

Linguistics Specificity of Universities Prospectuses: A Corpus-Based Genre Analysis

Aiman Qaiser

*MPhil Graduate, Department of English, KUST, Kohat
aimenqk50@gmail.com*

Syed Shujaat Ali

*Assistant Professor, Kohat University of Science and Technology, Kohat
shujaatali@kust.edu.pk*

Naveed-Ur-Rehman Khattak

*Associate Professor, Air University, Islamabad
naveed.khattak@au.edu.pk*

Abstract

This study investigates the specific lexical choices found in university prospectuses while examining the linguistic structures that define their textual characteristics and determining the genre type by comparing it with similar text types to understand the unique language features of these documents. The necessary prospectuses were downloaded from the universities' online platforms to achieve this objective. A collection of text samples was assembled by examining various sections of university websites to create a representative corpus. The section titled Introduction, About Us, Vice Chancellor Messages, Vision, Mission, and Facilities. A quantitative approach was employed; lexical and linguistic features were quantified to compare with registers' specific dimensions scores as studied in Biber, and for framing the textual dimensions. Functions of lexical specificity were interpreted to describe the workings of tools of lexical specificity. The data was analyzed with the help of software like MAT and AntConc. The investigation of the data utilized theoretical frameworks, including Hyland's (2009) specificity model alongside Biber's (1993) multidimensional analysis model. The research indicates that university prospectuses frequently employ lexical items that predominantly include self-reference terms, hedging expressions, directional words, reporting verbs, and lexical bundles. Through Biber's (1993) multidimensional analysis, the research findings indicate that university prospectuses align most closely with the learned exposition text type, encompassing registers such as official documents, press reports, and academic prose, thus establishing it as a distinct genre.

Keywords: Lexical specificity, textual dimension, University prospectuses, Multidimensional analysis, Corpus-based analysis

Introduction

A prospectus is described as "a printed booklet advertising a school or university to potential parents or students or giving details of a share offer for the benefit of investors" in the Online Oxford Dictionary. Institutional prospectuses serve as the primary medium for reaching a large audience. University prospectus language has been studied from the viewpoints of CDA, discourse analysis, genre analysis, and marketization in higher education. From a linguistic and cross-language standpoint, fewer studies have been done. Critical discourse analysis (CDA) was used by Hui (2009) to examine the wording in the opening pages of an online university brochure. Online prospectuses utilize promotional language to serve various purposes, including informing, luring, and convincing students. Previous research believed it to be a promotional tool. Such variety may be seen in online university prospectuses. Due to the many purposes, the language of the online prospectus serves to evoke interest in students. An essential part of informing the campus community and the broader public about the university's overall mission is available on the institution's website. In this study, the language features of university prospectuses on the websites of Pakistani institutions are examined, and the significant roles played by the linguistic details are noted.

Statement of the Research Problem

With the help of its linguistic components, such as words or phrases, language communicates meaning through its relationships. Corpus-based genre analysis across various domains, including weblogs, newspapers, business emails, lecture language, and job application letters, has already been carried out. The multiple perspectives from which they have tried to view the issue have been finding functional properties of the corpus, using the present tense, analyzing along micro- and macro-levels, exploring the communicative functions, and exploring the socio-cultural features of writing styles from the past. Almost the same research has been carried out on prospectuses. Still, their focus was on the abovementioned perspectives, except for examining the lexical items used to build public image and advertisement. Besides, those studies that analyzed online university prospectuses textually focused on critical discourse analysis (CDA) and the exploration of non-narrative features without touching aspects like textual analysis, lexical specificity, and genre classification. To fill this gap left by the earlier studies and to term this research unique, this research focuses on studying university prospectuses to find out

the specifics of the language used, the lexical items or the identities that are present, the textual dimensions, the genre types of the text, and how these lexical items are used as tools for promotion and advertisement. This research intends to address the gap identified in earlier studies.

Research Objectives

The basic aim of this research is to find the lexical specificity evident in the messages of the Vice-chancellors, which are available on the concerned universities' websites, and then to explore text dimensions present in these messages. A secondary objective is to recognize the 'prospectus' as a distinct genre within the field of education.

Research Questions

1. What is the lexical specificity reflected in the writing of the Vice-Chancellors' messages available on the websites of those universities?
2. What are the textual dimensions present in the Vice-Chancellors' messages on the websites of their universities?
3. How does the 'prospectus' qualify to be recognized as a distinct genre within the field of education?

Literature Review

A genre of text refers to the way language is composed or communicated. Texts are categorized into various genres based on the purpose of language use. A genre is a text with specific main characteristics regarding structure, style, and content, which writers emphasize when composing in that genre. Bhatia's (2014) analysis of genre involves the examination of organized linguistics conducted in standardized academic or professional contexts. According to Dudley-Evans (2002), the objective of genre analysis is educational, as it provides an adjustable framework based on genre analysis that guides us in the structure, appropriate use of language, and order for writing and discussing assignments. Out of the scientific writing courses for international graduate students, Cortes (2011) studied a couple of English genres. One semester involved corpus-based research, while the other was limited to genre. In his 2002 paper, Bhatia analyzes various aspects of the genre, stating that it either addresses the regulated exchange of

ideas within a discourse community or creates and organizes different teaching approaches and strategies used in classroom activities. The term 'genre' encompasses spoken as well as written discourse. The written genre comprises written texts and scripts, whereas the spoken genre involves verbal communication and conversation, engaging speakers and interlocutors.

The Internet is a web framework where various documents are identified through URLs, allowing users to access them online. Utilizing the Internet is very common when teaching English for academic purposes. Yang (2013) focused on a university website and compiled a corpus from the 'About Us' section. Askehave and Nielsen (2005) examined digital genres from a multimodal perspective.

Several characteristics, such as quantity, quality, representation, irretrievability, certainty, growth, and documentation, can describe a corpus. Conrad (2008) identified three key traits of corpus-based research. He raised questions about the significance of corpus-based research for language educators, arguing that integrating linguistic corpora and computer technology greatly influences language learning. Hussain (2013) conducted a corpus-based study on instant messaging in Pakistan. Hyatt (2005) presented a corpus-based analysis of evaluation editorials for master's level assignments, focusing on the conventional elements of language used in teacher critiques. Irshad and Anwar (2021) explored a multidimensional analysis of written discourse in Pakistani English. According to Khattak and Shehzad (2018), the term "specificity" in academic English refers to how academics and researchers use language to develop the specialized "discourse competencies" necessary for acceptance into a particular "discourse community." Amjad and Shakir (2014) investigated linguistic features in online university prospectuses that convey information.

Nawaz and Shakir (2014) conducted a corpus-based study on the importance of non-narrative features in the online university prospectuses of Pakistan. Zahra and Shakir (2015) used a corpus to compare explicit and context-dependent references in these prospectuses. Arshad (2014) identified the grammatical and communicative roles of the semantic category of verbs, specifically private and persuasive verbs, in her corpus-based study of private verbs in an online university prospectus. Cortes (2004) defined lexical bundles as extended collocations, sequences of three or more words that statistically co-occur in a register, or chunks of language of varying lengths with remarkable formal, functional, and statistical attributes (Saadatara, Kiany, &

Talebzadeh, 2023). The study by Zhou and Zhang (2024) looks at the present use of the pronoun 'we' to create our identity, express our views in academic circles, and change our use in the future. In their research, Choudhri, Zahra, and Shehzad (2022) examined the linguistic specificity in the introduction sections of selected Pakistani academic books through a corpus-based approach. Gul and Khattak (2020) conducted a corpus-based analysis of Vice Chancellors' Messages on the Websites of Universities in Pakistan, utilizing linguistic specificity.

Research carried out in earlier studies has explored the use of corpus-based methods for analyzing genres, and various genres have been explored in different ways. The researchers used various corpus techniques to analyze other kinds of data. Research done on the prospectuses of universities concerning language has been carried out through genre analysis. However, the 'about us' section, statements from the Vice Chancellor or Rector, and sections like 'mission' or 'vision' and 'introductions' have not been the focus of such study. This official document has significance and is composed in a formal tone. The current study aims to investigate these sections of online university prospectuses found on university websites. The objective is to identify the distinct linguistic features in the specified sections of university prospectuses.

Research Methodology

For methodology and materials, this study used a quantitative corpus-based research design to investigate the linguistic specificity of prospectuses of different universities. Only institutions with the most discipline and students were selected for the study. The theoretical framework adopted was Hyland's (2009) model of lexical specificity, according to which linguistic features can be classified into six dimensions, namely directives, self-mention, hedging, lexical bundles, and reporting verbs. In self-mention, 1st person pronouns like 'we,' 'our,' and 'I' and 2nd person pronouns like 'you' and 'you're' are used to establish an authorial presence and engage the reader. Imperatives and obligation modal verbs (must, should) are used to instruct the reader to take specific actions. The use of verbs like develop, offer, study, etc. is done to report on the work of others and convey the author's attitude towards the information reported. Recurrent sequences of words called lexical bundles (e.g., "state of the art," "wide range of") are used to convey specific meanings and functions. These linguistic features are examined to identify the lexical specificity of prospectuses of universities and to probe their textual dimensions, genre type, and the use of language for promotion.

For corpus creation, it comprised 25 prospectuses downloaded from the official websites of the Pakistani universities approved by the Higher Education Commission (HEC) of Pakistan. The technique used to select the prospectuses was a simple random technique. The researcher carefully selected and examined a limited selection of texts to build a sample corpus for this study. University prospectuses are downloaded in PDF format from their websites, and the relevant sections were copied. Only the relevant sections of the prospectuses, i.e., Introduction, About Us, Vice Chancellor Messages, Vision, Mission, and Facilities, were chosen as part of the corpus.

The 'specificity model' developed by Hyland (2009) was applied to data analysis. The data was analyzed using corpus analysis tools and software, including AntConc (3.4.4) and MAT. Frequency lists and concordance lines were generated to identify and examine the linguistic features of interest. Following the analysis, the data were analyzed and discussed regarding the results of several sample studies in each of the lexical categories described above.

Data Analysis and Discussion

The following is an in-depth analysis of the data and a discussion of the findings:

Self-Mention

According to Hyland's (2009) 'Lexical Specificity,' the first category is Self-Mention. According to this category, writers introduce themselves by using specific pronouns, intruding into the text, and establishing a connection with the piece's arguments and its target audience. According to Wu et al. (2015), 'Self-mention' is a strong rhetorical technique for building authorial identity in study papers. Based on the authors' self-mention strategies, there are three methods to measure an author's identity: the detached self (they and them), the individual self (I and me), and the collective self (we, our, us).

Table 1*Pronoun Usage Frequency in the Corpus*

| Sr no | Pronouns | Types of pronouns | Frequency | Percentage used in the corpus |
|-------|----------|---------------------------------|-----------|-------------------------------|
| 1. | I | 1 st Person Singular | 81 | 0.19% |
| 2. | Me | 1 st Person Singular | 10 | 0.023% |
| 3. | My | 1 st Person Singular | 21 | 0.05% |
| 4. | We | 1 st Person Plural | 100 | 0.23% |
| 5. | You | 2 nd Person | 100 | 0.23% |
| 6. | Your | 2 nd Person | 65 | 0.15% |
| 7. | Our | 3 rd Person | 100 | 0.23% |
| 8. | Us | 3 rd Person | 18 | 0.041% |
| 9. | Them | 3 rd Person | 61 | 0.14% |
| 10. | They | 3 rd Person | 32 | 0.074% |

Table 1 indicates that the most frequently used pronouns out of all the pronouns have been ‘we’ as 1st person plural pronoun, ‘you’ as 2nd person pronoun, and ‘your’ as 3rd person pronoun. They have been used a hundred times, and their percentage is 23%. The second most frequently occurring pronoun is ‘I,’ which is the 1st person singular pronoun and occurs 81 times in the corpus. The pronoun ‘me’ is the least used as a 1st person singular pronoun and has occurred only 10 times in the corpus.

The pronouns "we," "you," and "our" are used the most frequently compared to the other pronouns. The word "we" in the message signifies that the authority, speaking on behalf of their institution, wants to be close to and express support for their group. "We" are referred to collectively, and when it is used in phrases like "we invite you to explore campus" or "we help them," it conveys an open invitation to the audience to benefit from the staff members' unrestricted assistance. The author uses " We " to make their points to the reader more general. This instills a sense of community in the reader, causing them to react to the text as though it were written specifically for them and fulfilling the writing's duty of expressing the 'collective self' (Wu et al. 2015). The second-person personal pronoun "you" is categorized as indicating the

presence of a conversation between the communicators. As mentioned in the sentence "...we invite you to explore....," using "you" in the direct approach to the intended readers helps establish a link of shared interest. Similarly, the text tries to persuade by promising that "you will feel pride." The third-person possessive pronoun's plural form is represented by the lexeme "our." Another attempt by the author to convey the participation of employees or students in the intended message is the term "our." The libraries also provide the students with free access to a world of books through digital library access," the statement reads in this case. The pronoun "our" is employed to denote ownership by the current students and to guarantee the applicant's right to unrestricted use of the online library.

The first-person pronoun "I" is another in the universities' prospectuses and is used often throughout the corpus. Since the university is an institution in and of itself, "I" here can be used by any institution's authority, such as the chancellor, vice chancellor, registrar, chairman, director of the department, or any other office. Public relations or admission offices typically prepare prospectuses. The current corpus individual self is called "I" in the following situations (Wu et al., 2015).

The pronoun "Your" appears the third-highest number of instances in the corpus. It's a way of saying something belongs to you or the person you're speaking to. Because it is the second-person possessive pronoun, the writer formed a direct relationship with the reader and made him an integral part of his message. The writer assumes that the message's recipient is a member of the institute by using the phrase "So I wish you success in your years at SZABIST," which connects with the recipient psychologically. According to Hyland (2009), introducing oneself in a book allows the author to claim their work and express their opinions to the audience, revealing their viewpoint. First and third-person pronouns are frequently used in university prospectuses, according to an analysis of the self-mentioning in the corpus of their messages: first person (plural) "we," "you," and "our" are used 0.23% of the time, first person (singular) "I," 0.19% of the time, and second person "your," 0.15% of the time. Therefore, the usage of these pronouns relates to the significance of the individual and collective self in the prospectuses for universities that are posted on university websites.

Directives

According to Hyland (2009), the second type includes directives. A directive is defined as "something that serves to direct, guide, and usually impel towards an action or goal" in the Merriam-Webster dictionary. Derivatives often start with action verbs and have many imperative phrases. Examples are inviting, demanding, recommending, begging, requesting, expressing a wish, permitting, etc. English has two ways to communicate directives: obligation modals and imperatives. These two categories of instructions are examined and addressed in two distinct tables.

Table 2

Directives Usage Frequency in the Corpus

| Sr no | Directives | Frequency | Percentage used in the corpus |
|-------|------------|-----------|-------------------------------|
| 1. | Consider | 02 | 5.88% |
| 2. | Must | 23 | 67.64% |
| 3. | Note | 02 | 5.88% |
| 4. | Ought | 01 | 2.94% |
| 5. | Should | 14 | 41.47% |
| 6. | Could be | 01 | 2.94% |

Table 2 reveals the frequency and percentage of directives used in the corpus. The most frequently used directive is 'must,' which occurs 23 times and represents 67.64% of the total directives used. This shows that the corpus uses a strong tone of obligation or necessity. The second most common directive is 'should,' which occurs 14 times, making up 41.47% of the total directives. The least frequently used directives are 'consider,' 'note,' 'ought,' and 'could be,' each occurring only once or twice. These directives are the reason for the relatively smaller percentages, ranging from 2.94% to 5.88%.

The modal verb "must" indicate that a specific action is required. The word appears most frequently in the text (0.053%) for directions. The word's lexical precision is supposed to convey the author's steadfast adherence to the rules. The verb "must" have an imperative connotation.

The sample corpus's rules and regulations section is where it is primarily located. According to Hyland (2009), there are three basic ways that imperatives interact with their readers: physically, cognitively, and textually. According to Hyland (2009), textual actions direct the reader to a different section of the text or a related text.

Table 3

Imperatives Usage Frequency in the Corpus

| Sr no | Imperatives | Frequency | Percentage in the corpus |
|-------|-------------|-----------|--------------------------|
| 1. | Study | 57 | 0.13% |
| 2. | Prepare | 9 | 0.02% |
| 3. | Plan | 17 | 0.04% |
| 4. | Note | 1 | 0.002% |
| 5. | List | 2 | 0.004% |
| 6. | Inform | 1 | 0.002% |
| 7. | Copy | 1 | 0.002% |

The above table presents the frequency and percentage of imperatives that are used in the corpus. The most frequently used imperative is 'study,' which occurs 57 times and makes 0.13% of the total words in the corpus. This suggests that the corpus lays stress on academic instruction or guidance. The second and third most used imperatives are 'plan,' which occurs 17 times, i.e., 0.04% of the total, and 'prepare,' which appears 9 times, making up 0.02% of the total, respectively. The least frequently used imperatives are 'note,' 'list,' 'inform,' and 'copy,' each occurring only once and representing a negligible percentage (0.002%) in the whole corpus.

Physical Act

According to Hyland (2009), a physical act asks someone to perform something. Verbs that direct the reader to take specific action in the real world are known as physical action verbs. The verbs add, attach, assemble, assign, and many more are often used in the current corpus to denote physical actions.

Table 4*Imperatives Usage Frequency in the Corpus*

| Sr no | Imperative | Frequency | Percentage |
|-------|------------|-----------|------------|
| 1. | Prepare | 9 | 0.02% |
| 2. | Make | 57 | 0.13% |
| 3. | Maintain | 10 | 0.023% |
| 4. | Inform | 1 | 0.002% |
| 5. | Facilitate | 26 | 0.06% |
| 6. | Establish | 6 | 0.01% |
| 7. | Direct | 11 | 0.025% |
| 8. | Close | 11 | 0.025% |
| 9. | Center | 61 | 0.14% |
| 10. | Carry | 5 | 0.011% |
| 11. | Apply | 14 | 0.032% |
| 12. | Advice | 13 | 0.30% |
| 13. | Add | 5 | 0.011% |
| 14. | Open | 36 | 0.08% |
| 15. | Manage | 2 | 0.004% |
| 16. | Balance | 5 | 0.011% |

This table presents the frequency and percentage of imperatives that are used in the corpus. The most frequently used imperatives are 'center,' which occurs 61 times and makes up 0.14% of the total corpus, 'make,' which occurs 57 times, i.e., 0.13%, and 'prepare,' which does not occur as much as it happened in the previous table but 'open' which occurs 36, i.e., 0.08% occurs more frequently, indicating a focus on actions related to positioning, creating, and initiating processes. The table also shows some imperatives related to management, organization, and guidance: 'manage,' 'balance,' 'maintain,' and 'direct.' The imperatives that occur the least are 'manage,' 'list,' not present in this table, 'inform,' and others occur only once or just a few times, standing for negligible percentages (0.002% to 0.011%) in the corpus.

Cognitive Act

Readers are taught how to understand an argument through cognitive acts that immerse them in the text and prompt them to note, concede, or reflect on a particular argument or assertion.

Table 5

Cognitive Act Usage Frequency in the Corpus

| Sr no | Cognitive act | Frequency | Percentage |
|-------|---------------|-----------|------------|
| 1. | Stake | 1 | 0.002% |
| 2. | Review | 3 | 0.006% |
| 3. | Purpose | 35 | 0.081% |
| 4. | Program | 37 | 0.085% |
| 5. | Observe | 1 | 0.002% |
| 6. | Process | 44 | 0.10% |
| 7. | Evaluate | 1 | 0.002% |
| 8. | Develop | 42 | 0.09% |
| 9. | Assume | 3 | 0.006% |

This table presents the frequency and percentage of cognitive acts used in the corpus. The most frequently used cognitive acts are ‘process,’ which occurs 44 times, i.e., 0.10%; ‘develop,’ 42 times, i.e., 0.09%; and ‘program,’ which appears 37 times, i.e., 0.085%. These cognitive acts emphasize mental operations about planning, growth, and execution. The table also reveals some cognitive acts of analysis, assessment, and consideration, such as ‘evaluate,’ ‘review,’ and ‘purpose,’ which occur less frequently. The least commonly used cognitive acts are ‘stake,’ ‘observe,’ and ‘evaluate,’ each occurring only once, representing a negligible percentage (0.002%) in the whole corpus.

Various frequencies of these imperatives were found across the present corpus. These three imperatives were closely analyzed, and the findings indicated that words like "center," "make," and "open," which had frequencies of 0.14%, 0.13%, and 0.8%, respectively, are indicative of physical activities. The university prospectuses in the current study utilize

imperatives in a polite tone to urge readers to choose a particular course of action; an imperative is followed by a kind reminder to engage in any physical activity. Additionally, the corpus analysis shows that several often-used phrases, such as "program," "process," and "purpose," which have frequencies of 0.085%, 0.10%, and 0.081% respectively, suggest cognitive activities. The analysis of imperatives in the current corpus led to the discovery of a few examples of textual actions. The terms "plan" and "study," with percentages of 0.04% and 0.13%, respectively, indicate textual behaviors. Therefore, directions are one of their distinguishing characteristics, as shown by examining the many forms of imperatives included in university prospectuses.

Reporting Verbs

Reporting verbs make the category number 4 of Hyland's (2009) model of linguistic distinctiveness. The use of reporting verbs in academic writing is to discuss work done by other people. Reporting verbs are used in academic writing to discuss other people's work. Based on the category of speech they stand for; reporting verbs may be used in past and present tenses. Reporting verbs can be employed in the present or past tense, depending on the type of speech they reflect. It employs the past tense to describe previous events and the present tense to describe ongoing activities. The frequency with which reporting verbs are used in the universities' prospectuses is displayed in the table below.

Table 6

Reporting Verb Usage Frequency in the Corpus

| Sr no | Reporting verbs | Frequency | Percentage | Sr no | Reporting Verbs | Frequency | Percentage |
|-------|-----------------|-----------|------------|-------|-----------------|-----------|------------|
| 1. | Acknowledge | 1 | 0.002% | 15. | Mention | 3 | 0.007% |
| 2. | Added | 3 | 0.007% | 16. | Offered | 12 | 0.027% |
| 3. | Admitted | 9 | 0.02% | 17. | Propose | 1 | 0.002% |
| 4. | Allowed | 9 | 0.02% | 18. | Realize | 2 | 0.005% |
| 5. | Announced | 1 | 0.002% | 19. | Recommended | 3 | 0.007% |

| | | | | | | | |
|-----|------------|----|--------|-----|----------|---|--------|
| 6. | Asked | 1 | 0.002% | 20. | Report | 7 | 0.016% |
| 7. | Blessed | 1 | 0.002% | 21. | Reported | 2 | 0.005% |
| 8. | Considered | 6 | 0.014% | 22. | Request | 3 | 0.007% |
| 9. | Develop | 32 | 0.074% | 23. | Said | 4 | 0.009% |
| 10. | Discuss | 1 | 0.002% | 24. | Say | 2 | 0.005% |
| 11. | Explain | 1 | 0.002% | 25. | Show | 1 | 0.002% |

This table shows how often and commonly reporting verbs appear in the corpus. The most frequently used reporting verbs are 'develop,' which occurs 32 times, amounting to 0.074%, 'admitted' and 'allowed,' which occur 9 times each, amounting to 0.02%, and 'offered,' which appears 12 times, amounting to 0.027% of the total. These reporting verbs emphasize communication concerning growth, acknowledgment, and permission. The table also reveals some reporting verbs for speech, thought, and action: 'say,' 'said,' 'report,' and 'recommended.' The least frequently used reporting verbs are 'acknowledge,' 'announced,' 'asked,' and 'blessed,' in addition to some occurring only once, representing a negligible percentage (0.002%) of the corpus.

The university's prospectuses commonly utilize various reporting verbs in their communications, according to the analysis of reporting verbs in the current study. These verbs include develop, offered, admitted, and allowed, which are used 0.074%, 0.027%, and 0.02% more frequently. Hyland (2009) claims that reporting verbs employed in soft subjects, such as the humanities and social sciences, vary from those used in challenging disciplines. To make their arguments more powerful and persuasive, the reporting verbs employed in this study are comparable to those used in hard sciences, i.e., "offered and allowed." Universities' prospectuses of communications frequently employ reporting verbs like those used in the hard sciences and are thus a distinctive linguistic element of this form of writing.

Lexical Bundles

Hyland (2009) identifies lexical bundles as a key category. These bundles are words commonly used together in English, often called 'chunks.' They are also described by Biber

(1991) as "the most frequent recurring sequences of words in a given register." Based on their structural patterns and roles in speech, one may distinguish between distinct sorts of these bundles. Additionally, some studies have demonstrated that lexical bundles differ in discourse roles (such as conveying stance, discourse organization, or referential meanings) (Hyland, 2008).

Table 7

Lexical 4-Word Bundles Usage Frequency in the Corpus

| Sr No. | Type | Frequency | Percentage |
|---------------|--------------------------------------|------------------|-------------------|
| 1 | ‘state of the art’ | 29 | 0.067% |
| 2 | ‘of engineering and technology’ | 11 | 0.025% |
| 3 | ‘of the university of’ | 11 | 0.025% |
| 4 | ‘the higher education commission’ | 10 | 0.023% |
| 5 | ‘the university of Turbat’ | 10 | 0.023% |
| 6 | ‘a wide range of’ | 9 | 0.021% |
| 7 | ‘is one of the’ | 9 | 0.021% |
| 8 | ‘the university of Baluchistan’ | 9 | 0.021% |
| 9 | ‘the university of Sindh’ | 9 | 0.021% |
| 10 | ‘university of engineering and’ | 9 | 0.021% |
| 11 | ‘by the higher education’ | 8 | 0.018% |
| 12 | ‘chancellor of the university’ | 8 | 0.018% |
| 13 | ‘higher education commission’ | 8 | 0.018% |
| 14 | ‘keeping in view the’ | 8 | 0.018% |
| 15 | ‘male and female students’ | 8 | 0.018% |
| 16 | ‘Mehran university of engineering’ | 8 | 0.018% |
| 17 | ‘the university of education’ | 8 | 0.018% |
| 18 | ‘activities at appropriate national’ | 7 | 0.016% |
| 19 | ‘and promotion of its’ | 7 | 0.016% |
| 20 | ‘at appropriate national and’ | 7 | 0.016% |
| 21 | ‘campus of the university’ | 7 | 0.016% |

| | | | |
|----|--|---|--------|
| 22 | 'education in the country' | 7 | 0.016% |
| 23 | 'its activities at appropriate' | 7 | 0.016% |
| 24 | 'mandate of the club' | 7 | 0.016% |
| 25 | 'national and international forums' | 7 | 0.016% |
| 26 | 'of its activities at' | 7 | 0.016% |
| 27 | 'of the club includes' | 7 | 0.016% |
| 28 | 'one of the most' | 7 | 0.016% |
| 29 | 'promotion of its activities' | 7 | 0.016% |
| 30 | 'teaching learning and research' | 7 | 0.016% |
| 31 | 'and staff of the' | 6 | 0.014% |
| 32 | 'cost of Rs billion' | 6 | 0.014% |
| 33 | 'education commission of Pakistan' | 6 | 0.014% |
| 34 | 'engineering sciences and technology' | 6 | 0.014% |
| 35 | 'establishment of fata university' | 6 | 0.014% |
| 36 | 'facility to the students' | 6 | 0.014% |
| 37 | 'higher education commission of' | 6 | 0.014% |
| 38 | 'higher learning in the' | 6 | 0.014% |
| 39 | 'institution of higher learning' | 6 | 0.014% |
| 40 | 'on first come first' | 6 | 0.014% |
| 41 | 'short span of time' | 6 | 0.014% |
| 42 | 'the faculty and staff' | 6 | 0.014% |
| 43 | 'the students and faculty' | 6 | 0.014% |
| 44 | 'the students and staff' | 6 | 0.014% |
| 45 | 'the vice chancellor started' | 6 | 0.014% |
| 46 | 'and co-curricular activities' | 5 | 0.012% |
| 47 | 'as well as in' | 5 | 0.012% |
| 48 | 'education commission HEC Pakistan' | 5 | 0.012% |
| 49 | 'fata university was approved' | 5 | 0.012% |

| | | | |
|-----------|-------------------------------------|---|--------|
| 50 | 'library online information center' | 5 | 0.012% |
| 51 | 'medical and health sciences' | 5 | 0.012% |
| 52 | 'networking and promotion of' | 5 | 0.012% |
| 53 | 'social networking and promotion' | 5 | 0.012% |
| 54 | 'students of the university' | 5 | 0.012% |
| 55 | 'the top universities of' | 5 | 0.012% |
| 56 | 'a large number of' | 4 | 0.009% |
| 57 | 'a university of international' | 4 | 0.009% |
| 58 | 'administered tribal areas fata' | 4 | 0.009% |
| 59 | 'aim is to provide' | 4 | 0.009% |
| 60 | 'all walks of life' | 4 | 0.009% |

This table presents the frequency and percentage of 4-word lexical bundles used in the corpus. The most frequently used 4-word bundle is 'state of the art,' which occurs 29 times, i.e., 0.067%, and emphasizes innovation and modernity. Some other well-noted examples of 4-word bundles are 'of engineering and technology,' which occurs 11 times, i.e., 0.025%, 'of the university of,' which occurs 11 times, i.e., 0.025%, and 'the higher education commission,' which occurs 10 times, i.e., 0.023%. These words strongly emphasize educational institutions, technology, and higher education governance. The table includes some 4-word lexical bundles related to university life, research, and administration: ' university of engineering and,' 'chancellor of the university,' and 'teaching learning and research.' Although they occur less frequently, these bundles suggest a tone of academic and administrative activities. The least often used 4-word bundles come out to be a university of international', 'administered tribal areas face,' and 'aim is to provide,' each of which occurs 4 times, i.e., 0.009% of the total and represents a relatively minor percentage of the total such bundles used in the corpus.

Conclusion

In response to the first and second research questions, and hence to find out the lexical specificity reflected in the messages of the Vice Chancellors on the websites of the universities, and the textual dimensions present in the messages of the Vice-chancellors, the language of the

university prospectus was examined through Hyland's model of specificity, focusing on various linguistic features that highlight its specificity. For instance, the use of self-referential language in these documents often includes first-person pronouns like 'I,' 'we,' and 'our,' as well as second-person pronouns such as 'you' and 'your,' and third-person pronouns like 'they' and 'them.' The findings indicate that the pronoun I typically represents authority figures such as the Vice-Chancellor, Registrar, Dean, or Provost. When 'I' is used in the context of university prospectuses, it is not merely subjective, as Public Relations or Admissions Offices usually craft these documents. The pronoun 'we' also conveys authority on behalf of a group and fosters a sense of solidarity. The use of 'you' serves a persuasive purpose, reassuring the reader, while 'your' directly engages the reader, making them feel included in the message. The pronoun 'they' adopts a more objective stance, mainly when outlining instructions or explaining rules and regulations.

Another defining characteristic of this writing style is the use of directives to instruct the reader to take specific actions. Terms like must and should frequently appear in university prospectuses. The term should is often found in sections detailing rules, regulations, or responsibilities, while it must emphasize the necessity of specific actions, categorizing them as cognitive acts.

Additionally, reporting verbs are a notable linguistic feature in this corpus, with commonly used verbs including develop, offer, and study. The verb develop indicates cognitive acts, reflecting the author's attitude toward the reported information and serving as a persuasive tool to reassure the readers that not only is ensuring their academic success their goal, but also inculcation of moral values in them. Lexical bundles are unique in their usage across various discourses and play a vital role in language study. The phrase "wide range of" indicates a collection of diverse items. "Is one of the" highlights something specific within a sentence that stands out from others. "Male and female students" addresses the context where gender discrimination is relevant, considering societal expectations. "All walks of life" refers to different professions and social strata. The university prospectus serves as an official document packed with information. It is an official statement that informs students or the public about a specific topic, just like a press release. These official announcements are often found on websites like news outlets. The method of dissemination is through broadcasting, and websites facilitate easy

access. In response to the last research question about whether the 'prospectus' can be recognized as a distinct genre within the field of education, the university prospectus is both an official document and a form of academic writing. The text is presented in written form, as it can be read on websites and is published similarly to a press release or report.

References

- Amjad, T., & Shakir, A. (2014). Study of information generating linguistic features in online university prospectuses. *Research on Humanities and Social Sciences*, 4(25), 122-128. <https://www.iiste.org/Journals/index.php/RHSS/article/viewFile/17562/17760>
- Arshad, A., & Shakir, A. (2014). Corpus-based study of private and public verbs in online university prospectus. *Corpus*, 4(25), 24-28. <https://www.iiste.org/Journals/index.php/RHSS/article/view/17549>
- Askehave, I., & Nielsen, A. E. (2005, January). What are the characteristics of digital genres?- Genre theory from a multi-modal perspective. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences* (pp. 98a-98a). IEEE.
- Bhatia, V. (2002). Applied genre analysis: A multi-perspective model. *Ibérica: Revista de la Asociación Europea de Lenguas para fines específicos (AELFE)*, (4), 3-19. doi: 10.1075/iberia.4.02bha
- Bhatia, V. K. (2014). *Analysing genre: Language use in professional settings*. Routledge.
- Biber, D. (1991). *Variation across speech and writing*. Cambridge university press. doi: <https://doi.org/10.1017/CBO9780511621024>
- Biber, D. (1993). Using register-diversified corpora for general language studies. *Computational Linguistics*, 19(2), 219-241. doi: 10.3115/981574.981576
- Choudhri, S. (2021). A Corpus-based Study of Analyzing the Lexical Specificity in the Preface Section of the Selected Pakistani Academic Textbooks. *University of Chitral Journal of Linguistics and Literature*, 5(II), 268-294. DOI: <https://doi.org/10.33195/t4gmws87>
- Conrad, S. (2008). Corpus variety: Corpus linguistics, language variation, and language teaching. In *How to use corpora in language teaching* (pp. 67-85). John Benjamins Publishing Company.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23, 397-423. DOI: 10.1016/j.esp.2003.12.001

- Cortes, V. (2011). Genre analysis in the academic writing class: with or without corpora?. *Quaderns de Filologia-Estudis Lingüístics*, 16, 65-80.
- Dudley-Evans, T. (2002). Genre analysis: An approach to text analysis for ESP. In *Advances in written text analysis* (pp. 233-242). Routledge.
- Gul, H., & Khattak, N. (2020). A corpus-based genre analysis of linguistic specificity of VC messages on university websites in Pakistan. *Erevna Journal of Linguistics and Literature*, 4(2), 13-34. (No DOI or URL available)
- Hui, K. L. (2009). *Higher education in a globalized market: A comparative discourse study of university prospectuses in Hong Kong and the United Kingdom* [Master's thesis]. University of Edinburgh. <http://hdl.handle.net/1842/3631>
- Hussain, M. N. (2013). *Language of text messages: A corpus-based linguistic analysis of SMS in Pakistan* [Doctoral dissertation]. International Islamic University Islamabad. <http://142.54.178.187:9060/xmlui/handle/123456789/3544>
- Hyatt, D. F. (2005). 'Yes, a very good point!' A critical genre analysis of a corpus of feedback commentaries on Master of Education assignments. *Teaching in Higher Education*, 10(3), 339-353. DOI: 10.1080/13562510500122292
- Hyland, K. (2008). 'Small bits of textual material': A discourse analysis of Swales' writing. *English for Specific Purposes*, 27(2), 143-160. DOI: 10.1016/j.esp.2007.10.001
- Hyland, K. (2009). Writing in the disciplines: Research evidence for specificity. *Taiwan International ESP Journal*, 1(1), 5-22. (No DOI or URL available)
- Irshad, I., & Anwar, B. (2021). A multidimensional analysis of Pakistani English written discourse. *Linguistics and Literature Review*, 7(1), 69-82. DOI: 10.32350/llr.71.06
- Khattak, N. U. R., & Shehzad, W. (2018). A corpus-based study of specificity of academic verbs in introduction sections of PhD dissertations in English studies in Pakistan. *Dialogue*, 14(1), 63-74. https://www.qurtuba.edu.pk/thedialogue/The%20Dialogue/14_1/06-63-%2074,%20Naveed,%20AWKUM.pdf

- Nawaz, N., & Shakir, A. (2014). Significance of non-narrative features in online university prospectus of Pakistan: A corpus-based investigation. *Significance*, 4(19). <https://iiste.org/Journals/index.php/RHSS/article/view/15772>
- Saadatara, A., Kiany, G., & Talebzadeh, H. (2023). Bundles to beat the band in high-stakes tests: Pedagogical applications of an exploratory investigation of lexical bundles across band scores of the IELTS writing component. *Journal of English for Academic Purposes*, 61, 101208. DOI: 10.1016/j.jeap.2022.101208
- Yang, W. (2013). 'Why choose us?' texts in university websites: A genre analysis. *Taiwan International ESP Journal*, 5(1), 45-80.
- Wu, G., & Zhu, Y. (2015). Self-mention and authorial identity construction in English and Chinese research articles. *Linguistics and the Human Sciences*, 10(2), 133-158. DOI: 10.1558/lhs.v10.2.28557
- Zahra, F. T., & Shakir, A. (2015). A corpus-based study of explicit versus situation-dependent reference in online university prospectus. *Research on Humanities and Social Sciences*, 5(7), 128-141. <https://www.iiste.org/Journals/index.php/RHSS/article/view/21524>
- Zhou, Y., & Zhang, Y. (2024). Visualization analysis of research on self-mention WE: Current status and future trends (2008-2022). *Arab World English Journal*, 15(3), 283-296. DOI: <https://dx.doi.org/10.24093/awej/vol15no3.17>