

# A Corpus-Based Analysis of Epistemic Stance in AI-Generated Instructional Content

Rashad Ahmed

Jacksonville State University

## Abstract

AI systems generate educational content, yet their rhetorical characteristics remain insufficiently understood. This study examines epistemic stance construction through hedging and boosting devices in AI-generated discourse based on one hundred fifty AI-generated texts sampled from educational, professional, and conversational domains. Using a systematic corpus-based mixed-methods analysis that combined quantitative frequency measurements with qualitative functional interpretation, this investigation reveals significant asymmetry in rhetorical positioning. Educational texts displayed a mean hedging frequency substantially higher than booster deployment, creating a hedge-to-booster ratio that far exceeds patterns documented in human pedagogical discourse. Qualitative analysis identified numerous instances of inappropriate hedging in foundational content where confident presentation would better support learning, alongside relatively sparse and structurally formulaic use of boosters. These findings suggest that AI systems overgeneralize cautious patterns learned from training data to contexts that call for assertive instructional guidance. They also show critical limitations in current AI-based tools and establish frameworks for developing AI-mediated learning that supports rather than undermines knowledge construction in digital communities.

**Keywords:** artificial intelligence, LLMs, ChatGPT, hedging, boosting, epistemic stance, corpus linguistics

- **Received:** September 8, 2025
- **Accepted:** December 14, 2025
- **Published:** December 16, 2025

DOI: 10.56540/jesaf.v4i2.125

---

To cite this article (APA):

Ahmed, R. (2025). Ahmed, R. (2025). A Corpus-Based Analysis of Epistemic Stance in AI-Generated Instructional Content. *Journal of English Studies in Arabia Felix*, 4 (2), 80-97. DOI: 10.56540/jesaf.v4i2.125

## Introduction

AI systems now occupy a central position in educational ecosystems, where they generate explanations, feedback, assessments, and learning resources across formal, workplace, and informal learning contexts (Ouyang & Jiao, 2021; Yan et al., 2024). Recent systematic reviews document rapid growth in research and practice around generative AI and large language models in education, including applications for feedback generation, question and task design, tutoring, assessment, and content creation (Garzon et al., 2025; Lucas et al., 2024; Nzenwata et al., 2024). Complementing these more recent syntheses, a decade long review of AI in education from 2010 to 2020 maps the proliferation of intelligent tutoring, analytics, and personalized learning applications across formal and informal settings and highlights both opportunities and persistent challenges for pedagogy and assessment (Zhai et al., 2021). At the same time, these reviews foreground concerns about opacity, reliability, and misalignment between AI output and pedagogical intent (Saleh et al., 2025; Yan et al., 2024).

Unlike human teachers who calibrate certainty and tentativeness based on expertise, audience, and context, large language models operate through pattern matching over massive training corpora. This can yield stable stylistic tendencies that are not tuned to pedagogical needs. Initial studies of AI-generated essays and explanations show that such systems differ systematically from human writers in their use of hedging, boosters, and engagement markers, with some work finding more hedging and fewer engagement resources in AI texts and others showing overconfident or decontextualized claims (Almulla, 2025; Mo & Crosthwaite, 2025). These patterns raise concerns about whether AI systems present foundational knowledge with sufficient clarity and confidence and whether they signal uncertainty appropriately in areas where knowledge is genuinely contested.

Central to these issues is epistemic stance, understood as the expression of commitment, certainty, and attitude toward propositional content. Research in applied linguistics and rhetoric has long emphasized that epistemic stance in educational discourse is not a peripheral feature but a core resource for constructing teacher authority, guiding learner interpretation, and balancing confidence with openness to inquiry (Garrison, 2011). Two key families of epistemic markers are hedges and boosters. Hedges reduce the force of propositions and mark tentativeness through modal verbs, epistemic adverbs, predicates, and phrasal constructions (Hyland, 1998; Crompton, 1997). Boosters strengthen propositions and signal high commitment through intensifiers, assertive verbs, and emphatic constructions (Myers, 1989). In effective pedagogical discourse, these resources are deployed in a calibrated way: cautious and negotiable where knowledge is under development, confident and assertive where knowledge is foundational and well established.

Corpus studies of academic and pedagogical writing indicate that hedging generally outnumbers boosting, but that the balance between the two is more moderate and varies by genre and communicative purpose, including scientific articles, textbooks, and psychology books (Karami & Lohran Poor, 2020; Myers, 1989). Research articles and high-stakes scholarly genres tend to hedge more heavily than textbooks and instructional materials, where authors often present core concepts more assertively to support learner confidence (Hyland, 2005). When AI systems trained largely on research articles, web discourse, and user-generated content generate instructional texts, they may import hedging patterns suited to expert debate into contexts where novice learners need clearer epistemic guidance.

A growing body of work now examines stance and style in AI generated writing and help interactions. Studies of AI mediated help with dialogues, and tutoring responses show that large language models often rely on formulaic politeness and indirectness that can blur the main instructional point or dilute feedback (Ranade et al., 2025). Ranade et al. also demonstrate that human in the loop prompt design, informed by rhetorical analysis, can steer models toward more useful and context appropriate responses, yet they do not quantify how such prompts shape the balance between hedging and boosting. These findings reinforce the need for corpus-based

analyses that link observable stance patterns in AI outputs to their functional appropriateness in instructional contexts.

This study addresses this gap by conducting a corpus-based analysis of hedging and boosting in AI-generated texts across educational, professional, and conversational domains. Drawing on corpus methods and epistemic stance theory, it is deemed important to examine how AI systems deploy hedges and boosters, how these patterns differ across domains, and how epistemic positioning in educational outputs aligns with principles of effective pedagogical communication. This study therefore aims to quantify hedging and boosting frequencies and ratios in AI-generated texts across educational, professional, and conversational domains. It also examines which specific hedging and boosting devices dominate these texts and how they cluster.

### Research Questions

This study is guided by three research questions:

**RQ1:** What patterns characterize the deployment of hedging and boosting devices in AI-generated content across educational, professional, and conversational domains?

**RQ2:** How do epistemic marker frequencies and distributions in AI-generated educational content compare with patterns documented in human pedagogical discourse?

**RQ3:** In what contexts does AI epistemic positioning align with or diverge from principles of effective pedagogical communication?

## Literature Review

### Corpus Linguistic Approaches to Discourse Analysis

Corpus linguistics offers a well-developed toolkit for identifying systematic patterns in language use by combining quantitative frequency analysis with qualitative interpretation of function in context. Biber and Conrad (2019) outline corpus-based approaches that link lexicogrammatical features with communicative functions across registers and genres, emphasizing the importance of representative sampling, clear annotation procedures, and context-sensitive interpretation. McEnery and Hardie (2012) similarly argue that corpus methods can reveal subtle yet robust discourse patterns that are not apparent in individual texts, particularly when investigating stance, evaluation, and interpersonal meaning. These methods have been widely used to examine epistemic markers such as modal verbs, stance adverbs, and metadiscourse across academic genres, textbooks, and digital communication (Karami & Lohran Poor, 2020).

More recently, corpus techniques have been applied to AI-generated texts, including studies that compare human and AI essays in terms of hedging, engagement, and lexical density, and work that quantifies stylistic choices across different chatbots and model versions (Almulla, 2025; Mo & Crosthwaite, 2025). This emerging line of research indicates that corpus-based approaches are well suited to capturing stable rhetorical tendencies in large language model outputs. The present study builds on this foundation by applying systematic corpus methods to AI-generated educational discourse specifically, enabling comparison across domains and against established pedagogical norms.

### Epistemic Stance and Pedagogical Discourse

The theoretical foundation for understanding epistemic stance in discourse is rooted in extensive scholarship within applied linguistics and educational communication. Hyland's (1998) works on stance and meta-discourse positions, hedges and boosters as core resources through which writers manage knowledge claims, align with readers, and project an authorial voice. Hedges express tentativeness, possibility, or limitation, while boosters convey conviction and assertiveness. In his later work, Hyland (2005) expanded this model to include meta-discourse, illustrating how epistemic markers function as interactive elements that shape reader

interpretation and foster writer-reader engagement, both critical to effective educational communication.

Myers' (1989) analysis of scientific writing shows that research articles often combine extensive hedging with strategic boosts to navigate between caution and persuasion in scholarly communication. However, his work also demonstrated that hedging strategies are employed differently in scientific research articles versus textbooks, with educational materials requiring greater certainty and assertiveness to support learners' cognitive development. This distinction becomes especially critical in online learning environments, where textual communication must compensate for the lack of face-to-face interaction and immediate clarification.

Crompton (1997) further argues that hedges perform multiple simultaneous functions, including epistemic caution, politeness, and community alignment. Corpus studies have extended these frameworks to textbooks and pedagogical materials. For instance, work on psychology textbooks and disciplinary introductions finds that authors use hedges and boosters to present complex material in ways that remain authoritative while acknowledging uncertainty where appropriate, often with lower hedge density and more targeted booster use than in research articles (Karami & Lohran Poor, 2020). The balance of hedging and boosting in pedagogical discourse is thus tightly linked to epistemic and instructional goals: foundational concepts are usually presented with higher commitment, while contested or rapidly evolving knowledge is framed more cautiously.

### Digital Learning Environments and Epistemic Authority

The rise of online learning communities has reshaped traditional discourse dynamics. The shift to online and blended learning has foregrounded how textual and multimodal resources construct teaching presence and epistemic authority. Garrison's (2011) Community of Inquiry framework emphasized the role of teaching presence in sustaining both cognitive and social dimensions of learning. This framework highlights the importance of balancing authority with collaborative engagement in instructional communication, presenting unique challenges for managing epistemic stance in digital settings. In text-based environments where paralinguistic cues are diminished, linguistic markers of stance, certainty, and engagement carry greater communicative weight.

Kop and Hill's (2008) discussion of networked and connectivist learning emphasizes that effective online pedagogy requires a balance between authoritative guidance and learner-driven exploration. In such contexts, epistemic stance is part of the signaling system that tells learners when they are encountering settled knowledge that can be relied upon and when they are entering zones of uncertainty where critique and inquiry are central. This framework is particularly relevant when evaluating AI-generated educational content, which must navigate these same tensions without the flexibility and contextual awareness of human instructors.

#### AI in Educational Contexts: Emerging Research

Recent syntheses of generative AI in education document rapid expansion of applications and research. Systematic reviews report diverse uses of large language models for tutoring, feedback, assessment, content generation, and student support across higher education, K-12, and professional learning (Garzon et al., 2025; Nzenwata et al., 2024). These reviews also highlight practical and ethical challenges, including transparency, reliability, bias, and the need for human oversight (Saleh et al., 2025; Yan et al., 2024). Within this broader landscape, several strands of work relate directly to epistemic stance.

Studies of AI assisted feedback and tutoring have shown that tools are most effective when they are designed to complement rather than replace teacher judgment, particularly for conceptual understanding and safe application of knowledge (Holstein et al., 2019). Other work has documented that AI-generated explanations may oversimplify or misrepresent scientific findings, raise or lower confidence inappropriately, or gloss over critical details, particularly in high-stakes domains such as medicine (Lucas et al., 2024). These findings suggest that epistemic

stance in AI output is not merely a stylistic issue but one with direct implications for conceptual understanding and safe application of knowledge.

In parallel, research explicitly examining stance and engagement in AI texts has accelerated. Almulla (2025) compares hedging devices and engagement markers in AI-generated and human-written essays and finds that AI texts differ significantly in the distribution and clustering of hedges and engagement resources. Mo and Crosthwaite (2025) analyze stance and engagement features in texts produced by several large language models and report that models tend to favor cautious and non-confrontational formulations, with relatively few strong authorial commitments. Complementing this work, Jiang and Hyland (2025) compared engagement markers in ChatGPT-generated and student-written argumentative essays, finding that ChatGPT essays contained substantially fewer engagement markers, suggesting that AI texts may lack the interpersonal dimension characteristic of human academic discourse. In addition, Ranade et al. (2025) examine rhetorical moves in AI generated explanations and help responses, showing that large language models often combine hedged epistemic claims with authoritative directives, a pattern that can obscure the status of knowledge and the basis for recommendations. Eguchi and Kyle (2023) adds a computational perspective, showing that stance can be reliably identified and measured at scale, which is relevant for evaluating AI outputs.

Despite this progress, the specific problem of epistemic stance in AI generated educational texts has received limited corpus-based attention. Existing studies tend to focus on essays, assessments, or generic conversational outputs rather than on the kinds of short instructional explanations and summaries that learners encounter through educational platforms and AI assistants. Prior work shows that epistemic stance is central to effective pedagogical communication, that corpus methods are well suited to investigating stance at scale, and that AI generated texts have distinctive stance profiles compared with human writing. However, systematic corpus-based analyses of hedging and boosting in AI outputs explicitly designed for educational purposes are still lacking, and little is known about how educational AI outputs compare with AI discourse in other domains. Accordingly, this study quantifies hedging and boosting frequencies and ratios in AI generated texts across educational, professional, and conversational domains. It examines which specific hedging and boosting devices dominate these texts and how they cluster. It also evaluates the functional appropriateness of epistemic positioning in educational texts, with particular attention to differences between foundational content and contested content.

## Methodology

This investigation employed systematic corpus-based discourse analysis, combining quantitative and qualitative methods. The design followed established corpus linguistic procedures for sampling, annotation, and analysis, adapted to focus on epistemic stance markers in AI-generated discourse (Biber & Conrad, 2019). The quantitative component measured frequencies of hedging and boosting devices and compared distributions across domains. The qualitative component examined whether epistemic positioning aligned with pedagogical appropriateness in context.

### Corpus Construction and Sampling

The corpus comprised one hundred fifty AI-generated texts collected between March and May 2024. To capture variation across AI applications, the researcher used stratified sampling of three domains: educational, professional, and conversational. Educational texts ( $n = 50$ ) included AI-generated explanations, summaries, and instructional responses from widely used learning platforms that integrate conversational agents and explanation features. Professional texts ( $n = 50$ ) consisted of AI-generated content from corporate training environments, including onboarding assistants, automated feedback systems, and productivity-enhancing tools deployed in professional development contexts. Conversational texts ( $n = 50$ ) were sampled from general-

purpose conversational AI systems responding to user queries about educational content in informal settings.

Inclusion criteria required that each text be generated by an AI system in response to a user prompt, contain informational or instructional content relevant to learning, and fall between 100 to 300 words to ensure comparable length and sufficient context. This length range was selected to maintain consistency across samples and to allow for meaningful analysis of epistemic markers within compact discourse units while representing typical response lengths in educational AI applications. Within each domain, sampling covered multiple subject areas and difficulty levels, including mathematics, science, history, and language arts. Each text was anonymized and catalogued with metadata indicating domain, platform type, subject area, and collection date, enabling contextual analysis and cross-domain comparison.

The AI systems responsible for generating the texts included widely used large language models such as ChatGPT (GPT-3.5 and GPT-4 variants) and Google Gemini, as well as proprietary AI tools embedded within specific learning platforms. Representative prompts across domains included educational prompts such as "Explain how photosynthesis works" or "What is the Pythagorean theorem and how do I use it"; professional prompts such as "Describe best practices for conducting performance reviews" or "Explain the company's data security protocols"; and conversational prompts such as "Help me understand climate change" or "What should I know about investing in stocks."

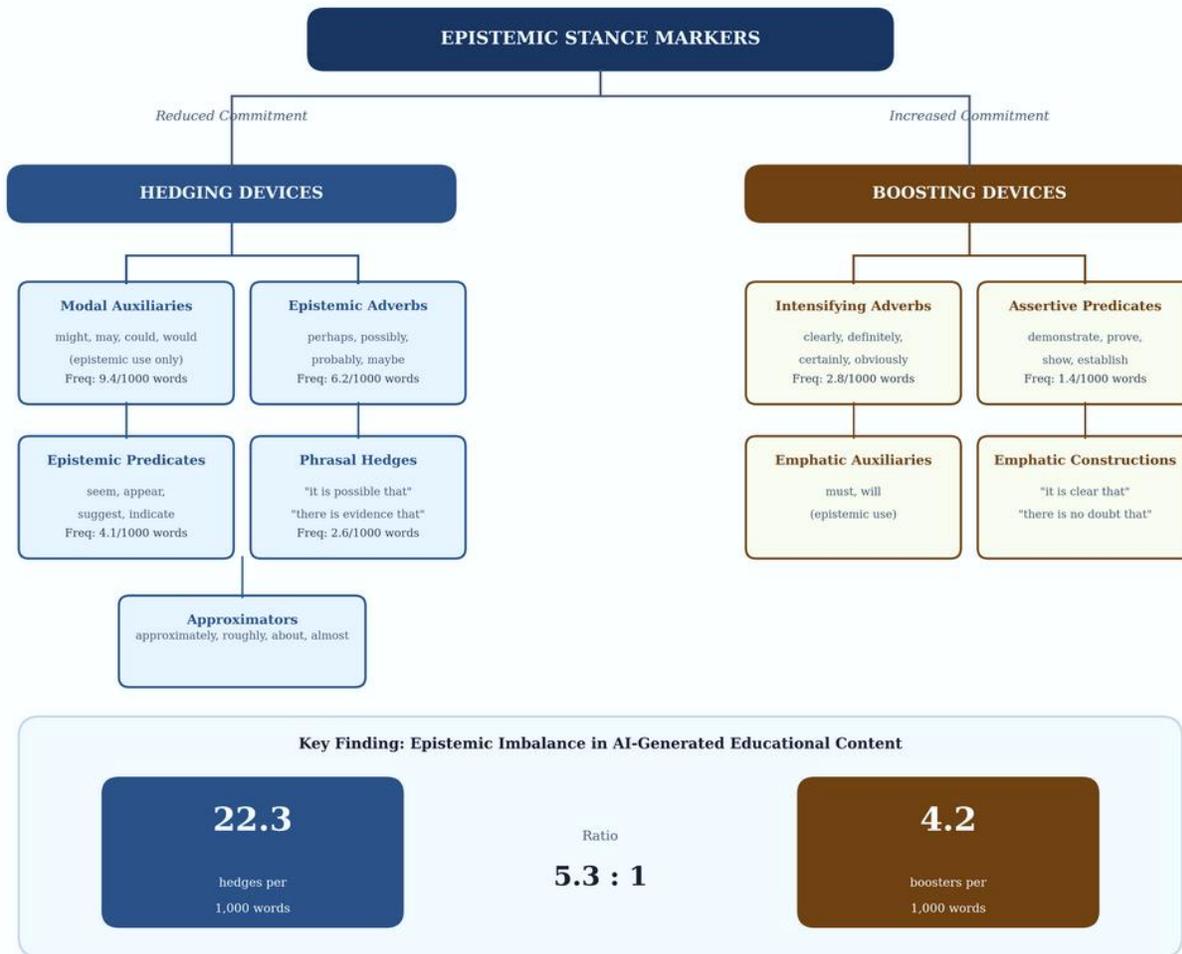
### Analytical Framework

The study employed an annotation scheme grounded in Hyland's (2005) stance model and expanded it with additional categories drawn from recent corpus-based investigations of hedging and boosting in both human and AI texts. The hedging inventory included modal verbs such as *might*, *may*, *could*, and *would* when used epistemically; epistemic adverbs such as *perhaps*, *possibly*, *probably*, and *maybe*; epistemic predicates such as *seem*, *appear*, *suggest*, and *indicate*; phrasal hedges such as "it is possible that" and "there is evidence that"; and approximators such as *approximately*, *roughly*, and *about*. The boosting inventory included intensifying adverbs such as *clearly*, *definitely*, *certainly*, and *obviously*; assertive predicates such as *demonstrate*, *prove*, *show*, and *establish*; emphatic auxiliaries such as *must* and *will* when used epistemically; and emphatic constructions such as "it is clear that" and "there is no doubt that." Figure 1 provides a visual overview of this classification framework, illustrating the hierarchical relationship between stance marker categories and their associated frequencies in educational texts.

Ambiguous cases were resolved through contextual analysis. For example, "may" expressing permission rather than epistemic possibility was excluded from hedging counts. Two trained coders, both holding graduate degrees in applied linguistics, manually annotated all instances of hedging and boosting in the corpus. A pilot phase on thirty practice texts not included in the main dataset was used to refine the coding manual and establish procedures. Inter-rater reliability on the pilot texts, calculated using Cohen's kappa, reached 0.82, which indicates substantial agreement. Disagreements were resolved through discussion and consensus, and the final coding scheme was then applied to the full corpus.

Frequencies were calculated as counts per one thousand words to permit comparison across texts of different lengths. For each domain, mean frequencies, standard deviations, and 95 percent confidence intervals were computed for hedges and boosters separately. Hedge-to-booster ratios were calculated from domain means; because ratios involve propagated uncertainty from both numerator and denominator, the error bars in Figure 1 represent approximate confidence intervals derived from bootstrap resampling of texts within each domain (1,000 iterations). Chi-square tests were used to examine differences in the distribution of epistemic marker frequencies across domains, and Cramér's *V* was reported as a measure of effect size. Because hedge-to-booster ratios are derived measures, ratio comparisons were interpreted descriptively rather than through inferential tests on the ratios themselves. Post-hoc comparisons

used Bonferroni-adjusted significance thresholds to control multiple comparisons. Frequencies were also calculated for specific device types, including modal auxiliaries, epistemic adverbials, epistemic predicates, phrasal hedges, intensifying adverbs, and assertive predicates. Additionally, clustering of multiple epistemic markers was analyzed within sentences, and collocational analysis was conducted using pointwise mutual information to identify systematic co-occurrence patterns.



<sup>1</sup>Figure 1. Analytical Framework for Epistemic Device Classification

For the qualitative component, sixty texts were selected for detailed analysis, twenty from each domain. Selection employed maximum variation sampling to capture texts displaying diverse hedging and boosting patterns, including unusually high or low hedge-to-booster ratios or particularly dense clustering of epistemic markers. Within these texts, each epistemic marker in its clause and sentence context was examined and coded with its functional appropriateness in relation to content status and pedagogical context. Markers were categorized as appropriate,

<sup>1</sup> Taxonomy of epistemic stance markers adapted from Hyland (2005) with corpus-derived examples. Hedging devices (left branch) signal reduced commitment to propositions, while boosting devices (right branch) signal increased commitment. Frequencies represent mean values from educational texts (n = 50) in the AI-generated corpus).

inappropriate, or ambiguous based on whether their epistemic stance aligned with the epistemic status of the content and the pedagogical goals of communication.

For educational texts, particular attention was given to whether foundational content such as basic mathematical formulas, scientific laws, and widely accepted historical facts was presented with confident language or hedged unnecessarily, and whether uncertain or speculative content was appropriately qualified. Two coders independently evaluated appropriateness for a subset of texts, achieving substantial agreement (Cohen's kappa= 0.79). Analytic memos documented recurring patterns and informed the interpretation of quantitative findings. The qualitative analysis aimed to contextualize the quantitative patterns by examining how epistemic devices functioned in actual instructional contexts and identifying specific cases where stance misalignment might compromise pedagogical effectiveness.

## Results and Discussion

Analysis of epistemic device deployment in AI-generated content revealed systematic patterns demonstrating both current capabilities and fundamental limitations in machine-mediated pedagogical discourse. These patterns exhibited consistent tendencies across educational contexts while showing notable variations reflecting differential sensitivity to pedagogical requirements. Across the corpus, hedging exceeded boosting in all domains, but the extent of this asymmetry differed markedly. Educational texts showed the strongest imbalance, with mean hedging frequency of 22.3 per one thousand words (SD= 8.1, 95% CI [19.1, 25.5]) and mean boosting frequency of 4.2 per one thousand words (SD= 2.4, 95% CI [3.5, 4.9]).

This corresponded to a hedge-to-booster ratio of approximately 5.3 to 1. This imbalance appears markedly stronger than the more moderate patterns reported in human pedagogical discourse and suggests that current systems may adopt hedging and boosting practices that are misaligned with effective instructional communication. These patterns raise the possibility that AI generated explanations might sometimes introduce unnecessary epistemic uncertainty into foundational content, which could make it harder for learners to distinguish well established knowledge from genuinely contestable claims.

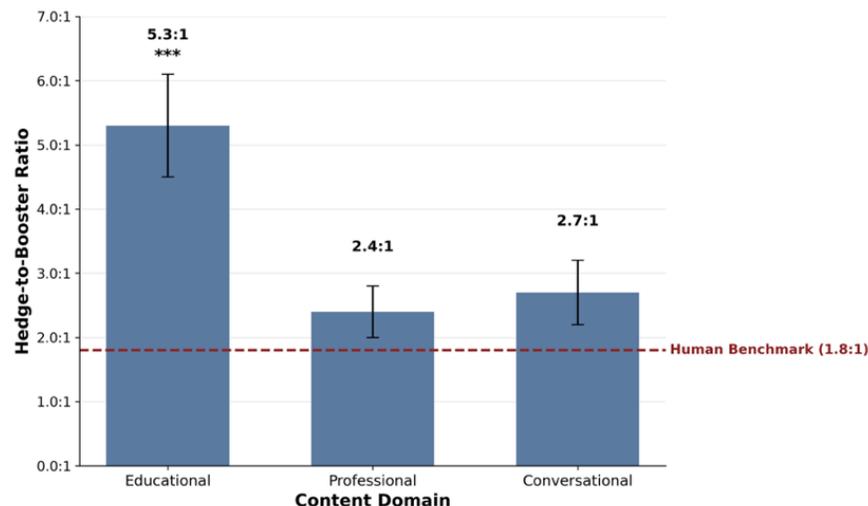


Figure 2. *Hedge-to-Booster Ratios Across Content Domains*

Bars in figure 2 show mean hedge-to-booster ratios for educational, professional, and conversational AI-generated texts in this corpus, with 95 percent confidence intervals. Each bar represents 50 texts per domain. The horizontal dashed line at 1.8 to 1 indicates an approximate hedge-to-booster ratio derived from published counts in human-authored pedagogical texts, specifically research articles and psychology textbooks analyzed by Myers (1989) and Karami and

Lohran Poor (2020). Because these genres differ from AI-generated instructional explanations in audience, length, and interactive affordances, this benchmark is used as a heuristic comparison rather than a definitive standard. Asterisks indicate chi-square comparisons for the educational domain, \*\*\*  $p < .001$ .

Professional texts displayed hedging at 16.4 per one thousand words ( $SD = 6.2$ , 95% CI [14.6, 18.2]) and boosting at 6.8 per one thousand words ( $SD = 3.8$ , 95% CI [5.7, 7.9]), yielding a ratio of about 2.4 to 1. This more moderate pattern approaches ranges observed in effective educational communication. Conversational texts had hedging at 17.4 per one thousand words ( $SD = 9.5$ , 95% CI [14.7, 20.1]) and boosting at 6.4 per one thousand words ( $SD = 4.1$ , 95% CI [5.2, 7.6]), for a ratio of about 2.7 to 1. The high standard deviations, particularly in conversational texts, indicate substantial within-category variation. This variability likely reflects differences in prompt specificity, topic complexity, and AI system versions. Further analysis revealed that GPT-4 output showed slightly lower hedge-to-booster ratios (mean = 4.1 to 1) compared to GPT-3.5 outputs (mean = 5.8 to 1) in educational contexts, suggesting some improvement in newer models, though both remained substantially above the approximate hedge-to-booster ratios reported for human-authored pedagogical texts (around 1.5 to 2.0 to 1 in the studies summarized by Myers, 1989, and Karami & Lohran Poor, 2020). Given the rapid pace of model development, these findings reflect system behavior at the time of data collection (March to May 2024) and may not generalize to subsequent model versions or fine-tuned variants. Hedge use clearly exceeds boosting across all three domains, but the imbalance is most pronounced in educational texts; as shown in Figure 2, educational texts exhibit a hedge-to-booster ratio of 5.3 to 1 compared with 2.4 to 1 in professional texts and 2.7 to 1 in conversational texts.

Chi-square analysis indicated that the distribution of hedges and boosters differed significantly across domains ( $\chi^2 = 47.3$ ,  $p < .001$ , Cramér's  $V = .31$ ), representing a medium effect size by conventional standards (Cohen, 1988). Educational content exhibited the most extreme asymmetry, raising serious concerns about pedagogical appropriateness. Bonferroni-corrected post-hoc pairwise comparisons revealed significant differences between all content type pairs: educational versus professional ( $p$  less than .001), educational versus conversational ( $p$  less than .001), and professional versus conversational ( $p = .012$ ).

Table 1. Frequency Distribution of Epistemic Devices by Content Type (per 1,000 words)

Content Type	Hedges M(SD)	Boosters M(SD)	Ratio	Modal Aux.	Epist. Adv.	Epist. Pred.	Phrasal H.	Intens. Adv.	Assert. Pred.
Educational	22.3 (8.1)	4.2 (2.4)	5.3:1	9.4	6.2	4.1	2.6	2.8	1.4
Professional	16.4 (6.2)	6.8 (3.8)	2.4:1	7.1	4.3	3.2	1.8	4.2	2.6
Conversational	17.4 (9.5)	6.4 (4.1)	2.7:1	8.2	5.1	2.9	1.2	3.8	2.6

Note. Modal Aux. = modal auxiliaries; Epist. Adv. = epistemic adverbials; Epist. Pred. = epistemic predicates; Phrasal H. = phrasal hedges; Intens. Adv. = intensifying adverbs; Assert. Pred. = assertive predicates. Frequencies are per one thousand words.

### Device-Specific Analysis and Contextual Patterns

Modal auxiliaries dominated hedging strategies across educational content types, accounting for approximately two fifths of all hedging instances. Educational texts averaged 9.4 modal hedges per one thousand words. The most frequent modal items were "might" (4.2 per one thousand words), "may" (3.8 per one thousand words), and "could" (3.1 per one thousand words). Qualitative analysis of the intensive sample revealed systematic misapplication of these

devices in educational contexts where confident presentation would better serve learning objectives. Educational content frequently employed modal qualification when explaining well-established procedures, for example "This equation might help calculate area"; scientific principles, for example "Gravity may affect object motion"; and historical facts, for example "World War II could have ended in 1945." Such patterns potentially undermine student confidence in foundational knowledge.

Epistemic adverbials demonstrated moderate deployment with "probably" (2.1 per one thousand words) and "possibly" (1.8 per one thousand words) leading frequency distributions. However, contextual analysis revealed problematic deployment patterns where probability markers appeared in explanations of established conceptual relationships that students need to master confidently. Mathematical procedures received hedging, for example "The formula probably works by..."; while scientific laws were presented uncertainly, for example "Evolution possibly explains species development." Professional and conversational texts showed a similar profile of device types but with lower overall hedging frequency and slightly greater reliance on epistemic predicates and approximators in some contexts.

Table 2. Most Frequent Epistemic Hedges with Contextual Distribution

Hedges	Freq./1000	Educational %	Professional %	Conversational %
might	4.2	31	22	47
may	3.8	28	35	37
could	3.1	25	33	42
probably	2.1	15	41	44
possibly	1.8	33	28	39

Table 3. Most Frequent Booster Devices with Contextual Distribution

Boosters	Freq./1000	Educational %	Professional %	Conversational %
clearly	1.4	18	42	40
certainly	0.9	22	34	44
definitely	0.7	15	38	47
shows	0.6	35	31	34
demonstrates	0.4	52	29	19

*Note.* Percentages indicate distribution across content types for each device.

The cross-domain distribution reveals that educational content uses hedges more frequently in technical contexts ("might," "possibly") while conversational content shows higher use of these same devices in explanatory contexts. Notably, "demonstrates," a strong assertive predicate, appears disproportionately in educational content despite overall low booster frequencies, suggesting recognition of where assertiveness is needed but insufficient deployment.

Among boosters, intensifying adverbials dominate deployment patterns at frequencies substantially lower than those characteristic of effective pedagogical discourse. Intensifying adverbs were the most common boosters across all domains, though much less frequent than modal verbs. Educational texts averaged 2.8 intensifying adverbs per one thousand words. "Clearly" emerged as the most frequent booster at 1.4 per one thousand words, followed by "certainly" at 0.9 per one thousand words and "definitely" at 0.7 per one thousand words. Assertive predicates such as "show" and "demonstrate" were relatively rare in educational texts, appearing at about 1.4 per one thousand words total, with "demonstrates" and "proves" each appearing fewer than 0.5 times per one thousand words. This suggests systematic reluctance to make strong knowledge claims even when pedagogically appropriate. Professional and conversational texts exhibited somewhat higher booster frequencies, particularly for assertive verbs in evaluative or

persuasive passages, though even in these domains, hedges remained more frequent than boosters.

### Clustering Analysis and Collocational Patterns

One of the most distinctive features of the AI corpus was the clustering of multiple epistemic markers within individual clauses and sentences. Analysis of immediate linguistic contexts (within three-word windows) revealed distinctive clustering behaviors differentiating AI from effective human discourse. Hedging devices frequently co-occurred within individual explanations, creating cumulative tentative effects exceeding conventional discourse norms. The researcher observed 127 instances of double hedging sequences such as "might possibly" or "could perhaps" and 89 cases where multiple epistemic adverbs or predicates appeared together, for example "probably somewhat likely" or "it seems to appear that." In addition, phrasal hedges such as "it is possible that" were frequently combined with modal verbs in the same clause. These clusters created a cumulative effect of tentativeness that went beyond what single hedges typically convey in human pedagogical writing.

Collocational analysis using pointwise mutual information (PMI) scores revealed systematic co-occurrence patterns. Calculations were based on observed versus expected co-occurrence frequencies within a five-word span. Hedges in educational texts were strongly associated with key disciplinary terms and foundational concepts. Modal auxiliaries demonstrated strong positive associations with technical terminology (PMI = 4.7, occurring together 3.2 times more often than expected by chance) and foundational concepts (PMI = 5.2, 4.1 times more often than expected), contexts where confident presentation typically supports learning effectiveness. For example, statements of definitional relationships in mathematics and physics were often framed with "may" or "might," and references to well-known historical dates were sometimes prefaced by "probably" or "possibly."

Boosters showed weak associations with established knowledge contexts (PMI = 2.1, only 1.4 times more often than expected), suggesting missed opportunities for appropriate confidence building. In educational texts, boosters appeared disproportionately in concluding or summary sentences, such as closing statements that used "clearly" or "definitely" to restate points, rather than within the central explanatory segments where learners encounter new concepts.

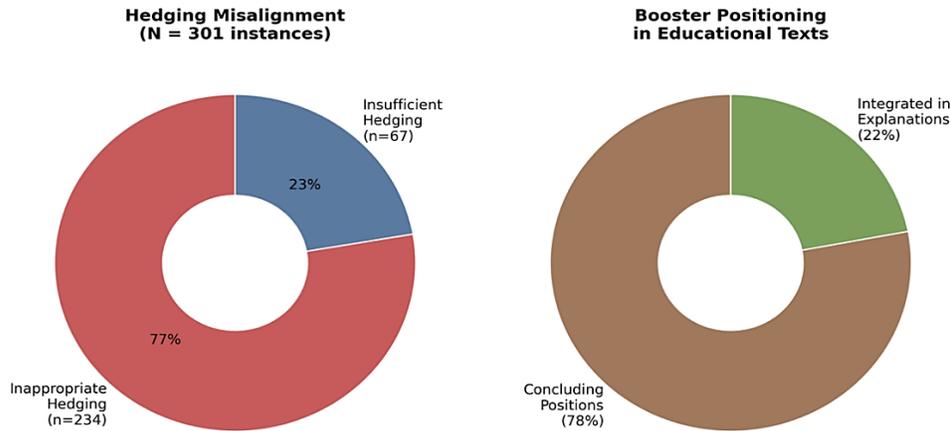
Educational content exhibited distinct problematic collocational patterns where hedging devices co-occurred with fundamental concepts requiring confident presentation. Analysis of the intensive qualitative sample revealed 127 instances where mathematical formulas received modal qualifications, 89 instances where scientific laws were presented tentatively, and 156 instances where historical facts were inappropriately hedged. These patterns demonstrate systematic misalignment between epistemic positioning and content certainty in educational contexts.

### Pedagogical Appropriateness Analysis

Qualitative analysis of contextual appropriateness revealed systematic patterns where epistemic positioning compromised rather than supported educational effectiveness. Qualitative evaluation of the sixty focal texts identified a large number of epistemic markers whose functional role appeared misaligned with pedagogical needs. Educational content addressing established scientific principles frequently employed unnecessary hedging potentially confusing learners about fundamental concepts. Mathematical relationships received inappropriate uncertainty markers, for example "This formula might help..." or "The theorem could apply..."; while well-established historical events were hedged unnecessarily, for example "The American Revolution probably occurred..." or "Shakespeare may have written Hamlet."

Through systematic review of the intensive qualitative sample, coders identified 234 instances of inappropriate hedging in foundational content where confident presentation would better serve pedagogical objectives. These instances were validated through consensus coding and included cases where: established scientific facts were hedged, mathematical procedures

were qualified unnecessarily, historical events of high certainty were presented tentatively, and definitional content received epistemic caution. The imbalance between inappropriate and insufficient hedging suggests that the dominant pattern in these AI outputs is overextension of cautious stance rather than occasional lapses into unjustified certainty. Figure 3 illustrates the distribution of epistemic misalignment in educational texts. Inappropriate hedging, where foundational content was presented with unwarranted tentativeness, accounted for 77% of misalignment instances ( $n = 234$ ), while insufficient hedging in genuinely uncertain contexts represented only 23% ( $n = 67$ ). Concurrently, booster deployment showed a formulaic pattern, with 78% of occurrences confined to concluding sentences rather than integrated throughout explanatory sequences.



<sup>2</sup>Figure 3. *Epistemic Marker Misalignment and Booster Positioning in Educational Texts*

Conversely, only 67 instances of insufficient hedging were documented in contexts requiring acknowledgment of genuine uncertainty or ongoing scholarly debate, typically in passages that described speculative theories, emerging research, or contested interpretations in history and the social sciences. In these cases, AI texts used boosters or unqualified declaratives where a degree of caution would have been more appropriate. This asymmetry suggests systematic bias toward excessive caution that may undermine rather than support student learning progression.

The clustered hedging structures contributed to this pattern by amplifying tentativeness around concepts that, in a pedagogical context, would normally be presented as reliable anchors for further learning. The deployment of boosters revealed formulaic rather than strategic positioning. The analysis also showed that boosters in educational texts were deployed in a structurally formulaic way. Approximately four fifths (78%) of booster occurrences in educational texts appeared in final sentences of paragraphs or responses, often in generic formulations such as "this clearly shows that" or "this definitely means that," while the earlier explanatory segments made little use of boosters even when introducing central definitions or principles. Common

<sup>2</sup> Left panel displays the distribution of hedging misalignment instances identified through qualitative analysis of the intensive sample ( $N = 301$ ). Inappropriate hedging refers to cases where foundational content was hedged unnecessarily; insufficient hedging refers to cases where genuinely uncertain content lacked appropriate qualification. Right panel shows the positional distribution of booster devices, with the majority appearing in paragraph-final or response-final positions).

patterns included "clearly shows" and "definitely demonstrates" in research summaries, suggesting recognition of structural positions where boosters typically appear without sophisticated understanding of their pedagogical functions in building student confidence and guiding learning progression.

These quantitative and qualitative findings collectively address the three research questions. RQ1 is addressed through documentation of systematic hedge-to-booster imbalances (5.3 to 1 in educational content, 2.4 to 1 in professional content, and 2.7 to 1 in conversational content) and device-specific frequency patterns showing modal auxiliaries as the dominant hedging strategy and intensifying adverbs as the most common but still infrequent boosting devices. RQ2 is addressed through comparison with established pedagogical discourse norms, revealing that AI educational texts substantially exceed typical hedge-to-booster ratios documented in human textbooks and instructional materials, and through identification of contextual misapplications where AI diverges from effective human strategies. RQ3 is addressed through qualitative analysis revealing 234 instances of inappropriate hedging in foundational content and formulaic booster deployment that compromises pedagogical effectiveness, a misalignment between epistemic positioning and instructional needs.

The findings indicate that current AI systems display a strong and systematic preference for hedging over boosting in educational outputs, with hedge-to-booster ratios that exceed those reported for human writers in comparable genres (Karami & Lohran Poor, 2020; Myers, 1989). From the perspective of stance theory, this suggests that models have internalized hedging as a default strategy for managing risk and error rather than as a carefully calibrated resource tuned to epistemic and interpersonal contexts. The systematic patterns revealed through this comprehensive analysis highlight a fundamental disconnect between AI rhetorical strategies and effective pedagogical communication in online learning communities.

The following discussion interprets these quantitative and qualitative patterns through the lens of stance theory and pedagogical discourse research, examining their implications for understanding how AI systems construct epistemic positioning in educational contexts and what adjustments might better align AI-generated content with principles of effective instructional communication.

### **Theoretical Interpretation and Implications**

These findings resonate with earlier theoretical frameworks established by Hyland (1998, 2005), who emphasized the importance of epistemic stance in academic writing as a means of managing knowledge claims and guiding learner interpretation. Hyland's account of stance and meta-discourse foregrounds the ways in which hedges and boosters reflect not only degrees of certainty but also social relationships, disciplinary norms, and expectations about debate and consensus. The pronounced epistemic imbalances observed, particularly the overuse of hedging and underuse of boosting, appear to reflect limitations in how current AI systems calibrate confidence levels in educational discourse. The corpus patterns reported here imply that AI systems are drawing more heavily on rhetorical models of research-based or cautious expository prose than on models of authoritative instructional discourse. These issues extend beyond stylistic preferences, touching fundamental concerns about knowledge authority, student confidence, and instructional scaffolding in digital education.

The clustering of hedges and their attachment to foundational content reinforce this interpretation: models appear to stack tentative markers in ways that would be unlikely for experienced human educators. The contrast between the present findings and recent work on hedging and engagement in AI essays is also informative. Studies of argumentative essays find that AI-generated texts use hedging somewhat differently from human writers but often with comparable overall density (Almulla, 2025). In our educational corpus, hedge density is more extreme, and the functional misalignments are more pronounced, particularly in how hedges are attached to basic, definitional statements. This suggests that when AI output is framed as

instructional explanation rather than as argument, models may still fall back on general-purpose hedging strategies that are poorly suited to pedagogical goals.

### Interpretation of Hedging Patterns

The excessive hedging found in AI-generated content suggests insufficient differentiation between contexts requiring caution and those demanding confidence. As Crompton (1997) argued, hedging is not merely a marker of uncertainty but a strategic rhetorical tool that must be deployed judiciously in educational contexts to avoid undermining learner trust in well-established knowledge. This study confirms that AI systems frequently hedge even when presenting foundational concepts such as mathematical formulas or historical facts, where clarity and confidence are pedagogically essential. Such patterns risk confusing learners and weakening their conceptual foundations, a concern that echoes Holstein et al, who highlight the importance of tools that support teacher AI complementarity rather than shifting epistemic authority entirely to the system (Holstein et al., 2019).

The clustering of multiple hedging devices within single explanations proves particularly problematic. Rather than using hedges strategically, AI systems appear to employ them additively, resulting in compounded uncertainty that serves no clear pedagogical purpose. This behavior contrasts sharply with the nuanced deployment described by Myers (1989), who showed that effective educational writing uses hedging selectively to balance authority with acknowledgment of complexity. The AI tendency to stack hedges, exemplified by phrases like "might possibly indicate," suggests surface-level mimicry of academic register without functional understanding of rhetorical purpose.

One plausible interpretation, though not directly tested in this study, is that current training paradigms prioritize caution to avoid confident errors, inadvertently producing excessive tentativeness across all content types. If AI systems are optimized to minimize confident falsehoods, they may default to hedging as a protective strategy even when addressing established knowledge. This hypothesis warrants empirical investigation in future research examining training objectives and their rhetorical consequences. The observation that GPT-4 shows marginal improvement over GPT-3.5 (4.1 to 1 versus 5.8 to 1 hedge-to-booster ratios) suggests that newer models may be developing somewhat better calibration, though they remain far from optimal pedagogical standards.

### Interpretation of Booster Patterns

Equally concerning is the limited and formulaic use of boosters. Boosters serve essential functions in reinforcing key ideas and building learner confidence, particularly in online environments where immediate human feedback is unavailable (Garrison, 2011). Yet AI systems tend to reserve assertive language for concluding statements, missing opportunities to scaffold understanding throughout explanatory sequences. This pattern aligns with findings by Holstein et al. (2019), who noted that AI often fails to align rhetorical strategies with instructional goals, leading to miscommunication and reduced learning outcomes.

The weak collocational association between boosters and established knowledge contexts (PMI= 2.1) suggests that AI systems do not recognize where assertiveness serves pedagogical purposes. When foundational concepts require confident presentation to support student mastery, AI-generated texts frequently fail to provide that confidence. This represents a significant limitation in AI's capacity to replicate effective teaching presence, a construct (Garrison, 2011) identifies as central to successful online learning. The finding that boosters appear primarily in concluding formulas (78% of occurrences) rather than integrated throughout explanations suggests that models recognize structural conventions of academic writing without understanding the functional purpose of strategic assertiveness in building learner confidence progressively through instructional sequences.

## Pedagogical Implications

From a pedagogical perspective, the most serious concern is the systematic hedging of foundational content. If AI-generated explanations routinely present basic mathematical relations, scientific laws, or well-established historical facts as tentative possibilities, learners may infer that these concepts are far less stable than they actually are. This can undermine the development of coherent mental models and weaken learners' sense of epistemic security about core disciplinary structures. Teaching presence in online learning involves establishing clear structures of knowledge and guiding learners through progressively more complex understanding (Garrison, 2011). Over-cautious AI output disrupts this process by blurring the distinction between settled knowledge and contested issues.

At the same time, the relatively small number of cases with insufficient hedging (67 instances compared to 234 instances of inappropriate hedging) shows that AI models are also capable of overstatement when dealing with speculative or controversial material. Together, these patterns point to a lack of systematic alignment between epistemic stance and the actual status of knowledge claims. Recent reviews of AI in education stress the importance of human oversight, explicit policy, and professional development to ensure that AI tools support rather than weaken pedagogical quality (Garzon et al., 2025; Yan et al., 2024). The present findings provide a concrete dimension of quality that educators and instructional designers can monitor: the calibration of hedging and boosting in AI-generated materials.

These rhetorical patterns may influence not only how students learn specific content but also how they perceive knowledge reliability itself. Persistent exposure to hedged explanations of established facts could cultivate unwarranted uncertainty about foundational knowledge, while insufficient assertiveness in explanatory development may deprive students of the confidence scaffolding that they require. This concern extends beyond individual learning outcomes to broader epistemological questions about how AI shapes learners' understanding of certainty, authority, and knowledge construction in educational contexts.

## Implications for AI Development and Educational Practice

For AI developers, the results identify epistemic stance as a target for model design and post-processing in educational applications. Several complementary strategies are possible. First, training data and fine-tuning procedures for educational systems could incorporate corpora of high-quality instructional texts where hedging and boosting are used in ways that align with pedagogical norms. Such corpora would include textbooks, worked examples, and instructor explanations in which foundational content is stated confidently, and uncertainty is appropriately marked. This approach would require curating specialized training data that represents effective pedagogical communication rather than relying solely on web-scraped corpora that blend research articles, informal discussion, and diverse rhetorical contexts.

Second, knowledge-aware systems could integrate external resources that encode information about the epistemic status of propositions, such as whether a statement concerns a definitional relation, a widely replicated empirical finding, or an active area of research. This information could guide stance selection, prompting models to avoid hedging for definitional content while encouraging more cautious formulations for speculative claims. Such systems might leverage knowledge graphs or expert-annotated databases that classify propositions by their epistemic status within disciplines.

Third, post-generation filters could detect and adjust problematic patterns in output. For example, heuristic rules could flag multiple hedges within a single clause for review, identify hedges attached to canonical formulas or historical facts, or prompt replacement of overly tentative verbs and adverbs with more appropriate alternatives. These interventions would not require deep semantic understanding but could nonetheless reduce the most disruptive mismatches between stance and content. Rule-based systems could be developed to recognize

patterns such as modal verbs co-occurring with mathematical operators or scientific laws and trigger either automatic revision or human review.

For educators and instructional designers, the study highlights the need for careful review of AI-generated content before it is used in teaching. Practical implications include the need for guidelines that encourage teachers to check AI-generated explanations for unnecessary tentativeness in foundational content and for overconfident claims in areas where uncertainty should be made visible. Professional development programs should equip educators to recognize epistemic imbalances and make targeted adjustments. When using AI-generated materials, instructors should be prepared to manually edit hedge-to-booster ratios, removing unnecessary hedges from explanations of established procedures and adding boosters to reinforce key concepts where appropriate.

Institutional adoption of AI in education should be accompanied by quality assurance protocols that assess rhetorical appropriateness alongside content accuracy. Review processes should evaluate whether epistemic positioning aligns with learning objectives and content certainty. Additionally, students may benefit from explicit instruction about how hedges and boosters function in academic discourse, enabling them to critically evaluate AI-generated content and develop their own rhetorical competence. Transparency with learners about AI's involvement in content generation can support appropriate critical awareness.

### Limitations

Several constraints define the scope of these findings. The corpus comprises exclusively English-language texts collected during a three-month period in 2024. The focus on texts between 100 and 300 words excludes examination of epistemic stance in longer instructional sequences. The sampling strategy prioritized breadth across domains rather than depth within specific disciplines, leaving open whether AI systems calibrate stance differently across subject areas or educational levels. While the functional coding schema achieved substantial intercoder reliability (Cohen's kappa of 0.82 for device identification and 0.79 for appropriateness judgments), determinations of pedagogical appropriateness involve interpretive judgment. Additionally, comparison of AI outputs to benchmarks from Myers (1989) and Karami and Lohran Poor (2020) involves cross-genre comparison between contemporary digital materials and print-era texts that differ in audience and purpose. More importantly, this investigation did not incorporate learner reception data. Claims about pedagogical consequences remain theoretically grounded but untested through experimental manipulation or learning outcome assessment. Studies in the future should address these gaps through experimental designs that measure effects of varying epistemic profiles on comprehension, self-efficacy, and knowledge retention. Establishing contemporary benchmarks through analysis of human-authored digital learning content would strengthen validity of observed disparities. Incorporating learner demographic variables and prior knowledge levels would also clarify how different populations interpret epistemic markers in AI-generated educational content.

### Conclusion

This study indicates that, in this corpus, AI-generated educational texts display a pronounced imbalance between hedging and boosting, with substantially more hedges and relatively few boosters compared with patterns reported in human authored pedagogical writing and the other AI domains examined. The issue is not only quantitative but also functional: hedges frequently introduce unnecessary doubt into well-established content, while boosters tend to appear in closing segments rather than supporting explanations throughout. These tendencies suggest that AI-generated instructional texts should not be treated as neutral replacements for human explanations, because the expression of certainty and doubt shapes how learners assess, trust, and organize knowledge. To address this, epistemic stance should be treated as a core dimension of quality, alongside factual accuracy, when evaluating and deploying educational AI systems. Institutions can incorporate stance-sensitive criteria into their review of AI-generated

materials, and teachers need preparation to recognize and revise rhetorically misaligned AI outputs. Future research should examine cross-linguistic patterns, track changes across model versions, evaluate training or prompt engineering strategies for stance calibration, and investigate how learners respond to different hedge-to-booster profiles in instructional texts. As AI becomes more deeply integrated into educational practice, attending to clear and appropriately calibrated epistemic communication will be essential for supporting effective learning.

### Disclosure Statement

I (the author of this paper) hereby declare that research ethics and citing principles have been considered in all the stages of this paper. I take full responsibility for the content of the paper in case of a dispute. I confirm that the manuscript has been created by the author(s) and not an AI tool/ Large Language Model (LLM).

### Conflict of interest:

I know of no conflict of interest associated with this publication.

### Funding:

There has been no significant financial support for this work that could have influenced its outcome.

## References

- Almulla, N. (2025). The use of hedging devices and engagement markers in AI generated and human written essays: A corpus-based comparison. *Open Journal of Modern Linguistics*, 15(5), 754–772. <https://doi.org/10.4236/ojml.2025.155044>
- Biber, D., & Conrad, S. (2019). *Register, genre, and style* (2nd ed.). Cambridge University Press.
- Crompton, P. (1997). Hedging in academic writing: Some theoretical problems. *English for Specific Purposes*, 16(4), 271-287. [https://doi.org/10.1016/S0889-4906\(97\)00007-0](https://doi.org/10.1016/S0889-4906(97)00007-0)
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Eguchi, M., & Kyle, K. (2023). Span identification of epistemic stance-taking in academic written English. In E. Kochmar, J. Burstein, A. Horbach, R. Laarmann-Quante, N. Madnani, A. Tack, V. Yaneva, Z. Yuan, & T. Zesch (Eds.), *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 429-442). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.bea-1.35>
- Garrison, D. R. (2011). *E-learning in the twenty-first century: A framework for research and practice* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203838761>
- Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic review of artificial intelligence in education: Trends, benefits, and challenges. *Multimodal Technologies and Interaction*, 9(8), Article 84. <https://doi.org/10.3390/mti9080084>
- Holstein, K., McLaren, B. M., & Alevan, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics*, 6(2), 27–52. <https://doi.org/10.18608/jla.2019.62.3>
- Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.
- Karami, Y., & Lohran Poor, M. (2020). A comparative corpus-based analysis of hedging, boosting, and self-mention metadiscourse markers in Persian and English psychology books. *International Journal of Research in English Education*, 5(2), 13-28.
- Jiang, F., & Hyland, K. (2025). Does ChatGPT write like a student? Engagement markers in argumentative essays. *Written Communication*. Advance online publication. <https://doi.org/10.1177/07410883251328311>

- Kop, R., & Hill, A. (2008). Connectivism: Learning theory of the future or vestige of the past. *International Review of Research in Open and Distance Learning*, 9(3), 1-13. <https://doi.org/10.19173/irrodl.v9i3.523>
- Lucas, H. C., Upperman, J. S., & Robinson, J. R. (2024). A systematic review of large language models and their implications in medical education. *Medical education*, 58(11), 1276-1285.
- McEney, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- Mo, Z., & Crosthwaite, P. (2025). Exploring the affordances of generative AI large language models for stance and engagement in academic writing. *Journal of English for Academic Purposes*, 75, 101499. <https://doi.org/10.1016/j.jeap.2025.101499>
- Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, 10(1), 1-35. <https://doi.org/10.1093/applin/10.1.1>
- Nzenwata, U. J., Barn-Nzekwe, C. L., Ojelabi, E. O., Oduware, O., Atalor, P. E., Yisau, Y., Adeyela, A. T., Nwanguma, E. C., Emokiniovo, E., & Osisanya, O. A. (2024). A systematic review of generative AI in education. *Journal of Computer Sciences and Applications*, 12(1), 25–30. <https://doi.org/10.12691/jcsa-12-1-4>
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers & Education: Artificial Intelligence*, 2, Article 100020. <https://doi.org/10.1016/j.caeai.2021.100020>
- Ranade, N., Saravia, M., & Johri, A. (2025). Using rhetorical strategies to design prompts: A human-in-the-loop approach to make AI useful. *AI & Society*, 40, 711-732. <https://doi.org/10.1007/s00146-024-01905-3>
- Saleh, Y., Abu Talib, M., Nasir, O., & Dakalbab, F. (2025). Evaluating large language models: A systematic review of efficiency, applications, and future directions. *Frontiers in Computer Science*, 7, Article 1523699. <https://doi.org/10.3389/fcomp.2025.1523699>
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90-112. <https://doi.org/10.1111/bjet.13370>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... (2021). A review of artificial intelligence (AI) in education from 2010-2020. *Complexity*, Volume 2021, Article ID 8812542. <https://doi.org/10.1155/2021/8812542>

### Author

**Dr. Rashad Ahmed** is an associate professor of TESOL and Linguistics. His research covers a wide range of areas, including Artificial Intelligence in language learning, Computer Assisted Language Learning (CALL), Language Assessment, Digital Literacy, TESOL Methods, Sociolinguistics, and First-Year Composition.