A MIXED-METHODS APPROACH TO ANALYZING WRITING CENTER SESSION NOTES

Zoë DeKruif & Jamie Smith | Iowa State University

Linguistic corpus analysis is often an overlooked research method in writing center studies. This methodology has the potential to reveal countless patterns in datasets, but frequently lacks important details. Pairing corpus analysis with inductive coding—a qualitative approach—provides a comprehensive view of both overarching themes and specific information. This paper utilized this mixed-methods approach to explore the types of feedback that writing consultants provide to students during sessions at Iowa State University's writing center. Session notes, written by a consultant during a writing session, contain an abundance of information surrounding the inner workings of writing centers, but few studies have recognized them as viable data sources. For the quantitative analysis, this study utilized AntConc to derive frequencies of commonly occurring words and n-grams in session notes. The qualitative analysis consisted of a process of inductively coding the data to identify commonly occurring themes and define them based on their linguistic realizations. By creating an initial coding guide, completing several rounds of session note annotations, and adjusting the guide as needed, inductive coding provided a level of context and detail that was instrumental in understanding the characteristics of writing center session notes.

INTRODUCTION

Corpus analysis is not often used in the realm of writing center session note analysis. However, it is an advantageous way to access a hidden wealth of information. Using corpus analysis by itself was intriguing, but we wanted to consider a more in-depth approach to make clearer connections. To achieve this goal, we wanted to pair corpus analysis with inductive coding, the detailed process of creating a coding guide and using it to annotate data. Together, these two techniques can provide an extensive look at many types of data.

We had the opportunity to put this methodology in motion through a research project at Iowa State University. The Iowa State University Honors Program provides a unique opportunity to first-year undergraduate students through the First-Year Honors Mentor Program (FHMP). We were matched to a faculty mentor, a graduate student mentor, and their corresponding project, "ProWrite: Biometric technology for improving college students' writing process." ProWrite is a research project sponsored by the National Science Foundation with the goal to improve undergraduate students' writing. ProWrite plans to provide individualized and accurate process-based feedback to writers by utilizing unobtrusive, deployable, biometric technology

(including coordinated keystroke logging and eye-tracking) to capture the moment-by-moment details of students' writing processes.

Within ProWrite, we have our own subproject in which we have utilized a mixed-methods approach to complete a needs analysis of the campus Writing and Media Center (WMC). The purpose of our project is to provide ProWrite with an accurate understanding of the WMC's operations: How does the WMC assist students with their writing, and what type of help are they providing? Specifically, we were interested in investigating the types of process-focused feedback and product-focused feedback provided by consultants at the WMC.

We utilized a variety of resources in order to answer this question and achieve our purpose of analyzing the WMC. Iowa State's WMC is a free resource for students looking for aid with any type of media, the most popular being technical papers from undergraduate classes ("Fall 2020 Review"). Students have the option to schedule a session with consultants, who are trained undergraduate and graduate students who are proficient in writing. Iowa State's Writing and Media Center helps students become stronger, more confident communicators. The WMC aims to inspire students to develop in their writing by promoting the values of creativity, critical thinking, and lifelong learning. This involves helping with all forms of communication (e.g., essays and multimedia presentations), and assisting with all stages of the composition process (e.g., brainstorming or revising) ("Writing and Media Center: Iowa State University").

After the one-hour session is complete, the consultant writes a session note that contains a description of what the consultant and student did within the session. The following is an example of a session note from the center:

> [Student Name] had an online appointment to work on his psychology paper for a class he is taking at Kirkwood. He wanted to focus on flow and transitions. We read through his paper out loud and made several phrasing changes to long and wordy sentences. We also worked on the conclusion paragraph. [Student Name] plans to make format changes to match the APA style guide before turning it in. ("Writing and Media Center: Iowa State University")

After the director of the WMC granted us access to over 6,000 session notes from the last two years, we used the previously mentioned mixed-methods approach to analyze the data. First, we used AntConc (Version 3.5.8) computer software, a widely used corpus analysis application, to quantitatively analyze the entire collection of data. This covered the corpus analysis portion of our approach. To gain a deeper analysis, we also qualitatively analyzed portions of the data to develop a coding guide that could enable us to organize session notes by the type of writing advice the session focused on. These two techniques, corpus analysis and inductive coding, gave us the perfect way to explore our session note data.

WMC SESSION NOTES

Session notes are instrumental in understanding the WMC's role in the student writing process. One resource that analyzes session notes is the article "It's All in the Notes: What Session Notes Can Tell Us About the Work of Writing Centers" by Genie N. Giaimo,

Joseph J. Cheatle, Candace K. Hastings, and Christine Modey. This paper argues that session notes contain a plethora of information and insight about the inner workings of writing centers, but researchers have not had an effective way to analyze such a large number of notes. To demonstrate a methodology for exploring session notes, they chose to use Voyant Tools, a freely available online corpus tool with a variety of functions. By primarily using word frequency lists, Giaimo and colleagues aim to demonstrate that session notes contain a deep well of information regarding the processes used within writing centers as well as how writing advice is focused and delivered. They also make claims about what specifically occurs in sessions and any areas of concern there may be. For example, "grammar" was a high-frequency word, suggesting it was heavily discussed and worked on in sessions. Giaimo and colleagues noted this as a concern from a writing center leadership standpoint since their primary goal for writing and media centers is to promote working on higher-order concerns, rather than lower-order concerns.

Overall, this methodology provides the basis of our session note analysis, but with a few changes. Along with a brief quantitative word analysis modeled in this article, we also wanted to analyze the data qualitatively. This qualitative view would provide us with more detailed information than information provided by a word list alone could, and it would help us make more detailed claims backed up by strong, contextual evidence. After selecting our desired methodology, we began collecting our session notes to do some preliminary analysis.

The WMC has access to data about these session notes that has been previously collected. For example, consultants fill out a checklist containing categories to accompany their session notes in order to mark what general areas the consultants and students worked on in the session. However, some of this provided information was vague and led to confusion. Roughly 4,000 of the 6,000 notes collected for this study were marked with the term "revising." "Revising" was too broad of a category to analyze these notes by, since it was clear the term was used loosely in two-thirds of the data; instead, we wanted to know what and how exactly the consultants and students were revising. This lack of specificity is seen in several other categories such as "polishing," "organizing ideas," and "drafting," all of which do not give detailed examples of what is being worked on. Without this information, it is difficult to come to reliable conclusions, so we decided to investigate the session notes in a more detailed manner without relying on this vague information.

To effectively draw conclusions from session notes, we obtained a large sum of them from the WMC. The 6,425 session notes used in our analysis were collected from every WMC consultation occurring between August 1, 2019, and February 17, 2021. We used several methods to analyze these notes, including a corpus tool named AntConc and the detailed process of inductive coding.

ANTCONC: QUANTITATIVE DATA

As discussed previously, corpus analysis has incredible research potential. In their paper about corpus linguistic analysis benefits, Jo Mackiewicz and Isabelle Thompson say, "Through corpus linguistic analysis, writing center researchers can begin to ask and answer

new research questions." The general purpose of corpus linguistic analysis is to discover patterns in language (Krieger). Many types of patterns can be revealed through this process: frequently used words, how different n-grams are used in context, etc. Once this information is clear and visible to the researcher, it is considerably easier to discern different theories as to why the data reflects these patterns. Our research in session note analysis aimed to do exactly that.

Our first method to analyze these session notes was using a corpus tool, AntConc. AntConc is described as "a freeware software programme for working with language corpora using a graphical user interface. Within AntConc are several 'tools' that support linguistic analysis by enabling the user to—for example—search corpora, to generate lists of words in corpora, and to browse 'concordances' of word use in corpora" ("Introduction to AntConc"). Giaimo et al. used a similar corpus tool named Voyant. Voyant has several tools that are similar to the ones provided by AntConc. We first attempted to use Voyant for our quantitative analysis, but our corpus was incompatible with the software due to the large number of text files we collected. Instead, we turned to AntConc, which was able to efficiently handle our gathered texts. Of the mentioned toolkit, we favored a few of the tools for our brief analysis. Similarly to Giamo and colleagues, we acquired a list of the corpus terms by frequency. This was a brief overview of the high-frequency words that provided us with preliminary ideas to think about. Though these word lists are interesting, they lack an important feature that we are looking for: context. To obtain context, we decided to utilize the n-grams tool—a tool that allowed

us to search for phrases that have commonalities through specific words—to identify some frequently used n-grams that would give a comprehensive look at what was being done in the WMC. Upon finding some commonly used n-grams, we found that AntConc's Concordance Tool was incredibly useful for seeing the context of these phrases. This tool allowed us to type in any given n-gram and see the words that appeared on either side of the inserted phrase. For instance, we searched for the commonly used four-gram "came to the WMC" and were able to see all the complete sentences that it appeared in, like "[Student name] *came to the WMC* to develop a personal statement" ("Writing and Media Center: Iowa State University"). This allowed us to elicit quantitative data about the entirety of our data while also determining how the phrase was used in context to come to the most informed conclusions. AntConc was very useful in our quantitative analysis by providing us with the common topics and themes found in writing center sessions, but it still left us with unanswered questions regarding the specific content that students and consultants work on. We needed information about students' writing concerns and what measures the consultant took to help the student. It was clear that an even more detailed look at the session notes was necessary to accomplish this, so we moved into a qualitative analysis of the data, also known as inductive coding.

INDUCTIVE CODING: QUALITATIVE DATA

To further analyze the WMC session notes, we decided to utilize inductive coding, a way of analyzing data qualitatively that compliments

the quantitative data from AntConc. Due to our preliminary work with AntConc, we already had several established themes found in the session notes. Our next step was to qualitatively narrow these themes into detailed categories for the strongest analysis.

When developing our qualitative analysis and inductive coding phases, we followed a methodology common to genre analysis that would allow us to identify the specific themes and information in session notes that were not found by corpus analysis (Cotos). Following this methodology, our qualitative analysis unfolded into four phases: (I) qualitative analysis of a small set of session notes to develop a tentative coding guide, (II) multiple rounds of pilot coding to further inform and refine the development of the coding guide and protocol, (III) consultation with an expert of the target discourse community ("Writing and Media Center: Iowa State University") to finalize the framework and protocol, and (IV) annotation for coding categories of a large set of the data.

In order to develop a tentative coding guide for Phase I, three annotators (both authors and the graduate student mentor) randomly selected a set of 30 session notes and individually developed themes for each session note. The three of us then collaboratively discussed the commonalities between our themes, narrowed them down into specific categories, and developed a tentative coding guide from them. In total, our coding guide has three levels of categories: primary categories, sub-categories, and micro-categories. We also used a secondary code to determine if the consultant physically worked on the paper with the student or if they just gave out suggestions on what to do at a later time.

Through AntConc, we found many vague themes that gave us an idea of how we would like to specifically categorize session notes through inductive coding. For example, the data reflected indistinct ideas like "revising" and "polishing" as common areas of struggle for students, but we didn't know how this assistance was specifically provided by the writing consultants (e.g., brainstorming or outlining). These themes allowed us to create a coding guide that encompassed these general ideas but broke them down into smaller categories for a more detailed analysis.

Phase II then consisted of four rounds of pilot coding. In the first round of coding, the two of us and our graduate student mentor individually coded the same randomly selected 30 session notes utilized in Phase I. We then went through any disagreements we had, refining our coding guide as we went in order to provide further specification. Our next step was to analyze our coding guide by measuring the reliability of our annotators, also known as how consistently they coded session notes individually in a content analysis approach. We did this to ensure that our annotators can consistently code session notes, thus making our coding guide reliable. We measured our reliability with Krippendorff's alpha. This is a formula used to determine interrater reliability that measures disagreement rather than agreement (Krippendorff). With this process in mind, we completed three additional rounds of coding with a new set of 30–50 session notes each round. After each round, we calculated the reliability using an R script (i.e., an algorithm for statistical computing in the R programming language) and discussed the reasons we did not reach our desired reliability, further specifying our coding guide

via micro-categories to increase our reliability, allowing us, as the annotators, less room for conflict. We reached an acceptable reliability rate of 78% in the fourth round of coding, allowing us to move on to the next phase.

In Phase III, although our graduate student mentor was already considered an expert member of our target discourse community as a former WMC consultant and Graduate Assistant Director, we chose to have a meeting with another WMC consultant to discuss our coding guide and how to make it best represent the functions of the WMC. The consultant we met with thought our coding guide was set up very well and covered what the WMC does effectively, but then pointed out one problematic code, the "Audience" code. She noted that discussing the point of view and general audience would likely be done in every session but may not be recorded in every session note. This violated the assumption of independence for this code, meaning that the code would not be able to stand alone, nor would it be accurately reported. Because of these factors, the "Audience" code was subsequently removed from our coding guide.

In Phase IV, our success in the previous phases allowed us to begin the next process of coding 300 session notes. Each person's corpus of session notes overlapped each other so that reliability could still be regularly checked while still accumulating a large set of data.

Implementing inductive coding into our analysis of WMC session notes was incredibly valuable to our overall findings because it provided us with qualitative data. Unlike our quantitative data from AntConc, inductive coding revealed new, detailed information regarding the content of the session notes that would not have been illuminated if we had only focused on AntConc. Additionally, this qualitative approach clarified many of the broad themes from the corpus analysis that were originally too vague, improving the specificity in our overall findings.

CONCLUSION

Combining AntConc with inductive coding provided us with two diverse methods to analyze our data. This was an important step that allowed us to explore other means of data analysis and discover parts of the project that would not have been found otherwise. Specifically for ProWrite, since we were still missing important information, we decided that combining qualitative and quantitative analysis would provide us with a well-rounded methodology that was firm enough to build conclusions upon.

Session notes provide a large repository of information regarding the processes used and the discussed content in writing and media centers, but it is a difficult task to analyze this data. Quantitative measures are reasonably effective at determining the broad topics found in session notes, such as lower-order versus higher-order concerns and large categories like "Revising" or "Polishing." By using AntConc, we were able to investigate these categories with more context and gain a preliminary understanding of the session note content. However, this quantitative view was lacking in detail, and we were unable to make detailed conclusions based on this information. We then turned to inductive coding, where we created and implemented a coding guide. This coding guide allowed us to annotate the session notes in greater detail and see the specific comments made by consultants. After doing several rounds of reliability checks, we annotated the remaining session notes with our guide

and documented our findings. Overall, both corpus analysis and inductive coding proved to be useful independently, but together they made for a comprehensive review of our data. Our mixed-methods approach has allowed us to highlight the importance of WMC session notes and provided an exciting pathway to continue expanding upon our findings.

ACKNOWLEDGMENTS

WORKS CITED

Anthony, Laurence. "AntConc Homepage." *Laurence Anthony's AntConc*, www.laurenceanthony.net/software/antconc/. Accessed 1 April 2021.

Cotos, Elena. "Move Analysis." *The Encyclopedia of Applied Linguistics*, 7 June 2018, pp. 1–8, https://doi.org/10.1002/9781405198431.wbeal1485.

"Fall Review 2020." *Iowa State University Writing and Media Center,* 13 April 2021, https://www.wmc.dso.iastate.edu/sites/default/files/uploads/WMC%20Fall%202020%20in%20Review.pdf.

Giaimo, Genie M., et al. "It's All in the Notes: What Session Notes Can Tell Us About the Work of Writing Centers." *Journal of Writing Analytics*, vol. 2, 2018, pp. 225-256, https://wac.colostate.edu/docs/jwa/vol2/giaimo.pdf.

"Introduction to AntConc." *Introduction to AntConc – Computational Analysis of Catalogue Data*, https://cataloguelegacies.github.io/antconc.github.io/01-introduction/index.html. Accessed 1 April 2021.

Krieger, Daniel. "Corpus Linguistics: What It Is and How It Can Be Applied to Teaching." *The Internet TESL Journal*, vol. IX, no. 3, 2003, http://iteslj.org/Articles/Krieger-Corpus.html.

Krippendorff, Klaus. "Computing Krippendorff's alpha-reliability." *Departmental Papers (ASC)*, 2011, https://repository.upenn.edu/asc_papers/43.

Mackiewicz, Jo, and Isabelle Thompson. "Adding Quantitative Corpus- Driven Analysis to Qualitative Discourse Analysis: Determining the Aboutness of Writing Center Talk." *The Writing Center Journal*, vol. 35, no. 3, 2016, pp. 187–225.

"Writing and Media Center: Iowa State University." *Writing and Media Center | Iowa State University*, www.wmc.dso.iastate.edu/.