

Computer Assisted Language Learning

ISSN: 0958-8221 (Print) 1744-3210 (Online) Journal homepage: www.tandfonline.com/journals/ncal20

Enhancing EFL argumentative writing through an AI-powered corpus: impact on learner writing proficiency

Lihang Zhu, Yanfang Su, Simin Xu & Kanglong Liu

To cite this article: Lihang Zhu, Yanfang Su, Simin Xu & Kanglong Liu (16 Dec 2025): Enhancing EFL argumentative writing through an AI-powered corpus: impact on learner writing proficiency, Computer Assisted Language Learning, DOI: 10.1080/09588221.2025.2599152

To link to this article: https://doi.org/10.1080/09588221.2025.2599152

Published online: 16 Dec 2025.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ncal20

COMPUTER ASSISTED LANGUAGE LEARNING
https://doi.org/10.1080/09588221.2025.2599152

Routledge
Taylor & Francis Group

Check for updates

Enhancing EFL argumentative writing through an AI-powered corpus: impact on learner writing proficiency

Lihang Zhu[a], Yanfang Su[b], Simin Xu[c] and Kanglong Liu[c,c]

[a]College of Foreign Languages and Cultures, Xiamen University, Xiamen City, Fujian Province, China[b]Department of Translation, Lingnan University, Hong Kong, China; [c]Department of Language Science and Technology, The Hong Kong Polytechnic University, Kowloon, Hong Kong SAR, China

ABSTRACT

Argumentative writing skills are fundamental to academic literacy in higher education. However, English as a Foreign Language (EFL) learners often face significant challenges in producing coherent and well-structured argumentative texts. While previous research has explored the potential of corpus-based instruction and ChatGPT-assisted learning in enhancing writing skills, limited attention has been paid to the role of AI-powered corpus tools in improving students' argumentative writing performance. This study investigated the effectiveness of corpus-based language pedagogy (CBLP) integrated with an AI-powered corpus platform to enhance students' argumentative writing skills. A quasi-experimental design was employed, with 24 undergraduate students from a Chinese university participating in the study. All participants completed the training sessions and writing tasks, and their performance data were analyzed. The results revealed a statistically significant improvement in students' overall argumentative writing performance over time. Additionally, a qualitative analysis of student interviews indicated positive attitudes toward the AI-powered corpus tool, along with the challenges in its application. The study concludes by discussing the pedagogical implications of AI-powered corpus tools for argumentative writing instruction and by offering suggestions for future research.

ARTICLE HISTORY
Received 28 February 2025
Accepted 30 November 2025

KEYWORDS
AI-powered corpus; argumentative writing; corpus-based language pedagogy; corpus-based writing instruction

1. Introduction

Argumentative writing is a key component of academic literacy, particularly in higher education settings. Through argumentative writing, students develop analytical and communicative skills (Andrews, 2009), which are essential for academic and professional success (Preiss et al., 2013; Qin & Karabacak, 2010).

Despite its importance, argumentative writing poses challenges for language learners at the linguistic, structural, and organizational levels (Hirose, 2003; Su et al., 2021). Researchers have explored various tools and strategies to address these

CONTACT Kanglong Liu klliu@polyu.edu.hk Department of Language Science and Technology, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China
© 2025 Informa UK Limited, trading as Taylor & Francis Group

challenges, with corpora emerging as one of the most effective tools. Corpora provide authentic, evidence-based resources that can enhance vocabulary, improve the use of collocations, and refine grammar and sentence structures in writing (Crosthwaite & Steeples, 2022; Flowerdew, 2015; Vannestål & Lindquist, 2007; Wu & Yeh, 2025). As a result, corpora have been widely adopted to support students' acquisition of lexico-grammatical patterns in different genres, including argumentative writing (Flowerdew, 2010; Karpenko-Seccombe, 2020). Annotated rhetorical moves, i.e., text segments serving particular communicative purposes (Wu et al., 2024), in corpora have also been proposed to help students understand different argument structures (Karpenko-Seccombe, 2024). However, corpora have inherent limitations. Corpus users need to engage with large volumes of data directly. As a result, users can feel overwhelmed by the amount of information, especially when they lack the technical and cognitive skills needed to successfully query, navigate, and analyze the corpus results. Moreover, the unsorted corpus content may not always align with students' needs (Crosthwaite & Baisa, 2024; Crosthwaite & Boulton, 2023).

Given these limitations, recent advancements in AI systems, such as ChatGPT, open new avenues for enhancing argumentative writing instruction. While AI systems also rely on large datasets like corpus tools, they are trained to develop predictive models, which then process and present information to users in a simpler way that caters to their needs. Research has demonstrated that AI-powered chatbots can effectively improve students' writing skills by providing immediate feedback on reasoning and evidence integration (Guo et al., 2022; Su et al., 2023). Furthermore, combining AI-powered chatbots with teacher support in classrooms has been found to boost student motivation (Chiu et al., 2023). However, AI-powered tools also have limitations. The effectiveness of ChatGPT is highly dependent on the quality of user prompts (Yeadon et al., 2023), and concerns remain regarding the reliability and accuracy of its generated content (Zhou & Hou, 2024).

To address these concerns, integrating corpora with AI tools offers a promising solution. Corpora provide EFL learners with authentic language data, including vocabulary, grammar patterns, and examples in context, which support the development of language proficiency. AI tools enable students to analyze large volumes of corpus data efficiently and respond to a wide range of queries, facilitating deeper engagement with the language. In this context, we developed an AI-powered corpus platform that combines corpus functionalities with AI assistance. This study investigated the potential of using this AI-powered corpus platform in corpus-based writing instruction to enhance students' argumentative writing. The findings offer valuable insights into the innovation of corpus tools and corpus-based language pedagogy in the AI era.

2. Literature review

2.1. Current practices in argumentative writing instruction

Argumentative writing is a critical skill in higher education, and its importance has been widely recognized and examined from multiple perspectives across various academic disciplines (Uccelli et al., 2012). Despite its significance, students often

face various challenges. Producing high-quality argumentative texts requires a solid foundation in the linguistic (Abdelrahim & Abdelrahim, 2020; Liu & Braine, 2005), structural and organizational (Hirose, 2003; Qin & Karabacak, 2010; Zhu, 2001), and dialogical aspects of argumentation (Su et al., 2021). To address the challenges faced by students in argumentative writing, various instructional methods have been proposed.

To improve students' argumentative writing, a range of frameworks of argument structure have been used in instruction. A widely adopted entry-level model is the CER framework proposed by McNeill and Krajcik (2011), comprising Claim, Evidence, and Reasoning. Designed originally for primary students, it typically adds Rebuttal at the middle-school level (McNeill & Martin, 2011). The CER framework has also been applied in L2 writing instruction (Su et al., 2021). Beyond CER, a more sophisticated approach is Toulmin's model (1958), which has been highly influential in teaching argumentative writing (Karbach, 1987; Kneupper, 1978; Yang & Pan, 2023). Toulmin's model (1958) consists of six components, including claim, grounds, warrant, backing, qualifier, and rebuttal, among which claim, grounds, and warrants are the most basic elements. Specifically, claim is the main position of the arguments, grounds are the evidence or facts that support the claim, and warrant is the underlying assumptions that link claim with grounds. The other three elements are secondary: backing supplies the warrant, rebuttal recognizes conditions or counter-cases that could undermine the claim, and qualifier is used to adjust the strength or scope of the claim's wording. Despite the widespread use of Toulmin's model, studies show that students struggle with its secondary elements. Qin and Karabacak (2010) found that Chinese EFL learners handle the basic components—claims and data—relatively well but often neglect more complex ones such as counterarguments and rebuttals. Hu and Liu (2025) further noted that acquiring argument structures is a lengthy process and suggested that the lower frequency of the secondary elements stems from their inherent complexity.

Another approach, collaborative writing, shifts students' focus to both structural and dialogical aspects of argumentation. Collaborative argumentation, in both written and oral forms, encourages students to consider multiple perspectives, reconcile opposing views, and construct integrated arguments, thereby fostering critical reasoning skills (Chin & Osborne, 2010; Su et al., 2021). However, collaborative writing also presents challenges. It often prioritizes structural integrity and audience awareness over linguistic accuracy, which can lead to difficulties with grammar, lexical variety, and academic phrasing. Moreover, identifying suitable collaborators and managing interpersonal dynamics, such as coordination issues and affective tensions, can limit its effectiveness, ultimately reducing both motivation and task efficiency (Chen & Lee, 2022).

2.2. Corpus use in language instruction

Corpora have been regarded as a valuable resource in language teaching and learning since the 1980s (Lee & Swales, 2006; Yoon & Jo, 2014). They can be integrated into language instruction through both indirect and direct approaches (Boulton & Vyatkina, 2021). In the indirect approach, corpus-derived information

is used as learning material. The direct approach, often referred to as data-driven learning (DDL), is defined as 'the hands-on use of authentic corpus data (concordances) by advanced, sophisticated foreign or second language learners in higher education for inductive, self-directed language learning of advanced usage' (Boulton, 2011). However, researchers have noted that most DDL studies have been conducted by teacher-researchers, with limited adoption of corpus tools by non-researcher teachers (Ma et al., 2024c; Pérez-Paredes, 2022). Furthermore, much of the research has focused on the affordances of corpus tools themselves rather than on their integration into broader language learning curricula (Pérez-Paredes, 2022). To address these limitations, Ma et al. (2022) advocated Corpus-Based Language Pedagogy (CBLP). Unlike DDL, which emphasizes learners' inductive 'self-exploration' of language data, CBLP focuses on the structured integration of corpora into teaching practices (Ma et al., 2024c). It highlights the critical role of teachers in using corpus tools to design and deliver curriculum-based language instruction. Ma et al. (2022) proposed a four-step CBLP approach, which was later expanded into a five-step approach in subsequent research (Ma et al., 2024c). The four-step CBLP begins with assessing students' prior knowledge and raising awareness of key language features, followed by corpus analysis to observe language patterns. Next, teachers guide students in summarizing these patterns. Finally, students consolidate their language skills through output-based exercises.

Early research on corpus use in language instruction has consistently demonstrated the effectiveness of DDL, particularly in enhancing learners' acquisition of collocations and vocabulary (Liu & Jiang, 2009; Pérez-Paredes, 2022). More recent studies have moved beyond focusing solely on lexical-grammatical patterns. Cotos et al. (2017) found that exploring corpus data can help novice writers foster the awareness of the linguistic features that realize genre-specific communicative purposes, improve structural organization, and develop procedural knowledge of genre. Furthermore, corpus consultation can promote autonomous learning (Bernardini, 2000). By encouraging students to independently engage with authentic language data, corpus-based instruction fosters self-directed learning while sustaining students' interest and motivation (Tekin & Soruç, 2016). Ma et al. (2024a) examined the effectiveness of a CBLP-supported pronunciation course targeting both segmental and suprasegmental features to improve students' pronunciation accuracy and fluency. Significant improvements were observed in reading accuracy, phonetic precision, and pronunciation fluency. Moreover, students reported positive perceptions of the CBLP-supported approach.

Despite the positive findings on corpus use in language instruction, several challenges persist. While corpus-based methods have been employed to raise students' awareness of linguistic features by enabling them to analyze authentic texts and identify language patterns, their application also has limitations. Researchers argue that corpus-based instruction primarily focuses on enhancing lower-level language patterns (Pérez-Paredes, 2022), such as lexical bundles (Eriksson, 2012; Flowerdew, 2015) and reporting verbs (Bloch, 2009), and is less frequently used for developing higher-order skills such as argument structure and reasoning (Charles, 2011). Although some studies have explored corpora's potential for analyzing rhetorical patterns (Karpenko-Seccombe, 2024), empirical evidence remains limited. Moreover,

practical obstacles such as limited access to suitable corpora and the complexity of instructional design hinder the integration of corpora (Charles, 2011). Furthermore, the effectiveness of integration is contingent on stakeholders' skills: learners must be able to analyze data and apply insights (Boulton, 2009), while teachers need the capacity to develop suitable materials and provide scaffolding (Karpenko-Seccombe, 2024).

With technological advancements, chatbots have emerged as valuable tools for supporting dialogical engagement in argumentative writing while addressing challenges in peer collaboration. Unlike peer-based approaches, chatbots provide neutral, consistent, and accessible learning partners, reducing affective conflicts and allowing students to practice argumentation skills flexibly in autonomous or asynchronous settings (Guo et al., 2022). However, traditional chatbots rely on pre-programmed responses, which limits their adaptability (Rakshit et al., 2018). Retrieval-based systems ensure consistency but remain constrained by the scope of their databases (Guo et al., 2023). In contrast, generative AI chatbots, such as GPT-based models, enhance interactivity by dynamically generating context-aware responses, enabling more personalized engagement (Su et al., 2023). Research has shown that ChatGPT can improve students' English writing skills (Li et al., 2024), as well as their argumentation skills, critical thinking awareness, and collaborative tendencies (Darmawansah et al., 2025). However, their effectiveness depends heavily on the quality of user input, necessitating well-crafted prompts (Dang et al., 2022). Moreover, generative models may produce irrelevant or fabricated content, raising concerns about reliability (Su et al., 2023). There is also the risk that students may become overly reliant on AI tools without engaging in critical evaluation (Seo, 2024). This unpredictability underscores the need for careful adoption and guided use to maximize their benefits in academic writing.

In summary, while corpus-based methods, collaborative writing, and chatbot-assisted instruction each offer distinct advantages for teaching argumentative writing, they also have inherent limitations. To address these challenges, this study introduces an AI-powered corpus platform within a corpus-based instructional framework, aiming to enhance students' argumentative writing while mitigating the limitations of these individual approaches.

2.3. Purpose of the study

Building on prior research, the integration of ChatGPT with corpus tools shows considerable potential for teaching and learning argumentative writing. However, research on the effects of AI-powered corpus tools on students' development in argumentative writing remains limited. This study seeks to address this gap by examining students' performance and perceptions when using AI-powered corpus tools to learn argumentative writing. The following research questions guide this study:

1. How does the use of an AI-powered corpus affect students' development in argumentative writing over time?

2. How does the use of an AI-powered corpus affect students' argumentative writing performance across content-related and language-related dimensions?
3. How do students perceive the use of an AI-powered corpus platform in learning argumentative writing?

3. Methods

3.1. Participants

A total of 35 undergraduate students (13 male, 22 female) from a Chinese university voluntarily participated in the study. All participants were non-English majors, with 20 majoring in humanities and social sciences (e.g. Chinese, Japanese, history) and 15 majoring in engineering and technology (e.g. civil engineering, computer science, chemistry). Their scores on the university's English Language Placement Test ranged from 70 to 80, corresponding to the B1 proficiency level according to the Common European Framework of Reference for Languages (CEFR). A pre-study background survey confirmed that none of the participants had prior experience with corpus tools or formal instruction in English argumentative writing. Of the initial 35 participants, 24 (8 male, 16 female) completed all training sessions and writing tasks, providing valid data for analysis. These final participants retained a similar distribution of majors, with 14 students from humanities and social sciences and 10 from engineering and technology. Eleven students withdrew due to scheduling conflicts during final examinations.

3.2. The AI-powered corpus platform

The AI-powered corpus platform developed by the research team integrates generative AI (GPT-4) with corpus data to support the teaching and learning of argumentative writing for EFL learners. The corpus comprises two types of data, totaling 905,221 words (see Table 1). The first component includes 296 long expert articles from academic journals, newspapers, and magazines (e.g. the opinion columns of *Nature*, *Science*, *The New York Times*, and *The Economist*). These articles present complex and complete argumentative structures, demonstrating how reliable evidence is used to support arguments. The second component consists of 176 shorter exemplar argumentative essays (e.g. high-scoring TOEFL and IELTS writing

Table 1. Corpus data of the platform.

	Long expert articles		Short exemplar essays		
Topic	Number of articles	Mean words	Number of essays	Mean words	Total words
Education	40	902	52	318	52643
Government	47	828	21	324	45719
Technology	49	1166	15	333	62136
Environment	59	1019	8	265	62213
Sociology	58	1084	71	327	81718
Health	43	889	9	294	40892

samples), which serve as models for standardized test writing (Hartwell & Aull, 2022).

The platform provides three core functions: Keyword, N-gram, and Top-K. Each function focuses on a different aspect of argumentative writing. Users can search within the corpus or employ GenAI for additional insights.

The Keyword function allows users to look up word meanings and usages, helping them understand how specific words are employed in context (see Figure 1). All sentences containing the keyword are displayed on the right-hand side of the interface. In the top-left section, titled 'Corpus Search', the retrieval process employs a hybrid NLP pipeline. When a search is initiated, the corpus component first performs keyword-based indexing and syntactic filtering to retrieve a set of raw, unannotated sentences containing the query word. Following this retrieval, the Large Language Model (LLM) component takes over as the analysis engine. The LLM processes these raw sentences to conduct the fine-grained linguistic analysis, including tagging part-of-speech (POS), disambiguating word senses, and generating example sentences for each identified sense. In this architecture, the corpus provides the authentic contextual data, while the LLM provides the analytical intelligence to parse and explain it. The results of this LLM analysis are then displayed in a table within 'Corpus Search' section. In the bottom-left section, titled 'AI Search', another table presents results generated by GenAI independently, without referencing corpus data.

The N-gram function enables users to explore n-grams, i.e., 'a sequence of n-words' (Jurafsky & Martin, 2025), as they occur within specific argument structures. After the user enters a keyword and specifies the topic and argument structure, the system will extract all n-grams containing that keyword from the matched corpus and rank them in descending order of frequency (see Figure 2). This function bridges the gap between lexical knowledge and genre-appropriate application (Hyland, 2008). By exposing learners to these patterns, the function supports the internalization of

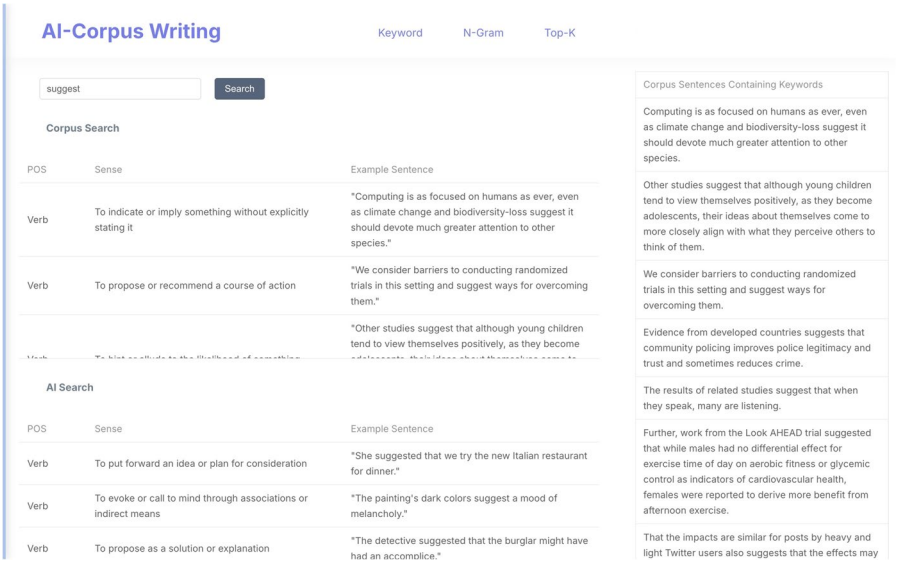

Figure 1. Keyword function.

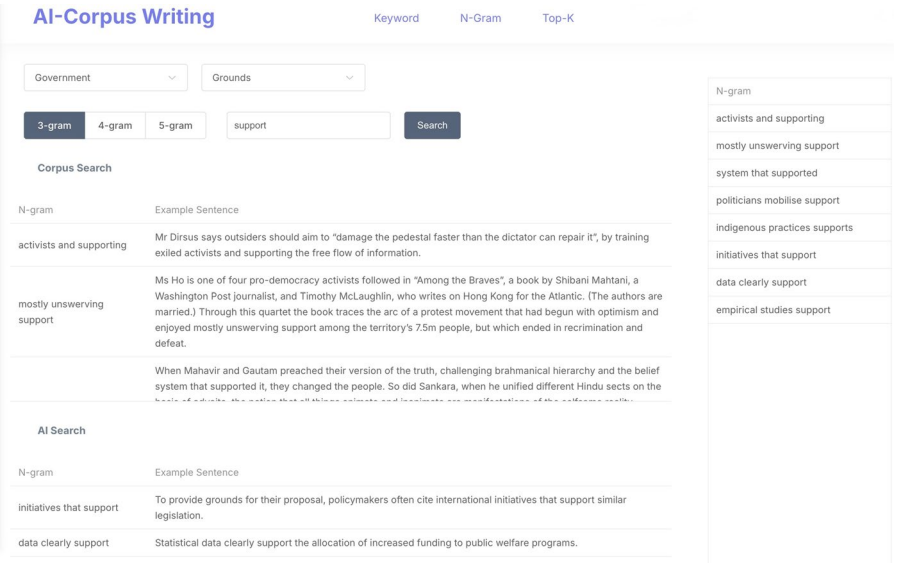

Figure 2. N-Gram function.

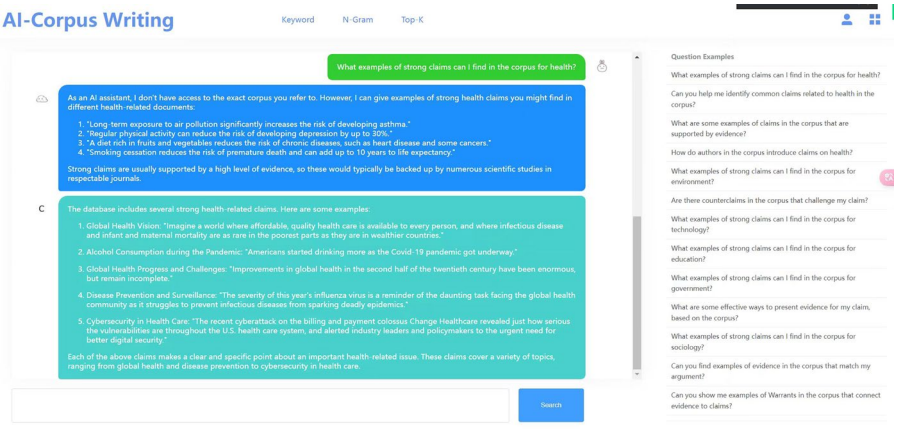

Figure 3. Top-K function.

native-like phrasing and reduces unnatural expressions (Daskalovska, 2013) and enhances textual cohesion by meeting established reader expectations (Halliday & Hasan, 1976).

The Top-K function (see Figure 3) illustrates an AI-based corpus search system designed to address students' queries by integrating two distinct response methods. The interface is structured into several key areas. On the left-hand side, the blue section presents AI-generated answers that rely solely on the AI's internal knowledge base, which provide diverse argumentative perspectives and maintain relevance to the student's query, especially when the small-scale corpus lacks sufficient data. In contrast, the Turquoise section presents answers generated through corpus-based Retrieval-Augmented Generation (RAG) framework (Lewis et al., 2021), which enhance a large language model by first retrieving relevant information from an

external knowledge source before generating a response. This corpus-based RAG system retrieves relevant information from the platform's corpus and synthesizes it into a coherent, contextually specific response. The Top-K function provides exemplary texts of different argumentative structures based on Toulmin's model of argumentation. To guide user interaction, a list of related 'Question Examples' is provided to guide users in formulating queries and to facilitate effective interaction with the system. By combining corpus-based insights with the diversity of AI-generated output, the system enables students to construct arguments that are both contextually relevant and grounded in real-world evidence, thereby fostering a deeper understanding of argumentation strategies.

3.3. Procedure

The study was conducted over a six-week period. Before the training course of Week 1, the participants received a detailed briefing on the research objectives, and informed consent was obtained. They then completed a pre-test and a background survey to assess their prior experience with corpus tools, AI technologies, and English argumentative writing. Following the pre-test in Week 1, the teacher introduced Toulmin's argumentative writing model and the functions of the platform.

From Week 1 to Week 5, the course was delivered online *via* Tencent Meeting. The online mode was chosen to accommodate participants from different campuses of the university, ensuring they could participate without being limited by time or location. Moreover, during the online sessions, students were asked to share their answers and the teacher monitored students' practices, offering guidance and feedback as needed to ensure effective use of the tools. The teaching schedules for each week focused on a different aspect of argumentative writing (see Figure 4). The argumentative writing course was designed based on the four-step CBLP model proposed by Ma et al. (2022) (see Figure 5). In the first step of each class, students' knowledge gaps in argumentative

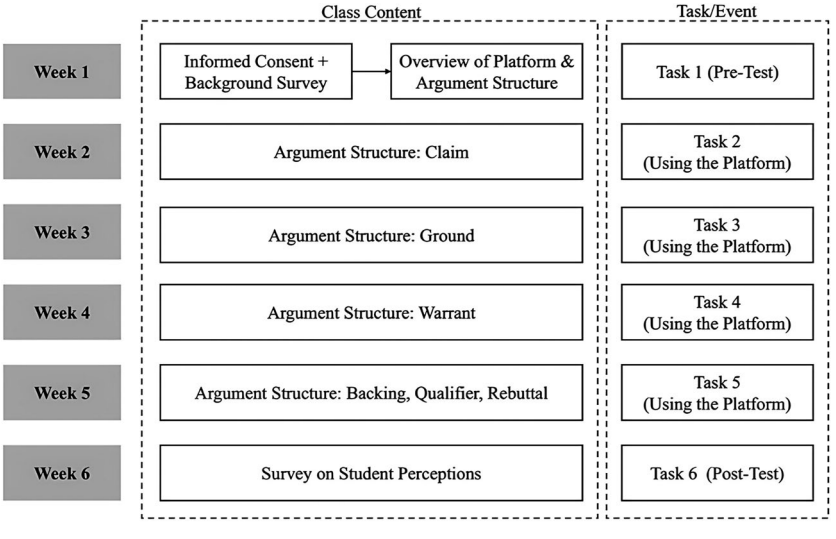

Figure 4. Research procedure.



Figure 5. The four-stage CBLP design.

writing were assessed using an instructional handout, and their awareness of the specific argumentative structures to be learned was raised. The second step involved hands-on corpus search practices. The teacher used model essays to explain argument structure. Students engaged in inductive discovery to summarize the key language features and structural elements of argumentative writing. The third step was an output exercise, during which students completed an argumentative writing task and received teacher feedback on the aspects of argumentative writing covered in the lesson. At the end of Weeks 2, 3, and 4, participants completed an argumentative writing task similar to the pre-test using the platform. Finally, students reflected on their use of the corpus and refined their argument development.

Upon completing the writing course in Week 5, 24 participants took a post-test, while 11 participants withdrew from the study due to final examinations. All five writing tasks were drawn from IELTS Task 2 and addressed different topics. For the pre-test and post-test, students completed a 45-minute timed writing task without external tools or reference materials. In contrast, the three practice tasks (Tasks 2, 3, and 4) allowed students to use the AI-powered corpus platform, and these tasks were not time-restricted. To ensure that students completed the tasks as instructed, they were asked to record their screens during all writing tasks. In Week 6, participants completed a questionnaire about their perceptions of using the platform. Nine participants, selected using a maximum variation sampling strategy, were invited to participate in semi-structured interviews to capture diverse developmental trajectories throughout the experiment (see Table 2 for participant profiles).

3.4. Data collection and analysis

Multiple sources of data were collected to answer the research questions, including 120 argumentative essays from the pre-test, post-test, and three writing tasks, reflection notes, and audio-recordings of nine interviews.

Table 2. Profile of focal participants.

Pseudo-name	Gender	Major	English argumentative writing experience	Corpus experience	AI-supported writing experience	Pre-Test	Post-Test
Danielle	Female	Chinese	No experience	No experience	Often use	20	22.5
Jerald	Male	Civil Engineering'	Have heard about basic argument structure but could not identify or develop proper organizational structure.	No experience	Occasionally use	18.5	23
Vicky	Female	Chemical Engineering	Have heard about basic argument structure but could not identify or develop proper organizational structure.	No experience	Occasionally use	14	25
Cici	Female	Engineering Experimental Program	No experience	No experience	Often use	19	18.5
Wendy	Female	Chinese	Little knowledge of argumentative writing, no experience in argumentative writing	No experience	Often use (not for argumentative writing)	17	23.5
Maya	Female	Journalism & Communication	Little knowledge of argumentative writing, no writing experience in argumentative writing	No experience	Occasionally use	20.5	22.5
Cyber	Male	Physics	Have heard about basic argument structure but could not identify or develop proper organizational structure.	No experience	Occasionally use	14.5	23.5
Charlie	Male	Japanese	No experience	No experience	Rarely use	14.5	24.5
Young	Female	Chemistry	No experience	No experience	Occasionally use	17.5	24.5

3.4.1. Argumentative writing

Writing topics were carefully selected from the official IELTS Writing Test (Task 2) with a recommended word limit of 250–500 words. To ensure consistency in familiarity and difficulty across participants, the tasks were reviewed by two experienced English teachers. Additionally, the five writing tasks covered diverse topics to minimize potential practice effects. The evaluation rubrics for the argumentative writing tasks were adapted from Lam et al. (2017) (see the Appendix 1). These rubrics assess content and language aspects of argumentative writing. Content-related criteria include task response, evidence support, rebuttal, and organization (the logical progression and coherence among different structural elements), while language-related criteria cover language use (the richness, variety, and

appropriateness of language), and grammar (accuracy of grammatical structure and spelling). Each criterion is scored on a 5-point scale, yielding a maximum possible score of 30 for each essay. Two experienced teachers independently evaluated the students' writing, assigning scores for each criterion according to the rubrics. Prior to the formal scoring process, five essays were randomly selected for calibration, during which the raters reviewed the essays together, discussing and resolving any discrepancies in their interpretation of the rubrics. Following this calibration, the two raters independently evaluated the remaining essays. The five writing tasks achieved an overall ICC of 0.766, indicating good reliability. To further minimize the influence of individual rater bias, the final score for each essay was obtained by averaging the ratings from both evaluators.

To answer the first research question concerning the influence of the AI-powered corpus platform on students' development in argumentative writing over time, descriptive statistics of students' overall writing performance in the five writing tasks were first presented. Following this, a one-way repeated measures ANOVA was conducted to examine differences in overall scores across the writing tasks. Before performing the ANOVA, the assumptions of normality ($p > .05$) and sphericity ($W = 0.685$, $p > .05$) were tested and satisfied. The analysis was performed using Python, with students' overall writing scores as the dependent variable and time points (the five writing tasks) as the independent variable. Subsequently, post-hoc tests were conducted to assess the effect size between two time points, where Cohen's d values of 0.2, 0.5, and 0.8 represent small, medium, and large effect sizes, respectively.

To address the second research question on the impact of AI-powered corpus on students' writing performance (see Appendix 2 and 3), descriptive statistics were calculated to summarize students' performance across different dimensions in the five writing tasks. Given that the assumption of normality was not met for most dimensions, non-parametric Friedman tests were used to assess differences across the writing tasks. Dunn's post-hoc tests were performed to evaluate the effect size between two time points, with a Bonferroni correction applied to control for multiple comparisons.

3.4.2. Interviews

To investigate students' perceptions of using the AI-powered corpus platform (RQ3), three-stage semi-structured interviews were conducted with nine focal students, following the suggestions of Seidman(2006, p. 16–19), to explore their prior experiences, key challenges in using AI-powered corpora for learning argumentative writing, and their reflections and suggestions. The interviews, conducted *via* Tencent Meeting in Chinese, were audio-recorded and transcribed verbatim. Thematic analysis was then applied to the reflection notes and interview transcripts, following the six-step procedure outlined by Braun and Clarke (2006). Initially, the second author coded the data by reviewing the transcripts, identifying meaningful segments, and generating preliminary codes. The third author independently reviewed and verified the coded data to ensure consistency with the extracts. Any discrepancies in coding were discussed and resolved. After finalizing the coding framework, related codes

were grouped and refined through iterative discussions to identify themes. Each theme was then clearly defined and assigned a concise label.

4. Findings

4.1. Development of students' argumentative writing over time

Table 3 presents the descriptive statistics of students' overall performance across different writing tasks, showing a significant improvement in their average scores from 19.23 in the pre-test to 23.15 in the post-test. A repeated measures ANOVA confirmed a significant main effect of time, $F(4, 92) = 9.302$, $p < .001$, $\eta^2_p = 0.194$, indicating that students' overall performance in argumentative writing improved significantly over time, accounting for 19.4% of the variance in scores. This represents a medium effect size, suggesting that students' writing scores differed significantly across time points. Pairwise comparisons with Bonferroni correction were conducted to examine specific differences in writing performance between time points. The results showed significant improvements from Task 1 (pre-test) to Task 4 ($d^z = 0.982$, $t = 4.812$, $p < .001$) and Task 5 (post-test) ($d^z = 1.075$, $t = 5.266$, $p < .001$), both with large effect sizes indicating substantial gains over time. Moreover, the post-test showed significant improvement compared to Task 2 ($d^z = 0.584$, $t = 2.863$, $p = .009$) and Task 3 ($d^z = 0.887$, $t = 4.343$, $p < .001$), with medium to large effect sizes. Overall, the results indicate a significant improvement in students' argumentative writing following the CBLP-supported argumentative writing course.

4.2. Impact on the subdimensions of argumentative writing

In addressing the second research question, learners' writing performance was evaluated based on the six subdimensions of argumentative writing. Table 4 presents the descriptive statistics of students' performance across these subdimensions over the learning period. The data showed a slight upward trend in the organizational and structural aspects of argumentative writing, namely organization, task response, evidence, and rebuttal, with some fluctuations observed across different writing tasks.

Table 3. Descriptive statistics of the overall performance.

Task	Mean	SD	Min	Max
1	19.23	3.44	13.5	26
2	20.71	3.54	14.5	27.5
3	20.75	2.56	15.5	25.5
4	22.58	2.85	15	26
5	23.15	2.08	18.5	26

Table 4. Descriptive statistics of the subdimensions.

	Task response		Evidence		Rebuttal		Organization		Language Use		Grammar	
Task	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
1	3.52	0.54	3.4	0.64	2.56	0.74	3.25	0.71	3.17	0.7	3.33	0.79
2	3.56	0.54	3.42	0.84	3.19	0.66	3.4	0.71	3.48	0.68	3.67	0.56
3	3.75	0.64	3.42	0.56	2.94	0.5	3.5	0.49	3.48	0.38	3.63	0.56
4	4.15	0.58	3.35	0.63	3.52	0.45	3.46	0.75	4	0.33	4.13	0.47
5	3.75	0.51	3.69	0.57	3.17	0.65	3.6	0.69	4.19	0.6	4.31	0.73

In contrast, students' performance in the language aspects, including language use and grammar, demonstrated a more consistent improvement throughout the learning period.

Friedman tests were conducted to determine whether students' performance on different dimensions of argumentative writing improved significantly over the learning period. For task response, results showed a significant effect of time on students' task response, $\chi^2(4) = 18.34$, $p = .001$, with a weak to moderate effect size (Kendall's $W = 0.19$). Post-hoc pairwise comparisons using Dunn's test with Bonferroni correction revealed marginal improvement in Task 4 compared with the pre-test ($p = .008$) and Task 2 ($p = .014$), but neither remained statistically significant after applying the Bonferroni-adjusted $\alpha = 0.005$. Regarding evidence use in argumentative writing, the results of the Friedman test indicated no significant differences, $\chi^2(4) = 7.799$, $p = .099 > .05$, suggesting that students' performance did not change significantly during the learning period. For rebuttal, the results of the Friedman test indicated a statistically significant effect of time on rebuttal responses, with a moderate effect size (Kendall's $W = 0.28$). Post-hoc pairwise comparisons using Dunn's test with Bonferroni correction revealed significant improvement in Task 4 compared with the pre-test ($p < .001$), small yet nonsignificant improvement in Task 4 compared with Task 3 ($p = .013$), andin the post-test compared with the pre-test ($p = .039$). For organization, time did not have any significant effect on students' performance, $\chi^2(4) = 2.602$, $p = .626$.

Regarding language-related dimensions, the results of the Friedman test revealed a more pronounced improvement in students' performance. For language use, the results indicated a statistically significant effect of time, $\chi^2(4) = 40.34$, $p < .001$, with a strong effect size (Kendall's $W = 0.42$). Post-hoc pairwise comparisons revealed significant improvement in the post-test compared with the pre-test ($p < .001$), Task 2 ($p < .001$), and Task 3 ($p < .005$). Finally, for grammar, the results indicated a statistically significant effect of time, $\chi^2(4) = 31.36$, $p < .001$, with a moderate to strong effect size (Kendall's $W = 0.33$). Dunn's post-hoc tests revealed significant improvement in the post-test compared with the pre-test ($p < .001$), Task 2 ($p < .005$), and Task 3 ($p < .005$).

Overall, students demonstrated improvements across most subdimensions of argumentative writing, with language aspects showing more noticeable improvement than content aspects.

4.3. Students' perceptions

Based on the thematic analysis of the interview transcripts, three major themes emerged: benefits, challenges, and suggestions for their argumentative writing learning experience. Within each theme, some sub-themes were identified (see Table 5).

The incorporation of AI-powered corpus tools into argumentative writing instruction elicited mixed reactions from students, highlighting both benefits and challenges. All students acknowledged the approach as beneficial, particularly for enhancing their understanding of argument structure and its logical flow. Wendy noted, 'I now understand how to make my arguments clearer and more persuasive, how to support them with factual evidence, and how to cite authoritative sources.' Similarly, Jerald

Table 5. Students' perceptions.

Themes	Subthemes	Excerpt
Benefits	(1) Providing access to authentic, reliable, and practical learning materials	In general, the keyword function and N-gram provided many authentic examples. (Dannielle)
	(2) Facilitating in-depth understanding of argument structure	The platform helped me with structure, such as proposing claims and improving logic (Jerald)
	(3) Facilitating the learning of language patterns across various structural elements	The platform provides me with clear suggestions after I input relevant keywords and the structural element I need, which helps refine my language effectively. (Wendy)
	(4) Providing opportunities for self-directed learning of argumentative writing	Compared with traditional learning methods, I think learning through the AI-powered corpus platform is better because it allows us to search for the structure we want, find the content we need, and look up information on our own. (Charlie)
	(5) Facilitating idea-generation	The points provided by AI can spark my own ideas, which may help me write more. (Vicky)
Challenges	(1) Having limited English proficiency	For beginners, the platform's exclusive use of English without translation might be a challenge. (Cyber)
	(2) Needing more time and practice for fully understanding and integrating	At the beginning, I have tried to explore the platform with the platform instruction, but the results are not very noticeable. With the teacher's support, I was gradually able to take more factors into account from the second to the fifth writing task. (Vicky)
	(3) Sometimes struggling to find useful resources in the corpus	Some of the retrieved information is not useful. (Jerald)
Suggestions	(1) Translanguaging in instruction	It would be helpful to use a mix of Chinese and English in teaching will using the platform. (Cici)
	(2) Incorporating a translation feature for corpus search results	The platform may include some translations and provide explanations for certain words in the search results. (Charlie)
	(3) Expanding data sources	It might be even better with more data. (Cici)

emphasized that the course helped him improve the structure of his essays, particularly in formulating claims and enhancing logical flow. Second, all interviewees mentioned that the platform enables students to 'notice,' 'imitate,' and eventually acquire advanced language patterns commonly used across various structural elements. The N-gram function allows students to analyze high-frequency n-grams across different structural elements of argumentative writing. Cici mentioned that, after analyzing these n-grams, she deliberately adopted the patterns in her own texts, for example, replacing first-person stance markers (e.g. 'I think') with impersonal claim frames (e.g. 'it is widely argued that …') when proposing claims. Third, students appreciated having access to authentic, reliable, and practical learning materials. They valued how the platform integrated theory with practice (as mentioned by Wendy), offering a more structured and evidence-based approach to writing (as noted by Young). Charlie added that while purely AI-generated responses might sometimes produce errors or misleading content, the corpus component ensures data reliability, making it a more trustworthy resource for their writing. Despite concerns about the potential unreliability of AI-generated information, as mentioned by three students, they still found the AI function of the platform useful for generating ideas and inspiration in their writing.

Furthermore, students highlighted long-term benefits during the interviews, including the promotion of self-directed learning. Unlike traditional methods, as Charlie

noted, the AI-powered corpus enables them to search for relevant structures and content, making the writing process more flexible and personalized. In addition, this learning approach enhanced students' interest in writing. Cici, Cyber, and Dannielle remarked that it made writing more engaging and enjoyable compared to conventional teaching approaches.

Despite its advantages, some students encountered difficulties, particularly due to language barriers. Limited English proficiency made it challenging for them to fully understand English-medium instructions. In particular, beginners often struggled to comprehend advanced language patterns and specialized vocabulary in the corpus. To address this challenge, students suggested incorporating translanguaging into instruction, allowing a mix of English and Chinese. Some also recommended adding a built-in translation feature to the corpus search results. Cyber further suggested providing clearer instructions for navigating and using the platform effectively.

Another challenge was the limited time available to fully practice argumentative writing skills using the platform. Vicky reported initial difficulties in applying instructional methods but noted improvement with practice. Charlie recommended incorporating additional exercises in full essay analysis after learning individual argument components, as this would help students develop a deeper and more holistic understanding of argumentative writing structures. Moreover, although the AI-powered corpus was generally easy to use, students sometimes found the search results irrelevant or lacking in examples, highlighting the need for improved search algorithms and an expanded data repository.

5. Discussion and implications

The study adopted a mixed-methods approach to examine the impact of an AI-powered corpus platform on learners' argumentative writing performance. The findings indicated that the platform had a positive overall influence on students' writing. This aligns with Cheung and Crosthwaite (2025) argument that integrating GenAI with a corpus can enhance students' writing efficiency by providing organized, relevant, and easily accessible writing support. Notably, students demonstrated significant improvement in their writing performance in the post-test (completed without any reference tools) compared to Task 2 and Task 3, where they were allowed to use the AI-powered corpus. This suggests that the argumentative writing knowledge and skills acquired during the CBLP, together with their engagement with the platform, had been internalized and transferred to subsequent writing tasks.

Students demonstrated noticeable improvements, particularly in language use and grammar. Traditional corpus-based teaching research has also emphasized learning lexico-grammatical patterns (Cotos et al., 2017). However, its effectiveness is influenced by learners' ability to analyze and apply corpus data (Boulton, 2009). Integrating GenAI into the corpus platform helps reduce this barrier, thereby enhancing the efficiency of learning language use (Cheung & Crosthwaite, 2025). Furthermore, the AI-powered corpus contributed to students' awareness of genre-specific language features. As noted by all focal participants, the platform's N-gram function enables students to search for, notice, imitate, and acquire advanced language patterns across structural elements (see Figure 2). By engaging with corpus data, students developed

a clearer understanding of how specific linguistic choices serve rhetorical purposes (Lu et al., 2020). This finding aligns with Mizumoto et al.'s (2017) research on the pedagogical value of linking linguistic patterns to rhetorical functions in genre-based writing instruction.

In addition to language use, students reported a deeper understanding of the structural aspects of argumentative writing, consistent with prior research demonstrating that corpus tools can enhance genre awareness and structural comprehension (Cotos et al., 2017). Interview data also indicated that corpus-based learning increases interest in writing and fosters self-directed learning, which aligns with previous findings (Chambers, 2005; O'Sullivan, 2007). This suggests that corpus-based activities enable students to take greater control of their learning process by engaging with the corpus, analyzing patterns, and making independent writing choices (Yoon & Jo, 2014).

Despite the overall improvement, some subdimensions of students' argumentative writing did not show significant improvement. For example, students' evidence use did not show significant improvement over the course of the experiment. One possible reason is that, although novice students may understand the basic rhetorical features of a genre, they often struggle to identify higher-order rhetorical issues without sufficient practice or guidance (Cotos et al., 2017). This aligns with previous research suggesting that the development of evidence use is a gradual and challenging process (Jin et al., 2019; Shi et al., 2022). To enhance the effectiveness and sophistication of evidence use, the Top-K function of the platform could be adapted in the future to provide more guided interaction with students, which can scaffold their analysis, evaluation, integration, and application of evidence step-by-step (Guo et al., 2025). Another possible reason may be students' limited English proficiency in analyzing and applying corpus content. To optimize the impact of corpus-based instruction, teachers could incorporate translanguaging strategies to help students draw on their full linguistic repertoire.

This study also found that the students showed no significant improvement in organization skills. To enhance students' organization skills, an integrated approach combining top-down and bottom-up methods is recommended, as noted in previous corpus-based writing research (Charles, 2007; Cotos et al., 2017; Lu et al., 2020). For example, bottom-up practice enables students to discover authentic language patterns across different argument structures through corpus consultation. Students can also work with complete annotated argumentative texts to understand how different structures (e.g. claim, warrant, grounding, backing, and rebuttal) are organized and compare them across different essays (Lu et al., 2020), as suggested by Charlie.

Interviews revealed that the majority of student feedback on the corpus's functionality was positive, with many praising its dual role as both a reference tool and a partner for self-directed learning. Students also suggested that AI-powered corpus platforms offer advantages over traditional corpus tools, particularly in terms of idea generation and user accessibility. In our study, students reported that the Top-K function helped them generate ideas and explore different perspectives on their writing topics. Unlike conventional corpus tools, which rely solely on existing corpus content, the AI-powered corpus tool provides an interactive writing environment in

which students not only receive information but also engage in critical dialogues. This aligns with the view of writing as a 'social act' (Hyland, 2005, p. 87). In addition, while the integration of AI functions is expected to lower the barriers to corpus consultation (Cheung & Crosthwaite, 2025), students still reported challenges in analyzing the purely English data. Moreover, the limited size of the corpus was identified as another constraint. To address these issues, the platform could integrate a built-in translation function to facilitate students' analysis of search results and expand the corpus size to better support learning.

6. Conclusion

Overall, the findings suggest that AI-assisted corpus tools have the potential to support the development of argumentative writing, particularly by improving language use and increasing awareness of rhetorical functions. The integration of AI-powered corpus tools into writing instruction not only sparked student interest but also fostered self-directed learning. Furthermore, the AI-powered corpus platform proved to be more flexible and accessible than traditional corpus tools, making it a versatile resource for writing instruction. These results underscore the value of exploring new approaches and incorporating AI- and corpus-based tools into writing pedagogy.

However, this study has several limitations. First, the small sample size and quasi-experimental design may limit the internal validity and generalizability of the findings. More rigorous experimental research with a larger sample size could be conducted in the future. Second, given the relatively short duration of this study, more extensive longitudinal research is necessary to assess the long-term effects of incorporating AI-powered corpus tools into writing instruction. Third, while the study primarily focused on students' performance and perceptions, future research could investigate how students interact with the AI-generated and corpus-derived content of the platform through methods such as eye-tracking and keylogging. Examining the relationship between engagement patterns and learning outcomes could provide further pedagogical insights into how the platform influences students' learning.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by The Hong Kong Polytechnic University, P0051009, TDG22-25/VTL-6.

Notes on contributors

Lihang Zhu is a lecturer at the College of Foreign Languages and Cultures, Xiamen University, and a doctoral student in the Department of Language Science and Technology at The Hong Kong Polytechnic University. Her research focuses on Second Language Acquisition (SLA), and she has extensive experience in writing instruction and language teaching, with a strong practical background in writing pedagogy. Email: lihang.zhu@connect.polyu.hk

Yanfang Su is currently an Assistant Professor in the Department of Translation, Lingnan University. Her research interests include corpus linguistics, corpus-based language studies, and computer-assisted language learning. She has previously published in linguistic journals such as *Language Learning & Technology* and *System* and contributed book chapters on corpus-based language teaching and learning. Email: yanfangsu@ln.edu.hk

Simin Xu is currently a PhD student in the Department of Language Science and Technology at The Hong Kong Polytechnic University. Her research interests include language teaching and learning, corpus linguistics, corpus-based translation studies, and learning technology. Email: siminxu.xu@connect.polyu.hk

Kanglong Liu is an Associate Professor in the Department of Language Science and Technology, The Hong Kong Polytechnic University. His research interests encompass corpus linguistics, language teaching and learning, learning technology and translation studies. He has published in various scholarly journals, including *Lingua*, *Plos One*, *Literary and Linguistic Computing*, *Corpora*, and *International Journal of Corpus Linguistics*. He is the author of monograph '*Corpus-Assisted Translation Teaching: Issues and Challenges*' (Springer 2021) and the editor of four volumes: '*Dream of the Red Chamber Literary and Translation Perspectives*' (Routledge 2022), '*Translation and Interpreting in the Age of Covid-19*' (Springer 2023), '*East Asian Perspectives on Corpora in Interpreting Studies*' (Routledge 2023), and '*Translation Studies in the Age of Artificial Intelligence*' (Routledge 2025). Currently, he serves as the Associate Editor of *Translation Quarterly*, the official publication of the Hong Kong Translation Society. Email: kl.liu@polyu.edu.hk.

ORCID

Lihang Zhu http://orcid.org/0009-0000-9820-6047
Simin Xu http://orcid.org/0009-0000-9463-5557
Kanglong Liu http://orcid.org/0000-0003-3962-2563

References

Abdelrahim, A. A. M., & Abdelrahim, M. A. M. (2020). Teaching and assessing metadiscoursal features in argumentative writing: A professional development training for EFL teachers. *International Journal of Applied Linguistics*, *30*(1), 70–91. https://doi.org/10.1111/ijal.12264

Andrews, R. (2009). *Argumentation in Higher Education*. Routledge.

Bernardini, S. (2000). Systematising serendipity: Proposals for concordancing large corpora with language learners. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective*. (pp. 225–235) Peter Lang.

Bloch, J. (2009). The design of an online concordancing program for teaching about reporting verbs. *Language Learning & Technology*, *13*(1), 59–78. https://doi.org/10.64152/10125/44168

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, *21*(1), 37–54. https://doi.org/10.1017/S0958344009000068

Boulton, A. (2011). Data-driven learning: The perpetual enigma. In S. Goźdź-Roszkowski (Ed.), *Explorations across languages and corpora* (pp. 563–580). Peter Lang.

Boulton, A., & Vyatkina, N. (2021). Thirty years of data-driven learning: Taking stock and charting new directions over time. *Language Learning & Technology*, *25*(3), 66–89. https://doi.org/10.64152/10125/73450

Braun, V., & Clarke, V. (2006). *Thematic analysis: A practical guide*. SAGE.

Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology*, *9*(2), 111–125. https://doi.org/10.64152/10125/44022

Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes*, *6*(4), 289–302. https://doi.org/10.1016/j.jeap.2007.09.009

Charles, M. (2011). The Routledge handbook of corpus linguistics. *System*, *39*(1), 113–115. https://doi.org/10.1016/j.system.2011.01.004

Chen, X., & Lee, I. (2022). Conflicts in peer interaction of collaborative writing – A case study in an EFL context. *Journal of Second Language Writing*, *58*, 100910. https://doi.org/10.1016/j.jslw.2022.100910

Cheung, L., & Crosthwaite, P. (2025). CorpusChat: Integrating corpus linguistics and generative AI for academic writing development. *Computer Assisted Language Learning*, 1–27. https://doi.org/10.1080/09588221.2025.2506480

Chin, C., & Osborne, J. (2010). Students' questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. *Journal of Research in Science Teaching*, *47*(7), 883–908. https://doi.org/10.1002/tea.20385

Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, *32*(7), 1–17. https://doi.org/10.1080/10494820.2023.2172044

Cotos, E., Link, S., & Huffman, S. (2017). Effects of DDL technology on genre learning. *Language Learning & Technology*, *21*(3), 104–130. https://doi.org/10.64152/10125/44623

Crosthwaite, P., & Baisa, V. (2024). A user-friendly corpus tool for disciplinary data-driven learning: Introducing CorpusMate. *International Journal of Corpus Linguistics*, *29*(4), 595–610. https://doi.org/10.1075/ijcl.23056.cro

Crosthwaite, P., & Boulton, A. (2023). DDL is dead? Long live DDL! Expanding the boundaries of data-driven learning. In H. Tyne, M. Bilger, L. Buscail, M. Leray, N. Curry & C. Pérez-Sabater (Eds.), *Discovering language: Learning and affordance*. Peter Lang. https://hal.science/hal-03506624

Crosthwaite, P., & Steeples, B. (2022). Data-driven learning with younger learners: Exploring corpus-assisted development of the passive voice for science writing with female secondary school students. *Computer Assisted Language Learning*, *37*(5-6), 1166–1197. https://doi.org/10.1080/09588221.2022.2068615

Dang, H., Mecke, L., Lehmann, F., Goller, S., & Buschek, D. (2022). How to prompt? Opportunities and challenges of zero-and few-shot learning for human-AI interaction in creative applications of generative models. arXiv preprint arXiv:2209.01390. https://doi.org/10.48550/arXiv.2209.01390

Daskalovska, N. (2013). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, *28*(2), 130–144. https://doi.org/10.1080/09588221.2013.803982

Darmawansah, D., Rachman, D., Febiyani, F., & Hwang, G.-J. (2025). ChatGPT-supported collaborative argumentation: Integrating collaboration script and argument mapping to enhance EFL students' argumentation skills. *Education and Information Technologies*, *30*, 3803-3827. https://doi.org/10.1007/s10639-024-12986-4

Eriksson, A. (2012). Pedagogical perspectives on bundles: Teaching bundles to doctoral students of biochemistry. In Thomas, J. & Boulton, A. (Eds.), *Input, process and product: Developments in teaching and language corpora* (pp. 195-211). Masaryk University Press.

Flowerdew, L. (2010). Using corpora for writing instruction. In *The Routledge handbook of corpus linguistics*. (pp. 444–457). Routledge.

Flowerdew, L. (2015). Corpus-based research and pedagogy in EAP: From lexis to genre. *Language Teaching*, *48*(1), 99–116. https://doi.org/10.1017/S0261444813000037

Guo, K., Li, D., Wang, J., & Chu, S. K. W. (2025). From retrieval to generative models: A design-based research approach to developing a chatbot for argumentative writing. *Innovation in Language Learning and Teaching*, 1–15. https://doi.org/10.1080/17501229.2025.2511273

Guo, K., Li, Y., Li, Y., & Kai, S. (2023). Understanding EFL students' chatbot-assisted argumentative writing: An activity theory perspective. *Education and Information Technologies*, *29*, 1-20. https://doi.org/10.1007/s10639-023-12230-5

Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing*, *54*, 100666. https://doi.org/10.1016/j.asw.2022.100666
Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. (1st ed.) Routledge.
Hartwell, K., & Aull, L. (2022). Constructs of argumentative writing in assessment tools. *Assessing Writing*, *54*, 100675. https://doi.org/10.1016/j.asw.2022.100675
Hirose, K. (2003). Comparing L1 and L2 organizational patterns in the argumentative writing of Japanese EFL students. *Journal of Second Language Writing*, *12*(2), 181–209. https://doi.org/10.1016/s1060-3743(03)00015-8
Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Bloomsbury Publishing.
Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching*, *41*(4), 543–562. https://doi.org/10.1017/S0261444808005235
Halliday, M. A. K., & Hasan, R. (1976). *Coherence in English*. Routledge.
Hu, X., & Liu, Y. (2025). Exploring argument structure development in EFL learners' argumentative writing. *The Asia-Pacific Education Researcher*, *34*(4), 1311–1320. https://doi.org/10.1007/s40299-024-00944-0
Jin, T., Shi, Z., & Lu, X. (2019). From novice storytellers to persuasive arguers: Learner use of evidence in oral argumentation. *TESOL Quarterly*, *53*(4), 1151–1161. https://doi.org/10.1002/tesq.541
Jurafsky, D., & Martin, J. H. (2025). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition with language models*. (3rd ed.). https://web.stanford.edu/~jurafsky/slp3/
Karpenko-Seccombe, T. (2020). *Academic Writing with Corpora: A Resource Book for Data-Driven Learning*. (1st ed.) Routledge. https://doi.org/10.4324/9780429059926
Karpenko-Seccombe, T. (2024). Practical use of corpora in teaching argumentation. In P. Crosthwaite (Ed.), *Corpora for Language Learning: Bridging the research-practice divide* (pp. 111–130). Routledge. https://doi.org/10.4324/9781003413301-9
Karbach, J. (1987). Using Toulmin's model of argumentation. *Journal of Teaching Writing*, *6*(1), 81–92.
Kneupper, C. W. (1978). Teaching argument: An introduction to the Toulmin model. *College Composition and Communication*, *29*(3), 237–241. https://doi.org/10.58680/ccc197816301
Lam, Y. W., Hew, K. F., & Chiu, K. F. (2017). Improving argumentative writing: Effects of a blended learning approach and gamification. *Language Learning & Technology*, *22*(1), 97–118. https://doi.org/10.64152/10125/44583
Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, *25*(1), 56–75. https://doi.org/10.1016/j.esp.2005.02.010
Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2021, April 12). Retrieval-augmented generation for knowledge-intensive NLP tasks. ArXiv.org. https://doi.org/10.48550/arXiv.2005.11401
Li, J., Zong, H., Wu, E., Wu, R., Peng, Z., Zhao, J., Yang, L., Xie, H., & Shen, B. (2024). Exploring the potential of artificial intelligence to enhance the writing of English academic papers by non-native English-speaking medical students - the educational application of ChatGPT. *BMC Medical Education*, *24*, 1–8. https://doi.org/10.1186/s12909-024-05738-y
Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, *93*(1), 61–78. https://doi.org/10.1111/j.1540-4781.2009.00828.x
Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, *33*, 623–635. https://doi.org/10.1016/j.system.2005.02.002
Lu, X., Casal, J. E., & Liu, Y. (2020). The rhetorical functions of syntactically complex sentences in social science research article introductions. *Journal of English for Academic Purposes*, *44*, 100832. https://doi.org/10.1016/j.jeap.2019.10083210.1016/j.jeap.2019.100832

Ma, Q., Mei, F., & Qian, B. (2024a). Exploring EFL students' pronunciation learning supported by corpus-based language pedagogy. *Computer Assisted Language Learning*, 1–27. https://doi.org/10.1080/09588221.2024.2432965

Ma, Q., Yuan, R., Cheung, L. M. E., & Yang, J. (2024c). Teacher paths for developing corpus-based language pedagogy: A case study. *Computer Assisted Language Learning*, *37*(3), 461–492. https://doi.org/10.1080/09588221.2022.2040537

Ma, Q., Tang, J., & Lin, S. (2021). The development of corpus-based language pedagogy for TESOL teachers: A two-step training approach facilitated by online collaboration. *Computer Assisted Language Learning*, *35*(4), 1–30. https://doi.org/10.1080/09588221.2021.1895225

Ma, Q., Tang, J., & Lin, S. (2022). The development of corpus-based language pedagogy for TESOL teachers: A two-step training approach facilitated by online collaboration. *Computer Assisted Language Learning*, *35*(9), 2731–2760. https://doi.org/10.1080/09588221.2021.1895225

McNeill, K. L., & Krajcik, J. (2011). *Supporting grade 5–8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing*. Pearson Education.

McNeill, K. L., & Martin, D. M. (2011). Claims, evidence, and reasoning. *Science and Children*, *48*(8), 52–56.

Mizumoto, A., Hamatani, S., & Imao, Y. (2017). Applying the bundle–move connection approach to the development of an online writing support tool for research articles. *Language Learning*, *67*(4), 885–921. https://doi.org/10.1111/lang.12250

O'Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, *19*(3), 269–286. https://doi.org/10.1017/S095834400700033X

Preiss, D. D., Castillo, J. C., Grigorenko, E. L., & Manzi, J. (2013). Argumentative writing and academic achievement: A longitudinal study. *Learning and Individual Differences*, *28*, 204–211. https://doi.org/10.1016/j.lindif.2012.12.013

Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning*, *35*(1-2), 36–61. https://doi.org/10.1080/09588221.2019.1667832

Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing. *System*, *38*(3), 444–456. https://doi.org/10.1016/j.system.2010.06.012

Rakshit, G., Bowden, K. K., Reed, L., Misra, A., & Walker, M. (2018). Debbie, the debate bot of the future. In M. Eskenazi, L. Devillers, & J. Mariani (Eds.), *Advanced social interaction with agents: 8th international workshop on spoken dialog systems* (pp. 45–52) Springer International Publishing.

Seidman, I. (2006). *Interviewing as qualitative research.* (3rd ed.) Teachers College Press.

Seo, J.-Y. (2024). Exploring the Educational Potential of ChatGPT: AI-Assisted Narrative Writing for EFL College Students. *Language Teaching Research Quarterly*, *43*, 1–21. https://doi.org/10.32038/ltrq.2024.43.01

Shi, Z., Liu, F., Lai, C., & Jin, T. (2022). Enhancing the use of evidence in argumentative writing through collaborative processing of content-based automated writing evaluation feedback. *Language Learning & Technology*, *26*(2), 106–128. https://doi.org/10.64152/10125/73481

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, *57*, 100752–100752. https://doi.org/10.1016/j.asw.2023.100752

Su, Y., Liu, K., Lai, C., & Jin, T. (2021). The progression of collaborative argumentation among English learners: A qualitative study. *System*, *98*, 102471. https://doi.org/10.1016/j.system.2021.102471

Tekin, B., & Soruç, A. (2016). Using corpus-assisted learning activities to assist vocabulary development in English. *The Turkish Online Journal of Educational Technology*,1270-1282.

Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.

Uccelli, P., Dobbs, C. L., & Scott, J. (2012). Mastering academic language. *Written Communication*, *30*(1), 36–62. https://doi.org/10.1177/0741088312469013

Vannestål, M. E., & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL*, *19*(3), 329–350. https://doi.org/10.1017/S0958344007000638

Wu, J., Zhao, C. G., Lu, X., & Jin, T. (2024). A rhetorical function and phraseological analysis of commentaries on visuals. *English for Specific Purposes*, *73*, 33–45. https://doi.org/10.1016/j.esp.2023.09.001

Wu, Y.-J., & Yeh, H.-C. (2025). (Ariel) Using Corpora to enhance learners' synonymous verb collocations. *Educational Technology & Society*, *28*(1), 144–160. https://www.jstor.org/stable/48810712

Yang, R., & Pan, H. (2023). Whole-to-part argumentation instruction: An action research study aimed at improving Chinese college students' English argumentative writing based on the Toulmin model. *Sage Open*, *13*(4), 1–15. https://doi.org/10.1177/21582440231207738

Yeadon, W., Inyang, O. O., Mizouri, A., Peach, A., & Testrow, C. P. (2023). The death of the short-form physics essay in the coming AI revolution. *Physics Education*, *58*(3), 035027. https://doi.org/10.1088/1361-6552/acc5cf

Yoon, H., & Jo, J. W. (2014). Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning & Technology*, *18*(1), 96–117. https://doi.org/10.64152/10125/44356

Zhou, C., & Hou, F. (2024). Can AI empower L2 education? Exploring its influence on the behavioural, cognitive and emotional engagement of EFL teachers and language learners. *European Journal of Education*, *59*(4), 1-13. https://doi.org/10.1111/ejed.12750

Zhu, W. (2001). Performing argumentative writing in English: Difficulties, processes, and strategies. *TESL Canada Journal*, *19*(1), 34-50. https://doi.org/10.18806/tesl.v19i1.918

Appendix 1. Assessment rubrics for argumentative writing

Dimension	Indicator
Task Response	**5 points**: Fully addresses the task, presents a clear position, and provides thorough and relevant support.
	4 points: Adequately addresses the task, presents a clear position, and provides sufficient support.
	3 points: Generally addresses the task, presents a somewhat clear position, and provides basic support.
	2 points: Partially addresses the task, presents an unclear position, and provides limited support.
	1 point: Fails to address the task, presents a vague position, and provides minimal or no support.
Organization	**5 points**: Clear and logical structure, with well-organized claims, evidence, and transitions.
	4 points: Mostly clear and logical structure, with organized claims, evidence, and transitions.
	3 points: Basic structure, with somewhat organized claims, evidence, and transitions.
	2 points: Weak structure, with disorganized claims, evidence, and transitions.
	1 point: Poor structure, with unclear or missing claims, evidence, and transitions.
Evidence supports	**5 points**: Provides strong, specific, and relevant evidence to support claims effectively.
	4 points: Provides sufficient, specific, and relevant evidence to support claims.
	3 points: Provides basic evidence that is somewhat specific and relevant to support claims.
	2 points: Provides limited evidence that is vague or less relevant to support claims.
	1 point: Provides little or no evidence, or evidence is irrelevant to the claims.
Rebuttal and response	**5 points**: Effectively identifies and rebuts counterarguments with strong and logical responses.
	4 points: Adequately identifies and rebuts counterarguments with reasonable responses.
	3 points: Identifies and rebuts counterarguments with basic responses.
	2 points: Attempts to identify and rebut counterarguments but responses are weak or unclear.
	1 point: Fails to identify or rebut counterarguments, or responses are missing.
Language use	**5 points**: Uses accurate, varied, and fluent language with a wide range of vocabulary and sentence structures.
	4 points: Uses mostly accurate, varied, and fluent language with a good range of vocabulary and sentence structures.
	3 points: Uses generally accurate language with a limited range of vocabulary and sentence structures.
	2 points: Uses language with frequent inaccuracies, limited vocabulary, and repetitive sentence structures.
	1 point: Uses language with significant inaccuracies, very limited vocabulary, and poor sentence structures.
Grammar and spelling	**5 points**: Demonstrates excellent grammar and spelling with minimal or no errors.
	4 points: Demonstrates good grammar and spelling with occasional minor errors.
	3 points: Demonstrates basic grammar and spelling with some noticeable errors.
	2 points: Demonstrates weak grammar and spelling with frequent errors.
	1 point: Demonstrates poor grammar and spelling with numerous errors.

Appendix 2: Writing topics

1. (Pre-test) Using a computer every day can have more negative than positive effects on young children. Do you agree or disagree.
2. Some people think that the climate change could have a negative effect on business. Other people think that climate change could create business more opportunities. Discuss both views and give your own opinion.
3. Some people argue that college students should be permitted to use mobile phones in class, while others believe it should be prohibited. Discuss both views and give your own opinion on the matter.
4. Some people believe that social media is becoming more and more influential in people's mental health and that it is a negative development.
   To what extent do you agree or disagree with this statement?
5. (Post-test) To what extent do you agree or disagree that artificial intelligence will eventually replace human jobs and possibly even human beings? Discuss the implications of AI advancement on the workforce and society.

Appendix 3: Interview outline

Module 1: Learning background

1. **Experience with Argumentative Writing**

   - Could you share your experiences learning English writing?
   - Have you ever learned how to write an English argumentative essay before?
   - Which part of the writing process do you find most challenging?

2. **Experience with Corpus Tools**

   - Do you know what a corpus is?
   - If you've used a corpus before, how did you use it to improve your English writing?

3. **Experience with AI Tools**

   - Have you used AI tools during the writing process? What AI tools do you mainly use?
   - How have you used AI to assist with your English writing? What do you think are its strengths and weaknesses?

Module 2: Critical incidents

1. **Recalling Initial Writing**

   - Do you remember the process of completing your first writing task without using the platform? How did you prepare? What difficulties did you encounter, and what kind of support did you wish to have?

2. **Impact of Training Sessions**

   - How did the training sessions help you? Did the sessions reflect in your writing (Tasks 2, 3, 4, and 5)?
   - During the training, what factors influenced your behavior, cognition, and emotions when working on Tasks 2, 3, 4, and 5? (e.g. in relation to writing performance or score fluctuations.)

3. **Perceptions of the AI-Assisted Writing Platform**

   - Do you think the AI-assisted writing platform helped you?
   - Which features of the platform did you use? Which feature did you use the most, and which did you find the most useful?
   - What are your thoughts on the results you found using the platform? Were you able to understand and apply the data provided by the platform?
   - Did you use the AI-generated examples to assist your writing? How and Why did you use them? How did you feel about the process?

Module 3: Overall impressions

- How was your overall experience with this writing course?
- What is your overall impression of the AI-powered corpus platform?
- Do you have any suggestions for improving the writing course or AI-powered corpus platform?