

Vol. 13, No. 3
pp. 139-173
July &
August 2022

A Corpus-based Analysis of Noun Phrase Complexity in Research Article Part-genres in Applied Linguistics and Clinical Medicine

Rajab Esfandiari ^{*1}  & Mohammad Ahmadi ² 

Abstract

Complexity measures in academic writing have experienced a shift from clausal to phrasal indices in recent years. Drawing on a subset of Biber et al.'s (2011) hypothesized stages of writing development, we explored phrasal complexity across sections (part-genres) of research articles (RAs) in applied linguistics and clinical medicine. A 389,332-word corpus consisting of 80 randomly selected RAs from leading journals in applied linguistics and clinical medicine was compiled for the purposes of the present study. One-way analysis of variance (ANOVA) and independent-samples *t*-test, as implemented in SPSS (version 25), were employed to find differences across the RA sections and between two groups of academic writers. The findings indicated that RAs in clinical medicine relied more heavily on noun phrase modifiers in all sections than those in applied linguistics, suggesting that the distributional pattern of these linguistic expressions is discipline-independent. The implications of the distributional pattern of phrasal complexity are discussed in relation to L2 writing pedagogy and the development of genre-based, discipline-specific academic writing.

Keywords: academic disciplines, applied linguistics, clinical medicine, part-genres, phrasal complexity, research articles

Received: 19 October 2021
Received in revised form: 6 February 2022
Accepted: 25 February 2022

¹ Corresponding Author, PhD. Imam Khomeini International University;

Email: esfandiari@hum.ikiu.ac.ir, ORCID ID; <https://orcid.org/000000022305762X>

² PhD. Imam Khomeini International University; ORCID ID; <https://orcid.org/0000000271428093>

1. Introduction

Research articles (RAs), as the most important sub-register of professional academic writing (Biber & Gray, 2010), have received a great amount of attention, as they provide a valuable academic means for exploring genre variation and disciplinary studies. Following Swales (1990), a considerable proportion of previous studies set out to investigate the rhetorical organization of the academic texts in terms of moves and steps in RAs across a wide range of disciplines as well as the linguistic means which are used to realize them within the framework of genre-based pedagogy.

A genre-based approach to writing can open up better understanding for linguistic specificity of the texts as a whole and their sub-sections in particular. Although the text has an overall function in its entirety, the rhetorical sections that build up the texts serve their own functions as well. The most salient types of text items that typify a specific rhetorical unit in an RA leads the readers of a particular discourse community to interpret the rhetorical function of a given unit (Moreno & Swales, 2018). When it comes to RAs, the rhetorical units are conventionally Introduction, Methods, Results, and Discussion (IMRD) (Swales, 1990). In the words of Casal et al. (2021), “investigations into the rhetorical stages of IMRD part-genres can highlight their distinct communicative purposes (p. 2).

While previous research studies on textual features of RA have broadened our knowledge of their disciplinary variations, they regard RA as rigid genres with similar distribution of language features across its different sections (e.g., Dahl, 2004). Surprisingly, there have been very few attempts to analyze disciplinary writing variations in terms of RA internal structures. As Basturkmen (2009) pointed out, student writers “often struggle to understand the forms and functions of sections of the research report” (as cited in Casal et al., 2021, p. 2). Accordingly, familiarity with function-form mapping particularly in disciplinary writing performance could enhance novice writers’ academic writing literacy.

Given that disciplines differ in their lexico-grammatical features, one of the major goals of the current study centers on the investigation of noun phrase complexity in hard science and soft science. In this study, the researchers follow Becher and Trowler’s (2001) classification of academic discipline based on two continua of applied/pure and hard/soft science, as “it is capable of making more subtle distinctions compared to its competitors” (p. 35). Following Hyland (2008), we choose the disciplines that “represent a cross-section of academic practice” (p.

8). Accordingly, two disciplines of medicine (hard applied) and applied linguistics (soft applied) were chosen for investigation in the current study. Other disciplines could also have been chosen to represent different dimensions of academic practice; however, we chose those that are established disciplines and have clearly defined boundaries compared to those that loosely fit into their domains. For example, one might argue that economics is more representative of hard pure than soft pure, especially if it is compared with, for example, anthropology (Becher & Trowler's, 2001).

Drawing on Swalesian genre theory for EAP pedagogy, some studies put a part-genre perspective on RAs and investigated sub-registers of Introduction (e.g., Samraj, 2005), Methods (e.g., Lim, 2006), Results (e.g., Basturkmen, 2009), or Discussion (Parkinson, 2011) within or across a number of disciplines. However, these studies produce a fragmented knowledge of these disciplines in their entirety (Kanoksilapatham, 2015) of the constructions and language features. Thus, there has been little research to explore linguistic differences across text-internal sections of RAs. Therefore, using a subset of Biber et al.'s (2011) hypothesized stages of writing development, we intend to explore noun phrase complexity across sections of RAs in two disciplines of applied linguistics and clinical medicine. This study, accordingly, seeks to address the following two research questions.

- 1 Are there any significant differences in the frequency of noun phrase modifiers between two writer groups in applied linguistics and clinical medicine?
- 2 Are there any significant differences in the frequency of noun phrase modifiers across sections of RAs in applied linguistics and clinical medicine?

2. Literature Review

2.1. Noun Phrase Complexity as a Hallmark of Advanced Academic Writing

Noun phrase complexity has received extensive treatment in academic writing since as early as 1960s when Rulon Wells (as cited in Biber et al., 2011) stated that in academic writing the use of nouns is preferred to that of verbs. Similarly, Halliday and Martin (1993) characterized academic register as a synoptic style of writing which is contrasted with dynamic style of writing. Synoptic style of expression occurs when the utterance, which is constructed as a series of interdependent clauses, is reconstructed as groups of words and phrases. Extract 1 indicates how

dynamic style of expression changes to synoptic one (verbs and nouns are underlined and in bold type).

Text Extract 1:

(1) Since some university professors require that the students summarize several essays during the semester, they prepare the way for article abstract writing in graduate programs.

(2) Some university professors' requirement of essay summary writing during the semester results in the students' preparation for article abstract writing in graduate programs.

Nominal style of writing in academic register is informed by the demand of packing more information into fewer words as far as possible. Such factors as economy of space, economy of expression, and reading efficiency contribute to nominal style of writing (Biber & Gray, 2010). Halliday and Martin (1993) argue that there are two major functions of expressing meaning through nominal phrases rather than clausal structures: First, the meaning that can be conveyed as a clause can be compressed into a phrase which can subsequently function as a part for the next clause. Second, information can be packed into phrases to express things/nouns instead of dealing with the tension between things and actions which are commonly found in clausal constituents.

Recently, traditional measures of writing complexity have been undermined on the grounds that they reflect the characteristic of spoken language rather than written language (Biber et al., 2011). Biber and Gray (2016) state that many of the complexity features commonly found in spoken language are nonexistent or are very rare in written language. For example, finite dependent clauses functioning as constituents in other clauses and finite clauses functioning as verb complements are much more common in conversation than in academic writing. Complexity in academic writing is constructed in noun phrases but in conversation it is realized in clausal embedding or subordination (Biber et al., 2011).

A number of studies have examined noun phrase modifiers (Table 1) introduced by Biber et al. (2011) in different academic contexts. Parkinson and Musgrave (2014) investigated academic writing produced by two groups of graduate L2 learners (i.e., MA academic writers to represent more proficient group and EAP writers to represent less proficient group). The results revealed that the EAP group relied heavily on attributive adjectives which are hypothesized as being an early

stage of writing development. However, the MA group's performance was much closer to those of expert academic writers in published RAs. Lan and Sun (2019) also performed a similar study and found that first-year Chinese students were much less likely to use noun phrase modifiers than the expert writers of academic RAs. While these studies contribute significantly to advancing our understanding of the importance of noun phrase modifiers in academic writing, results are fragmented and do not give a complete picture of complexity in academic register without considering disciplinary differences (Staples & Reppen, 2016).

Table 1

Biber et al.'s (2011) Hypothesized Stages of Writing Development

Stage	Grammatical Structure	Examples from our corpus
2	Attributive adjectives Relative clauses	<u>Quantitative</u> measures Universities <u>that provided the corpus</u>
3	Nouns as pre-modifiers Possessive noun as pre-modifiers Of phrase (concrete/locative meanings) Prepositions as noun post-modifiers other than of (concrete/locative meanings)	<u>Complexity</u> measures <u>Learners'</u> development Group <u>of learners</u> Clauses <u>in the sentences</u>
4	ed-participle as post-modifiers ing-participle as post-modifiers Attributive adjectives, nouns as pre-modifiers Of phrase (abstract meanings) Prepositions as noun post-modifiers other than of (abstract meanings)	Features <u>identified through manual check</u> Measures <u>targeting different components</u> <u>Overall writing quality</u> Perception of <u>progress</u> A wealth of research <u>on writing development</u>
5	Preposition + nonfinite complement clause Complement clauses controlled by nouns Appositive noun phrases Multiple prepositional phrases as post-modifiers, with levels of embedding	Difficulty <u>in using these properties</u> Tasks that the learners accomplish EFL (<u>English as a foreign language</u>) <u>Repetition of structures in sentences with discourse-level coherence within writing productions.</u>

2.2. Writing Complexity across Academic Disciplines

Attempts to define academic disciplines can be examined from at least five perspectives, namely, philosophical, anthropological, sociological, historical, and management perspectives (see Krishnan, 2009). These perspectives differ in the emphasis they put on the nature and theory of knowledge (epistemology), cultural

practices, labor and professionalization, historical conditions and development of disciplines, and the relationship between disciplinary division of knowledge with relation to education and market demands (Abbott, 2001).

From among different research paradigms outlined above, philosophical perspective for knowledge classification seems the most pertinent to the current study's choice of applied/pure linguistic studies, since it excludes such factors as cultural identity, power structures, professionalization, the influence of discipline founders, and past success of disciplinary organization (Krishnan, 2009). Probably, the only paradigm in knowledge classification, which takes into account epistemological features of disciplines, philosophical perspective mainly concerns the nature and organization of knowledge (Russel, 2002). Conventional division of knowledge into pure and practical is rooted in philosophical paradigm (Beecher & Trowler, 2001).

Although academic writing is characterized by objective and accurate reflection of reality, it "becomes persuasive when it employs social and linguistic conventions that colleagues find convincing" (Hyland, 1999, p.99). These conventions are informed by discipline/register into which academic texts fit. This means that texts are shaped based on the recipients' sociocultural expectations, lexico-grammatical preferences, agreed-upon rules, and their sensitivity and orientation (Hyland, 2015). This recipient-oriented approach to writing reflects differences among disciplines in terms of content word classes, nouns and pronouns, semantic classes of verbs, verb phrases, discourse connectors, and dependent clauses (see Biber, 2006 & Gray, 2015).

There is a growing body of research on linguistic variations across academic disciplines (e.g., Biber, 2006; Hyland, 2010; Hyland, 2006). Researchers often comment that disciplines draw on different lexico-grammatical features to achieve their communicative purposes. It seems that this impression is well founded, since "disciplines differ in their epistemological beliefs, research practices, and knowledge structures" (Gray, 2015, p.1). Linguistic variations across disciplines emerge as a result of different expectations of the members of discourse community (Hyland, 1998). Similarly, Charles (2003) states that variation in linguistic features is due to inherent differences between the disciplines in research practices and the construction of knowledge.

The extent to which disciplinary variations may inform the use and distribution of language features has triggered a number of empirical studies investigating syntactic complexity across disciplines. For example, Lu et al. (2021) explored the

relationship between syntactic complexity and rhetorical organizations of RA introductions among four social and engineering disciplines. The results obtained from their study indicated that there existed statistically significant disciplinary variations with regard to the realization of moves and steps in RA introductions as assessed by syntactic complexity metrics. When concluding their study, they noted that disciplinary variations play an important role in form-function mapping and genre-based pedagogy.

In a similar study, Casal et al. (2021) analyzed syntactic complexity across three social science disciplines. The findings revealed the significant effect of discipline types on eight syntactic complexity measures. The findings also showed that although some measures are useful indicators of academic writing proficiency, they may not be appropriate means for exploring disciplinary variations.

2.3. Macro Structure of RAs

Although there is not a uniform format to apply to all scientific journals for publishing RAs, IMRD is probably the most established organizational structure of RAs (Swales, 1990). This format is self-explanatory (Ruiying & Allison, 2003) and identifies the sections by their overall functions (Swales, 2004). Introduction section provides the background, identifies the gap, and introduces the means for contributing to new knowledge (Swales, 1990). Methods is the section that functions as a thread to bind a particular research method with previous research procedures and to bind itself with other key sections particularly the Introduction and Results (Lim, 2006). Results is the section where the authors present their findings and seek to establish their importance (Ruiying & Allison, 2003). In the Discussion section, the authors make claims about how their results relate and contribute to disciplinary knowledge (Basturkmen, 2009).

As a functional category within the genre of RAs, IMRD has been studied from different perspectives. While Swales (1990) claimed that IMRD macrostructure is the robust format of RAs, Ruiying and Allison (2003) argue that IMRD structure cannot be taken for granted in that the headings do not always reflect the rhetorical functions of each section explicitly, and that there are variations among the authors in the use of the established IMRD format in RAs. Similar concerns have been echoed by Mondal et al. (2019), who stated that traditional IMRD format does not capture the macro structure of RAs, and they proposed “IaMRDS” format which is

the acronym for “Introduction with aim, Materials and Methods, Results, Discussion, and Conclusion”. In spite of the criticisms levelled against IMRD, it is a well-established tradition in RA studies (Ruiying & Allison, 2003). Thus, this study employed IMRD format for investigating the distribution of noun phrase modifiers across macro structures of RAs in applied linguistics and clinical medicine.

3. Methodology

3.1. Construction of the Corpus

The corpora compiled in this study include the corpus of RAs in applied linguistics (henceforth AL) and the corpus of RAs in clinical medicine (henceforth CM). All RAs followed IMRD format and were published between 2018 and 2020. In other words, we only chose those RAs in which IMRD headings were explicitly labeled. Those RAs with the merged Introduction and Literature review, Results and Discussion, and Discussion and Conclusion were excluded from the analysis. Special editions such as book reviews, forums, meta-analyses, etc., were excluded from the study, as they differ from RAs in their communicative purposes, which may motivate the variations in the use of different linguistic means. Volumes, issues, and articles were chosen randomly. Before the main phase of analysis, we performed data cleansing that included correcting inaccurate data, and removing tables, titles, headings, figures, and footnotes. This was done using a special computer program written in Python environment.

The articles were randomly selected from peer-reviewed journals of *Applied Linguistics*, *Language Learning*, *TESOL Quarterly*, *Modern Language Journal*, and *English for Specific Purposes* in applied linguistics and from *Molecular Psychiatry*, *American Journal of Psychiatry*, *British Journal of Psychiatry*, *Schizophrenia Research*, and *Journal of Clinical Psychiatry* in medicine. The inclusion of these journals was based on the criteria of their publication history, their index in the major bibliographic databases, and h-index which is defined as the number of the researchers’ papers (h) that has been cited at least (h) times (Barnes, 2017). That is, a researcher who has published 20 articles each with 20 citations has an h-index of 20. The advantages of using the h-index over the previous journal impact factor (JIF) measures are that the researchers and the journals’ credibility is assessed based on both their productivity and their recognition among the communities of scholars. In addition, the h-index is a transparent measure of scholarly impact that minimizes

biases which may occur due to idiosyncratic over-citations (Ruscio, 2016).

Totally, the researchers compiled a corpus of 389,332 words from the leading journals in applied linguistics and clinical medicine. Thirty RAs from applied linguistics journals and 50 RAs from clinical medicine journals were selected. Descriptive features of the two corpora are presented in Table 2.

Table 2.
Descriptive Statistics of two Corpora of Applied Linguistics and Clinical Medicine

Statistics	Group	Introduction	Methods	Results	Discussion	Total
Words	Applied linguistics	50,150	50,085	47,436	40,932	188,603
	Clinical medicine	28,343	60,725	48,148	63,513	200,729
Mean	Applied linguistics	1671.67	1669.50	1581.20	1411.45	6286.77
	Clinical medicine	566.86	1214.50	962.96	1270.26	4014.58
Standard deviation	Applied linguistics	708.82	730.79	946.67	537.09	1667.51
	Clinical medicine	204.64	362.44	410.02	345.68	579.11

3.2. Grammatical Features of Interest

The present study aimed to investigate 15 noun phrase modifiers as identified by Biber et al. (2011). Noun phrase modification features are obtained from developmental stages of syntactic complexity proposed by Biber et al. (2011). The developmental index entails five stages which are categorized based on three grammatical types: Finite dependent clauses, nonfinite dependent clauses, and dependent phrases. In this study, the purpose was to examine (1) finite dependent clauses including relative clauses as noun modifiers, complement clauses controlled by nouns; (2) nonfinite dependent clauses including, ing- and ed-participles as noun post-modifiers, and preposition + nonfinite complement clauses as post-modifiers; and (3) dependent phrases including attributive adjectives, participles, nouns as pre-modifiers, possessive nouns, of phrases as noun post-modifiers, other prepositional phrases as noun post-modifiers, adjectives, noun as pre-modifiers, appositives, and

multiple prepositional phrases as noun post-modifiers.

3.3. Identification of Noun Phrase Modifiers

The current study sought to explore noun phrase modifiers identified by Biber et al., (2011). They include attributive adjectives, participles, nouns as pre-modifiers, possessive nouns, of phrases as noun post-modifiers, other prepositional phrases as noun post-modifiers, adjectives, noun as pre-modifiers, appositives, and multiple prepositional phrases as noun post-modifiers. These lexico-grammatical features were identified through automatic analysis and manual check of the data. Automatic analysis was employed, because it allowed for the investigation of large corpora with high degree of reliability. Automatic extraction tools, however, are not able to identify particular language features such as prepositional phrases, or zero relativizers. As a result, by employing automatic and manual check of the data, we were able to identify and extract all noun phrase modifiers proposed by Biber et al. (2011).

Automatic analysis of the data was carried out using a computer program called Stanford Core NLP Version 3.9.2. It is a tool for natural language processing in Java and Python environment. Stanford Core NLP has a number of features such as text tokenization, stemming and lemmatization, POS tagging, constituency parsing, named entities, sentiment analysis and so forth. Stanford Core NLP is normally used to assign parts of speech to the words, which is a necessary stage before the main phase of analysis. Then, using a special Perl program, it was possible to extract and count the number of noun phrase modifiers. Depending on the text types (native or non-native), the accuracy of Stanford Core NLP is reported to be between 97.21, and 97.67 (Manning, 2015).

3.4. Data Analysis

Table 3 is an example of how NP modifiers in our corpus were identified by the script which was written in Python environment using NLP: The imaginary situation evoked by task was also found to bring about different means of learner involvement. As shown in Table 3, sentence was the unit of analysis in our study. In other words, the text has been tokenized into sentences for further analysis which were dependency parsing and constituent processing.

Table 3*Noun Phrase Modifiers Identified by Automatic Extraction Tool*

Noun Phrase	Count	Text
Noun + noun	1	1- learner involvement
Adjective + noun	2	1- imaginary situation
Adjective + noun	2	2- different means
Noun + past participle	1	1- situation evoked
Noun + <i>of</i> preposition	1	1- means of

Our further investigation involved the qualitative check of problematic target forms which included those features with low accuracy rate in automatic processing (i.e., NP post modifiers and zero relativizers). Two experienced coders (the present researchers) conducted the manual check of the data. As the first step in the analysis, 10 percent of the corpus was rated by the coders independently in order to identify the linguistic features of interest. The inter-coder agreement was estimated to be 0.8734. Then, after resolving uncertain cases, the remaining texts were coded and the reliability rate of nearly 95% was achieved which was high according to inter-coder agreement reports of other studies (e.g., Covington et al., 2006). The remaining differences were further discussed until a complete agreement was reached.

After identification of noun phrase modifiers, they were normalized to 1000 words. This practice has the advantage of allowing researchers to compare the texts of different lengths (Biber & Barbieri, 2007) and to employ parametric tests (Biber et al., 2011). The normalized data were put into SPSS (Version 25). Multiple one-way analysis of variance (ANOVA) procedures were run to find differences across different sections of RAs in each of the identified features. It was followed by Pairwise comparisons for all significant results. Independent-samples *t*-tests were also used to compare the frequency of occurrence of noun phrase modifiers between two writer groups in applied linguistics and clinical medicine. Given multiple tests involved, Bonferroni correction was employed to adjust for the *p*-values.

4. Results

From among 15 pre- and -post noun phrase modifiers identified in Biber et al.'s (2011) hypothesized stages of writing development, the researchers chose six modifiers of “relative clauses”, “ed-clause post-modifiers”, “ing-clause post-

modifiers”, “prepositional phrases”, “attributive adjectives”, and “noun pre-modifiers” for cross-comparisons across sections. There are two reasons why these six modifiers from among the others were chosen. First, some of them did not distribute normally across the sections in two disciplines, making it impossible to employ parametric tests for cross comparisons. Second, the researchers decided not to include those extremely infrequent features like “appositive noun phrases” (3 per 1000 words) as identified by Biber et al. (1999).

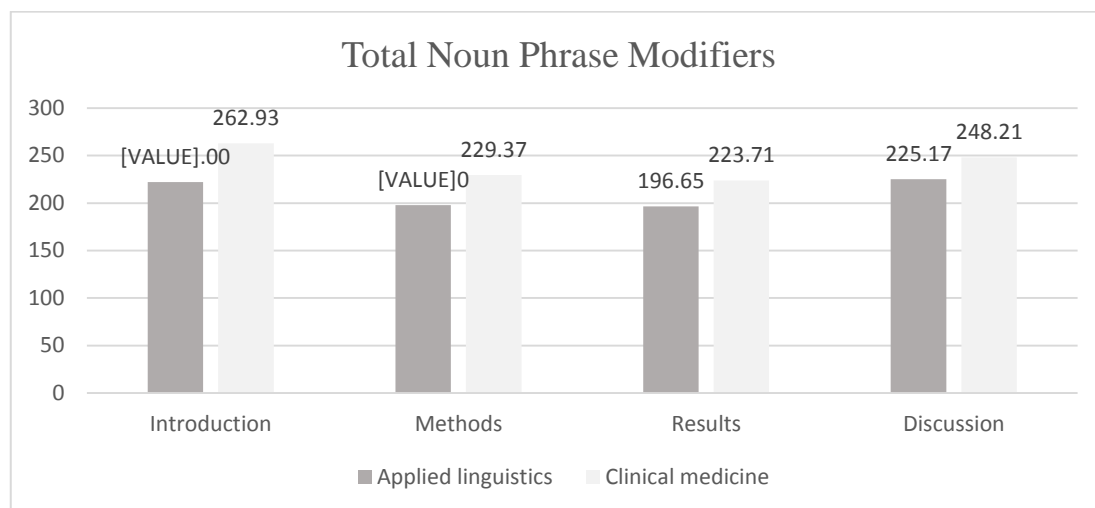
In order to answer the first research question, the distribution of noun phrase modifiers across different sections of research articles, namely, introduction, method, result, and discussion (IMRD) was compared between applied linguistics and clinical medicine. The results indicated that noun phrase modifiers are not evenly distributed across sections in both disciplines. Introduction and discussion noticeably embody a larger number of modifiers compared to the two other sections in two disciplines (See Appendix A for more detail). In applied linguistics, discussion placed the heaviest reliance noun phrase modifiers followed by introduction, methods and results.

However, in clinical medicine, it was introduction that included the largest number of noun phrase modifiers, followed by discussion, methods, and results respectively. Independent samples *t*-tests were used to check whether the differences between two disciplines in the use of noun phrase modifiers were statistically significant and to determine the effect size (eta squared). Because we used five tests on the same dataset simultaneously, Bonferroni post-hoc adjustment was used to adjust the alpha level, which was set at $p < 0.01$ after correction ($0.05/5 = 0.01$).

As shown in Figure 1, irrespective of the sections, the RA writers in clinical medicine used noun-modifying features of all types much more frequently than those in applied linguistics did. Normalized mean frequencies of modifiers in each section in two groups of writers are as follows: Introduction, 262.93 vs 222.00; methods, 229.37 vs 197.70; results, 223.71 vs 196.65; discussion, 248.21 vs 225.17.

Figure 1

Normalized Mean Frequency of Distribution of Total Noun Phrase Modifiers in Two Groups of Writers



The results obtained from independent-samples *t*-tests, as shown in Table 4, suggest that the groups differed significantly in the use of noun phrase modifiers in all IMRD sections of RAs in applied linguistics and clinical medicine. RAs in clinical medicine relied on the larger number of noun phrase modifiers in Introduction (Mean difference = 40.93; $t(78) = 10.20$, $p = 0.000$), Methods (Mean difference = 31.67; $t(78) = 6.98$, $p = 0.000$), Results (Mean difference = 27.06; $t(78) = 4.53$, $p = 0.000$), Discussion (Mean difference = 23.04; $t(78) = 2.79$, $p = 0.006$), and total (Mean difference = 30.67; $t(78) = 6.61$, $p = 0.000$).

Table 4

Independent-samples t-test to Compare Distribution of Noun Phrase Modifiers in Applied Linguistics and Clinical Medicine

Section	Group	Mean	Sig. (2-tailed)	Eta squared
Introduction	Applied Linguistics	222.00	0.000	0.23
	Clinical medicine	262.93	0.000	
Methods	Applied Linguistics	197.70	0.000	0.36
	Clinical medicine	229.37	0.000	
Results	Applied Linguistics	196.65	0.000	0.21
	Clinical medicine	223.71	0.000	
Discussion	Applied Linguistics	225.17	0.000	0.09
	Clinical medicine	248.21	0.006	
Total	Applied Linguistics	210.38	0.000	0.38
	Clinical medicine	241.05	0.000	

In order to answer the second research question which seeks to check the difference in the frequency of noun phrase modifiers in each section, the researchers employed six one-way ANOVA procedures (one for each modifier) in each discipline. Since multiple comparisons were used, Bonferroni post-hoc adjustment was used to adjust the alpha level which is set at $p < 0.0083$ after correction ($0.05/6 = 0.0083$). After within-discipline analysis, the researchers also performed between-discipline analysis by comparing the frequency and types of noun phrase modifiers in each section. The results obtained from one-way ANOVA (Table 5) revealed statistically significant differences for two of six non phrase modification types in applied linguistics: attributive adjectives and relative clauses (See Appendix B for more details). Other noun modifiers were not statistically significant with regard to their frequency difference across sections. In clinical medicine, on the other hand, attributive adjectives, noun pre-modifiers, relative clauses, and prepositional phrases differed statistically significantly with regard to their frequency across sections.

Table 5

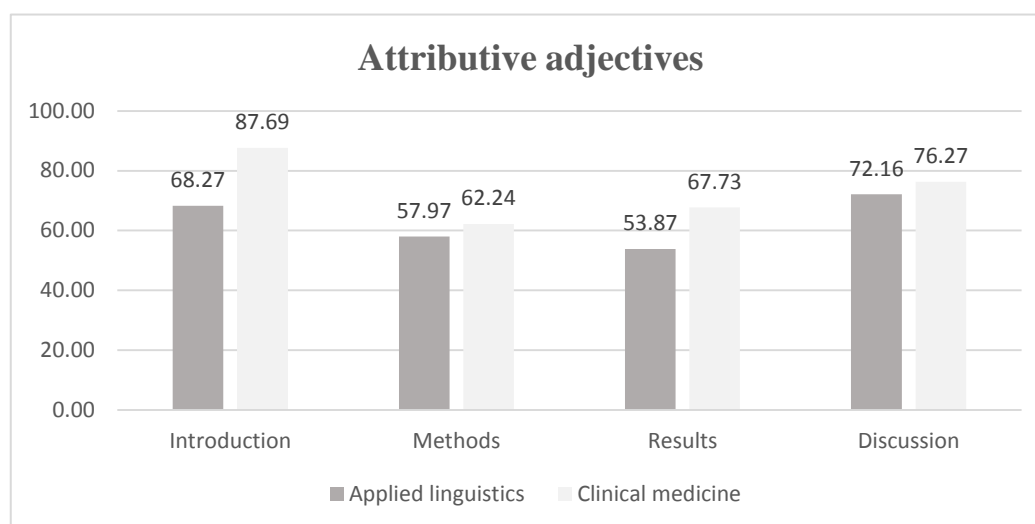
Differences in the Mean Frequencies of Noun Modifiers across IMRD Sections of RAs in Two Disciplines of Applied Linguistics and Clinical Medicine

Group	Grammatical Structure	F	Sig.	η^2
Applied linguistics	Attributive adjectives	6.00	0.001*	0.13
Clinical medicine		23.56	0.000*	0.27
Applied linguistics	Noun pre-modifiers	0.83	0.476	0.02
Clinical medicine		20.00	0.000*	0.14
Applied linguistics	Relative clauses	5.90	0.001*	0.13
Clinical medicine		6.71	0.000*	0.09
Applied linguistics	Prepositional phrases	2.39	0.072	0.06
Clinical medicine		12.96	0.000*	0.17
Applied linguistics	ed-clause post-modifiers	1.39	0.211	0.02
Clinical medicine		1.53	0.206	0.02

*The results are significant at 0.0083.

The first lexico-grammatical feature, which demonstrated a statistically significant difference across the sections in both disciplines, was attributive adjectives. The results of post-hoc Tukey HSD tests concerning its different patterns of distribution are represented in Appendix B and Appendix C. The mean frequencies of attributive adjectives in introduction and discussion sections were higher than those of methods and results in both disciplines (Figure 2). As table 5 demonstrates, in applied linguistics, the frequency of attributive adjectives in discussion and results sections differed significantly ($p = 0.002$, < 0.0083).

Figure 2
Normalized Mean Frequency Distribution of Attributive Adjectives in Two Groups of Writers



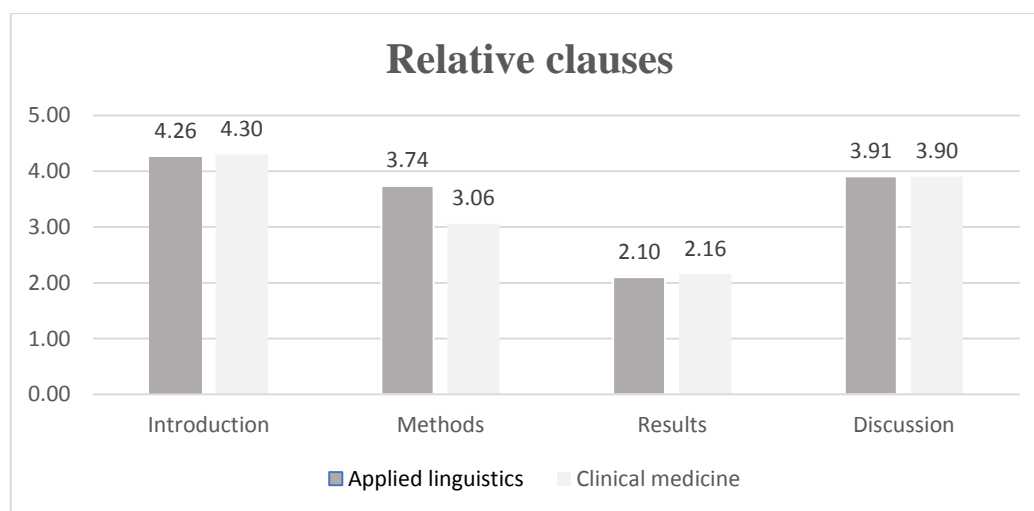
When it comes to clinical medicine, the introduction section differed significantly from other three sections (introduction vs methods, $p = 0.000$, introduction vs results, $p = 0.000$, introduction vs discussions, $p = 0.003$). Mean frequency of attributive adjectives in methods section, in the same vein, differed significantly from introduction and discussion ($p = 0.000$). The mean frequency in results section differed significantly only from introduction section ($p = 0.000$).

Relative clauses were the second statistically significant feature in applied linguistics. As shown in Figure 3, in applied linguistics and clinical medicine, relative clauses were used more frequently in Introduction followed by Discussion,

Methods and Results. However, the only sections that reached significance in terms of frequency of relative clauses were introduction and results ($p = 0.001$). In clinical medicine, introduction vs results ($p = 0.000$), and the results vs discussion were statistically significant ($p = 0.005$).

Figure 3

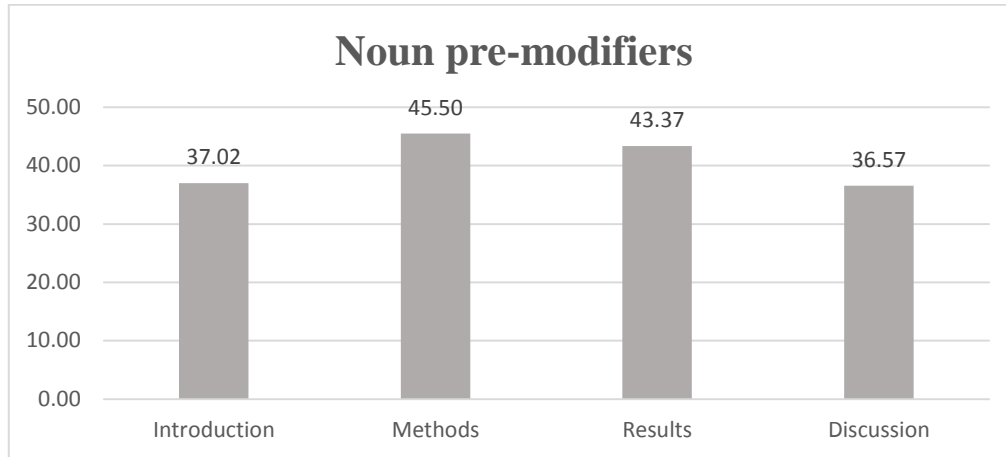
Normalized Mean Score of Distribution of Relative Clauses in Two Groups of Writers



In clinical medicine noun pre-modifiers were used unevenly across sections (Figure 4), with Methods and Results displaying more frequent uses of this lexicogrammatical feature than Introduction and Discussion. The difference between introduction vs methods ($p = 0.001$), and methods vs discussion ($p = 0.000$) reached significance.

Figure 4

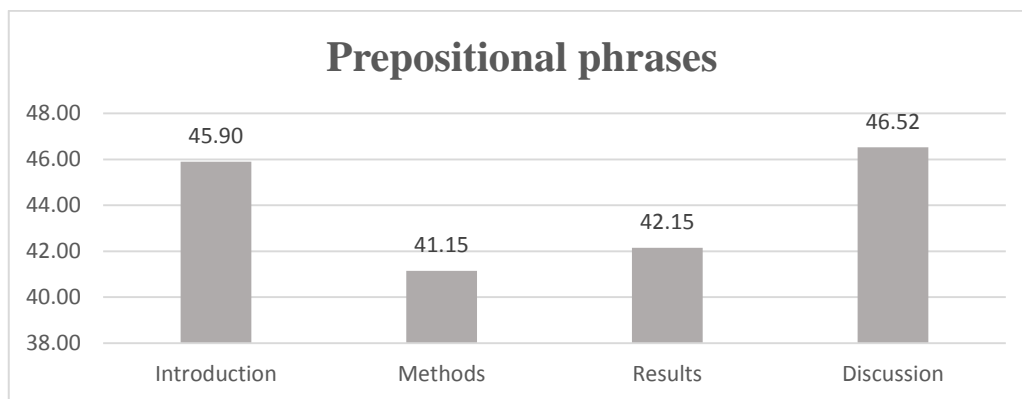
Normalized Mean Score of Distribution of Noun Pre-Modifiers in Clinical Medicine Group.



The last significant feature in clinical medicine was prepositional phrases. According to Figure 5, the use of prepositional phrases in Introduction and Discussion was denser than that of Methods and Results. The difference between introduction vs methods ($p = 0.001$), introduction vs results ($p = 0.000$), methods vs discussion ($p = 0.000$, and results vs discussion ($p = 0.000$) reached significance.

Figure 5

Normalized Mean Score of Distribution of Prepositional Phrases in Clinical Medicine Group.



5. Discussion

The present study has shown that RAs in clinical medicine placed heavier reliance on noun phrase modifiers than RAs in applied linguistics did. Unlike hard sciences, soft sciences are characterized by evaluative model of discourse where the persuasion is more interpretive and less explicit (Hyland, 2008). In soft sciences, the researchers rarely deal with new phenomena in the discussion of their experiences, because they discuss the phenomena that are already familiar to us, reflecting little need for new technical vocabulary and complex grammatical constructions (Biber & Gray, 2016). In the same vein, Gardner et al. (2019) argue that “hard sciences are more informational, while Humanities disciplines are more involved” (p. 3). This difference suggests that the writers in humanities disciplines tend to bolster their claims by more elaborated discussions through greater use of clausal features, while phrasal features used in hard sciences (especially pre-modifiers) ensure the communication of a great deal of technical information more concisely (Staples & Reppen, 2016).

The results of current study lend support to those of Biber and Gray (2016), who documented that science research writing is an outlier with regard to structural compression style of discourse in the continuum of academic disciplines with the other end being humanities. Therefore, noun phrase modifiers are not considered textual features characterizing the genre of academic writing generally; rather, they are “discipline-specific writing conventions” (See Baratta, 2010) with certain disciplines such as hard sciences being structurally more compact than others.

The second finding of the study was that the frequency of attributive adjectives differed significantly from section to section in clinical medicine (with the exception of method vs results); however, in applied linguistics, the only significant difference was between results and discussion. The difference in the frequency of attributive adjectives across RA sections may be related to discourse function and chief aim of each section. The very fact that the introduction and discussion sections in both disciplines contain a larger number of normalized attributive adjectives is not against our expectations, since as Biber et al. (1999) argued, attributive adjectives in academic writing are mainly used for descriptive, evaluative, topical, and relational functions. Unlike methods and results that follow a more straightforward and explicit paradigm, introduction and discussion are evaluative in nature.

The similar frequency of occurrence of attributive adjectives in introduction and

discussion sections is directly related to their discourse functions which are mutually interdependent. The discussion picks up where the introduction leaves off (Annesley, 2010). Introduction and Discussion are directly related by means of the research questions formulated in Introduction, and the authors usually find it an effective strategy to use the same wording to reiterate their purpose in the discussion (Bavdekar, 2015). Attributive adjectives, with their diverse language functions, facilitate the link between the two sections.

Relative clause was the next noun phrase modifier whose distribution was significantly different across IMRD sections in two groups. In applied linguistics, relative clauses were distributed significantly unevenly in introduction and results. In the same vein, in clinical medicine, the normalized frequency of relative clauses differed significantly between introduction and results. In order to account for different frequency counts of relative clauses across RA sections, we need to get rid of traditional categorization of “defining vs non-defining” or “restrictive vs non-restrictive” relative clauses” (see Tse & Hyland, 2010). Restrictive (or defining) relative clauses serve to restrict the possible interpretations of the reference of the noun phrase while non-restrictive (non-defining) relative clauses tend to add information to preceding noun phrases (Cowan, 2008). While this categorization seems to depict surface grammatical structure of relative clauses, it hardly explains their frequency in academic writing, more specifically when it comes to IMRD sections.

In order to have a clearer picture of distributional pattern of relative clauses in the current study, the researchers employed functional-rhetorical characterization of relative clauses as proposed by Tse and Hyland (2010), where the use of relative clauses in academic writing is explained in relation to the nouns modified and type of modification. Although their categorization was originally used to assess journal descriptions as a distinct genre, they suggested that their model be used to assess any kinds of academic texts. A distinct subcategory of the model is the evaluative function of relative clauses which “contributes to the meaning of the modified entity both in terms of its objective property as well as the value judgements ascribed to it” (p.12).

A closer analysis of noun phrases in our corpus revealed that nearly 43 percent of noun phrases in introduction section of applied linguistics RAs were modified, while for methods, results and discussion they were 35, 33, and 40 percent,

respectively. In clinical medicine, however, the ratio of modified noun phrases to total noun phrases for IMRD section is 0.56, 0.43, 0.38, and 0.52, respectively. Relative scarcity of relative clauses in results section in relation to other sections could be expected, since results is the space where the writers try to present the findings of their studies in much the same way as they are without any personal judgements that readers may think of clear, understandable, and unequivocal evidence obtained from statistical analyses.

On the other hand, introduction section helps readers to move from general research topic with the aim of centrality claim, to gap identification and description of the organizational structure of the current paper (Swales, 1990). Like journal description genres, RA introduction texts establish scope and specialty of the topics to be discussed or welcomed. Introduction, after abstract, is the second part that editors and reviewers read and it determines whether they are likely to continue reading or not (Abrahamson, 2008). Accordingly, introduction may mirror “hidden promotional element of academic texts” (Tse & Hyland, 2010, p.23) in a highly competitive field of activity. For academic writers who want to sell their academic work to their readership, to delimit the scope of their research and subsequently to establish and occupy the niche (Swales, 1990), relative clauses are invaluable tools which “ensure semantic clarity and textual variety” (Cho & Lee, 2016). The evaluative function of relative clauses enables writers to both establish the niche and create the positive attitude in readers to the things discussed (Tse & Hyland, 2010). This is a unique function of relative clauses, which may not be fulfilled by other noun phrase modifiers. Extract 2 shows how academic writers used relative clauses to establish niche and positive attitudes in readers.

Text Extract 2

In particular, this study focuses on learning of the meanings of new technical words that appear in a written context with explicit clues.
(Applied linguistics)

Identifying potentially malleable environmental factors that may alter the developmental course of heritable mental health problems is an important step toward guiding prevention strategies.
(Clinical medicine)

Abstract nouns are distinctive features of advanced academic writing (Biber, 2006; Biber et al. 1998). Taking macrostructure of research articles into account,

the researchers found more abstract nouns in introduction section compared to results section which normally relies on concrete nouns (in current study, abstract nouns were used nearly twice as frequently in introduction section as in results section). Evidently, unlike concrete nouns, abstract nouns generally require more clarification. While abstract nouns can be modified by pre-modifiers like attributive adjectives, and noun pre-modifiers, relative clauses merit special attention since they “open up more space for detailed description” (Tse & Hyland, 2010, p.4), something which seems necessary for establishing and occupying the niche. Against this backdrop, it may safely be argued that relative clauses in RAs are “primed” to occur in introduction section in RAs (see Hoey, 2005).

In clinical medicine, the difference in the mean frequency of noun pre-modifiers between introduction vs methods, and methods vs discussion was statistically significant. In other words, pre-modifying nouns were higher in methods and results sections than introduction and discussion. The results obtained from this section of our analyses ran contrary to our initial expectations. Significantly higher proportion of noun pre-modifiers in methods section compared to introduction and discussion sections in our study could be analyzed in relation to the rhetorical purpose of these sections in RAs. Introduction and method sections are understood to be argumentative (Hood, 2004), while the method section is supposed to be expository (Martinez, 2003). Expository texts expose readers to the facts, reasons and evidences of a particular topic. It follows a pattern of development that encompasses examples, analyses, processes, classifications and so forth (Richards & Schmidt, 2002). It is the processes, facts, and evidences that are foregrounded in methods section and, as opposed to introduction and discussion sections, there is little room for individuals to establish their own ‘niche of expertise’ (Tse & Hyland, 2010, p.13) and individualize the unique standing of their field of inquiry.

Biber et al. (1999) state that noun pre-modifiers are strongly favored in academic writing, because they can establish versatile meaning relations. Qualitative analysis of noun + noun sequences in methods section of our corpus revealed that a vast majority of the sequences were section-specific noun phrases with some others being discipline-specific, that is, noun phrases that are commonly found in a particular discipline, or a section of RAs, confirming Hyland and Tse’s (2007) argument that “all disciplines shape words for their own uses” (p.240). In either case, noun sequences associated with each other by means of content relations facilitate the smoothness of procedures, analyses, and information exchange, which

are supposed to be defining rhetorical features of methods sections. Extract 3 demonstrates section-specific and discipline-specific functions of noun + noun sequences employed by academic writers in clinical medicine.

Text Extract 3

The schizophrenia data was imputed using the SHAPEIT and IMPUTE2 software programs.

Genomic principal component scores were also covaried to control for population stratification in the ZHH-FE.

Extract 3 shows how noun pre-modifiers are used in method sections of clinical medicine RAs to illustrate the means, procedures, phenomena, and objectives. The first two are representative of discipline-specific noun sequences, and the others highlight section-specific noun phrases. Brevity and implicitness are two defining features of noun + noun sequences. Although, due to lack of function words, noun sequences could be a potential source of ambiguity for readership (Biber et al. 1999), the academics, because of their familiarity with the genre, do not usually have difficulty deciphering the meaning relations. On the other hand, the density of information communicated by noun sequences makes them ideal language features to be used in methods section. In sum, noun sequences in methods section assist writers in “convincing the readership of the validity of the means employed to obtain findings” (Lim, 2006) while meeting the challenge of limited space requirement imposed by journals in the new age of science.

While noun pre-modifiers were used in introduction and discussion sections as well, because of the argumentative nature of the introduction and discussion sections (Hood, 2010) in RAs, other types of pre-modifiers (ed-participial modifiers, ing-participial modifiers, and general adjectives) were used as frequently as, or even more frequently than, noun pre-modifiers. In other words, different rhetorical-communicative functions of the two sections explain their different degrees of reliance on noun pre-modifiers. These findings confirm those of Hong et al. (2017), who reported similar results in International Business Management RAs.

Post-modifying prepositional phrases are the last significant language features whose average distributional difference was significant across sections in clinical medicine. As was expected, higher mean frequencies of prepositional phrases were found in introduction and discussion sections than in methods and results. Introduction and results are the sections where the researchers underscore the new

knowledge generated by their study, which may contribute to the field. As a result, these two sections could be placed in the argumentative end of the cline and results and discussions in the expository end of the cline.

Because of the role of prepositional phrases as post-modifiers, the higher proportion of post-modifying prepositional phrases in introduction and discussion sections in relation to methods and results may be accounted for, using the principle of end-weight, which is the tendency to place longer, more complex structures towards the ends of a clause. Biber et al. (1999) maintain that the principle of end-weight is closely related to postponement of direct object before the object predicative. They further argue that postponement occurs particularly in the registers with highest degree of phrasal complexity. Introduction and discussion, as our study and other studies showed, were more complex in terms of noun modifying phrasal features compared to methods and results sections. As a result, uneven distribution of post-modifying prepositional phrases across IMRD sections of medical research articles can be justified.

6. Conclusion

The results of the first research question indicated that, overall, the academic writers in clinical medicine used noun phrase modifiers more frequently in all sections of RAs than the writers in applied linguistics did. Within-disciplinary analysis of distribution of noun phrase modifiers showed that, in applied linguistics, there were significant differences among the sections in the use of the phrasal modifiers of attributive adjectives and relative clauses. Other noun phrases were not statistically significant. In clinical medicine, modifying features of attributive adjectives, noun pre-modifiers, relative clauses, and prepositional phrases reached significance across sections. Between-section analyses using post-hoc Tukey HSD tests revealed that, both in applied linguistics and clinical medicine, attributives adjectives and relative clauses were used more frequently in introduction and discussion sections than in methods and results sections.

This study examined academic genres in two disciplines of applied linguistics and clinical medicine. The results revealed that irrespective of discipline, noun phrase modifiers were not distributed evenly across IMRD sections of research articles, with introduction being the most complex, followed by discussion, results, and methods. This indicates that the variation of noun phrase modifiers across

sections is discipline-independent, and is due to functional attributes of the sections. Thus, RAs are not a rigid genre with similar distribution of noun phrase modifiers across sections; rather, sub-registers of Introduction, Methods, Results, and Discussion inform the pattern of phrasal distribution across disciplines.

The present study showed that academic writers in clinical medicine placed more reliance on noun phrase modifiers than their counterparts in applied linguistics did. Therefore, it seems that noun phrase complexity in academic writing is influenced by the factor of discipline with those in hard/applied end of the cline being more complex than those in soft/pure end of the cline. Although the two disciplines fulfill the communicative purpose of reporting the research and belong to the same genre of academic research articles, the communicative purpose may be too general for thorough investigation of noun phrase complexity. Accordingly, various sub-genres that are produced in the context of different disciplinary courses (Samraj, 2004) may have some impact on the noun phrase complexity of a text constructed by academic writers.

The results of the study may have important implications for language pedagogy. Our study has shown that hard science relies on noun phrase modifiers more heavily than soft science. Moreover, the pattern of reliance of the two disciplines on noun phrase modifiers was different. Soft science employed almost equal number of pre- and postmodifiers, while hard science utilized far larger number of pre-modifiers for modifying noun phrases. This suggests that hard science is, by its very nature, more compact than soft science. L2 academic writers in the field of hard sciences may need to be explicitly and contextually aware of the ways by which pre-modifiers contribute to compact discourse style.

The use of nouns as pre-modifiers, in particular, seems to be somewhat challenging especially for novice academic writers who need to work out diverse meaning relations between the modifier and the modified. Since most noun sequences are discipline-specific (Elliot, 2019), it might be useful to raise the students' awareness of these lexico-grammatical features in academic writing and contextually explore the rhetorical functions that these sequences fulfill within disciplines.

Uneven distribution of noun phrase modifiers across RA sections (introduction and discussion place heavier reliance on noun phrase modifiers than methods and results), and the use of section-specific modifiers stress the genre-based approach to academic writing courses which underline genre conventions of

IMRD sections in RAs. Three-phase teaching-learning approach which consists of deconstruction of “sample” “expert” text stage, joint construction stage, and individual construction stage (Martin, 2009) may prove promising in genre-based teaching of writing whereby learners can enhance their knowledge of text types in terms of language and text features.

In addition to language pedagogy, the findings of the present study may benefit syllabus designers and curriculum developers. As our study showed, two disciplines of clinical medicine and applied linguistics deployed different patterns of lexico-grammatical resources for complexifying academic texts. As Hyland (2002) pointed out, common-core features of academic prose differ considerably in frequency, expression, and function across disciplines. This highlights the importance of designing discipline-specific academic writing courses (writing in disciplines) rather than general EAP courses for undergraduate/high school students.

The present study had two limitations that should be considered when interpreting the findings. This study explored the distribution of noun phrase modifiers across IMRD sections of RAs in applied linguistics and clinical medicine. However, IMRD is only one format for investigating text-intrinsic language features (those features which are attributed to macro-structures of the texts) of RAs. Future studies may employ other formats (see Mondal et al., 2019) that accurately capture the macrostructures of RAs. The present study investigated only one dimension of language complexity (phrasal complexity). It might be an interesting area of research for future studies to examine other dimensions such as morphological complexity, collocational complexity, etc. and compare them across text-intrinsic sections of RAs across disciplines.

References

- Abbott, A. (2001). *Chaos of disciplines*. University of Chicago Press.
- Ahlstrom, D. (2017). How to publish in academic journals: Writing a strong and organized introduction section. *Journal of Eastern European and Central Asian Research*, 4(2), 1–9.
- Annesley, T. M. (2010). The Discussion section: Your closing argument. *Clinical Chemistry* 56(11), 1671–1674.
- Baratta, A. M. (2010). Nominalization development across an undergraduate academic degree program. *Journal of Pragmatics*, 42(4), 1017–1036.
- Barnes, C. (2017). The h-index debate: an introduction for librarians. *The Journal of Academic Librarianship*, 43(6), 487–494.
- Basturkmen, H. (2009). Commenting on results in published research articles and masters dissertations in Language Teaching. *Journal of English for Academic Purposes*, 8(4), 241–251.
- Bavdekar, S. B. (2015). Writing introduction: Laying the foundations of a research paper. *Journal of the Association of Physicians of India*, 63 (7), 44–6.
- Becher, T., & Trowler, P.R. (2001). *Academic tribes and territories*. Open University Press.
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5(2), 97–116.
- Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, 9(1), 2–20.
- Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5–35.

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. Longman.
- Casal, J. E., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). Syntactic complexity across academic research article part-genres: A cross-disciplinary perspective. *Journal of English for Academic Purposes*, 52, 100996.
- Charles, M. (2003). 'This mystery...': a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes*, 2(4), 313–326.
- Cho, D. W., & Lee, K. (2016). English relative clauses in science and engineering journal papers: A comparative corpus-based study for pedagogical purposes. *Ampersand*, 3, 61–70.
- Covington, M. A., He, C., Brown, C., Naçi, L. & Brown, J. (2006). *How complex is that sentence? A proposed revision of the Rosenberg and Abbeduto D-Level scale*. CASPR Research Report 2006-01. The University of Georgia, Artificial Intelligence Center.
- Cowan, R. (2008). *The teacher's grammar of English*. Cambridge University Press.
- Dahl, T. (2004). Textual metadiscourse in research articles: a marker of national culture or of academic discipline? *Journal of Pragmatics*, 36(10), 1807–1825.
- Elliott, T. (2019). *Variation in use of nouns as nominal premodifiers in advanced student writing across academic disciplines* [Master's thesis, Iowa State University].
- 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.
- Gray, B. (2015). *Linguistic variation in research articles*. John Benjamins Publishing Company.
- Halliday, M. A. K. (1993). Some grammatical problems in scientific English. In M. A. K. Halliday, & J. R. Martin (Eds.), *Writing science: Literacy and discursive power* (pp. 76–94). The Falmer Press.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.

- Hong, A. L., Hua, T. K., & Mengyu, H. (2017). A Corpus-based Collocational Analysis of Noun Premodification Types in Academic Writing. *3L: Southeast Asian Journal of English Language Studies*, 23(1), 115–131.
- Hood, S. (2010). *Appraising research: Taking a stance in academic writing*. Palgrave MacMillan.
- Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins Publishing.
- Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, 20(3), 341–367.
- Hyland, K. (2002). Genre: Language, context, and literacy. *Annual Review of Applied Linguistics*, 22(1), 113–135.
- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4–21.
- Hyland, K. (2010). Constructing proximity: Relating to readers in popular and professional science. *Journal of English for Academic Purposes*, 9(2), 116–127.
- Hyland, K. (2015). Genre, discipline and identity. *Journal of English for Academic Purposes*, 19, 32–43.
- Kanoksilapatham, B. (2015). Distinguishing textual features characterizing structural variation in research articles across three engineering sub-discipline corpora. *English for Specific Purposes*, 37, 74–86.
- Krishnan, A. (2009). *What are academic disciplines? Some observations on the disciplinarity vs. interdisciplinarity debate*. National Centre for Research Methods.
- Lan, G., Lucas, K., & Sun, Y. (2019). Does L2 writing proficiency influence noun phrase complexity? A case analysis of argumentative essays written by Chinese students in a first-year composition course. *System*, 85, 102–116.
- Lim, J. M. H. (2006). Method sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, 25(3), 282–309.

- Lu, X., Casal, J. E., Liu, Y., Kisselev, O., & Yoon, J. (2021). The relationship between syntactic complexity and rhetorical move-steps in research article introductions: Variation among four social science and engineering disciplines. *Journal of English for Academic Purposes*, 54, 1-13. <https://doi.org/10.1016/j.jeap.2021.101006>.
- Manning, C. D. (2015). Computational linguistics and deep learning. *Computational Linguistics*, 41(4), 701–707.
- Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education*, 20(1), 10–21.
- 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.
- Mondal, H., Mondal, S., & Saha, K. (2019). What to Write in Each Segment of an Original Article?. *Indian Journal of Vascular and Endovascular Surgery*, 6(3), 221–221.
- Moreno, A. I., & Swales, J. M. (2018). Strengthening move analysis methodology towards bridging the function-form gap. *English for Specific Purposes*, 50, 40–63.
- Parkinson, J. (2011). The Discussion section as argument: The language used to prove knowledge claims. *English for Specific Purposes*, 30(3), 164–175.
- Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for Academic Purposes students. *Journal of English for Academic Purposes*, 14, 48–59.
- Richards, J. C., & Schmidt, R. (2002). *Longman dictionary of applied linguistics and language teaching*. Longman.
- Ruiying, Y., & Allison, D. (2003). Research articles in applied linguistics: Moving from results to conclusions. *English for Specific Purposes*, 22(4), 365–385.
- Ruscio, J. (2016). Taking advantage of citation measures of scholarly impact: Hip Hip h Index! *Perspectives on Psychological Science*, 11(6), 905–908.
- Russell, D. R. (2002). *Writing in the academic disciplines: A curricular history*. SIU Press.

- Samraj, B. (2004). Discourse features of the student-produced academic research paper: variations across disciplinary courses. *Journal of English for Academic Purposes*, 3(1), 5-22.
- Samraj, B. (2005). An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for Specific Purposes*, 24(2), 141–156.
- Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexicogrammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, 32, 17–35.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Tse, P., & Hyland, K. (2010). Claiming a territory: Relative clauses in journal descriptions. *Journal of Pragmatics*, 42(7), 1880–1889.
- Yin, S., Gao, Y., & Lu, X. (2021). Syntactic complexity of research article part-genres: Differences between emerging and expert international publication writers. *System*, 97, 102427.

Appendices

Appendix A

Mean Raw Frequency and Mean Normalized frequency of Noun Phrase Modifiers across Sections

Language feature	Group	No. of texts	Introduction	Methods	Results	Discussion
Attributive adjectives	Applied linguistics	30	112.60*	96.30*	87.63*	98.59*
	Clinical medicine	50	68.27**	57.97	53.87**	72.16**
Relative clauses	Applied linguistics	30	49.71*	75.99*	67.73*	96.88*
	Clinical medicine	50	87.69**	62.24**	65.22**	76.27**
Noun pre-modifiers	Applied linguistics	30	7.73*	6.24*	3.32*	5.34*
	Clinical medicine	50	4.26**	3.74**	2.10**	3.91**
Possessives	Applied linguistics	30	2.44*	3.72*	2.16*	4.95*
	Clinical medicine	50	4.30**	3.06**	2.08**	3.90**
Of prepositional phrases	Applied linguistics	30	50.40*	48.80*	51.96*	39.65*
	Clinical medicine	50	29.08**	29.23**	32.86**	29.02**
Prepositional phrases other than of	Applied linguistics	30	20.99*	55.26*	43.37*	46.45*
	Clinical medicine	50	37.02**	45.50**	41.76**	36.57**
ed-participle as post-modifiers	Applied linguistics	30	10.38*	8.55*	7.87*	7.75*
	Clinical medicine	50	6.21**	5.12**	4.98**	5.67**
ing-participle as post-modifiers	Applied linguistics	30	4.10*	5.78*	6.23*	7.76*
	Clinical medicine	50	7.23**	4.76**	6.00**	6.11**
Attributive adjectives, nouns as	Applied linguistics	30	62.43*	53.07*	45.41*	46.87*
	Clinical medicine	50	36.51**	31.79**	28.82**	34.30**
Prepositional phrases other than of	Applied linguistics	30	17.94*	31.25*	28.07*	40.94*
	Clinical medicine	50	31.64**	25.73**	27.03**	32.23**
ed-participle as post-modifiers	Applied linguistics	30	62.93*	57.18*	61.27*	56.68*
	Clinical medicine	50	37.36**	34.25**	38.75**	41.48**
ing-participle as post-modifiers	Applied linguistics	30	26.02*	49.98*	42.15*	59.09*
	Clinical medicine	50	45.90**	41.15**	40.59**	46.52**
Attributive adjectives, nouns as	Applied linguistics	30	6.30*	10.80*	11.34*	6.97*
	Clinical medicine	50	5.83**	6.47**	7.17**	5.10**
ing-participle as post-modifiers	Applied linguistics	30	3.28*	8.23*	6.38*	6.94*
	Clinical medicine	50	5.78**	6.78**	6.14**	5.46**
Attributive adjectives, nouns as	Applied linguistics	30	7.10*	5.68*	5.80*	4.78*
	Clinical medicine	50	4.46**	3.40**	3.67**	3.50**
Attributive adjectives, nouns as	Applied linguistics	30	2.09*	5.43*	3.24*	5.16*
	Clinical medicine	50	3.69**	4.47**	3.12**	4.06**
Attributive adjectives, nouns as	Applied linguistics	30	17.37*	14.69*	13.14*	13.36*
	Clinical medicine	50	10.11**	8.80**	8.31**	9.78**

Language feature	Group	No. of texts	Introduction	Methods	Results	Discussion
pre-modifiers	Clinical medicine	50	10.63*	21.82*	17.10*	21.71*
			18.76**	17.97**	16.47**	17.09**
Preposition+nonfinite complement clause	Applied linguistics	30	2.38*	4.11*	3.37*	5.87*
			4.20**	3.83**	3.25**	4.62**
	Clinical medicine	50	2.38*	4.11*	3.37*	5.87*
			4.20**	3.83**	3.25**	4.62**
Complement clauses controlled by nouns	Applied linguistics	30	2.63*	1.49*	1.49*	2.54*
			1.49**	0.89**	0.94**	1.86**
	Clinical medicine	50	0.48*	0.38*	0.19*	1.31*
			0.84**	0.31**	0.18**	1.03**
Appositive phrases	noun	Applied linguistics	16.90*	14.11*	12.86*	12.58*
			10.11**	8.45**	8.13**	9.21**
	Clinical medicine	50	6.30*	11.99*	8.45*	12.86*
			11.12**	9.87**	8.14**	10.12**
Multiple prepositional phrases	Applied linguistics	30	6.87*	6.28*	6.01*	6.23*
			4.11**	3.76	3.80**	4.56**
	Clinical medicine	50	2.70*	4.99*	3.87*	5.37*
			4.76**	3.70**	3.73**	4.23**
Total	Applied linguistics	30	373.300*	330.73*	313.88*	309.01*
			222.00**	197.70**	196.65**	225.17**
	Clinical medicine	50	149.04*	279.06*	231.31*	315.29*
			262.93**	229.37**	223.71**	248.21**

Note. * Raw frequency, ** Normalized frequency per 1000 words.

Appendix B

Pairwise Tukey Comparisons of Noun Phrase Modifiers across Sections in Applied Linguistics RAs

Language features	Sections		Mean Difference	Sig.
	Introduction	Methods		
Attributive adjectives	Introduction	Methods	10.297	0.165
		Results	14.398	0.022
		Discussion	-3.892	0.860
	Methods	Introduction	-10.297	0.165
		Results	4.100	0.841
		Discussion	-14.189	0.025
	Results	Introduction	-14.398	0.022
		Methods	-4.100	0.841
		Discussion	-18.289*	0.002
	Discussion	Introduction	3.892	0.860
		Methods	14.189	0.025
		Results	18.289*	0.002

Language features	Sections	Mean Difference	Sig.		
Relative clauses	Introduction	Methods	0.520	0.787	
		Results	2.155*	0.001	
		Discussion	0.352	0.922	
	Methods	Introduction	-0.520	0.787	
		Results	1.636	0.021	
		Discussion	-0.168	0.990	
	Results	Introduction	-2.155*	0.001	
		Methods	-1.636	0.021	
		Discussion	-1.804	0.009	
		Discussion	Introduction	-0.352	0.922
			Methods	0.168	0.990
			Results	1.804	0.008

*The results are significant at 0.0083

Appendix C

Pairwise Tukey Comparisons of Noun Phrase Modifiers across Sections in Clinical Medicine RAs

Language features	Sections	Mean Difference	Sig.		
Attributive adjectives	Introduction	Methods	25.329*	0.000	
		Results	19.958*	0.000	
		Discussion	11.420*	0.003	
	Methods	Introduction	-25.329*	0.000	
		Results	-5.371	0.343	
		Discussion	-13.908*	0.000	
	Results	Introduction	-19.958*	0.000	
		Methods	5.371	0.343	
		Discussion	-8.53794	0.043	
		Discussion	Introduction	-11.420*	0.003
			Methods	13.908*	0.000
			Results	8.53794	0.043
Noun pre-modifiers	Introduction	Methods	-8.557*	0.001	
		Results	-6.347	0.026	
		Discussion	0.445	0.997	
	Methods	Introduction	8.557*	0.001	
		Results	2.209	0.758	
		Discussion	9.001*	0.000	
	Results	Introduction	6.34792	0.026	
		Methods	-2.209	0.758	
		Discussion	6.79270	0.015	
		Discussion	Introduction	-0.445	0.997
			Methods	-9.001*	0.000
			Results	-6.792	0.015

Language features	Sections		Mean Difference	Sig.
Relative clauses	Introduction	Methods	1.230	0.085
		Results	2.138*	0.000
		Discussion	0.396	0.870
	Methods	Introduction	-1.230	0.085
		Results	0.909	0.298
		Discussion	-0.833	0.375
	Results	Introduction	-2.138*	0.000
		Methods	-0.909	0.298
		Discussion	-1.741	0.008
Discussion	Introduction	-0.396	0.870	
	Methods	0.833	0.375	
	Results	1.741	0.008	
Prepositional phrases	Introduction	Methods	10.461*	0.001
		Results	11.319*	0.000
		Discussion	-1.209	0.967
	Methods	Introduction	-10.461*	0.001
		Results	0.858	0.988
		Discussion	-11.670*	0.000
	Results	Introduction	-11.319*	0.000
		Methods	-0.858	0.988
		Discussion	-12.528*	0.000
Discussion	Introduction	1.209	0.967	
	Methods	11.670*	0.000	
	Results	12.528*	0.000	

About the Authors

Rajab Esfandiari is an associate professor of English Language Teaching at Imam Khomeini International University in Qazvin, Iran. His areas of interest, and specialty, include teaching and researching L2 writing, the construction of rating scales, and EAP teaching and testing. He is widely published in international journals, including *Assessing Writing*, *Journal of English for Academic Purposes*, *JALT Journal*, the *JALT CALL Journal*, *Journal of Teacher Education for Sustainability*, *Language Testing in Asia*, the *Journal of Asia TEFL*, and *TESL-EJ*. His most recent paper was published in *Southern African Linguistics and Applied Language Learning*.

Mohammad Ahmadi is a visiting professor in the Department of English Language, Faculty of Humanities at Imam Khomeini International University, in Qazvin, Iran. His research interest areas include corpus linguistics, second language writing, computational analysis of natural language, and formulaic language.