Macrothink Institute™

International Journal of Linguistics
ISSN 1948-5425
2026, Vol. 18, No. 3

Can AI Write Like a Scholar? A Comparative Analysis of Human vs. ChatGPT-Generated Research Article Introductions and Literature Reviews

Donatella Malavasi
Dept. of Studies on Language and Culture, University of Modena and Reggio Emilia
Largo S. Eufemia 19, 41121 Modena, Italy
E-mail: donatella.malavasi@unimore.it

Received: March 29, 2026 Accepted: April 30, 2026 Published: June 28, 2026
doi:10.5296/ijl.v18i3.23861 URL: https://doi.org/10.5296/ijl.v18i3.23861

Abstract

The rapid progression of technology and the proliferation of Artificial Intelligence (AI) have increasingly influenced many fields, ranging from information technology through to tourism, education, management, and medicine. Specifically, in the current academic context, where the 'publish or perish' imperative forces constant research and publication, AI technologies can assist academics in writing and improve the efficiency and quality of texts (Dwivedi et al., 2023; Nguyen et al., 2024). However, despite the spectrum of capabilities offered by these tools, growing concerns have been raised about the erosion of critical thinking skills and the production of substandard content. While the impact of the use of AI has received scholarly attention in student academic writing, the effects of its utilization by academics in research articles remain largely unexplored. Therefore, the aim of this study is to investigate how GenAI tools, particularly ChatGPT, could impact expert writers' research papers in their introduction and literature review sections. The present research compares authentic human-authored literature reviews and their counterparts generated by ChatGPT in response to specific prompts. The focus is on the linguistic, discursive, and rhetorical patterns that (dis)similarly characterize a set of human- vs. AI-generated introductions and literature reviews in the disciplinary field of linguistics. A combination of well-consolidated quantitative and qualitative methodological approaches, including corpus linguistics and discourse analysis, is adopted to assess whether and how ChatGPT chatbots emulate human language use in research article introductions and literature reviews.

Keywords: Academic writing, Human-authored texts, AI-generated texts, ChatGPT, Research articles, Introductions, Literature reviews

1. Introduction

The rapid evolution of technology and the advent of Artificial Intelligence (AI) have revolutionized many sectors of life, including medicine, management, tourism, and academia. Among its several applications, it is difficult to ignore the considerable impact that AI and Large Language Models (LLMs) have had on writing. In particular, Generative AI (GenAI), which is based on LLMs like GPT-3, GPT-4 and GPT-5, "can assist writers throughout the entire writing process, from idea generation and outlining to drafting and revising manuscripts (Enriquez et al., 2023)" (Nguyen et al., 2024, p. 850). Thanks to these models, "the sophistication and human-like quality of texts [...] have notably increased", making the task of distinguishing LLM-authored texts rather complicated (Ji et al., 2025).

Remarkably, in education, these technological innovations have transformed how academic writing is produced and research carried out (Khalifa and Albadawy, 2024; Nguyen et al., 2024; Chukwuere, 2025). Furthermore, as a consequence of the pressure to publish in academia, the so-called 'publish or perish' norm, GenAI systems have increasingly served as an invaluable aid in facilitating drafting text, improving content and structure, refining language, as well as generating content (Dwivedi et al., 2023; Nguyen et al., 2024). However, despite its potential to enhance academic writing, several negative consequences have been associated with the use of AI (Ariyaratne et al., 2023; Dergaa et al., 2023; Hsu, 2023; Costa et al., 2024; Cohen and Moher, 2025). These include growing concerns about research authenticity, the erosion of critical thinking skills, the production of substandard content, and "AI's ability to "hallucinate", provide incorrect or fabricated information" (van Niekerk et al., 2025). The impact and negative consequences of AI technologies on writing have been mainly examined in student writing (Marzuki et al., 2023; Nguyen et al., 2024; van Niekerk et al., 2025), while more limited consideration has been devoted to the effects that AI use may have on scholars' production (Jenko et al., 2024; Ngwenyama and Rowe, 2024; Schryen et al., 2024; Mozelius and Humble, 2024).

Thus, by concentrating on the prototypical genre of expert academic writing, namely the Research Article (RA), the present study aims to compare authentic human-authored introduction and literature review sections and their counterparts generated by a specific AI platform, ChatGPT, in response to particular prompts. ChatGPT, which is the most frequently used GenAI technology (Durak et al., 2025), "is a conversational interface built by OpenAI on top of Generative Pre-trained Transformer (GPT) models (e.g., GPT-3.5, GPT-4, and GPT-5), which produce contextually appropriate responses to user prompts" (Georgiou, 2025). The focus of the study on RA introductions and literature reviews, which are commonly combined into one section, is motivated by AI's undisputed role in facilitating "the extraction and analysis of information from existing literature, synthesising these findings into coherent overviews" (Khalifa and Albadawy, 2024). The introductory sections of RAs are indeed written to provide an overview of the relevant literature, while introducing readers to both the topic and the purpose of the research (Galvan and Galvan, 2017). The comparative analysis, which is intended to be exploratory, sets out to add to the limited research on the language and discourse choices that (dis)similarly characterize AI- and human-authored texts (Berber Sardinha, 2023; Amirjalili et al., 2024; Durak et al., 2025; Georgiou, 2025). Specifically, this paper, which does not aim to assess the content quality of texts, addresses the current research gap guided by the following research questions:

RQ1: What are the linguistic and discursive tools that are characteristic of human- and ChatGPT-generated texts?

RQ2: Do ChatGPT- and human-written introduction and literature review sections meet similar standards of argumentative reasoning?

RQ3: To what extent can ChatGPT generate content that linguistically and discursively mimics human writing?

2. Materials and Methodology

For the exploration of the linguistic, discursive, and rhetorical patterns that characterize human- vs. ChatGPT-generated texts, two small yet comparable corpora were compiled specifically for this research. The first database includes authentic human-authored introduction and/or literature review sections drawn from 30 research articles in the disciplinary field of linguistics. The RAs chosen for this study were published in 2010 and, in a few cases, in 2011, namely well before ChatGPT's release in November 2022. The papers were retrieved from three highly reputable, peer-reviewed international journals: *Discourse Studies, English for Academic Purposes,* and *English for Specific Purposes*. For each journal, ten research articles containing keywords and the "Introduction" and/or "Literature Review" sections were selected. The AI corpus was created using ChatGPT Model 5.2. (February 2026 version). The same prompt was run multiple times to generate the outputs, i.e., "Write a *-word introduction/literature review section for a research article on **. Include citations". To maintain consistency with the human dataset, the target length or number of words (*) was based on the average length of the sample of texts taken from each journal, and the keywords appearing in each research article (**) were included in the prompt. No examples were provided to ensure independent content generation, and the materials created by ChatGPT were collected without any human post-editing, correction, or revision. The design of the two corpora compiled for this study appears in Table 1.

Table 1. Corpus structure and size

Journals	Human-authored Introductions/Literature Reviews: no. of running words	ChatGPT-generated Introductions/Literature Reviews: no. of running words
Discourse Studies (DS)	15,740	13,948
English for Academic Purposes (EAP)	13,489	15,250
English for Specific Purposes (ESP)	17,746	14,137
Total no. of running words	46,975	43,335

The final corpora are intentionally limited in size as this is conceived as a preliminary investigation designed to focus on close reading of the texts collected and test the feasibility of the approach adopted. As for the research methods, in order to assess whether and how ChatGPT chatbots emulate human language use in introductions and literature reviews, quantitative and qualitative approaches were adopted, including corpus linguistics and discourse analysis. The combination of these two methodologies allowed, on the one hand, to extract quantitative information about statistically significant linguistic features (Sinclair,

2004), and on the other hand, to examine them qualitatively in their context of use (Partington, 2010; Baker et al., 2008; Taylor and Marchi, 2018). In more detail, with the support of corpus linguistics tools and the software suite AntConc 3.5.8 (Anthony, 2019), a selection of both content and grammar words among the 40 top keywords of each corpus vs. the other was first identified and then examined in their concordance lines, linguistic context, and phraseology. The analysis considered keywords as they are word forms with frequencies that are higher or lower in one corpus vs. the other, and it specifically focused on those elements that were assumed to be pivotal in creating well-structured and argued accounts of the literature. The present study, which is meant to be a pilot study, is not intended to be exhaustive, but rather to serve as a starting point for further inquiry. A more extensive set of materials and the use of other LLM-based technologies would certainly strengthen the generalizability of the study findings.

3. Results

3.1 Human-Authored Texts: An Analysis of Keywords

In order to gain insights into the linguistic and discursive similarities and differences between human- vs. ChatGPT-created introductions and literature reviews, the study concentrates first on the keywords in the corpus of texts written by expert scholars. The lexical and grammar words considered in this phase of the study are listed in Table 2 with their rank, raw frequency, and keyness.

Table 2. Human- vs. ChatGPT-generated RA selections: keywords

Rank	Frequency	Keyness	Keyword
2	1,981	+200.91	of
3	99	+129.52	p
6	123	+97.55	we
7	134	+85.74	other
8	72	+85.08	see
9	690	+84.95	is
13	50	+65.39	paper
16	127	+58.88	e
19	107	+54.42	one
23	70	+43.34	some
24	108	+42.4	g
30	43	+38.45	will
34	49	+34.98	our
35	209	+34.37	which
37	92	+34.04	different

A particularly interesting feature that emerges from the keyword list above is the preponderance of function items compared to lexical words. Specifically, worthy of note is the high keyness of the preposition *of*. An inspection of its concordance lines reveals that *of* is recurrently used both to introduce nominal post-modifiers that describe previous or current studies (e.g. *the use of, an analysis of,* and *the study of* in 1 and 2, my italics and bold as in the other quotations), and to clarify a research topic or establish its importance (e.g. *in terms of* and *the importance of* in 3 and 4).

(1) *Myers* (1994, pp. 92–96) *investigated* ***the use of*** *foreign languages* (German, French, Russian, Dutch, Maori and Japanese) in advertisements for native speakers of English, as well as ***the use of*** *English in advertisements* for native Dutch, German and Spanish speakers, while *Kelly-Holmes* (2005) *provided* ***an*** *extensive* ***analysis of*** *foreign language use* in advertising in Europe. On the whole, what is evident is that, although ***the study of*** *foreign languages* in advertising is not limited to English, its use in advertising has been studied relatively more than the use of other languages, perhaps reflecting that English has been assumed to be the most frequently used foreign language in advertising messages aimed at non-English speaking target groups (see Bhatia, 1992, Martin, 2002, Piller, 2003). (ESP)

(2) *This paper focuses on* ***the use of*** *wikis* in the course of Effective Communication in English at Stockholm University; it aims to describe how the course wiki was used to teach writing for academic and professional purposes, and to analyse what impact using the wiki had on the writer–reader relationship. (ESP)

(3) As Miller (1984) in her pioneering and often cited article on 'Genre as social action' demonstrated, genres emerge from the requirements of recurrent rhetorical situations, asking for an adequate response. The RA abstract is such a response, and, by nature, a social one. From this viewpoint the generic character of the RA abstract is no longer *defined* ***in terms of*** *specific lexis and syntax, but rather as a case of interaction between individuals*, acting in a social, institutional context. (EAP)

(4) In analysing data for network histories, we draw on several key SNA notions, including strength/weakness and durability/decay. We continue to emphasise ***the importance of*** *the local and transnational nature of multilingual scholars' networks* (Lillis and Curry, 2006a, Lillis and Curry, 2010, Chapter 3) and introduce, in the context of academic productivity, the characteristics of network formality and informality. (ESP)

As such, the phraseology of *of* signals considerable precision in the discussion of the specific study described in the RA and similar research previously conducted (1 and 2), while also elaborating and claiming the centrality of the research area or topic(s) under consideration (3 and 4). In addition to the preposition *of*, other keywords were found to be repeatedly deployed to synthesize the relevant literature within a defined field of research, and link or position the sources against each other. Evidence from the corpus suggests that forms such as *p, other, different, see, e,* and *g* were frequently resorted to by scholars to cite and link the sources of background information, and as such lay the foundations for their papers. While *p*, which abbreviates page, is used for quotations that reproduce words verbatim from another work, *see, other, different,* and *e.g.* introduce citations or draw connections between previous studies, and thus contribute to making the paper well-supported (5-8).

(5) Rundbald suggests that impersonal forms of authorial reference, such as passive verbs and metonymic expressions, enable writers to *"signal credibility, reliability, objectivity, and ultimately authority to their readers and the research community"* (2007, ***p.*** 251). (EAP)

(6) The passive voice was one of the first grammatical resources to be investigated in early studies of scientific and technical English (***see*** *Barber, 1962; White, 1974*). The aim was to provide evidence that specialised registers are characterised by distinctive lexico-grammatical features as compared to general English. (EAP)

(7) A few early studies present *a* ***different*** *perspective, arguing that academic writing is actually more nominal than verbal*. For example, Wells (1960) discusses the 'nominal' style of written discourse (contrasted with the 'verbal' style of speech), and Halliday and Martin (1993/1996) focus on the importance of nouns and nominalization in academic written discourse (cf. Fang, Schleppegrell, & Cox, 2006). ***Other*** *studies have gone further, showing that* clausal subordination is typical of the 'verbal' style of speech, while academic writing relies on phrasal modifiers instead of dependent clauses. (EAP)

(8) According to Schegloff (1992: x), Sacks' interest in categorizational topics waned post-1968 in favour of an emphasis on sequential organization. Nevertheless, Sacks' work on these topics inspired the development of 'MCD analysis' during the 1970s and 1980s (***e.g.*** *Atkinson and Drew, 1979; Drew, 1978; Jayyusi, 1984; Payne, 1976; Sharrock, 1974; Watson, 1978, 1983*) and, as it later became known, 'membership categorization analysis' (***e.g.*** *Eglin and Hester, 1992, 2003; Hester and Eglin, 1992, 1997; Watson, 1997*). (DS)

Apart from serving the function of weaving together previous studies and summarizing the relevant literature within a particular research area, the abbreviations *i.e.* and *e.g.*, the relative pronoun *which*, as well as the phrases *in other words* and *on the other hand* manifest the RA authors' willingness to clarify, exemplify, and explain a topic, as well as to establish contrastive relations between the elements of the ongoing discourse (9-13).

(9) A second problem lies in the focus of frameworks, such as metadiscourse, on the overt *"manifestation" of interpersonality,* ***i.e.*** *markers appearing on the surface of language* (Hyland, 2005a, p. 28). As a consequence, not all the grammatical structures which are part of the inventory of resources used to convey objectivity can be accounted for by applying current analytical schemes. (EAP)

(10) Three main orientations can be identified in the literature:

'Phraseological' approaches (***e.g.,*** *Cowie, 1998*) focus on the non-compositionality of certain expressions, defining formulaicity in terms of either the degree to which the meaning of a word combination is predictable from the meaning of its parts or the degree to which words with similar meanings can be substituted into the phrase. Non-compositional phrases include idioms (***e.g.,*** *kick the bucket, spill the beans*) and certain collocations (***e.g.,*** *curry favour, French window*). The 'formal idioms' (Fillmore, Kay, & O'Connor, 1988) of construction grammar (***e.g.*** *what's NP doing Y; the ADJ-er the ADJ-er*) can also be included in this category as items which cannot be easily understood and/or produced without specific learning. (ESP)

(11) It is important to emphasize here that the notion of rhetorical organization underlying this study comes from a rather different tradition than, for example, that of *the Rhetorical Structure Theory* (RST) (Grimes, 1975; Mann and Thompson, 1986, 1988), *in* ***which*** *a descriptive framework is proposed* to describe the relations among clauses in a text such as cohesion and coherence relations. (DS)

(12) This was evidently so since any explicit marker of confusion or request for clarification from the participants was lacking. ***In other words***, Markman (2006) found that the participants seemed to leave behind an understanding of turn-taking as being based on

temporality. Markman (2006) also studied how participants allocated turns in chat. (DS)

(13) Utilizing an RT approach for analyzing English discourse markers such as so, however, and well, Blakemore (2002) argues that they are linguistic expressions which contribute to the inferential processes involved in utterance understanding, and that they reflect the relationship between linguistic meaning and context in discourse. ***On the other hand,*** from an interactive perspective, conversation by definition involves two or more people. (DS)

As illustrated by the examples above, the RA sections collected for this study are characterized by an extensive use of constructions that explicitly signal relations between previous studies or research findings and elaborate on a topic or an approach. This shows the scholars' attempts to overview the literature in a systematic and argumentative way, and guide the reader through the unfolding literature review. While stating current knowledge in a field, the RA authors also enter the text to announce their study, and state how it extends, builds upon, and departs from previous research. Personal forms of self-reference, such as *we* and *our*, are accompanied by more impersonal ones, including *paper* and *it*. The first person pronoun *we* and phrases such as *our study* and *this paper* were found to collocate with a variety of research, cognitive, and discourse verbs, also in their future forms (*will*), to indicate the focus, aim, and argument of the paper. This is exemplified by clusters or n-grams such as *we examine, we believe, we focus on, we propose, we will answer, we will scrutinize, our study investigates,* and *this paper focuses on*, to name but a few (see 14-17).

(14) While recognizing that other factors play a role in securing English-medium publishing, as noted above, *in this article* ***we focus on*** how scholars gain access to and participate in different types of networks, and what publications—in particular, English-medium—result from such participation. In so doing, a key aim is to challenge the emphasis within much ESP/EAP research and pedagogy that foregrounds individual competence (Lillis and Curry, 2006a, Lillis and Curry, 2010) and to identify salient characteristics of network participation that support or lead to English-medium publication. (ESP)

(15) First, through an analysis of the corpus as a whole, ***we examine*** citations in order to see to what extent the geolinguistic context of publication affects citation practices, and thus, the disciplinary communit(y)ies being invoked. *Next,* ***we examine*** the ways in which different types of citation are used in relation to work published in English, local national languages, and other languages in two subsets of the corpus containing work published by Portuguese and Spanish context scholars. (EAP)

(16) ***Our study investigates*** differences in the use of this/these as determiners and pronouns because prescriptivism claims that problems of clarity can be solved by simply using demonstratives as determiners rather than pronouns. Biber, Johansson, Leech, Conrad, and Finegan (1999) point out that using demonstratives as determiners allows for more noun phrase modification, which can either efficiently give readers more information or further clarify what the coreferent is. (ESP)

(17) ***This paper*** *is concerned with* the analysis of on-site problem-oriented discourse in the international construction industry, and its realization through critical lexical items. Specifically, ***we will answer*** *the questions:* How do the English interactions recorded in a Japanese company in Hong Kong (hereafter referred to as 'the HK data') compare at the lexicogrammatical level to everyday English and business English? How can such lexicogrammatical items be interpreted to shed light on the context they reflexively constitute? (ESP)

The authors' presence in the text is also marked by their evaluation of what is discussed at centre stage (*is*), and by the forms *one* and *some* that respectively aim to single out specific studies or topics under consideration, and introduce tentativeness or caution. In more detail, in addition to its use for topic elaboration and interpretation (18), *is* appears in such affective statements as *it is important, it is difficult* (19), and *it is essential* that foreground the writer's viewpoint and direct readers to particular interpretations.

(18) *Discourse markers – that* ***is****, words or phrases* which are normally used to mark boundaries in conversation between one topic or subtopics and the next1 – can be frequently found in spoken discourse and play an important role in controlling the pragmatic nature of an interaction (Schiffrin, 1987). (DS)

(19) The main focus was perceptions of equity and access in English language academic publishing. However, while questionnaire surveys are useful for collecting large quantities of data relatively economically, they also have well-known limitations. ***It is difficult****,* for example, to probe into the reasons prompting particular responses. For this reason, a second more ethnographically-oriented phase of research was undertaken focussing on both a more detailed elicitation of attitudes to English-language publication and senior Spanish academics' writing/publishing practices in English. (ESP)

Finally, while *one* was selected by the RA authors to highlight a particular approach, study, or element under discussion (20), *some* was used to tentatively refer to previous research or assess components of discourse (21-22).

(20) Indeed, Gerritsen et al. (2000, p. 28) have also suggested that the comprehensibility of advertising messages may be relatively unimportant to marketers, as advertising aims not only at conveying a message, but also at creating brand awareness and positive attitudes to the brand among the target group. However, *some scholars maintain* that the comprehensibility of the advertising message on the part of its target group is key to achieving the desired communicative effect (e.g., Pieters & van Raaij, 1992). ***One*** *study* whose findings seem to confirm this latter assumption *to some exten*t is Gerritsen et al.'s (2000) investigation, which provides empirical evidence that people's comprehension of English in advertising messages is positively correlated with their appreciation of the English used. (ESP)

(21) Most network studies have taken a synchronous view. However, ***some*** *research has explored* networks diachronically, with time spans ranging from six months to 10 years (Suitor, Wellman, & Morgan, 1997). (ESP)

(22) It is also argued that ***some*** *of their uses seem to perform a function similar to that of emphatics* if they promote a claim to the status of the obvious once another claim is accepted (Ex. 2), or if they add assertiveness to a claim by suggesting a stronger term than the one used in the apodosis (Ex. 3). (EAP)

The results discussed in this section suggest that the human-authored texts considered in the study exhibit a wide range of citations, formulae that connect previous works, constructions for topic elaboration and clarification, as well as expressions of evaluation. The analysis of these language and discourse tools in context confirms their valuable contribution to making the introductions and literature reviews well-supported, systematic, and argumentative.

3.2 ChatGPT-Generated Texts: An Analysis of Keywords

In order to identify the specificities of ChatGPT-generated texts, a selection of keywords of

these materials vs. the human-authored content were identified and analyzed in their linguistic context. The keywords examined in this second phase of the study are listed in Table 3, together with their rank, raw frequency, and keyness.

Table 3. ChatGPT- vs. human-generated RA selections: keywords

Rank	Frequency	Keyness	Keyword
1	483	+236.8	research
2	110	+151.58	gaps
8	164	+85.72	often
9	105	+80.69	literature
10	93	+69.23	review
24	110	+51.91	particularly
28	107	+49.05	understanding
31	206	+47.63	studies
33	113	+45.33	while
34	36	45	rationale
37	37	+41.19	remain

A preliminary inspection of the list above, which intentionally does not include the keywords of the research articles used to develop the prompt, suggests that lexical items outnumber function ones. Starting with the most key content word, namely *research*, the analysis of its concordance reveals that several repeated constructions are utilized to provide an overview of the existing knowledge on a topic and justify the current study. Quite formulaic expressions appearing in the corpus of AI texts consist of *research* being pre-modified by adjectives such as *empirical, existing, extensive, recent, substantial,* and followed by reporting verbs like *suggest, demonstrate, show, highlight,* and *focus on* (23-25).

(23) ***Research suggests*** that self-mention in abstracts serves several rhetorical functions, such as stating research aims (This study investigates…), outlining methodology (We analyzed…), and claiming contributions (We argue that…). However, the extent to which self-mention is used varies widely across disciplines and time periods. (ChatGPT - EAP)

(24) ***Empirical research*** *on textbook discourse* ***has*** *increasingly* ***focused on*** macro-organization and rhetorical patterning. ***Studies*** *using move analysis* ***have shown*** *that* textbook chapters across disciplines display both stability and variation in their macro-move

structures (Nesi & Gardner, 2012). (ChatGPT - DS)

(25) ***Research Gaps*** *and* ***Rationale*** *for the Present Study*

> ***Despite substantial research*** *on* academic language, legal discourse, and identity, ***several gaps remain***. First, *much of the literature treats* legal language either as a technical system or as an object of critique, with less attention to how students actively negotiate linguistic and social identities through engagement with legal discourse. Second, *while sociological and ethnographic* ***studies have illuminated*** the socializing effects of legal education, fewer studies have integrated insights from EAP and sociolinguistics to examine the micro-level linguistic practices through which identity is constructed. Third, ***existing research*** *often* ***focuses on*** professional legal discourse rather than academic legal education, leaving undergraduate and early-stage students underexplored. [...] *The present study addresses these* ***gaps*** by examining academic language practices in legal education as sites of linguistic and social identity formation. (ChatGPT - DS)

Similarly, the lexical item *studies* in combination with boosting and tentative *verba dicendi* (e.g., *show, demonstrate, suggest, illuminate,* and *indicate*, 24-27) introduces what has already been researched on a specific topic, and points out areas where further research is needed (*few* or *fewer studies have examined* or *investigated*, 28).

(26) *Corpus-assisted discourse* ***studies have shown*** *that* evaluative lexis in news texts is not randomly distributed but systematically associated with particular topics, actors, and genres (Bednarek, 2006). (ChatGPT - DS)

(27) ***Studies indicate*** *that* computing introductions often rely heavily on technical terminology, nominalization, and concise clause structures, reflecting the need for precision, objectivity, and logical progression (Hyland, 2008; Flowerdew, 2000). (ChatGPT - ESP)

(28) First, ***few studies have*** *systematically* ***investigated*** the relationship between comprehension of English-language advertisements and consumer attitudes, brand perception, and purchase intentions across different cultural contexts. (ChatGPT - ESP)

In the texts generated by ChatGPT, quite general references to previous *studies* and a few citations (26-28) were found to be 'counterbalanced' by considerable attention paid to the *gaps* in the literature. The following excerpts give evidence of the tendency, on the one hand, of the *review* to identify *gaps* and *establish/provide a rationale* for the study, and, on the other, of *the present study to address these gaps* (29-31). The focus on unexplored areas in the extant literature is confirmed by several ChatGPT outputs that exhibit a subsection specifically dedicated to *Research Gaps and Rationale for the Present Study* (25). The gaps are highlighted either through quite fixed concessive chunks, including *despite extensive research…, despite a/the growing body of research, despite substantial research,* followed by the statement *several gaps remain* (25), or through references to the current *review* that *identifies gaps* and *the present study* that *addresses these gaps* (29-31).

(29) *The* ***review*** *synthesizes research* from applied linguistics, higher education studies, and language policy analysis to provide a comprehensive understanding of the intersections between language, scholarly communication, and institutional support. *It also* ***identifies gaps in the literature*** *and* ***provides a rationale for*** studying ESP language provision and policy in multilingual academic contexts. (ChatGPT - ESP)

(30) ***The present study addresses these gaps*** by combining corpus linguistics with move analysis to investigate macro-move organization and colony-in-loops structures in textbook

discourse. By focusing on recurring rhetorical patterns across textbooks, it aims to contribute to a more nuanced understanding of educational genres and their discourse organization. (ChatGPT - DS)

(31) Following this, *the literature on academic and collaborative networks is reviewed*, with an emphasis on how network structures influence research output and visibility. Finally, ***the review identifies gaps in the current literature and establishes the rationale for the present study***, which investigates how English proficiency and network engagement intersect to affect research productivity. (ChatGPT - ESP)

In addition to gaps, important research areas or topics are mentioned to justify the study, stress its contribution, and introduce future research developments. This is evidenced by the patterns *understanding* a topic *is crucial, important, essential,* or *critical* (32 and 33), and *the/this study aims to contribute/contributes to a deeper/a more nuanced understanding* (34) or *aims to deepen our understanding*.

(32) Despite this growing body of work, diachronic analyses of abstracts remain relatively limited. Given the abstract's role in indexing, database searches, and citation practices, ***understanding*** *how its interpersonal features evolve over time* ***is particularly important***. (ChatGPT - EAP)

(33) By adopting a diachronic perspective, future research can shed light on how abstract writing practices respond to changing academic environments, offering valuable insights for genre theory, discourse analysis, and EAP pedagogy. ***Understanding*** *these developments* ***is essential*** *for* capturing the dynamic nature of academic writing and for supporting effective scholarly communication. (ChatGPT - EAP)

(34) The present study responds to these needs by examining children's arguments using an integrated framework of conversation analysis, ethnomethodology, and membership categorization analysis. By focusing on how children invoke categories, organize sequences, and display cultural knowledge in argument, ***the study aims to contribute to a deeper understanding*** *of* children's interactional competence and the social organization of children's culture. (ChatGPT - DS)

Furthermore, deficiencies and limitations of previous studies are also expressed through the use of the keywords *often* and *while* in portions of text where concepts, theories, and approaches are contrasted with one another (35-38). Mentions of different viewpoints, conflicting findings, and inconsistencies between research studies are indicators of the argumentative nature of the ChatGPT-generated texts.

(35) *Although conditionals are* ***often*** *treated in grammatical descriptions as* logical or semantic devices, *research increasingly suggests that* they perform important interpersonal and rhetorical functions in academic discourse. (ChatGPT - EAP)

(36) *Early research on CMC* ***often*** *focused on* the affordances and constraints of specific technologies, comparing mediated communication with face-to-face interaction in terms of richness, social presence, and information transfer (Short, Williams, & Christie, 1976; Daft & Lengel, 1986). *More recent work has shifted toward* interactional and practice-based perspectives, examining how users adapt communicative strategies to accomplish social and organizational tasks within technologically mediated environments (Herring, 2004; Hutchby, 2001). (ChatGPT - DS)

(37) Hyland (1999, 2009) distinguishes between integral citations, where the author's name is

grammatically integrated into the sentence, and non-integral citations, where the reference appears parenthetically. *These forms carry different rhetorical meanings: integral citations foreground researchers and their contributions,* ***while*** *non-integral citations emphasize information or findings.* The choice between these forms reflects writers' stance toward sources and their desired level of authorial visibility. (ChatGPT - EAP)

(38) ***While*** *much of the literature on academic discourse has focused on English as the dominant language of global academia, there is growing recognition of* the need to examine academic discourse in other languages, including Portuguese, in order to understand how epistemological traditions and disciplinary practices are linguistically instantiated beyond Anglophone contexts (Bennett, 2014; Salgado & Gonçalves, 2020). (ChatGPT - EAP)

Notably, specific gaps in the state of the art or important elements to be further explored are singled out through the use of the keyword *particularly*. The analysis of this adverb in its wider co-text suggests that it is utilized either to delimit or narrow the focus of the discussion (39) or to intensify the importance or significance of the topic under consideration (40).

(39) Research on non-Anglophone writers' use of demonstratives in disciplinary writing *remains underdeveloped,* ***particularly*** *in relation to* cohesion, rhetorical efficacy, and genre conformity (Hyland & Tse, 2004). (ChatGPT - ESP)

(40) Formulaic sequences *are* ***particularly*** *prominent* in institutional English, where they function as conventionalized routines for negotiation, instruction, or bureaucratic reporting. By facilitating rapid comprehension and predictable meaning, these sequences serve both cognitive and social functions in professional communication. (ChatGPT - ESP)

Evidence from the analysis of the ChatGPT texts suggests that they seem to be generated following a rather standardized pattern and by means of a set of conventional phraseological units that describe previous research or the literature, highlight gaps, explain the rationale for the study, and stress its contribution to the existing body of knowledge.

4. Discussion and Concluding Remarks

The comparative analysis of a sample of ChatGPT- and human-generated texts has offered significant insights into their specificities in terms of language and discourse construction. The study, grounded in quantitative and qualitative methods, has revealed that the sections written by expert scholars and those created by ChatGPT show similarities in their textual organization and content discussion. Indeed, both the authentic and LLM-based introductions and literature reviews encompass an overview of previous theories and approaches to the study of a certain topic, while foregrounding the missing part in the reviewed literature and presenting it as the research focus. However, a close reading of the texts collected and the study of the keywords of the two datasets have shown that the ChatGPT production is characterized by a more uniform structure and more manifest attention paid to the gaps in the body of research on a specific topic, which serve as the foundation for justifying the current study.

At a discourse and language level (RQ1), evidence from the analysis of the keywords of one corpus vs. the other has suggested that the ChatGPT texts tend to use more content words, impersonal constructions, and rather standardized, homogeneous phraseological patterns, whereas human writing appears to feature simpler function words, personal and more varied formulations. This difference, which is also confirmed by an in-depth reading of content, is deemed to be indicative of two different approaches in writing, with AI possibly generating more informational content, and expert scholars prioritizing clarity and simplicity (Georgiou, 2025). The tendency of AI to prefer lexical items might be ~~possibly~~ interpreted as a

consequence of its attempts to emulate academic writing, which is documented to prefer impersonality, complex noun phrases, formulaic sequences, and informational packing. Thus, in line with Georgiou's considerations (2025), AI outputs appear to be more standardized and 'academic' than human texts.

Furthermore, in terms of argumentative strategies being adopted (RQ2), the two sets of materials seem to similarly synthesize the relevant literature within a particular field of research and compare or connect previous studies to identify gaps or controversies and motivate the present research. Nevertheless, the quantitative and interpretive analysis reveals that the ChatGPT texts feature quite generic references to previous studies, few citations, and no quotations, while their authentic counterparts exhibit a higher number of sources, quotations, and expressions used for exemplification and clarification of a topic or a theory. The authors' presence in the text is peculiar to human-authored content, whereas markers of evaluation can be identified in both sets of documents.

Thus, it can be inferred that ChatGPT might help develop comprehensive overviews of existing research and identify trends, gaps, and emerging themes in every single field. However, in relation to RQ3, the results of this study and an evaluation of output quality suggest that, although ChatGPT-generated content might appear convincingly human, disparities exist between authentic human writing and LLM-authored content.

This study clearly presents some limits as it takes into account only a restricted number of research articles and some particular sections for each of the texts under analysis. Moreover, this analysis might be further developed by examining how other GenAI platforms, such as Copilot, Gemini, Perplexity, and DeepSeek, mimic human academic writing. Finally, for a more rigorous and robust analysis, case study methodology, specifically questionnaires with academics, might be employed to gain experts' perceptions on the writing quality of the AI-created texts as well as their adherence to academic standards.

References

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

Anthony, L. (2019). *AntConc (Version 3.5.8) [Computer Software].* Tokyo Waseda University.

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

Baker, P., Gabrielatos, C., Khosravinik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society, 19*(3), 273-306. https://doi.org/10.1177/0957926508088962

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

Chukwuere, J. E. (2025). The Use of ChatGPT in Higher Education: The Advantages and

Disadvantages. *The Indonesian Journal of Computer Science, 14*(5), 8355-8365. https://doi.org/10.33022/ijcs.v14i5.5009

Cohen, J. F., & Moher, D. (2025). Generative artificial intelligence and academic writing: friend or foe?. *Journal of Clinical Epidemiology, 179,* 111646. https://doi.org/10.1016/j.jclinepi.2024.111646

Costa, K., Ntsobi P. M., & Mfolo, L. (2024). Challenges, Benefits and Recommendations for Using Generative Artificial Intelligence in Academic Writing - A Case of ChatGPT. *Medicon Engineering Themes, 7*(4), 03-38. https://doi.org/10.55162/MCET.07.236

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. *Biology of Sport, 40*(2), 615-622. https://doi.org/10.5114/biolsport.2023.125623

Durak, H. Y., Eğin, F., & Onan, A. (2025). A Comparison of Human-Written Versus AI-Generated Text in Discussions at Educational Settings: Investigating Features for ChatGPT, Gemini and BingAI. *European Journal of Education, 60,* e70014. https://doi.org/10.1111/ejed.70014

Dwivedi, Y. K., *et al.*. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71,* 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Galvan, J. L., & Galvan, M. C. (2017). *Writing Literature Reviews. A Guide for Students of the Social and Behavioral Sciences.* New York: Routledge.

Georgiou, G. P. (2025). Differentiating Between Human-Written and AI-Generated Texts Using Automatically Extracted Linguistic Features. *Information, 16,* 979. https://doi.org/10.3390/info16110979

Hsu, H.-P. (2023). Can Generative Artificial Intelligence Write an Academic Journal Article? Opportunities, Challenges, and Implications. *The Irish Journal of Technology Enhanced Learning, 7*(2), 158-171. https://doi.org/10.22554/ijtel.v7i2.152

Jenko, N., Ariyaratne, S., Jeys, L., Evans, S., Iyengar, K. P., & Botchu, R. (2024). An evaluation of AI generated literature reviews in musculoskeletal radiology. *Surgeon, 22*(3),194-197. https://doi.org/10.1016/j.surge.2023.12.005

Ji, J., Li, R., Li, S., Guo, J., Qiu, W., Huang, Z., ... Lu, X. (2024). Detecting Machine-Generated Texts: Not Just "AI vs Humans" and Explainability is Complicated. *ArXiv, abs/2406.18259*. https://doi.org/10.48550/arXiv.2406.18259

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update, 5*, 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education, 10*(2). https://doi.org/10.1080/2331186X.2023.2236469

Mozelius, P., & Humble, N. (2024). On the use of generative AI for literature reviews: An exploration of tools and techniques. *Proceedings of the 23rd European Conference on Research Methodology for Business and Management Studies 23*(1), 161-168.

https://doi.org/10.34190/ecrm.23.1.2528

Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education, 49*(5), 847-864. https://doi.org/10.1080/03075079.2024.2323593

Ngwenyama, O., & Rowe, F. (2024). Should We Collaborate with AI to Conduct Literature Reviews? Changing Epistemic Values in a Flattening World. *Journal of the Association for Information Systems, 25*(1), 122-136. https://doi.org/10.17705/1jais.00869

Partington, A. (2010). Modern diachronic corpus-assisted discourse studies (MD-CADS) on UK newspapers: An overview of the project. *Corpora 5*(2), 83-108. https://doi.org/10.3366/cor.2010.0101

Schryen, G., Marrone, M., & Yang, J. (2024). Adopting Generative AI for Literature Reviews: An Epistemological Perspective. Retrieved from https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/f7b86866-1fdc-47e7-874c-e889128effa9/content. https://doi.org/10.24251/HICSS.2024.486

Sinclair, J. (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.

Taylor, C., & Marchi, A. (2018). *Corpus approaches to discourse: A critical review*. London: Routledge. https://doi.org/10.4324/9781315179346

van Niekerk, J., Delport, P. M. J., & Sutherland, I. (2025). Addressing the use of generative AI in academic writing. *Computers and Education: Artificial Intelligence, 8*, 2025, 100342. https://doi.org/10.1016/j.caeai.2024.100342

Copyrights