

PHẢN HỒI VIẾT TỰ ĐỘNG (AWE) TRONG GIẢNG DẠY KỸ NĂNG VIẾT TIẾNG ANH CHO SINH VIÊN KHÔNG CHUYÊN NGÔN NGỮ: TỔNG QUAN LÝ THUYẾT

NGUYỄN THỊ THỦY

Khoa Ngoại ngữ, Đại học Bách khoa Hà Nội

Nhận bài ngày 10/4/2026. Sửa chữa xong 20/5/2026. Duyệt đăng 27/5/2026.

Abstract

Automated Writing Evaluation (AWE) is an emerging technology being increasingly applied in EFL writing instruction for non-English-major students. This paper reviews the theoretical foundations of AWE application in writing instruction, including effective feedback models, types of corrective feedback, several learner- and technology-related theories, and prompt engineering principles for designing LLM-based AWE tools. On this basis, this paper offers pedagogical suggestions for implementing AWE that suit A2-level students in Vietnamese higher education.

Keywords: Automated writing evaluation, AWE, English writing skills, non-English-major students, prompt engineering.

1. Đặt vấn đề

Đối với người học ngôn ngữ nước ngoài, viết là một trong những kỹ năng phức tạp và khó tiếp thu nhất trong bốn kỹ năng nghe, nói, đọc, viết, đặc biệt là trong bối cảnh học thuật. Khó khăn này càng trở nên rõ rệt hơn đối với sinh viên (SV) không chuyên ngôn ngữ tại các trường đại học Việt Nam. Việc sử dụng tiếng Anh trong các bài viết đòi hỏi nhiều nỗ lực về kiến thức ngữ pháp và từ vựng, trong khi SV không chuyên thường có nền tảng còn hạn chế, chịu ảnh hưởng của tiếng mẹ đẻ và có định hướng học tập thiên về các bài thi chuẩn hóa. Điều này khiến việc thực hành viết trên lớp ít được ưu tiên rèn luyện.

Phản hồi của giáo viên (GV) là một trong những yếu tố quan trọng nhất giúp nâng cao kỹ năng viết của SV. Tuy nhiên, với các lớp học đông và thời lượng giảng dạy hạn chế, việc đưa ra phản hồi cá nhân hóa sau mỗi bài viết gần như là điều bất khả thi đối với giảng viên đại học. Đây chính là khoảng trống mà các công cụ phản hồi viết tự động (Automated Writing Evaluation - AWE) đang dần lấp đầy. Các công cụ này không chỉ giảm tải công việc cho GV mà còn cung cấp phản hồi tức thì, có thể lặp lại nhiều lần - những đặc tính mà phản hồi truyền thống khó có thể đáp ứng được. Trong những năm gần đây, sự xuất hiện và phát triển nhanh chóng của các Mô hình Ngôn ngữ Lớn (Large Language Models - LLM) như GPT-4 hay Gemini đã tạo ra một làn sóng AWE mới với khả năng cung cấp phản hồi linh hoạt, chi tiết và sát ngữ cảnh hơn so với các thế hệ trước. Khác với các hệ thống thương mại đòi đầu vốn dựa trên các thuật toán cố định, AWE thế hệ mới có thể được tùy biến dễ dàng thông qua việc thiết kế lại prompt - hướng dẫn cho AI. Điều này mở ra khả năng cá nhân hóa công cụ theo từng bối cảnh giảng dạy, từng nhóm đối tượng người học và từng yêu cầu cụ thể của khóa học. Mặc dù nghiên cứu về AWE trên thế giới đã phát triển mạnh trong hai thập kỷ qua, các công trình tập trung vào SV không chuyên ngôn ngữ tại Việt Nam, đặc biệt ở trình độ A2 vẫn còn khá ít. Phần lớn các nghiên cứu hiện có hoặc khảo sát người học ngành Ngôn ngữ với yêu cầu viết học thuật cao hoặc thực hiện trong bối cảnh giáo dục nước ngoài có đặc thù khác biệt so với Việt Nam. Trong bối cảnh

Email: thuy.nguyenthi3@hust.edu.vn

đó, việc tổng hợp các nền tảng lý thuyết liên quan đến AWE và xem xét chúng dưới góc độ giảng dạy cho SV không chuyên là một hướng nghiên cứu cần thiết. Bài viết này tổng hợp và phân tích các nền tảng lý thuyết chính cho việc thiết kế và triển khai AWE trong giảng dạy viết tiếng Anh cho SV không chuyên, đồng thời đưa ra một số gợi ý sơ phạm để tham khảo trong thực tiễn giảng dạy.

2. Nội dung nghiên cứu

2.1. Đặc điểm kỹ năng viết của sinh viên không chuyên và vai trò của phản hồi

Hyland (2003) mô tả viết là một quá trình nhận thức phức tạp, đòi hỏi người học phải đồng thời quản lý nhiều cấp độ của ngôn ngữ - từ vựng, ngữ pháp, cấu trúc câu cho đến cách tổ chức ý và quy ước thể loại. Đối với người học Tiếng Anh như một ngoại ngữ, gánh nặng này càng lớn hơn vì họ phải làm việc bằng một ngôn ngữ chưa thực sự nội hóa. Trong khi người bản ngữ có thể vận hành phần lớn các quy tắc ngôn ngữ ở mức tự động, người học EFL phải xử lý chúng một cách có ý thức và điều này tiêu tốn rất nhiều dung lượng nhận thức trong quá trình viết.

Trong bối cảnh Việt Nam, một số nghiên cứu đã ghi nhận những đặc điểm khá phổ biến trong bài viết của SV không chuyên. Le (2023) qua phân tích 246 bài viết của SV trình độ tiền trung cấp đã chỉ ra rằng các lỗi về thì, sự hòa hợp giữa chủ ngữ và động từ, mạo từ và giới từ là những lỗi xuất hiện thường xuyên nhất. Đa phần các lỗi này có thể được giải thích bởi sự khác biệt về cấu trúc giữa tiếng Việt và tiếng Anh cùng với việc người học chưa có đủ thời gian luyện tập để chuyển hóa kiến thức ngữ pháp thành kỹ năng sử dụng tự động. Cụ thể, tiếng Việt không biến đổi hình thái động từ theo thì không yêu cầu sự hòa hợp giữa chủ ngữ và động từ, không có hệ thống mạo từ tương đương với tiếng Anh và sử dụng giới từ theo cách khác biệt rõ rệt. Do vậy, khi viết tiếng Anh, SV Việt Nam có xu hướng bỏ sót hoặc dùng sai các yếu tố này.

Phản hồi kịp thời từ GV được công nhận là một yếu tố then chốt cho sự tiến bộ của SV trong kỹ năng viết. Nguyen và cộng sự (2021) khảo sát 97 SV đại học Việt Nam và cho biết phần đông SV đánh giá phản hồi viết của GV là không thể thiếu cho sự phát triển kỹ năng này. Nghiên cứu cũng cho thấy SV đặc biệt đánh giá cao những phản hồi chi tiết, có giải thích lý do của lỗi và gợi ý cách sửa hơn là phản hồi chỉ đánh dấu vị trí lỗi hoặc cho điểm chung chung. Tuy nhiên, với lớp học 30-40 SV và lịch giảng 2-3 tiết một tuần, GV rất khó có thể chữa bài chi tiết cho từng người sau mỗi bài viết. Thực tế thường thấy là GV chỉ chữa chung trên lớp một vài lỗi tiêu biểu hoặc chữa kỹ một số bài sau đó chia sẻ làm ví dụ cho cả lớp. Cách làm này tuy có giá trị nhưng không thay thế được phản hồi cá nhân hóa - thứ mà SV thực sự cần nếu muốn tiến bộ rõ rệt.

Khoảng cách giữa nhu cầu được nhận phản hồi và thực tế lớp học chính là điểm xuất phát cho sự cần thiết của các công cụ phản hồi tự động. AWE không nhằm thay thế GV mà bổ sung một kênh phản hồi cá nhân hóa, kịp thời và có thể lặp lại nhiều lần. Trong một mô hình lý tưởng, AWE đảm nhận phần phản hồi về ngôn ngữ ở các vòng bản nháp đầu tiên, giải phóng thời gian của GV để tập trung vào những vấn đề phức tạp hơn như nội dung, cấu trúc bài viết và giọng văn của SV.

2.2. Lý thuyết về phản hồi hiệu quả

Một trong những khung lý thuyết được trích dẫn nhiều nhất khi nói về phản hồi trong giáo dục là mô hình của Hattie và Timperley (2007). Trong công trình nổi tiếng *"The Power of Feedback"*, hai nhà nghiên cứu đã đề xuất một mô hình phản hồi hiệu quả dựa trên ba câu hỏi mà người học cần được trả lời trong quá trình học tập. Câu hỏi đầu tiên là Feed Up - "Tôi đang đi đến đâu?" đòi hỏi phản hồi phải làm rõ mục tiêu và tiêu chí đánh giá để người học hiểu một bài viết tốt cần đạt được những gì. Câu hỏi thứ hai là Feed Back - "Tôi đang ở đâu?" - yêu cầu phản hồi chỉ ra khoảng cách giữa thực trạng bài viết và tiêu chí kỳ vọng, kèm theo minh chứng cụ thể từ bài viết của người học. Câu hỏi thứ ba là Feed Forward - "Bước tiếp theo là gì?" đòi hỏi phản hồi phải đưa ra định hướng cải thiện cụ thể, giúp người học biết phải làm gì tiếp theo.

Hai nhà nghiên cứu nhấn mạnh rằng phản hồi chỉ thực sự hiệu quả khi đáp ứng được cả ba tầng

nói trên. Phản hồi chỉ liệt kê lỗi mà thiếu định hướng cải thiện - một dạng khá phổ biến trong thực tiễn giảng dạy được chứng minh là có tác động hạn chế đến tiến bộ dài hạn của người học. Khi người học chỉ biết mình sai ở đâu mà không biết cách sửa và không hiểu lý do sửa, họ có xu hướng lặp lại lỗi tương tự trong các bài viết tiếp theo. Mô hình này cũng cung cấp một khung sư phạm rõ ràng để đánh giá chất lượng của AWE: một công cụ AWE tốt cần làm được nhiều hơn là chỉ ra cái sai, nó cần kết nối được với tiêu chí bài viết và gợi mở hướng sửa cụ thể.

Trong dạy viết cho người học ngoại ngữ, một tranh luận kéo dài nhiều thập kỷ là về hiệu quả tương đối của hai hình thức phản hồi sửa lỗi. Phản hồi trực tiếp (Direct Corrective Feedback) là hình thức trong đó GV vừa chỉ ra lỗi vừa cung cấp dạng đúng cho người học. Hình thức này giúp SV sửa bài ngay và giảm tải nhận thức, đặc biệt phù hợp với người học trình độ thấp (Ellis và cộng sự, 2008). Phản hồi gián tiếp (Indirect Corrective Feedback) chỉ đánh dấu vị trí lỗi mà không cho biết dạng đúng để người học tự suy nghĩ và tìm cách sửa. Ferris (2002) cho rằng phản hồi gián tiếp thúc đẩy tư duy phản biện và mang lại hiệu quả tiếp thu ngôn ngữ tốt hơn về mặt lâu dài vì nó duy trì vai trò chủ động của người học trong quá trình xử lý lỗi. Theo lập luận này, khi người học phải tự nghĩ ra cách sửa, họ buộc phải kích hoạt và củng cố kiến thức ngữ pháp đã học, từ đó khả năng ghi nhớ và áp dụng sẽ tốt hơn. Tuy nhiên, với người học ở trình độ thấp, việc tự sửa lỗi thường gặp khó khăn do chưa có đủ kiến thức ngữ pháp để nhận ra cách sửa đúng. Khi đó, phản hồi gián tiếp có thể trở thành nguồn gây thất vọng và nản chí, dẫn đến hiệu ứng ngược lại với mục đích ban đầu. Một hướng giải quyết hợp lý là sử dụng phản hồi trực tiếp làm hình thức chính cho SV A2, đồng thời bổ sung một cơ chế bắt buộc người học ghi lại và tự diễn giải lỗi - chẳng hạn dưới dạng nhật ký sửa lỗi để hạn chế sự thụ động. Cách kết hợp này tận dụng được lợi thế giảm tải nhận thức của phản hồi trực tiếp, đồng thời vẫn duy trì sự tham gia chủ động của người học vào quá trình xử lý lỗi.

2.3. Công cụ phản hồi viết tự động (AWE) trong dạy kỹ năng viết

AWE là các hệ thống ứng dụng công nghệ xử lý ngôn ngữ tự nhiên để phân tích bài viết và cung cấp phản hồi tức thì cho người học mà không cần sự can thiệp của GV (Link và cộng sự, 2022). Lịch sử phát triển của AWE có thể được chia thành hai thế hệ rõ rệt. Thế hệ đầu tiên gồm các hệ thống thương mại như Criterion của ETS, Revision Assistant của Turnitin hay Grammarly - vốn dựa trên các thuật toán thống kê và mô hình hóa đặc trưng ngôn ngữ. Các hệ thống này có ưu điểm là độ ổn định cao và dễ tích hợp vào quy trình giảng dạy nhưng phản hồi thường mang tính khuôn mẫu, ít có khả năng giải thích lý do của lỗi và khó tùy biến theo bối cảnh giảng dạy cụ thể. Thế hệ AWE thứ hai dựa trên các Mô hình Ngôn ngữ Lớn hiện đại như GPT-4, Gemini hay Claude. Điểm khác biệt lớn nhất giữa hai thế hệ nằm ở khả năng tạo ra phản hồi tự nhiên, có thể giải thích lý do của lỗi và đề xuất cách sửa cụ thể theo từng câu, từng đoạn của người học điều mà các hệ thống cũ không làm được. Bên cạnh đó, AWE dựa trên LLM còn cho phép GV tùy biến công cụ thông qua việc thiết kế prompt, nghĩa là có thể điều chỉnh giọng điệu phản hồi, ngôn ngữ giao tiếp, danh mục lỗi ưu tiên và cấu trúc output theo nhu cầu cụ thể của khóa học. Đây là một bước tiến quan trọng vì nó chuyển AWE từ một sản phẩm đóng gói sẵn thành một công cụ có thể được GV tự thiết kế và điều chỉnh.

Nhiều nghiên cứu thực nghiệm đã ghi nhận tác động tích cực của AWE đối với kỹ năng viết của người học ngoại ngữ. Wang (2022) trong một nghiên cứu thực nghiệm có nhóm đối chứng với người học EFL tại Trung Quốc đã xác nhận rằng nhóm sử dụng AWE có cải thiện đáng kể về độ chính xác ngữ pháp và sự phong phú từ vựng so với nhóm đối chứng. Đáng chú ý, hiệu quả của AWE thể hiện rõ nhất ở khía cạnh độ chính xác ngôn ngữ đúng với mong đợi vì đây cũng là khía cạnh mà công cụ tự động xử lý tốt nhất. Bên cạnh tác động đến độ chính xác ngôn ngữ, AWE còn có vai trò hỗ trợ phát triển khả năng tự điều chỉnh học tập của người học. Khi nhận được phản hồi tức thì sau mỗi bài viết, người học có cơ hội rút ngắn chu kỳ tự đánh giá và sửa lỗi, từ đó hình thành thói quen kiểm tra bài viết của bản thân một cách có hệ thống hơn. Đây là một lợi ích vượt ra ngoài bản thân kỹ năng viết, có thể chuyển hóa sang các kỹ năng học tập độc lập khác. Bên cạnh các hiệu quả tích cực, AWE cũng có một số hạn chế cần được lưu ý. Trước hết là nguy cơ người học lệ thuộc quá mức vào công cụ, dẫn

đến giảm khả năng tự sửa lỗi độc lập (Link và cộng sự, 2022). Khi mọi lỗi đều có thể được công cụ phát hiện và sửa, người học có thể không còn động lực để tự đọc lại bài và phát triển ý thức về ngôn ngữ. Trong dài hạn, điều này có thể dẫn đến tình trạng người học chỉ biết viết khi có AWE bên cạnh nhưng khả năng viết độc lập lại không tiến bộ tương xứng. Một hạn chế khác là AWE chủ yếu mạnh về các lỗi bề mặt như ngữ pháp, chính tả và lỗi cơ học, trong khi các yếu tố cấp cao hơn như tính mạch lạc, tổ chức ý hay giọng văn vẫn là thách thức ngay cả với các LLM hiện đại. Điều này có nghĩa là AWE phù hợp nhất với các vòng bản nháp đầu tiên - khi mục tiêu chính là sửa lỗi ngôn ngữ và không thể thay thế hoàn toàn cho phản hồi của GV về nội dung và cấu trúc bài viết. Ngoài ra, phần lớn các nghiên cứu hiện có tập trung vào SV ngành Ngôn ngữ hoặc môi trường học thuật có yêu cầu viết cao. Hiệu quả của AWE đối với SV không chuyên trình độ thấp, đặc biệt trong bối cảnh Việt Nam vẫn còn cần thêm nhiều nghiên cứu thực nghiệm để khẳng định.

2.4. Một số lý thuyết về người học làm nền tảng cho thiết kế AWE

Bên cạnh các lý thuyết về phản hồi, việc thiết kế một công cụ AWE phù hợp với người học còn cần dựa trên các lý thuyết về nhận thức và cảm xúc của người học ngôn ngữ. Krashen (1982) trong Giả thuyết Đầu vào (Input Hypothesis) cho rằng người học tiếp thu ngôn ngữ hiệu quả nhất khi tiếp nhận đầu vào ở mức $i+1$, nghĩa là cao hơn trình độ hiện tại một chút đủ để tạo thử thách nhưng không đến mức gây nản chí. Áp dụng vào thiết kế AWE, điều này có nghĩa là phản hồi nên được viết bằng ngôn ngữ đơn giản, tránh thuật ngữ ngôn ngữ học phức tạp và số lượng lỗi báo cáo trong mỗi lần phản hồi cũng cần được giới hạn để không vượt quá khả năng xử lý của người học. Nếu phản hồi quá khó hiểu hoặc quá nhiều lỗi cùng lúc, người học sẽ không thể chuyển hóa được thông tin trong phản hồi thành kiến thức ngôn ngữ thực sự. Cũng theo Krashen, Giả thuyết Bộ lọc Cảm xúc (Affective Filter Hypothesis) chỉ ra rằng các yếu tố cảm xúc như lo lắng, thiếu tự tin có thể tạo ra một “bộ lọc” ngăn cản quá trình tiếp thu ngôn ngữ. Phản hồi mang tính chỉ trích nặng hoặc trình bày quá nhiều lỗi cùng lúc rất dễ kích hoạt bộ lọc này, đặc biệt với người học trình độ thấp - những người vốn đã có mức lo âu khi viết khá cao. Do đó, một bản phản hồi AWE tốt nên có một phần tóm tắt mang tính khích lệ ở đầu, ưu tiên hiển thị các lỗi quan trọng nhất và sử dụng giao diện trực quan, không gây áp lực. Một số công cụ AWE hiện đại còn tích hợp phần ghi nhận tiến bộ của người học - chẳng hạn so sánh số lỗi giữa bản nháp đầu và bản nháp sau nhằm tạo cảm giác thành tựu và duy trì động lực học. Sweller (1988) trong Lý thuyết Tải nhận thức (Cognitive Load Theory) cho rằng bộ nhớ làm việc của người học có dung lượng hạn chế. Khi lượng thông tin cần xử lý vượt quá dung lượng này, hiệu quả học tập sẽ giảm sút rõ rệt. Lý thuyết này phân biệt ba loại tải nhận thức: tải nội tại (intrinsic) liên quan đến độ phức tạp của nội dung học, tải ngoại lai (extraneous) phát sinh từ cách trình bày không hiệu quả và tải hữu ích (germane) phục vụ cho quá trình xây dựng kiến thức. Mục tiêu thiết kế tốt là giảm tải ngoại lai và tăng tải hữu ích, trong khi tải nội tại phụ thuộc vào bản chất của nội dung học. Trong bối cảnh AWE, một bản phản hồi quá dài, quá nhiều lỗi hoặc dùng ngôn ngữ phức tạp sẽ tạo ra tải nhận thức không phục vụ cho việc học.

Để giảm tải nhận thức cho người học, một số nguyên tắc thiết kế có thể được áp dụng. Trước hết, cần giới hạn số lượng lỗi hiển thị trong mỗi lần phản hồi, chẳng hạn tối đa 8 lỗi cho lần phân tích bản nháp đầu tiên và 5 lỗi cho lần so sánh hai bản nháp. Ngoài ra, có thể sử dụng hệ thống mã lỗi nhất quán như TENSE, SVA, ART, PREP, WORDFORM, SPELL, CAP, PUNC thay vì giải thích dài dòng cho từng lỗi. Cuối cùng, output cần được trình bày theo cấu trúc trực quan, dễ đọc, có sử dụng các dấu hiệu phân biệt như màu sắc, biểu tượng và phân loại theo mức độ ưu tiên của lỗi.

Khi xem xét AWE từ góc độ một sản phẩm công nghệ, Davis (1989) đề xuất Mô hình Chấp nhận Công nghệ (Technology Acceptance Model - TAM) cung cấp một khung lý thuyết hữu ích để đánh giá trải nghiệm người dùng. Theo TAM, hai yếu tố cốt lõi quyết định mức độ chấp nhận và sử dụng công nghệ của người dùng là Perceived Usefulness (nhận thức về tính hữu ích) và Perceived Ease of Use (nhận thức về tính dễ sử dụng). Hai yếu tố này ảnh hưởng đến thái độ và ý định sử dụng công nghệ của người dùng, từ đó quyết định hành vi sử dụng thực tế. Đối với việc đánh giá AWE, các câu

hỏi khảo sát thường được thiết kế xoay quanh việc SV có nhận thấy công cụ giúp phát hiện lỗi và cải thiện bài viết hay không, công cụ có dễ sử dụng và phản hồi có dễ hiểu hay không, cùng với thái độ chung và ý định tiếp tục sử dụng trong tương lai. Việc đo lường các yếu tố này đặc biệt quan trọng với SV không chuyên - những người có thể chưa quen với việc tự sử dụng công cụ AI vào học tập. Nếu công cụ được thiết kế tốt về mặt sư phạm nhưng người dùng không cảm thấy hữu ích hoặc dễ sử dụng, hiệu quả thực tế sẽ rất hạn chế.

2.5. Prompt Engineering trong thiết kế AWE dựa trên Mô hình Ngôn ngữ Lớn

Với sự phát triển của các Mô hình Ngôn ngữ Lớn, Prompt Engineering hay quá trình thiết kế có chủ đích các hướng dẫn để định hướng output của AI đã trở thành một kỹ năng sư phạm quan trọng (Mollick & Mollick, 2023). Khác với việc dùng các công cụ AWE thương mại có sẵn, Prompt Engineering cho phép giảng viên tự thiết kế phản hồi theo đặc điểm cụ thể của người học, tiêu chí bài viết và mục tiêu sư phạm của từng bài học. Đây là mức độ linh hoạt mà các công cụ AWE thế hệ cũ không cung cấp được. Đáng chú ý, Prompt Engineering không đòi hỏi giảng viên phải có nền tảng lập trình - bản chất của nó là viết hướng dẫn bằng ngôn ngữ tự nhiên cho AI, gần với việc soạn giáo án hơn là viết code. Mollick và Mollick (2023) đề xuất một số nguyên tắc thiết kế prompt trong giáo dục. Nguyên tắc đầu tiên là giao vai trò sư phạm rõ ràng cho AI, chẳng hạn “Bạn là GV dạy viết tiếng Anh cho SV đại học Việt Nam ở trình độ A2” để định hướng giọng điệu và mức độ phức tạp của phản hồi. Khi AI được giao một vai trò cụ thể, output sẽ nhất quán hơn và phù hợp với đối tượng người học hơn. Nguyên tắc thứ hai là giới hạn phạm vi nhiệm vụ, ví dụ “Chỉ nhận xét về ngữ pháp và lỗi cơ học, không nhận xét về nội dung hay ý tưởng” để đảm bảo phản hồi tập trung và không vượt ngoài tiêu chí đã định. Nguyên tắc thứ ba là quy định cấu trúc output cụ thể nhằm tạo ra phản hồi nhất quán, dễ đọc và dễ sử dụng trong nhật ký sửa lỗi của SV. Nguyên tắc thứ tư là đưa vào các ràng buộc để ngăn AI thực hiện những hành động sư phạm không mong muốn, ví dụ “Không viết lại toàn bộ câu của SV” nhằm duy trì vai trò chủ động của người học trong quá trình sửa lỗi.

Quy trình thiết kế prompt cho AWE thường gồm các bước: phân tích đặc điểm người học và mục tiêu khóa học, xác định danh mục lỗi ưu tiên, viết bản prompt sơ thảo, thử nghiệm với mẫu bài viết thực tế của SV, đánh giá chất lượng output theo các tiêu chí sư phạm đã định và tinh chỉnh prompt dựa trên kết quả đánh giá. Quy trình lặp này phù hợp với phương pháp Nghiên cứu thiết kế (Design-Based Research) trong nghiên cứu giáo dục, vốn nhấn mạnh sự lặp lại có chủ đích giữa thiết kế, triển khai, đánh giá và điều chỉnh. Tiêu chí đánh giá chất lượng prompt có thể bao gồm: phản hồi có đáp ứng được ba tầng Feed Up - Feed Back - Feed Forward hay không, có sử dụng ngôn ngữ phù hợp với trình độ người học hay không, có giữ được số lượng lỗi trong giới hạn cho phép hay không, và có duy trì được giọng điệu khích lệ hay không. Việc đánh giá nên có sự tham gia của nhiều giảng viên để đảm bảo tính khách quan.

3. Kết luận

Phản hồi viết tự động (AWE) là một hướng ứng dụng công nghệ có tiềm năng giải quyết một bài toán dai dẳng trong dạy viết tiếng Anh cho SV không chuyên: bài toán về phản hồi cá nhân hóa trong điều kiện lớp đông và thời lượng giảng dạy hạn chế. Các tài liệu được tổng hợp trong bài viết này cho thấy phản hồi hiệu quả cần đáp ứng đồng thời ba tầng Feed Up - Feed Back - Feed Forward; với người học trình độ thấp, phản hồi trực tiếp có lợi thế trong ngắn hạn nhưng cần được kết hợp với cơ chế nhật ký sửa lỗi để duy trì vai trò chủ động của người học; AWE đã được chứng minh có hiệu quả tích cực với độ chính xác ngữ pháp và phát triển tự điều chỉnh học tập nhưng cũng có những hạn chế về sự thụ động và giới hạn ở các lỗi bề mặt; đồng thời, thiết kế AWE cần tuân thủ các nguyên tắc về tải nhận thức và bộ lọc cảm xúc của người học. Trên cơ sở các tài liệu đã trình bày, một số định hướng có thể tham khảo cho việc triển khai AWE trong các trường đại học Việt Nam bao gồm: ưu tiên phản hồi trực tiếp cho SV A2 và kết hợp với nhật ký sửa lỗi bắt buộc; giới hạn số lượng lỗi hiển thị mỗi lần để kiểm soát tải nhận thức; thiết kế phản hồi theo cấu trúc Feed Up - Feed Back - Feed Forward; tích

hợp AWE trong mô hình phản hồi kết hợp - không thay thế GV; đánh giá trải nghiệm người dùng theo khung TAM; và sử dụng Prompt Engineering để tùy biến công cụ theo đặc thù khóa học. Các nghiên cứu thực nghiệm trong tương lai, đặc biệt với thiết kế tựa thực nghiệm, sẽ giúp lấp đầy các khoảng trống hiện có và cung cấp bằng chứng cụ thể cho hiệu quả của AWE trong bối cảnh đại học Việt Nam.

Tài liệu tham khảo

- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36(3), 353-371.
- Ferris, D. (2002). *Treatment of error in second language student writing*. University of Michigan Press.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Hyland, K. (2003). *Second language writing*. Cambridge University Press.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Le, T. T. D. (2023). Grammatical error analysis of EFL learners' English writing samples: The case of Vietnamese pre-intermediate students. *International Journal of TESOL & Education*, 3(4), 1-14.
- Link, S., Mehrzad, M., & Rahimi, M. (2022). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, 35(4), 605-634.
- Mollick, E. R., & Mollick, L. (2023). *Assigning AI: Seven approaches for students, with prompts*. The Wharton School Research Paper.
- Nguyen, N. L. T., Nguyen, B. T. T., & Hoang, G. T. L. (2021). Students' perceptions of teachers' written feedback on EFL writing in a Vietnamese tertiary context. *Language Related Research*, 12(5), 405-431.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.
- Wang, Z. (2022). Computer-assisted EFL writing and evaluations based on artificial intelligence: A case from a college reading and writing course. *Library Hi Tech*, 40(1), 80-97.