Does ChatGPT write like a student? Engagement markers in argumentative essays

Abstract

ChatGPT has created considerable anxiety among teachers concerned that students might turn to Artificial Intelligence (AI) programmes to write their assignments. This AI-powered large language model is able to create grammatically accurate and coherent texts, thus potentially enabling cheating and undermining literacy and critical thinking skills. This study seeks to explore the extent AI can mimic human-produced texts by comparing essays by ChatGPT and student writers. By analysing 145 essays from each group, we focus on the way writers relate to their readers with respect to the positions they advance in their texts by examining the frequency and types of engagement markers. The findings reveal that student essays are significantly richer in the quantity and variety of engagement features, producing a more interactive and persuasive discourse. The ChatGPT-generated essays exhibited fewer engagement markers, particularly questions and personal asides, indicating its limitations in building interactional arguments. We attribute the patterns in ChatGPT's output to the language data used to train the model and its underlying statistical algorithms. The study suggests a number of pedagogical implications for incorporating ChatGPT in writing instruction.

Introduction

The use of computational technologies in language learning is nothing new. For decades, conversational agents, commonly referred to as chatbots, have been applied to support teaching of foreign languages (Rudolph, Tan & Tan, 2023; Ji et al., 2023). However, recent advancements in machine learning have led to the development of Large Language Models (LLMs), such as ChatGPT (Chat Generative Pre-trained Transformer), which can produce human-like text by responding to user queries. LLMs like ChatGPT utilise sophisticated algorithms and natural language processing to simulate conversational interactions and generate various forms of content. They cannot, of course, engage in human-like cognitive processes or possess understanding, intention, or awareness (e.g. Gallagher, 2023; Byrd, 2023). The apparent "intelligence" exhibited by these models is thus an emergent property of statistical patterns identified in their training data, rather than a result of human-like reasoning or comprehension. Therefore, in this study we use the term "LLM" to describe these computational models and refer to programmes like ChatGPT.

Reactions to the role of such AI programmes in language learning and student writing are mixed. For some, the language modelling capabilities of ChatGPT can help scaffold students' language study (e.g. Kasneci et al., 2023; Kohnke et al., 2023). Others express concern that "the eerily humanlike chatbot" (Satariano & Kang, 2023) might make it difficult to distinguish GPT-generated and human-authored texts in assessing students' writing (Revell et al., 2023). The possible temptation to rely on AI when writing assignments can undermine students' development of critical thinking, problem-solving and literacy skills, as Laquintano, Schnitzler and Vee (2023) have argued. Although tools such as *GPTZero* and *AICheatCheck* have been developed to detect AI involvement in writing, these seem currently unable to make a reliable distinction (Adeshola & Adepoju, 2023; Scarfe et al., 2024).

We approach this question by exploring the performance of the model when asked to complete a writing task previously done by students. We focus on engagement expressions, which refers to the way "writers relate to their readers with respect to the positions advanced in the text" (Hyland, 2005c, p. 176). The analysis seeks to offer textual evidence to help identify ChatGPT-generated texts and to offer support for L2 students and teachers seeking to use the tool in classrooms. Before describing the methods and results, we introduce ChatGPT, argumentative writing and reader engagement.

ChatGPT and writing assistance

ChatGPT is based on machine learning algorithms that have been trained on a large and genre-distinct datasets of texts published on the internet on an array of topics (e.g., Reddit, Wikipedia, New York Times, etc). It is designed to learn the statistical patterns and relationships between words and phrases in collections of texts (Kabir, 2022) and represents a significant advance in natural language processing and artificial intelligence. A form of generative AI, ChatGPT uses specialized algorithms to find patterns within data sequences in order to respond to user prompts with images, texts or videos created by artificial intelligence.

Essentially, the model generates responses through a sophisticated sequence-processing mechanism, wherein it analyses the input text word by word, "predicting the next word in the sequence based on the context of the words that come before it" (Kumar, 2023). This predictive process is iterative; each newly predicted word serves as input for subsequent predictions, continuing until the desired textual output is achieved. Rather than processing language on a clause-by-clause basis, then, ChatGPT calculates probabilities across vast spans of text in its training data. This approach allows it to generate coherent text but means it does not "understand" context in the way humans do. Instead, it produces text based on statistical patterns learned from its training data,

which can lead to both impressively fluent output and unexpected limitations in adapting to specific contexts or tasks (Byrd, 2023; OpenAI, 2023).

ChatGPT therefore offers the potential to provide a range of applications in language learning from natural language processing to conversation generation, language translation, text summarization, grammar correction, paraphrasing, and more (Bin-Hady et al., 2023; Laquintano, Schnitzler & Vee, 2023; Pack & Maloney, 2023). Yet it is seen by students and researchers as most useful when assisting with writing: to plan, write, edit and polish academic texts (Ingley & Pack, 2023; Nordling, 2023). Here ChatGPT can serve as an initial sounding board for brainstorming ideas (Su et al., 2023) and provide corrective feedback on students' writing assignments (Godwin-Jones, 2022). Its potential to assist writing rests on a capacity to generate writing which is "typically coherent and grammatically correct" (Barrot, 2023, p. 2), and to refine the tone and style of a text (Ji et al., 2023). Based on these studies, it can be argued that ChatGPT is particularly useful for non-native English speakers to improve their academic writing skills.

However, despite these advantages, there are ongoing concerns, particularly about bias, hallucinations and sycophancy (e.g., Santurkar et al., 2023; Zhang et al., 2023). In educational contexts, most serious is perhaps the difficulty of distinguishing "whether a text is machine- or human-generated, presenting an additional major challenge to teachers and educators" (Kasneci et al., 2023, p. 6). Thus, comparing undergraduate exam scripts generated by ChatGPT to those by students, for example, Scarfe et al. (2024) found that 94% of AI submissions were undetected by readers. Even trained linguists are not particularly effective in spotting the differences with an average total positive identification rate of only 38.9% (Casal & Kessler, 2023).

One possible difference is the extent of interactional involvement ChatGPT invests in the texts it creates. So, while ChatGPT can generate reasoned and contextually

appropriate text, it lacks an inherent understanding of audience. Unlike human writers, who develop a mental model of their readers and adjust their writing accordingly (Hyland & Jiang, 2023), ChatGPT does not possess an intrinsic awareness of who might be reading its output. This limitation is a consequence of the fact that the model is trained on huge amounts of texts from diverse registers and genres, each with its own purposes, structures and audiences (Milano, 2023). This process bleaches out any specific audience and means that the model operates with a 'generic' target reader. Consequently, any audience-specific features in ChatGPT-generated text, particularly engagement markers, are incidental, reflecting patterns in the training data rather than a clear consideration of readers' needs. This "audience blindness" can result in output that, while grammatically conventional and topically relevant, may lack the nuanced involvement features that characterise effective human writing (Markey et al., 2024).

To put this plainly, the absence of a built-in model of audience means that ChatGPT cannot automatically adapt its rhetorical style, tone, or level of detail to suit different reader groups, nor can it anticipate and address potential reader questions or objections without specific prompting. Audience awareness is central to academic writing and control of interpersonal elements can be crucial to successful persuasion (Hyland, 2005a; Su et al., 2023). Such dialogic aspects of argument not only involve conveying an appropriate stance but acknowledging and addressing the role of readers (Hyland, 2004; Shahriari & Shadloo, 2019). Recent studies, however, suggest that ChatGPT-generated essays "exhibit reduced involvement and integration compared to their human counterparts" (Berber Sardinha, 2024, p.9), and "often read as dialogically closed, 'empty,' and 'fluffy'" (Markey et al., 2024, p.1). These findings, based on Biber's (1988) multidimensional analysis, provide useful insights into the general characteristics of AI-generated text. However, while this approach offers a broad perspective on textual features, it does not specifically focus on the nuances of interpersonal interaction within the essays. As such, there remains a need for more targeted research to systematically compare the interpersonal elements in AI-generated and human-written essays.

Our study sets out to provide textual evidence for human-AI differences in this regard by exploring the extent to which ChatGPT can generate argumentative content with the same form, frequency and function of reader engagement as humans. Our comparison provides insights into the development of more nuanced writing instruction methods that leverage the strengths of LLMs while addressing their limitations. Beyond this, our study points to broader issues around the limitations of the datasets this particular bot was trained on and the difficulties of designing prompts that communicate a context for any desired text.

Engagement and academic writing

Academic writing is now widely regarded as a socially-mediated persuasive endeavour where writers seek to engage readers by anticipating their readers' reactions (Hyland, 2004; Jiang & Ma, 2018). For Hyland, (2005b) engagement is

> an alignment dimension where writers acknowledge and connect to others, recognizing the presence of their readers, pulling them along with their argument, focusing their attention, acknowledging their uncertainties, including them as discourse participants, and guiding them to interpretations (p.176).

The dialogic nature of academic writing suggests that writers, in constructing an appropriately disciplinary voice, must represent their readers in the text in acceptable ways (Hyland & Jiang, 2023). This is intrinsically linked to the writer's awareness of their audience and the expectations of their disciplinary community. It involves not only acknowledging the reader's presence but also anticipating their potential questions, objections, and need for clarification.

This use of *engagement* differs from that proposed by Martin and White (2005) in their evaluation framework, where it is a writer-centred feature referring to the ways authors take a position towards other voices. Following Hyland (2005a), in contrast, we are concerned with how language is used to bring reader into a text in order to gain their

acceptance of ideas and head-off possible objections. Engagement in this paper therefore refers to the overt marking of what Thompson (2001) calls the "reader-in-the-text"; a practice of social engagement central to the success of academic argument.

The significance of engagement in academic persuasion has been demonstrated in a range of genres and contexts. They are found to play a key role in academic blogs (Zou & Hyland, 2020), doctoral confirmation reports (Jiang & Ma, 2018) and final year undergraduate reports (Hyland 2006), as well as the popularising genres of TED talks (Liu et al, 2017) and 3-minute theses presentations (Hyland & Zou, 2022). McGrath and Kuteeva (2012), for example, report higher than expected shared knowledge markers and reader references in pure mathematics articles while Hyland and Jiang (2016) show that the use of engagement markers has increased over time in research articles in both hard and soft disciplines. Comparisons have also been made in the ways patterns of engagement vary by genre and language, revealing how writers shape their texts to the expectations of different audiences. Therefore, Hyland (2004) found differences between expert texts and undergraduate dissertations and between popular and professional science articles (Hyland, 2010). In addition, while context and national culture can influence the use of engagement, L1 transfer and L2 proficiency may also have some bearing (Lafuente-Millán, 2014).

Typically, the way writers articulate arguments and initiate social engagement is shaped by their understanding of 'audience'. In academic contexts this is rarely a real person but an abstraction conjured up by the writer and based on his or her knowledge of the community for which the text is written. Thus audience comprises the writer's perception of the external circumstances that define a rhetorical context and influence the specific textual conventions employed. Thus Park (1982), for example, argues that audience exists in the writer's mind and shapes a text as "a complex set of conventions, estimations, implied responses and attitudes" (p.251).

Writers therefore navigate the complexities of engaging audience by drawing on the rhetorical and structural conventions of the genre and by ways of crafting arguments that are recognised and valued within their disciplinary communities. Hyland (2005c) argues that there are five main ways that authors overtly intrude into their texts to connect with readers directly. At certain points writers acknowledge an active audience using the following:

Table 1 Categories of reader engagement (Hyland, 2005c)

Category	Description	Example
Reader mentions	They bring readers into a discourse, normally through second person pronouns, particularly inclusive *we* which identifies the reader as someone who shares similar ways of seeing to the writer.	(1) As **we** can imagine, this has had a tremendous influence on sales in places such as fast-food restaurants where beefburgers are the main item on the menu. (Student essay[1])
Questions	They invite direct collusion because they address the reader as someone with interest in the issue the question raises and the good sense to follow the writer's response to it.	(2) Can we expect a scientist to bear this additional burden for the whole world**?** In truth no, it is unreasonable. (Student essay)
Appeals to shared knowledge	They are explicit signals asking readers to recognise something as familiar, apparent or accepted.	(3) **Traditionally**, participating in a lottery involved purchasing a physical ticket from an authorized. retailer. (ChatGPT essay)
Directives	They are instructions to the reader, mainly expressed through *imperatives* (such as *consider*, *note*), *obligation modals* (*need to*, *should*), and *predicative adjectives* (*it is important to understand*…), which direct readers a) to another part of the text or another text, b) to carry out some action in the real world, or c) to interpret an argument in certain ways.	(4) As IVF technologies continue to advance and become more integrated into the fabric of society, **it is vital to consider** the demographic trends they influence. (ChatGPT essay)
Personal asides	They are brief interjections where the writer speaks directly to the reader, often to share a personal thought, comment, or anecdote. These asides can create a conversational tone, add personality to the writing, and help to engage the reader.	(5) When this tunnel is finally completed (**hopefully in the near future)** it will be much easier to travel to and from Europe. (Student essay)

[1] All the examples are taken from our corpora of argumentative essays by British university students and ChatGPT discussed below.

As seen in Table 1, these features are the most explicit means at the writer's disposal to recognise their readers in the text, acknowledge their expectations of inclusion and respond to their possible objections and alternative interpretations (Hyland, 2005c). While inclusion of readers might sometimes be based on tacit assumptions and expressed implicitly through, say, choice of method, theory or data, explicit engagement features help to concretise the ways that writers intervene to "engage actively or position readers, focusing their attention, recognizing their uncertainties, including them as discourse participants and guiding them to interpretations" (Hyland, 2001, p. 552). They therefore carry important rhetorical meanings while managing the impression readers get of the writer (Hyland & Jiang, 2016).

With the growing influence of AI on academic writing, what is missing from these studies is the question of whether ChatGPT can produce texts with the same degree of nuance and variability of reader engagement.

Methodology

Data collection

As outlined above, we set out to compare the argumentative essays generated by ChatGPT with those written by British university students. For the latter we drew upon the Louvain Corpus of Native English Essays (LOCNESS), a collection of texts written by British and American university students[2]. From this corpus, we extracted 145 argumentative essays written by second-year students from British universities, which are designated as 'GCE A-level' quality. GCE A-level essays are characterised by clear and coherent argumentation, advanced language proficiency, and adherence to academic writing conventions, serving as a benchmark for high standards in educational assessments (Wilson, Child & Suto, 2017). The topics cover a range of subjects, including Britain's relationship with the EU, transport, boxing, the parliamentary system, computers and life, eating beef, the lottery, and fox hunting. These texts lack

---

[2] https://www.learnercorpusassociation.org/

references but are nevertheless written in a formal style, characterised by the use of academic language, structured argumentation and an objective tone.

For the ChatGPT corpus we wrote a prompt for each topic, following the student prompts as closely as possible. We recognised the sensitivity of ChatGPT to phrasing and took advice from the literature on effective AI prompts (e.g. Sarrion, 2023; Laquintano, Schnitzler & Vee, 2023). In particular, we ensured that ChatGPT was provided with a specific role, a realistic context and the rules and tone of the writing assignment. The process of prompt engineering typically involves recursive fine-tuning, which entails a cycle of iterative adjustments and evaluations. For instance, when addressing the topic of transport, we provided specific instructions to the model, refining these through trial and error to achieve this final prompt:

> "You are competent in academic writing. Write 16 argumentative essays with a persuasive writing style on the topic of transport. Each essay is about 500 words long."

For the AI generated essays, we used ChatGPT 4.0, a more advanced model than its predecessor in terms of its enhanced understanding and contextualization abilities, its expanded knowledge base and its improved language skills (Koubaa, 2023). The details of the two corpora are shown in Table 1.

Table 1 Corpus characteristics

Corpus	No. of texts	Tokens	SD[*]	Types
ChatGPT essays	145	72,819	33	7,164
British student essays	145	78,060	52	6,975

* SD = Standard deviation

GPT4 therefore has the capability to produce text outputs that exhibit human-level performance in various professional and academic benchmarks, and our analyses set out to test these claims.

Data analysis

The two corpora were part-of-speech tagged by *TagAnt* (Anthony, 2014) and then searched for Hyland's (2005c) engagement features using *AntConc* (Anthony, 2022). Overall we examined about 100 different items of reader engagement provided by Hyland (2005c) and Hyland and Jiang (2016) (see Appendix 1), and manually checked each concordance to establish that the feature were performing an engagement function by addressing readers directly. Most obviously, this involved eliminating non-addressee modals (*every scientist should be a good judge*) and interjections which were not personal asides (*the problem is obvious - there are too many cars on Britain's roads*). In addition, some features were easily located through a corpus word-search (*we*, *of course*) while others entailed a regular expression search (*imperatives*, *it is adj to + verb*).

The two authors worked independently and coded a random sample of 10% of engagement expressions, achieving an inter-rater reliability of 97%. Disagreements were resolved through discussion and consideration of other examples. For instance, disagreement arose coding the phrase 'you might wonder why...' in the introduction of an essay. The second author argued that this phrase should be classified as a reader mention, while the first author felt it should be considered a rhetorical question, which typically serves a different function of engagement. After discussing the context and reviewing similar examples, we agreed to code it as a reader mention and concluded that the primary function of the phrase was to directly address the reader and anticipates their thoughts.

We normalised the results to 1,000 words to compare the use of engagement across the two corpora and determined statistical significances in these differences by applying log-likelihood (*LL*) tests using Rayson's (2016) calculator. We followed the suggestion in that paper that an *LL* score of 3.8 or higher is significant at a cut-off p-value of 0.05. We also considered the effect size for log-likelihood tests (*%DIFF*), which indicates

the percent of the difference between the two normalized frequencies (see Gabrielatos, 2018 for more information about %DIFF).

Reader engagement by ChatGPT and British students
Overview

We identified 393 cases of engagement in the ChatGPT essays, averaging 5.40 cases per 1000 words, and 1,326 cases in the essays by British students, amounting to 16.99 cases per 1000 words. This shows significantly less use of engagement markers by ChatGPT in creating argumentative essays (*LL*= 471.98, *%DIFF*=68.23, *p*<0.001). Interestingly, Jiang and Hyland (2024) similarly identified significantly fewer 3-word stance bundles (e.g., *it is possible*, *this never is*, *in my opinion*) in the ChatGPT essays than used by human writers. Clearly, this does not tell us a great deal about the *quality* of the essays per se, as more interactional devices do not necessarily mean more effective texts. Hyland (2004) and Jiang and Ma (2018), for instance, found far fewer uses of engagement in student than professional writing. However, significantly fewer markers of engagement reveal a distinctive characteristic of the AI texts and indicates a gulf in the interactional positions taken in the two corpora.

As we have mentioned, a writer's rhetorical investment in engagement contributes to the impression of reader-awareness and recipient design in a text and helps construct an effective line of reasoning, establishing a connection with readers as in (6) and (7):

    (6) **We** all feel that **we** have a divine right to be on the road. **Why?**

                                 (Student essay)

    (7) As **we** navigate this digital social landscape, **it is crucial to** foster digital literacy and etiquette to ensure that **our** online interactions are respectful, authentic, and enriching.   (ChatGPT essay)

We should also point out here that the normalised frequency of engagement in both students and ChatGPT-generated essays is higher than that reported by Hyland and Jiang (2016) for research articles over time. The fact that LLMs such as ChatGPT "learn"

from a wide range of registers and genres gives them a broad understanding of context and an adaptability to different writing styles and genres (Milano, 2023; Wolfram, 2023). This adaptability allows them to tailor essays to a broader audience, using a more accessible tone compared to the more formal conventions of research articles. Our comparison with student writers, however, shows the programme was unable to mirror the engaging tone of the student texts.

Table 2 shows the distribution of engagement features in the two corpora and we can see here that the overall percentages of reader mention and directives in the ChatGPT texts aligns quite closely with the students' choices. This distribution also corresponds with frequencies for these items in the argumentative essays by EFL students studied by Yoon (2021).

Table 2 Frequency of engagement in the two corpora (normed frequency & % of total)

Engagement features	ChatGPT essays					British student essays				
	raw	normed*	SD*	DP*	%	raw	normed	SD	DP	%
Reader mention	160	2.20	0.07	0.15	40.71	499	6.39	0.36	0.28	37.63
Directives	134	1.84	0.05	0.18	34.10	478	6.12	0.26	0.35	36.05
Knowledge appeals	93	1.28	0.03	0.25	23.66	124	1.59	0.04	0.61	9.35
Questions	6	0.08	0.01	0.20	1.53	134	1.72	0.05	0.48	10.11
Personal asides	0	0.00	0.00	0.30	0.00	91	1.17	0.02	0.72	6.86
Total	393	5.40	0.31	0.22	100.00	1326	16.99	0.89	0.46	100.00

Note: normed = normed frequency per 1,000 words; SD = standard deviation; DP = deviation of proportions

Reader mentions and directives, of course, function to foster both a shared perspective and ensure the argument is understood. They create a more conversational tone and help make the argument more accessible as in (8) and (9).

(8) In order to solve **our** transport problems, **we need to** start taking a

more long term view.   (Student essay)

(9) Public policy and regulatory frameworks **must** be developed with an

emphasis on inclusivity.   (ChatGPT essay)

This rhetorical work is key to argumentative essays which aim to persuade readers of a point of view by addressing them directly and making. Therefore, reader mentions and directives play a significant role in distinguishing successful and unsuccessful essays (Lee & Deakin, 2016).

Following these features, preferences differ with knowledge appeals more frequent in the ChatGPT-generated essays and questions in the student essays. Below, we discuss these results in more detail.

Reader mention: soliciting solidarity

Explicitly referring to the reader is the most direct way for writers to appeal to their readers and create a sense of common ground for interpreting claims. Second person *you*, *your* and the inclusive first person *we*, *our*, *us* all signal writers' attempts to bring readers directly into the text and demonstrate a bond with them, claiming all participants are on the same wavelength and see things in a similar way. Both ChatGPT and student writers overwhelmingly prefer the inclusive forms (Table 3).

Table 3 Reader mentions in the two corpora (normed frequency & % of total)

	ChatGPT essays					British student essays				
	raw	normed	SD	DP	%	raw	normed	SD	DP	%
we/our/us	152	2.09	0.08	0.12	95.00	721	9.24	0.21	0.25	90.24
you/your	8	0.11	0.01	0.15	5.00	78	1.00	0.02	0.30	9.76

Second person is the most unequivocal acknowledgement of the reader's presence in the argument, a fact which may help account for its declining use in academic research articles over time (Hyland & Jiang, 2016). The students too may seem equivocal about the form, seeing it as implying a stark separation of writer and reader, "marking out the differences and perhaps emphasising the writer's relatively junior status compared with the teacher/reader" (Hyland, 2005b, p. 369). Novice writers are often uncertain about the engaging readers in such an explicitly direct and personal way. They perhaps

recognise it as characterising more intimate registers, too personal or informal for academic writing (Hyland & Jiang, 2017) while their textbooks, style guides and teachers generally advise them to avoid it.

Perhaps as a consequence of these personal implications, students mainly used *you* and *your* with a broader semantic reference, referring to people in general (similar to the indefinite pronoun *one*) rather than specific discourse participants:

(10) It is said that **you** can meet people through computers and have "relationships".   (Student essay)

(11) For instance, spelling is no longer as important as it was as **you** can simply use a "spellcheck" to correct **your** English, which is absurd.

(Student essay)

Here, *you* and *your* carry a more encompassing meaning than rhetorically focusing on an individual reader, seeking instead to engage with readers by recruiting them into a world of shared experiences. Interestingly, this rhetorical use was not found of readers in the ChatGPT essays.

Inclusive *we*, on the other hand, implies a shared understanding and collective goal. Although it is undoubtedly dialogic by considering the readers' perspective on an issue, *we* addresses readers from a position of authority, steering them through an argument towards a preferred conclusion. Reader pronouns therefore assert both authority and collegiality; facilitating a dialogue intended to persuade readers to agree with the author's claims as in (12) and (13). This is perhaps why this form of reader mention dominates the two corpora.

(12) For example, if **we** select for certain cognitive or physical traits, **we** could inadvertently narrow **our** understanding of what it means to be human.   (ChatGPT essay)

(13) This is why it could be said **we** don't use **our** brains as much.

(Student essay)

15

Here inclusive *we* draws readers in less personally by invoking a world of commonsense activity and suggesting a shared general knowledge with the writer in the collective exploration of a knotty problem.

Questions: raising a query

Questions are a key engagement feature because they presuppose and mark the presence of the reader whose attention is captured and selectively focused on key points in the writer's argument, ostensibly presenting an invitation for readers to orientate themselves and respond to the argument. We note that the frequency of questions used in these students essays is higher than Hyland (2005b) found in his corpus of final year students' research reports. We see this as a consequence of different genre constraints, where argumentative essays give writers more leeway in this regard, although the more formal reports give writers the freedom to raise questions about aspects of the methodology or other elements of the research (Gong et al., 2024).

The following examples from our student corpus show how questions offer writers a way of creating a sense of immediacy and engagement with the reader. However, questions obviously differ in the functions they perform in academic writing, conveying a range of meanings from naive puzzlement of limited knowledge to the confident anticipation of reaching an answer.

> (14) If Britain were to join "The Single Market", because of our well-
> known independence and head-strength, would we not just be "rocking
> the bout" so to say**?**   (Student essay)
> (15) Should they have the right to "buy" themselves a baby**?** I think so.
>   (Student essay)

Whatever the sense questions carry, they all invite direct collusion since the reader is addressed as someone with an interest in the issue raised by the question, the ability to recognise the value of asking it, and the good sense to follow the writer's response to it

16

(Hyland, 2002). Questions, then, are the strategy of dialogic involvement par excellence, serving up an invitation for readers to orientate themselves in a certain way to the argument presented and to enter a frame of discourse where they can be led to the writer's viewpoint (Hyland, 2002). The ChatGPT essays, in contrast, contained very few questions, rhetorical or otherwise and appears to have limitations in accurately identifying and interpreting such questions. Curry et al. (2024), for example, observed that ChatGPT sometimes fabricated questions in their corpus by adding question marks or question tags to declarative statements, resulting in inaccurate questions.

In addition, we see a considerable percentage of questions in the student essays combined with inclusive *we* pronouns as writers interjected questions on behalf of the intelligent reader who is brought into the text through this shared exploration of the topic (16 and 17).

> (16) But we ought to ask ourselves "What happens when the computer-
> orientated world collapses**?**" We would then have to use our brains.
>
> (Student essay)
>
> (17) But are we right to blame him**?** Let us consider that he has discovered
> a cure for cancer as a result of genetic engineering.   (Student essay)

Such a rhetorical use of questions implies a cooperative effort to address a common problem. This is a powerful engagement strategy as the writer brings the reader along towards a pre-determined conclusion. The passage is almost relentless in the way it positions the reader in relation to the writer and to the issue at hand, presupposing the reader's response as well as the reasonableness of the questions themselves.

Knowledge appeals: constructing sharedness

Less imposing than questions and less directly personal than reader pronouns, knowledge appeals are used "to position readers within the apparently naturalized and unproblematic boundaries of disciplinary understandings" (Hyland 2001, p. 566). By

raising the possibility of common knowledge, writers seek "endorsement of sources which are highly respected in the field and carry the status of objective facts" and emphasise their footing with the common issues by "showing awareness of these sources and by showing their relevance to their work" (Koutsantoni, 2004, p. 176). We get a taste of the effect of this strategy in the following extracts:

> (18) While the quest to preserve sovereignty is legitimate, it is also evident that in today's interconnected world, complete sovereignty in the **traditional** sense is more of an ideal than a practical reality.
>
> (ChatGPT essay)
>
> (19) It is now possible, therefore, that the UK opposes proposed legislation but that is voted through by the other states and becomes law here without our consent. This is **obviously** an infringement of our sovereignty already.
>
> (Student essay)

Readers are brought to agreement with the writer through the sleight of hand of building on what the writer suggests is already implicitly agreed. By explicitly referring to this assumed agreement, writers construct themselves and their reader as fellow travellers on the path of knowledge. Interestingly, knowledge appeals account for a higher percentage in the ChatGPT-generated essays, which indicates the model's use of its ability to access vast amounts of existing information and data to generate responses (Sarrion, 2023; Wolfram, 2023). ChatGPT's essays draw on forms it sees in a range of registers, genres and contexts to help align the text and the reader with commonly known facts and concepts, potentially leading to reader agreement. The reliance on shared knowledge in ChatGPT-generated essays implies that the model can produce content that resonates with existing information and ideas, enhancing the overall quality and credibility of the output.

Following Hyland and Jiang (2016), we further categorised *appeals to shared knowledge* according to their principal function as follows:

18

- **Logical reasoning** – concerned with the coherence of the argument, such as *obviously*, and *of course*.

    (20) Soon, **<u>of course</u>**, this will become even less of a barrier with the completion of the "Channel Tunnel".   (Student essay)

- **Routine conditions** – concerned with usual circumstances or behaviour of real-world objects, such as *normally* and *regularly*.

    (21) The main disadvantage with the railways is as the rail service and the bus service are **<u>normally</u>** owned by different companies...   (Student essay)

- **Familiarity with tradition** – concerned with usual community practices and beliefs, such as *common* and *traditionally*.

    (22) This is important because an incorrect number of chromosomes, a condition known as aneuploidy, is a **<u>common</u>** cause of miscarriages ...   (ChatGPT essay)

Table 4 shows that almost all knowledge appeals in the ChatGPT-generated essays refer to usual practices and typical beliefs.

Table 4 Knowledge appeals in the two corpora (normed frequency & % of total)

	ChatGPT essays				British student essays			
	normed	SD	DP	%	normed	SD	DP	%
Logical reasoning	0.00	0.00	0.00	0.00	0.54	0.02	0.35	33.87
Tradition & typicality	1.27	0.03	0.14	98.92	0.74	0.02	0.28	46.77
Routine conditions	0.01	0.00	0.18	1.08	0.31	0.01	0.32	19.35

Because the model is based on the patterns and data inputs it has been trained on and generates information by learning from a vast amount of text data, it incorporates knowledge, and ways of referring to it, which are prevalent in society (Nordling, 2023; Sarrion, 2023). Therefore, when generating text, ChatGPT is more likely to reference

shared knowledge that aligns with usual practices and beliefs in (23) and (24), as this is the information it has been exposed to and learned from.

(23) The need to develop new markets has become pressing, though it is a challenging prospect given the **established** tastes and demands of traditional European markets. (ChatGPT essay)

(24) **Traditionally**, the concept of family was often narrowly defined: a heterosexual couple with the ability to conceive naturally. (ChatGPT essay)

There were no appeals to logical reasoning in the ChatGPT essays, suggesting that ChatGPT may be more proficient at regurgitating factual information than synthesizing complex ideas or concepts. In this way, it perhaps endorses findings that ChatGPT struggles with content that requires higher-order thinking such as critical and analytical thinking (Rudolph et al., 2023).

Logical reasoning, however, comprises 33.9% of shared knowledge appeals in students' essays. By summoning such a shared logic, writers can avoid fallacies and ensure that their arguments are based on sound principles as in (25) and (26).

(25) This will **apparently** extend our free market economy to the whole of Europe, or at least to those countries who participate. (Student essay)

(26) An **obvious** problem with a single Europe **of course** would be the language barrier, should we learn a common language? (Student essay)

This type of rhetorical appeal enhances the integrity of the argument, as it relies on (writer invoked) common ground as widely accepted fact and rational progression, increasing the likelihood that readers will find the argument compelling and credible. Putting ideas into readers heads and claiming that thy already know and agree with them, moreover, helps address potential counterarguments and allows the writer to anticipate and refute opposing views, thereby strengthening the overall argument.

Directives: managing readers

Directives are utterances which instruct the reader to perform an action or to see things in a way determined by the writer (Hyland, 2001, 2005c) and remain a frequent device used to initiate reader participation in academic texts, comprising around 30% of all features in both corpora. According to Hyland (2005a), they are typically realised by an imperative (27); by a modal of obligation addressed to the reader (28); by a first person inclusive *let*-imperative (29); and by a predicative adjective expressing the writer's judgement of necessity/importance controlling a complement *to-* clause (30):

(27) **Take for example** Kalashnikov, the inventor of the Kalashnikov machine gun.   (Student essay)

(28) Measures **should** be put in place to support farmers through the change...   (ChatGPT essay)

(29) **Let us consider** how a professional boxer would feel. (Student essay)

(30**) it is vital to** consider the demographic trends they influence…

                                              (ChatGPT essay)

In each case there is a clear reader-oriented focus as the writer signals a recognition of the dialogic dimension of argumentative writing, intervening to direct the reader to some action or understanding.

Table 5 shows that modals are the preferred form in both the ChatGPT and student texts, signalling what the writer believes is either necessary or desirable, they carry a less imposing and commanding force than forms such as imperatives (Hyland, 2001; Jiang & Ma, 2018). As seen from the extracts, the obligation is typically tempered with less imposing modals such as *need to, have to*, often combined with inclusive *we*, especially in the students' essays.

Table 5 Directives in the two corpora (normed frequency & % of total)

	ChatGPT essays				British student essays			
	normed	SD	DP	%	normed	SD	DP	%

Modal + V.	1.79	0.04	0.13	97.01	5.70	0.20	0.26	93.10
Adj. to V.	0.05	0.00	0.16	2.99	0.27	0.02	0.31	4.39
Let us/let's	0.00	0.00	0.00	0.00	0.09	0.01	0.38	1.46
Imperatives	0.00	0.00	0.00	0.00	0.06	0.01	0.40	1.05

(31) We **have to** question whether independent action, such as Britain is undertaking in the Gulf at the moment, will be permitted. (Student essay)

(32) Genetic manipulation and gene technology is a vast field of study and it is growing increasingly larger, which is why we **need to** monitor its progress in order to limit the size of any potential damage. (Student essay)

(33) However, these advances **must** be managed carefully to safeguard patient data ...   (ChatGPT essay)

We can also see that besides obligation modals and complement to- clauses, students also make use of imperative and inclusive *let*-imperative forms. As Hyland and Jiang (2016) note, these two options impose far less on readers while bringing them into the process of considering and interpreting data as a partner:

(34) **Consider** a general election, in which three parties stand, X, Y, and Z, in a country of 5000 voters.   (Student essay)

(35) Now **let us consider** the friends or family of a boxer… (Student essay)

Personal asides: making a personal connection

By interrupting the ongoing discussion to offer a meta-comment on what has been said, asides comprise another means by which writers claim affinity with readers and suggest a joint exploration of the topic. As we saw in Table 2, there were 91 asides in students' essays, and their interactional effect can be seen in these examples:

(36) all of which seems rather un-DEMOCRATIC **(and I use the term in its correct meaning)** that is, taking power away from the people.

(Student essay)

(37) When this tunnel, which will run under the water between England and France, is finally completed **(hopefully in the near future)** it will be much easier to travel to and from Europe. (Student essay)

Here the material marked by parentheses is neither grammatically or rhetorically related to the surrounding sentence and adds little to the propositional development of the text. Instead, they provide a personal commentary on the ideas, integrated into the text to add depth and an individual touch, helping the reader to relate to the unfolding arguments engaging them more closely in the ideas.

There were, however, no personal asides in the ChatGPT-generated texts, and this is attributable to the design of the model, which prioritises coherence and conciseness (Adeshola & Adepoju, 2023). Asides, are additional comments which are eminently disposable as far as the content is concerned and are perhaps assumed by the AI model as disrupting the flow of an argument with the potential to confuse the reader. The AI model tends to maintain a formal and objective tone in academic writing, ensuring that the argument remains focused and persuasive. In this context, the interjection of an aside is a distraction which might undermine the overall effectiveness and readability of the essay.

It may be thought that because asides are, by definition, personal, they might seem to be a more writer-centred, stance-taking strategy where the writer presents an opinion on the subject. However, asides serve a more reader-aware function than this initial impression suggests. Rather than merely expressing the writer's personal views, they advance the argument by providing an explanatory gloss or example of what precedes it, indicating a more reader-oriented motivation for their use. By interrupting the discourse to address the reader in mid-flow, the writer recognises and responds to an active audience, offering a remark that is predominantly dialogic and interpersonal. The writer includes the audience in the text to strengthen a relationship at that juncture.

Essentially, these diversions draw readers into the discourse to emphasise a personal relationship, uniting writer and reader through honesty or a shared understanding of matters. In examples (38) and (39), for instance, both writers engage readers by explicitly referring to their own reflection and practice, thereby establishing the personal reliability of their discourse:

> (38) Although, there are plans to create computers which can programme themselves, **(which I, personally, feel is a very dangerous idea)** the human brain still very much controls the computer and still the ability to end the existence of computers at any given moment; thankfully, a power computers do not have over humans.   (Student essay)

> (39) Computers have been used as a means of keeping records, they have all but superseded handwritten text, **(in a few decades people may well be faced with a computer screen and keyboard in their General Studies exam),** they are used to transfer money across the globe, even to create artwork and to entertain.   (Student essay)

6. Discussion

This comparative analysis of ChatGPT-generated essays and those written by British students reveals significant differences in the use of engagement markers and rhetorical strategies to involve readers. The findings indicate that the students use overwhelmingly more engagement features, although the AI essays demonstrate a recognition of context and adaptability in mimicking various writing styles and genres.

Our findings align with recent research in the field of AI-generated text analysis. Markey et al. (2024), for example, observed that AI-generated texts exhibit reduced involvement and integration compared to their human counterparts. Similarly, Berber Sardinha (2024) found AI-generated texts to be more informationally dense but less

24

interactive than human-authored texts. Our results also support Jiang and Hyland's (2024) findings that AI-generated responses often lack the stance features characteristic of human writing. Collectively, these studies provide insights that can inform the detection of AI-generated writing in educational settings, where Scarfe et al. (2024) have demonstrated the increasing sophistication of AI in mimicking certain aspects of academic writing.

The stark differences in engagement we have found here indicates ChatGPT's inability to model an audience or anticipate reader needs. ChatGPT does not operate by guessing the possible occurrence of items at a clause-by-clause level but by calculating probabilities across a massive span of text in its training data, rendering it sluggish to respond to context. The much narrower standard deviations and smaller dispersion of proportions for features in our ChatGPT data (see Table 2) show the relative lack of variance in the LLM's use of these features. This result aligns with those of other studies that have quantitively compared LLM-generated to human-generated writing (e.g., Jiang & Hyland, 2024; Markey et al., 2024). Thus the LLMs' reliance on statistical patterns, rather than an understanding of reader expectations, results in texts that, while coherent and grammatical, lack the interactive and persuasive qualities that characterise successful argument. These are texts which are less effective in building rapport, addressing potential counterarguments, and guiding readers through complex ideas. Our findings, then, underscore the importance of human input in crafting engaging and audience-aware texts, especially in contexts where reader buy-in is paramount.

We are, however, aware of the sensitivity of ChatGPT to the prompts it is given. While specifying a clear context, writer identity, audience, word length and topic, the fact we requested it to generate 'argumentative essays' with a 'persuasive writing style' should have influenced the model's use of certain engagement features. The prompt was designed to mirror the task given to human writers as closely as possible, but these terms may have triggered specific patterns in ChatGPT's output. We have no evidence

for this, but, for instance, the term 'argumentative' might have encouraged the use of directives and appeals to shared knowledge, while 'persuasive' could have prompted more reader mentions. However, the significantly lower overall use of engagement features in ChatGPT's outputs compared to human-written ones suggests that the LLM's ability to incorporate these elements remains limited. Future research could explore how variations in prompt wording affect ChatGPT's use of engagement features, providing deeper insights into the relationship between prompt design and AI-generated text characteristics.

7. Conclusion

Although providing important textual evidence of the rhetorical differences between ChatGPT and human writing, one shortcoming of our study is a focus of interactional elements of academic writing, a feature of argument which the programme might be expected to have limitations. We are also aware that undergraduate students are not expert writers and might potentially overuse engagement markers while the calibre of the data used to train ChatGPT is a constraint on its responses. Although the model is trained on a sizable amount of text data, this data may only be broadly indicative of how language is used in this context as the training data may be skewed towards certain registers, demographics or subject areas.

It must be said, however, that we were very impressed by the ability of the large language model to generate a series of extended and coherent responses to the prompts we gave it, and by the statistical procedure it uses to organise and present points in a logical sequence. Nevertheless, our findings indicate that ChatGPT is less adept at injecting the text with a personal touch. It still lacks the ability to adopt a strong perspective on a topic and to engage in persuasive interactions to carry it through, thus neglecting aspects of argument that are highly valued in academic writing. This takes nothing away from our positive assessment of the essays it generated nor are we undervaluing the obvious power and affordances of ChatGPT for writing assistance.

More broadly, our findings extend Hyland's model of engagement in academic discourse. While ChatGPT-generated texts contained the categories of engagement proposed by Hyland, their frequency and distribution differ significantly from human-written texts. A striking example of this difference is the near-absence of personal asides in ChatGPT essays. This finding serves as a clear counterpoint to any tendency to anthropomorphize or "personify" the model. Despite generating coherent text, ChatGPT does not spontaneously insert personal comments or digressions as a human writer might, underscoring its nature as a statistical model rather than a sentient entity. Similarly, the limited use of questions in ChatGPT essays further demonstrates that certain engagement strategies, which Hyland identifies as important for reader involvement, do not figure significantly in the model's output. Equally, the relative reliance on knowledge appeals in ChatGPT texts suggests that it prioritises different engagement strategies. So while Hyland's engagement model is a good yardstick by which to assess the LLM's ability to produce human-like arguments, it may be necessary to modify the categories when seeking to characterise AI-generated content.

Overall, this study underscores the importance of rhetorical engagement in crafting effective argumentative essays. While ChatGPT shows potential in generating coherent and contextually appropriate essays, these students exhibited more ability in employing a greater number and range of rhetorical strategies to engage readers and create a more interactive and persuasive discourse. The findings of this study have several important pedagogical implications. First, the significant difference in the use of engagement markers between student and ChatGPT-generated essays highlights the necessity of explicitly teaching rhetorical strategies in academic writing courses. Teachers should place greater emphasis on instructing students about the use of these engagement features for by understanding and using these rhetorical devices, students can enhance the effectiveness of their academic writing skills.

Second, integrating AI tools like ChatGPT in the classroom could serve as a valuable teaching aid. By analysing AI-generated texts and comparing them with human-created texts, students can learn to identify the strengths and weaknesses of using engagement markers. This comparative approach allows students to critically assess AI-generated content and understand the nuances of human-authored texts. Furthermore, students can be encouraged to improve upon AI-generated drafts by adding personalized and contextually appropriate rhetorical features, thereby developing their skills in crafting more effective academic arguments.

Lastly, and perhaps most importantly, discussions of ChatGPT texts cannot avoid discussion of the ethical implications and potential limitations of relying on AI-generated content. Teachers should guide students in understanding that while AI can assist in generating ideas and drafting texts, crucial areas of academic writing such as critical thinking, personal engagement, and creativity remain, for the moment, outside of the abilities of LLMs. This balanced approach ensures that students can harness the benefits of AI while still prioritising the development of their unique voice and rhetorical skills in academic writing.

References

Adeshola, I., & Adepoju, A. P. (2023). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments*, 1–14. https://doi.org/10.1080/10494820.2023.2253858

Anthony, L. (2014). *TagAnt* (Version 1.2.0) [Computer software]. Waseda University. http://www.antlab.sci.waseda.ac.jp

Anthony, L. (2022). *AntConc (Version 4.0.11)* [Computer software]. https://www.laurenceanthony.net/software

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, 100745.

Berber Sardinha, T. (2024). AI-generated vs human-authored texts: A multidimensional comparison. *Applied Corpus Linguistics*, *4*(1), 100083.

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.

Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: a global perspective. *Library Hi Tech.* Advance online publication. https://doi.org/10.1108/LHT-05-2023-0200

Byrd, A. (2023). Truth-telling: Critical inquiries on LLMs and the corpus texts that train them. *Composition Studies*, *51*(1), 135-142.

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), 100068.

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.

Gabrielatos, C. (2018). Keyness analysis: nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review* (pp. 225-258). Routledge.

Gallagher, J. R. (2023). Lessons learned from machine learning researchers about the terms "artificial intelligence" and "machine learning". *Composition Studies*, *51*(1), 149-154.

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, *26*(2), 5-24.

Gong, Z., Liu, Y., & Liu, Y. (2024). A comparative study of research questions written by L1 English authors and Chinese EFL scholars. *Journal of English for Academic Purposes*, *69*, 101383. https://doi.org/10.1016/j.jeap.2024.101383

Hyland, K. (2001). Bringing in the reader: Addressee features in academic articles. *Written Communication*, *18*(4), 549–574.

Hyland, K. (2002). What do they mean? Questions in academic writing. *TEXT*, *22*(4), 529–557. https://doi.org/10.1515/text.2002.021

Hyland, K. (2004). Patterns of engagement: dialogic features and L2 undergraduate writing. In L. Ravelli & R. A. Ellis (Eds.), *Open linguistics series. Analysing academic writing: Contextualized frameworks* (pp. 5–23). Continuum.

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

Hyland, K. (2005b). Representing readers in writing: Student and expert practices. *Linguistics and Education*, *16*(4), 363–377. https://doi.org/10.1016/j.linged.2006.05.002

Hyland, K. (2005c). Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, *7*(2), 173–192. https://doi.org/10.1177/1461445605050365

Hyland, K. (2006). Representing readers in writing: student and expert practices. *Linguistics and Education.* 16: 363-377

Hyland, K. (2010). Constructing proximity: Relating to readers in popular and professional science. *Journal of English for Academic Purposes*, *9*(2), 116–127. https://doi.org/10.1016/j.jeap.2010.02.003

Hyland, K., & Jiang, F. (2016). "We must conclude that…": A diachronic study of academic engagement. *Journal of English for Academic Purposes*, *24*, 29–42.

Hyland, K., & Jiang, F. (. (2017). Is academic writing becoming more informal? *English for Specific Purposes*, *45*, 40–51.

Hyland, K., & Jiang, F. (2023). Interaction in written texts: A bibliometric study of published research. *Studies in Second Language Learning and Teaching*, *13*(4), 903–924. https://doi.org/10.14746/ssllt.40220

Hyland, K. & Zou, H. (2022). Pithy persuasion: Engagement in 3-minute theses. *Applied Linguistics,* 43, 1, pp 21–44

Ingley, S. J., & Pack, A. (2023). Leveraging AI tools to develop the writer rather than the writing. *Trends in Ecology & Evolution*, *38*(9), 785–787. https://doi.org/10.1016/j.tree.2023.05.007

Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, *55*(1), 48–63. https://doi.org/10.1080/15391523.2022.2142873

Jiang, F., & Hyland, K. (2024). Does ChatGPT Argue Like Students? Bundles in Argumentative Essays. *Applied Linguistics*, advance access.

Jiang, F., & Ma, X. (2018). 'As we can see': Reader engagement in PhD candidature confirmation reports. *Journal of English for Academic Purposes*, *35*, 1–15. https://doi.org/10.1016/j.jeap.2018.05.003

Kabir, A. A. (2022). *Learn ChatGPT: The future of learning*. Self-published.

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., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 1–9. https://doi.org/10.1016/j.lindif.2023.102274

Kohnke, L., Moorhouse, B. L., & Di Zou (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Koubaa, A. (2023): GPT-4 vs. GPT-3.5: A concise showdown. TechRxiv. Preprint.https://doi.org/10.36227/techrxiv.22312330.v1

Koutsantoni, D. (2004). Attitude, certainty and allusions to common knowledge in scientific research articles. *Journal of English for Academic Purposes*, *3*(2), 163–182. https://doi.org/10.1016/j.jeap.2003.08.001

Kumar, D. V. (2023). *How "ChatGPT" seems to look behind: Working of ChatGPT*. https://www.kaggle.com/discussions/general/373463

Lafuente‐Millán, E. (2014). Reader engagement across cultures, languages and contexts of publication in business research articles. *International Journal of Applied Linguistics*, *24*(2), 201–223. https://doi.org/10.1111/ijal.12019

Laquintano, T., Schnitzler, C., & Vee, A. (2023). TextGenEd: Teaching with text generation technologies. WAC Clearinghouse.

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

Liu, Z., A. Xu, M. Zhang, J. Mahmud and V. Sinha. 2017. 'Fostering user engagement: Rhetorical devices for applause generation learnt from ted talks,' *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*. ICWSM: 592-595.

Markey, B., Brown, D. W., Laudenbach, M., & Kohler, A. (2024). Dense and disconnected: Analyzing the sedimented style of ChatGPT-generated text at scale. *Written Communication*, 1–30.

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.

McGrath, L., & Kuteeva, M. (2012). Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices. *English for Specific Purposes*, *31*(3), 161–173. https://doi.org/10.1016/j.esp.2011.11.002

Milano, E. (2023). *How to ask ChatGPT anything and get the right answers: Learn to prompt AI LLMs effectively*. Self-published.

Nordling, L. (2023). How ChatGPT is transforming the postdoc experience. *Nature*, *622*(7983), 655–657. https://doi.org/10.1038/d41586-023-03235-8

OpenAI. (2023). *ChatGPT: Optimizing language models for dialogue.* https://openai.com/blog/chatgpt.

Pack, A., & Maloney, J. (2023). Using generative artificial intelligence for language education research: Insights from using OpenAI's ChatGPT. *TESOL Quarterly*, *57*(4), 1571–1582. https://doi.org/10.1002/tesq.3253

Park, D. B. (1982). The meanings of "audience". *College English*, *44*(3), 247–257.

Rayson, P. (2016). Log-likelihood spreadsheet. http://ucrel.lancs.ac.uk/llwizard.html

Revell, T., Yeadon, W., Cahilly-Bretzin, G., Clarke, I., Manning, G., Jones, J., Mulley, C., Pascual, R., Bradley, N., Thomas, D., & Leneghan, F. (2023). ChatGPT versus human essayists: An exploration of the impact of artificial intelligence for authorship and academic integrity in the humanities. Advance online publication. https://doi.org/10.21203/rs.3.rs-3483059/v1

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, *6*(1), 1–22. https://doi.org/10.37074/jalt.2023.6.1.9

Satariano, A., & Kang, C. (2023). How nations are losing a global race to tackle AI's harms. *The New York Times*. Retrieved from https://www.nytimes.com/2023/12/06/technology/ai-regulation-policies.html?searchResultPosition=3

Santurkar, S., Durmus, E., Ladhak, F., Lee, C., Liang, P., & Hashimoto, T. (2023). Whose opinions do language models reflect? In A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, & J. Scarlett (Eds.), *Proceedings of the 40th International Conference on Machine Learning* (Vol. 202, pp. 29971–30004). JMLR.

Sarrion, E. (2023). *Exploring the power of ChatGPT*. Apress. https://doi.org/10.1007/978-1-4842-9529-8

Scarfe, P., Watcham, K., Clarke, A., & Roesch, E. (2024). A real-world test of artificial intelligence infiltration of a university examinations system: A "Turing Test" case study. *PloS One*, *19*(6), e0305354.

Shahriari, H., & Shadloo, F. (2019). Interaction in argumentative essays: The case of engagement. *Discourse and Interaction*, *12*(1), 96–110. https://www.ceeol.com/search/article-detail?id=834675

Baum, J., & Villasenor, J. (2023). The politics of AI: ChatGPT and political bias.

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

Thompson, G. (2001). Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics*, *22*(1), 58–78.

Wilson, F., Child, S., & Suto, I. (2017). Assessing the transition between school and university: Differences in assessment between A level and university in English. *Arts and Humanities in Higher Education*, *16*(2), 188–208.

Wolfram, S. (2023). *What is ChatGPT doing and why does it work?* Self-published.

Yoon, H.-J. (2021). Interactions in EFL argumentative writing: effects of topic, L1 background, and L2 proficiency on interactional metadiscourse. *Reading and Writing*, *34*(3), 705–725. https://doi.org/10.1007/s11145-020-10085-7

Zhang, M., Press, O., Merrill, W., Liu, A., & Smith, N. A. (2023). How language model hallucinations can snowball. *arXiv Preprint. arXiv:2305.13534*.

Zou, H. & Hyland, K. (2020). "Think about how fascinating this is": Engagement in academic blogs across disciplines. *Journal of English for Academic Purposes*. 43. https://doi.org/10.1016/j.jeap.2019.100809

Appendix Engagement features

(1) Reader mentions

your

you

one's

the reader

we

us

our

reader

(2) Questions

?

(3) Appeals to shared knowledge

apparently

as a rule

common

commonly

conventional

conventionally

established

familiar

normally

obvious

obviously

of course

prevailing

prevalent

traditional

traditionally

typical

typically

usual

routinely

(4) Directives

add

allow

analyse

analyze

apply

arrange

assess

calculate

choose

classify

compare

connect

consult

contrast

define

demonstrate

determine

do not

develop

employ

ensure

estimate

evaluate	set
follow	should
go	show
have to	suppose
review	state
increase	think of
input	turn
insert	use
integrate	take
key	consider
let us	find
look at	imagine
mark	let
measure	let's
mount	note
must	notice
need to	assume
ought	think about
observe	recall
order	remember
pay	let us
picture	let's
prepare	let
recover	need to
refer	should
regard	ought to
remember	do not
remove	have to
see	must
select	has to

(regular expression query)

it is adj. to V.   **(5) Asides**

it_PP\sis_VBZ\s\w*_JJ\sto_TO\s\w*_V   incidentally

V   by the way

imperatives   (

(\(_\(\s|._SENT\s)\w*_VV   —