An Analysis of Structural Bundles Clustered within Rhetorical Moves of Qualitative and Quantitative Discussion Section of Research Articles

Mohammed S. Mohammed
Department of Foreign Languages, Faculty of Modern Languages and Communication, University Putra Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan, Malaysia. Corresponding Author Email: mohammedsadaa6@gmail.com

Mohd A. Bin Abdul Jabar
Department of Foreign Languages, Faculty of Modern Languages and Communication, University Putra Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan, Malaysia. Email: azid@upm.edu.my

Hazlina A. Halim
Department of Foreign Languages, Faculty of Modern Languages and Communication (FBMK), Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia Email: hazlina_ah@upm.edu.my.

Zalina M. Kasim
Department of English, Faculty of Modern Languages and Communication, University Putra Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan, Malaysia. Email: zalina_mk@upm.edu.my

Abstract
One of the most significant activities that writers can engage in to share their knowledge and establish a voice within the disciplinary community is writing academic research articles. To have a greater understanding of this commonly practiced genre, exploring the linguistic realisations which are used to classify the rhetorical structures of academic research articles is beneficial. This study's main objective was to analyse structural bundles clustered in rhetorical moves in the Discussion parts of the specialised corpora of qualitative and quantitative datasets of Applied Linguistics Research articles (RAs), wherein each corpus contains Discussion sections of 50 RAs, as well as structural elements of those lexical bundles. For rhetorical move identification, the data were analysed based on Ruiying and Allison's (2003) model. For lexical bundle analysis, the corpus-driven approach was used to extract

four-word lexical bundles. Hence, the data were analysed using Biber et al.'s (1999) structural taxonomy. The analysis of Discussion sections identified some discrepancies in terms of lexical bundle frequency, which has also shown that some moves incorporated more bundles as compared to others. In terms of similarity of usage, the two groups of RA writers significantly used phrasal more than clausal. RA writers used more NP-based bundles at the expense of other phrasal sub-categories. The results of the current study provides insightful knowledge of genre convention as well as cluster of lexical bundles in Qualitative and Quantitative research articles. This knowledge enables authors, especially novice ones, to dive into their academic writing approach.

Keywords: Genre-Based Analysis, Corpus Driven Approach, Rhetorical Move, Structural Bundles, Discussion Section

Introduction

Genre analysis has sparked a growing interest in the rhetorical moves, to spread academic information within a particular discourse community (Tessuto, 2015; Lu et al., 2021), as well as use of formulaic sequences in academic discourse (Kashiha, 2014; Nam, 2017). In the space of academic communities, the fellows use research articles as a channel to express their ideas and feelings to others (Henceforth, RAs) and it is arguably the vital genre in academic discourse (Mohammad & Sadat, 2014; Li, 2020). Noticeably, RA is a preferred academic genre meant to exchange and disseminate knowledge amongst members of the academic community. It is also considered as an indicator of academic achievement; and for researchers/scholars to participate in the academic world, they have to have their studies published therein (Dobakhti, 2011).

However, each academic discipline has its own peculiar clusters which characterise it, which are valued in genres of those disciplines in particular (Vaseghi, 2016). Academic writers need to be familiar or acquire those clusters and other specific norms and conventions in the targeted genre which differs from the regular academic classroom writing to achieve the studies' significance and the worthiness of their studies' attention (Dobakhti, 2011). Whilst describing the features of successful academic writing, Hyland (2005b) has this to say, the ability of writers to offer a credible presentation of themselves and their work, by claiming solidarity with readers, evaluating their material and acknowledging alternative views. It is also pertinent to note that the authors' claims may be questioned or rejected at any stage by the members of discourse community if they do not reach the expectation manners of the discourse community level (Mohammad & Sadat, 2014) because considering writing as a socially situated activity which is used for a purpose and is conducted mainly for academicians of a specific discourse community. This means that, authors required to be awareness of both rhetorical shape and linguistic realizations associated with the research article in their fields.

As an influential knowledge of genres in academic discourse, RAs have been receiving notable interest towards genre analysis studies from researchers of the second language writing field. Previous studies have revealed that a handful of approaches have tried to describe the communicative units in specific genres and suggest a framework of studying rhetorical organization in diverse types of text. Amongst these approaches was the rhetorical analysis 'move-steps' proposed by Swales (1981), which recommends a hierarchical classification comprising moves and steps. Essentially, a 'move' can further be divided into steps and sub-steps to provide an account of more details of a rhetorical organization.

Multiple models and frameworks, for example, Swales (1990); Bhatia (1993); Hyland (2000); Nwogu (1997); Ruiying & Allison (2003) have been proposed by many scholars, researchers

and experts alike to analyse and describe RAs schematic structure such as (abstracts, introductions, literature review, methods, results, discussion, and conclusion). Within genre analyses (ESP), 'earlier studies may be summarised into two groups following their concentration of target objective'. The first group focuses on the structure of dissertations, research papers or other specialized writings, for instance, Hong (2019); Tessuto (2015); Amnuai (2019); Lu et al (2021) and the second group is on specific features of RAs, such as the use of 'hedging, modality, reporting verbs and function-form connections'. For example, Hyland (2000); Omidian et al (2018); Casal (2020) mentioned but a few (Ruiying & Allison, 2004).

In addition to conducting a genre analysis, this study is aimed at identifying the recurrent word sequences that are utilized to present the communicative function of "move boundaries". It is imperative to note that move boundaries could be identified based on the function of move performance, as well as that of the linguistic clues (Alamri, 2017). There is an increasing evidence that natural language, body of corpus research in particular, is occupied by formulaic sequences as Kashiha (2014) argued that a great proportion of discourse are established through the utilisation of these expressions. Although recent proliferation of genre analysis studies exists owing to the high demand of research articles by the researchers for one reason or the other, a move variation in the text's structure in addition to accompanying lexical bundles suitable to each move and steps, are necessary knowledge for researchers to publish in the international community successfully.

Studies consistently shown that natural language is invaded with formulaic sequences (FSs). They are vital to the development of academic discourse. Lexical bundles, which are components of FSs, are collections of word forms which occur frequently within a natural discourse and are quantified statistically by their frequency of recurrence. These bundles are commonly found in rhetorical moves of different sections of academic discourse. They do contribute in identifying members of a particular discourses' community. This means that language users could associate with a particular group by using formulaic sequences, such as a disciplinary community (Wray, 2002). Wray (2006) as cited in Ädel & Erman (2012) said "when we speak, we select particular turns of phrase that we perceive to be associated with certain values, styles and groups." Lexical bundles are statistically defined as collections of word forms that frequently co-occur in normal language; they greatly improve fluency in speech and writing. (Shin, 2019). Cortes, (2006) believes that repeated use of these bundles in a specific register or discipline is an indication of language use proficiency. On the contrary, that such clusters are absent indicates that newcomers to that academic community lack fluency (Hyland, 2012).

Publication of research articles in reputable journals offers numerous benefits for researchers such as knowledge sharing, career advancement, prestige, securing of research grant, as well as its capacity to transfer a great deal of information in the academic world. Recent developments have shown that publication is part of a requirement for postgraduate students' graduation of universities (Dobakhti, 2011). The awareness of the qualitative and quantitative designs is a matter of paramount importance especially to the novice writers. Creswell (2003) stresses that qualitative and quantitative designs are 'different in the knowledge claim that they make, the main purpose that they follow, the research questions that they impose, the data that they collect, and the methods that they use to analyze data'. Dobakhti (2011) stresses that they also differ in the use of rhetorical moves and other lexical patterning. In their attempt to accomplish their aim of journal publications, novice and apprentice writers find it difficult to produce a research paper which will be publishable in a

scholarly journal for their lack of knowledge of discourse conventions peculiar to a particular research design because the qualitative and quantitative research articles differ in some essential aspects. To date, few discussions or no published research have been conducted on the effect of research design on the rhetorical moves and lexical bundles of qualitative and quantitative RAs.

Admittedly, numerous studies have examined the rhetorical structures of RAs in different disciplines. For example, Nwogu (1997) Medicine; Yang (2001) Applied linguistics; Posteguillo (1999) Computer science; Kanoksilapatham (2005) Biochemistry; Paydari (2017) Political science; Musa et al (2015) Engineering; Peacock (2002) Seven disciplines, which comprise Language and Linguistics, Physics, Environmental Science, Biology, Public and Social Administration, Business and Law. These studies tend to focus almost exclusively on the organisational patterns of RA sections. RA has also been the subject of numerous studies, either in its entirety or in a particular section, within the framework of IMRDC (Introduction, Method, Result, Discussion and Conclusion). The most prominent amongst them have concentrated on the RAs' 'Introduction section' (e.g., Swales, 1981, 1990; Ozturk, 2007), 'Results section' (eg. Brett, 1994; Williams, 1999), 'Method section' (eg. Lim, 2006; Musa et al., 2015), 'Discussion section' (e.g., Holmes, 2000; Dobakhti, 2011). Others who were interested in studying all sections of the RAs include (Nwogu, 1997; Pho, 2008; Shi, 2014).

In another line of inquiry, the literature has revealed that previous studies have mostly compared lexical bundles used by native and non-native English speakers in academic writings (few amongst them are Nam, 2017; Shin, 2019; Chen & Baker, 2010; Ädel & Erman, 2012). Some studies focused on comparing lexical bundles across disciplines (e.g., Jalali, 2014; Güngör & Uysal, 2016; Tovar-viera, 2018; Lu et al., 2021), whilst others have particular interest in comparing novice and experts' use of lexical bundles in academic discourse (see, Jalali, 2014b; Wang, 2018). There are only a handful of studies which compare qualitative and quantitative RAs in a single discipline, especially applied linguistics as argued by (Ruiying, 2003) who described as under researched discipline. Even the ones available do not specifically study rhetorical moves and lexical bundles in the qualitative and quantitative RAs. For example, Dobakhti (2011) focused exclusively on generic structure and stance features. Numerous approaches to the study of lexical bundles have been used by scholars who opt for a corpus-based approach (see, Biber, 2009; Esfandiari & Barbary, 2017b) other studies favoured corpus-based approaches (see, Conrad Susan M., 1996) whilst others used frequency-based approaches (see, Chen & Baker, 2010; Leelasetakul, 2019). Studies that correlate genre-based and corpus-driven approaches are scarce in the existing literature.

Turning into the number of words as lexical bundles, a couple of studies which analysed the four-word lexical bundle, being the most frequently studied bundle (Chen & Baker, 2010; Kashiha, 2019) in academic discourse (see, Biber et al., 1999; Hyland 2008; Vo, 2019) have been presenting their objective arguments based on their choices. The current study also prioritises four-word bundles for obvious reasons, such as the fact that it subsumed three-word bundles in their structures, and it occurs more commonly and offers a clear range of structures and functions when compared with three- or five-word bundles (Hyland, 2008; Kashiha, 2019). Similarly, four-word clusters are significantly more common than five-word bundles in discourse (Hyland, 2008a; Cortes, 2004), making it easier to classify them and determine their context (Chen & Baker, 2010; Biber et al., 1999; Cortes, 2013; Tetyana & Lee (2017).

It is pertinent to note that although there are number of studies that investigated the rhetorical moves with lexical bundles and/or linguistic realizations in sections of published

RAs, those studies mainly focused on disciplinary variation (across various academic disciplines). To date, no studies were conducted to compare the move structure and lexical bundles of qualitative and quantitative Ras. Hence, this study hopes to resolve this neglect by investigating RM and LB in Applied Linguistics, which is argued to be an under-researched discipline (Ruiying, 2003).

Given the gap from the previous studies, this study aims to apply corpus-driven approach to investigate lexical bundles in rhetorical moves used in the Discussion sections of qualitative and quantitative RAs of applied linguistics through the structural taxonomy. The essence of investigating rhetorical move and lexical bundles in Discussion sections is that these sections pose further difficulties to the novice research writers as they commonly share some similar communicative functions in move/step structures (as Ruiying & Allison's (2003) model indicated), amidst their significance and obligatory status in most academic discourses (Alamri, 2017; Abdollahpour & Gholami, 2019; Hong, 2019). This study aimed at comparing the rhetorical moves found in the Discussion sections of qualitative and quantitative RAs, as well as comparing the lexical bundles clustered in the rhetorical moves of the Discussion sections of the two groups of RAs structurally.

Methodology
A. Study Design
The present study uses Ruiying and Allison's (2003) model for rhetorical move identification and categorisation, and corpus-driven approach to investigate lexical bundles in the two corpora. This corpus-driven study adopts mixed method research design to help achieve the research objective. This means that qualitative and quantitative methods were used for data collection and analysis in the study. This study comprises two major parts; each dealing with various research design. The first part deals with the analysis of rhetorical move of 50 RAs' Discussion sections. These RAs were analysed qualitatively. The second part, however, concerns with the lexical bundles clustered in the rhetorical moves identified in the 50 RAs' Discussion sections which were analysed quantitatively. The analysis in this part was conducted in two parts. The lexical bundles were analysed in the first part based on type and frequency of use. These lexical bundles were examined in the second part using Biber et al (1999) structural taxonomy. The focus of this part was to determine the lexical bundles' functions according to the rhetorical moves they were used in.

B. Corpus of the Study
A corpus refers to a collection of written and spoken texts that represent a particular areas of language use (Paltridge, 2006). In corpus linguistic studies, the researcher often analyses the texts collected to identify occurrences of specific linguistics features by means of a particular computer program. To Paltridge (2006), a corpus is of two different types: a general corpus and specialised corpus. Whilst the language use is represented at large so that the results could be generalizable in the former, the latter refers to a set of texts which represent a specific genre; for example, qualitative and quantitative RAs in a specific discipline or genre. Hence, this study relies on specialised corpora, which were systematically compiled for the purpose of extracting the linguistic features under investigation. Following the study aims to investigate the textual organization of the qualitative and quantitative RAs whilst exploring the lexical bundles associated with each identified rhetorical moves of Discussion sections, specialised corpora are needed to be compiled systematically to achieve the study's objectives. It is also noticeable that the present study is comparative in nature because it

focuses on examining the possible variations which exist between the two different methods of inquiry.

However, the corpora were chosen from five prestigious journals in the field of Applied Linguistics which were published between 2017 and 2022. The current study prefers journals that publish both qualitative and quantitative RAs, whilst rejecting one-sided journals. The study selected only 5 journals which include: Journal of English for Specific Purposes (ESP), Language Teaching Research (LTR), Applied Linguistics (APL), Journal of Pragmatics (JPR), and Journal of Sociolinguistics (JSL). Hence, the selection was based on research design it used, i.e. qualitative or quantitative. Thus, the utilisation of these corpora across the two different RAs provides valuable insights into the nature of qualitative and quantitative inquiry in Applied Linguistic discipline.

Following classification of all RAs in those five journals as either qualitative or quantitative (50 RAs for each), samples of 10 RAs from each qualitative journal and 10 RAs from each quantitative journal were chosen at random for the rhetorical move analysis, as shown in the following table:

Table 1
Summary of the corpus used in analysing Lexical Bundles

Journals	Qualitative		Quantitative	
	# Of RAs	Word count	# Of RAs	Word count
English for specific purpose	10	16737	10	18359
Applied linguistics	10	16337	10	18637
Language teaching and research	10	15595	10	19217
Journal of pragmatics	10	15852	10	21362
Journal of sociolinguistics	10	12947	10	21356
Total	50	77468	50	98931

C. Data Analysis

The RAs were extracted in PDF and were copied and pasted into the plain text files. In an attempt to clean up those document texts, any extra information in the RAs, which were irrelevant, were deleted. However, it should be noted that identifying the rhetorical move structure of the RAs discussion sections is a method of approaching genre. A genre is organized according to a set of communicative goals that are attained through Move.

D. Procedure for Rhetorical Move Identification

Following the fact that the notion of communicative purposes is central to moves identification in this study, it is imperative to note that lengths of move vary depending on the content the author wishes to express in the discourse using varieties of sentences, phrases or words. Since move lengths vary and the fact that the author uses several sentences (in a move) to express his intents, it is possible for a single move to perform multiple functions. When this situation occurred, Holmes (1997) suggested that the analysis should focus on the most salient function. Although the study was not focused on the RA sections, each RA section was read completely to uncover the general idea. To illustrate this point, many RAs in Qualitative corpora were read repeatedly and examined thoroughly to comprehend the Discussion. Then, the Qualitative (Quali) and Quantitative (Quanti) corpora were analysed separately using Ruiying and Allison's (2003) seven-move model/analytical framework of

move analysis. Given that the unit of moves and steps analyses is clause, according to Dobakhti (2011), it implies that the structure of moves and steps are realised by clause. With this in mind, the analysis of moves and steps in this study was limited to clause. Therefore, any grammatical unit lower than clause, for example, phrases and words were not deemed fit for the analysis. During the analysis, any sentence which realises more than one move or one step, the most dominant one was taken into account. Once again, the analysis in this study did not take the embedded moves or steps into account to avoid unwanted complication.

According to Alamri (2017); Ruiying and Allison's (2003) models are considered as comprehensive in terms of move analysis in the applied linguistics field. Although Pho's (2008a) model is the recent one compared with Ruiying and Allison's (2003) and regarded as all-inclusive by some researchers, it was not adopted simply because it provided options for the analysis of combined sections of Ras; for example, Result-Discussion and Discussion-Conclusion, which was not the aim of this study

Table 2
Ruiying and Allison's (2003) model for the move analysis of RA discussion section

Moves	Steps
Move 1 background information	
Move 2 reporting results	
Move 3 summarizing results	
Move 4 commenting on results	Step 1 interpreting results
	Step 2 comparing results with literature
	Step 3 accounting for results
	Step 4 evaluating results
Move 5 summarizing the study	
Move 6 evaluating the study	Step 1 indicating limitations
	Step 2 indicating significance/advantage
	Step 3 evaluating methodology
Move 7 deductions from the research	Step 1 making suggestions
	Step 2 recommending further research
	Step 3 drawing pedagogic implication

Result
A. Qualitative Analysis

To achieve and further clarify discourse functions which distinguish between qualitative and quantitative RAs, a qualitative analysis of the rhetorical moves in the context in which they were used was deemed necessary. We focused on how the bundles which were identified as being indicative of either qualitative or quantitative RAs were used and how they were organized in each rhetorical move in the Discussion sections at this point in the analysis.

Background information (Move 1).

Notable disparities between Discussions in qualitative and quantitative RAs were found after an analysis of the various bundles in this move. For instance, the majority of the research-oriented bundles in this move were used by authors in qualitative RAs to define

various aspects of the following noun's quality, form, and size and to explain their physical features. This is demonstrated in the following example:

1.   "In this study, the values manifested by the students during the design of the leaflet were expressed in the methods in which they applied, understood, explained and analysed knowledge (QL8)."

Reporting results (Move 2)
We discovered distinct bundles in the qualitative and quantitative RAs of this move, with qualitative RAs having a few more different move types and tokens compared with quantitative RAs. Whilst qualitative RA writers had the preference of using description sub-categories of research-oriented bundles to show qualities, extent and existence of the data, quantitative RA writers tended to use quantification sub-categories to indicate quantities, measures or proportions of the data. The following are the examples extracted from the corpora of qualitative and quantitative RAs:

2. "The purpose of this experiment was to further identify the nature of the meaning expressed by the voseo negative command in AS (QL28)."
3. "Although the size of the corpus and frequencies of Engagement resources were different in the two studies, the results were almost the same (QN2)."
Summarising results (Move 3)
A comparative analysis of the various bundles indicated that bundles specific to this move were only found in qualitative RAs. These bundles were found to be inferential sub-category of research-oriented bundles and were only three, which include it can be inferred, can be argued that it is unlikely, as seen in (4), (5) and (6). This indicates that qualitative RA writers used these bundles to draw inferences on their results, as following illustrative example shows:

4. "In sum, it can be argued that the DMAs' seemingly contradictory developments have been triggered by pragmatic inferences from the speech situation (QL25)."
5. "It is unlikely that the teacher education they had received would have led to an understanding of the various components of language competence (QL33)."

Commenting on results (Move 4)
Writers of RAs in these two research designs in this move were observed to make significant use of text-oriented bundles (e.g., on the other hands, the fact that the, in terms of the, findings of this study) whilst interpreting the results. However, quantitative RA writers were found to use more distinctive bundle tokens than qualitative RA writers in this move.
Although the two groups of writers preferred to use text-oriented type of bundles in this move, qualitative RAs writers had a preference for using the inferential subcategory, such as are more likely to, it appears that the, it may be that, over quantitative RAs writers to signal inferences and draw conclusion from the data. It indicates that authors of qualitative RAs temper their opinions and refrain from overgeneralizing their findings when interpreting their research findings, as illustrated in the following citations.

6. "demonstrate that academic writers are more likely to provide the precondition for the feasibility of a process or the validity of the proposition (QL46)."

7. "It appears that the teacher in Class 2 adopted a more inductive approach, whereas the teacher in Class 3 provided more elaborate contextualization of the target words (QL9)."

Summarizing the study (Move 5)

Examination of the bundles identified in Move 5 revealed that the number of bundles were concentrated more in qualitative than in quantitative RAs. It was discovered that writers of the qualitative RAs mainly apply specific bundles, which are mainly the comparative sub-category of text-oriented bundles (e.g., results are consistent with, are in line with), in which the qualitative RAs writers made use of to show the comparison of their results with the earlier ones. This result is consistent with Omidian, Shahriari & Siyanova-Chanturia's, (2018) findings.

8. "These results are in line with recent similar research on the learning of L2 (Serrano & Huang, 2018) and L1 vocabulary (Goossens et al., 2016) in classroom settings (QL27)."

Evaluating the study (Move 6)

The analysis of bundles exclusive to Move 6 revealed that certain identified bundles were exclusively found in quantitative RAs. We discovered that qualitative RAs writers in this move took a clear-cut stance when evaluating their results through the employment of description sub-category of research-oriented bundles such as findings of this study, the effectiveness of the, the extent to which, as illustrated in (9).

9. "It would be helpful to ascertain the extent to which students are held accountable for what they have read in English and how that accountability is measured (QN32)."

Deductions from the research (Move 7)

The analysis shows that the two groups of RA writers used more participant-oriented bundles, especially stance sub-category such as It seems likely that, has been suggested that, are more likely to, it can be concluded that to indicate the attitudes and evaluations of the writer. However, quantitative RA writers, unlike their quantitative RA counterparts, had a tendency to use more tokens but few types in this move. The following are some examples extracted from the two datasets that illustrate the use these bundles.

10. "It seems likely that the adult 'decoys' of the Perverted Justice data create different interactional patterns with these online offenders (QL17)."
11. "The higher proportion of Deontic and lower proportion of Dynamic matrix predicates demonstrate that academic writers are more likely to provide the precondition for the feasibility of a process (QN20)."

B. Quantitative Analysis

The identification procedures yielded 2174 bundle tokens and 532 bundle types, as shown in Table 3, which presents the information with regard to the types of bundles and the number of bundle tokens obtained in each set of RAs. However, having identified the lexical bundles, they were all assembled based on the discourse functions of rhetorical moves from which they were found.

Table 3
Structural features of Lexical Bundles in the discussion sections in both corpora

| Categories	Sub-Categories	Quali Corpus		Quanti Corpus	
	NP-based	Raw F	Rel. F	Raw F	Rel. F
	NP + of-phrase	152	19.29	205	15.60
	NP + other post-modifier	43	5.46	90	6.85
	Other NP PP-based	22	2.79	36	2.74
	PP + embedded of-phrase	114	14.47	113	8.60
	Other PP VP-based	135	17.13	387	29.45
	Copular be + N/Adj phrase	50	6.35	65	4.95
	1st Person pron + VP (+ Comp.)	18	2.28	4	0.31
	Passive verb + PP	12	1.51	30	2.28
	Anticipatory it + V/Adj phrase	49	6.22	105	7.99
Phrasal	Pron/NP phrase + be	20	2.54	27	2.05
	Sum	615	78.04	1062	80.82
	V/Adj + to-clause	24	3.05	60	4.57
	VP + that-clause	30	3.81	47	3.58
Clausal	Adverbial clause	61	7.74	54	4.10
	Sum	115	14.60	161	12.25
	Other expressions	58	7.36	91	6.93
Total		788	100	1314	100
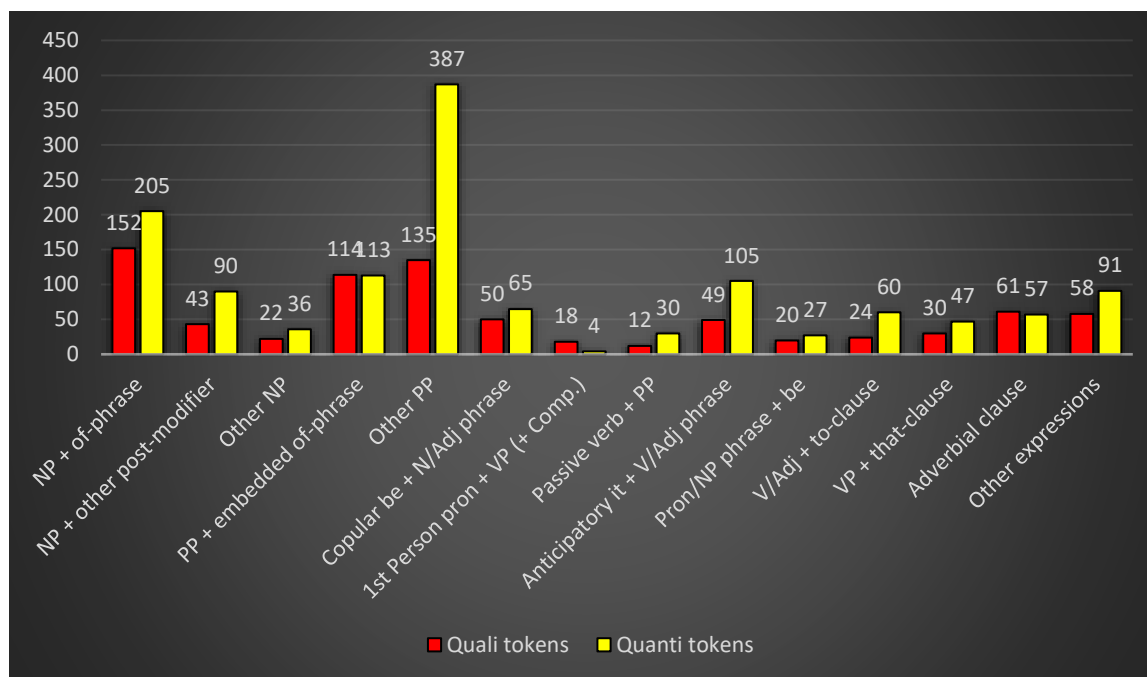
Figure 1. Distribution of structural tokens of target bundles of the discussion section in the two corpora

Table 3 and Figure 1 show that some similarities and differences manifest in the usage of structural patterns among the two datasets. In terms of similarity of usage, qualitative and quantitative article writers tend to use more phrasal than clausal bundles in the Discussion section. Although there are some differences in phrasal target bundles, NP-based with (27.3%) and (25.1%), PP-based (17%), whilst the latter had 387 (29.5%), and VP-based (17.6% and 6.3%), respectively. The clausal target bundles show virtually similar distribution in the two corpora, as proven in the table and figure above with 15.2% of qualitative, and 12.3% counterpart. The distribution shows that NP-based and PP-based were the largest groups in the target bundles. The majority of lexical bundles in the Discussion section were phrasal than clausal. Finally, the category 'other expressions' was the last structural category found in the two datasets. In terms of overuse of some miscellaneous bundles in this category, the results indicate the two datasets differ in some respects as each favours certain bundles from one another. For example, the bundles such as 'further research is needed', 'the situation was so', were frequently used in Quali corpora, whilst the bundles such as made errors with the, may be the case, the results suggest that were mostly used in Quanti corpora.

Discussion

The current study investigated the research design variation in the RA discussion section as exhibited in the use of lexical bundles in distinct rhetorical moves of this RAs section. The main aim of this study has been to investigate the extent to which authors of qualitative and quantitative RA discussion sections differ from one another in terms of the lexical bundles they typically use to achieve the communicative functions of the discussion sections. The study used a mixed-methods approach to analyse the patterns of variation found to explore these differences on a qualitative and quantitative level. The approaches adopted allowed us to identify number of differences between qualitative and quantitative RAs in terms of focus and practice. The patterns of variation were first discussed in terms of qualitative enquiry and

followed by a description of some of the key differences discovered in quantitative analyses of the study.

The discussion sections of these RAs were first compared based on the typical structures they employ in various discussion sections of RAs before considering possible distinctions between the qualitative and quantitative RAs. The analyses revealed that discussion section writers of quantitative RAs used considerably more bundle tokens in Move 4 (Commenting on results) than their qualitative counterparts. This result appears to demonstrate the formulaicity and conventionality of this move in the discussion sections of quantitative Ras. According to Hyland (2016), the norms and the community's expectations may heavily influence genre conventionality in which they are used. It might be claimed that the research design demands of quantitative RAs make writers of this nature to expend more rhetorical effort in commenting on their study findings, which caused the conventionalization of specific sequences in this move. This result is consistent with Dobakhti (2011), where she found that quantitative RA writers used stance features in commenting on results move.

In addition, it was also discovered that although the two groups of RA writers significantly used phrasal subcategories more than clausal subcategories, RA writers used more NP-based bundles at the expense of other phrasal subcategories, as well as their quantitative counterparts. This result was not taken by surprise, as it proves the assertion that English academic writings are structurally more phrasal-based than clausal-based (Vaseghi, 2016). In the same manner, the two corpora share some similarities in terms of frequency of use of NP-based bundles, as NP + of-phrase fragment was the most frequent, followed by NP + other post-modifier, then other NP/Pron fragment. PP-based is another phrasal group where the two datasets share some commonalities and differences. This group was found to be the most frequently used bundles, both in the two corpora. Thus, a statistically significant overuse of some bundles in Quanti corpora may be attributable to the fact that quantitative article writers are more familiar with specific bundles in expressing their intents. The last phrasal group of target bundles comprised the VP-based structure, a closer examination of the corpora indicated that the VP-based structures in the two datasets were commonly used to show logical relations between structural elements, writer's comparison and evaluation.

Clausal bundles were unlike Phrasal bundles which manifested more frequently. This type of bundle was found to be infrequent in both qualitative and quantitative corpora. They only made up about 15% and 12% of the whole bundles in the qualitative and qualitative corpora respectively. However, results show the differences between the two datasets, as the clausal bundles of the Quali dataset account for 15.2% of the total number of bundles in the Quali corpora, whereas clausal bundles of the Quanti dataset constitute 12.3% of the total number of bundles in the Quanti corpora.

The findings also revealed 'other expressions' which was the last structural category found in the two datasets. In terms of overuse of some miscellaneous bundles in this category, the results indicate the two datasets differ in some respects as each favours certain bundles from one another. For example, the bundles such as 'further research is needed' and 'the situation was so', were frequently used in Quali corpora, whilst the bundles such as 'made errors with the', 'may be the case' and 'the results suggest that' were mostly used in Quanti corpora.

It becomes clear that though they share some commonalities, these two groups' RAs differ significantly in some respects while express their communicative structures in different rhetorical moves. Admittedly, the results are based on small corpora which are rather speculative. Although the current study has raised several interesting distinctions, but larger

corpora are needed to establish extent in which the result can be generalised. Therefore, it can be argued this study can greatly help graduate students and researchers alike who wish to participate in RA publication but are yet to acquire the necessary knowledge and conventions of various research designs.

Conclusion

The present study demonstrates that participants from various academic fields place different priorities on how their research is represented in RA discussion sections. It was discovered through the synergy of genre analysis and corpus linguistics as methodologies that qualitative RA writers often ensure that the summary of their research is expressed clearly. However, quantitative RA writers consider it to be essential to advance their research through evaluation and drawing inferences from their findings. that although the two groups of RA writers significantly used phrasal more than clausal, RA writers used more NP-based bundles at the expense of other phrasal subcategories, as well as their quantitative counterparts. Clausal bundles were unlike phrasal bundles which manifested more frequently, this type of bundle was found to be infrequent in both qualitative and quantitative corpora, whereas the results indicate the two datasets differ in some respects as each favours certain bundles from one another. To develop our understanding of research design writing further, it is thought that highlighting the differences between the knowledge construction and reader communication styles used by writers in various academic fields has a significant potential to improve our comprehension of research design writing. It could be argued that studies of this nature can be extremely beneficial, particularly for graduate students and early-career researchers who want to take part in scientific activities in their area, but do not yet possess the necessary knowledge of such discipline. It is thought that investigating research design conventions through the prism of their linguistic realizations in discourse can aid in the creation of a thorough understanding of research design writing for language users in this particular group**.**

The Contribution of the Study

The findings from this study make noteworthy contributions to the extant literature in several ways. These contributions are seen from two broad perspectives: theoretical and practical contributions. The theoretical contributions consist of methodology building specialised corpora of English for academic purposes (EAP) and data refinement criteria.

The systematic procedures used in building the specialised corpora can easily be followed by EAP researchers to undertake empirical studies on several pedagogic aspects. This implies that useful sources of information will be available at a doorstep for the novice writers to write appropriately, which will be acceptable to the members of scientific community. In the same way, the specialised corpora developed in this study may be used by EAP researchers in carrying out empirical pedagogic-related research, namely, academic vocabularies, academic collocations, multi-word sequences et cetera, in the field of applied linguistics.

As far as practical perspective is concerned, the results of the present study offered valuable contributions to pedagogic practice. The knowledge of moves and steps enables students to use the conventional moves and steps thereby paying less effort in writing optional moves and steps in Discussion section of the research articles. Likewise, the knowledge of lexical bundles provides students with information pertaining to how lexical bundles are utilised in applied linguistic texts in respect to their structural features.

References

Alamri, B. M. (2017). *Connecting genre-based and corpus-driven approaches in research articles: A comparative study of moves and lexical bundles in Saudi and international journals* (Doctoral dissertation, The University of New Mexico).

Amnuai, W. (2019). Analyses of Rhetorical Moves and Linguistic Realizations in Accounting Research Article Abstracts Published in International and Thai-Based Journals. *SAGE Open*, 1–9.

Biber, D. (2009). A corpus-driven approach to formulaic language in English Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, 14(3), 275–311.

Biber, D., Johansson, S., Leech, G., Conrad, S. &, & Finegan, E. (1999). *Longman Grammar of spoken and written English*.

Chen, Y., & Baker, P. (2010). Lexical Bundles in L1 and L2 Academic Writing. *Language Learning & Technology*, 14(2), 30–49.

Susan, C. M. (1996). Investigating Academic Texts with Corpus-Based Techniques: An example from Biology. *Linguistics and Education*, 326, 299–326.

Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purposes*, 29, 21–42.

Gungor, F., & Uysal, H. H. (2016). A Comparative Analysis of Lexical Bundles Used by Native and Non-native Scholars. *English Language Teaching*, 9(6), 176-188.

Holmes, R. (1997). Genre analysis and the social sciences: An investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes*, 16(4), 321–337.

Hong, J. (2019a). Structural Use of Lexical Bundles in the Rhetorical Moves of L1 and L2 Academic Writing. *English Teaching*, 74(3), 29–54.

Hyland, K. (2005). Patterns of engagement: Dialogic features and L2 undergraduate writing. In R. A. E. Louise J. Ravell (Ed.), *Analysing Academic Writing* (pp. 5–23). London: Continuum.

Hyland, K. (2012). Bundles in Academic Discourse. *Annual Review of Applied Linguistics*, 32, 150–169.

Hyland, K. (2016). Journal of Second Language Writing Academic publishing and the myth of linguistic injustice. *Journal of Second Language Writing*, 31, 58–69.

Jalali, H. (2014a). A comparative study of anticipatory it lexical bundles in applied linguistics and analytical chemistry research articles. *Journal of Foreign Language Teaching and Translation Studies*, 3(1), 24-42.

Jalali, H. (2014b). Examining novices' selection of lexical bundles: The case of EFL postgraduate students in applied linguistics. *Journal of Applied Linguistics and Language Research*, 1(2), 1-11.

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, 24, 269–292.

Kashiha, H. (2014). Discourse Functions of Formulaic Sequences in Academic Speech across Two Disciplines. *GEMA Online Journal of Language Studies*, 14(6) 15–27.

Kashiha, H. (2019). An awareness of formulaic clusters in conclusion moves of Applied Linguistics research articles. *Journal of New Advances in English Language Teaching and Applied Linguistics*, 1(1), 1–18.

Koteyko, N. (2014). *Language and politics in post-Soviet Russia: A corpus assisted approach*. Springer.

Leelasetakul, M. (2019). Lexical Bundles used by Thai undergraduate learners of English.

Leila, D. (2011). *The discussion section of research articles in applied linguistics: Generic structure and stance features* (Doctoral dissertation, University of Malaya).

Li, X. (2020). Mediating cross-cultural differences in research article rhetorical moves in academic translation: A pilot corpus-based study of abstracts. *Lingua*, *238*, 102795.

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,* 52, 101006.

Mohammad, A., & Sadat, Z. (2014). Rhetorical Moves in Applied Linguistics Articles and their Corresponding Iranian Writer Identity. *Procedia - Social and Behavioral Sciences*, 98, 489–498.

Nam, D. (2017). Functional Distribution of Lexical Bundle in Native and Non-Native. *The Journal of Asia TEFL*, 14(4), 703–716.

Nwogu, K. N. (1997). The Medical Research Paper: Structure and Functions. *English for Specific Purposes*, 16(2), 119–138.

Omidian, T., Shahriari, H., & Siyanova-Chanturia, A. (2018). A cross-disciplinary investigation of multi-word expressions in the moves of research article abstracts. *Journal of English for Academic Purposes*, 36, 1–14.

Ozturk, I. (2007). The textual organisation of research article introductions in applied linguistics: Variability within a single discipline. *English for Specific Purposes*, 26, 25–38.

Paydari, S. S. (2017). *Rhetorical Moves and Discourse Structure in Political Science Research Articles abstracts of Iranian journals*. (Master thesis, Universiti Putra Malaysia).

Peacock, M. (2002). Communicative moves in the discussion section of research articles. *System*, 30, 479–497.

Pho, P. D. (2008). How can learning about the structure of research articles help international students. In *Conference proceedings of the 19th ISANA international education conference* (Vol. 2, p. 5).

Ruiying, Y., & Allison, D. (2003). Research articles in applied linguistics: moving from results to conclusions. *English for Specific Purposes*, 22, 365–385.

Ruiying, Y., & Allison, D. (2004). Research articles in applied linguistics: Structures from a functional perspective. *English for specific Purposes*, *23*(3), 264-279.

Salazar, D. (2014). *Lexical Bundles and Non-native Scientific Writing: Applying a Corpus-based to Language Teaching*. John Benjamins Publishing Company.

Shi, H. (2014). *Moves and lexical bundles: A contrastive study of English agricultural science research articles between Chinese journals and internationally published journals* (Doctoral dissertation, School of Foreign Languages Institute of Social Technology Suranaree University of Technology).

Shin, Y. K. (2019). Do native writers always have a head start over non-native writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes*, 40, 1–14.

Tessuto, G. (2015). Generic structure and rhetorical moves in English-language empirical law research articles: Sites of interdisciplinary and interdiscursive cross-over. *English for Specific Purposes*, 37, 13–26.

Tetyana, B., & 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.

Tovar-viera, R. (2018). Rhetorical Organization and Linguistic Realizations of Moves. *ICDEL Journal*, 3(1), 1–21.

Vaseghi. (2016). *Features of Lexical Bundles in Social Science Research Articles of native and Non-native experts' writers of English*. (Doctoral Dissertation, Universiti Putra Malaysia).

Vo, S. (2019). Use of lexical features in non-native academic writing. *Journal of Second Language Writing*, 44, 1–12.

Wang, Y. (2018). As Hill seems to suggest: Variability in formulaic sequences with interpersonal functions in L1 novice and expert academic writing. *Journal of English for Academic Purposes*, 33, 12–23.