| RESEARCH ARTICLE

Stylometric Analysis of AI Chatbot-Generated Emails: Are Students Losing Their Linguistic Fingerprint?

Jokha Al Hosni
Center for Preparatory Studies, Sultan Qaboos University, Muscat, Sultanate of Oman
Corresponding Author: Jokha Al Hosni, **E-mail**: alhosnij@squ.edu.om

| ABSTRACT

The integration of artificial intelligence (AI) in education, particularly through AI chatbots like ChatGPT, Copilot, Gemini, and Bing, has revolutionized second language teaching and learning. These chatbots, which utilize advanced natural language processing and machine learning algorithms, hold great potential. They can operate complex cognitive tasks, provide immediate feedback, enhance motivation and self-confidence, reduce anxiety and improve language skills. Despite these advantages, concerns about academic integrity, loss of unique voice, and lack of emotional depth have emerged. This research aims at conducting a stylometric analysis on 25 AI chatbot-generated emails produced by undergraduate English language learners at Sultan Qaboos University. By examining stylistic features like tone, lexical density, lexical diversity, choice of words, repetition, formality level and emotional depth, the study will provide insights into the strengths and limitations of AI chatbots in enhancing email writing skills while preserving the unique voice of student writing. Findings reveal that AI-generated emails are characterized by a repetitive structure, a high level of formality and politeness, a high lexical diversity and a lack of emotional depth and personal anecdotes. The outcomes of this research will shed light on the awareness and skills students need to acquire to optimize the use of AI chatbots without losing their linguistic fingerprint.

1. Introduction

The adoption of artificial intelligence (AI) in education has transformed the landscape of language teaching and learning, especially the use of generative AI in the form of AI chatbots, such as ChatGPT, Copilot, Gemini and Bing. These chatbots, which follow Natural Language Processing (NLP) and use advanced machine learning algorithms, can learn from users' previous inputs and they can recognize, infer and use human languages. In addition, they are capable of replicating human cognitive functions and dealing with more complex problems (Bashang & K, 2023). It is also believed that motivation and self confidence levels are generally higher when communicating with a machine (AI chatbot) rather than a human (Baskara, 2023; Silitonga et al., 2023; Song & Song, 2023; Songsiengchai , 2023; Wu & Yu, 2023; Haristiani , 2019). This is due to the reduced anxiety level and the ability to get immediate feedback and multiple versions of the output. Besides, the accuracy, human- likeness and social presence that feature AI chatbots have increased students' motivation to use them (Ebadi & Amini, 2022).

There is a high volume of research in the past two years on the use of AI chatbots in English language learning, which has highlighted the promising features that benefit language learners. AI-powered chatbots have shown effectiveness in improving specific language skills, including vocabulary (Özçelik & Ekşi, 2024; Zhang & Huang, 2024; Songsiengchai, 2023), grammar (Harunasari, 2023; Songsiengchai, 2023), speaking (Du & Daniel, 2024; Duong & Suppasetseree, 2024; Khasawneh, 2023; Rusmiyanto et al. , 2023; Songsiengchai, 2023), reading (Songsiengchai, 2023) and writing (Yuan et al., 2024; Baskara, 2023;

Harunasari, 2023; Hawanti, 2023; Kwon et al., 2023; Silitonga et al. , 2023; Yan, 2023). Using AI chatbots as writing assistants has been widely adopted. This is due to their expectational capacities to generate ideas, themes, topics and perspectives (Kasneci et al., 2023; Taecharungroj, 2023), reduce the time and efforts required to generate well-structured and coherent ideas (Lund et al., 2023; Yan, 2023), and proofread and edit student writing by suggesting corrections for spelling, grammar and syntax errors which helps improve the quality of their writing (Imran & Almusharraf, 2023).

Despite the highlighted potentials in these studies, they have also emphasized the need to investigate the limitations and concerns associated with the use of AI chatbots. Academic integrity and plagiarism have been one of the top concerns in higher education (Costa et al., 2024; Khalaf, 2024; Zeb et al., 2024; Barrot, 2023; King, 2023; Rudolph et al., 2023; Van Dis et al., 2023). Furthermore, using AI chatbots in learning can raise concerns about language barriers and cultural differences that may affect the quality and accuracy of the information provided as well as posing challenges related to privacy and bias (Xue et al., 2023). Pawar (2024) believed that the capacity of AI models is limited to the data they are trained on. While AI has made substantial strides in processing text data, it still struggles to fully comprehend the intricacies and subtleties of human language.

At the Centre for Preparatory Studies at Sultan Qaboos University, English language instructors have been observing a surge of AI chatbot-generated content turned in by their students in both the English foundation program and the ESP (English for Specific Purposes) credit courses. Student emails that are sent to their teachers is one obvious example. Teachers have shown concerns about the authenticity of these emails, fearing that students may be relying too heavily on AI tools instead of developing their own writing skills. This trend raises questions about academic integrity, the loss of a unique voice, and the lack of emotional depth in student writing, especially in email communications. Therefore, this research is an attempt to address these concerns by analyzing 25 emails generated by students using AI chatbots in order to understand the impact of AI chatbots on student writing and language development.

2. Literature Review
2.1 Characteristics of Chatbot-Generated vs. Human-Generated Texts
A couple of research papers have explored the stylistic features of ChatGPT- generated texts. Imran & Almusharraf (2023) conducted a systematic review of the 30 most relevant articles which investigated the role of ChatGPT as a writing assistant in academia. The review included scientific and scholarly articles, essays, creative writings and editorials. The findings revealed significant differences between ChatGPT and human-generated writings. For instance, ChatGPT has the ability to generate concise introductory passages, abstracts, and literature sections. However, it produces limited responses compared to human content. The researchers concluded that there are opportunities and challenges in adopting ChatGPT as a writing assistant. Although ChatGPT can be a promising tool for assessing and enhancing writing skills, researchers, teachers, and students cannot fully rely on ChatGPT for writing tasks as human control should be integrated to yield the optimal utilization of this tool.

AlAfnan & MohdZuki (2023) examined the stylistic features of case study, business correspondence, and academic writing ChatGPT-4-generated responses by exploring sentence length, paragraph structure, tense, voice, mood, pronouns, word choice , keywords density, lexical density, lexical diversity, and reading ease. The study revealed that ChatGPT-4 responses are generated in paragraphs of 2 to 4 sentences of 16 to 20 words each. The sentences are mainly generated in the declarative mood by using simple present tense in active voice using third-person singular pronouns. Technical words and abbreviations in business correspondences are used without explaining what they stand for. The lexical density and lexical diversity are high, and the reading ease is low. Similarly, keywords and lexical densities are high in ChatGPT-4 academic writing-generated responses , but the lexical diversity is low.

Casal & Kessler (2023) focused on research abstracts as the target genre for their comparative study. They believed that abstracts are a genre that may be particularly vulnerable to generative AI because of their constrained length requirements and highly formulaic language. Findings of their research showed that reviewers were unsuccessful in identifying AI versus human writing, despite employing multiple rationales to judge texts.

Mahyoob et al. (2023) proposed a framework for analyzing and measuring the ChatGPT capabilities in generating effective academic writing. A critical analysis of the content of some generated academic texts was conducted based on a proposed framework which consists of six principles: Relatedness, Adequacy, Limitation, Authenticity, Cognition, and Redundancy. The results suggested that despite ChatGPT's exceptional capabilities, many issues in academic writing are raised including information repetition, illogical reasoning, fake references, nonfactual inferences, hallucination, and lack of pragmatic interpretation. The findings indicated that, due to the large trained corpus, ChatGPT can produce human-like language with grammatically correct sentences and well-formed structure, but it fails to produce well-formed academic texts with authentic and reliable information.

2.2 Linguistic Fingerprint and Emotional Depth in AI Chatbot-Generated Texts

Another crucial aspect that should be analyzed in AI chatbot-generated texts is the linguistic fingerprint, which is a key focus of this research paper. Linguistic fingerprint is defined as the unique patterns of language use that each individual has. These patterns can include things like word choice, grammar, spelling, and punctuation. Linguistic fingerprint is a common term in forensic linguistics as it is used to help determine who wrote a document or to identify a suspect in a crime. Linguistic fingerprinting indicates that each human has a unique way of using language to communicate thoughts, ideas and opinions (Herbold et al. , 2023). In the era of generative AI where genuine writing can be lost, shedding light on linguistic fingerprint and the presence of the unique voice of individual writers in language learning is essential.

Few researchers have highlighted this topic in their research papers. Barrot (2023) believed that ChatGPT, as an AI chatbot, "may mimic a high degree of human intelligence and excel in many aspects of writing. However, this tool cannot totally replace humans because it lacks the emotional depth and life experiences that contribute to an individual's voice, identity, and distinctiveness in writing" (p.4). In his study, Barrot observed a certain level of template rigidity of writing as it follows the same structure when prompted to write an essay about different concepts. In a similar study, Johansson (2023) compared argumentative essays written by students for an English literature course with equivalent essays generated through ChatGPT. It investigated whether AI can meet the requirements of academic writing and the distinctiveness of voice in the generated text. Authorial presence, assertiveness and self-identification were explored. The results showed that ChatGPT can generate appropriate context-based texts, but assistance is required with factual accuracy and the nuanced characteristics of authorship which are present in human writing. The AI-generated text lacks the specificity, depth and accurate source referencing that are present in human-generated text. The study revealed that human-authored essays exhibit a more distinctive, nuanced, and intricate style. On the other hand, the AI-generated content, while using personal pronouns and a consistent voice, lacks individuality and nuance, often resorting to repetitive phrases. Human writing effectively uses sources, hedges, and boosters, creating a stronger sense of authorship and academic rigor.

Herbold et al. (2023) also explored the linguistic devices that are characteristic of student versus AI-generated content (using ChatGPT- 3, 3.5 and 4). The key findings showed that ChatGPT-4 considerably outperforms ChatGPT-3 and 3.5 in terms of logical structure, language complexity, vocabulary richness and text linking. Regarding the writing styles, humans and generative AI models differ significantly. For instance, more nominalizations and higher sentence complexity are used in the GPT models, whereas the students' writing makes more use of modal and epistemic constructions (which tend to convey the speaker attitude). ChatGPT-4 has a significantly higher lexical diversity than the students' writings. However, students' writing conveyed students' unique voice and distinctive style.

In a quantitative study conducted by Liao et al. (2023), the researchers analyzed the difference between ChatGPT-generated and human-generated medical texts. The results showed that medical texts written by humans were more concrete, more diverse, and typically contained more useful information. However, medical texts generated by ChatGPT had an emphasis on fluency and logic and usually expressed general terminologies rather than specific effective information related to the context of the problem.

In her analysis of empirical research on AI's impact on the emotional dimensions of human communication, Oritsegbemi (2023) concluded that AI cannot fully replicate the complexity and subjectivity of human intelligence and emotion, due to its lack of innate feelings and understanding of abstract concepts.

No research yet has explored the stylistic features of AI chatbot-generated emails produced by English language learners in higher education. I believe email writing is an important genre to investigate as it serves as a critical tool for communication in both academic and professional settings. In addition, students usually have a genuine reason to communicate with their teachers regarding absences, grades, or other academic concerns which makes email writing a practical and relevant area of study. Understanding how AI chatbots influence language use in this context can significantly impact language learning strategies and pedagogical approaches. Analyzing AI chatbot-generated emails using stylometric analysis can help explore the style of these emails and how learner identity as second language learners has been affected. If learners interact frequently with AI-generated content, they might adopt certain patterns or styles from AI chatbots which can potentially homogenize their language use. This could affect their ability to develop their linguistic fingerprint to express themselves uniquely. Therefore, this research aims to investigate the influence of AI chatbots like ChatGPT on the email writing skills of undergraduate English language learners at Sultan Qaboos University through conducting a stylometric analysis of chatbot-generated emails. The focus is on specific aspects such as tone, choice of words, lexical density and diversity and formality and politeness levels used. By focusing on these aspects, the research aims to provide insights into the strengths and limitations of AI chatbots in enhancing email writing skills while maintaining the unique voice of student writing. The outcomes of this research will shed light on the awareness and skills students need to acquire to optimize the use chatbots without losing their linguistic fingerprint.

3. Methodology
3.1 Participants
The participants of this research were 10 English Language instructors from the Centre for Preparatory Studies at Sultan Qaboos University who shared samples of student AI chatbot- generated emails. The 10 teachers taught different foundation and ESP credit courses to undergraduate students in Spring 2024, so the collected sample emails were produced by students with different English proficiency levels and different sciences and humanities majors. The table below summarizes the number of sample emails from each category and the average English proficiency level.

Table 1 Collected Sample Emails and the English Proficiency

Category of courses	English proficiency level (CEFR)	Number of sample emails
English foundation program (FP)	B1	10
English for ESP (Humanities majors)	B2	8
English for ESP (Sciences majors)	B2	7

3.2 Research Instruments
This research study used a mixed method where both quantitative and qualitative data analysis were used. A short questionnaire was shared with teachers to fill it in with information regarding the course they taught in Spring 2024 and to share the key characteristics they think feature the chatbot-generated emails produced by their students. The last question was to upload screenshots of student emails (up to 5) that participants believe were generated using an AI chatbot like ChatGPT. To maintain confidentiality, participants were asked to hide the student's name and university ID number. 25 emails were collected. The data analysis phase of the research adopted the stylometric analysis approach to find some stylistic patterns. Stylometry uses statistical methods to analyze stylistic features like vocabulary choices, sentence length, and punctuation usage. It is mainly used for author identification and has broader applications such as tracking changes in an author's style, analyzing genre characteristics, and examining social media posts to understand demographics and motivations (Yadav, 2024). This research examined various stylometric features of the collected emails such as tone, lexical density, lexical diversity, choice of words, repetition, formality level and emotional depth to explore how using AI-powered chatbots has affected the writing style and the linguistic fingerprint of the target group of students. *Nvivo*, a qualitative data analysis software was used to run the stylometric analysis of the emails' texts as well as using readabilityformulas.com (Free readability assessments tool) and Google Sheets to record some key text statistics. The analysis results are presented and discussed in the section below.

4. Results and Discussion
As shown in Figure 1, when asking the 10 participating teachers to select the key characteristics they think feature the emails that are produced by their students using an AI chatbot, the top two characteristics were: too formal and lexically dense and the next three were lengthy, have monotonous tone and lack authenticity. In the "other" category, participants added: "too perfect", "robotic", "complex structure", "vocab not right for the context" and "polite and clear". The high level of formality and politeness, the high lexical density/diversity and the monotonous tone have all been evident in the analyzed sample emails.
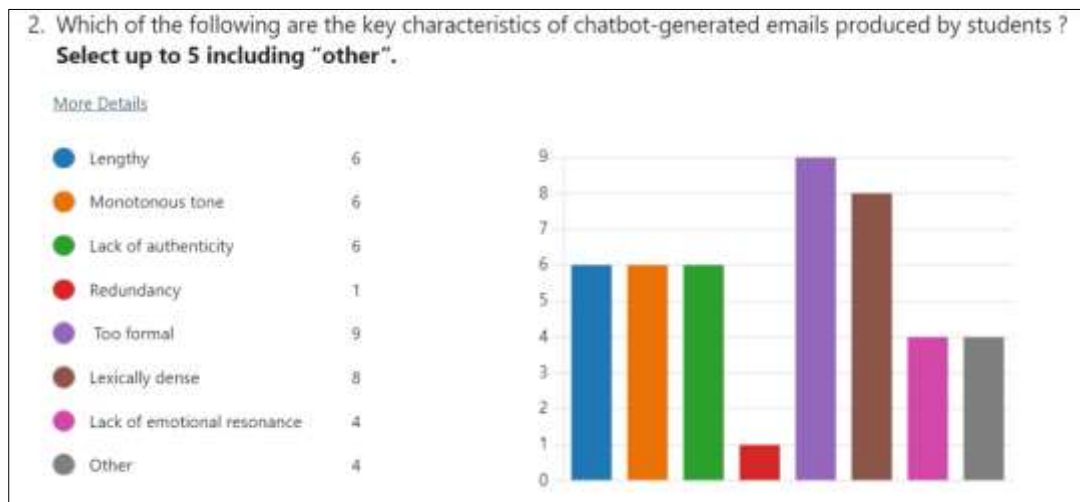

Figure 1 Key Characteristics of Chatbot-Generated Emails

The screenshot below is taken from *Nvivo* software that shows trends in the 25 emails after uploading and coding the emails' texts.

Figure 2 Qualitative Analysis of Sample Emails in *Nvivo*

First, out of the 25 emails, 13 emails (52%) were about excuses for missing classes, whereas 8 emails were about assignment queries. The rest were on other topics like course registration and office hours.

The first most apparent feature is the repetitive structure. Almost all emails had the following structure:

Polite greeting → Statement of issue → Request for assistance → Courteous closing

When exploring the stylistic features of these emails, it is interesting to find that there were 98 phrases of politeness in these 25 emails where few phrases were overused like " *I hope this email finds you well, I look forward to…, I kindly request…, I appreciate your…*". For example, "*I hope this email/message finds you well*" was written 17 times in 17 emails (68%) and "I *hope you understand / I appreciate your understanding*" was used 22 times in 16 emails (64%). There is also one recurring phrase in 6 emails related to the topic of absence which is " *I will catch up on / make up for the missed material*".

The word cloud and the table below illustrate the most common words in the collected sample emails. It is evident that the top 10 words are related to the pilotless phrases and the topic of the emails (absence and attendance).

Figure 3 Word Cloud of the Most Common Words Written in the Emails

Table 2 Top 24 Words Written in the Emails

No.	Word	Count	Similar Words
1	hope	36	hope
2	understanding	34	understand, understanding
3	regards	22	regarding, regards
4	well	22	well
5	attend	20	attend, attendance, attending
6	finds	19	find, finds
7	class	18	class, classes
8	absence	17	absence, absences
9	thank	17	thank, thanks
10	today	17	today
11	best	16	best
12	apologize	15	apologies, apologize
13	appreciate	15	appreciate, appreciated, appreciation
14	matter	13	matter, matters
15	sincerely	12	sincere, sincerely
16	academic	12	academic, academically
17	due	12	due
18	email	12	email
19	grades	12	grade, grades
20	make	12	make
21	wanted	12	want, wanted
2	course	11	course, courses
23	writing	11	write, writing
24	time	11	time

It is also noticeable that 8 emails had some big statements that convey broad ideas which sound unnecessary. Some examples of these statements are :

"I understand the importance of class attendance, and I realize the negative impact of my absence…"
"I fully realize that academic standards require commitment and seriousness…"
"I understand that accurate referencing is crucial…"
"In an effort to create a more comfortable and balanced learning environment for all students…"
"Considering the importance of these courses for my academic standing…"
"Education is a priority for me, and I am committed to staying on track with the curriculum…"

As for the tone , all the emails had a polite, formal, and apologetic tone. Overall, the tone was characterized by a blend of politeness, formality, and humbleness, with a focus on providing explanations and seeking assistance from the recipient (teacher). However, more than 60% of the emails lacked specific details or personal anecdotes which made them less engaging. They lacked emotional depth and urgency that might be present in a more personalized email.

The line charts below show the lexical density and lexical diversity of the emails . According to readabilityformulas.com, lexical density is used to "measures the proportion of content words (nouns, verbs, adjectives, adverbs) to the total number of words in a text, giving a sense of how "packed" the text is with information". (Para.2). A high lexical density means that most words in the text are content words. This indicates a more informative text. Academic or formal texts often have a higher lexical density compared to casual or conversational texts. Lexical diversity , on the other hand, refers to the number of unique words in a text to the number of the total words. A range of 60- 69% shows a moderate and balanced lexical diversity where writers have a balance between varied vocabulary and repeated words.
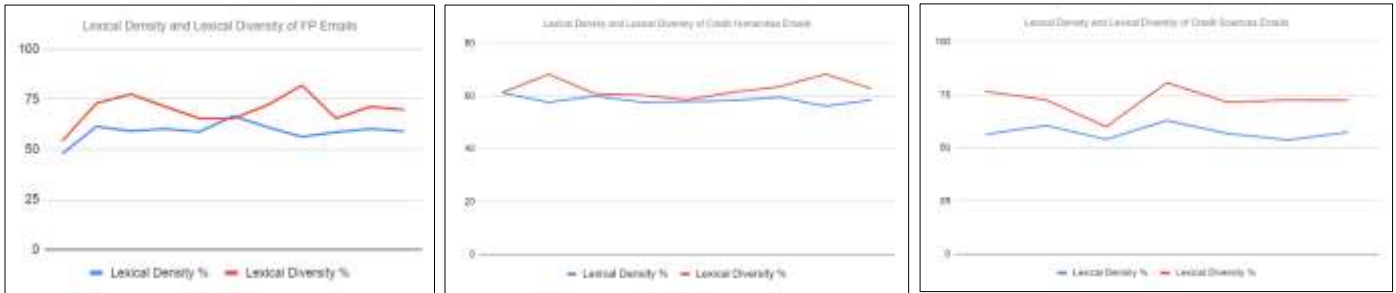
Figure 4 Lexical Density and Lexical Diversity of Emails

Due to the small sample size which is relatively uneven in each group , it is not possible to make a fair comparison regarding the lexical features of emails in the three groups. However, it is obvious from table 3 below that the level of lexical density of all emails is moderate. This is common in everyday formal communications where texts are "well-rounded with varied vocabulary but without excessive repetition. It balances unique terms and common language" . (readabilityformulas.com, n.d). However, the lexical diversity is moderate to high which indicates a relatively high diversity that is usually used in literacy fiction, descriptive non-fiction and high-quality feature journalism.

It is also evident that the FK (Flesch-Kincaid) and FRE (Flesch Reading Ease), used to indicate the readability level of an English text considering the sentence complexity and choice of words, are high. The results are generally high considering the actual writing skills of students in these groups, especially in the FP. In addition, all sample emails are spelling-error free and grammar is almost perfect. However, some emails had punctuation issues where full stops were missing and other punctuation marks like commas and colons are misused.

Table 3 Lexical Density, Lexical Diversity , FK

Emails group	Lexical density %	Scale	Lexical diversity %	Scale	FK	FRE
Foundation Program	59.08		69.85	High diversity	9.5	57.8
Credit Humanities	58.49	Moderate density	62.78	Moderate diversity	10.3	51.3
Credit Sciences	57.45		72.4	High diversity	8.8	60.7

Based on the main results outlined above, the following findings are highlighted:
1. It is apparent that teachers are capable of identifying AI chatbot-generated emails produced by their students. They were able to identify the key characteristics of chatbot-generated emails as revealed in the results of this research which are: high levels of formality and politeness, monotonous tone, lack of authenticity and high lexical density and diversity which seems to be used interchangeably among teachers meaning " a high lexical load".
2. The tone of chatbot-generated emails was all formal, polite, explanatory and submissive. Students showed a high level of politeness and formality which categorize the formulaic language of AI chatbot where particular phrases are used. Although professional and courteous tone is appropriate for formal communications (student- teacher communication), this tone can sometimes come across as overly polished or generic.
3. Most emails followed the same structure including some particular phrases which indicate following a template produced by an AI chatbot that leads to uniformity. The overuse of " *I hope this email finds you well*" and " *I appreciate your understanding*" was evident as explained earlier.

4. The high percentage (52%) of absence-related emails generated by AI indicates that students might be relying on AI chatbots to craft explanations for their absences. This reliance might suggest a need for more personalized communication.
5. It has been observed that these emails had a moderate lexical density of around 60% which is appropriate for the genre (formal emails). However, the level of lexical diversity was high (around 70%) as these emails used a wide range of unique vocabulary compared to the total number of words in the text. Similarly, sentence structures are generally of advanced level. Texts were spelling-error and grammar-error free. This indicates that the use of AI chatbots helps students produce communication that is technically proficient and suited to formal contexts, but might not be indicative of the student's actual language proficiency level. Punctuation was not as good as grammar and spelling where students occasionally had missing full stops. This could probably be a copy-paste issue.
6. The majority of the emails lacked emotional resonance as they did not provide specific personal details or anecdotes which made them monotonous, less engaging and sometimes sound robotic as one of the participating teachers described. The texts tended to be generic and contained some general unnecessary statements which gives the feeling that they are lengthy.
7. To address the key question in this research: Are students losing their linguistic fingerprint? The findings suggest that students may indeed risk losing their unique voice when relying on AI chatbots. The standardized format, highly formal language, and the use of the same polite expressions, coupled with a lack of personal details and emotional depth, render the generated texts strikingly similar. This finding aligns with the existing literature, which indicates that while AI tools are adept at producing coherent and contextually appropriate text, they often fall short in capturing an individual's distinct voice, experiences, and personal nuances unless explicitly prompted to do so.

5. Conclusion
This research paper explored the impact of AI-generated content on email writing among undergraduate English language learners at Sultan Qaboos University. By analyzing 25 student-generated emails using AI chatbots, several key findings and implications have emerged. AI chatbots have proven effective in assisting students with generating well-structured and grammatically-correct emails with moderate to high levels of lexical density which is beneficial for academic and professional communications. However, this reliance on AI tools has also resulted in repetitive structures and formulaic language which has raised concerns about the authenticity and uniqueness of student writing. The lack of personal anecdotes and emotional depth in these communications suggests that students may be missing out on developing their own distinctive voices which may lead to less engaging and personalized content. Therefore, the findings of this research highlight the need for a balanced approach to using AI chatbots in language learning. While these tools can significantly aid in improving the technical and linguistic aspects of writing, educators must ensure that students do not become overly dependent on them. AI chatbots should serve as writing tutor rather than a ghostwriter, so it is important to ensure that students are using AI chatbots mindfully to enhance their language proficiency. Clear student guidelines on the appropriate use of AI chatbots should be developed. These guidelines should clearly define what constitutes acceptable use of AI chatbots and what crosses the line into ghostwriting. It is also crucial to encourage and train students to engage in critical thinking, creativity, and personal expression in their writing to maintain their linguistic fingerprints and unique voices. By doing so, educators can leverage the benefits of AI by ensuring that students continue to develop their writing skills as well as developing their distinctive tone and shaping their linguistic fingerprint.

This research study has some limitations that should be acknowledged. The sample size of 25 emails may not fully represent the diversity of student writing across all language proficiency levels and disciplines within the English foundation program and ESP credit courses at the Centre for Preparatory Studies. Secondly, the study did not compare the AI-generated emails with emails produced by students without the assistance of AI tools, which may pose a limitation in understanding the unique impact of AI tools on student writing. However, the study provides valuable insights into the stylistic characteristics of AI chatbot-generated content in student emails within the specified educational context. It highlights potential concerns regarding authenticity, originality, unique voice, and emotional depth in student communication facilitated by AI chatbots. Future research could expand on these findings by including a broader range of student writing samples, exploring additional stylistic features and comparing them with emails produced by students without the assistance of AI chatbots.

References

[1] AlAfnan, M. A., & MohdZuki, S. F. (2023). Do Artificial Intelligence Chatbots Have a Writing Style? An Investigation into the Stylistic Features of ChatGPT-4 . *Journal of Artificial Intelligence and Technology*, *3*(3), 85–94. https://doi.org/10.37965/jait.2023.0267

[2] Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. Assessing Writing, 57, 100745. https://doi.org/10.1016/j.asw.2023.100745

[3] Bashang, S & K. Puttaana. (2023). Brain Duplication Techniques of Artificial Intelligence: A Comprehensive Review. *International Journal for Research in Engineering Application & Management* (IJREAM), 9(5), 80-88

[4] Baskara, F. R. (2023). Integrating ChatGPT into EFL writing instruction: Benefits and challenges. *International Journal of Education and Learning*, *5*(1), 44-55.

[5] Casal, J. E., & Kessler, M. (2023). Can linguists distinguish between ChatGPT/AI and human writing?: A study of research ethics and academic publishing. *Research Methods in Applied Linguistics*, *2*(3), 100068.

[6] Costa, R., Costa, A. L., & Carvalho, A. A. (2024). Use of ChatGPT in Higher Education: A Study with Graduate Students. In *Digital Transformation in Higher Education Institutions* (pp. 121-137). Cham: Springer Nature Switzerland.

[7] Due, Jinming, & Daniel, B. (2024). A Systematic Review of AI-Powered Chatbots in EFL Speaking Practice: Transforming Language Education. *Computers and Education: Artificial Intelligence*, 100230.

[8] Duong, T., & Suppasetseree, S. (2024). The Effects of an Artificial Intelligence Voice Chatbot on Improving Vietnamese Undergraduate Students' English Speaking Skills. *International Journal of Learning, Teaching and Educational Research*, *23*(3), 293-321.

[9] Ebadi, S., & Amini, A. (2024). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. *Interactive Learning Environments*, *32*(2), 655-673.

[10] Haristiani, N. (2019). Artificial Intelligence (AI) Chatbot as Language Learning Medium: An inquiry. Journal of Physics: Conference Series. https://doi.org/10.1088/1742- 6596/1387/1/012020.

[11] Harunasari, S. Y. (2023). Examining the effectiveness of AI-integrated approach in EFL writing: A case of ChatGPT. *International Journal of Progressive Sciences and Technology (IJPSAT)*, *39*(2), 357-368.

[12] Hawanti, S., & Zubaydulloevna, K. M. (2023). AI chatbot-based learning: alleviating students' anxiety in english writing classroom. *Bulletin of Social Informatics Theory and Application*, *7*(2), 182-192.

[13] Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific reports*, *13*(1), 18617.

[14] Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, *15*(4), ep464.

[15] Johansson, I. R. (2023). A Tale of Two Texts, a Robot, and Authorship: a comparison between a human-written and a ChatGPT-generated text.

[16] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274.

[17] Khalaf, M. A. (2024). Does attitude towards plagiarism predict aigiarism using ChatGPT?. *AI and Ethics*, 1-12.

[18] Khasawneh, M. A. S. (2023). Integrating Ai-Based Virtual Conversation Partners in Enhancing Speaking Skills In Foreign Languages: Insights From University Students. *Journal of Southwest Jiaotong University*, *58*(5).

[19] Kwon, S. K., Shin, D., & Lee, Y. (2023). The application of chatbot as an L2 writing practice tool. *Language Learning & Technology*, 27(1), 1–19. https://doi.org/10125/73541

[20] Liao, W., Liu, Z., Dai, H., Xu, S., Wu, Z., Zhang, Y., ... & Li, X. (2023). Differentiating ChatGPT-generated and human-written medical texts: quantitative study. *JMIR Medical Education*, *9*(1), e48904.

[21] Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, *74*(5), 570-581.

[22] Mahyoob, M., Al-Garaady, J., & Alblwi, A. (2023). A proposed framework for human-like language processing of ChatGPT in academic writing. *International Journal of Emerging Technologies in Learning (iJET)*, *18*(14).

[23] Oritsegbemi, O. (2023). Human Intelligence versus AI: Implications for Emotional Aspects of Human Communication. *Journal of Advanced Research in Social Sciences*, *6*(2), 76-85.

[24] Özçelik, N., & Ekşi, G. (2024). Cultivating writing skills: the role of ChatGPT as a learning assistant—a case study. *Smart Learning Environments*, *11*(1), 10.

[25] Pawar, S. (2024). Deciphering the language of AI: From text data to ChatGPT's generative power. *Medium*. https://medium.com/@sureshkumar.pawar/deciphering-the-language-of-ai-from-text-data-to-chatgpts-generative-power-aba29cdb56de

[26] Rudolph, J., Ismail, M. F. B. M., & Popenici, S. (2024). Higher Education's Generative Artificial Intelligence Paradox: The Meaning of Chatbot Mania. *Journal of University Teaching and Learning Practice*, *21*(6), 1-35. https://doi.org/10.53761/54fs5e77

[27] Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A., & Sari, M. N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. *Journal on Education*, *6*(1), 750-757.

[28] Silitonga, L. M., Hawanti, S., Aziez, F., Furqon, M., Zain, D. S. M., Anjarani, S., & Wu, T. T. (2023). The impact of AI chatbot-based learning on students' motivation in english writing classroom. In *International Conference on Innovative Technologies and Learning* (pp. 542-549). Cham: Springer Nature Switzerland.

[29] Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, 1260843.

[30] Songsiengchai, S., Sereerat, B. O., & Watananimitgul, W. (2023). Leveraging Artificial Intelligence (AI): Chat GPT for Effective English Language Learning among Thai Students. *English Language Teaching*, *16*(11), 1-68.

[31] Taecharungroj, V. (2023). "What can ChatGPT do?" Analyzing early reactions to the innovative AI chatbot on Twitter. *Big Data and Cognitive Computing*, *7*(1), 35.

[32] Thorpe, L. (2023). What are Lexical Density and Lexical Diversity? *Readability Formulas. https://readabilityformulas.com/what-are-lexical-density-and-lexical-diversity/*

[33] Van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. Nature, 614(7947), 224-226. https://doi.org/10.1038/d41586-023-00288-7

[34] Wu, R., & Yu, Z. (2024). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10-33.

[35] Xue, J., Wang, Y. C., Wei, C., Liu, X., Woo, J., & Kuo, C. C. J. (2023). Bias and fairness in chatbots: An overview. *arXiv preprint arXiv:2309.08836*.

[36] Yadav, A. (2024). What is Stylometric Analysis? *LinkedIn*. https://www.linkedin.com/pulse/what-stylometric-analysis-akansha-yadav-3aoec/

[37] Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*(11), 13943-13967.

[38] Yuan, Y., Li, H., & Sawaengdist, A. (2024). The impact of ChatGPT on learners in English academic writing: opportunities and challenges in education. *Language Learning in Higher Education*, *14*(1), 41-56.

[39] Zeb, A., Ullah, R., & Karim, R. (2024). Exploring the role of ChatGPT in higher education: opportunities, challenges and ethical considerations. *The International Journal of Information and Learning Technology*, *41*(1), 99-111.

[40] Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, *10*(3).