Seniority in Writing Studies: A Corpus Analysis

William Marcellino, *The RAND Corporation*

Structured Abstract

- **Technique Identification**: This new quantitative technique study introduces RAND-Lex and its unique contribution to the analysis of an edited collection on seniority in writing studies.
- **Value Contribution**: Demonstrated is the value of using one integrated platform for analyzing lexical features, determining stance/style, detecting themes through topic modeling, and conducting machine learning through exploratory clustering and classification.
- **Technique Application**: Presented is a RAND-Lex analysis of an edited collection by senior scholars that allows sub-group analysis of four fields of study: General Writing Studies; Creative Writing; Professional, Technical, Business, and Scientific Writing; and Rhetoric and Composition.
- **Results**: Presented are findings from a variety of writing analytic methods: lexical analysis, lexicogrammatical analysis, and auto- clustering. The study identifies unique language use patterns in each of the four fields that, in turn, suggests how writing studies currently conceptualizes the emerging field of seniority studies.
- **Directions for Further Research**: The study concludes with emphasis on topic-specific corpora and additional pedagogical applications.

Keywords: corpus analysis, lexical analysis, machine learning, RAND-Lex, stance/sentiment analysis, topic modeling, writing analytics

DOI: 10.37514/JWA-J.2019.3.1.09

1.0 Technique Identification

In this new quantitative technique study, I present an illustrative use example for RAND-Lex, a text analytics and machine learning tool suite developed at the RAND Corporation. At the present writing, the suite has been used in a wide variety of public policy analyses (Collins, Eberhart, Marcellino, Davis, & Roth, 2018; Helmus et al., 2018; Marcellino & Magnusson, 2019; Marcellino, Smith, Paul, & Skrabala, 2017). It has yet to be used on a corpus from the field of writing studies, and so a new application study is presented in this study.

RAND-Lex is a holistic approach to text as data, borrowing from multiple disciplinary traditions including corpus linguistics (McEnery & Wilson, 2001), rhetoric and functional linguistics (Halliday, 1994; Richards, 1991), and computer sciences (Blei, Ng, & Jordan, 2003; Breiman, 2001; Jain, 2010). By bringing a range of text analysis and machine learning methods into a single, user-friendly webservice, RAND-Lex designers seek to democratize these methods and make them broadly available within the humanities and social sciences. Our goal as designers is to make rigorous, scalable analytics available to researchers and teachers regardless of their background in statistics and computer science. Our focus on holism reflects a model of language in which different language features are mutually entailed, from the finest grains of lexical items, to a lexicogrammar, to the human recognizable topics (Allison, Heuser, Jockers, Moretti, & Witmore, 2011). This is, of course, only one of many possible models, and don't argue that there is a hierarchy of worth or complexity between different kinds of linguistic features. Rather, there may be value in thinking about distinct feature sets that can be described or manipulated with different analytical tests and algorithms. In seeking a holistic rather than a behavioral approach, RAND-Lex shares commonalities with situated (Gee, 2004) and translingual (Horner & Tetreault, 2017) language perspectives.

While my primary purpose in this new quantitative technique study is to highlight RAND-Lex as an example of how more holistic and accessible writing software tools may benefit writing analytics as a field, the particular illustrative case study on seniority in writing will be of interest to readers in writing studies, seniority studies, and writing analytics . Being able to analyze the lexical choices of senior writers in different writing disciplines, or compare their stylistic moves, helps us better understand those fields from a particular kind of reflective posture informed by their most senior voices.

To better situate RAND-Lex's value as a holistic approach to analyzing language, I begin with a review of illustrative examples of text analysis software, borrowing the above tripartite model of linguistic feature sets.

1.1 Lexical Analysis

Text analytics is an active area of writing practice, with applications for writing instruction, writing assessment, writing analytics, and digital humanities research (see Lang, Aull, & Marcellino, this issue). Analyzing writers' lexical choices is important because it signals what texts are primarily *about*: if words like "*rose*," "*hoe*," "*soil*," and "*fertilizer*" are over-represented in a corpus, it is not hard to guess that the text collection is about gardening. Lexical analysis

tools that draw from corpus linguistics have been most broadly available for both research and writing instruction. Wordsmith Tools (Scott, 2008) and AntConc (Anthony, 2004) are among the more longstanding software tools publicly available for a range of statistical tests on word frequencies and associations (1996 and 2005, respectively). Newer platforms have brought increasingly sophisticated lexical toolsets to bear, such as Tool for the Automatic Analysis of Lexical Sophistication (TAALES) (Crossley & Kyle, 2018). TAALES goes beyond frequency and association testing to also assess lexical sophistication through a variety of measures. Where earlier software was broadly descriptive of lexical features, newer software like TAALES brings a narrower but more sophisticated focus specifically on writing assessment.

1.2 Lexicogrammatical Analysis

Analyzing lexicogrammatical choices (essentially stylistics) is a way to make visible a writer's stance in text. David Kaufer and Suguru Ishizaki of Carnegie Mellon University have pioneered the use of software to detect and analyze stance and style in text corpora through their DocuScope software (Kaufer, Geisler, Vlachos, & Ishizaki, 2006). Where lexical analysis is best suited to *content extraction* (what is the text about?), lexicogrammatical analysis is suited for *function extraction* (what are the pragmatic moves in the text—the writer's stance inscribed in the text?). Where lexical analysis works through word counts and word distances, lexicogrammatical analysis of a writer's stance works through taxonomic counts: statistical testing on the frequencies and distributions of different categories of language: certainty markers, hedging language, types of affect, social relationships, values, and so on. In a recent application, DocuScope Classroom (Helberg et al., 2018) provides a writing instruction variant of the software meant to give students visibility over their stylistic choices, and thus enhance their genre performances. Lexicogrammatical analysis software is relatively rare, and the nearest comparison to DocuScope is Linguistic Inquiry and Word Count (LIWC) software (Pennebaker, Francis, & Booth, 2001). LIWC also uses a linguistic taxonomy, but one drawn from psycholinguistic theory. Whereas software like DocuScope is text focused, the creators of LWIC purport to be able to make inferences about the intentions, personality traits, and interior cognition of authors.

1.3 Human Recognizable Topics

Latent Dirichlet Allocation (LDA) topic modeling has had a broad impact in digital humanities, by discovering latent thematic features in large text collections. This modeling technique has been dominated by MALLET, an LDA topic modeling tool from UMASS Amherst (Meeks & Weingart, 2012). Topic modeling is in some ways the most human interpretable of the levels of text analysis described here; that is, topic modeling yields lists of words (topics) that appear in the same context, and thus are topically related and further interpreted by human readers. A researcher interested in French crime fiction might be unsurprised to find that a topic model built on a corpus of genre examples includes themes such as private investigators or guns, but might not have already been aware of topics such as twilight or bourgeoise (Schöch, 2015). Topic

modeling is valuable because it helps empirically define the content and boundaries of types of writing and is interpretable because it is an example of distant machine reading matching close human reading.

1.4 Supervised Classification and Unsupervised Clustering

In addition to descriptive analytics around the three feature sets of language discussed above, both supervised classification and unsupervised clustering machine learning methods have been increasingly brought to bear on writing analytics, prominently in writing assessment. Rather than describing text collections, these algorithms put texts (e.g., whole papers, sections of papers, comments/feedback) into piles based on similarity. Supervised machine learning involves learning from a labeled dataset provided by human coders or graders (hence "supervised"): the algorithm learns the statistical relationship between features and the label and can then label or "classify" new examples. Supervised learning methods that depend on training data have figured prominently in automated writing evaluation (AWE) applications that require scale (Burstein, Tetreault, & Mandini, 2013; Dikli, 2006). Unsupervised clustering approaches do not require humans to provide training data to learn from—instead the algorithm "clusters" documents based on similarity to each other and dissimilarity to the rest of the data set. Unsupervised clustering has also been applied to AWE (Chen, Liu, Lee, & Chang, 2010), and has the potential to inform program assessment and institutional level data-mining (Bowers, 2007). It is important to note that classification and clustering can learn using lexical, lexicogrammatical, and thematic feature sets. Among the newest innovations of AWE is an emphasis on formative assessment (Burstein et al., 2018).

2.0 Value Contribution

2.1 Holistic Language Modeling

Given a holistic, multi-feature model of language, a broadly useful and robust analytic platform would require multiple approaches across a range of disciplinary sources. My colleagues and I built RAND-Lex originally for public policy research across a range of areas, primarily in public health and national security, by investigating a variety of disciplines to borrow existing text analysis and machine learning methods: lexical analysis from corpus linguistics, stance analysis from digital rhetorics and functional linguistics, topic modeling from computer sciences, and text classification/clustering from machine learning.[1] RAND-Lex's unique value is that all of those approaches are in one toolset that is designed for usability and the democratization of these methods beyond technical experts in those disciplines. While different language features are distinct, their entailment means that the deepest and most valuable understandings of text must account for complex interactions between those levels. For example, writing assessment that

[1] While machine learning is an application of computer science, there is value in differentiating core machine learning tasks such as clustering and classification, which are much less well-known within the writing instruction and digital humanities world than topic modeling.

focuses only on the stylistic moves that define genre may fall short in accounting for complex witting situations wherein content reflects cultural and material circumstances (Hester, 2007). Beigman Klebanov et al. (2018) make this very point in their study of generalization inferences of utility value scoring:

> We believe that our findings emphasize the need to study the applicability conditions for an automated system if it is going to be tasked with evaluating writing produced in a variety of institutional settings. Thus, it is possible that shifts to a new institution (and, thus, a new student population), a new subject matter course, or a new variant of the original task may result in systematic changes in the textual features that render the original system inapplicable to the new context. (p. 326)

Just as in the public policy world, the academic world needs a broadly holistic and usable toolset to deal with language data. As noted above, there are existing tools available for any single one of the above approaches—in terms of each capability, RAND-Lex is duplicative—but there is no other single environment from which to conduct analytics, switching seamlessly from one language feature set to the next, free to use whatever approach(es) match the question and ends. Further, by removing technical barriers to entry, my hope is that writing faculty and researchers can credibly and confidently use powerful, scalable analysis, regardless of their statistical expertise or familiarity with coding.

2.2 Multilingual Approach

In addition to breadth and usability, RAND-Lex offers a unique capability to use stance features in languages other than English. RAND-Lex's stance comparison module for English language corpora was built on Kaufer & Ishizaki's (2006) DocuScope dictionaries. The higher-order taxonomy in DocuScope reflects a theory-driven model of discourse pragmatics and function: humans use language in durable and repeated ways to achieve pragmatic ends. These moves include marking social distance or closeness, voicing certainty or uncertainty, indexing a range of emotions, citing sources of public authority, referring to the concrete world or to abstractions, and other socio-cultural language constructions. In building RAND-Lex, my colleagues and I hypothesized that this taxonomy was portable across other languages and have at this time built working native-language dictionaries using the same taxonomy in Russian and Arabic, with Mandarin Chinese in development. Further, our development team at RAND have incorporated stance as a feature in our machine learning modules: in addition to statistical corpus analysis of stance in those three languages, RAND-Lex users can cluster and classify texts using supervised and unsupervised algorithms. Standard approaches to text-based classification and clustering use word similarity, while RAND-Lex allows users to select custom weights of words and stance. This function is a significant affordance: it gives machines a way to recognize relationships between texts not just using the "aboutness" of lexis, but also using the functional lexicogrammar of genre and purpose—and do so in multiple languages.

2.2.1. Traditions. While writing analytics is in many ways a nascent research specialization, the corpus and text analysis tools used are well-proofed and long-standing in use for systematic empirical investigation of real-world language use (Biber, Douglas, Conrad, & Reppen, 1998; McEnery, Xiao, & Tono, 2006). In particular, drawing on distinct traditions allows for analysis using lexical, lexicogrammatical, and thematic features of texts. This ability to examine different dimensions of language is critical because of the role they play in genre performance. As Aull (2015) has demonstrated, this kind of analysis is especially important to writing studies where evidence can—thanks to new tools—be gathered rapidly at the lexical and grammatical level.

2.2.2 Transparency and fairness. RAND-Lex is aligned with recent calls for attention to evidence of fairness in quantitative analytic methodologies in writing studies (Poe & Elliot, 2019). Because of its transparency, RAND-Lex is able to reduce construct-irrelevant variance and related threats to validity inferences. Especially important here is the use of multiple statistical features that allow modeling—and thus, yielding multiple ways to draw inferences (AERA, APA, NCME, 2014).

2.2.3 Multidisciplinary interpretative frameworks. As is the case with any tool, RAND-Lex has limitations. Ultimately, the interpretation of test results is a human task. Human reading of machine reading is very powerful, and rigorous work with RAND-Lex or any similar set of tools requires human contextual knowledge to make meaning of the patterns described by the software. For example, only a human researcher with cultural and political high-context knowledge can make the interpretive leap to see that an overabundance of words around *government*, *rights*, *privacy*, and so on—accompanied by a lack of words around *science*, *policy*, *medical*, and similar topics—related to an on-line talk around New York City's soda ban reflects a rhetorical framework for a contested issue. This cluster has been interpreted as being about government overreach, not public health (Huang, 2016). Hence, there is significance in emphasizing multidisciplinary in terms of interpretative frameworks. As I will demonstrate in the analysis that follows, the interpretative framework that is used leverages the experiences of the two writing studies researchers—Norbert Elliot and Alice Horning—who curated the edited collection that serves as the corpus of the present study. During the analysis, Elliot served as subject matter expert who collaborated with me to deliberate the heuristic interpretations presented in § 4. An abbreviated set of these interpretations is provided in the introduction to the edited collection.

3.0. Technique Applications

3.1 Origin of the Corpus

The corpus used in this illustrative study is drawn from a forthcoming edited collection, *Talking Back: Senior Scholars and Their Colleagues Deliberate the Past, Present, and Future of Writing Studies* (in press). Edited by Elliot and Horning, the collection features new deliberations by 26 senior scholars (in terms of both experience and age) in writing studies with common research dissemination experiences: a sense of responsibility for advancing a profession; a passion for

programs of research dedicated to advancing opportunities for others; and a reflective sense of their work, accompanied by humility for their contributions. These chapters are accompanied by 23 respondents, most of whom are early or mid-career researchers. The editors are explicit in their aim for the edited collection:

> Our aim in this book is thus straightforward: to document a reflective vision of senior colleagues, approaching or passing the age of retirement, on the ways their unique programs of research have influenced our discipline; and to spark the imagination of their successors in charting future directions for writing studies in which difference, not homogeneity, is the aim.

Encouraged by the editors, I then designed a new quantitative technique study to increase the body of knowledge associated with writing studies, seniority studies, and writing analytics.

3.2 Creating Categories

To structure the analysis, I used the Classification of Instructional Programs (CIP) codes that resulted from the Visibility Project—an organizational technique also used by editors Elliot and Horning. In 2010, this Visibility Project successfully obtained assignment of a code series to rhetoric and composition/writing studies in the federal system used by the National Center for Education Statistics to gather US postsecondary data (Phelps & Ackerman, 2010). The categories are shown in the Appendix.

This system allows a taxonomy of the field that can in turn be used for analysis of how language within each field functions. For General Writing, I followed the editors and grouped Chapters 3, 6, 15, 16, and 21. For Creative Writing, I grouped Chapters 7 and 9. For Professional, Technical, Business, and Scientific Writing, I grouped Chapters 1 and 5. For Rhetoric and Composition, I grouped Chapters 2, 4, 8, 10, 11, 12, 13, 14, 17, 18, 19, 20, and 22. The analysis was intended to provide a heuristic under which additional questions could be raised about the topics covered in this volume.

3.3 Limits Based in Size

To illustrate the value of RAND-Lex as a holistic tools-suite for text analysis in writing studies and to better understand seniority in writing studies, I applied as many of the tests in RAND-Lex as possible to the *Talking Back* corpus. Because of both data and type sparsity, I could not build a topic model nor conduct human-supervised classification. I was, however, able to apply the lexical and lexicogrammatical tests in RAND-Lex. While I was not able to build a human-supervised model, I used an unsupervised machine learning function to create sensible groupings of these chunks based on intrinsic similarity.

3.3.1 Lexical design. For lexical features, I conducted two descriptive statistical tests: keyness testing and collocate extraction. Keyness testing finds conspicuously over-present words, while collocate analysis identifies words that occur near each other non-randomly, sometimes through habitual turns of phrase, but also indicating semantic preference, discourse

prosody, abstract concepts, or entities. RAND-Lex incorporates a number of different statistical tests to ensure rigor, including confidence measures, effect-size measures, and strength of association measures. Log-likelihood is a common confidence measure for keyness testing (Dunning, 1993), and is used in AntConc, Wordsmith Tools, and RAND-Lex. For this measure, a LL score of 10.83 (99.9th percentile; 0.1 percent level; $p < 0.001$) (Potts & Baker, 2012) was used as a threshold for significance. However, as corpora have become increasingly large in the "Big Data" age, confidence measures become less meaningful—at very large numbers, all differences are significant. Thus, effect-size measures are critical so that researchers can also ask a key question: "How meaningful is this difference?" Thus, RAND-Lex includes "percent difference" as an effect-size measure (Gabrielatos & Marchi, 2012): an easily interpreted +100 to -100% scale, with 0% being identical rates of occurrence of a keyword relative to a baseline refence corpus, then ranging from 100% more-present or less-present as positive or negative numbers. I also used two distinct measures to extract two- and three-word collocates: log-likelihood and pointwise mutual information (PMI; Church & Hanks, 1990). While log-likelihood measures confidence that word associations are non-random, PMI measures the strength of those word associations.

3.3.2 Lexicogrammatical design. To understand lexicogrammatical moves in the corpus, I conducted stance analysis, looking for pragmatic function: group counts of words or phrases in a high-order taxonomy of rhetorical and function purpose. For this analysis, I used two statistical tests: a pairwise ANOVA test of variance, "Tukey's Honest Statistical Difference" (HSD) test (Abdi & Williams, 2010), and exploratory factor analysis (EFA; Biber, Conrad, Reppen, Byrd, & Helt,2002).

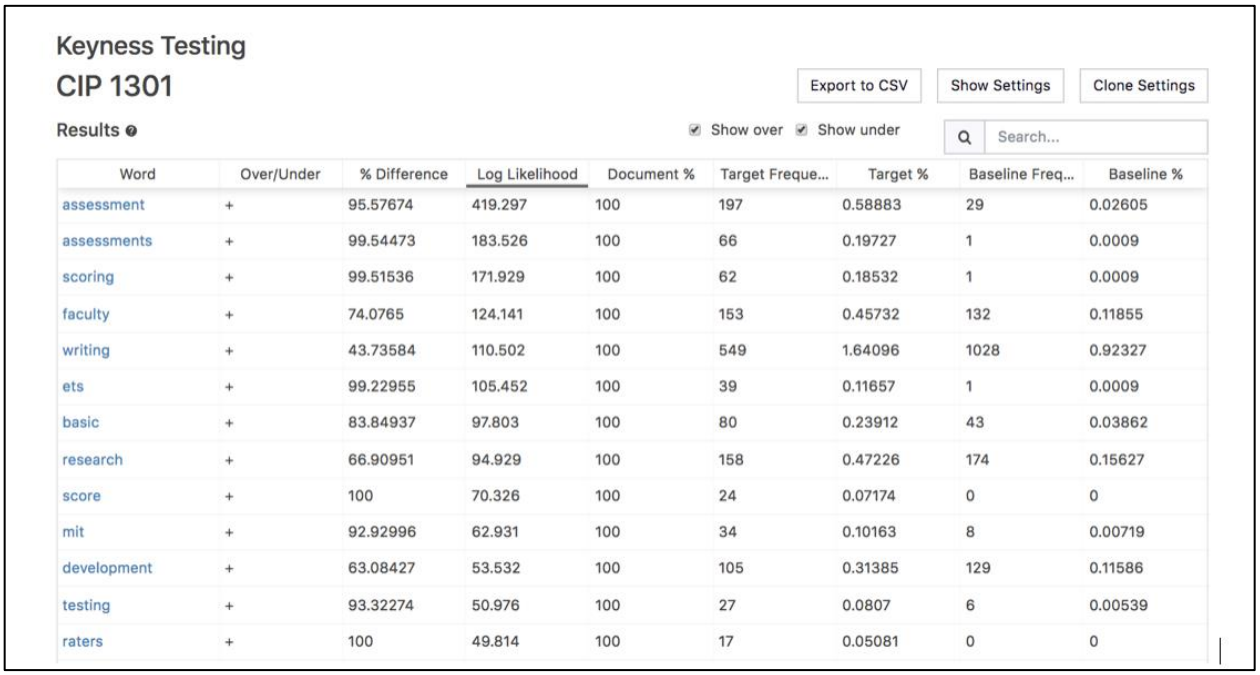
3.3.3 Auto-clustering. To demonstrate unsupervised machine learning (clustering) of documents, I used RAND-Lex's Auto-clustering (unsupervised machine learning) function to create sensible groupings of chapter chunks based on intrinsic similarity. To give the algorithm enough documents to work with, I artificially chunked chapters into 27 total smaller documents of approximately 6,000–8,000 words—enough pieces to be able to create a three-level cluster hierarchy. A typical use of clustering might be giving the algorithm a mixed collection of documents from a large, unsorted archive of hundreds of thousands, even millions of documents—in which the algorithm detects language features (words, stance) entailed in genres, and could thus group this large collection in human-recognizable ways: fiction in this cluster, news reporting in that cluster, and then technical documents in still a third. RAND-Lex uses a proprietary version of K-means clustering, a long-standing but still useful approach to data clustering (Jain, 2010).

4.0 Results

The analysis below was prepared in consultation with a subject matter expert and, as such, the analysis is deliberately reflective. As noted above, only a human researcher with cultural, high-context knowledge can make the heuristic interpretations described below.

4.1 Lexical Analysis: Keyness Testing

Keyness testing works by comparing word found frequencies and expected frequencies—in this analysis, for each CIP code identified in the Appendix—against the others. The screenshot in Figure 1 shows the results of one such test in RAND-Lex. The Over/Under column represents whether a word is over- or under-present (e.g., a negative "missing" keyword), as well as the % Difference (effect size) and Log-Likelihood score (confidence measure).[2] Analysts are able to sort results by clicking on any header, thus allowing for quick comparisons of keywords by confidence (LL) or by how large the difference is (% Difference).



Keyness Testing

CIP 1301

Export to CSV | Show Settings | Clone Settings

Results ☑ Show over ☑ Show under Search...

Word	Over/Under	% Difference	Log Likelihood	Document %	Target Freque...	Target %	Baseline Freq...	Baseline %
assessment	+	95.57674	419.297	100	197	0.58883	29	0.02605
assessments	+	99.54473	183.526	100	66	0.19727	1	0.0009
scoring	+	99.51536	171.929	100	62	0.18532	1	0.0009
faculty	+	74.0765	124.141	100	153	0.45732	132	0.11855
writing	+	43.73584	110.502	100	549	1.64096	1028	0.92327
ets	+	99.22955	105.452	100	39	0.11657	1	0.0009
basic	+	83.84937	97.803	100	80	0.23912	43	0.03862
research	+	66.90951	94.929	100	158	0.47226	174	0.15627
score	+	100	70.326	100	24	0.07174	0	0
mit	+	92.92996	62.931	100	34	0.10163	8	0.00719
development	+	63.08427	53.532	100	105	0.31385	129	0.11586
testing	+	93.32274	50.976	100	27	0.0807	6	0.00539
raters	+	100	49.814	100	17	0.05081	0	0

Figure 1. Keyness testing: CIP 1301 vs. Corpus.

4.1.1 General Writing. Under General Writing (CIP 1301, shown in Figure 1), presence of the terms *assessment*, *scoring*, *ETS*, *testing*, and *raters* reveals the central role that writing assessment research plays in writing studies. Words such as *faculty*, *development*, and *research* reveal a scholarly set of concerns that may be distinct from the often decontextualized environment of writing assessment as used for admission, placement, and progression.

[2] LL scores above 10.88

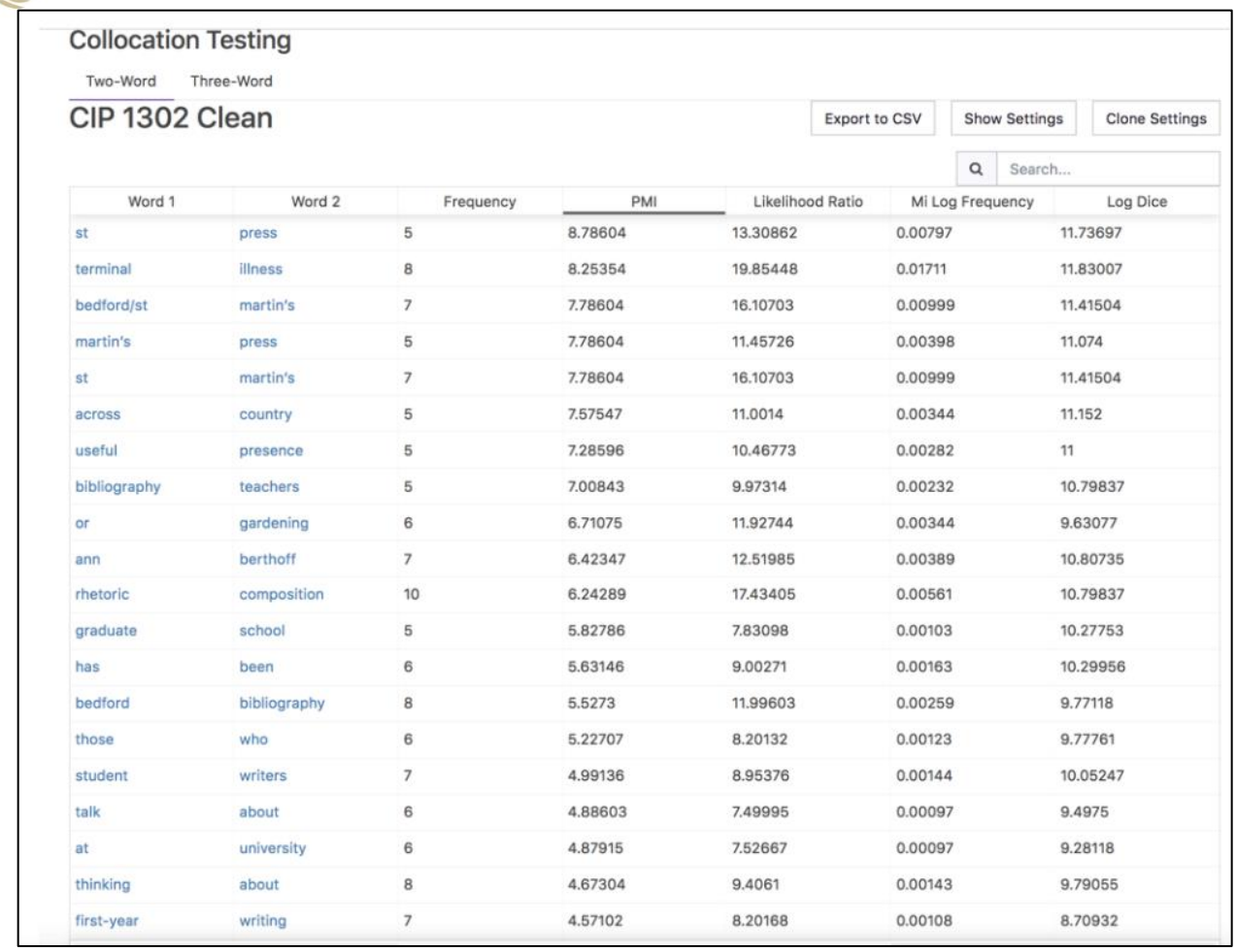
4.1.2 Creative Writing. Under the category of Creative Writing (not shown) are terms such as *experience*s, *hope,* and *ideas*, as well as *illness* and *without*—the language of promise and loss. Also under this category, the expected terms of *publishing* and *textbooks* are accompanied by *wisdom* and *success*—a reminder that, when done well, professional editing is about wisdom and that textbooks informed by such wisdom help students find success, not merely follow rules.

4.1.3 Professional Writing. While Professional, Technical, Business, and Scientific Writing (not shown) would, as expected, capture over-present words such as *technical* and *communication*, it is also interesting that these words are accompanied by the terms *discipline* and *field*, reminding us of the challenges to identity that technical communication scholars have encountered from humanistic study.

4.1.4 Rhetoric and Composition. Finally, under the Rhetoric and Composition category (not shown), attention to communication structures themselves is visible in the most conspicuously over-present term *language*, as well as *narrative*, and *languages*, in addition to research of those structures through professional organizations (with terms such as *NWP*, *WAC*, and *WPA*). Concerns over identity and racism are also visible in keywords such as *black*, *identity*, *African*, *negro*, and *racism.* The vocabulary and disciplinary concerns of rhetoric as a field are also present: *composing*, *reading*, *hospitality*, *rhetorics*, and *authoring.* Finally, the use of the terms *linguistic* and *raciolinguistics* suggests the deeply socio-cultural attitude toward language that is the hallmark of contemporary writing studies.

4.2 Lexical Analysis: Collocate Analysis

The screenshot provided in Figure 2 shows the results screen of a collocation analysis for the Creative Writing (CIP 1302) chapters. The screen has a two-word and three-word tab. As is the case with other results screens, there are buttons to show the user settings, create a duplicate analysis (helpful for users who want to vary a single parameter at a time on the same dataset and work empirically to explore), and export results to a comma-separated values (CSV) spreadsheet for further analysis and sharing outside of RAND-Lex. This screen shows two association measures, LL and PMI. In general terms, PMI scores under 3 indicate loose connections (Thanopoulos, Fakotakis, & Kokkinakis, 2002), whereas the highest natural occurring PMI score in FROWN corpus of 1990's English is 16, for *bum bum*—a remarkably unusual word combination.



Collocation Testing

Two-Word Three-Word

CIP 1302 Clean

Export to CSV | Show Settings | Clone Settings

Search...

Word 1	Word 2	Frequency	PMI	Likelihood Ratio	Mi Log Frequency	Log Dice
st	press	5	8.78604	13.30862	0.00797	11.73697
terminal	illness	8	8.25354	19.85448	0.01711	11.83007
bedford/st	martin's	7	7.78604	16.10703	0.00999	11.41504
martin's	press	5	7.78604	11.45726	0.00398	11.074
st	martin's	7	7.78604	16.10703	0.00999	11.41504
across	country	5	7.57547	11.0014	0.00344	11.152
useful	presence	5	7.28596	10.46773	0.00282	11
bibliography	teachers	5	7.00843	9.97314	0.00232	10.79837
or	gardening	6	6.71075	11.92744	0.00344	9.63077
ann	berthoff	7	6.42347	12.51985	0.00389	10.80735
rhetoric	composition	10	6.24289	17.43405	0.00561	10.79837
graduate	school	5	5.82786	7.83098	0.00103	10.27753
has	been	6	5.63146	9.00271	0.00163	10.29956
bedford	bibliography	8	5.5273	11.99603	0.00259	9.77118
those	who	6	5.22707	8.20132	0.00123	9.77761
student	writers	7	4.99136	8.95376	0.00144	10.05247
talk	about	6	4.88603	7.49995	0.00097	9.4975
at	university	6	4.87915	7.52667	0.00097	9.28118
thinking	about	8	4.67304	9.4061	0.00143	9.79055
first-year	writing	7	4.57102	8.20168	0.00108	8.70932

Figure 2. Two-word collocates from CIP 1302.

4.2.1 General Writing. Under General Writing (not shown), use of the paired terms *qualitative* and *quantitative* calls to mind the mixed-methods perspective that is characteristic of writing studies research. As expected, assessment language is also visible here: *outcomes statement*, *testing service*, *educational testing*. Paired terms *automated software* and *automated scoring* recall the promise and perils associated with machine approaches to assessment. Terms *reflective* and *practitioners* recall the sense of self-awareness characteristic of writing studies scholars. The paring of *active* and *wisdom* reflects the influence of cultural anthropologist and seniority scholar Mary Catherine Bateson (2010) on chapter authors—and her attention to growth in old age. Finally, collocate extraction, so useful for identifying abstractions, suggests the goods prized in this discipline, such as *critical thinking* and *social justice*.

4.2.2 Creative Writing. In the Creative Writing analysis shown in Figure 2, *terminal* was paired with *illness*, capturing the loss that Eli Goldblatt, Jessica Restaino and Paige Davis Arrington (with Ann E. Berthoff) describe as the knowledge paradox of end-of-life experiences. From Joan Feinberg's chapter, evident is the name of the press—*Bedford/St. Martin's*—that sponsored the *Bedford Bibliography*, the long-standing imprint that serves as a well-known guide

to research in writing studies. The term *useful presence* is a dominant theme of Feinberg in that the collocate stands as motto of the Bedford story. As Feinberg writes, "This is the Bedford story—to be 'a useful presence on the college composition scene.' We defined ourselves as a different kind of publisher, seeing ourselves as part of the disciplines we published for."

4.2.3 Professional Writing. Under the category of Professional, Technical, Business, and Scientific Writing (not shown), are expected pairs—*professional communication* and *technical communication*. Additional important combinations are: *pedagogical* and *theoretical*; and *social* and *justice*. Emphasis on the humanistic tradition in technical writing, manifested in theory-based pedagogies, has long been central to technical communication, as Jo Allen and Hugh Burns demonstrate in their chapters. The focus here is deepened and made specific by a renewed emphasis on social justice, understood as actions and practices that help to alleviate oppression, a direction for future research made clear by Michele Eble and Ann N. Amicucci.

4.2.4 Rhetoric and Composition. Under the Rhetoric and Composition category (not shown), there is attention paid to Bateson and her concept of the age of active wisdom. *Voices* is the term pared with *self*, suggesting the importance of reflection as central to the creation of voice. *Artificial intelligence*, *intimate machines*, and *digital technologies* all reflect a technological awareness in the discipline.

4.3. Lexicogrammatical Analysis: Stance Analysis

As noted above, RAND-Lex uses the expert dictionaries from DocuScope, developed by David Kaufer and Suguru Ishizaki at Carnegie Mellon University (Kaufer, Geisler, Vlachos, & Ishizaki, 2006). RAND-Lex measures and visualizes quantity differences for stance features, such as social closeness or distance language. Accompanying identification of these features, the tool uses statistical measures for analysis of variance (ANOVA) and exploratory factor analysis (EFA). Figure 3 highlights a small section of a fairly long results page. There are, in fact, 121 language categories (LCs) in the DocuScope dictionaries used. These 121 most granular LCs are grouped within higher order categories, so for example *anger*, *sadness*, *regret*, *apology*, *fear*, *general negativity*, and *general positivity* are all within the larger category of *Emotion*. In the Figure 3 screenshot, readers can see two social LCs (Social Closeness and Social Distance language) and the Asides LC within the higher category *Narrative*. The range of a given LC's use is displayed by box and whisker plots, color-coded by text corpus (in this case the four CIP codes and the FROWN corpus as a general reference point for English). Each box shows the interquartile range (from 25th to 75th), with a dot for the median. The whiskers show maximums and minimums for the LC. All of those numbers can be read directly via mouse-over on the screen or, if exported, in a CSV file. In essence, shorter boxes indicate less variability between documents, and longer ones more.

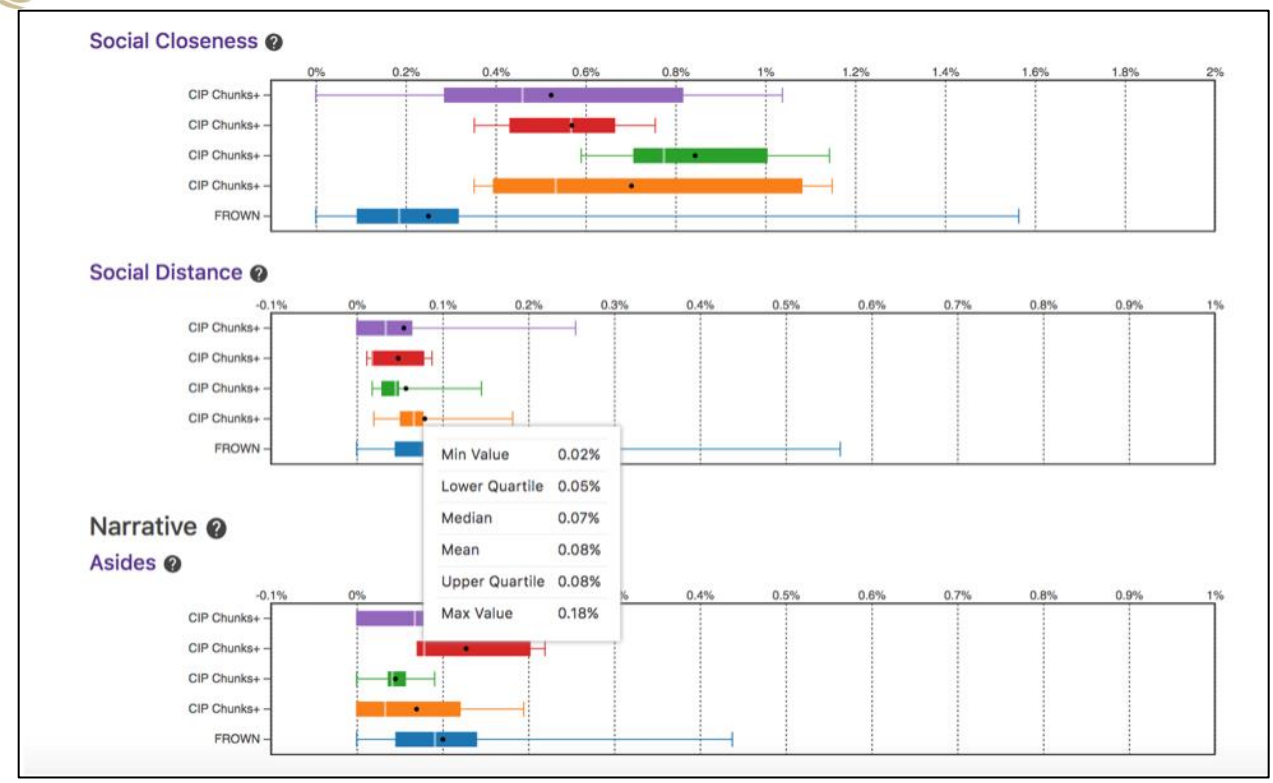


Figure 3. Stance analysis summary screen.

I also used exploratory factor analysis (EFA) to discover textual latencies (rhetorical strategies or genre features) in the corpus through the co-variance of style: purposeful language moves entailed in accomplishing rhetorical tasks. Because the Creative Writing category was so small, it was not included in the factor analysis.

4.3.1 General Writing. General Writing was distinct from the other categories in two factors. First, in articulating the importance of their programs of research, these authors combine *curiosity* ("inquiries into," "study"), *generalization* ("every" campus), *insistence* ("need to"), *causal reasoning backwards* ("because," "stem from"), and *looking ahead* ("gateway to," "in order to"). Second, these authors framed problem explication with *intensity* ("any and all," "even the") and *specifiers* ("several," "a range of"). Focus under this category is on applied contexts such as educational measurement (Doug Baldwin and Devon Tomasulo; Edward M. White and Sherry Rankins-Robertson), faculty development (William Condon and Mike Truong), writing placement (Rebecca Williams Mlynarczyk and Sean Molloy), and communication across disciplines (Les Perelman and Suzanne Lane). It is therefore notable that EFA analysis revealed *inquisitiveness*, *demand*, *causal logic*, and *directed action*—the essence of applied research.

4.3.2 Professional Writing. Professional, Technical, Business, and Scientific Writing was distinct in its co-variance of values language, both of *innovation* ("innovative applications," "new fields") and *public sphere virtues* ("collegiality," "sustainability"). Here again, authors are concerned with the location of their field (the professionalization of technical and professional

communication as envisioned by Jo Allen) and its imperative toward ethical conduct (as voiced by Hugh Burns and Ann N. Amicucci)—rhetorical ends that require specific kinds of language.

4.3.3 Rhetoric and Composition. Rhetoric and Composition had a unique factor of what I have labeled *critical complaint*: *negative emotion* ("frustrated," "at risk," "rigid"), *public sphere vice* ("racism," "imperialism"), and *linguistic references* ("language," "rhetorics of"). The covariance of LCs EFA detected reveals a critical posture towards understanding that oppression can be found at the level of word choice, and freedom is found in linguistic diversity. This posture is indicative of chapters on sociolinguistics (Akua Duku Anokye and Patricia Friedrich), language difference (Min-Zhan Lu, Bruce Horner, and Dylan Dryer), and raciolinguistics (Geneva Smitherman and Shenika Hankerson).

Figure 4 illustrates the stance analysis annotation results screen. Stance analysis can shift from the macro-view summary of the findings shown above, down to feature-rich examples from the text, an inherently mixed-method approach involving both computer quantification and human interpretative analysis.

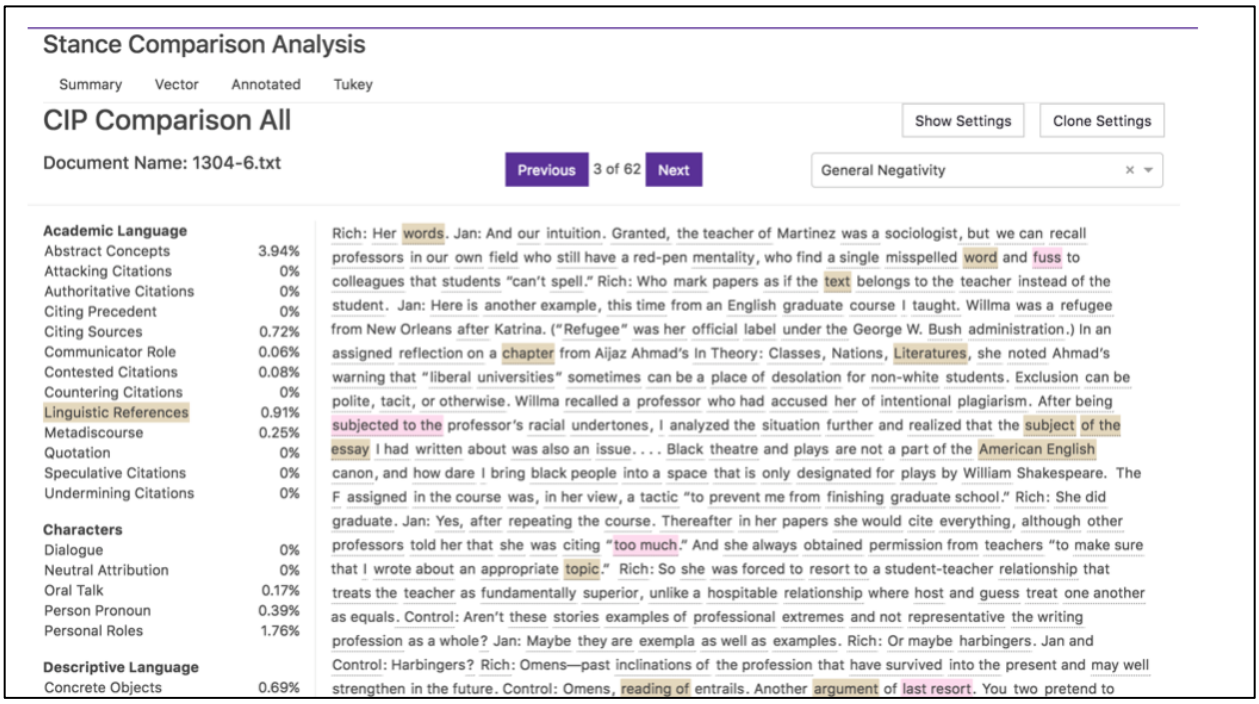


Figure 4. Stance analysis annotated screen.

The passage shown is from a chapter by Janis and Richard Haswell. In this section of the chapter, the authors recollect the case of a graduate student who has suffered loss following Hurricane Katrina in New Orleans. In an assigned reflection on a chapter from Aijaz Ahmad's *In Theory: Classes, Nations, Literatures* (1992), the student noted Ahmad's warning that "liberal universities" sometimes can be a place of desolation for non-white students. The student had been accused on intentional plagiarism. Notable here are the identified linguistic references: *words*, *text*, *literature*, *American English*, *topic*, and *argument*. Putting these linguistic references

in context with that particular section of the chapter, one inference is clear: Discrimination in the academy is likely to take place at the level of textual selection and canon formation. The authors note that such textual discrimination may well be "omens—past inclinations of the profession that have survived into the present and may well strengthen in the future."

4.4 Auto-Clustering

In order to have enough documents to work with, I cut the corpus into approximately 6,000–8,000 word chunks of text from the book chapters (an arbitrary size), and then ran the clustering algorithm twice: first using just words ("terms" in computational parlance), and then a second time using a mix of words and stance, weighed 30% to words and 70% to stance (a user-adjustable setting in the software). The word-only attempt produced only noise—chunks from the work were randomly grouped. However, in the second attempt when I added stance, I found two sensible clusters, both from Rhetoric & Composition (CIP 1304). Figure 5 illustrates a cluster of six CIP 1304 documents.

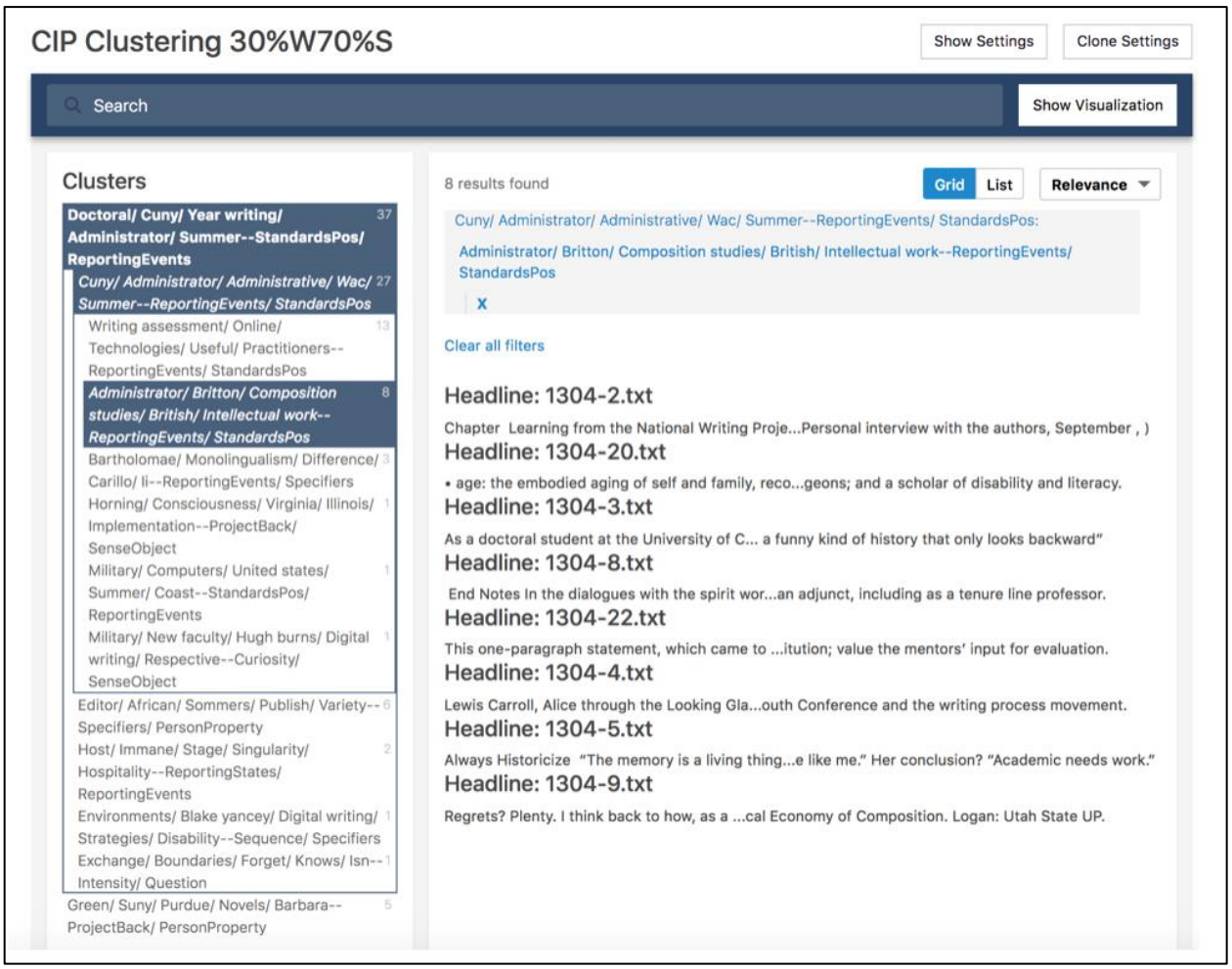


Figure 5. Auto-clustering results screen.

The heading for each cluster shows the top words/stance LCs from that cluster, so that the analyst can better understand why the documents are clustered (irrespective of the weighting). This iteration was 70% weighted towards stance, and so the LCs for *positive public standards of behavior* and **reporting events** in narration are the basis for the algorithm's two clusters. When looking at feature-rich samples, it's clear how rhetoric and composition scholars—in this case, Kathleen Blake Yancey—often position historical occurrences within their pursuit of positive values:

> And more specifically, as this chapter explains, seniors, **responding** to new exigencies from declines in memory to unfamiliar tasks, may need to be even more *flexible* and communal as composers, *flexible* in *adapting to* new needs and sometimes with technologies also new to them, communal in writing to and with others and in **calling on** *the help of* others to *facilitate* **composing**. (boldface and italics added for emphasis)

In responding to these exigencies, Yancey notes, older writers continue to exercise agency as their identities shift—from middle age, to late middle age, to seniority—which may be particularly challenging given the other shifts in roles seniors experience. The positive stance is evident.

As is the case with the other techniques, auto-clustering may well help us understand rhetorical frameworks for new areas of research, such as seniority studies.

5.0. Directions for Further Research

RAND-Lex was developed as a general-purpose research tool for public policy research, meant to answer questions like how people across the globe support or oppose ISIS over social media (Bodine-Baron, Helmus, Magnuson, & Winkelman, 2016) or how a mental health and wellness campaign impacts local discourse (Collins et al., 2018). Applications for RAND-Lex in writing studies and writing analytics are just beginning.

As this new quantitative technique study demonstrates, the benefits of working with genre-specific corpora are only now becoming clear. Dryer (2019) has used a corpus consisting of 1,818 research articles from the last ten years published in the leading writing studies journals. In his analysis of keywords, he found that four emergent terms—*communication*, *context*, *English*, and *language*—bring to our attention "a decade's worth of hard-earned insights" as well as "disciplinary problems requiring attention" (p. 241). As he concludes, the analysis "has revealed what we have meant—whatever the level of consciousness—with the words that distinguish writing studies from other academic disciplines" (p. 249). Similarly, the analysis has also illuminated why some conversations may rely on the same keywords but still do not intersect, thus revealing a collective disciplinary discourse of long-standing and divisive entrenchments. Dryer's study has much in common with the case study presented here in terms of allowing a fine-grained analysis of topics in writing studies—in our case, seniority studies—that, in turn,

allow us to examine language patterns and their use. As new specializations continue to emerge within writing studies, corpus analysis will continue to play a significant role.

Pedagogically, I hope to soon find a writing analytics-relevant case to explore autocoding (human supervised classification) applications. An obvious use case would be in assessment: a reliable, accurate algorithmic approach for assessing student writing. However, whether or not assessments of whole texts are possible (or desirable), there is potential for algorithmic approaches to triaging specific and constrained functional sections of genre performances. If genres do impose lexicogrammatical constraints (Allison et al., 2011), then it is quite possible that stock sections in different kinds of texts taught could *be assessed during drafting* to help students and instructors most efficiently leverage peer and expert revision guidance attention.

While RAND-Lex can be used for instruction, it lacks custom workflows and custom UI designed specifically for writing analytics. An example would be comparing a student essay to a model corpus to look for style variance: Does an argumentative essay have so much subjective talk that its genre-profile looks closer to a personal reflection? Currently in RAND-Lex this requires an instructor to run an analysis for each student's essay. A truly user-friendly design would work at the level of the assignment: "Computer, here is a model corpus, here are essays from my class—please give me a detailed report for each one to help inform my revision meetings with them." Along with my fellow designers, I hope in future releases of RAND-Lex to include such workflows so that instructors spend as little overhead as possible on the mechanics of uploading data and running analyses—and more time on interpretation and analysis.

Author Biography

Bill Marcellino is a Senior Behavioral Scientist at the RAND Corporation, a non-profit public policy research institution. Bill is a Professor of Text Analytics within the Frederick S. Pardee RAND Graduate School and teaches qualitative research methods within John Hopkin's Global Security Studies program. He also serves on the editorial review board of *The Journal of Writing Analytics*.

References

Abdi, H., & Williams, L. J. (2010). Tukey's honestly significant difference (HSD) test. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 1–5). Thousand Oaks, CA: Sage.

Ahmad, A. (1992). *In theory: Classes, nations, literatures.* UK: Verso.

Allison, S. D., Heuser, R., Jockers, M. L., Moretti, F., & Witmore, M. (2011). *Quantitative formalism: An experiment.* Stanford Literary Lab.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anthony, L. (2004). *Overview of the "AntConc" and "AntMover" text analysis tools*. Invited lecture hosted by the University of Michigan, English Language Institute (ELI).

Aull, L. (2015). Linguistic attention in rhetorical genre studies and first-year writing. *Composition Forum, 31*. Retrieved from http://compositionforum.com/issue/31/linguistic-attention.php

Bateson, M. C. (2010). *Composing a further life*: *The age of active wisdom.* New York: Knopf.

Beigman Klebanov, B., Prinisky, S., Burstein, J., Gyawali, B., Harackiewicz, J., & Thoman, D. (2018). Utility-value score: A case study in system generalization for writing analytics. *The Journal of Writing Analytics*, *2*, 314–328. Retrieved from https://wac.colostate.edu/docs/jwa/vol2/klebanov.pdf

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly*, *36*(1), 9–48.

Biber, D., Douglas, B., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge, UK: Cambridge University Press.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bodine-Baron, E., Helmus, T. C., Magnuson, M., & Winkelman, Z. (2016). *Examining ISIS support and opposition networks on Twitter*. Santa Monica, CA: RAND Corporation.

Bowers, A. J. (2007). *Grades and data driven decision making: Issues of variance and student patterns* (Doctoral dissertation). Retrieved from ERIC database. (ED538574)

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Burstein, J., Elliot, N., Beigman Klebanov, B., Madnani, N., Napolitano, D., Schwartz, M., Houghton, P., & Molloy, H. (2018). Writing Mentor™: Writing progress using self-regulated writing support. *The Journal of Writing Analytics*, *2*, 285–313. Retrieved from https://wac.colostate.edu/docs/jwa/vol2/bursteinetal.pdf

Burstein, J., Tetreault, J., & Madnani, N. (2013). The e-rater® automated essay scoring system. In M. D. Shermis & J. Burstein (Eds.), *Handbook for automated essay scoring* (pp. 55–67). New York, NY: Routledge.

Chen, Y. Y., Liu, C. L., Lee, C. H., & Chang, T. H. (2010). An unsupervised automated essay scoring system. *IEEE Intelligent systems*, *25*(5), 61–67.

Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, *16*(1), 22–29.

Collins, R. L., Eberhart, N. K., Marcellino, W., Davis, L., & Roth, E. (2018). *Evaluating Los Angeles County's mental health community engagement campaign*. Santa Monica, CA: RAND Corporation.

Crossley, S. A., & Kyle, K. (2018). Assessing writing with the tool for the automatic analysis of lexical sophistication (TAALES). *Assessing Writing*, *38*, 46–50.

Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning and Assessment*, *5*(1). Retrieved from https://ejournals.bc.edu/index.php/jtla/article/view/1640/1489

Dryer, D. B. (2019). Divided by primes: Competing meanings among writing studies' keywords. *College English*, *81*, 214–255.

Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics, 19*(1), 61–74.

Elliot, N., & Horning, A. (Eds.). (in press). *Talking back: Senior scholars and their colleagues deliberate the past, present, and future of writing studies*. Logan, UT: Utah State University Press.

Gabrielatos, C., & Marchi, A. (2012). Keyness: Appropriate metrics and practical issues. CADS International Conference 2012, University of Bologna, Italy.

Gee, J. P. (2004). *Situated language and learning: A critique of traditional schooling*. UK: Routledge.

Halliday, M. A. (1994). *An introduction to functional grammar*. UK: Edward Arnold.

Helberg, A., Poznahovska, M., Ishizaki, S., Kaufer, D., Werner, N., & Wetzel, D. (2018). Teaching textual awareness with DocuScope: Using corpus-driven tools and reflection to support students' written decision-making. *Assessing Writing, 38*, 40–45.

Helmus, T. C., Bodine-Baron, E., Radin, A., Magnuson, M., Mendelsohn, J., Marcellino, W., Bega, A., & Winkelman, Z. (2018). *Russian social media influence: Understanding Russian propaganda in Eastern Europe*. Santa Monica, CA: RAND Corporation.

Hester, V. (2007). When pragmatics precede pedagogy. *Journal of Writing Assessment, 3*(2), 123–144.

Horner, B., & Tetreault, L. (Eds.). (2017). *Crossing divides: Exploring translingual writing pedagogies and programs*. Logan, UT: University Press of Colorado.

Huang, C. Y. (2016). *Improving diet in low income neighborhoods*. Santa Monica, CA: RAND Corporation.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, 31*(8), 651–666.

Kaufer, D., Geisler, C., Vlachos, P., & Ishizaki, S. (2006). Mining textual knowledge for writing education and research: The DocuScope project. In L. van Waes, M. Leijten, & C. Neuwirth (Eds.), *Writing and digital media* (pp. 115–129). Amsterdam: Elsevier.

Lang, S., Aull, L., Marcellino, B. (2019). A taxonomy for writing analytics. *The Journal of Writing Analytics, 3*.

Marcellino, W. M., & Magnuson, M. (2019). ISIS versus the United States: Rhetorical battle in the Middle East. In J. Ridolfo & W. Hart-Davidson (Eds.), *Rhet Ops: Rhetoric and Information Warfare* (pp. 125–141). Pittsburgh, PA: University of Pittsburgh Press.

Marcellino, W., Smith, M. L., Paul, C., & Skrabala, L. (2017). *Monitoring social media: Lessons for future department of defense social media analysis in support of information operations*. Santa Monica, CA: RAND Corporation.

McEnery, A. M., & Wilson, A. (2001). *Corpus linguistics: An introduction*. Edinburgh: Edinburgh University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book.* New York: Routledge.

Meeks, E., & Weingart, S. B. (2012). The digital humanities contribution to topic modeling. *Journal of Digital Humanities, 2*(1), 1–6.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC 2001*. Mahwah: Lawrence Erlbaum Associates.

Phelps, L. W., & Ackerman, J. W. (2010). Making the case for disciplinarity in rhetoric, composition, and writing studies: The Visibility Project. *College Composition and Communication, 62*, 180–215.

Poe, M, & Elliot, N. (2019). Evidence of fairness: Twenty-five years of research in *Assessing Writing*, *42*, *Assessing Writing*. 100418.

Potts, A., & Baker, P. (2012). Does semantic tagging identify cultural change in British and American English? *International Journal of Corpus Linguistics, 17*, 295–324.

Richards, I. A. (1991). Context theory of meaning and types of context. In A. Berthoff (Ed.), *Richards on rhetoric* (pp. 111–117). New York: Oxford University Press.

Schöch, C. (2015). Topic modeling French crime fiction. In *DH 2015 Global Digital Humanities Conference abstracts*.

Scott, M. (2008). WordSmith Tools version 5 [Computer software]. Liverpool: Lexical Analysis Software.

Thanopoulos, A., Fakotakis, N., & Kokkinakis, G. (2002). Comparative evaluation of collocation extraction metrics. Proceedings of *International Conference on Language Resources and Evaluation*, 620–625.

Appendix: Chapters Mapped to CIP Code 23.13: Rhetoric and Composition/Writing Studies (Elliot & Horning, 2020, reprinted with permission)

23.1301: Writing, General
Definition: Writing for applied and liberal arts purposes. Includes instruction in writing and document design in multiple genres, modes, and media; writing technologies; research, evaluation, and use of information; editing and publishing; theories and processes of composing; rhetorical theories, traditions, and analysis; communication across audiences, contexts, and cultures; and practical applications for professional, technical, organizational, academic, and public settings
Chapter 3. Douglas Baldwin. "The Times, They Are A-Changin": One Assessment Specialist's Reflections on the Evolution of Research and Policy in Large-Scale Assessments at One Assessment Company Response: Devon Tomasulo. "You Better Start Swimmin' Or You'll Sink Like A Stone": How Assessment Keeps Changin' Chapter 6. William Condon. Assessment as a By-Product of Ongoing Research: Identifying, Describing, and Nourishing a Campus Culture of Teaching and Learning Response: Mike Truong. From Assessment as Research to Empirical Education Chapter 15. Rebecca Williams Mlynarczyk. Rethinking Basic Writing: Reflections on Language, Education, and Opportunity Response: Sean Molloy. A Reckoning for Basic Writing Chapter 16. Les Perelman. Contact Zones Across the Disciplines Response: Suzanne Lane. Writing Research Across Disciplinary Boundaries Chapter 21. Edward M. White. Fifty Years of Curriculum Changes: Looking In and Looking Out in College Writing Classes Response: Sherry Rankins-Robertson. Looking Back to Move Forward: A Response to Edward M. White

23.1302: Creative Writing
Definition: Process and techniques of original composition in various literary forms such as the short story, poetry, the novel, and others. Includes instruction in technical and editorial skills, criticism, and the marketing of finished manuscripts.
Chapter 7. Joan Feinberg. A Bedford Story: Taking the Measure of a Publisher Response: Leasa Burton. On Being Useful Chapter 9. Eli Goldblatt. Writing Wisdom: A Meditative Quilt Response: Jessica Restaino. Doors, Walls, and the Paradox of Not Knowing Response: Paige Davis Arrington, with Ann E. Berthoff. Legacy and Invitation
23.1303: Professional, Technical, Business, and Scientific Writing
Definition: Professional, technical, business, and scientific writing; and that prepares individuals for academic positions or for professional careers as writers, editors, researchers, and related careers in business, government, non-profits, and the professions. Includes instruction in theories of rhetoric, writing, and digital literacy; document design, production, and management; visual rhetoric and multimedia composition; documentation development; usability testing; web writing; and publishing in print and electronic media.
Chapter 1. Jo Allen. The Professionalization and Future of Technical and Professional Communication Response: Michele F. Eble. Turning Towards Social Justice Approaches to Technical and Professional Communication Chapter 5. Hugh Burns. Intimate Machines: Cultivating Wisdom in Elder Gardens Response: Ann N. Amicucci. Toward a Research Agenda for Digital Intimacy

23.1304: Rhetoric and Composition

Definition: Humanistic and scientific study of rhetoric, composition, literacy, and language/linguistic theories and their practical and pedagogical applications. Includes instruction in historical and contemporary rhetoric/composition theories; composition and criticism of written, visual, and mixed-media texts; analysis of literacy practices in cultural and cross-cultural contexts; and writing program administration.

Chapter 2. Akua Duku Anokye. Talking Brought Me Here: Sociolinguistics and African American Life
Response: Patricia Friedrich. Still Talking: Embracing Varieties, World Englishes, and the Power of Words
Chapter 4. Judy Buchanan and Richard Sterling. Learning from the National Writing Project as a Kindergarten–University Partnership: Talking Back and Forth
Response: Anne Elrod Whitney. Talking Back and Forth Between Memory and Legacy in the National Writing Project
Chapter 8. Cinthia Gannett and John C. Brereton. Framing and Facing Histories of Rhetoric and Composition: Composition-Rhetoric in the Time of the Dartmouth Conference
Response: Katherine E. Tirabassi. History Has Moved Through Us
Chapter 10. Janis Haswell and Richard Haswell. "Bottomless Mysteries" on the Margins: A Dream Interview
Response: Stacey Pigg. Toward Open Exchanges in a Networked World
Chapter 11. Douglas Hesse. Aging through the Thirty-Year Rise of Professionalized Writing Administration
Response: Eliana Schonberg, Embracing the Accidental Trajectory
Chapter 12. Alice S. Horning. Reading Old and New: An Autobiography and an Argument
Response: Ellen Carillo. Talking Back (to Alice Horning)
Chapter 13. Min-Zhan Lu and Bruce Horner. Rewriting the Language(s) of Language Differences in Writing
Response: Dylan B. Dryer. Not Trajectory but Translation: Talking Back With and To Min-Zhan Lu and Bruce Horner
Chapter 14. Donald McQuade. Starting from Scratch: Practicing and Teaching the Work of Words
Response: Eric Heltzel. The Goal of Teaching is to Become Obsolete
Chapter 17. Louise Wetherbee Phelps. Identity Work: Continuities and Transformations in the Senior Years
Response: Elisabeth L. Miller. Reading Identity Work through a Disability Lens: Care, Bodies, and Time
Chapter 18. Geneva Smitherman. Raciolinguistics and the "Mis-education of the Negro"–And You Too: Race, Language, and the Elder in "Post-Racial" America
Response: Shenika Hankerson. "I Love My African American Language. And Yours": Toward a Raciolinguistic Vision in Writing Studies
Chapter 19. Martha A. Townsend. Valuing New Approaches for Tenure and Promotion for WAC/WID Scholar/Administrators: Advice for Higher Education and the Writing Studies Community