การวิเคราะห์อัตถภาคของบทคัดย่อบทความวิจัยที่สร้างโดยปัญญาประดิษฐ์และบทคัดย่อที่ตีพิมพ์ในวารสารวิชาการสาขาภาษาศาสตร์ประยุกต์

RHETORICAL MOVES OF AI-GENERATED RESEARCH ARTICLE ABSTRACTS AND PUBLISHED RESEARCH ARTICLE ABSTRACTS IN APPLIED LINGUISTICS

ศาสตรา มาพร[1*] บรรจง บุรินประโคน[2] อินธิสาร ไชยสุข[3] และ อัญญารัตน์ นาถธีระพงษ์[3]

[1]กองส่งเสริมการวิจัยและบริการวิชาการ มหาวิทยาลัยมหาสารคาม

[2]สาขาวิชาภาษาไทยเพื่อการสื่อสาร คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏมหาสารคาม

[3]ภาควิชาภาษาตะวันตกและภาษาศาสตร์ คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยมหาสารคาม

Sattra Maporn[1*], Banjong Burinprakhon[2], Intisarn Chaiyasuk[3], Anyarat Nattheeraphong[3]

[1]Division of Research Facilitation and Dissemination, Mahasarakham University

[2]Department of Thai for Communication, Faculty of Humanities and Social Sciences, Rajabhat Mahasarakham University

[3]The Department of Western Languages and Linguistics, Faculty of Humanities and Social Sciences, Mahasarakham University

Email : sattra.m21@gmail.com

บทคัดย่อ

บทคัดย่อเป็นส่วนสำคัญในการตีพิมพ์บทความวิจัยในวารสารวิชาการโดยมีเนื้อหาสรุปบทความเพื่อช่วยให้ผู้อ่านประเมินความสำคัญของงานได้ โดยยุคปัจจุบันผู้เขียนอาจใช้เทคโนโลยีช่วย เช่น *ChatGPT* จากบริษัท *OpenAI* เพื่อช่วยในการสร้างส่วนบทคัดย่อของบทความวิจัย โดยงานวิจัยนี้มีวัตถุประสงค์เพื่อระบุรูปแบบอัตถภาคในบทคัดย่อของบทความวิจัยที่สร้างโดย *ChatGPT* โดยเปรียบเทียบกับผลงานที่เขียนโดยมนุษย์จากวารสารที่มีชื่อเสียงในด้านภาษาศาสตร์ประยุกต์ ข้อมูลประกอบด้วยบทคัดย่อของบทความวิจัย 60 บทคัดย่อที่มาจากสองแหล่งข้อมูล โดยใช้รูปแบบการวิเคราะห์อัตถภาคของ Hyland (2000) เพื่อวิเคราะห์ความถี่และค่าร้อยละ ผลการวิจัยพบว่า บทคัดย่อที่สร้างโดย AI มักใช้รูปแบบ P-M-Pr-C ในขณะที่บทคัดย่อที่เขียนโดยมนุษย์มักยึดถือรูปแบบ I-P-M-Pr-C ส่วนความแตกต่างในรูปแบบการเขียนนี้เสนอให้เห็นว่าอาจมีความแตกต่างในเลือกใช้รูปแบบของภาษาระหว่างข้อความที่สร้างขึ้นโดยปัญญาประดิษฐ์และผู้เขียนมนุษย์ในบริบทของการเขียนบทคัดย่อ

คำสำคัญ: อัตถภาค, *ChatGPT*, บทคัดย่อที่สร้างโดยปัญญาประดิษฐ์, บทคัดย่อบทความวิจัย

---

[1] กองส่งเสริมการวิจัยและบริการวิชาการ มหาวิทยาลัยมหาสารคาม

[2] สาขาวิชาภาษาไทยเพื่อการสื่อสาร คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏมหาสารคาม

[3] ภาควิชาภาษาตะวันตกและภาษาศาสตร์ คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยมหาสารคาม

Abstract

Abstracts play a crucial role in scientific publications, serving as summaries of research articles that aid readers in evaluating the significance of the work. Authors may utilize contemporary technologies, such as *OpenAI's ChatGPT*, to assist in generating the abstract portion of research articles. This study aims to identify the rhetorical patterns employed in research article abstracts generated by *OpenAI's ChatGPT* in comparison to those crafted by human authors from highly reputable journals in applied linguistics. The dataset comprises 60 research article abstracts sourced from two distinct corpora, employing the linguistic model proposed by Hyland (2000). Frequencies and percentages of rhetorical moves were analyzed. The findings reveal that abstracts generated by AI commonly utilize the P-M-Pr-C pattern, whereas human-authored abstracts predominantly adhere to the I-P-M-Pr-C patterns. This distinction in writing patterns suggests that there may be differences in writing style and rhetorical choices between texts produced by artificial intelligence and those authored by humans in the context of abstract composition.

Keyword: Rhetorical moves, *ChatGPT*, AI-generated research article abstracts, research article abstracts

Introduction

The AI-generated text is not a new issue in our world as Marti (2022) claims that the root product of AI-generated text appeared from 1950s and 1960s. At the present time, artificial intelligent (AI) can generate new content, which at the same time received a lot of controversies. It can answer a question, write a poem, fiction, and books. AI can also pass the medical licensing examination (PePeau-Wilson, 2023). In addition, language generated by AI has entered scientific community. The technology can surely assist researcher to compose a part of research article. Anderson et al. (2023) pointed out that user should be aware of false research methodologies and references which will be rejected immediately by publisher. In addition, language generated by AI has entered scientific community. Stokel-Walker (2023) reported that four manuscripts, which is the first draft for academic publications, have credited *chatGPT* as an author and then scientists were involved in unpleasant result.

As a part of research article, an abstract is one of the parts that technologies might assist researchers. It is a 250–300 words summary of the entire research article and is regarded as a first step in convincing an editor, reviewers, or a reader to read the article. An article might be rejected if the abstract is written in unclear patterns (Menon et al., 2020). This study tries to focus on it because it is the crucial part and it is also considered as a genre which has been studied for several decades.

Objectives

1. To identify the rhetorical move patterns of research article abstracts in applied linguistics generated by humans.

2 To identify the move frequency and patterns of AI-generated text, as well as the special characteristics of AI text.

Literature review

particularly in research article abstracts, is closely intertwined with linguistic research, as evidenced by studies conducted by Amnuai (2019), Kaya and Yağız (2020), Phonhan (2021), El-Dakhs (2018), Kanafani (2022), Kitjaroenpaiboon (2021), Maporn et al. (2023), Sukhapabsuk (2020), and Zand-Moghadam (2022). These investigations centered on research article abstracts found in both international and national databases. The results showed that the linguistics

research abstracts had common move patterns including P-M-Pr-C and I-P-M-Pr-C. Among those researches, there is a slight difference because of the difference in data used in each research as well as different research approach. However, according to current academic database, the differences in patterns between AI-generated text and published research article abstracts which possibly written by humans seems to be the gap of review literature.

In the aspect of academic article and technology assisted writing, several studies analyzed the AI-generated text (Dergaa et al., 2023). Some researches tried to prove the efficiency of AI compared with humans in case of scientific article texts (Salvagno et al., 2023; Macdonald, 2023), analysis of fabrications and plagiarism of AI-generated texts (Elali & Rachid, 2023). Elali & Rachid (2023) have shown that AI-generated text can be compared with human writing and it is not easy to detect plagiarism by normal plagiarism detector. However, the research claims that AI-technologies typically have unique writing style and verb used. Compare AI assay written to human-written essays (Herbold et al., 2023), Analysis of persuasive discourse of AI-generated text (Hinton & Wagemans, 2023), Discourse and perceptual analysis of AI-synthesized texts on coherence and cohesion (Bun, 2020; Lee et al., 2018), Comparing scientific abstracts generated by ChatGPT to real abstracts (Gao et al., 2023), anaysis of ChatGPT of potential to assist writing (Kumar, 2023), conducting academic research with the AI (Dönmez et al.,2023), linguistic ambiguity analysis in ChatGPT (Ortega-Martín, 2023).

To detect the writing style, genre analysis can be adopted. The analysis of AI-generated research article abstracts using linguistics framework may be the possible gap of the review studies as mentioned. Therefore, this research has two objectives. First, this research tried to identify the rhetorical move patterns of research article abstracts in applied linguistics generated by human. The second objective was to identify move frequency and patterns of AI generated-text as well as special characteristics of AI text.
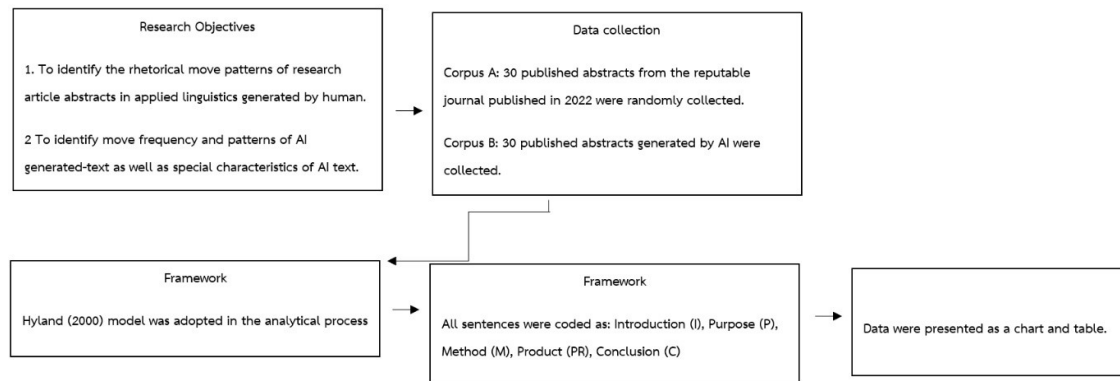
Research Objectives	Data collection
1. To identify the rhetorical move patterns of research article abstracts in applied linguistics generated by human.	Corpus A: 30 published abstracts from the reputable journal published in 2022 were randomly collected.
2 To identify move frequency and patterns of AI generated-text as well as special characteristics of AI text.	Corpus B: 30 published abstracts generated by AI were collected.

Framework	Framework	
Hyland (2000) model was adopted in the analytical process	All sentences were coded as: Introduction (I), Purpose (P), Method (M), Product (PR), Conclusion (C)	Data were presented as a chart and table.

Figure 1 Research design

This study used linguistics model of Hyland (2000). The five-move model has been widely used in several researches. It was mostly used in analysis of research article abstracts which were published in academic journals, thesis, and research grant etc. This research looks at the possibility to use this model to AI-generated text.

Data Collection

Two corpora including 60 research article abstracts in total were created. The corpus A refers to 30 abstracts from two randomly high reputation journals in applied linguistics indexed Scopus and Web of Science. Two journals were selected randomly from the category Language & Linguistics in the first quartile of *Scopus*. Data collected from corpus A were retrieved from *Modern Language Journal* during published in 2022. All data in corpus A were considered as human generated abstracts. The corpus B were 30 research article abstracts collected from *OpenAI's ChatGPT* version 3.0 on May 23rd, 2023. The researcher used command "*write research article abstract in the field of applied linguistics*". The command was sent without any restrictions. Artificial intelligent provided text freely. After researcher received each generated text, the conversation was refreshed every time to prevent the recognized conversation from AI.

According to Flowerdew (2004) and Biber (2006), the size of a corpus can be determined based on the diversity of grammatical features it encompasses. If the target features contain a frequently occurring grammatical structure, a smaller corpus size may be sufficient. Previous studies have employed the similar size of data, so a total of 60 abstracts were used in this study seems to be a reasonable number.

According to Biber et al. (2007), a top-down approach to analyze the discourse structure of texts were adopted. Firstly, each sentence was carefully examined to detect its rhetorical meaning. Secondly, at least one sentence was assigned to each move. In cases of multiple moves were encapsulated within a single sentence, the sentence was coded at multiple moves according to phase order from left to right. Thirdly, given texts in multiple sentences can convey the same communicative functions, a message may be conveyed through one sentence, several sentences, one paragraph, or multiple paragraphs. The cut-off frequencies for obligatory, conventional, and optional moves and steps are set at 90%, 60%, and below 60%, respectively.

Hyland (2000) framework

Introduction (I): Establishes context of the paper and motivates the research or discussion

Purpose (P): Indicates purpose, thesis or hypothesis, outlines the intention behind the paper

Method (M): Provides information on design, procedures, assumptions, approach, data, etc.

Product (Pr): States main findings or results, the argument, or what was accomplished.

Conclusion (C): Interprets or extends results beyond scope of paper, draws inferences, points to applications or wider implications

Inter-rater reliability

According to Holmes (1997), subjectivity among researchers is inevitable during analysis. To mitigate this issue, inter-rater reliability was employed in this study. Two coders were involved: the researcher served as the first coder, while a university lecturer was assigned as the second coder. The two coders collaboratively coded 10% of the entire data, including 3 items from corpus A and 3 items from corpus B. The coding process consisted of two sessions. In the first session, individual ratings were conducted with the guidance of a specific framework. The researcher initially described the methodology, providing detailed instructions and protocols for the coding process. The second session involved a discussion following the

individual rating phase. Ultimately, inter-rater reliability was determined by calculating the percentage of agreement. In this study, the percentage of agreement exceeded 80%. This process was approved by the university ethics committee under number 406-435/2022, expiring in December 2023.

Results

Table 1. Word length and number of sentences between human generated and AI – generated abstract

Topic	Corpus A	Corpus B
Avg number of words	187.30	170.53
Avg. number of sentences	7.13	7.57

Table 1 indicates that AI-generated abstracts tend to write shorter sentences compared to those written by humans. The AI tends to use more sentences but fewer words than humans. This suggests that humans may use more words to provide additional clarification in their abstracts.
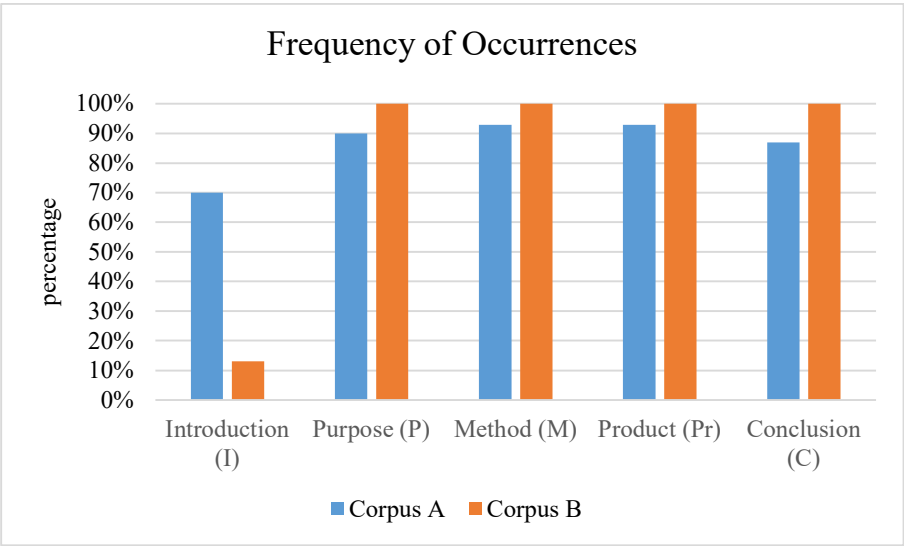

Figure 2 Frequency of occurrences

The percentage of five moves among two datasets were depicted in Figure 2. Differences in the frequency of appearance were observed between the two corpora.

Specifically, for the introduction (I), AI-generated applied linguistics abstracts accounted for 13%, whereas in applied linguistics abstracts generated by humans, it constituted 70%. The frequency of distribution was consistent across both corpora for Purpose (P), Method (M), and Product (Pr), this study indicated that AI has generated four obligatory moves similar to humans written abstract.
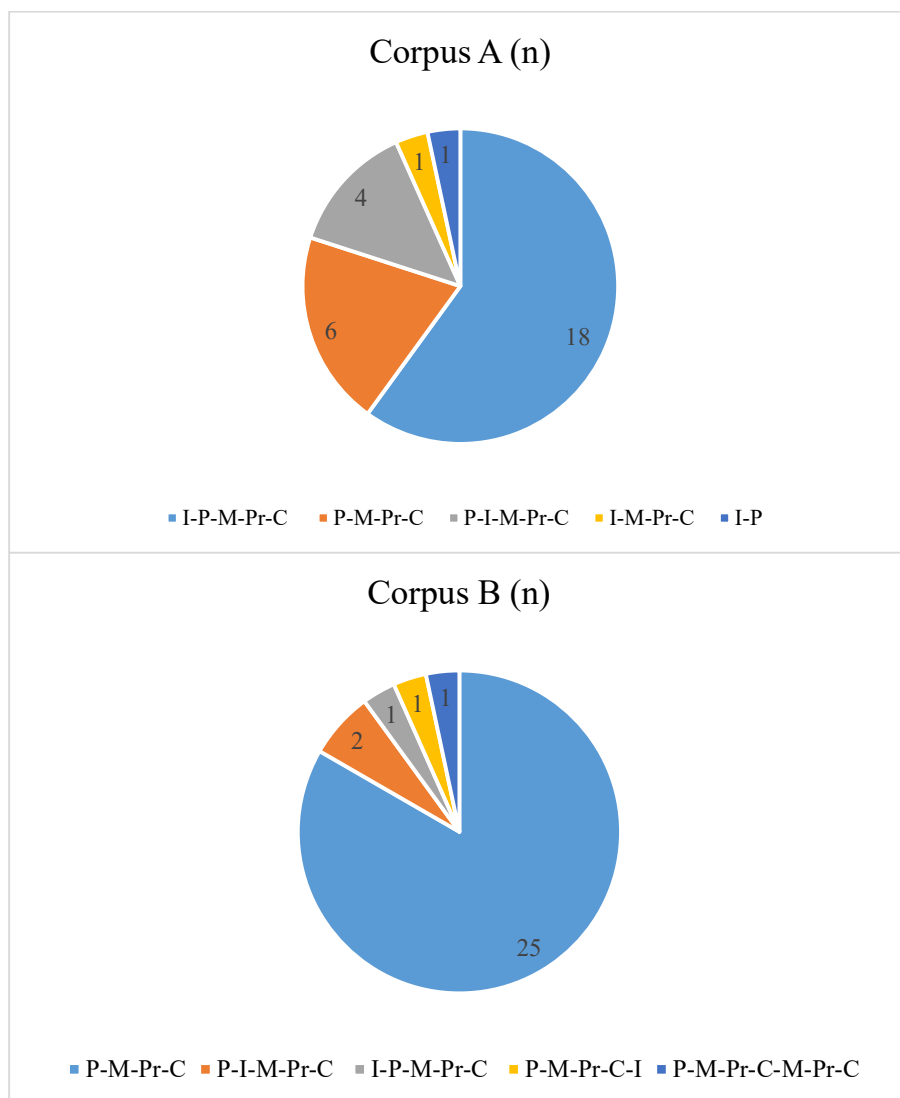

Figure 3 Rhetorical move patterns

Figure 3 presents the rhetorical moves found in research article abstracts from two corpora. The most preferred pattern in AI-generated abstracts was P-M-Pr-C (77%), which involves presenting the aim of research, followed by the methodology, the results, and the conclusion. The second most occurred pattern was P-I-M-Pr-C (17%), which starts with a statement of purpose, followed by details in the background, then the methodology, the results, and the conclusion. In contrast, most preferred abstracts generated by the human was I-P-M-Pr-C (40%), where the patterns were presented in linear sequences involving an introduction followed by the purpose, methodology, product, and conclusion. From the comparing two datasets reveals that AI-generated abstracts exhibit a more fixed pattern, while human-generated abstracts display more diversity. This data suggests that AI may offer assistance in producing abstracts with P-M-Pr-C structures, which may be beneficial in real writing scenarios. However, humans have more flexibility patterns, which may depend on the context and goals of the writer.

Table 2. Repetitive moves

Moves	Corpus A (n)	Corpus B (n)
Introduction (I)	1	-
Purpose (P)	1	1
Method (M)	2	-
Product (Pr)	-	2
Conclusion (C)	-	2

Some repetitive moves were identified in this study. The findings were presented in Table 2, which revealed that both AI-generated abstracts and human abstracts showed repetitive moves. Specifically, repetitive moves occurred five times in the AI-generated abstracts corpus and four times in the human-generated corpus. These results indicate that artificial intelligence produced repetitive moves at almost similar frequency to humans, as evidenced by the data from this study. Therefore, this seems no significantly differences.

Organization of Moves (Sequence, Linearity)

At the beginning of the research, researcher expected linear move from AI-generated text while more of non-linear for human generated text. In this section, author will discuss how well these five moves follow a linear order according to the Hyland (2000) model. Table 3 values and linearity of research article abstract of the two corpora in this study.

Table 3 Linearity of moves of this study

Linearity	Corpus A (n)	Corpus B (n)
Linear	I-P-M-Pr-C (18)	I-P-M-Pr-C (1)
Semi Linear	P-M-Pr-C (6) I-P (1) I-M-Pr-C (1)	P-M-Pr-C (25)
Non - Linear	P-I-M-Pr-C (4)	P-I-M-Pr-C (2) P-M-Pr-C-I (1) P-M-Pr-C-M-Pr-C (1)

Table 3 shows that AI-generated research article abstracts tend to use semi-linear P-M-Pr-C patterns more than 70%, while human-generated research article abstracts tend to use linear patterns I-P-M-P-C. The data indicates that human-generated research article abstracts show greater diversity in terms of patterns, whereas AI-generated abstracts are more restricted, with a single pattern comprising over half of the data. The result has rejected the null hypothesis of researcher on linearity. The result shows the number of non-linear were 4 items which were not different between two sets.

Discussion

The results of the study indicate that artificial intelligent *OPENAI's chatGPT* provide four move structure which appear more than 80%. The Introduction (I) were found very rare in AI generated abstract while human has generated this move more than 60%. Both corpora were different at frequency of move. The I-P-M-Pr-C patterns were mostly found in human generated abstracts while AI has generated P-M-Pr-C almost 80%. This study concludes that *chatGPT* may be a tool for assisting writers in applied linguistics in terms of generating the P-M-Pr-C pattern. It is beneficial to researcher who want to use P-M-Pr-C because of the pattern were common and also may get published, confirmed by previous research that utilized data from published research article abstracts in applied linguistics (Zand-Moghadam & Zhaleh, K., 2022; El-Dakhs, 2018). However, there were caution. There was a pattern which cannot be found in the human-generated corpus of research article abstract from this study. It was P-M Pr-C-M-Pr-C, which the author should be careful when using research abstract patterns suggested by AI.

For this reasons, further research need to explore the potential applications of AI in this field and to determine whether it can help writers. However, the effectiveness of AI may vary depending on the field of study.

Theoretical implications

This study implies that AI-generated texts use the same rhetorical patterns in applied linguistics research article abstracts which might be written by human (Zand-Moghadam & Zhaleh, K., 2022; El-Dakhs, 2018). Therefore, it can theorical assist human writing in the field in term of forming rhetorical patterns for researchers.

Practical implications

According to this study, AI-generated texts can help human writers for creating research article abstracts. This study indicates that AI-generated texts can structure the research article abstract without further command prompt even though this study cannot demonstrate the quality of abstracts.

Suggestion for further study

The study of lexical bundles of texts produced by AI is the one that is most recommended for future research since it expands our understanding of how AI chooses words. The academician will also know more about the dialogue regarding AI. Studying discourse analysis of AI from multiple disciplines is another recommendation. Lee et al. (2021) have evaluation quality of AI generated text. However, the research has conducted before *ChatGPT* have appeared, so it is interesting to evaluate again.

Conclusion

The present study employed discourse analysis to define the rhetorical moves in research article abstracts. A comparison between those authored by humans and those generated artificially by AI. The results showed both similarities and differences in move patterns, distinctions in obligatory and conventional moves. Moreover, human-written abstracts showed diversity, maybe serving the purpose of attracting readers and fostering engagement with the associated articles. However, it is imperative to acknowledge the study's inherent limitations, which include a restricted dataset and the lack of constraints on the AI' "prompt code" used for AI-generated abstracts. Such constraints may have improved a comprehensive analysis of AI's potential of abstract composition. Future questions might focus on additional aspects such as lexical selection and writing comprehensiveness between human and AI-generated abstracts.

References

Amnuai, W. (2019). **Analyses of rhetorical moves and linguistic realizations in accounting**.
 research article abstracts published in international and Thai-based journals, SAGE
 Open, 1-9.

Anderson , N.,   Belavy , D., Perle , S., Hendricks , S., Hespanhol , L.,Verhagen, E. & Memon,
 A. (2023). **AI did not write this manuscript, or did it? Can we trick the AI text
 detector into generated texts? The potential future of ChatGPT and AI in Sports
 & Exercise Medicine manuscript generation.** BMJ Open Sport & Exercise Medicine,
 9,e001568.

Biber, D. (2006). **University language: A corpus-based study of spoken and written
 registers.** John Benjamins.

Biber, D., Connor, U. & Upton, T. (2007). **Discourse on the Move: using corpus analysis to
 describe discourse structure.** John Benjamins Publishing Company.

Dergaa, I., Chamari, K., Zmijewski, P., Ben Saad, H. (2023). **From human writing to artificial
 intelligence generated text: examining the prospects and potential threats of
 ChatGPT in academic writing.** Biol Sport, 40(2), 615–622.

El-Dakhs, D. A.S. (2018). **Comparative genre analysis of research article abstracts in more
 and less prestigious journals: Linguistics journals in focus.** Research in Language, 16(1),
 47-63.

Elali, F. & Rachid, L. (2023). **AI-generated research paper fabrication and plagiarism in
 the scientific community.** Patterns, 4, 1-4.

Flowerdew, J. (2004). **The discursive construction of a world-class city.** Discourse &
 Society, 15(5), 579-605.

Gao, C. A., Howard, F.M., Markov, N.S., Dyer, E.C., Ramesh, S., Luo, Y., & Pearson, A.T. (2023).
 **Comparing scientific abstracts generated by ChatGPT to real abstracts with
 detectors and blinded human reviewers.** NPJ Digital Medicine, 6 (75), 1-5.

Hyland, K. (2000). **Disciplinary discourses: Social interactions in academic writing.** The
 University of Michigan Press.

Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). **AI, write an essay
 for me: A large-scale comparison of human-written versus ChatGPT-generated
 essays.** arXiv:2304.14276v1.

Hinton, M. & Wagemans, J. (2023). **How persuasive is AI-generated argumentation? An analysis of the quality of an argumentative text produced by the GPT-3 AI text generator.** Argument & Computation, 14, 59-74.

Holmes, J., (1997). **Review of women, men, and politeness.** Language, 73(2),395. DOI: 10.2307/416031.

Kanafani, A. (2022). **A move analysis of abstracts in a Scopus-indexed applied linguistics journal from different quartiles.** RETORIKA: Journal Ilmu Bahasa, 8(1), 1-12.

Kaya, F. & Yagiz, O. (2020). **Move analysis of research article abstracts in the field of ELT: A comparative study.** Journal of language and linguistic studies, 16(1), 390-404.

Kitjaroenpaiboon, W. (2021). **Rhetorical move and multidimensional analyses of applied linguistics research abstracts.** Journal of Language and Culture, 40(2), 137-165.

Kumar, A. (2023). **Analysis of ChatGPT Tool to Assess the Potential of its Utility for Academic Writing in Biomedical Domain.** BEMS Reports,9(1), 24-30.

Lee, J. H., Hu, T., & Han, H. (2018). **Discourse Analysis of the Artificial Intelligence Assistant.** International Journal of Advanced Science and Technology, 117, 67-76.

Maporn, S., Chaiyasuk, I. & Nattheeraphong, A. (2023). **Rhetorical moves of applied linguistics research article abstracts on Scopus-Indexed journals: Contrastive analysis among the three research approaches.** International Journal of Sociologies and Anthropologies Science Reviews (IJSASR), 3(2), 97-104.

Marti, G. (2022). **From data to trade: A machine learning approach to quantitative trading.** SSRN, 4315362.

Macdonald, C., Adeloye, D., Sheikh, A. & Rudan, I. (2023). **Can ChatGPT draft a research article? An example of population-level vaccine effectiveness analysis.** J Glob Health, 13, 01003.

Menon, V., Varadharajan, N., Praharaj, S.K., & Ameen, S. (2020). **Why do manuscripts get rejected? a content analysis of rejection reports from the Indian Journal of Psychological Medicine.** Indian Journal of Psychological Medicine, 44(1), 59-65.

Ortega-Martín, M., García-Sierra, O., Ardoiz, A., Álvarez, J., Armenteros, J. & Alonso, A. (2023). **Linguistic ambiguity analysis in ChatGPT.** arXiv:2302.06426v2 [cs.CL] 20 Feb 2023

PePeau-Wilson, M. (2023). **AI pass U.S. medical licensing exam.** MedpageToday.

Phonhan, P. (2021). **Generic structure of research article abstracts in technical education: A move- based study.** Journal of Liberal Arts, Prince of Sonkla University, 13(1), 282-306.

Salvagno, M., Taccone, F. & Gerli, A. (2023). **Can artificial intelligence help for scientific writing?.** Perspective, 27(75), 1-5.

Stokel-Walker, C. (2023). **ChatGPT listed as author on research papers: many scientists disapprove.** Nature, 613, 620-621.

Sukhapabsuk, D. (2020). **A genre analysis of linguistics research article abstracts: A comparative study with Hyland's moves.** Journal of Language and Linguistics, 39(2), 137-155.

Zand-Moghadam, A. & Zhaleh, K. (2022). **Comparative rhetorical move analysis of qualitative, quantitative, and mixed methods research article abstracts in Iranian vs. international applied linguistics journals.** Journal of Modern Research in English Language Studies, 9(3), 25-47.