International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 08 Issue 09 September 2025
DOI: 10.47191/ijcsrr/V8-i9-29, Impact Factor: 8.048
IJCSRR @ 2025
www.ijcsrr.org

AI-Driven Writing Instruction and College EFL Learners' Writing Proficiency: A Complex Dynamic Systems Perspective

Jian Wang[1*], Xiaohong Zhang[2], Xinli Ke[3]
[1]School of Foreign Languages and Cultures, Geely University of China
[2]School of Foreign Languages and Cultures, Panzhihua University
[3]Department of Foreign Languages, Southwest Jiaotong University Hope College

ABSTRACT: Generative artificial intelligence (GenAI) technology is transforming language education, particularly in the realm of writing instruction. Since the introduction of OpenAI's ChatGPT, extensive research has demonstrated the efficacy of incorporating GenAI into writing instruction to boost learners' writing proficiency. These studies have typically employed a "ladder" perspective to assess learners' writing development through comparing the outcomes of pre- and post-intervention tests. However, this approach often overlooks the dynamic progression of writing competence. To address this gap, the present study investigated the developmental trajectory of tertiary-level English learners' writing proficiency through the framework of Complex Dynamic Systems Theory (CDST). Over a 13-week AI-driven writing program, thirteen students participated and underwent seven writing assessments. The comparison between learners' initial and final writing proficiency states revealed significant improvements in overall writing proficiency, as well as in the dimensions of writing complexity and accuracy, though not in fluency. Nonetheless, analysis of the writing outcomes indicated fluctuations in overall writing proficiency, complexity, fluency, and accuracy across the seven tests. Additionally, learners displayed individual variability in their developmental trajectories across all aspects of writing. The study also identified trade-offs among writing complexity, fluency, and accuracy throughout the instructional process. These findings provide empirical support for CDST within an AI-driven teaching context and offer valuable insights for enhancing writing instruction.

KEYWORDS: AI-driven instruction, Complex Dynamic Systems Theory, Developmental trajectory, English writing, writing proficiency.

INTRODUCTION

The swift progress in deep learning, natural language processing, and the advent of large language models (LLMs) like the GPT series have endowed generative artificial intelligence (GenAI) with remarkable proficiency in both linguistic creation and comprehension. This technology mimics human innovation to craft a spectrum of superior and varied textual outputs (Hagos et al., 2024). The unveiling of OpenAI's ChatGPT in 2022, a chatbot predicated on LLM, has brought to light the transformative potential of AI across societal domains, with a pronounced effect on educational practices (Giannakos et al., 2024). The essence of GenAI is encapsulated in its ability to generate language, comprehend context, adapt to individual needs, and offer immediate responses—qualities that lay the groundwork for its integration into educational settings (Bahroun et al., 2023).

Within the sphere of language education, GenAI has been steadily gaining prominence as a crucial instrument, providing both students and educators with unprecedented and innovative experiences in pedagogical and didactic engagements (Xiao et al., 2025). Hitherto, scholars and practitioners in the domain of second language (L2) acquisition have evinced substantial enthusiasm for AI-augmented language instruction, with a particular focus on the pedagogy of L2 writing (e.g., Guo et al., 2024; Wang et al., 2024). An extensive corpus of scholarly inquiry has scrutinized the effectiveness of AI-augmented writing instruction in bolstering L2 learners' writing proficiency. Despite this, the conclusions drawn from these studies exhibit a significant degree of divergence. A preponderance of empirical studies posit that AI-assisted writing instruction is instrumental in enhancing learners' holistic writing proficiency (e.g., Aladini et al., 2025; Gayed et al., 2022) as well as their analytical writing skills in lexical richness, syntactic complexity, content development, and structural coherence (Boudouaia et al., 2024; Tsai et al., 2024). Conversely, a minority of studies has discerned no salient divergence in the efficacy of AI-assisted versus traditional instruction in elevating learners' writing proficiency (e.g., Escalante et al., 2023; Ironsi & Solomon Ironsi, 2025). Research has also shown that AI-generated writing

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 08 Issue 09 September 2025
DOI: 10.47191/ijcsrr/V8-i9-29, Impact Factor: 8.048
IJCSRR @ 2025
www.ijcsrr.org

feedback can significantly expand learners' lexical repertoire and improve grammatical accuracy in their written work. However, the impact of AI on aspects of content selection and discourse (e.g., coherence and cohesion) appears to be more limited (Luo et al., 2025). Moreover, extant studies have predominantly utilized quasi-experimental frameworks to probe the overarching impact of integrating AI within writing instruction on learners' writing proficiency, with scant attention accorded to the dynamic developmental implications of AI-driven writing instruction on learners' writing competence. In light of this, the present study adopted a longitudinal research approach and, through the lens of the Complex Dynamic Systems Theory (CDST), examined the impact of AI-driven English-as-a-foreign-language (EFL) writing instruction on learners' writing proficiency. It also evaluated learners' developmental trajectories in overall performance and analytical skills related to writing complexity, fluency, and accuracy.

LITERATURE REVIEW

A. Feasibility of Integrating GenAI into EFL Writing Instruction

The advent of GenAI has heralded a transformative era in the realm of foreign language education (Thorne, 2024), particularly in the domain of writing instruction (Kim et al., 2025). This technology is poised to revolutionize the way teachers approach L2 writing instruction, with its applications spanning across the entire writing process: from pre-writing to post-writing stages (Fields, 2024; Yuan et al., 2024). In the pre-writing phase, GenAI can be harnessed to produce comprehensible input materials that cater to various aspects of writing tasks, including content generation, textual organization, lexical choices, and syntactic patterns (Nguyen et al., 2024; Trang & Barrot, 2024). These elaborately crafted materials not only facilitate learners' comprehension of the writing task but also empower them to allocate cognitive and linguistic resources more judiciously during the formal writing process, thus helping learners generate higher-quality texts (Abulaiti, 2024; Naznin et al., 2025). GenAI can also generate tailored writing recommendations for learners, foster their creative ideation, broaden their conceptual frameworks, and assist them in mastering diverse writing styles and techniques (ElSayary, 2024; Kim et al., 2025). During the actual writing and subsequent revision stages, GenAI's potential is equally remarkable. It can provide immediate feedback and error correction to students, enabling swift identification and rectification of grammatical, spelling, and structural errors (Cai et al., 2025). This potentially enhances the precision of linguistic expression. GenAI tools like ChatGPT can significantly bolster students' capabilities in text summarization, content revision, language editing, and proofreading (Guo & Li, 2024; Liu et al., 2024), thereby augmenting both the efficiency and quality of writing. Another upside of GenAI lies in its capacity to deliver personalized instruction that is finely tuned to the varying language proficiency levels of learners, effectively addressing a broad spectrum of educational requirements (Xiao et al., 2025; Zhao, 2024). Besides, by streamlining tasks like grading essays, designing teaching materials and generating teaching cases, GenAI can substantially ease teachers' workload, allowing them to concentrate more on deploying sophisticated pedagogical strategies (Giannakos et al., 2024).

B. Empirical Research Related to Integrating GenAI into EFL Writing Instruction

As AI technology advances, an increasing number of researchers are conducting empirical studies on integrating GenAI into L2 writing instruction. Few scholars have focused on GenAI's performance in automated essay scoring. Shin and Lee (2024) instructed ChatGPT to evaluate student essays based on four dimensions: task completion, content, organization, and language use. Their comparison of ChatGPT's performance with that of human raters revealed a significant alignment between the two. This finding highlights the potential of GenAI as an automated essay scoring tool, which offers considerable advantages in terms of operational efficiency, scalability, and consistency in evaluation. In addition to providing automated scores based on established criteria, GenAI tools can offer learners prompt and effective feedback. They excel in identifying language errors, suggesting corrections, and providing language refinement strategies, thereby increasing learners' writing quality (Shin & Lee, 2024).

The bulk of research has employed a quasi-experimental design to determine the empowering effects of GenAI on L2 learners' writing proficiency by comparing changes before and after the intervention. The assessment of students' writing proficiency is generally conducted through writing tests and quantitative textual indices. Quantitative textual indices are predominantly measured using Skehan's (1996) CAF framework (i.e., complexity, accuracy, and fluency). For instance, Gayed et al. (2022) explored the use of AI KAKU, which is based on the GPT-2 language model, in a writing intervention and found that the experimental group exhibited significantly higher syntactic complexity than the control group. Boudouaia et al. (2024) conducted a 10-week quasi-experimental study to evaluate the efficacy of writing feedback provided by ChatGPT in enhancing English writing proficiency. Their results indicated that the experimental group achieved significantly better outcomes in writing scores and four specific dimensions—task achievement, discourse coherence, grammar, and vocabulary—compared to the control group. These findings substantiate the effectiveness of ChatGPT-facilitated writing feedback in improving learners' writing performance. Similarly, Tsai et al. (2024) examined the efficacy of ChatGPT-assisted writing revision, confirming that the treatment significantly improved students' writing quality, with notable enhancements in vocabulary, grammar, organization, and content.

Some researchers have also zoomed in on learners' affective and attitudinal responses to the incorporation of GenAI technology into L2 writing instruction. For example, Boudouaia et al. (2024) reported a generally favorable disposition among learners, highlighting their behavioral intentions to continue using GenAI technology. Teng (2024) found that GenAI-supported writing significantly enhanced English learners' motivation, self-efficacy, and engagement in writing tasks. However, the reception of GenAI technology is not uniformly positive. Some studies have identified negative sentiments among learners, such as concerns about ChatGPT's limitations in recognizing deep structures and pragmatic errors (Al-Garaady & Mahyoob, 2023), as well as the increased cognitive load it places on learners (Woo et al., 2024).

Additionally, a growing number of scholars have begun to focus on "human-AI interaction" and "human-AI collaboration" in learners' writing process. For instance, Bai and Wei (2024) utilized a three-stage writing task—composing, comparison, and rewriting—to explore how EFL writers perceived and assimilated ChatGPT's reformulations as corrective feedback. Their analysis included examining the instances where these reformulations were integrated into the learners' subsequent drafts. The findings revealed that the reformulations drew learners' attention to discrepancies in their initial drafts, particularly in terms of lexical selection, thereby encouraging them to incorporate changes suggested by ChatGPT into their revised work. Lee (2024) delved into the dynamics of human-AI interaction, uncovering that learners engaged in a cyclical process of planning, translating, and reviewing when working collaboratively with ChatGPT on writing tasks. This study underscored the pivotal role of scaffolding provided by teachers and peers in leveraging GenAI technologies effectively, which could enhance learners' emotional reassurance and metacognitive awareness. Hwang et al. (2024) scrutinized learners' behaviors during the application of ChatGPT-assisted writing revisions. Their research indicated that learners predominantly concentrated on surface-level features of writing (e.g., grammar and punctuation), while often overlooking more sophisticated elements like content, organization, and coherence. The study emphasized the importance of integrating prompt-based writing instruction into English writing pedagogy, suggesting that effective learner-AI interaction could significantly impact the enhancement of writing quality.

C. Research Gap

To date, research on incorporating GenAI tools in L2 writing instruction has yielded findings with significant implications for writing pedagogy. Despite this, substantial research gaps persist that require further examination. Notably, existing studies have predominantly focused on the role of ChatGPT, often overlooking the potential contributions of other AI tools. As AI technology continues to advance, a diverse array of tools or LLMs has emerged. In China, for example, tools such as DeepSeek, Kimi, Doubao, and ERNIE Bot are gaining traction alongside ChatGPT. The pedagogical effectiveness of incorporating these tools into L2 writing instruction remains an open question. Furthermore, many studies have employed quasi-experimental designs to assess the impact of AI integration on learners' writing performance. These studies typically involve pre- and post-tests, analyzing students' scores and changes across various writing dimensions (e.g., the CAF indices) before and after the intervention. This approach, however, tends to oversimplify language development as a linear progression. In reality, language proficiency is multifaceted, and its development is not a straightforward, stepwise improvement; it may involve periods of regression, stagnation, and fluctuation (Larsen-Freeman, 1997). Consequently, the dynamic process through which learners utilize AI tools to enhance their writing has been largely neglected. To address these issues, this study adopted a longitudinal research design and integrated the GenAI tool Kimi into EFL writing instruction from a complex dynamic systems perspective, aiming to elucidate the intricate developmental trajectory of learners' writing proficiency.

THEORETICAL FRAMEWORK: COMPLEX DYNAMIC SYSTEMS THEORY

CDST serves as an overarching conceptual framework that encompasses a multitude of interrelated theories, including "chaos/complexity theory," "dynamic systems theory," and "complexity theory" (Khomeijani Farahani et al., 2020). This theoretical paradigm is widely applied across a spectrum of academic disciplines (Hiver & Al-Hoorie, 2016). With its roots in physics and mathematics, it fundamentally investigates the dynamic behaviors inherent to complex systems and the intricate interactions among their constituent subsystems (Guastello et al., 2008). CDST posits that systems are comprised of numerous interconnected variables or parameters that are constantly changing (Aspås & Ugedal, 2023). In contrast to conventional models of information processing that emphasize straightforward input-output causality, CDST places a greater emphasis on the interplay among a system's components and its dynamic relationship with the external environment (Warren et al., 1998).

Larsen-Freeman (1997) pioneered the application of CDST to applied linguistics, proposing that language acquisition extends beyond cognitive-psychological or sociocultural phenomena to encompass a dynamic process involving the continuous interaction of diverse resources. These resources, spanning cognitive, psychological, and sociocultural domains, interact across multiple levels and dimensions, contributing to the complexity of language development. From the CDST perspective, the conventional notion of "L2 acquisition" is more aptly termed "L2 development." This shift in terminology from "acquisition" to "development" underscores the fluid and fluctuating nature of language proficiency, which encompasses both progression and regression. Within this framework, language acquisition and attrition are seen as intertwined and equally significant facets of the developmental trajectory. Larsen-Freeman (1997) argues that "language development" captures the non-linear nature of language learning, thereby encapsulating phenomena such as "language degradation" and "language loss." This perspective more accurately mirrors learners' active construction of their linguistic systems and the diverse ways in which language is utilized. The L2 system is composed of various subsystems, including phonetics, vocabulary, and grammar, each of which contains smaller, interrelated components (de Bot et al., 2007). These subsystems interact in complex, dynamic, and interdependent ways throughout the stages of L2 development. According to CDST, the language learning process is characterized by dynamism, complexity, non-linearity, chaos, unpredictability, sensitivity to initial conditions, openness, self-organization, sensitivity, and adaptability (Larsen-Freeman, 2016).

RESEARCH QUESTIONS

This study aimed to investigate the effects of AI-driven writing instruction on the writing proficiency of Chinese tertiary-level EFL learners, as well as the developmental trajectories of their writing skills. To achieve this, four research questions were formulated:

(1) Does incorporating the GenAI tool Kimi into EFL writing instruction enhance learners' writing proficiency?
(2) What is the dynamic trajectory of EFL learners' average writing proficiency in an AI-augmented teaching context?
(3) Are there individual differences in the developmental pathways of writing proficiency?
(4) How do learners' writing complexity, fluency, and accuracy dynamically interact with each other?

METHODS

A. Participants and Procedure

This study involved 16 third-year students from a Chinese college, all majoring in non-English disciplines, with ages ranging from 20 to 22 years (SD=0.45). In June 2024, these participants had partaken in the College English Test-Band 4 (CET-4), a standardized English proficiency test for Chinese university students. They achieved an average score of 449.31 out of 710 (SD=21.56) and self-assessed their English writing proficiency as low, with an average writing score of 56.31 out of 106.5 (SD=7.33). None of them were enrolled in English courses or participating in English training programs, either in-person or online, during the study period. They expressed a strong desire for targeted English writing instruction and were highly motivated to enhance their writing skills. Although all students reported using GenAI, it was solely for accessing materials unrelated to English learning, and none had utilized AI for English writing purposes. Due to absences, three students missed one or two tests, resulting in incomplete data, which was subsequently excluded from analysis.

This longitudinal study spanned a 13-week timeframe from late September to mid-December 2024. Throughout this period, the second researcher convened with participants on 13 separate occasions within the university's computer lab. In Week 1, a pretest was conducted to gauge the students' baseline writing proficiency. Participants were informed of the study's objectives and provided with a written consent form. They were assured of anonymity, confidentiality, and the right to participate voluntarily or withdraw at any time. In Weeks 2, 4, 6, 8, 10, and 12, students were allotted 40 minutes to compose an essay devoid of any online or offline resources for ideation or language refinement. They subsequently submitted their essays through Kimi's interface and directed Kimi—a GenAI tool developed by the China-based technology company Moonshot AI—to provide constructive

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 08 Issue 09 September 2025
DOI: 10.47191/ijcsrr/V8-i9-29, Impact Factor: 8.048
IJCSRR @ 2025
www.ijcsrr.org

feedback, pinpointing strengths and areas for improvement across content, language, mechanics (e.g., punctuation and capitalization), and organization. The researcher was consistently accessible during this phase to clarify any queries or concerns raised by the participants. Each session, including essay writing and interaction with Kimi, extended for approximately 90 minutes. During the remaining six weeks, a weekly writing test was administered, with participants completing the tests in the lab under conditions that precluded the use of additional writing aids or online/offline resources. The researcher presided over the entire process, with each test allocated 40 minutes.

B. Instrument: English Writing Tests

The study comprised six in-class writing exercises and seven writing tests, all derived from authentic IELTS exam materials and focusing on the argumentative essay genre. The writing tests featured seven distinct topics, a strategic choice made for two primary reasons: firstly, to mitigate student fatigue, and secondly, to emulate natural learning conditions, thereby assessing learners' adaptability to varying tasks—an approach consistent with CDST. For comparative purposes, the same two topics (A and B) were employed in both the initial and final tests. In the initial test, half of the students were assigned Topic A, while the other half tackled Topic B; in the final test, the groups exchanged topics, ensuring the difficulty level remained consistent across both tests.

Two experienced English writing teachers, previously involved in grading essays for large-scale standardized English tests, were recruited to evaluate the participants' essays. To maintain anonymity and objectivity, students' names were removed prior to submission to the graders. Both graders independently rated the English essays from seven distinct tests using Jacobs et al.'s (1981) rating rubrics. In instances of substantial scoring discrepancies, the graders engaged in constructive discussions to reach a consensus. Following the grading process, the researchers utilized SPSS 27.0 to calculate the inter-rater reliability. The Pearson correlation coefficients for the scores assigned by the two graders ranged from 0.792 to 0.867, demonstrating a high level of inter-rater reliability.

C. Measures of Writing Complexity, Fluency, and Accuracy

In this study, words per minute (WPM) was employed as a metric to assess writing fluency (Tabari et al., 2024). Complexity, a multifaceted construct, encompasses both lexical and syntactic dimensions. Following Plakans et al.'s (2019) research, syntactic complexity was gauged using the mean length of T-units (MLT). The underlying rationale for this metric is the assumption that longer T-units signify more complex syntactic structures. To facilitate this evaluation, the L2 Syntactic Complexity Analyzer (Lu, 2010) was used. For lexical complexity, the study utilized the U-index, a variant of the type/token ratio (TTR; Jarvis, 2002). Unlike the TTR, which can be distorted by text length—where longer texts are more likely to include repeated words—the U-index effectively mitigates this limitation. The L2 Lexical Complexity Analyzer (Lu, 2012) was employed to compute this measure. To prevent the software from classifying misspelled words as instances of complex or novel vocabulary (Wang & Wang, 2020), researchers reviewed and corrected all misspellings before conducting the lexical complexity analysis. Accuracy was determined using the formula proposed by Bai and Ye (2018), in which a flawless essay received a score of 100 points, with one point deducted for each error per 100 words. Thus, the accuracy score was calculated as 100 minus the number of errors per 100 words.

D. Data Analysis

Statistical analysis was conducted using SPSS version 27.0. We began by employing descriptive statistics to explore the means and standard deviations of both the total scores and CAF scores of essays produced in each writing test. Given the small sample size, the normal distribution of all data was assessed using Shapiro-Wilk tests, as understanding data distribution is essential for determining whether to use parametric or non-parametric analyses for within-group comparisons. The threshold for statistical significance was established at a $p < .05$ level. To visually depict the dynamic changes of participants' writing skills, line graphs were generated using Excel.

FINDINGS

A. Results of Descriptive Analyses and Paired-Samples T-Test

We conducted descriptive analyses to explore the general trends and variations in essay scores, along with the four CAF indices, for all participants across the seven tests. As shown in Table 1, a comparison of the students' initial (Time 1) and final (Time 7) states indicates an overall upward trajectory for all variables. This finding suggests a noticeable enhancement in students' overall writing proficiency, as well as in their lexical and syntactic complexity, accuracy, and fluency, after 13 weeks of instruction. However, it is crucial to acknowledge that the values of these variables fluctuated throughout the period from Time 1 to Time 7, rather than following a strictly linear upward path.

Table 1. Descriptive statistics

Variable	T1		T2		T3		T4		T5		T6		T7	
	Mean	***SD***	**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***
Score	69.42	2.54	73.15	2.366	76.35	2.641	75.38	3.083	78.27	3.113	78.54	3.326	81.77	2.969
U-index	17.51	2.35	17.23	1.722	18.86	2.058	20.28	4.166	17.27	3.431	18.34	3.177	22.92	2.375
MTL	11.12	1.61	13.46	2.553	13.53	2.527	12.75	2.981	13.97	2.202	17.05	3.626	16.17	1.514
Accuracy	80.34	4.81	81.52	4.72	82.78	4.10	83.33	4.90	85.41	4.95	86.57	4.36	88.31	4.35
Fluency	7.70	0.889	8.35	0.637	8.28	0.81	8.11	0.773	8.1	0.485	8.38	0.686	8.09	0.509

SD, standard deviation.

Subsequently, we examined whether there were statistically significant differences in students' writing scores and CAF indices between the Time 1 and Time 7 tests. The Shapiro-Wilk test results indicated that the W values for all groups ranged from 0.924 to 0.984, with all p-values exceeding 0.05, suggesting that our data sets were normally distributed. Consequently, we employed a paired-samples t test to assess the differences between the outcomes of the two tests. As presented in Table 2, the mean difference in students' writing scores was 12.350, with the mean score for the Time 7 test being significantly higher than that for the Time 1 test ($t=-9.425$, $p<.01$, Cohen's $d=0.459$). Additionally, the essays from the Time 7 test demonstrated significant improvements in CAF quantitative textual indices, specifically in lexical complexity ($t=-1.426$, $p<.05$, Cohen's $d=0.218$), syntactic complexity ($t=-2.862$, $p<.05$, Cohen's $d=0.306$), and accuracy ($t=-6.498$, $p<.01$, Cohen's $d=0.523$). However, no significant change was observed in fluency between the essays from the two tests ($t=-1.584$, $p>.05$, Cohen's $d=0.105$).

Table 2. Results of paired-samples t test

Variables	MD (T1-T7)	*SD*	*SE*	95% Confidence Interval		*t*	*df*	*p*	*Cohen's d*
				Lower Bound	Upper Bound				
Score	-12.350	4.723	1.310	-15.200	-9.492	-9.425	12	.000	0.459
Lexical Complexity	-5.410	4.419	1.226	-2.148	3.193	-1.426	12	.026	0.218
Syntactic Complexity	-5.050	2.371	0.658	-3.315	-0.449	-2.862	12	.014	0.306
Accuracy	-7.970	4.426	1.228	-10.651	-5.302	-6.498	12	.000	0.523
Fluency	-0.390	0.901	0.250	-0.940	0.149	-1.584	12	.139	0.105

MD, mean difference; SE, standard error; df, degree of freedom.

B. Dynamic Changes of the Average Writing Scores and CAF over Time

To effectively illustrate the dynamic changes of students' mean scores and the mean values of the CAF indices across seven tests, we employed Excel to create line charts. Figures 1A to 1C depict a gradual and overall upward trend in both writing scores and

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 08 Issue 09 September 2025
DOI: 10.47191/ijcsrr/V8-i9-29, Impact Factor: 8.048
IJCSRR @ 2025
www.ijcsrr.org

complexity among the 13 students, interspersed with minor fluctuations. Notably, Figure 1D demonstrates a consistent linear improvement in writing accuracy, suggesting that the integration of GenAI in writing instruction has effectively enhanced students' awareness of linguistic precision, resulting in more accurate text production. Meanwhile, Figure 1E reveals significant variations in writing fluency, with a marked improvement between the first and second tests. However, following this initial increase, fluency generally declined, except for a slight recovery noted in the sixth test.

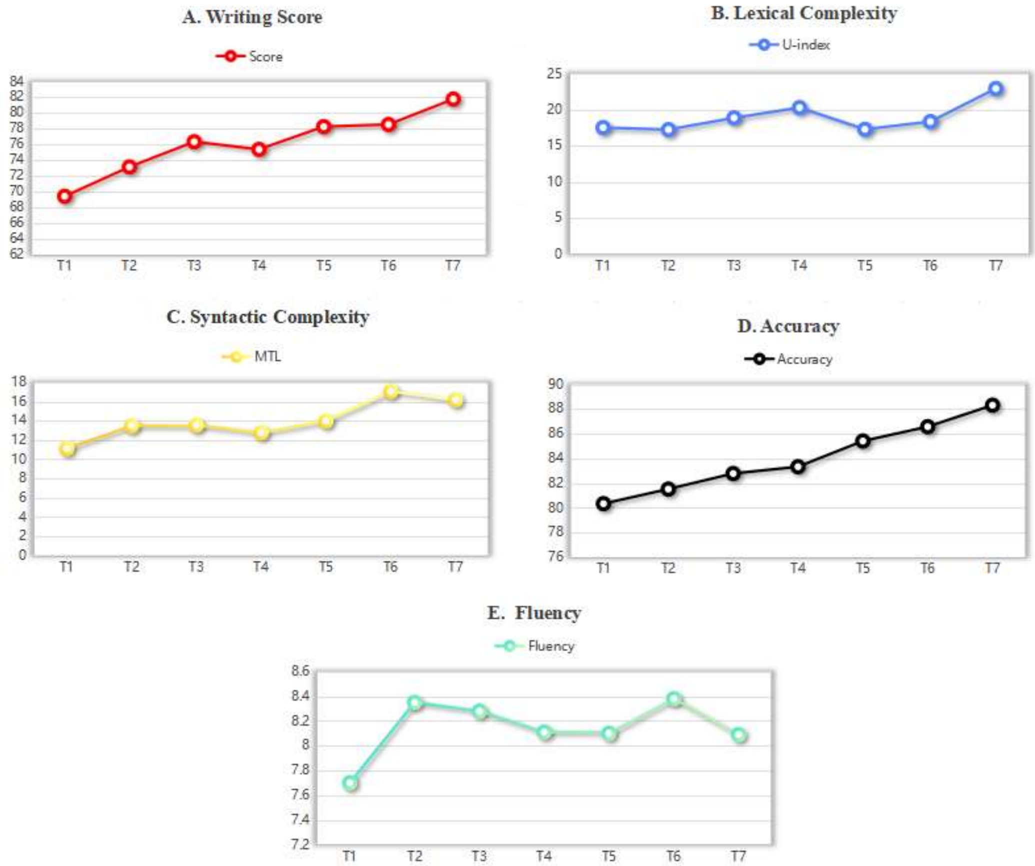

Figure 1. Changes in average writing score and CAF indices

C. Dynamic Changes of Individual Students' Writing Proficiency

Upon a comprehensive analysis of the developmental trajectories in writing performance among 13 students, a discernible pattern emerged, revealing that their writing scores predominantly followed four distinct paths. Recognizing the complexity involved in capturing the nuanced development of these students' writing abilities within a single line chart, we chose to present the representative trajectories of four students—designated as S1, S2, S3, and S4—in the ensuing line charts.

Figure 2A illustrates the distinct developmental trajectories of writing scores for the four students. S1 exhibited a rapid, step-like linear increase in writing scores, whereas S2 showed a gradual and consistent linear improvement. In contrast to S1 and S2, the writing scores of S3 and S4 demonstrated fluctuations over the course of seven tests. Initially, S3's scores rose steadily during the first three tests but then experienced a significant decline in the subsequent two tests. However, the scores rebounded sharply in the final two tests, with the last test score notably exceeding that of the first. Conversely, S4's scores increased consistently over the first four tests, peaking at the Time 4 test. Following this peak, there was a pronounced decline, with the score at the Time 7 test falling below the initial level.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 08 Issue 09 September 2025
DOI: 10.47191/ijcsrr/V8-i9-29, Impact Factor: 8.048
IJCSRR @ 2025
www.ijcsrr.org

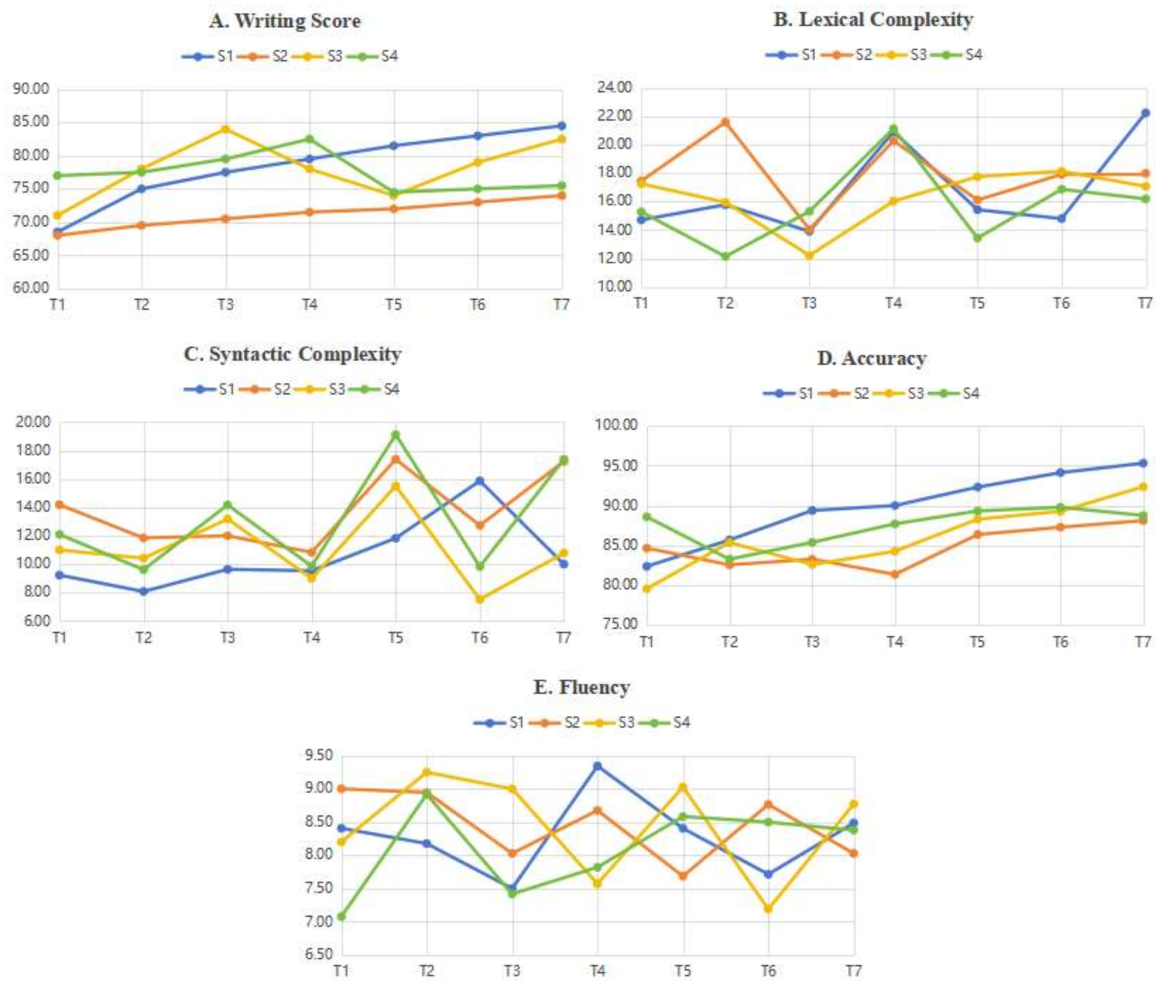

Figure 2. Changes in individuals' writing score and CAF indices

Figure 2B portrays the fluctuating development of lexical complexity among the four students over the course of seven tests. Throughout these assessments, their writing complexity exhibited significant peaks and troughs, at times falling below baseline levels. Despite these fluctuations, the students' lexical complexity generally trended back towards their initial state. Notably, S1 demonstrated a marked improvement in lexical complexity by the Time 7 test compared to the Time 1 test.

Figure 2C reveals that the syntactic complexity of these students, akin to their lexical complexity, exhibited significant fluctuations. Specifically, S2 and S4 followed comparable patterns, with their syntactic complexity ultimately surpassing the initial level after experiencing several oscillations. In contrast, S3's syntactic complexity reached two low points during the Time 4 and Time 6 tests, both of which fell below the initial level. Notably, in the Time 6 test, S3's complexity dropped significantly below the starting point before returning to baseline in the final test. Meanwhile, S1 demonstrated a consistent increase in syntactic complexity over the first six tests, but experienced a precipitous decline in the last test, concluding slightly above the initial state.

Figure 2D delineates the dynamic progression of writing accuracy among the four students. S1 manifested a linear upward trend in this dimension, while the remaining three encountered minor undulations yet generally maintained a stable upward trajectory.

Figure 2E shows the significant variation in writing fluency experienced by the four students, which contrasts with other quantitative textual indices. S3 and S4 demonstrated an undulating developmental pattern, ultimately achieving fluency levels that surpassed their initial states. Conversely, S1's writing fluency reverted to its baseline level, while S2 ended with a fluency level lower than their starting point.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 08 Issue 09 September 2025
DOI: 10.47191/ijcsrr/V8-i9-29, Impact Factor: 8.048
IJCSRR @ 2025
www.ijcsrr.org

D. Dynamic Interplay among CAF over Time

A meticulous comparative analysis is indispensable to elucidate the intricate dynamics among writing complexity, fluency, and accuracy. Given the diverse computational techniques for various indices and the heterogeneity in data, standardizing the data is an essential initial step. To address this necessity, we employed SPSS 27.0 to convert the original data into standardized Z-scores, following the approach by Larsen-Freeman (2006). This standardization process enabled a consistent comparison across all indices.

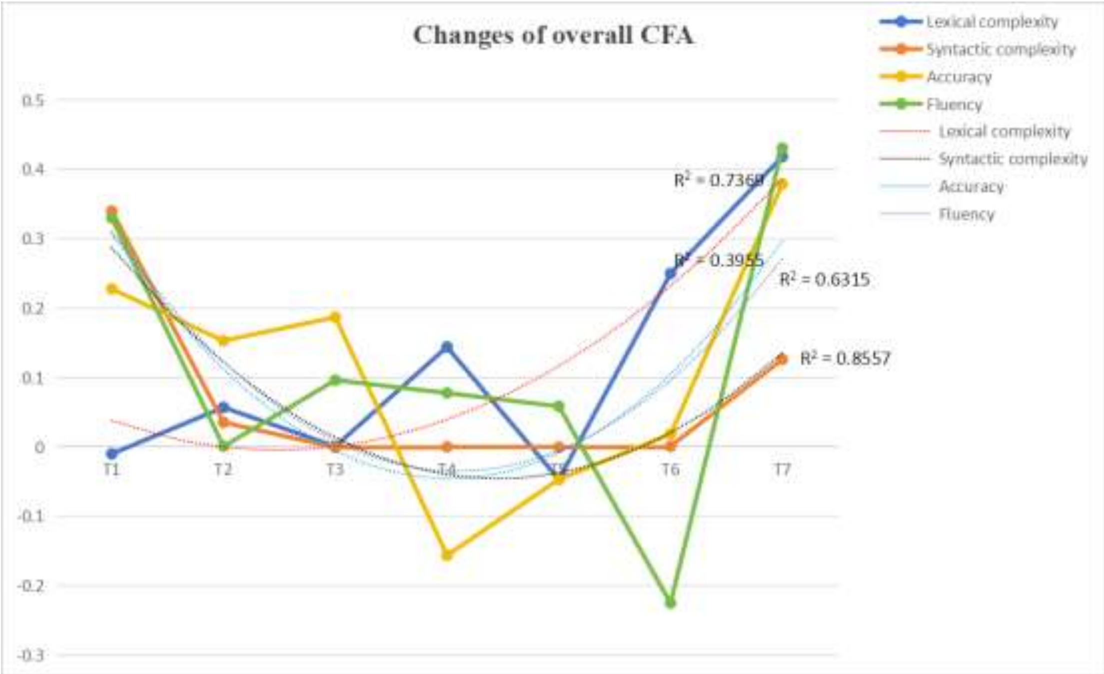

Figure 3. Interrelationships between writing complexity, fluency, and accuracy over time

Figure 3 presents the standardized line graphs and trend lines for the constructs of writing complexity, fluency, and accuracy. The graphical representation indicates that these three elements were in a state of dynamic equilibrium, exhibiting oscillatory patterns. For instance, in the Time 1 test, syntactic complexity reached its peak, followed closely by fluency, with accuracy ranking third and lexical complexity at the lowest point. In contrast, the Time 4 test showed lexical complexity rising to the highest level, with fluency in second place, syntactic complexity in third, and accuracy at the lowest. By the Time 7 test, fluency attained its highest point, with lexical complexity closely following in second place, while syntactic complexity and accuracy were in third and fourth positions, respectively. As depicted in Figure 3, the peaks and troughs of these indices never coincided, indicating that the development of writing complexity, fluency, and accuracy was in a constant cycle of fluctuation.

DISCUSSION

A. Impact of AI-Enhanced Writing Instruction on Writing Proficiency

The present study employed a longitudinal research design to investigate the impact of integrating the GenAI tool Kimi into tertiary-level EFL writing instruction. The analysis of the initial and final states of the participating students' English writing proficiency indicates a significant improvement in their average writing scores. This outcome aligns with findings from previous studies (e.g., Boudouaia et al., 2024; Tsai et al., 2024; Zhao, 2024), which also demonstrated the effectiveness of AI-integrated instruction in enhancing learners' writing quality. GenAI tools utilize advanced algorithms and language models to provide tailored writing suggestions and constructive feedback to language learners (Barrot, 2023). These tools are particularly advantageous in areas such as grammar checking, word selection, sentence refinement, organizational coherence, and content creation, which are essential components of a well-crafted essay (Pitukwong & Saraiwang, 2024). By engaging with AI tools for writing practice, students are exposed to polished essays and receive personalized guidance, helping them identify weaknesses in their writing and comprehend the insights and strategies offered by AI. Incorporating this extensive knowledge into subsequent writing tasks or assessments can potentially result in a marked enhancement of their writing abilities. Nonetheless, further rigorous research is necessary to substantiate these explanations.

Our research revealed significant improvements in participants' writing complexity and accuracy. However, no notable differences in writing fluency were observed between the initial and final tests. AI tools are particularly adept at identifying language errors (Mitra & Banerjee, 2023), and by providing corrective feedback, they enable learners to recognize and effectively rectify mistakes in their essays. Consequently, learners may become more vigilant about the errors in their essays and amend them more effectively in practical writing scenarios. The findings on fluency and complexity are partially in line with those of previous studies. For instance, Gayed et al. (2022) reported a significant enhancement in MLT—a measure of syntactic complexity—following treatment, yet no substantial changes in fluency or lexical complexity were noted. Conversely, Seo (2024) found significant improvements in both writing fluency and MLT between pretest and post-test essays. Collectively, these studies suggest a considerable increase in MLT, potentially due to AI tools' ability to assist learners in moving beyond habitual sentence structures by offering immediate grammatical feedback and structural enhancement suggestions (Linares Carrasquer, 2025). When AI tools detect a series of short sentences in students' essays, they may suggest combining clauses and provide illustrative examples. With consistent exposure to such constructive feedback, students are likely to adopt AI's writing style, thereby gradually enhancing their syntactic complexity. The inconsistent findings regarding lexical complexity and fluency may be attributed to the different indices utilized across studies. For example, Gayed et al. (2022) employed lexical diversity to assess lexical complexity, whereas Seo (2024) considered essay length and T-unit count as indicators of writing fluency. In our study, the U index was used to evaluate lexical complexity, and WPM served as a measure of writing fluency. Given these disparate research findings, further investigation is warranted.

B. Developmental Trajectories of Writing Proficiency in AI-Empowered Teaching Context

This study adopted a CDST perspective to explore the developmental trajectories of learners' writing proficiency. We found that the average writing scores of the 13 participants showed a general upward trend. This finding echoes those of existing research works (e.g., Boudouaia et al., 2024; Jamshed et al., 2024; Tsai et al., 2024; Zhao, 2024). While most existing studies have utilized a quasi-experimental design, comparing results from a single pretest and post-test, they adopt a simplistic "developmental ladder" perspective (Fischer et al., 2002), suggesting a linear stepwise progression in language acquisition. However, language learning is a complex process, akin to what Fischer et al. (2002) describes as a constructed "web," intricately interwoven in multiple directions. In our study, although learners' mean writing scores showed a general upward trend influenced by the writing assistance and support from Kimi, there were fluctuations across the seven tests. Beyond accuracy, we observed notable variations in the average fluency and complexity indices among all students, indicating a nonlinear trajectory in the development of these skills. Specifically, students' average writing fluency exhibited alternating phases of progression and regression throughout the 13-week instructional period. Conversely, students' average writing accuracy demonstrated a linear development, suggesting that the integration of AI into EFL writing instruction can consistently enhance learners' ability to convey their ideas accurately. The diverse developmental pathways of writing scores and CAF indices provide empirical support for CDST, which posits that the journey of language acquisition is inherently complex and fraught with twists and turns (Larsen-Freeman, 2006).

Our research also highlights the diversity in developmental trajectories among participants, revealing patterns of stability, singular significant shifts, or multiple fluctuations. In terms of developmental trends, most participants exhibited upward growth, while others experienced stagnation or regression. Among those with upward trajectories, some demonstrated steady progress, others reached a sharp peak before experiencing declines with oscillations, and still others ascended in a spiral pattern. These findings corroborate Larsen-Freeman's (2006) dynamic CAF model, which describes oral and written proficiency development among Chinese EFL learners in traditional educational settings. The outcomes of language development result from a complex interplay of factors, which, although minor in isolation, interact synergistically rather than merely cumulatively, leading to significant variations in learning behaviors and outcomes (Bai & Ye, 2018; Marchman & Thal, 2005). For instance, participants in our study likely varied in initial writing proficiency, vocabulary and syntax mastery, perceptions of AI technology, familiarity with AI-assisted writing, and cognitive processing approaches, all of which might have substantially influenced their writing performance and development across the CAF dimensions. However, further research is needed to explicate the effects of these factors.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 08 Issue 09 September 2025
DOI: 10.47191/ijcsrr/V8-i9-29, Impact Factor: 8.048
IJCSRR @ 2025
www.ijcsrr.org

C. Interplay among Writing Complexity, Fluency, and Accuracy in the AI-Driven Teaching Context

The findings indicate that these three elements existed in a dynamic equilibrium, characterized by a competitive relationship. Across the seven tests, an increase in one aspect often corresponded with a decrease in the others, rather than any single aspect maintaining a consistently high or low level. This competition is likely related to the limitations of human cognitive capacity, as clarified by the Cognitive Load Theory, which emphasizes the finite capacity of working memory during learning (Sweller, 1988). When learning tasks exceed this capacity, trade-offs become inevitable. The learning system, with its complexity and sensitivity due to interconnected components (Larsen-Freeman, 1997), forms a complex system where fluency, accuracy, and complexity are interdependent and mutually restrictive. Constraints on short-term memory, attention, and time mean that focusing on one aspect naturally diverts attention from others (Cowan, 2000). For instance, students who prioritize language accuracy during a writing test may avoid experimenting with complex sentences or vocabulary, potentially compromising linguistic complexity to ensure clarity and precision. Conversely, students who emphasize language complexity might find themselves preoccupied with sophisticated vocabulary and syntax, which can slow down their writing and reduce fluency.

CONCLUSION, IMPLICATIONS AND LIMITATIONS

This study explored the development of EFL learners' writing skills within an AI-driven instructional context through the lens of a CDST perspective. The integration of AI into writing instruction was found to significantly enhance learners' overall writing proficiency, as well as improve their analytical skills related to writing complexity, fluency, and accuracy. However, the progression of learners' writing development was nonlinear, exhibiting fluctuations throughout the instructional period. Furthermore, individual variability was evident in learners' developmental trajectories across all aspects of writing. The study also identified a dynamic equilibrium among writing complexity, fluency, and accuracy, with competition and trade-offs occurring among these elements during the teaching process.

These findings carry significant pedagogical implications for EFL writing instruction. Our research highlights the substantial impact of AI technology on improving learners' writing proficiency, particularly in the areas of complexity, fluency, and accuracy. This suggests that writing instructors can effectively incorporate AI tools as a supplement to traditional teacher-led instruction. The robust capabilities of Generative AI (GenAI) in text generation and corrective feedback provision can be leveraged to support learners throughout the writing process. Specifically, AI-generated feedback, including error correction suggestions and exemplary essays, can activate students' metalinguistic knowledge and enable them to discern differences between superior and subpar essays, particularly in terms of lexical choices and syntactic structures. It is crucial for teachers to recognize that the development of writing skills, even in an AI-driven learning environment, is not linear and may exhibit fluctuations. Teachers can communicate to students that learning is a long-term process where persistence is key to achieving desired outcomes. Encouraging students to view AI tools as enduring learning "companions" can optimize their use of AI-generated resources. Additionally, due to individual differences in learners' writing development trajectories, it is advisable for teachers to tailor instructional strategies to align with each student's learning path and pace. Ultimately, teachers should acknowledge the presence of competition and trade-offs among various dimensions of writing (i.e., complexity, accuracy, and fluency). By continuously monitoring students' development across these aspects, teachers can promptly identify and address deficiencies in writing to foster comprehensive improvement in students' writing skills.

Despite its valuable implications, this study has several limitations. First, the absence of a control group limits the ability to compare the developmental trajectories of learners' writing proficiency across diverse teaching contexts (e.g., traditional versus AI-driven writing instruction). Second, the study relied on previously published research practices and selected only one quantitative textual feature to measure writing complexity, fluency, and accuracy, which might not fully represent these aspects. Third, the nonlinearity in the writing development trajectory observed in this study might be influenced by topic variation; repeating identical writing tasks could yield different results. Finally, due to space limitations, the study focused on group characteristics when analyzing the dynamic relationships among fluency, accuracy, and complexity, without addressing individual differences. Future research could refine these aspects for further exploration.

FUNDING STATEMENT

This article represents the interim research outcomes of the 2024 research project of the Sichuan Provincial Private Education Association titled "The Impact of Artificial Intelligence Technology on the Developmental Trajectories of College Students'

English Writing Proficiency from the Perspective of the Complex Dynamic Systems Theory" (**Grant No.: MBXH24YB191**) and the 2025 "Three-Campus Initiative" teaching reform program of Geely University of China titled "Study on Human-AI Collaborative English Writing Assessment Model Driven by Artificial Intelligence" (**Grant No.: 2025SGXYJG133**).

REFERENCES

1. Abulaiti, M. (2024). Higher Education in the Era of AI. In G. Morris (Ed.), *Developments and future trends in transnational higher education leadership* (pp. 244-265). IGI Global.
2. Aladini, A., Ismail, S. M., Khasawneh, M. A. S., & Shakibaei, G. (2025). Self-directed writing development across computer/AI-based tasks: Unraveling the traces on L2 writing outcomes, growth mindfulness, and grammatical knowledge. *Computers in Human Behavior Reports*, *17*, 100566. https://doi.org/10.1016/j.chbr.2024.100566
3. Al-Garaady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journals*, *9*, 3-17. https://doi.org/10.24093/awej/call9.1
4. Aspås, K. S., & Ugedal, S. (2023). *Investigating emergent dynamicity and development of complexity and accuracy in writing of Norwegian L2 learners of English: A Complex Dynamic Systems Theory perspective* (Master's thesis). Nord Universitet.
5. Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability*, *15*(17), 12983. https://doi.org/10.3390/su151712983
6. Bai, L., & Wei, Y. (2024). Exploring EFL learners' integration and perceptions of ChatGPT's text revisions: A three-stage writing task study. *IEEE Transactions on Learning Technologies*, *17*, 2161-2172. https://doi.org/10.1109/TLT.2024.3491864
7. Bai, L.F., & Ye, S.F. (2018).The dynamic development of English-L2 writing competence. *Modern Foreign Languages*, *41*(3): 354-366.
8. Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, 100745. https://doi.org/10.1016/j.asw.2023.100745
9. Boudouaia, A., Mouas, S., & Kouider, B. (2024). A study on ChatGPT-4 as an innovative approach to enhancing English as a foreign language writing learning. *Journal of Educational Computing Research*, *62*(6), 1509-1537. https://doi.org/10.1177/07356331241247465
10. Cai, B., He, Z., Fu, H., Zheng, Y., & Song, Y. (2025). Reducing English major students' writing errors with an automated writing evaluation system: Evidence from eye-tracking technology. *IEEE Transactions on Learning Technologies*, 1-18. https://doi.org/10.1109/TLT.2025.3547321
11. Cowan, N. (2000). Processing limits of selective attention and working memory: Potential implications for interpreting. *Interpreting*, *5*(2), 117-146. https://doi.org/10.1075/intp.5.2.05cow
12. de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*(1), 7–21. https://doi.org/10.1017/S1366728906002732
13. ElSayary, A. (2024). An investigation of teachers' perceptions of using ChatGPT as a supporting tool for teaching and learning in the digital era. *Journal of Computer Assisted Learning*, *40*(3), 931-945. https://doi.org/10.1111/jcal.12926
14. Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, *20*(1), 57. https://doi.org/10.1186/s41239-023-00425-2
15. Fields, Z. (2024). *Impacts of generative AI on creativity in higher education*. IGI Global.
16. Fischer, K. W., Yan, Z., & Stewart, J. (2002). Adult cognitive development: Dynamics in the developmental web. In J. Valsiner & K. J. Conolly (Eds.), *Handbook of developmental psychology* (pp. 491-516). Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781848608306.n21
17. Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing Assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, *3*, 100055. https://doi.org/10.1016/j.caeai.2022.100055

18. Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järveläh, S., Mavrikis, M., & Rienties, B. (2024). The promise and challenges of generative AI in education. *Behaviour & Information Technology*, 1-27. https://doi.org/10.1080/0144929x.2024.2394886
19. Guastello, S. J., Koopmans, M., & Pincus, D. (2008). *Chaos and complexity in psychology: The theory of nonlinear dynamical systems*. Cambridge University Press.
20. Guo, K., & Li, D. (2024). Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System*, *124*, 103362. https://doi.org/10.1016/j.system.2024.103362
21. Guo, K., Pan, M., Li, Y., & Lai, C. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The Internet and Higher Education*, *63*, 100962. https://doi.org/10.1016/j.iheduc.2024.100962
22. Hwang, M., Jeens, R., & Lee, H. K. (2024). Exploring learner prompting behavior and its effect on ChatGPT-assisted English writing revision. *The Asia-Pacific Education Researcher*, 1-11. https://doi.org/10.1007/s40299-024-00930-6
23. Jamshed, M., Manjur Ahmed, A. S. M., Sarfaraj, M., & Warda, W. U. (2024). The impact of ChatGPT on English language learners' writing skills: An assessment of AI feedback on mobile. *International Journal of Interactive Mobile Technologies*, *18*(19), 18-36. https://doi.org/10.3991/ijim.v18i19.50361
24. Hagos, D. H., Battle, R., & Rawat, D. B. (2024). Recent advances in generative AI and large language models: Current status, challenges, and perspectives. *IEEE Transactions on Artificial Intelligence*, *5*(12), 5873-5893. https://doi.org/10.1109/TAI.2024.3444742
25. Hiver, P., & Al-Al-Hoorie, A. H. (2016). A dynamic ensemble for second language research: Putting complexity theory into practice. *The Modern Language Journal*, *100*(4), 741-756. https://doi.org/10.1111/modl.12347
26. Ironsi, C. S., & Solomon Ironsi, S. (2025). Experimental evidence for the efficacy of generative AI in improving students' writing skills. *Quality Assurance in Education*, *33*(2), 237-252. https://doi.org/10.1108/QAE-04-2024-0065
27. Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.
28. Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, *19*(1), 57-84. https://doi.org/10.1191/0265532202lt220oa
29. Khomeijani Farahani, A.A., Rezaee, A.A., & Moshtaghi Zonouz, R. (2020). Exploring the development of writing complexity, accuracy, and fluency in relation to the motivational trajectories: A dynamically-oriented case study. *English Teaching & Learning*, *44*, 81-100. https://doi.org/10.1007/s42321-019-00040-3
30. Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies*, *30*(1), 1265-1300. https://doi.org/10.1007/s10639-024-12878-7
31. Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*(2), 141-165. https://doi.org/10.1093/applin/18.2.141
32. Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, *27*(4), 590-619. https://doi.org/10.1093/applin/aml029
33. Larsen-Freeman, D. (2016). Classroom-oriented research from a complex systems perspective. *Studies in Second Language Learning and Teaching*, *6*(3), 377-393. https://doi.org/10.14746/ssllt.2016.6.3.2
34. Luo, M., Hu, X., & Zhong, C. (2025). The collaboration of AI and teacher in feedback provision and its impact on EFL learner's argumentative writing. *Education and Information Technologies*, 1-21. https://doi.org/10.1007/s10639-025-13488-7
35. Lee, Y. J. (2024). Can my writing be polished further? When ChatGPT meets human touch. *ELT Journal*, *78*(4), 401-413. https://doi.org/10.1093/elt/ccae039
36. Linares Carrasquer, J. (2025). *Interactive language learning application using artificial intelligence* (Doctoral dissertation). Universitat Politècnica de València.
37. Liu, Y., Park, J., & McMinn, S. (2024). Using generative artificial intelligence/ChatGPT for academic communication: Students' perspectives. *International Journal of Applied Linguistics*, *34*(4), 1437-1461. https://doi.org/10.1111/ijal.12574

38. Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474-496. https://doi.org/10.1075/ijcl.15.4.02lu
39. Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, *96*(2), 190-208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x
40. Marchman, V., & Thal, D. (2005). Words and grammar. In M. Tomasello & D. I. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates* (pp. 141-164). Lawrence Erlbaum Associates Publishers.
41. Mitra, N., & Banerjee, A. (2023). AI in subconscious language learning for error remediation. In *International conference on cyber intelligence and information retrieval* (pp. 81-92). Springer Nature Singapore. https://doi.org/10.1007/978-981-97-3594-5_7
42. Naznin, K., Al Mahmud, A., Nguyen, M. T., & Chua, C. (2025). ChatGPT integration in higher education for personalized learning, academic writing, and coding tasks: A systematic review. *Computers*, *14*(2), 53. https://doi.org/10.3390/computers14020053
43. Nguyen, L. Q., Le, H. V., & Nguyen, P. T. (2024). A mixed-methods study on the use of ChatGPT in the pre-writing stage: EFL learners' utilization patterns, affective engagement, and writing performance. *Education and Information Technologies*, 1-24. https://doi.org/10.1007/s10639-024-13231-8
44. Pitukwong, K., & Saraiwang, S. (2024). Exploring the effectiveness of digital writing tools on Thai EFL students' writing. *Contemporary Educational Technology*, *16*(3), ep519. https://doi.org/10.30935/cedtech/14808
45. Plakans, L., Gebril, A., & Bilki, Z. (2019). Shaping a score: Complexity, accuracy, and fluency in integrated writing performances. *Language Testing*, *36*(2), 161-179. https://doi.org/10.1177/0265532216669537
46. Seo, J. Y. (2024). The intersection of AI and language assessment: A study on the reliability of ChatGPT in grading IELTS writing task 2. *Language Teaching Research Quarterly*, *43*, 1-21. https://doi.org/10.32038/ltrq.2024.43.01
47. Shin, D., & Lee, J. H. (2024). Exploratory study on the potential of ChatGPT as a rater of second language writing. *Education and Information Technologies*, 1-23. https://doi.org/10.1007/s10639-024-12817-6
48. Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, *17*(1), 38-62. https://doi.org/10.1093/applin/17.1.38
49. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257-285. https://doi.org/10.1016/0364-0213(88)90023-7
50. Tabari, M. A., Khajavy, G. H., & Goetze, J. (2024). Mapping the interactions between task sequencing, anxiety, and enjoyment in L2 writing development. *Journal of Second Language Writing*, *65*, 101116. https://doi.org/10.1016/j.jslw.2024.101116
51. Teng, M. F. (2024). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence*, *7*, 100270. https://doi.org/10.1016/j.caeai.2024.100270
52. Thorne, S. L. (2024). Generative artificial intelligence, co-evolution, and language education. *Modern Language Journal*, *108*(2), 567-572. https://doi.org/10.1111/modl.12932
53. Trang, N. H., & Barrot, J. S. (2024). Pre-writing and post-writing explicit instruction: Differential effects on L2 learners' writing accuracy. *RELC Journal*, *55*(3), 688-704. https://doi.org/10.1177/00336882231184270
54. Tsai, C. Y., Lin, Y. T., & Brown, I. K. (2024). Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges. *Education and Information Technologies*, 1-19. https://doi.org/10.1007/s10639-024-12722-y
55. Wang, C., Li, Z., & Bonk, C. (2024). Understanding self-directed learning in AI-Assisted writing: A mixed methods study of postsecondary learners. *Computers and Education: Artificial Intelligence*, *6*, 100247. https://doi.org/10.1016/j.caeai.2024.100247
56. Wang, J., & Wang, X.F. (2020). An analysis of the predictive power of quantitative lexical features on machine-graded essay scores. *Language Education*, *8*(3): 26-32, 39.
57. Warren, K., Franklin, C., & Streeter, C. L. (1998). New directions in systems theory: Chaos and complexity. *Social Work*, *43*(4), 357-372. https://doi.org/10.1093/sw/43.4.357

58. Woo, D. J., Wang, D., Guo, K., & Susanto, H. (2024). Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process. *Education and Information Technologies*, 1-28. https://doi.org/10.1007/s10639-024-12819-4
59. Xiao, F., Zhu, S., & Xin, W. (2025). Exploring the landscape of generative AI (ChatGPT)-powered writing instruction in English as a foreign language education: A scoping review. *ECNU Review of Education*, 20965311241310881. https://doi.org/10.1177/20965311241310881
60. Yuan, Y., Li, H., & Sawaengdist, A. (2024). The impact of ChatGPT on learners in English academic writing: Opportunities and challenges in education. *Language Learning in Higher Education*, *14*(1), 41-56. https://doi.org/10.1515/cercles-2023-0006
61. Zhao, D. (2024). The impact of AI-enhanced natural language processing tools on writing proficiency: An analysis of language precision, content summarization, and creative writing facilitation. *Education and Information Technologies*, 1-32. https://doi.org/10.1007/s10639-024-13145-5