A Multi-Dimensional Analysis of Conclusions in Research Articles of Linguistics

Ya Xu iD

School of Foreign Studies, Northwestern Polytechnical University, China

Yi Zhang ✉

School of Foreign Studies, Northwestern Polytechnical University, China

Abstract:

The research article section, especially the introduction, has been a focus of scholarly discourse research for many years while linguistic strategies in the conclusion of research articles remain understudied. Moreover, most of the previous studies discussed linguistic features from the perspective of single-language strategies. Given this, this paper adopted an MDA (multi-dimensional analysis) method (Biber, 1988) to analyze the distribution of 67 linguistics features in the conclusions of 200 RAs (Research Articles) with the aid of the corpus tool MAT (Multi-dimensional Analysis Tagger) devised by Nini. The result showed that the dimension scores directed to the negative polarity for Dimension 1 (Mean=-15.80, SD=5.40), Dimension 2 (Mean=-2.61, SD=2.08) and Dimension 4 (Mean=-1.62, SD=2.74), and positive polarity for Dimension 3 (Mean=7.33, SD=2.85), Dimension 5 (Mean=5.47, SD=3.05), which indicated that conclusions of linguistic RAs are presented as informational-dense, relatively context-independent, less explicitly persuasive, highly technical, and abstract. Besides, the main linguistic features that contributed to the language variation of the RAs conclusion writing in Linguistics are Nouns, Attributive adjectives, Present tenses, Past participial WHIZ deletions, Phrasal coordination, Nominalization, Pied piping constructions, Infinitive TO, Possibility modals (e.g. may, might), Suasive verbs and Agentless passives. The study revealed the specific distribution of linguistic features of conclusion writing in RAs, highlighting the nature of informativeness and the abstractness of academic writing. This study may have some implications for writing academic papers, especially for graduates studying linguistics.

Keywords: *Multi-dimensional analysis, Research articles, Conclusions, Linguistic variation, Linguistics.*

Introduction

Academic discourse is an important way for the dissemination and exchange of scientific knowledge (Halliday & Martin, 1993; Hyland, 2004; Matalene, 1985). With the rise of functional linguistics, researchers have gradually realized that it not only conveys conceptual information and produces credible texts, but also expresses rich interpersonal meanings, achieving social interaction between authors and readers (Halliday & Martin 1993; Hyland 2004; Li Zhanzi 2001). Different disciplines construct disciplinary culture and knowledge through different discourse conventions, such as the expression of author positions and participation, as well as the structure of arguments (Hyland & Bondi, 2006). In terms of the disciplinary characteristics of academic discourse,

interdisciplinary research is particularly in-depth. Examples of written academic discourse that reveal interdisciplinary differences in previous studies are research articles (e.g. Basturkman, 2012; Cao&Hu, 2014; Hyland, 1999; Sotesbury, 2006), textbooks (e.g. Jalilifar et al., 2017; Moore, 2002; Parodi, 2014), student and second language writing (e.g. Crossley et al., e.g.), doctoral and master's theses (e.g. Helan, 2008; Paltridge, 2002; Samraj and Monk, 2008; Shaw and Sun, 2020; Xie and Ma, 2021). In short, rhetorical preferences among different disciplines have been well-established in literature. Compared with interdisciplinary research, research on language variations of linguistics is understudied comprehensively.

Currently, research on the linguistic features of academic discourse has applied two main analytical frameworks (Biber, 2019). On the one hand, it examines one or several language attributes from a micro perspective. Some representative research language features include lexical bundles (Bychkovska & Lee, 2017), collocation use (Chen, 2019), cohesion and discourse markers (Yang & Sun, 2012). On the other hand, a few studies have adopted a comprehensive framework, mainly a multi-dimensional analysis (MDA) framework, combining macro and micro-level analysis to study language differences between different disciplines and describe accurate language choices in specific contexts (Biber, 2006; Kosome, 2005; Hardy&R'Omer, 2013). Gray (2013) contended that although research focusing on a small number of language features can indeed reveal the high-frequency lexical and grammatical structures used in a certain register, using MDA methods can simultaneously identify and classify up to 67 language features, allowing for a more comprehensive examination of the register style and related comprehensive influencing factors of the text.

The traditional IMRD (Introduction-Method-Result-Discussion) model of research articles treats the discussion and conclusion as inseparable closing parts, potentially undermining the unique communicative function of the conclusion as a separate subcategory (Swales, 1990). And recent research suggests that the conclusion section of RAs is crucial in its own right. Yang and Allison (2003) examined the macrostructure of research articles and found that the discussion and conclusion sections serve different purposes. The discussion section focuses on specific results and comments, while the conclusion section provides an overall evaluation of the research findings (Yang & Allison, 2003). Based on this distinction, it has been suggested that the conclusion section should be treated as a separate entity, rather than merged with the discussion section. Furthermore, Lin and Evans (2012) argue that the introduction (I), literature review (L), and result-discussion (RD) sections of contemporary RAs are all important independent parts. They emphasize the need for academia to pay closer attention to these sections, including the conclusion, which plays a critical role in modern academic writing. The conclusion, especially when presented as a separate section, highlights research findings, contributions, and future directions, thereby achieving the authors' communicative purpose (Lin & Evans, 2012). These communicative purposes underscore the significance of RA conclusions. Despite their importance, research on RAs conclusions has been relatively neglected. Therefore, this study aims to address this gap and examine the linguistic changes in RA conclusions of linguistics, complementing the existing literature on this topic.

This study will adopt the MDA approach to investigate the linguistic features of conclusion in RAs of Linguistics. Through MDA, the register and language features of linguistics RAs will be revealed, highlighting the importance of studying conclusions as an independent section.

MDA

Biber (1988) first proposed the MDA method, which involves quantitative and qualitative comparative analysis of various dimensions of registers (spoken and written language). This method examines the co-occurrence patterns of 67 linguistic features, resulting in the extraction of seven functional dimensions. Each dimension serves different communicative functions and

can predict the register differences between spoken and written language. This approach, known as multi-feature/multi-dimensional analysis (MFA/MDA), has been widely used in corpus research to analyze register variation (Biber, 1988). The seven dimensions of Biber (1988) include dimension 1: "Interactivity/Informative expression". Positive values of factors in this dimension indicate strong interactivity of the text, and negative values indicate strong information transmission. Dimension 2: "Narrative/Non-narrative concern", the factor values of narrative texts such as novels are mostly positive, and the negative linguistic features indicate that the discourse is non-narrative. Dimension 3: "Clear and context-dependent reference", that is, the degree of text dependence on the environment is defined. The positive factor value indicates that the degree of text dependence is low, and the opposite indicates that the degree of dependency is high. Dimension 4: "Explicit persuasive representation". Positive values in this dimension indicate that the persuasive degree of the labeled text is large, while negative values indicate that the persuasive degree of the labeled text is small. Dimension 5: "Information abstract/non-abstract style". The value of the linguistic feature factor in this dimension is positive, indicating that the discourse style is abstract, professional and formal, and negative value indicates that the discourse is a non-abstract style. Dimension 6: "Fineness of immediate information organization", which mainly distinguishes between real-time and non-real-time information transmission texts; Dimension 7: "Academic model expression" (Biber,1988).

Each dimension generally contains the language features that the factor load is positive and the factor load is negative, and the positive load and the negative load are mutually complementary to a large extent (Biber,1988). "These co-occurring features are used together because they serve the same function" (Biber, 1988). Thus, positive load co-occurrence has the opposite communicative function as negative load.

MD analysis has been applied to different corpora and has demonstrated the prevalence of variation dimensions in the language (Friginal & Weigle, 2014). For example, studies have examined linguistic differences between abstracts published in the United States and Iraq (Friginal & Mustafa, 2017), as well as differences between abstracts written by native British and Chinese authors across academic fields (Cao & Xiao, 2013). These studies have revealed the influence of cultural differences on microscopic linguistic features and have implications for abstract writing across disciplines. Interdisciplinary MD analysis has also attracted attention, studying language variation in different disciplines and identifying new dimensions of language variation influenced by disciplinary norms and purposes (Gray, 2013; Gardner et al., 2019).

There have been generally two ways to conduct an MD analysis. One is to carry out a new factor analysis (also called full MD analysis) as Biber did in the late 1980s. Researchers first calculate the normalized frequencies of numerous linguistic features and then conduct a factor analysis on those frequency values to reach several new factors (or dimensions in MD analysis). New factors are then interpreted and named by researchers through qualitative interpretations of corresponding features. This way is widely used for identifying the co-occurrence patterns of specialized or newly born registers, for example, non-western languages including Portuguese (Sardinha et.al, 2014), Chinese (Zhang, 2012) and Pop Songs (Dutra, 2014) among others. The other approach is called additive MD analysis. This approach applies Biber's (1988) dimensions directly and hence omits the factor extraction. Studies using additive MD analysis calculate the dimension scores of a text based on the mean frequencies for variables (i.e., linguistic features) provided in Biber (1988) and then contrast the obtained dimension scores to the scores of registers in Biber's (1988) dimension scale to evaluate how the text under discussion is related to the wide variety of discourses investigated in Biber (1988).

The present study applied the latter method as it is comparatively less demanding for researchers to conduct an MD analysis in comparison with a full MD analysis. Another benefit brought by

this method is that other discourses that are investigated in the same way, along with the discourses in Biber's (1988) dimension scale, can provide rich reference information for a more fine-grained analysis.

Methodology

Research Question

Given the pivotal role of conclusion sections in RAs and the methodological significance of MDA, this study aims to investigate the linguistic features that characterize conclusion sections in RAs. Specifically, the following research questions will be addressed:

RQ1: Based on Biber's (1988) MDA framework, what are the functional representations of conclusion sections of linguistic RAs regarding five dimensions?

RQ2: What are the main linguistic features of conclusion sections employed in each dimension in linguistic RAs?

Corpus

The corpus used in this study consists of conclusions from 200 research articles in linguistics. Those research articles were selected from four representative journals published from 2018 to 2022 evenly through stratified sampling. The four journals are the Journal of English for Academic purpose, Journal of Pragmatics, Studies in Second Language Acquisition, and Journal of Sociolinguistics. The impact factor of these journals suggests that it has a high citation rate in the field, which may be indicative of the journal's high quality and are highly respected in linguistic fields.

Journal articles were downloaded as .pdf files and saved as .txt files, and the conclusion sections were then extracted and saved as .txt files. Each text was stored independently, titled with information such as publishing journal, publishing year and serial number. For example, EAP 1-2022-V202-4 indicates that the text is the conclusion extracted from the fourth paper in Volume 202 of the Journal of English for Academic purpose published in 2022. Next, the conclusion was then analyzed by Multidimensional Analysis Tagger (MAT) as a whole document. The details of the corpus are summarized in Table 1. It includes 50 conclusions from each journal, with the Journal of English for Academic Purpose containing a total of 30,544 tokens, the Journal of Pragmatics with 27,901 tokens, Studies in Second Language Acquisition with 24,834 tokens, and the Journal of Sociolinguistics with 29,937 tokens.

Table 1. Descriptive Statistics of the Corpus

Journal	Journal of English for Academic purpose (EAP)	Journal of Pragmatics (JOP)	Studies in Second Language Acquisition (SLA)	Journal of Sociolinguistics (JOS)	Sum
Number of texts	50	50	50	50	200
Number of word types	3932	4244	3319	4820	16315
Number of word tokens	30544	27901	24834	29937	113216

Analysis

The corpus of conclusions in linguistics RAs was analyzed by using Nini's (2015) MAT, which applies Biber's (1988) MD analysis to assess register variation in speech and writing. The MAT has been investigated for reliability and applicability by Nini (2019) and has also been used by other researchers that contribute to the scholarship with some very illuminating findings (e.g., Crosthwaite, 2016; Ren & Lu, 2021). MAT first conducted a grammatical analysis of each text using the Stanford Tagger (Toutanova, Klein, Manning, & Singer, 2003) and then identified the linguistic features based on Biber's (1988) dimensions from each grammatically

tagged text. The frequency counts of these linguistic features were normalized to occurrences per 100 words to account for text length variations. These features were then used to generate scores for Biber's (1988) six functional dimensions for each text. Table 2 provides an overview of the six dimensions from Biber (1988) that were employed in the program. In this study, we focus on the first five dimensions proposed by Biber because the data of the seventh dimension is thin and often omitted in practice (Biber,1988) and the sixth dimension discusses the features of Online Informational Elaboration, which seems irrelevant to academic writing.

According to Nini (2019), the MAT produces a tab-delimited file titled 'corpus_statistics.txt' with frequency counts, as well as a tab-delimited file including the z-scores of the linguistic variables. Besides, a tab-delimited text file containing the dimension scores is displayed. It is interesting that MAT also produces a graph that illustrates the location of the analyzed corpus concerning Biber's (1989) eight text types.

Table 2. A Summary of Biber's (1988) Dimensions (Nini, 2015)

Dimension	Description
D1	Involved vs. Informational Discourse: A high score indicates the text is affective and interactional, e.g., a casual conversation; the text presents many verbs and pronouns (among other features). A low score indicates the text is informationally dense, e.g., academic prose; the text presents many nouns, long words, and adjectives (among other features).
D2	Narrative vs. Non-Narrative Concerns: A high score indicates the text is narrative, e.g., a novel. The text presents many past tenses and third-person pronouns (among other features). A low score indicates the text is non-narrative.
D3	Context-Independent vs. Context-Dependent Discourse: A high score indicates the text is context-independent, e.g., academic prose; the text presents many nominalizations (among other features). A low score indicates the text is context-dependent, e.g., a sports broadcast; the text presents many adverbs (among other features).
D4	Overt Expression of Persuasion: A high score indicates that the text explicitly marks the author's point of view and their assessment of likelihood and/or certainty, e.g., a professional letter; the text presents many modal verbs (among other features).
D5	Abstract vs. Non-Abstract Information: A high score indicates the text is highly technical, abstract, or formal, e.g., scientific discourse; the text presents many passive clauses and conjuncts (among other features).

Results and Discussion

Table 3 presents descriptive statistics for each dimension and the genre that the corpus are closest to on each dimension overall. The corpus shows that the dimensions scores of the negative polarity for Dimension 1 (Mean=-15.80, SD=5.40), Dimension 2 (Mean=-2.61, SD=2.08) and Dimension 4 (Mean=-1.62, SD=2.74), and positive polarity for Dimension 3 (Mean=7.33, SD=2.85), Dimension 5 (Mean=5.47, SD=3.05) and Dimension 6 (Mean=1.23, SD=2.11). According to the result of MAT, the conclusions in linguistics RAs are closest to scientific exposition, a type of text characterized by low scores on D1 (informationally dense) and high scores on D3 (more context-independent) and D5 (more technical, abstract, or formal), with such characterizing genres as official documents and academic prose (Nini, 2015). Texts of this type

are "typically informational expositions that are formal and focused on conveying information and very technical" (Nini, 2015).

Dimension 1: Involved vs. Informational Discourse

The mean of the first dimensions scores was -15.08(See Table 3), indicating that the texts tended to locate toward the negative, namely the informational direction of the continuum. This reveals the high informational features of the conclusion in RAs. The primary linguistic characteristics of the negatived along D1 are the frequent use of nouns, prepositional phrases, high type/token ratios, attributive adjectives, place adverbials, passives and longer words (Biber, 1988), all of which contribute to the highly informational nature of academic discourse. According to the z-score generated (See Table 4), long word length (Mean Z-Score = 2.71), nouns (Mean Z-Score = 1.51) and attributive adjectives (Mean Z-Score = 2.48) contribute greatly to the informational feature of conclusions in RAs.

Table 3. Descriptive Statistics and Closest Genres of the Corpus

Dimension	Mean	SD	Closet Genre
Dimension1	-15.08	5.40	Press reportage
Dimension2	-2.61	2.08	Academic prose
Dimension3	7.33	2.85	Official documents
Dimension4	-1.62	2.74	Press reportage
Dimension5	5.47	3.05	Academic prose

Table 4. Mean Scores of Linguistic Features in Dimension 1

	Linguistic features	Mean Z-Score
Positive	Private verbs (PRIV)	-0.3
	Subordinator that deletion (THATD)	-0.46
	Contractions (CONT)	-0.73
	Present tense verb (VPRT)	-0.93
	2nd person pronouns (SPP2)	-0.71
	Do as pro-verb (PROD)	-0.75
	Analytic negation (XX0)	-0.54
	Demonstrative pronouns (DEMP)	-0.2
	Emphatics (EMPH)	-0.07
	1st person pronouns (FPP1)	-0.68
	Pronoun it (PIT)	-0.56
	Be as main verb (BEMA)	-1.74
	Causative adverbial (CAUS)	-0.17
	Discourse particles (DPAR)	-0.52
	Indefinite pronouns (INPR)	-0.64
	Hedges (HDG)	-0.42
	Amplifiers (AMP)	-0.61
	Sentence relatives (SERE)	3.35
	Direct WH-questions (WHQU)	-0.11
	Possibility modals (POMD)	0.83
	Independent clause coordination (ANDC)	0.63
	Wh-clauses (WHCL)	0.45
	Stranded preposition（STPR）	-0.57

	Nouns (NN)	1.51
Negative	Average Word Length (AWL)	2.71
	Prepositions (PIN)	0.55
	Type/token ratio (TTR)	-0.93
	Attributive adjectives (JJ)	2.48
	Agentless passives (PASS)	0.46

(1) Moreover, the Austinian(JJ) approach(NN) enables us to account for(PIN) a number(NN) of puzzling(JJ) phenomena(NN) characteristic of(PIN) verbal(JJ) comprehension(NN). For instance, it enables us to argue that ironizing is a socially constituted (JJ) communicative(JJ) practice(NN) and that ironic(JJ) utterances(NN), rather than being manifestations(NN) of(PIN) our 'imaginative(JJ) engagement' (Lepore and Stone, 2015: 177–180), are speech actions done as conforming to a certain(JJ) procedure(NN). (JOP 2-2022-V201-5)

(2) Its advantage(NN) lies in(PIN) presenting the students with(PIN) an authentic(JJ) text(NN) in(PIN) a way(NN) that encourages them to relate the target word(NN) back to its context(NN) and then perform various(JJ) types(NN) of(PIN) tasks(NN), such(JJ) as gap-filling(JJ) and close(JJ) activity(NN). It also promotes developing their discovery strategies(NN) by(PIN) seeking the reasons(NN) why particular(JJ) words(NN) go together or not. (EAP 32-2019-V42-5)

Note. NN=noun, JJ=attributive adjectives, PIN=Prepositions

In example (1), the author employed 7 nouns, 7 attributive adjectives and 3 prepositions to illustrate the function of the Austinian approach to explain some puzzling phenomena in the verb comprehension with an instance of "Irony". In example (2), it shows the practical use of nouns, attributive adjectives and Prepositions in the articles of EAP. We can see the structure "(JJ+) NN+PIN" appears frequently in the examples, such as "various (JJ) types (NN) of (PIN) tasks" and "manifestations (NN) of (PIN)". Besides, the phenomenon of "JJ+JJ+NN" shows the writing skills to condense information, such as "socially-constituted (JJ) communicative (JJ) practice (NN)", which means using fewer words to express the same meaning as accurately as possible.

Dimension 2: Narrative vs. Non-Narrative Concerns

For Dimension 2, the mean dimension score was -2.61. Low scores along the negative pole in D2 mark non-narrative concerns, such as expository and descriptive discourse; and features that demonstrate considerable negative weight can be present tense and attributive adjectives (Biber, 1988), as shown in examples (1) and (2). The Conclusions in RAs revealed the overall tendency of non-narrative characteristics. This result is somewhat reminiscent of Gray's (2015) "Human Focus and Non-Human Focus" (D3), in which she pointed out that the discipline of applied linguistics is about humans and language. This helps to explain some non-narrative features in linguistics.

Table 5. Mean Scores of Linguistic Features in Dimension 2

	Linguistic features	Mean $_{\text{Z-Score}}$
Positive	Past tense verbs (VBD)	-0.85
	Third person pronouns (TPP3)	-0.83
	Perfect aspect verbs (PEAS)	-0.61
	Public verbs (PUBV)	-0.18
	Synthetic negation (SYNE)	-0.59
	Present participial clauses (PRESP)	0.56

Negative	Present tense verbs (VPRT)	-0.93
	Attributive adjectives (JJ)	2.48
	Past participial WHIZ deletions (WZPAST)	0.3
	Word length (AWL)	2.71

(3)   Nevertheless, while the results from the Faroese verbal guise tests are(VPRT) dissimilar to the pattern found(WZPAST) in comparable verbal guise tests in Denmark (Kristiansen 2009), they resemble(VPRT) the ones found(WZPAST) in a series of verbal guise tests in Western Norway (JOS-46-2018-V22-I3-2).

Note. VPRT =Present tense verbs, WZPAST=Past participial WHIZ deletions

In this example, only present tenses and Past participial WHIZ deletions were used (e.g., are dissimilar to, found in) and attributive adjectives frequently occurred (e.g., comparable, Western). These features reflect a more frequent use of elaborated nominal referents, which is commonly observed in non-narrative types of discourse.

Dimension 3: Explicit vs. Situation-Dependent Reference

The mean dimension score of D3 was 7.33, which indicates the discourse is explicit and context-independent. Features that have positive weights include WH-relative clauses, pied piping constructions, phrasal coordination, and nominalizations. The co-occurrence of these features indicates the explicit and elaborated identification of referents in a text (Biber, 1988). Since referentially explicit discourse is usually integrated and informational language is used, this aligns with the previous analysis of Dimension 1. According to Biber (1988), high scores along this dimension are featured by 'text-internal reference.' The highly nominalized information and explicit referent may be related to the overall purposes of the conclusions.

Table 6. Mean Scores of Linguistic Features in Dimension 3

	Linguistic features	Mean $_{Z\text{-Score}}$
Positive	WH relative clauses on object position (WHOBJ)	-0.68
	Pied piping constructions (PIRE)	0.88
	WH relative clauses on subject positions (WHSUB)	-0.59
	Phrasal coordination (PHC)	3.97
	Nominalizations (NOMZ)	2.00
Negative	Time adverbials (TIME)	-1.04
	Place adverbials (PLACE)	-0.41
	Adverbs (RB)	-1.93

(4)   Exploring publication(NOMZ) practices through the lens of trajectory shows that in the context of what appears to be an inexorable drive for ever-more rigid evaluation(NOMZ) regimes globally (Curry & Lillis, 2017), opportunities(NOMZ) for women scholars' agency arise at different points in time; thus scholars engage with evaluation(NOMZ) regimes, aspects of which(PIRE) align with their interest and(PHC) desires, whilst also enacting practices which(WHSUB) challenge such regimes. (EAP 47-2018-V32-5)

Note. PIRE=Pied piping constructions, NOMZ=Nominalizations.

Nominalization (e.g., evaluation, opportunities) and the frequent co-occurrence of relative clauses (e.g., which challenge) explicitly indicate that discourse with stronger clarity is informational-oriented (Biber, 1988: 110), as shown in Example 4, trying to integrate and concisely present the research findings. Writers are good at using relative clauses to compress information into sentences, and readers

understand their references through guiding words, thus increasing the readability of the text.

Dimension 4: Overt Expression of Persuasion

The mean dimension score of the conclusions in Dimension 4 was -1.62. Low scores toward the negative side in Dimension 4 suggest the low degree to which persuasion is marked overtly. Persuasion is expressed by using infinitives, suasive verbs, and modal verbs (Biber, 1988). According to Table 7, it can be seen that writers of linguistics tend to use Possibility modals (e.g., can, may, might, could), Infinitives(to) and Suasive verbs(e.g., agree, allow) to convince the readers of the value of the study and they will avoid the frequent use of Necessity modals(ought, should, must), Prediction modals (will, would, shall) and Conditional subordination (if and unless). Generally speaking, conclusions explored in this study have a relatively high frequency of using possibility modals (e.g., might, may), which are considered hedged remarks to emphasize the implications of the research. Writers from social sciences can present knowledge as a common understanding and promote tolerance in readers' resulting in realizing the purpose of emphasizing the value of the study (Hyland, 2008).

Table 7. Mean Scores of Linguistic Features in Dimension 4

	Linguistic features	Mean $_{Z\text{-Score}}$
Positive	Infinitives (TO)	0.39
	Prediction modals (PRMD)	-0.79
	Suasive verbs (SUAV)	0.41
	Conditional subordination (COND)	-0.80
	Necessity modals (NEMD)	-0.27
	Split auxiliaries (SPAU)	-0.54
	Possibility modals (POMD)	0.83

(5) Based on the findings and methodological features of this study, we recommend(SUAV) that feedback be provided immediately after learners' initial exposure to the linguistic feature and at the beginning of an instruction cycle. We argue that errors should(NEMD) be addressed before they are proceduralized in the L2 system. We also argue that should(NEMD) CF be accompanied by instruction, they should(NEMD) be delivered in tandem rather than separately, so as to(TO) achieve maximal effects for both (SLA 9-2022-V44-I1-1).

(6) In addition, the analysis of the data described here suggests(SUAV) important implications for the ways in which speakers mobilize linguistic resources in identity performance and how social networks reveal the influence of linguistic ideologies. While the data described here can(POMD) not fully confirm these observations, it may(POMD) motivate additional research in these areas...These participants are more likely to(TO) encounter raciolinguistic ideologies that function to(TO) restrict the linguistic choices available to speakers in the city. (JOS 2-2022-V26-I5-2)

Note. TO=infinitive TO, POMD=Possibility modals, SUAV=Suasive verbs, NEMD=Necessity modals

Examples 5 and 6 show the frequently-used necessity modals "should" and possibility modals "may" and "can" in the conclusions. In example 5, those writers use "should" to provide some suggestions about teaching based on their findings, which shows their confidence and high expectations about changes to make in practical classrooms. It can be seen in example 6 that writers use possibility modals and Suasive verbs to try to avoid blunt assertions when emphasizing the implication of the research findings.

Dimension 5: Abstract vs. non-abstract information

Positive dimension scores (5.47) in Dimension 5 indicate technical, abstract, and formal

discourse. Conjuncts, agentless passives, past participial clauses, by passives, past participial WHIZ deletion relatives, and other adverbial subordinators are key features contributing to high positive scores in Dimension 5 (Biber, 1988). In terms of the linguistic features along D5, conclusion texts from linguistics use many conjuncts (Mean Z-Score=4.7), other adverbial subordination (Mean Z-Score=1.17) and Past participial clauses (Mean Z-Score=1.44), emphasizing the objective description of findings.

Table 8. Mean Scores of Linguistic Features in Dimension 5

	Linguistic features	Mean $_{Z\text{-Score}}$
Positive	Conjuncts (CONJ)	4.7
	Agentless passives (PASS)	0.46
	Past participial clauses (PASTP)	1.44
	BY-passives (BYPA)	0.48
	Past participial WHIZ deletions (WZPAST)	0.3
	Other adverbial subordinators (OSUB)	1.17
	Predicative adjectives (PRED)	0.94
Negative	Type-token ratio (TTR)	-0.93

(7) The findings also indicate that the antecedent's degree of individuation is relevant(PRED) to the pronoun choice: Generic he is(PASS) used more with antecedents that have a high degree of individuation, i.e.(CONJ) definite and indefinite NPs, whereas(OSUB) singular they are more frequent(PRED) with antecedents that are low(PRED) in individuation, i.e.(CONJ) quantificational NPs and indefinite pronouns. (EAP 38-2019-V38-9)

Note. PASS=Agentless passives, PRED=Predicative adjectives, OSUB=Other adverbial subordinators, PRED=Predicative adjectives.

In example 7, passives are used to underline the research findings rather than the people who conducted the research. The utilization of conjunctive and adverbial subordinator (e.g., which) here explains the general finding with comparison and examples. Those linguistic features combined contribute to the abstractness and formality of conclusion writing.

Conclusion

This study uses MDA to study the linguistic characteristics in the conclusions of linguistic RAs of four representative journals published from 2018 to 2022. Our research results contribute to revealing linguistic features of conclusion writing in linguistics, which may have some reference significance for the teaching and writing of academic papers for students of linguistics majors.

The MDA of the conclusion in linguistics RAs shows that there are some substantial findings on linguistic characteristics. The MD analysis of conclusions in this study is based on the first five of the seven dimensions of Biber(1988) (D1: Discourses involved versus information discourses; D2: Narrative and non-narrative concerns; D3: Context-independent and context-dependent discourse; D4: Public expression of persuasion; D5: Abstract and non-abstract information), which is due to the use of a multi-dimensional analysis tool (Nini, 2015). Based on the dimension score and analysis of the conclusion, the following key findings were made in this study. Regarding these two research questions, the research results show that in general, the conclusions of RAs in linguistics are presented as information-dense, context-independent, less persuasive, highly technical and abstract. This also shows the nature of informativeness and abstractness in academic writing. To achieve this effect, the main linguistic features employed by linguistics writers are Nouns, Attributive adjectives, Present tenses, Past participial WHIZ deletions, Phrasal

coordination, Nominalization, Pied piping constructions, Infinitive TO, Possibility modals (e.g. may, might), Suasive verbs and Agentless passives. The frequent use of those linguistic devices collectively contributes to the unique linguistic features of the RAs conclusion writing of linguistics.

The study has several limitations, some of which constitute useful avenues for future research. Firstly, the compiled corpus can be expanded in various ways to enrich the research scope. For example, the corpus can be expanded by including RAs from other disciplines like Economics and Education. Second, this study only focuses on the conclusion part of the RAs, and other parts of the RAs like the Discussion and Introduction are also worthy of further exploration, so as to more comprehensively recognize the linguistic commonalities and differences in those sections, and help us to understand the linguistic characteristics better.

Conflict of Interests

No conflict of interest.

References

Askehave, I., & Swales, J. M. (2001). Genre identification and communicative purpose: A problem and a possible solution. *Applied Linguistics, 22*(2), 195-212. https://doi.org/10.1093/applin/22.2.195

Biber, D. (1988). Variation across speech and writing. Cambridge: Cambridge University Press. Biber, D. (1989). *A typology of English texts. Linguistics, 27*(1), 3-43.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers.* Amsterdam: John Benjamins.

Biber, D., & Gray, B. (2013). *Identifying multi-dimensional patterns of variation across registers.*

Biber, D. (2019). Text-linguistic approaches to register variation. *Register Studies, 1*(1), 42-75. https://doi.org/10.1075/RS.18007.BIB

Bruce, I. (2009). Results sections in sociology and organic chemistry articles: A genre analysis.

English for Specific Purposes, 28(2), 105-124. https://doi.org/10.1016/j.esp.2008.12.005

Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes, 30,* 38-52. https://doi.org/10.1016/j.jeap.2017.10.008

Chen, W. (2019). Profiling collocations in EFL writing of Chinese tertiary learners. *RELC Journal, 50*(1), 53-70. https://doi.org/10.1177/0033688217716507

Cortes, V. (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

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes, 22,* 166-178. https://doi.org/10.1016/j.jeap.2016.04.005

Friginal, E., & Mustafa, S. S. (2017). A comparison of US-based and Iraqi English research article abstracts using corpora. *Journal of English for Academic Purposes, 25,* 45-57. https://doi.org/10.1016/j.jeap.2016.11.004

Friginal, E., & Weigle, S. (2014). Exploring multiple profiles of L2 writing using multi-dimensional analysis. *Journal of Second Language Writing, 26*(2), 80-95. https://doi.org/10.1016/j.jslw.2014.09.007

Gao, X. (2016). A cross-disciplinary corpus-based study on English and Chinese native speakers' use of linking adverbials in academic writing. *Journal of English for Academic Purposes, 24*, 14-28. https://doi.org/10.1016/j.jeap.2016.08.002

Gardner, S., Nesi, H., & Biber, D. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics, 40*(4), 646-674. https://doi.org/10.1093/applin/amy005

Gray, B. (2013). More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles. *Corpora, 8*(2), 153-181. https://doi.org/10.3366/cor.2013.0039

Jin, B. (2018). A multidimensional analysis of research article discussion sections in the field of chemical engineering. *IEEE Transactions on Professional Communication, 61*(3), 242-256. https://doi.org/10.1109/TPC.2018.2817002

Jin, G., Li, C., & Sun, Y. (2020). Exploring the macrostructure of research articles in economics. *IEEE Transactions on Professional Communication, 63*(3), 227-243. https://doi.org/10.1109/TPC.2020.3014535

Kwan, B. C., Chan, H., & Lam, C. (2012). Evaluating prior scholarship in literature reviews of research articles: A comparative study of practices in two research paradigms. *English for Specific Purposes, 31*(3), 188-201. https://doi.org/10.1016/j.esp.2012.02.003

Le, T. P., & Harrington, M. (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, L., & Evans, S. (2012). Structural patterns in empirical research articles: A cross-disciplinary study. *English for Specific Purposes, 31*(3), 150-160. https://doi.org/10.1016/j.esp.2011.10.002

Martínez, I. A. (2003). Aspects of theme in the method and discussion sections of biology journal articles in English. *Journal of English for Academic Purposes, 2*(2), 103-123. https://doi.org/10.1016/S1475-1585(03)00003-1

Moreno, Ana I. (2021). Seling research in RA discussion sections through English and Spanish: An intercultural rhetoric approach. *English for Specific Purposes, 63,* 1-17. https://doi.org/10.1016/j.esp.2021.02.002

Nini, A. (2015). Multidimensional analysis tagger (version 1.3) – Manual. Retrieved from http://sites.google.com/site/multidimensionaltagger

Nini, A. (2019). The multi-dimensional analysis tagger. In T. Berber Sardinha, & M. Veirano Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 67-94). London.

Ren, C. W., & Lu, X. F. (2021). A multi-dimensional analysis of the Management's Discussion and Analysis narratives in Chinese and American corporate annual reports. *English for Specific Purposes, 64,* 84-99. https://doi.org/10.1016/j.esp.2020.12.004

Sheldon, E. (2011). Rhetorical differences in RA introductions written by English L1 and L2 and Castilian Spanish L1 writers. *Journal of English for Academic Purposes, 10*(4), 238-251. https://doi.org/10.1016/j.jeap.2011.08.004

Sheldon, E. (2018). Dialogic spaces of knowledge construction in research article conclusion sections written by English L1, English L2 and Spanish L1 writers. *Ibérica, 35,* 13-40.

Swales, J. (1990). *Genre analysis: English in academic and research settings.* Cambridge: Cambridge University Press.

Toutanova, K., Klein, D., Manning, C., & Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. In Proceedings of the *2003 conference of the North American chapter of the association for computational linguistics on human language technology* (pp. 252-259). East Stroudsburg, PA: Association for Computational Linguistics.

Xie, Q. & Ma, Y. (2021). A multi-dimensional comparative analysis of English abstracts of Chinese and foreign doctoral dissertations. *Journal of Foreign languages, 04,* 39-44. https://doi.org/10.1177/21582440221145669

Yang, R. Y., & Allison, D. (2003). Research articles in applied linguistics: Moving from results to conclusions. *English for Specific Purposes, 22*(4), 365-385. https://doi.org/10.1016/S0889-4906(02)00026-1

Yang, W. X., & Sun, Y. (2012). The Use of Cohesive Devices in Argumentative Writing by Chinese EFL Learners at Different Proficiency Levels. *Linguistics and Education, 23,* 31-48. https://doi.org/10.1016/j.linged.2011.09.004

Ye, Y. P. (2019). Macrostructures and rhetorical moves in energy engineering research articles written by Chinese expert writers. *Journal of English for Academic Purposes, 38,* 48-61. https://doi.org/10.1016/j.jeap.2019.01.007

Appendix

A Summary of Key Linguistic Features in Each Dimension. The following content is adapted from Biber (1988)

	Dimension 1: Involved vs. informational discourses	Dimension 2: Narrative vs. non-narrative concerns	Dimension 3: Explicit vs. situation-dependent reference	Dimension 4: Overt expression of persuasion	Dimension 5: Abstract vs. non-abstract information	Dimension 6: Online informational elaboration
Positive Weights	•Private verbs •THAT deletion •Contractions •Present tense verbs •2nd person pronouns •DO as pro-verb •Analytic negation •Demonstrative pronouns •General emphatics •1st person pronouns •Pronoun IT •BE as main verb •Causative subordination •Discourse particles •Indefinite pronouns •General hedges •Amplifiers •Sentence relatives •WH questions •Possibility modals •Non-phrasal coordination •WH clauses •Final prepositions •Adverbs •Conditional subordination	•Past tense verbs •Third person pronouns •Perfect aspect verbs •Public verbs •Synthetic negation •Present participial clauses	•WH relative clauses on object position •Pied piping constructions •WH relative clauses on subject positions •Phrasal coordination •Nominalizations	•Infinitives •Prediction modals •Suasive verbs •Conditional subordination •Necessity modals •Split auxiliaries •Possibility modals	•Conjuncts •Agentless passives •Past participial clauses •BY-passives •Past participial WHIZ deletions •Other adverbial subordinators •Predicative adjectives	•THAT clauses as verb complements •Demonstratives •THAT relative clauses on object positions •THAT clauses as adj. complements •Final prepositions •Existential THERE •Demonstrative pronouns •WH relative clauses on object positions
Negative Weights	•Nouns •Word length •Prepositions •Type/token ratio •Attributive adjectives •Place adverbials •Agentless passives •Past participial WHIZ deletions •Present participial WHIZ deletions	•Present tense verbs •Attributive adjectives •Past participial WHIZ deletions •Word length	•Time adverbials •Place adverbials •Adverbs	No negative features	•Type-token ratio	•Phrasal coordination