

Disassembling Lexical Bundles Clustered within Rhetorical Moves in Discussion Sections of Qualitative and Quantitative Articles: A Comparative Analysis

Mohammed S. Mohammed¹, Mohd A. Bin Abdul Jabar²,
Hazlina A. Halim³ and Zalina M. Kasim⁴

1- Department of English Language, Art College,
University of Anbar, 31001, Ramadi, Iraq.
Mohammedsadaa6@gmail.com

2- Department of Foreign Languages, Faculty of
Modern Languages and Communication, University
Putra Malaysia, 43400, UPM Serdang, Selangor Darul
Ehsan, Malaysia. azid@upm.edu.my

3- Department of Foreign Languages, Faculty of
Modern Languages and Communication, University Putra
Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan,
Malaysia. hazlina_ah@upm.edu.my

4- Department of English, Faculty of Modern Languages
and Communication, University Putra Malaysia, 43400,
UPM Serdang, Selangor Darul Ehsan, Malaysia.
zalina_mk@upm.edu.my

Corresponding author: Mohammed S. Mohammed¹,
Mohammedsadaa6@gmail.com

ABSTRACT

Producing academic research papers stands as a paramount undertaking for researchers seeking to disseminate their expertise and establish their presence within their respective disciplinary communities. Exploring the linguistic attributes embedded within the rhetorical structures of academic research papers aids in enhancing our comprehension of this frequently employed genre. Consequently, the principal aim of this investigation was to scrutinize the functional elements of lexical bundles clustered within the rhetorical moves in the Discussion sections of specialized corpora containing qualitative and quantitative research articles (RAs) within the domain of Applied Linguistics. Each corpus encompassed a total of 50 Discussion sections. The data

underwent analysis based on the methodology outlined by Ruiying and Allison (2003) for discerning rhetorical moves. A corpus-driven methodology was employed to identify four-word lexical bundles and the functional taxonomy introduced by Salazar (2014) was employed to dissect these bundles. The scrutiny of the Discussion sections brought to light variations in the frequency of lexical bundles, highlighting those certain moves exhibited a greater prevalence of bundles compared to others. Notably, both qualitative and quantitative article authors demonstrated a similar pattern of preference, favouring text-oriented bundles over functional bundle categories in the Discussion sections. This study has contributed to a deeper understanding of lexical bundles and genre conventions within both qualitative and quantitative research papers, thereby equipping authors, especially those in the early stages of their academic journey, with enhanced knowledge and skills essential for proficient academic writing.

Introduction

Growing interest in the rhetorical moves used in academic discourse to apply formulaic sequences (Kashiha, 2014; Nam, 2017) and disseminate knowledge within a certain discourse community has been ignited by genre analysis (Tessuto, 2015; Lu, Casal, Liu, Kisselev, & Yoon, 2021). Research articles (Henceforth, RAs) are undoubtedly far more important type in academic discourse because they allow fellows to express their ideas and opinions to others in academic communities (Mohammad & Sadat, 2014; Li, 2020). The preferred academic genre known as RA was created to promote knowledge sharing among academics. It is also regarded as a measure of academic success. In order to join the academic community, researchers and academics must publish their work there (Dobakhti, 2011).

Nevertheless, it is essential to recognize that each academic discipline possesses distinctive clusters of norms and conventions that define its identity. Among these, the ones highly valued within a particular genre of academic disciplines are those that find frequent application (Vaseghi, 2016). In order to underscore the importance of their research and establish its worthiness as a subject for scholarly discourse, academic authors must acquaint themselves with, or acquire an understanding of, the

linguistic patterns and established norms specific to the genre they are targeting. These distinctions set the genre apart from other forms of academic writing, such as conventional classroom assignments (Dobakhti, 2011).

As articulated by Hyland (2005b), one of the hallmarks of exemplary academic writing is the author's capacity to craft a persuasive presentation of both them and their work. This involves aligning with readers, critically analyzing their research materials, and acknowledging alternative viewpoints. Furthermore, it is crucial to acknowledge that claims made by authors in research articles may face scrutiny or dismissal at any stage by members of the academic discourse community if authors do not meet the community's expectations (Mohammad & Sadat, 2014). This is because writing is an intentional, socially situated activity primarily intended for audiences within a specific discourse community. Authors must acquaint themselves with the linguistic characteristics and rhetorical strategies commonly employed in research articles pertaining to their respective fields.

As RA studies constitute a significant genre of information in academic discourse, researchers, particularly those working in the field of second language writing, have shown rising interest in them. In accordance with the existing body of knowledge, several methodologies have endeavored to define the communicative components within specific genres and provide a structure for scrutinizing the rhetorical organization found in diverse forms of written expression. Among these approaches is Swales' (1981) "move-steps" rhetorical analysis, which proposes a hierarchical categorization comprising moves and steps. Essentially, to accommodate more intricate rhetorical arrangements, a "move" can be dissected further into steps and sub-steps.

A variety of models and frameworks proposed by academics, professionals, and researchers, including those by Ruiying & Allison, (2004), Hyland (2000), Nwogu (1997), Bhatia (1993) and Swales (1990), have been used to analyse and describe the schematic structure of RAs, which consists of abstracts, introductions, literature reviews, methods, results, discussion, and conclusions. Based on their diverse emphasis, prior works in ESP genre analyses might be divided into two types. The structure of dissertations, RAs, and other professional writings is the focus of one group;

examples include Lu et al. (2021), Hong (2019), Amnuai (2019), and Tessuto (2015), while another category focuses on features specific to RAs like function-form connections, hedging, modality, and reporting verbs; such as Casal (2020), Omidian et al. (2018) Ruiying & Allison, (2004) and Hyland (2000).

This study aims to discover the recurring sequences of words utilised in conveying the discourse function of "move borders" in addition to doing a genre analysis. The purpose of the move and the language cues can be utilised to establish the move's bounds (Alamri, 2017). An expanding corpus of research underscores the prevalence of formulaic sequences in natural language, with Kashiha (2014) asserting that a considerable portion of discourse is constructed using these expressions. While there has been a recent surge in genre analysis studies, driven by the heightened demand for research articles across various research domains, achieving successful publication in the international academic arena necessitates researchers' comprehension of variations in the structural components of text, including the associated lexical bundles specific to each move and step.

The invasion of formulaic sequences (FSs) into natural language has often been demonstrated through research. They are essential to the growth of scholarly debate. Lexical bundles, which make up FSs, are groups of word forms that often collocate in real-world speech and are rated according to how frequently they do occur. These bundles are frequently seen in the rhetorical moves used in different academic discourse parts. They do aid in locating participants in a certain discourse's community. This implies that language users can identify with a specific community, such as a disciplinary one, by adopting formulaic sequences (Wray, 2002). According to Wray (2006), as cited in Del and Erman (2012), we use specific expressions when we speak that we believe are connected to specific ideals, styles, and groups. Lexical bundles are groups of words that commonly appear together in real language and greatly increase the fluency of both speech and writing (Shin, 2019). Repeated usage of these bundles in a particular register or field, according to Cortes (2006), is a sign of proficiency in language use. The absence of particular word sequences, on the other hand, suggests that a novice to that academic

community is lacking fluency of academic writing (Hyland, 2012).

The publication of research articles in recognised journals has many advantages for researchers, including the ability to disseminate a lot of knowledge within the academic community, professional promotion, status, and grant security. Recent proceedings have demonstrated that postgraduate students must publish research RAs before receiving their university degrees (Dobakhti, 2011). The understanding of qualitative and quantitative designs is crucial, especially for beginning authors. According to Creswell (2003), there are differences between the research designs of qualitative and quantitative studies in terms of the knowledge sort, claims made, main objectives sort, the research questions imposed, the data gathered, and the methods employed to analyse the data. Dobakhti (2011) emphasises that their usage of rhetorical moves and other lexical patterns also varies. Considering the significant distinctions between qualitative and quantitative research articles (RAs) in various aspects, inexperienced and budding authors often encounter difficulties in crafting a research article suitable for publication in scholarly journals due to their limited grasp of the specific discourse conventions associated with a particular research approach. Surprisingly, there has been limited or virtually no exploration in published research concerning the influence of research design on the rhetorical tactics and vocabulary patterns employed by qualitative and quantitative RAs.

Numerous research endeavors have delved into the rhetorical components of research articles (RAs) across various academic disciplines. For instance, Paydari (2017) explored political science, while Musa, Khamis, & Zanariah (2015) delved into engineering. Kanoksilapatham (2005) examined biochemistry, Peacock (2002) conducted investigations across seven diverse fields encompassing language and linguistics, physics, environmental science, biology, public and social administration, business, and law. Yang (2001) focused on applied linguistics, Posteguillo (1999) ventured into computer science, and Nwogu (1997) probed into medicine. These studies have primarily centered their attention on the structural organization of various sections within RAs.

Additionally, some research inquiries have employed the IMRDC (Introduction, Method, Result, Discussion, and

Conclusion) framework as their focal point, either in its entirety or in parts thereof. A predominant focus has been placed on the "Introduction," "Results," "Method," and "Discussion" sections of RAs. Notable examples include the works of Swales (1981, 1990), Ozturk, Brett, Williams, and Musa et al. (2015). Furthermore, there have been investigations conducted by scholars such as Holmes (2000), Dobakhti (2011), Nwogu (1997), Pho (2008), Shi (2014), and others who have shown interest in examining RAs across a wide spectrum of subject areas.

According to the existing scholarly literature, numerous prior investigations have scrutinized multi-word expressions in academic writing, with a predominant focus on both native and non-native English speakers. A selection of exemplar studies includes those conducted by Nam (2017), Shin (2019), Chen and Baker (2010), and Del and Erman (2012). Some of these studies, such as those by Jalali (2014), Güngör and Uysal (2016), Tovar-Viera (2018), and Lu et al. (2021), have engaged in cross-disciplinary comparisons of lexical bundles, while others have specifically explored the contrast between novice and expert utilization of lexical bundles in academic discourse, as exemplified by Jalali (2014b) and Wang (2018).

Surprisingly, there has been a scarcity of research comparing qualitative and quantitative research articles within a single field, notably within the under-researched domain of applied linguistics, as noted by Ruiying (2003). Furthermore, even the limited existing studies in this area have not focused on the rhetorical strategies and lexical bundles employed within qualitative and quantitative research articles. For instance, Dobakhti (2011) solely concentrated on general structural and stance-related characteristics.

In terms of the methodologies employed for analyzing lexical bundles, researchers have employed a variety of approaches, including corpus-based techniques (Biber, 2009; Esfandiari & Barbary, 2017b), frequency-based methods (Conrad Susan M., 1996), and other corpus-driven methods (e.g., Chen & Baker, 2010; Leelasetakul, 2019). Comparatively, there are limited studies in the literature that have compared genre-based and corpus-driven analytical techniques.

Regarding the length of lexical bundles, a few investigations have concentrated on four-word bundles, which represent the most extensively researched bundle type in academic discourse (Kashiha, 2019; Chen & Baker, 2010; also see Biber et al., 1999; Hyland, 2008; Vo, 2019). The current study aligns with this focus on four-word bundles. Notably, these four-word bundles incorporate three-word bundles into their structures more frequently, offering a broader array of structures and functions compared to three- or five-word bundles, as indicated by Hyland (2008) and Kashiha (2019). Similar assertions by Cortes (2004) and Hyland (2008) posit that four-word bundles tend to be more prevalent in spoken discourse than five- or six-word strings, making them more accessible for categorization and contextual comprehension (Chen & Baker, 2010; Biber et al., 1999; Cortes, 2013; Tetyana & Lee, 2017).

It's noteworthy that while numerous studies have investigated the use of rhetorical moves in research articles (RAs) along with lexical bundles and linguistic realizations, the majority of these studies have primarily emphasized disciplinary variations across different academic fields. The comparative analysis of move structures and lexical bundles within qualitative and quantitative RAs has yet to be empirically explored. Consequently, this study aims to fill this research gap by investigating rhetorical strategies and lexical bundles in the field of Applied Linguistics, which is perceived as a less-explored domain (Ruiying, 2003).

This study adopts a corpus-driven approach through a functional taxonomy to investigate the lexical bundles employed within the rhetorical moves of Discussion sections in qualitative and quantitative research articles in applied linguistics. The primary goal is to address the gap in prior research by examining how these sections, despite their significance and prevalence in academic discourse, present notable challenges for inexperienced RA writers due to their shared communicative functions within move and step structures (as identified by Ruiying & Allison, 2003, as cited in Alamri, 2017, Abdollahpour & Gholami, 2019; Hong, 2019). Specifically, our study compares the rhetorical moves utilized in the Discussion sections of qualitative and quantitative RAs, as well as the functional attributes of lexical bundles found within these sections in both groups of research articles.

Methodology

A. Study Design

The research design for this comparative mixed-methods study involves analyzing the discussion sections of qualitative and quantitative research articles to examine the use of lexical bundles and rhetorical moves. A corpus of articles will be compiled, and a computational linguistic tool (AntConc program) was utilized to identify lexical bundles and quantitatively analyze their frequency. Qualitative analysis was also conducted to identify and categorize the rhetorical moves employed. The research design integrates both qualitative and quantitative approaches, allowing for a comprehensive understanding of the linguistic patterns and rhetorical strategies used in each type of research. The findings were interpreted, compared, and contrasted to provide insights into the similarities and differences between qualitative and quantitative research discourse.

B. Corpora of the Study

In accordance with Kotevko (2014), a corpus signifies an assemblage of textual materials. Moreover, the term "corpus" denotes a compilation of written and spoken texts that encompass various linguistic subdomains, as articulated by Paltridge (2006). Within the domain of corpus linguistics research, practitioners frequently employ specialized computer software to scrutinize the collected texts with the aim of identifying instances of specific linguistic characteristics. According to Paltridge (2006), two distinct categories of corpora exist: generic corpora and specialized corpora. The former pertains to a set of texts that exemplify a particular genre, such as qualitative and quantitative research articles within a specific academic field or genre. Conversely, the latter alludes to a collection of publications that epitomize a specific genre. This study employs specialized corpora that were meticulously curated to extract the linguistic components under investigation.

In order to accomplish the study's objectives, the creation of specialized corpora demands careful consideration, primarily to investigate the textual structure of qualitative and quantitative research articles and to scrutinize the recurring lexical patterns associated with each recognized rhetorical move found within the

Discussion sections. Furthermore, it is evident that the current research adopts a comparative approach, focusing on the analysis of potential disparities between these two distinct modes of inquiry.

The articles constituting the corpora were selected from five renowned applied linguistics journals with high-impact factors, encompassing publications from the period between 2017 and 2022. The present analysis prefers journals that publish both qualitative and quantitative research articles, while journals exclusively featuring one approach were excluded. The five journals selected for this study are the Journal of English for Specific Purposes (ESP), Language Teaching Research (LTR), Applied Linguistics (APL), Journal of Pragmatics (JPR), and Journal of Sociolinguistics (JSL). The choice of research design, either qualitative or quantitative, determined the inclusion of research articles in the corpora. Consequently, the utilization of these corpora spanning both research approaches yields valuable insights into the characteristics of qualitative and quantitative inquiry within the domain of Applied Linguistics.

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

All the articles were copied and pasted into the plain text files after being retrieved from the PDF. Any further, unnecessary information (such as page numbers, headings and sub-headings, annotations, graphs, tables and figures) in the RAs were eliminated in an effort to tidy up those

document texts. It is worth emphasizing that the identification of the rhetorical move structure within the Discussion sections of qualitative and quantitative research articles constitutes a genre-based approach. The quantity of communicative objectives that are reached by the use of Move, or communicational units, determines how a genre is constructed.

D. Procedure for Rhetorical Move Identification

Considering that the concept of communicative intent plays a pivotal role in move identification within this study, it is crucial to acknowledge that move lengths fluctuate depending on the content an author seeks to convey in their discourse, employing an array of sentences, phrases, or words. It is entirely possible for a single move to fulfil multiple objectives, given the variation in move lengths and the author's use of numerous words within a move to clarify their intentions. In such instances, Holmes (1997) recommended that the analysis should prioritize the most significant function. Although the entire RA portions were not the subject of the study, each RA part was reviewed in its entirety to ascertain the overarching theme. To prove this claim, numerous RAs from the Qualitative (Quali) corpus were read numerous times and closely scrutinised in order to fully understand the Discussion. The seven-move model/analytical framework created by Ruiying and Allison (2003) was then used to analyse the Qualitative (Quali) and Quantitative (Quanti) corpora individually. Dobakhti (2011) suggests that the structure of movements and steps is realised by clause given that clause is the unit of analysis for moves and steps. The study's investigation of moves and steps was restricted to clause. As a result, any grammatical unit beneath clause was ruled unsuitable for investigation. Any statement that makes more than one move or step was subject to examination, with the most important one being taken into consideration. Again, to avoid unneeded complications, the embedded moves or steps were not taken into account throughout the analysis for this study. Alamri (2017) asserts that Ruiying and Allison's (2003) models are complete with regard to move analysis in the field of applied linguistics. Despite being more current than Ruiying and Allison's (2003) model and being viewed as all-inclusive by certain researchers (such as: Alamri 2017; Suherdi et al, 2020), Pho's (2008a) model provided alternatives for the examination of RAs' combined sections

as such, it is not adopted. Result-Discussion and Discussion-Conclusion, for instance, which were not the goal of this study.

TABLE 2 RUIYING AND ALLISON'S (2003) MODEL FOR 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

E. Procedure for Lexical Bundles Identification

It is noticeable that number of studies have set various criteria for the lexical bundles' identification. According to Shi (2014), identification of lexical bundles are mainly based on four criteria such as: move label, cut-off frequency, distribution threshold and length of bundles. The present study adopts these criteria in identifying the lexical bundles.

The move label criterion categorizes specific linguistic moves or discourse functions exhibited by frequently occurring word sets or phrases. Its purpose is to analyze the distribution and usage of lexical bundles within a specific genre or context. Meanwhile, the cut-off

frequency establishes a threshold below which lexical bundles are deemed infrequent or insignificant for inclusion in the analysis. This aids in setting a baseline frequency, ensuring the inclusion of only common or recurrent bundles. Similarly, the distribution threshold encompasses the number of texts in which lexical bundles occur across a given corpus to be considered valid across. Lastly, the length of bundles pertains to the number of words or tokens that comprise a specific lexical bundle or phrase.

The present study focuses on identifying multi-word sequences that indicate the function of the word label while exempting those that do not designate this function. To examine the functional features of lexical bundles, a specialized corpus was employed, leading to the establishment of a cut-off frequency of 10 times per million words (PMWs) to identify common lexical bundles at move boundaries, a method supported by Salazar (2014). Furthermore, a distribution threshold of three texts/files was set, excluding multi-word sequences that occurred in fewer texts/files from the analysis. Based on previous research indicating the prevalence and clear functions of four-word units (Cortes, 2006; Biber & Barbieri, 2007; Hyland, 2008; Vo, 2019), the study selected four-word formulaic clusters as the unit of analysis. Four-word sequences offer a richer display of bundle functions, occurring more commonly and providing a clear range of structures and functions compared to five or six-word sequences. It is worth noting that four-word bundles also subsume three-word bundles. For identification of the most prevalent four-word clusters, the Wordsmith Tools software was used, yielding 532 bundle types and 2174 bundle tokens.

For discourse function determination, only Concordance and Cluster/N-grams tools of AntConc were used in this study. They perform the function of extracting words or phrases such as lexical bundles, which are used contextually in the corpus and find pattern in language use. The concordance tool was used to determine the specific functions the lexical bundles perform. The lexical bundles were searched and KWIC were automatically generated to illustrate the point. After generating the KWIC, the researcher then took the co-text into account, i.e., what comes before and after the keywords searched (lexical bundles), to uncover their communicative functions.

The lexical bundles were analysed or classified based on communicative function using Salazar's (2014) functional taxonomy. Salazar's (2014) taxonomy has three broad classifications—research-oriented, text-oriented and participant-oriented— with their accompanying sub-categories. She proposed five sub-categories under the research-oriented category, which consist of location, procedure, quantification, description and grouping, text-oriented category contains 9 sub-categories such as additive, comparative, inferential, causative, structuring, framing, citation, generalization and objective while participant-oriented category encompasses three sub-categories including stance, engagement and engagement. Besides, there exist many functional taxonomies in the literature, but Salazar's (2014) is preferred as it is found to be more comprehensive and relevant, considering the fact that the study is focused on academic written genre, whereas other functional taxonomies cover different registers.

Table 3 Functional taxonomy of target bundles (adopted from Salazar 2014)

Research-oriented bundles Help writers to structure their activities and experiences of the real world	Text-oriented bundles Concerned with the organization of the text and its meaning as a message or argument.	Participant-oriented bundles Focused on the writer or reader of the text
Location Indicate place, extremity and direction at the site, the tip of, on the left	Additive Establish additive links between elements on the other hand, in addition to, in concert with	Stance Convey the writer's attitudes and evaluations is likely to, is necessary for, it is possible that, it is clear
Procedure Indicate events, actions and methods the onset of, was carried out, used to identify	Comparative Compare and contrast different elements as compared with, in contrast to, significantly different from	Engagement Acknowledgment Address readers directly it should be noted that, see Figure 1, as seen in
Quantification Indicate measures, quantities, proportions	Inferential Signal inferences and conclusions drawn from data	Recognize people or institutions that have

and changes thereof	found to be, these results suggest	participated in or
total volume of, a large	that, we	contributed
number of, the ratio of,	conclude that	to the study
a decrease in	Causative	a gift from, kindly
Description	Mark cause and effect relations	provided by
Indicate quality, degree	between	
and existence	elements	
the appearance of, the	as a result of, is caused by, by virtue of	
extent of, the presence	Structuring	
of	Text-reflexive markers that organize	
Grouping	stretches of discourse or direct the	
Indicate groups,	reader	
categories,	elsewhere in text as described	
parts and order	previously, as shown in figure, in the	
a wide range of, this	materials and methods section	
type of, the sequence	Framing	
of,	Situate arguments by specifying	
a portion of	limiting	
	conditions	
	in the case of, with respect to, on the	
	basis of,	
	in the presence of, with the exception	
	of	
	Citation	
	Cite sources and supporting data	
	it has been proposed that, as reported	
	previously,	
	studies have shown that	
	Generalization	
	Signal generally accepted facts or	
	statements	
	little is known about, is thought to be	
	Objective	
	Introduce the writer's aims	
	we asked whether, to show that, in	
	order to	

Results

A. Qualitative analysis

It was deemed essential to conduct a qualitative examination of the identified rhetorical moves within their respective contexts to attain a nuanced understanding of

the discursive functions that differentiate qualitative from quantitative research articles (RAs). The primary focus of this phase of analysis centered on investigating the utilization and functions of the bundles that were observed as a characteristic of either qualitative or quantitative RAs within each rhetorical move found in the Discussion sections.

Move 1: Background information

The study of the various bundles revealed significant differences between Discussions in qualitative and quantitative RAs in this step. For instance, it was found that writers in qualitative RAs used the bulk of the research-oriented bundles to describe the following noun's physical characteristics as well as to go into detail about various aspects of its quality, form, and size. The example below serves as an illustration of this:

"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 analyzed knowledge (QL8)."

Move 2: Reporting results

In this move, disparities in the bundles employed in qualitative and quantitative research articles (RAs) became evident. Qualitative RAs exhibited a slightly higher variety of distinctive move types and tokens compared to their quantitative counterparts. Furthermore, authors of qualitative RAs displayed a predilection for employing bundles that leaned towards descriptive language, thereby elucidating the attributes, extent, and presence of their data. On the other hand, within the realm of quantitative RAs, quantification emerged as a frequently employed sub-category, facilitating the articulation of data quantities, measurements, or proportions. The following examples are drawn from corpora encompassing both qualitative and quantitative RAs:

"The purpose of this experiment was to further identify the nature of the meaning expressed by the voseo negative command in AS (QL28)."

"Although the size of the corpus and frequencies of Engagement resources were different in the two studies, the results were almost the same (QN2)."

Move 3: Summarising results

Only qualitative RAs had bundles specific to this move, according to an analysis of the various bundles. There were only three of these bundles, which comprise it can be inferred, can be argued that, it is unlikely that as can be seen in (4) and (5), that these bundles are an inferential subcategory of research-oriented bundles (6). This implies that writers of qualitative RAs used these bundles to deduce conclusions about their findings, as seen by the illustrative scenario that follows:

“It can be inferred, that the values and attributes associated with SMU's brand image and branding discourse have found some resonance with individuals (QL37).”

“In sum, it can be argued that the DMAs’ seemingly contradictory developments have been triggered by pragmatic inferences from the speech situation (QL25).”

“It is unlikely that the teacher education they had received would have led to an understanding of the various components of language competence (QL33).”

Move 4: Commenting on results

It was observed that the RA authors used text-oriented bundles in both of the study designs employed this move such as the fact that the, findings of this study, on the other hand, in terms of the, to interpret their findings. On the contrary, it was discovered that in this step, writers of quantitative RAs used more distinguishing bundle tokens than writers of qualitative RAs. The qualitative RA writers preferred using inferential subcategories, such as: it appears that the, it may be that, are more likely to over the quantitative RA writers to signal implications and draw conclusions from the findings, despite the fact that both groups of writers favoured using text-oriented type of bundles in this move. As seen in the following citations, it implies that authors of qualitative RAs typically do not overgeneralize their findings and temper their judgements when interpreting them.

“Demonstrate that academic writers are more likely to provide the precondition for the feasibility of a process or the validity of the proposition (QL46).”

“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)."

Move 5: Summarizing the study

According to the analysis of the bundles found in this move, Quantitative RAs did not contain as many of the specified bundles as qualitative RAs did. It was found that authors of qualitative RAs mostly used a few bundles, namely the comparative sub-category of text-oriented bundles (results are consistent with, are in line with, for example), in order to compare their results to those of the prior studies. This outcome is in line with the findings from Omidian, Shahriari, and Siyanova-Chanturia (2018).

"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)."

Move 6: Evaluating the study

Certain discovered bundles were only present in quantitative RAs, according to the study of Move 6-specific bundles. By using the description sub-category of research-oriented bundles, such as: the effectiveness of the, the extent to which, such findings of this study, qualitative RA writers in this step took a more direct posture when analysing the results of their studies as shown in (10).

"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)."

Move 7: Deductions from the research

The findings show that both sets of RA authors used more participant-oriented bundles, especially stance sub-categories such as: it has been suggested, it can be concluded, it seems likely that, are more likely to, to represent the writer's opinions and evaluations. Contrary to their quantitative RA peers, writers of quantitative RA tended to employ more tokens but fewer types in this action. The following examples that show how to use these bundles were taken from the two datasets.

"It seems likely that the adult 'decoys' of the Perverted Justice data create different interactional patterns with these online offenders (QL17)."

"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

As illustrated in Table 3, which provides a breakdown of bundle types and the corresponding count of bundle tokens in each set of research articles, our identification techniques yielded a total of 2,174 bundle tokens across 532 distinct bundle types. Subsequently, after identifying these lexical bundles, they were categorized based on their alignment with the communicative objectives of the rhetorical moves from which they originated.

Table 3 Functional features of lexical bundles in the Discussion sections in both corpora

Categories	Sub-Categories	Quali Corpus		Quanti Corpus	
		Types	Tokens	Types	Tokens
Research-oriented Bundles	Location	13	51	17	150
	Procedure	18	72	27	96
	Quantification	16	52	25	133
	Description	30	121	22	85
	Grouping	6	22	8	24
	Sum	83	318	99	488
Text-oriented Bundles	Additive	10	36	9	49
	Comparative	16	45	17	57
	Inferential	28	68	51	210
	Causative	8	29	15	82
	Structuring	15	64	18	98
	Framing	17	72	23	111
	Citation	9	36	33	112
	Generalization	4	12	5	15
	Objective	6	16	8	23
	Sum	113	372	179	757
Participant-oriented Bundles	Stance	15	46	29	143
	Engagement	4	7	10	43
	Acknowledgement	0	0	0	0
	Sum	19	53	39	186
Total		215	743	317	1431

Figure 1 Distribution of Functional Types of the Categories of Bundles in the Discussion Section in the Two Corpora

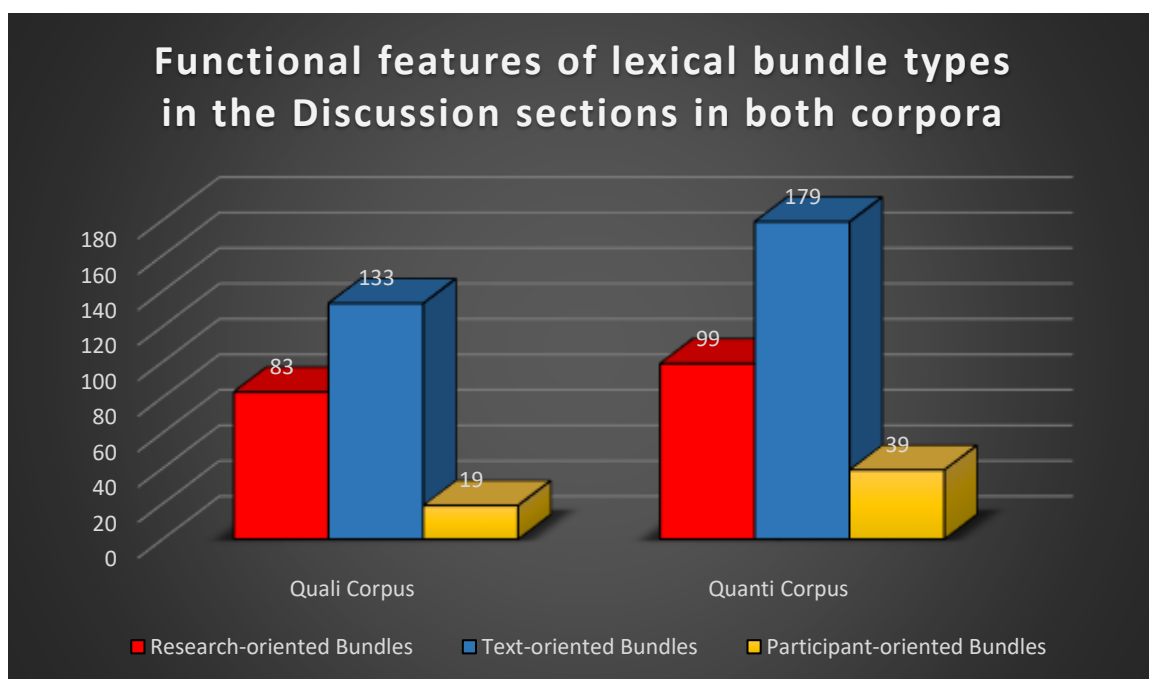
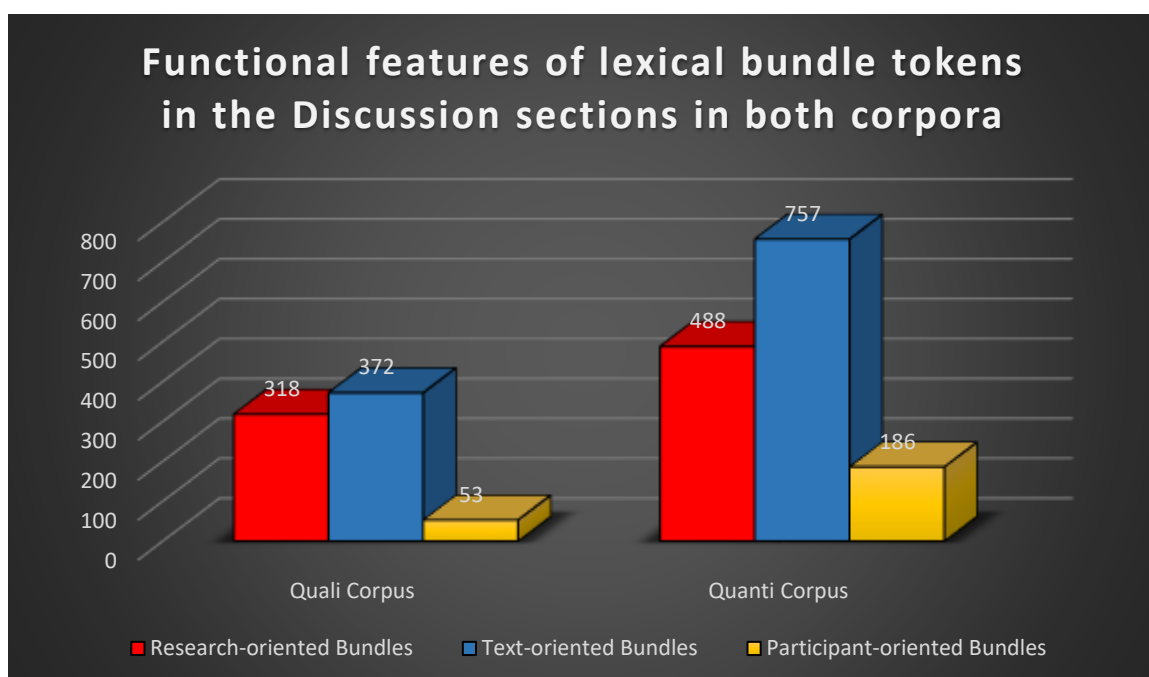


Figure 2 Distribution of Functional Tokens of the Categories of Bundles in the Discussion Section in the Two Corpora



Based on the provided data, the distribution of functional categories in the texts of qualitative and quantitative corpora was analyzed. It was observed that text-oriented bundles were the most frequently used category, accounting for approximately 53% of the target bundle tokens in both the qualitative and quantitative corpora.

Research-oriented bundles constituted about 39% and 34% of the target bundle tokens in the two datasets, while participant-oriented bundles were the least frequently used, making up 8% and 13% of the target bundles in the respective datasets.

Furthermore, when comparing the number of types and tokens between the qualitative and quantitative corpora, the quantitative corpus had a higher count. It contained 317 types and 1431 tokens, which was significantly higher than the qualitative corpus, which had 215 types and 743 tokens. This indicates a larger sample size and variety of functional bundles in the quantitative corpus compared to the qualitative counterpart.

Thus, the data analysis reveals the distribution of functional categories in the two corpora, highlighting the prevalence of text-oriented bundles and the differences in the number of types and tokens between the qualitative and quantitative datasets.

Discussion

The present study delved into the variability in research design as reflected in the discussion sections of research articles (RAs) by examining the use of lexical bundles. The primary objective was to assess the extent to which authors of qualitative and quantitative RAs employed lexical bundles in achieving the communicative goals of their discussion sections. A mixed-methods approach, encompassing both statistical and qualitative analyses, was employed to discern patterns of variation between these two types of RAs.

The initial comparative analysis was grounded in the distinct structural elements typically found in the various moves within RA discussion sections, with the aim of uncovering potential disparities between qualitative and quantitative RAs in the context of these sections. It was observed that, in Move 4 (Commenting on Results), writers of quantitative RAs employed a substantially higher number of bundle tokens, signifying the conventional nature of this approach within quantitative RAs. This aligns with Hyland's (2016) observation that genre usage is influenced by societal norms and expectations. It can be argued that the demands of quantitative RAs, driven by their research design, necessitate a more skillful application of rhetorical

techniques when commenting on study findings, resulting in the standardization of certain structures. This finding is consistent with Dobakhti's (2011) discovery that writers of quantitative RAs commonly use posture features when commenting on results.

Furthermore, the study revealed that qualitative RA writers, despite relying more heavily on text-oriented bundles compared to the other two groups of RA writers, preferred descriptive bundles over other text-oriented sub-categories. In contrast, their quantitative counterparts exhibited a preference for employing quantification bundles. This distinction is understandable, given that qualitative research primarily focuses on the description, explanation, and interpretation of data, whereas quantitative research aims to statistically quantify data for specific purposes.

In terms of move-specific analysis, quantitative RA writers were found to use significantly more bundle tokens but fewer bundle types in fulfilling the communicative functions of Move 5 (Summarizing the Study) compared to qualitative RA writers. This suggests that qualitative RA writers may exhibit less diversity in their bundle usage, emphasizing the conventional acceptance of specific bundles for achieving this move.

The study also highlighted that lexical bundles utilized to accomplish the discourse function of Move 3 (Summarizing the Results) were exclusive to qualitative RAs, while quantitative RA writers exclusively employed certain bundles to realize the communicative purposes of Move 6 (Evaluating the Study). This implies that authors of qualitative and quantitative RAs prioritize different aspects when presenting their research. Qualitative RA writers place greater emphasis on the uniqueness of their study summary as a distinguishing feature, while quantitative RA writers tend to concentrate on evaluating the study's findings. These findings align with those reported by Dobakhti (2011) and Omidian et al. (2018).

In summary, it is evident that, despite some commonalities, qualitative and quantitative RAs exhibit significant differences in how they fulfill their communicative roles through various rhetorical moves. It is important to note that these results are based on a relatively small corpus and may warrant further

investigation with larger datasets to assess the generalizability of these findings. Consequently, this study can offer valuable insights to graduate students and researchers aiming to contribute to RA publications while navigating the conventions associated with different research designs.

Conclusion

The present study sheds light on the preferences of authors and how they tailor their presentations within the discussion sections of research articles (RAs) across diverse academic disciplines. It was observed that qualitative RA writers often prioritize conveying the essence of their research by employing a methodological fusion of corpus linguistics and genre analysis. Conversely, quantitative RA authors place significant emphasis on evaluating their study and drawing conclusions to advance its trajectory. Although both groups of RA writers made extensive use of text-oriented bundles, a notable distinction emerged: qualitative RA writers exhibited a proclivity for employing descriptive bundles over other text-oriented sub-categories, while their quantitative counterparts leaned towards the utilization of quantification bundles.

Unveiling these disparities is expected to significantly enhance our understanding of how authors in various academic domains generate knowledge and engage with readers. Such insights hold the potential to facilitate the effective composition of research design papers. Moreover, these findings are particularly valuable for emerging researchers and graduate students seeking to participate in scholarly writing within their respective fields, especially when they may not possess the requisite disciplinary expertise and skills. For this specific cohort of language users, a comprehensive grasp of research design writing can be achieved by examining research design standards through the lens of their linguistic expressions in discourse.

References

- Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31, 81–91.

- 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.
- 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., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26(3), 263–286.
<http://doi.org/10.1016/j.esp.2006.08.003>
- 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.
- Conrad Susan M. (1996). Investigating Academic Texts With Corpus-Based Techniques : An Example From Biology. *Linguistics and Education*, 326, 299–326.
- Cortes, V. (2006). Teaching lexical bundles in the disciplines : An example from a writing intensive history class, 17, 391–406.
- Dobakhti, L. (2011). The Discussion Section of Research Articles in *Applied Linguistics : Generic Structure and Stance Features*.
- 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.
- Güngör, F., & Uysal, H. H. (2016). A Comparative Analysis of Lexical Bundles Used by Native and Non-native Scholars, 9(6).
- 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.
- Hong, J. (2019b). 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. (2008a). Academic clusters: text patterning in published and postgraduate writing, 18(1).
- 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 Hassan Jalali English Language Center , Isfahan University of Technology, 3(1).
- Jalali, H. (2014b). Examining Novices ' Selection of Lexical Bundles : The Case of EFL Postgraduate Students in Applied Linguistics, 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(June), 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*. Palgrave Macmillan.
- Leelasetakul, M. (2019). Lexical Bundles used by Thai undergraduate learners of English.
- Li, X. (2020). ScienceDirect Mediating cross-cultural differences in research article rhetorical moves in academic translation : A pilot corpus-based study of abstracts. *Lingua*.
- 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.
- Musa, N. F., Khamis, N., & Zanariah, J. (2015). The Structure of Method Section in Engineering Research Articles The

- Structure of Method Section in Engineering Research Articles. *Asian Social Sciences*, 11(17).
- 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.
- Paltridge, B. (2006). *Discourse Analysis: An Introduction*. Continium publishing group.
- Paydari, S. S. (2017). Rhetorical Moves and Discourse Structure in Political Science Research Articles abstracts of Iranian journals. *Universiti Putra Malaysia*.
- Peacock, M. (2002). Communicative moves in the discussion section of research articles. *System*, 30, 479–497.
- Pho, D. P. (2008). How can Learning about the Structure of Research Articles Help International Students ?, 1–11.
- Posteguillo, S. (1999). The Schematic Structure of Computer Science Research Articles. *English for Specific Purposes*, 18(2), 139–160.
- 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). E NGLISH FOR Research articles in applied linguistics : structures from a functional perspective q, 23, 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. *Suranaree University of Technology*.
- Shin, Y. K. (2019). Do native writers always have a head start over nonnative writers ? The use of lexical bundles in college students ’ essays. *Journal of English for Academic Purposes*, 40, 1–14.
- Suherdi, D., Kurniawan, E., & Lubis, A. H. (2020). A genre analysis of research article ‘findings and discussion’ sections written by Indonesian undergraduate EFL students. *Indonesian Journal of Applied Linguistics*, 10(1), 59-72.

- 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. *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.
- Williams, I. A. (1999). Results Sections of Medical Research Articles: Analysis of Rhetorical Categories for Pedagogical Purposes. *English for Specific Purposes*, 18(4), 347–366.
- Wray, A. (2002). Formulaic Language and the Lexicon.
- Yang, R. (2001). A genre analysis of research articles in applied linguistics. *National University of Singapore*.