

ARTIFICIAL INTELLIGENCE IN ENGLISH WRITING FEEDBACK: A SYSTEMATIC REVIEW (2020 - 2025)

Nguyen Thi Linh Nga

Email: linhnga.nt88@gmail.com

Hanoi Star Primary & Secondary School
Lot T1, Trung Hoa - Nhan Chinh Urban Area,
Yen Hoa ward, Hanoi, Vietnam

Received: 04/8/2025

Revised: 16/9/2025

Accepted: 10/10/2025

Published: 20/11/2025

Abstract: In recent years, the integration of Artificial Intelligence into education has expanded significantly, with growing applications in teaching English writing as a foreign language (EFL) and as a second language (ESL). While prior reviews have focused mainly on traditional automated writing evaluation systems, less attention has been given to the pedagogical implications of generative artificial intelligence. This study conducts a systematic review of eighteen peer-reviewed articles published from 2020 to 2025 that examines the use of automated evaluation tools and artificial intelligence in providing feedback on English writing for language learners. The selection and analysis of the studies were conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta - Analyses) framework to ensure methodological transparency and rigor. The findings demonstrate that Artificial Intelligence enhances writing instruction through immediate and individual feedback, expanded idea generation, and support for learner autonomy. However, limitations remain in addressing higher-order writing skills such as coherence, argumentation, and critical thinking. These results emphasize the necessity of complementing Artificial-Intelligence-driven tools with teacher and peer feedback. By clarifying both the strengths and constraints of artificial intelligence, this study not only extends previous reviews but also provides concrete pedagogical, research, and policy implications for effectively integrating both traditional automated writing evaluation and generative artificial intelligence into English writing pedagogy.

Keywords: *Artificial Intelligence, automated writing evaluation, generative artificial intelligence, academic English writing, systematic review.*

TRÍ TUỆ NHÂN TẠO TRONG PHẢN HỒI VIẾT TIẾNG ANH: MỘT PHÂN TÍCH TỔNG QUAN HỆ THỐNG (2020 - 2025)

Nguyễn Thị Linh Nga

Email: linhnga.nt88@gmail.com

Trường Liên cấp Trung học cơ sở,
Tiểu học Ngôi Sao Hà Nội
Lô T1, khu Đô thị Trung Hòa - Nhân Chính,
phường Yên Hòa, Hà Nội, Việt Nam

Nhận bài: 04/8/2025

Chỉnh sửa xong: 16/9/2025

Chấp nhận đăng: 10/10/2025

Xuất bản: 20/11/2025

Tóm tắt: Trong những năm gần đây, việc tích hợp trí tuệ nhân tạo vào giáo dục đã trở nên phổ biến, với ngày càng nhiều ứng dụng trong giảng dạy kỹ năng viết tiếng Anh như một ngoại ngữ (EFL) và tiếng Anh như một ngôn ngữ thứ hai (ESL). Trong khi các tổng quan trước đây chủ yếu tập trung vào những hệ thống đánh giá viết tự động truyền thống, nghiên cứu này tiến hành một tổng hợp 18 công trình đã được bình duyệt, công bố trong giai đoạn 2020 - 2025, về việc sử dụng công cụ đánh giá tự động và công cụ trí tuệ nhân tạo trong việc đánh giá phản hồi bài viết tiếng Anh cho người học. Việc lựa chọn và phân tích các nghiên cứu được lựa chọn và phân tích tuân theo khung PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) nhằm bảo đảm tính minh bạch phương pháp và độ chặt chẽ học thuật. Kết quả cho thấy, trí tuệ nhân tạo góp phần nâng cao hiệu quả dạy viết thông qua việc cung cấp phản hồi tức thì và cá nhân hóa, mở rộng gợi ý phát triển ý tưởng cũng như hỗ trợ năng lực tự chủ của người học. Tuy nhiên, vẫn còn một số hạn chế trong việc xử lý các kỹ năng bậc cao như tính mạch lạc, lập luận và tư duy phản biện. Những phát hiện này nhấn mạnh sự cần thiết của việc kết hợp công nghệ với phản hồi từ giáo viên và bạn học. Bằng cách làm rõ điểm mạnh và hạn chế của trí tuệ nhân tạo, nghiên cứu này không chỉ mở rộng các tổng quan trước đó mà còn đưa ra những hàm ý cụ thể về mặt nghiên cứu, sự phạm và chính sách nhằm tích hợp hiệu quả cả hệ thống đánh giá viết tự động truyền thống và trí tuệ nhân tạo tạo sinh vào giảng dạy viết tiếng Anh.

Từ khóa: *Trí tuệ nhân tạo, phản hồi viết tự động, trí tuệ nhân tạo tạo sinh, viết tiếng Anh học thuật, tổng quan hệ thống.*

1. Đặt vấn đề

Trong xu thế toàn cầu hóa và hội nhập quốc tế, năng lực viết học thuật bằng tiếng Anh trở thành yêu cầu then chốt đối với người học ngoại ngữ. Kỹ năng này không chỉ đòi hỏi sự nắm vững về ngữ pháp và từ vựng, mà còn bao gồm khả năng phát triển ý tưởng phức tạp dựa trên tư duy phản biện và lập luận logic (Zhang, Aubrey, Huang, & Chiu, 2025; Teng, 2024). Một trong những yếu tố quan trọng hỗ trợ quá trình này là phản hồi sửa lỗi viết (Written corrective feedback), vốn giúp người học nhận diện, hiểu và điều chỉnh những điểm hạn chế để cải thiện chất lượng bài viết (Corder, 1981; Bitchener & Ferris, 2012). Nhờ vậy, phản hồi sửa lỗi viết (WCF) từ lâu đã được xem là một quy trình sư phạm cốt lõi trong dạy viết và được nhiều giáo viên chú trọng.

Tuy nhiên, phản hồi của giáo viên thường bị giới hạn bởi áp lực khối lượng công việc, thời gian và sự thiếu nhất quán (Ferris, 2006). Trước bối cảnh đó, sự phát triển của công nghệ đã mở ra những hướng tiếp cận mới trong việc hỗ trợ phản hồi truyền thống. Các hệ thống phản hồi tự động như Criterion, Grammarly cùng với các công cụ tạo sinh văn bản như ChatGPT và Gemini, đang được xem là giải pháp tiềm năng để nâng cao chất lượng giảng dạy viết (Schunk & Zimmerman, 2007; Barrot, 2021; Ho, 2024). Trên cơ sở ấy, việc nghiên cứu hiệu quả sư phạm của công cụ trí tuệ nhân tạo trong phản hồi viết tiếng Anh trở nên cần thiết. Bài viết này tiến hành một tổng quan hệ thống gồm 18 nghiên cứu gần đây về ứng dụng trí tuệ nhân tạo trong phản hồi viết học thuật cho người học tiếng Anh nhằm trả lời hai câu hỏi: 1) Những khía cạnh nào của kỹ năng viết thường được các công cụ trí tuệ nhân tạo xử lý? 2) Hiệu quả sư phạm của việc sử dụng công cụ trí tuệ nhân tạo trong phản hồi viết tiếng Anh được thể hiện ra sao? Đồng thời, bài viết cũng chỉ ra những khoảng trống nghiên cứu và đề xuất cho những nghiên cứu tiếp theo.

2. Phương pháp nghiên cứu

Phương pháp nghiên cứu của bài viết tuân theo quy

trình tổng quan hệ thống dựa trên khung PRISMA, kết hợp với phân tích định tính, dữ liệu được truy xuất từ hai cơ sở Scopus và Web of Science. Toàn bộ bản thảo cũng được hiệu đính bằng sự hỗ trợ của công cụ trí tuệ nhân tạo ChatGPT nhằm nâng cao tính mạch lạc và học thuật. Các từ khóa được xác định bao gồm các nhóm: “Artificial intelligence”, “Automated written evaluation”, “EFL” (xem Bảng 1) được sàng lọc từ tiêu đề, tóm tắt, nội dung và bảng biểu.

Bảng 1: Hệ thống từ khóa tìm kiếm

Các từ khóa
Nhóm 1: “Artificial intelligence” OR “AI” OR “Generative AI” OR “Genai” AND
Nhóm 2: “Automated writing evaluation” OR “Writing feedback” OR “Writing assessment” OR “Writing correction” AND
Nhóm 3: “EFL” OR “ESL” OR “Second language writing” OR “Foreign language writing”

Xác định và lựa chọn các nghiên cứu phù hợp:

Người viết lựa chọn các nghiên cứu phù hợp dựa trên các tiêu chí sau: Tất cả các nghiên cứu được viết bằng tiếng Anh; Nghiên cứu về việc sử dụng trí tuệ nhân tạo trong phản hồi viết cho người học tiếng Anh; Nghiên cứu được công bố từ năm 2020 đến năm 2025. Các nghiên cứu trùng lặp giữa các hệ thống cơ sở dữ liệu điện tử, các nghiên cứu không thể tìm thấy hoặc hạn chế truy cập toàn văn đều được loại bỏ khỏi quá trình phân tích.

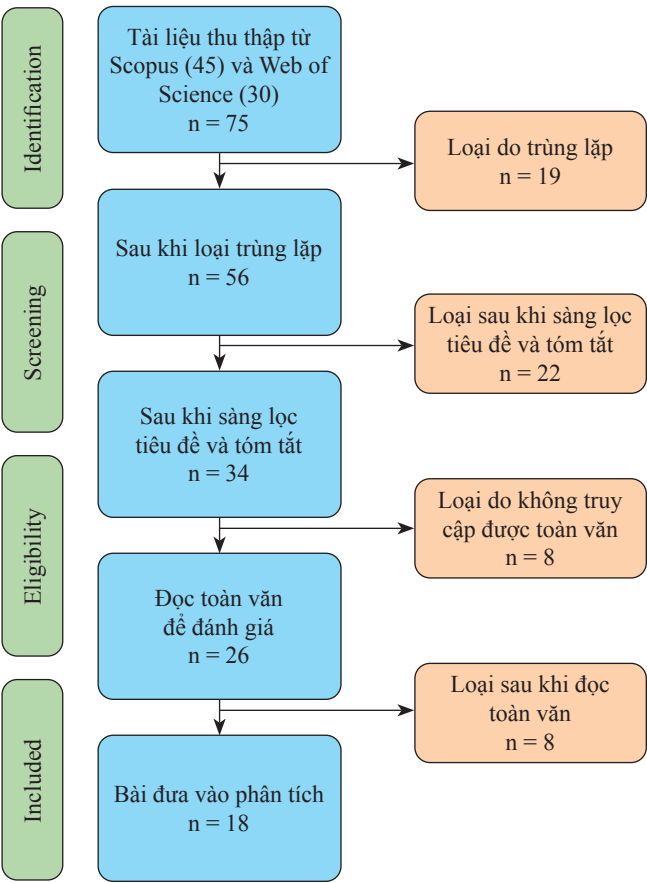
Tác giả cho rằng, các tiêu chí này có thể chịu ảnh hưởng nhất định từ yếu tố chủ quan. Tuy nhiên, phần lớn nội dung trong Bảng 2 đã được tham khảo và điều chỉnh dựa trên các tiêu chí được áp dụng trong những nghiên cứu tổng quan hệ thống trước đây (Ví dụ: Shi, 2024; Teng, 2024). Vì vậy, các tiêu chí này vẫn được xem là phù hợp cho các nghiên

Bảng 2: Tiêu chí lựa chọn các nghiên cứu phù hợp

Tiêu chí	Lựa chọn	Loại bỏ
Ngôn ngữ	Tiếng Anh	Không phải tiếng Anh.
Loại tài liệu	Bài báo học thuật đã bình duyệt.	Tài liệu khác (sách, báo cáo, luận văn, kỷ yếu chưa bình duyệt).
Nội dung	Thảo luận về việc sử dụng trí tuệ nhân tạo trong phản hồi viết tiếng Anh.	Không thảo luận về việc sử dụng trí tuệ nhân tạo trong phản hồi viết tiếng Anh.

Tiêu chí	Lựa chọn	Loại bỏ
Bối cảnh nghiên cứu	Bối cảnh học tiếng Anh như ngôn ngữ thứ hai (English as a second language) hoặc tiếng Anh như một ngoại ngữ (English as a Foreign Language).	Không phải bối cảnh học tiếng Anh như ngôn ngữ thứ hai hoặc tiếng Anh như một ngoại ngữ.
Năm công bố	2020 – 2025.	Trước năm 2020.
Loại nghiên cứu	Thực nghiệm (Empirical).	Phi thực nghiệm (Ví dụ: tổng quan, lí thuyết, khái niệm).
Khả năng truy cập	Truy cập mở (Open-access).	Không truy cập mở; cần trả phí để truy cập.

cứ tiếp theo với khả năng bổ sung hoặc điều chỉnh ở một số khía cạnh, chẳng hạn như yêu cầu về số lượng trích dẫn tối thiểu của bài báo nhằm nâng cao chất lượng và độ tin cậy của tổng quan.



Hình 1: Biểu đồ PRISMA thể hiện quy trình lựa chọn tài liệu cho tổng quan hệ thống

Biểu đồ trên minh họa quy trình thu thập, sàng lọc và lựa chọn tài liệu theo hướng dẫn của PRISMA. Tổng cộng 75 tài liệu đã được truy xuất từ hai cơ sở dữ liệu chính là Scopus và Web of Science thông qua các từ khóa liên quan đến *trí tuệ nhân tạo* và *phản hồi viết* trong bối cảnh người học tiếng Anh như ngoại

ngữ hoặc ngôn ngữ thứ hai. Sau khi loại bỏ 19 tài liệu trùng lặp, còn lại 56 tài liệu được đưa vào bước sàng lọc tiêu đề và tóm tắt, 34 bài được giữ lại và 22 bài bị loại do không liên quan trực tiếp đến việc sử dụng trí tuệ nhân tạo trong phản hồi kĩ năng viết học thuật bằng tiếng Anh. Trong số 34 bài đủ điều kiện để đọc toàn văn, 8 bài không thể truy cập được (do giới hạn truy cập hoặc lỗi đường dẫn), còn lại 26 bài được đọc toàn văn. Kết quả, 8 bài không đáp ứng tiêu chí chọn lọc về đối tượng, phương pháp và nội dung nghiên cứu nên bị loại. Cuối cùng, 18 bài báo được đưa vào phân tích hệ thống trong nghiên cứu này (xem phần tài liệu tham khảo).

3. Kết quả nghiên cứu

Sau quá trình tổng hợp 18 nghiên cứu trong giai đoạn 2020 - 2025, kết quả phân tích nội dung và cho thấy các đặc điểm nổi bật trong xu hướng ứng dụng công cụ trí tuệ thông minh trong phản hồi viết tiếng Anh nhằm trả lời hai câu hỏi nghiên cứu đã đặt ra.

3.1. Các khía cạnh được xử lí

Một trong những năng lực nổi bật của công cụ phản hồi viết tự động là khả năng phản hồi đa dạng về mặt hình thức. Các nghiên cứu thực nghiệm cho thấy, ChatGPT và Grammarly có hiệu quả đáng kể trong việc nhận diện và cải thiện những nội dung phổ biến như sự hòa hợp chủ ngữ - động từ, thì và thể, số nhiều, cũng như việc sử dụng mạo từ (Alsaweed & Aljebreen, 2024). Bên cạnh việc phát hiện lỗi ngữ pháp, công cụ trí tuệ nhân tạo còn thể hiện thế mạnh trong xử lí các khía cạnh liên quan đến từ vựng và phong cách diễn đạt. Nghiên cứu của Mekheimer (2025) cho thấy Grammarly không chỉ phản hồi về việc lựa chọn từ chưa phù hợp hoặc dùng sai giới từ mà còn có khả năng đề xuất từ thay thế và tinh chỉnh văn phong học thuật. Điều này khuyến khích người học thử nghiệm các cách diễn đạt đa dạng và linh hoạt hơn, đồng thời phát triển khả năng viết với sắc thái học thuật rõ nét. Trong khi đó, Chen, Xu, Zhang

& Liu (2023) phát triển một hệ thống viết đa chiến lược tích hợp học sâu, có khả năng đưa ra phản hồi đa chiều về vốn từ vựng, cấu trúc câu và sự đa dạng trong diễn đạt. Kết quả này cho thấy công cụ trí tuệ nhân tạo không chỉ dừng lại ở mức sửa lỗi cục bộ, mà còn mở rộng phạm vi hỗ trợ, giúp người học tiếp cận nhiều chuẩn mực diễn đạt khác nhau và hình thành phong cách viết cá nhân gắn với yêu cầu học thuật.

Một lĩnh vực khác mà phản hồi tự động xử lý hiệu quả là những vấn đề về mặt cơ học (Mechanical errors) trong kỹ năng diễn đạt ở cấp độ câu. Các công cụ như Grammarly và ChatGPT thường xuyên xác định được những yếu tố nhỏ nhưng gây ảnh hưởng đến chất lượng văn bản, bao gồm việc sử dụng dấu câu chưa chính xác, câu quá dài, hoặc trật tự từ thiếu tự nhiên. Các nghiên cứu thực nghiệm và quan sát trong lớp học cho thấy các công cụ còn giúp phát hiện những chi tiết đôi khi bị bỏ sót bởi giáo viên như chính tả hoặc câu chưa hoàn thiện, câu chắp vá (Ho, 2024; Shen, Shi, Guo, Xu & Tian, 2023; Li, Wang & Huang, 2024). Điều này cho phép công cụ trí tuệ nhân tạo đảm nhận vai trò như một “người biên tập kỹ thuật” song hành cùng giáo viên, đảm bảo văn bản đạt chuẩn chỉnh về hình thức, đồng thời rèn luyện cho người học kỹ năng tự rà soát và điều chỉnh trong quá trình viết.

Bên cạnh việc cải thiện ngôn ngữ bề mặt, công cụ phản hồi trí tuệ nhân tạo còn cho thấy tiềm năng trong việc hỗ trợ người học cải thiện tính mạch lạc và phát triển nội dung bài viết. Các hệ thống như ChatGPT thường đưa ra gợi ý về từ nối, cấu trúc chuyển tiếp và cách tổ chức ý, qua đó giúp văn bản trở nên gắn kết và liên mạch hơn (Ho, 2024; Kurt & Kurt, 2024). Shen và các cộng sự (2023) cũng chỉ ra rằng, phản hồi do máy tính tạo ra khuyến khích người học chú ý hơn đến sự liên kết chặt chẽ giữa các câu và đoạn, nhờ đó cải thiện tính mạch lạc tổng thể của bài viết. Ngoài ra, một số nghiên cứu còn ghi nhận khả năng của công cụ trí tuệ nhân tạo trong việc đề xuất ý tưởng và mở rộng nội dung, đặc biệt ở giai đoạn viết nháp, khi người học thường gặp khó khăn trong việc phát triển luận điểm (Li, Huang, Wu & Whipple, 2024; Zhang và các cộng sự, 2025). Chẳng hạn, ChatGPT có thể đưa ra gợi ý ví dụ minh họa, cách triển khai luận điểm phụ, hoặc các khía cạnh liên quan để người học tham khảo, từ đó khuyến khích quá trình tư duy phản biện và làm giàu nội dung bài viết. Có thể thấy, công cụ trí tuệ nhân tạo có thể đóng vai trò hỗ trợ quan trọng, giúp người học vừa duy trì sự gắn kết mạch lạc vừa phát triển nội dung phong phú và đa chiều hơn.

3.2. Hiệu quả sư phạm

a. Tính tức thời và khả năng tiếp cận

Một trong những ưu điểm nổi bật nhất của công cụ phản hồi trí tuệ nhân tạo là tính tức thời và khả năng tiếp cận, tạo ra sự khác biệt so với phản hồi truyền thống từ giáo viên hoặc bạn học. Nhiều nghiên cứu chỉ ra rằng ChatGPT có thể cung cấp nhận xét trong vài giây, cho phép người học chỉnh sửa ngay trong quá trình viết thay vì phải chờ nhiều ngày để nhận phản hồi (Ho, 2024; Kurt & Kurt, 2024). Ở bậc Sau đại học, Grammarly cho thấy hiệu quả khi hỗ trợ phát hiện và điều chỉnh lỗi bề mặt trong giai đoạn viết nháp (Mekheimer, 2025). Các nghiên cứu về công cụ phản hồi viết tự động khẳng định tính tức thời không chỉ giúp giảm tải cho giáo viên mà còn duy trì động lực học tập của sinh viên (Barrot, 2021; Chang, Li, Huang & Whitfield, 2021). Bằng chứng từ những lớp học đồng tại Oman và Ả Rập Xê Út càng củng cố luận điểm này, cho thấy sinh viên đánh giá cao sự kịp thời trong phản hồi khi giáo viên khó có thể hỗ trợ đồng đều cho tất cả (Al Ghaithi & Behforouz, 2025; Alsalem, 2024). Các nghiên cứu gần đây cũng cho rằng tốc độ phản hồi là yếu tố dự báo quan trọng về mức độ chấp nhận công cụ trí tuệ nhân tạo của người học (Shen và các cộng sự, 2023; Teng, 2024).

b. Độ bao quát và tính toàn diện

Công cụ trí tuệ nhân tạo cũng được ghi nhận ở khả năng cung cấp phản hồi đa dạng trên nhiều phương diện: Ngữ pháp, từ vựng, tổ chức và nội dung. So với giáo viên - những người thường tập trung vào một số khía cạnh như phát triển ý nghĩa hoặc tính thuyết phục - ChatGPT thường đưa ra nhận xét rộng hơn, chạm đến cả lỗi cục bộ và toàn văn bản (ElEbyary, Shabara & Boraie, 2024; Ho, 2024; Kurt & Kurt, 2024). Các hệ thống mới như DeepSeek-R1 hay nền tảng đa chiến lược tại Trung Quốc cũng cho thấy khả năng phản hồi đa tiêu chí, bao gồm tính mạch lạc, sự đa dạng từ vựng và cấu trúc logic (Chen và các cộng sự, 2023). Các tổng quan hệ thống cũng nhất quán với nhận định này, coi phạm vi bao quát rộng là điểm mạnh đặc trưng của trí tuệ nhân tạo so với phản hồi truyền thống (Shen và các cộng sự, 2023; Teng, 2024).

c. Quá trình chỉnh sửa nâng cao độ chính xác

Nhiều bằng chứng khẳng định phản hồi tự động có vai trò quan trọng trong việc khuyến khích chỉnh sửa và cải thiện độ chính xác ngôn ngữ. Thí nghiệm với Grammarly cho thấy người học tham gia nhiều vòng chỉnh sửa hơn và tiến bộ rõ rệt về ngữ pháp, dấu câu và cơ chế viết (Mekheimer, 2025). Các nghiên cứu đối chứng cũng cho thấy sinh viên dùng ChatGPT có

cải thiện độ chính xác cao hơn nhóm không dùng công cụ phản hồi tự động (Li, Wang & Huang, 2024). Các nghiên cứu trước về công cụ phản hồi tự động cũng ghi nhận lợi ích cho cả người học trình độ thấp (cải thiện độ chính xác) lẫn cao (tăng vốn từ) (Barrot, 2021). Ngoài ra, phản hồi tự động giúp đẩy nhanh quá trình tự chỉnh sửa, cho phép người học chú ý hiệu quả hơn đến hình thức ngôn ngữ (Shen và các cộng sự, 2023; Teng, 2024). Tại Oman, phản hồi của EditGPT thúc đẩy sinh viên viết lại nhiều hơn và cải thiện chất lượng bản thảo, dù phản hồi giáo viên vẫn thuyết phục hơn (Al Ghaithi & Behforouz, 2025). Ở Trung Quốc, các hệ thống kết hợp trí tuệ nhân tạo và giáo viên cũng khẳng định vai trò củng cố hành vi chỉnh sửa khi phản hồi tự động đi kèm hướng dẫn sự phạm (Zhang và các cộng sự, 2025).

d. Động lực và sự tham gia

Ngoài khía cạnh ngôn ngữ, phản hồi sử dụng trí tuệ nhân tạo còn tác động tích cực đến động lực và sự tham gia của người học. Sinh viên mô tả ChatGPT như một “người bạn đồng hành”, giúp viết bớt áp lực và thú vị hơn (Kurt & Kurt, 2024; Ho, 2024). Trong bối cảnh đại học Trung Quốc, phản hồi tự động được ghi nhận làm giảm sự nản chí và tăng động lực chỉnh sửa (Li và các cộng sự, 2024). Ở bậc sau đại học, người học dùng Grammarly báo cáo tự tin hơn và ít lo lắng hơn, đồng thời được khuyến khích thử nghiệm vốn từ và cấu trúc mới (Mekheimer, 2025). Các nghiên cứu về hybrid feedback cũng cho thấy không chỉ cải thiện điểm số mà còn tăng thái độ tích cực của sinh viên (Zhang và các cộng sự, 2025). Bằng chứng từ Oman bổ sung rằng, sinh viên sẵn sàng viết nhiều bản thảo hơn khi có phản hồi tự động, duy trì sự gắn kết (Al Ghaithi & Behforouz, 2025). Các tổng quan khẳng định động lực và hỗ trợ cảm xúc là yếu tố cốt lõi trong việc chấp nhận trí tuệ nhân tạo vì công cụ này giảm bớt lo sợ bị phán xét thường gắn với phản hồi từ con người (Shen và các cộng sự, 2023; Teng, 2024).

e. Sự hỗ trợ với phản hồi của con người

Các bằng chứng nhất quán cho thấy trí tuệ nhân tạo phát huy hiệu quả tối đa khi bổ sung cho phản hồi của giáo viên và bạn học, thay vì thay thế hoàn toàn. Các nghiên cứu so sánh cho thấy phản hồi kết hợp giữa ChatGPT và giáo viên giúp sinh viên cải thiện toàn diện hơn, bao gồm cả kỹ năng ngôn ngữ và tư duy bậc cao (Zhang và các cộng sự, 2025). Giáo viên cũng thừa nhận lợi ích của trí tuệ nhân tạo trong việc tăng hiệu quả nhưng nhấn mạnh cần thận trọng trong bối cảnh thi cử, đánh giá (Alsalem,

2024). Nghiên cứu tại Oman cũng cho thấy công cụ trí tuệ hữu ích nhưng phản hồi giáo viên vẫn vượt trội hơn về tính phù hợp và chiều sâu (Al Ghaithi & Behforouz, 2025). Các tổng quan hệ thống khẳng định công cụ chữa lỗi tự động nên được coi là “vòng phản hồi đầu tiên”, sau đó cần bổ sung của giáo viên để đảm bảo bối cảnh và định hướng (Shen và các cộng sự, 2023; Teng, 2024). Ngay cả các nghiên cứu về độ tin cậy cao của công cụ chữa lỗi tự động cũng nhấn mạnh rằng phân tích ở cấp độ diễn ngôn vẫn cần giám sát của con người (Ho, 2024).

4. Thảo luận

4.1. Những hạn chế và rủi ro

Mặc dù các công cụ phản hồi tự động được ghi nhận với nhiều lợi ích, song các nghiên cứu cũng chỉ ra không ít hạn chế. Trước hết, phản hồi do trí tuệ nhân tạo tạo ra đôi khi thiếu ổn định và chưa nhất quán, chất lượng còn phụ thuộc đáng kể vào cách thức đặt câu lệnh (Alsaweed & Aljebreen, 2024; Ho, 2024; Kurt & Kurt, 2024). Một số hệ thống đưa ra những gợi ý khác nhau cho cùng một dạng diễn đạt hoặc thay đổi giữa các lần sử dụng, khiến người học dễ rơi vào tình trạng hoang mang. Ngoài ra, hiệu suất chỉnh sửa cũng không đồng đều: Các lỗi ngữ pháp, từ vựng và cơ học như dấu câu thường được xử lý khá chính xác, trong khi những vấn đề ở cấp độ diễn ngôn và ý nghĩa lại ít được cải thiện hoặc cải thiện chưa triệt để. Chẳng hạn, ChatGPT thường gặp khó khăn trong việc nâng cao tính mạch lạc giữa các đoạn, làm rõ những ý tưởng mơ hồ hoặc củng cố lập luận chưa hợp lý (Alsaweed & Aljebreen, 2024; Kurt & Kurt, 2024). Các báo cáo từ giáo viên cũng cho thấy những công cụ như CoGrader hay Grammarly vẫn chưa đáp ứng được kì vọng trong việc phát hiện và hỗ trợ điều chỉnh các vấn đề liên quan đến tính gắn kết, logic hoặc tổ chức văn bản ở cấp độ toàn cục (Alsalem, 2024; Al Ghaithi & Behforouz, 2025). Những hạn chế này phản ánh đặc trưng của trí tuệ nhân tạo: Mạnh trong xử lý dữ liệu ngôn ngữ bề mặt nhưng chưa đủ khả năng thay thế các phán đoán sự phạm vốn gắn liền với bối cảnh, mục tiêu giao tiếp và ý đồ diễn ngôn. Do đó, sự giám sát và định hướng của giáo viên vẫn là yếu tố then chốt để đảm bảo chuẩn mực học thuật cũng như giá trị sự phạm của phản hồi.

Một vấn đề khác được nhiều nghiên cứu nhấn mạnh là nguy cơ lệ thuộc quá mức vào công cụ trí tuệ nhân tạo, dẫn đến suy giảm khả năng tự chỉnh sửa và hạn chế sự phát triển tư duy phản biện của người học. Mekheimer (2025) cảnh báo rằng, sinh viên có

xu hướng dựa dẫm vào các gợi ý của Grammarly và bỏ qua việc tự rà soát bản thảo, từ đó làm giảm tính chủ động trong học tập. Tương tự, Kurt & Kurt (2024) ghi nhận lo ngại rằng phản hồi của ChatGPT mặc dù hữu ích trong ngắn hạn có thể khiến người viết “mất đi tiếng nói cá nhân” khi quá tin tưởng vào nội dung do trí tuệ nhân tạo đề xuất. Bên cạnh đó, các nghiên cứu cũng đặt ra những câu hỏi đạo đức liên quan đến quyền sở hữu trí tuệ và sự công bằng học thuật. Ho (2024) và Alsalem (2024) đều nhấn mạnh rằng, việc sử dụng công cụ trí tuệ nhân tạo trong lớp học đòi hỏi phải có hướng dẫn rõ ràng nhất là trong bối cảnh thi cử và đánh giá chính thức. Ngoài ra, Zhang và các cộng sự (2025) cũng lưu ý rằng, phản hồi lai (hybrid feedback) chỉ phát huy hiệu quả khi học viên được khuyến khích khai thác trí tuệ nhân tạo như một công cụ hỗ trợ, thay vì thay thế hoàn toàn phản hồi của giáo viên hoặc bạn học. Những phát hiện này cho thấy vấn đề không chỉ dừng lại ở hiệu quả ngôn ngữ mà còn liên quan chặt chẽ đến đạo đức học thuật và sự hình thành năng lực viết bền vững.

Tóm lại, công cụ trí tuệ nhân tạo trong phản hồi viết tiếng Anh mang lại những lợi ích đáng kể về độ chính xác, tốc độ và phạm vi hỗ trợ, đặc biệt trong việc xử lý lỗi bề mặt và nâng cao phong cách diễn đạt. Tuy nhiên, ở những kỹ năng viết bậc cao như tổ chức văn bản, lập luận và phát triển ý tưởng, phản hồi từ giáo viên và bạn học cũng như chính người học vẫn giữ vai trò không thể thay thế. Do đó, trí tuệ nhân tạo nên được coi là một thành tố bổ sung quan trọng trong hệ sinh thái phản hồi thay vì một giải pháp thay thế hoàn toàn cho sự can thiệp sư phạm của con người.

4.2. Khoảng trống nghiên cứu và đề xuất

Mặc dù số lượng nghiên cứu về phản hồi có sự hỗ trợ của trí tuệ nhân tạo trong viết tiếng Anh ngày càng gia tăng, vẫn còn tồn tại nhiều khoảng trống quan trọng.

Thứ nhất, phần lớn các nghiên cứu hiện nay chủ yếu dựa trên báo cáo tự thuật hoặc cảm nhận của sinh viên, với rất ít sự đối chiếu từ giáo viên hoặc dữ liệu quá trình lớp học, khiến việc tìm hiểu cách tiếp nhận thực tế và sự tích hợp phản hồi tự động trong giảng dạy chưa được khai thác đầy đủ (Ho, 2024; Kurt & Kurt, 2024).

Thứ hai, nhiều can thiệp mang tính ngắn hạn và tập trung hẹp, thường chỉ giới hạn trong một hoặc hai nhiệm vụ viết hoặc một thể loại cụ thể, làm hạn chế khả năng rút ra kết luận về tác động dài hạn và khả năng chuyển giao trong nhiều bối cảnh viết học

thuật khác nhau (ElEbyary, Shabara & Boraie, 2024; Strobl, Menke-Bazhutkina, Abel & Michel, 2024).

Thứ ba, trong khi các cải thiện về độ chính xác được ghi nhận rõ ràng, vẫn còn ít nghiên cứu so sánh tác động của phản hồi tự động đối với các kỹ năng bậc cao như lập luận, tính mạch lạc và tư duy phản biện (Li và các cộng sự, 2024; Zhang và các cộng sự, 2025).

Một khoảng trống khác liên quan đến các biến số của người học. Dù một số nghiên cứu nhấn mạnh vai trò của nhận thức siêu nhận thức, kỹ năng đưa câu lệnh và năng lực tiếp nhận phản hồi trong việc quyết định hiệu quả sử dụng trí tuệ nhân tạo, nhưng rất ít nghiên cứu kiểm chứng các yếu tố này như biến số điều tiết trong các thiết kế thực nghiệm (Chen và các cộng sự, 2023; Alsaweed & Aljebreen, 2024). Bên cạnh đó, phần lớn nghiên cứu chỉ đánh giá một hệ thống trí tuệ nhân tạo duy nhất, dẫn đến việc thiếu bằng chứng so sánh trực tiếp giữa các nền tảng khác nhau (Ví dụ: ChatGPT, Grammarly, DeepSeek, CoGrader) (Alsalem, 2024; Al Ghaithi & Behforouz, 2025). Điều này hạn chế khả năng xác định công cụ nào phù hợp nhất với từng bối cảnh học tập hoặc mục đích viết cụ thể.

Ngoài ra, những lo ngại về tính hợp lệ và độ tin cậy cũng chưa được giải quyết triệt để. Một số bằng chứng cho thấy các mô hình tiên tiến như ChatGPT-4 có thể đạt độ tin cậy chấm điểm tương đương với giáo viên, nhưng sự phù hợp với thang đo và độ chính xác ở cấp độ diễn ngôn vẫn chưa nhất quán (Li và các cộng sự, 2024). Hơn nữa, dù các cách tiếp cận kết hợp thường cho thấy ưu thế vượt trội so với phản hồi chỉ từ công cụ trí tuệ nhân tạo, lĩnh vực này vẫn thiếu hướng dẫn chi tiết về cách phối hợp hiệu quả giữa phản hồi tự động, giáo viên và bạn học trong các tình huống lớp học thực tế (Alsalem, 2024; Zhang và các cộng sự, 2025). Cuối cùng, các vấn đề đạo đức và công bằng - bao gồm sự phụ thuộc quá mức vào trí tuệ nhân tạo, việc duy trì tính sở hữu của người học đối với bài viết, cũng như sự khác biệt trong khả năng tiếp cận ở những bối cảnh giáo dục có nguồn lực không đồng đều - thường được thừa nhận nhưng hiếm khi được khảo sát thực nghiệm (Mekheimer, 2025).

Tóm lại, để khắc phục những khoảng trống này, các nghiên cứu tương lai cần áp dụng thiết kế dài hạn, đa thể loại và phương pháp hỗn hợp, nhằm không chỉ đo lường độ chính xác mà còn theo dõi sự phát triển kỹ năng viết bậc cao, kết hợp cả góc nhìn của người học và giáo viên cũng như làm rõ những hệ quả về đạo đức, động lực và sự phạm của

phản hồi tự động trong các bối cảnh học tiếng Anh đa dạng. Ngoài ra, việc xây dựng các mô hình phản hồi kết hợp giữa trí tuệ nhân tạo, giáo viên và bạn học cần được nghiên cứu sâu hơn, nhằm tìm ra sự phân bổ hợp lý về vai trò và mức độ can thiệp. Song song với đó, đào tạo năng lực sử dụng trí tuệ nhân tạo (Artificial Intelligence literacy) cho giáo viên và người học cũng cần được xem xét như một hướng đi chiến lược, giúp họ biết cách khai thác, đánh giá và tích hợp trí tuệ nhân tạo một cách hiệu quả, có đạo đức và phù hợp với mục tiêu sư phạm.

5. Kết luận

Tổng quan các nghiên cứu cho thấy, việc ứng dụng trí tuệ nhân tạo trong phản hồi kỹ năng viết tiếng Anh mang lại nhiều lợi ích nổi bật, đặc biệt trong khả năng xử lý nhanh chóng các lỗi ngôn ngữ bề mặt, cung cấp phản hồi tức thì và góp phần duy trì động lực học tập của người học. Tuy nhiên, bằng chứng cũng khẳng định rằng, các công cụ trí tuệ nhân tạo chưa đủ hiệu quả trong việc hỗ trợ các khía cạnh viết ở bậc cao như tổ chức ý tưởng, lập luận

hay tư duy phản biện, những lĩnh vực vẫn đòi hỏi sự hướng dẫn trực tiếp từ giáo viên và sự hợp tác của bạn học. Chính vì vậy, mô hình phản hồi kết hợp giữa công cụ đánh giá tự động và phản hồi của con người được xem là giải pháp tối ưu vừa tận dụng được thế mạnh công nghệ vừa đảm bảo chiều sâu sư phạm và tính phát triển toàn diện của kỹ năng viết.

Để hiện thực hóa tiềm năng này, cần triển khai đồng bộ nhiều giải pháp mang tính chiến lược, bao gồm: đổi mới chương trình đào tạo nhằm tích hợp hiệu quả công nghệ trí tuệ nhân tạo vào giảng dạy viết; bồi dưỡng năng lực khai thác và giám sát công cụ trí tuệ nhân tạo cho giáo viên cũng như người học; xây dựng bộ tiêu chí chuẩn để đánh giá chất lượng phản hồi tự động; đồng thời mở rộng các nghiên cứu thực nghiệm dài hạn, đa bối cảnh và đa đối tượng. Những định hướng này không chỉ góp phần nâng cao chất lượng giảng dạy viết học thuật mà còn đảm bảo sự phát triển bền vững và có trách nhiệm của việc ứng dụng trí tuệ nhân tạo trong giáo dục ngoại ngữ.

Tài liệu tham khảo

- Al Ghaithi, A. & Behforouz, B. (2025). Effects of interaction with AI-assisted writing evaluation on EFL students' writing performance. *Knowledge Management & E-Learning*, 17(1), 95–111. <https://doi.org/10.34105/j.kmel.2025.17.009>
- Alsalem, M. S. (2024). EFL teachers' perceptions of the use of an AI grading tool (CoGrader) in English writing assessment at Saudi universities: An Activity Theory perspective. *Cogent Education*, 11(1), 2430865. <https://doi.org/10.1080/2331186X.2024.2430865>
- Alsaweed, W. & Aljebreen, S. (2024). Investigating the accuracy of ChatGPT as a writing error correction tool. *International Journal of Computer-Assisted Language Learning and Teaching*, 14(2), 1–15. <https://doi.org/10.4018/IJCALLT.364847>
- Barrot, J. S. (2021). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 34(5–6), 707–729. <https://doi.org/10.1080/09588221.2021.1936071>
- Bitchener, J. & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge.
- Chang, T., Li, Y., Huang, H. & Whitfield, B. (2021). Exploring EFL students' writing performance and their acceptance of AI-based automated writing feedback. In *Proceedings of the 2021 5th International Conference on Education and E-Learning (ICEDL 2021)* (pp. 130–135). ACM. <https://doi.org/10.1145/3459043.3459065>
- Chen, B., Bao, L., Zhang, R., Zhang, J., Liu, F., Wang, S. & Li, M. (2024). A multi-strategy computer-assisted EFL writing learning system with deep learning incorporated and its effects on learning: A writing feedback perspective. *Journal of Educational Computing Research*, 62(1), 122–148. <https://doi.org/10.1177/07356331231189294>
- Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford University Press.
- ElEbyary, K. & Shabara, R. (2024). ChatGPT-generated corrective feedback: Does it do what it says on the tin? *Teaching English with Technology*, 24(3), 45–63. <https://doi.org/10.56297/vaca6841/BFFO7057/MYEH4562>
- ElEbyary, K., Shabara, R. & Boraie, D. (2024). The differential role of AI-operated WCF in L2 students' noticing of errors and its impact on writing scores. *Language Testing in Asia*, 14, 312. <https://doi.org/10.1186/s40468-024-00312-1>
- Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81–104). Cambridge University Press.
- Ho, S. L. (2024). Exploring the use of ChatGPT as a tool for written corrective feedback in an EFL classroom. *Journal of Asia TEFL*, 21(2), 397–416. <https://doi.org/10.18823/asiatefl.2024.21.2.8.397>

- Kim, M. & Chon, Y. (2025). The impact of self-revision, machine translation, and ChatGPT on L2 writing: Raters' assessments, linguistic complexity, and error correction. *Assessing Writing*, 55, 100950. <https://doi.org/10.1016/j.asw.2025.100950>
- Kurt, G. & Kurt, Y. (2024). Enhancing L2 writing skills: ChatGPT as an automated feedback tool. *Journal of Information Technology Education: Research*, 23, 185–204. <https://doi.org/10.28945/5370>
- Li, J., Huang, J., Wu, W. & Whipple, P. B. (2024). Evaluating the role of ChatGPT in enhancing EFL writing assessments in classroom settings: A preliminary investigation. *Humanities and Social Sciences Communications*, 11, 755. <https://doi.org/10.1057/s41599-024-03755-2>
- Li, M. (2024). Leveraging ChatGPT for second language writing feedback and assessment. *International Journal of Computer-Assisted Language Learning and Teaching*, 14(1), 1–18. <https://doi.org/10.4018/IJCALLT.360382>
- Mekheimer, M. A. A. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency, and writing quality. *Discover Education*, 4, 170. <https://doi.org/10.1007/s44217-025-00602-7>
- Schunk, D. H. & Zimmerman, B. J. (Eds.). (2007). *Motivation and self-regulated learning: Theory, research, and applications*. Lawrence Erlbaum Associates.
- Shen, C., Shi, P., Guo, J., Xu, S. & Tian, J. (2023). From process to product: Writing engagement and performance of EFL learners under computer-generated feedback instruction. *Frontiers in Psychology*, 14, 1258286. <https://doi.org/10.3389/fpsyg.2023.1258286>
- Shi, H. & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, 36(2), 187–209. <https://doi.org/10.1017/S0958344023000265>
- Teng, M. F. (2024). A systematic review of ChatGPT for English as a Foreign Language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies*, 6(3), 36–57. <https://doi.org/10.58304/ijts.20240304>
- Teng, M. F. (2025). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *European Journal of Education*. Advance online publication. <https://doi.org/10.1111/ejed.12811>
- Zhang, J. & Zhang, L. J. (2025). Integrating various types of feedback in L2 writing instruction: Teachers' and students' perspectives. *International Journal of Applied Linguistics*. Advance online publication. <https://doi.org/10.1111/ijal.12822>
- Zhang, Z., Aubrey, S., Huang, X. & Chiu, T. K. F. (2025). The role of generative AI and hybrid feedback in improving L2 writing skills: A comparative study. *Innovation in Language Learning and Teaching*. Advance online publication. <https://doi.org/10.1080/17501229.2025.2503890>