INDONESIAN JOURNAL OF APPLIED LINGUISTICS
Vol. 15 No. 3, January 2026, pp. 561-572

Available online at:
https://ijal.upi.edu/index.php/ijal/article/view/101

https://doi.org/10.17509/shgzwd41

# Metadiscourse in ChatGPT-generated and human-written research articles in linguistics

**Loan Thi To Pham**
*Faculty of English, Thuongmai University, Hanoi, Vietnam*

**ABSTRACT**

ChatGPT has served as a reference for researchers seeking to generate scholarly articles since its emergence in 2022. Its increasing use in academic writing necessitates exploring how ChatGPT generates discourse compared with human authors. This study investigates the use of interactive and interactional metadiscourse in 100 human-authored and 100 ChatGPT-generated English-language linguistics research articles. Using AntConc, metadiscourse markers were identified and analyzed, drawing on Hyland's (2005) model, revealing both convergences and divergences between the two corpora. Transitions occur the most frequently among interactive markers, reflecting ChatGPT's ability to emulate human discourse organization. Evidentials are commonly used in both datasets; however, the striking difference is that the AI corpus often contains inaccurate and fictitious citations, raising concerns about the trustworthiness of its content. Moreover, the frequent use of frame markers in ChatGPT texts, compared to a human corpus, indicates a strong reliance on formulaic structure, leading to a lack of sophisticated argumentation. The sparse use of code glosses and endophoric markers in ChatGPT also suggests limited references within the text. Interactionally, hedges, boosters, and engagement markers are predominant in both corpora as conventions in academic writing. However, the ChatGPT corpus lacks self-mentions and fewer attitude markers, which typically reflect the impersonal nature of AI-produced texts. These findings offer practical pedagogical implications and significant insights into quality control, academic integrity, and plagiarism detection for all stakeholders amid the rapid advancement of AI.

**Keywords:** ChatGPT-generated texts; human-written texts; linguistics; metadiscourse; research articles

| **Received:** | **Revised:** |
|---|---|
| 14 July 2025 | 18 November 2025 |
| **Accepted:** | **Published:** |
| 28 January 2026 | 30 January 2026 |

**How to cite (in APA style):**
Pham, L.T.T. (2026). Metadiscourse in ChatGPT-generated and human-written research articles in linguistics. *Indonesian Journal of Applied Linguistics*, *15*(2), 561-572. https://doi.org/10.17509/shgzwd41

## INTRODUCTION

The rapid advancement of artificial intelligence, especially the emergence of large language models such as ChatGPT, has brought both opportunities and challenges in education (Adeshola & Adepoju, 2024; Baig et al., 2024; Kasneci et al., 2023; Jiang & Hyland, 2025b; Rahman & Watanobe, 2023). One prominent opportunity is its affordances in generating ideas and removing writer's block for researchers and students (Cotton et al., 2023). However, the use of ChatGPT in academia has prompted major concerns regarding the risks of plagiarism and academic integrity (Bin-Nashwan et al., 2023; Lo, 2023). To this end, higher education institutions worldwide have implemented different measures to address these challenges, including either a ban on ChatGPT access and AI tools (Tlili et al., 2023) or guidelines and workshops to use it ethically in academic works (Cotton et al., 2023). Apart from institutional actions, these issues can also be addressed by exploring how AI and human writers differ in discourse in general and the way authors organize the texts, position themselves, and engage readers through metadiscourse in particular (Hyland, 2005; Jiang & Hyland, 2025a). Thus, an increasing number of studies have focused on the distinguishing features between sophisticated

Email: loanptt@tmu.edu.vn

ChatGPT-generated and human-written discourse with an emphasis on how to use it responsibly while maintaining academic honesty (Gao et al., 2023; Liao et al., 2023; Mindner et al., 2023). Previous studies have addressed the differentiations of ChatGPT as compared to human-written texts, noting that the former generally focus on fluency, logic, and general terminologies while the latter concentrates more on concrete, diverse, and informative texts (Liao et al., 2023).

Additionally, as discovered by Gao et al. (2023), AI-generated texts are characterized as vague and formulaic, resulting in a lack of details and in-depth information. To address these concerns and reinforce academic integrity in education, Alam et al. (2024) explore seven techniques for detecting ChatGPT-generated and student-authored texts, including discourse particles, conversational cues, grammar accuracy and clarity, formulaic genre structure, paragraph numbering, transitional expressions, and self-acknowledgment. These characteristics emphasize the notable distinctions between AI-generated and human-written texts, particularly the typical presence of variability and personalization in human writing, elements that are largely absent in ChatGPT texts. Similarly, Jiang and Hyland (2025b) also note that although ChatGPT-created texts are well-organized, they lack the rhetorical engagement demonstrated evidently in students' essays, an aspect that AI hardly replicates. Specifically, metadiscourse elements such as hedges, boosters, and attitude markers are more frequently used in students' essays, reflecting humans' frequent use of personalized and interactive language. These features accommodate a dialogic relationship with readers, which is considered a critical condition in academic writing. However, AI-generated texts have the tendency to employ fewer rhetorical moves, leading to an impersonal and less engaging tone.

Metadiscourse plays a critical role in enhancing the clarity, persuasiveness, and reader engagement in academic writing (Appel & McKay, 2025; Ho & Li, 2018; Hyland, 2005; Qiu et al., 2024). Its core function in fostering writer-reader interaction is well documented in various academic genres such as research articles (Saz-Rubi, 2011; Hu & Cao, 2015; Hyland, 2005; Hyland & Jiang, 2018; Mu et al., 2015), undergraduate student essays (Appel & McKay, 2025, Ho & Li, 2018; Cheng & Steffensen, 1996; Lee & Deakin, 2016), postgraduate master and PhD theses (Akbas, 2012; Ayieta Ondondo, 2020; Akoto, 2020; Deng et al., 2021; Lee & Casal, 2014; Lin, 2005), textbooks and coursebooks (Crismore, 1984; Hyland, 1999). Given that these studies deepened our understanding of metadiscourse in conventional human-authored texts, research remains limited in how such features are adopted in AI-generated texts.

This article is an attempt to examine the distinctive features of interactive and interactional metadiscourse in AI and human-created scholarly research, focusing on how the two groups organize their texts, construct authorial stance, and engage with readers. The metadiscourse preferences identified in both corpora are closely linked to ethical and scholarly concerns because they demonstrate the authorial voice and conformity to academic conventions. Therefore, this study aims to inform educators, teachers, reviewers, editors, and policymakers about the necessary empirical evidence to differentiate between human-written and AI-generated research texts. This investigation is critical in the context of growing reliance on AI tools in the academic community and increasing ethical concerns to uphold scholarly standards.

**Metadiscourse in Research Articles**

Over the past two decades, Hyland's (2005) metadiscourse model, characterized as a tool for facilitating writer-reader interaction, has attracted considerable scholarly attention. Metadiscourse is perceived as the writer's intention to establish interaction with both the textual content and the audience. It is classified into two categories, including interactive and interactional factors, in which the first component guides the reader through the texts while the second communicates the writer's stance and establishes a relationship, thereby fostering engagement with the readers (Hyland, 2005). Interactive resources enable readers to interpret the information flow with the elements depicted in Table 1.

**Table 1.**
*Elements of Interactive Resources (Hyland, 2005)*

| No. | Elements | Descriptions | Examples |
|---|---|---|---|
| 1. | Transition | Signal relationships between ideas, including addition, contrast, and consequence | *In addition, However, Therefore,* |
| 2. | Frame markers | Organize the structure of the text | *First, Second, In conclusion,* |
| 3. | Endophoric markers | Help readers refer to the other part of the same text for relevant information | *As noted above, See table 1…* |
| 4. | Evidentials | Indicate the source of information outside the text, mostly citation | *According to Smith (2020)…* |
| 5. | Code glosses | Clarify the information | *For example, namely, such as.* |

On the other hand, interactional resources reflect the writer's efforts to actively engage the readers in their discussion. Table 2 displays the details of interactional metadiscourse.

**Table 2**
*Interactional Metadiscourse (Hyland, 2005)*

| No. | Elements | Descriptions | Examples |
|---|---|---|---|
| 1. | Hedges | Express caution or uncertainty, withholding full commitments | *Might Perhaps It seems…* |
| 2. | Boosters | Express certainty, emphasizing the writer's confidence in proposition | *Clearly, Obviously, In fact,* |
| 3. | Attitude markers | State writers' attitude towards the proposition, including evaluation, emotion, and stance | *Unfortunately, Surprisingly,* |
| 4. | Engagement markers | Address the readers, aiming to involve them in the discourse | *Consider Note that As you can see,* |
| 5. | Self-mentions | Refer to the authors of the text | *I, We, My, Our* |

Extant literature indicates that metadiscourse has been extensively investigated in research article genre concerning the differentiations between native and non-native authors' use of metadiscourse (Ajideh et al., 2024; Eghtesadi & Navidinia, 2009; Gholami & Ilghami, 2016); metadiscourse variations across diverse academic disciplines (Bondi, 2005; Dahl, 2004; Khedri et al., 2013); metadiscourse distinctions among varied academic types such as research articles, master's theses, doctoral theses (Alharbi, 2021; Amnuai et al., 2023; Kawase, 2015). Recently, the rapid development of artificial intelligence (AI) has triggered other issues relating to AI, particularly ChatGPT-created and human-authored discourse. Given that conventional studies have particularly focused on research articles written by human scholars, the emergence of ChatGPT necessitates more studies on variables that redefine the practices of discourse. One of the key aspects focuses on how these two types of texts created by ChatGPT and human writers differ in the use of metadiscourse in academic writing (Amirjalili et al., 2024; Jiang & Hyland, 2025a; Jiang & Hyland, 2025b). Results show that while ChatGPT-created essays maintain coherent structure and logical organization, they demonstrate a relatively low utilization of interactional metadiscourse markers, namely hedges, boosters, and attitude markers. The tone, accordingly, is more impersonal and objective, leading to the lack of rhetorical engagement with readers. In contrast, essays written by students adopt more stance markers and personalized expressions to promote reader-interaction.

However, existing studies focus primarily on essays and some limited research on research article abstracts between human-written and AI-generated texts (Gao et al., 2023; Zhang & Zhang, 2025a; Zhang & Zhang, 2025b). Focusing solely on abstracts may hinder the exploration of important patterns of metadiscourse use across the full structure of research articles. To develop a more holistic understanding, further investigation is needed to uncover the metadiscursive features in ChatGPT-generated and human-written research articles. This study is, therefore, an attempt to offer more insights into academic writing in the transformative era of AI.

## ChatGPT and Its Use of Interactive and Interactional Metadiscourse in Academic Writing

ChatGPT (Chat Generative Pre-Trained Transformer) is trained on an extensive corpus of Internet-based texts, thereby exhibiting remarkable capability to generate logically coherent, contextually relevant, and grammatically correct discourse (Briana, 2024; Curry et al., 2024; Moorley & Watson, 2023). These positive features have positioned ChatGPT as an effective tool in educational and professional discourse (Cotton et al., 2023). In a scholarly setting, it has a significant influence on academic writing (Warschauer et al., 2023) due to its ability to edit, revise, and create academic texts (Su et al., 2023). However, critical concerns arise regarding how strictly its generated texts conform to established academic discourse conventions, including metadiscourse features, characterized by the language to organize a text, convey the writer's stance, and engage the reader. While ChatGPT exhibits the mastery of generating texts with appropriate use of hedging, formal tone, and logical transitions, it falls short of the rhetorical awareness that typically defines human writing. Noticeably, ChatGPT predominantly employs interactive metadiscourse markers, including transitional words such as *however* and *therefore* (Jiang & Hyland, 2025b). Additionally, Jiang and Hyland (2025a) discovered that ChatGPT has a preference for less complex syntactic structures with the particular use of *determiner + Noun* pattern and heavy reliance on anaphoric references, whereas students' writing focuses more on appropriate utilization of syntactic structures and cataphoric references.

Interactionally, ChatGPT-generated texts tend to underuse hedges, boosters, self-mentions, and engagement markers. In contrast, students' texts employ a wider range and greater number of interactional resources to manage stance, engage

readers, and establish a persuasive authorial voice. These devices are critically vital in academic writing as they facilitate writer-reader rapport in the co-construction of meaning, while demonstrating writers' modesty or confidence in their propositions. In line with these arguments, Mo & Crosthwaite's (2025) study reveals that human-written essays use a greater number and wider range of stance and engagement features in comparison with AI-generated discourse. This is evident in the common use of hedges (e.g., *would, perhaps, seems*) and boosters (e.g., *indeed, in fact, clearly*) in human writing texts, whereas those generated by AI tend to maintain a higher frequency of attitude markers, particularly with over-reliance on the word "*significant*". Furthermore, its output texts are typically repetitive expressions along with unusual or inconsistent language usage (Cotton et al., 2023). As such, human writers are more capable of engaging their readers in the communication process, an aspect that ChatGPT algorithms still struggle to achieve (Berber Sardinha, 2024). Curry et al. (2024) postulate that although AI is designed to generate human-like text and emulate human-like intelligence, it is unable to meet the analytical standards of human writers. Moreover, as argued by Korteling et al. (2021), AI tools are unlikely to completely replicate human intelligence in the near future. This suggests that neither the language they produce nor the tasks they perform can entirely substitute for human language and cognitive capacities.

This study aims to investigate how ChatGPT-generated and human-authored linguistics research articles align with or deviate from each other in the use of metadiscourse. To date, metadiscourse in the linguistics field has been extensively documented and empirically investigated; therefore, the findings in this discipline will make an enormous contribution to the body of literature. Consequently, the results shed light on the evolving discourse of academic writing by analyzing the strengths and shortcomings of ChatGPT's use of interactive and interactional resources. The research also offers practical significance since its comparative findings serve as a source of reference to enhance the natural features and adaptability of AI-generated texts. The research seeks the answers to the following questions:

1. What are the similarities and differences in the use of interactive metadiscourse between human-written and ChatGPT-generated research articles?
2. How do human-written and ChatGPT-generated research articles differ in their use of interactional metadiscourse in expressing authorial stance and engaging readers?

## METHODS

### Corpus

A research article is an academic genre that employs a great density of metadiscourse to demonstrate a writer's guide, express their stance, and engage readers through the text (Hyland, 2005). This study aims to investigate the distinctions in the use of metadiscourse between two corpora of research articles produced by humans and those generated by ChatGPT. The first corpus (coded from H1 to H100) comprises 100 research articles derived from two well-established journals in the area of linguistics and language, namely *English for Specific Purposes* and *Journal of English for Academic Purposes,* in the period of 2020-2024. This period was chosen to ensure consistency in disciplinary norms and academic writing conventions, thereby minimizing the differences in metadiscourse use.

The second corpus (coded from C1 to C100) includes 100 research articles generated by ChatGPT with the same topics, ensuring that the article's title, content, and structure resemble those in the human-authored texts. Accordingly, prompts were created to produce articles that conform to standard academic traditions with required length, including title, abstract, introduction, methodology, results, discussion, and conclusion. Subsequently, these texts were carefully reviewed to ensure alignment with the structure and content of human-written counterparts. ChatGPT-generated articles were kept unchanged to preserve the authenticity of the corpus. Li et al. (2023) suggested three common prompting templates, including example, context, and query. As examples and queries are less suitable for generating research articles, this study exclusively focuses on context prompts. For example:

> *"You are a researcher in linguistics. Prepare a research article for a peer-reviewed journal about the topic: Engagement in doctoral dissertation discussion sections written by native English speakers. The article should be 5000 words and comply with APA 7th style."*

The same template was repeatedly used across different topics to maintain a consistent output. However, one of the inherently limited features of ChatGPT is its inability to generate lengthy academic articles (Tudino & Qin, 2024); therefore, the collected data for each AI-generated article actually falls short of the required 5000 words. Typically, they range from only 1000 to 2000 words, resulting in a discrepancy in length between AI-generated and human-written corpora. Given that the differences in text length may affect frequencies of metadiscourse, the comparison still remains feasible since all the analyses were normalized per 10,000 words. Additionally, the current study aims at metadiscourse choices between the two corpora rather than ultimate word counts. The basic

information of the two corpora is presented in Table 3.

**Table 3.**
*Information of the Corpora*

| Corpus | Number of articles | Number of words |
|---|---|---|
| Human-written corpus | 100 | 729,535 |
| ChatGPT corpus | 100 | 233,561 |

**Data analysis**

This study adopted AntConc software (Anthony, 2024) to analyze the interactive and interactional metadiscourse in the two corpora, including ChatGPT-generated and human-authored research articles. AntConc facilitates keyword analysis, concordance generation, collocation patterns, and word frequency. Hyland's (2005) framework, comprising interactive and interactional metadiscourse, was used for analysis. Specifically, interactive resources include transitions, frame markers, endophoric markers, evidentials, and code glosses, while interactional metadiscourse is composed of hedges, boosters, attitude markers, engagement markers, and self-mentions. All the articles of the two corpora were converted into plain texts and uploaded into AntConc for analysis. The frequency of each metadiscourse marker was calculated as 10,000 words.

All articles from both corpora were converted into plain text format and uploaded into AntConc for analysis. The frequency of each metadiscourse marker was normalized to occurrences per 10,000 words. To ensure the reliability, a cross-checking phase conducted by two researchers was carried out to affirm the accuracy. Discrepancies were then resolved, resulting in 96% inter-rater agreement.

## FINDINGS

**Interactive Metadiscourse in The Two Corpora**

Both ChatGPT-generated and human-authored research articles utilize a number of interactive markers to facilitate readers in navigating and understanding the flow of the discourse. The comparative result below reveals that transition markers are used the most frequently, whereas evidentials occur the least. Table 4 demonstrates the frequency, percentage, and normalized frequency (per 10,000 words) of interactive metadiscourse markers, including transitions, evidentials, frame markers, code glosses, and endophoric markers in 100 research articles written by humans and 100 generated by ChatGPT.

**Table 4.**
*Interactive Metadiscourse Features in Human-Authored and ChatGPT-Generated Research Articles*

| | Human-authored research articles | | | ChatGPT-generated research articles | | |
|---|---|---|---|---|---|---|
| | Frequency | Percentage | Per 10,000 words | Frequency | Percentage | Per 10,000 words |
| Transitions | 1565 | 28.95% | 21.5 | 1021 | 29.47% | 43.73 |
| Evidentials | 1221 | 22.59% | 16.7 | 806 | 23.26% | 34.50 |
| Frame markers | 1022 | 18.91% | 14.0 | 623 | 17.98% | 26.67 |
| Code gloss | 992 | 18.36% | 13.6 | 612 | 17.66% | 26.20 |
| Endophoric markers | 605 | 11.19% | 8.3 | 403 | 11.63% | 17.26 |
| Total | 5405 | 100.00% | 74.1 | 3465 | 100.00% | 148.36 |

As seen from Table 4, transition markers were predominantly used in both corpora, accounting for 28.95% and 29.47% in human-authored and ChatGPT-generated texts, respectively. Notably, the normalized frequency per 10,000 words shows that transitions appeared 21.5 times in human articles and 43.73 times in ChatGPT texts, indicating that ChatGPT employs this type of discourse marker at a rate twice as high as human authors. Most commonly used transitions include: *However, Therefore, In addition, As a result, Accordingly, Furthermore, etc.,* as in the following examples:

1. ***In addition,*** *they differ in the degree to which they chose such in-between linguistic behaviors depending on where they are in their academic career (e.g., as bachelor's thesis writers, as master's thesis writers, as doctoral dissertation writers, or as conference abstract writers) (H85)*
2. ***However,*** *the limitations of automated feedback such as lack of contextual nuance and generic advice underscore the need for teacher mediation. (C7)*
3. ***Therefore,*** *evaluating their potential for supporting incidental learning is essential. (C35)*
4. ***As a result****, there are 304 categorisations in EN and 177 in ES, as shown in Table 4. (H13)*
5. *Phrasal structures,* ***on the other hand****, reflect more advanced writing proficiency and are associated with academic writing (Norris & Ortega, 2009; Ruan, 2018). (C98)*

Evidentials made up the second-highest percentage, comprising 22.5% in human texts and 23.26% in ChatGPT-generated articles. However, the frequency per 10,000 words was notably higher in the ChatGPT corpus. This suggests that both human writers and ChatGPT are well aware of the role of citing evidence or external sources to support their arguments, as in examples (6) and (7).

6. ***Hyland (2004) emphasizes*** *that academic conflict is often hedged to preserve interpersonal harmony and disciplinary decorum (C15)*
7. ***According to Biber et al.*** *(1999), the three key criteria for generating lexical bundles are the length of word combinations, the frequency threshold, and the range of distribution. (H19)*

Frame markers were relatively common in the two corpora, constituting 17.98% in ChatGPT texts as compared to only 18.91% in human texts, making them approximately two times more frequent per 10,000 words (14.0 in the human corpus versus 26.67 in the ChatGPT corpus). This suggests that AI relies heavily on markers to facilitate transitions, clarify structure, and guide the flow of information throughout the discourse, as in Examples 8 and 9.

8. *The motivation for this study stems from two key considerations.* ***First****, exemplification is inherently multifunctional—it serves both referential and interpersonal purposes.* ***Second****, the way exemplification is realized linguistically and rhetorically may vary across academic disciplines, shaped by disciplinary conventions and epistemologies. (C30)*
9. ***In conclusion,*** *"this" is a powerful cohesive device in engineering semi-popularization articles, especially when functioning as an encapsulator. (C10)*

In the same vein, human-authored articles used frame markers to manage the flow of ideas and the coherence of academic discourse. Sequential markers such as "*First, Second, Next, Then, In addition"* were employed to ensure the logical order of information (Example 10), while concluding markers like "*To conclude, In summary, Ultimately, In short"* function as the signal to synthesize the key points and close the discussion (Example 11).

10. ***First****, master's thesis writers in the reported study use high numbers of 3rd-person NPs (N ¼ 270; 47%) and inanimate NPs (N ¼ 124; 22%) in their work.* ***Second****, they use more 1st-person plurals than singulars (N ¼ 140; 24% vs. N ¼ 39; 7%). (H85)*
11. ***In conclusion,*** *it is hoped that the present study will shed light on the field of academic writing research in a way that provides insights for curriculum designers, instructors, and researchers … (H15).*

Finally, code glosses and endophoric markers occurred less frequently in the two corpora, reflecting a lower emphasis on explanation, clarification, elaboration, and internal referencing. Code glosses help the reader better understand the text through typical expressions, such as *in other words, such as, for example, that is, i.e., namely, like,*" etc. (Example 12). Endophoric markers enable the readers to refer to information elsewhere within the text, using phrases like *as mentioned above, see Table 5, or in the next section* (Example 13).

12. ***For instance,*** *hard sciences may prefer concise factual exemplars, while soft disciplines may use more elaborative and argumentative examples. (C30)*
13. *Purpose bundles were the most frequent bundles, with 30 different types and 388 occurrences* ***(see Table 5)****. (H19)*

However, phrases such as *"see Table 5, see Figure 1"* are rarely found in ChatGPT corpus due to its current limitations, which exclude images, figures, and tables in its generated texts (Bubeck et al., 2023). As a result, the generated texts lack rhetorical complexity, leading to the presentation of merely descriptive information rather than multimodal communication.

**Interactional Metadiscourse in The Two Corpora**

A total of 2746 metadiscourse instances were discovered in human-authored articles, whereas 1169 cases were identified in the ChatGPT corpus, as illustrated in Table 5.

**Table 5.**
*Interactional Metadiscourse Features in Human-Authored and ChatGPT-Generated Research Articles*

| | Human-authored research articles | | | ChatGPT-generated research articles | | |
|---|---|---|---|---|---|---|
| | **Frequency** | **Percentage** | **Per 10,000 words** | **Frequency** | **Percentage** | **Per 10,000 words** |
| Hedges | 1499 | 54.58% | 20.55 | 801 | 68.52% | 34.29 |
| Boosters | 656 | 23.89% | 8.99 | 223 | 19.08% | 9.55 |
| Engagement markers | 230 | 8.38% | 3.15 | 90 | 7.70% | 3.85 |
| Attitude markers | 205 | 7.47% | 2.81 | 55 | 4.70% | 2.35 |
| Self-mentions | 156 | 5.68% | 2.14 | 0 | 0.00% | 0.00 |
| Total | 2746 | 100.00% | 37.64 | 1169 | 100% | 50.04 |

It is noticeable that hedges were the most frequently used markers in both human and AI texts (54.58% in human texts versus 68.52% in ChatGPT texts), indicating that both corpora tend to soften claims and possibly avoid strong claims to align with academic norms of caution, as in Examples 14-17.

14. *We* ***might*** *find an explanation for these divergences by looking at the cultural contexts in which the HTPT are received. (H13)*

15. ***It seems*** *that improved language skills may have helped students better understand the content of their courses. (H45)*
16. *Future research* ***might*** *explore more disciplines, examine longitudinal changes, or investigate reader responses to different conclusion styles. (C53)*
17. *The findings* ***suggest*** *the need for explicit instruction in engagement strategies to enhance rhetorical awareness and persuasive strength in EFL academic writing. (C79)*

While hedges function to express caution or uncertainty, boosters demonstrate writers' confidence, thereby facilitating the strength of propositions. Among the metadiscourse markers, boosters ranked the second-highest frequency after hedges in both corpora, with the human-written corpus (23.89%) exhibiting a slightly higher percentage than that of ChatGPT (19.08%). By using boosters, human researchers establish the authority and attempt to convince readers of the significance of their arguments (Examples 18, 19, 20).

18. ***In fact,*** *the one recurrent pattern identified, composed of an imperative 'verb þ complement', does not occur in all the strategies described. (Cook, 2001; Janoschka, 2004). (H13)*
19. *These results indicate that gender did not* ***clearly*** *relate to challenges experienced by students in this EMI context, with the exception of listening. (H25)*
20. *… and (ii) that while there* ***certainly*** *is overlap between the groups in terms of the three most frequent fillers, the novice writers make more frequent use of important (it is important to) and importance (the importance of the), similar to findings in studies such as Granger and Larsson (2021) and Larsson (2019a). (H48)*

However, a significant divergence exists between the rhetorical purpose of articles written by humans and those generated by ChatGPT. For example, the use of the booster *"obviously"* in Examples 21 and 22, drawn from ChatGPT and a human-written corpus. However, they differ in that example 21 signals the writer's certainty about the claim as common sense or widely accepted knowledge. In contrast, example 22, written by a human researcher, employs *"obviously"* to assert the proposition involving Hyland's notion of threat and Tang and John's notion of power. This means human writers are more aware of intertextuality in a scholarly context to foreground claims in a discourse community.

21. ***Obviously****, being able to navigate these exchanges effectively can enhance one's perceived credibility and authority in the field. (C52)*
22. ***Obviously****, Hyland's notion of threat and Tang and John's notion of power are the two edges of the author authority double sword: to make an argument places authors in a position to claim expertise while at the same time exposes them to scrutiny by colleagues. (H85)*

An engagement marker serves as a tool to involve the readers directly in the discussion, making the argument more persuasive and collaborative. The data reveals that the human corpus used this type of metadiscourse 230 times (8.38%, 3.15 per 10,000 words), whereas ChatGPT employed 90 instances (7.70%, 3.85 per 10,000 words). As such, there is a tendency that human writers use engagement markers slightly more frequently than ChatGPT, possibly because they are more attuned to audience expectations, as in Examples 23 and 24.

23. ***As can be seen*** *in the example in Table 2, the point is made complete in the combination of Findings and Nub. (H50)*
24. *It is interesting to* ***note that*** *the functions of a large majority of bundles (i.e. 91%), that is, number 1–5 in the table, reflected IMRD organisation and were closely related to the five moves of research article abstracts proposed in Dos Santos (1996) and Hyland (2004). (H19)*

The lower percentage of ChatGPT corpus in the use of engagement markers may be due to the fact that its content is generated to simulate information that prioritizes objectivity over direct interaction with readers. Ultimately, attitude markers and self-mentions represent the least frequently used types of metadiscourse in both corpora. Human researchers convey their affective stance through attitude markers at a rate of 7.47% (2.81 per 10,000 words). In ChatGPT-produced texts, this marker accounts for 4.7% (2.35 per 10,000 words), as shown in Examples 25-26.

25. ***Importantly,*** *the teachers also acknowledged the tool's affordances for enhancing self-regulated genre instruction to novice L2 writers (H7)*
26. ***Not surprisingly,*** *a closer look at these eight measures using pairwise comparisons for each category showed that measures based on T-units vs. clauses were strongly correlated, as shown in Figs. 3 and 4 below displaying CN. T vs. CN.C and CP. T and CP.C, respectively. (H8)*
27. *The text analysis shed light on Fulbright writers' rhetorical practices in their statements, but the investigation into raters' perspectives provided additional information concerning what the raters reacted to, and* ***importantly****, what they deemed effective or ineffective when reviewing statements (RQ#2). (H14)*
28. ***Interestingly****, some students internalized their struggles, attributing academic setbacks to personal inadequacy rather than systemic linguistic challenges. (C25)*

The above examples demonstrate how human researchers utilize attitude markers to express their affective and evaluative stance as a way to highlight

the significance of the claims. However, the low frequency of attitude markers in ChatGPT articles suggests that AI tends to adopt a neutral tone, reflecting its fact-based content, which inherently leads to its limitations in emulating humans' rhetorical nuances.

One of the most notable divergences is the complete absence of self-mentions metadiscourse (for example, *I, we, my, our*) in the ChatGPT corpus. In contrast, the human corpus employs 156 instances, constituting 5.68% of total features (2.14 per 10,000 words). The result reveals that human authors tend to use self-mentions to establish authorial stance and assert their responsibility in claiming the argument, while impersonal tone is one typical trait of ChatGPT texts. Examples 29-30 illustrate this point.

29. *Against these backdrops, **I** report a series of studies carried out at Xi'an International Studies University (XISU) by my colleagues, students, and **me** on the use of self-referring expressions by Chinese AWE writers of different statuses in the general area of language studies. (H85)*
30. *With respect to formal realisation, evaluation in **our** strategies is observed to be "expressed lexically -rather than grammatically", which has been seen as characteristic of informational-persuasive web documents (Biber & Zhang, 2018, p.97) (H13)*

In example 29, the first-person singular pronoun "I" was used to signal authorial responsibility for reporting the details in which the phrase "my colleagues, students, and me" emphasizes the role of these agents in conducting a series of studies. Likewise, example 30 used the plural self-mention "our" to affirm the ownership of the strategies, indicating the author's effort to foster the readers' engagement and inclusivity. In short, the examples reinforce human writers' strategic deployment of self-mentions to establish the authorship and credibility, as well as involve the readers in the propositions. Such features are typically absent in ChatGPT texts, characterized by an impersonal and neutral style.

## DISCUSSIONS

The results from this study offer significant insights into the convergences and divergences of interactive and interactional metadiscourse in research articles written by human researchers and those generated by ChatGPT in terms of communicative purposes, audience-targeted awareness, and authorial presence. Pertaining to interactive metadiscourse, transitions constitute the highest percentage use of all types in both corpora, as they are the tools to link ideas together (Swales & Feak, 2012). Notably, the dominant usage of transitions in ChatGPT texts reflects the AI tool's ability to replicate the standard of humans' discourse patterns and write seemingly real scientific articles in the context that language processing and machine learning algorithms keep evolving steadily (Kung et al., 2023). Moreover, transitions (*However, Therefore*) are the easiest way for the algorithm to create the illusion of logical cohesion.

Furthermore, evidentials are among the commonly used markers in human and ChatGPT corpus at a relatively high rate, suggesting that AIs are making efforts to enhance credibility by referencing external sources (Hyland & Jiang, 2018). However, it should be noted that while human authors use evidentials with caution to situate the study within the existing literature, ChatGPT tends to integrate inaccurate and fictitious citations (Ariyaratne et al., 2023). This raises concerns about the authenticity and reliability of ChatGPT-generated texts. Although these texts may be seemingly authentic articles in terms of metadiscourse use, the accuracy of the information still remains questionable (Shen et al., 2023). Consequently, the inconsistent and counterfactual text of AI (Liu et al., 2022; Salge et al., 2022) poses a potential threat to scientific publication and research integrity (Ma et al., 2023) if it is adopted without careful scrutiny. In the same vein, frame markers occur twice as frequently in the ChatGPT corpus in comparison with human-authored texts, indicating AI's heavy reliance on structuring and formulaic expressions for clarity. Nevertheless, such overuse may lead to a lack of rhetorical sophistication in research articles.

On the contrary, code glosses and endophoric markers are among the least adopted metadiscourse in both corpora, with sparser use in ChatGPT. Endophoric expressions such as *"see Table 5, see Figure 1"* rarely occur in AI-generated texts, reflecting its limitations in incorporating intra-textual references, an important feature in academic discourse (Hyland, 2005). This finding goes in line with Bubeck et al.'s (2023) study, highlighting that the formulaic and surface-level nature of AI texts limits the use of multimodal elements such as images, charts, figures, and tables in their generated articles. Consequently, these texts lack the scholarly depth and quality that typically differentiates human-written articles from those generated by ChatGPT.

Regarding interactional metadiscourse, hedges, boosters, and engagement are among the most frequently used features in both corpora. This finding aligns with Jiang and Hyland's (2025b) result, in which these elements are used with high density in both human and ChatGPT texts as a norm in research article discourse to exhibit caution, assert confidence, and engage the readers in the study. It is noteworthy that attitude markers and self-mentions are the least frequent features in the two corpora. However, the lower frequency of

attitude markers and the absence of self-mentions in the AI corpus reflect the fact that AI undermines the authorial presence. This finding is compatible with previous studies, stressing that the nature of AI texts is impersonal and detached (Ariyaratne et al., 2023; Shen et al., 2023). The underlying explanation for these results might be that AI lacks domain-specific knowledge and expertise, making it unable to generate the exact, sophisticated scientific information (Liao et al., 2023). This feature limits the use of self-mentions such as "I, we, my, our" or attitude markers like "importantly, surprisingly" in an AI tool. Accordingly, the adoption of an impersonal and neutral tone becomes common in ChatGPT texts to minimize the risks of inaccuracies, which in turn results in omitting the presence of authorship. Moreover, it can be attributed to ChatGPT's tendency to conform to the long-established academic writing standard, which focuses on objective and impersonal norms, thereby avoiding the use of first-person pronouns. To this end, AI seems to lag behind in updating the recent trends toward the increasing use of first-person self-mentions in human-authored research articles, reflecting an individualized and personal style (Hyland & Jiang, 2017; Zhao, 2023). This highlights ChatGPT's lack of attention to the evolving rhetorical features of academic discourse. Therefore, its stance and engagement strategies should be adjusted to minimize formulaic patterns and enhance the rhetorical flexibility, thereby aligning with academic writing norms. However, this finding should be interpreted with caution since the result may vary when prompts request a personal and reflexive style, potentially leading to different distributions of self-mentions. This, therefore, necessitates further research to draw a more comprehensive conclusion about ChatGPT's use of self-mentions.

Consequently, the findings of this study offer valuable insights to distinguish between human-written and AI-generated texts for quality assurance, plagiarism detection, and research integrity in the context of the increasingly evolving advancement of AI and growing concerns about plagiarism (Bin-Nashwan et al., 2023; Cotton et al., 2023). Additionally, since AI-generated texts have heavy reliance on transition markers, teachers should focus less on instructing transition words that AI adopts easily and more on critical argumentation and positioning, such as self-mention or attitude, and areas that AI still struggles with.

## CONCLUSION

This study analyzes the comparative use of metadiscourse in 100 human-written and 100 ChatGPT-generated research articles. The findings indicate both corpora have strong reliance on transitions and evidentials, showing adherence to conventional structures and patterns of organizing information. However, the overuse of frame markers and the limited use of endophoric and code glosses in the ChatGPT corpus highlight the formulaic nature of AI discourse. Moreover, while both corpora share the convergences in frequent use of hedges, boosters, and engagement markers, ChatGPT corpus's tendency to use fewer attitude markers and self-mentions demonstrates its undermining of authorial voice and detached style, thereby ultimately weakening the rhetorical effectiveness of its generated texts. The results provide significant evidence for all stakeholders, including educators, editors, reviewers, and policymakers, to identify the distinction between human-authored and ChatGPT-produced texts to ensure credibility and academic integrity.

Despite its vital contribution to academic discourse in the rapid advancement of AI, this study has some limitations that call for further research. First, as the corpus generated from ChatGPT might vary according to prompts and versions, future studies could utilize various AI versions for extended investigation to validate the results of this study. Second, the limited samples of 100 human-written and 100 ChatGPT-produced research articles may affect the ability to generalize the overall findings. Other studies could focus on expanding the corpus in the dataset to enhance the validity of the results. Finally, this research solely investigates metadiscourse use in the field of linguistics; thus, other disciplines should be taken into account to explore any possible interdisciplinary variations.

In conclusion, this study sheds light on the metadiscourse features of ChatGPT-generated research articles and highlights major differences from human-authored texts. While AI provides researchers with significant support in terms of coherence and structure, it demonstrates some limitations in conveying the writer's stance and promoting reader engagement. Consequently, ChatGPT may serve as a good source of reference but can not substitute human writers, especially in contexts that require citation-heavy standards and interpersonal communication in published linguistics research.

## REFERENCES

Adeshola, I., & Adepoju, A. P. (2024). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments*, *32*(10), 6159–6172. https://doi.org/10.1080/10494820.2023.2253858

Ajideh, P., Ansarin, A.-A., Zohrabi, M., & Nouri, Z. (2024). An analysis of interactive and interactional metadiscourse: Native vs. non-native author dichotomy. *Teaching English as a Second Language Quarterly*, 43(2), 31–48.

Akbaş, E. (2012). Exploring metadiscourse in master's dissertation abstracts: Cultural and linguistic variations across postgraduate writers. *International Journal of Applied Linguistics and English Literature, 1*(1), 12–26. http://dx.doi.org/10.7575/ijalel.v.1n.1p.12

Akoto, O. Y. (2020). Metadiscourse within a discipline: A study of introduction and literature review chapters of sociology masters' theses. *Indonesian Journal of Applied Linguistics, 10*(2), 471–480. https://doi.org/10.17509/ijal.v10i2.28588

Alam, M. S., Asmawi, A., Haque, M. H. ul, Patwary, M. N., Ullah, M. M., & Fatema, S. (2024). Distinguishing between student-authored and ChatGPT-generated texts: A preliminary exploration of human evaluation techniques. *Iraqi Journal for Computer Science and Mathematics, 5*(3), 287–304. https://doi.org/10.52866/ijcsm.2024.05.03.016

Alharbi, S. H. (2021). An investigation of metadiscourse features in applied linguistics academic research articles and master's dissertations. *Advances in Language and Literary Studies, 12*(1), 46. http://dx.doi.org/10.7575/aiac.alls.v.12n.1.p.46

Amirjalili, F., Neysani, M., & Nikbakht, A. (2024). Exploring the boundaries of authorship: A comparative analysis of AI-generated text and human academic writing in English literature. *Frontiers in Education, 9*, 1347421. https://doi.org/10.3389/feduc.2024.1347421

Amnuai, W., Wimuttisuksuntorn, W., & Wuttikanokkarn, T. (2023). Rhetorical moves and metadiscourse in English abstracts of research articles and master's theses. *Journal of English Language and Linguistics, 4*(2), 29–46. https://doi.org/10.14456/jel.2023.13

Anthony, L. (2024). *AntConc (Version 4.3.1) [Computer Software]*. Tokyo, Japan: Waseda University.

Appel, R., & McKay, R. (2025). The value of interactional metadiscourse in university level writing: Differences between high and low performing undergraduate business students. *English for Specific Purposes, 79*, 30–42. https://doi.org/10.1016/j.esp.2025.03.002

Ariyaratne, S., Iyengar, K. P., Nischal, N., & et al. (2023). A comparison of ChatGPT-generated articles with human-written articles. *Skeletal Radiology, 52*(9), 1755–1758. https://doi.org/10.1007/s00256-023-04340-5

Ayieta Ondondo, E. (2020). Interactional metadiscourse in doctoral thesis writing. *Applied Linguistics Research Journal*, *4*(4), 100–113. https://doi.org/10.14744/alrj.2020.92053

Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in the higher education: A systematic literature review and research challenges. *International Journal of Educational Research, 127*, 102411. https://doi.org/10.1016/j.ijer.2024.102411

Berber Sardinha, T. (2024). AI-generated vs human-authored texts: A multidimensional comparison. *Applied Corpus Linguistics*, 4(1), 100083. https://doi.org/10.1016/j.acorp.2023.100083

Bin-Nashwan, S. A., Al-Dhaen, E. S., & Al-Khalifa, K. N. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society, 75*, 102370. https://doi.org/10.1016/j.techsoc.2023.102370

Bondi, M. (2005). Metadiscursive practices in academic discourse: Variation across genres and disciplines. In M. Bondi & G. Diani (Eds.), *Dialogue within discourse communities: Metadiscursive perspectives on academic genres* (pp. 3–30). Niemeyer. https://doi.org/10.1515/9783110933222.3

Briana, J. C. D. (2024). Is ChatGPT-Produced Text Authentic? A Contrastive Analysis of Cohesive Markers in Human and AI-Generated Text. *Journal of English and Applied Linguistics*, *3*(2), 38-54. https://doi.org/10.59588/2961-3094.1120

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J. A., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S. M., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). *Sparks of artificial general intelligence: Early experiments with GPT-4*. arXiv. https://arxiv.org/abs/2303.12712

Casal, J. E., & Kessler, M. (2023). Can linguists distinguish between ChatGPT/AI and human writing?: A study of research ethics and academic publishing. *Research Methods in Applied Linguistics, 2*(3), Article 100068. https://doi.org/10.1016/j.rmal.2023.100068

Cheng, X., & Steffensen, M. S. (1996). Metadiscourse: A technique for improving student writing. *Research in the Teaching of English, 30*(2), 149–181.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228-339. https://doi.org/10.1080/14703297.2023.2190148

Crismore, A. (1984). The Rhetoric of Textbooks: Metadiscourse. *Journal of Curriculum Studies*, *16*(3), 279–296. https://doi.org/10.1080/0022027840160306

Curry, N., Baker, P., & Brookes, G. (2024). Generative AI for corpus approaches to discourse studies: A critical evaluation of ChatGPT. *Applied Corpus Linguistics, 4*(1), 100082. https://doi.org/10.1016/j.acorp.2023.100082

Dahl, T. (2004). Textual metadiscourse in research articles: A marker of national culture or of academic discipline? *Journal of Pragmatics, 36*(10), 1807–1825. https://doi.org/10.1016/j.pragma.2004.05.004

Deng, L., Bagheri, F., & Gao, X. (2021). Exploring the interactive and interactional metadiscourse in doctoral dissertation writing: A diachronic study. *Scientometrics, 126*(8), 7223–7250. https://doi.org/10.1007/s11192-021-04064-0

Eghtesadi, A. R., & Navidinia, H. (2009). A comparative study of metadiscourse use in research articles written by native and non-native speakers: Is audience taken into account? *The Journal of Asia TEFL, 6*(3), 157–176.

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. https://doi.org/10.1038/s41746-023-00784-7

Gholami, J., & Ilghami, R. (2016). Metadiscourse markers in biological research articles and journal impact factor: Non-native writers vs. native writers. *Biochemistry and Molecular Biology Education, 44*(4), 349–360. https://doi.org/10.1002/bmb.20961

Ho, V., & Li, C. (2018). The use of metadiscourse and persuasion: An analysis of first year university students' timed argumentative essays. *Journal of English for Academic Purposes, 33*, 53–68. https://doi.org/10.1016/j.jeap.2018.02.001

Hu, G., & Cao, F. (2015). Disciplinary and paradigmatic influences on interactional metadiscourse in research articles. *English for Specific Purposes, 39*, 12–25. https://doi.org/10.1016/j.esp.2015.03.002

Hyland, K. (1999). Talking to students: Metadiscourse in introductory coursebooks. *English for Specific Purposes, 18*(1), 3–26. https://doi.org/10.1016/S0889-4906(97)00025-2

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing.* London: Continuum.

Hyland, K., & Jiang, F. (2017). Is academic writing becoming more informal? *English for Specific Purposes, 45*, 40–51. https://doi.org/10.1016/j.esp.2016.09.001

Hyland, K., & Jiang, F. (2018). "In this paper we suggest": Changing patterns of disciplinary metadiscourse. *English for Specific Purposes, 51*, 18–30. https://doi.org/10.1016/j.esp.2018.02.001

Hyland, K., & Jiang, F. (2024). Metadiscourse: the evolution of an approach to texts. *Text & Talk, 44*(3), 411-433. https://doi.org/10.1515/text-2021-0156.

Jiang, F. (K.), & Hyland, K. (2025a). Metadiscursive nouns in academic argument: ChatGPT vs student practices. *Journal of English for Academic Purposes, 75*, 101514. https://doi.org/10.1016/j.jeap.2025.101514

Jiang, F. (K.) & Hyland, K. (2025b). Rhetorical distinctions: Comparing metadiscourse in essays by ChatGPT and students. *English for Specific Purposes, 79*, 17-29. https://doi.org/10.1016/j.esp.2025.03.001

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kawase, T. (2015). Metadiscourse in the introductions of PhD theses and research articles. *Journal of English for Academic Purposes, 20*, 114–124. https://doi.org/10.1016/j.jeap.2015.08.006

Khedri, M., Chan, S. H., & Tan, B. H. (2013). Cross-disciplinary and cross-linguistic perspectives on metadiscourse in academic writing. *Southern African Linguistics and Applied Language Studies, 31*(1), 129–138. https://doi.org/10.2989/16073614.2013.793957

Korteling, J. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human versus artificial intelligence. *Frontiers in Artificial Intelligence, 4*, 622364. https://doi.org/10.3389/frai.2021.622364

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health, 2*(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

Lee, J. J., & Casal, J. E. (2014). Metadiscourse in results and discussion chapters: A cross-linguistic analysis of English and Spanish thesis writers in engineering. *System, 46*, 39–54. https://doi.org/10.1016/j.system.2014.07.009

Lee, J. J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less successful argumentative essays. *Journal of Second Language Writing, 33*, 21–34. https://doi.org/10.1016/j.jslw.2016.06.004

Li, Y. (2023). *A practical survey on zero-shot prompt design for in-context learning*. In R.

Mitkov & G. Angelova (Eds.), *Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing (RANLP 2023)* (pp. 641–647).

Liao, W., Liu, Z., Dai, H., Xu, S., Wu, Z., Zhang, Y., Huang, X., Zhu, D., Cai, H., Li, Q., Liu, T., & Li, X. (2023). Differentiating ChatGPT-generated and human-written medical texts: Quantitative study. *JMIR Medical education, 9*, e48904. https://doi.org/10.2196/48904

Lin, C. (2005). Metadiscourse in academic writing: Aninvestigation of graduate students' MA theses in Taiwan. *Taiwan Journal of TESOL, 2*(1), 1-66

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Ma, Y., Liu, J., Yi, F., Cheng, Q., Huang, Y., Lu, W., & Liu, X. (2023). AI vs. Human -- Differentiation Analysis of Scientific Content Generation. https://doi.org/10.48550/arXiv.2301.10416

Mindner, L., Schlippe, T., Schaaff, K. (2023). Classification of Human- and AI-Generated Texts: Investigating Features for ChatGPT. In: Schlippe, T., Cheng, E.C.K., Wang, T. (eds) *Artificial Intelligence in Education Technologies: New Development and Innovative Practices* (pp.152-170). *AIET 2023. Lecture Notes on Data Engineering and Communications Technologies,* vol 190. Springer, Singapore. https://doi.org/10.1007/978-981-99-7947-9_12

Mo, Z., & Crosthwaite, P. (2025). Exploring the affordances of generative AI large language models for stance and engagement in academic writing. *Journal of English for Academic Purposes, 75*, 101499. https://doi.org/10.1016/j.jeap.2025.101499

Moorley, C., & Watson, R. (2023). Writing with ChatGPT: An illustration of its capacity, limitations, and implications for academic writers. *Perspectives on Medical Education, 12*(1), 261–270. https://doi.org/10.5334/pme.1072

Mu, C., Zhang, L. J., Ehrich, J., & Hong, H. (2015). The use of metadiscourse for knowledge construction in Chinese and English research articles. *Journal of English for Academic Purposes, 20*, 135–148. https://doi.org/10.1016/j.jeap.2015.09.003

Qiu, X., Wang, Y. (A.), Dartey, E. A., & Kim, M. (2024). Interactional metadiscourse in expert and student disciplinary writing: Exploring intrageneric and functional variation. *English for Specific Purposes, 73*, 124–140. https://doi.org/10.1016/j.esp.2023.10.007

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for Education and Research: Opportunities, Threats, and Strategies. *Applied Sciences*, *13*(9), 5783. https://doi.org/10.3390/app13095783

Saz-Rubio, M. M. del. (2011). A pragmatic approach to the macro-structure and metadiscoursal features of research article introductions in the field of Agricultural Sciences. *English for Specific Purposes, 30*(4), 258–271. https://doi.org/10.1016/j.esp.2011.03.002

Shen, Y., Heacock, L., Elias, J., Hentel, K. D., Reig, B., Shih, G., & Moy, L. (2023). ChatGPT and other large language models are double-edged swords. *Radiology, 26,* 230163. https://doi.org/10.1148/radiol.230163

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, 100752. https://doi.org/10.1016/j.asw.2023.100752

Swales, J. M., & Feak, C. B. 2012. *Academic Writing for Graduate Students: Essential Tasks and Skills (3rd ed.).* University of Michigan Press.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), 15

Tudino, G. & Qin, Y. (2024). A corpus-driven comparative analysis of AI in academic discourse: Investigating ChatGPT-generated academic texts in social sciences. *Lingua, 312*, 103838. https://doi.org/10.1016/j.lingua.2024.103838

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing, 62,* 101071. https://doi.org/10.1016/j.jslw.2023.101071.

Zhang, M., & Zhang, J. (2025a). Reflexivity in human-written and ChatGPT-generated English research article abstracts: A comparison of metadiscourse. *Applied Linguistics,* amaf032. Advance online publication. https://doi.org/10.1093/applin/amaf032

Zhang, M., & Zhang, J. (2025b). Disciplinary variation of metadiscourse: A comparison of human-written and ChatGPT-generated English research article abstracts. *Journal of English for Academic Purposes, 76,* 101540. https://doi.org/10.1016/j.jeap.2025.101540

Zhao, N. (2023). A corpus-based comparison study of first-person pronoun *we* in English-language abstracts. *Journal of English for Academic Purposes, 63,* 101244. https://doi.org/10.1016/j.jeap.2023.101244.