International Journal of Linguistics, Literature and Translation
ISSN: 2617-0299 (Online); ISSN: 2708-0099 (Print)
DOI: 10.32996/ijllt
Journal Homepage: www.al-kindipublisher.com/index.php/ijllt

IJLLT
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

AI-Human Writing Divide: Pedagogical Considerations

Aisah Jamayesah A. Hadji Jamel
Assistant Professor IV, Institute of Science Education, Mindanao State University-Main Campus, Marawi City, Philippines
Corresponding Author: Aisah Jamayesah A. Hadji Jamel **E-mail**: aisahjamayesah.hadjijamel@msumain.edu.ph

| ABSTRACT

The emergence of Artificial Intelligence, or AI, has rapidly transformed the panoramic landscape of academic writing by defying the traditional ideas of authorship and creativity. This study examined the linguistic distinctions between AI-generated and human-written persuasive essays by analyzing five texts from each category. Utilizing computational tools like AntConc and UAM Corpus tool, along with qualitative methods, the study investigated key features of writing: lexical diversity, sentence complexity, nominalization, and the use of modals, epistemic, and discourse markers. Results indicate that human-written texts (HWT) demonstrated significantly higher lexical diversity, as measured by Type-Token Ratio (TTR), suggesting richer and varied vocabulary. Contrastingly, AI-generated texts exhibited a lower TTR but showed higher lexical diversity in keyword analysis, stressing methodological differences in assessing vocabulary richness. Sentence complexity analysis revealed that AI-generated essays tend to have longer sentences with more complex syntactic structures, while human texts contain a greater number of shorter sentences, indicating a more direct communication. Nominalization patterns differ between two corpora: human writers used process nominalizations, highlighting focus on actions and processes relevant to the nature of persuasive essays, while AI-generated texts preferred quality nominalizations, stressing attributes and descriptions. Furthermore, this study also found out that AI-generated texts overused the two modals 'can' and 'must' compared to human texts. In contrast, human writers used a broader range of epistemic and discourse markers, which made their essays natural and had a nuanced expression of certainty, source attribution, and argument coherence. These findings emphasize the need for language teachers to emphasize the flexible and context-appropriate use of linguistic markers, expand students' modal verb repertoire, and deepen their understanding of nominalization types to enhance sophistication in academic writing. This study also highlights the importance of nurturing creativity and critical thinking in writing instruction in light of AI's growing role. In conclusion, the AI-Human writing divide presents both challenges and opportunities for academic writing instruction.

| KEYWORDS

AI, ChatGPT, writing, writing pedagogy, corpus linguistics, lexical diversity, sentence complexity, nominalization, modals, discourse markers

| ARTICLE INFORMATION

ACCEPTED: 15 November 2024 **PUBLISHED:** 07 December 2025 **DOI:** 10.32996/ijllt.2025.8.12.12

1. Introduction

ChatGPT raises important concerns and questions since its early demonstration release on November 30, 2022 (Marr, 2023). This is a reckoning of a class of machine learning Natural Language Processing models known as large language models (LLM), which are a type of Artificial Intelligence (AI) that can recognize and generate texts based on large-scale data and information. In academia, the issue of using ChatGPT as a learning tool, especially in writing, has divided the opinions of educators and others. For some, this may revolutionize and positively impact the learning process of students and simplify the work for teachers to instruct language and other subjects. Others believe that ChatGPT undermines education and is basically a so-called 'high-tech plagiarism, as labeled by Noam Chomsky (EduKitchen, 2023) in one of his interviews, even coming to a point where these people are advocating for legislative measures to ban its use (CNBC, 2023).

Based mostly on anecdotal experiences, many people today perceive AI, specifically ChatGPT, as a machine that can 'effectively' replace humans in performing language and writing tasks. However, based on prior research studies already conducted, there is no comprehensive and extensive linguistic analysis that thoroughly compares text generated by AI and humans across different genres. This type of study would help stakeholders comprehend how well current AI systems, such as ChatGPT, can emulate human language in specific registers.

This study seeks to compare and analyze five (5) AI-generated and five (5) human-written persuasive texts in terms of lexical diversity, sentence complexity, nominalization, presence of modals, epistemic, and discourse markers. Specifically, this study sought to answer the following questions:

1. How do AI-generated persuasive essays compare with those written by students?
2. What do these results imply for language pedagogy?

This study aimed to compare these two types of texts to assess whether an AI-generated text can truly emulate or be indistinguishable from a naturally human-generated text. Another goal of this study is to lay out pedagogical considerations that language educators and policymakers may consider when using Artificial Intelligence (AI), such as ChatGPT, for effective instruction and optimal learning outcomes.

2. Literature Review

The growing intersection of AI and academic writing has sparked complex dialogues that transcend the conventional boundaries of authorship and creativity. This section synthesizes different literature and theoretical frameworks to provide an overview of the panoramic landscape, mainly focusing on the comparative analysis of AI- and human-generated texts in the context of persuasive writing.

2.1. Large Language Model

A Large Language Model (LLM) is an advanced artificial intelligence system trained to understand and generate text in a human-like manner (The University of Alabama at Birmingham, 2024). These models are trained on an extensive body of text, referred to as a corpus, which enables them to provide contextually appropriate responses to various prompts. The most advanced LLMs have been trained on a corpus consisting of publicly available texts on the Internet, which necessitates a huge amount of processing and money.

ChatGPT is a large language model (LLM) trained by OpenAI, an Artificial intelligence (AI) research and deployment company, released in a free research preview on November 30th, 2022, to obtain users' feedback and learn about its strengths and weaknesses (De Angelis et. al., 2023). ChatGPT is one of its kind since it is a one-of-a-kind chatbot optimized for dialog, especially for human-like conversations. In the last five years, the trajectory of LLMs has accelerated, and their performance in performing certain complex tasks has been impressive.

With the increasing popularity of AI-powered language models, ChatGPT has emerged as a potential tool to help students with academic writing. Its ability to comprehend natural language and generate coherent and relevant responses has made it an attractive option for writers seeking to streamline their writing processes (Mondal & Mondal, 2023).

2.2. Persuasive Writing

The goal of a persuasive essay is to persuade the reader to adopt the writer's point of view and convince them to take a specific course of action (Azis & Ahmad, 2017). Over the years, various models of persuasive writing instruction have been formulated to make writing tasks less daunting for students. For instance, in Malaysian schools, the five-paragraph essay model is used to teach students how to write (Smith, 2008). By identifying paragraphs (introduction paragraph, 3-body paragraphs, and conclusion), students learn how to present, develop, and summarize their ideas.

Rhetorical models for essay writing are more focused on the art of rhetoric, which is the art of speaking persuasively (Smith, 2008), by developing persuasive arguments within persuasive composition. One of the models built based on the rhetorical approach is Toulmin's Model (Toulmin, 2003), which was also composed based on the rhetorical approach to arguments. It is considered the definitive model for developing persuasive arguments owing to its practicality, accuracy, ability to generate new ideas, focus on audience, and flexibility (Nimehchisalem and Mukundan, 2011; Karbach, 1987).

Toulmin's model best captures the features of persuasive writing, as it provides scaffolding for substantial reasoning and critical thinking. It was developed by Stephen Toulmin, who identified the elements of a persuasive argument. Toulmin's model is specifically tailored to persuasive writing. It emphasizes the arguers making their claims and supporting the claims with evidence before finally concluding them (Hormazabal, 2007). The model is shown in Figure 1.

Figure 1. Toulmin's Model of Arguments

Figure 1 shows the six elements of Toulmin's Model. The primary components used to compose the basic foundation of an argument are claims, grounds, and warrants (Toulmin, 2003). A claim is the basic purpose of an argument (Karbach, 1987). It is the statement being argued (Toulmin, 2003) and the main point of an argument that the writer attempts to have accepted by the other person (Rafik-Galea et al., 2008; Wood, 2001). This is the central assertion bolstered by the grounds (facts or evidence). The Warrant bridges the Claim and Grounds and is often reinforced by additional evidence termed as Backing. Rebuttals or counterarguments recognize alternate viewpoints, whereas qualifiers refine the claim's ambit. Collectively, these elements craft a coherent and persuasive argument (Sampson and Clark, 2008).

2.3. Corpus Linguistics

Corpus linguistics is a relatively new and fast-changing discipline. As computer resources, particularly web-based ones, develop, sophisticated corpus investigations are within the reach of the ordinary translator, language learner, or linguist. Essentially, it is a methodology or set of methodologies rather than a theory of language description (Hunston, 2006). Essentially, corpus linguistics refers to the following:

- looking at naturally occurring language:
- looking at relatively large amounts of such language.
- observing relative frequencies, either in raw form or mediated through statistical operations; and
- observing patterns of association, either between a feature and a text type or between groups of words.

Corpus linguistics relies on computational methods to analyze linguistic data and has important applications in the field of language technology and AI. One application of corpus linguistics is the development of natural language processing (NLP) systems (Sajjad, 2023). With large datasets available for analysis, researchers can identify patterns and regularities in practical language use that would be difficult to detect in smaller datasets.

3. Methodology

This study employed a comparative analysis with mixed quantitative and qualitative methods as its research design. For the quantitative analysis, two corpus analysis tools were used to identify the following variables: lexical diversity, sentence complexity, nominalization, presence of modals, epistemic, and discourse markers. These are AntConc version 4.2.4 and the UAM Corpus Tool. The latter was used to annotate the presence of modals, epistemic, and discourse markers because AntConc does not have a built-in annotation feature or a sentence complexity analyzer.

To identify the lexical diversity of both corpora, the Type-Token Ratio (TTR) and keyword list feature of AntConc were used. TTR is the total number of different words used (types) divided by the total number of words in the segment (tokens) (Center for Advanced Research on Language Acquisition, 2024). However, it should be noted that the TTR is sensitive to the length of the texts, which might affect the ratio or percentage of the lexical diversity of both corpora.

For sentence complexity, the automatic sentence analysis feature of the UAM Corpus Tool was used. The presence of nominalization, modals, epistemic, and discourse markers was identified using the annotation feature of the UAM Corpus Tool.

For qualitative analysis, a close reading and examination of both corpora were used to identify patterns and linguistic devices that were not captured by quantitative measures. Pedagogical implications are provided based on the outcomes of the computational linguistic analysis.

3.1 Sample Data

The study used five (5) human-written texts and five (5) AI-generated texts. A comparative analysis of both corpora was conducted. To facilitate concise referencing throughout the analysis, the corpus of five human-written texts is labeled HWT, and the corpus of five AI-generated texts is labeled AGT in the remainder of the paper.

The sample data of human-written texts were written by five random Grade 10 students at a Junior High School in Mindanao State University-Main Campus in Marawi City. These students were non-native English speakers. The researcher ensured that these texts were not assisted by any AI tools since the activity was done spontaneously during their English class as part of their discussion on persuasive writing. The texts were originally handwritten on a single paper. During the said activity, the researcher provided a writing prompt to the students on the topic they should work on. The exact writing prompt used was as follows:

You are about to write a persuasive write-up to the Secretary-General of the United Nations calling for a ceasefire in the Gaza Strip with at least three paragraphs with 5-7 sentences each.

The same writing prompt was used to input ChatGPT. The researcher asked ChatGPT to provide five different texts using the same prompt, which became the corpus of AI-generated texts. Table 1 presents the core statistics of both corpora.

Table 1. Core Statistics of HWT and AGT

Corpus	Word Token	Word Types	Sentences in text	Average Word Length
HWT	2,713	934	161	4.88
AGT	2,085	451	75	5.75

3.2 Limitations to consider

This study employed a smaller sample of data for the computational linguistic analysis. Although this study allowed for an in-depth analysis, it may not be fully accepted as representative of the broad spectrum of AI-generated and human-written persuasive texts. Therefore, the findings should be interpreted with caution and considered indicative rather than definitive. Future research with larger sample sizes could provide more robust and generalizable data.

The AI-generated texts in this study were produced using ChatGPT, which is based on the GPT-3.5 architecture. It is important to note that AI models can have inherent biases based on their training data and algorithms. These biases may influence the linguistic features and content of generated texts.

3.3. Ethical Considerations

This study adhered to strict ethical guidelines to protect the rights and privacy of the participants. Two key ethical considerations were addressed. Informed consent was obtained from each student whose texts were used in this study. Students were provided with an explanation of the study's purpose and how their texts would be used in the study.

The anonymity of student writers was a major concern throughout this study. Several actions were taken to safeguard the participants' privacy. All identifying information about the participants, such as their names, was removed before analysis. Moreover, each text was assigned a unique code (HWT1, HWT2, HWT3, HWT4, and HWT5) throughout the remainder of the study. These codes were consistently used in all analyses and discussions to prevent any possibility of identification.

4. Results and Discussions

The first problem sought to compare AGT and HWT in terms of lexical diversity, sentence complexity, nominalizations, and the presence of modals, epistemic, and discourse markers.

4.1. Lexical Diversity

To identify lexical diversity, the study employed the Type-Token Ratio (TTR) and the keyword list feature of AntConc. To obtain the TTR, the total number of word types was divided by the total number of word tokens in the corpus and then multiplied by 100 to obtain the percentage. Table 2 presents the calculated TTR for both corpora.

Table 2. Calculated TTR of HWT and AGT

Corpus	Types	Tokens	TTR
HWT	934	2,713	34.43%
AGT	451	2,085	21.63%

The table above reveals that human-generated texts have a higher Type-Token Ratio than AI-generated texts. This substantial difference of 12.8 percent indicates a notable disparity in lexical diversity between the two text types. This further suggests that human-generated texts demonstrate higher lexical diversity and potentially richer vocabulary. The lower TTR in AI-generated texts might indicate that ChatGPT tends to reuse words more frequently, resulting in a less varied vocabulary because of the probabilistic model it follows.

The higher lexical diversity of HWT might suggest that the writers tend to draw their persuasive writing from personal observations and emotions, or that they consciously vary the words for stylistic purposes. It should also be noted that the utilization of TTR is sensitive to the length of the texts. Because the HWT corpus consists of longer texts, it could affect the TTR calculation.

Another way to identify lexical diversity is to conduct a keyword analysis of both corpora using the keyword analysis feature of AntConc. According to Xu et. al. (2018), keyword list feature in AntConc is a tool that shows which words are unusually frequent (or infrequent) in the corpus in comparison with the words in a reference corpus. This allowed the researcher to identify the characteristic words in the corpus. To compare the results of both corpora, the AGT was loaded first as the target corpus, followed by the HWT as the reference corpus. After obtaining the results, the researcher swapped the two corpora, with HWT becoming the target corpus and AGT the reference corpus. The results were then compared. Table 3 provides the results of the first procedure (AGT as the target corpus and HWT as the reference corpus).

Table 3. Keyword List of AGT as target corpus

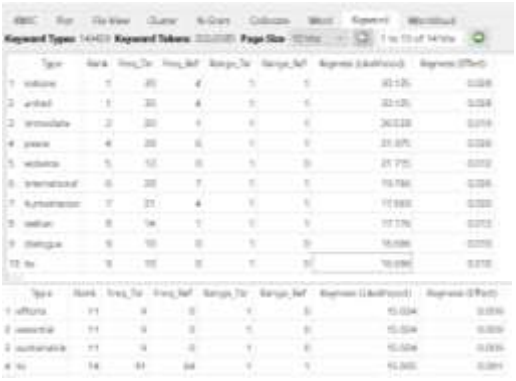

Table 3 reveals that there are fourteen (14) keyword types identified by AntConc, where AGT acted as the target corpus and HWT as the reference corpus. The word 'Nations' ranked first with a keyness (Likelihood) of 31.135 and a keyness (Effect) score of 0.028. It was followed by the words united, immediate, peace, violence, and international. Table 4 presents the results where HWT acted as the target corpus and AGT as the reference corpus.

Table 3. Keyword List of HWT as target corpus

KWIC Plot File View Cluster N-Gram Collocate Word Keyword Wordcloud

Keyword Types 6/875 Keyword Tokens 130/2713 Page Size 10 hits 1 to 6 of 6 hits

	Type	Rank	Freq_Tar	Freq_Ref	Range_Tar	Range_Ref	Keyness (Likelihood)	Keyness (Effect)
1	are	1	48	4	1	1	28.570	0.031
2	they	2	24	0	1	0	27.460	0.018
3	israel	3	19	0	1	0	21.724	0.014
4	more	4	15	0	1	0	17.148	0.011
5	palestine	4	15	0	1	0	17.148	0.011
6	will	6	14	0	1	0	15.996	0.010

The table above reveals that only six (6) word types were identified by AntConc as keywords. The word 'are' ranked first in the list with a keyness (likelihood) score of 28.570 and keyness (effect) score of 0.031. Words such as are, they, Israel, more, Palestine, and will, were also included in the list.

It is evident that there is a significant difference between Tables 3 and 4 in terms of the number of keyword types. Based on the two results, the keyword list feature implies that AI-generated texts have higher lexical diversity than human-written texts. This result did not align with the previous result of TTR, where HWT had higher diversity than AGT. This emphasizes the complexity of the lexical analysis. This discrepancy suggests different interpretations. First, this may indicate that TTR and keyword analyses measure different aspects of lexical richness.

TTR focuses on overall vocabulary variety, whereas keyword analysis compares the frequency of words between corpora. The TTR is also sensitive to text length, whereas the keyword analysis is not affected by this condition. It might also be affected by the use of function and content words. The calculation of TTR included all words, whereas the keyword analysis focused more on the content words. According to Scott and Tribble (2006), a text might have a high TTR due to varied function words but fewer keywords.

4.2. Sentence Complexity

The next variable examined in this study was sentence complexity. Sentence complexity refers to the structural intricacies of sentences in a text. According to Lu (2010), sentence complexity is characterized by the number and type of clauses, depth of embedding, and variety of syntactic structures used within sentences. More complex sentences typically contain multiple clauses, subordination, and diverse syntactic patterns.

To identify the sentence complexity of both corpora, the automatic sentence analyzer of the UAM Corpus Tool was used. Table 5 presents the results.

Table 4. Sentence Complexity Results of Two Corpora

Corpus	Sentences in text	Average Sentence Length
HWT	161	16.6
AGT	75	27.5

The table above reveals that HWT contains more than twice (161 sentences) as many sentences as AGT (75 sentences). This may imply that the students who wrote these texts used fewer but more sentences. In contrast, AGT gained a higher percentage of average sentence length, implying that ChatGPT tends to produce more complex sentences. Based on the corpus, ChatGPT provided more descriptive details with more clauses and intricate syntactical structures. This made its sentences longer, which might reflect its training on formal or academic texts. Below are some extracts from sentences produced by ChatGPT:

"The current escalation of violence between Israeli forces and Palestinian groups has reached a critical point, leading to devastating humanitarian consequences for civilians on both sides." [AGT1]

"It is imperative that the United Nations, as the beacon of global peace and security, takes decisive action to call for an immediate ceasefire and create conditions conducive to dialogue and reconciliation." [AGT2]

"Therefore, I urge you, as Secretary-General, to mobilize the international community in support of a ceasefire and to engage with all stakeholders to advance peacebuilding efforts in Gaza." [AGT3]

The shorter average sentence length of HWT demonstrates that the students used simpler and more concise sentences. This may indicate the readability of the text corpus compared to the texts in the AGT. Crossley et. al. (2011) asserted that shorter sentences are often associated with clearer, more direct communication. Below are some extracts from the HWT corpus:

"Millions of people are going on with their lives and fulfilling their daily jobs for survival." [HWT2]

"Where is justice in all of this? The healthcare facilities are not excluded from the abomination of the genocide due to destruction and overabundance." [HW4]

"This resolution was the root of all the conflicts and suffering that have followed that proclamation. Years have passed, and the Zionists were clearly not satisfied with the land that they were given for free."

Student writers appear to use more varied sentence structures, alternating between short and long sentences for stylistic effect. On the other hand, longer AGTs are more likely to be informationally dense but offer less readability to users because of the use of more formal words.

4.3. Nominalizations

Nominalization, or the process of creating nouns from adjectives or verbs, is a characteristic of academic writing (Bychkovska & Lee, 2023). By removing the agent of the action, nominalization allows writers to shift the focus from human participants to abstract objects that carry the meaning of actions or processes (Biber et al. 1999). Halliday (1994) classified five types of nominalizations: nominalization of process, nominalization of quality, nominalization of circumstances, nominalization of relator, and nominalization of zero.

Process nominalization is often realized by transferring a verb to a noun. Quality nominalization refers to any nominalized word or word group that is derived from an adjective. Circumstance nominalization is often realized by the transference of a prepositional phrase to a noun. Relator nominalization refers to any nominalized word or word group derived from a conjunction. Zero nominalization is often realized by transferring zero to a noun (Yue & Zhang, 2019).

In both corpora, nominalization was examined and compared using the UAM Corpus Tool. All nominalizations across both corpora were annotated. After annotation, the UAM Corpus Tool generated statistical results for each dataset. Table 6 presents these results for nominalization.

Table 5. Nominalization Results of AGT and HWT from UAM Corpus Tool

	Texts/sample corpus (		Texts/sample corpus (	
Feature	N	Percent	N	Percent
NOMINALIZATION	N=84		N=112	
process	46	54.76%	74	66.07%
quality	30	35.71%	37	33.04%
circumstance	7	8.33%	1	0.89%
relator	0	0.00%	0	0.00%
zero	1	1.19%	0	0.00%

Table 6 shows the number of nominalizations and the most common type of nominalization observed in the corpora. The first sample refers to the AWT, and the second sample refers to the HWT. It can be gleaned from the table that HWT has a higher number of nominalizations with a total of 112, while AWT has only 84. The majority of nominalizations used in both samples were process nominalizations, which is the conversion of a verb to a noun, followed by quality nominalization, which is the transference of an adjective to a noun.

It can also be gleaned that process nominalization is most common among the five samples of the HWT corpus, while quality and circumstance nominalization are more common among the AGT corpus. Neither sample employed relator nominalization, and there was one instance of zero nominalization in the AGT corpus.

The high prevalence of process nominalization signals that the texts in HWT frequently present actions or ongoing processes. In the HWT corpus, the student writers typically focused on different processes, which can be observed through the presence of process nominalized words such as ceasefire, intervention, interconnection, advocacy, communication, transportation, and decision. This result implies that the HWT corpus emphasizes the process and needed actions for a ceasefire in the Gaza Strip, since this was the main topic or writing prompt that they wrote about.

In contrast, the AGT corpus has a higher prevalence of quality nominalization than the HWT corpus. This indicates that the AI-generated texts from ChatGPT focused on the properties or qualities of the topic it was 'fed.' The presence of words in this sample such as depth, tension, stability, dignity, difference, urgency, hostility, etc. suggests that ChatGPT emphasized attributes or descriptions of the ongoing war in Gaza, Palestine and how these descriptions affect the actions and decisions of the Secretary General of the United Nations.

The small number of circumstance nominalizations, zero nominalization, and zero presence of relator nominalization suggest further emphasis on those types of nominalizations in language and grammar pedagogy. When teaching academic writing, it is important to discuss the significance of integrating various types of nominalizations into students' academic writing.

The high frequency of process and quality nominalizations in both corpora suggests that the texts tend to focus on actions, events, and attributes. This could be indicative of an academic and formal writing style, as the genre is persuasive writing.

4.4. Modals, epistemic, and discourse markers

This section discusses the linguistic markers observed in both corpora. These include modals, epistemic, and discourse markers. According to Herbold et al.. al. (2023), modals and epistemic markers signal the commitment of the writers to their statements. Modal verbs and expressions such as may, might, can, could, should, and might indicate the degree of certainty, possibility, and necessity of an action or state. They help convey the writers' stance and level of commitment to their claims.

On the other hand, epistemic markers refer to words and phrases that indicate certainty and/or the source of knowledge, such as certainly, possibly, and according to. In this study, phrases such as 'I think', 'I believe', 'In my opinion', etc., were accepted as epistemic, as adopted in the study of Hautli-Janisz et al.. al. (2022). These epistemic markers establish the credibility of writers, which can make their writing more trustworthy.

Finally, discourse markers (i.e., however, moreover, in addition to, etc.) connect ideas and guide the readers through the texts. They help craft a logical flow throughout the text, making the arguments easy to follow and comprehend. They aid in the overall coherence and persuasiveness of texts. Table 7 shows the results of the two corpora (AGT and HWT) in terms of the number of modals, epistemic, and discourse markers.

Table 6. Obtained results of number of modals, epistemic and discourse markers from UAM Corpus Tool

	Texts/sample corpus (		Texts/sample corpus (	
Feature	N	Percent	N	Percent
MARKERS-TYPE	N=64		N=94	
modals	20	31.25%	13	13.83%
epistemic	22	34.38%	45	47.87%
discourse	22	34.38%	36	38.30%

Table 7 shows the number of modals, epistemic, and discourse markers observed across the two corpora. The first sample refers to the AGT corpus (n = 64), and the second sample refers to the HWT corpus (n = 94). From the table, it is observed that HWT corpus contains more markers than the AGT corpus.

There is a frequent occurrence of modals (31.25%) in the AGT corpus, as compared to the HWT corpus (13.83%). In fact, only two modals were used interchangeably throughout the two corpora: can and must. The modal can obtain 14 occurrences (70%) in the AGT corpus, while it was only used three times (23.08%) in the HWT corpus. Must occurred six times (30%) in the AGT

corpus and only once (7.69%) in the HWT corpus. This underuse of modals in the two corpora signals a need to emphasize the importance of modals in establishing writers' stance and commitment to persuading readers to accept and act upon the claim in their writing.

In the case of epistemic markers, the HWT corpus heavily relied on these markers, as it obtained 47.87%, compared to the AGT corpus, which only had 34.38% of the usage of epistemic markers. This indicates that the HWT corpus highlights the expression of degrees of certainty and sourcing information. Below are some of the extracts from the HWT corpus that utilized epistemic markers:

"Your assertion of the Hamas attack arguably has an interconnection to historical events and contemporary actions" [HWT1]

"Allow me this opportunity to directly address an urgent matter regarding the ongoing conflict between Israel and Palestine. The escalation of death toll within these countries, specifically in Palestine, has caused them nothing but grievance and loss." [HWT2]

"Today, Palestine only owns two regions, the illegally occupied West Bank and the Gaza Strip. The Gaza Strip was described by Prime Minister David Cameron as "an open-air prison." [HWT4]

These extracts manifest student writers' use of epistemic markers to establish credibility and trustworthiness in their writing. They also used these markers to specify information about the situation in Gaza and be certain about their intention in writing a plea for ceasefire to the Secretary-General of the UN.

Discourse markers were more prevalent in the HWT corpus (38.30%) than in the AGT corpus (34.38%). Some common discourse markers used in both corpora were moreover, however, in conclusion, and therefore. Even though there is only 3.92% difference in terms of using discourse markers, it should be noted that the HWT corpus has a balanced dispersion of its discourse markers throughout the corpus, while the AGT corpus, which came from ChatGPT, has a rigid format of using discourse markers as it strictly abides by the prompt it was given to.

Compared to the AGT corpus, the HWT has a natural flow of information as it uses these markers to connect ideas and discussions pertaining to the topic of persuading the Secretary General of the United Nations to consider calling for a ceasefire.

4.5. Implications to Language Pedagogy

The results of this study provide a clear distinction in terms of lexical diversity, sentence complexity, nominalizations, and the presence of modals, epistemic markers, and discourse markers between human-written texts (HWT) and AI-generated text (AGT) corpora.

In terms of lexical diversity, it is still difficult to pinpoint whether one is more diverse than the other, since it depends on the perspective of the researcher and the method they utilize in the exploration of this concept. In this study, TTR and AntConc's keyword list were employed to determine the lexical diversity of both corpora, and both methods provided different answers. In the calculation of TTR, the HWT corpus manifests higher lexical diversity, while the AGT corpus has higher lexical diversity when both corpora are examined using a keyword list.

In language pedagogy, this may suggest that ChatGPT offers a highly structured format, resulting in the overuse of words in every persuasive essay it provided (i.e., overabundance of in conclusion at the last paragraph and almost identical beginnings and middle paragraphs). The corpus of human-written texts has a more diverse use of words since the students consciously feel the urge to change the words in their paragraphs, resulting in lexical diversity. Language educators should remind their students who often use ChatGPT or any other LLMs about this observation, especially if they work on highly academic writing such as persuasive writing.

The HWT corpus is also higher in terms of sentence complexity than the AGT corpus. Even though there was a writing prompt that both students and ChatGPT should follow, it is crucial to recognize that student writers tend to write longer than ChatGPT. ChatGPT was rigid in the number of paragraphs it provided, while students did not abide by the guidelines provided by the teacher gave them to do. The ChatGPT has the same number of paragraphs which is consistently four paragraphs, while each student writer wrote a different number of paragraphs in his/her persuasive essays.

This implies that ChatGPT strictly follows the guidelines that a user tells it to do, which is highly evident and common among Artificial Intelligence such as ChatGPT, as opposed to human-written essays, which are looser in representing the guidelines on the linguistic surface.

This result has several implications. The first is the human authenticity in writing academic texts. The fact that students deviated from the guidelines, whereas ChatGPT adhered strictly to the rules, highlights the importance of teaching students to balance following instructions and expressing their creativity and opinion, which can emphasize personal expression and authenticity in writing. The second is complexity and elaboration. The higher sentence complexity and longer texts produced by students indicate that language teaching should continually emphasize the use of different sentence types, such as simple, compound, complex, and compound complex. Finally, this could open up opportunities to teach writing strategies, AI literacy, and ethical use of Artificial Intelligence such as ChatGPT.

The third variable examined in this study was nominalization. From the previous section, it was revealed that the HWT corpus had a higher usage of process nominalization, while AGT focused more on using quality nominalization. There was little use of circumstance nominalization and almost no use of relator and zero nominalization.

These results suggest several implications for language pedagogy. Since nominalization is not frequently used by students, it is important for language educators to focus on teaching students the types of nominalizations and their significance in academic writing, such as in persuasive writing, so that students can develop a more nuanced understanding of how to use language effectively, especially in academic writing. This could also lead to more sophisticated and varied uses of language. With the prevalence of LLMs among students, teachers can use these AIs to tackle important issues of using nominalization and to compare both corpora for educational purposes.

The last set of variables explored in this study is the presence of modals, epistemic markers, and discourse markers. From the previous section, it was revealed that modals were more prevalent in the AGT corpus. However, there were only two modals that ChatGPT kept on utilizing: modals like can and must. In using epistemic markers, the HWT corpus had more instances of using this type of marker in persuasive writing. Student writers used these markers to express certainty and provide sources of information. In the case of discourse markers, the HWT corpus showed a balanced dispersion of markers (i.e., moreover, therefore, in conclusion, etc.) throughout the texts. It was also revealed that the HWT corpus used more discourse markers when connecting ideas to make logical links. Thus, the study infers that the HWT corpus demonstrates a more natural flow of information using these markers.

Several implications for language pedagogy come to mind. First, the importance of using modals, epistemic, and discourse markers in writing persuasive essays should be emphasized. As can and must were the only modals observed in both corpora, teachers should take this opportunity to expand the range of modal verbs used beyond just 'can' and '' must.' While both corpora used discourse markers, the more balanced dispersion in human writing suggests that there is a need to teach the flexible and context-appropriate use of these markers. Emphasizing the significance of epistemic markers in human writing to provide specific information would develop students' critical thinking and research skills when writing an academic essay.

5. Conclusions

This study warrants that there is really such a thing as the AI-Human writing divide, where these two types of writing possess differences and unique characteristics. This divide offers important insights into the nature of both types and has implications for education, assessment, and the future of writing. Educators need to change how they approach writing assessments as more and more students employ the help of these Artificial Intelligences such as the most common one, ChatGPT, in making academic write-ups. Teachers should also develop the skill to immediately detect what is AI-generated and what is human-made.

Based on the results of this study, it can be concluded that both types of writers consist of distinctions in their approach to academic writing differently. ChatGPT demonstrated a more structural pattern and repetitive word usage, resulting in more rigid writing with lower creativity. On the other hand, human-written texts, from those who wrote these persuasive texts, contained more diverse word usage, exhibited higher sentence complexity, and used more process nominalizations and markers like epistemic and discourse markers.

As a recommendation, educators should take action to change how they assess students' writing, considering AI prevalence. They can take this opportunity to develop students' critical thinking and writing skills without the need to fully prohibit students from getting help from an AI. Educators should also focus on developing skills that are distinctly human, such as creativity, personal expression, and critical thinking, while also educating students on how to effectively use and integrate AI in their

pedagogical methods. This divide may someday be an initial step toward revolutionizing language pedagogy through a harmonious link between AI and human capabilities.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
ORCID iD: 0009-0005-6621-6722
Publisher's Note: All claims expressed in this article are solely those of the author and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Azis, F. I. & Ahmad, K. U. (2017). Persuasive writing: How students argue. *Sains Humanika*. 9: 4-2 (2017), 19-32. e-ISSN ISSN: 2289-6996.
[2] Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English.* Longman.
[3] Bychkovska, T. & Lee, J. J. (2023). Nominalization in high- and low-rated L2 undergraduate writing. *International Journal of English for Academic Purposes*. ISSN 2634-4610.
[4] Center for Advanced Research on Language
[5] Acquisition. (2024). "Lexical complexity (TTR) on oral and written versions of a narrative task." *University of Minnesota.* Retrieved from: https://carla.umn.edu/learnerlanguage/per/comp/activity4.html
[6] CNBC. (2023). "Italy became the first Western country to ban ChatGPT. Here's what other countries are doing. Retrieved from: https://www.cnbc.com/2023/04/04/italy-has-banned-chatgpt-heres-what-other-countries-are-doing.html.
[7] De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. *Frontiers in public health*, 11, 1166120. https://doi.org/10.3389/fpubh.2023.1166120.
[8] EduKitchen. (2023). Chomsky on ChatGPT, Education, Russia, and the unvaccinated [Video]. YouTube. https://www.youtube.com/watch?v=IgxzcOugvEI&t=1182s
[9] Halliday M A K. (1994). *An Introduction to Functional Grammar (2nd edition).* London: Edward Arnold.
[10] Herbold, S., Hautli-Janisz, A., Heuer, U. et al. (2023) A large-scale comparison of human-written versus ChatGPT-generated essays. *Sci Rep 13*, 18617 (2023). https://doi.org/10.1038/s41598-023-45644-9
[11] Hormazabal, R. D. (2007). Argumentative writing strategies and perceptions of writing academic by EFL college students, *Literatura y Linguistica*, 18, 253-282.
[12] Hunston, S. (2006). *Encyclopedia of Language & Linguistics* (Second Edition).
[13] Karbach, J. (1987). Using Toulmin's Model Of Argumentation. *Journal of Teaching Writing,* 6(1), 81-91.
[14] Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496.
[15] Marr, B. (2023). "A short history of ChatGPT: How we got to where we are today". Retrieved from: https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/
[16] Mondal, H., & Mondal, S. (2023). ChatGPT in academic writing: Maximizing its benefits and minimizing the risks. *Indian journal of ophthalmology*, 71(12), 3600–3606. https://doi.org/10.4103/IJO.IJO_718_23
[17] Nimehchisalem, V., & Mukundan, J. (2011). Determining the evaluative criteria of an argumentative writing scale. *English Language Teaching*, 4(1), 58-70.
[18] Rafik-Galea S., Zainuddin, S. Z., & Galea, P. V. (2008). Learning to Think Critically The Toulmin Way. *13th Seminar International Conference on Thinking.* 2008. Sweden.
[19] Sajjad, H. M. Linguistics, Language Models, and Artificial Intelligence. *Medium.* Retrieved from: https://medium.com/@hmsajjad/on-the-intersection-of-linguistics-language-models-and-artificial-intelligence-an-overview-d3c89c67690d
[20] Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education*, 92(3), 447–472
[21] Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education.* John Benjamins.
[22] Smith, S. N. (2008). The art of rhetoric: Understanding the how of persuasive writing. Retrieved from: http://www.ic.arizona.edu/ic/snsmith/rhetanal/AristotleChartForm.pdf
[23] The University of Alabama at Birmingham. (2024). "Artificial Intelligence (AI) in Education". *UAB Libraries.* Retrieved from: https://guides.library.uab.edu/ai/llm
[24] Toulmin, S. (2003). The uses of argument. Cambridge, England: Cambridge University Press.
[25] Wood, N. V. (2001). *Perspectives on argument.* (Third Edition). Upper Saddle River, New Jersey: Prentice Hall.
[26] Xu, L., Chen, P., Liu Fang, Zhang, Z., Zhang Y. 'Keyword List'. *AntConc Manual.* Retrieved from: https://antconc-manual.readthedocs.io/en/latest/keyword_list.html
[27] Yue, L. & Zhang, Y. Realization of Nominalization Functions in Abstracts. *International Journal of Language and Linguistics*. Vol. 6, No. 4. doi:10.30845/ijll.v6n4p22