NHẬN THỨC CỦA SINH VIÊN VỀ RỦI RO CHUYÊN NGÀNH KHI SỬ DỤNG AI TẠO SINH TRONG PHÂN TÍCH NGÔN NGỮ: KHẢO SÁT TẠI TRƯỜNG NGOẠI NGỮ, ĐẠI HỌC THÁI NGUYÊN

NGUYỄN THỊ HẢI ANH
Trường Ngoại ngữ, Đại học Thái Nguyên

Nhận bài ngày 20/9/2026. Sửa chữa xong 20/10/2026. Duyệt đăng 03/11/2026.

Abstract

This study investigates students' perceptions of discipline-specific risks associated with using generative artificial intelligence in linguistic analysis. Data were collected from 207 students who had used generative AI in the Introduction to Linguistics course at the School of Foreign Languages, Thai Nguyen University. The data were analyzed using descriptive statistics, group-comparison tests, and rank-based correlation analysis. The findings indicate that most students recognized several limitations of AI: 59.9% identified the risk of inaccurate information or fabricated rules, 59.4% recognized the risk of incorrect analyses of complex or ambiguous Vietnamese structures, and 58.5% acknowledged AI's limitations in interpreting culturally specific nuances. Risk perceptions were positively associated with students' perceived need for assessment rubrics, discipline-specific prompting guidance, and structured support for integrating AI use with critical thinking. However, self-reported perceptions do not demonstrate students' actual ability to detect errors in AI-generated output. The study therefore recommends designing learning tasks that require students to verify, justify, and defend their analytical decisions using linguistic evidence.

Keywords: *Discipline-specific risks, generative AI, Introduction to Linguistics, linguistic analysis, student perceptions.*

1. Đặt vấn đề

Các trí tuệ nhân tạo (AI) tạo sinh đang được sinh viên (SV) sử dụng để giải thích khái niệm, tìm ví dụ và hỗ trợ xử lý nhiệm vụ học tập. Nhiều nghiên cứu tại Việt Nam đều ghi nhận rằng, việc sử dụng ChatGPT ngày càng phổ biến, đồng thời cảnh báo nguy cơ tiếp nhận nội dung thiếu kiểm soát (Đặng Văn Em và cộng sự, 2024; Nguyễn Thị Phương Hoa, 2024). Do đó, năng lực AI không chỉ bao gồm kỹ năng viết câu lệnh mà còn đòi hỏi tư duy kiểm chứng độ tin cậy, nhận thức giới hạn công nghệ và trách nhiệm giải trình đối với kết quả đầu ra (Long & Magerko, 2020; Miao et al., 2024). Trong phân tích ngôn ngữ, yêu cầu này đặc biệt quan trọng. Trên thực tế, một câu trả lời trôi chảy vẫn có thể chứa định nghĩa gần đúng, quy tắc bị khái quát hóa quá mức hoặc phương án phân tích phiến diện (Ji et al., 2023; Mahowald et al., 2024). Giá trị giáo dục của AI vì thế phụ thuộc vào kiến thức nền và cách nhiệm vụ buộc người học đối chiếu, giải thích và sửa kết quả (Forchini & Murphy, 2025; Loock & Holt, 2024). Dù vậy, khả năng nhận diện các rủi ro mang tính đặc thù trong phân tích ngôn ngữ ở người học hiện chưa được nghiên cứu đầy đủ. Mục tiêu nghiên cứu là khảo sát nhận thức của SV về ba nhóm rủi ro: tri thức ngôn ngữ học, phân tích cấu trúc - ngữ nghĩa và sắc thái văn hóa - ngữ dụng; đồng thời xem xét sự khác biệt theo đặc điểm người học và mối liên hệ với nhu cầu hỗ trợ. Phân tích chỉ dừng lại ở mức độ mô tả - khám phá, không đóng vai trò là công cụ đo lường năng lực thực hành hay cơ sở suy luận tác động của AI lên kết quả học tập.

2. Phương pháp nghiên cứu

Nghiên cứu áp dụng thiết kế khảo sát cắt ngang trong giai đoạn đánh giá thực trạng của đề tài tích hợp AI tạo sinh vào học phần Dẫn luận ngôn ngữ tại Trường Ngoại ngữ, Đại học Thái Nguyên. Việc

Email: nguyenhaianh.sfl@tnu.edu.vn

báo cáo kết quả tuân thủ các nguyên tắc minh bạch theo chuẩn STROBE đối với nghiên cứu quan sát (von Elm et al., 2007). Phiếu khảo sát tiếng Việt được giảng viên (GV) giảng dạy học phần phát hành trực tuyến qua Google Forms từ ngày 16 đến ngày 23/3/2026. Trong số 224 phiếu thu được, 207 SV xác nhận đã sử dụng AI trong học phần và hoàn thành nhánh đánh giá rủi ro; 17 người chưa sử dụng AI trả lời một nhánh riêng. Các phân tích của bài viết tập trung vào 207 trường hợp sử dụng AI.

Dữ liệu được thu thập dựa trên phương thức chọn mẫu thuận tiện và tự nguyện. Do không dựng khung chọn mẫu phân lớp và không ghi nhận số lượng được mời, nghiên cứu không tính toán tỷ lệ phản hồi. Mẫu nghiên cứu có sự mất cân đối về giới tính và chuyên ngành, vì vậy các kết quả chỉ phản ánh thực trạng tại một cơ sở đào tạo cụ thể. Ngoài ra, dữ liệu không thu thập mã lớp nên nghiên cứu không thể đánh giá sự phụ thuộc giữa các SV trong cùng môi trường học tập.

Dữ liệu bảng hỏi gồm thông tin nền, kinh nghiệm sử dụng AI, lợi ích và rủi ro cảm nhận, thực hành kiểm chứng và nhu cầu hỗ trợ. Các chỉ báo chuyên ngành được phân tích trong bài là: AI có thể tạo thông tin sai hoặc quy tắc không có thật; AI có thể phân tích sai câu tiếng Việt phức tạp hoặc đa nghĩa; và AI có thể không hiểu đầy đủ sắc thái văn hóa đặc thù. Ba nội dung khác được sử dụng để mở rộng bối cảnh bao gồm: giảm tra cứu tài liệu, giảm tư duy độc lập và nguy cơ đạo văn. Thang đo Likert năm mức được sử dụng để đo các mục và phân tích riêng, không cộng thành điểm tổng hợp. Trong nghiên cứu, cách xử lý này phù hợp với nguyên tắc thận trọng khi diễn giải dữ liệu Likert, đặc biệt khi từng mục không được thiết kế như một thang đo thống nhất (Sullivan & Artino, 2013).

Các dữ liệu thu thập được kiểm tra về giá trị thiếu, bản ghi trùng hoàn toàn và logic rẽ nhánh. Kết quả, không phát hiện bản ghi trùng hoàn toàn hoặc dữ liệu thiếu ở các biến chính. Thống kê mô tả sử dụng tần số, tỷ lệ phần trăm, trung vị và khoảng tứ phân vị. Hai mức "Đồng ý" và "Hoàn toàn đồng ý" được gộp tính tỷ lệ tán thành song vẫn bảo lưu phân bố gốc. Phân tích so sánh nhóm áp dụng kiểm định Mann-Whitney U hoặc Kruskal-Wallis tùy theo số nhóm, với kích thước ảnh hưởng đo bằng hệ số tương quan hạng-nhị phân hoặc epsilon bình phương. Spearman và Kendall tau-b được sử dụng để xam xét mối liên hệ giữa từng rủi ro và nhu cầu hỗ trợ. Sai số loại I phát sinh từ các phép kiểm định đa, bội được kiểm soát bằng phương pháp hiệu chỉnh Holm (Holm, 1979). Các phân tích này mang tính khám phá, không được diễn giải theo hướng nhân quả.

Nghiên cứu thu được 154 phản hồi có nội dung từ phần câu hỏi mở. Dữ liệu câu hỏi mở chỉ được tổng hợp các nhóm đề xuất chính mà không phát triển thành hợp phần định tính độc lập do chưa thực hiện mã hóa đa người và tính toán độ nhất trí liên mã hóa. Nghiên cứu được Hội đồng của Đại học Thái Nguyên phê duyệt trong khuôn khổ đề tài. Người học được thông báo về mục đích, cách sử dụng dữ liệu, tính ẩn danh và quyền tham gia tự nguyện. Việc tham gia không ảnh hưởng đến điểm học phần hay quyền lợi học tập của người học.

3. Kết quả nghiên cứu

3.1. Trong mẫu khảo sát gồm 207 SV có sử dụng AI, nữ giới chiếm tỷ lệ 93,2% và SV chuyên ngành Ngôn ngữ Trung Quốc chiếm 63,3%. Kết quả, 51,7% SV tự đánh giá năng lực số ở mức khá (18,8% tốt/rất tốt); tỷ lệ sử dụng AI thường xuyên và thỉnh thoảng lần lượt là 34,8% và 55,6%. Cơ cấu này cho thấy, AI đã trở thành công cụ quen thuộc trong nhóm khảo sát. Dù vậy, sự chênh lệch cơ cấu về giới tính và ngành học là rào cản đối với việc suy rộng kết quả ra phạm vi tổng thể.

Bảng 1: Đặc điểm nhóm SV có sử dụng AI (N = 207)

Biến	Nhóm	n	%
Giới tính	Nữ	193	93,2
	Nam	11	5,3
	Không muốn trả lời	3	1,4
Ngành học	Ngôn ngữ Trung Quốc	131	63,3
	Song ngữ	71	34,3
	Ngành khác	5	2,4

Năng lực số	Trung bình trở xuống	61	29,5
	Khá	107	51,7
	Tốt hoặc rất tốt	39	18,8
Kinh nghiệm AI	Rất ít	20	9,7
	Thỉnh thoảng	115	55,6
	Thường xuyên	72	34,8

Tỷ lệ tán thành đối với 6 rủi ro dao động từ 49,3% đến 59,9%, trong khi đó, tỷ lệ không đồng ý đều thấp (bảng 2). Các chỉ báo chuyên ngành có mức tán thành gần nhau, lần lượt là 59,9% (nguy cơ sinh thông tin/quy tắc giả), 59,4% (lỗi phân tích câu phức/đa nghĩa) và 58,5% (hạn chế xử lý sắc thái văn hóa). Kết quả cho thấy, phần lớn người tham gia đã nhận thức rằng độ trôi chảy của đầu ra không bảo đảm độ chính xác chuyên môn. Tuy nhiên, tỷ lệ trung lập ở ba chỉ báo vẫn từ 34,8% đến 39,1%. Điều này phản ánh nhận thức chưa đồng đều hoặc sự thiếu chắc chắn khi người học tự đánh giá. Tỷ lệ tán thành không được xem là bằng chứng rằng, SV có thể phát hiện và giải thích lỗi trong một nhiệm vụ thực tế.

Bảng 2: Nhận thức về rủi ro khi sử dụng AI

Rủi ro	Không đồng ý n (%)	Trung lập n (%)	Đồng ý n (%)	Trung vị (IQR)
Thông tin sai hoặc quy tắc không có thật	11 (5,3)	72 (34,8)	124 (59,9)	4 (3-4)
Giảm tra cứu giáo trình và nguồn khoa học	16 (7,7)	89 (43,0)	102 (49,3)	3 (3-4)
Giảm tư duy độc lập và sự tinh tế ngôn ngữ	13 (6,3)	78 (37,7)	116 (56,0)	4 (3-4)
Không hiểu sắc thái văn hóa đặc thù	5 (2,4)	81 (39,1)	121 (58,5)	4 (3-4)
Nguy cơ đạo văn khi không ghi nhận nguồn AI	8 (3,9)	82 (39,6)	117 (56,5)	4 (3-4)
Phân tích sai câu phức tạp hoặc đa nghĩa	5 (2,4)	79 (38,2)	123 (59,4)	4 (3-4)

Cần lưu ý rằng lựa chọn trung lập không thể coi là biểu hiện của sự thiếu nhận thức ở SV. Thái độ này xuất phát từ việc người học chưa trải qua lỗi cụ thể, thiếu tiêu chí đánh giá hoặc chưa hiểu rõ nội dung câu hỏi. Dù vậy, tỷ lệ trên 33% chọn mức giữa tại cả ba chỉ báo chuyên ngành phản ánh nhận thức rủi ro của một phần lớn SV vẫn chưa đạt tới mức độ phán đoán khẳng định. Khoảng trung lập này gợi mở hàm ý sư phạm: thay vì chỉ cảnh báo rủi ro từ AI, GV cần xây dựng tình huống thực hành giúp người học quan sát lỗi, đối chiếu phương án và luận giải căn cứ. Sự lệch pha giữa nhận thức tự báo cáo và biểu hiện thực tế giải thích cho xu hướng chuyển dịch sang đánh giá năng lực AI bằng kiểm tra kiến thức hoặc bài tập tình huống trong các nghiên cứu gần đây (Jin et al., 2025).

Mặc dù chênh lệch với hai rủi ro còn lại rất nhỏ, rủi ro tri thức được tán thành cao nhất. Ở cấp độ nhập môn, người học dễ bị thuyết phục bởi các quy tắc "có vẻ đúng" do AI tạo ra nhờ vỏ bọc thuật ngữ chuyên môn và lập luận có vẻ hợp lý. Sai lệch có thể xuất hiện khi AI trộn lẫn hệ thuật ngữ, biến khuynh hướng thành quy tắc bắt buộc hoặc đưa ra nguồn không thể xác minh. Các giáo trình ngôn ngữ học đều đặt khái niệm trong quan hệ với cấp độ phân tích và hệ thống thuật ngữ tương ứng (Mai Ngọc Chừ và cộng sự 2020; Nguyễn Thiện Giáp và cộng sự 2020). Vì vậy, hoạt động học tập cần yêu cầu SV xác định nguồn của quy tắc, giới hạn phạm vi áp dụng và tìm phản ví dụ, thay vì chỉ chấp nhận một định nghĩa có hình thức hợp lý.

Về mặt cấu trúc - ngữ nghĩa, kết quả 59,4% cho thấy, SV đã lưu ý đến nguy cơ AI xử lý phiến diện câu phức tạp hoặc đa nghĩa. Do ranh giới thành phần hoặc quan hệ phụ thuộc không duy nhất, một câu có thể cho phép nhiều cách phân tích. Chẳng hạn, câu "Tôi nhìn thấy người đàn ông với chiếc ống nhòm" dung chứa hai khả năng kết hợp: phương tiện quan sát của chủ thể (người nói) hoặc đặc điểm nhận diện của đối tượng (người đàn ông). Một đầu ra chỉ nêu một cách hiểu không nhất thiết sai hoàn toàn, nhưng chưa đầy đủ nếu bỏ qua khả năng còn lại. Việc phân tích tiếng Việt cần xem xét đồng thời cấu trúc, chức năng và ngữ cảnh (Diệp Quang Ban, 2012). Do đó, đánh giá đầu ra AI trong ngôn ngữ học nên chuyển từ lựa chọn "đúng/sai" đơn giản sang yêu cầu nêu phương án, bằng chứng và điều kiện ngữ cảnh.

58,5% SV tán thành rủi ro văn hóa - ngữ dụng. Mặc dù AI giải thích tốt về nghĩa hiển ngôn nhưng

lại có thể bỏ qua quan hệ xã hội, mức độ lịch sự hoặc chức năng của phát ngôn. Chẳng hạn, câu "Anh về đấy à?" có hình thức nghi vấn nhưng trong một số tình huống chủ yếu thực hiện chức năng chào hỏi. Nghĩa ngữ dụng chỉ được xác định khi đặt phát ngôn trong quan hệ giữa người nói, người nghe và hoàn cảnh giao tiếp (Nguyễn Đức Dân, 1998). Vì vậy, khi sử dụng AI để phân tích xưng hô, hàm ý hoặc hành vi ngôn ngữ, SV cần cung cấp ngữ cảnh và kiểm tra xem công cụ có giải thích được ai nói, nói với ai, trong quan hệ nào và nhằm mục đích gì.

Sự tán thành cũng được ghi nhận khá cao cho các rủi ro còn lại: 49,3% đối với giảm tra cứu giáo trình và nguồn khoa học; 56,0% đối với giảm tư duy độc lập và sự tinh tế ngôn ngữ; 56,5% đối với nguy cơ đạo văn khi không ghi nhận việc sử dụng AI. Các tỷ lệ trên cho thấy, nhận thức của người học không chỉ tập trung vào độ chính xác của nội dung mà còn liên quan đến thói quen học tập và trách nhiệm học thuật. Kết quả này tương đồng với những quan sát tại Việt Nam về việc người học đánh giá cao tính tiện lợi của AI nhưng còn hạn chế trong kiểm soát nội dung hoặc đánh giá phê phán (Đặng Văn Em và cộng sự, 2024; Nguyễn Thị Phương Hoa, 2024).

Phân tích so sánh giữa các nhóm theo ngành học, năng lực công nghệ số và kinh nghiệm AI không cho thấy sự khác biệt có ý nghĩa thống kê sau hiệu chỉnh Holm đối sánh bội; chỉ số kích thước ảnh hưởng ở tất cả các phép kiểm định đều thuộc khoảng rất nhỏ đến không đáng kể (bảng 3). Kết quả này không chứng minh các nhóm hoàn toàn giống nhau, bởi một số nhóm có dung lượng nhỏ và mẫu không cân đối. Trong phạm vi dữ liệu hiện có, nhận thức về ba rủi ro chuyên ngành tương đối phổ biến ở các nhóm được khảo sát. Vì vậy, nội dung hướng dẫn kiểm chứng nên được xem là yêu cầu chung của học phần. Tuy vậy, nghiên cứu tiếp theo vẫn cần mẫu cân bằng hơn để đánh giá các khác biệt tinh tế.

Bảng 3: So sánh nhận thức rủi ro theo đặc điểm người học

Đặc điểm	Rủi ro	p thô	p Holm	Kích thước ảnh hưởng
Ngành học	Tri thức	0,746	1,000	|rrb| = 0,025
	Cấu trúc - nghĩa	0,859	1,000	|rrb| = 0,013
	Văn hóa - ngữ dụng	0,444	1,000	|rrb| = 0,058
Năng lực số	Tri thức	0,794	1,000	$\varepsilon^2 = 0,000$
	Cấu trúc - nghĩa	0,185	1,000	$\varepsilon^2 = 0,007$
	Văn hóa - ngữ dụng	0,284	1,000	$\varepsilon^2 = 0,003$
Kinh nghiệm AI	Tri thức	0,205	1,000	$\varepsilon^2 = 0,006$
	Cấu trúc - nghĩa	0,504	1,000	$\varepsilon^2 = 0,000$
	Văn hóa - ngữ dụng	0,687	1,000	$\varepsilon^2 = 0,000$

Sự vắng mặt của các khác biệt có ý nghĩa thống kê cần được lý giải thông qua đặc điểm phân bố của cấu trúc mẫu. Sự áp đảo của nhóm nữ và ngành Ngôn ngữ Trung Quốc tạo ra các phân nhóm đối sánh dung lượng nhỏ, làm hạn chế khả năng nhận biết các khác biệt có kích thước ảnh hưởng nhỏ. Vì vậy, kết quả phù hợp nhất với nhận định rằng, dữ liệu hiện có chưa cung cấp bằng chứng về sự khác biệt, chứ không phải chứng minh ngành học, năng lực số hay kinh nghiệm AI hoàn toàn không liên quan đến nhận thức rủi ro. Năng lực công nghệ số có thể giúp người học thao tác công cụ thuận lợi hơn nhưng không tự động bảo đảm khả năng đánh giá một phân tích cú pháp, ngữ nghĩa hoặc ngữ dụng. Tri thức chuyên ngành và kinh nghiệm kiểm chứng vẫn là điều kiện riêng cần được hình thành trong quá trình học tập.

Nhận thức rủi ro có tương quan thuận với ba nhu cầu hỗ trợ (bảng 4). Hệ số Spearman dao động từ 0,243 đến 0,481 và Kendall tau-b từ 0,224 đến 0,447; các mối liên hệ vẫn có ý nghĩa thống kê sau hiệu chỉnh Holm. Mức tương quan cao nhất xuất hiện giữa nhận thức rủi ro cấu trúc - ngữ nghĩa và nhu cầu kết hợp AI với tư duy phản biện ($\rho = 0,481$; $\tau b = 0,447$). Như vậy, SV nhận thức rủi ro cao hơn cũng có xu hướng mong muốn được hỗ trợ rõ hơn. Tuy nhiên, tương quan không cho phép kết luận nhận thức rủi ro tạo ra nhu cầu hỗ trợ hoặc ngược lại.

Bảng 4: Tương quan giữa nhận thức rủi ro và nhu cầu hỗ trợ

Rủi ro	Nhu cầu hỗ trợ	Spearman ρ	Kendall τ-b
Tri thức	Rubric	0,399	0,374
	Prompting	0,400	0,374
	AI và tư duy phản biện	0,385	0,353
Cấu trúc - nghĩa	Rubric	0,409	0,386
	Prompting	0,458	0,433
	AI và tư duy phản biện	0,481	0,447
Văn hóa - ngữ dụng	Rubric	0,243	0,224
	Prompting	0,288	0,265
	AI và tư duy phản biện	0,324	0,297

Mức tương quan từ yếu đến trung bình cho thấy nhu cầu hỗ trợ còn chịu ảnh hưởng của những yếu tố khác chưa được đo lường, chẳng hạn kinh nghiệm gặp lỗi AI, mức độ khó của nhiệm vụ, sự hướng dẫn của GV hoặc mức tự tin của người học. Dù vậy, xu hướng đồng biến có giá trị đối với thiết kế học phần: cảnh báo rủi ro không nhất thiết làm SV né tránh AI mà có thể đi cùng mong muốn được hướng dẫn sử dụng công cụ bài bản hơn. Trong ba dạng hỗ trợ, rubric giúp cụ thể hóa tiêu chí đánh giá; hướng dẫn prompting hỗ trợ xác định yêu cầu và ngữ cảnh; còn quy trình kết hợp AI với tư duy phản biện giúp người học kiểm tra đầu ra sau khi công cụ đã trả lời. Ba hình thức này có chức năng bổ sung cho nhau, không nên được xem như những giải pháp thay thế hoàn toàn. Kết quả trên có thể được diễn giải từ góc độ cảnh giác nhận thức: người tiếp nhận cần đánh giá cả độ đáng tin của nguồn và tính hợp lý của nội dung được truyền đạt (Sperber et al., 2010). Trong tương tác với AI, sự trôi chảy và giọng điệu chắc chắn dễ làm giảm động lực kiểm tra. Bởi vậy, nhận thức rằng AI "có thể sai" là điều kiện khởi đầu nhưng chưa đủ để chứng minh người học biết AI sai ở đâu, vì sao sai và cần sửa như thế nào. Phân biệt này cũng phù hợp với các công trình cho thấy tri thức nền và hướng dẫn có cấu trúc ảnh hưởng đến mức độ tham gia phê phán của người học khi sử dụng AI để phân tích ngôn ngữ (Forchini & Murphy, 2025; Loock & Holt, 2024).

3.2. Từ góc độ chuyên ngành, năng lực AI trong ngôn ngữ học có thể được hiểu là khả năng khai thác, thẩm định và hiệu chỉnh đầu ra AI bằng hệ thống khái niệm, ngữ liệu và phương pháp lập luận của ngành. Khả năng này khác với sự thành thạo công nghệ nói chung. Một SV có thể đặt câu lệnh chi tiết nhưng vẫn khó phát hiện việc AI gán sai chức năng cú pháp hoặc bỏ qua hàm ý giao tiếp; ngược lại, người có tri thức ngôn ngữ học vẫn cần biết cách cung cấp ngữ cảnh, lưu vết đầu ra và tổ chức kiểm chứng. Cách hiểu này không nhằm xây dựng một thang đo mới từ ba mục hỏi của nghiên cứu mà dùng để xác định yêu cầu của nhiệm vụ học tập. Năng lực chỉ có thể được chứng minh khi người học thực hiện việc phát hiện, giải thích và sửa lỗi trên đầu ra cụ thể, thay vì chỉ tán thành một phát biểu về rủi ro.

Các phản hồi mở chủ yếu đề nghị GV hướng dẫn cách đặt câu lệnh theo yêu cầu môn học, cung cấp nguồn để đối chiếu, chỉ ra cách nhận biết thông tin sai và tổ chức bài tập buộc người học tự lập luận. Từ kết quả định lượng và những gợi ý trên, GV có thể thử nghiệm nhiệm vụ bốn bước: xác định hiện tượng và ngữ cảnh; lưu lại câu lệnh cùng đầu ra AI; đối chiếu với giáo trình, ngữ liệu, phản ví dụ và phương án thay thế; sau đó chỉnh sửa và giải trình lựa chọn cuối cùng. Rubric có thể tập trung vào độ chính xác khái niệm, sự nhất quán thuật ngữ, căn cứ ngữ liệu, khả năng xem xét nhiều cách hiểu và tính minh bạch của nguồn. Đây là đề xuất thiết kế dựa trên dữ liệu khảo sát, chưa phải một can thiệp đã được chứng minh hiệu quả.

Khi triển khai, GV không nên đánh giá việc sử dụng AI chỉ qua sản phẩm cuối cùng. Một câu trả lời đúng có thể xuất phát từ việc sao chép, trong khi một phương án chưa hoàn chỉnh vẫn cho thấy quá trình lập luận có giá trị nếu SV nhận ra điểm bất định và biết tìm thêm bằng chứng. Vì vậy, hồ sơ nhiệm vụ cần giữ được dấu vết của câu lệnh, phiên bản đầu ra, nguồn đối chiếu, những điểm SV đồng ý hoặc bác bỏ và lý do chỉnh sửa. Cách tổ chức này giúp chuyển trọng tâm từ việc cấm hay khuyến khích AI một cách chung chung sang đánh giá chất lượng quyết định học thuật của người học. Với các hiện tượng đa nghĩa hoặc phụ thuộc ngữ cảnh, GV cũng có thể yêu cầu SV trình bày ít nhất hai phương án phân tích, xác định điều kiện khiến mỗi phương án phù hợp và nêu rõ dữ liệu còn thiếu. Những yêu cầu

đó trực tiếp đáp ứng ba nhóm rủi ro được nhận diện trong khảo sát nhưng cần được kiểm chứng bằng nghiên cứu can thiệp có bài làm thực tế, tiêu chí chấm độc lập và thiết kế trước - sau.

Nghiên cứu có một số giới hạn. Mẫu thuận tiện chỉ được thu thập tại một cơ sở, nữ chiếm tỷ lệ rất cao và SV Ngôn ngữ Trung Quốc chiếm đa số. Mỗi rủi ro được đo bằng một mục hỏi riêng; dữ liệu hoàn toàn dựa trên tự báo cáo và không lưu prompt, phiên bản mô hình hoặc đầu ra AI thực tế. Câu hỏi mở chưa được mã hóa độc lập. Vì vậy, kết quả không cho phép suy rộng cho SV nói chung, đánh giá năng lực phát hiện lỗi hoặc khẳng định AI ảnh hưởng đến kết quả học tập. Nghiên cứu tiếp theo cần sử dụng công cụ đa mục đã thẩm định, nhiệm vụ có cài lỗi kiểm soát, chấm độc lập và thiết kế trước - sau hoặc nhóm so sánh.

4. Kết luận

Khảo sát 207 SV sử dụng AI trong học phần Dẫn luận ngôn ngữ cho thấy nhận thức về ba rủi ro chuyên ngành đạt mức tương đối cao và khá gần nhau: 59,9% nhận thức nguy cơ thông tin sai hoặc quy tắc không có thật; 59,4% nhận thức nguy cơ phân tích sai câu phức tạp hoặc đa nghĩa; 58,5% nhận thức hạn chế trong xử lý sắc thái văn hóa. Tỷ lệ trung lập còn lớn cho thấy nhận thức này chưa hình thành đồng đều. Các khác biệt theo ngành học, năng lực số và kinh nghiệm AI không có ý nghĩa thống kê sau hiệu chỉnh Holm, trong khi nhận thức rủi ro có tương quan thuận với nhu cầu hỗ trợ về rubric, prompting và tư duy phản biện. Đóng góp của nghiên cứu nằm ở việc đặt rủi ro AI trong yêu cầu cụ thể của phân tích ngôn ngữ, nơi một đầu ra cần được đánh giá bằng khái niệm, ngữ liệu, phương án giải thích và ngữ cảnh. Tuy nhiên, nhận thức tự báo cáo không đồng nghĩa với năng lực phát hiện và giải thích lỗi. Vì thế, việc tích hợp AI vào học phần cần chuyển trọng tâm từ tiếp nhận câu trả lời sang kiểm chứng và bảo vệ phương án phân tích. Quy trình bốn bước được đề xuất có thể làm cơ sở cho các nghiên cứu can thiệp tiếp theo nhưng hiệu quả của nó chỉ có thể được xác định bằng nhiệm vụ thực hành và thiết kế thực nghiệm phù hợp.

Tài liệu tham khảo

Diệp Quang Ban (2012). *Giáo trình Ngữ pháp tiếng Việt: Theo định hướng ngữ pháp chức năng*. NXB Đại học Sư phạm.

Đặng Văn Em, Nguyễn Đình Loan Phương, Nguyễn Thị Hảo (2024). Thực trạng ứng dụng ChatGPT trong việc học tập, nghiên cứu của sinh viên Đại học Quốc gia Thành phố Hồ Chí Minh. *Tạp chí Giáo dục, 24*(1), 36-41. https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/1212.

Forchini, P., & Murphy, A. C. (2025). iThink, therefore iCheck: Critical engagement with ChatGPT in linguistic analysis and learning. *Applied Corpus Linguistics, 5*(3), 100169. https://doi.org/10.1016/j.acorp.2025.100169.

Hoàng, V. H. (2025). The interplay between Vietnamese EFL students' artificial intelligence literacy and tool usage for academic writing. *TNU Journal of Science and Technology, 230*(4), 393-400. https://doi.org/10.34238/tnu-jst.12221.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), Article 248. https://doi.org/10.1145/3571730.

Jin, Y., Martinez-Maldonado, R., Gašević, D., & Yan, L. (2025). GLAT: The generative AI literacy assessment test. *Computers and Education: Artificial Intelligence, 9*, 100436. https://doi.org/10.1016/j.caeai.2025.100436.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727.

Loock, R., & Holt, B. (2024). Augmented linguistic analysis skills: Machine translation and generative AI as pedagogical aids for analyzing complex English compounds. *Technology in Language Teaching & Learning, 6*(3), 1-27. https://doi.org/10.29140/tltl.v6n3.1489.

Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2024). Dissociating language and thought in large language models. *Trends in Cognitive Sciences, 28*(6), 517-540. https://doi.org/10.1016/j.tics.2024.01.011.

Mai Ngọc Chừ, Vũ Đức Nghiệu, Hoàng Trọng Phiến (2020). *Cơ sở ngôn ngữ học và tiếng Việt*. NXB Giáo dục Việt Nam.

Miao, F., Shiohira, K., & Lao, N. (2024). *AI competency framework for students*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391105.

Nguyễn Đức Dân (1998). *Ngữ dụng học*. NXB Giáo dục.

Nguyễn Thiện Giáp, Đoàn Thiện Thuật, Nguyễn Minh Thuyết (2020). *Dẫn luận ngôn ngữ học* (Tái bản lần thứ 20). NXB Giáo dục Việt Nam.

Nguyễn Thị Phương Hoa (2024). Thực trạng sử dụng ChatGPT để luyện viết của sinh viên. *Tạp chí Khoa học và Công nghệ Đại học Thái Nguyên, 229*(12), 260-269. https://doi.org/10.34238/tnu-jst.10606.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359-393. https://doi.org/10.1111/j.1468-0017.2010.01394.x.

Sullivan, G. M., & Artino, A. R., Jr. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education, 5*(4), 541-542. https://doi.org/10.4300/JGME-5-4-18.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Annals of Internal Medicine, 147*(8), 573-577. https://doi.org/10.7326/0003-4819-147-8-200710160-00010.