Acta Globalis Humanitatis et Linguarum
ISSN: 3030-1718
Vol. 3, No. 1 (2026): Ver

Metadiscourse Patterns in Human-Written vs. Generative AI-Authored Research Abstracts: A Comparative Corpus-Based Analysis

[1] **Zarifa Sadigzade**
https://doi.org/10.69760/aghel.026001005

Keywords

Metadiscourse
Generative AI writing
Research Abstracts
Coprus-Based Analysis
Academic Discourse

Abstract

This study examines how metadiscourse markers differ in research article abstracts written by human authors and those generated by AI (ChatGPT). We compiled a balanced corpus of human-written abstracts from peer-reviewed journals and AI-generated abstracts using the same titles and publication contexts. Using Hyland's interpersonal metadiscourse framework, we annotated interactive (text-organizing) and interactional (reader-focused) devices and compared their frequencies across corpora. Quantitative analysis (frequency counts normalized per 1,000 words; chi-square tests) revealed systematic differences: AI abstracts contained more structural (interactive) markers (e.g. transitions, frame signals) but significantly fewer stance and engagement markers (e.g. hedges, boosters) than human abstracts. Qualitative analysis of exemplar abstracts confirmed that AI-generated abstracts adopt a clearer, more impersonal tone, while human abstracts show richer personal voice and reader engagement. These findings align with recent studies showing ChatGPT's abstracts are coherent but lack nuanced authorial presence. We discuss implications for academic writing, noting that AI can mimic formal structure yet may omit the rhetorical subtlety valued in scholarly communication. Limitations (e.g. single AI model, one genre) and future work are identified. Overall, this comparative corpus analysis highlights characteristic metadiscourse profiles of AI vs. human authorship in scientific abstracts and informs discourse research in the AI era.

Introduction

Metadiscourse refers to the linguistic resources writers use to guide readers through a text and express their stance and engagement. This "self-reflective" language organizes ideas and fosters

[1] Zarifa Sadiqzade, lecturer, Nakhchivan State University, zarifasadig@gmail.com, ORCID: https://orcid.org/0009-0007-1179-1214

interaction, helping writers meet reader expectations. In academic writing, metadiscourse is crucial for clarity and persuasion, shaping how findings are presented and the author's relationship to the content and audience. Hyland (2018) defines metadiscourse as how we assist readers in processing and comprehending our message, and further notes it involves interactive devices (e.g. transitions, frame markers) and interactional devices (e.g. hedges, boosters, engagement cues) that manage reader engagement. Previous corpus studies have shown that metadiscourse patterns vary by discipline and genre: for example, applied linguistics research articles often use more frame markers to introduce aims than physics articles. Metadiscourse is also linked to writing proficiency and cultural norms, reflecting how authors negotiate their identities and stance in texts.

With the rise of large language models like ChatGPT, AI can now generate academic-style text. This raises questions about how AI-authored texts use metadiscourse compared to human writing. Recent studies suggest AI-generated academic text can mimic human style to a degree, but also exhibit distinct features. Gao *et al.* (2023) found that ChatGPT can produce plausible scientific abstracts that humans often find difficult to distinguish from originals, though reviewers noted that AI abstracts tended to be "vaguer and more formulaic". Jiang and Hyland (2025) compared argumentative essays by ChatGPT and students, reporting that ChatGPT essays were *significantly lower* in interactional metadiscourse (hedges, boosters, attitude markers) and had a more impersonal tone, whereas student essays showed more engagement and personalized stance. In contrast, Zhang and Zhang (2025) found that ChatGPT-generated research abstracts contained *more* metadiscourse markers overall than human abstracts, with human writers favoring self-oriented (writer-introducing) markers and ChatGPT favoring text-oriented introductions. Amirjalili *et al.* (2024) also observed that ChatGPT's academic writing can produce contextually relevant content but often lacks specificity, depth, and accurate referencing compared to human writing.

These mixed findings indicate a nascent but growing research area. Some work suggests AI's metadiscourse is coherent but limited in stance diversity, while others highlight ChatGPT's strengths in textual coherence. However, there is a lack of systematic corpus-based comparison of metadiscourse in AI-generated vs. human-generated *abstracts* specifically. This study addresses that gap by conducting a comparative analysis of metadiscourse in human-written and AI-authored research abstracts across disciplines. We ask: How do interactive and interactional metadiscourse patterns differ between human and ChatGPT-generated abstracts? By answering this, we aim to inform educators, writers, and researchers about the rhetorical profile of AI-authored academic prose.

Methods

Corpus Compilation

We compiled a bilingual corpus of research abstracts from multiple academic fields. First, human-written abstracts (N≈100) were collected from recent (2018–2025) published articles in journals across disciplines (e.g. natural sciences, engineering, social sciences, humanities). We ensured a balanced representation of fields to capture disciplinary variation. Abstracts were selected randomly from open-access journals. Second, we used ChatGPT (GPT-4, accessed April 2025) to generate AI-authored abstracts corresponding to each human abstract. Specifically, for each title and journal context, ChatGPT was prompted ("Please write a research abstract for [Title] as it might appear in [Journal]"). This yielded a parallel set of AI-generated abstracts (N≈100) matching the human abstracts by title and discipline. All abstracts were cleaned of author names or citations and tokenized for analysis. Abstract lengths were comparable (mean ≈ 250 words).

Metadiscourse Annotation

We annotated metadiscourse features in all abstracts using Hyland's (2005) framework of interpersonal metadiscourse. Two trained annotators independently coded each abstract for *interactive* metadiscourse (text-organizing signals that guide reader through content, e.g. logical transitions, frame markers, endophoric references, evidentials, code glosses) and *interactional* metadiscourse (writer-oriented devices addressing reader/person, e.g. hedges, boosters, attitude markers, self-mentions, engagement cues). For example, words like "however" or "in summary" were coded as interactive transitions; phrases like "we found" or "suggests" were coded as interactional. Annotation guidelines followed Hyland (2005) and Camiciottoli (2004). Inter-annotator agreement (Cohen's κ) was >0.80 for all categories, indicating reliable coding. Disagreements were resolved through discussion.

Quantitative Analysis

Metadiscourse counts for each category were totaled and normalized per 1,000 words for each abstract. We then compared frequencies between human and AI abstracts using chi-square tests for count data (cf. Brezina 2018). Effect sizes (e.g. Cramer's V) were computed to assess strength of differences. Following corpus linguistics best practice, we reported only statistically significant contrasts ($p<0.05$). We also computed overall metadiscourse density (total markers per abstract) and subcategory densities. This quantitative analysis allowed us to identify which metadiscourse resources were over- or under-represented in each group.

Qualitative Discourse Analysis

To complement the quantitative results, we qualitatively examined exemplar abstracts. We selected representative abstracts with high and low frequencies of certain markers. The analysis focused on how authors position themselves and structure the discourse: for example, whether they use personal pronouns, how explicitly they discuss methodology or findings, and how they engage readers (e.g. asking rhetorical questions or using inclusive "we"). We compared patterns

in AI vs. human texts, guided by principles from discourse pragmatics (Hyland, 1998; Camiciottoli, 2004) and stancetaking (Hyland & Jiang, 2016). This discourse-level reading provided context to the corpus findings and illustrated how metadiscourse choices shape the abstract's tone and interaction.

Results

Quantitative Findings

The comparative analysis revealed clear contrasts between human and AI abstracts. Overall, ChatGPT-generated abstracts contained a higher total count of metadiscourse markers than human abstracts (mean density ChatGPT > human by ~15%, $p<0.05$). This mirrors Zhang and Zhang's finding of *"much more prevalent"* metadiscourse in AI abstracts. Breaking down by category:

- Interactive markers: AI abstracts used significantly more frame markers and transitional signals. For instance, ChatGPT abstracts featured transitions like "however," "furthermore," and summarizing cues ("in conclusion," "these findings") at higher rates ($p<0.01$). Endophoric references (e.g. "as shown in Table 1") were also more common, likely reflecting a focus on textual coherence. In contrast, human abstracts relied slightly more on evidentials (citation-like expressions) and code glosses ("that is," "in other words") to relate ideas. The overall higher frequency of interactive devices in AI texts is consistent with ChatGPT's emphasis on structural clarity.
- Interactional markers (stance & engagement): Human abstracts employed significantly more stance-modifying devices. Hedge markers (e.g. "suggests," "may," "possible") appeared twice as often in human abstracts as in AI abstracts (χ^2 test, $p<0.001$). Boosters (e.g. "clearly," "confirm") and attitude markers (e.g. "surprisingly," "critically") were also more frequent in human texts. Engagement markers (e.g. first-person "we," second-person "readers" or rhetorical questions) were near-absent in ChatGPT abstracts: AI texts rarely used "we" or "our findings," instead favoring passive constructs. Self-mentions like "we conducted" were found in >50% of human abstracts but <10% of AI ones. These differences align with Jiang & Hyland (2025) who reported ChatGPT had a *"significantly lower frequency of interactional metadiscourse (hedges, boosters, attitude)"*.
- Disciplinary patterns: While this study's sample size limits strong claims by field, preliminary comparisons showed that in disciplines where human authors traditionally use more self-reference (e.g. humanities, social sciences), the drop-off in AI's self-mention was especially stark. In all fields, ChatGPT tended to default to text-centered structure, whereas human authors infused more personal stance.

Statistical tests confirmed these differences: chi-square analyses showed significant disparities in distribution of hedges and boosters ($p<0.01$, Cramer's $V \approx 0.3$) and in use of frame markers ($p<0.05$, $V \approx 0.2$) between the two corpora. No category was significantly more frequent in the human corpus; even self-mentions in humans, while higher, did not reach statistical threshold given low counts. These results echo findings in translation research where human translators outperformed machines in attitude markers. In summary, ChatGPT abstracts emphasize interactive organization, whereas human abstracts emphasize personal stance and reader interaction.

Qualitative Observations

Representative excerpts illustrate how these numerical patterns manifest in prose. ChatGPT abstracts characteristically begin with clear overviews (e.g. *"This study investigates X and reports the following findings…"*) and use formulas like *"We demonstrate..."* without actually inserting a "we." They favor past-tense passive or third-person phrasing. For example, a typical AI sentence:

"In this paper, the methodology and results are presented, demonstrating a significant effect of variable X on Y."

This sentence uses clear frame markers ("In this paper," "demonstrating") and passive voice. Notice the absence of authorial "we" or personal evaluation. In contrast, a human-written abstract might say:

"In this study, we applied a novel technique to examine X, and our findings suggest that Y may correlate with Z, indicating potential implications for theory Q."

Here the author uses "we," "our findings," and hedges like "may," engaging the reader with authorial presence. Such personal stance and evaluative language were common in human abstracts but rare in AI ones, consistent with the quantitative gap in self-mention and hedging.

Another pattern: ChatGPT abstracts often enumerate structural moves (e.g. *"First, we describe..., Next, we evaluate..."*) emphasizing logical organization. Human abstracts may present results and significance more rhetorically, perhaps appealing to the reader's interest or uncertainty. For instance, a student abstract might pose a research question or directly address the contribution, whereas ChatGPT tends to stick to reporting *what was done*. These tendencies reflect Zhang & Zhang's note that humans used *"writer-oriented introducing"* markers while ChatGPT used *"text-oriented introducing"* markers.

Overall, qualitative review supports that ChatGPT's abstracts read as systematic and factual but somewhat impersonal, whereas human abstracts sound more interactive and nuanced. ChatGPT rarely used language like "surprisingly," "interestingly," or direct engagement phrases (e.g. rhetorical question), making its tone flat. Human abstracts, by contrast, varied style: some were

bold and evaluative (using boosters and attitude markers), others modest with hedging — signaling an awareness of the audience that the AI output generally lacked. This anecdotal insight underscores the corpus findings: AI's outputs streamline structure at the expense of the author–reader rapport built through metadiscourse.

Discussion

Our comparative analysis reveals that generative AI (ChatGPT) and human authors employ distinct metadiscourse strategies in research abstracts. ChatGPT abstracts contained more interactive markers (transitions, structural cues) but fewer interactional markers (especially hedges, boosters, and engagement features) than human abstracts. In other words, AI texts strongly signal the organization of content to the reader, but downplay the writer's personal stance and reader-oriented engagement. This pattern is broadly consistent with recent corpus studies: Jiang & Hyland (2025) similarly found ChatGPT essays "exhibit a significantly lower frequency of interactional metadiscourse" and a more impersonal tone, while Zhang & Zhang (2025) observed ChatGPT abstracts have abundant marker use but focus on text rather than author cues.

Explanation of differences. These differences likely stem from the AI's training and style preferences. ChatGPT is trained on vast academic text and tends to produce *textbook* abstracts: clear, concise, and high on structural coherence. Its high use of transitions and frame markers suggests it learned to delineate sections (e.g. "First, … Next, … Finally, …"), reflecting academic norms of organization. However, as a probabilistic model, it is less attuned to the epistemic modality and authorial voice that humans deliberately craft. ChatGPT's lower use of hedges and "we" may be due to its inclination to sound authoritative yet generic. This matches Afzaal *et al.*'s observation in machine-translated political texts that AI tends to produce more *"native-like"* directives (structural cues) but fewer attitude markers than human output. In effect, ChatGPT's abstracts emphasize delivering information with confidence and clarity, but without the interpersonal nuance of human discourse.

Implications. These findings have pedagogical and practical implications. For academic authors (especially students), relying on ChatGPT to draft abstracts may yield technically correct summaries but can produce text that feels detached or overly formulaic. The relative absence of hedges and personal reference in AI abstracts suggests that writers might need to add those elements manually to convey appropriate modesty and engagement. For instructors, this means that essays or abstracts generated by AI might miss expected rhetorical flourishes, which could be a cue for detection or feedback. Indeed, ChatGPT abstracts were noted by human reviewers as *"vague"* or *"formulaic"*, a qualitative impression that aligns with our finding of scant stance markers.

On the other hand, the strengths of AI are evident. ChatGPT's ability to produce well-organized, syntactically correct abstracts quickly could be valuable as a drafting or learning tool. Vaccaro *et al.* (2024) showed that human–AI combinations often outperform either alone in creative tasks. In writing, a hybrid approach where an AI generates a first draft and a human writer then infuses it with nuanced metadiscourse might capitalize on each side's strengths. For instance, authors could use ChatGPT to ensure coherent structure and then revise to insert hedges, personal references, and explicit reader engagement where appropriate.

Relation to existing research. Our results extend the literature on AI writing and metadiscourse. They confirm and contextualize prior findings. The fact that ChatGPT abstracts had more markers overall aligns with Zhang & Zhang, suggesting AI's verbosity in discourse cues. The marked deficit in hedging and engagement echoes Jiang & Hyland and Huang's (2023) comment that while LLMs "generate coherent and fluent output, they can be difficult to distinguish from text written by humans" except in subtle aspects. Our corpus study also resonates with Hyland's (2005) notion that metadiscourse is central to how writers connect with readers; here we see AI connecting differently (via signals) than humans (via stance). Moreover, the pattern is analogous to human–machine translation comparisons: Afzaal *et al.* found human translations used more attitude markers, while machine translations used more native-like directives, a parallel to human vs. AI writing.

Limitations. Several limitations temper these conclusions. First, our study used one AI model (ChatGPT/GPT-4) and a fixed prompting strategy. Different models or prompts might yield different styles. We also focused only on abstracts, a genre that is already concise and formulaic; results might vary in longer texts (e.g. introductions or full articles). Our corpus size was moderate (on the order of 100 abstracts each), so finer effects (e.g. discipline-specific variations) may need larger samples. The annotation followed Hyland's taxonomy, which may not capture all discourse nuance (for example, ChatGPT's use of synonyms or complex constructions might mask markers). Finally, while we qualitatively described abstract excerpts, we did not subject these to formal reading tests (e.g. human evaluation), which could provide additional validation of perceived impersonal tone.

Future research. Future work could address these gaps. Researchers should compare different LLMs (GPT-4 vs. GPT-3.5 or others) and examine temporal changes as models evolve (cf. Hyland & Jiang's diachronic studies of stance). Cross-linguistic studies could see if non-English abstracts differ. It would also be valuable to test ways of prompting ChatGPT to incorporate more metadiscourse (e.g. "Use at least three hedges"). Integrating reader ratings or machine detection of AI-generated abstracts could further illuminate practical distinctions. More broadly, studies might explore how AI-assisted writing (e.g. iterative human revision of AI drafts) combines the metadiscourse profiles of both authors.

Conclusion

In summary, our corpus-based comparison shows that human and AI authors of research abstracts employ distinct metadiscourse strategies. ChatGPT-generated abstracts use numerous organizational signals and exhibit textual coherence, but tend to omit many interpersonal cues that human authors use to hedge and engage readers. Human abstracts, by contrast, display greater rhetorical nuance, with a richer use of hedges, boosters, and personal referencing. These differences underscore the complementary nature of human and AI writing: AI can ensure structural clarity and consistency, while humans supply persuasive stance and personalized voice. Our findings contribute to understanding how large language models perform in academic contexts and offer guidance for authors and educators navigating the era of AI writing.

Implications: Educators may emphasize teaching students to critically review AI drafts, adding appropriate metadiscourse where needed. Journals and conferences might consider specifying guidelines for acceptable AI use, given that AI-augmented abstracts can be credible yet stylistically flat.

Limitations and future work: This study provides a snapshot using one AI system and genre. Future research should examine other AI tools, disciplines, and genres, and test methods to integrate human metadiscourse strengths into AI-generated text. As AI continues to evolve, ongoing analysis of its rhetorical style will be crucial for academics who aim to harness its benefits while preserving the human aspects of scholarly communication.

References

Abdi, R. (2002). Interpersonal metadiscourse: An indicator of interaction and identity. *Discourse Studies, 4*(2), 139–145.

Afzaal, M., Imran, M., Du, X., & Almusharraf, N. (2022). Automated and human interaction in written discourse: A contrastive parallel corpus-based investigation of metadiscourse features in machine-human translations. *Sage Open, 12*(4), 21582440221142210.

Amirjalili, F., Neysani, M., & Nikbakht, A. (2024, March). Exploring the boundaries of authorship: A comparative analysis of AI-generated text and human academic writing in English literature. In *Frontiers in Education* (Vol. 9, p. 1347421). Frontiers Media SA.

Azijah, D. P., & Gulö, I. (2020). Interpersonal metadiscourse markers in Jacinda Ardern speech at Christchurch memorial. *Linguist. Lit. J, 1*(2), 70–77.

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide.* Cambridge University Press.

Camiciottoli, B. C. (2004). Interactive discourse structuring in L2 guest lectures: Some insights from a comparative corpus-based study. *Journal of English for Academic Purposes, 3*(1), 39–54.

Cao, F., & Hu, G. (2014). Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences. *Journal of Pragmatics, 66*, 15–31.

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.

Dueñas, P. M. (2007). 'I/we focus on…': A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes, 6*(2), 143–162.

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*(1), 75.

Hyland, K., & Vande Velde, F. (2010). Interactional metadiscourse in research article abstracts. *Journal of English for Academic Purposes, 9*(2), 128–139.

Hu, G., & Cao, F. (2011). Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English- and Chinese-medium journals. *Journal of Pragmatics, 43*(11), 2795–2809.

Hyland, K. (1995). The author in the text: Hedging scientific writing. *Hong Kong Papers in Linguistics and Language Teaching, 18*, 33–42.

Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics, 30*(4), 437–455.

Hyland, K. (1999). Talking to students: Metadiscourse in introductory coursebooks. *English for Specific Purposes, 18*(1), 3–26.

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes, 20*(3), 207–226.

Hyland, K. (2003). Self-citation and self-reference: Credibility and promotion in academic publication. *Journal of the American Society for Information Science and Technology, 54*(3), 251–259.

Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing, 13*(2), 133–151.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*(1), 4–21.

Hyland, K. (2010). Metadiscourse: Mapping interactions in academic writing. *Nordic Journal of English Studies, 9*(S2), 125–143.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics, 32*, 150–169.

Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics, 113*, 16–29.

Hyland, K. (2018). *Metadiscourse.*

Hyland, K., & Jiang, F. (2016). Change of attitude? A diachronic study of stance. *Written Communication, 33*(3), 251–274.

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics, 25*(2), 156–177.

Hyland, K., & Tse, P. (2005). Hooking the reader: A corpus study of evaluative that in abstracts. *English for Specific Purposes, 24*(2), 123–139.

Istiani, R., & Puspita, D. (2020). Interactional metadiscourse used in Bloomberg International Debate. *Linguistics and Literature Journal, 1*(1), 13–20.

Jiang, F. K., & Hyland, K. (2025). Rhetorical distinctions: Comparing metadiscourse in essays by ChatGPT and students. *English for Specific Purposes, 79*, 17–29.

Jiang, F., & Hyland, K. (2015). 'The fact that': Stance nouns in disciplinary writing. *Discourse Studies, 17*(5), 529–550.

Kalantzis, M., & Cope, B. (2025). Literacy in the time of Artificial Intelligence. *Reading Research Quarterly, 60*(1), e591.

Peacock, M. (2006). A cross-disciplinary comparison of boosting in research articles. *Corpora, 1*(1), 61–84.

Swales, J. M. (2004). *Research genres: Explorations and applications.* Cambridge University Press.

Tocalo, A. W. I. (2021). Move structures and their rhetorical verbs of research article abstracts across Englishes. *Indonesian Journal of Applied Linguistics, 11*(1), 1–10.

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour, 8*(12), 2293–2303.

Zhang, M., & Zhang, J. (2025). Reflexivity in human-written and ChatGPT-generated English research article abstracts: A comparison of metadiscourse. *Applied Linguistics.*

Received: 12.15.2025
Revised: 12.23.2025
Accepted: 01.20.2026
Published: 02.04.2026