Liberal Journal of Language & Literature Review
Print ISSN: 3006-5887
Online ISSN: 3006-5895
https://llrjournal.com/index.php/11

Linguistic Choices and Relevance of Essays Generated Through AI Tools: A Corpus-Assisted Analysis

[1]Hafiza Suman Khan
[2]Saqlain Hassan

[1]M.Phil. Corpus Linguistics, Riphah International University, Islamabad. Email: ksumank876@gmail.com
[2]Assistant Professor, Department of English Linguistics and Literature, Riphah International University, Pakistan. Email: saqlain.hassan@riphah.edu.pk
ORCID: https://orcid.org/0000-0001-8861-6800

Liberal Journal of Language & Literature Review
Print ISSN: 3006-5887
Online ISSN: 3006-5895

Abstract

This study aims to identify AI-based linguistic choices in various text types within the academic genre and to investigate the relevance of AI-generated text in performing the communicative function across informative and persuasive texts, employing a corpus-assisted mixed-methods research paradigm. The corpus for this study comprises AI-generated essays, classified into informative and persuasive text types within the academic genre. It uses Lexical Specificity by Hyland (2009) as a theoretical framework to synthesise the findings. The study's findings reveal that although AI generates genre-compatible content, a significant linguistic deviation from norms has been found in both the selected text types of the academic genre. However, AI tools generate more relevant content in informative writing than in persuasive writing, as evidenced by text-type-compatible linguistic choices. Thus, the study concludes on the reliability of the AI-generated content in informative writing within the academic domain. The study's findings could help developers enhance the outputs of such tools in the future, enabling them to be utilised by students in the academic domain, particularly in linguistics.

Keywords*:* Artificial Intelligence, Corpus Linguistics, Linguistic Choices, Academics, Genre Analysis

1. INTRODUCTION

Academic writing, a cornerstone of scholarly communication, relies on precise and contextually rich language to convey complex ideas and research findings effectively (Swales & Feak, 2012). It requires a good linguistic aptitude and intense exposure to the subject matter, rhetoric, and target audience (Hyland, 2009). The most significant form of writing is founded on the understanding of the context in terms of which writers create systematic arguments that are aware of his or her field of disciplines. In this way, academic writing courses place a higher level of focus on three key areas: critical thinking and analysis, practical research, and the synthesis of ideas. The other areas include the development of academic integrity in terms of originality and the absence of plagiarism in assignments (Teng & Wang, 2023).

Historically, this has been left to the human writers whose linguistic decisions are

shaped by their years of experience and knowledge of disciplinary norms. Nevertheless, the sphere of academic writing has undergone significant changes with the rapid evolution of artificial intelligence (AI). The AIs have now been improved on by large Language Models (LLMs), including GPT (Radford et al., 2019) and LaMDA (Bard) (Thoppilan et al., 2022).

The AI-based tools of this type are typically created through Natural Language Processing (NLP), which is trained on large amounts of human-written text (Nazari et al., 2021; Ginting & Barella, 2022; Perkins, 2023). Being trained on large datasets through thorough pretraining processes, they possess the capability to generate human-like text in a variety of genres and communicative functions (Shahriar & Hayawi, 2023).

Genre plays an important role in both academic and professional communication, serving as a framework for understanding and producing texts within specific contexts (Brommer, 2019). Swales (1990) defines genre as "a class of communicative events, the members of which share some set of communicative purposes which are recognized by the expert members of the parent discourse community" (p. 58) Although Swales describes genre as a "fuzzy concept" in his masterpiece on English for Special Purposes and academic language (Swales, 1990, p. 33) and views genres as instruments of social action, he nonetheless recognizes that "there may be pedagogical value in sensitizing students to rhetorical effects and to the rhetorical structures that tend to recur in the genre-specific texts" (Swales, 1990, p. 213).

This is because recurring features of genre, such as textual structure, content, and style, create specific expectations regarding the task participants and the type of social activity expected to undertake. Furthermore, these tasks align with the broader expectations of their discourse community, with genre serving as the primary means of indexing and symbolically codifying these expectations (Matsagouras & Tsiplakou, 202X). In this regard, Hyland (2002) highlighted the concept of disciplinary variation, asserting that students should be taught the specific literacy skills required for each discipline, which he labelled as language specificity.

Specificity in language refers to the distinguishing features that differentiate various entities, nouns, or referents within a given context (Hyland, 2008). Hyland (2008)

emphasises that this variability forms the foundation of disciplinary specificity. Consequently, each discourse community develops its own communicative norms, which are of significant interest due to their connection to the concept of genre analysis (Hyland, 2009). Thus, genre competence—the ability to recognise and generate texts according to established conventions—is important for both human writers and AI systems like ChatGPT, which generate texts based on statistical probabilities (Zhao et al., 2023).

Therefore, as AI continues to develop in creating increasingly human-like texts, a crucial question arises: how do these models perform across different text types within the academic genre, and do they conform to normative genre definitions? Given the potential of LLMs like ChatGPT in various fields, there is a pressing need to investigate the linguistic features of AI and the relevance of its generated content, so that the strengths and weaknesses in possible applications across different text types are understood. Thus, this corpus-assisted study aims to examine the linguistic choices of AI, such as verbs and pronouns, across various text types within academic essays and assess the relevance of AI-generated text (in the form of essays) in fulfilling its communicative function.

The study examines the language use and communicative effectiveness of AI-generated texts across various academic genres. It will determine how AI, using large language models such as GPT and LaMDA, selects linguistic units, including verbs, nouns, and stylistic patterns, and whether these choices are consistent with the conventions and communicative intentions of various types of academic texts. Although AI is increasingly being incorporated into academic writing, there is limited knowledge of its capacity to meet the expectations of the discipline and to fulfil the desired communicative roles. The research helps to fill this gap, as it examines AI-generated academic texts, their potential, and their limitations. The results can also be utilised to enhance AI writing tools, inform teachers and researchers, and inform the design of effective AI-enhanced academic writing practices.

2. LITERATURE REVIEW

The corpus-based research aims to analyse the linguistic decisions of AI tools, including pronouns and modal auxiliaries, across various text types in the academic genre, and determine how well AI-generated text can be applied to the purposes for

which it is intended. In this way, this section reviews past studies to understand the effectiveness of AI-generated text in various fields, such as academia. It synthesises the findings, with particular attention to the gaps and opportunities in future studies.

2.1 Artificial Intelligence

Artificial Intelligence (AI) was first described as "the science and engineering of making intelligent machines" in 1956 (McCarthy, 2007, p. 2). It has slowly developed into algorithms and intelligent machines throughout several decades of the 20th century. They can reason and adjust in response to settings and rule sets that resemble human intelligence (McCarthy, 2007)

Wang (2019) further extended the definition of artificial intelligence (AI) to encompass the ability to perform cognitive tasks. Such as learning and problem-solving, using modern technological advancements like neural networks, machine learning, and natural language processing (Zawacki-Richter et al., 2019). AI systems and chatbots represent a hopeful technological advancement that can boost productivity and efficiency in everyday tasks (Cotton et al., 2023).

Artificial intelligence is currently being utilised in various fields, including business, research, the arts, and academia, to enhance productivity and user experience.

2.2 Role of Artificial Intelligence in Writing

The development of advanced language models has extended their applications beyond conventional tasks. Now, it has touched upon the artistic area of verse and creative writing. The study has exemplified this expansion: *"Most Language Models Can Be Poets Too."* It elucidates how models can be prepared to generate captivating and engaging content in poetry. In this study, Rous et al. (2022) explore various methods of constrained text generation, utilising lexical, semantic, and phonetic constraints to bias vocabulary choices, thereby enabling models to generate text according to a predefined form and style of poetry.

Artificial intelligence (AI) has significantly impacted various fields of society, and educational writing is one of them. According to Smith and Johnson (2021), the use of AI-based writing assistance in educational contexts has enhanced academic writing ability and skills. Studies have shown that the advent of artificial intelligence (AI) writing help has had a substantial impact on the academic writing skills of

university teachers (Mohamed, 2024). These tools include: plagiarism detectors, research enhancement tools, and spell and grammar checks. Moreover, in academia, AI-based Writing Assistants have transformed into one of the essential tools that give results quickly to rapid reviewers, such as university faculty members, concerning their research works. These innovative applications locate and resolve spelling and grammar errors (Khabib, 2022).

In a study, Johnson and Smith (2019) observed that AI-enabled writing assistance applications used for grammar and spelling checks enhanced a person's accuracy while saving considerable amounts of time. Furthermore, study findings by Thompson et al. (2020) suggest that faculty members who use AI writing assistance tend to submit more readable papers with more positive feedback from reviewers. To this end, a study investigated the effectiveness of AI writing tools in academic writing for non-native postgraduate learners. A cohort of 120 students was enrolled into either an experimental group that used an AI tool, Grammarly, or a non-AI control group. This randomised controlled trial design was then used to analyse engagement, self-efficacy, and academic emotions through pre- and post-tests. The AI tool users reported improvements in engagement, emotion, behaviour, cognition, self-efficacy, and positive emotions in the academy setting, but a decline in negative ones. Finally, the research suggests that AI writing tools can be leveraged to achieve better learning outcomes and enhance technology acceptance in academic writing, whether through formative feedback or automated assistance (Nazari et al., 2021).

In addition, a study conducted by Imran and Almusharraf (2023) systematically investigated the role of ChatGPT as a writing assistant at higher education levels. Their focus included the use of ChatGPT in academic writing from December 2022 to May 2023. Findings highlighted that ChatGPT has a high capacity to provide support to students and educators through improved efficiency, idea generation, language facilitation, and editing.

The rapid advancements in AI have now made these tools capable of producing genre-specific texts. The study conducted by Brommer and colleagues in 2024 also investigated the genre competence of AI applications, such as ChatGPT, to determine whether they can create, revise, analyse, and enhance specific genre-based constructions. The pilot study included tasks such as classification, analysis,

summarisation, and generation of texts in six genres: congratulatory letters, condolence letters, book or film reviews, discussion essays, job advertisements, and promotional slogans. The findings indicate that ChatGPT is very good at generating and categorising texts, but has a more difficult time revising and truncating them.

Moreover, Ma et al. (2023) compared AI-generated scientific text to human-written scientific text, with the primary concern being the intervention of AI writing assistants in the generation of scientific text, a well-established genre of writing. The findings showed that AI-written texts were fluently translated and sensible; however, they were overly superficial and imprecise in factual content and not as well-written as the human counterpart.

Speaking more precisely, Choudhri et al. (2022) used the lexical specificity model developed by Hyland (2009) to analyze the prefaces of the Pakistani academic textbooks in soft sciences. In their work, they paid special attention to two dimensions of the Hyland model self-mention and hedges. They identified through a corpus-based study based on the AntConc that first-person pronouns of the type " I " were frequently employed by textbook authors to define authorial identity and produce a feeling of contact with readers. On the same note, the inclusion of hedging devices such as would and will indicated the attempt to show modesty and invite the reader to interpret, thereby increasing communicative appeal. The study demonstrates how the categories employed by Hyland can serve as aids in comprehending the persuasive and dialogic roles of academic discourse, revealing how authors manage to balance authoritative voice and audience interest.

On the same note, Qaiser et al. (2025) also used the model developed by Hyland (2009) when examining the linguistic patterns of university prospectuses, especially the messages of Vice-Chancellor and other informative sections such as the about us and mission sections. This paper has dealt with each of the five dimensions of the model of Hyland self-mention, directives, hedges, reporting verbs, and lexical bundles. Their results point out the role of self-mention (e.g., we, our, you) to create a common institutional voice and develop a feeling of community with potential students. Such directives like must and should were observed to underline institutional policies and necessities, whereas hedges were used to make claims soft and to assert institutional modesty. It was also observed in the study how the

reporting verbs and lexical bundles were used strategically to increase the credibility and guarantee the coherence of the text. In that way, the multidimensional lens of this model allows studying in detail the functions of language.

To sum up, this systematic literature review examine the role, limitations, and potential developments of AI, as well as in the linguistic aspects and relevance or choices resulting in the quality of such text in academic writing. In this regard, the research on artificial intelligence (AI) to write academic papers and create contents reflects significant progress and usage, yet it is yet to be described in a more specific manner in terms of linguistic options in AI texts and their applicability based on various genres as examined under corpus-supported approaches.

3. METHODOLOGY

This corpus-based study attempts to investigate the linguistic options of AI including pronouns and modal auxiliaries in different text types in the academic genre and evaluate the applicability of AI-generated text in carrying out its communicative roles in various text types. The study is situated within the context of corpus analysis, a methodological approach that integrates quantitative and qualitative analysis of linguistic features and the interpretation of texts.

3.1 Corpus Info

This study corpus is a collection of AI-generated essays produced during the use of different tools: ChatGPT, Perplexity AI, and Gemini and organized based on the types of texts or communicative functions of the academic genre. The texts are categorized and labeled as informative and persuasive to facilitate a systematic comparison between the types of texts.

3.2 Research Design

This study employs a corpus-based mixed-methods inquiry into the linguistic patterns and the relevance of AI-generated scholarly writings. The study aims at analyzing how the big language models, specifically ChatGPT, Perplexity AI, and Gemini, utilize linguistic decisions, like pronouns in various communicative functions or academic texts, among other things. The present study employs a non-random sampling method for data collection. The types of texts chosen to conduct the study include informative and persuasive essays created by three distinct AI applications, that is, ChatGPT, Gemini, and Perplexity AI. Therefore, this annotated corpus, created

to facilitate cross-text type comparisons and assess the effectiveness of AI-generated texts in fulfilling their communicative role, has been stored on Google Drive.

The reason behind choosing to analyze informative and persuasive types of texts is that they play a significant role in determining written communication, especially in academic fields. These two types of texts are crucial academic skills: the ability to inform with factual evidence and the ability to persuade through argumentation. Therefore, the knowledge of the functioning of these text types on the linguistic level could be applied to guarantee certain advancements in AI technologies that will ensure the reliability and accuracy of the final product in the future applications and chatbots.

3.3 Data Analysis Tool

Sketch Engine (Kilgarriff et al., 2014) provides the opportunity to conduct linguistic analysis of text generated by an AI. The strategy is primarily aimed at recognizing, gathering, and identifying patterns of linguistic decisions, as represented by the use of pronouns and modal auxiliaries in informative and persuasive texts within the academic genre. The tool/instrument facilitates a thorough analysis of the language patterns necessary for the research.

3.4 Theoretical Framework

This study employs the Lexical Specificity theoretical framework by Hyland (2009) according to the need of the study for genre analysis. Hyland further categorizes the genre for analysis into five that are: self-mentioning, hedges, directives, reporting verbs, and lexical bundles. The study focuses, therefore, especially on self-mentioning and directives, as the focus is on how Artificial Intelligence uses language for fulfilling the communication functions in academic writing across informative and persuasive text types.

4. RESULTS AND DISCUSSION

This corpus-assisted study aims to identify the linguistic choices AI uses in various text types within the academic genre and examine the relevance of AI-generated text in performing the communicative functions across different types. This section presents the results obtained using Sketch Engine, which has been employed to provide the required list of linguistic choices, specifically pronouns and modal Auxiliaries, along with their frequencies in persuasive and informative essays,

respectively generated through three different AI tools, i.e., ChatGPT, Gemini, and Perplexity AI, on a variety of topics. These lexical selections with the lens of the chosen theoretical framework of 'Lexical Specificity' by Hyland (2009) have been analyzed below.

4.1 Results

This section represents the results drawn from Sketch Engine after analyzing the informative and persuasive essays generated through three different AI tools: ChatGPT, Gemini, and Perplexity AI. These results have been analyzed and discussed in detail in the discussion section of the study.

Figure 1: Modal Auxiliary Verb, 'Can' Used by AI in the Informative Genre

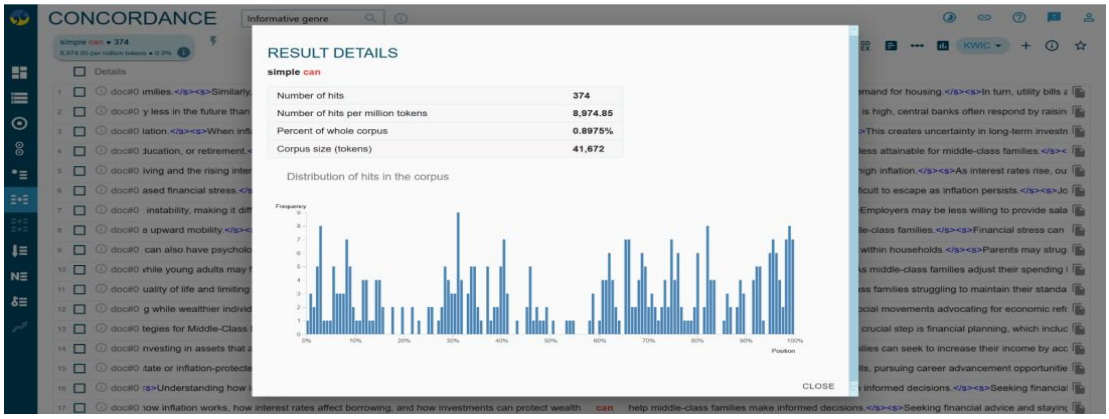

Figure 2: Modal Auxiliary Verb, 'Can' Used by AI in the Persuasive Genre

4.2 Discussion

This section answers the research questions under the theoretical framework, 'Lexical Specificity' by Hyland (2009). The analysis is conducted based on the results drawn from Sketch Engine, which provides the list of linguistic choices in persuasive and informative essays generated through three AI tools: ChatGPT, Gemini, and Perplexity AI.

4.2.1 Analysis of Essays Generated Through AI Tools in the Informative Genre

4.2.1.1 Self-Mentioning

The first category of Hyland's (2009) concept of 'Lexical Specificity' is Self-Mention. It refers to the practice of writers referring to themselves using specific pronouns. This allows them to actively engage with the text, establish a presence within their arguments, and build a connection with the intended readers (Gul & Khattak, 2021). Similarly, Wu et al. (2015) consider self-mention a powerful rhetorical tool for constructing authorial identity in the text. Table 1 below represents the frequencies and the types of self-mentioning pronouns used in the informative essays.

Table 1: ***List of Pronouns Generated Through AI in Informative Genre***

Pronoun	Person	Frequency (ChatGPT)	Frequency (Gemini)	Frequency (Perplexity AI)	Overall Frequency (Informative Genre)
We	First	06	69	52	127
Our	Person	11	76	13	100
Us		01	27	05	33
Your	Second	01	01	03	05
You	Person	02	01	-	03
Their		62	49	40	151
It		47	104	63	214
Its		25	30	35	90
They		13	23	12	48
Them		12	11	14	37
Themselves	Third	02	03	03	08
Itself	Person	03	04	-	07
One		02	02	01	05
His		01	-	01	02

The analysis of the corpus reveals that the most frequently used pronouns across all three AI tools in the informative essays are the third-person pronouns, specifically "It" and "Their", which appear 214 and 151 times, respectively. This reflects the objective nature of informative essays, where topics, concepts, or data points are often

discussed without personal involvement. Similarly, the use of the pronoun *"Its"* with a moderately higher frequency, *90* times, further emphasizes the reliance of the AI tools on third-person references while dealing with the informative text type.

However, despite the AI's preference for objectivity in the informative text type of the academic genre, the use of first-person pronouns in the datasets, such as "We" and "Our," is also noticeable, with *"We"* appearing *127* times and *"Our" 100* times. According to Hyland (2009), self-mentioning in a text enables a writer to assert ownership of their work and express their personal opinion to readers, thereby revealing the author's perspective. However, the informative text type requires a more objective tone, where information is reported without the writer's personal involvement. Thus, this frequent use of personal pronouns in the corpus deviates from the expected objectivity of the informative text type.

Thus, the distribution of pronoun types in the corpus, i.e., AI-generated informative texts, reveals distinct patterns in linguistic choices, shaped by the communicative functions of the text type. In informative writing, third-person pronouns dominate, reflecting AI's tendency to explain concepts, cognitive processes, and information in an objectively formal tone of academic writing. It indicates an inclination to explain concepts with a clear focus on objects or subjects rather than on the author or reader. Therefore, it effectively performs the communication function through appropriate linguistic choices, making the essays more relevant in the academic domain. However, a noticeable deviation is observed in the moderately high use of first-person pronouns.

4.2.1.2 Directives

According to Hyland (2009), another important category is directives. In English grammar, directives are sentences that instruct or encourage the reader to take action. They often feature a significant number of imperative sentences, typically starting with action verbs. Directives can involve inviting, commanding, suggesting, pleading, requesting, expressing wishes, or granting permission (Gul & Khattak, 2021). In English, they are conveyed through two primary forms: imperatives and modal Auxiliaries. This study focuses on one form of directives, i.e., modal auxiliaries.

4.2.1.3 Modal Auxiliaries

The modal Auxiliaries from the corpus, i.e., informative essays generated through

different AI tools, ChatGPT, Gemini, and Perplexity AI, have been presented in Table 2.

Table 2: ***Modal Auxiliaries in the Informative Essays Generated Through AI Tools***

Modal Auxiliary	Frequency (ChatGPT)	Frequency (Gemini)	Frequency (Perplexity AI)
Have	0.63%	0.47%	0.38%
Can	0.85%	0.71%	1.1%
Could	0.022%	0.014%	0.051%
May	0.19%	0.097%	0.094%
Might	0.015%	0.062%	0.0072%
Must	0.13%	0.035%	0.087%
Should	0.037%	0.0069%	0.15%
Will	0.067%	0.16%	0.094%

Generally, in the informative text, the use of modal auxiliaries reflects an emphasis on conveying factual information and expressing abilities, possibilities, or certainties.

Among the various modal auxiliaries, it is noted in Table 2 that *"Can"* emerges as the most frequently used modal in the corpus, with the highest frequency. This is probably because most informative texts in academic writing are expected to establish themselves as showing potential or capacity or permission, as "can" is often associated with capability, potential, or permission (Sujatna et al., 2019), which is essential in informative writing. Closely following *"can"* is the modal auxiliary *"have,"* in second place to frequency, mostly serving in perfect-tense constructions to fix or narrate events in the past. The frequent occurrence of this modality is in tune with the tendency of the text type to provide accurate and detailed information.

In contrast, the modals, which express uncertainty or possibility or hypothetical situations such as *"Could," "Might,"* and *"May,"* have shown a significantly lower frequency in the corpus because informative texts are largely concerned with fact accuracy and objective statements rather than speculations. Thus, the reduced number of the modals in this corpus suggests that AI models will avoid ambiguity when creating informative texts. Similarly, necessity or obligation modals (Sujatna et al., 2019) like Must and Should have not been used in corpus very often to

show strong recommendations, rules, or guidelines. In general, their usage within the text refers to the cases when the authoritative statements are required, in particular, in the instructional or technical texts.

In general, the fact that modal auxiliaries are used in informative texts indicates that the models of the AI are oriented to the objective tone of the text type in the academic genre. By prioritising the use of modals that convey certainty and factuality and by eschewing the use of speculative language, the AI models have managed to keep their output coherent and clear.

4.2.2 Analysis of the Essays Generated Through AI Tools in the Persuasive Genre

Persuasion involves conveying a message to a receiver with the intent of influencing their beliefs, opinions, and behavior through the use of argumentation (Bonta, 2008). It typically blends emotional appeal with intellectual reasoning. Additionally, it functions as a form of influence that predisposes rather than imposes, meaning that while it does not compel individuals to act in a specific way, it presents logical, emotional, and cultural arguments to encourage a particular course of action. (Larson 2003: 18). This section analyzes the AI-generated essays in the persuasive text-type under Hyland's (2009) Lexical Specificity model.

4.2.2.1 Self-Mentioning

In persuasive writing, the use of pronouns plays an important role in constructing authorial identity, engaging readers, and emphasizing the arguments. Thus, the pronouns used by different AI tools such as ChatGPT, Gemini, and Perplexity AI have been listed in the table below. They have been analyzed under Hyland's (2009) lexical specificity model.

Table 3: ***List of Pronouns Generated Through AI in the Persuasive Genre***

Pronoun	Person	Frequency (ChatGPT)	Frequency (Gemini)	Frequency (Perplexity AI)	Overall Frequency (Persuasive Genre)
We		37	93	96	226
Our		28	131	26	185
Us	First Person	03	45	05	53
Your		01	09	04	14

You	Second Person	02	02	01	05
It		79	124	99	302
Their		71	55	71	197
They		20	41	23	84
Its		24	20	38	82
Them		21	09	26	56
Themselves	Third Person	05	01	05	11
Itself		01	03	-	04

The analysis of the corpus shows that the most frequently used pronoun across all AI tools in the persuasive text is *"It,"* with a total of *302* occurrences. This high frequency indicates a reliance on objective discussions, likely involving abstract concepts, arguments, or evidence. On one hand, Table 3 shows that ChatGPT has used *"It" 79* times, while Gemini and Perplexity AI have used it even more frequently (124 and 99 times, respectively).

Unlike informative writing, where the primary goal is to explain and present facts objectively (Swales & Feak, 2012), persuasive writing seeks to convince and influence the readers (Bonta, 2008). Thus, the language is adapted accordingly. However, the extensive use of third-person pronouns in the data set suggests a deviation from the expected persuasive style. Instead of emphasizing the author's stance or appealing directly to the audience, the AI models focus more on presenting objective facts and external arguments. Other third-person pronouns used in the persuasive essays are: *"Their"* (197 occurrences), *"They"* (84 occurrences), and *"Them"* (56 occurrences). Thus, the AI's preference for third-person references over first-person involvement diminishes the subjective and authoritative tone often characteristic of effective persuasion.

This is because, generally, the first-person plural pronouns convey the shared sentiments between the writer and the readers. These mutual feelings are intended to positively influence readers by persuading and convincing them of the message's content. The use of "we" allows the author to generalize their statements, building a sense of solidarity with the readers (Gul & Khattak, 2021). This sense of connection encourages readers to engage with the text as though it directly addresses them,

thereby fulfilling the role of representing the 'collective self' in the writing (Wu et al., 2015). Furthermore, these First-person pronouns are used to create a sense of emotional connection between the writer and the audience. By using these pronouns, communicators share their perspectives and experiences, establishing a deeper bond that allows readers to feel more engaged and connected (Wu et al., 2015). Resultantly, it leads to the fulfillment of the communicative effect in the persuasive text, i.e., to persuade by appealing to the emotions targeted.

4.2.2.2 Directives

4.2.2.2.1 Modal Auxiliaries

The modal Auxiliaries from the corpus, i.e., persuasive essays generated through different AI tools, ChatGPT, Gemini, and Perplexity AI, have been presented in Table 4.

Table 4: ***Modal Auxiliaries in the Persuasive Essays Generated Through AI Tools***

Modal Auxiliary	Frequency (ChatGPT)	Frequency (Gemini)	Frequency (Perplexity AI)
Can	0.67%	0.55%	1.3%
Could	0.044%	0.02%	0.069%
May	0.087%	0.06%	0.1%
Might	-	0.053%	0.046%
Must	0.31%	0.06%	0.12%
Should	0.19%	0.0066%	0.13%
Will	-	0.13%	-

In persuasive writing, the selection of modal auxiliaries serves a different purpose than in informative writing. While informative texts prioritize factual clarity and objectivity (Swales & Feak, 2012), persuasive content focuses on influencing the audience's opinions or actions (Bonta, 2008).

Table 4 illustrates a relatively high use of *"Can"* across all AI models in the persuasive text type of academic writing, particularly in Perplexity AI (1.3%), followed by ChatGPT (0.67%) and Gemini (0.55%). This reflects a deviation of artificial intelligence to make text type-compatible linguistic choices. This is because *"Can"* is often associated with informative statements, where the information is presented as possibilities (Sujatna et al., 2019). However, the more frequent use of *"Can"* in the persuasive essays may confuse the readers about the possibilities

associated with *"can,"* resulting in the unfulfilled purpose of the text, to convince the readers. This doubt in language may significantly affect the decisions of the audience. Thus, it may undermine the assertiveness that is often characteristic of strong persuasive arguments.

On the flip side, the presence of *"Must"* and *"Should"* in the persuasive text is more relevant. *"Must"* conveys a strong sense of obligation or necessity (Sujatna et al., 2019), with ChatGPT using it the most (0.31%), while Perplexity AI (0.12%) and Gemini (0.06%) apply it more sparingly. This form of modal is commonly applied to authoritative arguments in order to underline the necessity of a certain action. Equally, Should is a device used to provide suggestions or promote certain opinions. The fact that ChatGPT is more inclined to use more direct and assertive language in persuasive messages is reflected by its frequency (0.19) in contrast to Perplexity AI (0.13) and Gemini (0.0066).

On the whole, modal auxiliaries choice in the persuasive text does not seem to be a characteristic feature to define the intent of persuasion. Although some modals such as Must and Should are applicable when it comes to stating a strong opinion and giving a call to action, the constant application of Can can be interpreted to mean a deficiency in assertiveness. This dependence on modal auxiliaries might reflect the fact that AI-written persuasive texts might be more inclined towards the cautious less forceful tone.

5. CONCLUSION

The goal of the present research was to define the linguistic decisions of the Artificial Intelligence (AI tools) in different types of texts of the academic genre and to discover the applicability of AI-generated text in executing the communicative functions. To this end, the corpus-assisted analysis examined AI-generated essays in the chosen texts, specifically those that were informative and persuasive. These essays were produced by AI tools, i.e., ChatGPT, Gemini, and Perplexity AI. Afterwards, they were analyzed under Hyland's (2009) model of linguistic specificity as a theoretical framework.

The study's findings reveal that AI tools generate academic genre-compatible content. However, when the cross-text type comparison was carried out, it was observed that the AI-generated informative essays display recognizable linguistic

patterns that coincide with the communicative function of the concerned type, while also highlighting certain minor deviations. The most remarkable finding is that there is a strong preference for third-person pronouns, especially "it" and "their", which appeared to have the highest frequencies. This reflects an objective position where informative writing centers on topics, concepts, and data rather than personal involvement.

Moving ahead, within the particular analysis of the modals under this theoretical approach, it is important to note that with the dominance of "Can" in the corpus, AI seems biased toward the expression of ability, possibilities, or permissions, which are generally inferred in informative writing. Similarly, the overwhelming abundance of the word "Have", particularly in its application in perfect tenses, clearly strengthens the informative text's warranty of accuracy and truthful reporting. In contrast, speculative modals such as "Could," "Might," and "May" are used less frequently, suggesting that AI tools prioritize certainty over ambiguity when generating informative content.

On the other hand, the analysis of the corpus, i.e., persuasive essays generated by artificial intelligence, shows that the most frequently used pronoun across all AI tools in the persuasive writing is "It". This high frequency indicates a reliance on objective discussions, likely involving abstract concepts, arguments, or evidence. Unlike the informative text, where the primary goal is to explain and present facts, persuasive writing seeks to convince and influence the readers. However, the extensive use of third-person pronouns suggests a deviation from the expected persuasive style, which typically emphasizes the author's stance and appeals directly to the audience.

The analysis further indicates that AI's selection of modal auxiliaries in the persuasive text deviates from persuasive writing. The persuasive content focuses on influencing opinions or actions (Bonta, 2008). However, the relatively high use of "Can" across all AI models reflects a text-type-incompatible linguistic choice. "Can" is associated with possibilities, which may weaken the assertiveness of persuasive arguments. In contrast, "Must" and "Should" convey strong obligation and recommendation, making them more relevant to persuasion.

Thus, after the cross-text type comparison of AI-generated texts (essays), it may be

concluded that artificial intelligence generates genre-compatible content (academic genre) but deviates in its linguistic choices across the text types, informative and persuasive within the academic genre. However, the deviation is more pronounced in the persuasive texts. This is largely due to an inappropriate use of pronouns. Additionally, the frequent use of the modal auxiliary *can* in persuasive texts introduces ambiguity, as it often fails to establish the writer's stance, thereby weakening the argumentative clarity.

REFERENCES

Bonta, R. (2008). Stylistics and Persuasion. *Stylistyka*, *17*, 223–244. https://czasopisma.uni.opole.pl/index.php/s/article/view/3663

Brommer, S. (2019). Empirically grounded language criticism: A contribution to the operationalization of the vague category of appropriateness. *Aptum: Zeitschrift für Sprachkritik und Sprachkultur, 15*(2), 123–133.

Brommer, S., Frick, K., Bursch, A., Rodrigues Crespo, M., & Schwerdtfeger, L. K. (2024). ChatGPT and its genre competence: An exploratory study. *Weizenbaum Journal of the Digital Society, 4*(4), w4.4.6. https://doi.org/10.34669/WI.WJDS/4.4.6

Choudhri, S., Zahra, T., & Shehzad, W. (2022). 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), 268–279. https://doi.org/10.33195/jll.v5iII.310

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Ginting, D., & Barella, Y. (2022). Academic writing centers and teaching academic writing at colleges: A literature review. *Journal of Education and Learning (EduLearn), 16*(3), 350–356.https://doi.org/10.11591/edulearn.v16i3.20473

Gul, H., & Khattak, N.-U.-R. (2021). A corpus-based genre analysis of the linguistic specificity of VC messages on the websites of universities in Pakistan. *Erevna: Journal of Linguistics & Literature, 4*(2), 14-34.

Hyland, K. (2008). Genre pedagogy: Language, literacy, and L2 writing instruction.

Journal of second language writing, 16(3), 148-164.

Hyland, K. (2009). *Academic Discourse: English in a Global Context*. Continuum.

Hyland, K. (2009). Writing in the disciplines: Research evidence for specificity. *Taiwan International ESP Journal, 1*(1), 5-22.

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at the higher education level: A systematic review of the literature. *Contemporary Educational Technology, 15*(4). https://doi.org/10.30935/cedtech/13605

Johnson, A., & Smith, B. (2019). The impact of AI-based writing assistants on grammar and spelling checks. *Journal of Writing Technology, 15*(2), 45-62.

Khabib, S. (2022). Introducing artificial intelligence (AI)-based digital writing assistants for teachers in writing scientific articles. *Teaching English as a Foreign Language Journal, 1*(2), 114-124.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography, 1*(1), 7–36. https://doi.org/10.1007/s40607-014-0009-9

Lametti, D. (2022). AI could be great for college essays. slate.com. https://slate.com/technology/2022/12/ chatgpt-college-essay-plagiarism.html

Larson, C. (2003). *Persuasiunea: Receptare și responsabilitate*. Polirom.

Matsagouras, E. G., & Tsiplakou, S. (202X). Who's afraid of genre? Genres, functions, text types, and their implications for a pedagogy of critical literacy. In S. Tsiplakou & X. Hadjioannou (Eds.), *Scientia Paedagogica Experimentalis - International Journal of Experimental Research in Education*, *45*(1), 71–90.

McCarthy, J. (2007). From here to human-level AI. *Artificial Intelligence, 171*(14-15), 1174-1182. https://doi.org/10.1016/j.artint.2007.10.009

Mohamed, A. M. (2024). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL Faculty Members. *Education and Information Technologies*, *29*(3), 3195-3217.

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence-powered Digital Writing Assistant in higher education: randomized controlled trial. *Heliyon, 7*(5). https://doi.org/10.1016/j.heliyon.2021.e07014

Perkins, M. (2023). Academic integrity considerations of AI Large Language Models

in the postpandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice, 20*(2), 3-16. https://doi.org/10.53761/1.20.02.07

Qaiser, A., Syed, S. A., & Khattak, N. U. R. (2025). Linguistics specificity of universities prospectuses: A corpus-based genre analysis. *Corporum: Research Journal of Corpus Linguistics*, 7(II), 34–46. http://journals.au.edu.pk/ojscrc/index.php/crc/Home

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, *1*(8), 9.

Rous, A., Basu, S., Moorthy, A., & Dubovoy, D. (2022). Most Language Models Can Be Poets Too: An AI Writing Assistant and Constrained Text Generation Studio. Retrieved from arXiv:2306.15926 https://arxiv.org/abs/2306.15926

Shahriar, S., & Hayawi, K. (2023). Let's Have a Chat! A Conversation with ChatGPT: Technology, Applications, and Limitations. *Artificial Intelligence and Applications*, *2*(1), 11–20. https://doi.org/10.47852/bonviewAIA3202939

Shen, Z., Zhou, M., & Zaib, K. (2024). Religious semiotics in performance and visual art: Symbolism in aboriginal dot painting, sichuan opera makeup, chinese traditional sculpture and shu embroidery. *European Journal for Philosophy of Religion*, *16*(3), 266-292.

Smith, B., & Johnson, A. (2021). Faculty approval and acceptance rates with the use of AI-based writing assistants. *Journal of Academic Publishing, 30*(4), 210-225.

Sujatna, M. L., Sujatna, E. T. S., & Pamungkas, K. (2019). Exploring the use of modal auxiliary verbs in Corpus of Contemporary of American English (COCA). *Sosiohumaniora: Jurnal Ilmu-ilmu Sosial dan Humaniora*, *21*(2), 166–172. https://doi.org/10.24198/sosiohumaniora.v21i2.19970

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Swales, J., & Feak, C. (2012). *Academic Writing for Graduate Students: Essential Tasks and Skills.* University of Michigan Press. https://doi.org/10.3998/mpub.2173936

Teng, M. F., & Wang, C. (2023). Assessing academic writing self-efficacy belief and writing performance in a foreign language context. *Foreign Language Annals, 56*(1), 144–169. https://doi.org/10.1111/flan.12638

Teng, M. F., Qin, C., & Wang, C. (2022). Validation of metacognitive academic writing strategies and the predictive effects on academic writing performance in a foreign language context. *Metacognition and Learning, 17*(1), 167–190. https://doi.org/10.1007/s11409-021-09278-4

Thompson, R., Smith, L., Ahmed, N., & Lee, J. (2020). Enhancing content quality through AI-based writing assistants. *Journal of Academic Writing, 25*(3), 78-95.

Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.-T., Jin, A., Bos, T., Baker, L., Du, Y., Li, Y., Lee, H., Zheng, H. S., Ghafouri, A., Menegali, M., Huang, Y., Krikun, M., Lepikhin, D., Qin, J., ... Le, Q. (2022). LaMDA: Language Models for Dialog Applications. *arXiv*. https://doi.org/10.48550/arXiv.2201.08239

Ullah, S., Saeed, S., Ahmad, I., Khan, F., & Naz, A. (2021). COVID-19 and Online Teaching Strategies: The Impact of Online Teaching-Learning on Students of Poor Socio-economic Backgrounds in Malakand Division Khyber Pakhtunkhwa. *Indian Journal of Economics and Business*, *20*(4), 1217-1224.

Ullah, S., Saeed, S., Khan, F., & Naz, A. (2021). Covid-19, Economic Lockdown, Treatment Interruptions And The Fear Of Survival Among HIV/AIDS Patients In Malakand Division, Khyber Pakhtunkhwa. *Webology (ISSN: 1735-188X)*, *18*(6).

Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, *10*(2), 1-37.

Wu, G., & Zhu, Y. (2015). Self-mention and authorial identity construction in English and Chinese research articles: A contrastive study. *Linguistics and the Human Sciences, 10*(2), 133–158. https://doi.org/10.1558/lhs.v10i2.28557

Zaib, K. (2022). THE ROLE OF MOTHER TONGUE IN LEARNING ENGLISH AT THE INTERMEDIATE LEVEL (A Case Study of Government Degree College Batkhela, District Malakand). *Pakistan Journal of Society, Education and Language (PJSEL)*, *8*(2), 248-253.

Zaib, K., Ahmed, S., & Al Murshidi, G. (2025). Exposed Online, Endangered Offline: The Intersection of Social Media Presence and Gendered Violence in Pakistan. *Liberal Journal of Language & Literature Review*, *3*(3), 1243-1265.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–

where are the educators?. *International Journal of Educational Technology in Higher Education*, *16*(1), 1-27.

Zhao, X., Sbaffi, L., & Cox, A. (2023). The digitization of writing in higher Education: Exploring the use of Wordtune as an AI writing assistant. https://doi.org/10.31219/osf.io/uzwy7