A Linguistic Comparison between Human- and AI- Generated Texts in Lebanon

Christelle Frangieh[1] and Manal Abdallah[2]
[1]Lebanese American University, Byblos, Lebanon. e-mail: christelle.frangieh@lau.edu.lb
[2]Balamand University, Balamand, Lebanon. e-mail: Manal.abdallah@fty.balamand.edu.lb

Abstract

On November 30, 2022, OpenAI launched ChatGPT- a leading chatbot that generates human-like output by using Natural Language Processing. While this novel technology has gained popularity worldwide, its integration in the educational field specifically has stirred controversies and mixed reactions. The major problem that many instructors have been recently facing is academic dishonesty in exams and written assignments, which led many institutions and professionals to redefine their policies and change their assessment methods. Considering the paucity of research exploring the language of ChatGPT, especially in Lebanon, the aim of this study is to provide an analysis of the linguistic features of AI-generated texts versus human-generated ones and highlight the divergence between the two. It is therefore essential to provide an insight into the capacities and limitations of AI in order to help instructors understand the platform the students use and spot, or at least question, any possible academic dishonesty. To derive the data corpus for this study, a total of 50 essays were collected from students in a private university in Lebanon and generated using ChatGPT 3.5 given the same instructions. Subsequently, a comparison between the AI- and human- generated writings was done through LIWC-22 (Linguistic Inquiry and Word Count)- a software for analysing word use- in order to investigate and interpret different variables such as pronouns, articles, positive and negative emotions, social words and other categories. Moreover, a comparison was made between different versions of ChatGPT- generated texts using the LMS (Language Matching Style) tool in LIWC to investigate the language of AI. The results have shown that, in terms of word frequencies, ChatGPT employed much more sophisticated words than students, who used simple and basic terminologies. Additionally, the detailed LIWC analysis ran for both has shown that AI tends to be more formal, objective and direct, while students' writing remains more emotional and subjective. Last but not least, the LMS and LIWC analysis of AI-texts have shown very close percentages, denoting the similarities between the AI-generated texts.

Keywords: ChatGPT, human-generated texts, LIWC, linguistic analysis, students

Introduction

"Words are the building blocks of human language, and the richness of human language enables expressions of intricate thoughts and feelings" (Sandler et al., 2024, p.16587). Machines with human capabilities have long impressed technologists. In this digital era, technology, specifically 'artificial intelligence', has become a substantial powerful and a main component in every sector such arts, medicine, business, and education. Several researchers have defined AI in the community of computer science. According to Russell and Norvig (2009), AI refers to computers that mimic cognitive functions that humans associate with the human mind, such as learning and problem-solving" (p. 2). Recent advances in AI have shown interest in its potential to improve the quality of students' education, which makes it welcomed by colleges and universities. However, AI is a loose umbrella term that refers to a collection of methods, capabilities, and limitations—many of which are often not explicitly articulated by researchers, education technology companies, or other AI developers (Gillani et al., 2023). Thus, yet helpful and time- and money- saving, the use of AI has become problematic and has stirred controversies when it comes to the academic life and

students' performance in writing. The major problem that many instructors have been recently facing is academic dishonesty in exams and written assignments, which led many institutions and professionals to redefine their policies and change their assessment methods.

The paucity of research exploring the language of ChatGPT makes this study significant; its importance additionally lies in the fact that it is the first one conducted in Lebanon that aims to provide an analysis of the linguistic features of AI-generated texts versus human-generated ones and highlights the divergence between the two with the use of LIWC. Therefore, this research aims at providing an insight into the capacities and limitations of AI in order to aid instructors understand the platform the students use and spot, or at least question, any possible academic dishonesty.

Literature Review

AI and Natural Language Processing

When analysing the linguistic features of AI- generated texts and human generated ones, it is essential to understand that the richness of human language lies in the way words are built together to enable expressions of intricate thoughts and feelings (Gillani et al. 2023). Starting from this idea and with the advances in generative LLMs (large language models), AI applications such as OpenAI's ChatGPT, Meta's LLaMA, Google's Gemini, Anthropic's Claude, and other programs have demonstrated proficiencies, thus signalling the beginning of new possibilities and challenges in NLP (natural language processing) and artificial intelligence (Chandler et al. 2024).

According to Beysolow (2018), NLP focuses on allowing computers to understand language as humans do: in a "natural" way. Typically, this would refer to tasks such as understanding the sentiment of text, recognizing speech, and generating responses to questions. "The massive uptake in the development and deployment of large-scale NLG (Natural Language Generation) systems in recent months has yielded an almost unprecedented worldwide discussion of the future of society. The ChatGPT service, which serves as Web front-end to GPT-3.51 and GPT-4, was the fastest-growing service in history to break the 100 million user milestone in January and had 1 billion visits by February 2023" (Herbold et al., 2023, p.18617). This fast-growing development of AI applications, and their use by university students created a need among scholars and researchers to study the difference between the human-generated texts and the computer-generated ones. However, the fact that this subject is novel had an influence on the number of studies written about the matter.

When ChatGPT is used, a linguistic examination demands the use of a software to help analyze and study the linguistic differences between humans, who are students in this case, and computers. The examination tool used in this study is LIWC software (Linguistic Inquiry and Word Count) which has the capacity of summarizing and coding the linguistic features in any written or oral text.

A study that examined the difference between ChatGPT generated essays and human generated essays was conducted by Herbold et al. (2023). Their results showed that the writing style of the AI models exhibits linguistic characteristics that are different from those of the human-written essays. The significance of these studies lies in the fact that they can be of a great support for educators and help them cope with the fast-growing power of AI. Understanding the differences between the human language and the computer one helps educators in developing their own teaching concepts and techniques to be able to utilize the AI concepts.

Another study was that of Sandler et al. (2024) who have used LWIC in a study that explores linguistic differences between human- and LLM-generated dialogues, using written dialogues generated by ChatGPT-3.5 as a companion to the Empathic Dialogues dataset. Their research employed Linguistic Inquiry and Word Count (LIWC) analysis, comparing ChatGPT-generated conversations with human conversations. Their findings showed that while human texts are more variable and authentic, ChatGPT demonstrates superior proficiency in areas such as social processes, analytical style, cognition, attentional focus, and positive emotional tone. The authors also concluded that ChatGPT scores higher than humans in linguistic proficiency and cognitive features such as analytical thinking, cognition, and attentional focus.

Why LWIC?

LIWC, a text analysis program that counts words in psychologically meaningful categories, is the software used in this study to analyse the difference between the students' writings and the ChatGPT ones. The main reason for that, and as expressed by Tausczik and Pennebaker (2010), is to detect meaning in a wide variety of experimental settings, including attentional focus, emotionality, social relationships, thinking styles, and individual differences (p.25), which are essential in understanding both the social and cognitive factors that lead constitute a speech.

Methodology

The Corpus and Procedure

A total of 50 texts were collected, generated, and analysed to highlight potential differences between AI- and Human-written texts. An initial sample of 32 writings were obtained from 32 undergraduate students from a private university in Lebanon taking a Communication Skills course. The participants were asked to write an academic one-sided argumentative essay on the topic of 'Online Shopping' of approximately 400 words. The collected texts were then scanned, and the irrelevant ones were excluded, resulting in a final sample of 25 students' writings. Later, the same instructions given to the students were fed into ChatGPT 3.5 to generate the same number of texts given with the same word count too. Subsequently, the entire corpus was analyzed using LIWC.

Statistical Analysis: LWIC

The Linguistic Inquiry and Word Count (LIWC) is a tool for language analysis that entails general linguistic analysis of 115 features. The software includes a dictionary that groups the words into conjunctions, articles, and words carrying positive and negative emotions. These words are even categorized into higher-level groups such as personal pronouns, impersonal pronouns, affect, cognition, and social processes (Tausczik & Pennebaker, 2010). They are also used in formulas to calculate summary variables such as emotional tone, analytical thinking, clout, and authenticity (Boyd et al., 2022). Using LIWC, the corpus of human- and AI- generated texts were both interpreted in order to compare them and further understand AI's language output. Among the 118 linguistic features, the categories that showed null results in human- and AI- generated texts were excluded from the study considering their irrelevancy to the topic and writing style some of which area assents, non-fluencies, netspeak, fillers, death and sexual vocabulary.

Findings and Analysis

LIWC analysis was divided into three main parts: (1) Main text processing, (2) Word frequencies, (3) and Language Style Matching (LSM).

Main Text Processing: Human-Generated Texts vs. AI-Generated Texts

In order to compare humans and ChatGPT writings, LIWC-22 analysis coded and summarized the linguistic dimensions into 118 categories. Subsequently, the obtained scores of AI- and human-generated texts were matched together, thus highlighting the differences between the two. Of the 118 categories, we examined the most relevant categories mainly related to the affect, social, and cognitive areas (see table 1).

Analytic Thinking. This measure shows the extent to which people employ words that denote formal, rational, and hierarchical thinking patterns. Analytical thinking in LIWC is measured through the use of prepositions, articles, big words, connectors, and less usage of emotional, subjective pronouns, and narrative language. The results of students' texts showed that human-generated texts (M=81.6, SD=14.5) are less analytical than ChatGPT-generated texts (M= 87.3, SD=5.4). Despite using more prepositions, articles and connectors in their writings, humans tend to be more emotional and subjective, and their use of big words is lower than AI's.

Clout. Leadership, confidence and social status are very often detected in writings. The results showed that human-generated texts (M=39.4, SD= 14) are almost similar to AI-generated texts (M=40.6, SD= 5.8). Both results imply a humble manifestation of confidence and authority in writing.

Authenticity. In writing, authenticity is the genuine or honest way of expression that is spontaneous and needs no self-regulation or filter. In LIWC, the use of the first-person pronoun and emotional words, and a less formal language are usually associated with authenticity. The findings revealed that both humans and ChatGPT have low scores, yet humans' score (M=22.2, SD=14.8) is greater than AI writing (M=16.6, SD=10.7). This can be also justified by the fact that humans' scores are higher for the use of the first-person pronoun and emotional words and have a less formal style. Consequently, humans remain more true and self-revealing in their writing compared to AI that remains more formal and detached.

WPS. This category refers to the mean average number of words per sentence. In fact, shorter sentences often refer to a clear, more direct communication, while longer sentences show more complex or detailed expressions. It is obvious that humans have more words in

their sentence (M=25.55, SD=7.09) than ChatGPT (M=19.69, SD=1.52). This implies that human writers tend to use more clauses, descriptive elements, and phrases within each sentence. ChatGPT texts seem to have a more straightforward, formal communication. This finding correlates with Alafnan and Mohdzuki (2023), who found that the average number of words per sentence in a text generated by ChatGPT-4.0 is between 16 and 19 words.

BW. Big words are words with more than six letters. A high rate of big words is often linked to less emotions and detachment in writing. It is clear that ChatGPT has a higher rate of big words (M=43.45, SD=2.23) compared to humans (M=23.77, SD=2.99), showing that AI produces less emotional writings, which also verifies the obtained numbers in the emotions dimension.

Emotional language. LIWC-22 provides a detailed analysis of this category, breaking it down into tone and emotion. While tone refers to directly or indirectly sentimental words, whether positive (happy, joy, birthday, marriage) or negative (sad, angry, funeral), emotions refer to the positive and negative subjective feelings of the writer. Humans tend to be more emotional in writing (M=50.22, SD=28) than ChatGPT (M= 46.6, SD=11.6), with a tendency to be more positive.

Social. Another dimension examined by LIWC is the social one. Generally, using social words is related to being outgoing and more socially connected with others. This category entails two major parts: (1) social behaviors such as prosocial behaviors (helping, caring), politeness (thank you, please), interpersonal conflict (argue, fight), moralization (bad, good) and communication (explain, talk, say), (2) social referents including any reference to friends (colleague, friend), family (mother, father), male (he, him) or female (she, her) people. The results showed that humans' writings (M=6.8, SD=1.79) were less social than ChatGPT (M=8.2, SD=0.7). Given that ChatGPT was initially designed to mimic humans and their interaction, the responses obtained appear to be more socially constructed.

Cognition. According to Sandler et al. (2024), cognition encompasses sub-categories such as memory words (reflecting people's beliefs, attention and references to their memories), certitude (reflecting certainty that often shows an insecurity or lack of true information), and all-or-none (reflecting a thinking style that is over-generalized and extreme). The findings revealed that cognition is higher for all areas in human-generated texts (M=13.82, SD= 3.10) than AI ones (M=12.8, SD=1.14).

Pronouns. The use of pronouns has several psychological meanings. The use of '*I*' and '*we*' pronouns refers to oneself and is associated with personal involvement and collective perspective. In the given argumentative essay, students used the first-person pronouns '*I*' (M=0.22, SD=0.8) and We (M=0.15, SD=0.19) while ChatGPT scored zero for both. Moreover, the use of the second-person pronouns '*you*' shows readers' engagement and interaction, and they are found only in human-generated texts (M=0.20, SD=0.56). As for the third-person pronouns, their use indicates reference to external factors, thoughts, characteristics of people or things, with the singular ones being more related to narrative and descriptive texts. Only human-generated texts used the third-person singular pronouns (M=0.5, SD=0.98), and scored higher for the plural ones (M=1.56, SD=1.05) than ChatGPT (M=1.20, SD=0.52). This clearly shows that humans tend to be less objective in their writings compared to AI.

Time. LIWC analysis clearly shows the temporal focus of attention. Both human and AI texts blend different time focuses in writing with a more emphasis on the present. This is normal considering the type of writing involved. Human score (M= 5.42, SD=1.24) was higher than ChatGPT score (M= 3.47, SD=0.66). This proves that humans are more mindful of their task and tend to focus more on their feelings.

Content vs. Style Words. According to Tausczik and Pennebaker (2010), content words are usually nouns, verbs, adverbs and adjectives, used to convey the meaning of a

communication or the message, whereas style words are pronouns, prepositions, articles, conjunctions, and auxiliary verbs, used to focus on the way people communicate and measure their social and psychological worlds. The findings revealed that both humans and AI tend to focus more on the style rather than the content of the writing, with humans' writing being linguistically richer than AI. $\sum$ (style words) = 38.61 > $\sum$ (content words) = 28.62 in human-generated texts, and $\sum$ (style words) = 31.26 > $\sum$ (content words) = 22.45 in AI-generated texts.

Table 1

Results of LIWC-22 Linguistic Analysis

Summary Variables	Human-Generated Texts		AI-Generated Texts	
	Mean	SD	Mean	SD
	n	n	n	n
Analytic	81.6	14.5	87.3	5.4
Clout	39.4	14	40.6	5.8
Authenticity	22.2	14.8	16.6	10.7
WPS	25.55	7.09	19.69	1.52
BW	23.77	2.99	43.45	2.23
Tone	50.22	28	46.6	11.6

Detailed Dimensions	Human-Generated Texts		AI-Generated Texts	
	%		%	
Affect	**4.13**	**1.57**	**2.94**	**0.7**
Pos. tone	2.9	1.5	2.33	0.52
Neg. tone	1.08	1.12	0.6	0.48
Emotion	0.72	0.71	0.64	0.33
Pos. emotions	0.33	0.34	0.30	0.21
Neg. emotions	0.32	0.46	0.33	0.29
Social	**6.8**	**1.79**	**8.2**	**0.7**
Social Behavior	2.1	0.8	3.6	0.79
Social referents	4.24	1.45	4.48	0.74
Cognition	**13.82**	**3.10**	**12.80**	**1.14**
All-or-None	1.16	0.80	0.04	0.13
Certitude	0.47	0.45	0.23	0.25
Memory	0.01	0.05	0	0
Linguistic	**68.19**	**2.17**	**54.82**	**2.37**
Pronouns	7.44	2.09	5.31	0.76
I pron.	0.02	0.08	0	0
We pron.	0.15	0.19	0	0
You pron.	0.20	0.56	0	0
He/she pron.	0.5	0.98	0	0
They pron.	1.56	1.05	1.20	0.52
Articles	8.71	1.9	5.48	0.67
Prepositions	15.08	1.83	13.8	1.19
Conjunctions	7.38	1.82	6.67	1.73
Auxiliary Verbs	7.42	1.69	4.67	1.17
Verbs	13.53	1.71	8.09	1.37
Adjectives	11.16	2.29	11.41	1.67
Adverbs	3.93	0.94	2.95	0.58
Time				
Focus Past	1.98	0.89	0.51	0.25
Focus Present	5.43	1.24	3.47	0.66
Focus Future	1.22	0.71	0.86	0.60

Word Frequencies

Word frequencies is another feature in LIWC-22 that enables researchers to statistically compare two corpora and identify the dominant words. First of all, running LIWC for both datasets shows that a total of 145 words are common between students in human-generated texts, while 295 words are common between AI-generated texts. This underscores the degree of similarity noticed across AI texts, and the diversity and uniqueness of human production.

Moreover, despite the difference in rates, both texts include common words such as online, shopping, product(s), purchase, items, e-commerce, and others. However, a careful consideration of the words reveals the advanced level of words in ChatGPT texts compared to students' versions. Words such as myriad, breaches, phishing, fraudulant, counterfeit, cumbersome, niche, robust, landscape and many others are unlikely to be appear in ESL/EFL human-generated writings. Additionally, AI clearly uses more academically appropriate words such as consumers instead of people, retailers instead of businesses, purchase instead of buy, considerable instead of huge, drawbacks instead of negatives and many other terms. On the other hand, we noticed that students tend to relate the topic to their own experience and the environment they live in (Lebanese and Lebanon), unlike AI which keeps the writing more general to match with different contexts (locally and worldwide). Undoubtedly, ChatGPT has the ability to generate instantly different terms by accessing many eletronic websites and dictionaries, while students only count on their repertoire and bulk of vocabulary. The two word clouds below (see Figure 1 and Figure 2) illustrate the differences between the two.

Figure 1

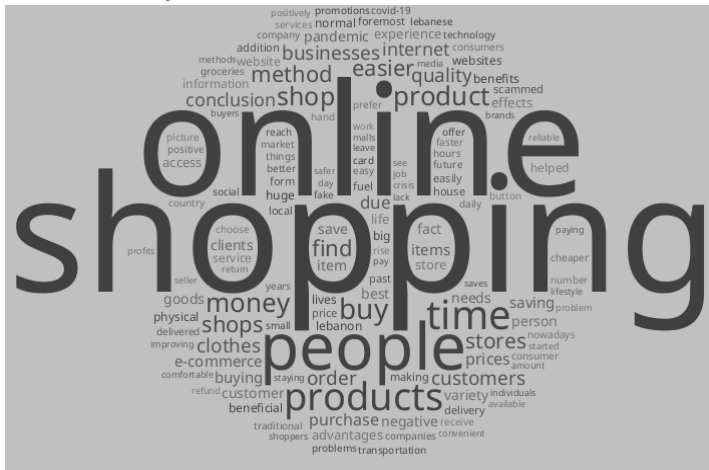*Word Cloud for Human-Generated Texts*

Figure 2

Word Cloud for AI-Generated Texts

Language Matching Style

In LIWC-22, LMS is a metric that measures the extent to which two samples are similar in terms of their writing style. The software typically focuses on the presence of function words such as pronouns, prepositions, auxiliary verbs, articles, and conjunctions, in order to show how people socially interact and structure their ideas rather than the message in their writing. In this paper, the LMS feature was used to look into the writing style of ChatGPT particularly and understand how different versions of the same topic are produced. LMS scores are quantitatively reported from 0 to 1. If the score is closer to 1, it indicates a strong matching, but if it is closer to 0 then it indicates a lower matching. A Pairwise Comparison for all ChatGPT texts shows that the average LMS score is 0.85 (SD=0.05). This high score clearly shows that all AI-generated texts are very much similar in terms of their language style, and the low standard deviation score also proves that all the results cluster around the mean.

Content Analysis of ChatGPT Texts

When it comes to any academic essay, students are expected to produce an introduction, one or several body paragraphs, and a conclusion. The essay that the students had to write was a 1-2-1 one-sided argumentative essay on the topic of e-commerce. This came after explaining and practicing argumentative writing in class. In the argumentative essay given to the students and ChatGPT, we noticed the inconsistency in students' writing style and ideas, as opposed to ChatGPT, which had 25 very similar versions.

A careful examination of the introduction shows that AI-generated texts began the same way, with background information and a thesis statement. For instance, there was one or two general opening statements that highlighted the power of online shopping nowadays ('has surged in popularity', 'has become an integral part of modern consumer behavior', 'has revolutionized numerous aspects of daily life', 'has revolutionized the retail industry', 'has transformed the retail industry'…). Moreover, the essays that argued against e-commerce included a paradox ('while the convenience and accessibility', 'however, despite its advantages'). Lastly, the thesis statement in all AI-generated texts seems to be common and repeated. The thesis began with 'This essay' followed by a reporting verb such as 'argues with/against…', 'explores', and then listed the main points that were expanded in the body parts. Almost all ChatGPT essays tackled the same points: convenience and product variety for the essays arguing with, and security risks and lack of physical product inspection for the essays arguing against. This goes in line with Herbold et al. (2023) who found that 'the initial sentences of each essay are also very similar starting with a general statement using the main concepts of the essay topics' (p.18617). Looking at the human-generated texts, we noticed diversity in ideas and inconsistency in the writing style across all essays. For example, while some essays focused on the worldwide pandemic to introduce this controversial topic, others defined this mode, underscored the paradox, provided a question, or even related the topic to the Lebanese context. As for the thesis statement, all the participants produced different thesis statements, disregarding whether they were correct or not, with most of them listing the body paragraph ideas.

Moving on to the body paragraphs, ChatGPT was asked to generate two body paragraphs as the students did. The findings proved that all AI texts began with a topic sentence that included the main idea of the first paragraph, introduced with a short phrase such as 'one of the primary/major/foremost/most compelling', 'a primary concern' etc. The second body paragraph follows the same style, with the exception of the topic sentence that actually adds to

the previous part. For example, almost all the second paragraphs began with: 'another significant advantage/disadvantage/benefits/drawbacks', 'In addition to X' etc. Both paragraphs continued with very similar explanation and exemplification to elaborate the main ideas. On the other hand, students' essays were not religiously structured the same way as ChatGPT's, but there were some common transitional words and expressions used to introduce the topic sentences such as 'first and foremost', 'moreover', 'another dis/advantage', 'to begin with', 'firstly', 'secondly', and some paragraphs did not even have any transitions. Furthermore, one can notice that AI's paragraphs were more structured and organized, while some students' writings were inconsistent, with some lacking unity, support or even coherence. Last but not least, the conclusions of ChatGPT texts started the same way with a thesis restatement and summary of the main ideas and ended with some recommendations. The results correlate with Herbold et al. (2023) who found that AI essays have identical beginnings of the conclusions. However, in students' essays, the conclusion was kept brief and, in many cases, underdeveloped. This shows that, since the conclusion is the last part of the essay, some students did not write it properly due to time constraints or maybe reduced efforts, something that can never be found in AI because it is known to be swift, instantaneous, and robotic.

Table 2

Extracts from AI- and Human-Generated Introductions

AI-Generated Introduction	Human-Generated Introduction
The advent of the internet has revolutionized numerous aspects of daily life, and perhaps none more so than the way consumers shop. Online shopping, characterized by its convenience and accessibility, has rapidly become a dominant force in the retail industry. *This essay argues that online shopping offers significant advantages, particularly in terms of convenience and product variety, making it a superior alternative to traditional shopping methods.*	Online shopping is one of the most known and used method when talking about shopping. It is a way of getting our needs via the internet without going to the shops and buying products. *Nowadays, this mode of shopping is becoming popular as it can affect people's lives positively because it is time-saving and very accessible.*

Table 3

Extracts from AI- and Human-Generated Body Paragraphs

AI-Generated Body Paragraph	Human-Generated Body Paragraph
One of the most compelling benefits of online shopping is its convenience. Unlike traditional shopping, which requires physical travel to stores, online shopping allows consumers to purchase items from the comfort of their homes. This aspect is particularly advantageous for individuals with busy schedules, limited mobility, or those living in remote areas. Online stores are open 24/7, enabling consumers to shop at any time that suits them best. Furthermore, online platforms often feature detailed product descriptions, customer reviews, and comparison tools, allowing shoppers to make informed decisions without the pressure of in-store sales tactics. This convenience not only saves time and effort but also enhances the overall shopping experience by providing a stress-free environment	*First and foremost, online shopping doesn't take as much time as the traditional shopping.* For instance, it takes less than 1 minute to purchase any online product while having to buy it from any store, even the nearest one, will take minimum 10 to 15 minutes. People who have a full-time job and who have a busy schedule tend to choose online shopping over going to stores due to this time factor. Thus, online shopping has a positive effect on people's lives when it comes to saving time.

Table 4

Extracts from AI- and Human-Generated Conclusions

AI-Generated Conclusion	Human-Generated Conclusion
In conclusion, online shopping offers considerable advantages, such as unparalleled convenience and access to a broader range of products. These benefits have made it an essential aspect of modern consumer behavior. To enhance the online shopping experience, consumers should adopt secure online practices, such as using reputable websites and monitoring their financial	In conclusion, I consider online shopping as a very effective way to buy our needs easily and in a short amount of time. However, online shopping still needs to be ameliorated. Therefore, it is it is important to acknowledge its downgrades in order to improve the industry.

statements for any unauthorized transactions. Retailers should focus on improving their online platforms by providing accurate product information, competitive pricing, and excellent customer service. By addressing these considerations, the online shopping experience can be further optimized, ensuring it remains a convenient and valuable option for consumers worldwide.

Conclusion

This aim of this paper was to compare the linguistic constituents in AI- and Human-generated texts using LIWC-22 software and distinguish between the writing style of both types. This is the first study in Lebanon to quantify and qualify the linguistic features produced by humans and AI.

In terms of the linguistic and stylistic features, the findings of this study correlate with Sandler et al. (2024) and show that, just as in dialogues, humans display more authenticity in their writings, while AI demonstrated more proficiency in social processes and analytical thinking. In fact, the texts produced by AI tend to be more formal, objective and direct. That was mainly noticeable through ChatGPT's use of big words, academic or formal language, and low usage of emotional language and pronouns. Moreover, surprisingly, the high scores of social processes in AI reveal that it can be even 'more human than humans' (Jakesch et al., 2023; Sandlers et al., 2024). Indeed, AI is capable of mimicking humans' writing style but not their emotions. The usage of more pronouns, the positive tone, and temporal focus in humans' writings made it more subjective and emotional. With regards to word frequencies, our results go in line with Alafnan and Mohdzuki's (2023) finding showing that AI includes high lexical density and low lexical diversity. ChatGPT uses more sophisticated terms that tend to be common for multiple versions on the same topic. Last but not least, when examining the content and structure of AI and humans' writings of the same essay, we found that AI is capable of generating very structured essays that have well-developed introductions, body paragraphs and conclusions. Similar to what Herbold et al. (2023) found, AI's texts rigidly realize the sought for model, whereas humans' texts are looser in representing the guidelines on the linguistic level. The consistency in AI offers considerable benefits in writings that require analytical precision. As noted above, the main reason for that is that computers in the end 'are not vulnerable to human variability such as individual differences and fatigue (Sandler et al., 2024, p. 12) as they are programmed to perform one single task: to instantaneously produce complete writings by accessing a wide array of online sources. However, AI's uniformity is also a major drawback, especially when human creativity, depth of thinking, and uniqueness are appreciated and required. In the end, it is in these subtleties- the impulsive, unstructured, and emotional distinctions that are not often analytical but entail cultural context, personal experience, and empathy- that human communication lies.

Several limitations emerged during the study. First, this paper has analyzed a total of 50 AI- and human-generated essays. As such, we cannot claim high generalizability. Thus, further research can expand the sample size and conduct a large-scale study to obtain more accurate findings. Second, the analysis of AI texts was only restricted to ChatGPT 3.5. Future studies can use newer models such as ChatGPT 4.0 or even other AI platforms as well, which might have more technically developed features. Third, LIWC-22 analysis includes 118 linguistic components, yet in this paper, some dimensions were specifically selected for the

analysis. Other works can explore the use of additional dimensions and interpret their psychological and social use in AI and human writings. Finally, the writing generated by AI and the participants is argumentative. Subsequent studies can look into other types of writings such as narrative texts, business reports, critiques, or responses.

References

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*, 85-94. Retrieved from: https://doi.org/10.37965/jait.2023.0267

Beysolow II, T. (2018). *Applied Natural Language Processing with Python*. Apress, Berkeley, CA. doi-org.uobezsecureaccess.idm.oclc.org/10.1007/978-1-4842-3733-5_1

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). *The Development and Psychometric Properties of LIWC-22*. University of Texas at Austin.

Chandler, C., Holmlund, T. B., Foltz, P. W., Cohen, A. C., & Elvevag, B. (2021). Extending the usefulness of the verbal memory test: The promise of machine learning. *Psychiatry Research, 297*, 113743. https://doi.org/10.1016/j.psychres.2021.113743

Gillani, N., Eynon, R., Chiabaut, C., & Finkel, K. (2023). Unpacking the "Black box" of AI in education. *Journal of Educational Technology & Society, 26*(1), 99-111. Retrieved from: https://www.jstor.org/stable/48707970.

Herbold, S., Hautli-Janisz, A., Ute, H., Zlata, K., & Alexander, T. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Sci Rep, 13*, 18617. Retrieved from: https://doi.org/10.1038/s41598-023-45644-9

Jakesch, M., Hancock, J. T., & Naaman, M. (2022). Human heuristics for AI-generated language are flawed. *Psychological and Cognitive Sciences, 120* (11), e2208839120. Retrieved from: https://doi.org/10.1073/pnas.2208839120

Russel, S. J., & Norvig, P. (2023). *Artificial Intelligence: A Modern Approach*. Pearson

Sandler, M., Choung, H., Ross, A., & David, P. (2024). A Linguistic Comparison between Human and ChatGPT-Generated Conversations. *arXiv preprint arXiv:2401*.16587.

Tausczik, Y. R., and Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24–54.