Combining rhetorical move analysis with multi-dimensional analysis

Research writing across disciplines

Bethany Gray[i], Elena Cotos[i] and Jordan Smith[ii]
[i]Iowa State University / [ii]University of North Texas

We carry out comprehensive form–function mapping in Introduction-Methods-Results-Discussion/Conclusion – structured research articles across 30 academic disciplines by merging move analysis (revealing rhetorical structure) and multi-dimensional (MD) analysis (modeling patterns of linguistic variation). These two analytic paradigms converge to map the communicative functions of text segments with patterns of functional linguistic variation, but the juxtaposition of these two approaches requires adaptations to the traditional MD methodology. The primary purpose of this chapter is to provide a detailed account of the methodological complexities of combining MD analysis and move analysis. We briefly present the results of the MD analysis, acknowledging that further methodological adjustments may be needed to arrive at an optimal multi-dimensional description of moves.

1. Introduction

Studies of expert academic research writing have often focused on discipline-specific or cross-disciplinary patterns of variation. Two types of analyses have been particularly common: investigations into (a) the linguistic characteristics of texts through analysis of lexical, phraseological, grammatical, and lexico-grammatical patterns of use; and (b) the rhetorical composition of texts through Swalesian (1981, 2004) move analysis. Comparatively less research, however, has combined these perspectives to study the underlying linguistic realizations of rhetorical moves (e.g., Kanoksilapatham 2007; Tankó 2017) or the move-related functions of particular linguistic features (e.g., Cortes 2013 on lexical bundles as "triggers" of moves and steps in research-article introductions). Such systematic connections between the rhetorical functions of texts and linguistic patterns of use bridge the "function-form gap" (Moreno & Swales 2018: 41) and contribute to the pedagogical

https://doi.org/10.1075/scl.95.06gra

aims of English for Academic Purposes (EAP), particularly for the teaching and learning of research-writing skills.

In this chapter, we report on a project aimed at extensive form–function mapping in Introduction-Methods-Results-Discussion/Conclusion (IMRD/C) – structured research articles across 30 disciplines. The analyses merge two analytic paradigms, both of which have a fundamental focus on communicative function but which approach the text in distinctive ways:

1. move analysis (Swales 1981, 2004), which considers the rhetorical composition of texts based on the communicative functions of segments of discourse; and
2. multi-dimensional (MD) analysis (Biber 1988), which considers the linguistic patterning of texts, examining how co-occurrence patterns of lexical, grammatical, and lexico-grammatical structures reflect shared communicative functions associated with those features.

The application of MD analysis to move-annotated corpora is complex, requiring several adaptations to traditional MD methodology. In part, this is because MD analysis was developed as a method for modeling variation *across* texts, particularly across texts covering a range of spoken and written registers (Biber 1988). In contrast, applying MD analysis to a move-annotated corpus considers variation *within* texts, focusing on variation across internal segments of texts. In addition, the texts belong to a single, relatively narrow domain (published academic research writing). This shift in the observational unit for MD analysis subsequently impacts multiple analytical stages:

1. establishing the move as the unit of observation, rather than the continuous text;
2. accounting for the issue of text length when the move is the unit of observation;
3. selecting appropriate linguistic variables to be included in the model given the issues of text length and the relatively narrow domain being modeled; and
4. choosing an appropriate statistical method that can best handle the resulting data structure.

In this chapter, our goal is to provide a detailed, step-by-step account of the methodological complexities of combining MD and move analysis, as previous studies carrying out similar analyses have not provided sufficient or transparent methodological details. We fully document each adaption used to resolve the methodological issues introduced above. We then briefly present the results of the analysis and evaluate to what extent these adaptations were successful, pointing to potential alternatives and future directions and discussing the broader implications in terms of replicable and reproducible form–function mapping research.

1.1 Move analysis

Move analysis originated in English for Specific Purposes (ESP) to offer a genre-based framework for teaching post-secondary students "the myriad communicative events that occur in the contemporary English-speaking academy" (Swales 1981: 1). Swales' (1981) framework encompasses two key concepts – *moves* and *steps* – that characterize genres as rhetorical actions practiced by discourse communities. The moves are discourse units, or text segments, performing coherent communicative goals (Swales 2004), while steps are rhetorical strategies that help to fulfill move goals.

Move analysis has evolved into a robust, corpus-based inquiry unfolding in four phases: move/step-schema development, trialing, validation, and analytic description (Cotos 2018). The first three phases enable researchers to generate frameworks with detailed and comprehensive inventories of move/step categories (including functional definitions, descriptions of rhetorical and content realizations, and examples of characteristic language choices) that can be applied reliably and that fully account for the genres being analyzed. The last phase entails manual annotation of corpus data for moves and steps, allowing for quantitative analyses that reveal the "anatomy" of genre texts[1] (Cotos, Huffman & Link 2017) and in turn inform genre-based writing pedagogy.

1.2 Investigating the linguistic characteristics of moves

A number of studies have established meaningful links between the use of grammatical or phraseological features and the moves of research articles (RAs), revealing important form–function relationships. Table 1 lists these studies, summarizing the part-genres and disciplines, moves, and features investigated in each in order to demonstrate the range of analyses that have been carried out. Table 1 shows that studies have typically been restricted to a single part-genre (Cortes 2013 on introductions; Omidian, Shahriari & Siyanova-Chanturia 2018 on abstracts), a single move (Durrant & Mathews-Aydınlı 2011; Le & Harrington 2015), and/or a single discipline (Kanoksilapatham 2007; Le & Harrington 2015; Marco 2000; Mizumoto, Hamatani & Imao 2017; Tankó 2017). In addition, there has been a particular focus on lexical/phraseological features, with six of the eight studies listed focusing on recurrent formulaic language (e.g., lexical bundles, clusters, collocational frameworks).

1. For example, analytic description can reveal prototypical and optional occurrences of rhetorical constituents, differential distributions of moves/steps, variation conditioned by disciplinary cultures, and clustering of disciplines based on epistemology and research tradition (Cotos et al. 2017).

Table 1. Survey of studies mapping the use of linguistic features to moves in research writing

Study	Genre (or part-genre)	Moves/Steps	Linguistic features
Cortes (2013)	Research article introductions in 13 disciplines (1,372 texts)	3 moves (and associated steps)	lexical bundles
Durrant and Mathews-Aydınlı (2011)	Introductions in student essays (96 texts) and research articles (94 texts) in 8 disciplines	1 move: "indicating structure"	recurrent formulaic language
Kanoksilapatham (2007)	Research articles in biochemistry (60 texts)	15 moves across IMRD sections	41 grammatical features (MD analysis)
Le and Harrington (2015)	Quantitative research articles in applied linguistics (124 texts)	1 move: "Commenting on results"	word clusters, keywords
Marco (2000)	Research articles in medicine (100 texts)	various moves	collocational frameworks
Mizumoto et al. (2017)	IMRD research articles in applied linguistics (1,000 texts)	25 moves across IMRD sections	lexical bundles
Omidian et al. (2018)	Research article abstracts in 6 disciplines (5,910 texts)	5 moves in abstracts	lexical bundles
Tankó (2017)	Research article abstracts in literary studies (135 texts)	8 moves in abstracts	syntactic and lexical complexity measures

In sum, two research directions have been common: (1) investigations of a limited set of linguistic features across a range of disciplines, with a particular focus on phraseological features, or (2) investigations of multiple linguistic features in a limited generic context within a single discipline. While both approaches contribute important insights into the rhetorical and linguistic features of academic writing, neither is able to offer a broader profile of the linguistic features used to realize the full range of moves in an extensive range of academic disciplines. One method that has been used to create such a profile is multi-dimensional analysis (e.g., Kanoksilapatham 2007; Gray 2015).

1.3 Multi-dimensional analyses of moves

MD analysis (Biber 1988) uses statistical factor analysis to determine co-occurrence patterns for a wide range of linguistic features with a goal of locating systematic patterns of variation across texts. Patterns of linguistic variation are interpreted functionally on the premise that features tend to co-occur because they contribute to a shared communicative function (Conrad & Biber 2001; Biber 2019). Key to the

MD methodology is the ability to consider an extensive set of linguistic variables, thus creating comprehensive descriptions of the target domains being analyzed. For example, surveying 23 MD analysis studies, Egbert and Staples (2019) found that MD studies on average account for 60 different linguistic features. The analyses typically include features from multiple linguistic levels, including lexical (e.g., semantic classes of nouns, adjectives, and verbs), grammatical (e.g., short passives, noun + *of*-phrases, verb complement clauses, stance adverbials, modal verbs), and discourse (e.g., hedges, amplifiers).

Factor analysis is then used to identify groups of features that tend to co-occur in the texts. Each factor forms a "dimension" with a positive and a negative pole, each of which is characterized by a group of co-occurring linguistic features. These two groups of features are in complementary distribution, meaning that a text that uses high frequencies of the positive features will tend to have lower frequencies for the negative features, and vice versa. The dimensions are then interpreted functionally to characterize the variation in the corpus, and texts can be described according to each dimension (see Conrad & Biber 2001 for a full description of MD analysis).

In the past three decades, the MD methodology has been productive in providing comprehensive profiles of the use of lexico-grammatical features across published academic texts, including across registers such as textbooks, RAs, and course packs (Biber 2006; Egbert 2015; Conrad 1996) and across RAs in different disciplines (Gray 2015). However, very little work has used MD analysis to study variation *within* texts (e.g., Biber & Finegan 2001 across IMRD sections of medical RAs). Kanoksilapatham (2007) is perhaps the first study to apply MD analysis to a move-annotated corpus, analyzing 60 biochemistry RAs that had been segmented into 5,617 move instances. Kanoksilapatham's study was an important development in the alignment of linguistic and rhetorical move analysis in that it includes all IMRD sections and considered a wide range of linguistic features.

However, we are not aware of any studies that have focused on describing the linguistic characteristics of moves across (1) all sections of RAs, (2) a wide range of disciplines, and (3) a full complement of linguistic features. Understanding patterns of language use that correspond to generic moves regardless of discipline facilitates a broader understanding of academic research writing as a whole, as well as informs effective pedagogical practices in contexts that bring together students from multiple disciplines. The current study fills this gap by applying MD analysis to a corpus of 900 RAs representing 30 disciplines, each of which has been systematically annotated for moves across all sections of the articles. In the next section, we describe the corpus design and compilation, as well as the move analysis procedures.

2. The ISURA corpus: Design, compilation, and move annotation

2.1 Design and compilation

The Iowa State University Research Article (ISURA) Corpus is a pedagogical corpus containing published RAs following an IMRD/C structure from 30 academic fields of study (Table 2). The disciplinary names/groupings indicated in Table 2 reflect departments/majors at Iowa State University, making the corpus design a "localize[d] classification to a particular EAP environment" (Krishnamurthy & Kosem 2007: 363). The corpus is stratified by discipline, with 30 texts per discipline for a total of 900 RAs (4,655,464 words).

Table 2. Disciplines represented in the ISURA Corpus

Social Sciences	
	Applied Linguistics
	Art and Design
	Curriculum and Instruction
	Economics
	Business
	Psychology
	Sociology
	Special Education
Engineering	
	Agricultural and Bio-Systems Engineering
	Chemical Engineering
	Environmental Engineering
	Mechanical Engineering
Natural and Applied Sciences	
	Agronomy
	Animal Science
	Bioinformatics
	Biomedical Sciences
	Biophysics
	Food Science
	Forestry
	Geological and Atmospheric Sciences
	Horticulture
	Immunobiology
	Meteorology
	Microbiology
	Molecular, Cellular and Developmental Biology
	Physics and Astronomy
	Plant Physiology

Table 2. (*continued*)

	Synthetic Chemistry
	Urban and Regional Planning
	Veterinary Medicine

The ISURA corpus was compiled as part of a larger project to develop a genre-based automated writing evaluation tool, the *Research Writing Tutor* (Cotos 2016), which provides writing support to upper-level students learning to write IMRD/C-structured research articles in their disciplines. Because of the need for the corpus texts to serve as pedagogic models, the corpus design specifically targeted articles across a wide range of disciplines that (1) report on empirical research, (2) follow an IMRD/C structure, (3) cover a range of topics and research methods within each discipline (sampled from three to five journals per discipline), and (4) represent high-quality research writing (based on quality evaluations by subject experts). For a description of the corpus and its creation, see Cotos, Huffman & Link (2015).

The ISURA corpus is thus representative of high-quality, empirical, IMRD/C research articles, but is necessarily restricted to disciplines that tend to use this organizational structure. The corpus cannot be said to be representative of the many RAs that do not follow an IMRD/C structure (Lin & Evans 2012), and may in particular underrepresent humanities research writing and qualitative research reports (which are less likely to use this structure; see Gray 2015: Chapter 4). The extent to which non-IMRD/C-structured RAs utilize the same moves and linguistic patterns of use as identified in this corpus is beyond the scope of this project and remains a topic for further investigation.

2.2 Development of the IMRD/C move/step framework

The corpus was analyzed in four phases: move/step-schema development, trialing, validation, and analytic description (Cotos 2018). The goal was to develop a move/step framework characteristic of IMRD/C sections that would be applicable across disciplines. Phase 1 (move/step-schema development) was carried out on a sample of 150 articles, following Biber, Connor, and Upton's (2007) recommendation to use inductive analysis to determine the global and local rhetorical purposes of text segments. During this process, some initial categories were combined and some were integrated as potential descriptors of functional, rhetorical, or content realizations of moves and steps. Equalization checks (Parodi 2010) addressed conceptual duplication while consolidating move/step categories.

Once the move/step IMRD/C schemas were formulated, they were trialed through pilot coding (Phase 2), in which the move/step models and their descriptors were refined and documented. The validation component (Phase 3) sought input from subject experts who reviewed the move/step categories and offered their assessments of the clarity of the model. Equalization checks at Phase 3 addressed ambiguity and impreciseness raised by the subject experts to arrive at the final move/step framework, which is summarized in Figures 1 to 4 (see Cotos et al. 2015 for further details).[2]

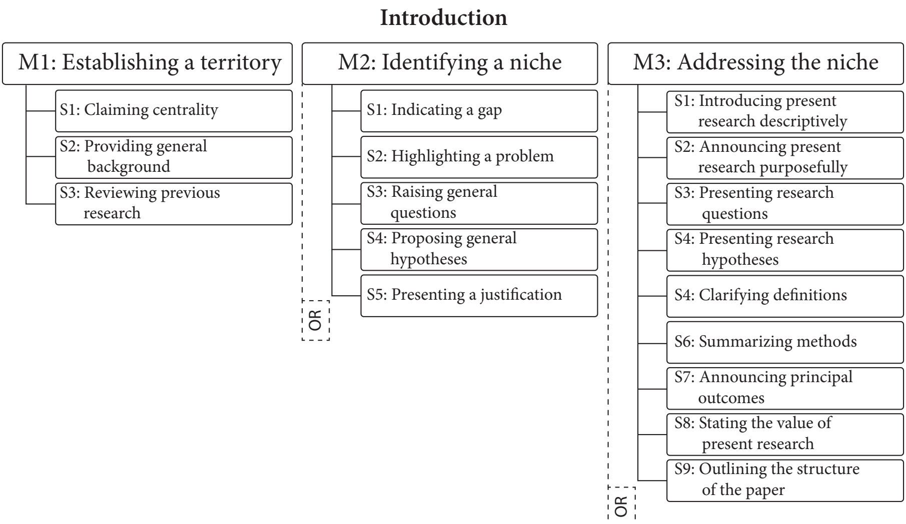

Figure 1. Move/step framework for RA introduction sections

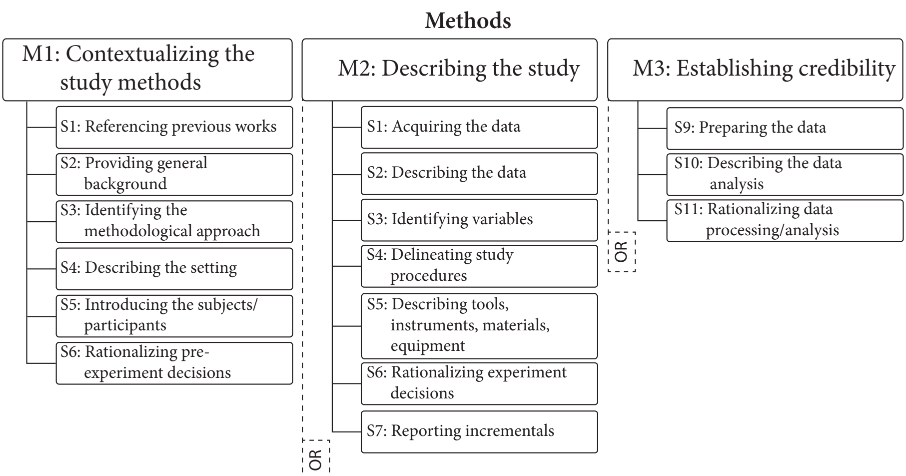

Figure 2. Move/step framework for RA methods sections

2. The moves/steps for introduction sections are based on Swales' Create-A-Research-Space model.

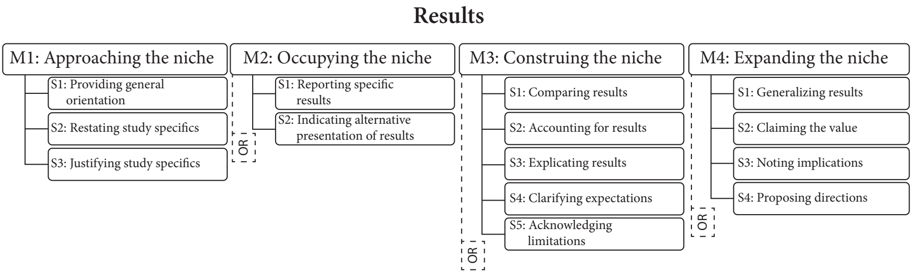

Figure 3. Move/step framework for RA results sections

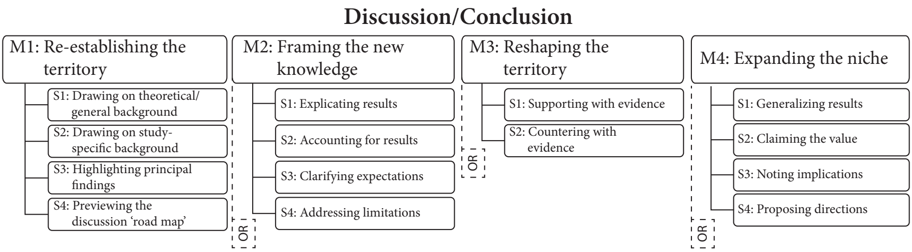

Figure 4. Move/step framework for RA discussion/conclusion sections

2.3 Annotating the ISURA corpus for moves and steps

Using a coding protocol,[3] three trained coders carried out corpus annotation (Phase 4) following the move/step framework summarized above. All sentences were annotated with XML-style coding for a move and respective step in Callisto, an open-source annotation software. Sentences that exhibited more than one rhetorical function were annotated with secondary move/step layers. To reduce subjective judgments and ensure the quality of annotated corpus data, concurrent accounts of coding reliability were conducted to monitor the level of agreement. Overall, the Intraclass Correlation Coefficient reliability estimates indicate relatively high agreement among the three annotators for moves (.86) as well as for steps (.80).

3. The protocol defined the unit of coding as a functional text segment, which could be a sentence, a clause, or a phrase contributing to the communicative intent of a functional step. It also clarified how to determine unit boundaries, how to code multifunctional text units, and what format to follow for reliability verifications.

3. Applying MD analysis to the move-annotated ISURA corpus: Methods

In this section, we describe the application of the MD methodology on the move-annotated ISURA corpus, focusing on preparing the corpus for analysis, obtaining data for a range of linguistic features, and carrying out the statistical analyses, specifically addressing the four issues established in Section 1.

3.1 Preparing the move-annotated ISURA corpus for analysis: Text segmentation and tagging

Because MD analysis relies on inferential statistical methods, quantitative data for features of interest is required for each text in a corpus ("Type B" designs in Biber & Jones 2009). Traditionally, the text (e.g., a whole RA) is the unit of observation because the goal is to describe variation across texts. In the present study, the goal is to describe variation among moves. Thus, the first challenge was to format the corpus to enable the retrieval of quantitative data at the move level (i.e., treating the move as the unit of observation).

A Python script was developed to split each RA into 14 new files corresponding to the 14 possible moves based on the XML mark-up (see Section 2). Figure 5 represents this schematically. Each of the resulting files contained all instances of a particular move from one RA. In order to maintain data independence, only the primary move was used for text segmentation (secondary moves were not considered; see Section 2.3).

Within each extracted file, we preserved the discourse structure of the original texts by inserting placeholder tags for intervening moves within that section, illustrated in Figure 6 using an example from Introduction, Move 2 (Identifying a niche). In this particular text, only the third and sixth sentences (Lines 8 and 14) were coded as Move 2; the remaining sentences are thus represented by placeholder tags indicating their move category.

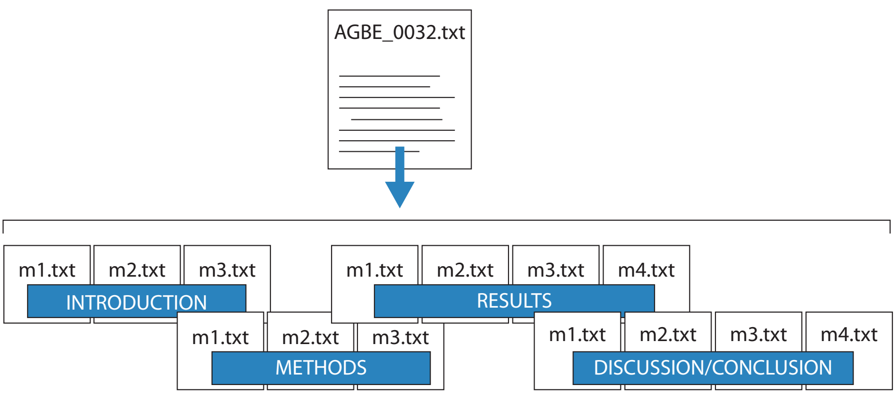

Figure 5. Schematic representation of text segmentation

```
<introductionsmlestablishingterritory>

<introductionsmlestablishingterritory>

Furthermore, technological revolutions in computer networking
allow for innovative research methodologies that have been
difficult, if not impossible, to conceptualize in the past.

<introductionsm3addressingniche>

<introductionsm3addressingniche>

This area is chosen for the known difficulty of tracking the
learner's cognition during the composing process (Miller,
2005).

<introductionsm3addressingniche>

<introductionsmlestablishingterritory>
```

Figure 6. Example segmented file for Introduction, Move 2 (Identifying a niche)

Table 3 shows the composition of the segmented corpus by move. The total number of files typically does not equal 900 (the number of RAs in the corpus) because not all RAs contained instances of each move. One consequence of segmenting the corpus in this way is that text length becomes much shorter. As will be discussed in Section 3.2, only files with 100 or more words were retained for this analysis. Thus, Table 3 documents the corpus composition when individual files with fewer than 100 words are excluded.

Table 3. Description of ISURA corpus segmented by move

Section	Move	Total no. of files	No. of files ≥ 100 words	Total words†	Mean length†	SD†
Introduction	Move 1	898	878	426,783	486.1	270.3
	Move 2	836	432	85,222	197.3	103.1
	Move 3	886	480	109,986	229.1	180.7
Methods	Move 1	815	523	216,469	413.9	368.6
	Move 2	900	895	905,448	1,011.7	591.8
	Move 3	631	355	122,393	344.8	259.0
Results	Move 1	839	688	353,378	513.6	390.0
	Move 2	900	898	858,173	955.7	651.8
	Move 3	776	536	202,270	377.4	274.7
	Move 4	271	60	10,414	173.6	110.8
Discussion	Move 1	899	879	652,429	742.2	458.8
	Move 2	872	783	366,839	468.5	302.3
	Move 3	677	253	45,152	178.5	79.3
	Move 4	843	493	149,111	302.5	226.2
Total	*All Moves*	*11,043*	*8,153*	*4,504,067*	–	–

† based on files ≥ 100 words

After text segmentation, all files were tagged for part-of-speech (POS) with the Biber tagger (Biber 1988; Biber et al. 1999; see also Biber 2019; Gray 2019). The Biber tagger annotates each word in a text with grammatical information (word form as well as some syntactic information) using probabilistic and rule-based algorithms. The accuracy of Biber tags on academic RAs has been shown to be quite high (see Gray 2015 for precision/recall rates by feature). Automatic scripts developed for academic RAs (Gray 2015) were run to correct select systematic errors. Of the 55 features analyzed for accuracy in Gray (2015), accuracy rates above 90% were achieved for 87% (precision) and 89% (recall) of the features respectively.

3.2 Obtaining rates of occurrence for linguistic features: Adapting to a specialized domain and short texts

One strength of MD analysis is the ability to simultaneously consider a wide range of features using specialized programs to generate large data files of quantitative data (per-text counts for each linguistic variable). We began with the Biber Tagcount program (see description in Gray 2019), which produces per-text normalized rates of occurrence (per 1,000 words) for over 140 linguistic variables including semantic categories of words (e.g., process nouns), grammatical structures (e.g., short passives, prepositions), and lexico-grammatical features (e.g., stance noun + *that*-clause). Notably, several features of particular importance to published academic writing have been integrated into Tagcount in recent years, such as noun + noun sequences and an expanded set of stance features.

Egbert and Staples (2019: 130) argue that principled variable selection for MD analysis is critical "because [linguistic variables] ultimately determine the success of the statistical procedure as well as the interpretability and specific interpretation of the underlying dimensions that emerge." As MD methods have been applied to increasingly specialized domains, researchers have expanded variable sets to include domain-specific features (e.g., politeness markers in call-center discourse in Friginal 2008; turn-taking patterns in university classroom talk in Csomay 2005). Based on previous literature on research writing and disciplinary variation, we developed a second program to quantify the use of additional features that may contribute to explaining variation across generic moves:

1. Linking adverbials (given their importance in marking logical relationships within a text)
2. Individual modal verbs (e.g., *can, could, may, might*, etc.)
3. Common N + preposition combinations (based on the importance of prepositional phrases as noun post-modifiers to structure/condense information in academic writing; Biber & Gray 2016).

As mentioned in Section 3.1, segmenting texts into moves results in some very short (in terms of number of words) observations. Short texts cause two problems for quantitative corpus data: normalized rates can be unreliable, particularly for rare features (as the normalized rates will be inflated when the feature does occur), and many features will not occur in short texts, thus resulting in a dataset with many zeros (which is problematic for inferential statistics) (Biber 1993; Biber & Gray 2013).

To mitigate this issue, we first excluded all texts with fewer than 100 words (Biber & Gray 2013), resulting in 8,153 texts to be included in the analysis (see Table 3). Second, we automatically excluded any variable with a frequency of 0 for more than 90% of the corpus texts.[4] Third, we ran pilot analyses comparing the use of specific individual features (e.g., separate counts for short and long passives) to the use of summed "composite" variables (e.g., all passives).[5] In most cases, piloting revealed that more general summed variables were better able to contribute to the analysis.

3.3 Conducting the statistical analyses

Following Biber (1988), Exploratory Factor Analysis (EFA) has been the dominant statistical technique in MD studies in order to prioritize patterns of covariance (see discussion in Egbert & Staples 2019). Because of the challenges resulting from shorter texts and a narrower target domain, we used Principal Components Analysis (PCA), a related data-reduction procedure. EFA and PCA are similar in many respects, but EFA focuses only on covariance and excludes the error variance, while PCA considers the total variance (Tabachnick & Fidell 2007). This statistical difference may mean that PCA is better-suited to data with many zeroes (Gary Ockey, personal communication). Piloting revealed that PCA and EFA accounted for a similar amount of variance in the data, with similar dimension structures. However, the PCA results were more interpretable, and thus were selected for the final analysis.

Forty-five features with communalities $\geq .20$ and component loadings $\geq .3$ on at least one component were retained for the analysis (list available upon request). Preliminary analyses confirm the factorability of the corpus, with the Kaiser-Olkin Measure of Sampling Adequacy (KMO = .664, mediocre) meeting the minimum requirements. Barlett's Test for Sphericity (approximate Chi-Square = 113,372.16,

4. Since the goal of MD analysis is to locate patterns of variation across the texts, a feature that does not occur in most texts cannot vary across texts, thus indicating that it is not important for modeling variation in the corpus.

5. Summed counts and individual counts for a particular feature or set of features should not be included in the same analysis to avoid multicollinearity, and piloting is commonly used to determine which type of variable to retain (Egbert & Staples 2019).

df = 990, $p < .000$) was significant, indicating sufficient correlations in the dataset (Tabachnick & Fidell 2007). Based on an examination of the scree plot (Appendix A) and a comparison of four- and five-component solutions, the four-component solution was selected as the most interpretable, accounting for 30.45% of the variance in the corpus. Table 4 lists the linguistic structure of each component (based on features with component loadings of ≥ ± .3). Positive-loading features tend to co-occur together, while negative-loading features also co-occur but in complementary distribution to the positive features.

Table 4. Summary of component structure: Features with loadings ≥ ± .3 (features that load on multiple factors are italicized on secondary components)

Component 1 – Interpretation and expansion vs. simple reportage		
Positive-loading features	Verbs:	verb *be* (.59)
		present tense (.33)
	Modals:	possibility modals (.64)
		prediction modals (.41)
		necessity modals (.35)
	Adverb(ial)s:	linking adverbials (.48)
		adverbs (.47)
	Stance:	adjective-based stance (lexical and grammatical) (.49)
	Adjectives:	predicative adjectives (.47)
		attributive adjectives (.36)
	Other:	pronoun *it* (.45)
Negative-loading features	Verbs:	past tense (–.71)
		passive voice (–.50)
	Nouns:	common nouns (–.36)
		proper nouns (–.36)
Component 2 – Abstraction / overt empiricism		
Positive-loading features	Verbs:	*present tense (.33)*
		infinitives (.31)
	Stance:	noun-based stance (lexical and grammatical) (.53)
		verb-based lexical stance (.45)
	Adjectives:	attributive adjectives (.39)
	Noun Structures:	nominalizations (.72)
		frequent N + prep combinations (.43)
	Noun Types:	process nouns (.51)
		abstract nouns (.51)
		cognition nouns (.48)
		human nouns(.31)
	Other:	word length (.65)
Negative-loading features	Verbs:	*passive voice (–.38)*
		past tense (–.31)
	Nouns:	proper nouns (–.54)

Component 3 – Procedural narration		
Positive-loading features	Verbs:	all verbs (.80)
		activity verbs (.50)
		mental verbs (.45)
		communication verbs (.32)
	Verb Phrase:	infinitive verbs (.61)
		progressive aspect (.43)
	Stance:	stance verb + *to*-clause (.42)
	Clauses:	verb + *wh*-clause (.31)
	Other:	*3rd person pronouns (.40)*
Negative-loading features	Prepositional phrases:	prepositions (–.53)
		N + *of* phrase (0.32)
Component 4 – Interpreting results vs. informational density		
Positive-loading features	Verbs:	communication verbs (.41)
		all verbs (.42)
		mental verbs (.39)
	Verb Phrase:	*present tense (.31)*
	Stance:	verb-based lexical stance (.47)
		communication verb + *that*-clause (.40)
		factive verb + *that*-clause (.31)
		noun-based stance total (.38)
	Pronouns:	3rd person pronouns (.41)
		demonstrative pronouns (.32)
	Adverb(ial)s:	*adverbs (–.30)*
		linking adverbials (.30)
Negative-loading features	Nouns:	common nouns (–.70)
		N + N sequences (–.67)
		concrete nouns (–.46)

3.4 Calculating dimension scores and interpreting dimensions

Dimension scores were calculated for each text in the corpus by taking standardized z-scores for each variable listed in Table 4, summing the scores for the positive features, summing the scores for the negative features, and taking the difference between the two (Egbert & Staples 2019; Conrad & Biber 2001). Features loading on multiple components are listed for interpretation but are only included in the dimension score for the component with the highest loading value (Egbert & Staples 2019). Dimension scores characterize to what extent the texts utilize the features associated with the dimensions with higher scores indicating the prevalence of positive-loading features and lower dimension scores indicating the prevalence of negative-loading features. Each move type is characterized along the dimensions by taking the mean dimension score for texts representing that move.

Functional interpretations (reflected in descriptive labels) of the dimensions take into account the shared communicative functions of the features loading on each dimension along with how they function in text excerpts and how the sub-corpora are distributed along the dimension (Conrad & Biber 2001). In Section 4, we briefly present our interpretation of the four dimensions and the distribution of moves along them to demonstrate the useful synergies of these two approaches, while also evaluating to what extent our methodological adaptations result in a meaningful multi-dimensional description of moves in research writing.

4. Dimensions of variation across rhetorical moves

One-way ANOVAs reveal significant differences (reported alongside Figures 7–10) between moves for all four dimensions, with R^2 values indicating that 10% to 52% of the variance in dimension scores can be accounted for by the move category. The four dimensions are interpreted as follows:

Dimension 1. Interpretation and Expansion vs. Simple Reportage
Dimension 2. Abstraction / Overt Empiricism
Dimension 3. Procedural Narration
Dimension 4. Interpreting Results vs. Informational Density

In this section, we point out key features, their shared functions, and how those relate to the distribution of moves along the dimensions. Throughout, readers are directed to Table 4 for the complete listing of features that contributed to each dimension.

4.1 Dimension 1: Interpretation and expansion vs. simple reportage

Dimension 1 is posited to reflect interpretation and expansion on the one hand and simple reportage on the other. Positive features (see Table 4) on this dimension contribute to involved discourse in which topics are defined, characterized, and evaluated (e.g., verb *be*, present tense, predicative and attributive adjectives, adjective-based stance), relationships are explicitly stated (linking adverbials), and possibilities, predictions, and necessities are put forward (modal verbs).[6]

Example (1) illustrates these features in an excerpt from a Discussion, Move 2 (framing new knowledge) in chemical engineering. Positive features are bolded,

6. This interpretation is supported by previous MD models that revealed similar co-occurrence patterns in academic writing, including Dimension 1 (academic involvement and elaboration) in Gray (2015) and Dimension 2 (definition and evaluation of new concepts) in Egbert (2015).

illustrating how they function to take a topic and explicitly label conclusions, interpretations, and evaluations, as well as state relationships between propositions.

(1) **Consequently**, **it is conceivable** that a **similar** acyl intermediate **may be readily** formed on Pd under our reaction conditions. <discconcl_m1_reestablishing territory>
Thus, in this case, **it is** more **likely** that the hydroxyalkyl species **is** the **dominant** intermediate in **furfural** hydrogenation. **Therefore**, a mechanism for **furfural** conversion over Pd catalysts **can** be proposed, as shown in Scheme 4. **In contrast** to furfural, methyl pentanal **does** not have a **conjugated** system and **therefore** formation of a hydroxyalkyl species **would** not **be favorable**.

Negative features point to simple reportage in which actions are described (but not evaluated) through the use of past tense and passive voice. Common and proper nouns have been associated with informational density (see Gray 2015, Dimension 1); however, this does not seem to fully explain the co-occurrence patterns here. While common and proper nouns loaded on the negative pole of this dimension, other features that typically combine with nouns to create nominal density are not present (e.g., attributive adjectives, prepositions). Instead, the co-occurrence of these nouns with past tense and passive voice creates discourse in which nouns represent basic concepts/entities, and past tense and passive voice function to report what was done with those concepts/entities. In a sense, this dimension may indirectly reflect agency, being the strongest indicator of researchers as actors in the research process (despite the fact that the agents are not explicitly specified).[7]

Example (2), which comes from the Methods section of an agronomy article (Move 2, describing the study), illustrates how these negative features create simple reportage (negative features are bolded, with passives underlined):

(2) **Experiments** **were performed** using C. pepo L. **seeds** that were routinely **purchased** from **Nunhems Zaden** BV (**Haelen**, The **Netherlands**).
Seeds **were germinated** in **Petri dishes** on wet **glass fiber filter paper** and **covered** with **aluminum foil**, to exclude **light**, at 20 **degrees** C for 7 **days**.

The distribution of moves along Dimension 1 (Figure 7) further supports these interpretations. Moves with the highest positive scores on this dimension include Discussion and Results moves associated with generalizing, stating the value and implications of findings, and explicating results: Discussion Moves 2 (framing the new knowledge) and 4 (establishing additional territory); Results Moves 3 (construing the niche) and 4 (expanding the niche). Move 2 from Introductions (identifying the niche) also has a relatively high positive Dimension 1 score; this move

7. Our thanks to Doug Biber, who pointed this interpretation out to us.

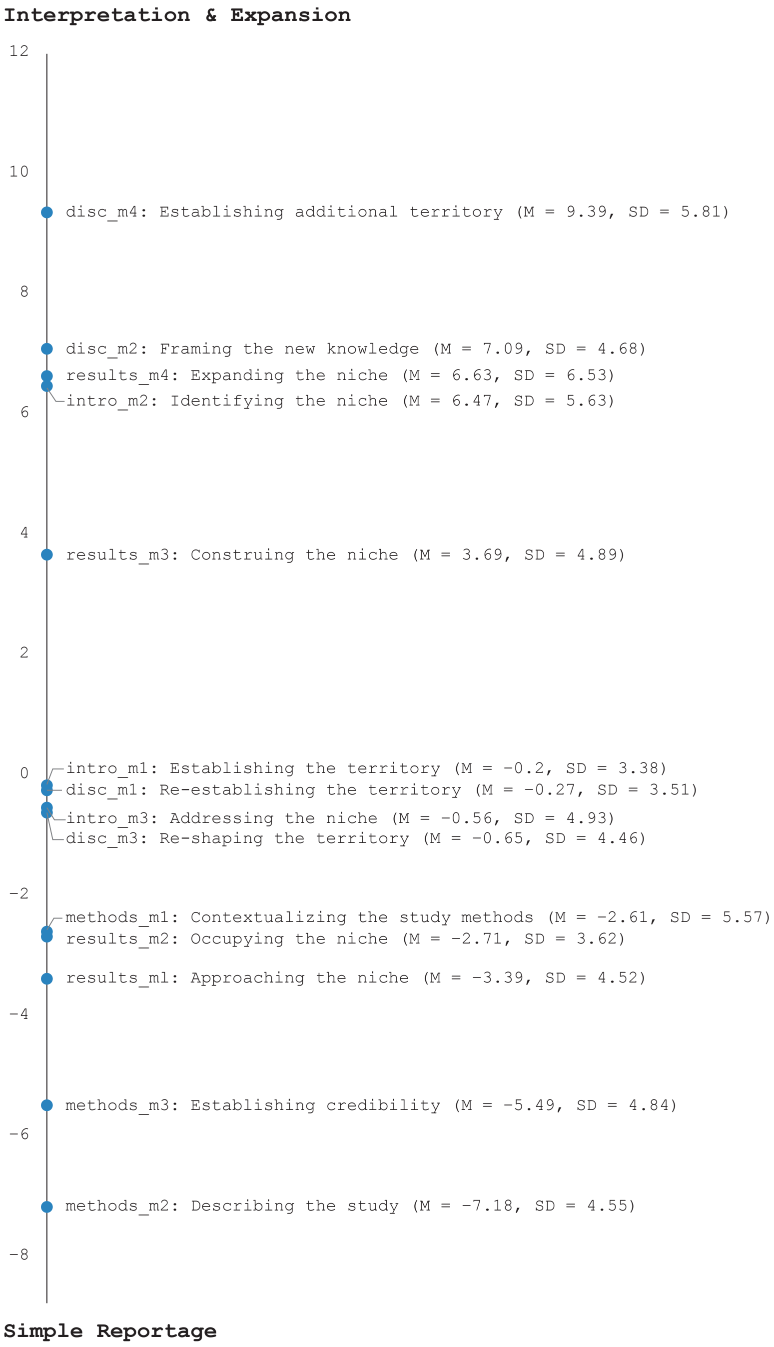

Figure 7. Distribution of moves along Dimension 1: Interpretation and Expansion vs. Simple Reportage ($F = 689.7$, d.f. $= 13$, $p < .000$, $R^2 = .52$)

in particular functions to locate a gap, highlight a problem, propose hypotheses, and justify why the reported research fills that gap (see the steps listed in Figure 1). These rhetorical functions call for language that not only provides definitions and characterizations, but which also evaluates and elaborates concepts, research gaps, and findings.

In contrast, moves with the lowest negative scores on Dimension 1 include all three Methods moves in which negative Dimension 1 features are used to succinctly report data, data collection, procedures, and analysis methods – reporting what was done to carry out the study. Two Results moves also rely extensively on the negative Dimension 1 features. Results Move 1 (approaching the niche) involves a general restatement of the study methods or topic, not dissimilar to the Methods moves. However, Results Move 2 represents a shift away from reporting methods/procedures, as simple reportage is used to present the specific results of the study. Importantly, the function of this move is to deliver the study results objectively, stating the facts (not the interpretations or meanings of those results). This is in direct contrast to Results, Move 3 (construing the niche), in which those findings are compared, explicated, clarified, and explained. It is particularly noteworthy that Dimension 1 differentiates these two Results moves so clearly (Move 3 has a high positive Dimension 1 score, while Move 2 has a negative score), demonstrating that MD analysis is an appropriate method for identifying function-form mappings.

4.2 Dimension 2: Abstraction / Overt empiricism

Dimension 2 reflects abstraction and overt empiricism. Abstract, process, and cognition nouns, longer words, and nominalizations serve to package abstract concepts into empirical pursuits, placing them into nominal slots that can be modified and expanded upon (e.g., through attributive adjectives, N + prep constructions).[8] Example (3) comes from an Introduction, Move 2 (identifying the niche) in business. Abstract, process and cognition nouns are **bolded**, nominalizations are underlined, attributive adjectives are *italicized*, and common noun + prep combinations are double-underlined. Placing these abstract, process, and cognition concepts into nominal units (often nominalized forms) overtly frames concepts in technical ways that can be investigated empirically.

8. Like Dimension 1, the interpretation of Dimension 2 is further supported by Gray's (2015) Dimension 4 (academese or overt empiricism) and Egbert's (2015) Dimension 5 (abstract observation).

(3) One **issue** in the **debate** was raised by REF, who suggested that the **relationships** between some personality **traits** and **job performance** may not be linear, as generally assumed in the literature. Because the linearity **assumption** of predictor-criterion **relationships** underlies the *common* **practice** of *top-down* **selection** in **employment selection** and is the basis of utility **analysis** (REFS)…

No interpretation for the negative pole of Dimension 2 is proposed, as only three features loaded and two of those had higher loadings on another dimension and were thus not used to calculate the dimension score. Thus, negative Dimension 2 scores indicate a below-the-mean use of the positive features, rather than a frequent use of the negative features. To give a brief example, consider Example (4) from Results, Move 2 (occupying the niche). The negative Dimension 2 score reflects a lack of use of the positive features for this dimension.

(4) In both miSmad1 lines tested, chromosome number did not deviate from the expected 40, and karyotypes were indistinguishable on a gross *morphologic* **level** compared with the AinV18 *parental* cells…

Figure 8 displays the distribution of moves along Dimension 2. The highest positive dimension score occurs for Discussion, Move 4 (establishing additional territory), followed by Introduction, Moves 2 (identifying the niche) and 3 (addressing the niche). The purpose of Discussion, Move 4 is to generalize and claim the value of results, state implications, and propose new directions. To accomplish these functions, the concepts, entities, and processes under investigation are generalized into nominal units in order to position them as entities to be evaluated and to serve as topics in future research, as illustrated in Example (5).

(5) The **use** of seed size facilitates the **interpretation** of the **patterns** found for grazing and climate, and makes them testable in other regions with *different* floras. *Future* **research** on Mediterranean grasslands should include other *grazing-dependent* plant **traits**, apart from seed size (REF) and combine *spatial* **variability** with *interannual* **variability**…

Introduction, Moves 2 (identifying the niche) and 3 (addressing the nice) are associated with announcing the purpose and topic of the research, raising gaps/problems, and justifying the research. The high positive Dimension 2 scores for these Introduction moves shows these features being used to establish abstract concepts and issues in technical ways in order to position them for empirical inquiry.

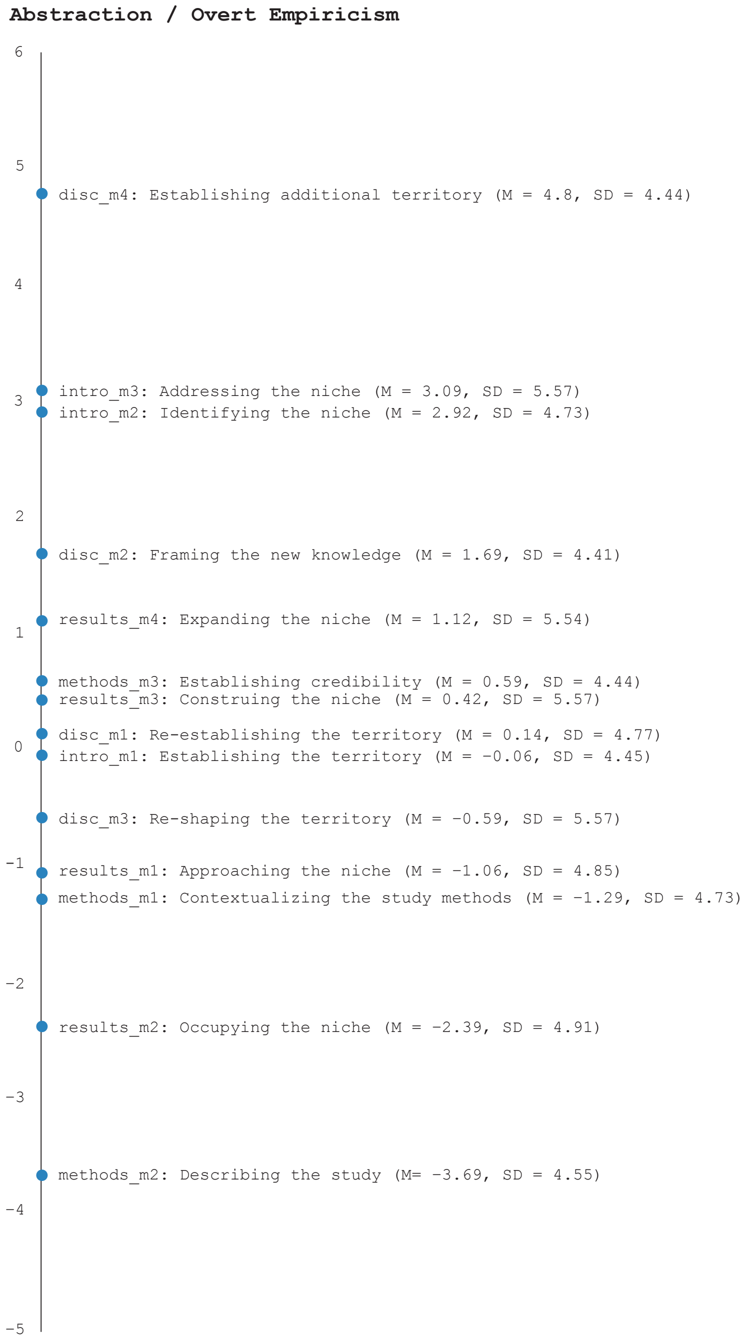

Figure 8. Distribution of moves along Dimension 2: Abstraction / Overt Empiricism ($F = 127.29$, d.f. = 13, $p < .000$, $R^2 = .18$)

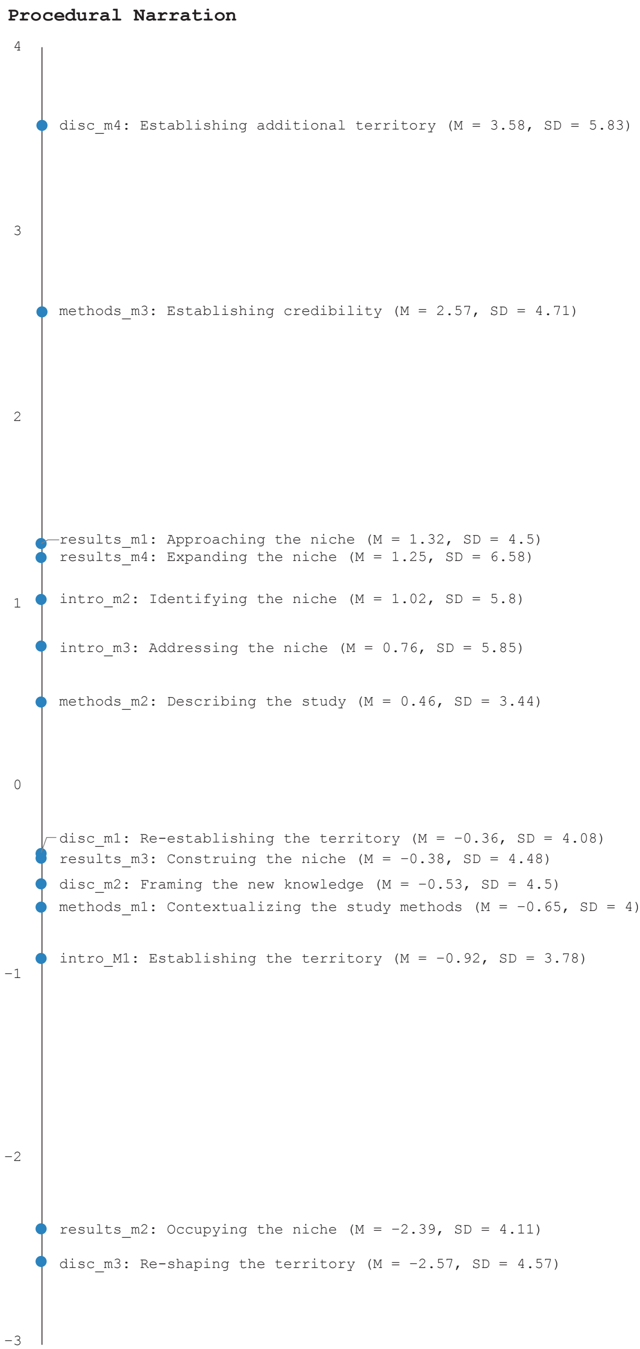

Figure 9. Distribution of moves along Dimension 3: Procedural Narration ($F = 72.39$, d.f. = 13, $p < .000$, $R^2 = .10$)

4.3 Dimension 3: Procedural narration

On first examination, the interpretation for Dimension 3, labeled "Procedural Narration," seems quite clear. Positive-loading features include many verb and verb-phrase features, including activity, mental, and communication verbs, as well as total verbs. Other verbal features include progressives and infinitives, stance verb + *to*-clauses, and verb + *wh*-clauses. The dimension seems to be capturing the narration of study procedures.[9]

However, in looking at the distribution of moves along this dimension, and examining text excerpts, this interpretation is less clear. The two moves with the highest positive Dimension 3 scores (Figure 9) are Discussion, Move 4 (establishing additional territory) and Methods, Move 3 (establishing credibility). Methods, Move 3 makes sense with narration, as this move describes and rationalizes data preparation and analysis. Example (6) illustrates Move 3 in agronomy (positive features in bold, with infinitives and progressives underlined). Verbs (particularly activity and mental verbs) are used to describe procedures, *ing*-clauses are using to provide the method for carrying out an action, and *to*-clauses are used to provide the reasons for procedures.

(6) Data **were analyzed** separately for each location **<u>using</u>** the SAS Proc Mixed procedure (REFS) with treatment and site as fixed effects and year and replication as random effects. Means **were compared** **<u>using</u>** the SAS least square means with Tukey adjustment at $P = 0.05$ **<u>to test</u>** comparisons across means.

Yet other Methods moves where we might expect to find procedural narration have low positive or even negative dimension scores. The interpretation is further complicated when looking at text excerpts from Discussion, Move 4 (establishing additional territory), which generalizes results, notes implications, and proposes new directions. It is difficult to see a dense use of Dimension 3 features in text excerpts from this move. In examining the text excerpts, it seems possible that the distribution of moves may be the result of shorter text lengths. Discussion, Move 4 occurs in 94% of the RAs in the corpus, but only 493 (55%) of those contain at least 100 words and are included in the analysis. Of those included in the analysis, the mean text length is 302.5 words (but with a high SD of 226.2). Thus, it may be that even a single occurrence of some of the Dimension 3 features may be resulting in inflated normalized counts, thus leading to high z-scores and subsequently high dimension scores. Thus, the difficulty in explaining the distribution of moves along

9. This interpretation of Dimension 3 is also supported by its alignment with previous MD models of disciplinary writing, sharing features with Gray's (2015) Dimension 2 (contextualized narration) and Egbert's (2015) Dimension 4 (colloquial narration).

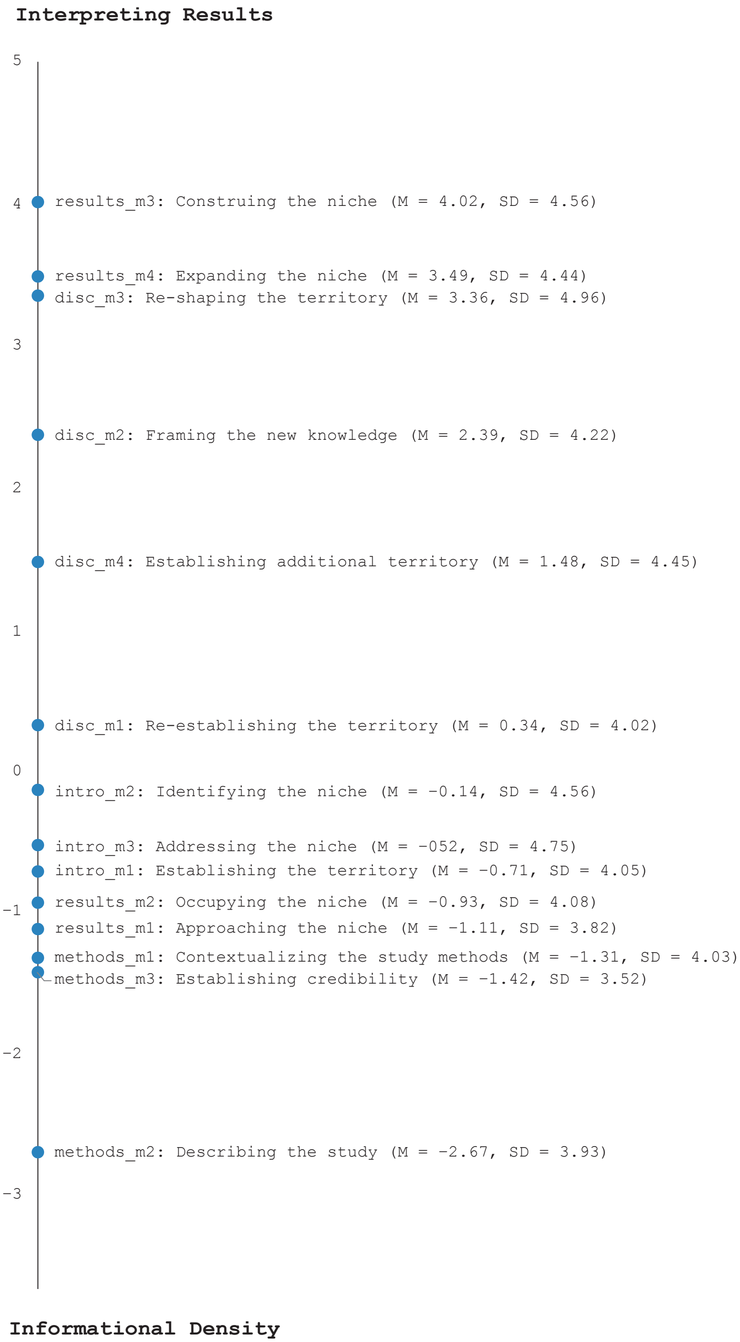

Figure 10. Distribution of moves along Dimension 4: Interpreting Results vs. Informational Density ($F = 123.7$, d.f. = 13, $p < .000$, $R^2 = .17$)

this dimension may indicate that the measures we have taken to address the issue of text length may not be sufficient.

Only two features loaded on the negative pole of Dimension 3, prepositions and noun + *of*-phrases. Such features have been associated with nominal packaging and noun-phrase density, which can be seen in Example (7) (biomedical sciences, Results, Move 2). However, with only two features, an interpretative label is not recommended.

(7) **Analysis** of hematopoietic cell lineages **in** the secondary lymphoid organs of miR-146a-null mice **by** flow cytometry revealed massive myeloproliferation (~10-fold increase **in** the number of CD11b cells), whereas no significant change **in** the absolute number of B and T cells was found (FIGS).

4.4 Dimension 4: Interpreting results vs. informational density

Positive features on Dimension 4 include communication and mental verbs and present tense. More stance features load here than any other dimension, including verb- and noun-based lexical stance along with communication and factive verbs + *that*-clauses. Third person and demonstrative pronouns, adverbs, and linking adverbials also load positively. Taken together with the distribution of moves (Figure 10) along this dimension (which shows the frequent use of these features in multiple Results and Discussion moves), the positive pole is labeled "Interpreting Results". Example (8) illustrates the use of these features in a Results, Move 3 (construing the niche) in sociology, where the focus is on comparing, accounting for, and explicating results, as well as clarifying expectations and acknowledging limitations. *That*-clauses report conclusions and findings, and provide the means for the authors to announce their observations and statements of limitations.

(8) **Again**, we **note** that caution **is** in order due to the relatively small number of cases **involved**…We **find** that the two-way interaction between centrality and **having** multiple cross-gender friendships **is** significant and positive in the high-segregation schools for both forms of aggression, as we **hypothesized** (**see** TABLE in the online supplement)…**This confirms** that having multiple cross-gender friendships **is** central to the concept of a gender bridge.

Results, Moves 4 (expanding the niche) and 3 (construing the niche), and Discussion, Moves 3 (reshaping the territory) and 2 (framing the new knowledge) also work toward the function of interpreting and explaining results discursively.

An important validation of this dimension is the fact that related moves in different sections have similar dimension scores. For example, the steps for Results, Move 4 and Discussion, Move 2 (see Figures 3–4) are functionally quite similar (accounting for results, explicating results, clarifying expectations, and acknowledging limitations). Thus, the MD methodology is able to capture linguistic similarities that can be clearly attributed to communicative function. These types of findings are particularly important for pedagogical contexts, where similarities in the moves and in the linguistic structures can be taught efficiently.

On the negative pole of Dimension 4, three features loaded, including common nouns, N + N sequences, and concrete nouns. These features have been widely associated with nominal density and the packaging of information into noun phrases (Biber & Gray 2016). Figure 10 shows that Methods and Results moves that focus on factual statements without elaboration have the lowest negative scores: Methods, Moves 2 (describing the study), 3 (establishing credibility), and 1 (contextualizing the study methods), along with Results, Moves 1 (approaching the niche) and 2 (occupying the niche).

Example (9) (Methods, Move 2 in special education) shows common nouns, N + N sequences, and concrete nouns being used to specify the participants and entities being studied. N + N sequences concisely convey precision and specificity. For example, the noun phrase "student's special education classification" conveys a great deal of information in just four words, clearly identifying what the "classification" is for, and who receives it. Using alternative structures to express the same information is much less efficient (c.f., "classifications given to students regarding their special education needs).

(9) All **participants** received special **education** reading **instruction** in **resource rooms**, except 3 who were served in a self-contained **classroom**. A **researcher** met individually with each participating **teacher** to obtain **information** on **students'** special **education classification**, **language status**, and **race/ethnicity**.

5. Applying and evaluating the MD analysis of moves

The MD analysis of moves in IMRD/C-structured RAs in 30 disciplines has shown the potential of a paired approach to provide rich and nuanced descriptions of language use, mapping the communicative functions of text segments with patterns of functional linguistic variation through these two complementary methodologies. The MD analysis enabled us to consider a wide range of features to arrive at comprehensive linguistic profiles of these moves. Without the data reduction techniques of MD analysis and the focus on functional interpretations, it would be difficult to make sense of patterns of use for so many features or to apply these descriptions to genre-based pedagogy (discussed in Section 5.2).

On the other hand, the move analysis also contributed meaningfully to the interpretation of the MD analysis. Since the interpretation stage of an MD analysis seeks to identify the communicative functions underlying patterns of co-occurring features, pairing interpretations of the functions of features with the distribution of moves on the dimensions both facilitates and validates the interpretations. That is, because the moves represent segments of text grouped according to their communicative functions, interpretations of the shared functions of co-occurring linguistic features that are common in those moves can become more transparent and triangulated.

In the remainder of the chapter, we return to the methodological challenges that accompany the pairing of MD analysis and move analysis, evaluating to what extent our adaptations resulted in a meaningful MD solution and identifying issues that may benefit from further methodological experimentation. We then briefly consider the potential applications and implications of findings such as these, demonstrating why continued efforts to customize the MD methodology for this type of corpus data are a worthwhile endeavor.

5.1 Evaluation of the dimensions and process

One of our goals in this project is to critically assess whether our adjustments to an established methodology (MD analysis) were successful and sufficient in addressing the issues that motivated those adaptations. As the analysis above showed, the MD analysis was successful in that the dimensions clearly differentiated among move types, and clear functional interpretations were possible for each of the four dimensions. However, as we considered the structure/composition of the dimensions and distribution of moves along them, it became apparent that further experimentation may help to optimize the MD-analysis method for use with move-annotated corpus data.

For example, we believe that MD analysis has the potential to better account for the variance in the ISURA corpus, particularly given the specialized domain of texts being analyzed. Our analysis accounted for 30.45% of the variance, which is below average for MD analysis studies (Based on a review of 23 MD studies, Egbert & Staples 2019: 132 report that on average, MD studies accounted for 38% of the variance). Another less-than-ideal outcome from the MD analysis was the degree of overlap between features that loaded on the dimensions. Although we minimized the impact of multiple loadings by only using a feature to calculate the dimension score for the component with the highest loading, this resulted in fewer features per dimension than other MD analyses we have conducted. Caution should be used when interpreting dimensions with multiple overlapping features. A second issue related to interpretation is problematic: explaining why Discussion, Move 4 has the highest positive score on Dimensions 1–3, despite a seeming mismatch in the purpose of that move and the functional interpretations of some dimensions.

Two additional outcomes were unexpected given our experience with previous MD analyses: the lack of interpretable negative poles on multiple dimensions, and the need to use primarily more general, composite variables. While there are established, robust MD analyses with a dimension with insufficient negative-loading features to create a functional interpretation, we have not seen many studies in which this was the case on multiple dimensions. One hypothesis is that this occurred due to the narrower domain represented by the corpus. Because of this narrower domain, we originally predicted that more specific linguistic variables would best be able to capture variation across these texts (as the variation might be more nuanced). For example, we generated counts for individual modal verbs (e.g., *may, might, could*, etc.), with the thought that particular modals may be related to the more specific rhetorical functions of moves. Instead, it seems that text length trumped specificity: shorter texts combined with more specific features simply led to more zeros in the data, which decreased the ability of the statistical procedures to work properly.

Thus, we believe that further work is needed to experiment with possible methodological adaptations to better align MD methods with the nature of move-annotated data, particularly with respect to text length. For example, if text length were not an issue, we would not lose data in having to exclude texts < 100 words, which would result in better representation of less common and/or shorter moves. One possibility is to use Multiple Correspondence Analysis (MCA), which is a data-reduction technique that works with nominal data. Clark and Grieve (2017) have applied MCA to MD analyses of Twitter data, in which normalized rates of occurrence are not reliable due to extremely short texts. In their study, they record linguistic data as presence/absence: any feature with a raw frequency of 0 is "absent," while any feature with a raw frequency of 1 or more is "present" in a text. By turning frequencies into "presence" vs. "absence," the use of a feature is converted into a nominal scale. MCA is then used in a similar way to EFA or PCA to generate dimensions of variation. This new adaptation of MD analysis may work well for move-annotated data, allowing us to keep all observations in our corpus, and to have increased specificity in the features included in the analysis. Perhaps most importantly, it may help to ensure that dimension scores are more reliable and not unduly impacted by issues of text length.

5.2 Implications and applications

Despite the limitations discussed above, the analysis demonstrates the potential applicability of pairing these methods to reveal richer, more nuanced descriptions of language that have clear pedagogical implications. Using dimensions, we can create linguistic profiles for each move type. For example, given the distributions of the moves along the dimensions, we might characterize Introduction, Move 2 (identifying the niche) as having the following linguistic profile (where + indicates highly positive dimension scores, – indicates low negative scores, and +– indicates a neutral dimension score):

+ Interpretation & Expansion (Dim 1)	+ Abstract/Overt Empiricism (Dim 2)
– Procedural Narration (Dim 3)	+– Nominal Density (Dim 4)

Profiles can then be used to help novice research writers notice discourse patterns in their own fields and to make connections between what they want to do with their writing and the linguistic resources available to carry out those functions. A further application of these findings arises from the comparative nature of MD analysis, which enables us to not only locate how moves are different from one another linguistically but also to identify different moves that have shared linguistic patterns that can typically be traced back to shared/similar rhetorical functions.

At the most basic level, descriptions of the linguistic realizations of rhetorical moves in published RAs can be applied in advanced EAP contexts to the teaching of reading and writing of academic research. However, a range of additional applications of this type of knowledge are also possible, including applied natural language processing tools for scientific writing (e.g., *AWSuM* autosuggests most frequent lexical bundles related to a particular move in applied linguistics articles, see Mizumoto et al. 2017) and automated writing analysis tools (e.g., *Mover* automatically identifies the moves in abstracts for information technology RAs, see Anthony & Lashkia 2003).

Once the MD method has been further refined to address the issues raised in this chapter, we anticipate that the results of the final MD analysis will find immediate application in the development of a new form of feedback by the *Research Writing Tutor*, a writing platform with an automated writing evaluation module that generates move/step-based and discipline-specific feedback on all the sections of the research article in 30 disciplines (Cotos 2016). By integrating knowledge of the linguistic profiles of moves from the MD analysis, we hope to enable more accurate and more precise feedback for writers.

Acknowledgements

We would like to thank Sowmya Vajjala for her assistance in developing the Python script to segment the texts.

References

Anthony, Laurence & Lashkia, George V. 2003. Mover: A machine learning tool to assist in the reading and writing of technical papers. *IEEE Transactions on Professional Communication* 46: 185–193. https://doi.org/10.1109/TPC.2003.816789

Biber, Douglas. 1988. *Variation across Speech and Writing*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511621024

Biber, Douglas. 1993. Representativeness in corpus design. *Literary and Linguistic Computing* 8(4): 243–257. https://doi.org/10.1093/llc/8.4.243

Biber, Douglas. 2006. *University Language: A Corpus-Based Study of Spoken and Written Registers* [Studies in Corpus Linguistics 23]. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.23

Biber, Douglas. 2019. Multi-dimensional analysis: A historical synopsis. In *Multi-Dimensional Analysis: Research Methods and Current Issues*, Tony Berber Sardinha & Marcia Veirano Pinto (eds), 11–26. London: Bloomsbury.

Biber, Douglas, Connor, Ulla & Upton, Thomas A. 2007. *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure* [Studies in Corpus Linguistics 28]. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.28

Biber, Douglas & Finegan, Edward. 2001. Intra-textual variation within medical research articles. In *Variation in English: Multi-dimensional Studies*, Susan Conrad & Douglas Biber (eds), 108–137. London: Longman.

Biber, Douglas & Gray, Bethany. 2013. *Discourse Characteristics of Writing and Speaking Task Types on the TOEFL iBT Test: A Lexico-Grammatical Analysis*. TOEFL iBT Research Report (TOEFL iBT-19). Princeton NJ: Educational Testing Service.

Biber, Douglas & Gray, Bethany. 2016. *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511920776

Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Biber, Douglas & Jones, James. 2009. Quantitative methods in corpus linguistics. In *Corpus Linguistics: An International Handbook*, Anke Lüdeling & Merja Kytö (eds), 1286–1304. Berlin: Walter de Gruyter.

Clark, Isobelle & Grieve, Jack. 2017. Dimensions of abusive language on Twitter. *Proceedings of the First Workshop on Abusive Language Online* 1: 1–10. https://doi.org/10.18653/v1/W17-3001

Conrad, Susan & Biber, Douglas. 2001. *Variation in English: Multi-dimensional Studies*. London: Longman.

Conrad, Susan. 1996. Investigating academic texts with corpus-based techniques: An example from biology. *Linguistics and Education* 8: 299–326. https://doi.org/10.1016/S0898-5898(96)90025-X

Cortes, Viviana. 2013. The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes* 12: 33–43. https://doi.org/10.1016/j.jeap.2012.11.002

Cotos, Elena, Huffman, Sarah & Link, Stephanie. 2015. Furthering and applying move/step constructs: Technology-driven marshalling of Swalesian genre theory for EAP pedagogy. *Journal of English for Academic Purposes* 19: 52–72. https://doi.org/10.1016/j.jeap.2015.05.004

Cotos, Elena. 2016. Computer-assisted research writing in the disciplines. In *Adaptive Educational Technologies for Literacy Instruction*, Scott Crossley & Danielle McNamara (eds), 225–242. New York NY: Routledge.

Cotos, Elena. 2018. Move analysis. In *Encyclopedia of Applied Linguistics*, Carol A. Chapelle (ed.). Oxford: Wiley-Blackwell.

Cotos, Elena, Huffman, Sarah & Link, Stephanie. 2017. A Move/Step model for Methods sections: Demonstrating rigour and credibility. *English for Specific Purposes* 46: 90–106. https://doi.org/10.1016/j.esp.2017.01.001

Csomay, Eniko. 2005. Linguistic variation within university classroom talk: A corpus-based perspective. *Linguistics and Education* 15: 243–274. https://doi.org/10.1016/j.linged.2005.03.001

Durrant, Philip & Mathews-Aydınlı, Julie. 2011. A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes* 30: 58–72. https://doi.org/10.1016/j.esp.2010.05.002

Egbert, Jesse. 2015. Publication type and discipline variation in published academic writing: Investigating statistical interaction in corpus data. *International Journal of Corpus Linguistics* 20(1): 1–29. https://doi.org/10.1075/ijcl.20.1.01egb

Egbert, Jesse & Staples, Shelley. 2019. Doing multi-dimensional analysis in SPSS, SAS, and R. In *Multi-Dimensional Analysis: Research Methods and Current Issues*, Tony Berber Sardinha & Marcia Veirano Pinto (eds), 125–144. London: Bloomsbury.

Friginal, Eric. 2008. Linguistic variation in the discourse of outsourced call centers. *Discourse Studies* 10(6): 715–736. https://doi.org/10.1177/1461445608096570

Gray, Bethany. 2015. *Linguistic Variation in Research Articles: When Discipline Tells Only Part of the Story* [Studies in Corpus Linguistics 71]. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.71

Gray, Bethany. 2019. Tagging and counting linguistic features for multi-dimensional analysis. In *Multi-Dimensional Analysis: Research Methods and Current Issues*, Tony Berber Sardinha & Marcia Veirano Pinto (eds), 43–66. London: Bloomsbury.

Kanoksilapatham, Budsaba. 2007. Rhetorical moves in biochemistry research articles. In *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure* [Studies in Corpus Linguistics 28], Douglas Biber, Ulla Connor & Thomas Upton (eds), 73–120. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.28.06kan

Krishnamurthy, Ramesh & Kosem, Iztok. 2007. Issues in creating a corpus for EAP pedagogy and research. *Journal of English for Academic Purposes* 6(4): 356–373. https://doi.org/10.1016/j.jeap.2007.09.003

Le, Thi Ngoc Phuong & Harrington, Michael. 2015. Phraseology used to comment on results in the discussion section of applied linguistics quantitative research articles. *English for Specific Purposes* 39: 45–61. https://doi.org/10.1016/j.esp.2015.03.003

Lin, Ling & Evans, Stephen. 2012. Structural patterns in empirical research article: A cross-disciplinary study. *English for Specific Purposes* 31: 150–160. https://doi.org/10.1016/j.esp.2011.10.002

Marco, Maria J. L. 2000. Collocational frameworks in medical research papers: A genre-based study. *English for Specific Purposes* 19: 63–86. https://doi.org/10.1016/S0889-4906(98)00013-1
Mizumoto, Atsushi, Hamatani, Sawako & Imao, Yasuhiro. 2017. Applying the bundle–move connection approach to the development of an online writing support tool for research articles. *Language Learning* 67(4): 885–921. https://doi.org/10.1111/lang.12250
Moreno, Ana I. & Swales, John M. 2018. Strengthening move analysis methodology toward bridging the function-form gap. *English for Specific Purposes* 50: 40–63. https://doi.org/10.1016/j.esp.2017.11.006
Omidian, Taha, Shahriari, Hesamoddin & Siyanova-Chanturia, Anna. 2018. A cross-disciplinary investigation of multi-word expressions in the moves of research article abstracts. *Journal of English for Academic Purposes* 36: 1–14. https://doi.org/10.1016/j.jeap.2018.08.002
Parodi, Giovanni. 2010. The rhetorical organization of the textbook genre across disciplines: A 'colony-in-loops'? *Discourse Studies* 12(2): 195–222. https://doi.org/10.1177/1461445609356500
Swales, John. 1981. *Aspects of Articles Introductions* [Aston ESP Reports, No. 1]. Birmingham: The University of Aston in Birmingham.
Swales, John. 2004. *Research Genres*. Cambridge: CUP. https://doi.org/10.1017/CBO9781139524827
Tabachnick, Barbara G. & Fidell, Linda S. 2007. *Using Multivariate Statistics*, 5th edn. Boston MA: Allyn & Bacon/Pearson Education.
Tankó, Gyula. 2017. Literary research article abstracts: An analysis of rhetorical moves and their linguistic realizations. *Journal of English for Academic Purposes* 27: 42–55. https://doi.org/10.1016/j.jeap.2017.04.003

Appendix A. Scree plot for Principal Components Analysis