

ĐỌC MỞ RỘNG VỚI VĂN BẢN CHUYÊN NGÀNH ĐƯỢC TRÍ TUỆ NHÂN TẠO ĐIỀU CHỈNH: MỘT NGHIÊN CỨU HÀNH ĐỘNG VỀ PHẢN HỒI CỦA SINH VIÊN KHÔNG CHUYÊN TIẾNG ANH TẠI VIỆT NAM

Phạm Thu Hiền¹

Email: pthien@uneti.edu.vn. ORCID: 0009-0007-5859-2052

Ngày tòa soạn nhận được bài báo: 15/06/2026

Ngày phản biện đánh giá: 22/06/2026

Ngày bài báo được duyệt đăng: 14/09/2026

DOI: 10.59266/houjs.2026.1434

Tóm tắt: Nghiên cứu hành động lớp học này tìm hiểu phản hồi của sinh viên không chuyên tiếng Anh đối với việc học từ vựng chuyên ngành qua đọc mở rộng có hướng dẫn với văn bản được trí tuệ nhân tạo (AI) điều chỉnh và các yếu tố giải thích phản hồi đó. Đối tượng là 40 sinh viên tham gia hai chu kỳ can thiệp; dữ liệu thu thập đa nguồn (bảng hỏi, nhật ký đọc, quan sát, phỏng vấn). Tất cả tiêu chí tự báo cáo đều tăng ở chu kỳ 2, cao nhất ở sự tự tin, tính hữu ích và ý định tiếp tục đọc; ghi nhớ từ vựng tăng khiêm tốn hơn. Phỏng vấn cho thấy độ dễ tiếp cận văn bản và tính liên quan với nội dung học trên lớp là hai yếu tố chính lý giải phản hồi tích cực. Do thiết kế trong nhóm không đối chứng, cỡ mẫu nhỏ và không có bài kiểm tra từ vựng chuẩn hóa, các phát hiện chỉ nên được hiểu là xu hướng liên hệ trong một bối cảnh cụ thể. Đóng góp chính là chỉ ra sự khác biệt giữa trải nghiệm tích cực của sinh viên và mức độ ghi nhớ từ vựng chuyên ngành tự cảm nhận.

Từ khóa: đọc mở rộng có hướng dẫn, nghiên cứu hành động lớp học, từ vựng chuyên ngành, văn bản được AI điều chỉnh

I. Đặt vấn đề

Từ vựng chuyên ngành giúp người học hiểu các tài liệu thuộc lĩnh vực của mình, từ giáo trình đến tài liệu nghề nghiệp. Gablasova (2014) cho thấy người học có thể tiếp thu từ vựng chuyên ngành qua đọc văn bản giáo khoa, còn Lacková (2025) cho thấy kho ngữ liệu (corpus) giúp người học nhận diện và sử dụng từ

chuyên ngành trong nhiều ngữ cảnh. Tại Việt Nam, khó khăn về thuật ngữ chuyên ngành cũng được ghi nhận là rào cản ngôn ngữ nổi bật đối với sinh viên (Nguyễn, 2024), trong khi công cụ công nghệ đã cho thấy tiềm năng hỗ trợ sinh viên không chuyên ghi nhớ từ vựng tiếng Anh tốt hơn (Vũ, 2026).

Tuy nhiên, người học trình độ thấp thường khó tiếp cận văn bản chuyên ngành

¹ Trường Đại học Kinh tế - Kỹ thuật Công nghiệp, Hà Nội, Việt Nam

do nhiều thuật ngữ và cấu trúc phức tạp, điều này cản trở cả việc hiểu nội dung lẫn hứng thú đọc. Đọc mở rộng (ER) tạo cơ hội tiếp xúc thường xuyên với ngôn ngữ trong văn bản có ý nghĩa; Suk (2017) cho thấy ER cải thiện đọc hiểu, tốc độ đọc và tiếp thu từ vựng, song chỉ xét từ vựng nói chung, chưa xem xét từ vựng chuyên ngành.

Trí tuệ nhân tạo tạo sinh (Generative AI - GenAI) mở ra khả năng điều chỉnh văn bản theo trình độ người học, hỗ trợ hiệu quả các nhiệm vụ học từ vựng tiếp nhận và sản sinh (Ptasznik và cộng sự, 2024) và hỗ trợ giáo viên phát triển tài liệu đọc cho lớp EFL (Xin, 2024). Tuy nhiên, ít nghiên cứu tập trung vào dùng GenAI đơn giản hóa văn bản chuyên ngành để hỗ trợ học từ vựng qua đọc mở rộng và còn ít bằng chứng về cách người học thực sự tiếp nhận loại tài liệu này trong lớp học đại học tại Việt Nam.

Khoảng trống nghiên cứu nằm ở giao điểm của ba yếu tố: phản hồi của người học đối với văn bản chuyên ngành được GenAI điều chỉnh trong một hoạt động đọc mở rộng thực tế. Nghiên cứu hiện tại cung cấp bằng chứng thăm dò, đa nguồn dữ liệu về khoảng trống đó tại một trường đại học Việt Nam, không nhằm khẳng định hiệu quả của GenAI mà mô tả có hệ thống các mô thức phản hồi của người học.

Nghiên cứu này sử dụng nghiên cứu hành động lớp học (CAR) để tìm hiểu phản hồi của sinh viên không chuyên tại Trường Đại học Kinh tế Kỹ thuật Công nghiệp (UNETI) đối với hoạt động đọc mở rộng với văn bản chuyên ngành được GenAI đơn giản hóa. Vì sinh viên được yêu cầu đọc tối thiểu một văn bản mỗi tuần, hoạt động được triển khai dưới hình

thức đọc mở rộng có hướng dẫn, khác với đọc mở rộng tự nguyện hoàn toàn (Mục III.2). Hai câu hỏi nghiên cứu được đặt ra:

RQ1: Sinh viên phản hồi như thế nào đối với việc học từ vựng chuyên ngành thông qua đọc với văn bản được GenAI đơn giản hóa?

RQ2: Vì sao sinh viên phản hồi như vậy?

II. Cơ sở lý thuyết

Đọc mở rộng (ER) là đọc số lượng lớn văn bản vừa trình độ người học, nhằm tiếp xúc ngôn ngữ tự nhiên hơn là phân tích cấu trúc câu (Suk, 2017). Việc tiếp thu từ vựng chuyên ngành qua đọc thường được lý giải như học từ vựng ngẫu nhiên: người học «nhặt» nghĩa từ trong quá trình xử lý nội dung văn bản, không chủ đích ghi nhớ. Theo giả thuyết Đầu vào có thể hiểu được (Krashen, 1985), việc tiếp thu chỉ hiệu quả khi đầu vào ở mức $i+1$, vừa thách thức vừa trong khả năng xử lý của người học; vượt ngưỡng này, người học bị quá tải nhận thức (Sweller, 1988) và hạn chế ghi nhớ từ mới. Khung lý thuyết này giúp lý giải vì sao văn bản chuyên ngành nguyên bản là rào cản đặc thù với người học trình độ thấp.

Gablasova (2014) so sánh việc tiếp thu từ kỹ thuật của học sinh Slovak khi đọc bằng L1 và L2, cho thấy nhóm đọc L2 nhớ được ít nghĩa từ hơn, hiểu nông hơn và quên nhanh hơn. Ở hướng đọc mở rộng, Suk (2017) cho thấy đọc mở rộng cải thiện đáng kể tốc độ đọc, đọc hiểu và tiếp thu từ vựng so với phương pháp truyền thống, dù đo từ vựng nói chung, không phân biệt từ chuyên ngành.

GenAI giúp giải quyết vấn đề quá tải nhận thức bằng cách điều chỉnh độ khó

văn bản. Çelik và cộng sự (2024) cho thấy văn bản blog được ChatGPT đơn giản hóa cải thiện đọc hiểu và suy luận ở sinh viên EFL Thổ Nhĩ Kỳ. Aldamen và cộng sự (2026), kết hợp trình độ người học với phiên bản văn bản (gốc và được AI điều chỉnh theo khung ACTFL), phát hiện người học trình độ thấp cải thiện rõ điểm suy luận và tìm thông tin khi đọc văn bản điều chỉnh, trong khi người học trình độ cao không khác biệt đáng kể; tuy nhiên cả hai nghiên cứu đo đọc hiểu và suy luận như cấu trúc độc lập, không xem việc học từ vựng chuyên ngành là kết quả riêng.

AI cũng được dùng như công cụ hỗ trợ từ vựng trực tiếp: Ptasznik và cộng sự (2024) cho thấy ChatGPT vượt trội hơn từ điển Longman về độ chính xác và tốc độ trong nhiệm vụ từ vựng tiếp nhận và sản sinh, dù nhiệm vụ tách biệt khỏi ngữ cảnh đọc và đối tượng là từ vựng chung. Gần chủ đề ESP hơn, Liu và Ma (2025) mô tả một nghiên cứu trường hợp về học từ vựng chuyên ngành qua corpus được ChatGPT hỗ trợ, cải thiện rõ nhất ở người học trình độ thấp, song cỡ mẫu nhỏ khó khái quát hóa.

Ba hướng nghiên cứu trên cùng cho thấy khoảng cách giữa độ khó văn bản và trình độ người học là biến số trung tâm, nhưng đo lường các kết quả khác nhau trên các thiết kế không thể so sánh trực tiếp. Renandya và cộng sự (2025) lập luận rằng GenAI mở ra cơ hội mở rộng quy mô và cá nhân hóa văn bản đọc mở rộng, luận điểm mà nghiên cứu này góp phần kiểm chứng bằng dữ liệu thực nghiệm ở bối cảnh ESP tại Việt Nam. Nghiên cứu hành động lớp học được lựa chọn vì phù hợp để nắm bắt trải nghiệm và phản hồi định tính của sinh viên trong bối cảnh lớp học thực

tế, nơi các thiết kế thực nghiệm nghiêm ngặt khó triển khai.

III. Phương pháp nghiên cứu

Nghiên cứu sử dụng nghiên cứu hành động lớp học (CAR) trong 10 tuần, gồm hai chu kỳ liên tiếp (tuần 1-5 và tuần 6-10), nhằm tìm hiểu và lý giải phản hồi của sinh viên đối với việc học từ vựng chuyên ngành qua đọc mở rộng có hướng dẫn bằng văn bản được AI điều chỉnh (RQ1, RQ2). CAR phù hợp vì mục tiêu không phải kiểm định hiệu quả trong điều kiện thực nghiệm kiểm soát, mà là tìm hiểu quá trình triển khai, phản hồi người học và các điều chỉnh sư phạm cần thiết, qua bốn bước mỗi chu kỳ: lập kế hoạch, thực hiện, quan sát, phản ánh. Kết quả cuối Chu kỳ 1 định hướng trực tiếp việc điều chỉnh cho Chu kỳ 2.

3.1. Bối cảnh và người tham gia

Nghiên cứu thực hiện tại UNETI với 40 sinh viên không chuyên tiếng Anh, năm thứ nhất, 18-20 tuổi, thuộc một lớp nguyên vẹn học giáo trình Market Leader. Giáo viên đồng thời là nhà nghiên cứu; vai trò kép này tiềm ẩn nguy cơ thiên lệch khi quan sát và diễn giải dữ liệu, nên nghiên cứu đã tam giác hoá dữ liệu qua 4 phương pháp thu thập để giảm thiểu rủi ro này. Trình độ đầu vào của lớp không đồng đều, dao động A2-B1 theo CEFR (23 sinh viên B1, 17 sinh viên A2) và phỏng vấn có tính đến sự đa dạng này. Sinh viên tham gia tự nguyện, được thông báo mục đích nghiên cứu và có quyền rút lui bất cứ lúc nào. Nội dung Market Leader không bị thay đổi; can thiệp diễn ra ngoài giờ học, với điểm kết nối duy nhất là chủ đề và từ vựng, bài đọc giúp sinh viên gặp trước thuật ngữ chuyên ngành trước khi học chính thức trên lớp.

3.2. Quy trình can thiệp

Đầu mỗi tuần, sinh viên nhận ba bài đọc, mỗi bài khoảng 300-400 từ, lấy từ nguồn bên ngoài giáo trình (báo chí kinh doanh, blog chuyên ngành, bài viết phổ thông về kinh tế), chọn theo chủ đề và từ vựng trùng với nội dung sắp học trong Market Leader tuần đó, không trích đoạn từ chính giáo trình, nhằm tránh vi phạm bản quyền và đảm bảo tiếp xúc thuật ngữ chuyên ngành trong nhiều ngữ cảnh.

Văn bản sau đó được ChatGPT điều chỉnh độ khó, đơn giản hóa từ vựng ngoài nhóm chuyên ngành và cấu trúc câu phức theo biên độ trình độ A2-B1 của lớp, trong khi giữ nguyên thuật ngữ chuyên ngành và nội dung chuyên môn cốt lõi. Prompt được sử dụng thống nhất qua các tuần: *“Hãy điều chỉnh văn bản sau cho phù hợp với người học tiếng Anh trình độ A2-B1 (CEFR). Đơn giản hóa từ vựng thông thường và cấu trúc câu phức, nhưng giữ nguyên các thuật ngữ chuyên ngành sau: [danh sách từ mục tiêu]. Giữ nguyên nội dung chuyên môn cốt lõi. Độ dài văn bản khoảng 300-400 từ.”* Danh sách từ mục tiêu được xây dựng từ từ vựng chuyên ngành trong giáo trình Market Leader của tuần tương ứng.

Giáo viên kiểm tra thủ công từng văn bản về độ chính xác, tính mạch lạc và độ phù hợp trước khi phát. Sinh viên đọc tối thiểu một văn bản mỗi tuần và được khuyến khích đọc thêm; quy định này tạo ra một yếu tố bắt buộc, khác với đọc mở rộng hoàn toàn tự nguyện, nên nghiên cứu xem đây là “đọc mở rộng có hướng dẫn”. Sau khi đọc, sinh viên hoàn thành nhật ký đọc ngắn, ghi nhận mức độ hoàn thành, độ khó cảm nhận, việc dùng từ điển/công cụ dịch và một từ hoặc cụm từ ghi nhớ được.

3.3. Thu thập và phân tích dữ liệu

Bốn nguồn dữ liệu được dùng để tam giác hóa: (1) bảng hỏi Likert cuối chu kỳ, đo khả năng tiếp cận văn bản, độ khó từ vựng, tự tin, hứng thú, hữu ích và ý định tiếp tục đọc; (2) nhật ký đọc hàng tuần; (3) quan sát lớp học, ghi nhận biểu hiện quen thuộc với từ vựng chuyên ngành; và (4) phỏng vấn bán cấu trúc cuối chu kỳ với cùng 18 sinh viên đại diện mức độ tham gia và trình độ đầu vào. Bảng hỏi do nhóm nghiên cứu xây dựng dựa trên các khung: Đầu vào có thể hiểu được (Krashen, 1985), Tải nhận thức (Sweller, 1988), tự hiệu năng (Bandura, 1977), phát triển hứng thú (Hidi và Renninger, 2006) và Chấp nhận Công nghệ (Davis, 1989). Bảng hỏi và nhật ký phân tích thống kê mô tả; phỏng vấn và quan sát phân tích chủ đề. Sự hội tụ giữa các nguồn củng cố; khác biệt được làm rõ qua phỏng vấn. Với 40 sinh viên, không nhóm đôi chứng và hai chu kỳ ngắn, kết quả mang tính mô tả, cần thận trọng khi diễn giải và không nên suy rộng thêm.

IV. Kết quả nghiên cứu

Phần này trình bày kết quả theo hai câu hỏi nghiên cứu. RQ1 được trả lời qua bảng hỏi, nhật ký đọc và quan sát lớp học; RQ2 được trả lời chủ yếu qua phỏng vấn, đối chiếu ngược lại với ba nguồn trên. Cùng một nhóm 40 sinh viên tham gia xuyên suốt hai chu kỳ, không có nhóm đối chứng, các số liệu dưới đây phản ánh thay đổi trong nội bộ một nhóm (thiết kế trong nhóm, within-subject), trình bày bằng thống kê mô tả.

4.1. Phản hồi tự báo cáo của sinh viên (RQ1)

Bảng 1. Phản hồi tự báo cáo của sinh viên (N = 40)

Tiêu chí	Chu kỳ 1 M (SD)	Chu kỳ 2 M (SD)	Chênh lệch
Khả năng tiếp cận văn bản	3,42 (0,71)	3,77 (0,64)	+0,35
Tự đánh giá tiếp xúc, ghi nhớ từ vựng	3,51 (0,68)	3,81 (0,61)	+0,30
Tự tin và cảm nhận hữu ích	3,38 (0,76)	3,82 (0,65)	+0,44
Hứng thú và tự chủ	3,21 (0,82)	3,56 (0,74)	+0,35
Ý định tiếp tục đọc	3,35 (0,79)	3,79 (0,67)	+0,44

Ghi chú. Thang Likert 5 mức (1 = hoàn toàn không đồng ý, 5 = hoàn toàn đồng ý). Số liệu là thống kê mô tả; nghiên cứu không thực hiện kiểm định thống kê suy luận (xem Mục III.3 và Mục IV.4).

Cả năm tiêu chí đều tăng ở Chu kỳ 2, thấp nhất ở tự đánh giá ghi nhớ từ vựng (+0,30), cao nhất ở tự tin/cảm nhận hữu ích và ý định tiếp tục đọc (đều +0,44). Với N = 40 và không có kiểm định thống kê suy luận, chênh lệch dưới 0,5 điểm nên được xem là xu hướng mô tả cần diễn giải thận trọng, phản ánh cảm nhận tự báo cáo chứ không phải năng lực từ vựng khách quan.

Ở cấp độ từng câu hỏi, tỷ lệ đồng ý rằng văn bản phù hợp trình độ tăng từ 55% lên 75%; tỷ lệ đồng ý bài đọc giúp gặp lại từ vựng trong nhiều ngữ cảnh tăng từ 60% lên 77,5%; tỷ lệ đồng ý về khả năng ghi nhớ từ chuyên ngành tăng thấp hơn, từ 57,5% lên 70%, nhất quán với Bảng 1.

4.2. Mức độ tham gia, biểu hiện trên lớp (RQ1)

Nhật ký đọc cho thấy tỷ lệ hoàn thành tối thiểu một văn bản mỗi tuần tăng từ 87,5% lên 92,5%; tỷ lệ đọc thêm ngoài yêu cầu tăng từ 27,5% lên 40%. Tỷ lệ mô tả bài đọc khó hoặc rất khó giảm từ 37,5% xuống 22,5%, trong khi tỷ lệ dùng từ điển hoặc công cụ dịch giảm ít hơn, từ 72,5% xuống 62,5%, một điểm phân kỳ được lý giải thêm qua phỏng vấn.

Quan sát lớp học do giáo viên, đồng thời là nhà nghiên cứu, thực hiện, không

có quan sát viên độc lập, nên kết quả là bằng chứng bổ trợ. Tỷ lệ sinh viên chủ động nhận diện hoặc nhắc lại thuật ngữ đã gặp khi học nội dung tương ứng trên lớp tăng từ 35% lên 52,5%. Tỷ lệ sử dụng thuật ngữ trong phát biểu chỉ được ghi nhận ở Chu kỳ 2, đạt 32,5%, vì chỉ số này bổ sung sau Chu kỳ 1. Khoảng cách giữa nhận diện thụ động và sử dụng chủ động là điểm phân kỳ thứ hai, phù hợp với mức tăng khiêm tốn ở tiêu chí ghi nhớ từ vựng.

4.3. Lý do phía sau phản hồi của sinh viên (RQ2)

Phỏng vấn được thực hiện với cùng 18 trong 40 sinh viên ở cả hai chu kỳ, chọn có chủ đích theo mức độ tham gia và trình độ (A2, B1), xem tiêu chí chọn mẫu chi tiết ở Mục III.3. Bốn chủ đề được xác định qua phân tích chủ đề: (1) văn bản dễ tiếp cận hơn, (2) gặp lại từ chuyên ngành trong ngữ cảnh mới, (3) văn phụ thuộc từ điển hoặc công cụ dịch, và (4) động lực đọc phụ thuộc mức độ phù hợp của chủ đề.

14/18 sinh viên cho rằng văn bản dễ đọc hơn văn bản nguyên bản họ từng gặp. 12/18 cho biết gặp lại thuật ngữ trong nhiều văn bản giúp nhận ra từ nhanh hơn khi học Market Leader. Về điểm phân kỳ công cụ dịch, 11/18 vẫn dùng công cụ, nhưng với lý do đã đổi khác: ở Chu kỳ 2, nhiều sinh

viên dùng công cụ chủ yếu để xác nhận nghĩa chính xác của thuật ngữ cụ thể, thay vì để hiểu toàn bộ nội dung như trước. Sinh viên cũng nhắc tính liên quan với Market Leader như yếu tố đi kèm việc đọc thường xuyên hơn, và bài đọc dễ bị bỏ dở khi văn bản dài hoặc chủ đề kém hấp dẫn.

4.4. Đối chiếu dữ liệu và giới hạn

Bốn nguồn dữ liệu hội tụ ở xu hướng tích cực hơn tại Chu kỳ 2, với phỏng vấn lý giải nhất quán cho hai điểm phân kỳ đã nêu, cho thấy sự quen thuộc với từ vựng không đồng nghĩa với khả năng sử dụng từ.

Bốn giới hạn cần nhấn mạnh: (1) thiết kế trong nhóm (within-subject) không đối chứng khiến các thay đổi quan sát được có thể chịu ảnh hưởng của hiệu ứng làm quen quy trình và phát triển ngôn ngữ tự nhiên, nên các phát biểu chỉ mô tả xu hướng liên hệ, không phải quan hệ nhân quả; (2) nghiên cứu đo phản hồi và trải nghiệm, không đo thụ đắc từ vựng bằng bài kiểm tra chuẩn hóa; (3) dữ liệu quan sát và mã hóa phỏng vấn dựa trên một giáo viên duy nhất, đồng thời là nhà nghiên cứu, nên mang tính chủ quan nhất định dù đã áp dụng tam giác hóa; (4) nghiên cứu chỉ sử dụng thống kê mô tả, không có kiểm định thống kê suy luận, do tính chất thăm dò và cỡ mẫu nhỏ của một nghiên cứu hành động lớp học.

Những kết quả và giới hạn này là cơ sở để phần Bàn luận tiếp tục đối chiếu với các nghiên cứu trước về đọc mở rộng và văn bản được điều chỉnh bằng AI.

V. Thảo luận

Kết quả nghiên cứu cho thấy một bức tranh nhất quán nhưng không đồng đều: đọc mở rộng có hướng dẫn với văn bản được AI điều chỉnh có liên hệ với

mức tăng ở hầu hết các khía cạnh trải nghiệm của sinh viên, nhưng mức tăng lớn nhất nằm ở động lực và cảm nhận hữu ích, không phải ở chỉ số gần nhất với “học từ vựng”. Phát hiện này phù hợp với khung lý thuyết Đầu vào có thể hiểu được của Krashen và Cognitive Load Theory: khi văn bản nằm trong vùng $i+1$, người học có đủ dư địa nhận thức để xử lý từ vựng thay vì chỉ giải mã cú pháp, nhưng lợi ích về thái độ và thói quen đọc dường như bộc lộ nhanh hơn lợi ích ghi nhớ từ vựng. Kết quả này cũng cùng hướng với phát hiện của Suk (2017) về đọc mở rộng nói chung (xem Mục II), dù thước đo và thiết kế của hai nghiên cứu không thể so sánh trực tiếp.

Khoảng cách giữa nhận diện thụ động (52,5%) và sử dụng chủ động (32,5%) thuật ngữ trên lớp củng cố thêm cách đọc này, phù hợp với hiện tượng khoảng cách thụ động-chủ động đã được ghi nhận rộng rãi trong thụ đắc từ vựng L2. Kết quả này cũng gần với phát hiện của Liu và Ma (2025) khi nghiên cứu việc dùng GPT hỗ trợ học từ vựng theo hướng corpus cho người học trình độ thấp, gợi ý rằng công cụ AI có thể hỗ trợ tốt việc nhận diện từ vựng chuyên ngành hơn là chuyển hóa thành khả năng sử dụng chủ động, dù thiết kế của Liu và Ma can thiệp vào công cụ tra cứu chứ không phải văn bản đọc.

Điểm phân kỳ về công cụ dịch, nơi tỷ lệ sử dụng giảm chậm hơn độ khó cảm nhận, có thể soi chiếu qua phát hiện của Ptasznik và cộng sự (2024) khi so sánh ChatGPT với từ điển Longman, cho thấy công cụ AI vượt trội đặc biệt ở nhiệm vụ sản sinh từ vựng. Nếu sinh viên trong nghiên cứu này chuyển từ dùng công cụ

dịch để hiểu nghĩa chung sang xác nhận thuật ngữ cụ thể, điều đó phần nào phản ánh xu hướng công cụ AI ngày càng dùng cho nhiệm vụ từ vựng chính xác hơn, dù đây vẫn là khả năng cần kiểm chứng thêm.

Kết quả về mức độ hưởng lợi khác nhau tùy độ khó văn bản cũng gợi nhắc phát hiện của Aldamen và cộng sự (2026), dùng thiết kế 2×2 so sánh trình độ và phiên bản văn bản: người học trình độ Intermediate-Low cải thiện đáng kể trên văn bản được AI điều chỉnh, trong khi Advanced-Low không khác biệt đáng kể. Đây là cơ sở lý thuyết khả dĩ để giải thích vì sao lớp học trình độ không đồng đều A2-B1 trong nghiên cứu hiện tại có thể có mức hưởng lợi khác nhau giữa hai nhóm, dù nghiên cứu hiện tại chưa tách riêng phân tích theo từng nhóm trình độ (xem Giới hạn). Tương tự, Çelik và cộng sự (2024) phát hiện văn bản blog được ChatGPT đơn giản hóa cải thiện đọc hiểu và suy luận nhưng không làm giảm lo âu khi đọc, cho thấy lợi ích nhận thức không nhất thiết chuyển hóa đồng đều sang khía cạnh cảm xúc hoặc động lực, điểm tương đồng với việc nghiên cứu hiện tại ghi nhận mức tăng động lực rõ hơn mức tăng cảm nhận ghi nhớ từ vựng.

Thiết kế trong nhóm (within-subject) không đối chứng, không đo từ vựng khách quan khiến các liên hệ trên không thể tách khỏi ảnh hưởng của việc làm quen với quy trình hoặc phát triển ngôn ngữ tự nhiên song song. Hướng nghiên cứu tiếp theo nên bổ sung nhóm đối chứng, đo từ vựng trước-sau và phân tích riêng theo nhóm trình độ A2/B1, theo

hướng thiết kế của Aldamen và cộng sự (2026). Đóng góp chính của nghiên cứu là mô tả có hệ thống các mô thức phản hồi khác biệt giữa động lực và cảm nhận ghi nhớ từ vựng, trong một bối cảnh tiếng Anh chuyên ngành (ESP) thực tế tại Việt Nam, nơi bằng chứng về văn bản AI điều chỉnh còn hạn chế.

VI. Kết luận và kiến nghị

Nghiên cứu hành động cho thấy đọc mở rộng có hướng dẫn với văn bản chuyên ngành được GenAI đơn giản hóa có liên hệ với mức tăng rõ nhất về động lực, cảm nhận hữu ích và mức độ tham gia của sinh viên (RQ1), trong khi mức tăng về ghi nhớ từ vựng khiêm tốn hơn; tỷ lệ sử dụng chủ động thuật ngữ (32,5%) vẫn thấp hơn nhận diện thụ động (52,5%). Đối với RQ2, phản hồi tích cực chủ yếu được lý giải bởi khả năng tiếp cận văn bản, hứng thú và cơ hội gặp lại từ vựng trong nhiều ngữ cảnh. So với các nghiên cứu trước chủ yếu đánh giá đọc hiểu hoặc kết quả từ vựng bằng công cụ khách quan, nghiên cứu này bổ sung bằng chứng định tính đa nguồn về trải nghiệm đọc trong lớp học tiếng Anh thực tế tại Việt Nam. Tuy nhiên, do không có nhóm đối chứng, cỡ mẫu nhỏ và không sử dụng bài kiểm tra từ vựng chuẩn hóa, các kết quả chỉ nên được xem là xu hướng trong bối cảnh nghiên cứu. Nghiên cứu tiếp theo cần sử dụng thiết kế đối chứng, thước đo chuẩn hóa và phân tích theo trình độ A2/B1.

Tài liệu tham khảo

- Aldamen, H., Al-Deaibes, M., Alnemrat, A., & Alsharefeen, R. (2026). Using AI to adapt reading input for mixed-proficiency EFL learners. *Frontiers in Education, 11*, 1737903. <https://doi.org/10.3389/feduc.2026.1737903>

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Çelik, F., Yangın Ersanlı, C., & Arslanbay, G. (2024). Does AI simplification of authentic blog texts improve reading comprehension, inferencing, and anxiety? *International Review of Research in Open and Distributed Learning*, 25(3), 287-303. <https://doi.org/10.19173/irrodl.v25i3.7779>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Gablasova, D. (2014). Learning and retaining specialized vocabulary from textbook reading: Comparison of learning outcomes through L1 and L2. *The Modern Language Journal*, 98(4), 976-991. <https://doi.org/10.1111/modl.12150>
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111-127. https://doi.org/10.1207/s15326985ep4102_4
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Lacková, M. (2025). Teaching ESP vocabulary with the application of corpus-based materials at universities: A rapid review of literature. *Arab World English Journal*, 16(4), 273-288. <https://doi.org/10.24093/awej/voll6no4.16>
- Liu, J., & Ma, Q. (2025). Supporting low-proficiency L2 learners' vocabulary learning with custom GPT-scaffolded corpus-based language pedagogy: A case study. *Computer Assisted Language Learning*. Advance online publication. <https://doi.org/10.1080/09588221.2025.2539152>
- Nguyễn, T. T. N. (2024). Những thách thức ngôn ngữ đối với sinh viên trong các lớp học môn chuyên ngành bằng tiếng Anh. *Tạp chí Khoa học Trường Đại học Mở Hà Nội*, 19-26. <https://doi.org/10.59266/houjs.2024.508>
- Ptasznik, B., Wolfer, S., & Lew, R. (2024). A learners' dictionary versus ChatGPT in receptive and productive lexical tasks. *International Journal of Lexicography*, 37(3), 322-336. <https://doi.org/10.1093/ijl/eca011>
- Renandya, W. A., Floris, F. D., & Zhou, J. (2025). Extensive reading in the age of artificial intelligence. *International Journal of TESOL Studies*, 250902, 1-12. <https://doi.org/10.58304/ijts.250902>
- Suk, N. (2017). The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. *Reading Research Quarterly*, 52(1), 73-89. <https://doi.org/10.1002/rrq.152>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4
- Vũ, T. M. Q. (2026). Sử dụng trò chơi trực tuyến để cải thiện từ vựng tiếng Anh cho sinh viên không chuyên năm thứ nhất tại Trường Đại học Mở Hà Nội. *Tạp chí Khoa học Trường Đại học Mở Hà Nội*, Số đặc biệt 11A, 178. <https://doi.org/10.59266/houjs.2026.1204>
- Xin, J. J. (2024). Investigating EFL teachers' use of generative AI to develop reading materials: A practice and perception study. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/13621688241303321>

AI-ADAPTED SPECIALIZED TEXTS FOR GUIDED EXTENSIVE READING: A CLASSROOM ACTION RESEARCH ON NON-ENGLISH-MAJOR STUDENTS' RESPONSES IN VIETNAM

Pham Thu Hien¹

Abstract: *This classroom action research investigated non-English-major students' responses to learning specialized vocabulary through guided extensive reading (at least one text per week) using AI-adjusted texts and examined factors underlying responses. Forty students participated in two intervention cycles. Data were collected through questionnaires, reading logs, classroom observations, and semi-structured interviews with 18 students in both cycles. All self-reported measures increased in Cycle 2, with the largest gains in confidence, perceived usefulness, and intention to continue reading, whereas vocabulary retention increased more modestly. Interviews indicated that text accessibility and relevance to classroom learning were the factors explaining positive responses. Given the uncontrolled within-subject design, small sample, absence of inferential testing, and lack of a standardized vocabulary test, findings should be interpreted as associative trends within this context. The study provides empirical evidence from Vietnamese university EFL classrooms that positive reading experiences may not translate into comparable gains in perceived specialized vocabulary retention.*

Keywords: *AI-adjusted texts, classroom action research, guided extensive reading, specialized vocabulary*

¹ University of Economics - Technology for Industries, Hanoi, Vietnam