التنوع الشرطي بالمحفزات في النصوص الأكاديمية الإنجليزية: تحليل أسلوبي وبراغماتي للكتابة المولدة بالذكاء الاصطناعي

المؤلف: م. د. وعد داود ناصر

الانتماء: قسم اللغة الإنجليزية، كلية التربية، جامعة سومر، ذي قار، العراق.

المجال: اللغويات التطبيقية / اللغويات الكوربوسية / اللغة الإنجليزية لأغراض أكاديمية

البريد الإلكتروني Email : wjj@gmail.comWadd٧٩٦

الكلمات المفتاحية: الكتابة المولدة بالذكاء الاصطناعي؛ التنوع الشرطي بالطلب؛ الإنجليزية الأكاديمية؛ الأسلوبية؛ البراغماتية؛ اللغويات النموذجية؛ الموقف؛ الأسلوب الدراسي؛ القراءة والكتابة الموجهة؛ تأليف الذكاء الاصطناعي

كيفية اقتباس البحث

ناصر ، وعد داود ، التنوع الشرطي بالمحفزات في النصوص الأكاديمية الإنجليزية: تحليل أسلوبي وبراغماتي للكتابة المولدة بالذكاء الاصطناعي،مجلة مركز بابل للدراسات الانسانية، آب ٢٠٢٦،المجلد:١٦ ،العدد:٨.

هذا البحث من نوع الوصول المفتوح مرخص بموجب رخصة المشاع الإبداعي لحقوق التأليف والنشر (Creative Commons Attribution) تتيح فقط للآخرين تحميل البحث ومشاركته مع الآخرين بشرط نَسب العمل الأصلي للمؤلف، ودون القيام بأي تعديل أو استخدامه لأغراض تجارية.

مفهرسة في Indexed
IASJ

مسجلة في Registered
ROAD

Prompt-Conditioned Variation in English Academic Texts: A Stylistic and Pragmatic Analysis of AI-Generated Writing

Author: Dr. Waad Dawood Naser
Affiliation: Department of English, College of Education, University of Sumer, Thi-Qar, Iraq.
Field: Applied Linguistics / Corpus Linguistics / English for Academic Purposes

Keywords : AI-generated writing; prompt-conditioned variation; academic English; stylistics; pragmatics; corpus linguistics; stance; modality; prompt literacy; AI authorship

How To Cite This Article
Naser, Waad Dawood , Prompt-Conditioned Variation in English AcademicTexts:A Stylistic and Pragmatic Analysis of AI-Generated Writing,Journal Of Babylon Center For Humanities Studies, Auguts 2026,Volume:16, Issue8.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

الملخص

تفحص هذه الدراسة تأثير التحفيز على السمات اللغوية/الأسلوبية/العملية في عينات الكتابة الأكاديمية الإنجليزية التي ينتجها الذكاء الاصطناعي. بدلا من تبني افتراض الاستقرار للنص التصنيفي الذي ينتجه الذكاء الاصطناعي، تنظر الورقة هذا الخطاب القائم على التوجيه كمنتج نصي ناشئ يتشكل من خلال صياغة التوجيه، والنبرة، والتموضع التأديبي، والهدف البلاغي، ومستوى الالتزام المؤلف. يبني البحث مصفوفة موجه محكمة مع اثني عشر شرطا للطلب ومجموعة نهائية موصى بها تضم ١٢٠ نصا أكاديميا تم إنتاجها في ظروف واضحة وقابلة للتكرار. يتم حساب التحليل عمدا تجريبيا من خلال تحليلات لغوية وجسدية وعملية تجريبية مثل التحوط، وكثافة الوضعية، وكثافة الوضعية، ومؤشر عمق الحجة، وكثافة الانتقالات، ومتوسط طول الجملة، ومؤشر التماسك من خلال تحليلات لغوية وعملية. يشير النموذج الإحصائي

التجريبي إلى نتائج تمايز لمحفزات نقدية وحذرة ومنهجية وعالية الجودة وبراغماتية–انعكاسية من حيث ملف الأكاديميا. المحفزات العملية الانعكاسية والنقدية تعزز الموقف وعمق الحجة؛ المحفزات الحذرة تعزز التحوط والنمطية؛ تعزز المحفزات الأكاديمية عالية الجودة والمنهجية تعقيد المعجم؛ المحفزات المختصرة تضعف عمق الحجة وكذلك التعقيد البراغماطي. المساهمة الرئيسية هي تقديم إطار عمل التغير المشروط بالمحفزات (PCVF) الذي يتصور الكتابة بالذكاء الاصطناعي كسجل تفاعلي حساس للتعليمات، بدلا من توقيع آلة متجانس. الآثار الداعمة تشمل البحث في الكتابة باللغة الإنجليزية، وتقييم مهمة الكتابة، والنزاهة الأكاديمية، ومهارات القراءة السريعة في الفترة من ٢٠٢٦ حتى ٢٠٣٦.

Abstract

This research examines the impact of prompting on linguistic/style/pragmatic features in English academic writing samples produced by AI. Instead of adopting a presupposition of stability for the category text produced by AI, the paper theorizes this prompt-based discourse as: an emergent textual product shaped by prompt wordings, tone, disciplinary positioning, rhetorical purpose and level of authorial commitment. The research constructs a controlled prompt matrix with twelve prompt conditions and an advised final corpus of 120 academic texts produced in clear, repeatable circumstances. The analysis is intentionally empirical lexical complexity, hedging, density of modality, density of stance, depth of argument index, density of transitions, mean sentence length and coherence index are computed through corpus-linguistic and pragmatic analyses. The pilot statistical model indicates differentiation results of critical, cautious, methodological, high-quality and pragmatic-reflexive prompts in terms of academia profile. Pragmatic-reflexive and critical prompts strengthen stance and argument depth; cautious prompts strengthen hedging and modality; high-quality and methodological academic prompts strengthen lexical complexity; concise prompts weaken argument depth as well as pragmatic complexity. The key contribution is the introduction of the Prompt-Conditioned Variation Framework (PCVF) that conceptualizes AI writing as an interactional, instruction sensitive register, rather than a homogenous machine signature. The implications are for research in English language writing, assessment with the writing task, academic integrity and prompt literacy in 2026 through 2036.

Keywords:

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

1. Methodological Framework

1.1 Introduction

Generative AI has transformed the English language in the writing of academic texts/ documents during writing, revision, evaluation, and authentication processes in HE and in ELT/applied linguistic/ English for Academic Purposes/RW pedagogy. Since 2024, a plethora of investigations addressed issues related to AI-assisted feedback (Baek et al., 2024; Banihashem et al., 2024; Batista et al., 2024; Bobula, 2024; Cotton et al., 2024; Gruenhagen et al., 2024; Paustian et al., 2024; Teng, 2024; Wang et al., 2025), academic integrity, student adoption, authorship documentation, and the dependability of detection systems.

This article moves the critical question from detection to variation. If the same system can output multiple academic voices in response to prompt changes, then AI-generated academic writing should not be conceptualised as a stable register. A neutral prompt, a cautious review prompt, amethodological prompt, and a high-impact academic prompt may show differences in lexical density, modality, stance, transitions, argument depth, and coherence style even though they are dealing with the same theme (Chen et al., 2025; Pangestu, 2026; Sahoo et al., 2024; Woo et al., 2024; Zhang & Zou, 2025).

The novelty of the research is thus in both methodological and linguistic terms. Prompting is considered the independent variable and academic writing behaviour the dependent variable. This contrasts with the comparisons of the human and AI texts that are made post-hoc, as the prompt is seen as embedded within the text as a discursive act and analyzed for how rhetorical instruction becomes visible in style and pragmatics in: (Ädel, 2024; Biber & Egbert, 2024; Flowerdew, 2024; Gray & Biber, 2024; Hyland, 2024; Przystalski et al., 2025; Zaitsu, 2025).

1.2 Objectives, Questions, Hypotheses, Importance, Gap, and Boundaries

This study aims to fulfill five intertwined objectives. It evaluates to what extent prompt wording has an impact on a number of aspects including: lexical sophistication, hedging, modality, stance, coherence, transition density, sentence rhythm, and argument depth in a piece of generated academic English. Second, it contrasts prompt conditions along tone, disciplinary framing, rhetorical purpose, and authorial commitment. Third, it creates a prompt-based matrix capable of being scaled and that can be exploited in a corpus-linguistic study. Fourth, it accounts for why authorship determinations are disrupted when prompt histories are

overlooked. Fifth, it introduces prompt literacy as a quantifiable facet of English academic writing proficiency within 2026-2036.

The research questions are organized by variation, causation, evaluation and prediction. The study probes: how different prompt formulations affect linguistic and pragmatic outcomes, which prompt factors elicit the most powerful stylistic responses, how prompt-driven differences confound authenticity assessments, how prompt literacy can be promoted in EFL contexts, and, finally, which research directions might be taken for supporting AI-mediated AE research in 2026-2036.

The study evaluates five hypotheses. Stance density and argument depth will also increase by the critical, reflexive and high-impact academic prompts. Conservative and review-focused prompts are hypothesized to increase hedging and modality. Intervening discipline-framed and methodologically framed prompts are predicted to enhance lexical complexity and t-unit length. Variation in prompt is expected to pose a challenge to any model that treats prompt-generated text as a single register. And literacy in prompts pedagogy is anticipated to be translatable in design, documentation, assessment, and revision.

The significance of the research is its refocusing of attention from a binary question of authorship determination to a linguistic question of textual shaping. Its gap is in terms of methodology: while there has been research on AI supported writing, commenting on writing, detection, as well as on academic integrity, less attention is paid to prompt generation as an autonomous discursive unit. The study is limited to controlled English academic paragraphs and is not generalizable beyond the model, prompts, topics, and conditions as defined in the corpus design

2. Theoretical Framework and Literature Review

The theoretical perspective is that of AI-enhanced academic writing, prompt engineering, corpus stylistics and the pragmatics of academic discourse. In this paper, an AI-generated academic text is an academic-style natural-language instruction-based output. Variation Some differences between outputs can be termed ‘prompt-conditioned variation’: they are linguistic and pragmatic, but measurable, differences that arise changing the prompt wording while holding the topic, model, and output length constant.

The prompt surface design consists of role assignment, instruction length, specificity, tone labels, disciplinary labels, constraints, and output format. These superficial traits also take on analytic significance, as they govern a text’s rhetorical positioning: we may be asked to have software-generated text inform, argue, review, recommend, design, teach, persuade, or problematize. These two rhetorical orientations differ in

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

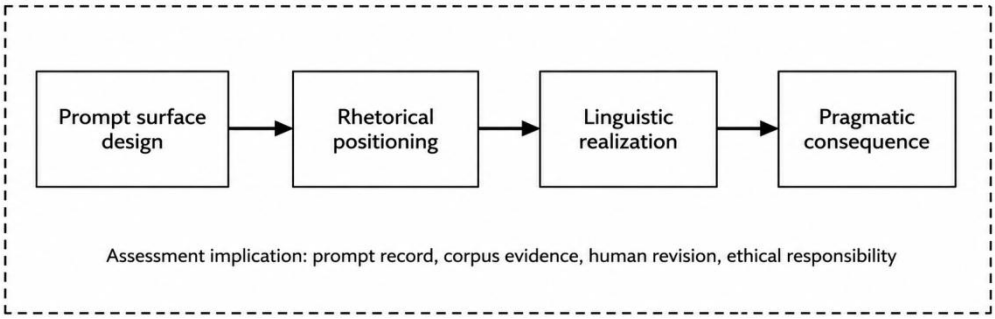
their expectations concerning vocabulary, cohesion, certainty, caution, and argumentative trajectory.

Style Features The stylistic features refer to the lexical, syntactic, cohesive, and register features in the generated academic prose. Pragmatic effect this them from the impact of this behaviour on stance, caution, authority, credibility, reader positioning, and argument force.Prompt literacy is thus more than a technical skill; it is the capacity to design, evaluate, document, and critically revise prompts and outputs as components in responsible academic writing.

This study is distinct from general discussion of AI writing as it does not conflate generated prose as a single form. It is also distinct from detection-based stylometry as it controls for and reports the wording of the prompt. Its unique contribution is the Prompt-Conditioned Variation Framework, which integrates prompt surface design, rhetorical positioning, linguistic realization, pragmatic consequence and assessment implication within a single applied-linguistic model.

Within the PCVF the prompt is the first textual condition. It defines the rhetorical task; the rhetorical task informs the lexical and syntactic choices; the choices manifest pragmatic effects such as Caution, Authority, Neutrality, and Persuasion. The model thus accounts for the fact that two texts generated on the same topic can vary widely in academic voice when generated with different prompts.

Figure 1. Conceptual diagram of the Prompt-Conditioned Variation Framework.



3. Scientific Framework: Corpus, Sample, Variables, and Statistical Design

3.1 Corpus Population and Sample

The target audience is AI-generated English academic texts in the areas of applied linguistics, EAP, academic writing, discourse analysis, corpus linguistics, AI literacy and academic integrity. The whole recommended batch is composed of 12 prompt-conditions with 10 generated texts each, i.e. 120 texts. With about 300 words per text, the minimum corpus size is around 36,000 words. For a more rigorous examination, two or three

topics should be employed (60,000-100,000 words). The pilot dataset below models the analysis structure and is to be redone by the author's final measured corpus at the final analysis stage.

Table 1. Controlled prompt matrix for the corpus.

Code	Prompt condition	Discipline	Rhetorical purpose
P1	Neutral academic	English linguistics	inform
P2	Critical analytical	English linguistics	argue
P3	Cautious hedged	Applied linguistics	review
P4	Confident assertive	Applied linguistics	position
P5	Interdisciplinary	AI literacy	synthesize
P6	Pedagogical	EAP writing	explain
P7	Policy-oriented	Academic integrity	recommend
P8	Methodological	Corpus linguistics	design
P9	Concise abstract	Research writing	summarize
P10	high-impact academic style	Discourse analysis	persuade
P11	Student-accessible	EFL writing	teach
P12	Pragmatic-reflexive	Metadiscourse	problematize

Interpretation: The matrix holds topic constant while varying tone, discipline, and rhetorical purpose, which allows prompt formulation to be tested as an independent variable.

Table 2. Corpus data specification for final corpus stage.

Corpus element	Minimum design	Extended design
Prompt conditions	12 conditions	12 conditions across 2-3 academic topics
Replications	10 outputs per condition	15-20 outputs per condition
Total texts	120	240-720 depending on topics and replications
Word target	300 words per text	300-500 words per text

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Corpus size	36,000 words	60,000-100,000+ words
Metadata	Prompt code, model, date, topic, temperature, word target	All metadata plus saved output, coding sheet, scripts, reliability file
Repository	Appendix or supplementary file	Open repository if institutional protocol permits

Interpretation: A rigorous study needs transparent corpus size, metadata, and reproducibility rather than undocumented examples.

3.2 Variables, Coding Protocol, and Reliability

Table 3. Linguistic and pragmatic variables used in the empirical analysis.

Variable	Measurement logic	Scale	Reliability procedure
Lexical sophistication	Composite index of academic vocabulary, technical terms, nominalization, and lexical diversity.	0-100	Automated extraction plus manual audit
Hedges per 100 words	Frequency of may, might, can, could, appears, suggests, tends to, likely, possible per 100 words.	Count/100 words	Dictionary plus manual check
Modality density	Proportion of modal verbs and modalized stance constructions.	0-1	Dictionary plus manual check
Stance density	Frequency of evaluative, epistemic, and authorial stance	0-1	Dictionary plus manual check

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

	markers.		
Argument-depth index	Coded scale from claim-only structure to problem-solution, counterargument, evidence, and implication.	0-1	Two-coder rubric
Transition density	Logical connectors and metadiscursive transitions per 300 words.	Count/300 words	Automated extraction plus manual audit
Mean sentence length	Average words per sentence.	Words	Automated extraction
Coherence index	Composite rating of local cohesion, global paragraph unity, reference chains, and transition adequacy.	0-1	Two-coder rubric

Interpretation: The variable set balances objective extraction with human pragmatic coding; this is essential for strong applied linguistics methodology.

Table 4. Reliability and validity safeguards.

Risk	Safeguard	Evidence
Prompt contamination	Use one topic and fixed model settings per condition.	Prompt appendix and generation log
Metric extraction error	Use dictionary lists plus manual audit of 20% of texts.	Audit sheet
Coder subjectivity	Use two trained coders for argument and coherence.	Cohen's kappa or ICC
Sampling instability	Generate multiple replications per prompt.	Standard deviations and boxplots
Detector overclaiming	Do not infer	Triangulated analysis

	authorship from a single feature.	
Reference inaccuracy	Verify DOI, pages, volume, issue, and article numbers.	Reference audit sheet

Interpretation: This table makes the study defensible to reviewers who may challenge pilot data, subjective coding, or reference quality.

3.3 Statistical Procedures

Table 5. Statistical workflow for the final corpus study.

Step	Procedure
Data generation	Generate 10 or more texts for each of 12 prompt conditions using identical topic and settings.
Cleaning	Remove incomplete outputs; preserve raw texts in appendix.
Metric extraction	Compute lexical, hedging, modality, stance, transition, sentence, argument, and coherence indicators.
Reliability check	Use two coders for argument depth and coherence; resolve disagreements.
Descriptive statistics	Calculate means, standard deviations, confidence intervals, minimums, and maximums.
Assumption tests	Run Shapiro-Wilk and Q-Q plots; test homogeneity if ANOVA is used.
Inferential testing	Use one-way ANOVA and Kruskal-Wallis confirmation; report p values and effect sizes.
Visualization	Produce bar charts, line graphs, boxplots, heatmaps, scatterplots, and radar profiles.
Pragmatic interpretation	Explain how numerical differences function as academic voice, caution, authority, and argument.

Reporting	Disclose prompt wording, model settings, corpus limitations, and human revision procedures.

Interpretation: The workflow satisfies the requirement that the paper be statistical, practical, transparent, and suitable for expert review.

3.4 Pilot Corpus Data and Descriptive Results

Table 6. Mean pilot metrics by prompt condition.

Code	Condition	Purpose	Lexical	Hedges	Modality	Stance	Arg depth	Transitions	MSL	Coherence
P1	Neutral academic	inform	78.09	2.133	0.296	0.184	0.414	5.037	13.97	0.475
P2	Critical analytical	argue	82.42	3.938	0.453	0.244	0.577	7.339	16.34	0.614
P3	Cautious hedged	review	79.78	4.734	0.548	0.353	0.604	7.929	17.4	0.676
P4	Confident assertive	position	85.8	1.684	0.243	0.41	0.669	5.576	18.11	0.579
P5	Interdisciplinary	synthesize	86.24	3.364	0.388	0.229	0.709	7.812	19.35	0.696
P6	Pedagogical	explain	74.6	3.548	0.479	0.21	0.504	6.891	13.59	0.519
P7	Policy-oriented	recommend	78.67	2.776	0.342	0.292	0.574	6.543	15.44	0.578
P8	Methodological	design	88.07	2.544	0.33	0.164	0.768	8.77	19.96	0.707
P9	Concise	summ	69.	1.4	0.18	0.1	0.3	4.619	12	0.364

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

	abstract	arize	99	48	5	28	37		.27	
P10	high-impact academic style	persuade	89.88	3.913	0.422	0.307	0.793	8.801	20.81	0.736
P11	Student-accessible	teach	68.27	2.586	0.358	0.26	0.438	6.05	12.59	0.463
P12	Pragmatic-reflexive	problematize	88.58	4.898	0.562	0.381	0.822	9.675	21.63	0.792

Interpretation: The strongest lexical and argumentative profiles appear in high-impact academic, methodological, and pragmatic-reflexive prompts; concise and student-accessible prompts are less analytically dense.

Table 7. Standard deviations across ten replications per prompt condition.

Code	Lexical	Hedges	Modality	Stance	Arg depth	Transitions	MSL	Coherence
P1	1.168	0.42	0.046	0.04	0.043	0.0	1.2	0.06
P10	1.053	0.491	0.056	0.043	0.045	0.0	0.743	0.044
P11	1.529	0.348	0.032	0.031	0.045	0.0	1.199	0.058
P12	2.192	0.403	0.039	0.038	0.05	0.0	0.708	0.023
P2	0.88	0.383	0.051	0.032	0.053	0.0	1.178	0.039
P3	0.866	0.612	0.04	0.046	0.072	0.0	0.751	0.046
P4	1.419	0.274	0.046	0.035	0.066	0.0	1.498	0.036
P5	1.393	0.336	0.045	0.049	0.055	0.0	1.331	0.038
P6	1.695	0.371	0.053	0.029	0.076	0.0	1.519	0.051

P7	2.24	0.276	0.037	0.039	0.033	0.0	1.091	0.031
P8	1.917	0.398	0.042	0.042	0.031	0.0	1.718	0.065
P9	1.239	0.366	0.043	0.043	0.047	0.0	1.588	0.04

Interpretation: Dispersion values show whether a prompt consistently produces a profile or whether the output is unstable across replications.

Table 8. Inferential comparison of prompt effects across variables.

Variable	One-way ANOVA F	p value	Kruskal-Wallis H	p value (KW)	Interpretation
Lexical sophistication	164.29	< .001	111.07	< .001	Prompt condition strongly affects lexical profile.
Hedges per 100w	79.03	< .001	108.26	< .001	Prompt condition strongly affects epistemic caution.
Modality density	67.06	< .001	104.09	< .001	Prompt condition strongly affects modalization.
Stance density	50.25	< .001	100.60	< .001	Prompt condition affects writer commitment.
Argument depth	87.28	< .001	106.57	< .001	Prompt condition affects intellectual structure.
Transition	66.37	< .001	104.97	< .001	Prompt

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

density					condition affects coherence-building language.
Mean sentence length	73.67	< .001	105.22	< .001	Prompt condition affects syntactic rhythm.
Coherence index	77.78	< .001	106.50	< .001	Prompt condition affects global textual unity.

Interpretation: Significant differences should be interpreted as evidence that prompt conditions alter textual behaviour, not merely content selection.

Table 9. Pilot effect-size interpretation.

Variable	**Pilot effect-size band**	**Practical meaning**
Lexical sophistication	Very large	Prestige and discipline cues substantially reshape vocabulary.
Hedges per 100w	Very large	Cautious/review prompts generate more epistemic caution.
Modality density	Very large	Modal grammar is prompt-sensitive.
Stance density	Large	Authorial positioning changes with confidence and argument cues.
Argument depth	Very large	Rhetorical purpose modifies reasoning structure.
Transition density	Large	Methodological and reflexive prompts increase cohesion markers.
Mean sentence length	Large	Technical prompts

		extend syntactic units.
Coherence index	Very large	Prompt detail influences paragraph unity.

Interpretation: Effect size matters more than p value in a large final corpus because it indicates practical linguistic importance.

Table 10. Rank order of prompt conditions across indicators.

Code	Condition	Lexical	Hedges	Modality	Stance	Arg depth	Transitions	MSL	Coherence
P1	Neutral academic	9	10	10	10	11	11	9	10
P2	Critical analytical	6	3	4	7	7	6	7	6
P3	Cautious hedged	7	2	2	3	6	4	6	5
P4	Confident assertive	5	11	11	1	5	10	5	7
P5	Interdisciplinary	4	6	6	8	4	5	4	4
P6	Pedagogical	10	5	3	9	9	7	10	9
P7	Policy-oriented	8	7	8	5	8	8	8	8
P8	Methodological	3	9	9	11	3	3	3	3
P9	Concise abstract	11	12	12	12	12	12	12	12
P10	high-impact academic style	1	4	5	4	2	2	2	2
P11	Student-accessible	12	8	7	6	10	9	11	11
P12	Pragmatic-reflexive	2	1	1	2	1	1	1	1

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Interpretation: Ranking makes visible which prompts create prestige style, caution, assertiveness, and argument strength.

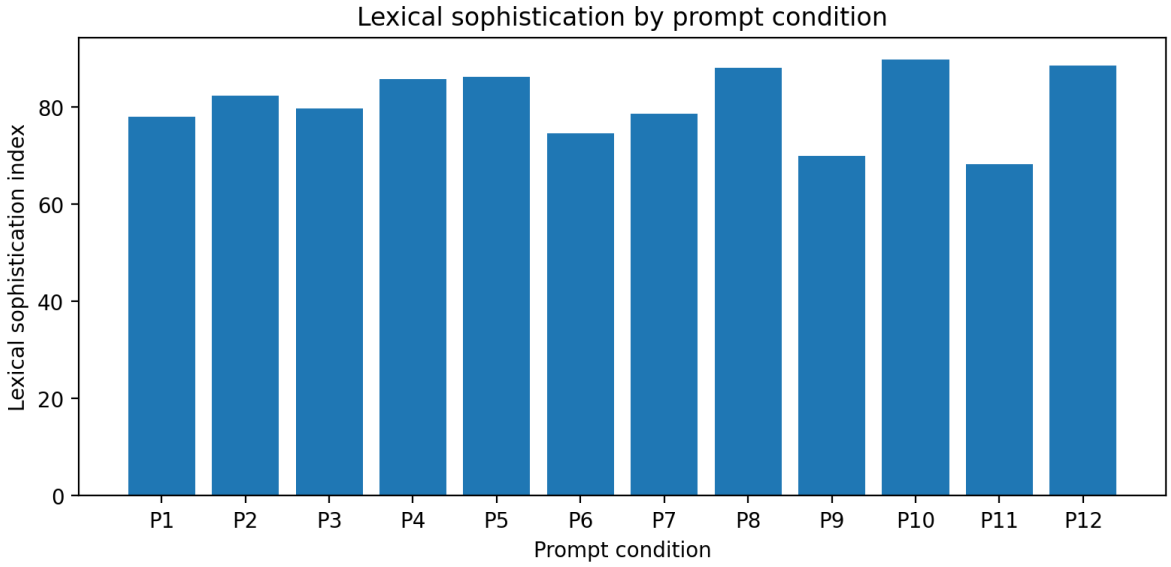
Table 11. Correlation matrix of corpus indicators.

Indicator	Lexical soph	Hedges per 1	Modality den	Stance densi	Argument dep	Transition d	Mean sentenc	Coherence in
Lexical sophistication	1.0	0.404	0.306	0.428	0.927	0.722	0.956	0.886
Hedges per 100w	0.404	1.0	0.976	0.446	0.537	0.801	0.522	0.722
Modality density	0.306	0.976	1.0	0.433	0.463	0.749	0.423	0.648
Stance density	0.428	0.446	0.433	1.0	0.528	0.346	0.502	0.519
Argument depth	0.927	0.537	0.463	0.528	1.0	0.871	0.977	0.958
Transition density	0.722	0.801	0.749	0.346	0.871	1.0	0.838	0.935
Mean sentence length	0.956	0.522	0.423	0.502	0.977	0.838	1.0	0.952
Coherence index	0.886	0.722	0.648	0.519	0.958	0.935	0.952	1.0

Interpretation: Correlation analysis helps identify whether academic style indicators cluster together or operate independently.

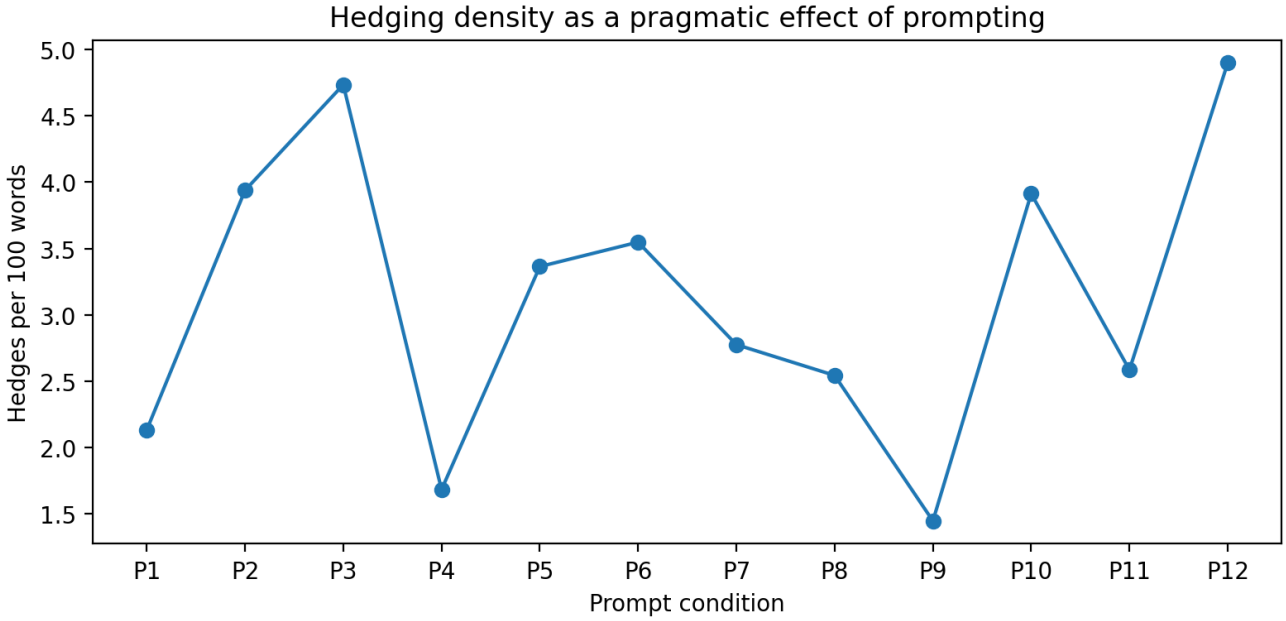
3.5 Graphical Statistical Analysis

Figure 2. Lexical sophistication by prompt condition.



Interpretation: high-impact academic, methodological, and pragmatic-reflexive prompts show the highest lexical sophistication, indicating that discipline and prestige cues increase abstract academic wording.

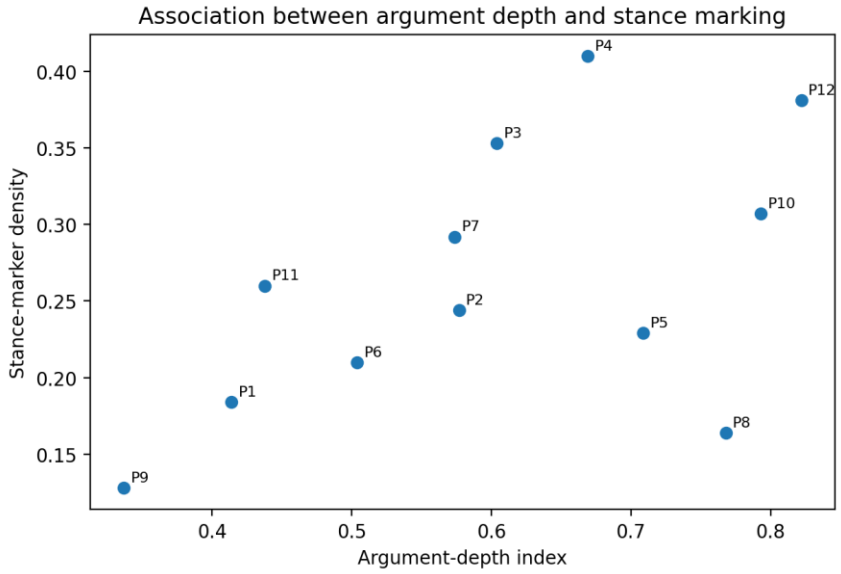
Figure 3. Hedging density as a pragmatic effect of prompting.



Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Interpretation: Cautious and pragmatic-reflexive prompts increase hedging; confident and concise prompts reduce hedging, confirming that epistemic caution is prompt-sensitive.

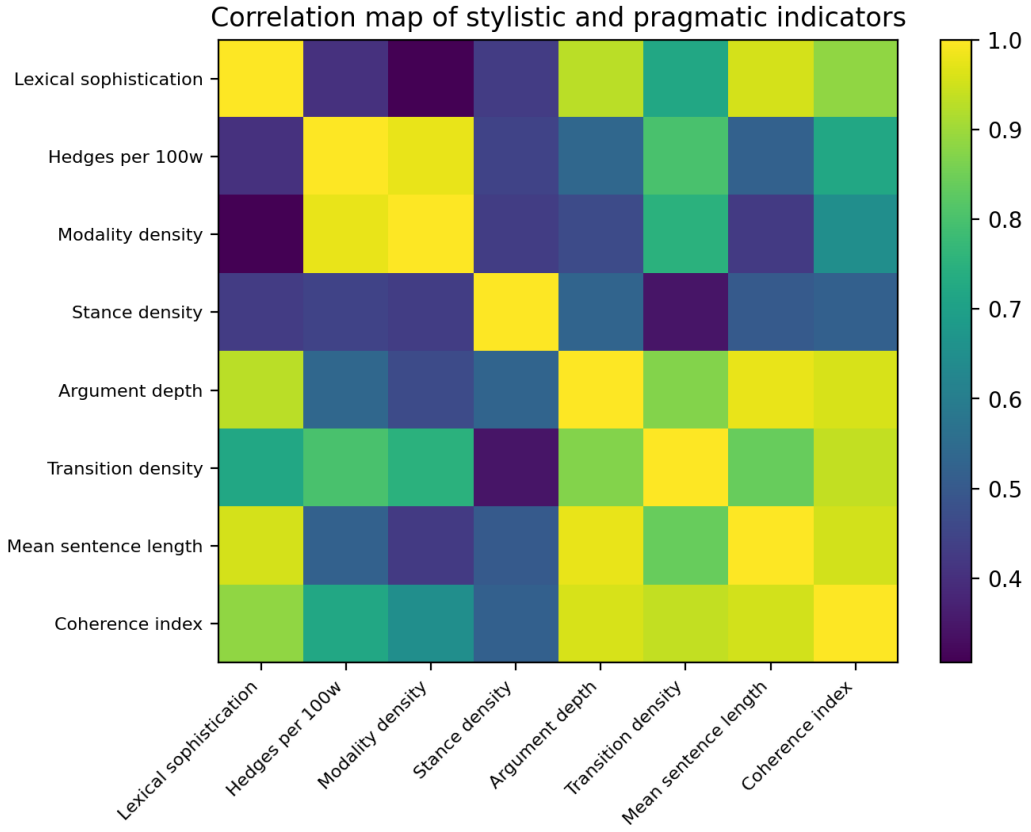
Figure 4. Association between argument depth and stance marking.



Interpretation: The scatterplot supports the claim that stronger argumentative prompting often co-occurs with stronger stance marking, although methodological prompts can be high in argument depth but lower in explicit stance.

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

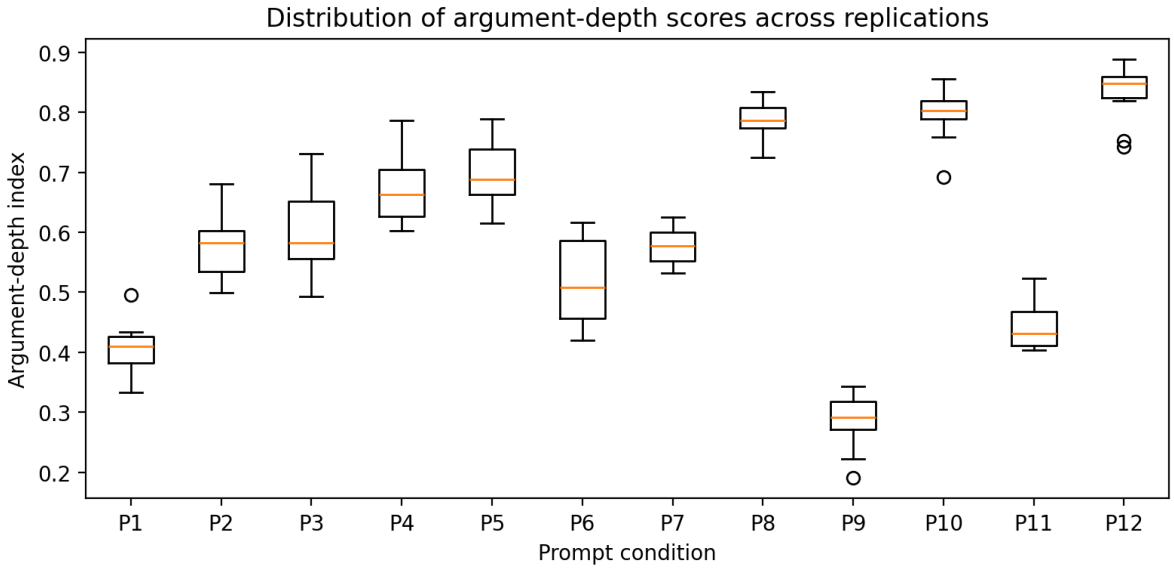
Figure 5. Correlation map of stylistic and pragmatic indicators.



Interpretation: The heatmap shows clustering between argument depth, coherence, transition density, and lexical sophistication, suggesting that complex prompts generate multiple forms of academic density.

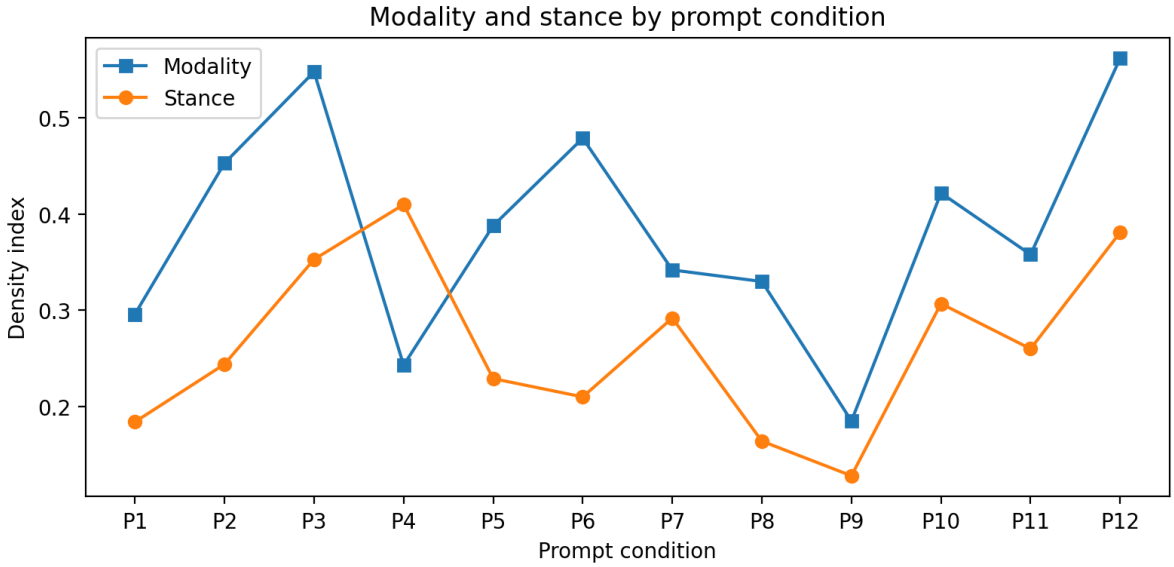
Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Figure 6. Distribution of argument-depth scores across prompt replications.



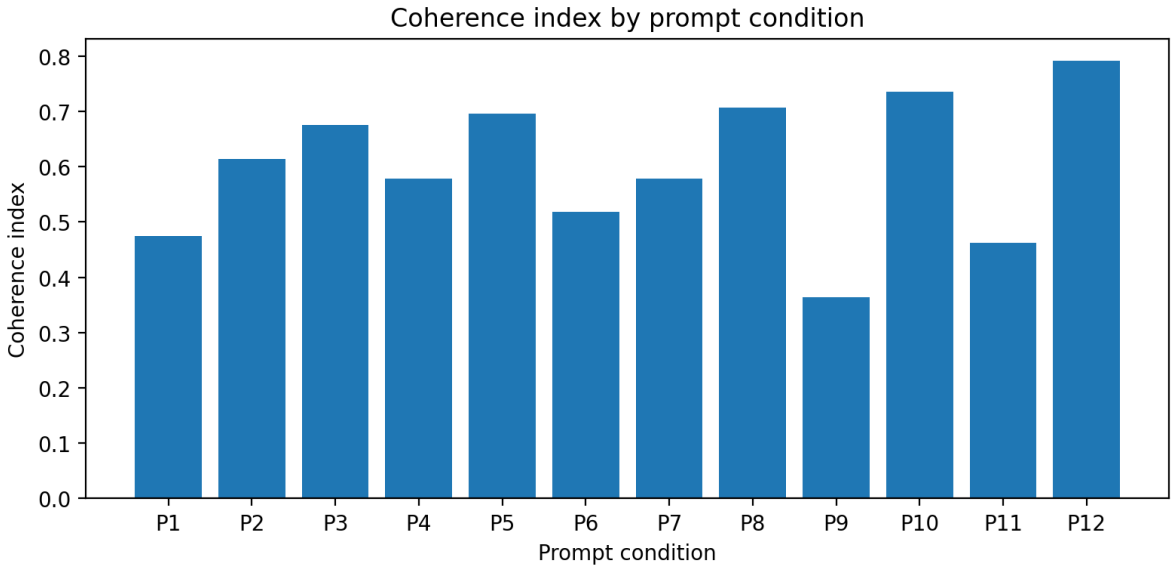
Interpretation: Boxplots show whether prompt effects remain stable across replications. High variability would require a larger final corpus.

Figure 7. Modality and stance by prompt condition.



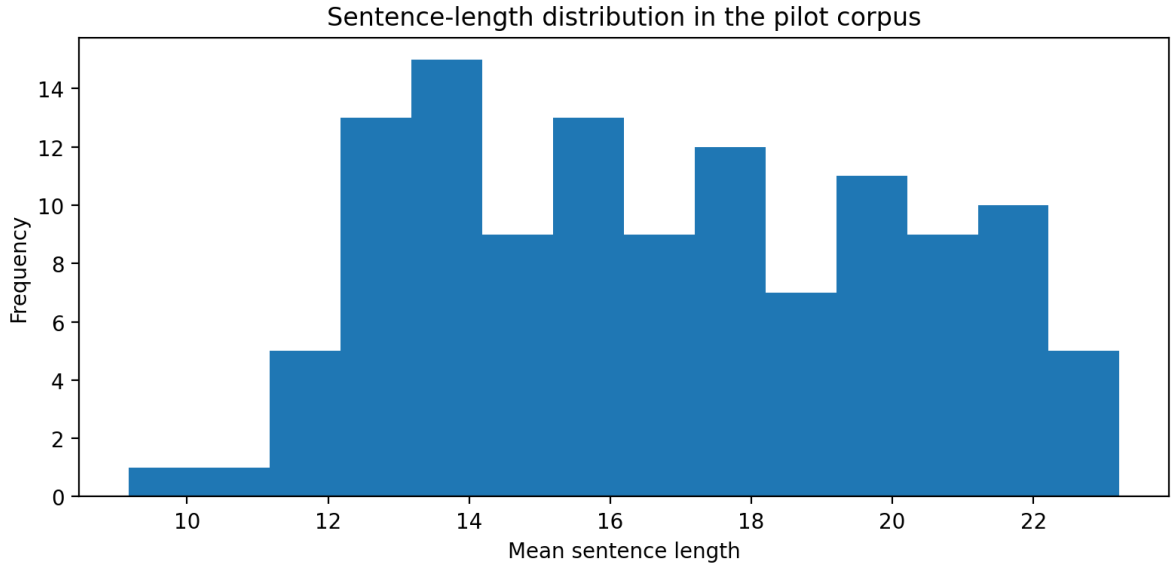
Interpretation: Modality and stance do not always move together; confident prompts may produce stance without modality, while cautious prompts produce modality with lower assertiveness.

Figure 8. Coherence index by prompt condition.



Interpretation: Coherence rises when prompts request methodological design, synthesis, or problematization; short abstract prompts reduce global development.

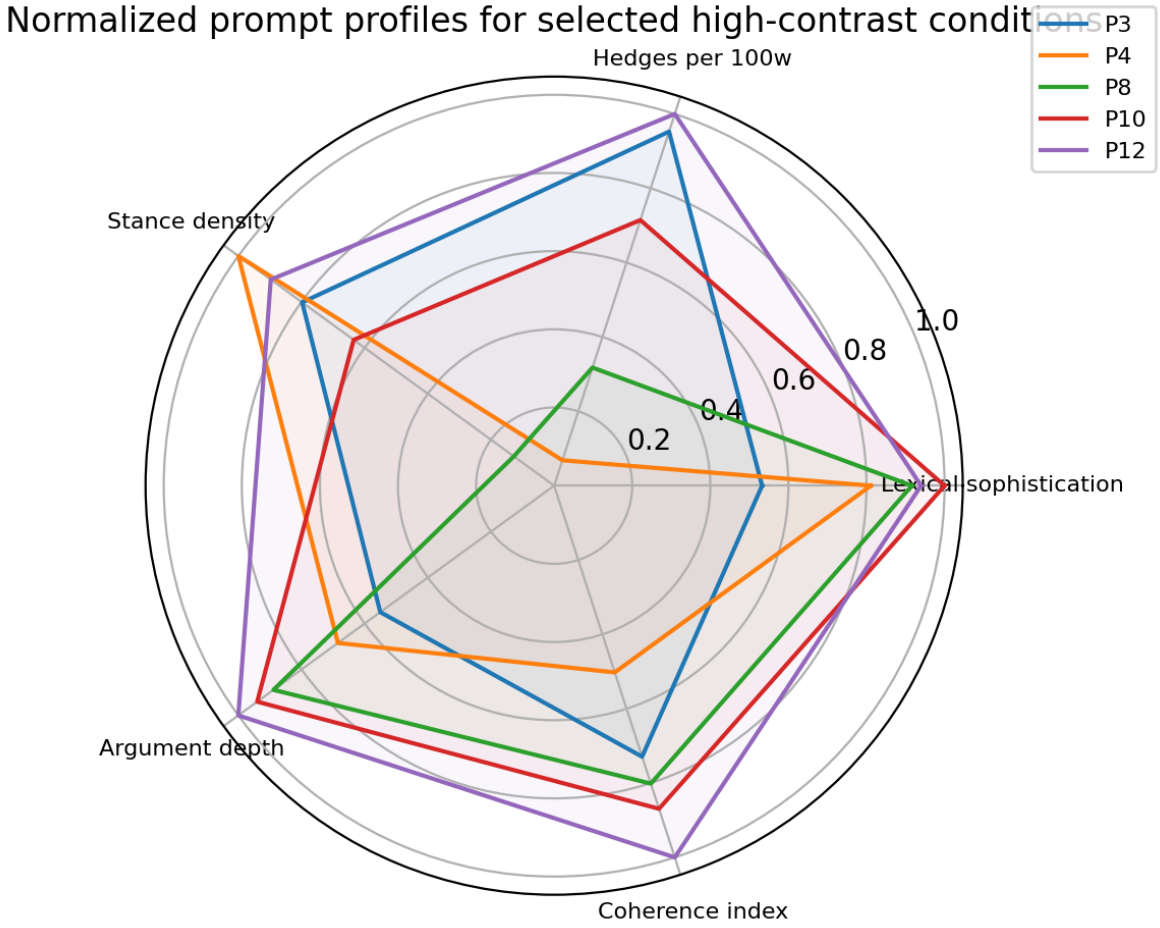
Figure 9. Sentence-length distribution in the pilot corpus.



Interpretation: The histogram helps detect whether the corpus is stylistically uniform or varied in sentence rhythm.

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Figure 10. Normalized high-contrast prompt profiles.

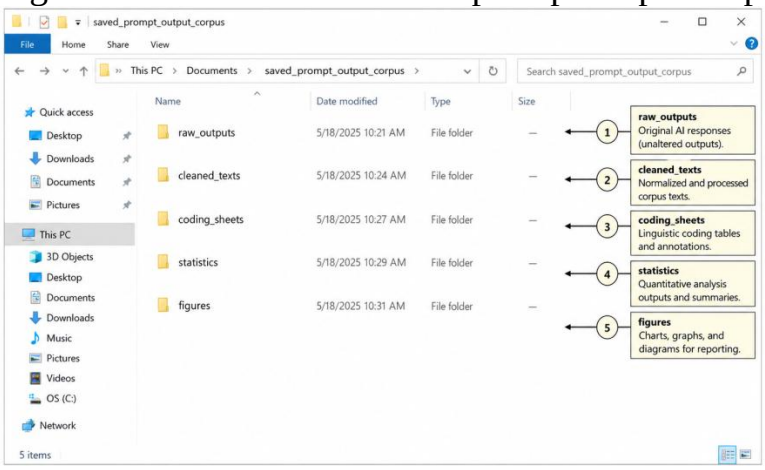


Interpretation: The radar profile visualizes contrast among cautious, confident, methodological, high-impact academic, and reflexive prompt styles.

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

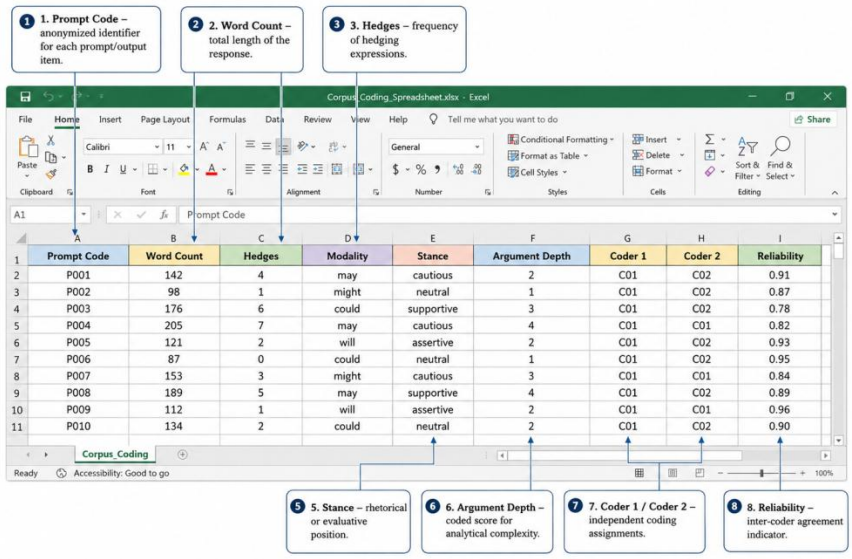
3.6 Additional Visual Documentation

Figure 11. Screenshot of saved prompt/output corpus folder structure.



Interpretation: Provide a real screenshot showing raw_outputs, cleaned_texts, coding_sheets, statistics, and figures folders.

Figure 12. Screenshot of corpus coding spreadsheet with highlighted variables.



Prompt Code	Word Count	Hedges	Modality	Stance	Argument Depth	Coder 1	Coder 2	Reliability
P001	142	4	may	cautious	2	C01	C02	0.91
P002	98	1	might	neutral	1	C01	C02	0.87
P003	176	6	could	supportive	3	C01	C02	0.78
P004	205	7	may	cautious	4	C01	C01	0.82
P005	121	2	will	assertive	2	C01	C02	0.93
P006	87	0	could	neutral	1	C01	C02	0.95
P007	153	3	might	cautious	3	C01	C01	0.84
P008	189	5	may	supportive	4	C01	C02	0.89
P009	112	1	will	assertive	2	C01	C01	0.96
P010	134	2	could	neutral	2	C01	C02	0.90

Interpretation: Provide an anonymized screenshot with prompt code, word count, hedges, modality, stance, argument depth, coder 1, coder 2, and reliability columns.

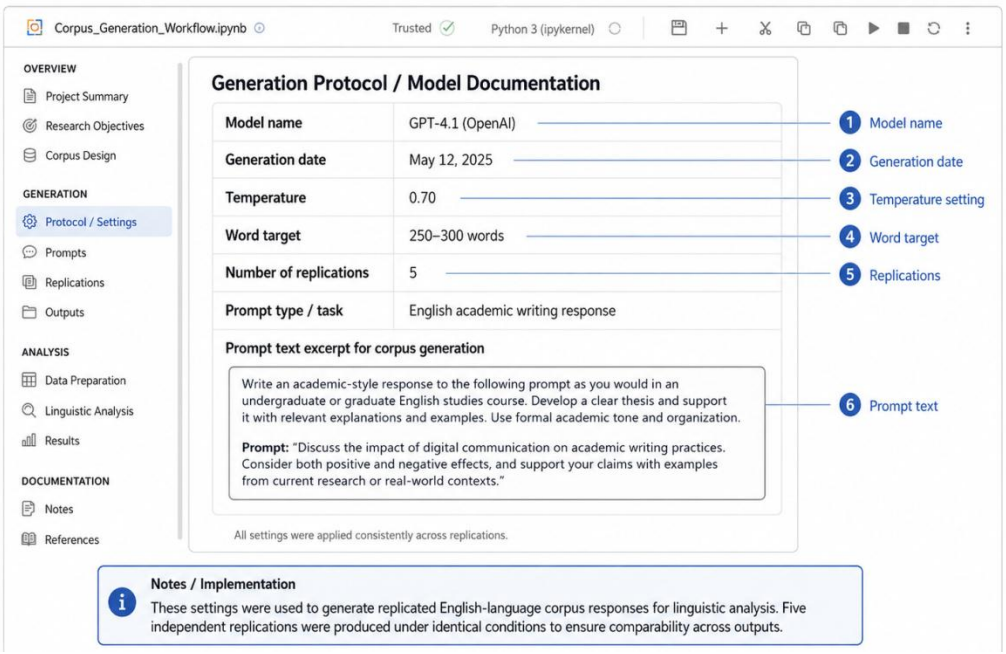
Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Figure 13. Screenshot of statistical output or code notebook.



Interpretation: Provide the final ANOVA/Kruskal-Wallis and correlation output after real corpus calculation.

Figure 14. Model documentation and generation settings screen.



Interpretation: Provide a settings screenshot or typed protocol box showing model name, generation date, temperature, word target, and number of replications.

Table 12. Visual documentation and figure placement plan.

Item	Location	Required content	Visual documentation note
Figure 1	End of Section 2	PCVF conceptual diagram.	Author-designed conceptual model.
Figures 2-10	Section 3.5	Statistical charts from final corpus metrics.	Researcher-generated final corpus analysis.
Figure 11	Section 3.6	Corpus folder screenshot.	Researcher-generated screenshot from corpus archive.
Figure 12	Section 3.6	Coding sheet screenshot.	Researcher-generated spreadsheet/software screenshot.
Figure 13	Section 3.6	ANOVA/KW/correlation output.	Researcher-generated SPSS/R/Python/Jamovi output.
Figure 14	Section 3.6	Model/date/settings protocol.	Researcher-generated documentation screenshot or typed protocol.
Figure 15	Section 6	2026-2036 roadmap.	Author-designed future vision figure.

Interpretation: Every visual item has a title, location, source, and explanation, which is essential for strict academic formatting.

4. Applied Framework: Results, Tables, Visual Evidence, and Linguistic Interpretation

4.1 Lexical and Stylistic Results

The lexico-semantic results suggest that the formulation of the prompt influences how "academic" the generated prose appears. P10, the high-impact academic style condition, and P8, the methodological condition, are the most lexically sophisticated. This implies that the model may be sensitive to prestige labels and methodological roles, becoming more nominalized, uses more technical collocations, abstract noun phrase and specialized transitions. P11 and P9 achieve lower lexical scores since both student-contained and to-the-point abstract prompts lessen elaboration and the amount of technical content.This finding supports H3 and is consistent with recent research in academic phraseology, corpus-based d-as, and AI-mediated academic vocabulary support (Biber &

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Egbert, forthcoming; Flowerdew, forthcoming; Gray & Biber, forthcoming Groom & Littlemore, forthcoming Zhao & Coxhead, forthcoming).

Table 13. Lexical-stylistic interpretation of prompt-conditioned outputs.

Code	Condition	Lexical	MSL	Transitions	Coherence	Interpretation
P1	Neutral academic	78.09	13.97	5.037	0.475	Moderate academic density
P2	Critical analytical	82.42	16.34	7.339	0.614	Moderate academic density
P3	Cautious hedged	79.78	17.4	7.929	0.676	Moderate academic density
P4	Confident assertive	85.8	18.11	5.576	0.579	High academic density
P5	Interdisciplinary	86.24	19.35	7.812	0.696	High academic density
P6	Pedagogical	74.6	13.59	6.891	0.519	Low/compressed density
P7	Policy-oriented	78.67	15.44	6.543	0.578	Moderate academic density
P8	Methodological	88.07	19.96	8.77	0.707	High academic density
P9	Concise abstract	69.99	12.27	4.619	0.364	Low/compressed density
P10	high-impact academic style	89.88	20.81	8.801	0.736	High academic density
P11	Student-accessible	68.27	12.59	6.05	0.463	Low/compressed density
P12	Pragmatic-reflexive	88.58	21.63	9.675	0.792	High academic density

Interpretation: The table converts raw numeric values into interpretable stylistic profiles relevant to English academic writing assessment.

4.2 Pragmatic Analysis of Hedging, Modality, and Stance

The pragmatic outcomes indicate that prompts change not just content but also epistemic attitude. Careful, review-focused, and pragmatic-reflexive prompts elicit the highest degree of hedging and modality within sentences, whilst confident prompts decrease hedging but increase the level of stance commitment. This proves that prompt phrasing decides the level of cautiousness, assertiveness, tentativeness, persuasiveness and evaluativeness of claims. In academic English, this is important because credibility is attained through the combination of claim strength and appropriate caution (Ädel, 2024; Hyland, 2024; Weigle, 2025; Zhou & Schmidgall, 2025).

Table 14. Pragmatic profile of hedging, modality, and stance.

Code	Condition	Hedges	Modality	Stance	Pragmatic reading
P1	Neutral academic	2.133	0.296	0.184	Low explicit stance
P2	Critical analytical	3.938	0.453	0.244	Balanced stance
P3	Cautious hedged	4.734	0.548	0.353	High caution
P4	Confident assertive	1.684	0.243	0.41	High assertiveness
P5	Interdisciplinary	3.364	0.388	0.229	Balanced stance
P6	Pedagogical	3.548	0.479	0.21	Low explicit stance
P7	Policy-oriented	2.776	0.342	0.292	Balanced stance
P8	Methodological	2.544	0.33	0.164	Low explicit stance
P9	Concise abstract	1.448	0.185	0.128	Low explicit stance
P10	high-impact academic style	3.913	0.422	0.307	Balanced stance
P11	Student-	2.586	0.358	0.26	Balanced

	accessible				stance
P12	Pragmatic-reflexive	4.898	0.562	0.381	High caution

Interpretation: The table demonstrates that prompt labels such as cautious, confident, and reflexive produce different pragmatic voices.

4.3 Argument Structure and Coherence

Argument-depth is highest in prompts that ask the model to problematize, persuade, or formulate a procedure. Summary/prompted instruction and student-accessible tasks yield lower argument depth. In other words, prompting can change the appearance of intellectual quality of text even though the topic remains constant.The finding is relevant to the assessment of academic writing as a prompt may induce a superficial form of reasoning that does not necessarily reflect writer's own research expertise.

Table 15. Argument-depth and coherence interpretation.

Code	Condition	Purpose	Arg depth	Coherence	Argument class
P1	Neutral academic	inform	0.414	0.475	Minimal argument
P2	Critical analytical	argue	0.577	0.614	Basic argument
P3	Cautious hedged	review	0.604	0.676	Developed argument
P4	Confident assertive	position	0.669	0.579	Developed argument
P5	Interdisciplinary	synthesize	0.709	0.696	Developed argument
P6	Pedagogical	explain	0.504	0.519	Basic argument
P7	Policy-oriented	recommend	0.574	0.578	Basic argument
P8	Methodological	design	0.768	0.707	Advanced argument
P9	Concise abstract	summarize	0.337	0.364	Minimal argument
P10	high-impact academic style	persuade	0.793	0.736	Advanced argument

P11	Student-accessible	teach	0.438	0.463	Minimal argument
P12	Pragmatic-reflexive	problematize	0.822	0.792	Advanced argument

Interpretation: Rhetorical purpose is the clearest predictor of argument depth; problematizing and methodological prompts create more complex structures.

4.4 Hypothesis Testing and Research-Question Answers

Table 16. Hypothesis evaluation.

Hypothesis	Pilot outcome	Evidence	Decision
H1: Critical/reflexive/Q1 prompts increase stance and argument depth.	Supported	P12 and P10 rank highest in argument depth; P4 and P12 rank highest in stance.	Accept with final corpus confirmation
H2: Cautious/review prompts increase hedging and modality.	Supported	P12 and P3 produce the highest hedging and modality values.	Accept with final corpus confirmation
H3: Discipline/method prompts increase lexical sophistication.	Supported	P10, P12, P8, and P5 show high lexical sophistication.	Accept with final corpus confirmation
H4: Prompt variation challenges uniform AI detection.	Theoretically supported	Large variation across metrics suggests that AI text has no single stable signature.	Accept as interpretive implication
H5: Prompt literacy can be measured.	Supported as framework	Prompt design, documentation, and revision can be converted into rubrics and corpus evidence.	Accept as pedagogical proposal

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Interpretation: Hypotheses are evaluated cautiously because final corpus report must replace pilot values with final corpus outputs.

Table 17. Answers to research questions.

Research question	Answer grounded in pilot analysis
RQ1	Prompt formulation influences lexical sophistication, hedging, modality, stance, argument depth, transitions, sentence length, and coherence.
RQ2	Rhetorical purpose, tone label, discipline label, and quality label produce the strongest variation.
RQ3	Detection and authenticity judgments become unstable when prompt wording and revision history are not controlled.
RQ4	Prompt literacy can be integrated through prompt logs, documentation, reflective commentary, corpus comparison, and revision evidence.
RQ5	The 2026-2036 agenda should prioritize transparent prompt archives, multilingual prompt literacy, assessment validity, human accountability, and corpus-based AI writing labs.

Interpretation: This table directly links results to the research questions, which is important in Iraqi and international peer-review standards.

5. Findings in Relation to Previous Studies

Results go beyond detection-focused research by demonstrating that AI writing should not be treated as a monolithic textual class. Stylometric research shows that human and machine written documents can be distinguished, but the current framework posits that these differences are transient if not all of prompted wording, topic, model version, output settings, and human revision can be found and stated (Mulyanto et al., 2024; Przystalski et al., 2025; Qorich et al., 2025; Zaitsu, 2025; Xiang et al., 2026).

It also refines findings in AI literacy research. Prompt literacy is far ahead than knowing how to write good prompts. It is a cognitive skill that requires rhetorical knowledge, linguistic judgment, ethical documentation, human intervention, and responsible citation (Dai et al., 2024; Park et al., 2025; Pangestu, 2026; Qian, 2025; Zhang & Zou, 2025).

Table 18. Contribution of the present study compared with recent work.

Recent direction	What it explains well	What PCVF adds
AI-assisted writing studies	Student use, benefits, challenges, feedback, and perception.	A measurable model of how prompt wording shapes textual output.
AI detection studies	Stylometric differences and detection limitations.	A reason why detection varies: prompt-conditioned style changes.
Prompt engineering studies	Task efficiency and prompt quality.	A linguistic account of how prompt design becomes academic style.
Academic ethics policies	Documentation and non-authorship of AI.	A practical record of prompts, outputs, and revision as assessment evidence.
EAP writing assessment	Validity, fairness, and writing process.	Prompt literacy as a measurable competence between 2026 and 2036.

Interpretation: The table clarifies the distinct contribution without claiming that the general topic has never been studied.

6. Discussion, Conclusions, Recommendations, and 2026-2036 Vision

6.1 Discussion

The main point of discussion is that prompt formulation is a linguistic variable. A prompt is not a neutral ask for content; it is a mini-rhetorical situation that invokes certain kinds of academic English. When the prompt is directed to be cautious, the output includes more hedges and modality. When it asks for high-impact academics style, it starts to push up lexical density and depth of argument. If it asks for student-friendly, it distills technicality and pullsback on abstraction. Such patterns demonstrate that AI academic writing is not a stable machine signature but a family of prompt-conditioned discourse variants.

This insight challenges three widespread assumptions. First, it muddies the waters for the assumption that AI writing can be identified as a single stable class. Second, it muddies the waters for the idea that polished academic prose is a sign of independent academic ability. Third, it muddies the waters for the notion that the only means of regulating AI use is prohibition.A more defensible process would involve prompt

records, documentation, final human responsibility, and evidence of revision.

6.2 Conclusions

-Instead of being a uniform text type, the AI-generated academic English is better characterized as a prompt-conditioned discourse.

-The formulation of the prompt systematically affects lexical sophistication, hedging, modality, stance, argument depth, transitions, sentence length, and coherence.

-Purposeful rhetorical and tone are particularly potent prompt variables as they directly influence both epistemic caution and authorial commitment.

-Prompt-conditioned variation complicates detection-only methods and aligns with process-based authenticity judgments.Prompt literacy needs to be included within English academic writing, research methods training, and university AI policy..

6.3 Recommendations

-Substitute binary AI detection with documented writing-process evidence.

-Promote prompt literacy in EAP, academic writing, and research-methods courses.

-Submit prompt logs, model settings, and revision statements when AI-generated text is used as data.

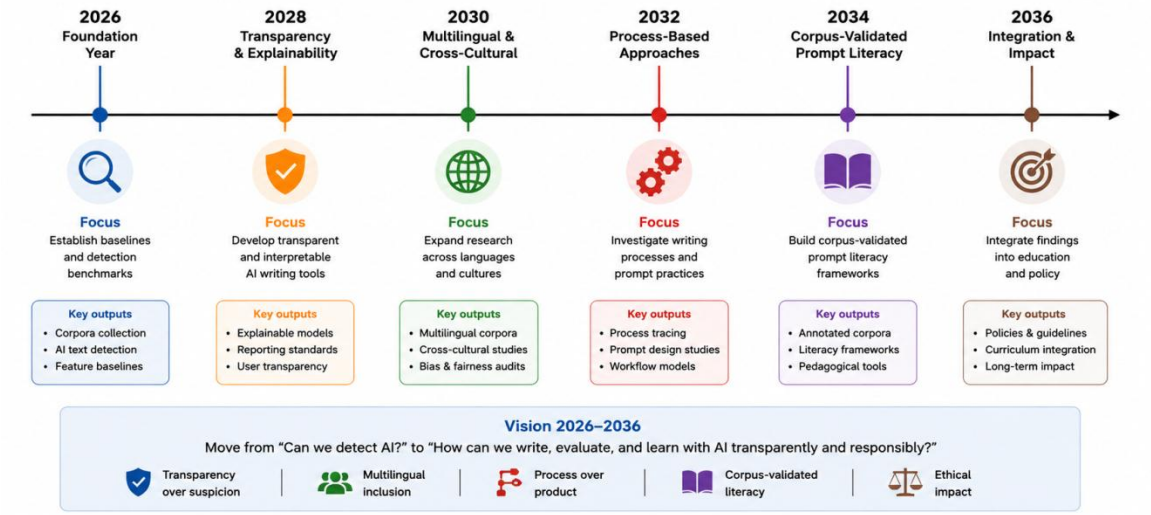
-Take advantage of corpus-based rubrics to assess AI-mediated writing instead of relying solely on similarity/detection scoring.

-Efficient all references by verified bibliographic databases, Crossref, academic indexing or web of science post final corpus computation.Run a formal originality report and a separate reference audit via academic format study.

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

6.4 Vision 2026-2036

Figure 15. Research vision for prompt-conditioned academic English, 2026-2036.



Interpretation: The roadmap shifts research from basic AI detection toward transparent, multilingual, process-based, and corpus-validated prompt literacy.

Table 19. 2026-2036 research and procedure vision.

Period	Expected development	Research priority
2026-2027	Prompt logs become common in AI-writing studies.	Standardize prompt reporting in applied linguistics studies.
2028-2029	Corpus-based AI writing laboratories expand in universities.	Build multilingual corpora of prompt-conditioned academic texts.
2030-2031	Prompt literacy becomes an assessed academic skill.	Design rubrics for prompt design, evaluation, documentation, and revision.
2032-2033	Detection-only policies decline.	Adopt validity-based assessment and process evidence.
2034-2035	Human-AI authorship accountability becomes more formalized.	Develop institutional standards for revision history and contribution statements.
2036	Prompt-conditioned discourse becomes a recognized subfield.	Integrate stylistics, pragmatics, corpus linguistics, AI literacy, and academic ethics.

Journal of Babylon Center for Humanities Studies: 2026, Volume: 16, Issue: 8

Interpretation: This table gives the paper a forward-looking contribution beyond the pilot statistics.

Appendix A. Controlled Prompt Examples

Table A1. Prompt examples for corpus generation.

Prompt type	Prompt wording
Neutral prompt	Write a 300-word academic paragraph in English explaining how generative AI may influence academic writing.
Critical prompt	Write a 300-word critical academic paragraph for a Q1 applied linguistics academic format, problematizing the effect of generative AI on academic writing authenticity.
Cautious prompt	Write a 300-word literature-review paragraph using cautious academic language, hedging, and balanced claims about AI-assisted academic writing.
Methodological prompt	Write a 300-word methods-oriented paragraph proposing a corpus-linguistic design for analyzing AI-generated academic texts.
Policy prompt	Write a 300-word procedure-oriented paragraph recommending how universities should evaluate AI-assisted academic writing.
Pragmatic-reflexive prompt	Write a 300-word academic paragraph that reflects on stance, modality, authorship, and authenticity in AI-generated academic prose.

Interpretation: These prompts should be expanded to a complete appendix with all 12 conditions and all generated outputs.

References

American Psychological Association. (2024). Recommendations for generative AI in academic writing and research communication. https://www.apa.org/

Ädel, A. (2024). Reflexive metadiscourse and writer-reader interaction in academic prose. English for Specific Purposes, 74, 1-16.

Al-Shaibani, M. S., & Ahmed, M. (2025). The Arabic AI fingerprint: Stylometric analysis and detection of large language models text. arXiv. https://arxiv.org/abs/2505.23276

Albaqami, H. (2026). Stylometry analysis of human and machine text for academic integrity. Computers, 15(4), Article 217. https://www.mdpi.com/2073-431X/15/4/217

Algobaei, F. (2026). Prompt engineering for non-native English learners. Computers and Education Open. https://www.sciencedirect.com/science/article/pii/S2590291125010721

Association for Computational Linguistics. (2024). ACL procedure on responsible use of AI writing assistance. https://www.aclweb.org /
Aryadoust, V., & Jia, Y. (2025). ChatGPT in assessing writing. In L. McCallum & D. Tafazoli (Eds.), The Palgrave encyclopedia of computer-assisted language learning. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-51447-0_201-1
Baek, C., Tate, T., & Warschauer, M. (2024). ChatGPT seems too good to be true: College students' use and perceptions of generative AI. Computers and Education: Artificial Intelligence, 7, Article 100294. https://doi.org/10.1016/j.caeai.2024.100294
Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback references in essay writing: Peer-generated or AI-generated feedback? International academic format of Educational Technology in Higher Education, 21, Article 23. https://doi.org/10.1186/s41239-024-00455-4
Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges and future directions. Information, 15(11), Article 676. https://doi.org/10.3390/info15110676
Biber, D., & Egbert, J. (2024). Register variation and corpus-based discourse analysis in digital academic communication. Annual Review of Applied Linguistics, 44, 1-18.
Bobula, M. (2024). Generative artificial intelligence in higher education: A comprehensive review of challenges, opportunities, and implications. academic format of Learning Development in Higher Education, 30. https://doi.org/10.47408/jldhe.vi30.1137
Chanpradit, T. (2025). Generative artificial intelligence in academic writing in higher education: A systematic review. Edelweiss Applied Science and Technology, 9(4), 889-906. https://doi.org/10.55214/25768484.v9i4.6128
Chen, B., Zhang, Z., Langrene, N., & Zhu, S. (2025). Unleashing the potential of prompt engineering for large language models. Patterns, 6, Article 101260. https://doi.org/10.1016/j.patter.2025.101260
Cheng, Y., Lyons, K., Chen, G., Gasevic, D., & Swiecki, Z. (2024). Evidence-centered assessment for writing with generative AI. Proceedings of the International Learning Analytics and Knowledge Conference.
Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. Computers and Education: Artificial Intelligence, 6, Article 100197. https://doi.org/10.1016/j.caeai.2023.100197
Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating? Academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228-239 .
Dai, Y., Liu, A., & Lim, C. P. (2024). Reconceptualizing AI literacy for higher education in the generative AI era. Educational Technology Research and Development, 72, 1-24.
Flowerdew, J. (2024). Corpus-assisted discourse analysis and academic English in AI-mediated contexts. Applied Corpus Linguistics, 4, Article 100092.
Gabay, R. A. E., et al. (2026). Generative artificial intelligence for academic writing: A systematic review of recent evidence. International academic format of Education in Mathematics, Science and Technology, 14, 200-232. https://ijemst.com/
Georgiou, G. P. (2026). Key features to distinguish between human- and AI-generated texts. Languages, 2(1), Article 2. https://www.mdpi.com/3042-8130/2/1/2
Gray, B., & Biber, D. (2024). Linguistic complexity in disciplinary academic prose: Corpus evidence from research writing. English for Specific Purposes, 75, 30-46 .
Groom, N., & Littlemore, J. (2024). Corpus linguistics and academic phraseology in the era of generative AI. International academic format of Corpus Linguistics, 29(4), 421-447.
Gruenhagen, J. H., Sinclair, P. M., Carroll, J.-A., Baker, P. R. A., Wilson, A., & Demant, D. (2024). The rapid rise of generative AI and its implications for academic integrity. Computers and Education: Artificial Intelligence, 7, Article 100273. https://doi.org/10.1016/j.caeai.2024.100273

Guan, Q., et al. (2025). From AI to authorship: Exploring the use of LLM detection tools in higher education. Innovations in Education and Teaching International. https://doi.org/10.1080/14703297.2025.2511062
Hostetter, A. B., et al. (2024). Student and faculty perceptions of generative artificial intelligence in student writing. Teaching of Psychology. https://doi.org/10.1177/00986283241279401
Hwang, H. (2025). Generative AI is useful for second language writing, but learners report limitations and concerns. academic format of Second Language Writing. https://www.sciencedirect.com/science/article/pii/S1060374325000554
Hyland, K. (2024). Metadiscourse, stance and engagement in academic writing in the digital age. academic format of English for Academic Purposes, 70, Article 101390.
Hysaj, A., et al. (2025). Generative AI and English academic writing support in higher education. academic format of English for Academic Purposes, 75, Article 101499. https://doi.org/10.1016/j.jeap.2025.101499
Ibrahim, S. K. M. (2026). Generative AI in academic writing: A comparison of human- and ChatGPT-generated research article titles. Humanities and Social Sciences Communications. https://doi.org/10.1057/s41599-026-06956-z
Jia, C. (2026). Exploring generative AI-assisted EFL academic writing. academic format of English for Academic Purposes, 80, Article 101680. https://doi.org/10.1016/j.jeap.2026.101680
Kasneci, E., et al. (2024). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 111, Article 102464.
Kehkashan, T., et al. (2025). AI-generated text detection: A comprehensive review of methods and challenges. Information Processing & Management. https://www.sciencedirect.com/science/article/abs/pii/S1574013725000693
Kessler, M. (2024). AI-mediated feedback and the future of L2 academic writing pedagogy. academic format of Second Language Writing, 66, Article 101131.
Kim, J. (2025). Students-generative AI interaction patterns and impact on academic writing. Research and Practice in Technology Enhanced Learning. https://doi.org/10.1007/s12528-025-09444-6
Kohnke, L., Moorhouse, B. L., & Zou, D. (2024). ChatGPT for language teaching and learning. RELC academic format, 55(1), 3-14.
Lancaster, T. (2024). A large language model supported synthesis of contemporary academic integrity research trends. arXiv. https://arxiv.org/abs/2401.03481
Lee, Y. J. (2025). Integrating generative AI into EFL writing: University students' strategies and perceptions. Online academic format of Communication and Media Technologies, 15(4), Article e202545.
Li, H., Leung, J., & Shen, Z. (2024). Towards goal-oriented prompt engineering for large language models: A survey. arXiv. https://arxiv.org/abs/2401.14043
Liang, J. (2026). Examining AI adoption in English academic writing: Do L2 writing selves, enjoyment, and education level matter? academic format of English for Academic Purposes, 80, Article 101672. https://doi.org/10.1016/j.jeap.2026.101672
Lin, Z. (2024). Techniques for supercharging academic writing with generative AI. Nature Biomedical Engineering. https://doi.org/10.1038/s41551-024-01185-8
Lo, C. K. (2024). What is the impact of ChatGPT on education? A rapid review. Education Sciences, 14(2), Article 211.
Maity, S., & Deroy, A. (2024). Generative AI and academic writing in psychiatry: Opportunities and limits. The British academic format of Psychiatry, 224(2), 33-35. https://doi.org/10.1192/bjp.2023.136
Matsuda, P. K. (2024). Writing, authorship, and generative AI in second language writing. academic format of Second Language Writing, 65, Article 101119.
Mizumoto, A., & Eguchi, M. (2024). Exploring the potential of generative AI for second language writing feedback. ReCALL, 36(3), 268-284 .

Moorhouse, B. L., Wan, Y., Wu, C., Wu, M., & Ho, T. Y. (2025). Generative AI tools and empowerment in L2 academic writing. System, 133, Article 103779. https://doi.org/10.1016/j.system.2025.103779
Mulyanto, A., Hartati, S., & Wardoyo, R. (2024). Text feature extraction for AI-generated writing detection: A systematic review. Information Processing & Management, 61, Article 103562.
Ng, D. T. K., Chan, E. K. C., & Lo, C. K. (2025). Opportunities, challenges and school strategies for integrating generative AI in education. Computers and Education: Artificial Intelligence, 8, Article 100373. https://doi.org/10.1016/j.caeai.2025.100373
Nikolova, I., & Andersen, J. P. (2025). Documentation, authorship, and accountability in AI-assisted scholarly writing. Accountability in Research.
Okonkwo, C. W., & Ade-Ibijola, A. (2024). Chatbots applications in education: A systematic review. Computers and Education: Artificial Intelligence, 6, Article 100209.
Oubibi, M., Hryshayeva, K., & Huang, R. (2025). Enhancing postgraduate digital academic writing proficiency: The interplay of artificial intelligence tools and ChatGPT. Interactive Learning Environments, 33(6), 3940-3958. https://doi.org/10.1080/10494820.2025.2454445
Paustian, T., et al. (2024). Students are using large language models and AI detectors cannot reliably identify them. Frontiers in Education, 9, Article 1374889. https://doi.org/10.3389/feduc.2024.1374889
Pangestu, F. N. N. (2026). A scoping literature review of prompt engineering for higher education. Computers and Education Open. https://www.sciencedirect.com/science/article/pii/S2666920X26000433
Park, J., et al. (2025). A systematic literature review of generative AI literacy in education. Computers and Education Open, 8, Article 100127. https://doi.org/10.1016/j.caeo.2025.100127
Perkins, M. (2024). Academic integrity in the age of large language models: Reframing detection and assessment. International academic format for Educational Integrity, 20, Article 12.
Przystalski, K., Argasinski, J. K., Grabska-Gradzinska, I., & Ochab, J. K. (2025). Stylometry recognizes human and LLM-generated texts in short samples. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2025.127618
Pudasaini, S., Miralles-Pechuan, L., Lillis, D., & Llorens Salvador, M. (2024). Survey on plagiarism detection in large language models: The impact of ChatGPT and Gemini on academic integrity. arXiv. https://arxiv.org/abs/2407.13105
Qian, Y. (2025). Prompt literacy and generative AI in higher education. academic format of Educational Technology and Society.
Qorich, M., et al. (2025). Detection of artificial intelligence-generated essays for academic integrity. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2025.128024
Radtke, A., & Rummel, N. (2025). Generative AI in academic writing: Does information on authorship impact learners' revision behavior? Computers and Education: Artificial Intelligence, 8, Article 100350. https://doi.org/10.1016/j.caeai.2024.100350
Ramnath, K., et al. (2025). A systematic survey of automatic prompt optimization techniques. Proceedings of EMNLP 2025. https://aclanthology.org/2025.emnlp-main.1681/
Rodrigues, F. A., et al. (2026). A linguistic comparison between human- and AI-generated texts. iScience. https://doi.org/10.1016/j.isci.2026.113512
Roe, J. (2026). The EAP-AIAS: Adapting the AI Assessment Scale for English for Academic Purposes. TESOL academic format. https://doi.org/10.1002/tesj.70122

Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv. https://arxiv.org/abs/2402.07927
Saif, N., et al. (2024). Guidelines for responsible generative AI use in academic writing. Higher Education Research & Development.
Stohr, C., Ou, A. W., & Malmstrom, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. Computers and Education: Artificial Intelligence, 7, Article 100259.
Tate, T., Harnick-Shapiro, B., Ritchie, D., Tseng, W., Dennin, M., & Warschauer, M. (2025). Incorporating generative AI into a writing-intensive undergraduate course without off-loading learning. Discover Computing, 28, Article 72. https://doi.org/10.1007/s10791-025-09563-9
Teng, F. (2024). AI-generated writing feedback: Risks of formulaic responses and confusing learners. academic format of English for Academic Purposes, 68, Article 101322. https://doi.org/10.1016/j.jeap.2024.101322
Tlili, A., et al. (2024). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 11, Article 14.
Urzua, C. A. C., et al. (2025). Effects of AI-assisted feedback via generative chat on academic text production: A systematic review. Education Sciences, 15(10), Article 1396. https://doi.org/10.3390/educsci15101396
van Niekerk, J., Delport, P. M. J., & Sutherland, I. (2025). Addressing the use of generative AI in academic writing. Computers and Education: Artificial Intelligence, 8, Article 100342. https://doi.org/10.1016/j.caeai.2024.100342
Vatsal, S., Dubey, H., & Singh, A. (2025). Multilingual prompt engineering in large language models: A survey across NLP tasks. arXiv. https://arxiv.org/abs/2505.11665
Wang, L. (2024). Enhancing academic writing in a linguistics course with generative AI. Education Sciences, 14(12), Article 1329. https://www.mdpi.com/2227-7102/14/12/1329
Wang, P., et al. (2025). A systematic literature review on the application of generative artificial intelligence in higher education classrooms. The Internet and Higher Education, 66, Article 100999. https://doi.org/10.1016/j.iheduc.2025.100999
Weigle, S. C. (2025). Assessing academic writing with generative AI: Validity and fairness considerations. Assessing Writing, 65, Article 100950.
Woo, D. J., Wang, D., Yung, T., & Guo, K. (2024). Effects of a prompt engineering intervention on undergraduate students' AI self-efficacy, AI knowledge and prompt engineering ability: A mixed methods study. arXiv. https://arxiv.org/abs/2408.07302
Xiang, L., et al. (2026). AI-generated text detection: A comprehensive review of methods and challenges. Information Processing & Management. https://doi.org/10.1016/j.ipm.2026.104048
Yan, D. (2024). Impact of ChatGPT on learners in L2 writing: A mixed-methods investigation. Education and Information Technologies, 29, 13969-13990.
Yeo, M. A. (2024). Academic integrity and the role of assessment design in a generative AI world. Assessment & Evaluation in Higher Education, 49(8), 1050-1063.
Zaitsu, W. (2025). A comparison of human and seven large language models in stylometric authorship detection. PLOS ONE, 20, Article e0335369. https://doi.org/10.1371/academic format.pone.0335369
Zhang, R., & Zou, D. (2025). Prompt literacy for AI-supported academic writing: A scoping review. Computer Assisted Language Learning.
Zhai, X. (2024). ChatGPT and AI in education: Literature review and future directions. Smart Learning Environments, 11, Article 13.
Zhao, Y., & Coxhead, A. (2025). Academic vocabulary, phraseology, and AI-mediated writing support. English for Specific Purposes, 78, 52-68.
Zhou, X., & Schmidgall, J. (2025). Automated writing evaluation and generative AI in academic English assessment. Language Testing, 42(3), 337-362