JOALL (JOURNAL OF APPLIED LINGUISTICS AND LITERATURE)
Vol. 10 No. 2, August 2025
ISSN (print): 2502-7816; ISSN (online): 2503-524X
Available online at https://ejournal.unib.ac.id/index.php/joall/article/view/39244
http://doi.org/10.33369/joall.v10i2.39244

Lexical Bundle Structure and Function of ChatGPT-Generated Essay: Corpus-Based Study of Advanced Learners in Indonesia

[1*]Syafina Azalia, [2]Pratomo Widodo, [3]Pangesti Wiedarti, [4]Titik Sudartinah

[1,2,3,4]Applied Linguistics Major, State University of Yogyakarta, INDONESIA
Jl. Colombo No.1, Karang Malang, Caturtunggal, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55281

ARTICLE INFO

Article history:
Received: January 25th, 2025
Revised: May 5th, 2025
Accepted: May 7th, 2025

Keywords:
Academic Writing
Linguistic Features
Lexical Bundles
ChatGPT

Conflict of interest:
Linguistic Structure of AI-Generated Text

Funding information:
Type here

Correspondence:
Syafina Azalia
syafinaazalia.2023@student.uny.ac.id

ABSTRACT

Numerous studies have primarily focused on investigating the linguistic structures of academic writing, such as theses, dissertations, and articles from reputable journals authored by individuals ranging from novices to experts. However, few studies have examined the linguistic structures of texts written by AI, particularly those produced by ChatGPT. This study aims to analyze the lexical bundle structure and function variations of 10 argumentative essays generated by ChatGPT, which were composed based on prompts provided by advanced learners in Indonesia, specifically graduate students of applied linguistics from Yogyakarta University. The objective is to explore the lexical bundle structure and function of ChatGPT-generated essays to understand how the AI employs various linguistic forms and constructs relationships between sentences. For data collection, this study utilized the corpus tool AntConc's N-gram feature to refine the four-word lexical bundles from the collected essays. The analysis used the frameworks of Hyland (2008b) and Biber et al. (1999). The result showed the Noun Phrase + pattern as the most frequent lexical bundle structure found in ChatGPT-generated essays. This structure has been used primarily to assert ideas and positions on the discussed topics (Stance Features), marking that ChatGPT has a strong ability in presenting and elaborating ideas. However, the least frequently found lexical bundle function is text construction that can link one idea to others to create the connection. Therefore, this study suggested that ChatGPT-generated essays can assist in effectively conveying the writer's perspectives but need attention or development in creating cohesive links between ideas in written discourse.

How to cite (APA Style):
Azalia, S., Widodo, P., Wiedarti, P., & Sudartinah, T. (2025). Lexical Bundle Structure and Function of ChatGPT-Generated Essay: Corpus-Based Study of Advanced Learners in Indonesia. *JOALL (Journal of Applied Linguistics and Literature), 10*(2), 515–533.
https://doi.org/10.33369/joall.v10i2.39244

INTRODUCTION

Many students encountered significant challenges in producing academic writing (Bisriyah, 2022). They have difficulties in organizing ideas coherently, limited vocabulary, grammatical errors, and struggles with maintaining academic tone and formality (Aldabbus & Almansouri, 2022; Alqasham et al., 2021; Alsariera & Yunus, 2023; Anindita, 2024; Fauzan et al., 2022). These problems often stem from inadequate writing practice, exposure to academic texts, and insufficient feedback (Fitriyah et al., 2023; Lee, 2021). Hence, several scholars and experts must be innovative in selecting instructional lessons and resources, as well as teaching approaches that make the writing class more fascinating, thrilling, and pleasurable (Limbong et al., 2023).

One notable endeavour that is mostly proposed by academics is the integration of digital tools to assist students' efforts in producing credible writing text (Rahim et al., 2023). Digital tools offer innovative ways to support and improve students' writing abilities (Wahab et al., 2024; Zakaria et al., 2016). Ramadhan et al. (2024) pointed out several tools, including grammar checkers, citation managers, plagiarism detectors, and collaborative platforms to help students produce more structured, academically compliant writing. Not only that, since its launch in 2022, ChatGPT, a chatbot powered by Generative Artificial Intelligence (Gen AI), has been popular in educational settings due to its capabilities in writing assistance (Bom, 2023; Punar Özçelik & Yangın Ekşi, 2024; Rababah et al., 2024).

Wu (2024), Gayed et al. (2022), and Irfan et al. (2020) identified that ChatGPT can provide students with personalized writing guidance and feedback to help them clarify their thoughts and improve their expression skills. It also enhances various writing constructs such as coherence, cohesion, grammar, syntax, and writing conventions, allowing students to refine their skills effectively (Barrot, 2023; Chia et al., 2024; Mahapatra, 2024). Additionally, ChatGPT supports brainstorming and topic generation that facilitate students to develop structured arguments and refine their essays (Hwang & Chang, 2023; Jowarder, 2023). With its capabilities, ChatGPT can facilitate users in writing essays (Hayawi et al., 2024; Ranade et al., 2024). Therefore, ChatGPT can support the technical aspects of writing and

contribute to the development of higher-order thinking skills such as analysis, synthesis, and argumentation.

However, excessive reliance on ChatGPT raises concerns about its potential impact on human creativity in writing (Aljuaid, 2024; Rahayu et al., 2024; Soelistiyowati et al., 2024) and its threat to academic integrity(Husnaini & Madhani, 2024). This condition becomes a core issue in educational institutions. Educators face the potential misuse of the ChatGPT tool because students might use it to complete assignments or generate academic essays without genuine understanding or effort (Hayawi et al., 2024). Hence, the ChatGPT-generated text has become a debated issue because it is believed to be an unoriginal text produced by humans.

Even though several studies stated that ChatGPT can mimic human-written text, other previous studies revealed that ChatGPT-generated and human-written texts are different regarding their linguistic pattern (Markowitz et al., 2023; Mizumoto et al., 2024; Sattra et al., 2024). Markey et al. (2024) analized the lexico-grammatical patterns in the proposal introduction of students at a private midwestern university, it was found that generated text tends to be denser in information compared to those written by humans. Regarding text organization, Sattra et al. (2024) compared the rhetorical moves of human and ChatGPT-written text, showing that both utilized different rhetorical moves. Furthermore, the study of Harunasari et al. (2023) found that the written text produced by AI usually lacks vibrancy and appears rigid. It is also aligned with the study of Ranade et al. (2024), revealing that AI-generated texts often suffer from limited specificity and factual inaccuracies. Even through a critical analysis, AI technologies typically have unique writing styles.

Due to this case, the present study intends to extend the investigation of other linguistic features (e.g., rhetorical structure, stylistic, discourse feature, and syntactic pattern) on ChatGPT-generated essays. It concerns lexical bundles. Lexical bundle, widely known as phraseology, refers to word sequences that are essential in building academic discourse (Jalilifar & Niamadpour, 2024). Khamkhien (2021) stated that understanding lexical bundles of natural language may facilitate people in understanding information more efficiently. Given its importance, numerous studies have explored the English lexical bundles in academic works, including reputable articles (Acorinti & Bocorny, 2024; Hong & Hua, 2018; Wijaya Kusuma et al., 2022), thesis, and essay (Ulfa & Muthalib, 2020; Yang & Fang, 2021) to find the appropriate structure and function of natural language. Unfortunately, the study of lexical bundles in ChatGPT-generated essays remains underexplored.

In this study, lexical bundle analysis focuses on four-word bundles or four-word sequences in the essay produced by ChatGPT. four-word bundles are more prevalent than three-word bundles and five-word bundles. Four-word bundles are typically more semantically complete and easier to classify

by discourse function (Biber, 2004; Hong & Hua, 2018). Additionally, the function of four-word bundles is higher than other-word bundles (Malik, 2023; Wing Yee Siu et al., 2024). Moreover, five-word bundles are much less common and often serve as extensions of core four-word bundles (Yang & Fang, 2021). Based on these precedents, this study selects four-word bundles as the core unit of analysis to explore how ChatGPT constructs academic discourse in terms of essays.

The present study systematically analyzes lexical bundles of argumentative essays produced by ChatGPT 4.0 that advanced Indonesian learners have prompted. These essays reflect typical academic writing tasks that require structured reasoning, evidence-based claims, and persuasive discourse. Using Biber et al. (1999) to identify the structural patterns of lexical bundles and Hyland's (2008b) framework to classify their discourse functions, this research explores how ChatGPT utilizes phraseological features to construct argumentative discourse.

The findings offer valuable insights into the phraseological patterns of AI-generated argumentative texts. It contributes to broader discussions on AI-assisted academic writing and discourse analysis. Specifically, the study uncovers how ChatGPT organizes arguments, expresses a stance, and structures information through recurrent four-word lexical bundles. These results are particularly relevant for educators, linguists, and researchers in language education and technology, providing a critical lens on the textual characteristics of AI-generated writing. By highlighting ChatGPT's linguistic tendencies, this study also informs pedagogical strategies for evaluating AI-generated texts and guiding students in their responsible and effective use.

METHOD

Research Design

This study employed qualitative methodology in terms of text analysis to explore lexical bundles of text produced by ChatGPT. This design was used because this study intended to describe the specific pattern and emergent features of lexical bundles on the studied object. The qualitative method facilitated the researcher in discovering the recurring sequences of words and phrases to present valuable knowledge of the structure and style of language production. This study concentrated exclusively on 4-word bundles by utilizing Hyland (2008b) framework. The choice of 4-word bundles is grounded in previous research, which has demonstrated that they are more prevalent in academic writing and offer valuable insights into discourse structure (Long et al., 2024). Additionally, 4-word bundles are widely used in linguistic analysis because they balance frequency and functional

significance. Hence, they are more suitable for systematic investigation than shorter or longer sequences.

The data and data sources used in this study were argumentative essays created by ChatGPT, based on instructions provided by advanced learners, specifically graduate students from the Applied Linguistics Department in Yogyakarta. This study employed purposive sampling to select participants with sufficient academic writing experience and familiarity with constructing argumentative texts. Graduate students in this department were chosen because they are regularly engaged in academic writing tasks and are trained in linguistic analysis, which ensured the quality and relevance of the input prompts used to generate the essays. While the scope is limited to one institution, this decision allowed for controlled conditions and consistent participant profiles. Future studies are encouraged to expand the sample to include learners from other universities and academic disciplines to enhance generalizability.

Instruments and Procedures

This study employed a corpus tool, AntConc, to generate the data. AntConc version 4.3.1 was used because of its ease of use, which can facilitate the researcher in obtaining data from an argumentative essay. Moreover, AntConc is a tool designed to conduct research in corpus linguistics.

In data collection, this study first generated argumentative essays from Graduate Students of the Applied Linguistics Department concentrating on Foreign Language Education at Yogyakarta State University. The sample consisted of 10 selected participants to ensure a representative dataset. These essays were produced in response to the following prompt: "Write an argumentative essay on the topic The Advantages and Disadvantages of ChatGPT by using English and no more than 500 words".

Secondly, this study utilized the N-gram feature on AntConc with specifications on four-word bundles by commanding AntConc to process the N-gram size to 4. This study concentrated on 4-word lexical bundles because the lower frequency of 4-word combinations generally hinders meaningful analysis (Yoo & Shin, 2022). This rationale also corroborated the findings of the Hyland (2008a) study indicated that four-word bundles were markedly more numerous and demonstrated a greater diversity of structures and functions than three-word bundles.

Thirdly, because of the limited corpus size in this work, the analysis concentrated on lexical bundles with a frequency above two. This methodology corresponds with Biber & Barbieri (2007), as referenced by Ulfa

& Muthalib (2020), who observed that studies comprising fewer than 40,000 words characterize lexical bundle as word combinations appearing at least three times across a minimum of three distinct texts. This study focused on four-word lexical bundles that appeared three or more times in data obtained from argumentative writings.

Data Analysis Procedures

After collecting data, this study conducted data analysis through several stages. First, this study analyzed the structure of lexical bundles and classified their patterns based on Biber et al. (1999). Then, to discover the lexical bundle functions, this study utilized Hyland's (2008b) theory that characterized the lexical bundle functions into three classifications, namely research-oriented function, text-oriented function, and participant-oriented function. Lastly, the authors displayed lexical bundle structure and function frequency found from the argumentative essay generated by graduate students of applied linguistics.

FINDINGS

To address the study's objective of identifying and analysing lexical bundle structure and function in ChatGPT-generated argumentative essays, this section presents an overview of the lexical bundle occurrences derived from the corpus.

The total tokens obtained in this study from ten argumentative essays by advanced learners were 6.182. This indicates that each essay averaged around 500–600 words in length. Below are the emergences of four words lexical bundles found in the essay, following the rule of Biber & Barbieri (2007) for a smaller corpus-based study.

Table 1. The emergence of four-word lexical bundles for three and more than three frequencies

No	Type	Freq	Range
1	is its ability to	8	8
2	and problem solving skills	7	7
3	of ChatGPT is its	7	6
4	one of the most	7	6
5	thinking and problem solving	7	7
25	the development of critical	4	4
26	the integration of ChatGPT	4	3
27	use of ChatGPT in	4	4
28	artificial intelligence ai technologies	3	3
29	can serve as a	3	2

6	critical thinking and problem	6	6
7	its ability to provide	6	5
8	the use of ChatGPT	6	5
9	ChatGPT can serve as	5	5
10	ChatGPT is its ability	5	5
11	has the potential to	5	5
12	is the potential for	5	5
13	most significant advantages of	5	5
14	of ChatGPT in education	5	3
15	of ChatGPT in Indonesia	5	2
16	of the most significant	5	5
17	the most significant advantages	5	5
18	advantages of ChatGPT is	4	4
19	ai tools like ChatGPT	4	3
20	development of critical thinking	4	4
21	hinder the development of	4	4
22	of artificial intelligence ai	4	4
23	of using ChatGPT in	4	2
24	significant advantages of ChatGPT	4	4
30	can serve as an	3	3
31	ChatGPT can be a	3	3
32	ChatGPT has the potential	3	3
33	drawbacks of ChatGPT in	3	2
34	especially in remote areas	3	3
35	in conclusion ChatGPT has	3	3
36	in rural and remote	3	3
37	indonesia can harness the	3	3
38	lead to the dissemination	3	3
39	of ChatGPT for education	3	2
40	of critical thinking and	3	3
41	one of the main	3	2
42	one of the primary	3	3
43	over reliance on ai	3	3
44	rural and remote areas	3	3
45	such as over reliance	3	3
46	to the dissemination of	3	3
47	using ChatGPT in education	3	2
48	using ChatGPT in indonesia	3	1
49	whenever they need it	3	3
50	with its ability to	3	2
51	with limited access to	3	3

As illustrated in Table 1 above, the total four-word lexical bundles appeared three times and more than three times in the small corpus data of the argumentative essay, with 51 occurrences. However, not all of the lexical bundles above can be included in the lexical bundle structure based on Biber's theory. Hence, this study utilized the exclusion criteria formulated by Salazar (2011) to classify the lexical structure. After sorting the lexical bundles for exclusion, below are the lexical bundle structures analyzed in this study.

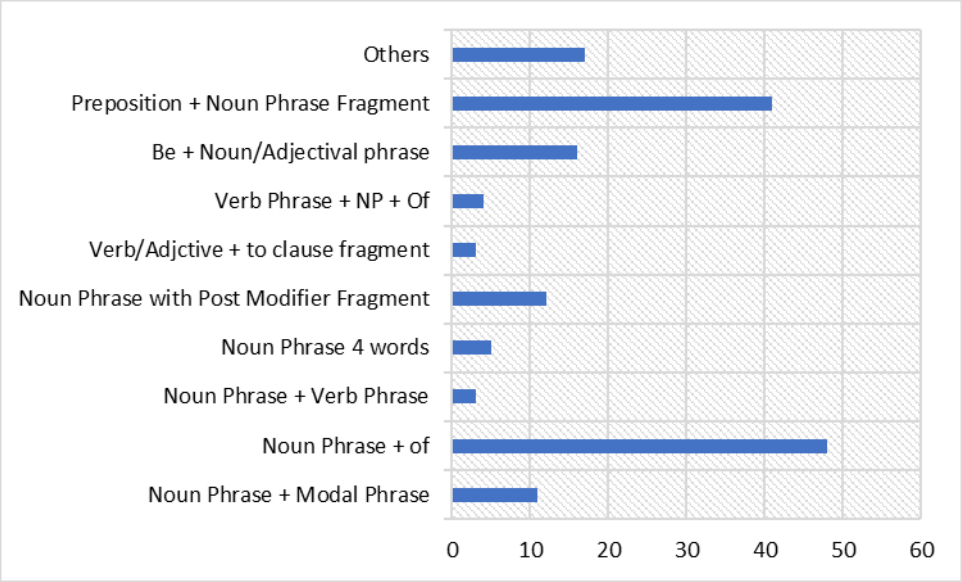

Figure 1. The Finding of Lexical Bundles Structure

After sorting out the lexical bundles according to the criteria of Salazar (2011), forty-one clusters could be studied structurally. Figure 1 above demonstrates that the 4-word lexical bundles can be classified into several patterns. The structure of noun phrases varies by incorporating other structures such as post-modifier fragments, other post-modifier fragments, prepositions *of* verb phrases, and modal phrases. The extended structure of Noun phrases that had not been mentioned by the framework of Bieber and Hyland was found in this study, such as *NP + Modal Verb* (e.g., *chatgpt can be a, ChatGPT can serve as, Indonesia can harness the*) and *NP 4-words* referred to the multiple words included in Noun Phrase (*the most significant advantages* {occurred 5 times}). Additionally, the table above also depicts that the structure of NP + of was the most frequent lexical bundle, and the other structure, NP + VP, was the least frequent lexical bundle found in 10 argumentative essays generated by ChatGPT.

Lexical Bundle Function

Hyland (2008b) classifies the lexical bundles of academic writing into three classifications: research-oriented, text-oriented, and participant-

oriented. Hyland (2008b) examined research-oriented functions by focusing on lexical bundle structures that denote time and place, text procedures, quantification, and description, enabling authors to depict their actions and experiences within real-world situations. For text-oriented functions, Hyland identified lexical bundles that aid writers in structuring their texts, including the establishment of additive or contrastive relationships, the indication of referential or clausal connections, the provision of structural signals, and the framing of arguments through the specification of limiting conditions. Conversely, participant-oriented bundles address on how articulate the writer's opinions and evaluations (stance expression) and immediately interact with readers (engagement features).

Table 2. Occurrences of Lexical Bundles in Ten Argumentative Essays

Function	**Research-Oriented Bundle**	**19**	**Text-Oriented Bundle**	**7**	**Participant-Oriented Bundle**	**15**
Sub-category	Description	5	Framing Signal	3	Engagement Features	2
	Location	1	Resultative Signal	1	Stance Features	13
	Procedure	2	Structuring Signal	1		
	Quantification	4	Transition Signal	2		
	Topic	7				

Table 2 illustrates the distribution and frequency of lexical bundle functions in argumentative essays generated by ChatGPT. The data highlighted that research-oriented bundles constituted the most dominant category, with 19 occurrences. Within this category, topic-related bundles appeared most frequently (7 instances), underscoring their crucial role in introducing or elaborating on specific research areas. Then, Participant-oriented bundles followed, totaling 15 instances, with Stance Features emerging as the most significant sub-function (13 instances). This dominance of stance features indicated a strong tendency for writers to express evaluative attitudes and perspectives within their arguments, reflecting the subjective and judgmental tone commonly associated with argumentative discourse. Engagement Features, while less frequent (2 instances), serve to connect with readers, signaling a deliberate effort to engage the audience. Lastly, text-oriented bundles, though the least frequent with only 7 occurrences, demonstrated their importance in structuring and organizing arguments. Among these, Framing Signals (3 instances) were the most common, providing a framework for the discussion, followed by Transition Signals (2 instances), Resultative Signals (1 instance), and Structuring Signals (1 instance), which help to guide readers through the argumentative flow.

Table 3. The functional and structural 4-word lexical bundles in the Argumentative Essay generated by ChatGPT

Research-Oriented Bundles Sub-Category	**Structure**	**Example**
Description	Noun Phrase+of	the development of critical
	preposition + noun phrase fragment	of artificial intelligence ai
Location	preposition + noun phrase fragment	in rural and remote
Procedure	preposition + noun phrase fragment	to the dissemination of
	Verb/Adjctive + to clause fragment	lead to the dissemination
Quantification	Noun Phrase+of	one of the most
	Noun Phrase 4 words	the most significant advantages
Topic	Noun Phrase+of	the integration of ChatGPT
	Noun Phrase with Post Modifier Fragment	using ChatGPT in education
	preposition + noun phrase fragment	of ChatGPT for education
Text-Oriented Bundles Sub-Category		
Framing Signal	preposition + noun phrase fragment	with limited access to
	Others	especially in remote areas
Resultative Signal	Verb Phrase + NP + Of	hinder the development of
Structuring Signal	preposition + noun phrase fragment	in conclusion ChatGPT has
	Noun Phrase + of	one of the main
	Others	such as over reliance
Participant-Oriented Bundles Sub-Category		
Engagement Features	Noun Phrase + Modal Phrase	indonesia can harness the
	Others	whenever they need it
Stance Features	Be + noun/adjective phrase	is the potential for
	Noun Phrase + Modal Phrase	ChatGPT can serve as
	Noun Phrase + of	advantages of ChatGPT is
	Noun Phrase + Verb Phrase	ChatGPT has the potential
	Noun Phrase with Post Modifier Fragment	its ability to provide
	Others	over reliance on ai
	preposition + noun phrase fragment	with its ability to

As illustrated in Table 3, each category encompassed various structures. It demonstrated the flexibility and multifunctionality of lexical

bundle patterns in academic writing. It means that a single structural pattern can serve multiple purposes depending on its contextual use. For instance, in the Research-Oriented Bundles category, *NP + of* is utilized across multiple subcategories such as Description, Quantification, and Topic. This versatility underscored how lexical bundles were not confined to a singular role but rather adapted to different functions based on the surrounding discourse. Similarly, patterns like *Preposition + NP Fragment* appeared in both the Location subcategory under Research-Oriented Bundles and the Framing Signal subcategory under Text-Oriented Bundles to showcase their adaptability in fulfilling diverse textual needs. This diverse application of structural patterns across functions illustrates the dynamic and context-dependent nature of lexical bundles in academic texts that enable writers to achieve clarity, coherence, and rhetorical effectiveness in their arguments.

DISCUSSION

ChatGPT has developed and reached its advancement to facilitate the educators, including students, to finish their assignments. ChatGPT 4.0 which becomes the latest version of ChatGPT, showed its superior performance, particularly in academic and professional contexts (Watari et al., 2023). This version incorporates sophisticated language processing abilities that allow users to produce well-organized, coherent, and contextually appropriate responses. Its adeptness in handling intricate questions and delivering comprehensive explanations has established it as a vital tool for individuals needing support in research, academic writing, and troubleshooting. Consequently, ChatGPT users can easily leverage its capabilities to simplify and enhance their academic and professional workflows.

Understanding the natural language of ChatGPT is certainly essential to view and differentiate the language structures and features predominantly used by ChatGPT as an advanced technology today. Educators can use this knowledge to verify students' assignments and to build educators' critical thinking in correcting the assignments. Consequently, lexical bundles found in this study reflected the linguistic features of text generated by ChatGPT to demonstrate how ChatGPT used language and constructed sentences as well as connected one sentence to others.

Linguistic features, such as lexical bundles, may not constitute a complete structural unit within a text, but they are crucial components of textual structure (Biber et al., 1999). This study analyzed 4-word lexical bundles, which provide a balance between specificity and frequency while delivering meaningful contextual insights. These bundles encompass various structural and functional categories, allowing for a thorough analysis of how ChatGPT utilizes language to generate cohesive and contextually relevant responses (Long et al., 2024).

Based on the data result above, the language structure of four-words lexical bundles consisted of ten patterns with the extended patterns like *Noun Phrase (NP) + Modal Phrase*, *NP 4 words,* and *Verb Phrase (VP) + NP + Of* that were found 11 times, 3 times, dan 4 times respectively. This study still included this pattern into various structural lexical bundles because its occurrence surpassed the rule of Salazar (2011) who also investigated lexical bundle features based on Bieber's framework. Moreover, Hong & Hua (2018) investigated lexical bundles by using a corpus tool in the field of International Business Management presented several patterns that are similar to the present study, including *NP + Modal Phrase*, *NP 4 words,* and *VP + NP + Of* in their study.

The most prevalent lexical bundle structures found in this study were NP + of with 48 occurrences, which was followed by *Preposition + NP Fragment* with 41 occurrences. The high result of *NP + of* the pattern was aligned with the study of Ulfa & Muthalib (2020) that investigated the argumentative essays produced by the undergraduate students in Malaysia. Additionally, this study result was different from other studies analyzing argumentative essays produced by the participants from China (Yang & Fang, 2021) who used *anticipatory it* patterns and Korea (Kim & Kessler, 2022; Yoo & Shin, 2022) who mostly used clausal bundles containing Verb Phrase (VP). However, this study could not assert that the reason for these differences was influenced by language learners' nationality because the study participants used ChatGPT to create the argumentative essay.

Regarding lexical bundle function, the highest rank of lexical bundle function was stance features included in participant-oriented function, primarily used to demonstrate the writers' attitudes and evaluations toward the discussed topic. Meanwhile, the text-oriented functions used to support structuring and linking the discourse were the least frequently found in this study. The lexical bundle functioning as stance features in the data results above showed how ChatGT-generated essays were rich in delivering the writers' stance or position toward the topic. It was also seen in various lexical bundle pattern types, such as *Be + Noun/Adjective phrase, NP + Modal Phrase, NP + of, NP + VP, NP with Post Modifier Fragment, Preposition + NP fragment, and Other Expressions*. More than half of the lexical bundle structure types occurred in the stance feature, with *Be + Noun/Adjective phrase* (16 occurrences) as the highest stance feature.

The ChatGPT-generated essays reflected a strong emphasis on presenting and elaborating on research content, with less attention to discourse structuring. This finding corresponded with the research of Markey et al. (2024) investigated the ChatGPT-Generated Text. They explained that texts constructed by ChatGPT are denser than texts authored by humans but can come across as dialogically closed, lacking depth, and relatively

insubstantial. It means that without adequate discourse structuring, such as transitions, framing signals, or conclusions, the essay might fail to effectively guide the reader through the argument, diminishing its overall impact and clarity.

The lack of effective discourse organization diminished the essay's persuasive efficacy, as a coherent argument is essential for establishing credibility and captivating the audience. Kim & Kessler (2022) analyzed the comparison of lexical bundle functions between low and high-proficiency students, discovering that lower-proficiency groups were more inclined to utilize attitude expressions to demonstrate their stances rather than discourse organizers for managing one argument to others. Without these structures, the audience may struggle to evaluate the assertions' truth and relevance, reducing their overall efficacy (Liu et al., 2021). Furthermore, Mizumoto et al. (2024) revealed that essays generated by ChatGPT were predominantly factual and less personal, as they relied on information from websites that frequently lack the variety of registers commonly present in student writings.

The connections between arguments are essential for grasping the logical progression and contextual basis of claims. While detailed research content is essential, its effectiveness relies heavily on how it is organized and presented. Overlooking the role of structuring elements may result in a text that, although informative, lacks the cohesion needed to deliver a compelling and comprehensive argument. Therefore, this study advocated several previous findings suggesting that the text provided by ChatGPT may lack depth and precision necessary for more intricate aspects of language learning.

CONCLUSION

Overall, this study found that the most frequent structural pattern in ChatGPT-generated texts was *the Noun Phrase + of*, with the most common lexical bundle function being the Stance Features subcategory related to participant-oriented functions. However, ChatGPT-generated essays have sentences that are low in clarity because they often lack the cohesive elements necessary to guide readers effectively through arguments. This suggests that although ChatGPT improves writing efficiency, its reliance on specific lexical structures may restrict the complexity and nuance needed for advanced language acquisition and compelling academic writing.

One notable limitation of this study is its focus on a specific academic genre and linguistic context, which may not fully represent the broader applications of AI-assisted writing. Since the study primarily examines ChatGPT-generated texts within a controlled academic framework, its findings may not generalize to other forms of writing, such as creative, professional, or technical compositions. Moreover, the linguistic characteristics of AI-generated essays can vary based on several factors,

including prompt specificity, user expertise, and iterative refinements, which were not extensively explored in this study. As a result, while the findings highlight recurring patterns in ChatGPT-generated academic writing, they do not necessarily capture the full range of stylistic and structural variations that may emerge in different writing scenarios.

Future studies interested in AI writing tools, such as ChatGPT, may investigate other text genres to analyze the evolution of ChatGPT-generated texts in response to user interventions and develop methodologies to augment its rhetorical efficacy through human-AI collaboration. Investigating the role of human agency in shaping AI-assisted writing outcomes may inform the development of pedagogical strategies that integrate AI tools into pedagogical activities. Eventually, such research could contribute to creating more sophisticated AI-assisted writing models that balance automation with human creativity and critical thinking

REFERENCES

Acorinti, W. H., & Bocorny, A. E. P. (2024). Lexical Bundles in Computer Science Academic Texts Produced by Novice and Expert Writers. *Journal of the Undergraduate Linguistics Association of Britain*, *2*(2), 189–210. https://doi.org/https://doi.org/10.5281/zenodo.12738429

Aldabbus, S., & Almansouri, E. (2022). Academic Writing Difficulties Encountered by University EFL Learners. *British Journal of English Linguistics*, *10*, 1–11. https://www.eajournals.org/

Aljuaid, H. (2024). The Impact of Artificial Intelligence Tools on Academic Writing Instruction in Higher Education: A Systematic Review. *Arab World English Journal*, *1*(1), 26–55. https://doi.org/10.24093/awej/chatgpt.2

Alqasham, F. H., Al-Ahdal, A. A. M. H., & Babekir, A. H. S. (2021). Coherence and Cohesion in Saudi EFL Learners' Essay Writing: A Study at a Tertiary-level Institution. *Asian EFL Journal Research Articles*, *28*(1), 8–25. https://www.researchgate.net/profile/Arif-Al-Ahdal/publication/349256060_Coherence_and_Cohesion_in_Saudi_EFL_Learners'_Essay_Writing_A_Study_at_a_Tertiary-level_Institution/links/60270fc545851589399ee195/Coherence-and-Cohesion-in-Saudi-EFL-Learners-Essay-Writing-A-Study-at-a-Tertiary-level-Institution.pdf

Alsariera, A. H., & Yunus, K. (2023). Issues and Challenges in the Use of Coherence Among Jordanian EFL Students in Writing Academic Essays. *Journal of Nusantara Studies (JONUS)*, *8*(2), 141–156. https://doi.org/10.24200/jonus.vol8iss2pp141-156

Anindita, W. K. (2024). Cohesion and Coherence Problems Among Non-Native English Students' Writing Essays. *Lensa: Kajian Kebahasaan,*

Kesusastraan, Dan Budaya, *14*(1), 58. https://doi.org/10.26714/lensa.14.1.2024.58-79

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

Biber, D. (2004). If you look at ...: Lexical Bundles in University Teaching and Textbooks. *Applied Linguistics*, *25*(3), 371–405. https://doi.org/10.1093/applin/25.3.371

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, *26*(3), 263–286. https://doi.org/10.1016/j.esp.2006.08.003

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written english*. Longman.

Bisriyah, M. (2022). EFL University Students' Difficulties in the Essay Writing Process. *Scope: Journal of English Language Teaching*, *7*(1), 66. https://doi.org/10.30998/scope.v7i1.13793

Bom, H. S. H. (2023). Exploring the Opportunities and Challenges of ChatGPT in Academic Writing: a Roundtable Discussion. In *Nuclear Medicine and Molecular Imaging* (Vol. 57, Issue 4, pp. 165–167). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s13139-023-00809-2

Chia, C. S. C., Jonathan, P., Owen, H., & Kit-Mun, L. (2024). Graduate Students' Perception and Use of ChatGPT as a Learning Tool to Develop Writing Skills. *International Journal of TESOL Studies*. https://doi.org/10.58304/ijts.20240308

Fauzan, U., Hasanah, N., & Hadijah, S. (2022). The Undergraduate Students' Difficulties in Writing Thesis Proposal. *Indonesian Journal of EFL and Linguistics*, *7*(1), 175. https://doi.org/10.21462/ijefl.v7i1.515

Fitriyah, I., Ningrum, A. S. B., & Gozali, I. (2023). An Investigation of Written Corrective Feedback in EFL Writing Assessment: How Teachers' Feedback Practices Meet Students' Expectations. *International Journal of Language Testing*, *14*(1), 166–184. https://doi.org/10.22034/IJLT.2023.411616.1275

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing Assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, *3*. https://doi.org/10.1016/j.caeai.2022.100055

Harunasari, S. Y., Tinggi, S., & Asing, B. (2023). Examining the Effectiveness of AI-integrated Approach in EFL Writing: A Case of ChatGPT. *International Journal of Progressive Sciences and Technologies (IJPSAT*, *39*(2), 357–368.

Hayawi, K., Shahriar, S., & Mathew, S. S. (2024). The imitation game: Detecting human and AI-generated texts in the era of ChatGPT and BARD. *Journal of Information Science*. https://doi.org/10.1177/01655515241227531

Hong, A. L., & Hua, T. K. (2018). Specificity in English for Academic Purposes (EAP): A corpus analysis of lexical bundles in academic writing. In *3L: Language, Linguistics, Literature* (Vol. 24, Issue 2, pp. 82–94). Penerbit Universiti Kebangsaan Malaysia. https://doi.org/10.17576/3L-2018-2402-07

Husnaini, M., & Madhani, L. M. (2024). Perspektif Mahasiswa terhadap ChatGPT dalam Menyelesaikan Tugas Kuliah. *Journal of Education Research*, *5*(3), 2655–2664. https://doi.org/10.37985/jer.v5i3.1047

Hwang, G. J., & Chang, C. Y. (2023). A review of opportunities and challenges of chatbots in education. In *Interactive Learning Environments* (Vol. 31, Issue 7, pp. 4099–4112). Routledge. https://doi.org/10.1080/10494820.2021.1952615

Hyland, K. (2008a). Academic clusters: text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, *18*(1), 41–62. https://doi.org/10.1111/j.1473-4192.2008.00178.x

Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Irfan, I., Sofendi, S., & Vianty, M. (2020). Technological knowledge application on academic writing English education study program students. *English Review: Journal of English Education*, *9*(1), 157–166. https://doi.org/10.25134/erjee.v9i1.3788

Jalilifar, A., & Niamadpour, V. (2024). Teaching English as a Second Language Quarterly (TESLQ) (Formerly Journal of Teaching Language Skills) Online Book Reviews in Sciences: Structural and Functional Differences in Lexical Bundles. *Formerly Journal of Teaching Language Skills)*, *43*(4), 87–116. https://doi.org/10.22099/tesl.2024.49990.3280

Jowarder, M. I. (2023). The Influence of ChatGPT on Social Science Students: Insights Drawn from Undergraduate Students in the United States. *Indonesian Journal of Innovation and Applied Sciences (IJIAS)*, *3*(2), 194–200. https://doi.org/10.47540/ijias.v3i2.878

Khamkhien, A. (2021). Functional Patterns of Lexical Bundles and Limitations in Academic Writing by Thai L2 English Learners. *Journal of English Language Teaching and Linguistics*, *6*(3), 607. https://doi.org/10.21462/jeltl.v6i3.629

Kim, S., & Kessler, M. (2022). Examining L2 English University students' uses of lexical bundles and their relationship to writing quality. *Assessing Writing*, *51*. https://doi.org/10.1016/j.asw.2021.100589

Lee, I. (2021). The Development of Feedback Literacy for Writing Teachers. *TESOL Quarterly*, *55*(3), 1048–1059. https://doi.org/10.1002/tesq.3012

Limbong, S., Alberto, & Frans Tandibura. (2023). Lecturer's Personal and Social Competence Toward Students' Writing Proficiency in EFL Classroom. *Journal of English Education and Teaching*, *7*(4), 914–927. https://doi.org/10.33369/jeet.7.4.914-927

Liu, X., Ou, J., Song, Y., & Jiang, X. (2021). Exploring Discourse Structures for Argument Impact Classification. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, 3958–3969. https://doi.org/10.18653/V1/2021.ACL-LONG.306

Long, B., Shao, S., Yao, J., & Zhang, J. (2024). Structural and Functional Usages of Lexical Bundle in Native and Non-Native English Sociology Academic Writings: A Comparative Study Based on MICUSP. *English Language Teaching and Linguistics Studies*, *6*(4), p11. https://doi.org/10.22158/eltls.v6n4p11

Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: a mixed methods intervention study. *Smart Learning Environments*, *11*(1). https://doi.org/10.1186/s40561-024-00295-9

Malik, Z. A. (2023). A Corpus-based Study of the Lexical Bundles in the Ph. D Linguistics Dissertations' Abstracts in Pakistani Universities. *Pakistan Journal of Humanities and Social Sciences*, *11*(2), 2317–2325. https://doi.org/10.52131/pjhss.2023.1102.0524

Markey, B., Brown, D. W., Laudenbach, M., & Kohler, A. (2024). Dense and Disconnected: Analyzing the Sedimented Style of ChatGPT-Generated Text at Scale. *Written Communication*. https://doi.org/10.1177/07410883241263528

Markowitz, D. M., Hancock, J. T., & Bailenson, J. N. (2023). Linguistic Markers of Inherently False AI Communication and Intentionally False Human Communication: Evidence from Hotel Reviews. *Journal of Language and Social Psychology*. https://doi.org/https://doi.org/10.31234/osf.io/mnyz8

Mizumoto, A., Yasuda, S., & Tamura, Y. (2024). Identifying ChatGPT-generated texts in EFL students' writing: Through comparative analysis of linguistic fingerprints. *Applied Corpus Linguistics*, *4*(3). https://doi.org/10.1016/j.acorp.2024.100106

Punar Özçelik, N., & Yangın Ekşi, G. (2024). Cultivating writing skills: the role of ChatGPT as a learning assistant—a case study. *Smart Learning Environments*, *11*(1). https://doi.org/10.1186/s40561-024-00296-8

Rababah, L. M., Rababah, M. A., & Al-Khawaldeh, N. N. (2024). Graduate Students' ChatGPT Experience and Perspectives during Thesis Writing.

International Journal of Engineering Pedagogy, *14*(3), 22–35. https://doi.org/10.3991/ijep.v14i3.48395

Rahayu, Weda, S., Muliati, & De Vega, N. (2024). Artificial Intelligence in writing instruction: A self-determination theory perspective. *XLinguae*, *17*(1), 234–244. https://doi.org/10.18355/XL.2024.17.01.16

Rahim, W. ur, Emon, M. M. H., & Haque, T. Al. (2023). Advancements in Tertiary-Level Writing: Insights From Bangladeshi Students and Their Evolving Writing Practices. *World Languages, Literature and Cultural Studies*, *2*(2), 26–32. https://doi.org/10.26480/wllcs.02.2023.26.32

Ramadhan, R. A., Abdul Rahim, Darmanto, & Paris, A. S. (2024). Enhancing Academic Writing through Digital Tools: A Systematic Review. *Journal of Language and Literature Studies*, *4*(4), 707–770. https://doi.org/10.36312/jolls.v4i4.2342

Ranade, N., Saravia, M., & Johri, A. (2024). Using rhetorical strategies to design prompts: a human-in-the-loop approach to make AI useful. *AI and Society*. https://doi.org/10.1007/s00146-024-01905-3

Salazar, D. J. L. (2011). *Lexical bundles in scientific English: A corpus-based study of native and non-native writing*. www.tdx.cat

Sattra, M., Banjong, B., Intisarn, C., & Anyarat, N. (2024). Rhetorical Moves of AI-Generated Research Article Abstracts and Published Research Article Abstracts in Applied Linguistics. *Journal of Humanities and Social Sciences*, *8*(1). https://so06.tci-thaijo.org/index.php/husojournalpnru/article/view/265142

Soelistiyowati, E., Nugky Permatasari, A., Nugroho, V. A., Shalsabillah, A., & Kusmulyadi, F. (2024). *Enhancing College Learners' Writing Skills through The Integration of ChatGPT: Strategies, Benefits, And Concerns*. *8*(1). https://ojs.unm.ac.id/eralingua

Ulfa, N., & Muthalib, K. A. (2020). Lexical Bundles In Students' Essay Writing. *English Education Journal (EEJ)*, *11*(3), 367–379.

Wahab, A., Pd, M., Sidik, A. S., Ningsih, F., & Rahim, R. (2024). *Exploring The Use Of Digital Writing Tools To Enhance Creative Writing In The Second Semester Of The English Department, Qamarul Huda Badaruddin University*. *4*(01). https://ejournal.hamjahdiha.org/index.php/pena/article/view/77/77

Watari, T., Takagi, S., Sakaguchi, K., Nishizaki, Y., Shimizu, T., Yamamoto, Y., & Tokuda, Y. (2023). Performance Comparison of ChatGPT-4 and Japanese Medical Residents in the General Medicine In-Training Examination: Comparison Study. *JMIR Medical Education*, *9*(1). https://doi.org/10.2196/52202

Wijaya Kusuma, F., Rukmini, D., & Faridi, A. (2022). The Realization of Lexical Bundle Structures and Functions in English Education Journal. /

English Education Journal, *12*(3), 437–446. http://journal.unnes.ac.id/sju/index.php/eej

Wing Yee Siu, B., Afzaal, M., Saleh Aldayel, H., & Curle, S. (2024). Unlocking the Mysteries of Academic Writing: A Corpus-based Analysis of Lexical Bundles in L2 English for Engineering Students. *SAGE Open*, *14*(4). https://doi.org/10.1177/21582440241299997

Wu, Y. (2024). Study on the Impact of Utilizing ChatGPT and Other AI Tools for Feedback in EAP Writing Classrooms on the Discursive Writing Performance of English Major Students. *Transactions on Social Science, Education and Humanities Research*, *4*, 143–150. https://doi.org/10.62051/4se95x52

Yang, W., & Fang, Q. (2021). The Structures and Functions of Lexical Bundles in Argumentative Essays by Chinese EFL Students at the Tertiary Level. *International Journal of TESOL Studies*. https://doi.org/10.46451/ijts.2021.09.03

Yoo, I. W., & Shin, Y. K. (2022). English lexical bundles in a learner corpus of argumentative essays written by Korean university students. *Corpora*, *17*(Supplement), 23–42. https://doi.org/10.3366/cor.2022.0245

Zakaria, S. M., Yunus, M. Md., Nazri, N. M., & Shah, P. M. (2016). Students' Experience of Using Storybird in Writing ESL Narrative Text. *Creative Education*, *07*(15), 2107–2120. https://doi.org/10.4236/ce.2016.715210