

PHẢN HỒI SỬA LỖI TỰ ĐỘNG TRONG HỌC VIẾT NGÔN NGỮ: TỔNG QUAN HỆ THỐNG NGHIÊN CỨU THỰC NGHIỆM

TRẦN THỦY VINH
NGUYỄN TRẦN VIỆT HỒNG
Trường Đại học Quốc tế Hồng Bàng

Nhận bài ngày 18/7/2026. Sửa chữa xong 10/8/2026. Duyệt đăng 15/8/2026.

Abstract

This systematic review examines the development and application of automated writing evaluation (AWE) in language writing instruction. AWE systems are increasingly used to reduce teachers' workloads and provide learners with more frequent feedback. However, systematic syntheses of high-quality empirical research in this field remain limited. To address this gap, the review analyzes empirical studies published over the past decade, focusing on research methodologies, participants' proficiency levels, and the benefits and challenges of using AWE. The findings indicate that AWE can improve overall writing performance, linguistic accuracy, lexical diversity, and syntactic complexity, particularly across multiple drafts. Nevertheless, AWE primarily addresses lower-order linguistic features and is less effective in providing feedback on content, organization, meaning, and communicative intent. The review therefore emphasizes that AWE should complement rather than replace teacher feedback. Future research should adopt more rigorous methodologies and examine genre-related differences, long-term effects, cognitive and affective engagement, effective pedagogical integration, and the development of advanced AWE tools capable of providing higher-order feedback.

Keywords: Automated feedback, automated writing evaluation, language writing, systematic review, writing performance.

1. Đặt vấn đề

Trong dạy và học viết ngoại ngữ, phản hồi sửa lỗi bằng văn bản là rất cần thiết để nâng cao kỹ năng viết và trình độ ngôn ngữ của người học (L.Xiaosa and K.Ping, 2025); (B. Yoonkyeong and J. YeonJoo, 2024). Trong những thập kỷ qua, cùng với những tiến bộ vượt bậc trong công nghệ, các công cụ phản hồi sửa lỗi bài viết tự động (AWE) đã ra đời và ngày càng được ứng dụng rộng rãi (L.Xiaosa and K.Ping, 2025). Các hệ thống AWE có khả năng đánh giá bằng máy tính, chấm điểm tức thì và đưa ra phản hồi cá nhân hóa dưới dạng văn bản (G. T. L. Hoang and N. Storch, 2024). Nhiều nghiên cứu cho thấy AWE giúp giảm tải cho người dạy khi xử lý các vấn đề viết ở bậc thấp (low-order concerns - LOC) như chính tả, ngữ pháp và cấu trúc câu, tạo điều kiện cho người dạy tập trung nhiều hơn vào các vấn đề bậc cao (high-order concerns- HOC), chẳng hạn như nội dung, lập luận và bố cục (L.Xiaosa and K.Ping, 2025); (S.Koltovskaia, 2023). Hơn nữa, AWE còn tạo cơ hội sửa bài qua nhiều vòng, tăng tần suất phản hồi và nuôi dưỡng tính tự chủ của người học (J. Woodworth and K. Barkaoui; (G. T. L. Hoang and N. Storch, 2024); (S. Koltovskaia, 2023). Tuy nhiên, hiệu quả của AWE đối với sự phát triển kỹ năng viết một cách toàn diện và lâu dài vẫn chưa được khẳng định thống nhất do các kết quả nghiên cứu còn khác biệt. Một số nghiên cứu tại Philippines và Hàn Quốc cho thấy AWE giúp người học tăng mức độ sử dụng chính xác từ vựng và ngữ pháp sau khóa học (J.S. Barrot (2023); (K. Lim, J. Song, and J. Park), nhưng chưa làm rõ mức độ duy trì hiệu quả về dài hạn. Về phía người học, phản hồi đối với AWE thường tích cực bởi tính liên tục và tức thì của hệ thống (J. Woodworth and K. Barkaoui), và nhiều người học cho biết khả năng viết của bản thân được cải thiện, cũng như mức độ lo lắng khi viết giảm mặc dù phản hồi từ hệ thống đôi khi còn mơ hồ và một vài điểm sửa lỗi không cần thiết (J. Woodworth and K. Barkaoui; D. Gilbert and G. Jason. Bên cạnh đó, giáo viên (GV) cũng nhìn nhận AWE là công cụ hỗ trợ hữu ích, song cũng

Email: vinhtt@hiu.vn; viethong3989@gmail.com

nhấn mạnh việc kết hợp phản hồi của giáo viên với AWE là cần thiết để phát huy tối đa ưu điểm của công cụ (S.Koltovskaia, 2023). Chính vì có nhiều kết quả trái chiều trong các nghiên cứu về AWE, cùng với xu hướng phát triển của các hệ thống AI tạo sinh, việc tích hợp có hệ thống các bằng chứng thực nghiệm hiện có là điều cần thiết để định hướng nghiên cứu và triển khai AWE trong thực tiễn. Dù các tổng hợp gần đây có ghi nhận sự gia tăng nhanh của các nghiên cứu về AWE, nhưng vẫn thiếu một đánh giá có hệ thống chuyên sâu cung cấp tổng hợp toàn diện về phương pháp luận và các kết quả thực nghiệm liên quan đến hiệu quả cũng như cách triển khai AWE trong bối cảnh dạy–học khác nhau. Vì vậy, bài viết này nhằm lấp khoảng trống đó bằng cách tiến hành tổng quan hệ thống các nghiên cứu thực nghiệm về AWE trong 10 năm qua, phân tích và đối chiếu các kết quả học tập khi sử dụng AWE, đồng thời đề xuất những hướng nghiên cứu tiếp theo trong bối cảnh AWE tiếp tục được phát triển và ứng dụng rộng rãi.

2. Mục tiêu nghiên cứu

Nghiên cứu tổng quan hệ thống này nhằm phân tích xu hướng nghiên cứu trong các công trình thực nghiệm về phản hồi sửa lỗi tự động (AWE) giai đoạn 2016–2025, qua đó trả lời các câu hỏi sau: 1) Những phương pháp thu thập dữ liệu nào được sử dụng phổ biến trong các nghiên cứu thực nghiệm về AWE trong học viết ngoại ngữ (2016–2025)?; 2) Đối tượng chủ yếu và kết quả nghiên cứu thực nghiệm nổi bật về hiệu quả sử dụng AWE trong 10 năm qua là gì?; 3) Vai trò của AWE, bao gồm lợi ích và thách thức, trong học viết ngoại ngữ được thể hiện như thế nào trong các nghiên cứu thực nghiệm?

3. Cơ sở lý thuyết

3.1. Tổng quan về AWE

AWE bắt nguồn từ những năm 1960s gắn với sự phát triển của các dự án chấm bài viết bằng máy tính tại Hoa Kỳ, tiêu biểu là dự án “Essay Grade”, nhằm đánh giá những bài viết đã được con người chấm trước đó (Y. Karatay and L. Karatay, 2024). Khi cung cấp phản hồi sửa lỗi tự động, AWE đã vận dụng sự kết hợp của nhiều hướng tiếp cận, bao gồm chấm điểm và đánh giá bài viết, các thuật toán xử lý ngôn ngữ và trí tuệ nhân tạo (Y. Karatay and L. Karatay, 2024). Trong vài thập kỷ qua, trọng tâm nghiên cứu đã chuyển từ việc kiểm định độ tin cậy của AWE sang khảo sát cách thức AWE được triển khai trong môi trường lớp học và tác động của nó đến người học (L.Xiaosa and K.Ping, 2025). Sự chuyển đổi này đã dẫn đến sự ra đời của các hệ thống như Criterion, My Access! và Grammarly, hướng đến đánh giá quá trình thay vì chỉ đánh giá tổng kết dựa trên sản phẩm cuối cùng (L.Xiaosa and K.Ping, 2025). Trong những năm gần đây, AWE được ứng dụng ngày càng phổ biến trong các lớp học, đặc biệt là lớp học ngôn ngữ, nhờ khả năng cung cấp phản hồi nhanh, tự động và tạo điều kiện cho người học chỉnh sửa, nộp lại nhiều lần. Trên cơ sở đó, nhiều nghiên cứu thực nghiệm đã được thực hiện nhằm làm rõ hơn ảnh hưởng của AWE đối với sự phát triển kỹ năng viết và năng lực ngôn ngữ của người học.

3.2. Các nghiên cứu tổng quan trước đây về sử dụng AWE

Sự phát triển nhanh của công nghệ, đặc biệt là các công cụ AWE, trong những năm gần đây đã thu hút sự quan tâm đáng kể của giới nghiên cứu. Trong khoảng năm năm trở lại đây, nhiều công trình tổng quan đã được thực hiện nhằm hệ thống hóa kết quả và xu hướng nghiên cứu liên quan đến AWE. Tiêu biểu, Fan và Ma (2022) tiến hành một tổng quan hệ thống về tác động của AWE đối với chất lượng bài viết của người học, phân tích 22 nghiên cứu công bố trong giai đoạn 2005–2020 (N. Fan and Y. Ma, 2022). Kết quả cho thấy AWE có thể cải thiện kỹ năng viết của sinh viên (SV) trong một số điều kiện nhất định, chẳng hạn khi được sử dụng như một biện pháp can thiệp trong nghiên cứu thực nghiệm (N. Fan and Y. Ma, 2022). Đồng thời, nhờ phản hồi tức thì, AWE còn góp phần thúc đẩy người học chỉnh sửa bài viết thường xuyên hơn và tăng mức độ tham gia vào quá trình viết (N. Fan and Y. Ma, 2022). Cũng trong năm 2022, Fu và cộng sự thực hiện một tổng quan nghiên cứu, phân tích 48 bài báo về AWE nhằm nhận diện xu hướng nghiên cứu, các loại phản hồi và lợi ích của AWE (Q. Fu, D. Zou, H. Xie, and G. Cheng, 2022). Kết quả cho thấy AWE có tác động tích cực đến kỹ năng viết, giảm lo âu khi viết, và tăng mức độ tham gia của người học trong quá trình chỉnh sửa cũng như trong suốt khóa học. Nghiên cứu đồng thời ghi nhận hai hướng tiếp cận phổ biến nhất là phương pháp thuần định lượng và phương pháp kết hợp định lượng–định tính. Về khách thể, phần lớn nghiên cứu tập trung vào người học trình độ trung cấp đến cao cấp, chủ yếu là SV. Ở một góc độ khác, Zhai và Ma (2022) khảo sát mức độ chấp nhận

AWE của SV trong học tập (N. Zhai and X. Ma, 2022), và kết quả cho thấy đa số người học có thái độ tích cực đối với việc ứng dụng AWE trong học viết. Barrot (2023) tiến hành một phân tích đo lường thư mục (bibliometric analysis) nhằm khảo sát tác động của AWE, cùng xu hướng và các chủ đề nghiên cứu liên quan trong giai đoạn 2002–2021 (J. Barot, 2023). Tương tự các tổng quan trước, tác giả cho thấy khách thể nghiên cứu chủ yếu là SV ở trình độ trung cấp và các nghiên cứu thuần định lượng được sử dụng phổ biến nhất trong giai đoạn này. Kết quả cũng ghi nhận AWE góp phần cải thiện vốn từ vựng và ngữ pháp của người học sau một thời gian sử dụng. Ding và Zou (2024) thực hiện một tổng quan hệ thống về tác động của ba công cụ AWE phổ biến gồm Criterion, Pigai và Grammarly (L. Ding and D. Zou, 2024). Kết quả cho thấy hiệu quả của từng công cụ phụ thuộc vào mức độ phù hợp với đối tượng sử dụng; chẳng hạn, Grammarly có xu hướng cá nhân hóa phản hồi nhiều hơn Criterion, trong khi Pigai được thiết kế chủ yếu cho người học tại Trung Quốc. Mặc dù đã có một số công trình tổng hợp về sự phát triển của AWE trong những năm gần đây, các nghiên cứu này chủ yếu tập trung vào lợi ích của từng công cụ cụ thể; trong khi đó, một tổng quan hệ thống chuyên biệt về các nghiên cứu thực nghiệm trong thập kỷ vừa qua vẫn chưa được thực hiện. Vì vậy, bài viết này nhằm lấp khoảng trống nêu trên bằng cách tổng hợp các nghiên cứu thực nghiệm trên thế giới giai đoạn 2016–2025, qua đó phân tích xu hướng và đặc điểm khách thể nghiên cứu, đồng thời làm rõ vai trò của AWE, bao gồm lợi ích và thách thức, trong học viết ngoại ngữ. Trên cơ sở đó, nghiên cứu góp phần gợi mở các hướng tiếp cận toàn diện hơn nhằm tối ưu hóa việc ứng dụng AWE trong dạy và học viết ngoại ngữ.

4. Phương pháp nghiên cứu

Nghiên cứu này được thực hiện theo thiết kế tổng quan có hệ thống của PRISMA 2020 (M. Page et al, 2021), nhằm tổng hợp thông tin về xu hướng thu thập dữ liệu trong các nghiên cứu thực nghiệm trong 10 năm qua ở các nước trên thế giới; đồng thời qua các kết quả nghiên cứu đó, định hướng cho các nghiên cứu thực nghiệm trong tương lai về AWE ở bậc đại học.

4.1. Công cụ và nền tảng thu thập dữ liệu

Nhóm tác giả quyết định chọn Scopus và Web of Science là 2 nền tảng thu thập dữ liệu chính, vì đây là các nền tảng có dữ liệu bài báo đáng tin cậy và chuyên về ngôn ngữ học ứng dụng, giảng dạy ngôn ngữ, đặc thù ngôn ngữ nói chung (Y. Karatay and L. Karatay, 2024). Ngoài ra, các nền tảng này được sử dụng rộng rãi trên thế giới và cách thiết lập chỉ mục dễ hiểu và thân thiện với người sử dụng.

Từ khóa tìm kiếm được xây dựng theo hai nhóm: (i) nhóm thuật ngữ về AWE và (ii) nhóm thuật ngữ về bối cảnh học ngôn ngữ. Cụ thể, nhóm (i) gồm “automated writing evaluation” OR “AWE” OR “automated essay scoring” OR “automated feedback” OR “computer-assisted feedback”; nhóm (ii) gồm “second language” OR “English as a second language” OR “English as a foreign language”. Hai nhóm từ khóa được kết hợp bằng toán tử AND để truy xuất các bài báo liên quan.

4.2. Tiêu chí lựa chọn

Nhóm tác giả thu được 82 bài báo sau tìm kiếm ban đầu trên 2 nền tảng, trong đó có 25 bài từ Scopus và 57 bài từ Web of Science. Sau bước phân tích thứ hai, nhóm tác giả đã loại bỏ 7 bài trùng từ hai hệ thống. Ở bước kế tiếp, nhóm tác giả đọc tiêu đề và tóm tắt của 75 bài còn lại và tiếp tục loại bỏ các bài không liên quan. Các tiêu chí được bao gồm trong quá trình phân tích như sau: Bài báo được bình duyệt khoa học và công bố từ năm 2016 đến tháng 9/2025. Bài báo là nghiên cứu thực nghiệm, sử dụng phương pháp định lượng, định tính, hoặc kết hợp cả hai. Bài báo có kết quả nghiên cứu liên quan kỹ năng viết ngôn ngữ, phản hồi từ người học, hoặc chất lượng bài viết. Bài báo được viết bằng tiếng Anh và có toàn văn. Các bài báo là nghiên cứu lý thuyết, báo tường, hoặc tóm tắt, hoặc kỹ yếu, hoặc chưa qua bình duyệt và không có kết quả nghiên cứu, sẽ không được bao gồm trong bài nghiên cứu tổng quan hệ thống này. Sau phân tích theo tiêu chí, tổng số bài còn lại là 36 bài. Hình 1 minh họa quá trình thu thập và phân tích dựa theo hướng dẫn từ PRISMA 2020 (M. Page et al, 2021).

4.3. Phân tích số liệu

Nghiên cứu tổng quan này sử dụng phương pháp miêu tả số liệu và phân tích theo chủ đề. Miêu tả số liệu là phương pháp phổ biến được sử dụng trong các nghiên cứu định lượng liên quan đến

sử dụng công nghệ trong giáo dục (J. S. Barrot, 2023), và phân tích theo chủ đề có thể giúp các nhà nghiên cứu đi sâu vào vấn đề và hiểu sâu hơn về từng khía cạnh của vấn đề được nghiên cứu (Y. Karatay and L. Karatay, 2024). Nhóm tác giả đọc và phân tích sâu các bài báo, sau đó cùng đối chiếu số liệu và thống nhất về các chủ đề lớn cũng như chủ đề nhỏ hơn.

5. Kết quả và thảo luận

5.1. Các phương pháp thu thập dữ liệu phổ biến trong nghiên cứu thực nghiệm về AWE trong học viết giai đoạn 2016–2025

Các nghiên cứu thực nghiệm về AWE trong bối cảnh lớp học chủ yếu sử dụng bốn hướng phương pháp: thuần định lượng, thuần định tính, kết hợp định lượng–định tính và nghiên cứu theo chiều dọc (longitudinal research). Bảng 1 tóm tắt số lượng nghiên cứu theo từng hướng trong giai đoạn 2016–2025.

Bảng 1: Thống kê số lượng nghiên cứu theo phương pháp

Phương pháp nghiên cứu	Số lượng bài nghiên cứu thực nghiệm	Phương pháp thu thập dữ liệu
Thuần định lượng	10	Kiểm tra trước và sau can thiệp Bảng hỏi +Thang đo Likert Ghi hình màn hình Theo dõi chuyển động mắt Nhật ký học tập Nhớ lại có định hướng Theo dõi và hệ thống hoá hành vi sửa bài Phòng vấn bán cấu trúc
Thuần định tính	6	
Kết hợp định lượng và định tính	16	
Nghiên cứu nhiều trường hợp điển hình	4	

Số liệu cho thấy phương pháp kết hợp định lượng–định tính là hướng tiếp cận phổ biến nhất trong 10 năm qua, tiếp đến là nghiên cứu thuần định lượng. Kết quả này phản ánh xu hướng ưu tiên thiết kế nghiên cứu hỗn hợp nhằm vừa đo lường hiệu quả, vừa làm rõ quá trình học viết khi sử dụng AWE, đồng thời tương đồng với tổng quan của Fu và cộng sự (2022). Về thu thập dữ liệu, các nghiên cứu sử dụng nhiều nguồn dữ liệu khác nhau; trong đó, bài kiểm tra trước–sau can thiệp và phỏng vấn bán cấu trúc là hai phương pháp được áp dụng thường xuyên nhất.

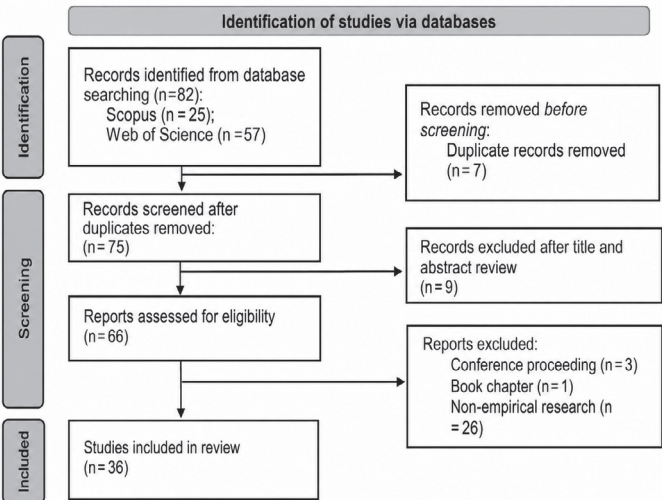
5.2. Đối tượng nghiên cứu chủ yếu và các kết quả thực nghiệm nổi bật về hiệu quả sử dụng AWE trong 10 năm qua

Bảng 2 trình bày số lượng nghiên cứu theo từng năm, đồng thời tổng hợp thông tin về công cụ AWE, bối cảnh triển khai, quy mô mẫu và trình độ khách thể.

Bảng 2: Tổng hợp các nghiên cứu thực nghiệm về AWE giai đoạn 2016–2025

Năm-số lượng	Nơi thực hiện nghiên cứu	Công cụ AWE	Số lượng khách thể	Trình độ khách thể
2016 - 1	Đài Loan	Criterion	66 SV	Nhiều trình độ (A2-B2)
2017 - 1	Trung Quốc	Criterion	2 SV	Không rõ
2018 - 2	Mỹ Trung Quốc	Criterion PIGAI	82 SV 2 SV	Nhiều trình độ Không rõ
2019 - 1	Saudi Arabia	My Access	6 SV	Trung cấp
2020 - 3	Trung Quốc Mỹ Trung Quốc	PIGAI Grammarly PIGAI	11 giảng viên 17 SV 3 SV	- Không rõ Không rõ

Hình 1: Biểu đồ quy trình chọn lọc theo PRISMA 2020



2021 - 3	Trung Quốc Trung Quốc Mỹ	PIGAI Paperrater Grammarly	520 SV 161 SV 6 SV	Không rõ
2022 - 4	Trung Quốc Trung Quốc Afghanistan Trung Quốc	Write&Improve PIGAI Grammarly PIGAI	25 SV 65 SV 66 SV 4 SV	Trung cấp Trung cấp Trung cấp Trung cấp
2023 - 6	Mỹ Hàn Quốc Philippine Ai Cập Trung Quốc Iran	Grammarly Neural AWE Write&Improve New developed Writingmentor	15 giảng viên 4094 SV 103 SV 2 giáo viên 4 SV	- Trung cấp Trung cấp - Trung cấp
2024 - 6	Nhật Hàn Quốc Hàn Quốc Việt Nam Việt Nam Hàn Quốc	W-Writing ChatGPT ChatGPT Criterion Criterion Write&Improve	101 SV 465 học sinh cấp 3 50 bài viết 75 SV 2 SV 72 học sinh cấp 3	Trung cấp
9/2025 - 6	Trung Quốc Trung Quốc Trung Quốc Trung Quốc Thái Lan Không xác định	Learnalytics YoudaoWriting AWE-peer PIGAI Grammarly ChatGPT	4 SV 3 SV 68 SV 7 SV 27 SV 120 bài viết	Chuyên ngành kỹ thuật 3 trình độ - - - Sơ cấp

Thông tin từ bảng 2 cho thấy số lượng nghiên cứu về AWE tăng dần từ năm 2016 và gia tăng rõ rệt trong 5 năm gần đây. Xu hướng này có thể được lý giải bởi sự phát triển mạnh của công nghệ số và sự lan tỏa của trí tuệ nhân tạo trong nhiều lĩnh vực, trong đó có giáo dục ngôn ngữ. Đáng chú ý, chỉ trong 9 tháng đầu năm 2025 đã ghi nhận 6 nghiên cứu triển khai AWE trong lớp học. Các bối cảnh nghiên cứu tương đối đa dạng, bao gồm cả quốc gia phát triển (như Trung Quốc, Hoa Kỳ) và quốc gia đang phát triển, trong đó có Việt Nam. Những công cụ được sử dụng phổ biến gồm Criterion, Pigai, Write&Improve và Grammarly, phù hợp với nhận định của Ding và Zou (2024) rằng dù có nhiều công cụ AWE, một số nền tảng vẫn nổi trội nhờ tính năng và mức độ phù hợp với người dùng. Về khách thể, phần lớn nghiên cứu tập trung vào SV bậc Đại học với trình độ trung cấp, tương đồng với tổng quan của Barrot (2023) về xu hướng ưu tiên nhóm người học này. Tuy nhiên, tổng quan hiện tại cũng ghi nhận AWE đã được áp dụng cho một số chuyên ngành ngoài ngôn ngữ, chẳng hạn khối kỹ thuật, cho thấy tiềm năng mở rộng của AWE sang các lĩnh vực đào tạo khác.

5.3. Vai trò của AWE trong học viết ngoại ngữ

5.3.1. Lợi ích của AWE

Các nghiên cứu cho thấy nhờ tính tức thời trong cung cấp phản hồi, AWE góp phần giảm tải khối lượng chấm bài cho GV, qua đó giúp GV tập trung nhiều hơn vào các vấn đề bậc cao (HOC) như nội dung, bố cục và phát triển ý tưởng (S. Koltovskaia, 2023). Đồng thời, AWE có khả năng cung cấp phản hồi tức thì, mang tính cá nhân hóa và có thể truy cập mọi lúc, mọi nơi (L.Xiaosa and K.Ping, 2025); (J. Woodworth and K. Barkaoui). Rõ ràng, tính tức thời này giải quyết một thách thức lâu dài trong việc giảng dạy viết: tính chất tốn thời gian của việc cung cấp phản hồi chi tiết. GV, thường bị hạn chế bởi khối lượng công việc nên hiếm khi có thể cung cấp phản hồi thường xuyên và chi tiết cho các bản thảo bài viết của người học, nhưng AWE làm cho việc này trở nên khả thi và có thể phát triển. Phản hồi từ AWE thường được mô tả là nhất quán, tương đối khách quan và kèm theo giải thích chi tiết (J.S. Barrot, 2023); D. Gilbert and G. Jason, 2023. Barrot (2023) nhấn mạnh rằng AWE có thể cải thiện độ chính xác ngôn ngữ và giảm lỗi, đặc biệt ở cấp độ câu (ngữ pháp, lựa chọn từ, cấu trúc câu), đồng thời thúc đẩy tính tự chủ và học tập tự định hướng của người học (D. Gilbert and G. Jason, 2023). Những tố chất này phù hợp với sự nhấn mạnh của các phương pháp sư phạm đương đại vào việc lấy người học làm trung tâm, trao quyền cho SV chủ động

trong quá trình phát triển kỹ năng viết của mình. Nhờ phản hồi nhanh từ AWE, người học có xu hướng viết và chỉnh sửa thường xuyên hơn mà không quá phụ thuộc vào người dạy, qua đó tăng mức độ tham gia và động lực trong quá trình học viết (K. Lim, J. Song, and J. Park).

Ngoài ra, AWE còn hỗ trợ mở rộng vốn từ vựng thông qua gợi ý từ đồng nghĩa, trái nghĩa và các lựa chọn diễn đạt thay thế (J.S. Barrot (2023)., đồng thời nội dung này cũng được Barrot (2023) ghi nhận trong tổng quan của mình. Bên cạnh đó, một số nghiên cứu ghi nhận AWE có thể góp phần giảm lo âu ngoại ngữ và sự e ngại khi viết, nhất là đối với những người học có mức lo lắng cao; xu hướng này phù hợp với các tổng quan của Fan và Ma (2022) cũng như Fu và cộng sự (2022), vì người học thường cảm thấy ít bị áp lực hơn khi nhận phản hồi từ một cỗ máy khách quan thay vì sự phán xét trực tiếp từ con người, điều này giúp tạo ra một không gian học tập an toàn về mặt tâm lý. Sự phát triển này cho thấy một tương lai nơi AWE có thể mở rộng tiện ích vượt ra ngoài việc sửa lỗi kỹ thuật sang hỗ trợ về mặt tâm lý học ngoại ngữ, có khả năng định hình lại bối cảnh của phương pháp giảng dạy viết.

5.3.2. Thách thức khi sử dụng AWE

Các nghiên cứu cho rằng hệ thống AWE vẫn có thể mắc lỗi, bao gồm bỏ sót lỗi hoặc đưa ra phản hồi chưa chính xác, từ đó làm giảm mức độ tin tưởng của người học và dẫn đến những chỉnh sửa kém phù hợp (G. T. L. Hoang and N. Storch, 2024). Bên cạnh đó, điểm số do AWE tạo ra đôi khi không nhất quán hoặc khác với đánh giá của GV, và AWE thường thiếu khả năng đưa ra phản hồi có chiều sâu về nội dung (L.Xiaosa and K.Ping, 2025). Nhìn chung, AWE tập trung chủ yếu vào các đặc điểm bề mặt như ngữ pháp, chính tả và cấu trúc câu, trong khi ít hiệu quả đối với các yếu tố bậc cao như tính mạch lạc, lập luận, tính thuyết phục và sáng tạo (L.Xiaosa and K.Ping, 2025); (S. Koltovskaia, 2023); nhận định này phù hợp với tổng quan của Fan và Ma (2022), khẳng định vai trò hỗ trợ của AWE trong phản hồi sửa lỗi. Việc thiếu sót về cấp độ LOC phản ánh một số hạn chế cơ bản của hầu hết các hệ thống AWE, vốn dựa vào thuật toán nhận dạng mẫu và thiếu sự nhạy bén về ngữ dụng cần thiết cho việc đánh giá ở cấp độ diễn ngôn. Kết quả khẳng định vai trò không thể thay thế của giáo viên trong việc giải quyết các khía cạnh tổng quát của việc viết vượt quá khả năng phân tích của các công cụ AWE hiện nay.

Ngoài hạn chế về năng lực phản hồi, nhiều nghiên cứu thực nghiệm có thời gian can thiệp ngắn, vì vậy bằng chứng về khả năng chuyển giao kiến thức (áp dụng kỹ năng vào các nhiệm vụ/hoàn cảnh viết khác nhau) còn hạn chế và chưa được làm rõ. Barrot (2023) cũng lưu ý rằng người học có thể đánh giá phản hồi của AWE là chung chung, lặp lại, mơ hồ hoặc thiếu rõ ràng; đặc biệt, người học trình độ thấp có thể thiếu nền tảng để thẩm định mức độ phù hợp của gợi ý sửa lỗi hoặc đề xuất từ hệ thống (J. S. Barrot (2023). Kết quả này tương đồng với các nhận định trong tổng quan của Fan và Ma (2022), xác nhận rằng việc sử dụng AWE trong giáo dục cần có sự cân nhắc về trình độ của người học. Sự cân nhắc này giữ vai trò quan trọng trong việc giúp người học phát triển kỹ năng tự điều chỉnh và tự chủ trong học tập, hoặc sẽ làm người học hoàn toàn phụ thuộc vào AWE.

6. Kết luận

Bài viết tổng hợp bằng chứng thực nghiệm về việc sử dụng các công cụ phản hồi sửa lỗi tự động (AWE) trong dạy và học viết ngoại ngữ. Kết quả cho thấy AWE hỗ trợ rõ rệt việc cải thiện các khía cạnh bậc thấp của năng lực viết, đặc biệt là độ chính xác từ vựng và ngữ pháp, đồng thời góp phần thúc đẩy tính tự chủ của người học khi được tích hợp một cách phù hợp vào thực tiễn sư phạm. Phần lớn người học trong các nghiên cứu được tổng hợp có phản hồi tích cực đối với việc sử dụng AWE. Tuy nhiên, AWE vẫn còn hạn chế trong việc hỗ trợ các kỹ năng bậc cao như phát triển ý tưởng, tổ chức bài viết và cung cấp phản hồi theo ngữ cảnh, cho thấy nhu cầu áp dụng các chiến lược dạy học kết hợp giữa phản hồi của người dạy và công cụ tự động. Bên cạnh đó, sự khác biệt về thiết kế nghiên cứu, mức độ minh bạch trong báo cáo và thời lượng can thiệp ngắn ở nhiều nghiên cứu nhấn mạnh yêu cầu nâng cao tính nghiêm ngặt và khả năng lặp lại trong các nghiên cứu tương lai. Trong bối cảnh công nghệ AWE tiếp tục phát triển, các hướng nghiên cứu tiếp theo nên tập trung vào tác động dài hạn, các vấn đề đạo đức, cũng như tiềm năng của AWE trong hỗ trợ phát triển kỹ năng viết siêu nhận thức ở nhiều nhóm người học, bao gồm học sinh tiểu học và trung học.

Kết quả tổng quan nghiên cứu cũng đề xuất một số ứng dụng thực tiễn trong môi trường giáo dục ngày

nay. Thứ nhất, lãnh đạo nhà trường (các cấp) cần hỗ trợ cơ sở vật chất và kết nối mạng ổn định, tạo điều kiện cho người dạy và người học sử dụng các công cụ AWE trong học tập và giảng dạy. Thứ hai, nhà trường cần tổ chức các khóa đào tạo kỹ năng công nghệ số cho người dạy và người học trước khi sử dụng AWE. Thứ ba, người dạy có thể tích hợp AWE vào giảng dạy viết ngoại ngữ, cũng như các chuyên ngành khác như công nghệ kỹ thuật, để AWE hỗ trợ sửa lỗi từ vựng và ngữ pháp của người học trước, và dành nhiều thời gian hơn cho các vấn đề ngôn ngữ cao hơn, như nội dung và sự liên kết các ý tưởng trong bài viết. AWE có thể hỗ trợ làm giảm khối lượng công việc chấm bài của GV và giảng viên, đặc biệt là trong các lớp học đông học sinh SV. Cuối cùng, hiệu quả của AWE phụ thuộc rất nhiều vào cách thức triển khai. Các nhà giáo dục nên xem xét các yếu tố như trình độ của học sinh SV, mục tiêu chương trình giảng dạy và thời điểm giảng dạy, để tối ưu hóa lợi ích của công cụ và giảm thiểu tối đa sự phụ thuộc vào công cụ của người dạy cũng như người học. Việc áp dụng AWE phù hợp trình độ sẽ hỗ trợ người học, cụ thể là SV, phát triển được tính tự chủ trong học viết ngoại ngữ nói riêng, và học tập toàn diện.

Bên cạnh những đóng góp nêu trên, bài tổng quan này vẫn tồn tại một số hạn chế. Thứ nhất, chiến lược tìm kiếm chỉ giới hạn ở các ấn phẩm tiếng Anh, do đó có thể bỏ sót các nghiên cứu liên quan được công bố bằng ngôn ngữ khác. Thứ hai, tiêu chí lựa chọn tập trung vào các nghiên cứu thực nghiệm đã bình duyệt, nên có khả năng loại trừ những nguồn dữ liệu có giá trị như luận văn, báo cáo thực hành hoặc tài liệu không chính thức. Cuối cùng, thiên lệch xuất bản có thể ảnh hưởng đến kết quả tổng hợp, vì các nghiên cứu báo cáo hiệu quả tích cực thường có xu hướng được công bố nhiều hơn so với các nghiên cứu có kết quả không đáng kể hoặc tiêu cực. Do đó, các nghiên cứu trong tương lai cần cân nhắc các hạn chế này để phản ánh đầy đủ hơn bức tranh bằng chứng và tối ưu hóa tiềm năng ứng dụng AWE trong giảng dạy ngôn ngữ.

Tài liệu tham khảo

- Bae, Y., & Jung, Y. (2024). The effect of feedback types on learner engagement and L2 writing development. *English Teaching*, 79(3), 159–180. <https://doi.org/10.15858/engtea.79.3.202409.159>.
- Barrot, J. S. (2023). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 36(4), 584–607. <https://doi.org/10.1080/09588221.2021.1936071>.
- Barrot, J. S. (2024). Trends in automated writing evaluation systems research for teaching, learning, and assessment: A bibliometric analysis. *Education and Information Technologies*, 29(6), 7155–7179. <https://doi.org/10.1007/s10639-023-12083-y>.
- Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies*, 29(11), 14151–14203. <https://doi.org/10.1007/s10639-023-12402-3>.
- Dizon, G., & Gold, J. (2023). Exploring the effects of Grammarly on EFL students' foreign language anxiety and learner autonomy. *The JALT CALL Journal*, 19(3), 299–316. <https://doi.org/10.29140/jaltcall.v19n3.1049>.
- Fan, N., & Ma, Y. (2022). The effects of automated writing evaluation (AWE) feedback on students' English writing quality: A systematic literature review. *Language Teaching Research Quarterly*, 28, 53–73. <https://doi.org/10.32038/ltrq.2022.28.03>.
- Fu, Q.-K., Zou, D., Xie, H., & Cheng, G. (2022). A review of AWE feedback: Types, learning outcomes, and implications. *Computer Assisted Language Learning*, 37(1–2), 179–221. <https://doi.org/10.1080/09588221.2022.2033787>.
- Hoang, G. T. L., & Storch, N. (2024). How do EFL learners process and uptake Criterion automated corrective feedback? Insights from two case studies. *Iranian Journal of Language Teaching Research*, 12(1), 1–20. <https://doi.org/10.30466/ijltr.2024.121420>.
- Karatay, Y., & Karatay, L. (2024). Automated writing evaluation use in second language classrooms: A research synthesis. *System*, 123, Article 103332. <https://doi.org/10.1016/j.system.2024.103332>.
- Koltovskaia, S. (2023). Postsecondary L2 writing teachers' use and perceptions of Grammarly as a complement to their feedback. *ReCALL*, 35(3), 290–304. <https://doi.org/10.1017/S0958344022000179>.
- Li, X., & Ke, P. (2025). How L2 student writers engage with automated feedback: A longitudinal perspective. *Assessing Writing*, 64, Article 100919. <https://doi.org/10.1016/j.asw.2025.100919>.
- Lim, K. T., Song, J., & Park, J. (2023). Neural automated writing evaluation for Korean L2 writing. *Natural Language Engineering*, 29(5), 1341–1363. <https://doi.org/10.1017/S1351324922000298>.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>.
- Woodworth, J., & Barkaoui, K. (2020). Perspectives on using automated writing evaluation systems to provide written corrective feedback in the ESL classroom. *TESL Canada Journal*, 37(2), 234–247. <https://doi.org/10.18806/tesl.v37i2.1340>.
- Zhai, N., & Ma, X. (2022). Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. *Computer Assisted Language Learning*, 35(9), 2817–2842. <https://doi.org/10.1080/09588221.2021.1897019>.