Language Use in Academic Contexts: A Multidimensional Analysis of Business and Engineering Student Writing

Yoo Lae Kim

A Dissertation

In the Department

of

Education

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Education)

at Concordia University

Montréal, Québec, Canada

April 2025

CONCORDIA UNIVERSITY SCHOOL

OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: *Yoo Lae Kim*

Entitled: *Language Use in Academic Contexts: A Multidimensional Analysis of Business and Engineering Student Writing*

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (EDUCATION)

compiles with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair

Dr. Julie Corrigan

_____ External Examiner

Dr. Larissa Goulart

_____ Arm's Length Examiner

Dr. Heike Neumann

_____ Examiner

Dr. Sara Kennedy

_____ Examiner

Dr. Pavel Trofimovich

_____ Thesis Supervisor

Dr. Kim McDonough

Approved by _____

Dr. Walcir Cardoso, Graduate Program Director

March 27, 2025 _____

Pascale Sicotte, Faculty of Arts and Science Dean

ABSTRACT

Language Use in Academic Contexts: A Multidimensional Analysis of Business and Engineering Student Writing

Yoo Lae Kim, Ph.D.

Concordia University, 2025

This dissertation investigated linguistic variation in student writing across Business and Engineering disciplines through Multidimensional Analysis (MDA). Analyzing a corpus of business case studies and engineering proposals, the study identified five dimensions of variation: Informational Density and Elaborated Discourse, Interactive and Situated Discourse, Narrative Focus, Interpersonal and Dialogic Discourse, and Descriptive and Stative Discourse. These dimensions revealed distinct communicative purposes and rhetorical strategies reflecting disciplinary conventions.

Business case studies paid attention to clarity and situational relevance by employing interactive language to engage stakeholders and present actionable solutions. In contrast, Engineering proposals emphasized technical precision and informational density, relying on nominalizations, dense noun phrases, and explicit references to describe systems and processes. Both disciplines also demonstrated shared linguistic features that underline the importance of clarity and coherence.

The findings have significant implications for writing pedagogy. Discipline-specific instruction should address the unique linguistic demands of each field while fostering foundational skills transferable across contexts. For Business education, this involves enhancing students' ability to construct persuasive and audience-aware arguments. Engineering writing

instruction should give focus to the technical language and precision. The study further raises questions for future research, including the integration of evolving technologies and the exploration of linguistic variation in interdisciplinary genres. By providing a comprehensive framework for analyzing linguistic variation, this research bridges the gap between academic instruction and professional communication, equipping students to meet the demands of their respective fields.

Acknowledgements

This dissertation is a reflection of the incredible support I have been fortunate to receive from my committee, family, and friends throughout this challenging yet rewarding journey.

First and foremost, I want to express my heartfelt thanks to my supervisor, Dr. Kim McDonough. Working with you over these past few years has been an incredible privilege. I will always cherish the mentorship you have shared with me, and I am grateful for the opportunities and support you have provided every step of the way. To my committee members – Dr. Pavel Trofimovich, Dr. Sara Kennedy, Dr. Heike Neumann, and Dr. Larissa Goulart – thank you for your time, expertise, and constant support. Your insights have shaped this dissertation in meaningful ways, and I am so grateful to have you on this journey with me.

A special thanks to the Concordia Applied Linguistics (CAL) lab members for your encouragement and kindness over the years. I also want to say thanks to Wonder Women ladies for being my laughter and my safe space. Your endless support and our shared moments of joy have been a part, and I feel so lucky to have you along with this journey. I also want to thank my wonderful friends in Montreal and Korea, who have cheered me on every step of the way. Your constant encouragement has been a source of strength during the ups and downs of this journey. Every chat and every moment of laughter has brought so much light into my life. I feel lucky to have each of you in my corner, and your friendship has made this experience more meaningful.

I would also like to say thank you to Dr. Bill Crawford, who has now been a part of this PhD journey. Working with you during that beautiful summer in Flagstaff was a truly meaningful

experience, and I am so grateful for the opportunity to learn from you. Our research inspired this dissertation and laid the groundwork for it, and I deeply appreciate the impact you've had on my work.

Finally, my deepest gratitude goes to my family. Mom and Dad, thank you for believing in me, supporting me unconditionally, and always being there, even from afar. Your love and encouragement have been my greatest source of strength. And to my brother – thank you for being silly (that's what brother is for, you know), which has cheered me up.

Table of Contents

List of Tables

List of Figures

Glossary

Figures 1 and 2 outline the key concepts central to this dissertation by illustrating the interconnections between student writing, linguistic variation, and the analytical methods used to study these elements. Student writing, situated within academic and disciplinary contexts, involves producing texts that respond to the varied academic demands faced by students. Key to understanding the texts is the linguistic variation that is shaped by factors such as genre, register, and situational characteristics. These aspects, including communicative purposes, audience, and settings, influence language use.

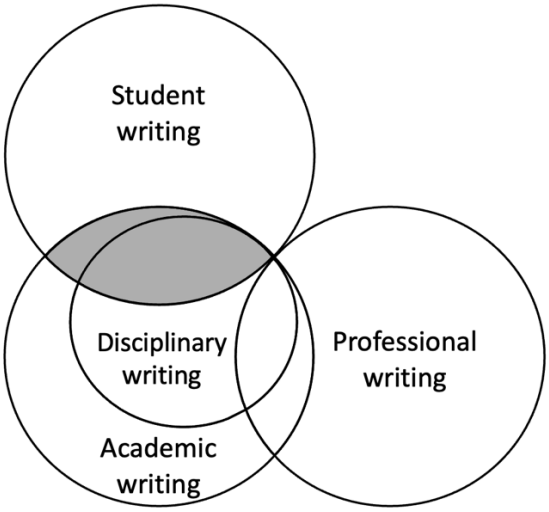

Figure 1. The relationship between concepts in writing

The colored area in Figure 1 represents the overlap between academic, disciplinary, and student writing, which emphasizes how these writing contexts converge in student disciplinary writing. This overlap highlights the multifaceted nature of student writing, shaped by academic rigor and discipline-specific expectations.

Figure 2 also illustrates how the lexico-grammatical approach is used as analytical framework for examining linguistic variation across contexts. By focusing on both the linguistic features and the contexts in which language is used, these approaches enable a detailed analysis across different contexts drive variation in student writing. A glossary following the figures provides precise definitions of key terms, offering further clarification and helping the interpretation of these concepts.

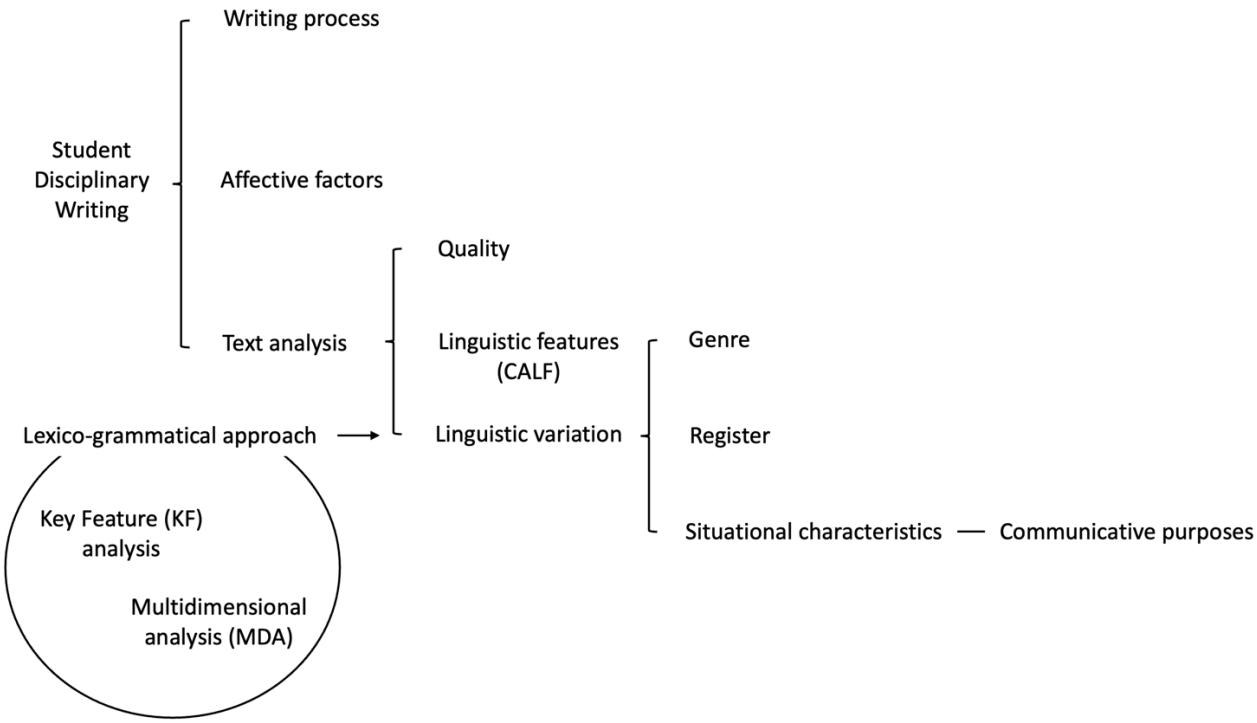

Figure 2. The relationship between concepts in student disciplinary writing

Academic writing: A formal style of writing used in scholarly communication which includes both student and professional writing. Academic writing involves various genres such as assignments, exams, dissertations, research papers, and journal articles, and can be characterized by its tone, structured approach, and following conventions within the field

(Hyland, 2009; Swales, 1990). This dissertation will focus on student disciplinary texts, which can be a foundation for developing the skills necessary for future academic and professional success.

Communicative purpose: The primary intention or goal that a writer seeks to achieve through language use such as informing, persuading, or instructing the audience. It shapes the content, structure, and style of a text to effectively deliver the intended message (Swales, 1990).

Disciplinary writing: The specific conventions, styles, and formats used within different academic disciplines. Each discipline has its norms including terminology, structure, methodology, and evidence types to communicate that follows the standards of that field (Beaufort, 2008).

Genre: The conventional structures, purposes, and features that define a particular type of communication or text. Genres are typically associated with specific formats, expectations, and goals such as a research article, lab report, email, or essay. Each genre has distinctive characteristics that guide how content is organized, what language is appropriate, and what the readers expect in terms of style and purpose (Swales, 1990).

Key Feature (KF) analysis: A quantitative method for comparing text varieties by identifying distinguishing lexico-grammatical features using Cohen's *d* effect size measures. The analysis provides a simpler approach focused on individual texts and functionally motivated linguistic features (Egbert & Biber, 2023).

Lexico-grammatical approach: An analytical approach that emphasizes the interconnectedness of lexis (vocabulary) and grammar in constructing meaning with texts. It examines how words and grammatical structures convey specific meanings in different contexts (Biber, 2019).

Linguistic features: The specific elements of language used within a text including both grammatical structures and vocabulary. These features are often analyzed to understand how different texts function and can be categorized within the CALF (Complexity, Accuracy, Lexis, and Fluency) framework (Wolfe-Quintero et al., 1998). For example, complexity refers to sentence structure such as syntactic variety and sentence length; lexis covers word choice, diversity, and sophistication. Together, these elements contribute to the overall meaning, effectiveness, and proficiency of a text.

Linguistic variation: The differences in language use that can be observed across different contexts that were influenced by factors such as social background, geographical region, or communicative situation. This variation occurs across multiple dimensions, reflecting how language adapts based on the situational characteristics (Biber & Finegan, 2014).

Multidimensional Analysis (MDA): A statistical method in linguistics that analyzes and interprets sets of language data across multiple dimensions to discover patterns of language variation. It is used in register analysis to identify and compare linguistic features across different contexts and provide a detailed understanding of language use (Biber, 1998).

Register variation: The differences in language use that occur based on situational context, purpose, and audience, often categorized into different registers such as formal vs. informal or academic vs. conversational (Biber, 1995).

Register: It is concerned with the linguistic choices that vary based on the situation or context of communication, including factors like formality, the relationship between speakers or writer/reader, and the mode of communication (e.g., spoken vs. written). This can vary widely even within a single genre, for example, a scientific report may be written in a highly formal register when published in an academic journal, while a less formal register may be used if the same information is presented in a blog post or lecture (Biber, 1995).

Situational characteristics: Contextual factors that influence language use in communication such as communicative purpose, participant, and setting. These characteristics shape the linguistic choices made in a given context (Biber & Conrad, 2019).

Student writing: The range of written texts produced by students in academic contexts, such as essays, research papers, lab reports, and exams. It includes both informal and formal writings that may or may not be evaluated by instructors, peers, or external reviewers. Student writing can extend to published work such as journal articles or conference papers (Hyland, 2000).

Syntactic complexity: The range and sophistication of sentence structures in spoken or written language. It is typically through various metrics such as the length of clauses, the number of subordinate clauses, and the overall variety and depth of sentence constructions. While syntactic complexity is often discussed alongside lexical complexity (the richness of vocabulary), the two are distinct yet complementary aspects of language development and proficiency. In academic writing, higher syntactic complexity often correlates with advanced writing skills, as it demonstrates the writer's ability to construct more intricate sentences. However, it is important to note that syntactic complexity alone does not guarantee clarity or effectiveness in

communication; it must be balanced with accuracy and appropriateness of lexis and grammar (Ortega, 2003).

Chapter 1. Introduction

1.1 Overview of Academic Writing Research

Academic writing is crucial in higher education as a primary means for constructing, sharing, and evaluating knowledge. Its significance lies in its ability to convey complex ideas clearly and systematically, which allows for critical engagement and scholarly dialogue (Hyland, 2009). As the standard means of communication in academia, writing plays a role in knowledge dissemination and intellectual exchange. These demands present challenges for all students, regardless of their linguistic or cultural backgrounds, who need to adapt to complex linguistic and disciplinary norms for effective participation in academic discourse.

Research on academic writing encompasses multiple domains including cognitive processes involved in writing, skill development, and the effectiveness of instructional methods. For example, Flower and Hayes (1981) explored the writing process through their cognitive models, which outline the mental activities involved in writing such as planning, translating, and reviewing. Additionally, Swales (1990) introduced a genre-based approach focusing on the conventions and structures of different types of academic texts, while Gee (1999) emphasized the role of social contexts and practices in shaping writing. Complementing these perspectives, the Academic Literacies framework, proposed by Lea and Street (1998), argued the negotiation of power, identity, and meaning within academic communities and offered a sociocultural lens through which to understand the challenges students face. For example, Fernsten and Reda (2011) emphasized the importance of addressing students' self-perceptions as writers, especially those struggling with negative writer identities. Authors argued that many students

internalize labels like “bad writers,” which can lead to resistance or fear of failure in academic writing tasks. Thus, educators can provide a foundation for more effective engagement with academic discourse to invite students to challenge these negative perceptions.

One of the key challenges for students is balancing clarity and complexity, as the need to express complex ideas often lead to dense and inaccessible prose (Swales & Feak, 2004). Furthermore, students need to navigate disciplinary-specific conventions to adapt their style to align with distinct norms (Hyland, 2000). For example, writing in the Humanities usually emphasizes interpretative and argumentative approaches, whereas the Sciences demand empirical precision and methodological rigor. The globalized nature of higher education adds another layer of complexity that requires students to write for interdisciplinary and multicultural audiences.

In this light, this dissertation focuses on student writing in higher education, with particular attention to linguistic and structural variations across disciplines and text types. By analyzing language use in student writing, this study aims to contribute to our understanding of how students navigate the linguistic demands of their respective fields. Findings gained from this investigation can inform more effective instructional strategies and enhance academic writing practices across educational and professional contexts.

1.2 Importance of Student Writing

Writing is an essential academic skill that supports learning, personal development, and professional preparation. Through writing assignments such as essays, research papers, and reports, students can engage deeply with the subject matter, organize their thoughts, and

present evidence-based arguments (Hyland, 2005). This process not only facilitates knowledge retention but fosters intellectual growth (Bean & Melzer, 2021). Writing also cultivates critical thinking by requiring students to analyze, evaluate, and synthesize information, develop coherent arguments, and assess counterarguments (Birkenstein & Graff, 2018). These skills enable students to articulate ideas clearly and persuasively, which are crucial for participating in academic discourse and professional communication (Leki, 2001).

The importance of writing extends beyond academia, as effective written communication is highly valued in professional fields such as business, engineering, law, and medicine. For instance, creating business proposals or technical reports requires precision and clarity to deliver ideas effectively and meet professional expectations (Swales & Feak, 2004). Writing also works as a bridge between academic and professional competencies to make students translate their skills into workplace success. For example. Business writing emphasizes persuasive rhetoric and clarity, while Engineering writing demands precision and technical accuracy, preparing students for distinct professional challenges. To further expand, writing assignments such as reflective essays and portfolios also help students to critically evaluate their own learning experiences and make connections between theory and practice. This reflective aspect fosters self-awareness and encourages lifelong learning skills that are important in both academic and professional contexts (Moon, 2006). Additionally, collaborative writing tasks such as group reports or joint research proposals teach students the value of teamwork, negotiation, and shared responsibility, where skills are highly prized in the workplace (Storch, 2019). By emphasizing these broader dimensions, academic writing contributes significantly to holistic student development.

Research further argues the multifaceted role of writing. Graham et al. (2012) emphasized that writing is not only a tool for learning but also a mechanism to enhance reading comprehension and content mastery across disciplines. Zhu (2004) noted the significance of discipline-specific writing tasks, which prepare students to address varied audiences and purposes effectively in contexts like business and engineering, where writing is central to communication. Considering the significant role of writing in higher education, a detailed examination of the linguistic elements that influence student writing can lead to more targeted and effective writing instruction across text types and disciplines. For instance, lexical choices play a role in effective writing. A rich and varied vocabulary allows students to express their ideas more precisely and persuasively. As Nation and Nation (2001) noted, a strong vocabulary enhances clarity and impact to make students write with sophistication and appropriateness. Recognizing these fundamental elements provides a solid foundation for evaluating and enhancing academic writing practices. However, it is also necessary to address specific challenges and gaps in the current understanding of academic writing, leading to the rationale for this study, which seeks to address these challenges and offer actionable insights to the field.

1.3 Key Aspects of Academic Writing

Understanding the significance of writing in higher education requires an exploration of its aspects such as accuracy, fluency, complexity, coherence, audience awareness, genre, and register. Each of these aspects poses unique challenges for students as they develop their writing skills. Studies (e.g., Flower & Hayes, 1981; Swales, 1990; Paltridge, 2004) revealed the multifaceted nature of academic writing and its dependence on both cognitive and sociocultural factors.

Linguistic accuracy in effective writing plays a role, as errors can obscure meaning and weaken arguments (Polio, 1997). Fluency, which involves logical flow and coherence, ensures that arguments are presented clearly and persuasively. Research by Flower and Hayes (1981) suggested that fluency depends on cognitive strategies for organizing ideas, while Paltridge (2004) argued fluency also requires anticipating audience needs and effectively structuring content. Complexity often demonstrated through varied sentence structures and sophisticated language use reflects a writer's capacity to engage with complex ideas and present them effectively (Biber et al., 2011). Coherence also makes sure that ideas are logically connected and easy to follow. According to Halliday and Hasan (1976), coherence depends on the effective use of cohesive devices such as conjunctions, lexical repetition, and pronoun referencing. Without coherence, even accurate and complex writing can fail to communicate its intended message effectively. Audience awareness is another key aspect of academic writing. Writing for a specific audience involves making appropriate linguistic and structural choices to meet the expectations of the readership. This is especially important in higher education, where students are often required to write for varied audiences, including peers, instructors, and broader academic communities. Audience awareness often requires students to strike a balance between simplicity for accessibility and sophistication for scholarly rigor. Another essential dimension of academic writing is argumentation. Effective academic writing often involves constructing well-reasoned arguments supported by evidence. Walton's (2013) argumentation schemes guide students in constructing arguments that are logical, evidence-based, and adaptable to various disciplines. By focusing on patterns of reasoning and incorporating audience awareness,

students can foster their critical thinking and analytical skills that are relevant to academic and professional writing.

Academic writing is also deeply embedded in the sociocultural practices of academic communities. Gee (1999) argued for the role of social context in shaping writing practices, while Paltridge (2004) explored how discourse communities and the political dimensions of text production influence writing. The Academic Literacies framework (Lea & Street, 1998) extends these perspectives by emphasizing the negotiation of power, identity, and institutional norms within academic writing. Students need to develop a repertoire of linguistic and rhetorical strategies that allow them to participate meaningfully in these academic contexts. Finally, understanding genre and register is important in academic writing. Genre knowledge enables students to meet the structural and rhetorical conventions of specific text types such as essays, reports, or research articles (Swales, 1990). Register, which includes the level of formality, tone, and choice of vocabulary, makes that writing aligns with disciplinary and situational expectations.

1.4 Academic Writing in Business and Engineering

In the Business and Engineering disciplines, writing is integral to developing both academic and professional skills. However, the expectations for writing in these fields differ due to their distinct communicative purposes and conventions. In Business programs, students often engage with professional genres such as business case reports, proposals, and memos. Drawing on Gardner and Nesi's (2013) classification of genres, this dissertation adopts the term "case study" to refer to the type of Business writing students produced. Case studies typically involve

analyzing scenarios to provide structured recommendations that require clearly, specialized terminology, and an ability to adopt professional roles (Nathan, 2013). Moreover, Business writing also frequently requires students to analyze market trends, interpret financial data, and propose strategic initiatives. These tasks demand a synthesis of technical knowledge and persuasive communication skills. For instance, a business plan needs to present detailed data while also making a compelling case for investment, balancing rigor with rhetorical appeal. Zhu (2004) argued that such assignments are often designed to initiate students into the real business world with problem-solving, decision-making, and teamwork skills.

In contrast, Engineering writing bridges technical proficiency and communication skills to articulate complex technical ideas clearly and logically. This study adopts Gardner and Nesi's (2013) classification of genres to categorize Engineering student writing as "proposals," focusing on detailed plans and persuasive arguments for future actions. Engineering proposals and technical reports usually require precision, clarity, and accuracy to deliver technical data effectively and without ambiguity (Wheeler & McDonald, 2000). Gardner (2016) observed that Engineering writing includes a diverse range of genres, from methodology recounts to design specifications, reflecting the multidisciplinary nature of the field. Engineering assignments often demand the integration of technical data with management strategies, particularly at advanced levels of study, where professional oriented tasks become more prominent.

Hyland (2008) emphasized that genres are socially situated and shaped by disciplinary conventions and require students to adapt their writing to meet the expectations of their specific academic communities. This shows the importance of genre-based pedagogies, which make explicit the structures and rhetorical strategies necessary for effective communication

within particular disciplines. In Business writing, for example, the persuasive and results-oriented nature of case studies reflects the community's emphasis on strategic decision-making. Conversely, Engineering writing's reliance on technical accuracy and methodological rigor stems from its goal to provide unambiguous solutions to practical problems. Zhu (2004) also noted that writing in these disciplines often mirrors professional tasks that prepares students for real-world challenges by fostering skills in audience awareness, evidence-based reasoning, and precise communication.

Understanding these genre-specific demands is essential for developing instructional strategies that address the needs of students in such disciplines. Targeted support can help students advance their writing skills by balancing technical precision with effective communication, ensuring they are well-prepared for both academic and professional challenges.

1.5 Rationale for the Study

The rationale for this study is grounded in the need to further explore and understand the complexities of linguistic variation in student writing across different disciplines and text types. Despite extensive research on academic writing, gaps remain in our understanding of how linguistic features vary between academic fields and the specific demands placed on student writers. This study, therefore, seeks to fill these gaps by investigating the distinctive linguistic characteristics that differentiate student writing across disciplines and text types.

Understanding these differences is important for multiple reasons. First, it can inform the development of more effective, discipline-specific writing instruction. For example, while

Humanities writing often emphasizes interpretative approaches, Science writing prioritizes empirical evidence and methodological precision (Becher, 1994). By recognizing these differences, educators can provide more targeted support that enables students to meet the particular demands of their respective fields. Additionally, acknowledging linguistic variation allows for more reliable assessment practices that account for the norms and expectations of different disciplines. This makes sure that evaluations are fair and reflective of the diverse writing contexts students encounter.

This study contributes to theoretical knowledge by offering perspectives on the dynamic relationship between language use and disciplinary practices. It explores how language functions within specific academic contexts, evolves in response to disciplinary demands and adapts to meet these expectations. These theoretical views enhance our understanding of academic discourse and its role in intellectual engagement and knowledge production. Practically, this study addresses the needs of learners navigating diverse linguistic demands. By identifying and describing disciplinary variation, the findings can inform the development of instructional materials targeted to specific text types and disciplines. For instance, discipline-specific writing guides and workshops can equip students with strategies to confront the particular challenges of their fields. Furthermore, findings can help educators design assessment tools that accurately capture students' writing abilities within their academic and professional contexts. This study investigates the importance of linguistic variation in enhancing both academic writing and the quality of higher education instruction.

1.6 Organization of the Dissertation

The first chapter has provided an overview of the dissertation with an introduction to the background and rationale of the study. Chapter 2 contextualizes the present study by reviewing related literature on linguistic variation in academic writing and identifying gaps in existing research that the current study aims to address. Chapter 3 outlines the methodological approach, including corpus description and analysis. Chapter 4 presents the results of the analyses with the linguistic variations identified in different academic disciplines and text types. Chapter 5 discusses the findings in the context of the existing literature and explores their implications for learning academic writing and its instruction. Finally, Chapter 6 concludes the dissertation by summarizing the main findings, discussing their broader implications, and providing recommendations for future research.

Chapter 2. Literature Review

2.1 Introduction

This chapter provides an overview of previous research on linguistic variation in writing, register variation, and Multidimensional analysis (MDA). Section 2.2 examines key studies on linguistic features in writing, followed by Section 2.3, which discusses linguistic variation within various contexts. Section 2.4 introduces the MDA and relevant studies. Section 2.5 identifies gaps in the prior research to be addressed in the dissertation. Finally, Section 2.6 summarizes the chapter.

2.2 Linguistic Features in Writing

Linguistic features are fundamental in shaping written texts, particularly in academic contexts where clarity, fluency, and communicative effectiveness are critical. The analysis of these features provides an understanding of a writer's proficiency and the overall quality of the text. In academic writing, linguistic features are often categorized into four main aspects: complexity, accuracy, lexis, and fluency (CALF). This framework offers a comprehensive approach to evaluating writing performance, enabling researchers and educators to assess both surface-level accuracy and deeper structural sophistication (Biber et al., 2011; Ortega, 2003; Wolfe-Quintero et al., 1998).

While the traditional CALF model has provided perspectives in writing, especially regarding syntactic complexity, more recent approaches have extended the scope to include lexico-grammatical features. By analyzing lexico-grammatical features, researchers gain a deeper understanding of how linguistic elements influence the structure and communicative

effectiveness of written texts. This approach allows for more detailed assessments of writing proficiency and can inform targeted instructional practices aimed at developing specific linguistic competencies (Hyland, 2004).

Recent studies emphasized the importance of genre-specific linguistic features in academic writing, revealing how distinct academic genres, such as research articles, essays, and lab reports, follow structural and linguistic conventions (Swales, 1990; Hyland, 2005). For instance, academic texts in the sciences frequently use nouns and complex noun phrases to focus on processes and results (Biber & Gray, 2010). Conversely, texts in the humanities often employ verb-based constructions and 1st person pronouns to deliver stance and engage readers. Furthermore, linguistic feature analysis has practical implications for writing assessment and instruction. Crossley and McNamara (2012) demonstrated that features such as lexical sophistication and cohesion significantly correlated with essay scores on standardized tests. These findings showed the predictive power of linguistic features in evaluating writing quality as perceived by human raters.

Linguistic features across different contexts can be examined using MDA, which provides a broader perspective on language variation. MDA employs statistical analyses of text corpora to identify patterns of co-occurring linguistic features, interpreted as underlying dimensions of variation. This approach has been particularly effective in diverse contexts. For example, Nesi (2008) identified how student writing varies across disciplines and academic levels and showed significant genre-based linguistic differences. Similarly, Zhang (2023) compared L1 English speakers and L2 English learners' argumentative essays, indicating differences in linguistic

features like modality and personal pronoun use. These applications illustrate how MDA can capture the complex relationship between linguistic patterns and contexts.

In summary, linguistic features in writing encompass a diverse range of elements from word choices and grammatical structures to genre-specific conventions. The integration of measurement variables with corpus-based analysis provides a comprehensive understanding and evaluation of writing in academic contexts. This approach can contribute to the development of more effective writing instruction and assessment practices that address the complexities of language use in academic settings.

2.2.1 Syntactic Complexity

Syntactic complexity is often seen as a marker of writing proficiency and development. It refers to the sophistication and variety of sentence structures within a text. Numerous studies have examined how syntactic complexity can reflect a writer's ability to construct more complexed sentences, thereby indicating a higher level of writing proficiency (Norris & Ortega, 2009). It is commonly measured through various indices, such as the number of clauses per sentence, T-units, and dependant clauses, to evaluate the structural depth of a writer's composition. For example, Lu's (2010) syntactic complexity analyzer has become a widely used tool in this area, offering a way to measure aspects such as the average length of T-units and the frequency of subordinate clauses. These measures have been particularly useful in tracking linguistic development over time and helping researchers understand how students' syntactic complexity evolves as they gain proficiency.

However, while syntactic complexity is often associated with writing proficiency, it does not always directly correlate with writing quality or communicative effectiveness. Writers may produce structurally complex sentences that are grammatically correct but lack coherence, clarity, or appropriateness to the context, which leads to less effective overall compositions (Ortega, 2012). This suggests that while greater syntactic complexity may be indicative of advanced writing skills, it should be evaluated in conjunction with other linguistic features such as coherence, lexical richness, and appropriateness to the context (Biber et al., 2016).

Recent research has expanded the scope of syntactic complexity to include substructural elements like phrasal complexity, which focuses on the modification of noun phrases and the use of embedded structures (Biber et al., 2011; Lan et al., 2019). For example, Biber et al. (2011) and Kyle and Crossley (2018) explored how syntactic features differentiate levels of proficiency and found that proficient writers tended to use more complex noun phrases, higher densities of noun modifiers, and more elaborate syntactic constructions, which can contribute to the overall structural sophistication of their writing.

Despite the findings, syntactic complexity measures are limited in scope and may not capture the full range of linguistic features present in written texts. While these measures offer a detailed view of sentence structures and syntactic maturity, they often focus narrowly on grammatical structures without fully accounting for other aspects of language, such as lexical richness, discourse cohesion, and pragmatic appropriateness (Bulté & Housen, 2012; Norris & Ortega, 2009). For instance, syntactic complexity primarily emphasizes the arrangement and variety of clauses and sentence structures, but it does not sufficiently address how vocabulary usage (lexical complexity) contributes to meaning, how sentences are linked together to create

coherent discourse, or how language is adapted to suit different communicative purposes (Larsen-Freeman, 2006). Therefore, integrating these elements through the analysis of lexico-grammatical features, where vocabulary and grammar are examined together, could provide a more detailed understanding of language use and offer a broader perspective on how language operates within texts (Hyland, 2008).

2.2.2 Lexico-Grammatical Features

The study of lexico-grammatical features encompassing both vocabulary and grammar has gained increasing attention in research in academic writing (e.g., Biber & Gray, 2016; Hyland & Tse, 2012). Lexico-grammatical analysis examines how lexical choices and grammatical structures interact to fulfill communicative purposes and convey meanings in specific contexts. These features are important in academic writing as they contribute to the precision, clarity, and complexity necessary for scholarly communication. For instance, lexical richness, which includes the use of domain-specific vocabulary, allows students to express intended ideas and demonstrate expertise within a discipline. Similarly, grammatical structures such as nominalization, passivation, and complex sentences play a key role in achieving conciseness, formality, and logical organization.

Biber and Gray (2010, 2013) explored how lexico-grammatical features vary across different registers and genres, showing that there was an increasing use of compressed noun phrases and nominalizations in academic writing to achieve informational density and precision. Their analysis revealed that written academic texts, especially in formal contexts, show higher levels of syntactic complexity with frequent use of noun phrase modification, attributive

adjectives, and nominalizations to reflect the need for formal and dense language in such discourse. In contrast, spoken registers relied more on finite verbs and exhibited less syntactic complexity with coordination over subordination.

Extending this line of research, Staples and Reppen (2016) focused on the pedagogical implications of lexico-grammatical analysis by examining how L1 and L2 writers develop their academic writing skills, particularly in first-year university courses. They analyzed the use of phrasal and clausal complexity and found that while L1 English writers tended to demonstrate greater lexical diversity and more frequent use of implicit stance markers through phrasal complexity, L2 writers often had more repetitive lexical items and more overt stance markers like verb complement clauses (e.g., *think that, believe that*).

The same study also showed significant differences between genres, for example, rhetorical analysis essays involved fewer premodifying nouns than long argument essays, which required greater informational density and complexity. Moreover, L1 background also impacted the use of lexico-grammatical features, with L1 Chinese and L1 Arabic writers using more noun premodifiers than L1 English writers but showing a narrower range of lexical choices overall. These findings suggest that academic writing development can be influenced by disciplinary conventions, genre conventions, and specific communicative purposes (Hyland, 2004; Swales, 1990).

Understanding how linguistic features function in written texts provides important perspectives into academic writing. Lexico-grammatical analysis allows researchers to capture the interplay between vocabulary and grammar with a deeper understanding of how language

is used to construct meaning to meet communicative purposes and disciplinary conventions. However, a deeper exploration is needed to fully understand how language use adapts to different contexts, genres, and audiences. Integrating various aspects of language can contribute to a wider perspective of writing.

2.3 Linguistic Variation

Linguistic variation refers to how language use differs across various contexts, affected by variables such as genre, register, and situational characteristics. From a text-linguistic perspective, these variations are important to understanding how language functions in specific settings (Gray & Biber, 2012; Egbert et al., 2024). In addition, the interplay of genre, academic level, and discipline significantly influences linguistic patterns in student writing (Gardner et al., 2019). For example, disciplinary conventions shape the frequency and type of lexical bundles used in texts, which emerge empirically rather than being predefined (Durrant, 2015). The findings from Gardner et al. (2019) and Durrant (2015) show the importance of understanding how disciplinary conventions influence linguistic strategies. This section explores the role of linguistic variation in shaping written texts, especially in relation to genres, registers, and communicative purposes.

2.3.1 Genre and Register

Genres and registers reflect linguistic variation based on distinct ways that language is structured and used. Genres are generally defined by the conventions and expectations associated with specific text types, such as essays, lab reports, or research papers, which are shaped by their communicative purposes and rhetorical structures (Biber & Conrad, 2019; Nesi

& Gardner, 2012). For example, scientific reports tended to emphasize objectivity by using passive voice constructions to focus on processes rather than the researchers, whereas argumentative essays in the humanities may employ more direct language to engage readers and present claims.

Recent studies have shown that genres in student writing exhibit substantial internal variation. For example, Goulart et al. (2022) found that even within a single genre, student texts often combine multiple communicative purposes such as argumentation, explanation, and procedural recounts based on the specific assignment and discipline. This suggests that genres vary in response to the distinct demands of each discipline. Hyland (2022) supports this perspective that genre-based teaching approaches help students understand the conventions and expectations across different contexts to enhance their ability to adapt language use to meet these demands.

In contrast, register variation is defined by situational characteristics such as audience, setting, and communicative purpose. Registers influence linguistic choices, determining how formal, technical, or interactive the language is. Academic writing, for instance, tends to feature complex syntax, precise vocabulary, and a formal tone, while conversational language tends to be more informal and interactive (Biber, 1988; Egbert et al., 2024). According to Egbert et al. (2024), registers should not be seen as rigid categories, but as dynamic constructs shaped by the interplay of linguistic features and situational factors, which reflect functional goals that adapt to contextual demands. For example, Hyland (2004) noted academic genres in the sciences often use nominalizations, passive structures, and long noun phrases to convey

precision, whereas genres in the humanities may employ more varied forms and engage the reader more directly to fulfill interpretive or persuasive purposes (Hyland, 2000).

Additionally, Lancaster (2016) examined how disciplinary context influences the use of stance markers, such as hedges, boosters, and attitude markers, in student writing across the social and natural sciences. The study conducted a focused analysis of these specific linguistic features to explore how students convey certainty, caution, or attitude in their texts. The findings showed that social science texts used more hedges (e.g., might, could) and boosters (e.g., clearly, undoubtedly) to present arguments and to critically engage with literature. In contrast, natural science texts exhibited a more assertive style with fewer hedges, consistent with disciplinary norms favoring confident reporting of empirical results. These findings further demonstrate how genre and register choices are closely tied to the rhetorical expectations of each discipline.

2.3.2 Analytical Approaches for Exploring Linguistic Variation

The study of linguistic variation within genres, registers, and communicative purposes relies on diverse analytical tools, each offering valuable perspectives on the complexity of language use. These tools vary in their scope and focus, from traditional descriptive methods to more advanced approaches such as KF analysis and MDA.

Traditional methods, such as type/token ratio and measures of syntactic complexity, have long been used to examine individual linguistic features. For example, Staples et al. (2016) analyzed student writing using the British Academic Written English (BAWE) corpus to investigate grammatical complexity. They found that as students progressed academically, their

writing demonstrated increased phrasal complexity (e.g., nominalizations, premodifying nouns) and reduced clausal complexity (e.g., finite dependent clauses). These findings illustrate how disciplinary conventions, and communicative purposes can shape language use.

Building on traditional approaches, KF analysis, introduced by Egbert and Biber (2023), identifies distinguishing lexico-grammatical features across texts within different communicative purposes. This approach uses effect sizes (i.e., Cohen's *d*) to pinpoint features that are not only frequent but also contextually significant. For instance, Kim et al. (2025) applied KF analysis to student writing, showing that narrative essays were characterized by frequent use of 1st person pronouns, activity verbs, and mental verbs based on their dynamic and personal communicative purpose. In contrast, descriptive essays showed higher prevalence of attributive adjectives, concrete nouns, and indefinite articles, emphasizing specificity and tangibility. KF analysis complements traditional methods by focusing on contextually salient features and offer meaningful views on how linguistic variation aligns with functional demands.

While KF analysis excels at identifying isolated distinguishing features, it is limited in capturing relationships among co-occurring linguistic elements. To address this, researchers have been using MDA, a more comprehensive framework that examines clusters of linguistic features as part of co-occurring sets. Biber's (1988) foundational work on MDA, based on a wide range of spoken and written registers, revealed dimensions of variation, such as the distinction between involved and informational production or narrative and non-narrative discourse. These dimensions provide a framework for understanding how linguistic features goes along with the functional and situational characteristics of texts.

MDA has been particularly effective in exploring linguistic variation across contexts and disciplines. For example, Gray and Biber (2012) focused on professional academic writing in the natural and social sciences, demonstrating that scientific writing prioritized informational density and objectivity, reducing personal stance markers. In contrast, social sciences texts employed interactional features (e.g., hedges and boosters) to engage readers and emphasize the author's arguments. Goulart et al. (2020), analyzing student writing, further explored the interplay between register and genre, showing how situational factors like audience and communicative purpose influence linguistic choices across different contexts. These studies indicate MDA's strength in capturing the multifaceted nature of linguistic variation.

2.4 Multidimensional Analysis and Linguistic Variation

In the field of text linguistics, understanding linguistic variation across different genres and registers requires comprehensive analytical approaches that move beyond the examination of individual features. One such method is MDA, introduced by Biber (1988), offering a more holistic view of how language varies based on situational characteristics such as communicative purposes, settings, and audiences.

2.4.1 Overview of MDA

MDA is a corpus-based method that extends traditional lexico-grammatical analysis by focusing on how linguistic features function together in texts. Rather than having specific elements such as sentence length or word frequency, MDA examines groups of linguistic features that frequently co-occur, known as dimensions. These dimensions are identified

through factor analysis, which groups together linguistic features that appear in similar contexts to reveal functional patterns across texts (Biber, 1988).

Each dimension within the MDA typically reflects a contrast between linguistic features associated with different communicative functions. For example, in Dimension 1 in Biber's framework (see Table 1) contrasts involved production (e.g., frequent use of 1st person pronouns, contractions, and present tense verbs) with informational production (e.g., high use of nouns, prepositions, and attributive adjectives). Texts with high scores on the informational end of the spectrum such as academic prose, which has positive loadings, tended to prioritize factual and precise information over personal engagement, whereas texts on the involved end, like conversational speech, which has negative loadings, were more interactive and relational.

Table 1. Dimension Description (Biber, 1988)

Dimension 1: Informational vs. Involved		
Feature	Example	Factor loading
Positive features (informational)		
Nouns	case, community	0.80
Prepositions	in, for, of	0.54
Attributive adjectives		0.47
…		
Negative features (involved)		
Private verbs	believe, think, know	-0.96
that-deletions	I think ϕ he went	-0.91
Contractions	can't, don't	-0.90
Present tense verbs	is, wants, likes	-0.86

Biber (1992) expanded on the principles of MDA by focusing on its foundation in the co-occurrence of linguistic features, which reflect shared communicative functions. MDA dimensions are conceived as continuous scales, enabling analysis of linguistic variation across

registers. For instance, Dimension 2, which distinguishes narrative from non-narrative discourse, identifies narrative texts by features such as past tense verbs and 3rd person pronouns, while non-narrative texts use present tense verbs and elaborative structures to support exposition and argumentation. Similarly, Dimension 3 describe the contrast between elaborated references, often found in academic prose, and situation-dependent references typical of conversational speech. These dimensions collectively enable a comprehensive analysis of how language adapts to differing communicative purposes and genres, which offer perspectives on the interaction between linguistic features and their functional contexts.

To elaborate further on the Biber's dimensions, Dimension 1 (Involved vs. Informational Production) distinguishes between texts that are interactive and those that are more informational. Texts with positive loadings on this dimension are typical of conversational speech that contain features such as 1st person pronouns, present tense verbs, and contractions. These features reflect the involved nature of face-to-face conversations, which score the highest within this dimension (*M* = 35.3). Prepared speeches also align with involved production even though to a lesser degree (*M* = 2.2), as they retain some interactive elements despite being scripted. In contrast, texts with negative loadings, usually common in academic writing, include features such as nouns, long words, and prepositional phrases, which indicate a more informational focus. For example, academic prose (*M* = -14.9) and official documents (*M* = -18.1) are positioned towards the informational end of the range showing their focus on detailed and factual content rather than interpersonal communication (see Figure 3).

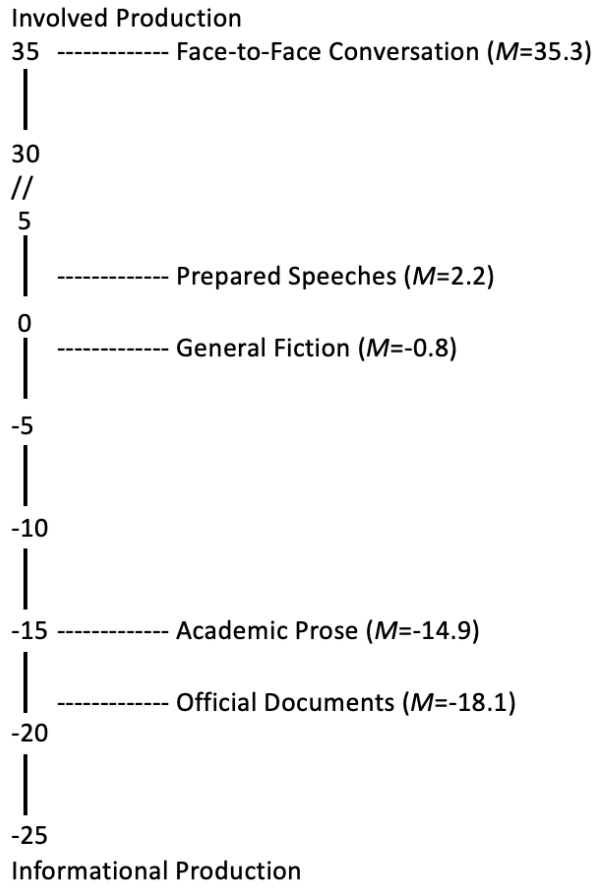

Figure 3. Distribution of the registers in Biber's (1988) Dimension 1

Dimension 2 (Narrative vs. Non-Narrative Discourse) captures the degree to which texts are narrative, often marked by past tense verbs and 3rd person pronouns found in stories and personal anecdotes. Texts with positive loadings in this dimension are typically narrative, while non-narrative texts that focus on exposition and argumentation tend to have negative loadings. Dimension 3 (Elaborated vs. Situated reference) differentiates texts based on the clarity and specificity of references. Texts with positive loadings use explicit reference which involves full noun phrases and other linguistic markers that clarify meaning common in academic essays. In contrast, situation-dependent reference is based on context and shared knowledge, as seen in

spoken conversations that often have negative loadings. Texts vary in their use of persuasive language within Dimension 4 (Overt Expression of Persuasion), having features such as modals, infinitives, and adverbials that indicate necessity, possibility, and persuasion. Texts with positive loadings like academic and argumentative essays score high on this dimension due to their frequent use of these persuasive elements. On the other hand, texts with negative loadings may be more descriptive or neutral without overt persuasive intent. Lastly, Dimension 5 (Abstract vs. Non-Abstract Style) contrasts texts that convey abstract, theoretical information with those that provide concrete, specific details. Texts with positive loadings on this dimension are characterized by the use of abstract nouns and complex sentences that are common in scientific and technical writing, whereas texts focused on concrete and specific details like instructional manuals might show negative loadings.

2.4.2 Application of MDA to Linguistic Variation

MDA can follow two primary methodological approaches: the use of predefined dimensions (e.g., Biber, 1988) and a data-driven approach in which new dimensions are extracted through factor analysis of a dataset. The predefined approach allows for direct comparison with Biber's findings on register variation in English and is often used when researchers wish to position their texts relative to a known linguistic baseline (e.g., Crossley et al., 2014; Friginal & Weigle, 2014). In contrast, the data-driven approach is typically adopted in studies aiming to reveal context-specific dimensions, such as discipline-based or learner-specific registers (e.g., Gardner et al., 2018; Tasker, 2019; Jin, 2018). These studies extract new dimensions from co-occurrence patterns in the dataset itself, allowing for greater sensitivity to the communicative functions relevant to the corpus at hand. While the predefined approach

offers comparability and generalizability, the data-driven method yields greater ecological validity and interpretability, especially when studying student writing, emerging genres, or multilingual corpora.

Building on Biber's initial framework, subsequent studies have adopted MDA depending on their research goals and the nature of their corpora. Research employing MDA has demonstrated its effectiveness in distinguishing between different genres, registers, and disciplines by identifying patterns of co-occurring linguistic features that reflect specific communicative purposes. This approach revealed how texts produced in various contexts differ significantly in terms of syntactic structures, vocabulary choices, and overall linguistic strategies, with a robust framework for analyzing linguistic variation.

One significant application of MDA has been the analysis of disciplinary writing. Research comparing the linguistic features of texts from the hard sciences (e.g., physics, biology) with those from the social sciences and humanities has revealed notable differences in how language is used to achieve different communicative purposes. For example, Gray and Biber (2012) found that science texts tended to be more informationally dense, using longer noun phrases, nominalizations, and attributive adjectives to pack information into fewer words. In contrast, humanities texts often included more interactional features such as 1st person pronouns, stance markers, and modal verbs to engage with readers and present arguments persuasively. Adding to this body of research, Jin (2018) also used MDA to explore linguistic variation in discussion sections of chemical engineering research articles, focusing on how high-impact and low-impact articles differ in their linguistic features. The author identified six dimensions of variation, including "involvement and interactivity" and "informational density,"

which revealed that high-impact articles used more evaluative and interactive linguistic features. These findings indicate the role of disciplinary conventions and communicative purposes in shaping linguistic choices in such fields.

Hardy and Friginal (2016) applied a data-driven MDA to student writing using the Michigan Corpus of Upper-Level Student Papers (MICUSP), identifying four dimensions that reflected variation across genres such as research papers, reports, and creative writing. Their analysis showed a continuum from personal and narrative discourse to highly technical and informational styles. Similarly, Hardy and Römer (2013) also conducted a data-driven MDA of successful upper-level student writing across 16 academic disciplines in the MICUSP. They identified four emergent dimensions including the contrast between involved, academic narrative versus descriptive, informational discourse, and the expression of opinions and mental processes. Their findings revealed disciplinary distinctions, for example, Humanities disciplines like Philosophy and Education relied more on involved and narrative discourse with frequent use of personal pronouns, stance verbs, and mental process verbs, while disciplines in the Biological and Physical Sciences had more descriptive and informational features such as nominalizations, technical language, and dense noun phrases. These dimensions showed how different disciplines emphasize varying linguistic strategies depending on their communicative purposes and provided further evidence of the role disciplinary variation plays in academic writing.

The research by Gardner et al. (2018) further advanced MDA by analyzing the British Academic Written English (BAWE) corpus through four new dimensions that incorporated situational variables such as disciplines, level, and genre. Their study revealed that student writing becomes progressively more informational and less involved over time, with disciplines

like Physical Sciences showing greater procedural density and technicality, while Humanities maintained narrative and stance-focused elements. These dimensions, ranging from Compressed Procedural Information to Informational Density, have informed EAP pedagogy by indicating which features might form a common academic core and which are discipline specific.

The shift toward examining register variation in student writing, Larsson et al. (2021) analyzed both L1 and L2 students' argumentative essays and research papers. Their MDA analysis revealed that registers was the most influential factor shaping linguistic features such as stance, involvement, and factual density, with significant variations observed between genres. Complementary to this, Moran's (2013) interdisciplinary study on undergraduate writing in Psychology and Chemistry combined corpus analysis with interviews and writing task classifications. Her findings emphasized the discrepancy between instructors' expectations and students' understanding of disciplinary writing, as well as key linguistic differences across registers that the greater presence of procedural, technical language in Chemistry vs. more interpretive and human-centered language in Psychology.

Tasker (2019) contributed a department-level perspective by examining register variation within undergraduate English coursework. Using a bottom-up MDA alongside complexity analysis, the study identified dimensions of variation derived from the corpus itself. Findings showed that differences in student writing were shaped more by task type and subdisciplinary focus (e.g., literature analysis vs. rhetoric) than by course level. This complicates traditional assumptions that student writing becomes uniformly more technical or abstract over time, which emphasize the curricular and situational context in shaping linguistic performance. Extending this inquiry beyond single-department or national contexts, Ansarifar et al. (2025)

explored linguistic variation in thesis abstracts written by MA and PhD students from three different L1 backgrounds: Persian, Chinese, and English. Using MDA, they identified four dimensions of linguistic variation such as extended procedural discourse vs. descriptive informational discourse and human-focused informational density. Their study revealed that MA and PhD students differed significantly on certain dimensions, particularly in the extent to which they used persuasive and human-focused features. Furthermore, the analysis showed notable differences across L1 backgrounds, with Persian and English abstracts aligning more closely compared to Chinese abstracts. These findings suggest that variation in academic writing is shaped by an interplay of disciplinary norms, educational levels, and linguistic-cultural backgrounds, offering critical perspectives into how writers navigate academic discourse globally.

Overall, the application of MDA to linguistic variation provides a wider perspective on how linguistic features co-occur to fulfill specific communicative purposes across genres, registers, and disciplines. Identifying dimensions of variation has demonstrated how academic writing varies based on disciplinary conventions, academic levels, and language backgrounds. Notably, both predefined approaches (e.g., using Biber's dimensions) and bottom-up (or data-driven) analyses have yielded converging insights, emphasizing consistent patterns such as the contrast between involved vs. informational discourse, or narrative vs. technical styles. These complementary findings reinforce the value of MDA in capturing functional linguistic variation in academic writing. Consequently, this understanding can inform the development of targeted writing instruction that addresses the specific linguistic and rhetorical requirements of various disciplines, ultimately supporting students in developing the proficiency needed to succeed in

their academic and further pursuits by helping them to navigate and adapt to the diverse linguistic expectations of their fields.

2.5 Addressing Confounds in Academic Writing Research

A recurring challenge in the analysis of academic writing is the presence of confounding variables that complicate the interpretation of linguistic variation. Studies investigating register or disciplinary differences often grapple with overlapping influences such as genres, disciplinary norms, task design, and assessment context (Conrad, 2013; Nesi & Gardner, 2012; Hardy & Römer, 2013). For instance, a Business case study and an Engineering proposal may differ not only in discipline but also in genre conventions, rhetorical purpose, and even task format. Similarly, writing produced under times exam conditions may exhibit different features, such as reduced lexical diversity or syntactic complexity, compared to take-home assignments or portfolio work (Friginal & Weigle, 2014; Crossley et al., 2014). The nature of task prompts, whether open-ended or highly structured, also significantly influences student responses, particularly in L2 contexts (Larsson et al., 2021). These overlapping variables make it difficult to isolate the influence of discipline or genre alone, calling for cautious interpretation and clear methodological framing. This study acknowledges these complexities and interprets linguistic variation not as the result of a single factor, but as the product of an interaction among disciplinary expectations, communicative purposes, and task design.

2.6 Research Gaps and Limitations

Despite extensive research on linguistic variation in academic writing, several gaps need further investigation. One notable gap lies in the limited focus to date on disciplines such as

Business and Engineering. While much of the existing research has concentrated on academic genres in fields like Humanities and Sciences (e.g., Hyland, 2004; Biber & Gray, 2016), there is less understanding of how linguistic features function in contexts with distinct communicative purposes. Business and Engineering, in particular, require precise and clear communication to effectively convey complex information and technical concepts. In Business courses, writing tasks often require analytical and persuasive skills that reflect real-world professional scenarios (Zhu, 2004). These tasks emphasize practical applications, preparing students for workplace communication. Similarly, Conrad and Newberry (2011) argued the importance of bridging the gap between academic training and industry expectations by addressing communication skills explicitly in the curriculum. However, how genres and communicative purposes influence lexico-grammatical features in these fields has not been thoroughly explored.

Additionally, although MDA has been successfully applied to distinguish between disciplines and genres (Biber, 1988; Hardy & Römer, 2013), few studies have investigated how specific genres of proposals, and case studies within Business and Engineering adapt linguistic strategies to meet their specific communicative purposes. This represents a gap in understanding the intra-disciplinary variations and the role of genre-specific conventions in shaping linguistic choices. Another gap is the lack of cross-disciplinary comparisons in the study of lexico-grammatical features across disciplines. While there have been many studies on linguistic variation within individual disciplines, there is a need for more comparative studies that examine how these features function across different professional domains (Hyland, 2012). Such comparisons can reveal unique and shared linguistic practices to enhance our understanding of disciplinary writing and inform targeted writing instruction.

Moreover, the existing literature often overlooks the impact of genre and communicative purpose on linguistic variation in student writing. Given that genres impact the organization of content and language use, exploring how genre conventions influence lexico-grammatical choices in Business and Engineering can provide a better understanding of disciplinary writing practices. Therefore, addressing these gaps, this dissertation aims to apply MDA to both Business and Engineering texts, specifically focusing on student writing across different genres. By examining how lexico-grammatical features reflect the communicative purposes and expectations of each genre and discipline, this research seeks to contribute to a deeper understanding of linguistic variation in disciplinary student writing. This study also contributes to the debate by examining how discipline-specific writing instruction can benefit from understanding shared and unique linguistic features in Business and Engineering writing (Zhu, 2004; Goulart, 2021). The findings aim to inform targeted pedagogical strategies that bridge general academic skills and discipline-specific demands.

2.7 Chapter Summary

This chapter has provided a comprehensive overview of research on register variation, linguistic variation, and the use of MDA to explore how linguistic features vary across different contexts. Studies reviewed in this chapter revealed that genres, registers, disciplines, and communicative purposes often shape linguistic features in academic writing. For example, procedural registers such as lab reports and research articles in the natural sciences tended to emphasize informational density by using premodifying nouns and nominalizations. In contrast, texts like essays and critiques in the humanities relied more on interactional features to engage readers and present arguments persuasively.

This chapter also highlighted the role of MDA as an analytical framework for identifying patterns of linguistic variation across genres, registers, and disciplines. Research employing MDA has shown significant distinctions between disciplines; for example, the natural sciences tended to prioritize informational content and showed reduced personal stances, while the humanities often featured more involved and narrative discourse. Additionally, studies examining genre variation have illustrated how specific genres adapt linguistic features to meet their communicative purposes, whether by providing detailed procedural or presenting arguments in a more interactive format.

Despite these findings, further exploration of linguistic variation is necessary to gain a deeper understanding of how language functions across different contexts, particularly within specialized disciplines such as Business and Engineering. The existing studies left a gap in knowledge regarding how lexico-grammatical features can be employed in particular fields to meet specific communicative purposes. Addressing this gap, the following chapter will outline the corpus and methods used to apply MDA to a comparative analysis of Business and Engineering texts. This analysis aims to elucidate how these fields use lexico-grammatical features to fulfill their communicative purposes and stick to disciplinary conventions.

Chapter 3. Method

3.1 Introduction

This chapter outlines the methodological framework used in this dissertation to explore the linguistic variation in Business and Engineering student writing. The chapter begins by presenting the research questions guiding this study and then provides a detailed description of the corpus used for analysis, including the selection criteria for the texts and their relevance to the research goals. The chapter further explains the application of MDA, an analytical framework used to identify patterns of linguistic variation across different contexts. In addition, this section discusses how MDA has been employed to examine the interaction between linguistic features and communicative purposes, focusing on the ways that Business and Engineering text adapt their linguistic choices to meet the specific demands of their disciplines.

Additionally, the chapter provides an overview of the annotation process, which includes tagging specific linguistic features in the corpus for further analysis. The methods for evaluating the annotations are discussed to ensure that the findings are both reliable and valid. By providing these details, the chapter offers the foundation for the comparative analysis that follows, which will describe distinctive linguistic variations of Business and Engineering student writing and contribute to a deeper understanding of how language functions across different disciplines.

3.2 Research Questions

To address the need for a detailed description of linguistic variation across different text types and disciplines, this dissertation focuses on exploring the underlying dimensions and

associated features in student writing across Business case studies and Engineering proposals. Specifically, the study investigates how linguistic features manifest across these text types and disciplines, revealing patterns of variation shaped by their respective communicative purpose and field.

The first research objective is to identify the underlying dimensions that characterize the linguistic features of Business case studies and Engineering proposals. This involves using MDA to examine clusters of co-occurring linguistic features within the corpus and interpret these clusters as dimensions that reflect the functional demands and contextual requirements of the texts. The second objective is to compare the dimensional differences between Business and Engineering student writing. This comparative analysis shows how shared and distinct linguistic features are employed differently across disciplines to meet their specific demands.

Building on these objectives, this dissertation employs MDA as the primary methodological framework to analyze linguistic variation across text types and disciplines. The following research questions are designed to guide this investigation:

1. What underlying dimensions and associated features emerge in student writing across Business case studies and Engineering proposals?
2. What are the dimensional differences between Business and Engineering student writing?

By addressing these questions, the study seeks to contribute to a deeper understanding of disciplinary writing practices and provide practical perspectives on developing targeted pedagogical resources to enhance students' writing.

3.3 Corpus Design and Data Collection

The corpus for this dissertation was collected from two different disciplines, namely Business and Engineering, comprising texts written by undergraduate students at an English-medium university in Montreal. It includes a total of 77 texts from the Business department and 137 texts from the Engineering department. These texts were written as either part of students' assignments or exams.

Texts from the Business department (McDonough et al., 2021; Neumann et al., 2019) were part of the final examination of a business communications course at a university. This required course taught students core business theories and critical analysis of business texts and case studies. For the final examination, students were asked to read four case studies and respond to seven given prompts within three hours. The last case study was selected for this analysis due to its requirement for a longer response, examining the influence of Jeff Bezos' leadership style on Amazon's success, and referencing leadership framework in business. Similarly, the Engineering texts were part of students' coursework. This mandatory technical writing and communication course aimed to instruct students in composing technical and scientific papers, including abstracts and reports, and in communicating technical information through oral presentations. The selected assignment for this dissertation was a proposal significant for categorizing and classifying sources based on individual contributions to comprehending the problem addressed in their final project. This task required each student to review prior research to establish context and guidance for their subsequent work, aiming to understand the issue better and identify gaps in current knowledge or resolve conflicting findings. Table 2 provides a detailed description of the texts from each discipline. The analysis

process involved a systematic examination of the linguistic features present in the texts, focusing on identifying patterns and variations across the two disciplines using dimensions.

Table 2. Corpus of Study

	Number of texts	Number of words	Mean text length	SD
Business	77	36222	470.41	100.93
Engineering	137	129344	944.11	269.90
Total	214	165566	773.67	

3.4 Data Coding

All texts were grammatically annotated using the Biber Tagger (Biber, 1988), which assigns labels for lexico-grammatical features to individual words within the texts (see Appendix A). It employs a comprehensive set of tags to annotate various linguistic features across multiple categories. These categories include lexical items (e.g., common nouns, pronouns, main verbs, auxiliary verbs, attributive adjectives, and adverbs (manner, time, degree)), grammatical structures (e.g., tense, aspect), discourse features (e.g., stance markers, hedges, boosters, connectives), and syntactic complexity features (e.g., subordination, nominalization, coordination). For instance, the sentence *Resolution determines the level of detail and clarity in the virtual environment* was tagged as below:

Resolution ^nn+nom+++=Resolution
determines ^vbz++++=determines
the ^ati++++=the
level ^nn++++=level
of ^in++++=of
detail ^nn++++=detail

and ^cc++++=and

clarity ^nn+nom+++=clarity

in ^in++++=in

the ^ati++++=the

virtual ^jj+atrb+++=virtual

environment ^nn+nom+++=environment

The automated tagging process enables the extraction of a larger and detailed set of linguistic data from the corpus than manual annotations. However, certain issues may arise with tags, such as *that*, which can function as a determiner, relative clause, or complement clause (Gary, 2019). To address these potential inaccuracies, all annotations were thoroughly reviewed to ensure the accuracy of the final data. Given the manageable size of the corpus for this study, a manual review was conducted for all texts. In instances where inaccuracies were identified, necessary corrections were applied to enhance the reliability of the annotations. For example, several errors were found in the sentence *It becomes clear that AI machine learning is essential for improving the usability*. Specifically, *AI* was incorrectly tagged as a main verb but was corrected to a proper noun; *learning* was tagged as a postnominal modifier but was corrected to a common noun; and *essential* was identified as a noun but was corrected to an adjective (see Table 3).

Table 3. Example of tag correction

Word	Initial annotation	Corrected annotation
It	3rd person pronoun	
becomes	Verb	
clear	Adjective	
that	Dependent clause	
AI	Main verb	Proper noun
machine	Common noun	

learning	Postnominal modifier	Common noun
is	Main verb	
essential	Common noun	Adjective
for	Preposition	
improving	Postnominal modifier	
the	Article	
usability	nominalization	

3.5 Data Analysis

The linguistic features analyzed in this study were selected based on their established relevance in prior MDA research, such as those featured in Biber's (1988) foundational tagset and subsequent extensions (e.g., Friginal & Weigle, 2014; Gardner et al., 2018; Goulart, 2021). These features represent a broad range of grammatical and lexico-grammatical categories, including nouns, verb types, modals, pronouns, determiners, stance markers, and indicators of syntactic and informational complexity. Feature selection was guided by both theoretical considerations and empirical suitability. To make sure meaningful variation across the corpus, features were evaluated for frequency and distribution, with low-frequency or invariant items excluded to prevent distortion in factor analysis. Statistical checks were applied to assess the appropriateness of variables that items with multicollinearity ($r > .80$) or low commonality scores ($h^2 < .18$) were removed, as they contribute minimal shared variance or introduce instability into factor extraction. After this filtering process, 93 features were retained for analysis. This final set reflects both theoretical significance and empirical robustness, providing a reliable foundation for identifying functionally interpretable dimensions in this study.

Building on this feature set, the analysis followed a systemic approach to ensure comparability and interpretability across texts. Following annotation, the frequency of each

feature was calculated and normalized per 1,000 words to allow comparison across texts. This normalization process aligns with Biber's methodology to make sure that features contribute equally to the analysis. Factor analysis was then employed to identify dimensions of variation, with features grouped based on their co-occurrence patterns. Studies using the MDA approach can either apply predefined dimensions from Biber's framework or others (e.g., Biber & Conrad, 2019; Hardy & Römer, 2013) or perform factor analysis on the linguistic features within their dataset to identify and derive their dimensions (e.g., Weigle & Friginal, 2015). In this study, dimensions were derived using factor analysis to make sure that the identified dimensions were grounded in the linguistic patterns within the analyzed corpus.

This systematic approach makes the feature scores in standard deviation units, giving all features equivalent weights when computing the dimension scores (Biber, 1988). Then, dimension scores (or factor scores) were calculated for each text by adding up the scores of features with prominent loadings on that dimension. Only features with loadings greater than 0.30 on a factor were considered meaningful enough to be included in the dimension scores. The threshold was set at 0.30, slightly lower than the standard 0.35 initially used by Biber (1988), to capture a broader range of linguistic features that may still contribute significantly to the variation within the corpus. By using a slightly less conservative threshold, to the analysis accounts for subtle but potentially relevant co-occurrence patterns that may otherwise be excluded, which is particularly useful with smaller or more specialized datasets. Additionally, the mean dimension scores were calculated for each discipline. These scores characterize the given disciplines, enable comparisons between the two disciplines, and provide a detailed interpretation of the underlying dimensions.

By using MDA, the prevalence of specific linguistic features such as nominalizations (e.g., the frequent use of abstract nouns), hedges (e.g., might, could), and other rhetorical strategies can be identified to compare language use across two chosen academic disciplines and genres. For instance, business-related texts might emphasize more dynamic language such as hedges and modal verbs to suggest possibilities or recommendations (Swales, 1990; Hyland, 2005). In contrast, in disciplines like engineering, writers may frequently use technical terms, nouns, and nominalized structures that reflect a more abstract and objective discourse (Hyland, 2000; Biber et al., 2007). Additionally, to address sub-questions, a focus on grammatical structures such as passive and active voice can reveal how each discipline approaches language. Business texts are likely to prefer the active voice to attribute actions directly to agents to enhance clarity and persuasiveness (Thompson & Thetela, 1995), while engineering texts might have passive constructions to emphasize processes or results to keep an impersonal and precise tone.

Moreover, discipline-specific patterns can be explored to understand how these features are distributed. For example, business communication may employ more concrete language when discussing real-world applications to involve stakeholders and convey practical solutions, whereas engineering texts often lean toward abstract descriptions to explain theoretical models or designs to reflect their emphasis on precision and technicality (Hyland, 2006). Finally, connecting these linguistic choices to the communicative purposes of each discipline can provide further understanding. In business, the need to persuade or recommend actions may keep more interactive language (Charles, 2006), while in engineering, the focus on precision and reporting technical information may result in a more static and information-dense linguistic style (Bhatia, 1993).

Applying MDA can identify key dimensions of linguistic variation within a corpus of texts from both disciplines. For example, the dimensions of involved versus informational production (Dimension 1) and narrative versus non-narrative (Dimension 2) might reveal substantial differences between business and engineering texts. Dimension 1 could show that business texts are more involved with features like personal pronouns and present tense verbs to create a sense of interaction (Biber, 1988) while engineering texts might score higher on the informational side with more complex noun phrases and prepositional phrases to focus on the content (Biber et al., 2007). In Dimension 2, narrative elements may be less dominant in both fields compared to disciplines such as Humanities. However, business writing could still maintain some narrative elements (e.g., case studies), whereas engineering texts might focus more on reporting objective data (Hyland, 2000). Dimension 3 examines the clarity and specificity of reference and can further explain the differences between the two disciplines. Business texts might rely more on context and shared knowledge to convey meaning, while engineering texts may feature more explicit references and highly specific terminology due to the technical nature of the field (Hyland, 2014). Lastly, other dimensions such as persuasive language or the balance between abstract and concrete information may also significantly differentiate business and engineering texts. Business writings are likely to employ more modals and stance markers to influence decision-making (Hyland, 2005), whereas engineering texts may be more detached and objective to emphasize factual and abstract information (Swales, 1990).

To answer the first research question, MDA was applied to the corpus of student texts to identify the underlying dimensions of linguistic variation such as informational density, stance, or narrative focus. This was achieved by analyzing a range of lexico-grammatical features (e.g.,

nominalizations, premodifying nouns, adjectives, adverbs) across the texts. These features were extracted and statistically analyzed to reveal the primary dimensions that characterize linguistic strategies in student writing across Business and Engineering disciplines. The entire corpus was analyzed together to identify overarching linguistic dimensions using the MDA framework for the first research question. This methodological choice aligns with prior MDA studies, which often analyze diverse text types collectively to discover shared and distinctive patterns of linguistic variation. Combining texts from both disciplines enables the identification of dimensions that broadly apply to student writing and provides a baseline for subsequent discipline-specific comparisons. This holistic approach situates the findings within broader discussions of linguistic variation in student writing while also facilitating a deeper exploration of the unique features specific to Business and Engineering writing in the second research question.

To explore the dimensional differences between Business and Engineering student writing for the second research question, the results were compared to identify dimensions that are more prominent in each discipline and analyze how the associated linguistic features differ. For example, Business writing may exhibit a stronger emphasis on persuasive stance markers and informational compression to support decision-making and stakeholder engagement, while Engineering writing may prioritize technical descriptions and procedural discourse to show precision and clarity. This comparison will represent how each discipline adapts its linguistic strategies to meet distinct communicative purposes, revealing both shared features and discipline-specific differences.

The comparative analysis for the second research question also addresses the communicative purposes underlying these linguistic differences. By examining how the identified dimensions align with the specific purpose of each text type and discipline, the study reveals the functional motivations behind the observed patterns. For instance, Business writing often prioritizes persuasion, decision-making, and situational engagement, which necessitate evaluative language and the use of linguistic features such as modals, stance markers, and temporal features. In contrast, Engineering writing focuses on factual accuracy, technical explanations, and procedural clarity, favoring features like nominalizations, dense noun phrases, and precise descriptions. Additionally, genre-specific conventions within each discipline will be analyzed to understand how they create the linguistic patterns. By connecting the dimensions to their respective communicative purposes, this analysis provides a deeper understanding of how Business and Engineering writing employs language to fulfill the demands of their disciplinary contexts.

3.6 Chapter Summary

This chapter outlined the methodology used to investigate linguistic variation across different disciplines and genres using MDA. The chapter began by stating the research questions that guide this dissertation regarding how linguistic features differentiate Business and Engineering texts and how MDA can reveal underlying dimensions of variation between these disciplines. The corpus design was described in detail with an explanation of the selection of student texts from Business and Engineering courses and each contributes to distinct genres: Business case studies and Engineering proposal.

Then the process of data annotation was explained by using Biber Tagger for grammatical annotation that categorizes lexical, grammatical, and syntactic complexity features across the texts. The chapter also addressed the steps taken to ensure the accuracy and reliability of the annotations. The data analysis involved calculating the frequency of linguistic features and standardizing these across texts with MDA identifying the significant dimensions that differentiate two disciplines. Finally, expected key findings were included which reflected the communicative purposes of each discipline and genre. The application of MDA is expected to reveal that business texts are more involved with more interactive language and narrative elements while engineering texts are more informational based on explicit references and technical terminology. The findings will be discussed in the following chapter.

Chapter 4. Results

4.1 Introduction

This chapter presents the findings of the study by addressing the two research questions: (1) what underlying dimensions and associated features emerge in student writing across Business case studies and Engineering proposals? and (2) what are the dimensional differences between Business and Engineering student writing? MDA identifies co-occurring linguistic features and groups them into underlying dimensions of variation. These dimensions reflect functional and rhetorical differences in the texts to reveal how linguistic features distinguish Business case studies and Engineering proposals. The findings further show the distinctive linguistic characteristics of each text type and illustrate their alignment with the communicative purposes and disciplinary conventions of Business and Engineering writing.

4.2 Identifying Linguistic Features and Dimensions in Student Writing

This section addresses the first research question: What underlying dimensions and associated features emerge in student writing across Business case studies and Engineering proposals? The analysis focused on the examination of the entire corpus to identify overarching dimensions and linguistic features that characterize student writing. This foundational analysis provides a comprehensive understanding of the corpus before comparing the disciplinary corpora in addressing the second research question.

Linguistic variable groups were identified based on their patterns of co-occurrence in the dataset. The dimensions were then analyzed to interpret the function associated with each dimension, considering the co-occurring features and their presence in individual texts. To

ensure data reliability, the commonality and factorability of the variables were evaluated. No variable showed a correlation score above 0.8. The Kaiser-Meyer-Olkin (KMO) test was used to measure factorability, and it revealed an overall KMO value of 0.79. Variables with commonality scores below 0.18 were excluded because such low scores indicate that these variables do not adequately share variance with the extracted factors. Retaining them could compromise the quality of the factor solution, as they may contribute more noise than meaningful information to the analysis. A five-factor solution was determined to be the most suitable, explaining 40% of the cumulative shared variance. Factors were rotated using Promax rotation to enhance interpretability.

While the analysis draws on the general methodological approach of Biber's (1988) MDA framework, as described previously the dimensions were identified through a data-driven (i.e., bottom-up) approach linguistic features and context of the dataset. Five dimensions emerged from the factor analysis: (1) Informational Density and Elaborated Discourse, (2) Interactive and Situated Discourse, (3) Narrative Focus, (4) Interpersonal and Dialogic Discourse, and (5) Descriptive and Stative Discourse. Unlike Biber's original five dimensions, which provide labels for both ends of the scale (e.g., Involved vs. Informational Production), the dimensions in this study have been named to reflect the dominant linguistic and functional patterns observed in the data. This choice was made because the analysis primarily exhibits one end of each dimension, where the most salient and meaningful linguistic features emerged.

The interpretation and naming of dimensions in this study followed established methodological and rhetorical conventions in MDA, as outlined in both foundational and contemporary works (e.g., Biber, 1988; Friginal & Hardy, 2014; Sardinha & Pinto, 2014). Drawing

on the principle that dimensions represent latent constructs revealed through the co-occurrence of linguistic features, the analysis examined the sets of features with the extent to prominent loadings on each factor. In line with Biber's (1988) original protocol, these groupings were interpreted through a functional lens based on their contribution to communicative purposes such as stance, involvement, or informational density. Where patterns aligned with previously established constructs related, though not always identical, naming conventions were adopted to preserve interpretive consistency while reflecting the specific characteristics of the corpus.

Dimension labels were also shaped by the genre and discipline characteristics of the dataset. For example, a dimension heavily loaded with nominalizations, dense noun phrases, and attributive adjectives was interpreted as Informational Density, a term consistent with prior studies as Gardner et al. (2018), Jin (2018), and Crossley et al. (2014). When dimensions showed less commonly observed co-occurrences, such as clusters of 2nd person pronouns, imperatives, and epistemic modals, the study applied contextual insights and prior interpretations (e.g., Hardy & Römer, 2013; Goulart, 2021) to craft functionally meaningful labels. As Friginal and Hardy (2014) and Sardinha and Pinto (2014) emphasized, interpreting MDA results is both a scientific and interpretive act that one grounded in statistical validity, yet deeply reliant on contextual awareness and disciplinary intuition. The study, therefore, approached this phase as an iterative process that involved consulting factor loadings, examining representative texts, and cross-referencing emerging patterns with findings from prior research. This process informed the labeling of each dimension in a way that captures its rhetorical function while preserving

the specific linguistic and genre-related characteristics of the academic texts under investigation.

Table 4 presents the linguistic features comprising each dimension with only features with factor loadings greater than +/- 0.30 included, which indicates moderate to strong contribution to their respective factors. Factor loadings for features across the corpus are provided in parentheses with underlined features showing overlap with other dimensions.

Table 4. Factor loadings (greater than +/- 0.30) on each dimension

Dimension 1: Informational Density and Elaborated Discourse	
Positive-loading features	Prepositions (0.95) Common nouns (0.94) Attributive adjectives (0.87) Nominalizations (0.84) Definite articles (0.81) Coordinating conjunctions (0.77) Verbs (0.76) Present tense (0.73) Present particle (0.73) Infinitive markers (0.70) Demonstrative determiners (0.69) Indefinite articles (0.68) Clausal coordination (0.67) Adjectives (0.65) Present Participial Postnominal clauses (0.64) Adverbs (0.63) Infinitives (0.63) Auxiliary verbs (0.60) Past Participial Postnominal clauses (0.58) Qualifiers (0.55) Conjunctions (0.55) Past tense (0.54) Cardinal numbers (0.53) Comparative adverbs (0.52) Possibility modals (0.49) Past participle (0.48) *be* as main verbs (0.47) Gerunds (0.45)

	Quantifiers (0.43) Phrasal coordination (0.42) Adverbial particles (0.40) Demonstrative pronouns (0.38) Subordinating conjunctions (0.36)
Negative-loading features	
Dimension 2: Interactive and Situated Discourse	
Positive-loading features	Negations (0.52) 3rd person pronouns (0.51) Main verbs (0.50) *wh* clauses (0.46) *be* as main verbs (0.45) Auxiliary verbs (0.45) Time adverbs (0.43) *that* verb complements (0.41) Predictive modals (0.40) Determiners (0.37) Nouns (0.35)
Negative-loading features	Nominalizations (-0.32) Attributive adjectives (-0.33)
Dimension 3: Narrative Focus	
Positive-loading features	Past tense auxiliary verbs (0.61) Past tense *be* as main verbs (0.57) Past tense main verbs (0.53) Cardinal numbers (0.39)
Negative-loading features	Verbs (-0.30)
Dimension 4: Interpersonal and Dialogic Discourse	
Positive-loading features	Negations (0.39) 3rd person pronouns (0.39)
Negative-loading features	
Dimension 5: Descriptive and Stative Discourse	
Positive-loading features	*be* as main verbs (0.43) Post determiners (0.35)
Negative-loading features	Public verbs (-0.31)

	Private verbs (-0.32)

In this study, positive loadings indicate the presence of linguistic features associated with a given dimension, while negative loadings reflect their relative absence rather than an opposing linguistic pattern. For example, positive loadings in Dimension 2 (Interactive and Situated Discourse) captured features that show interaction and situational references. Negative loadings on this dimension included features such as nominalizations and attributive adjectives, which are less characteristic of interactive and situated discourse. This approach ensured that dimension names offer a clear and precise interpretation of dominant linguistic characteristics based on the communicative purposes and disciplinary conventions of the analyzed text types. It also aligns with the study's focus on the functional significance of linguistic features in the corpus, emphasizing the interpretations rather than binary contrasts that may not meaningfully reflect the observed data.

Each dimension was named by interpreting their co-occurrence patterns in terms of functional roles. Dimension 1 (Informational Density and Elaborated Discourse) was defined by features such as prepositions, common nouns, attributive adjectives, and nominalizations, all of which contribute to precise and elaborated content. These features are typically found in formal and academic texts, which reflect a dense and structured style of communication. Additionally, high loadings for definite articles and coordinating conjunctions further emphasize the cohesive and information dense nature of this dimension, which is well-suited to contexts for clarity and detailed exposition.

Dimension 2 (Interactive and Situated Discourse) reflects the use of linguistic features such as negations, 3rd person pronouns, auxiliary verbs, and time adverbs, which are common in interactive and context-dependent communication. These features suggest an emphasis on immediacy and interpersonal interaction, characteristics often associated with spoken or informal discourse. Negative loadings for nominalizations and attributive adjectives further show the distinction between this dimension and more abstract or formal discourse styles, indicating its situated and dialogic nature. Dimension 3 (Narrative Focus) was defined by the prominence of past tense auxiliary verbs, past tense main verbs, and cardinal numbers, which represent a focus on recounting events and telling stories. These temporal markers align with the sequential structure of narratives, emphasizing the unfolding of actions or events in the past. The slight negative loadings of verbs in general suggests that this dimension prioritizes specific actions over broader linguistic diversity, which reinforce its narrative nature.

Dimension 4 (Interpersonal and Dialogic Discourse) emerged from the presence of features such as negations and 3rd person pronouns, which signal dialogic and interpersonal interaction. These features are commonly used to express interpersonal stance, engage with others, and reference shared participants within the discourse. This dimension captures the linguistic strategies employed in contexts requiring interaction or persuasion, such as conversations, discussions, or instructional texts. Finally, Dimension 5 (Descriptive and Stative Discourse) was characterized by the use of *be* as main verbs and post determiners, which aligns with descriptive and static expressions. These features are based on an emphasis on defining states or qualities, often found in observational or explanatory contexts. The negative loadings for public and private verbs contrast the stative focus of this dimension with more dynamic and

action-oriented discourse, further solidifying its descriptive and static nature. By examining the patterns of co-occurrence and their functional associations, the dimensions were defined to reflect the communicative purposes and variations within the dataset. Each dimension exhibits a unique linguistic variation that goes along with the situational and functional demands of the texts.

4.3 Detailed Analysis of Dimensions in Business and Engineering Writing

Building on the identification of linguistic features and dimensions in Section 4.2, this section continues to address the first research question by exploring the functional and rhetorical roles of dimensions within student writing of Business case studies and Engineering proposals. The dimensions identified through the MDA reflect linguistic variation across the entire corpus, including both Business case studies and Engineering proposals. Analyzing the entire corpus allowed for the identification of shared features that characterize student writing while providing a foundation for subsequent comparisons of discipline-specific patterns. This approach emphasizes revealing co-occurring linguistic features across diverse datasets, while also providing a foundation for subsequent exploration of discipline-specific patterns.

4.3.1 Dimension 1: Informational Density and Elaborated Discourse

This dimension is strongly characterized by features such as prepositions (0.95), common nouns (0.94), attributive adjectives (0.87), and nominalizations (0.84), which indicate a dense and elaborated discourse. These linguistic patterns can be typical of case studies and proposals, where the focus is on presenting detailed information and constructing precise, content-rich descriptions. For instance, the use of prepositions and nominalizations made writers to

compactly express complex ideas and relationships, as in sentences like “Herzberg argues that intrinsic factors motivate employees while hygiene factors lead to no-job dissatisfactions.” Similarly, attributive adjectives and common nouns were employed to add specificity and detail to facilitate the presentation of dense and objective information. The presence of definite articles (0.81) and coordinating conjunctions (0.77) suggests a need for cohesion and logical flow, which are important in creating clear and persuasive arguments in documents. Additionally, features such as present tense verbs (0.73) and infinitive markers (0.70) reflect the forwarding-looking and solution-oriented stance. These features support the analytical nature of the texts by focusing on current situations and actionable views. For instance, proposals often emphasize steps for implementation or analysis of challenges, as seen in “Hydroelectric power emerges as the most efficient means of electricity production, converting around 90% of available energy into electricity.” Overall, the patterns in this dimension are indicative of analytical business and technical writing, where the communicative purposes were to deliver detailed, objective, and actionable information. These texts were designed to inform decision-making processes while having clarity and precision in presenting ideas.

4.3.2 Dimension 2: Interactive and Situated Discourse

This dimension is characterized by features such as negations (0.52), 3rd person pronouns (0.51), and *wh* clauses (0.46), which are indicative of interactive and situational discourse. These features are relevant in case studies and proposals, where stakeholder analysis or situational discussions require interpersonal references and contextual framing. For example, 3rd person pronouns like “they” or “their” help to refer to specific stakeholders or groups within a given context, while negations like “not” are often used to show gaps, contradictions, or

limitations in data or proposals. The inclusion of auxiliary verbs (0.45) and predicative modals (0.40) reflects the speculative and evaluative tone, which is essential for discussing potential outcomes or assessing scenarios. For instance, sentences such as “Urban farms can feasibly be implemented profitably with the right strategies,” demonstrate how predictive modals like “can” are used to express possibilities, while auxiliary verbs like “be” establish factual grounding in a specific context. Similarly, time adverbs (0.43) like “recently” or “yesterday” provide temporal referencing, which is critical for discussing trends, performance evaluations, or situational updates within the discourse.

The negative loadings of features such as nominalizations (-0.32) and attributive adjectives (-0.33) further exhibit the interactive and situated nature of this dimension by contrasting it with the abstract and elaborated style seen in Dimension 1. These linguistic patterns suggest a more immediate and applied style of communication, where situational awareness and targeted recommendations are central to the communicative purposes. This aligns with the practical and applied focus of both business and engineering texts, which often involve real contexts and stakeholder engagement.

4.3.3 Dimension 3: Narrative Focus

This dimension is defined by features such as past tense auxiliary verbs (0.61), past tense *be* as main verbs (0.57), and past tense main verbs (0.53), which are central to recounting events or describing prior actions. In business and engineering contexts, these patterns are often seen in sections that summarize project histories, describe past performance, or narrate the sequence of events leading to the current problem. For example, in an engineering

proposal, sentences like “It was found that 81% had prior suicidal intentions, 66% had previously attempted, and 105 of the victims had serious mental health problems.” illustrate how past tense constructions are used to contextualize findings. Cardinal numbers (0.39) further enhance this narrative dimension by offering specific details about quantities, dates, or measurable outcomes that are important in building credibility and clarity in documents. The use of numerical specifics often allows business and engineering texts to substantiate claims and provide a robust foundation for current analysis or proposed solutions.

The slight negative loading of general verbs (-0.30) suggests that the dimension focuses on temporally anchored verbs over broader verbal usage, reinforcing its focus on recounting specific events or actions. This dimension, therefore, captures linguistic patterns associated with narrating and contextualizing prior actions or outcomes, which are essential for situating current analyses and proposals within a broader temporal framework.

4.3.4 Dimension 4: Interpersonal and Dialogic Discourse

This dimension features negations (0.39) and 3rd person pronouns (0.39), reflecting interactional and relational aspects of discourse. These linguistic features can be particularly relevant in case studies when discussing stakeholders’ perspectives or contrasting different approaches. For example, in the sentence, “Although some people may say treating employees with good welfare is the best way, it does not align with Amazon’s current approach,” the use of 3rd person pronouns (e.g., “some people”) and negation (e.g., “does not”) facilitates the discussion of differing perspectives, which is central to analyzing stakeholder views and providing balanced arguments. In engineering proposals, relational language is often used to

acknowledge collaborators or external entities, as see in “In Asian countries, the implementation of EPR systems has greatly improved their recycling rates.” The use of 3rd person pronouns (e.g., “their”) and relational phrases emphasizes the involvement of external entities. These examples illustrate the dialogic elements of communication that are important in collaborative and stakeholder-focused contexts.

This dimension captures the relational and dialogic elements necessary for effective communication in business and engineering environments. The linguistic patterns reflect an emphasis on addressing concerns, negotiating solutions, and acknowledging the roles of various entities, all of which are essential for fostering collaboration and achieving effective engagement.

4.3.5 Dimension 5: Descriptive and Stative Discourse

This dimension is characterized by features such as *be* as main verbs (0.43) and post determiners (0.35), which emphasize precise and explicit descriptions of states, attributes, or conditions. In business and engineering writing, these features are often central to describing system states, technical specifications, or solution outcomes with clarity and accuracy. For example, in engineering proposals, a sentence like “These signals are electrical pulses created by your body to trigger limb movement.” Illustrates the stative use of “be” to describe a technical process or condition. Similarly, in business contexts, sentences such as “Its many benefits include its short-form entertainment and accessible, informative content.” showcase the use of post determiners like “many” to quantify and qualify specific attributes clearly.

The negative loadings of features such as public verbs (-0.31) and private verbs (-0.32) further reinforce the stative and descriptive nature of this dimension by contrasting it with more dynamic and action-oriented discourse. Public and private verbs, often associated with interactional or evaluative language, are less prominent in this dimension, representing its focus on objective description over subjective evaluation or interpersonal engagement. This dimension notes the descriptive focus inherent in technical and analytical documents, where clarity, specificity, and precision are essential for ensuring that stakeholders and audiences fully understand the attributes, requirements, or implications of the subject matter. The linguistic patterns identified here are particularly well-suited to technical communication, which often requires explicit detailing of processes, designs, or outcomes.

4.3.6 Summary of Dimensions and Linguistic Features in Student Writing

This summarizes the findings for the first research question, which sought to identify the dimensions and linguistic features of student writing from both Business case studies and Engineering proposals. Five dimensions were identified through MDA, each reflecting distinct functional and rhetorical roles. Dimension 1 showed dense, content-rich language important for technical precision. Dimension 2 captured interpersonal engagement and situational framing, which are critical for applied contexts. Dimension 3 focused on recounting events to contextualize findings. Dimension 4 integrated relational elements to address stakeholders or contrasting viewpoints. Finally, Dimension 5 emphasized clarity and explicit descriptions essential for technical communication. Together, these dimensions illustrated the linguistic strategies and conventions employed in Business and Engineering writing, describing their alignment with disciplinary demands.

4.4 Dimension Comparison of Business case studies and Engineering proposals

This section addresses the second research question: What are the dimensional differences between Business case studies and Engineering proposals? By comparing the dimension scores of these two text types, the analysis aimed to see how their linguistic and functional characteristics diverge, which reflect the distinct communicative purposes and disciplinary conventions of Business and Engineering. Dimension scores were calculated for each text type by averaging the factor loadings across all texts within each group. To evaluate these differences, one-way ANOVAs were conducted to compare the mean dimension scores of Business case studies and Engineering proposals across the identified dimensions. This statistical approach was well-suited for identifying significant differences in mean scores while accounting for variability within groups and ensuring the reliability of the findings. Figure 4 illustrates the dimension scores for each corpus.

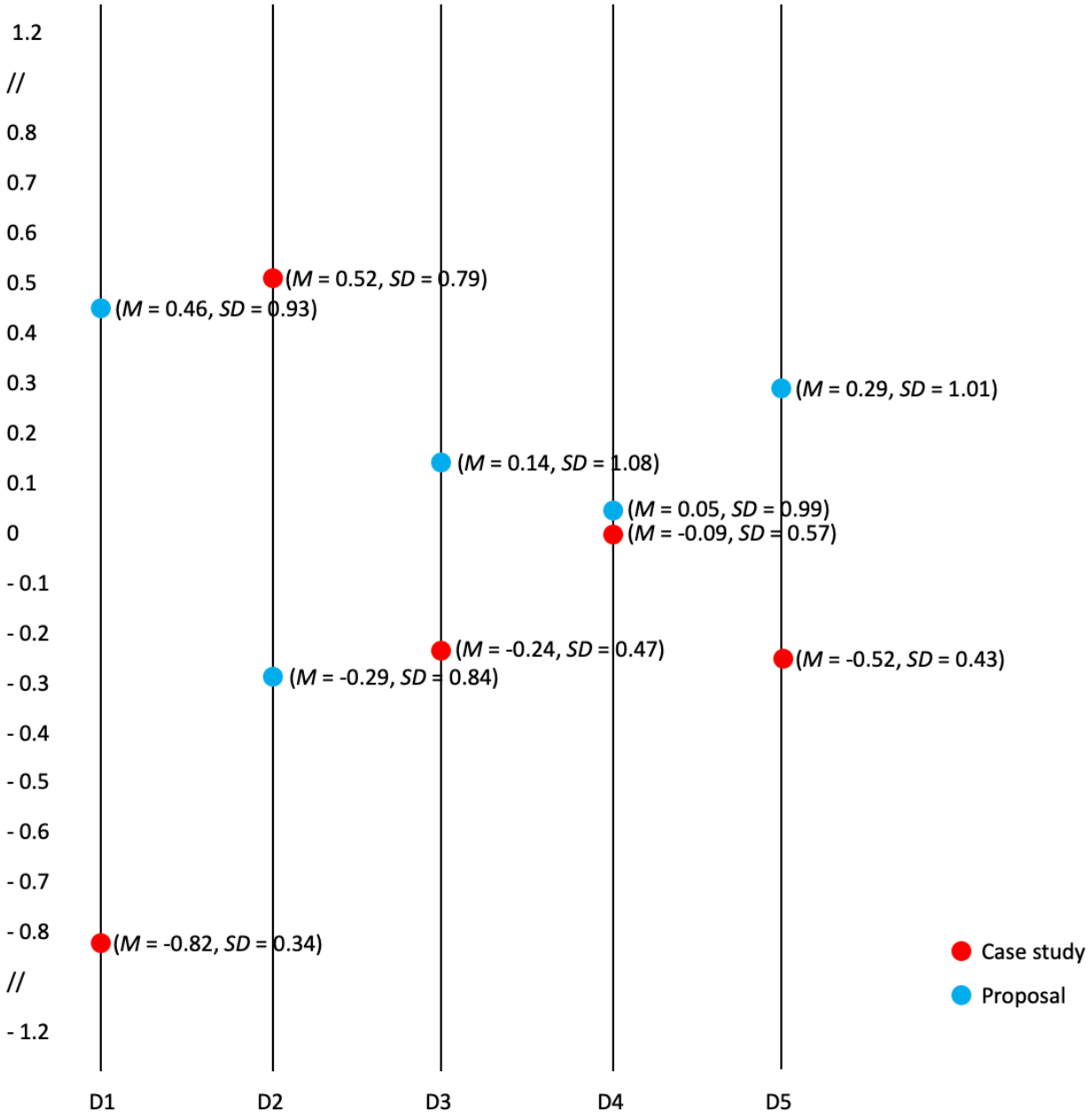

Figure 4. Dimension comparisons of Business case studies and Engineering proposals

As shown in Figure 4, there were clear differences for most dimensions between Business case studies and Engineering proposals, reflecting their distinct communicative purposes and disciplinary conventions. Engineering proposals demonstrated a greater emphasis on informational density and elaborated discourse (Dimension 1). The linguistic features

associated with this characteristic were used to highlight the technical precision required in Engineering communication, where conveying detailed and complex information is often paramount. In contrast, Business case studies relied more on interactive and situated discourse (Dimension 2), marked by interpersonal language and contextual framing that facilitates engagement with particular audiences and addresses authentic scenarios. These differences revealed how each text type employs targeted linguistic strategies to meet the communicative purposes of its discipline, showing the broader relationship between language use and disciplinary practices.

Dimension 3 (Narrative Focus) demonstrated a smaller difference, with both text types incorporating narrative elements, though Engineering proposals slightly favored recounting project histories and contextualizing findings to provide a technical backdrop for proposed solutions. This suggests that while narrative features are integral in both text types, their function aligns with the distinct disciplinary goals of situating findings within either practical or technical frameworks. Dimension 4 (Interpersonal and Dialogic Discourse) were very similar, indicating that both text types minimize dialogic features to maintain professionalism and analytical tone. This can be explained by the shared need for objectivity and precision in both text types, despite their differing communicative purposes.

Finally, Dimension 5 (Descriptive and Stative Discourse) showed larger differences, with Engineering proposals relying more on descriptive and stative features to articulate system specifications and material properties. This contrasts with Business case studies, which use descriptive features to qualify organizational challenges and articulate strategic recommendations. Collectively, these findings reinforce the distinct linguistic and functional

patterns of Business case studies and Engineering proposals, showing their targeted approaches to meet specific communicative and disciplinary conventions.

4.4.1 Statistical Validation of Dimension Comparison through ANOVA

The MDA analysis illustrated different linguistic patterns across five identified dimensions for the two corpora: Business case studies and Engineering proposals. To statistically validate these differences, five one-way ANOVAs were conducted to compare mean dimension scores across the two corpora. ANOVA was chosen over t-test due to its ability to analyze multiple dependent variables within a unified framework, reducing the risk of inflated Type 1 errors. Conducting separate t-tests for each dimension would increase the likelihood of false positives because each test would independently assess significance without considering the cumulative error rate. ANOVA mitigates this risk by evaluating all dimensions in a single model, which preserve the integrity of the statistical analysis.

Given that only two groups were being compared, post-hoc tests were unnecessary; however, to make sure the validity of the findings, the alpha level was adjusted using the Bonferroni correction. This adjustment accounted for the multiple comparisons conducted and reinforced the reliability of the results by minimizing the potential for spurious significance. The statistical validation confirmed the observed differences in linguistic dimensions were not due to random variation but rather reflected genuine distinctions in the communicative purposes employed by Business case studies and Engineering proposals. These findings indicate the functional linguistic choices that characterize each text type, further supporting the efficacy of MDA in distinguishing textual patterns across academic writing domains.

Significant differences were observed between Business case studies and Engineering proposals for Dimension 1 (Informational Density and Elaborated Discourse), with Engineering proposals scoring significantly higher ($F(1,212) = 135.8$, $p < 0.001$, $\eta^2 = 0.39$). This statistically validates the observed emphasis on dense noun phrases, nominalizations, and elaborated structures in Engineering proposals, showing their focus on conveying detailed technical information. Similarly, Dimension 2 (Interactive and Situated Discourse) showed a significant difference ($F(1,212) = 47.54$, $p < 0.001$, $\eta^2=0.18$), with Business case studies scoring higher, confirming their stronger reliance on interpersonal language and situational framing to engage with stakeholders and real-world contexts. For Dimension 3 (Narrative Focus), a smaller but significant difference was found ($F(1,212) = 9.04$, $p < 0.01$, $\eta^2=0.04$), indicating that both text types incorporate narrative elements, while Engineering proposals slightly favor recounting project histories and contextualizing findings. In contrast, no significant difference was found for Dimension 4 (Interpersonal and Dialogic Discourse) ($F(1,212) = 1.53$, $p = 0.218$, $\eta^2=0.007$), which suggests that both text types minimize dialogic features to maintain professionalism and analytical tone. Finally, Dimension 5 (Descriptive and Stative Discourse) revealed significant variation ($F(1,212) = 44.67$, $p < 0.001$, $\eta^2= 0.17$) with Engineering proposals relying more heavily on descriptive and stative features, consistent with their focus on system specifications and material properties. These results reinforce the distinct linguistic strategies used by Business case studies and Engineering proposals to fulfill their respective communicative purposes, with the ANOVA results providing robust statistical evidence for the observed patterns.

4.4.2 Functional and Rhetorical Perspectives across Dimensions

This section expands on the findings for the second research question by illustrating how Business and Engineering student writing differ across dimensions, with a particular focus on their functional and rhetorical views, illustrated by several excerpts from the texts. Engineering proposals scored notably higher on Informational Density and Elaborated Discourse (Dimension 1), emphasizing their reliance on detailed, precise, and technical descriptions. This shows their focus on delivering complex information through dense noun phrases, nominalizations, and elaborated structures, important for conveying system specifications, design details, and technical processes. For example, in Excerpt 1, the text shows the significant contribution of buildings to electricity consumption and CO2 emissions by presenting specific metrics (71% of total electricity and 40% of CO2 emissions). This precise quantification is followed by a discussion of AI-powered smart buildings, where technical details about interconnected sensors and their functions – optimizing energy efficiency, security, and functionality – reveal the elaborated discourse typical of engineering texts. Such descriptions make a systematic and technical understanding of complex systems.

Excerpt 1.

> Buildings consume approximately 71% of total electricity and are responsible for 40% of CO2 emissions into the atmosphere. To solve this problem, we need to find a balance between reducing energy usage while ensuring occupants' comfort and well-being. One of the most useful functions of AI in reducing energy consumption is through smart buildings. These structures are equipped with various interconnected sensors aimed at optimizing energy efficiency, security, and functionality.

In contrast, Business case studies scored lower on this dimension and suggested that they focus on other discursive features over lexical and syntactic complexity, often favoring more accessible and situationally relevant language to address broader audiences. For instance, in Excerpt 2, the discussion of Amazon's workplace relied on relatable, situational descriptions such as "bad working conditions" and "big power distance between subordinates and managers." These phrases convey critical issues in simpler terms without excessive technical detail. While less dense, the text emphasizes authentic scenarios and workplace dynamics to connect with the audiences. This approach reflects the business focus on practical implications and accessibility rather than technical elaboration.

Excerpt 2.

> Amazon has mostly hygiene factors like bad working conditions (burnout stated by Linda Duxbury); hard-working policies that should be prohibited according to Ray Williams; big power distance between subordinates and managers, and a big turnover from the employees. The only way Amazon uses motivation in a satisfactory way is by reward, as seen in their Nike shoes reward.

Conversely, Interactive and Situated Discourse (Dimension 2) is significantly more prominent in Business case studies with higher scores indicating a stronger emphasis on engagement with stakeholders and situational framing. Case studies frequently employ interpersonal language, temporal references, and situational framing to analyze and discuss real-world scenarios, client interactions, or customer feedback. For example, Excerpt 3 from one

student's case study on Amazon's leadership style illustrates this dimension through the statement, "Amazon sells stuff online, so customers never directly interact with employees." This situational framing reveals how Amazon separates its employee dissatisfaction from customer interactions, shielding its reputation while maintaining public satisfaction.

Excerpt 3.

> Employees at Amazon experience little to no interpersonal roles, as they sell stuff online, so customers never directly interact with [its] employees. This shows how Amazon covers up its dissatisfied employees, as customers see no emotion or encounter face-to-face interaction – it is all completed through technology.

On the other hand, Engineering proposals scored lower on this dimension, which shows their objective tone and reduced focus on situational engagement. Their primary aim was to propose solutions and elaborate on technical details in a formal and detached manner. For example, Excerpt 4 emphasized the technical feasibility and systematic solutions for recycling e-waste and minimized interpersonal or situational references. The contrast between these approaches demonstrates how Business case studies leverage interactive discourse to connect with stakeholders, whereas Engineering proposals focus on systematic problem-solving with a formal and technical tone.

Excerpt 4.

> Hydrometallurgical recycling strategies for the recovery of rare earth elements from consumer electronic scraps are key to addressing environmental concerns. Urban waste, such as consumer electronic scrap, has remarkable potential to meet growing demand for rare earth elements in hi-tech applications, in line with circular economy principles. Advances in recycling could reduce dependency on virgin materials and promote sustainable practices.

Narrative Focus (Dimension 3) scores were also significantly different, with Engineering proposals slightly favoring narrative elements. In proposals, past tense constructions and narrative sequences are often used to recount project histories, describe testing processes, or contextualize findings. For example, one proposal on Montreal's metro system recounts its historical and functional context (Excerpt 5). This narrative outlines the problem before introducing technical solutions. Similarly, another engineering text on virtual reality narrates its development in a sequential manner (Excerpt 6). This style emphasizes technical process while maintaining a narrative structure.

Excerpt 5.

> The Montreal metro system is one of the key modes of transportation for many residents on the island. Being a non-exhaust transit system, the metro does not use any fossil fuels and is a greener mode of transportations. However, the system does have its own flaws; despite it not emitting CO2 gases, it still does contribute to air pollution through the emission of fine air particulate, especially in the underground stations.

Excerpt 6.

> Virtual Reality (VR) technology has made significant progress in recent years, offering immersive experiences that were once only seen in science fiction. VR technology creates a simulated environment that users can interact with in a seemingly real way. This is achieved through a combination of hardware components, such as headsets with stereoscopic displays, motion sensors, and controllers, and software that processes input and renders the virtual environment in real time.

Business case studies also employed narrative elements but tended to use them differently – to analyze events and derive actionable perspectives. For instance, one analysis of Jeff Bezos' leadership style narrates the impact of Amazon's coercive workplace policies (Excerpt 7). This example links specific events to leadership theories and illustrates how such practices negatively affect employee well-being. Another case study elaborates on Amazon's punitive management practices (Excerpt 8). This narrative highlights the situational context of workplace practices and emphasizes their implications for employee morale and retention. These examples together show how Business texts use narrative structures to connect real-world practices with theoretical perspectives, contrasting with the more technical and process-focused narrative sequences in Engineering texts. Both approaches reflect the priorities of their respective fields – Business texts centering on actionable human outcomes, while Engineering texts emphasize systemic solutions.

Excerpt 7.

Jeff Bezos' leadership will ultimately lead to decline in Amazon's success based on Herzberg, Mintzberg, and French and Raven's theories. Bezos does not provide motivators but rather bad hygiene; shows poor interpersonal, informational, and decisional roles as a manager; shows strong coercive power, but little reward power, legitimate, referent, and expert power. For instance, during the 'Save Santa' emergency, employees were made to work overtime in the warehouse without proper food or rest, demonstrating coercive power rather than motivational leadership.

Excerpt 8.

The coercive power Bezos utilizes is evident in how he forces employees to work through personal crises such as cancer or miscarriages. Coercive power is not effective in the long run because it scares people into completing tasks without fostering engagement or loyalty. This is contrasted with referent or expert power, which could build trust and motivate employees more effectively.

Both text types exhibited minimal reliance on Interpersonal and Dialogic Discourse (Dimension 4), as evidenced similarly by low scores in this dimension. This suggests that both Business case studies and Engineering proposals avoid relational or conversational language, which aligns with their shared emphasis on professionalism and analysis rather than interaction or dialogue. While both text types prioritize objectivity and analytical depth, Business case studies realize this dimension through references to stakeholders and contrasting perspectives to build arguments and evaluate decisions. For example, one Business text contrasted Amazon's managerial strategies with Google's approach (Excerpt 9), illustrating how contrasting perspectives are used to explore the effectiveness of leadership styles. Additionally, another

Business text linked leadership decisions to broader societal impacts (Excerpt 10), such as employees and customers, were incorporated to contextualize managerial actions. These examples show the ways in which Business case studies engage with dialogic features without relying on interpersonal discourse, using such features to support their analytical and evaluative purposes.

Excerpt 9.

> Amazon focuses on the older age group who are ‘more interested in wages, vacation, and flexible work.’ Amazon focuses on these factors to ensure that employees stay even with Bezo’s brute management style. In contrast, Google provides ‘extrinsic motivating factors’ to keep their employees completely contented, creating a sense of community.

Excerpt 10.

> It will catch up to Amazon and customers are going to become more sensitive to how employees get treated and will therefore boycott Amazon, lowering Amazon’s revenues.

In contrast, Engineering proposals emphasized technical precision and procedural clarity, avoiding dialogic strategies as their focus is on solution-oriented communication rather than evaluative argumentation. For instance, one proposal focused on the technical requirements for platform screen doors (Excerpt 11), emphasizing design and implementation details. Another proposal discussed energy storage systems (Excerpt 12), centering on the technical aspects of the technology without incorporating interpersonal elements. These examples show how

Engineering proposals achieve their communicative purposes by presenting detailed and precise technical information in a manner that avoids interpersonal or dialogic framing.

Overall, these examples exhibited how both text types adapted their linguistic strategies to meet their respective communicative purposes, demonstrating variations in the realization of Dimension 4 without significant divergence.

Excerpt 11.

> To evaluate the needed PSD [Platform Screen Doors] for Montreal's metro system, we can take a look at the design created for Elizabeth's line of metros in London, which required a hybrid frame that would both be supported by the ceiling and by the floor, to simplify the supporting structure of the door frames.

Excerpt 12.

> Compressed Heat Energy Storage (CHEST) is a thermomechanical process meant for medium-long term storage with medium-large scale applications. It works by using electrical energy to compress a vapour organic fluid, produced by a heat source, then condensing and cooling the fluid by transferring it to the HES, usually pressurized water or a PCM.

Finally, Descriptive and Stative Discourse (Dimension 5) is more pronounced in Engineering proposals based on their technical orientation and the need to describe system states, specifications, and material properties with precision and clarity. For example, one

proposal describes the state of the metro infrastructure (Excerpt 13), focusing on static system issues that represent the descriptive nature of engineering discourse. In another example, an analysis of driver fatigue detection systems was explained (Excerpt 14) and this descriptive approach provides clarity on the system's operational states and technical specifications.

> Excerpt 13.
>
> The STM is outdated and has frequent interruptions. The disruptions stem from multiple reasons, most frequently from an unauthorized person or object in the tracks and these disruptions generally take a lot of time to remediate.

> Excerpt 14.
>
> Macro glances are defined by predetermined regions… often specific landmarks such as road signs. Identifying macro and micro glances is essential to understand the state of the agent; however, micro glances, such as eye blinking in a temporal context, remain a useful indicator for fatigue.

Business case studies, in contrast, score lower on this dimension, as their discourse often focuses on dynamic processes, strategic analyses, and recommendations rather than static descriptions. For instance, one case study focused on leadership improvement at Amazon and offered actionable recommendations (Excerpt 15). This approach emphasizes the need for dynamic adjustments in management strategies to achieve long-term success. Similarly, another discussion of Amazon's market position argues potential future scenarios (Excerpt 16), reflecting

the case study's focus on adaptability and foresight. These examples demonstrate how Business case studies prioritize strategic planning and adaptability over static descriptions, contrasting with the technical precision and descriptive focus of Engineering proposals.

Excerpt 15.

> According to Herzberg's theory on motivation, Bezos is doing it all wrong. Motivation leads to a happier workforce and boosts the image of your company. To have motivated employees, you must first take care of hygiene factors, such as wages and work conditions, and then provide motivational factors like responsibility. Satisfied employees tend to do extra work without coercion.

Excerpt 16.

> Amazon's laser focus on customer needs allowed it to outperform the stock market over time, but what if there was a shift in customer values, and they decided to wait an extra day to receive their purchases and shop elsewhere where bosses treat employees fairly?

Overall, these excerpts highlighted the distinct linguistic patterns employed by Business case studies and Engineering proposals with their respective communicative purposes. Business case studies focused on engagement, situational framing, and dynamic analyses, using interpersonal and dialogic features to address challenges, provide actionable perspectives, and connect with a broader range of stakeholders. In contrast, Engineering proposals emphasized technical precision, informational density, and descriptive clarity, relying on objective, stative

discourse to deliver complex specifications, detailed processes, and practical solutions to technically informed audiences.

4.4.3 Summary of Dimensional Patterns in Business and Engineering Writing

The findings for the second research question revealed distinct dimensional differences between Business case studies and Engineering proposals based on their communicative purposes and disciplinary conventions. For Dimension 1 (Informational Density and Elaborated Discourse), Engineering proposals demonstrated a significantly higher emphasis on technical precision through dense noun phrases, nominalizations, and elaborated structures, essential for conveying detailed technical information. In Dimension 2 (Interactive and Situated Discourse), Business case studies exhibited stronger engagement with situational and interpersonal elements, leveraging contextual framing and audience interaction to analyze real-world scenarios and engage stakeholders. Dimension 3 (Narrative Focus) showed that while both text types incorporated narrative elements, Engineering proposals used them to recount project histories and contextualize technical solutions, whereas Business case studies applied narratives to analyze events and derive actionable recommendations.

For Dimension 4 (Interpersonal and Dialogic Discourse), both text types minimized dialogic features to maintain professionalism and analytical tone, though Business case studies integrated dialogic strategies, such as contrasting perspectives, to support evaluative discussion. Finally, Dimension 5 (Descriptive and Stative Discourse) revealed that Engineering proposals relied heavily on descriptive and stative features to articulate system specifications and materials properties, whereas Business case studies emphasized dynamic processes and

strategic solutions. Overall, these findings illustrate how Business and Engineering texts adapt their linguistic strategies to align with their communicative purposes, indicating that language is shaped by disciplinary and functional demands.

4.5 Chapter Summary

This chapter explored the linguistic features and dimensional differences in student writing across Business case studies and Engineering proposals by addressing two research questions. The analysis revealed how these text types differ in their use of language based on their distinct communicative purposes and disciplinary conventions. To summarize the findings, Table 5 provides a summary of the five identified dimensions, their primary characteristics, and the comparative results for Business and Engineering texts.

The first research question examined the linguistic features and dimensions of the corpus. Five dimensions of linguistic variation were identified through MDA, each representing distinct functional and rhetorical roles. The first dimension (Informational Density and Elaborated Discourse) was characterized by dense, precise, and highly informative language. While present in both text types, this dimension was significantly more pronounced in Engineering proposals, which relied heavily on technical terminology, nominalizations, and structured descriptions to deliver complex information effectively.

The second dimension (Interactive and Situated Discourse) demonstrated the use of interpersonal and situational language, particularly in Business case studies. This dimension captured how case studies engage stakeholders and address real-world challenges through situational framing, temporal references, and interpersonal language. The third dimension

(Narrative Focus) illustrated the use of narrative elements in both text types to recount events and contextualize findings. Engineering proposals employed narratives to provide technical context, while Business case studies used them to link practical scenarios with theoretical perspectives.

The fourth dimension (Interpersonal and Dialogic Discourse) was presented in both text types that their professional and analytical tone was minimal. However, Business case studies incorporated more dialogic elements to present contrasting perspectives and address stakeholder considerations. Finally, the fifth dimension (Descriptive and Stative Discourse) was predominantly observed in Engineering proposals. This dimension emphasized explicit descriptions of system states, specifications, and material properties, which are essential for technical communication.

The second research question focused on the dimensional differences between Business and Engineering texts. Business case studies excelled in Interactive and Situated Discourse (Dimension 2) and Dialogic (Dimension 4) elements, emphasizing engagement and practical applications. On the other hand, Engineering proposals scored higher in Informational Density and Elaborated Discourse (Dimensions 1) and Descriptive and Stative Discourse (Dimensions 5), which indicates that their focus was on technical precision and detailed descriptions. Both text types incorporated Narrative Focus (Dimension 3), though with differing purposes: Business case studies analyzed workplace dynamics and leadership practices, while Engineering proposals contextualized technical findings.

Table 5. Summary of Dimensions and Comparisons

Dimension	Primary characteristics	Business case studies	Engineering proposals
1. Informational Density and Elaborated Discourse*	Dense noun phrases, nominalizations, technical precision, and structured descriptions	Lower alignment	Higher alignment
2. Interactive and Situated Discourse*	Interpersonal language, situational framing, and temporal references	Higher alignment	Lower alignment
3. Narrative Focus*	Use of narrative structures to contextualize findings and recount events	Moderate alignment	Moderate alignment
4. Interpersonal and Dialogic Discourse	Dialogic features, contrasting perspectives, stakeholder-focused language	Minimal alignment	Minimal alignment
5. Descriptive and Stative Discourse*	Descriptions of system states, specifications, and technical details	Lower alignment	Higher alignment

Note: The asterisk indicates a statistically significant difference.

Overall, the findings illustrated how Business case studies and Engineering proposals employ distinct linguistic strategies to fulfill their respective communicative purposes. Business case studies prioritize situational engagement, practical recommendations, and stakeholder connections, while Engineering proposals focused on technical precision, detailed descriptions, and systematic solutions aimed at specialized audiences. These differences highlight the alignment of each text type with its disciplinary conventions and objectives.

The next chapter will synthesize these findings with broader theoretical and practical implications. It will explore how the identified dimensions and observed differences align with existing literature on student writing. Furthermore, the chapter will consider how these perspectives can inform teaching practices, helping students develop the necessary linguistic

and rhetorical skills to excel in their respective fields. By integrating the results of this study with pedagogical and theoretical perspectives, Chapter 5 will provide a comprehensive analysis of the implications of these findings and propose directions for future research.

Chapter 5. Discussion

5.1 Introduction

This chapter builds upon the findings presented in the previous chapter by synthesizing the observed linguistic and dimensional differences in Business case studies and Engineering proposals. The discussion situates these results within the broader context of disciplinary writing and student writing, examining how these findings go along with existing theoretical frameworks and prior empirical studies adopting MDA analyses. The chapter also explores practical implications, particularly in education, where the perspectives can inform strategies to enhance students' writing skills for disciplinary-specific purposes. By addressing these broader implications and considering the limitations of the study, this chapter seeks to provide a comprehensive understanding of how linguistic strategies are to meet the communicative purposes of different disciplines. Finally, directions for future research will be proposed to make sure that the contribution extend to ongoing discussions in this field.

5.2 Overview of Findings

This study compared linguistic variation in student writing in Business and Engineering disciplines, focusing on case studies and proposals, respectively. Using MDA, five dimensions of linguistic variation emerged from the data: (1) Informational Density and Elaborated Discourse, (2) Interactive and Situated Discourse, (3) Narrative Focus, (4) Interpersonal and Dialogic Discourse, and (5) Descriptive and Stative Discourse. These dimensions revealed distinct linguistic patterns in each text type and demonstrated how language reflects disciplinary conventions and communicative purposes. Business case studies highlighted interpersonal

engagement and the use of analytical narratives to address real-world scenarios and propose solutions. In contrast, Engineering proposals prioritized technical precision, informational density, and descriptive clarity to convey detailed system specifications and processes. These findings illustrate how linguistic strategies go with the unique demands of Business and Engineering writing, shedding light on the alignment between language use and the objectives of the writing.

5.3 Similarities and Differences

5.3.1 Shared Linguistic Features

Despite the disciplinary differences, Business case studies and Engineering proposals share several linguistic features, particularly those that enhance informational density, such as dense noun phrases, nominalizations, and attributive adjectives, through common constellations of linguistic features. These features co-occur to construct meaningful arguments and enhance the clarity of complex ideas. This finding goes along with research by Biber and Gray (2010), which noted the prevalence of nominal structures in academic writing. Additionally, both text types incorporated narrative elements that function as cohesive devices, situating information within broader frameworks. For instance, narratives in Business case studies often recounted organizational challenges or leadership practices, while Engineering proposals used narratives to describe project histories or technical advancements. Gray and Biber (2012) observed that these narrative elements were essential for adding coherence and situational relevance to texts. These shared linguistic patterns show the importance of aligning

co-occurring features with communicative purposes, such as clarity and relevance across disciplines.

5.3.2 Distinct Linguistic Strategies

While Business and Engineering texts share some linguistic features, they exhibited distinct constellations of features that reflect their unique disciplinary norms. Business case studies were characterized by interactive and dialogic discourse, with frequent use of 3rd person pronouns, temporal markers, and evaluative language. These features are critical for analyzing real-world scenarios and proposing actionable solutions. Hyland (2005) argued that persuasion and adaptability are central to Business communication and these patterns of language enable writers to balance analytical rigor with audience engagement. In contrast, Engineering proposals adopted a more detached tone, which focuses on descriptive and stative discourse to detail technical specifications and procedures. Swales (1990) argued the importance of these features in ensuring the replicability and clarity of technical documents. Moreover, case studies integrated interpersonal discourse (e.g., hedging and boosting) to balance caution and assertiveness in their arguments (Hyland, 2000). Proposals, however, relied on a formal and objective tone, minimizing interpersonal elements to maintain their focus on factual accuracy. Staples and JoEtta (2022) emphasized that linguistic features such as nominalizations and passive constructions reflect the functional goals of disciplinary writing, particularly in STEM fields.

5.3.3 Contextual and Disciplinary Influences

The linguistic differences between Business case studies and Engineering proposals can be understood through the broader contextual and disciplinary norms of their respective fields. Business writing is often shaped by its emphasis on persuasion, adaptability, and stakeholder engagement, leading to the use of features that enhance interpersonal interaction and situational framing. Paltridge (2004) noted, genre conventions are influenced by their socio-cultural and professional contexts that align with the findings of this study. Business writing often pays attention to persuasion and adaptability and requires writers to engage with diverse stakeholders. Conversely, Engineering writing focuses on standardization and precision due to the technical and procedural focus on the discipline. These findings also go along with prior studies that argued the role of situational variables in shaping linguistic choices. For example, Gardner et al. (2019) observed that Business and Engineering writing reflect distinct audience expectations and communicative purposes. These findings are also supported by Staples and JoEtta (2022), who argued that situational and disciplinary contexts influence both the audience and linguistic features of texts. This study contributes to this body of work by illustrating how these conventions manifest in student writing and offers insights into the unique challenges and strategies of each genre.

5.3.4 Comparison with Prior MDA studies

To situate the findings of this study within the broader context of prior MDA research, it is essential to compare the linguistic patterns observed in Engineering proposals to those identified in other STEM fields in previous studies. Research by Biber and Gray (2010) and Staples et al. (2016) on academic writing in STEM disciplines showed the prevalence of dense informational discourse and nominalizations. This study supports those findings, as Engineering

proposals demonstrated a strong reliance on Informational Density and Elaborated Discourse (Dimension 1). Specifically, the frequent use of dense noun phrases, attributive adjectives, and nominalizations in this study reflects the technical precision necessary for conveying detailed specifications and system processes, consistent with patterns reported in STEM-focused MDA studies.

However, this study extends prior research by revealing how Engineering proposals specifically emphasize Descriptive and Stative Discourse (Dimension 5), which was not as prominently discussed in earlier studies. Unlike broader STEM texts analyzed by Hyland (2000) and Biber et al. (2011), which often include elements of dynamic analyses or methodological discussions, the Engineering proposals in this corpus focused on static descriptions of system states, technical specifications, and material properties. For example, features such as *be* as main verbs and post determiners were identified as central to Dimension 5 in this study, showing their role in achieving technical clarity.

In contrast to previous MDA studies on STEM fields, which often exhibit the narrative and procedural aspects of scientific writing (e.g., Gray & Biber, 2012), the Engineering proposals in this study demonstrated only moderate reliance on Narrative Focus (Dimension 3). While narratives were used to provide technical context, they were often secondary to descriptive and elaborative discourse. This contrast with findings in fields like biology or environmental science, where narratives are more prominent in telling experiences and results (e.g., Cortes, 2004). Jin (2018) extended this perspective by exploring the linguistic features of discussion sections in chemical engineering research articles. Similar to the findings in Engineering proposals in this study, the study revealed that technical precision and informational density were emphasized

based on the shared communicative purpose of providing clear and persuasive accounts of technical work. However, Jin also noted the importance of rhetorical features like evaluative language in high-impact articles, which were less prominent in the Engineering proposals examined here. This suggests that while Engineering proposals prioritize clarity and technical description in general, specific genres or disciplines may use additional strategies to engage their academic goals. Finally, the relative lack of Interactive and Situated Discourse (Dimension 2) in Engineering proposals compared to Business case studies shows another key difference from STEM-related studies that emphasize collaboration and interdisciplinary engagement (e.g., Swales, 1990). While collaborative aspects are often central to broader STEM texts, the proposals in this study displayed a more formal and detached tone due to their highly specialized communicative purposes.

In addition to comparisons with STEM fields, it is also valuable to examine how the findings for Business case studies align with or different from prior research on Business-related texts. Studies such as Hyland (2005) and Biber et al. (2007) have focused on the importance of audience engagement, practical applications, and situational framing in Business communication. This study reinforces those findings through the prominence of Interactive and Situated Discourse (Dimension 2) in Business case studies. Features such as interpersonal pronouns, temporal markers, and situational constructions were frequently observed, reflecting the practical and audience-focused nature of Business writing. For instance, this study builds on Hyland's (2005) work by demonstrating how Business texts actively incorporate negations and temporal adverbs to discuss organizational challenges and decision-making processes.

However, compared to previous studies on promotional or persuasive Business genres, such as advertisements or sales letters (e.g., Dos Santos, 2002), the Business case studies in this corpus showed a more restrained use of overt persuasive strategies. For example, while Interpersonal and Dialogic Discourse (Dimension 4) was more pronounced in Business case studies than in Engineering proposals, it was still relatively minimal overall. This finding contrasts with studies of Business genres that relied on dialogic language to engage audiences and negotiate meanings, which suggest that the case study format prioritizes analytical and reflective discussion over direct persuasion.

The Narrative Focus (Dimension 3) observed in Business case studies also shows similarities with previous findings by Swales and Rogers (1995) and Hyland (2000), which noted the use of narratives to contextualize organizational practices or leadership styles. However, this study adds another view by showing how narratives in Business texts are sued not only to recount events but also to draw connections between real-world scenarios and theoretical frameworks. These purposes distinguish Business case studies from promotional texts, which tended to use narratives primarily to engage the audience emotionally.

Finally, Informational Density and Elaborated Discourse (Dimension 1) was presented in Business case studies but less pronounced than in Engineering proposals. This finding reveals the need for Business texts to balance complexity and accessibility, particularly when addressing diverse audiences that include stakeholders, decision-makers, and non-specialists. This can be supported by Bhatia (1993) and Hyland (2005), who noted the hybrid nature of Business genres, which should combine analytical rigor with audience-centered clarity. Compared to Engineering

proposals, which emphasize technical precision, Business case studies employ these features selectively to ensure relatability and practical applicability.

Overall, the findings for Business case studies are consistent with prior research on Business communication but contribute additional perspectives on the specific linguistic strategies used in the case studies. By emphasizing situational engagement, stakeholder-focused language, and the integration of authentic and theoretical perspectives, Business case studies demonstrate a distinct approach that goes with the field's communicative purposes while differentiating itself from other Business genres.

5.4 Describing Dimensions

5.4.1 Dimension 1: Informational Density and Elaborated Discourse

Engineering proposals exhibited significantly higher scores on this dimension, characterized by dense noun phrases, nominalizations, and technical terms. These characteristics can be supported by Biber and Gray's (2010) findings on the centrality of informational density in technical and scientific writing. However, while informationally dense language is common across STEM fields, this study revealed the unique role of elaborated structures, such as nominalized clauses and attributive adjectives to convey precision and technical complexity specific to Engineering proposals. Compared to STEM research articles (Gray & Biber, 2012), which also emphasized dense informational content, Engineering proposals used these features to directly address practical and technical solutions rather than purely theoretical discussion.

Business case studies, on the other hand, employed informationally dense language but focused on accessibility and situational relevance. This aligns with Zhu's (2004) argument that Business writing often balances rigor with persuasive clarity, as seen in the selective use of specific terms combined with simplified structures to engage broader audiences. Unlike promotional genres in Business, which are often designed to be highly persuasive (Dos Santos, 2002), the case studies in this study reflected a more analytical tone while still maintaining audience engagement. These findings suggest that informational density in Business writing was strategically moderated to suit diverse audience groups.

5.4.2 Dimension 2: Interactive and Situated Discourse

This dimension was more prominent in Business case studies, showing the genre's emphasis on interpersonal engagement and situational analysis. Features such as 3rd person pronouns, negations, and temporal markers indicate a focus on real-world contexts and stakeholder interactions. Similar findings have been observed in Business reports and marketing texts (Hyland, 2005; Paltridge, 2012), where situational construction was used to connect with specific audiences and propose possible solutions.

By comparison, Engineering proposals demonstrated a detached and formal tone, consistent with the field's emphasis on objectivity and precision (Hyland, 2000). While situational framing was presented in some scientific and technical writing, such as environmental impact assessments (Cortes, 2004), it was less prevalent in Engineering proposals, which focused on descriptive precision over interpersonal engagement. This divergence shows how disciplinary conventions influence the deployment of situational and

interactive features, with Business texts leveraging these strategies for stakeholder engagement and Engineering texts emphasizing technical objectivity.

5.4.3. Dimension 3: Narrative Focus

Both text types incorporated narrative elements, but their purposes differed. In Business case studies, narratives were used to contextualize real-world scenarios and support situational analyses based on the functional adaptability of narratives discussed by Gardner et al. (2019). For example, narratives often linked organizational challenges with theoretical frameworks, providing actionable views for decision-making. In contrast, narratives in Engineering proposals were used to contextualize technical findings, often integrating features such as past tense constructions and quantifications to enhance coherence and technical relevance. These findings align with Gray and Biber's (2012) observations on the role of narrative in scientific and technical communication but show a greater emphasis on procedural and historical content in Engineering texts. Compared to research articles in fields like biology (Gardner et al., 2019), which may use narratives to describe experimental processes, Engineering proposals focused more on reducing technical histories and operational contexts.

5.4.4 Dimension 4: Interpersonal and Dialogic Discourse

Interpersonal and dialogic discourse was minimally presented in both text types, indicating their shared emphasis on professionalism and analytical rigor. However, Business case studies incorporated slightly more dialogic features, such as contrasting perspectives and stakeholder references, to build arguments and justify decisions. This finding follows Hyland's (2005) observations on the importance of persuasion and engagement in Business

communication. Not as usual Business promotional genres (Dos Santos, 2002), which relied heavily on overt persuasion strategies, the case studies in this corpus employed dialogic features selectively, often using hedges and boosters to balance assertiveness and caution. Engineering proposals, on the other hand, minimized dialogic features to maintain objectivity and focus on technical precision. This finding differs from some STEM-related genres, such as collaborative research reports (Staples et al., 2016), which often incorporate more dialogic elements to acknowledge interdisciplinary contributions. The relative absence of dialogic discourse in Engineering proposals reflects their narrower communicative purpose, which aimed at delivering clear technical solutions.

5.4.5 Dimension 5: Descriptive and Stative Discourse

Engineering proposals relied heavily on in this dimension to detail technical specifications and material properties. These findings are consistent with Swales' (1990) and Biber et al.'s (2011) observations on the descriptive demands of technical writing. However, this study extends prior research by identifying specific linguistic patterns, which contribute to the clarity and precision required in Engineering proposals. Compared to other genres, such as environmental science papers (Cortes, 2004), Engineering proposals demonstrated a stronger focus on static descriptions rather than dynamic processes. In contrast, Business case studies focused on dynamic processes and strategic analyses, scoring lower on this dimension. These findings align with Hyland's (2008) discussion of Business communication's emphasis on adaptability and forward-looking strategies. As opposed to promotional marketing genres, which may use descriptive discourse to show product attributes (Dos Santos, 2002), the case

studies in this study prioritized actionable recommendations and situational analyses over static descriptions.

5.5. Methodological Considerations

5.5.1 Absence of Negative Loadings for Factor Analysis

The limited presence of negative loadings for the linguistic features in this study, with the decision to filter loadings below an absolute value of 0.30 merits careful consideration. This pattern diverges from the findings of traditional MDA studies, where negative loadings are often used to represent opposing communicative functions within dimensions (e.g., Biber, 1988). In this study, the focus on features with stronger loadings (above +/- 0.30) appears to have resulted in cohesive clusters of positively correlated linguistic features, with only a few dimensions displaying minor oppositional tendencies. This methodological consideration reflects an effort to enhance the interpretability of the findings by prioritizing stronger and more robust associations. However, it also has implications for the functional interpretation of the dimensions. The absence or reduction of negative loadings suggests a limitation in capturing the full range of functional oppositions, as weaker correlations – which might include such contrasts – were excluded. Biber et al. (2007) noted that linguistic dimensions are often marked by opposing ends that reveal distinct communicative purposes, such as informational vs. interactive discourse. By having more on positively loaded features, this study emphasizes the dominant patterns in the dataset but may not fully explore the contrasts that could exist within the corpus.

Another factor contributing to the limited presence of negative loadings can be the nature of the linguistic features analyzed. The features included in this study, such as nouns, prepositions, and pronouns, are inherently non-negative in their frequencies and tended to co-occur within specific patterns of discourse. As a result, the data naturally favored the identification of positive associations rather than oppositional relationships. Similar findings were noted in Gray and Biber (2012), where the choice of features and corpus design influenced the prominence of specific loadings.

Compared to prior studies (e.g., Cortes, 2004; Staples et al., 2016), which analyzed more heterogeneous corpora across genres or disciplines, this study focused on two text types: Business case studies and Engineering proposals. This narrow scope may also contribute to the reduced variability in linguistic features, further limiting the emergence of oppositional patterns. For example, previous research on broader academic corpora often identified dimensions with contrasting functions, such as narrative vs. non-narrative or interactive vs. transactional discourse. However, the specialized and structured nature of the texts in this study may inherently prioritize cohesive linguistic strategies over functional oppositions.

Despite these limitations, the findings contribute valuable perspectives on the dominant linguistic patterns within the corpus. By emphasizing cohesive clusters of features, this study is similar to Hyland (2008), who exhibited the importance of analyzing disciplinary genres as reflective of shared communicative purposes. This approach allows for a more focused exploration of how specific gernes achieve their rhetorical objectives but shows the need for complementary analyses that could examine weaker correlations or expand the corpus scope to reveal additional contrast. Future research could address these limitations by incorporating a

broader range of genres, introducing a wider variety of linguistic features, or adjusting methodological threshold to explore more patterns. Such approaches could show how oppositional relationships manifest across diverse academic contexts and contribute to a more comprehensive understanding of linguistic variation.

5.5.2 Corpus Representativeness

The representativeness of the corpus is a critical consideration in a corpus-based study. The texts were selected to reflect Business and Engineering student writing; however, the corpus may not fully capture the diversity within these disciplines. Differences in academic levels, cultural backgrounds, and institutional contexts can influence linguistic variations. The study focused on specific text types – case studies and proposals – to analyze linguistic variation. While this targeted approach provided meaningful perspectives, it also limits the applicability of findings to other genres within these disciplines, such as research articles, reports, or promotional materials. Moreover, the corpus represented linguistic patterns at a specific time period and within a specific context. Writing conventions evolve and influenced by various standards, as well as technological advancements. Incorporating temporal and cross-cultural perspectives into future studies could provide deeper insights into the dynamic nature of disciplinary writing.

5.6 Implications

5.6.1 Implications for Writing Pedagogy

The findings note the importance of discipline-specific writing instruction, with the need for pedagogical strategies adapted to the linguistic and rhetorical demands of distinct

disciplines. For Business writing, instructors should focus on developing students' proficiency in interactive discourse and situational framing, enabling them to effectively engage with diverse audiences. In contrast, Engineering writing instruction should prioritize informational density and descriptive clarity to emphasize the construction of dense noun phrases and nominalizations to enhance the delivery of technical information.

Genre-based pedagogies, as advocated by Hyland (2008), can support students in navigating the specific conventions of their fields by having targeted instruction to discipline-specific language use. For example, in Engineering writing, technical terminology and precise descriptions are often crucial elements associated with Information Density and Elaborated Discourse (Dimension 1). Teaching students effectively use noun phrases and nominalizations (e.g., "optimization of energy efficiency" or "implementation of hydrometallurgical recycling strategies") can help them deliver complex ideas succinctly and meet the standards of professional communication. Pedagogical activities such as deconstructing sample texts to identify nominalization patterns or practicing their use in technical summaries can foster student's ability to recognizing content-dense and precise descriptions.

Similarly, in Business writing, features tied to Interactive and Situated Discourse (Dimension 2), such as temporal markers, 3rd person pronouns, and negations, were instrumental in engaging audiences and constructing situational analyses. For instance, activities that guide students to analyze stakeholder interactions using temporal markers (e.g., "recently", "persuasively") or frame contrasting perspectives with negations (e.g., "Amazon's strategy does not prioritize employee welfare") can enhance their skills to address real-world scenarios effectively. Gardner et al. (2019) emphasized the value of integrating situational

perspectives into writing instruction and suggested that targeted approaches to disciplinary needs can significantly enhance students' genre awareness and rhetorical competence. These targeted practices can make sure that Business students align their language choices with the communicative purposes of the text.

Moreover, corpus-based pedagogical tools, as suggested by Biber and Reppen (2002), can further refine writing instruction by providing data-driven perspectives on the linguistic features prevalent in specific genres. For example, instructors could use corpus analysis to identify frequent use of prepositions and attributive adjectives in Engineering proposals and design materials to help students practice constructing detailed and cohesive descriptions. Similarly, in Business writing, corpus-based exercises might focus on the most common persuasive phrases or strategies for stakeholder engagement. These approaches expose learners to authentic language patterns and help them internalize the linguistic structures they are most likely to encounter in academic and future contexts. By applying these strategies, instructors can address the linguistic and rhetorical demands of different disciplines. Providing students with opportunities to practice specific lexico-grammatical features within meaningful contexts ensures that writing instruction is not only theoretically grounded but also practically relevant to their academic and further success.

Writing programs can also blend in multimodal tasks grounded in authentic practices, incorporating the findings of this study into teaching strategies. For example, Dimension 1 (Informational Density and Elaborated Discourse), which features dense noun phrases, nominalizations, and technical descriptions, can be integrated into multimodal assignments for Engineering students. A task could involve students creating an infographic or technical video

presentation where they explain a complex engineering concept using concise, content-rich language supported by visual aids. This activity reinforces the linguistic precision required in technical communication and aligns with real-world practices where engineers need to communicate technical details clearly to diverse audiences.

Likewise, Dimension 2 (Interactive and Situated Discourse), prominent in Business case studies, could be incorporated into a multimodal task where students simulate stakeholder presentations. For instance, students could develop a PowerPoint presentation with accompanying oral commentary to analyze a business scenario. This would require the use of situational markers (e.g., temporal adverbs like “currently” or “previously”) and interpersonal language (e.g., 3rd person pronouns or references to audiences) to frame their analysis in a contextually relevant and engaging manner. Such tasks can allow students to practice adapting their language and rhetorical strategies to meet the needs of specific audiences.

Lim and Polio (2020) noted that multimodal assignments, which integrate textual, visual, and spoken elements, develop critical communication skills increasingly valued in both Business and Engineering fields. For example, using Descriptive and Stative Discourse (Dimension 5), Engineering students could create a narrated technical diagram that describes the state of a system or material property in precise detail. This task would require them to integrate stative constructions (e.g., “The material is highly conductive”) with descriptive visuals, reinforcing their ability to deliver accurate and clear technical information in a multimodal format. Paltridge’s (2004) argument complements this approach by emphasizing the socio-cultural dimensions of writing and the importance of aligning pedagogy with disciplinary norms. By embedding multimodal tasks targeted to specific dimensions and linguistic features, writing

programs can help students in learning the conventions of their respective fields and preparing them for the multimodal nature of professional communication.

5.6.2 Implications for Professional Communication

Although the analysis focused on student writing, the findings also have implications for professional communication. The distinct linguistic strategies observed in Business and Engineering writing reflect the communicative purposes of their respective professional fields. Understanding these differences can help practitioners formulate their communication strategies to meet audience expectations. For example, business professionals need to balance analytical rigor with persuasive engagement to address diverse stakeholders, while engineers should pay more attention to clarity and precision when detailing technical specifications. Crossley (2020) illustrated the importance of linguistic features such as lexical sophistication and cohesion in achieving high-quality professional communication. Therefore, incorporating discipline-specific features into professional training programs can foster the effectiveness of communication in both fields. Biber and Conrad (2005) argue for the explicit teaching of register characteristics to improve professional writing practices. Hyland and Tse (2007) further suggest that developing discipline-specific vocabularies is crucial for achieving communicative competence.

5.7 Theoretical Contributions

This study contributes to the theoretical understanding of linguistic variation by extending MDA to student writing in Business and Engineering disciplines. The identified dimensions revealed how linguistic features co-occur to fulfill specific communicative purposes,

enriching existing frameworks of genre and register analysis. The findings also described the adaptability of language use across contexts and illustrated how shared dimensions can manifest differently based on disciplinary conventions. Biber et al. (2012) argue for the inclusion of discourse-specific features to further enhance the scope of such studies. By integrating findings from Paltridge (2000) and Hyland (2008), this study demonstrates the dynamic interplay between situational variables and linguistic strategies and offers a perspective on professional variation.

5.8 Limitations and Future Research

While this study provides valuable insights into the lexico-grammatical features of student writing, it is not without limitations. The corpus was limited to student writing within two disciplines, which may not fully capture the diversity of writing practices in each one. Future research could expand the corpus to include texts from additional disciplines. Another limitation lies in the focus on two specific text types – case studies and proposals. Although these text types are representative of their respective disciplines, they do not include the full range of writing tasks that students or professionals undertake. Future research could broaden the scope to include additional types, for example, business marketing plans or engineering design reports to gain a more comprehensive understanding of disciplinary writing practices. This research would help clarify which lexico-grammatical features are globally associated with the dimensions versus features that are more closely aligned with specific genres.

Additionally, this study's reliance on quantitative methods through MDA could be complemented by qualitative approaches, such as interviews with students, instructors, or

professionals to understand the process behind linguistic choices. Combining the two methods could provide richer perspectives into the interplay between linguistic features, disciplinary conventions, and communicative purposes. Finally, future research could explore the pedagogical applications of these findings. Intervention studies, for instance, could assess how teaching specific dimensions of linguistic variation influence students' ability to meet disciplinary writing expectations. Longitudinal studies could also investigate how students develop their command of disciplinary genres over time and identify strategies to support their progression. By addressing these limitations and exploring further, we can deepen our understanding of linguistic variation in student writing and refine pedagogical practices to support student success in diverse disciplinary contexts.

Chapter 6. Conclusion

6.1 Introduction

This chapter concludes the dissertation by summarizing the findings, discussing their implications for academic writing pedagogy, and providing suggestions for future research. By applying MDA, this study investigated linguistic variations in student writing across Business case studies and Engineering proposals. The research not only identified key dimensions and associated linguistic features that reflect disciplinary conventions and communicative purposes but also provided a framework for designing effective and discipline-specific writing instruction. In what follows, the chapter revisits the broader themes introduced earlier in the dissertation and places the findings within the larger context of academic and professional communication challenges.

6.2 Summary of Findings

The study has identified five primary dimensions of linguistic variation that reveal key distinctions between Business and Engineering writing, which reflected how linguistic choices can align with their respective communicative purposes and disciplinary conventions. Dimension 1 (Informational Density and Elaborated Discourse) revealed that Engineering texts exhibit a higher degree of linguistic complexity with frequent use of nominalizations, prepositions, and complex noun phrases. These features emphasize technical precision and the need to convey detailed and information-dense content. In contrast, Business texts, while also informational, employed simpler and more accessible structures that focused on clarity over density, considering broader and often non-technical audiences.

Dimension 2 (Interactive and Situated Discourse) showed the strong emphasis in Business writing on situational and interpersonal construction. Pronouns, modals, and contextual references were frequently used to engage stakeholders and analyze authentic scenarios. In comparison, Engineering writing maintained an objective and detached tone, focusing on technical descriptions without engaging with situational framing or interpersonal discourse. Dimension 3 (Narrative Focus) revealed that both text types employed narrative elements, albeit with distinct purposes. Business texts used narratives to connect events to actionable solutions and leadership strategies, which aimed to contextualize practical applications. Conversely, Engineering proposals used narratives to recount technical processes, providing context for findings and solutions.

Furthermore, Dimension 4 (Interpersonal and Dialogic Discourse) found that Business case studies incorporated more dialogic features to evaluate decisions and argue strategies. This approach aligns with the needs of genres to address diverse audiences and build persuasive arguments. On the other hand, Engineering proposals minimized interpersonal language for an objective and formal tone to maintain technical rigor. Lastly, Dimension 5 (Descriptive and Stative Discourse) demonstrated reliance on descriptive language in both disciplines, though with differing emphases. Business texts used accessible and practical language to deliver solutions to specific challenges, while Engineering texts focused on precise technical descriptions to ensure clarity in detailing specifications and system properties.

6.3 Revisiting the Broader Context

Revisiting the key themes introduced in the previous chapter, this study argued the dynamic relationship between linguistic features, disciplinary conventions, communicative purposes, and the broader socio-academic landscape. These findings contribute to a comprehensive understanding of how student writing supports disciplinary engagement, professional preparation, and knowledge dissemination.

6.3.1 Implications for Globalized Academic Writing

The increasing globalization of higher education necessitates teaching writing strategies that are adaptable to interdisciplinary and multicultural contexts. As students and professionals operate in more diverse settings, the ability to communicate effectively across cultural and disciplinary boundaries becomes more important. Writing instructions, therefore, should give attention to linguistic variations with the flexibility in style, tone, genre, and audience expectations. The findings of this study can contribute to these needs by highlighting how linguistic strategies vary across Business and Engineering disciplines. For example, the distinct emphasis on Informational Density and Elaborated Discourse in Engineering texts showed the importance of teaching precision, technical terminology, and nominalized structures to students preparing for global technical communication. In contrast, the prominence of Interactive and Situated Discourse in Business texts points to the need for training students to incorporate interpersonal engagement and situational construction, skills critical for multicultural and interdisciplinary Business settings where diverse audiences present.

Moreover, this study illustrates how genre-specific linguistic features align with broader communicative purposes, which offered perspectives on writing instruction. For instance,

integrating narrative elements into Business writing pedagogy can help student connect practical solutions to theoretical perspectives, an approach consistent with the Narrative Focus dimension. Similarly, emphasizing descriptive and stative language in Engineering education can improve students' ability to articulate detailed specifications and system properties effectively.

By incorporating these dimensions into writing pedagogy, educators can better prepare students to navigate globalized contexts. The study's findings support the integration of corpus-based and genre-based approaches to address the distinct needs of disciplines and foster transferable skills for cross-cultural and interdisciplinary communication. As globalized academic writing continues to evolve, this research shows the need for pedagogies that address the linguistic patterns and rhetorical demands of specific fields while promoting adaptability to diverse contexts.

6.3.2 Challenges in Professional and Academic Writing

Emerging trends in education and the workplace can present both opportunities and challenges for writing instruction. The rise of interdisciplinary collaborations, technological integration, and the rapid digitization of communication require educators to reformulate traditional approaches to teaching writing. For example, interdisciplinary collaborations often necessitate flexibility in writing styles and the ability to address various audiences. A Business professional may need to communicate technical findings to non-professionals, while and an engineer might need to present project updates in a persuasive yet accessible format. Writing instruction can address these challenges by integrating exercises that simulate interdisciplinary tasks, such as reports or presentations for mixed audiences.

The increasing reliance on digital platforms for communication also poses unique challenges, as students should adapt their writing to fit different digital contexts. For instance, preparing succinct and engaging messages for professional emails, designing persuasive proposals for virtual presentations, or creating concise yet detailed content for project management tools all require different writing skills. Educators can incorporate multimodal tasks such as creating infographics, interactive reports, or video scripts to prepare students for the multimodal demands of the modern workplace.

Finally, the digitalization of communication also raises the stakes for professionalism, clarity, and tone. Miscommunication can have significant consequences in digital spaces such as email exchanges or collaborative platforms. To address this, writing programs can include modules on tone, audience adaptation, and the use of clear and structured language, helping students build the skills necessary to avoid ambiguity and foster effective professional communication. By addressing these emerging challenges with concrete strategies, writing instruction can equip students with abilities to succeed in dynamic academic and professional contexts.

6.4 Implications for Student Writing Instruction

The findings argue the importance of targeted and discipline-specific writing instruction to address the unique demands of Business and Engineering writing. At the same time, writing pedagogy need to balance these specialized needs with foundational skills applicable across disciplines. This raises questions about the role of general first-year writing courses in preparing students with transferable writing skills versus the need for discipline-specific courses later in

their academic journey. For Business education, instruction should focus on developing student's ability to craft persuasive and situationally adaptive communication, with an emphasis on audience awareness, contextual framing, and the use of evaluative language to address stakeholder needs effectively. Similarly, Engineering education should prioritize the teaching of technical language skills such as constructing dense noun phrases, employing nominalizations, and utilizing explicit references to ensure clarity and precision in technical descriptions. These discipline-specific needs show the value of Writing in the Disciplines (WID) frameworks, which advocate for writing instruction targeted to the communicative purposes of specific fields.

The dimension-based findings in this study reinforce the value of such WID approaches by offering fine-grained perspective on the linguistic and rhetorical demands of specific text types like case studies and proposals. For instructors and curriculum developers, the MDA offers a framework for understanding which linguistic features matter most in disciplinary writing and why. By giving attention to constructs such as Informational Density, Narrative Focus, and Dialogic Discourse, educators can better scaffold instruction around the communicative goals of each text type, helping students understand how linguistic choices signal precision, persuasion, or interpersonal stance. Instructors can also use these findings to design targeted feedback, model genre-specific language use, and encourage metalinguistic awareness through corpus-based writing activities. Moreover, English as Academic Purpose (EAP)/English for Specific Purposes (ESP) can integrate dimension-based teaching to prepare both L1 and L2 students for writing that aligns with real disciplinary expectations instead of relying on generic academic writing materials. This functional and data-driven perspective supports a more responsive and contextualized pedagogy, one that reflects how language actually works across different fields.

However, the increasing prevalence of artificial intelligence (AI) tools in academic contexts introduce both challenges and opportunities for writing instruction. On the one hand, AI-driven writing assistants can help students enhance their linguistic precision and stylistic choices by providing feedback on grammar, coherence, and tone. For example, AI-powered text generators can offer suggestions for improving sentence structures and generating ideas. On the other hand, reliance on AI tools raises concerns about authenticity, critical thinking, and the development of students' own writing abilities. Educators should integrate AI tools into the curriculum responsibly to make sure they are used to supplement rather than replace the writing process. Writing instruction might incorporate strategies for evaluating AI-generated content critically and adapting it to meet specific rhetorical goals, fostering students' ability to engage with AI as a collaborative tool.

Shared linguistic strategies across disciplines, such as clear and precision language, remain important for all students. General first-year writing courses can work as an introduction to these skills, preparing students to navigate a range of academic and professional writing contexts. For instance, teaching students how to adapt their tone and structure for different audiences or how to organize information clearly provides a foundation that supports both Business and Engineering communication. The integration of AI tools into first-year writing courses could focus on helping students understand the strengths and limitations of these technologies, developing their ability to critique and refine AI-generated content effectively. This approach also aligns with the goals of Writing Across the Curriculum (WAC) programs, which emphasize integrating writing instruction throughout a student's academic career. WAC programs could bridge the gap between first-year writing courses and discipline-specific need by

embedding writing instruction into subject-specific courses. For example, Engineering students could learn to refine their technical descriptions within a thermodynamic course, while Business students might practice persuasive strategies in a marketing class.

Ultimately, the findings can support a blended approach to writing pedagogy. While discipline-specific instruction addresses the specialized demands of fields like Business and Engineering, first-year writing and WAC courses ensure that students develop strong writing skills. By integrating AI literacy into these courses, writing instruction can better prepare students to meet the complex communication demands of academic and further contexts while equipping them with the adaptability required for interdisciplinary and globalized settings. Incorporating AI responsibly into the writing process can enhance students' capabilities while fostering critical thinking and creativity, and they remain active participants in the evolving landscape of written communication.

6.5 Future Research

While this study provided valuable perspectives to the understanding of linguistic variation, several areas warrant further exploration. Longitudinal studies could investigate how students' writing evolves as they transition from academic to professional contexts and provide other perspectives on the developmental trajectories of disciplinary conventions. Expanding the scope of analysis to include other disciplines, such as Medicine, Law, Humanities or Social Sciences, would offer a more comprehensive perspective on writing practices across disciplines. Additionally, research into hybrid genres, such as technical marketing materials or collaborative reports, could illuminate how writing adapts to interdisciplinary demands.

The influence of AI on education, professional writing, and corpus research also merits specific attention. In education, tools like AI-powered writing assistants are increasingly integrated into academic contexts, potentially reshaping how students approach writing tasks. Future studies could examine how these tools affect linguistic complexity, creativity, and adherence to disciplinary conventions. In terms of professional writing, AI technologies that automate technical descriptions, summarize complex information, or generate draft content could alter traditional communication practices. Understanding how writers incorporate or adapt to AI's capabilities and limitations would provide valuable perspectives on the evolving nature of communication.

For corpus research, AI presents unprecedented opportunities for data collection, analysis, and interpretation. Machine learning algorithms may enhance the annotation of linguistic features or identify emerging patterns across vast datasets and analysis process. Future studies could explore how AI-driven methodologies compare to traditional corpus approaches, particularly in discovering subtle linguistic patterns or addressing limitations such as sampling bias. Incorporating AI into corpus research could also enable real-time analysis of digital communication and capture dynamic changes in language use as they occur. Together, these considerations for future research can enrich our understanding of linguistic variation and support the development of adaptive and relevant writing pedagogy.

6.6. Conclusion

This dissertation has demonstrated the contribution of using MDA to discover linguistic variation in student writing, shedding light on the distinct communicative practices of Business

and Engineering disciplines. By identifying dimensions such as Informational Density, Interactive Discourse, and Narrative Focus, the study provided perspectives on the rhetorical and functional strategies employed by each discipline. These findings offer a framework for designing effective, discipline-specific writing instruction that aligns with the conventions and communicative purposes of Business and Engineering fields. Furthermore, this study contributes to a growing body of research on academic discourse by indicating the implications of linguistic variation for preparing students for academic success and professional communication. In doing so, it emphasizes the importance of integrating both foundational skills and disciplinary-specific strategies into writing instruction. The findings also point to the relevance of pedagogical trends, including genre-based and multimodal approaches, which help students to navigate complex and interdisciplinary writing contexts.

As the role of technology continues to expand, the study also argues the need to address challenges posed by AI and digital tools in both educational and professional communication. Having AI-driven writing technologies into pedagogy and research could enhance students' ability to engage with authentic practices and adapt to the rapid evolution of workplace communication. This consideration reflects a wider necessity for educators to balance tradition with innovation, which makes students to succeed in increasingly interconnected and technology-driven environments.

Finally, this dissertation notes the interplay between linguistic variation, disciplinary norms, and pedagogical strategies. By bridging the gap between research and practice, it offers a foundation for future studies on academic and professional writing across disciplines. The

findings show the importance of fostering adaptable, precise, and accessible communication skills that empower students to excel in diverse contexts.

References

Ansarifar, A., Shahriari, H., Staples, S., & Ghazanfari, M. (2025). A multi-dimensional analysis of thesis abstracts: Variation across academic levels and L1 backgrounds. *Journal of English for Academic Purposes*, *73*, 101465. https://doi.org/10.1016/j.jeap.2024.101465

Bean, J. C., & Melzer, D. (2021). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*. John Wiley & Sons.

Beaufort, A. (2008). *College writing and beyond: A new framework for university writing instruction*. University Press of Colorado.

Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher education*, *19*(2), 151-161. https://doi.org/10.1080/03075079412331382007

Bhatia, Vijay K., (1993): Analysing Genre-Language Use in Professional Settings, London, Longman.

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.

Biber, D. (1992). On the complexity of discourse complexity: A multidimensional analysis. *Discourse processes*, *15*(2), 133-163. https://doi.org/10.1080/01638539209544806

Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge University Press.

Biber, D. (1998). Corpus linguistics: Investigating language structure and use. *Cambridge University Press google schola*, *2*, 230-239.

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for academic purposes*, *5*(2), 97-116. https://doi.org/10.1016/j.jeap.2006.05.001

Biber, D. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.

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

Biber, D. (2019). Text-linguistic approaches to register variation. *Register studies*, *1*(1), 42-75.

Biber, D., & Conrad, S. (2005). Register variation: A corpus approach. *The handbook of discourse analysis*, 175-196. doi:10.1002/9780470753460

Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.

Biber, D., & Finegan, E. (2014). On the exploitation of computerized corpora in variation studies. In *English corpus linguistics* (pp. 204-220). Routledge.

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, *9*(1), 2-20. https://doi.org/10.1016/j.jeap.2010.01.001

Biber, D., & Gray, B. (2013). Discourse characteristics of writing and speaking task types on the TOEFL ibt® test: a lexico-grammatical analysis. *ETS Research Report Series*, *2013*(1), i-128. https://doi.org/10.1002/j.2333-8504.2013.tb02311.x

Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching?. *Studies in Second Language Acquisition*, *24*(2), 199-208. doi:10.1017/S0272263102002048

Biber, D., Connor, U., & Upton, T. (2007). Discourse analysis and corpus linguistics. *Discourse on the move: Using corpus analysis to describe discourse structure*, 1-20.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development?. *Tesol Quarterly*, *45*(1), 5-35. https://doi.org/10.5054/tq.2011.244483

Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, *37*(5), 639-668. https://doi.org/10.1093/applin/amu059

Birkenstein, C., & Graff, G. (2018). *They say/I say: The moves that matter in academic writing*. WW Norton & Company.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, *32*, 21.

Charles, M. (2003). 'This mystery…': a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes*, *2*(4), 313-326. https://doi.org/10.1016/S1475-1585(03)00048-1

Conrad, D., & Newberry, R. (2011). 24 Business communication skills: Attitudes of human resource managers versus business educators. *American Communication Journal*, *13*(1).

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for specific purposes*, *23*(4), 397-423. https://doi.org/10.1016/j.esp.2003.12.001

Crossley, S. A. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, *11*(3), 415-443. https://doi.org/10.17239/jowr-2020.11.03.01

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, *35*(2), 115-135. https://doi.org/10.1111/j.1467-9817.2010.01449.x

Dos Santos, V. P. (2002). Genre analysis of business letters of negotiation. *English for specific purposes*, *21*(2), 167-199. https://doi.org/10.1016/S0889-4906(00)00028-4

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, *38*(2), 165-193. https://doi.org/10.1093/applin/amv011

Egbert, J., & Biber, D. (2023). Key feature analysis: a simple, yet powerful method for comparing text varieties. *Corpora*, *18*(1), 121-133. http://doi.org/10.3366/cor.2023.0275

Egbert, J., Biber, D., Keller, D., & Gracheva, M. (2024). Register and the dual nature of functional correspondence: accounting for text-linguistic variation between registers, within registers, and without registers. *Corpus Linguistics and Linguistic Theory*, (0). https://doi.org/10.1515/cllt-2024-0011

Fernsten, L. A., & Reda, M. (2011). Helping students meet the challenges of academic writing. *Teaching in Higher Education*, *16*(2), 171-182. https://doi.org/10.1080/13562517.2010.507306

Ferris, D. R. (2003). Response to student writing: Implications for second language students. *Lawrence Earlbaum Associates*.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition and Communication, 32(4), 365–387. https://doi.org/10.2307/356600.

Gardner, S. (2016). A genre-instantiation approach to teaching English for specific academic purposes: Student writing in business, economics and engineering. *Writing and Pedagogy*, *8*(1), 117-144. 10.1558/wap.v8i1.27934

Gardner, S., & Nesi, H. (2013). A classification of genre families in university student writing. *Applied linguistics*, *34*(1), 25-52. https://doi.org/10.1093/applin/ams024

Gardner, S., Biber, D., & Nesi, H. (2015). MDA perspectives on Discipline and Level in the BAWE corpus. *Corpus Linguistics 2015*, *255*(229), 126.

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

Gee, J. P. (1999). Critical issues: Reading and the new literacy studies: Reframing the national academy of sciences report on reading. *Journal of literacy research*, *31*(3), 355-374. https://doi.org/10.1080/1086296990954805

Goulart, L. (2021). Register variation in L1 and L2 student writing: A multidimensional analysis. *Register Studies*, *3*(1), 115-143. https://doi.org/10.1075/rs.20012.gou

Goulart, L., Biber, D., & Reppen, R. (2022). In this essay, I will…: Examining variation of communicative purpose in student written genres. *Journal of English for Academic Purposes*, *59*, 101159. https://doi.org/10.1016/j.jeap.2022.101159

Goulart, L., Gray, B., Staples, S., Black, A., Shelton, A., Biber, D., ... & Wizner, S. (2020). Linguistic perspectives on register. *Annual Review of Linguistics*, *6*(1), 435-455. https://doi.org/10.1146/annurev-linguistics-011718-012644

Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and instruction. *Reading and writing*, *26*, 1-15. https://doi.org/10.1007/s11145-012-9395-2

Gray, B. (2019). Tagging and counting linguistic features for multi-dimensional analysis. *Multi-dimensional analysis: Research methods and current issues*, 43-66.

Gray, B., & Biber, D. (2012). Current Conceptions of Stance. In: Hyland, K., Guinda, C.S. (eds) Stance and Voice in Written Academic Genres. Palgrave Macmillan, London. https://doi.org/10.1057/9781137030825_2

Halliday, M. A. K. & Hasan R. (1976). Cohesion in English. London: Longman

Hardy, J. A., & Friginal, E. (2016). Genre variation in student writing: A multi-dimensional analysis. *Journal of English for Academic Purposes*, *22*, 119-131. https://doi.org/10.1016/j.jeap.2016.03.002

Hardy, J. A., & Römer, U. (2013). Revealing disciplinary variation in student writing: A multi-dimensional analysis of the Michigan Corpus of Upper-level Student Papers (MICUSP). *Corpora*, *8*(2), 183-207. https://doi.org/10.3366/cor.2013.0040

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, *30*(4), 461-473. https://doi.org/10.1093/applin/amp048

Hyland, K. (2000). Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language awareness*, *9*(4), 179-197. https://doi.org/10.1080/09658410008667145

Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of second language writing*, *13*(2), 133-151. https://doi.org/10.1016/j.jslw.2004.02.001

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse studies*, *7*(2), 173-192. https://doi.org/10.1177/1461445605050365

Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching*, *41*(4), 543-562. https://doi.org/10.1017/S0261444808005235

Hyland, K. (2012). Bundles in academic discourse. *Annual review of applied linguistics*, *32*, 150-169. https://doi.org/10.1017/S0267190512000037

Hyland, K. (2014). English for academic purposes. In *The Routledge companion to English studies* (pp. 392-404). Routledge.

Hyland, K. (2019). *Second language writing*. Cambridge university press.

Hyland, K. (2022). English for specific purposes: What is it and where is it taking us?. *Esp Today*, *10*(2), 202-220.

Hyland, K. L. (2009). Writing in the disciplines: Research evidence for specificity. *Taiwan international esp journal*.

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"?. *TESOL quarterly*, *41*(2), 235-253. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

Hyland, K., & Tse, P. (2012). 'She has received many honours': Identity construction in article bio statements. *Journal of English for academic purposes*, *11*(2), 155-165. https://doi.org/10.1016/j.jeap.2012.01.001

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

Kim, Y. L., Crawford, W. J., & McDonough, K. (2025). The role of communicative purpose in describing and interpreting lexico-grammatical variation in L2 writing. *Journal of Second Language Writing*, *67*, 101176. https://doi.org/10.1016/j.jslw.2025.101176

Kyle, K., & Crossley, S. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, *34*(4), 513-535. https://doi.org/10.1177/0265532217712554

Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal*, *102*(2), 333-349. https://doi.org/10.1111/modl.12468

Lan, G., Liu, Q., & Staples, S. (2019). Grammatical complexity:'What does it mean'and 'so what'for L2 writing classrooms?. *Journal of Second Language Writing*, *46*, 100673. https://doi.org/10.1016/j.jslw.2019.100673

Lancaster, Z. (2016). Expressing stance in undergraduate writing: Discipline-specific and general qualities. *Journal of English for Academic Purposes*, *23*, 16-30. https://doi.org/10.1016/j.jeap.2016.05.006

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied linguistics*, *27*(4), 590-619. https://doi.org/10.1093/applin/aml029

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, *30*(4), 579-589. https://doi.org/10.1093/applin/amp043

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in higher education*, *23*(2), 157-172. https://doi.org/10.1080/03075079812331380364

Leki, I. (2001). “A narrow thinking system”: Nonnative-English-speaking students in group projects across the curriculum. *TESOL quarterly*, *35*(1), 39-67. https://doi.org/10.2307/3587859

Lim, J., & Polio, C. (2020). Multimodal assignments in higher education: Implications for multimodal writing tasks for L2 writers. *Journal of Second Language Writing*, *47*, 100713. https://doi.org/10.1016/j.jslw.2020.100713

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International journal of corpus linguistics*, *15*(4), 474-496. https://doi.org/10.1075/ijcl.15.4.02lu

McDonough, K., Uludag, P. &, Neumann, H. (2021). Assessing case study exams in a business course: What role does language play? *Business and Professional Communication Quarterly, 82*(2), 116-134. https://doi.org/10.1177%2F23294906211012398

Moon, J. A. (2006). *Learning journals: A handbook for reflective practice and professional development*. Routledge. https://doi.org/10.4324/9780203969212

Nathan, P. (2013). Academic writing in the business school: The genre of the business case report. *Journal of English for Academic Purposes*, *12*(1), 57-68. https://doi.org/10.1016/j.jeap.2012.11.003

Nation, I. S., & Nation, I. S. P. (2001). *Learning vocabulary in another language* (Vol. 10, pp. 126-132). Cambridge: Cambridge university press.

Nesi, H. (2008, September). A multidimensional analysis of student writing across levels and disciplines. In *Taking the measure of applied linguistics: Proceedings of the BAAL Annual Conference, University of Swansea* (pp. 11-13).

Nesi, H. & S. Gardner (2012). Genres across the disciplines: Student writing in higher education. Cambridge: Cambridge University Press.

Neumann, H., Padden, N., & McDonough, K. (2019). Beyond English language proficiency scores: Understanding the academic performance of international undergraduate students during the first year of study. *Higher Education Research & Development*, *38*(2), 324-338. https://doi.org/10.1080/07294360.2018.1522621

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics*, *30*(4), 555-578. https://doi.org/10.1093/applin/amp044

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics*, *30*(4), 555-578. https://doi.org/10.1093/applin/amp044

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics*, *30*(4), 555-578. https://doi.org/10.1093/applin/amp044

Moran, K. E. (2013). *Exploring undergraduate disciplinary writing: Expectations and evidence in psychology and chemistry*. Georgia State University.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied linguistics*, *24*(4), 492-518. https://doi.org/10.1093/applin/24.4.492

Ortega, L. (2012). Interlanguage complexity. *Linguistic complexity: Second language acquisition, indigenization, contact*, *13*, 127.

Paltridge, B. (2000). Genre knowledge and teaching professional communication. *IEEE transactions on professional communication*, *43*(4), 397-401. 10.1109/47.888814

Paltridge, B. (2004). Academic writing. *Language teaching*, *37*(2), 87-105. doi:10.1017/S0261444804002216

Paltridge, B. (2012). Genre and English for specific purposes. *The handbook of English for specific purposes*, 347-366. https://doi.org/10.1002/9781118339855.ch18

Polio, C. G. (1997). Measures of linguistic accuracy in second language writing research. *Language learning*, *47*(1), 101-143. https://doi.org/10.1111/0023-8333.31997003

Sardinha, T. B., & Pinto, M. V. (Eds.). (2014). *Multi-dimensional analysis, 25 years on: A tribute to Douglas Biber* (Vol. 60). John Benjamins Publishing Company.

Staples, S., & JoEtta, A. (2022). Comparing the situational and linguistic characteristics of first year writing and engineering writing. *Applied Corpus Linguistics*, *2*(3), 100031. https://doi.org/10.1016/j.acorp.2022.100031

Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, *32*, 17-35. https://doi.org/10.1016/j.jslw.2016.02.002

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, *33*(2), 149-183. https://doi.org/10.1177/0741088316631527

Storch, N. (2019). Collaborative writing. *Language Teaching*, *52*(1), 40-59. https://doi.org/10.1017/S0261444818000320

Swales, J. M. (1990). *Genre analysis*. Cambridge university press.

Swales, J. M., & Feak, C. B. (2004). *Academic writing for graduate students: Essential tasks and skills* (Vol. 1). Ann Arbor, MI: University of Michigan Press.

Swales, J. M., & Feak, C. B. (2004). *Academic writing for graduate students: Essential tasks and skills* (Vol. 1). Ann Arbor, MI: University of Michigan Press.

Swales, J. M., & Rogers, P. S. (1995). Discourse and the projection of corporate culture: The mission statement. *Discourse & Society*, *6*(2), 223-242. https://doi.org/10.1177/0957926595006002005

Tasker, D. G. (2019). *Situational and linguistic variation in undergraduate English-department student writing* (Doctoral dissertation, Northern Arizona University).

Thompson, G., & Thetela, P. (1995). The sound of one hand clapping: The management of interaction in written discourse. *Text & Talk*, *15*(1), 103-128. https://doi.org/10.1515/text.1.1995.15.1.103

Walton, D. (2003). *Relevance in argumentation*. Routledge. https://doi.org/10.4324/9781410609441

Weigle, S. C., & Friginal, E. (2015). Linguistic dimensions of impromptu test essays compared with successful student disciplinary writing: Effects of language background, topic, and L2 proficiency. *Journal of English for Academic Purposes*, *18*, 25-39. https://doi.org/10.1016/j.jeap.2015.03.006

Wheeler, E., & McDonald, R. L. (2000). Writing in engineering courses. *Journal of Engineering Education*, *89*(4), 481-486. https://doi.org/10.1002/j.2168-9830.2000.tb00555.x

Wolfe-Quintero, K., S. Inagaki & H. Y. Kim (1998). Second language development in writing: Measures of fluency, accuracy and complexity. Honolulu, HI: Second Language Teaching & Curriculum Center, University of Hawai'i.

Zhang, Y. (2023). A Multidimensional Analysis of Language Use in English Argumentative Essays: An Evidence From Comparable Corpora. *SAGE Open*, *13*(3), 21582440231197088. https://doi.org/10.1177/21582440231197088

Zhu, W. (2004). Faculty views on the importance of writing, the nature of academic writing, and teaching and responding to writing in the disciplines. *Journal of second language Writing*, *13*(1), 29-48. https://doi.org/10.1016/j.jslw.2004.04.004

Zhu, W. (2004). Writing in business courses: An analysis of assignment types, their characteristics, and required skills. *English for Specific purposes*, *23*(2), 111-135. https://doi.org/10.1016/S0889-4906(02)00046-7

Appendices

Appendix A. Tag descriptions (Biber, 1988)

: +clp + + + colon + clause punctuation

; +clp + + + semi-colon + clause punctuation

? +clp + + + question mark + clause punctuation

! +clp + + + exclamation mark + clause punctuation

, + + + + comma

- + + + + dash

" + + + + double quote mark

, + + + + single quote mark

(+ + + + left parenthesis

) + + + + right parenthesis

$ + + + + dollar sign

% + + + + percent sign

&fa + + + + formula symbols

&fw + + + + foreign word

abl + ++ + pre-qualifier {rather, such)

abn + + + + pre-quantifier {all, half)

abx + + + + pre-quantifier / double conjunction (both)

ap + + + + post -determiner (many, more, most, only, other, own, same, ...)

aps + + + + (others)

at + + + + singular indefinite article (a, an)

ati + + + + singular definite article {the, no)

cc + + + + coordinating conjunction (and, but, or)

cc +cls + + + coordinating conjunction + clausal connector

cc +phrs + + + coordinating conjunction + phrasal connector

cc" + + + + multi-word coordinating conjunction (as well as)

cc + +neg + + coordinating conjunction + + negation (nor)

cd + + + + cardinal number (2, 3, 4, two, three, four, hundred, ...)

cd +date++ + cardinal number + date (year only)

cd 1 + + + + cardinal number (I, one)

cd 1s + + + + cardinal number (ones)

cds + + + + cardinal plural (tens, hundreds, thousands)

od + + + + ordinal number (1st, 2nd, first, second, ...)

cs +cnd + + + subordinating conjunction + conditional (if I unless)

cs +con + + + subordinating conjunction + concessive (although, though)

cs +cos + + + subordinating conjunction + causative (because)

cs +who + + + subordinating conjunction + WH word (whether)

cs +sub + + + subordinating conjunction + other (as, except, until, ...)

cs” + + + + multi-word subordinating conjunction (in that, so that, ...)

dt +dem + + + determiner + demonstrative (this, that, these, those modifying N)

dt +pdem + + + determiner + demonstrative pronoun (this, that, these, those)

dti + + + + singular or plural determiner (any, enough, some)

dt + + + + other singular determiner (another, each)

dtx + + ++ determiner/double conjunction (either)

ex + pex + + + existential there

in + + + + preposition

in +ppvb + + + preposition + prepositional verb (account for, join in,…)

pi + + + preposition + place marker (above, behind, beside, ...)

in " + + + + multi-word perposition (as to, away from, instead of, ...)

in +strn + + + preposition + stranded

jj +atrb + + + adjective + attributive function

jj +atrb + +xvbg + adjective + attributive function + + -ing form

jj +atrb + +xvbn + adjective + attributive function + + past participle form

jj +pred + + + adjective + predicative function

jj + + + + adjective + indeterminate function

jjb +atrb + + + attributive-only adjective + attributive (chief. entire)

jjr +atrb + + + comparative adjective + attributive function

jjr +pred + + + comparative adjective + predicative function

jjt +atrb + + + superlative adjective + attributive function

All modal forms can be marked as O in Field 5 {e.g., md +prd + + +0) to show that they are contracted forms (e.g., ‘ll, 've)

md+nec+++

md+pos+++

md +prd +++

md"++pmd"++

modal + necessity (ought, should, must)

modal + possibility (can, may, might, could)

modal + prediction (will, would, shall)

modal + + multi-word periphrastic modal (e.g., be going to)

nn++++ singular common noun

nn+nom +++ singular noun + nominalization

nvbg + + +xvbg + singular noun + + + -ing form

nn + + +xvbn + singular noun + + + past participle form

nns + + + + plural common noun

nns +nom + + + plural noun + nominalization

nnu + + + + unit of measurement (lb, kg, ...)

np + + + + singular proper noun

nps + + + + plural proper noun

npl + + + + locative noun

npt + + + + singular titular noun

npts + + + + plural titular noun

nr + + + + singular adverbial noun (east, west, today, home, ...)

nrs + + + + plural adverbial noun

In the following pronoun tags, be careful of the difference between the number 1, used to mark first person, and the letter I (i.e., lower case L), used to mark reflexives.

ppla +pp1 + + + first person subject pronoun + first person pronoun

ppla +pp1 + + +0 first person subject pronoun + lst person pro. + contracted

pplo +pp1 + + + first person object pronoun + first person pronoun

pp$ +pp1 + + + possessive determiner + first person pronoun {my, our)

ppl +pp1 + + + singular reflexive pronoun + first person pronoun (myself

ppls +ppI + + + plural reflexive pronoun + first person pronoun (ourselves)

pp2 +pp2 + + + second person pronoun + second person pronoun

pp$ +pp2 + + + possessive determiner + second person pronoun (your)

ppl +pp2 + + + singular reflexive pronoun + second person pronoun (yourself

pp3a +pp3 + + + third person subject pronoun + third person personal pronoun

pp3o +pp3 + + + third person object pronoun + third person personal pronoun

pp3 +pp3 + + +0 third person pronoun + 3rd person personal pro. + contracted

pp$ +pp3 +++ possessive + 3rd pers. personal pro. (his, her, their)

ppl +pp3 + + + sg. reflexive pronoun + 3rd pers. personal pro. (her/himself)

ppls +pp3 + + + pi. reflexive pronoun + 3rd pers. personal pro. (themselves)

pp3 +it + + + third person pronoun + third person impersonal pronoun (it)

pp$ +it + + + possessive determiner + third person impersonal pronoun (its)

pp$$ + + + + possessive pronoun (mine, yours, ...)

pn" + + + + multi-word nominal pronoun (no one, ...)

pn++++

All adverb forms can be marked as splt in Field 3 (e.g., rb +amp +splt + +) to

indicate that the adverb occurs within the auxiliary (e.g., they've probably been looking...).

rb + + + + general adverb

rb" + + + + multi-word adverb (at last, in general)

rb +cnj + + + adverb + conjunct (however, therefore, thus, ...)

rb + +neg + + neither

rb +amp + + + adverb + amplifier (absolutely, completely, entirely, ...)

rb +down + + + adverb + down toner (nearly, only, merely, ...)

rb +emph + + + adverb + emphatic (just, really, 50, ...)

rb +hdg + + + adverb + hedge (almost, maybe, ...)

rb" +hdg" + + + multi-word adverb + hedge (kind of, sort of)

rb +phrv + + + adverb + phrasal verb (get in, wrap up, ...)

rb +pi + + + adverb + place marker (abroad, ahead, far, upstream, ...)

rb +tm + + + adverb + time marker (afterwards, again, immediately, ...)

rb +dspt + + + adverb + discourse particle (anyway, well, ...)

rbr + + + + comparative adverb (better, quicker)

rbr +tm + + + comparative adverb + time marker (earlier, later, sooner, ...)

rn +pi + + + nominal adverb + place marker (here, there)

m +tm + + + nominal adverb + time marker (now, then)

m +dspt + + + nominal adverb + discourse particle (now)

rp + + + + adverbial particle (back, in, round, Up, ...)

rp +pi + + + adverbial particle + place marker (away, behind, out, ...)

tht +jcmp + + + that as dependent clause head + adjective complement

tht +ncmp + + + that- as dependent clause head + noun complement

tht +vcmp + + + that as dependent clause head + verb complement

tht +rel + + + that as dependent clause head + relative clause

ql + + + + qualifier + (as, less, more, too)

ql +amp + + + qualifier + amplifier (very)

ql +emph + + + qualifier + emphatic (most)

tht +reI +obj + + that as dep. clause head + relative clause + object position

tht +rei +subj + + that as dep. clause head + relative clause + subject position

to++++ infinitive marker

to" ++++ multi-word infinitive marker (in order to)

uh + + + + interjection/filler (hey, oh, ok, yes, erm ...)

vb + + + + base form of verb, excluding verbs in infinitive clauses (uninflected present tense, imperative)

vb + + +xvbn + base form of verb + + + past participle form (e.g., cut, hit, hurt,)

vb +be +aux + + base form of verb + be + auxiliary verb

vb +be +vrb + + base form of verb + be + main verb

vb +bem +aux + + verb + am + auxiliary verb

vb +bem +aux + +0 verb + am + auxiliary verb + + contracted ('m)

vb +bem +vrb + + verb + am + main verb

vb +bem +vrb + +0 verb + am + main verb + + contracted ('m)

vb +ber +aux + + verb + are + auxiliary verb

vb +ber +aux + +0 verb + are + auxiliary verb + + contracted ('re)

vb +ber +vrb + + verb + are + main verb

vb +ber +vrb + +0 verb + are + main verb + + contracted ('re)

vb +do +aux + + verb + do + auxiliary verb

vb +do +vrb + + verb + do + main verb

vb +hv +aux + + verb + have + auxiliary verb

vb +hv +aux + +0 verb + have + auxiliary verb + + contracted ('ve)

vb +hv +vrb + + verb + have + main verb

vb +hv +vrb + +0 verb + have + main verb + + contracted ('ye)

vb +seem + + + base form of verb + seem / appear

vb +vprv + + + base form of verb + private verb (believe, feel, think, ...)

vb +vprv +thtO + + base form of verb + private verb + that deletion **

vb +vpub + + + base form of verb + public verb (assert, complain, say, ...)

vb +vpub +thtO + + base form of verb + public verb + that deletion **

vb +vsua + + + base form of verb + suasive verb (ask, command, insist. ...)

All past tense verb forms (excluding were, was, did, had) are marked as either xvbn or xvbnx; only the sequences with xvbn are listed below.

vbd + + +xvbn + past tense verb + + + past participle form

vbd +bed +aux + + past tense verb + were + auxiliary verb

vbd +bed +vrb + + past tense verb + were + main verb

vbd +bedz +aux + + past tense verb + was + auxiliary verb

vbd +bedz +vrb + + pa.c:;t tense verb + was + main verb

vbd +dod +aux + + past tense verb + did + auxiliary verb

vbd +dod +vrb + + past tense verb + did + main verb

vbd +hvd +aux + + past tense verb + had + auxiliary verb

vbd +hvd +vrb + + past tense verb + had + main verb

vbd +seem + +xvbn + past tense verb + seem/appear

vbd +vprv + +xvbn + past tense + private verb (believe, feel, think, ...)

vbd +vprv +thtO +xvbn + past tense + private verb + that deletion **

vbd +vpub + +xvbn + past tense + public verb (assert, complain, say, ...)

vbd +v pub +thtO +xvbn + past tense + public verb + that deletion **

vbd +v sua + +xvbn + past tense + suasive verb (ask, command~ insist~ ...)

** the tag thtO marks the occurrence of a following that complement clause when the complementizer that has been deleted.

vbg + + +xvbg + present progressive verb + + + -ing form

vbg +beg + +xvbg + present progressive verb + being

vbg +beg +aux +xvbg + present progressive verb + being + auxiliary verb

vbg +hvg + +xvbg + present progressive verb + having

vbg +vprv + +xvbg + pres. prog. + private verb (believe, feel, think, ...)

vbg +vprv +thtO +xvbg + present progressive + private verb + that deletion **

vbg +v pub + +xvbg + pres. prog. + public verb (assert, complain, say, ...)

vbg +vpub +thtO +xvbg + present progressive + public verb + that deletion **

vbg +v sua + +xvbg + pres. prog. + suasive verb (ask, command, insist, ...)

vwbg + + +xvbg + present progressive postnominal modifier

vwbg +beg + +xvbg + present progressive postnominal modifier + being

vwbg +hvg + +xvbg + present progressive postnominal modifier + having

vwbg +vprv + +xvbg + present prog. postnom. modifier + private verb

vwbg +vpub + +xvbg + present prog. postnom. modifier + public verb

vbi + + + + base form of verb in infinitive clause

vbi +vprv + + + infinitive verb + private verb (believe, feel, think, ...)

vbi +vprv +thtO + + infinitive verb + private verb + that deletion **

vbi +vpub + + + infinitive verb + public verb (assert, complain, say, ...)

vbi +vpub +thtO + + infinitive verb + public verb + that deletion **

vbi +v sua + + + infinitive verb + suasive verb (ask, command, insist, ...)

vbz + + + + 3rd person singular verb

vbz +bez +aux + + 3rd person sg. verb + is + auxiliary verb

vbz +bez +aux + +0 3rd person sg. + is + auxiliary verb. + + contracted (IS)

vbz +bez +vrb + + 3rd person sg. verb + is + main verb

vbz +bez +vrb + -..0 3rd person sg. + is + main verb + + contracted (IS)

vbz +doz +aux + + 3rd person sg. verb + does + auxiliary verb

vbz +doz +vrb + + 3rd person sg. verb + does + main verb

vbz +hvz +aux + + 3rd person sg. verb + has + auxiliary verb

vbz +hvz +vrb + + 3rd person sg. verb + has + main verb

vbz +seem + + + 3rd person sg. verb + seem/ appear

vbz +vprv + + + 3rd person sg. + private verb (believe, feel, think, ...)

vbz +vprv +thtO + + 3rd person sg. + private verb + that deletion **

vbz +vpub + + + 3rd person sg. + public verb (assert, complain, say, ...)

vbz +vpub +thtO + + 3rd person sg. + public verb + that deletion **

vbz +v sua + + + 3rd person sg. + suasive verb (ask, command, insist, ...)

** the tag thtO marks the occurrence of a following that complement clause

the complmentizer that has been deleted.

All perfect aspect verb forms and passive verb forms are marked as either xvbn or xvbnx; only the sequences with xvbn are listed below.

vprf + + +xvbn + perfect aspect verb + + + past participle form

vprf + +thtO +xvbn + perfect aspect verb + + that deletion **

vprf +ben +aux +xvbn + perfect aspect verb + been + auxiliary verb

vprf +ben +vrb +xvbn + perfect aspect verb + been + main verb

vpsv + +agls +xvbn + main clause passive verb + + agentless passive

vpsv + +by +xvbn + main clause passive verb + + by passive

vwbn + + +xvbn + passive postnominal modifier + + + past participle form

vwbn +vprv + +xvbn + passive postnominal modifier + private verb

vwbn +vpub + +xvbn + passive postnominal modifier + public verb

vwbn +v sua + +xvbn + passive postnominal modifier + suasive verb

wdt +who + + + WH determiner + WH word (what, whatever, whichever, ...)

wdt +who +whcl + + WH determiner + WH word + WH clause

wdt +who +whq + + WH determiner + WH word + WH question

whp +reI +obj + + WH pronoun + relative clause + object position

whp +reI +pied + + WH pronoun + relative clause + object position with

prepositional fronting ('pied piping')

whp +rel +subj + + WH pronoun + relative clause + subject position

whp +who + + + WH pronoun + WH word (not a relative clause)

whp +who +whq + + WH pronoun + WH word + WH question

wrb +who + + + WH adverb (how, when, where, ...) + WH word

wrb +who +whcl + + WH adverb + \NH word + WH clause

wrb +who +whq + + WH adverb + WH word + WH question

x not + +not + + not + + negation

x not + +not + ...0 not + + negation + + contracted form (n't)

xvbn + + +xvbn + past participle form --indeterminate grammatical function

xvbg + + +xvbg + present participle form --indeterminate grammatical function

zz + + + + letter of the alphabet

Appendix B. R script

```
library(corrplot)
cor_matirx <- cor(data %>% select(where(is.numeric)), use = “pairwise.complete.obs”)
corrplot(cor_matrix, method = “circle”, type = “upper”, tl.cex = 0.8)
high_corr_vars <- which(abs(cor_matrix) > 0.8 & abs(cor_matrix) < 1, arr.ind = TRUE)
print(high_corr_vars)

factor_data <- data %>% select(where(is.numeric))
factor_data_imputed <- factor_data %>%
        mutate(across(everything(), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))
zero_variance_vars <- factor_data_imputed %>%
        select(where(~ var(.) == 0)) %>%
        colnames()
factor_data <- factor_data %>%
        select(-all_of(zero_variance_vars))
print(paste(“Removed zero-variance variables:”, paste(zero_variance_vars, collapse = “, “)))
```

```
factor_data <- factor_data %>%
    mutate(across(everything(), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))

library(psych)
kmo_results <- KMO(factor_data)
print(kmo_results)

fa_results <- fa(r = factor_data, nfactors = 5, rotate = "promax")
print(fa_results)

dimension_scores <- fa_results$scores
data <- cbind(data, dimension_scores)

#compute mean and standard deviation of dimension scores for each discipline
dimension_summary <- data %>%
    group_by(discipline) %>%
    summarize(across(starts_with("Dimension"), list(mean = mean, sd = sd), .names =
    "{.col}_{.fn}"))
print(dimension_summary)

#statistical comparison
anova_results <- data %>%
    pivot_longer(starts_with("Dimension"), names_to = "Dimension", values_to =
    "Score") %>%
    group_by(Dimension) %>%
    summarize(anova_p = summary(aov(Score ~ discipline, data =
    cur_data())))[[1]][["Pr(>F)"]][1])

anova_results <- updated_data %>%
```

```
        pivot_longer(starts_with("MR"), names_to = "Dimension", values_to = "Score") %>%
        group_by(Dimension) %>%
        summarize(
        p_value = summary(aov(Score~ discipline, data = cur_data()))[[1]][["Pr(>F)"]][1]
        )

anova_results <- anova_results %>%
        mutate(
        bonferroni_corrected = p_value * nrow(anova_results),
        significant = bonferroni_corrected < 0.05
        )
print(anova_results)
```