,

> [R]hetorical dexterity attempts to develop in writers the ability to negotiate the school literacies celebrated in the current social order in ways that are as ethical and meta-aware as possible. We begin this process by articulating the ways in which what they already know well may help them learn what is, as of yet, less familiar to them. (*The Way Literacy Lives* 18)

Building connections between what is known and what is new helps students value their literacy practices and also creates a foundation for thinking of those practices as situated within social contexts.

Our corpus study situates writing within the local context of City College first-year students. By mapping the rhetorical moves that students make in their argument-based essays, we were able to help students compare their own and their peers' rhetorical moves to the moves made by professional writers. Some scholars might argue that corpus studies do not, as Lynne Flowerdew describes these critiques in her study of the variation of cultural expression within academic genres, "consider the socio-cultural context as they deal with decontextualized corpus data" (321). Flowerdew goes on to say, however, that as "genre analysts are keen to emphasize, 'move structures' should not be seen as a rigid set of labels for coding text but instead should accept variations of the prototypical move structure patterning for a genre" (326).

Far from studying genre as a set of rigid guidelines, Anis Bawarshi and Mary Jo Reiff note in their resource guide to genre studies that by "arguing for genre as a centerpiece of literacy teaching, . . . genre scholars have debated the ways genre can be used to help students gain access to and select more effectively from the system of choices available to language users for the realization of meaning in specific contexts" (37). Their aims, in other words, are to make clear the available choices and to demystify the conventions of the genre, goals that genre theorists share with many basic writing instructors. In our study of rhetorical moves, our aim, as with portfolio assessment, was to take a closer, program-wide look at what students were doing, and to fold that knowledge back into the curriculum, faculty development, and pedagogy.

Multiple genre studies theorists have made the argument for the explicit teaching of genre as a means of demystifying the expectations for second language learners; the same argument applies to basic writing stu-

dents and to their instructors. As Ken Hyland notes in his explanation of genre pedagogy, "genre is a socially informed theory of language offering an authoritative pedagogy grounded in research on texts and contexts, strongly committed to empowering students to participate effectively in target situations" ("Genre-Based" 27). The study of rhetorical moves in corpora situates literacy within a social practice. In classroom corpus study activities, both undergraduate and graduate students look at how their peers, or their students, make use of, in our case, the rhetorical moves that writers use to introduce objections, to make concessions, and to offer counter arguments. In comparing these practices to the moves made in peer-reviewed publications, students and instructors situate their own literate practices within a broader social setting. While the most immediate social setting is academic, variations of the claim/ objection/ concession/ counter argument structure that we look at in this study are recognizable across a wide range of public-facing genres, including long-form journalism, op-eds, advocacy articles in print and online magazines, and arguments for and against institutional policy changes in a wide variety of organizations including universities, schools, and other non-profits, among other contexts.

Our aim was to make students aware of the rhetorical moves associated with conventional academic genres so that they are more familiar with the genre conventions of academic writing and to make explicit connections between the genres we study in the classroom and the genres that exist, in Mary Soliday's words, "in the wild." Teaching genre is a way of making the conventions explicit for basic writers and other students of composition, but this needs to be coupled with an awareness that genres shift over time and are responsive to social situations. There are, certainly, values that these genres express, and discussions of these values fit well within a broader focus on the shifting and contingent nature of genres. Basic writers, who are likely to be less familiar with conventional rhetorical moves than other students, might, like English language learners, leave our classes with a much better sense of academic genre expectations as well as an awareness of the ways in which genres reproduce social relations. Explicit instruction of genre conventions is similar to providing students with model essays or with grading rubrics, both of which are intended to make the instructor's expectations as transparent as possible. Overt instruction in genre expectations—asking students to find examples of a particular genre, making connections with them between various genre types, and looking at the same message written for different audiences—extends this transparency so that students are in

a better position to assess, evaluate, imitate, and reject genre conventions and the values they express.

In addition, corpus analysis for rhetorical moves offers new instructors of composition at every level methods for discovering, studying, and making use of the rhetorical moves that their students will need as they pursue their educations. Much current scholarship highlights the value of genre studies and demonstrates some of the conventions of those genres (Adler-Kassner; Hart-Davidson). For new graduate student instructors, it introduces them to a community of practice, which helps them learn more quickly about the ways in which students struggle with a particular form. And, as Lancaster shows, it offers a way to illuminate the values that are embedded in the genre—in this case, an emphasis on the importance of interpersonal relations in the academic argument, an emphasis that is largely overlooked in our field's discussion of argument. The corpus collection encouraged instructors to be more digitally active and to make connections, for students, between digitally driven, non-academic writing and academic writing. The explicit study of genre helps demystify the rhetorical moves that students will need to make in academic writing, but it doesn't do so in a socio-cultural vacuum. Instead, if they build their own corpus, corpus analysis offers instructors and students an opportunity to examine their essays in a highly local context.

Works Cited

Adler-Kassner, Linda, Irene Clark, Liane Robertson, Kara Taczak, and Kathleen Blake Yancey. "Assembling Knowledge: The Role of Threshold Concepts in Facilitating Transfer." *Critical Transitions: Writing and the Question of Transfer*, edited by Chris M. Anson and Jessie L. Moore, UP of Colorado, 2017, pp. 17-48.

Anthony, Laurence. *Laurence Anthony's AntConc.* www.laurenceanthony.net/software/antconc/. Accessed 19 Feb. 2018.

Aull, Laura. *First-Year University Writing: A Corpus-Based Study with Implications for Pedagogy.* Palgrave, 2015.

Bartholomae, David. "Inventing the University." *Journal of Basic Writing,* vol. 5, no. 1, 1986, pp. 4-23.

Bawarshi, Anis S., and Mary Jo Reiff. *Genre: An Introduction to History, Theory, Research, and Pedagogy.* Parlor, 2010.

Bazerman, Charles. "Genre and Cognitive Development: Beyond Writing to Learn." *Genre in a Changing World,* edited by Charles Bazerman, Adair Bonini, and Débora Figueiredo. Parlor, 2009, pp. 279-94.

Bizzell, Patricia. *Academic Discourse and Critical Consciousness.* U of Pittsburgh P, 1992.

Buell, Duncan, and Christopher Holcomb. "First-Year Composition as Big-Data: Natural Language Processing and Portfolio Assessment." The Conference on College Composition and Communication, 19 Mar. 2015, Tampa, FL. Conference Presentation.

Carter, Shannon. "Redefining Literacy as a Social Practice." *Journal of Basic Writing*, vol. 25, no. 2, 2006, pp. 94-125.

---. *The Way Literacy Lives: Rhetorical Dexterity and Basic Writing Instruction.* State U of New York P, 2008.

"City Facts." *The City College of New York.* 7 Oct. 2017, www.ccny.cuny.edu. Accessed 15 May 2018.

Delpit, Lisa P. "The Silenced Dialogue: Power and Pedagogy in Educating Other People's Children." *Harvard Educational Review,* vol. 58, no. 3, 1988, pp. 280-98.

Devitt, Amy. "Teaching Critical Genre Awareness." *Genre in a Changing World*, edited by Charles Bazerman, Adair Bonini, and Débora Figueiredo, Parlor, 2009, pp. 337-51.

"Economic Diversity and Student Outcomes at City College of New York." *The Upshot, The New York Times*, www.nytimes.com/interactive/projects/college-mobility/city-college-of-new-york. Accessed 11 Oct. 2017.

Elbow, Peter. *Writing Without Teachers.* Oxford UP, 1973.

Faigley, Lester, and Stephen Witte. "Analyzing Revision." *College Composition and Communication*, vol. 32, no. 4, 1981, pp. 400-14.

Flowerdew, Lynne. "An Integration of Corpus-Based and Genre-Based Approaches to Text Analysis in EAP/ESP: Countering Criticisms against Corpus-Based Methodologies." *English for Specific Purposes,* vol. 24, 2005, pp. 321-32.

Graff, Gerald, and Cathy Birkenstein. *They Say / I Say: The Moves that Matter in Academic Writing.* 3rd ed., Norton, 2014.

Hart-Davidson, Bill. "Genres are Enacted by Writers and Readers." *Naming What We Know: Threshold Concepts of Writing Studies*, edited by Linda Adler-Kassner and Elizabeth Wardle, Utah State UP, 2016, pp. 39-40.

Hyland, Ken. "Genre-Based Pedagogies: A Social Reponse to Process." *Journal of Second Language Writing,* vol.12, 2003, pp. 17-29.

Johnson-Eilola, Johndan. "The Database and the Essay: Understanding Composition as Articulation." *Writing New Media: Theory and Application for Expanding the Teaching of Composition,* edited by Anne Wysocki,

Johndan Johnson-Eilola, Cynthia L. Selfe, and Geoffrey Sirc, Utah State UP, 2004, pp. 199-236.

Kraemer, Don. "Fact and Theory and Value Judgment: What *They Say/We Say* of Basic Writing's Unhedged Good." *Journal of Basic Writing,* vol. 36, no. 1, 2017, pp. 5-24.

Lancaster, Zak. "Do Academics Really Write this Way? A Corpus Investigation of Moves and Templates in *They Say/I Say.*" *College Composition and Communication*, vol. 67, no. 3, 2016, pp. 437-64.

Miller, Carolyn. "Genre as Social Action." *Quarterly Journal of Speech,* vol. 70, 1984, pp. 151-67.

Peele, Thomas. "Blackboard CMS for Large Scale Data Collection." *Kairos: Rhetoric, Technology, and Pedagogy*, vol. 20, no. 2, 2016, praxis.technorhetoric.net/tiki-index.php?page=PraxisWiki%3A_%3ABlackboard+for+Data+Collection. Accessed 20 Feb. 2018.

---. "Cultivating Faculty Participation in Corpus-Driven Studies." The Conference on College Composition and Communication, March 15-18, 2017, Portland, OR. Conference Presentation.

---. "'Is This Too Polite?': The Limited Use of Rhetorical Moves in a First-Year Corpus." *Journal of Writing Analytics*, vol. 2, 2018. pp. 78-95.

Peele, Thomas, and Melissa Antinori. "iBooks Portfolios: Interface, Audiences, and the Making of Online Identities." *Journal of Basic Writing,* vol. 32, no. 2, 2013, pp. 28-50.

Reiff, Mary Jo, and Anis Bawarshi. "Teaching Discursive Resources: How Students Use Prior Genre Knowledge to Negotiate New Writing Contexts in First-Year Composition." *Written Communication,* vol. 28, no. 3, 2011, pp. 312-37.

Rose, Mike. *Lives on the Boundary: A Moving Account of the Struggles and Achievements of America's Educationally Underprepared.* Penguin, 2005.

Schnee, Emily, and Jamil Shakoor. "Self/Portrait of a Basic Writer: Broadening the Scope of Research on College Remediation." *Journal of Basic Writing,* vol. 35, no. 1, 2016, pp. 85-114.

Shaugnessy, Mina P. *Errors and Expectations: A Guide for the Teacher of Basic Writing.* Oxford UP, 1979.

Soliday, Mary. *Everyday Genres: Writing Assignments across the Disciplines.* Southern Illinois UP, 2011.

Swales, John. *Genre Analysis: English in Academic and Research Settings.* Cambridge UP, 1990.

Yancey, Kathleen Blake, Liane Robertson, and Kara Taczak. *Writing Across Contexts*. Utah State UP, 2015.