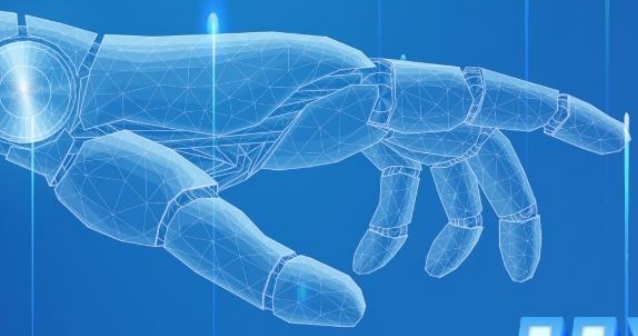
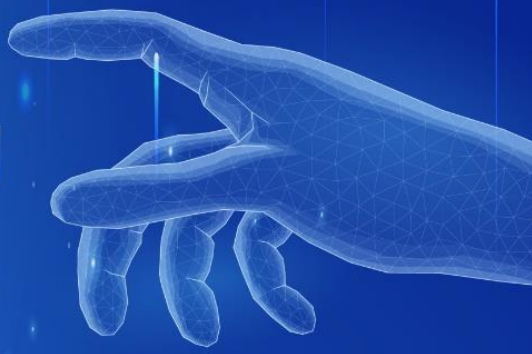
KỶ YẾU HỘI THẢO KHOA HỌC QUỐC GIA LẦN 3

NGÔN NGỮ HỌC TÍNH TOÁN

CÔNG NGHỆ VÌ CON NGƯỜI





#VCL 2025

Thành phố Hồ Chí Minh, ngày 27 tháng 11 năm 2025

Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh

TRƯỜNG ĐẠI HỌC NGOẠI NGỮ - TIN HỌC THÀNH PHỐ HỒ CHÍ MINH

HỘI NGÔN NGỮ HỌC VIỆT NAM

KỶ YẾU HỘI THẢO KHOA HỌC QUỐC GIA LẦN 3

NGÔN NGỮ HỌC TÍNH TOÁN: CÔNG NGHỆ VÌ CON NGƯỜI

NHÀ XUẤT BẢN TRI THỨC

2025

BAN TỔ CHỨC HỘI THẢO

Trưởng ban:

- *PGS.TS. Nguyễn Lân Trung,* Chủ tịch Hội Ngôn ngữ học Việt Nam
- *TS. Nguyễn Anh Tuấn,* Hiệu trưởng Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh

Phó Trưởng ban:

- *PGS.TS. Nguyễn Ngọc Vũ,* Phó Hiệu trưởng Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh (thường trực)
- *PGS.TS. Đặng Ngọc Lệ,* Phó Chủ tịch Hội Ngôn ngữ học Việt Nam, Chủ tịch Hội Ngôn ngữ học TP. Hồ Chí Minh

Các Ủy viên:

- *PGS.TS. Đinh Điền,* Giám đốc Trung tâm Ngôn ngữ học tính toán, Trường Đại học Khoa học Tự nhiên TP. Hồ Chí Minh
- *PGS.TS. Phan Văn Hòa,* Phó Chủ tịch Hội Ngôn ngữ học Việt Nam, Chủ tịch Hội Ngôn ngữ học TP. Đà Nẵng
- *PGS.TS. Trần Văn Lăng,* Trưởng phòng Tạp chí Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh
- *PGS.TS. Nguyễn Thanh Bình,* Phó Trưởng phòng Đào tạo Sau đại học - Khoa học Công nghệ Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh
- *TS. Phạm Văn Lam,* Tổng thư ký Hội Ngôn ngữ học Việt Nam
- *ThS. Nguyễn Thị Lan Hường*, Phó Chủ tịch - Tổng Thư ký phân hội VietCALL, Thành viên Hội Ngôn ngữ học Việt Nam
- *TS. Lương Thị Hiền,* Phó trưởng Bộ môn Ngôn ngữ Trường Đại học Sư phạm Hà Nội, Thành viên Hội Ngôn ngữ học Việt Nam
- *Đào Minh Phương,* Trưởng phòng Hành chính - Tổng hợp Hội Ngôn ngữ học Việt Nam
- *TS. Nguyễn Đức Châu,* Trưởng Bộ môn Biên - Phiên dịch Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh

- *TS. Trần Vũ Diễm Thúy,* Trưởng Bộ môn Ngôn ngữ học Ứng dụng Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh
- *TS. Nguyễn Đức Cường*, Trưởng Khoa CNTT Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh
- *TS. Võ Thị Hồng Tuyết,* Giảng viên Khoa CNTT Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh
- *TS. Trần Minh Thái,* Giảng viên Khoa CNTT Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh
- *TS. Trần Văn Tiếng,* Giảng viên Khoa Ngôn ngữ & VHPĐ Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh
- *TS. Hoàng Thị Kim Oanh,* Giảng viên Khoa Ngôn ngữ & VHPĐ Trường Đại học Ngoại ngữ - Tin học TP. Hồ Chí Minh

BAN BIÊN TẬP KỶ YẾU

Trưởng ban:

PGS.TS. Nguyễn Ngọc Vũ

Các Ủy viên:

TS. Phạm Văn Lam

ThS. Lê Thị Thảo

ThS. Nguyễn Thị Mỹ Hằng

Tchen Vân An

BIÊN TẬP KĨ THUẬT

Đào Minh Phương

MỤC LỤC

LỜI NÓI ĐẦU **7**

1. SỬ DỤNG TỪ TREND MẠNG XÃ HỘI TRONG TIÊU ĐỀ BÁO MẠNG ĐIỆN TỬ: GÓC NHÌN TỪ CÔNG CHÚNG **1**

Trần Văn Vỹ, Đỗ Thanh Huyền, Lại Thị Ninh, Nguyễn Minh Tâm 1

2. KHAI THÁC NGỮ LIỆU CHUYÊN NGÀNH PHÁP LUẬT BẰNG TRÍ TUỆ NHÂN TẠO (AI) TRONG GIẢNG DẠY NGOẠI NGỮ **20**

Nguyễn Thị Cẩm Nhung 20

3. MATH HANDWRITING CONVERSION USING AI APPROACHES: A LIGHTER ALTERNATIVE TO SCANNING **31**

Hien Nguyen Phuoc, Nghia Cao Trong, Linh, Nguyen Hoang Anh 31

4. KHÁM PHÁ CÁC PHƯƠNG PHÁP TÍCH HỢP KỸ THUẬT SỐ TRONG GIẢNG DẠY TIẾNG ANH CHO SINH VIÊN **37**

Nguyễn Thị Luyến, Nguyễn Thúy Ngọc 37

5. VAI TRÒ CỦA TRÍ TUỆ NHÂN TẠO TRONG SỰ HỘI NHẬP GIÁO DỤC CỦA VIỆT NAM VỚI THẾ GIỚI **51**

Phan Thị Ngọc Lệ 51

6. TỪ DỮ LIỆU ĐẾN ẨN DỤ: QUY TRÌNH DỰA TRÊN KHỐI NGỮ LIỆU ĐỂ PHÂN TÍCH BÀI VIẾT HỌC THUẬT TIẾNG ANH BẰNG ANTCONC VÀ MIPVU **64**

Trần Thế Phi 64

7. NHỮNG CÁCH XỬ LÝ NGỮ NGHĨA CỦA AI KHI DỊCH TỤC NGỮ VIỆT SANG TIẾNG HÀN **81**

Trần Văn Tiếng, Nguyễn Thị Kim Ngọc 81

8. ỨNG DỤNG HỘI HỌA AI TRONG GIẢNG DẠY PHÂN MÔN ĐỌC (Thực hiện qua phân tích khung cảnh thiên nhiên trong các tác phẩm) **93**

Nguyễn Thị Hoài Thương, Phạm Phương Linh 93

9. ĐÁNH GIÁ KHẢ NĂNG GÁN NHÃN NGỮ NGHĨA TỰ ĐỘNG CỦA MÔ HÌNH NGÔN NGỮ LỚN CHO MỤC TỪ DANH TỪ TIẾNG VIỆT **112**

Phan Thị Mỹ Trang, Đỗ Quốc Trí, Phan Văn Bá Hải, Đặng Ngọc Lệ, Đinh Điền, Trần Thị Minh Phượng 112

10. ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TẠO SINH THIẾT KẾ NỘI DUNG ÔN LUYỆN THỰC HÀNH TIẾNG VIỆT CHO HỌC SINH LỚP 8 ..... **128**

Đỗ Thị Thu Hương 128

11. ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TRONG DẠY HỌC VIẾT VĂN MIÊU TẢ PHONG CẢNH CHO HỌC SINH LỚP 5 143

Đỗ Lam Ngọc 143

12. PHÂN LOẠI PHÁT NGÔN THÙ GHÉT TIẾNG VIỆT BẰNG MÔ HÌNH NGÔN NGỮ LỚN ĐƯỢC TINH CHỈNH THEO HƯỚNG CHỈ DẪN 164

Lê Tỷ Khánh - Nguyễn Hoàng Minh Nhật - Lê Thanh 164

13. AN EXPLORATORY CASE STUDY ON THE IMPACT OF CHATGPT ON ARGUMENTATIVE THINKING IN IELTS WRITING TASK 2 174

Ngô Thị Thu Hiền, Trần Phương Anh 174

14. TIẾP CẬN PHÁT HIỆN TIN GIẢ BẰNG TRÍ TUỆ NHÂN TẠO CÓ GIẢI THÍCH 207

Đào Xuân Kỳ, Đinh Minh Hoà, Tiếu Phùng Mai Sương[2] 207

15. TRÍ TUỆ NHÂN TẠO VÀ ĐỔI MỚI ĐÀO TẠO NHÂN LỰC TIẾNG TRUNG CHUYÊN NGÀNH THƯƠNG MẠI Ở VIỆT NAM 220

Nguyễn Duy An 220

16. INTERGRATING AI TOOLS IN SPEAKING SKILLS: A CASE STUDY FROM TEACHER'S PERSPECTIVES ON PERSONALIZED INSTRUCTION 230

Phùng Thị Bích Ngọc , Pham Ngoc Diem 230

17. CẢI THIỆN KHẢ NĂNG ÁNH XẠ NGỮ NGHĨA OALD-WORDNET BẰNG KỸ THUẬT BIỂU DIỄN VECTOR TRI THỨC TỔNG HỢP 246

Phan Văn Bá Hải, Đỗ Quốc Trí, Phan Thị Mỹ Trang, Đinh Điền 246

18. A CORPUS-BASED STUDY ON THE USE OF EPISTEMIC MODALS IN TED TALKS ABOUT TECHNOLOGY 258

Tran Quoc Toan, Nguyen Ngo Linh Phung, Nguyen Thi Tu My, Tran Thi Huynh Tran 258

19. EVALUATIVE ADJECTIVES AND CONCEPTUAL METAPHORS IN MUSIC REVIEWS: A CORPUS-BASED ANALYSIS OF PITCHFORK REVIEWS..... 290

Vo Huu Khanh, Nguyen Cao Thang, Truong Pham Xuan Mai, Nguyen Tran Kim Thanh, Bui Vy Yen Ly 290

20. XÂY DỰNG NGỮ LIỆU HUẤN LUYỆN CHO MÔ HÌNH NHẬN DẠNG VĂN BIA HÁN NÔM 305

Trương Thị Vân Anh, Bùi Hồng Phúc, Đinh Điền 305

21. ỨNG DỤNG TRÍ TUỆ NHÂN TẠO AI TRONG GIẢNG DẠY TIẾNG VIỆT CHO HỌC VIÊN LÀO HỌC TIẾNG VIỆT TẠI HỌC VIỆN AN NINH NHÂN DÂN . 318

Tống Thị Hảo 318

22. XÂY DỰNG HỆ THỐNG HỖ TRỢ THÍCH ỨNG ĐỂ TỰ ĐỘNG TẠO HƯỚNG DẪN SỬ DỤNG PHẦN MỀM TỪ NGỮ CẢNH TRỰC QUAN VÀ CƠ SỞ TRI THỨC 325

Lê Đoàn Phương Uyên, Nguyễn Hồng Bửu Long 325

23. CHUYỂN ĐỔI SỐ TRONG ĐÀO TẠO TIẾNG VIỆT CHO HỌC VIÊN LÀO TẠI HỌC VIỆN AN NINH NHÂN DÂN: ỨNG DỤNG AI CHO HỌC LIỆU TƯƠNG TÁC VÀ CÁ NHÂN HÓA 337

Nguyễn Thụy Trang 337

24. ỨNG DỤNG CÔNG NGHỆ TRÍ TUỆ NHÂN TẠO (AI) TRONGHỌC PHẦN PHƯƠNG PHÁP PHÁT TRIỂN NGÔN NGỮ CHO TRẺ MẦM NON 349

Trần Thị Ban Mai 349

25. NOTEBOOKLM TRONG ĐỔI MỚI GIẢNG DẠY VÀ NGHIÊN CỨU NGOẠI NGỮ 361

Trần Ngọc Hà 361

26. BRIDGING AI AND CULTURAL HERITAGE: A DOMAIN-SPECIFIC CHATBOT FOR CLASSICAL SINO-VIETNAMESE NOVELS 372

Huynh Thanh Xuan, Pham Hoang Vu, Luu Thien Duc, U Cao Ky Long 372

27. AI IN LANGUAGE TEACHING: BRIDGING TECHNOLOGY AND COMMUNICATION 383

Phạm Ngọc Diễm 383

28. VAI TRÒ CỦA VIỆC ỨNG DỤNG AI TRONG DẠY HỌC BIÊN - PHIÊN DỊCH TIẾNG TRUNG QUỐC TẠI TRƯỜNG ĐẠI HỌC SƯ PHẠM HÀ NỘI 2 392

Nguyễn Thị Nga, Nguyễn Thị Lan Anh 392

29. ỨNG DỤNG LÝ THUYẾT NGÔN NGỮ HỌC CHỨC NĂNG HỆ THỐNG VÀO VIỆC THIẾT KẾ KHUNG PROMPT ĐỂ TẠO SINH NGỮ LIỆU GIẢNG DẠY TIẾNG ANH 403

LÊ THÀNH DANH - LÊ HỮU SỸ 403

30. XÂY DỰNG ĐỒ THỊ TRI THỨC CHO Y HỌC CỔ TRUYỀN VIỆT NAM TỪ TÁC PHẨM NAM DƯỢC THẦN HIỆU - TUỆ TĨNH 418

Nguyễn Thị Xuân Hiền, Phạm Trần Nhật Minh, Nguyễn Thành Giang, Nguyễn Hồng Bửu Long 418

31. APPLICATION OF AI IN AUTOMATICALLY TRANSLITERATING SINO - NÔMSCRIPT INTO VIETNAMESE NATIONAL SCRIPT 426

Tam Tong Trong, Linh Pham Van, Vinh Luong An 426

32. APPLICATION OF AI IN THE AUTOMATIC CONSTRUCTION OF A SINO-NOM – VIETNAMESE NATIONAL SCRIPT PARALLEL CORPUS..... 440

Phuc Bui Hong, Dinh Si Dien, Nguyen Thuy Trang, Truong Thi Van Anh 440

33. USING THE ENGLISH VIETNAMESE CORPUS OF CULTURAL TOURISM AS AN ESP MATERIAL..... 456

Phan Thi Thanh Thao..... 456

34. VIETNAMESE UNIVERSITY STUDENTS' PERCEPTIONS OF QUIZLET IN IELTS VOCABULARY LEARNING..... 476

Truong Thi Nhu Ngoc*, Nguyen Tu Nhan, Nguyen Minh Hung..... 476

35. ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TẠO SINH TRONG DẠY HỌC TIẾNG ANH NHƯ NGÔN NGỮ THỨ HAI..... 500

Nguyễn Thị Mỹ Lệ 500

36. HUMAN-IN-THE-LOOP STRATEGIES IN MACHINE TRANSLATION – CHALLENGES, TRENDS, AND FUTURE DIRECTIONS: A CASE STUDY AT APD..... 513

Nguyen Thi Nuong..... 513

37. PHÂN TÍCH CÁC MÔ HÌNH NGÔN NGỮ HỌC TÍNH TOÁN TRONG HỖ TRỢ XỬ LÝ NHẬP NHẰNG NGÔN NGỮ TỰ NHIÊN..... 533

Huỳnh Quang Đức 533

38. NÂNG CAO HIỆU QUẢ VÀ THÚC ĐẨY HIỂU BIẾT LIÊN VĂN HÓA QUA TÍCH HỢP NLP VÀO DẠY VÀ HỌC NGOẠI NGỮ..... 542

Trần Thị Hải Bình 542

39. MỨC ĐỘ SẴN SÀNG CỦA GIẢNG VIÊN EFL VIỆT NAM TRONG VIỆC TÍCH HỢP AI VÀO GIẢNG DẠY KỸ NĂNG NGHE..... 552

NGUYỄN KHÁNH LĨNH– ĐINH THÚY LY 552

40. GIAO THOA VĂN HÓA GIỮA TỪ HÁN VIỆT VÀ TỪ THUẦN VIỆT (QUA HÁT-ĐỐ VÀ HÁT- ĐỐI TRONG HÁT PHƯỜNG VẢI)..... 570

Đỗ Thị Kim Liên..... 570

41. THE IMPACT OF GAMIFIED FEATURES OF DUOLINGO ON LEARNING MOTIVATION: A STUDY OF ENGLISH-MAJOR STUDENTS' AI-ENHANCED LANGUAGE LEARNING..... 581

Le Ngoc Phi Yen, Tran Vu Diem Thuy 581

42. TRANSLATING EMOTIONS THROUGH METAPHOR: CROSS-CULTURAL INSIGHTS FROM TRUYỆN KIỀU INTO ENGLISH AND JAPANESE CLASSROOMS..... 600

Nguyễn Minh Hưng 600

43. ỨNG DỤNG TRÍ TUỆ NHÂN TẠO VÀO VIỆC BIÊN SOẠN TỪ ĐIỂN BÁCH KHOA (NGHIÊN CỨU TRƯỜNG HỢP BIÊN SOẠN TỪ ĐIỂN BÁCH KHOA VỀ NGHỀ THỦ CÔNG MỸ NGHỆ TẠI HÀ NỘI) 622

Ngô Thị Thu Hương 622

44. TỐI ƯU HÓA GIAO TIẾP ĐA NGÔN NGỮ VÀ CHẤT LƯỢNG DỊCH THUẬT CHUYÊN NGÀNH BẰNG XỬ LÝ NGÔN NGỮ TỰ NHIÊN (NLP) 636

Tô Ngọc Sơn 636

45. POTENTIAL BENEFITS OF THE METAVERSE IN CREATING A VIRTUAL REALITY LEARNING ENVIRONMENT FOR TEACHING ENGLISH COMMUNICATION 646

Nguyen Thi Ngoc Diep, Bui Thi Kim Loan 646

46. NGHIÊN CỨU MÔ HÌNH DẠY - HỌC 3 CHIỀU "GIẢNG VIÊN CHỦ ĐẠO – AI HỖ TRỢ - SINH VIÊN TỰ CHỦ" TẠI LỚP PHIÊN DỊCH CHUYÊN NGÀNH TRUNG - VIỆT 669

Lưu Cẩm Huệ 669

47. SỬ DỤNG AI VÀO CÁC HOẠT ĐỘNG HỌC TẬP DẠNG TRÒ CHƠI, NHẰM MỤC ĐÍCH NÂNG CAO ĐỘNG LỰC HỌC TẬP CHO SINH VIÊN TRONG MÔN BIÊN DỊCH 1 TẠI KHOA NGOẠI NGỮ, HUFLIT 683

Nguyễn Đức Châu 683

48. NHẬN DIỆN QUAN HỆ BAO THUỘC BẰNG CÁC DẤU HIỆU HÌNH THỨC 705

Phạm Văn Lam 705

LỜI NÓI ĐẦU

Hội thảo khoa học quốc gia lần thứ 3 do Hội Ngôn ngữ học Việt Nam phối hợp với Trường Đại học Ngoại ngữ - Tin học Thành phố Hồ Chí Minh tổ chức vào ngày 27 tháng 11 năm 2025 tại Trường Đại học Ngoại ngữ - Tin học Thành phố Hồ Chí Minh, 828 Sư Vạn Hạnh, Phường Hòa Hưng, Thành phố Hồ Chí Minh.

Với chủ đề *"Ngôn ngữ học tính toán: Công nghệ vì con người",* Hội thảo tạo diễn đàn trao đổi về vai trò của ngôn ngữ học tính toán trong việc phục vụ con người và bảo tồn văn hóa. Trong thời đại trí tuệ nhân tạo (AI) phát triển mạnh mẽ, công nghệ xử lý ngôn ngữ tự nhiên (NLP) không chỉ hỗ trợ tự động hóa mà còn giúp duy trì, phát triển bản sắc ngôn ngữ và văn hóa. Hội thảo sẽ tập trung vào các cách thức AI có thể phục vụ giáo dục, hỗ trợ giao tiếp đa ngôn ngữ, bảo tồn di sản văn hóa, và đảm bảo tính nhân bản trong các ứng dụng ngôn ngữ học.

Ban Tổ chức Hội thảo đã nhận được 77 báo cáo, sau khi biên tập đã lựa chọn được 48 báo cáo phù hợp với chủ đề và đáp ứng được yêu cầu về chất lượng.

48 báo cáo khoa học được xếp theo các nội dung chính của hội thảo như sau:

+ Ứng dụng Ngôn ngữ học Tính toán trong công tác bảo tồn và phục hồi các ngôn ngữ bản địa, ngôn ngữ thiểu số có nguy cơ mai một.
+ Vai trò của Xử lý Ngôn ngữ Tự nhiên (NLP) trong giao tiếp đa ngôn ngữ, nâng cao chất lượng dịch thuật tự động và chuyên ngành.
+ Ứng dụng AI trong cá nhân hóa hành trình học tập, xây dựng hệ thống đánh giá ngôn ngữ tự động và hỗ trợ học tập thích ứng.
+ Nghiên cứu về tính nhân bản và đạo đức trong phát triển các hệ thống NLP: đảm bảo định hướng nhân văn trong ứng dụng công nghệ.
+ NLP như một công cụ thúc đẩy giao thoa văn hóa và liên thông ngôn ngữ, hỗ trợ giao tiếp trong bối cảnh toàn cầu hóa.
+ Tích hợp công nghệ giọng nói và xử lý văn bản: nhận diện giọng nói, chuyển đổi văn bản – giọng nói, nâng cao hiệu quả ứng dụng thực tiễn trong giảng dạy và truyền thông.
+ Dịch máy nâng cao và các giải pháp kết hợp giữa công nghệ và yếu tố con người để đảm bảo tính chính xác và tính biểu cảm trong dịch thuật.
+ Khai thác NLP trong phân tích xã hội học ngôn ngữ trên môi trường số, bao gồm phân tích xu hướng ngôn ngữ, phát hiện chủ đề, và phân tích cảm xúc.
+ Ứng dụng NLP trong hỗ trợ người yếu thế, bao gồm các công cụ hỗ trợ người khiếm thính, khiếm thị và người gặp khó khăn trong đọc hiểu.
+ Vấn đề bảo mật dữ liệu và đạo đức công nghệ trong NLP: đảm bảo quyền riêng tư, xử lý dữ liệu công bằng, và giảm thiểu thiên lệch thuật toán trong các hệ thống ngôn ngữ học trí tuệ nhân tạo.

Kỷ yếu được hoàn thành là nhờ tâm sức của tất cả những người tham gia hội thảo, của Ban tổ chức và của Ban Biên tập. Tuy nhiên, vì nhiều lí do, chắc chắn cuốn kỷ yếu còn có những hạn chế, theo đó, sai sót là điều khó tránh khỏi.

Rất mong nhận được những góp ý chân tình của các nhà khoa học, quý thầy cô giảng viên cùng bạn đọc.

BAN BIÊN TẬP KỶ YẾU

SỬ DỤNG TỪ TREND MẠNG XÃ HỘI TRONG TIÊU ĐỀ BÁO MẠNG ĐIỆN TỬ: GÓC NHÌN TỪ CÔNG CHÚNG

Trần Văn Vỹ[1], Đỗ Thanh Huyền[2], Lại Thị Ninh[3], Nguyễn Minh Tâm[4]

Tóm tắt: Mối quan hệ cộng sinh giữa báo chí và mạng xã hội đã mang lại nhiều chuyển biến tích cực, đặc biệt trong việc thu hút sự chú ý của công chúng. Một xu hướng nổi bật trong những năm gần đây là việc sử dụng từ trend mạng xã hội trong tiêu đề báo mạng điện tử. Nghiên cứu này khảo sát 374 độc giả cho thấy, nhìn chung, công chúng có thái độ tích cực đối với thực trạng trên. Việc sử dụng từ trend không chỉ kích thích sự tò mò mà còn tạo cảm giác gần gũi, giúp độc giả dễ tiếp cận thông tin và gia tăng khả năng lựa chọn đọc bài báo. Tuy nhiên, mức độ tương tác sâu như bình luận hay thảo luận vẫn còn hạn chế, cho thấy hiệu quả của từ trend chưa thực sự tạo được gắn kết bền vững. Đáng chú ý, tần suất đọc báo mạng có tác động đáng kể đến mức độ chấp nhận và hài lòng của công chúng. Trên cơ sở kết quả này, nghiên cứu đưa ra các khuyến nghị giúp các cơ quan báo chí khai thác từ trend một cách hợp lý, không chỉ để thu hút độc giả mà còn để đảm bảo chất lượng nội dung và tính chuyên nghiệp. Điều này sẽ giúp xây dựng mối quan hệ bền vững với công chúng và thúc đẩy sự tương tác thực chất hơn.

Từ khóa: báo chí hiện đại, công chúng, mạng xã hội, xu hướng

I. GIỚI THIỆU

Trong bối cảnh truyền thông số phát triển mạnh mẽ, một hiện tượng nổi bật trên báo mạng điện tử là sự bùng nổ của các từ trend mạng xã hội trong tiêu đề bài báo. Những cụm từ như "phông bạt", "check var", "quay xe", "đào mộ", hay "lật mặt" - vốn xuất phát từ giao tiếp trên mạng xã hội - đã nhanh chóng len lỏi vào đời sống báo chí, trở thành "vũ khí" đắc lực để thu hút sự chú ý của độc giả. "Bắt trend" đúng thời điểm không chỉ làm cho tiêu đề báo chí trở nên hấp dẫn, gần gũi mà còn tạo ra sự kết nối mạnh mẽ với công chúng, đặc biệt là giới trẻ. Điều này được thể hiện qua sự phản ứng tích cực của công chúng như tương tác, bình luận, chia sẻ hay thậm chí đưa các từ trend này vào đời sống hàng ngày.

Tuy nhiên, sự lan tỏa của ngôn ngữ mạng xã hội trong báo chí cũng đặt ra không ít tranh cãi. Trong khi một số ý kiến cho rằng đây là sự sáng tạo ngôn ngữ cần thiết để bắt kịp thị hiếu công chúng thì nhiều người lại bày tỏ lo ngại về tính chuyên nghiệp và giá trị cốt lõi của báo chí. Liệu việc sử dụng từ trend có mang lại sự hài lòng của công chúng? Hay công chúng có những tương tác cụ thể nào đối với các tác phẩm báo chí sử dụng từ trend? Những câu hỏi này đang trở nên cấp thiết trong xu hướng truyền thông hiện đại.

[1] Học viên cao học tại Trường Đại học Khoa học Xã hội và Nhân văn – ĐHQG HN, đang công tác tại Trường Đại học Sư phạm, Đại học Đà Nẵng, *email: tvvy@ued.udn.vn*.
[2] Học viên cao học tại Trường Đại học Khoa học Xã hội và Nhân văn – ĐHQG HN, email: thanhhuyentntp@gmail.com
[3] Học viên cao học tại Trường Đại học Khoa học Xã hội và Nhân văn – ĐHQG HN, email: laininhsjc@gmail.com
[4] Học viên cao học tại Trường Đại học Khoa học Xã hội và Nhân văn – ĐHQG HN, email: minhtamnguyen310501@gmail.com

Xuất phát từ thực trạng kể trên, đề tài ***Sử dụng từ trend mạng xã hội trong tiêu đề báo điện tử: Góc nhìn từ công chúng*** được xây dựng với mục tiêu làm rõ phản hồi của công chúng trước hiện tượng này. Nghiên cứu không chỉ tập trung vào việc đánh giá tác động của từ trend đối với trải nghiệm của độc giả mà còn hướng đến đề xuất các khuyến nghị cân bằng giữa sáng tạo ngôn ngữ và tính chuyên nghiệp, qua đó góp phần nâng cao chất lượng báo chí.

II. TỔNG QUAN LÝ THUYẾT

1. Thuật ngữ "từ trend" mạng xã hội

Theo từ điển Oxford[5], trend được định nghĩa là xu hướng hoặc xu thế, chỉ một chiều hướng chung mà sự vật hay sự việc đang thay đổi hoặc phát triển theo đó. Trong bối cảnh hội nhập quốc tế, các trend được du nhập và lan tỏa thông qua nhiều hình thức khác nhau, đặc biệt trong các lĩnh vực như thời trang, ẩm thực, công nghệ, giải trí và phong cách sống. Trong đó, ngôn ngữ cũng là một khía cạnh không thể bỏ qua.

Ngôn ngữ hot trend được hiểu là những cụm từ hoặc cách diễn đạt xuất hiện và lan truyền nhanh chóng theo trào lưu, đặc biệt phổ biến trong giới trẻ. Những cụm từ này thường bắt nguồn từ các sản phẩm truyền thông như phim truyền hình, MV ca nhạc, clip trên mạng xã hội và dần dần thâm nhập vào đời sống hàng ngày [1]. Đặc điểm nổi bật của ngôn ngữ hot trend là được sử dụng lặp đi lặp lại trong một khoảng thời gian ngắn, tạo nên sự lan tỏa mạnh mẽ và rộng rãi.

Trên cơ sở đó, bài báo chuyển dịch "từ trend" là những từ hoặc cụm từ bắt nguồn từ mạng xã hội, mang tính hài hước, sáng tạo hoặc mang ý nghĩa đặc biệt. Những từ này khi được sử dụng liên tục và phổ biến, dần trở thành xu hướng ngôn ngữ không chỉ trong đời sống hàng ngày mà còn thâm nhập vào đời sống báo chí như một nét đặc trưng trong phong cách truyền tải thông tin.

Bảng 1. *Một số từ trend mạng xã hội được sử dụng trên báo chí*

TT	Từ/ Cụm từ	Nghĩa biểu đạt
1	Phông bạt	Phông bạt ám chỉ một lối sống thể hiện sự hào nhoáng, phô trương bên ngoài, thường để che giấu sự thật hoặc tạo ấn tượng không đúng với thực tế.
2	Check var	Xuất phát từ bóng đá (kiểm tra video hỗ trợ trọng tài), ám chỉ việc cần kiểm tra hay xác minh lại thông tin, sự việc để làm rõ đúng sai.
3	Sao kê	Sao kê mang nghĩa gốc là việc liệt kê các giao dịch tài chính. Tuy nhiên, sao kê còn ám chỉ việc minh bạch, công khai tài chính để chứng minh sự trong sạch.

[5] https://www.oxfordlearnersdictionaries.com/

4	Nóc nhà	Ám chỉ người vợ hoặc bạn gái trong gia đình/ cặp đôi. Nóc nhà mang nghĩa là người quyền lực nhất trong gia đình, dùng trong bối cảnh hài hước để nói về sự tôn trọng người phụ nữ.
5	Flex	"Flex" có nghĩa là phô trương, khoe mẽ một cách tự hào về bản thân hoặc điều gì đó ấn tượng.
6	Ra dẻ	Hành động tỏ vẻ, thể hiện, cố gắng làm nổi bật bản thân để chứng tỏ mình hơn người khác.
7	Pha ke	"Pha ke" hay "fa ke" cũng là cách đọc sai có chủ ý từ từ tiếng Anh "fake", mang nghĩa đồ giả, hàng nhái, hoặc điều gì đó không thật.
8	Cà khịa	Từ này mang nghĩa nói móc, trêu chọc hoặc khiêu khích một cách cố ý nhưng mang tính giải trí, hài hước, châm biếm nhẹ nhàng.
9	Ngã ngựa	Ngã ngựa được hiểu như sự thất bại, trong hoàn cảnh này được dùng với nghĩa bị mất chức, cách chức, giáng chức.
10	Tiểu tam	Chỉ người thứ ba xen vào mối quan hệ tình cảm, đặc biệt là kẻ phá vỡ hạnh phúc gia đình, mang nghĩa tiêu cực.

2. Lý thuyết sử dụng và hài lòng (Uses and Gratifications Theory)

Lý thuyết sử dụng và hài lòng (Uses and Gratifications Theory) là một trong những lý thuyết cốt lõi của nghiên cứu truyền thông, giúp giải mã động cơ sâu xa khiến cá nhân, nhóm hay xã hội nói chung tìm đến và sử dụng các phương tiện truyền thông, đồng thời khám phá những lợi ích họ nhận được từ đó. Manh nha xuất hiện từ những năm 1940, lý thuyết này đặt nền tảng cho việc hiểu rõ hơn về mối quan hệ giữa công chúng và truyền thông [2]. Không dừng lại ở lĩnh vực truyền thông, theo thời gian, lý thuyết này được ứng dụng rộng rãi trong các lĩnh vực khác như khoa học máy tính, khoa học chính trị và tâm lý học. Với tính linh hoạt và khả năng thích ứng cao, lý thuyết sử dụng và hài lòng trở thành công cụ đắc lực giúp các nhà nghiên cứu khám phá động cơ, nhu cầu cũng như mức độ thỏa mãn của công chúng khi họ tiếp cận và tương tác với các phương tiện truyền thông trong bối cảnh khác nhau [3], [4], [5].

Nhiều học giả đã làm sáng tỏ các khía cạnh quan trọng của lý thuyết này như mức độ phụ thuộc của cá nhân vào truyền thông [6] và khả năng đáp ứng nhu cầu của công chúng [7]. Những nhu cầu mà các học giả đề cập bao gồm nhu cầu cảm xúc, nhu cầu giao tiếp, nhu cầu tương tác xã hội, nhu cầu nhận thức và nhu cầu được giải trí. Đặc biệt, mô hình kỳ vọng của Palmgreen & Rayburn (1982) đã kết hợp lý thuyết sử dụng và hài lòng với lý thuyết giá trị kỳ vọng, hình thành nên hai khái niệm quan trọng: sự thỏa mãn tìm kiếm thông tin (GS) và sự thỏa mãn thực tế đạt được (GO). Những dẫn chứng trên cho thấy lý thuyết sử dụng và hài lòng không ngừng được củng cố, mở rộng và phát triển theo theo gian, trở thành một nền tảng vững

chắc để giải thích hành vi và động lực tiếp cận thông tin của công chúng trong bối cảnh truyền thông không ngừng thay đổi. Trong nghiên cứu của Katz et al. (1974), mô hình hành vi tiếp xúc với truyền thông được khái quát thông qua chuỗi các nhân tố: Nhân tố xã hội + Nhân tố tâm lý → Kỳ vọng truyền thông → Tiếp xúc truyền thông → Nhu cầu được thỏa mãn. Theo mô hình này, công chúng tiếp xúc với truyền thông chủ yếu để thỏa mãn nhu cầu cá nhân và mức độ thỏa mãn này phụ thuộc vào khả năng tiếp cận phương tiện truyền thông cũng như kinh nghiệm tiếp xúc với các nội dung truyền thông. Mức độ hài lòng hay không hài lòng của công chúng sẽ ảnh hưởng trực tiếp đến hành vi tiếp nhận thông tin trong tương lai, đồng thời điều chỉnh kỳ vọng và nhu cầu của họ trong những lần tiếp xúc tiếp theo.

Trong bối cảnh truyền thông hiện đại, lý thuyết sử dụng và hài lòng không chỉ là công cụ lý luận quan trọng mà còn có ý nghĩa thực tiễn trong việc định hướng hoạt động truyền thông. Hiện tượng sử dụng từ trend trong tiêu đề báo mạng điện tử không đơn thuần là chiến lược thu hút sự chú ý của công chúng mà còn phản ánh nỗ lực của các tòa soạn trong việc đáp ứng nhu cầu thông tin, giải trí và tương tác xã hội của độc giả. Lý thuyết này giúp khám phá phản hồi của công chúng đối với thực trạng trên, đồng thời đưa ra các khuyến nghị nhằm điều chỉnh cách thức truyền tải và nội dung thông tin. Qua đó, không chỉ nâng cao hiệu quả truyền thông mà còn từng bước gia tăng mức độ hài lòng của công chúng đối với nội dung được truyền tải.

3. Mối quan hệ giữa mạng xã hội và báo chí hiện đại

Trong bối cảnh truyền thông hiện đại, công chúng ngày càng phân mảnh, mạng xã hội đã trở thành kênh truyền thông chính yếu bên cạnh các phương tiện truyền thông truyền thống. Theo báo cáo của We Are Social 2024, số lượng người dùng mạng xã hội trên toàn cầu đạt 5,04 tỉ người, tăng 5,6% so với năm 2023 [9]. Sự thay đổi mạnh mẽ trong thói quen tiêu thụ tin tức của công chúng đã được nhiều học giả ghi nhận, với xu hướng chuyển dịch từ báo in và truyền hình sang các nền tảng trực tuyến [10], [11], [12], [13]. Đặc biệt, đại dịch Covid-19 được xem là một trong những tác nhân thúc đẩy quá trình này. Tại Việt Nam, nghiên cứu Tran (2022) chỉ ra rằng, 61% người tham gia khảo sát tiêu thụ tin tức thông qua các trang thông tin điện tử và nền tảng internet, trong khi chỉ 39% sử dụng báo mạng điện tử. Hai nguồn thông tin phổ biến nhất bao gồm các trang thông tin điện tử như Baomoi, Ybox.vn, Kenh14.vn và mạng xã hội như Facebook, Zalo, Youtube. Tương tự, ở các quốc gia châu Âu, nghiên cứu của Van Aelst et al. (2021) cho thấy công chúng ngày càng phụ thuộc vào mạng xã hội và internet để tiếp cận thông tin.

Mặt khác, mạng xã hội không chỉ cung cấp tin tức mà còn đóng vai trò quan trọng trong việc lan truyền thông tin một cách nhanh chóng. Với tính tức thời và khả năng lan tỏa rộng khắp, mạng xã hội giúp công chúng tiếp cận và chia sẻ thông tin dễ dàng hơn bao giờ hết [15]. Đặc biệt, nghiên cứu của Asur et al. (2021) đã chứng minh rằng các xu hướng (trend) trên mạng xã hội thường được tạo bởi sự cộng hưởng của nội dung với người dùng. Trong đó, nhiều xu hướng có nguồn gốc từ báo chí truyền thống, còn mạng xã hội đóng vai trò như một "bộ khuếch đại có chọn lọc", giúp các xu hướng lan rộng hơn. Như vậy, có thể thấy rằng, mạng xã hội vừa

là nơi khởi nguồn vừa là chất xúc tác quan trọng trong việc định hình và khuếch đại các xu hướng thông tin.

Mối quan hệ báo chí và mạng xã hội là một mối quan hệ cộng sinh, mạng xã hội vừa là nguồn tin vừa là kênh truyền tải hiệu quả của báo chí. Paulussen & Harder (2014) nhấn mạnh rằng, việc theo dõi mạng xã hội đã trở thành thói quen hàng ngày của các nhà báo để tìm kiếm nội dung và xu hướng mới. Mạng xã hội không chỉ giúp các nhà báo cập nhật thông tin mà còn hỗ trợ kết nối với công chúng, thu thập và xác thực thông tin [18]. Nhiều chủ đề và xu hướng xuất phát từ mạng xã hội đã được báo chí khai thác thành công, thu hút sự quan tâm từ công chúng.

Tuy nhiên, sử dụng mạng xã hội làm nguồn tin cũng mang lại nhiều rủi ro. McGregor (2019); Paulussen & Harder (2014); Weaver & Willnat (2016) cảnh báo về nguy cơ thông tin thiếu xác thực trên mạng xã hội, dẫn đến xung đột giữa tính chính xác và tốc độ đưa tin. Bên cạnh đó, sự phụ thuộc vào mạng xã hội cũng có thể làm giảm tính độc lập và chuyên nghiệp của báo chí.

Truyền thông xã hội không chỉ định hình cách báo chí tiếp cận công chúng mà còn ảnh hưởng đến cấu trúc và phong cách nội dung. Nghiêm Tuấn Anh (2024) cho rằng, báo chí hiện đại đã phải thay đổi để đáp ứng thói quen tiếp cận thông tin nhanh chóng của công chúng. Tiêu đề bài viết ngắn gọn, hấp dẫn hơn; nội dung được trình bày trực quan với sự tích hợp các hình thức đa phương tiện như audio, video, đồ họa, hoạt họa, v.v.

Ngoài ra, các nền tảng truyền thông mới cũng thúc đẩy sự tương tác theo theo gian thực. Báo chí không còn là kênh một chiều mà trở thành không gian truyền thông hai chiều, nơi công chúng có thể đánh giá bài viết qua biểu tượng cảm xúc, bình luận, chia sẻ, hoặc xếp hạng chất lượng bài viết [21], [22].

Như vậy, mạng xã hội đã và đang định hình lại báo chí hiện đại, không chỉ trong cách sản xuất mà còn ở cách báo chí tương tác với công chúng. Điều này đòi hỏi báo chí phải cân bằng giữa việc chạy theo xu hướng và duy trì các giá trị cốt lõi như tính chính xác, khách quan và chuyên nghiệp.

4. Vai trò của công chúng đối với truyền thông mạng xã hội và báo chí

Sự thay đổi thói quen tiêu thụ thông tin không chỉ định hình hành vi của công chúng mà còn đặt ra những câu hỏi quan trọng đối với các học giả truyền thông và tòa soạn báo chí: Trong bối cảnh mới này, công chúng hiện đại đóng vai trò gì đối với truyền thông xã hội và báo chí? Liệu họ có đơn thuần tiếp nhận thông tin, hay đã trở thành những chủ thể chủ động tham gia vào quá trình sản xuất và lan tỏa thông tin? Những câu hỏi này không chỉ mở ra góc nhìn sâu sắc về sự chuyển đổi trong mối quan hệ giữa báo chí và công chúng mà còn gợi ý về sự thay đổi mà nhà báo cần tiếp nhận, lắng nghe và tương tác với độc giả của mình.

Trong bối cảnh truyền thông xã hội phát triển mạnh mẽ, công chúng ngày càng chủ động trong việc lựa chọn thông tin tiếp nhận. Tuy nhiên, song song với sự chủ động đó, một xu hướng tiếp nhận tin tức khác đang nổi lên: sự tình cờ. Hiện tượng này được khái niệm hóa là "tiếp xúc tin tức ngẫu nhiên" (incidental news exposure - INE). Theo các nghiên cứu của

Antunovic et al. (2018); letcher & Nielsen (2018); Kim et al. (2013), INE xảy ra khi người dùng gặp tin tức một cách tình cờ trong lúc trực tuyến vì các mục đích khác, chẳng hạn như giải trí hay kết nối xã hội. Boczkowski et al. (2017) cũng nhận định rằng công chúng thường tiếp cận tin tức rời rạc, không theo một hệ thống cụ thể và dành rất ít thời gian cho từng mẩu tin. Điều này đặt ra thách thức lớn đối với nhà báo trong việc thu hút và giữ chân độc giả, đồng thời làm nổi bật tầm quan trọng của việc tối ưu hóa nội dung và tiêu đề phù hợp với thói quen tiếp nhận thông tin nhanh của công chúng hiện đại.

Sự trỗi dậy của các nền tảng mạng xã hội đã thúc đẩy sự phát triển mạnh mẽ của báo chí công dân (citizen journalism), cho phép công chúng tham gia trực tiếp vào quá trình sản xuất và lan tỏa tin tức. Báo chí công dân trở thành nguồn bổ trợ quan trọng cho báo chí chính thống (mainstream media), cung cấp thêm góc nhìn đa dạng và những thông tin kịp thời từ hiện nơi nhà báo chuyên nghiệp không sẵn có [27], [28], [29]. Tuy nhiên, dù đóng góp đáng kể vào dòng chảy thông tin và lan tỏa câu chuyên mới lạ trên mạng xã hội, báo chí công dân vẫn không thể thay thế được vai trò của báo chí chính thống.

Các học giả đã chỉ ra một số hạn chế đáng lo ngại về xu hướng này. Tính công bằng và cân bằng trong các sản phẩm tin tức từ báo chí công dân thường không được đảm bảo, do phần lớn các nguồn tin phụ thuộc vào thông tin trên internet hoặc những mối quan hệ cá nhân [27]. Bên cạnh đó, nguy cơ bóp béo thông tin, thiếu kiểm chứng và trình bày mang tính dẫn dắt, thậm chí tác động tiêu cực đến quan điểm công chúng cũng là vấn đề cần được quan tâm [30]. Điều này đặt ra yêu cầu về việc nâng cao nhận thức và khả năng đánh giá thông tin của người dùng trong bối cảnh truyền thông phức tạp.

Vai trò của công chúng truyền thông ngày nay không chỉ dừng lại ở việc tiêu thụ thông tin mà còn mở rộng sang việc lan tỏa các thông điệp truyền thông. Al-Rawi (2019) nhấn mạnh rằng công chúng đóng vai trò quan trọng trong việc tạo ra tin tức lan truyền (viral news) - những nội dung nổi bật thu hút sự chú ý giữa vô vàn thông tin khác. Vậy điều gì thúc đẩy công chúng tương tác mạnh mẽ hơn với tin tức trực tuyến?

Berger (2011) cho rằng cảm xúc là yếu tố then chốt trong việc thú đẩy sự quan tâm và hành vi chia sẻ tin tức. Nội dung khơi gợi những cảm xúc mãnh liệt như kinh ngạc (tích cực), tức giận hoặc lo lắng (tiêu cực), thường lan truyền nhanh hơn so với nội dung khác [33]. Không chỉ giới hạn trong tin tức, các sản phẩm truyền thông khác cũng khai thác mối quan hệ giữa cảm xúc và hành vi chia sẻ. Golan & Zaidner (2008) phân tích 360 mẫu quảng cáo lan truyền và phát hiện rằng các chủ đề như hài hước (91%), tình dục (28,1%) và bạo lực (14,4%) có khả năng thúc đẩy hành vi chia sẻ mạnh mẽ.

Ngoài cảm xúc, mối liên hệ cá nhân cũng đóng vai trò quan trọng. Karnowski et al. (2017) cho thấy rằng khi nội dung tin tức phù hợp với sở thích hoặc mối quan tâm cá nhân, công chúng sẽ dành nhiều thời gian tương tác hơn, kể cả khi tiếp xúc ban đầu chỉ là ngẫu nhiên. Hơn nữa, tin tức được chia sẻ bởi những người thân thiết trong mạng lưới xã hội thường dễ thu hút sự chú ý của công chúng hơn cả [35], [36] .

Động lực chia sẻ tin tức cũng cần được xem xét. Lee et al. (2011) khẳng định giá trị thông tin là yếu tố quan trọng nhất thúc đẩy ý định chia sẻ, trong khí đó, yếu tố giải trí không

có tác động đáng kể. Điều này cho thấy, để tăng cường tương tác, các cơ quan báo chí cần chú trọng vào việc cung cấp nội dung không chỉ hấp dẫn mà còn có giá trị thông tin cao, đáp ứng nhu cầu của công chúng.

Việc sử dụng từ trend trong tiêu đề báo mạng điện tử không chỉ nhằm thu hút sự chú ý mà còn đánh trúng vào yếu tố cảm xúc và kết nối cá nhân của công chúng. Ngôn ngữ trend phản ánh sự vận động tất yếu của ngôn ngữ, tạo ra những lớp từ mới được sử dụng rộng rãi trong một cộng đồng nhất định. Khi các cụm từ này xuất hiện trên báo chí, đặc biệt là trong tiêu đề - yếu tố đầu tiên tiếp xúc với độc giả - chúng có khả năng khơi gợi sự tò mò, hứng thú và cả cảm giác gần gũi. Tiêu đề hấp dẫn không chỉ quyết định liệu công chúng có tiếp tục đọc bài báo hay không mà còn tạo ra sự kết nối tức thì với trải nghiệm và ngôn ngữ đời sống thường ngày của họ. Do đó, việc bận dụng từ trend trong đặt tiêu đề không chỉ là một chiến thuật thu hút sự chú ý, mà còn là xu hướng hợp lý và kịp thời, giúp báo chí nhịp bước cùng sự thay đổi trong cách công chúng tiếp nhận và phản hồi thông tin.

III. CÂU HỎI NGHIÊN CỨU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

1. Câu hỏi nghiên cứu

Mặc dù phương thức cộng sinh giữa báo chí và mạng xã hội, cũng như vai trò của công chúng đối với tin tức trên môi trường số đã được thảo luận nhưng các nghiên cứu về việc sử dụng ngôn ngữ trend trong tiêu đề báo chí vẫn còn hạn chế, đặc biệt trong bối cảnh Việt Nam. Vì vậy, nghiên cứu này mang tính khám phá nhằm khảo sát mức độ chấp nhận và hài lòng của công chúng đối với xu hướng sử dụng từ trend trong tiêu đề báo mạng điện tử, đồng thời xem xét các yếu tố ảnh hưởng đến mức độ chấp nhận và hài lòng đó. Để đạt được mục tiêu nghiên cứu, bài báo đề xuất hai câu hỏi nghiên cứu chính:

RQ1: Công chúng có mức độ chấp nhận và hài lòng như thế nào đối với việc sử dụng từ trend trong tiêu đề báo mạng điện tử?

RQ2: Các yếu tố như đặc điểm nhân khẩu học, thói quen sử dụng mạng xã hội và tần suất đọc báo mạng điện tử có tạo ra sự khác biệt đáng kể trong mức độ chấp nhận và hài lòng hay không?

2. Phương pháp nghiên cứu

Nghiên cứu này nhằm khảo sát mức độ chấp nhận và hài lòng của công chúng đối với việc sử dụng từ trend trong tiêu đề báo mạng điện tử. Phương pháp nghiên cứu định lượng được áp dụng, cụ thể với kỹ thuật điều tra bản hỏi là công cụ nghiên cứu chính. Khảo sát được triển khai dưới hai hình thức trực tiếp và trực tuyến trong khoảng thời gian từ tháng 09/2024 đến tháng 10/2024 với chiến lược chọn mẫu ngẫu nhiên thuận tiện. Dung lượng mẫu là 374 đáp viên. Dữ liệu được xử lý và phân tích bằng phần mềm JASP phiên bản 0.18.3.0.

Bộ câu hỏi được xây dựng dựa trên lý thuyết sử dụng và hài lòng (Uses and Gratifications Theory), với bốn nhân tố chính: **Tiếp nhận (access), thái độ tích cực (positive attitude), thái độ tiêu cực (negative attitude) và tương tác (interaction)**. Các thang đo được

thiết kế theo thang Likert 5, trong đó: từ 1 = Rất không đồng ý đến 5 = Rất đồng ý. Chi tiết các thang đo cụ thể như sau:

- Tiếp nhận: gồm 5 biến quan sát, ví dụ: "tôi cảm thấy mình bị thu hút bởi các tiêu đề có sử dụng từ trend trên báo mạng điện tử".

- Thái độ tích cực: gồm 3 biến quan sát, ví dụ: "tôi cảm thấy hứng thú khi đọc tiêu đề trên báo mạng điện tử sử dụng từ trend".

- Thái độ tiêu cực: gồm 3 biến quan sát, ví dụ "tôi nhận thấy việc sử dụng từ trend trong các tiêu đề báo mạng điện tử khiến tôi khó chịu".

- Tương tác: gồm 5 biến quan sát, ví dụ, "tôi thường để lại bình luận dưới các bài báo có sử dụng từ trend trong tiêu đề".

Ngoài ra, nghiên cứu cũng thu thập một số thông tin nhân khẩu học, thói quen sử dụng mạng xã hội và tần suất đọc báo mạng điện tử của đáp viên nhằm kiểm tra mối quan hệ giữa các yếu tố này với mức độ chấp nhận và hài lòng của công chúng.

Trước khi tiến hành khảo sát chính thức, bộ câu hỏi được thử nghiệm với 30 khách thể ngẫu nhiên nhằm kiểm tra độ tin cậy bằng hệ số Cronbach's Alpha. Kết quả cho thấy, tất cả các thang đo đều có hệ số Cronbach's Alpha > 0,7 và các biến quan sát đều đạt giá trị Cronbach's Alpha > 0,3. Điều này khẳng định rằng các thang đo và biến quan sát đều có độ tin cậy cao, đảm bảo tính chính xác cho nghiên cứu. Kết quả chi tiết về hệ số Cronbach's Alpha được trình bày trong bảng dưới đây:

Bảng 2. *Kết quả kiểm định Cronbach's Alpha*

Thang đo	**Cronbach's Alpha**
Tiếp nhận	0,909
Thái độ tích cực	0,869
Thái độ tiêu cực	0,910
Tương tác	0,875

IV. KẾT QUẢ NGHIÊN CỨU

1. Đặc điểm của đối tượng nghiên cứu

Theo trình bày ở phần phương pháp nghiên cứu, số lượng mẫu khách thể được đưa vào nghiên cứu chính thức là 374 mẫu. Cơ cấu mẫu được thể hiện qua bảng số liệu dưới đây:

Bảng 3. *Đặc điểm khách thể nghiên cứu*

Phân loại	**Đặc điểm**	**Số lượng (N)**	**Tỷ lệ (%)**
Giới tính	Nam	109	29,1
	Nữ	265	70,9

Tuổi	Từ 18 đến 25 tuổi	350	93,6
	Từ 26 đến 35 tuổi	12	3,2
	Từ 36 đến 45 tuổi	12	3,2
Nghề nghiệp	Lao động có trình độ	31	8,3
	Bộ đội, công an, nhân viên quốc phòng	1	0,3
	Chủ doanh nghiệp, hộ kinh doanh	6	1,6
	Lãnh đạo, quản lý	7	1,9
	Nông/ Lâm/ Ngư nghiệp	1	0,3
	Nhân viên thời vụ/ Bán thời gian	2	0,5
	Học sinh/ Sinh viên	326	87,2
Trình độ học vấn	Thấp hơn THPT	3	0,8
	Tốt nghiệp THPT	88	23,5
	Trung cấp/ Đào tạo nghề	1	0,3
	Cao đẳng/ Đại học	257	68,7
	Sau đại học	21	5,6
Tần suất đọc báo mạng điện tử	Không chắc chắn	4	1,1
	Không bao giờ	4	1,1
	Hiếm khi	36	9,6
	Thỉnh thoảng	174	46,5
	Thường xuyên	143	38,2
	Luôn luôn	17	4,5
Thời gian sử dụng mạng xã hội	Dưới 1 giờ/ngày	31	8,3
	1 giờ đến dưới 3 giờ/ ngày	94	25,1
	3 giờ đến dưới 5 giờ/ ngày	117	31,3
	5 giờ đến dưới 8 giờ/ ngày	96	25,7
	Trên 8 giờ/ ngày	36	9,6

Dữ liệu ***bảng 3*** cho thấy sự phân bố đặc điểm nhân khẩu học và hành vi sử dụng mạng xã hội của các đáp viên rất đa dạng. Về giới tính, nữ chiếm đa số với 70,9%, trong khi nam giới chỉ chiếm 29,1%. Độ tuổi tham gia khảo sát chủ yếu nằm trong khoảng từ 18 đến 25 tuổi

(93,6%), phản ánh nhóm dân số trẻ là đối tượng chính, trong khi các nhóm tuổi từ 26 đến 35 và từ 36 đến 45 chỉ chiếm 3,2% mỗi nhóm. Nghề nghiệp của đáp viên tuy đa dạng nhưng tập trung chủ yếu vào học sinh/sinh viên, chiếm 87,2%, phù hợp với xu hướng người trẻ sử dụng mạng xã hội thường xuyên. Xét về trình độ học vấn, phần lớn đáp viên đã tốt nghiệp THPT trở lên, trong đó 68,7% có trình độ cao đẳng hoặc đại học, cho thấy nhóm khảo sát có mức độ học vấn tương đối cao. Đáng chú ý, 89,2% số người tham gia tự đánh giá mình có thói quen đọc tin tức từ báo mạng điện tử ở mức từ thỉnh thoảng trở lên. Về thời gian sử dụng mạng xã hội, nhóm sử dụng từ 3 đến 5 giờ mỗi ngày chiếm tỷ lệ lớn nhất (31,3%), tiếp theo là nhóm sử dụng từ 5 đến dưới 8 giờ mỗi ngày (25,7%). Như vậy, mẫu khảo sát mang tính đa dạng về đặc điểm nhân khẩu học, hành vi sử dụng mạng xã hội cũng như tần suất đọc báo mạng điện tử. Điều này tạo cơ sở để nghiên cứu phản ánh các xu hướng khác nhau của vấn đề nghiên cứu, đồng thời kỳ vọng mang lại những kết quả phong phú và toàn diện.

2. Mức độ chấp nhận và hài lòng của công chúng về việc sử dụng từ trend trong tiêu đề báo mạng điện tử

Để trả lời cho ***RQ1: Công chúng có mức độ chấp nhận và hài lòng như thế nào đối với việc sử dụng từ trend trong tiêu đề báo mạng điện tử?*** Sau khi thu thập dữ liệu, nghiên cứu tiến hành nhập liệu, xử lý bằng phần mềm JASP phiên bản 0.18.3.0 và đưa ra số liệu tổng quan về phản hồi của công chúng đối với thực trạng sử dụng từ trend trên báo mạng điện tử. Bằng việc thiết lập nhân số đại diện (trung bình cộng) cho mỗi yếu tố và chạy thống kê mô tả, kết quả được thể hiện qua bảng sau:

Bảng 4. *Mức độ chấp nhận và hài lòng của công chúng về việc sử dụng từ trend trong tiêu đề báo mạng điện tử*

Biến	**M**	**SD**	**SE**
Tiếp nhận	3,64	0,885	0,046
Thái độ tích cực	3,51	0,945	0,049
Thái độ tiêu cực	3,05	1,094	0,057
Tương tác	3,19	0,956	0,049

Việc sử dụng từ trend mạng xã hội trên các tiêu đề báo mạng điện tử nhận được phản hồi đa dạng từ phía công chúng. Cụ thể, mức độ tiếp nhận được đánh giá cao với giá trị trung bình 3,64. Thái độ tích cực của công chúng đối với việc sử dụng từ trend đạt giá trị trung bình 3,51, cho thấy sự ủng hộ vừa phải, trong khi thái độ tiêu cực có giá trị trung bình thấp hơn, ở mức 3,05, biểu thị mức độ phản đối không quá mạnh. Mức độ tương tác của công chúng được ghi nhận ở mức trung bình, với giá trị trung bình 3,2, cho thấy sự tham gia của công chúng đối với các nội dung này ở mức ổn định nhưng chưa nổi bật. Những kết quả này gợi ý rằng, việc sử dụng từ trend trên các tiêu đề báo mạng điện tử có tiềm năng thu hút sự quan tâm, song vẫn cần cân nhắc để tối ưu hóa sự tương tác và tránh phản ứng tiêu cực từ một bộ phận công chúng.

Theo Lê (2021), ngôn ngữ hot trend là sản phẩm của nhóm công chúng trẻ và là một hiện tượng ngôn ngữ tự nhiên, phản ánh quy luật phát triển của ngôn ngữ. Nhờ vậy, việc sử

dụng từ trend đã góp phần làm phong phú thêm vốn ngôn ngữ, làm cho ngôn ngữ mang hơi thở hiện đại và đa dạng, không đơn điệu, đồng thời tạo ra sự hấp dẫn trong quá trình giao tiếp. Nhà báo "bắt trend" với mong muốn thu hút sự chú ý và tăng cường tương tác của công chúng với tin tức. Tuy nhiên, từ kết quả nghiên cứu chỉ ra rằng mức độ tương tác của độc giả đối với tác phẩm báo chí sử dụng từ trend vẫn chưa được đạt như kỳ vọng, đồng thời tồn tại những thái độ tiêu cực. Để lý giải rõ hơn về những phản ứng này cần áp dụng các phương pháp nghiên cứu định tính như phỏng vấn sâu hoặc quan sát tham gia đối với cả công chúng và người làm báo nhằm hiểu rõ động cơ và thái độ của hai bên.

Từ góc nhìn chủ quan, một trong những nguyên nhân chính dẫn đến phản ứng tiêu cực và sự e ngại trong tương tác trực tuyến là nỗi lo về dấu chân kỹ thuật số [38]. Công chúng ngày càng quan tâm đến quyền riêng tử và bảo mật thông tin cá nhân. Khi tham gia thảo luận, người dùng không chỉ để lại thông tin cá nhân mà còn có thể đối mặt với nguy cơ bị giám sát hay lạm dụng thông tin. Mặt khác, môi trường trực tuyến luôn đối diện với nguy cơ bị tấn công, đặc biệt đối với những người có quan điểm trái chiều, điều này tạo nên bức tường vô hình cản trở sự tham gia thảo luận của công chúng đối với bài báo sử dụng từ trend.

Nhằm khám phá mối quan hệ giữa các biến độc lập với biến phụ thuộc, tác giả đã tiến hành phân tích tương quan (Correlation analysis). Theo Field (2009) phương pháp này sử dụng chỉ số Pearson's r, dao động từ -1 đến 1, để đánh giá mức độ tương quan tuyến tính giữa các cặp biến. Khi giá trị r tiến gần đến 1 hoặc -1, mối quan hệ tuyến tính giữa các biến càng mạnh, với r = 1 biểu thị tương quan dương hoàn hảo và r = -1 biểu thị tương quan âm hoàn hảo. Ngược lại, giá trị r gần 0 cho thấy mối quan hệ tuyến tính yếu hoặc không có. Mặt khác, giá trị p-value cũng được xem xét để xác định ý nghĩa thống kê của mối quan hệ. Nếu p-value < 0,05, mối tương quan tuyến tính giữa các cặp biến được coi là có ý nghĩa và ngược lại. Kết quả phân tích tương quan chi tiết được trình bày trong bảng dưới đây:

Bảng 5. *Tương quan giữa sự chấp nhận và hài lòng và các yếu tố đối với việc sử dụng từ trend trong tiêu đề báo mạng điện tử*

Biến		**1**	**2**	**3**	**4**	**5**
1	Pearson's r	-				
	p-value	-				
2	Pearson's r	0,836	-			
	p-value	<0,001	-			
3	Pearson's r	0,232	0,272	-		
	p-value	<0,001	<0,001	-		
4	Pearson's r	0,646	0,708	0,421	-	
	p-value	<0,001	<0,001	<0,001	-	
5	Pearson's r	0,828	0,865	0,637	0,865	-

	p-value	<0,001	<0,001	<0,001	<0,001	-
Ghi chú: 1: Tiếp nhận, 2: Thái độ tích cực, 3: Thái độ tiêu cực, 4: Tương tác, 5: Sự chấp nhận và hài lòng						

Kết quả phân tích tương quan Pearson cung cấp những hiểu biết sâu sắc về mối quan hệ giữa các biến độc lập (tiếp nhận, thái độ tích cực, thái độ tiêu cực và tương tác) và biến phụ thuộc (sự chấp nhận và hài lòng). Chỉ số tương quan dao động trong khoản từ 0,232 đến 0,865, thể hiện mức độ tương quan từ yếu đến rất mạnh. Trong đó, biến "tiếp nhận" cho thấy mối tương quan rất mạnh với "thái độ tích cực" ($r = 0,836$, $p < 0,001$) và "sự chấp nhận và hài lòng" ($r = 0,828$, $p < 0,001$). Điều này chứng tỏ rằng mức độ tiếp nhận thông tin chứa từ trend trên các tiêu đề báo có ảnh hưởng đáng kể đến việc hình thành thái độ tích cực và mức độ chấp nhận, hài lòng của công chúng. Tương tự, "thái độ tích cực" có mối tương quan rất mạnh với "sự chấp nhận và hài lòng" ($r = 0,865$, $p < 0,001$), cho thấy yếu tố này đóng vai trò quan trọng trong việc định hình cảm nhận của công chúng đối với việc sử dụng từ trend.

Ngoài ra, biến "tương tác" cũng thể hiện mối quan hệ mạnh mẽ với "tiếp nhận" ($r = 0,646$, $p < 0,001$), "thái độ tích cực" ($r = 0,708$, $p < 0,001$) và "sự chấp nhận và hài lòng" ($r = 0,865$, $p < 0,001$). Điều này phản ánh rằng mức độ tương tác cao thường gắn liền với khả năng tiếp nhận và thái độ tích cực của công chúng đối với các bài báo có sử dụng từ trend. Tuy nhiên, thái độ tiêu cực lại chỉ có mối tương quan yếu với tiếp nhận ($r = 0,232$, $p < 0,001$) và thái độ tích cực ($r = 0,272$, $p < 0,001$). Kết quả này cho thấy, mặc dù vẫn tồn tại một bộ phận công chúng có thái độ tiêu cực đối với việc sử dụng từ trend trên tiêu đề báo, nhưng mức độ ảnh hưởng của thái độ tiêu cực là không đáng kể. Điều này hàm ý rằng thái độ tiêu cực không phải là yếu tố chính ảnh hưởng đến sự tiếp nhận hoặc thái độ tích cực của công chúng.

Phân tích mô tả và tương quan chứng minh rằng, công chúng có xu hướng tiếp nhận tích cực đối với việc sử dụng từ trend trong tiêu đề báo mạng điện tử, với mức độ chấp nhận và hài lòng ở mức trung bình. Thái độ tích cực chiếm ưu thế, trong khi phản ứng tiêu cực không quá mạnh mẽ. Tuy nhiên, mức độ tương tác còn khá hạn chế, đặc biệt là các tương tác sâu như bình luận hay thảo luận. Mối tương quan chặt chẽ giữa sự tiếp nhận, thái độ tích cực và mức độ hài lòng cho thấy rằng từ trend giúp thu hút sự chú ý, nhưng chưa tạo được sự kết nối bền vững với độc giả. Điều này đặt ra yêu cầu cho nhà báo, phóng viên trong việc sử dụng từ trend một cách hợp lý để vừa thu hút công chúng, vừa tăng cường chất lượng nội dung và sự gắn kết.

3. Các yếu tố tạo ra sự khác biệt về mức độ chấp nhận và hài lòng trong việc sử dụng từ trend trong tiêu đề báo mạng điện tử

Để trả lời cho RQ2: Các yếu tố như đặc điểm nhân khẩu học, thói quen sử dụng mạng xã hội và tần suất đọc báo mạng điện tử có tạo ra sự khác biệt đáng kể trong mức độ chấp nhận và hài lòng hay không?, nghiên cứu đã thực hiện kiểm định Independent Sample T-Test và One-way ANOVA đối với các biến nhân khẩu học, thói quen sử dụng mạng xã hội và thói quen đọc báo mạng điện tử. So sánh giá trị trung bình giữa các biến được xem là khác biệt có ý nghĩa thông kê khi mức độ sai số cho phép nhỏ hơn 5% ($p < 0,05$)

Kết quả từ các kiểm định cho thấy không tồn tại sự khác biệt mang ý nghĩa thống kê giữa các nhóm khách thể theo đặc điểm nhân khẩu học như giới tính, độ tuổi, nghề nghiệp, trình độ học vấn hay thói quen sử dụng mạng xã hội về sự chấp nhận và hài lòng trong việc sử dụng từ trend trong tiêu đề báo mạng điện tử ($p > 0,05$). Điều đó chứng tỏ rằng, dù giới tính, độ tuổi, nghề nghiệp, trình độ học vấn hay thói quen sử dụng mạng xã hội như thế nào thì mức độ chấp nhận và hài lòng là như nhau. Tuy nhiên, tồn tại sự khác biệt mang ý nghĩa thống kê giữa các nhóm khách thể với tần suất đọc báo mạng điện tử khác nhau như sau:

Bảng 6. *Sự khác biệt về mức độ chấp nhận và hài lòng giữa các nhóm khách thể có tần suất đọc báo mạng điện tử khác nhau*

Biến	Tổng bình phương	Bậc tự do	Trung bình bình phương	F	p
Tiếp nhận	18,808	4	4,702	6,350	<0,001
Thái độ tích cực	16,035	4	4,009	4,663	0,001
Thái độ tiêu cực	22,519	4	5,630	4,897	<0,001
Tương tác	28,755	4	7,189	8,505	<0,001

Bảng 6 cho thấy có sự khác biệt đáng kể về mức độ chấp nhận và hài lòng giữa các nhóm khách thể với tần suất đọc báo mạng điện tử khác nhau. Cụ thể, đối với biến "tiếp nhận", giá trị F là 6,350 và $p < 0,001$, chứng tỏ sự khác biệt rõ rệt về khả năng tiếp nhận thông tin giữa các nhóm, qua đó ảnh hưởng đến mức độ chấp nhận và hài lòng đối với việc sử dụng từ trend trên các tiêu đề báo. Tương tự, biến "thái độ tích cực" có giá trị F là 4,009 với $p = 0,001$, khẳng định rằng những nhóm khách thể có tần suất đọc báo khác nhau cũng biểu hiện sự khác biệt ý nghĩa trong cách nhìn nhận tích cực về các tiêu đề báo chứa từ trend. Đối với biến "thái độ tiêu cực", giá trị F đạt 4,897 với $p < 0,001$, điều này cho thấy thái độ tiêu cực của công chúng đối với các tiêu đề báo có chứa từ trend cũng thay đổi đáng kể giữa các nhóm. Đặc biệt, biến "tương tác" có giá trị F cao nhất, đạt 8,505 với $p < 0,001$, cho thấy mức độ tương tác với các bài báo mạng có sử dụng từ trend khác nhau đáng kể tùy thuộc vào tần suất đọc báo.

Như vậy, kết quả nghiên cứu khẳng định rằng tần suất đọc báo mạng điện tử đóng vai trò quan trọng trong việc định hình thái độ và hành vi của công chúng đối với các tiêu đề sử dụng từ trend. Những người đọc báo thường xuyên có xu hướng chấp nhận, hài lòng và tương tác tích cực hơn, trong khi những người ít tiếp xúc với báo mạng có thể giữ thái độ thận trọng hoặc tiêu cực. Điều này gợi ý rằng các tòa soạn cần cân nhắc đến đối tượng công chúng mục tiêu khi áp dụng từ trend vào tiêu đề nhằm đảm bảo giá trị của báo chí chính thống.

V. THẢO LUẬN VÀ KẾT LUẬN

Ngôn ngữ hot trend hay từ trend trong báo chí là một hiện tượng mới cho đến nay vẫn chưa có nhiều nghiên cứu chuyên sâu về vấn đề này. Nghiên cứu này có thể được xem là một trong những nghiên cứu đầu tiên, mang tính chỉ dấu, nhằm mô tả phản ứng của công chúng về thực trạng sử dụng từ trend mạng xã hội trong tiêu đề báo mạng điện tử.

Kết quả nghiên cứu cho thấy, mức độ chấp nhận và hài lòng của công chúng ở mức trung bình. Trong đó, mức độ tiếp nhận có giá trị trung bình cao nhất (M = 3,64), tuy nhiên, mức độ tương tác vẫn còn hạn chế (M = 3,19), đặc biệt các tương tác sâu như bình luận hoặc thảo luận. Ngoài ra, một bộ phận công chúng thể hiện phản hồi tiêu cực đối với việc sử dụng từ trend trong tiêu đề báo mạng điện tử (M = 3,05). Đồng thời, nghiên cứu này cũng khám phá ra rằng tần suất đọc báo mạng điện tử đóng vai trò quan trọng, tạo nên sự khác biệt rõ rệt về mức độ chấp nhận và hài lòng của công chúng.

Trong mối quan hệ cộng sinh giữa báo chí và mạng xã hội, báo chí không chỉ sử dụng mạng xã hội như một kênh truyền tải thông tin hay nguồn khai thác tin tức mà còn khai thác chất liệu ngôn ngữ, với từ trend là một điển hình. Từ trend trong báo chí hiện đại đóng vai trò kích thích sự tò mò, tạo hứng thú và kết nối với công chúng. Báo chí thường sử dụng từ trend trong tiêu đề – phần quan trọng nhất, quyết định sự chú ý và quyết định công chúng có tiếp tục đọc hay không. Điều này cho thấy, từ trend không chỉ giúp tiếp cận công chúng mục tiêu hiệu quả mà còn thu hút và giữ chân độc giả đối với các tác phẩm báo chí. Bên cạnh đó, việc sử dụng từ trend còn góp phần tối ưu hóa công cụ tìm kiếm (SEO - Search Engine Optimization), không chỉ từ các kho dữ liệu đại chúng như Google, Cốc Cốc hay Bing mà còn trên các nền tảng mạng xã hội. Khi công chúng tìm kiếm một từ trend trên các công cụ này, các bài báo chứa từ trend sẽ dễ dàng xuất hiện, giúp tăng khả năng tiếp cận. Như vậy, từ trend không chỉ là công cụ ngôn ngữ mà còn là chiến lược hiệu quả để báo chí kết nối với công chúng.

Mức độ tương tác của công chúng đối với các tác phẩm báo chí sử dụng từ trend vẫn còn hạn chế, đặc biệt là các tương tác sâu như bình luận hay tham gia thảo luận. Tuy vậy, các từ trend được sử dụng trên báo chí vẫn được công chúng lan tỏa và áp dụng trong quá trình giao tiếp hàng ngày. Trong bối cảnh truyền thông đa dạng và phức tạp như hiện nay, trong một ngày công chúng tiếp nhận một lượng thông tin khổng lồ từ thế giới thực lẫn môi trường số. Điều này có thể dẫn đến hiện tượng bão hòa thông tin, người đọc chỉ dừng ở việc đọc thay vì dành thời gian tham gia thảo luận hay bình luận sâu sắc. Bên cạnh đó, các lo ngại về quyền riêng tử và sự an toàn trực tuyến cũng là yếu tố đáng lưu ý. Công chúng hiện nay rất quan tâm đến việc bảo vệ thông tin cá nhân trên mô trường số, lo ngại việc lộ lọt thông tin hay theo dõi. Chính những yếu tố này khiến họ e dè hơn khi tham gia bình luận hay thảo luận công khai.

Một phát hiện thú vị của nghiên cứu là các yếu tố nhân khẩu học và thời gian sử dụng mạng xã hội không tạo ra sự khác biệt có ý nghĩa thống kê về sự chấp nhận và hài lòng của công chúng đối với các tiêu đề sử dụng từ trend mạng xã hội. Dù nghiên cứu kỳ vọng rằng việc sử dụng và tương tác với mạng xã hội nhiều sẽ giúp công chúng trở nên quen thuộc với từ trend và từ đó dễ dàng chấp nhận nó trong tiêu đề bài báo, nhưng kết quả lại cho thấy điều này không đúng. Yếu tố duy nhất tạo ra sự khác biệt là tần suất đọc báo mạng điện tử. Từ chỉ dấu này có thể thấy mức độ quen thuộc với từ trend không phải là yếu tố quyết định trong việc công chúng chấp nhận sự xuất hiện của nó trong tiêu đề báo chí. Công chúng có thể nhận thức và hiểu các từ trend, nhưng đối với báo chí, họ đặt ra những yêu cầu cao hơn về chất lượng thông tin, tính chuyên nghiệp hay độ tin cậy của nguồn tin. Đối với công chúng đọc báo mạng điện tử thường xuyên họ có xu hướng dễ dàng chấp nhận việc sử dụng từ trend vì họ tin tưởng vào tính xác thực, cũng như sự nghiêm túc trong cách trình bày thông tin của báo chí. Do vậy, việc sử dụng

từ trend cần được cân nhắc một cách cẩn thận, nếu lạm dụng có thể đánh mất lòng tin của công chúng đối với chất lượng thông tin mà báo chí cung cấp.

Các nghiên cứu trong tương lai cần bổ sung các phương pháp nghiên cứu định tính như phỏng vấn sâu, phỏng vấn nhóm tập trung hay quan sát tham gia để giải thích rõ hơn về vai trò của từ trend trên báo chí và phản ứng của công chúng đối với thực trạng này, đặc biệt các phản ứng tiêu cực.

Từ kết quả nghiên cứu và những phân tích gợi mở trên, nghiên cứu đề xuất một số khuyến nghị đối với các cơ quan báo chí khi vận dụng từ trend trong hoạt động báo chí như sau:

Thứ nhất, việc sử dụng từ trend trên báo chí cần có quy định chặt chẽ. Trong môi trường cạnh tranh thông tin gay gắt, các toà soạn luôn tìm cách thu hút sự chú ý của công chúng và từ trend đã trở thành một phương tiện ngôn ngữ giúp gia tăng tương tác. Tuy nhiên, các tòa soạn cần xây dựng những quy định cụ thể liên quan đến việc sử dụng từ trend để đảm bảo quy tắc đạo đức, tuân thủ pháp luật, giữ gìn sự trong sáng của tiếng Việt và phù hợp với thuần phong mỹ tục. Đối với phóng viên, nhà báo, việc "bắt trend" không nên thực hiện một cách nóng vội hay chạy theo thị hiếu nhất thời của công chúng. Trước hết, người làm báo nên cân nhắc kỹ lưỡng, đặc biệt khi từ trend có thể mang tính công kích cá nhân hoặc không phù hợp với thị hiếu chung, để đảm bảo nội dung truyền tải giữ được tính chuyên nghiệp và vai trò định hướng xã hội của báo chí.

Thứ hai, theo dõi, chọn lọc và sử dụng phù hợp. Việc "bắt trend" kịp thời mang lại hiệu quả trong việc thu hút và níu chân công chúng trong tác phẩm báo chí. Tuy nhiên, việc sử dụng từ trend cần chọn lọc, tránh lạm dụng, chạy theo thị hiếu công chúng. Các nhà báo, phóng viên cần đảm bảo rằng, việc sử dụng từ trend không làm mất đi tính nghiêm túc, chất lượng thông tin và tính chuyên nghiệp của báo chí.

Thứ ba, nâng cao chất lượng nội dung và thông tin. Mặc dù từ trend có thể khiến công chúng thích thú ở các phản ứng đầu tiên, tuy nhiên, để giữ chân công chúng ở lại lâu hơn với tác phẩm thì chất lượng thông tin là yếu tố tiên quyết. Các cơ quan báo chí cần đảm bảo rằng thông tin được truyền tải phải mang lại giá trị cho công chúng và đáp ứng yêu cầu về mặt chất lượng. Báo chí cần thực hiện tốt vai trò cốt lõi của mình trong việc cung cấp thông tin chính xác, đa dạng và đáng tin cậy.

Cuối cùng, đảm bảo tính linh hoạt trong việc điều chỉnh nội dung phù hợp với công chúng. Công tác thực hành báo chí cần thực hiện phối hợp song hành cùng các nhiệm vụ nghiên cứu thăm dò để lắng nghe và phân tích những phản hồi của công chúng nhằm điều chỉnh sử dụng từ trend sao cho phù hợp. Đồng thời, các đơn vị cũng cần cân nhắc sự thay đổi trong sở thích, thói quen, sự nhạy bén của công chúng với ngôn ngữ và xu hướng mới sẽ giúp cơ quan báo chí tạo dựng sự kết nối đối với độc giả của mình, nhất là đối với độc giả trẻ.

TÀI LIỆU THAM KHẢO

[1] T. T. V. Lê, "Về một kiểu ngôn ngữ hot trend của giới trẻ hiện nay," *Từ điển học và Bách khoa thư*, vol. 1, no. 69, pp. 109–114, Jan. 2021.

[2] H. Cantril, "Professor quiz: A gratifications study," in *Radio research 1941*, New York: Duell, Solan & Pearce: In P.F. Lazarsfeld & F. Stanton (Eds.), 1942, pp. 34–35.

[3] A. Whiting and D. Williams, "Why people use social media: a uses and gratifications approach," *Qualitative Market Research: An International Journal*, vol. 16, no. 4, pp. 362–369, Aug. 2013, doi: 10.1108/QMR-06-2013-0041.

[4] R. A. Raji, O. A. Arikewuyo, A. S. Oladimeji Adeyemi, and M. R. Pahore, "Unveiling Social Gratifications Sought and Obtained from Social Media Utilization," *Jurnal The Messenger*, vol. 12, no. 2, p. 168, Jul. 2020, doi: 10.26623/themessenger.v12i2.1818.

[5] E. Katz, J. G. Blumler, and M. Gurevitch, "Uses and Gratifications Research," *Public Opin Q*, vol. 37, no. 4, p. 509, 1974, doi: 10.1086/268109.

[6] J. J. Galloway, "Audience uses and Gratifications: An Expectancy Model," *Communic Res*, vol. 8, no. 4, pp. 435–449, Oct. 1981, doi: 10.1177/009365028100800403.

[7] A. Lichtenstein and L. B. Rosenfeld, "Uses and misuses of gratifications research," *Communic Res*, vol. 10, no. 1, pp. 97–109, Jan. 1983, doi: 10.1177/009365083010001005.

[8] P. Palmgreen and J. D. Rayburn, "Gratifications sought and media exposure: An expectancy value model," *Communic Res*, vol. 9, no. 4, pp. 561–580, Oct. 1982, doi: 10.1177/009365082009004004.

[9] Phong Vân, "Thế giới có 5 tỉ người dùng mạng xã hội," *Báo Thanh Niên*, 2024.

[10] T. T. Y. M. Tran, "Online news consumption about COVID-19 pandemic a snapshot from Vietnam," *Journal of Science and Technology Issue on Information and Communications Technology*, pp. 96–102, Jun. 2022, doi: 10.31130/ud-jst.2022.258E.

[11] D. Y. Wohn and M. Ahmadi, "Motivations and habits of micro-news consumption on mobile social media," *Telematics and Informatics*, vol. 44, p. 101262, Nov. 2019, doi: 10.1016/j.tele.2019.101262.

[12] E. Yanardağoğlu, "'Just the way my generation reads the news': News consumption habits of youth in Turkey and the UK," *Global Media and Communication*, vol. 17, no. 2, pp. 149–166, Aug. 2021, doi: 10.1177/1742766520979729.

[13] L. Molyneux, "Mobile News Consumption," *Digital Journalism*, vol. 6, no. 5, pp. 634–650, May 2018, doi: 10.1080/21670811.2017.1334567.

[14] P. Van Aelst *et al.*, "Does a Crisis Change News Habits? A Comparative Study of the Effects of COVID-19 on News Media Use in 17 European Countries," *Digital*

Journalism, vol. 9, no. 9, pp. 1208–1238, Oct. 2021, doi: 10.1080/21670811.2021.1943481.

[15] K. Lerman and R. Ghosh, "Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 4, no. 1, pp. 90–97, May 2010, doi: 10.1609/icwsm.v4i1.14021.

[16] S. Asur, B. A. Huberman, G. Szabo, and C. Wang, "Trends in Social Media: Persistence and Decay," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 5, no. 1, pp. 434–437, Aug. 2021, doi: 10.1609/icwsm.v5i1.14167.

[17] S. Paulussen and R. A. Harder, "Social Media References in Newspapers," *Journalism Practice*, vol. 8, no. 5, pp. 542–551, Sep. 2014, doi: 10.1080/17512786.2014.894327.

[18] D. H. Weaver and L. Willnat, "Changes in U.S. Journalism," *Journalism Practice*, vol. 10, no. 7, pp. 844–855, Oct. 2016, doi: 10.1080/17512786.2016.1171162.

[19] S. C. McGregor, "Social media as public opinion: How journalists use social media to represent public opinion," *Journalism*, vol. 20, no. 8, pp. 1070–1086, Aug. 2019, doi: 10.1177/1464884919845458.

[20] T. A. Nghiêm, "Ảnh hưởng của truyền thông xã hội đến hoạt động báo chí," *Tạp chí Lý luận chính trị và Truyền thông điện tử*, 2024, Accessed: Nov. 23, 2024. [Online]. Available: https://www.phunuonline.com.vn/noi-lo-tu-dau-vet-so-a1471626.html

[21] M. T. Bastos, "Shares, Pins, and Tweets," *Journal Stud*, vol. 16, no. 3, pp. 305–325, May 2015, doi: 10.1080/1461670X.2014.891857.

[22] A. Ju, S. H. Jeong, and H. I. Chyi, "Will Social Media Save Newspapers?," *Journalism Practice*, vol. 8, no. 1, pp. 1–17, Jan. 2014, doi: 10.1080/17512786.2013.794022.

[23] Y. Kim, H.-T. Chen, and H. Gil de Zúñiga, "Stumbling upon news on the Internet: Effects of incidental news exposure and relative entertainment use on political engagement," *Comput Human Behav*, vol. 29, no. 6, pp. 2607–2614, Nov. 2013, doi: 10.1016/j.chb.2013.06.005.

[24] R. Fletcher and R. K. Nielsen, "Are people incidentally exposed to news on social media? A comparative analysis," *New Media Soc*, vol. 20, no. 7, pp. 2450–2468, Jul. 2018, doi: 10.1177/1461444817724170.

[25] D. Antunovic, P. Parsons, and T. R. Cooke, "'Checking' and googling: Stages of news consumption among young adults," *Journalism*, vol. 19, no. 5, pp. 632–648, May 2018, doi: 10.1177/1464884916663625.

[26] P. Boczkowski, E. Mitchelstein, and M. Matassi, "Incidental News: How Young People Consume News on Social Media," 2017. doi: 10.24251/HICSS.2017.217.

[27] Z. Reich, "How citizens create news stories," *Journal Stud*, vol. 9, no. 5, pp. 739–758, Oct. 2008, doi: 10.1080/14616700802207748.

[28] M. A. Salaudeen, "From Personal to Professional: Exploring the Influences on Journalists' Evaluation of Citizen Journalism Credibility," *Journalism Practice*, vol. 16, no. 10, pp. 2040–2063, Nov. 2022, doi: 10.1080/17512786.2021.1892517.

[29] Y. Kim and W. Lowrey, "Who are Citizen Journalists in the Social Media Environment?," *Digital Journalism*, vol. 3, no. 2, pp. 298–314, Mar. 2015, doi: 10.1080/21670811.2014.930245.

[30] G. Mythen, "Reframing risk? Citizen journalism and the transformation of news," *J Risk Res*, vol. 13, no. 1, pp. 45–58, Jan. 2010, doi: 10.1080/13669870903136159.

[31] A. Al-Rawi, "Viral News on Social Media," *Digital Journalism*, vol. 7, no. 1, pp. 63–79, Jan. 2019, doi: 10.1080/21670811.2017.1387062.

[32] J. Berger, "Arousal Increases Social Transmission of Information," *Psychol Sci*, vol. 22, no. 7, pp. 891–893, Jul. 2011, doi: 10.1177/0956797611413294.

[33] J. Berger and K. L. Milkman, "What Makes Online Content Viral?," *Journal of Marketing Research*, vol. 49, no. 2, pp. 192–205, Apr. 2012, doi: 10.1509/jmr.10.0353.

[34] G. J. Golan and L. Zaidner, "Creative Strategies in Viral Advertising: An Application of Taylor's Six-Segment Message Strategy Wheel," *Journal of Computer-Mediated Communication*, vol. 13, no. 4, pp. 959–972, Jul. 2008, doi: 10.1111/j.1083-6101.2008.00426.x.

[35] V. Karnowski, A. S. Kümpel, L. Leonhard, and D. J. Leiner, "From incidental news exposure to news engagement. How perceptions of the news post and news usage patterns influence engagement with news articles encountered on Facebook," *Comput Human Behav*, vol. 76, pp. 42–50, Nov. 2017, doi: 10.1016/j.chb.2017.06.041.

[36] J. Kaiser, T. R. Keller, and K. Kleinen-von Königslöw, "Incidental News Exposure on Facebook as a Social Experience: The Influence of Recommender and Media Cues on News Selection," *Communic Res*, vol. 48, no. 1, pp. 77–99, Feb. 2021, doi: 10.1177/0093650218803529.

[37] C. S. Lee, L. Ma, and D. H.-L. Goh, "Why Do People Share News in Social Media?," 2011, pp. 129–140. doi: 10.1007/978-3-642-23620-4_17.

[38] V. T. Phan, "Nỗi lo từ 'dấu vết số,'" *Báo Phụ nữ Thành phố Hồ Chí Minh*, 2022. Accessed: Nov. 23, 2024. [Online]. Available: https://www.phunuonline.com.vn/noi-lo-tu-dau-vet-so-a1471626.html

[39] A. Field, *Discovering Statistics Using SPSS*, 3rd Edition. London: Sage Publications Ltd., 2009.

Thông tin tác giả

Trần Văn Vỹ*: Học viên cao học tại Trường Đại học Khoa học Xã hội và Nhân văn – ĐHQG HN, đang công tác tại Trường Đại học Sư phạm, Đại học Đà Nẵng, email: tvvy@ued.udn.vn.

Đỗ Thanh Huyền: Học viên cao học tại Trường Đại học Khoa học Xã hội và Nhân văn – ĐHQG HN, email: thanhhuyentntp@gmail.com.

Lại Thị Ninh: Học viên cao học tại Trường Đại học Khoa học Xã hội và Nhân văn – ĐHQG HN, email: laininhsjc@gmail.com.

Nguyễn Minh Tâm: Học viên cao học tại Trường Đại học Khoa học Xã hội và Nhân văn – ĐHQG HN, email: minhtamnguyen310501@gmail.com.

KHAI THÁC NGỮ LIỆU CHUYÊN NGÀNH PHÁP LUẬT BẰNG TRÍ TUỆ NHÂN TẠO (AI) TRONG GIẢNG DẠY NGOẠI NGỮ

Nguyễn Thị Cẩm Nhung[1]

Tóm tắt

Việc giảng dạy ngôn ngữ chuyên ngành pháp luật đang đặt ra yêu cầu đổi mới mạnh mẽ nhằm đáp ứng nhu cầu hội nhập và chuyển đổi số trong giáo dục nghề nghiệp. Bài viết tập trung phân tích vai trò của trí tuệ nhân tạo (AI) trong việc khai thác, xử lý và tích hợp ngữ liệu chuyên ngành pháp luật vào quá trình giảng dạy ngôn ngữ tại các trường Công An Nhân dân (CAND). Qua đó, bài viết làm rõ những ứng dụng cụ thể của AI, đồng thời chỉ ra các cơ hội, thách thức và đề xuất giải pháp khả thi để nâng cao hiệu quả giảng dạy ngoại ngữ trong môi trường đặc thù của lực lượng vũ trang.

Từ khóa: *Trí tuệ nhân tạo, ngữ liệu chuyên ngành pháp luật, giảng dạy ngoại ngữ*

Abstract

Teaching specialized legal language requires strong innovation to meet the needs of integration and digital transformation in vocational education. The article analyzes the role of artificial intelligence (AI) in exploiting, processing, and integrating specialized legal into the language-teaching process at People's Public Security schools. Thereby, the article clarifies specific applications of AI while pointing out opportunities and challenges and proposing feasible solutions to improve the effectiveness of teaching foreign languages in the particular environment of the armed forces.

Keywords: *Artificial Intelligence (AI), legal corpus, foreign language teaching*

1. Giới thiệu

Nghiên cứu về Ngôn ngữ Chuyên ngành (Language for Specific Purposes-LSP) luôn nhấn mạnh tầm quan trọng của corpus (kho ngữ liệu) đích thực trong đào tạo(Hutchinson & Waters, 1987). Trong đào tạo tiếng Anh pháp lý (Legal English/English for Specific Purposes), sự thiếu hụt các bộ ngữ liệu lớn và phân loại chuyên sâu là rào cản chính. Các nghiên cứu gần đây trong lĩnh vực Luật học Điện toán (Computational Law) đã nhận định việc xử lý văn bản pháp luật cần các mô hình xử lý ngôn ngữ tự nhiên (Natural Language Processing-NLP) chuyên biệt để đối phó với sự mơ hồ và tính phức tạp về cấu trúc (Bench-Capon, 1997). Trong bối cảnh toàn cầu hóa và chuyển đổi số hiện nay, việc hợp tác quốc tế, trao đổi thông tin và tiếp cận tri thức pháp luật quốc tế trở nên quan trọng đối với lực lượng CAND. Năng lực sử dụng ngoại ngữ chuyên ngành, đặc biệt là tiếng Anh chuyên ngành pháp luật, đã trở thành yêu cầu thiết yếu, là tiêu chuẩn bắt buộc để cán bộ, học viên có thể nghiên cứu, học tập và thực

[1] *Th.S*

hiện nhiệm vụ hiệu quả trong môi trường hội nhập. Tuy nhiên, ngôn ngữ chuyên ngành pháp luật có tính đặc thù cao, việc dạy và học trở nên khó khăn hơn so với các lĩnh vực ngoại ngữ thông thường. Trước hết, hệ thống thuật ngữ pháp lý đòi hỏi người học phải nắm vững và hiểu sâu các khái niệm chuyên môn, đảm bảo tính chính xác tuyệt đối trong sử dụng. Bên cạnh đó, cấu trúc diễn ngôn của văn bản pháp luật thường phức tạp, bao gồm nhiều mệnh đề phụ, câu điều kiện và văn bản được thiết lập một cách chặt chẽ, chính xác. Ngoài ra, phong cách hành chính - pháp luật còn yêu cầu người sử dụng phải tuân thủ nghiêm ngặt các quy chuẩn về ngữ pháp, cú pháp, giọng văn và ngữ cảnh phù hợp với văn bản quy phạm pháp luật, văn thư hành chính hay văn bản tố tụng. Thực tế giảng dạy cho thấy việc khai thác và sử dụng ngữ liệu chuyên ngành pháp luật trong giảng dạy hiện vẫn còn nhiều hạn chế. Nguồn tư liệu pháp lý phục vụ đào tạo còn thiếu, các kho ngữ liệu (bao gồm văn bản luật, án lệ, và tài liệu nghiệp vụ) chưa được xây dựng và chọn lọc một cách có hệ thống. Việc thu thập và phân tích dữ liệu pháp luật đòi hỏi kỹ năng chuyên sâu về xử lý dữ liệu, trong khi nguồn nhân lực hiện tại của các cơ sở đào tạo CAND còn hạn chế về thời gian và công cụ hỗ trợ. Quá trình biên soạn học liệu từ ngữ liệu gốc cũng mát nhiều thời gian, nhiều công sức, vì vậy, việc đổi mới chương trình và phương pháp giảng dạy gặp khó khăn. Trong bối cảnh đó, trí tuệ nhân tạo (AI) nổi lên như một giải pháp đột phá, mở ra hướng đi mới cho việc khai thác và ứng dụng ngữ liệu chuyên ngành pháp luật trong giảng dạy ngoại ngữ. Với khả năng xử lý ngôn ngữ tự nhiên (NLP), AI có thể tự động hóa quá trình thu thập, phân loại, chú thích và phân tích ngữ liệu từ nhiều nguồn khác nhau với tốc độ nhanh và độ bao phủ rộng. Công nghệ AI, đặc biệt là các mô hình chuyển đổi (Transformer models) như ChatGPT, Copilot, và Gemini, đã chứng minh khả năng vượt trội trong việc tạo và xử lý ngôn ngữ tự nhiên (NLP). Trong giảng dạy, AI hỗ trợ giáo viên biên soạn học liệu chuyên biệt, tạo các bài tập như dịch thuật thuật ngữ pháp lý, tóm tắt văn bản luật hoặc mô phỏng tình huống giao tiếp pháp lý thực tế. Đồng thời, các hệ thống học tập thông minh ứng dụng AI còn có khả năng cá nhân hóa quá trình học tập, đánh giá năng lực từng học viên và đề xuất lộ trình, tài liệu hoặc bài tập phù hợp với nhu cầu và tiến độ riêng. Trong giáo dục ngoại ngữ, AI đã được chứng minh là có khả năng hỗ trợ Cá nhân hóa việc học (Personalized Learning) (Merino-Campos, C., 2025) và tạo phản hồi ngôn ngữ theo ngữ cảnh (Contextual Feedback) (Lee, S., Choe, H., Zou, D., & Jeon, J., 2025). Việc ứng dụng AI vào lĩnh vực luật học bắt đầu bằng các công cụ tra cứu thông minh, nhưng gần đây đã mở rộng sang việc phân tích và chuẩn hóa văn bản luật (Chalkidis et al., Findings 2020), đặt nền móng cho việc khai thác ngữ liệu pháp luật trong giảng dạy. Nhờ đó, AI không chỉ giúp nâng cao hiệu quả và chất lượng đào tạo ngoại ngữ chuyên ngành mà còn hỗ trợ cán bộ, học viên CAND nhanh chóng đạt được năng lực ngôn ngữ cần thiết, đáp ứng yêu cầu hội nhập quốc tế và phát

triển trong thời kỳ chuyển đổi số.

2. Vai trò của trí tuệ nhân tạo (AI) trong khai thác ngữ liệu pháp luật

2.1. Khai thác và xây dựng ngữ liệu tự động (Corpus construction)

Các nghiên cứu gần đây đã khẳng định vai trò của AI trong lĩnh vực ngôn ngữ học ứng dụng và giáo dục ngoại ngữ. Theo Zawacki-Richter et al. (2019), AI có thể hỗ trợ hiệu quả trong việc thu thập và phân tích ngữ liệu, tạo tiền đề cho các mô hình học tập thích ứng và cá nhân hóa. NLP giúp tự động hóa quy trình xử lý văn bản phức tạp, đặc biệt trong các lĩnh vực chuyên ngành như luật (Bertl, M., Price, S., & Draheim, D.,2026), y khoa hay tài chính (Oyewole, A. T, et al., 2024), lĩnh vực yêu cầu độ chính xác ngữ nghĩa cao. Ngoài ra, ứng dụng các công cụ như ChatGPT và Copilot trong giáo dục giúp giáo viên tiết kiệm thời gian biên soạn tài liệu, đồng thời tăng cường khả năng tương tác, sáng tạo và phản hồi tự động trong quá trình dạy học (Karataş, F., Abedi, F.Y., Ozek Gunyel, F. *et al.*, 2024). Trong lĩnh vực pháp luật, người tiên phong trong lĩnh vực luật học điện toán (Computational Legal Studies), Ashley (2017) đã nhận định AI có thể truy xuất, phân loại và phân tích hàng triệu văn bản pháp lý, từ luật quốc tế, hiệp ước cho đến án lệ, giúp hình thành nên các ngữ liệu pháp luật (legal corpora) phục vụ cho nghiên cứu và đào tạo. Các công cụ trí tuệ nhân tạo hiện nay như ChatGPT, GitHub Copilot hay các mô hình xử lý ngôn ngữ tự nhiên (NLP) tiên tiến như BERT, GPT, Claude và LLaMA có khả năng truy xuất, tổng hợp và xử lý dữ liệu pháp luật từ các nguồn mở như: United Nations Treaty Series, European Court of Human Rights Database hay UNCLOS Texts. Với năng lực học tập sâu (deep learning) và phân tích ngữ nghĩa chính xác, các hệ thống này có thể tự động thu thập và phân loại văn bản theo từng chủ đề như "tội phạm mạng", "quyền con người" hay "di cư quốc tế", đồng thời chú thích (annotation) cho từng loại văn bản dựa trên cấu trúc pháp lý, thuật ngữ chuyên biệt hoặc loại diễn ngôn như legal reasoning, argumentation hay ruling. Trên cơ sở đó, AI có thể tạo ra các ngữ liệu song ngữ có tính cập nhật cao, hỗ trợ giáo viên và học viên trong việc đối chiếu, phân tích và dịch thuật chuyên ngành pháp lý. Ví dụ, thay vì giáo viên phải tự tìm kiếm và đối chiếu các văn bản pháp luật về "Tội phạm mạng" (Cybercrime) giữa các quốc gia, một hệ thống AI có thể tự động xây dựng kho ngữ liệu song ngữ gồm hàng ngàn văn bản liên quan từ các nguồn công khai. Kho ngữ liệu này trở thành nguồn tư liệu quý giá để thiết kế bài giảng chuyên ngành, xây dựng bài tập dịch thuật pháp lý hoặc phát triển các kịch bản mô phỏng tranh luận pháp lý (mock legal hearings), qua đó góp phần nâng cao hiệu quả giảng dạy và năng lực sử dụng ngoại ngữ chuyên ngành của người học. Trước đây, việc biên soạn học liệu pháp luật chuyên ngành chủ yếu dựa vào năng lực thủ công của giáo viên, từ việc tìm kiếm, chọn lọc văn bản pháp luật, trích xuất thuật ngữ chuyên ngành cho đến thiết kế các bài tập tình huống thực hành. Quá trình này không chỉ tiêu tốn nhiều thời gian và công sức mà còn thiếu tính hệ thống và khả năng cập nhật thường xuyên. Nghiên cứu thực nghiệm của Assassi, T. (2025) đã làm nổi bật những thách thức của giáo viên giảng dạy tiếng Anh chuyên ngành không phải là người bản xứ trong việc tiếp cận và ứng dụng kiến thức chuyên ngành, đồng thời cho thấy trí tuệ nhân tạo (AI), đặc biệt là ChatGPT, có tiềm năng hỗ trợ phát triển chuyên môn như một công cụ bổ sung cho kiến thức

chuyên sâu. Vì vậy, sự xuất hiện của trí tuệ nhân tạo đã giúp khắc phục ba hạn chế lớn trong mô hình đào tạo truyền thống. Thứ nhất, AI có thể truy xuất và tổ chức một kho dữ liệu pháp luật phong phú, bảo đảm tính đa dạng và độ bao phủ cao, giúp giáo viên dễ dàng tiếp cận nguồn tư liệu đáng tin cậy. Thứ hai, nhờ công nghệ xử lý ngôn ngữ tự nhiên (NLP), AI có thể tự động chọn lọc, chú thích và phân tích ngữ liệu, giảm đáng kể khối lượng công việc thủ công của giáo viên trong khâu xử lý dữ liệu. Thứ ba, AI liên tục cập nhật các văn bản mới từ nhiều hệ thống pháp lý quốc tế, giúp học liệu luôn phù hợp với bối cảnh thực tiễn và tiến trình hội nhập pháp luật toàn cầu.

2.2. Tự động nhận diện và phân loại thuật ngữ pháp lý

Trong giảng dạy ngôn ngữ chuyên ngành pháp luật, việc nhận diện và hiểu chính xác các thuật ngữ pháp lý là yếu tố then chốt giúp học viên nâng cao khả năng đọc hiểu, phân tích và dịch thuật văn bản. Bài viết cũng xem xét cả các mô hình ngôn ngữ dành riêng cho pháp lý và các phương pháp thích ứng mô hình mục đích chung. Cuối cùng, nghiên cứu xác định 16 thách thức mở quan trọng, tập trung vào việc giảm thiểu sai lệch của AI, tăng cường tính mạnh mẽ và khả năng diễn giải, cũng như cải thiện khả năng giải thích để xử lý hiệu quả sự phức tạp trong ngôn ngữ và lý luận pháp lý. Chen, J., (2025) giới thiệu một mô hình trích xuất văn bản pháp luật tiên tiến, dựa trên xử lý ngôn ngữ mơ hồ và nhận diện ẩn dụ, được thiết kế riêng cho lĩnh vực quản trị môi trường trực tuyến. Mô hình sử dụng mạng Bi-LSTM hai chiều để xác định hành vi vi phạm bằng cách kết nối các quy định pháp luật với các phán quyết, đồng thời tích hợp mô-đun tự chú ý (self-attention) và kỹ thuật chuẩn hóa ngôn ngữ thứ hai để trích xuất thông tin pháp lý một cách hiệu quả, giúp nhận diện và phân loại nội dung vi phạm. Các công cụ AI hiện đại, đặc biệt là các hệ thống dựa trên kỹ thuật xử lý ngôn ngữ tự nhiên (NLP), đã chứng minh hiệu quả vượt trội trong việc tự động phát hiện, phân loại và giải thích thuật ngữ pháp lý trong văn bản. Thông qua NLP, AI không chỉ xác định các từ hoặc cụm từ quan trọng mà còn phân tích ngữ nghĩa và chức năng của chúng trong từng ngữ cảnh. Ví dụ, các thuật ngữ như *"territorial sea"* (lãnh hải), *"jurisdiction"* (thẩm quyền) hay *"due process"* (thủ tục pháp lý đúng) được gắn nhãn và liên kết với các thông tin ngữ pháp, cú pháp và ngữ nghĩa xung quanh. Thuật ngữ *"jurisdiction"* không chỉ được định nghĩa là phạm vi quyền lực pháp lý của một cơ quan mà AI còn phân tích cách thuật ngữ này kết nối với các mệnh đề khác, xác định vai trò trong câu và mối quan hệ với các điều khoản pháp luật liên quan. Ứng dụng này giúp học viên mở rộng vốn từ vựng chuyên ngành, đồng thời nắm vững các từ đồng nghĩa, từ liên quan và cách dùng thuật ngữ trong văn bản pháp luật thực tế. Ví dụ: Trong quá trình giảng dạy môn *English for Criminal Law* tại các trường CAND, giáo viên có thể sử dụng công cụ trí tuệ nhân tạo như ChatGPT, LegalBERT hoặc CaseText AI để hỗ trợ học viên phân tích và nhận diện thuật ngữ pháp lý trong các điều khoản của Bộ luật Hình sự Việt Nam. Chẳng hạn, khi nhập đoạn văn tiếng Anh tương ứng với điều luật: *"Any person who intentionally inflicts injury or causes harm to the health of another person shall be subject to criminal liability under Article 134 of the Penal Code,"* AI có thể tự động xác định và gắn nhãn các thuật ngữ pháp lý quan trọng như *"intentionally"* (yếu tố lỗi – lỗi cố ý), *"inflicts injury"* (hành vi gây thương tích), *"criminal liability"* (trách nhiệm hình sự), và *"Article 134"* (tham chiếu đến điều luật

cụ thể). Thông qua việc gắn nhãn và phân tích này, hệ thống có thể tự động tạo chú thích giải thích hoặc liên kết đến các điều khoản liên quan, giúp người học hiểu sâu hơn bối cảnh pháp lý của quy định. Cách tiếp cận này hỗ trợ học viên trong quá trình học tiếng Anh pháp lý hoặc tiếng Việt chuyên ngành luật tại các trường CAND, bởi vì họ không chỉ tiếp cận thuật ngữ đơn lẻ mà còn nắm được mối quan hệ ngữ nghĩa – pháp lý giữa hành vi, hoàn cảnh và hình phạt. Từ đó, năng lực đọc hiểu, phân tích và dịch thuật văn bản pháp luật được nâng cao rõ rệt.

2.3. Phân tích cấu trúc câu phức tạp

Các văn bản pháp lý thường được viết với cấu trúc cú pháp phức tạp, gồm nhiều mệnh đề phụ (subordinate clauses), mệnh đề điều kiện (conditional clauses), hoặc các thành phần bổ nghĩa dài. Điều này khiến việc hiểu đúng nội dung pháp lý trở nên khó khăn đối với người học. Trí tuệ nhân tạo (AI) có thể hỗ trợ bằng cách phân tách câu (sentence parsing), xác định mệnh đề chính và mệnh đề phụ (main vs. subordinate clause identification), và phân tích vai trò ngữ pháp của từng thành phần (syntactic role labeling). Nhờ đó, người học có thể hiểu rõ mối quan hệ ngữ pháp trong câu, chẳng hạn như điều kiện, nguyên nhân, hệ quả, hoặc giới hạn pháp lý của một quy định. Chẳng hạn, trong câu tiếng Anh "*A person who intentionally causes the death of another person shall be guilty of murder and shall be liable to imprisonment for life,"* AI có thể tự động phân tích cấu trúc và xác định mệnh đề chính là "*A person shall be liable to imprisonment for life,"* mệnh đề phụ "*who intentionally causes the death of another person*" là mệnh đề quan hệ giải thích cho chủ ngữ "*a person*." Bên cạnh đó, hệ thống cũng có thể gán nhãn vai trò ngữ pháp cho từng thành phần như đề (theme), thuyết (rheme), hoặc cụm giới từ chỉ hình phạt. Nhờ khả năng này, người học và giáo viên có thể hiểu rõ mối quan hệ ngữ pháp giữa các thành phần trong câu, nắm được ranh giới và điều kiện của quy định pháp luật, từ đó nâng cao hiệu quả trong việc đọc hiểu, phân tích và dịch thuật các văn bản pháp lý tiếng Anh. Điều này giúp rút ngắn thời gian phân tích, tăng độ chính xác khi dịch thuật, và nâng cao năng lực đọc hiểu văn bản pháp lý bằng tiếng Anh. Ariai và Demartini (2024) cung cấp tổng quan về các tác vụ xử lý ngôn ngữ tự nhiên dành riêng cho văn bản pháp lý. Trong phần về mô hình ngôn ngữ pháp lý (legal language models), các nhà nghiên cứu phân tích cả các mô hình ngôn ngữ đã phát triển và các phương pháp để điều chỉnh các mô hình ngôn ngữ chung cho lĩnh vực pháp lý. Vì vậy, ứng dụng AI không chỉ giúp đơn giản hóa quá trình xử lý ngôn ngữ pháp lý phức tạp mà còn góp phần đổi mới phương pháp giảng dạy ngoại ngữ chuyên ngành luật trong các cơ sở đào tạo, đặc biệt là trong môi trường giáo dục CAND.

2.4. Nhận diện mối quan hệ ngữ nghĩa giữa các mệnh đề và đoạn văn

AI còn đóng vai trò quan trọng trong việc nhận diện mối quan hệ ngữ nghĩa giữa các mệnh đề và đoạn văn trong văn bản pháp luật, một yếu tố then chốt giúp người học hiểu sâu hơn về logic và lập luận pháp lý. Vadlapati P. (2022) nhận định việc đơn giản hóa văn bản phức tạp giúp người đọc dễ dàng hiểu văn bản phức tạp hơn. Nhà nghiên cứu đề xuất phương pháp tận dụng mô hình ngôn ngữ lớn (Large Language Model-LLM) để thay thế các từ không phổ biến bằng các từ chuẩn, chia nhỏ các câu phức tạp thành các điểm chính dễ hiểu và tạo ra một bản tóm tắt ngắn gọn. Một quá trình xác thực được thực hiện bằng cách tận dụng LLM đã

chọn để đảm bảo rằng các đầu ra được đơn giản hóa vẫn giữ nguyên ý nghĩa và ngữ cảnh của văn bản gốc. Phương pháp này có nhiều ứng dụng tiềm năng trong giáo dục và diễn giải pháp lý về luật pháp và các quy định. Những phát hiện cho thấy tiềm năng của LLM trong việc cải thiện khả năng tiếp cận thông tin cho nhiều đối tượng độc giả khác nhau.Vì vậy, thông qua các mô hình xử lý ngôn ngữ tự nhiên (NLP) tiên tiến, AI có thể tự động xác định các quan hệ ngữ nghĩa như nguyên nhân – kết quả, điều kiện – kết quả, hoặc so sánh – đối chiếu giữa các phần của câu hoặc giữa các đoạn. Ví dụ, trong quy định *"If a person causes the death of another by negligence, he shall be guilty of manslaughter,"* AI có thể nhận diện đây là cấu trúc điều kiện – kết quả, trong đó mệnh đề điều kiện *"If a person causes the death of another by negligence"* thiết lập hoàn cảnh xảy ra hành vi phạm tội, còn mệnh đề chính *"he shall be guilty of manslaughter"* thể hiện hệ quả pháp lý. Tương tự, trong các đoạn văn dài hơn, AI có thể phát hiện mối quan hệ đối chiếu như trong câu *"While civil law seeks to compensate the victim, criminal law aims to punish the offender,"* cho thấy sự khác biệt về mục đích giữa hai lĩnh vực pháp luật. Khả năng này giúp học viên nhận ra cách thức lập luận được tổ chức trong các văn bản pháp lý, hiểu rõ cách mà các điều khoản được ràng buộc logic với nhau, từ đó phát triển tư duy phân tích và năng lực đọc hiểu chuyên sâu. Việc ứng dụng AI trong phân tích mối quan hệ ngữ nghĩa không chỉ nâng cao hiệu quả học tập mà còn giúp người học tiếp cận với tư duy pháp lý quốc tế một cách hệ thống và khoa học hơn.

2.5. Chú thích tự động (Annotation)

Một trong những ứng dụng nổi bật của trí tuệ nhân tạo trong khai thác ngữ liệu pháp luật là chú thích tự động (automatic annotation). Công nghệ này cho phép hệ thống AI tự động gắn nhãn, giải thích và mở rộng thông tin cho các thuật ngữ pháp lý ngay trong văn bản gốc, giúp người học hiểu nội dung một cách trực quan và sâu sắc hơn. Hughson, I. (2024) mô tả một tập hợp văn bản được mã hóa XML trong lĩnh vực pháp lý, được thu thập cho một dự án tóm tắt tự động bằng hai lớp chú thích riêng biệt: chú thích thủ công trạng thái tu từ của câu và một quy trình chú thích hoàn toàn tự động kết hợp nhiều bộ xử lý ngôn ngữ riêng lẻ. Thông qua kỹ thuật xử lý ngôn ngữ tự nhiên (NLP) và học máy (machine learning), AI có thể phát hiện những từ khóa pháp lý quan trọng, cung cấp định nghĩa, ví dụ sử dụng, hoặc dẫn chiếu đến các điều luật có liên quan. Thêm vào đó, AI cung cấp giải thích ngữ nghĩa và chú thích tự động giúp học viên tiết kiệm thời gian tra cứu và tăng hiệu quả học tập, đồng thời hỗ trợ giáo viên trong việc chuẩn bị tài liệu và thiết kế bài tập, góp phần nâng cao chất lượng giảng dạy ngoại ngữ chuyên ngành pháp luật, đặc biệt trong môi trường giáo dục của các trường CAND. Ví dụ, trong câu tiếng Anh *"Any person who intentionally causes serious bodily harm to another shall be liable to imprisonment,",* hệ thống AI có thể chú thích tự động cho các thuật ngữ pháp lý như *"intentionally"* (chủ ý, thể hiện yếu tố lỗi của hành vi), *"serious bodily harm"* (tổn hại nghiêm trọng về thể chất – một khái niệm pháp lý có thể khác nhau tùy hệ thống luật), và *"liable to imprisonment"* (chịu trách nhiệm hình sự bằng hình phạt tù). Ngoài việc giải thích nghĩa, AI còn có thể liên kết các cụm này đến những điều khoản tương ứng trong Bộ luật Hình sự hoặc các án lệ minh họa, giúp người học hiểu rõ cách sử dụng ngôn ngữ pháp lý trong thực tiễn. Nhờ khả năng chú thích tự động này, học viên không chỉ học từ vựng pháp lý một cách

ngữ cảnh hóa mà còn phát triển kỹ năng phân tích và hiểu lập luận pháp luật. Đồng thời, công cụ này cũng hỗ trợ giáo viên trong việc biên soạn giáo trình và thiết kế bài tập ứng dụng, giúp tiết kiệm thời gian, tăng tính chính xác và nâng cao hiệu quả giảng dạy ngoại ngữ chuyên ngành pháp luật, đặc biệt trong môi trường đào tạo của các trường CAND. Ngoài ra, giáo viên có thể yêu cầu AI hiển thị thêm chú thích (annotation) như: *"intentionally"*: deliberate or purposeful act, indicating mens rea (criminal intent); *"inflicts injury"*: physical act causing harm to another person; *"criminal liability"*: legal responsibility under criminal law. Việc học viên tương tác với văn bản qua hệ thống AI giúp họ không chỉ mở rộng vốn từ chuyên ngành mà còn hiểu sâu hơn mối quan hệ giữa hành vi – yếu tố lỗi – hậu quả pháp lý trong các điều khoản hình sự. Nhờ đó, quá trình giảng dạy và học tập ngoại ngữ chuyên ngành pháp luật tại các trường CAND trở nên sinh động, thực tiễn và gắn kết chặt chẽ hơn với yêu cầu đào tạo nghề nghiệp.

3. Cơ hội, thách thức và đề xuất giải pháp nâng cao hiệu quả giảng dạy ngoại ngữ chuyên ngành pháp luật bằng AI

3.1. Cơ hội và thách thức

Ứng dụng trí tuệ nhân tạo trong giảng dạy ngoại ngữ chuyên ngành pháp luật đang mở ra những cơ hội to lớn, đặc biệt trong bối cảnh giáo dục số hóa và hội nhập quốc tế, nơi yêu cầu học viên phải tiếp cận thông tin pháp luật một cách chính xác và hiệu quả. Trước hết, AI cho phép cá nhân hóa quá trình học tập, giúp mỗi học viên tiếp cận các ngữ liệu phù hợp với trình độ, mục tiêu học tập và năng lực ngôn ngữ riêng của mình. Nhờ các công cụ như ChatGPT, Grammarly hay DeepL, học viên có thể thực hành các kỹ năng đọc hiểu, viết và dịch thuật văn bản pháp lý với phản hồi tức thời, từ đó nâng cao khả năng tự học, rút ngắn thời gian phân tích và hiểu sâu các nội dung pháp luật phức tạp, đồng thời giảm bớt sự phụ thuộc vào sự hướng dẫn trực tiếp từ giáo viên. Bên cạnh đó, AI còn giúp mở rộng và làm phong phú nguồn tài nguyên học tập. Khả năng tự động thu thập, phân loại và phân tích các văn bản pháp luật quốc tế như UNCLOS, ICCPR hay các đạo luật hình sự của Hoa Kỳ tạo ra một kho dữ liệu thực tiễn và đa dạng. Điều này cho phép học viên tiếp cận trực tiếp với văn bản gốc, phân tích các cấu trúc pháp lý và thuật ngữ chuyên ngành, từ đó nâng cao khả năng đọc hiểu và rèn luyện kỹ năng phân tích các văn bản pháp luật chuyên sâu, đồng thời phát triển tư duy pháp lý một cách hệ thống. Ngoài ra, AI còn nâng cao năng lực nghiên cứu và giảng dạy của giáo viên. Nhờ khả năng phân tích diễn ngôn pháp lý và nhận diện các đặc trưng ngôn ngữ chuyên ngành, giáo viên có thể thiết kế các bài giảng tích hợp công nghệ số, xây dựng bài tập tương tác và chuẩn bị tài liệu học tập một cách nhanh chóng và hiệu quả. Việc ứng dụng AI không chỉ tiết kiệm thời gian mà còn giúp nâng cao chất lượng giảng dạy, tạo môi trường học tập hiện đại, sinh động và gắn kết với thực tiễn pháp luật, đồng thời thúc đẩy học viên phát triển kỹ năng ngoại ngữ chuyên ngành một cách toàn diện.

Mặc dù AI mang lại nhiều cơ hội trong giảng dạy ngoại ngữ chuyên ngành pháp luật, nhưng việc ứng dụng công nghệ này cũng đi kèm với nhiều thách thức đáng kể. Trước hết, vấn đề độ chính xác và tin cậy của dữ liệu pháp lý là một mối quan ngại lớn. Các công cụ AI, dù tiên tiến, đôi khi có thể tạo ra nội dung thiếu căn cứ pháp lý hoặc sai lệch thông tin, điều này

có thể dẫn đến hiểu nhầm hoặc áp dụng sai trong quá trình học tập và giảng dạy. Thứ hai, năng lực số của giáo viên và học viên còn hạn chế, gây khó khăn trong việc khai thác tối đa tiềm năng của các công cụ AI. Nhiều giáo viên chưa quen với việc tích hợp AI vào bài giảng, trong khi học viên có thể chưa đủ kỹ năng để vận dụng các tính năng nâng cao, dẫn đến việc sử dụng AI chưa thực sự hiệu quả. Thứ ba, vấn đề bảo mật thông tin và bản quyền tài liệu cũng là một thách thức quan trọng, đặc biệt khi sử dụng dữ liệu pháp luật đặc thù của lực lượng CAND. Việc chia sẻ hoặc xử lý dữ liệu nhạy cảm trên các nền tảng AI cần đảm bảo tuân thủ các quy định pháp lý và bảo vệ thông tin cá nhân. Ngoài ra, sự phụ thuộc quá nhiều vào AI có thể làm giảm khả năng tư duy phản biện và phân tích độc lập của học viên, vốn là những năng lực cốt lõi trong lĩnh vực pháp lý. Nếu học viên chỉ dựa vào các đề xuất hay phân tích của AI mà không rèn luyện kỹ năng tự suy luận, khả năng đánh giá và áp dụng pháp luật trong các tình huống thực tế sẽ bị hạn chế. Tóm lại, việc ứng dụng AI trong giảng dạy ngoại ngữ chuyên ngành pháp luật cần được cân nhắc kỹ lưỡng, vừa khai thác các lợi ích về công nghệ, vừa kiểm soát các rủi ro liên quan đến dữ liệu, năng lực người dùng và phát triển tư duy pháp lý độc lập.

3.2. Giải pháp nâng cao hiệu quả giảng dạy ngoại ngữ chuyên ngành pháp luật bằng AI

Để khai thác tối đa tiềm năng của trí tuệ nhân tạo (AI) trong giảng dạy ngoại ngữ chuyên ngành pháp luật, việc triển khai các giải pháp đồng bộ và khả thi là hết sức cần thiết. Trước hết, cần xây dựng một hệ thống ngữ liệu pháp lý số hóa riêng cho các trường CAND, nhằm đảm bảo các văn bản, thuật ngữ và tài liệu pháp luật được chuẩn hóa, chính xác và bảo mật, đồng thời phù hợp với đặc thù đào tạo của lực lượng vũ trang. Hệ thống này không chỉ tạo ra một cơ sở dữ liệu phong phú mà còn cho phép học viên tiếp cận trực tiếp với văn bản gốc, các ví dụ minh họa và các quy định pháp luật liên quan, từ đó nâng cao khả năng đọc hiểu, phân tích và dịch thuật văn bản pháp lý chuyên ngành. Thứ hai, đào tạo năng lực số và kỹ năng sử dụng AI cho cả giáo viên và học viên là một yếu tố then chốt để công nghệ được vận dụng hiệu quả. Giáo viên cần nắm vững cách khai thác các công cụ AI, từ việc nhận diện thuật ngữ pháp lý, phân tích cấu trúc câu, chú thích tự động đến phân tích mối quan hệ ngữ nghĩa giữa các mệnh đề và đoạn văn. Đồng thời, học viên cũng cần được hướng dẫn sử dụng AI một cách có chọn lọc, vừa tận dụng tiện ích công nghệ, vừa phát triển tư duy độc lập, khả năng phân tích và phản biện pháp lý, tránh phụ thuộc hoàn toàn vào công cụ số.

Thứ ba, kết hợp AI với phương pháp giảng dạy truyền thống theo mô hình Blended Learning sẽ tối ưu hóa hiệu quả học tập. Trong mô hình này, AI đóng vai trò hỗ trợ, cung cấp phản hồi tức thì, gợi ý phân tích hoặc chú thích, giúp học viên nắm bắt nội dung nhanh hơn. Đồng thời, giáo viên vẫn giữ vai trò định hướng, giải thích các khái niệm pháp lý phức tạp và tổ chức các hoạt động tương tác, thảo luận, tạo điều kiện để học viên phát triển kỹ năng toàn diện, từ đọc hiểu, phân tích đến trình bày quan điểm và giải thích pháp lý.

Cuối cùng, thiết lập cơ chế kiểm định và đánh giá chất lượng dữ liệu AI là cần thiết để bảo đảm tính chính thống, nhất quán và độ tin cậy của nội dung pháp lý được sử dụng trong giảng dạy. Điều này không chỉ đảm bảo các thông tin học viên tiếp cận phản ánh đúng luật

pháp và diễn ngôn pháp lý quốc tế, mà còn nâng cao uy tín của quá trình đào tạo và chất lượng học tập.

Như vậy, AI mang đến tiềm năng to lớn trong việc đổi mới phương pháp giảng dạy và học tập ngoại ngữ chuyên ngành pháp luật. Tuy nhiên, để phát huy tối đa hiệu quả, cần có chiến lược ứng dụng đồng bộ, kết hợp giữa công nghệ và con người, đồng thời chú trọng yếu tố đạo đức nghề nghiệp để các cơ sở đào tạo CAND có thể xây dựng một môi trường học tập hiện đại, an toàn, hiệu quả và phù hợp với yêu cầu thực tiễn của lực lượng vũ trang trong bối cảnh hội nhập và chuyển đổi số.

4. Kết luận

Trí tuệ nhân tạo (AI) đang mở ra những cơ hội vượt trội trong giảng dạy ngoại ngữ chuyên ngành pháp luật, giúp cá nhân hóa quá trình học tập, mở rộng nguồn tài nguyên học tập, nâng cao năng lực nghiên cứu và giảng dạy của giáo viên, đồng thời hỗ trợ học viên phát triển kỹ năng đọc hiểu, phân tích và dịch thuật văn bản pháp luật. Những ứng dụng như nhận diện thuật ngữ pháp lý, phân tích cấu trúc câu phức tạp, xác định mối quan hệ ngữ nghĩa và chú thích tự động đã minh chứng rõ ràng vai trò quan trọng của AI trong việc nâng cao chất lượng đào tạo ngoại ngữ chuyên ngành, đặc biệt trong môi trường đặc thù của các trường CAND. Tuy nhiên, bên cạnh các cơ hội, việc ứng dụng AI cũng đặt ra những thách thức về độ chính xác của dữ liệu pháp lý, năng lực số của giáo viên và học viên, bảo mật thông tin, bản quyền tài liệu và khả năng phát triển tư duy phản biện độc lập. Do đó, để khai thác tối đa tiềm năng của AI, cần có chiến lược ứng dụng đồng bộ, bao gồm xây dựng hệ thống ngữ liệu pháp lý số hóa, đào tạo kỹ năng sử dụng AI, kết hợp AI với phương pháp giảng dạy truyền thống theo mô hình Blended Learning và thiết lập cơ chế kiểm định chất lượng dữ liệu AI. Nhìn chung, AI không chỉ là công cụ hỗ trợ mà còn là yếu tố thúc đẩy đổi mới phương pháp giảng dạy ngoại ngữ chuyên ngành pháp luật. Khi được vận dụng một cách khoa học, có chọn lọc và kết hợp hài hòa với năng lực con người, AI sẽ góp phần xây dựng một môi trường học tập hiện đại, an toàn và hiệu quả, đáp ứng yêu cầu đào tạo chất lượng cao trong bối cảnh hội nhập và chuyển đổi số của lực lượng CAND.

Tài liệu tham khảo

Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. *Education Research International*, *2023*, 1–15. https://doi.org/10.1155/2023/4253331

Assassi, T. (2025). ESP practitioners' perspectives on the efficiency of using artificial intelligence 'ChatGPT' in needs analysis. *Computer Assisted Language Learning*, 1–32. https://doi.org/10.1080/09588221.2024.2428943

Ashley, K. D. (2017). Artificial intelligence and legal analytics: New tools for law practice in the digital age. Cambridge University Press.

https://doi.org/10.1017/9781316761380

Ariai, F., & Demartini, G. (2024). *Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models, and Challenges*. arXiv.

https://doi.org/10.48550/arXiv.2410.21306

Bertl, M., Price, S., & Draheim, D. (2026). Transforming legal texts into computational logic: Enhancing next generation public sector automation through explainable AI decision support. *International Journal of Cognitive Computing in Engineering*, *7*, 40– 57. https://doi.org/10.1016/j.ijcce.2025.07.003

Bench-Capon, T. Argument in Artificial Intelligence and Law. *Artificial Intelligence and Law* **5**, 249–261 (1997). https://doi.org/10.1023/A:1008242417011

Chen, J. (2025). BiLSTM-enhanced legal text extraction model using fuzzy logic and metaphor recognition. *PeerJ Computer Science*, *11*, e2697. https://doi.org/10.7717/peerj-cs.2697

Grover, C., Hachey, B., & Hughson, I. (2004). The HOLJ corpus: Supporting summarisation of legal texts. In *Proceedings of the 5th International Workshop on Linguistically Interpreted Corpora* (pp. 47–54). COLING.

Ilias Chalkidis, Manos Fergadiotis, Prodromos Malakasiotis, Nikolaos Aletras, and Ion Androutsopoulos. 2020. LEGAL-BERT: The Muppets straight out of Law School. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2898–2904, Online. Association for Computational Linguistics.

Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Yildirim, T., & Öztürk, M. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies, 29*(5), 19343– 19366. https://doi.org/10.1007/s10639-024-12574-6

Lee, S., Choe, H., Zou, D., & Jeon, J. (2025). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*, 1–25. https://doi.org/10.1080/10494820.2025.2498537

Merino-Campos, C. (2025). The Impact of Artificial Intelligence on Personalized Learning in Higher Education: A Systematic Review. *Trends in Higher Education*, *4*(2), 17. https://doi.org/10.3390/higheredu4020017

Oyewole, A. T., Adeoye, O. B., Addy, W. A., & Okoye, C. C. (2024). Automating financial reporting with natural language processing: A review and case analysis. *World Journal of Advanced Research and Reviews*, *21*(3), 575–589. https://doi.org/10.30574/wjarr.2024.21.3.0688

Shrivastav, D., Malathi. H, Kolaventi, S. S., & Patra, B. (2024). Integrating Natural Language Processing in Medical Information Science for Clinical Text Analysis. *Seminars in Medical Writing and Education*, *3*(1), 513. https://doi.org/10.56294/mw2024513

Singh, A., Joshi, A., Jiang, J. *et al.* A survey of classification tasks and approaches for legal contracts. *Artif Intell Rev* **58**, 380 (2025). https://doi.org/10.1007/s10462-025- 11359-8

Vadlapati P. EaseText: Enhancing Comprehension of Complex Sentences Using LLMs. J Artif Intell Mach Learn & Data Sci 2022, 1(1), 1560-1563. DOI: doi.org/10.51219/JAIMLD/praneeth-vadlapati/349

Zawacki-Richter, O., Marín, V.I., Bond, M. *et al.* Systematic review of research on artificial intelligence applications in higher education – where are the educators?. *Int J Educ Technol High Educ* **16**, 39 (2019). https://doi.org/10.1186/s41239-019-0171-0

MATH HANDWRITING CONVERSION USING AI APPROACHES: A LIGHTER ALTERNATIVE TO SCANNING

Hien Nguyen Phuoc[1], Nghia Cao Trong[2], Linh, Nguyen Hoang Anh[3]

Abstract

Handwritten mathematical expression recognition is a complex task due to the diverse symbols and two-dimensional spatial arrangements involved. Traditional scanning methods for digitizing handwritten math are often resource-intensive and lack real-time interactivity. This research explores AI-driven handwriting conversion as a lighter, more efficient alternative to scanning. Using the AZ Handwritten Alphabets dataset for initial training and a newly collected dataset of 1,000 handwritten mathematical inputs for evaluation, our model achieves an accuracy of 87%. A key practical challenge is the scarcity of large-scale, math-specific handwriting datasets, which limits recognition performance. We propose expanding training data with specialized math handwriting samples, optimizing neural network architectures to better capture spatial relationships, and incorporating adaptive learning techniques to improve accuracy and generalization. This study demonstrates the potential of AI-based handwriting conversion to provide real-time, accurate math digitization, with clear directions for future improvements.

1. Introduction

Digitizing handwritten mathematical expressions is essential for educational, scientific, and professional workflows. Traditional scanning methods convert paper-based math into digital formats but tend to be cumbersome, slow, and require significant hardware resources. AI-based handwriting recognition offers a promising lighter alternative that can process handwritten math in real-time on standard digital devices such as tablets and smartphones.

Mathematical handwriting recognition poses unique challenges compared to general text recognition due to the complex two-dimensional layout and the large variety of symbols. This research focuses on developing an AI model capable of converting handwritten math into digital text with high accuracy, using publicly available handwriting datasets and a custom-collected test set.

2. Related Work

Recent advances in AI, particularly deep learning, have significantly improved handwriting recognition across languages and scripts. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and attention mechanisms have been widely

[1] Student, High School for The Gifted, VNUHCM, Vietnam
[2] Student, High School for The Gifted, VNUHCM, Vietnam
[3] Faculty of Interdisciplinary Sciences, VNUHCM - University Of Science

adopted for text recognition tasks, achieving accuracy levels exceeding 90% on standard handwriting datasets.

However, mathematical handwriting recognition remains less explored due to the lack of large, diverse datasets and the complexity of spatial symbol relationships. The recently introduced MathWriting dataset, the largest online handwritten mathematical expression dataset to date, contains over 230,000 human-written samples and 400,000 synthetic ones, facilitating both online and offline handwritten math recognition research [1][2]. Advanced architectures such as encoder-decoder models with attention mechanisms and syntax-aware networks have demonstrated improved recognition rates on benchmark datasets like CROHME and HME100K [3][4].

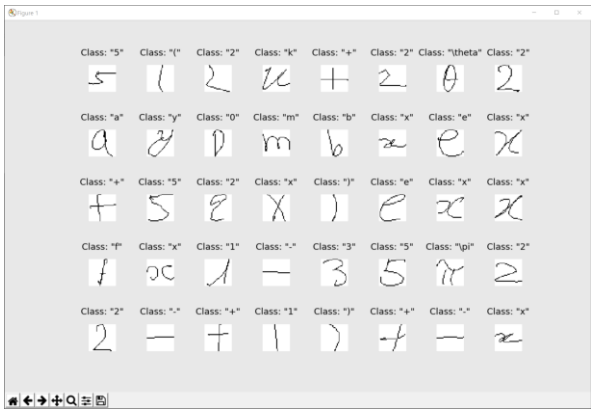


Fig 1. *CROHME dataset extractor*

Despite these advances, practical implementations often face challenges related to limited math-specific data and handwriting style variability.

3. Data and Methodology

3.1 Training Data

Our model is initially trained on the AZ Handwritten Alphabets dataset from Kaggle, which includes over 370,000 images of handwritten English alphabets in CSV format [5].

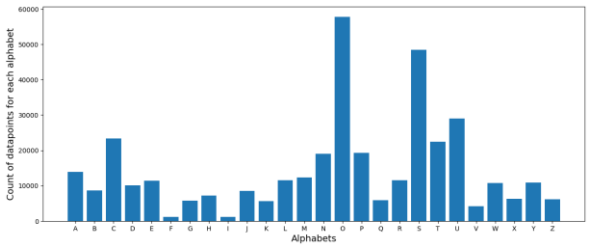


Fig 2. *Count of data points for each alphabet*

While this dataset provides a solid foundation for character recognition, it lacks mathematical symbols and two-dimensional structural information.

Fig 3. *Example of training data in grayscale.*

3.2 Evaluation Dataset

To evaluate the model's performance on math handwriting, we collected a new dataset of 1,000 handwritten mathematical inputs. This dataset consists of common mathematical terms and symbols written by diverse users, reflecting realistic handwriting variations.

3.3 Model Architecture

The model employs a deep learning architecture combining convolutional layers for feature extraction with sequence modeling components to interpret character sequences. Future iterations aim to integrate Capsule Networks or attention-based Transformer layers to improve spatial relationship understanding.

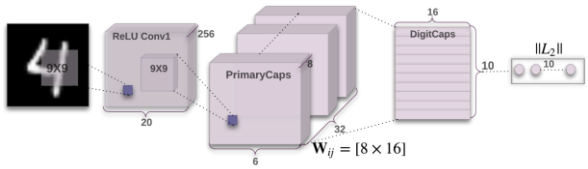


Fig 4. *A simple CapsNet with 3 layers.* [6]

4. Results and Evaluation

The model achieved an **accuracy of 87%** on the 1,000-sample test dataset, correctly recognizing the majority of handwritten math inputs. This performance demonstrates the model's practical viability as a math handwriting conversion tool.

Processing times ranged from 1 to 2 seconds per input, suitable for real-time applications on modern devices.

Sample	Correct quantity	Wrong quantity	Percentage of Correction
A	18	4	82%
B	4	0	100%

C	5	5	50%
D	6	0	100%
E	18	2	90%
F	1	0	100%
G	3	3	50%
H	2	0	100%
I	10	16	38%
J	0	0	NaN
K	0	0	NaN
L	8	2	80%
M	5	2	71%
N	20	1	95%
O	11	5	69%
P	4	2	67%
Q	1	0	100%
R	20	2	91%
S	2	1	67%
T	13	6	68%
U	6	0	100%
V	5	0	100%
W	0	0	NaN
X	1	0	100%
Y	4	0	100%
Z	0	0	NaN

5. Practical Challenges

A major limitation is the lack of large-scale, publicly available datasets specifically designed for handwritten mathematical expressions. Existing datasets primarily focus on

alphabetic or numeric handwriting and do not capture the complexity of math notation, including superscripts, fractions, and spatial symbol arrangements.

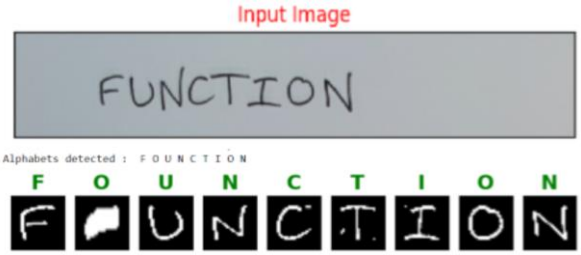


Fig 5. *Example of result*

This scarcity restricts the model's exposure to diverse math handwriting styles and limits its ability to generalize across different users and notation complexities [1][4].

6. Future Development Directions

To enhance the system's accuracy and robustness, we propose the following:

- Expand Training Data: Incorporate math-specific handwriting datasets such as MathWriting, which contains over 230,000 human-written and 400,000 synthetic math samples, to better represent the diversity of mathematical notation [1][2].
- Optimize Model Architectures: Develop and test architectures that explicitly model spatial relationships, such as Capsule Networks and Transformer-based attention mechanisms, to improve recognition of complex math layouts [3][4].
- Implement Adaptive Learning: Use test-time training or meta-learning approaches to enable the model to adapt dynamically to individual handwriting styles, improving personalized accuracy.
- Symbol-Specific Enhancements: Focus on targeted recognition improvements for frequently used mathematical symbols and structures through data augmentation and specialized training.

7. Conclusion

This study demonstrates that AI-based handwriting conversion provides a lighter, more efficient alternative to traditional scanning for digitizing handwritten mathematical expressions. With an 87% accuracy on a realistic test set, the approach shows promise for real-time applications in education and research.

Addressing the practical challenges of limited math handwriting data and refining model architectures will be crucial to advancing this technology. Our proposed development directions aim to guide future research toward more accurate and adaptable math handwriting recognition systems.

References

1. Fadeeva, A., et al. (2024). *MathWriting: A Dataset For Handwritten Mathematical Expression Recognition*. arXiv preprint arXiv:2404.10690. Retrieved from https://arxiv.org/abs/2404.10690

2. Patel, S. (2018). *A-Z Handwritten Alphabets in CSV Format*. Kaggle. Retrieved from https://openreview.net/forum?id=bxwWikAXSy&noteId=i2SNi4bTfK

3. Bian, X., et al. (2021). *Handwritten Mathematical Expression Recognition via Attention Aggregation based Bi-directional Mutual Learning*. arXiv preprint arXiv:2112.03603. Retrieved from https://arxiv.org/abs/2112.03603

4. Yuan, Y., Liu, X., Dikubab, W., et al. (2022). Syntax-Aware Network for Handwritten Mathematical Expression Recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Retrieved from https://openaccess.thecvf.com/content/CVPR2022/papers/Yuan_Syntax-Aware_Network_for_Handwritten_Mathematical_Expression_Recognition_CVPR_2022_paper.pdf

5. Fadeeva, A., et al. (2024). *MathWriting: A Dataset For Handwritten Mathematical Expression Recognition*. OpenReview. Retrieved from https://www.sciencedirect.com/science/article/pii/S1877050924007014

6. Sabour, S., Frosst, N. and Hinton, G. E.Sabour, S., Frosst, N. and Hinton, G. E. (2017) Dynamic Routing Between Capsules, arXiv.org. Available at: https://arxiv.org/abs/1710.09829

KHÁM PHÁ CÁC PHƯƠNG PHÁP TÍCH HỢP KỸ THUẬT SỐ TRONG GIẢNG DẠY TIẾNG ANH CHO SINH VIÊN

Nguyễn Thị Luyến[1], Nguyễn Thúy Ngọc[2]

Tóm tắt: Tích hợp kiến thức kỹ thuật số (IDL)vào quá trình dạy và học ELT (giảng dạy tiếng Anh (EFL)) đã trở thành một thách thức đáng kể đối với giảng viên (GV) ở thế kỷ 21. IDL đã thay đổi căn bản cách người học tiếp cận nền giáo dục của họ. Nhiều người học EFL, phụ thuộc vào các tiện ích tiên tiến và các công cụ dựa trên web để tạo và chia sẻ ý nghĩa. Sự thay đổi mô hình này ảnh hưởng đến chương trình giảng dạy sinh viên (SV) EFL tại Đại học Kinh tế- Kỹ Thuật Công nghiệp (UNETI). Các phương pháp được thực hiện trong nghiên cứu (NC) này sử dụng cách trình bày bằng đồ họa để đánh giá một cách chủ quan mô tả các cấp độ nâng cao của việc học EFL, nhằm xác định bảy lĩnh vực kỹ năng đóng vai trò quan trọng trong việc học EFL, bao gồm sự thành thạo về thiết bị và phần mềm, làm chủ dữ liệu và thông tin, giao tiếp hiệu quả và cộng tác, tạo nội dung bằng máy tính, đảm bảo an ninh trực tuyến, năng lực giải quyết vấn đề và các kỹ năng liên quan đến nghề nghiệp. Tóm lại, các tiêu chuẩn phổ cập DL do UNESCO thiết lập có thể được được áp dụng hiệu quả trong bối cảnh học EFL. Bảy lĩnh vực kỹ năng máy tính (PC) không thể thiếu trong quá trình học tập được tích hợp một cách có hệ thống vào chương trình dự bị, các bài tập, đảm bảo cách tiếp cận toàn diện và cập nhật về giáo dục EFL trong thời đại IDL.

Từ khóa: Tích hợp kiến thức kỹ thuật số, dạy và học tiếng Anh, sinh viên EFL, thực hành tiếng Anh, thời đại IDL

INVESTIGATING DIGITAL LITERACY METHODS FOR TEACHING ENGLISH TO EFL STUDENTS

Abstract: One of the biggest challenges facing educators in the 21st century is Integrating digital literacy (IDL) into the teaching and learning process in ELT (English Language Teaching) classrooms. Beyond just being a technical ability, IDL has completely changed how students approach their education. In reaction to this change, a large number of English language learners—including those in university now mostly rely on sophisticated software and web-based resources to produce and distribute meaning.This paradigm shift has been directed by the global IDL competency framework established by UNESCO, in addition to having an impact on the English curriculum for students at the University of Economics-Technology for Industries (UNETI) . This study uses a graphical depiction to describe the advanced degrees of English language proficiency in a subjective manner. It lists seven

[1] ThS., Khoa Ngoại Ngữ, Đại học Kinh tế- Kỹ thuật Công nghiệp (UNETI), Hà Nội, *Email: ntluyen.nn@uneti.edu.vn*

[2] ThS., Khoa Ngoại Ngữ, Đại học Kinh tế- Kỹ thuật Công nghiệp (UNETI), Hà Nội

essential higher education skill areas that are vital to UNETI EFLstudents' acquisition. These include competence with tools and software, data and information mastery, efficient communication and teamwork, computer-generated content, maintaining online security, problem-solving skills, and career-related competencies. In conclusion, studying English can benefit greatly from the application of the digital literacy standards set by UNESCO for the larger digital society. The preparatory, intermediate, and finishing exercises at UNETI methodically incorporate these seven computer abilities, which are essential to the learning process and guarantee a thorough and modern approach to teaching English in the digital age.

***Keywords**:* Integrating digital literacy methods, ELT, EFL students, English practices, digital age.

I.GIỚI THIỆU

IDL rất quan trọng đối với người học ngôn ngữ [1], cho phép họ tiếp cận nguồn tài liệu ngôn ngữ phong phú, tham gia vào trải nghiệm học tập năng động và tương tác, trở thành người giao tiếp hiệu quả và tự điều chỉnh theo nhu cầu của thị trường việc làm thế kỷ 21 [2]. Khi công nghệ (IT) hỗ trợ cách chúng ta học và giao tiếp, khả năng đọc viết IDLvẫn là một kỹ năng thiết yếu để người học ngôn ngữ phát triển, đã nổi lên như một kỹ năng nền tảng không thể thiếu đối với người học ngôn ngữ, đóng vai trò then chốt trong việc chuẩn bị cho họ trước những thách thức và cơ hội của giáo dục và truyền thông hiện đại [3]. IDL trao quyền cho người học ngôn ngữ khả năng điều hướng và khai thác các nguồn tài nguyên IDL phong phú hiện có giúp dễ dàng truy cập trực tuyến, trang bị cho người học những kỹ năng tìm kiếm, đánh giá và sử dụng nội dung số một cách hiệu quả. Hơn nữa, IDL thúc đẩy một môi trường học tập năng động và hấp dẫn [4]. Người học ngôn ngữ có thể tận dụng vô số công cụ và nền tảng IDL để nâng cao kỹ năng ngôn ngữ của mình. Từ các ứng dụng học ngôn ngữ tương tác và các khóa học ngôn ngữ trực tuyến đến cộng đồng trao đổi ngôn ngữ và tài nguyên đa phương tiện, IDL cho phép người học tiếp cận những trải nghiệm học ngôn ngữ đa dạng và phong phú, bao gồm thực hành ngôn ngữ theo thời gian thực, sử dụng nội dung đa phương tiện và tham gia vào cộng đồng ngôn ngữ ảo, tất cả đều góp phần giúp việc tiếp thu ngôn ngữ trở nên hiệu quả và thú vị hơn [5]. Hơn nữa, IDL khuyến khích người học ngôn ngữ trở thành người sáng tạo và người giao tiếp theo cách riêng của họ. Giao tiếp hiệu quả thường dựa vào nhiều phương tiện IDL khác nhau, bao gồm văn bản, hình ảnh, âm thanh và video.

Người học ngôn ngữ hiểu biết về IDL có thể tạo ra và chia sẻ những sáng tạo của riêng họ. Khả năng này rất quan trọng để thể hiện ý tưởng, tham gia giao tiếp đa văn hóa và giúp sử dụng ngôn ngữ thành thạo [6]. Ngoài ra, IDLgiúp người học ngôn ngữ phù hợp với nhu cầu của lực lượng lao động hiện đại.

II. CƠ SỞ LÝ THUYẾT

Những cải tiến mang tính đổi mới trong xã hội đã bắt đầu trong giảng dạy. Đồng ý với [7] về nhu cầu hợp tác đổi mới trong quá trình học tập liên quan đến sự gia tăng trình độ hiện

đại ở thế kỷ 21. Cho đến nay, những tiến bộ trong đổi mới dữ liệu và web về giảng dạy xuất phát từ nguồn dữ liệu máy tính là vô tận [8]. Theo cách này, khả năng chọn và sắp xếp dữ liệu nâng cao đặc biệt quan trọng vì đó không phải là tất cả dữ liệu có giá trị và hỗ trợ. Khả năng sàng lọc dữ liệu và thông tin tồn tại trong thế giới máy tính được gọi là trình độ IDL.

Kiến thức kỹ thuật số (DK) đang dần trở thành một phần thiết yếu trong hầu hết mọi khía cạnh của cuộc sống con người, trong đó có việc dạy EFL. Về cơ bản, các cơ sở giáo dục tập trung các khóa học của họ vào ba lĩnh vực: mức độ và hiểu biết về khả năng sử dụng PC, sử dụng kiến thức PC để tiến bộ trong một số khả năng cốt lõi nhất định. nhận biết và sử dụng các kỹ thuật nhất định để tạo ra các kỹ năng nâng cao. Những phản ánh về năng lực đã tập trung và dựa trên PC của GV trong quá trình chuẩn bị ban đầu [9], và dựa vào hiểu biết sâu sắc của GV EFL và SV về trình độ PC [10]. Vì vậy, có thể kết luận rằng DK là một kỹ năng hoặc khả năng cá nhân sử dụng IT một cách thích hợp để truy cập, quản lý, xây dựng kiến thức và giao tiếp với người khác phù hợp với bối cảnh.

Ngày nay, hầu hết SV đều rất hào hứng khám phá nhiều loại hình IDL. Do đó, các cơ quan chức năng trong lĩnh vực giáo dục đang cố gắng tìm ra những cách hiệu quả để sử dụng phương tiện IDL nhằm nâng cao khả năng học tập của SV [11]. Việc IDL vào quá trình dạy và học đang trở thành một thách thức đối với GV trong thế kỷ 21. Nói cách khác, IDL đã không chỉ là một kỹ thuật. những kỹ năng mà SV phải thành thạo, chủ yếu do GV kiểm soát [12].

Khung kỹ năng đọc viết IDL có thể áp dụng trực tuyến trên toàn cầu được gọi là Khung kiến thức kỹ thuật số toàn cầu (DLGF) được phát triển bởi UNESCO (Tổ chức Giáo dục, Khoa học và Văn hóa của Liên Hợp Quốc) vào năm 2018. Trình độ hiểu biết về IDL này được phát triển với hy vọng rằng các chỉ số về trình độ IDL có thể được điều chỉnh và áp dụng không chỉ cho cộng đồng hiểu biết về DL ở Châu Âu hoặc địa điểm NC này, nhưng cũng có thể được áp dụng ở bất kỳ nước nào trên thế giới. Vì vậy, tất cả cư dân trên thế giới đều có chỉ số kỹ năng chung như nhau về 1) kỹ năng số; 2) kiểm soát thông tin và dữ liệu; 3) giao tiếp và hợp tác; 4) tạo nội dung số, 5) bảo mật; 6) giải quyết vấn đề; và 7) các kỹ năng liên quan đến nghề nghiệp. Mỗi lĩnh vực kỹ năng đều có các chỉ số thực tế về trình độ IDL như một phần của quá trình học tập tiến bộ.

Ở Indonesia, NC về năng lực IDL đang được thực hiện ở một số trường đại học, đặc biệt là liên quan đến việc phân tích bảy yếu tố của năng lực kỹ thuật số do JISC phát triển và việc sử dụng e-learning trong giáo dục [13]. NC sâu hơn về thực hành DK ở cấp trung học được thực hiện bởi [14], họ đã khảo sát việc sử dụng DK của GV thuộc mọi hoàn cảnh. tạo ra các bài học EFL và phản ứng của SV đối với việc sử dụng phương tiện IDL theo khung lý thuyết ACOT (Lớp học ngày mai của Apple) do Apples Company phát triển. Ứng dụng IDL vào các quá trình học EFL khác do [15]. NC này được xây dựng dựa trên những quan sát của NNC, cho thấy rằng mặc dù GV và SV hiện là những người tích cực sử dụng IT nhưng SV và GV lại không muốn sử dụng IDL để nâng cao kỹ năng đọc viết.

Nhiều NC khác nhau liên quan đến DK và thực hành chuyên sâu DK trong quá trình học tập, đặc biệt là bằng EFL. với sự phát triển của IT, hầu hết các NC về DK ở đại học đều hướng

tới việc sử dụng học tập trực tuyến hoặc đề cập đến một số khía cạnh hoặc lĩnh vực kỹ năng của quá trình học tập. Nhưng trong NC này, các chỉ số về trình độ IDL không chỉ được sử dụng để bao gồm một kỹ năng mà còn bao gồm bảy lĩnh vực kỹ năng kể trên. Điểm mới của NC nằm ở chỉ số được áp dụng để khám phá các hoạt động phổ cập IDL sử dụng Khung tham chiếu toàn cầu về phổ cập IDL do UNESCO biên soạn và phát triển vào năm 2018.

III. PHƯƠNG PHÁP NGHIÊN CỨU

3.1. Thiết kế

NC định tính mô tả đặc biệt phù hợp cho mục đích kiểm tra toàn diện và giải thích chi tiết về bối cảnh kiến thức IDL tại UNETI. Cách tiếp cận này là công cụ mang lại sự hiểu biết sâu sắc và toàn diện về các thực tiễn, thách thức và cơ hội đa dạng liên quan đến DK. Thời gian NC kéo dài ba tháng, bắt đầu vào tháng 2 và kết thúc vào tháng 4 năm 2024, vì cần phải có khoảng thời gian NC sâu rộng để nắm bắt được bản chất năng động và đang phát triển của hoạt động xóa mù DL trong một khoảng thời gian đáng kể. Hơn nữa, NNC đang xem xét chỉ số phổ cập IDL do UNESCO thiết lập vào năm 2018, bao gồm bảy lĩnh vực kỹ năng riêng biệt, mỗi lĩnh vực được đặc trưng bởi các chỉ số cụ thể về khả năng đọc viết trong Bảng 1, đóng vai trò là điểm tham chiếu nền tảng cho NC.

Bảng 1. *Lĩnh vực năng lực và chỉ số năng lực về kiến thức IDL*

Số lượng	**Lĩnh vực năng lực**	**Chỉ số năng lực**
1	Phần cứng và phần mềm	-Truyền thông và hợp tác -Hoạt động vật lý của các thiết bị IDL
2	Dữ liệu thông tin và xóa mù IDL	-Duyệt, lọc dữ liệu, thông tin và nội dung số, -Thông tin, nội dung đánh giá -Quản lý dữ liệu, thông tin và nội dung số
3	Truyền thông và hợp tác	- Tương tác thông qua công nghệ số, -Chia sẻ qua công nghệ số - Tham gia vào quyền công dân thông qua công nghệ IDL. - Hợp tác thông qua công nghệ số Netiket -Quản lý danh tính IDL
4	Sáng tạo nội dung số	-Sáng tạo nội dung số -Tích hợp và phân tích nội dung IDL - Bản quyền và giấy phép - Lập trình
5	Bảo vệ	- Bảo vệ dữ liệu cá nhân và quyền riêng tư

		- Bảo vệ sức khỏe và hạnh phúc. - Bảo vệ môi trường
6	Giải quyết vấn đề	- Giải quyết vấn đề kỹ thuật -Xác định nhu cầu và đáp ứng công nghệ - Sử dụng công nghệ số một cách sáng tạo
7	Liên quan đến nghề nghiệp	-Vận hành công nghệ số cho một lĩnh vực nhất định
	Năng lực	- Giải thích và xử lý dữ liệu số

3.2. Đối tượng tham gia NC

Với sự tham gia của 16 GV tiếng Anh có kinh nghiệm, tất cả đều là thạc sỹ ngôn ngữ Anh, mà được lựa chọn dựa trên cam kết đã được thể hiện của họ trong việc triển khai các phương pháp học tập IDL hoàn toàn trong môi trường lớp học tương ứng của họ. Tiêu chí lựa chọn này rất quan trọng vì nó đảm bảo rằng NC thu hút sự tham gia của các GV có trình độ chuyên môn về phương pháp sư phạm IDL và do đó, có vị trí tốt để cung cấp những hiểu biết sâu sắc có giá trị về tình trạng hiểu biết về DL trong bối cảnh giáo dục EFL.

3.3. Công cụ và phương pháp thu thập dữ liệu

Để thu thập dữ liệu toàn diện và đáng tin cậy cho NC này, nhà NNC đã sử dụng các hướng dẫn phỏng vấn có cấu trúc như một trong những công cụ chính để tạo điều kiện cho các cuộc phỏng vấn sâu sắc với các bên liên quan chính. Những cuộc phỏng vấn này được chuẩn bị cẩn thận, liên quan đến việc đặt ra một loạt câu hỏi thích hợp để gợi ra những câu trả lời phong phú và đa sắc thái với các câu hỏi phỏng vấn đã được thông qua bằng cách kiểm tra trước và xác nhận nghiêm ngặt. Các cuộc phỏng vấn này được thực hiện qua mạng, nhấn mạnh khả năng thích ứng của các phương pháp NC hiện đại để phù hợp với thực tế của truyền thông IDL.

Sự tham gia có chủ ý của nhiều bên liên quan nhằm cung cấp cái nhìn toàn diện và toàn diện về thực hành DK trong bối cảnh UNETI. Ngoài các cuộc phỏng vấn có cấu trúc, các NNC còn khai thác tiện ích của các bảng quan sát như một công cụ thu thập dữ liệu bổ sung. Cách tiếp cận này yêu cầu các NNC phải trực tiếp vào lớp học để quan sát và ghi lại các phương pháp thực hành IDL và phương pháp sư phạm cụ thể đang được các GV sử dụng. Cách tiếp cận kép để quan sát cho phép hiểu biết toàn diện về thực tiễn học tập IDL, bao gồm cả khía cạnh vật lý và IDL.

3.4. Phân tích dữ liệu

Dữ liệu được thu thập cho NC này tạo thành một kho thông tin định tính có giá trị liên quan đến việc triển khai các phương pháp học tập IDL trong phạm vi giảng dạy EFL ở UNETI. Bộ dữ liệu này bao gồm nhiều hiểu biết sâu sắc, nắm bắt ý kiến, nhận thức và thái độ của các cá nhân khác nhau tham gia NC. Những cá nhân này, bao gồm GV tiếng Anh, quản trị viên và các bên liên quan khác có mối liên hệ chặt chẽ, đã cung cấp góc nhìn đa diện và toàn diện về sự

phức tạp của học tập IDL. Các NNC đã áp dụng phương pháp định tính toàn diện. Quá trình phân tích diễn ra qua một số giai đoạn được kết nối với nhau. Trong giai đoạn đầu này, dữ liệu định tính thô được tổ chức một cách có hệ thống và cô đọng thành các đơn vị có thể quản lý được, cho phép kiểm tra thông tin có cấu trúc và tập trung hơn.

Giai đoạn tiếp theo liên quan đến hiển thị dữ liệu. Dữ liệu cô đọng được trình bày ở định dạng dễ tiếp cận bằng trực quan, thường bao gồm biểu đồ, bảng biểu và các phương tiện trực quan khác.

IV. KẾT QUẢ VÀ THẢO LUẬN

4.1. Kết quả nghiên cứu

Sử dụng DK trong dạy EFL

Một số NC đã chỉ ra rằng GV sử dụng PC và thiết bị làm nguồn DK để dạy EFL. Điều này phù hợp với ý tưởng của [16] rằng việc học ngôn ngữ với sự hỗ trợ của PC có thể được sử dụng để thúc đẩy các kỹ năng IDL. Về việc tích hợp điện thoại thông minh vào việc học ngôn ngữ của GV, là một phương pháp giảng dạy sáng tạo trong giảng dạy EFL [17]. Trong thực tế giảng dạy EFL, hầu hết GV EFL tại UNETI đều tin rằng IT có thể làm cho quá trình học tập trở nên thú vị hơn và giúp SV đạt được mục tiêu học tập. NC tập trung vào mặt tích cực của việc tích hợp IT trong lớp học. Sự thành công của việc tích hợp IT trong lớp học được thể hiện ở sự hiểu biết tích cực của GV về việc sử dụng IT [18]. Khi thực hành kỹ năng số, GV cần nhận thức được vai trò thiết yếu của mình trong việc quản lý điều kiện lớp học. Bằng cách đóng những vai trò quan trọng này, GV có thể thực hành DK một cách hiệu quả. Phát hiện này phù hợp với các tài liệu hiện có cho rằng vai trò của GV có tác động đáng kể đến khả năng đọc viết và tính đa dạng về IDL [19]. Cả hai NNC đều sử dụng nhiều kỹ năng đọc (đọc, viết và DL). Để minh họa quan điểm của mình, các GV đã sử dụng PowerPoint, phim và bài hát để dạy tiếng Anh. Việc thiếu sự hỗ trợ của SV và sự phổ biến của các thiết bị đã tạo điều kiện thuận lợi cho việc sử dụng kiến thức một cách đa dạng. Theo tài liệu, kiến thức đa dạng không thể tách rời khỏi IDL [20] (McCord, 2015). Trong bài viết này, các GV đã phản ánh:

UNETI đã áp dụng đổi mới và dữ liệu trong việc giảng dạy. Vấn đề này có thể được chứng minh thông qua trong các trường học, tính toán các phòng PC có sẵn cho SV. UNETI cung cấp máy chiếu và một số thiết bị máy tính bàn có sẵn. Việc áp dụng đổi mới trong giảng dạy không phải như trong chương trình học mà còn trong môi trường trường học. Những khám phá này được các NNC phân tích và phỏng vấn 16 GV tiếng Anh.

Các GV bày tỏ rằng thực tế là trình độ thông thạo trên PC đã được kết nối trong quá trình học tập tại UNETI. Đặc biệt trong việc học EFL, việc thực hành thành thạo trên PC phải được thực hiện ở các phạm vi năng lực khác nhau. Thực hành đọc viết nâng cao được thực hiện trong quá trình học EFL được thực hiện từ đầu đến cuối. Năng lực thực hành giáo dục vi tính hóa liên quan đến giáo dục dữ liệu và thông tin hầu hết bao gồm các bài tập chuẩn bị và bài tập trung tâm trong quá trình chuẩn bị học EFL. Việc sử dụng IT đã giúp SV được hưởng lợi từ một môi trường nơi họ có thể củng cố kỹ năng ngôn ngữ của mình. Các GV EFL đồng ý rằng

trình độ IDL có thể cải thiện cả bốn kỹ năng; các GV bổ sung thêm việc dạy EFL bằng một bài hát có thể bao gồm cả bốn kỹ năng ngôn ngữ. Tuy nhiên, một số GV gợi ý rằng DK đặc biệt thúc đẩy khả năng nghe và nói nhiều hơn các kỹ năng khác,và chỉ ra kỹ thuật nghe nhìn và hình ảnh là hiệu quả nhất. Tóm lại, các GV trong NC này đã tích hợp IT để nâng cao kỹ năng EFL.

Việc sử dụng DK giống như các phương pháp thực hành IDL trong việc giảng dạy EFL qua năm giai đoạn. Đầu tiên, xác định các thành phần của DK cũng như các sản phẩm và thực hành IDL học thuật đích thực. Thứ hai, phân tích các mô tả về các thành phần DK liên quan đến các sản phẩm và thực hành DL học thuật đích thực. Thứ ba, tích hợp các mô tả về DK vào kết quả học tập. Thứ tư, đưa DK vào các sản phẩm học thuật. Thứ năm, sử dụng các phương pháp thực hành IDL mang tính học thuật đích thực trong quá trình học tập và đánh giá DK như những năng lực xuyên suốt chương trình học. Giai đoạn một, hai và ba được trình bày trong bảng 2 dưới đây.

Bảng 2. *Xác định thực hành IDL trong dạy học EFL*

	THÀNH PHẦN KỸ THUẬT SỐ (KTS)				
Các giai đoạn	**KẾT QUẢ DK DỰ KIẾN (khả năng ...)**	**SẢN PHẨM DK**	**KINH NGHIỆM HỌC TẬP**	**PHÂN LOẠI DK CỦA Bloom**	**MÔ TẢ DK**
GIAI ĐOẠN 1: Thu nhận Kiến thức					
Truy cập thông tin	Xác định nguồn thông tin, truy xuất thông tin và thu thập thông tin			1) Ghi nhớ (Sao chép, Xác định, Tìm kiếm, v.v.)	1.Truy cập thông tin: : xác định nguồn thông tin bao gồm việc thu thập và truy xuất thông tin
Quản lý thông tin	Sắp xếp và phân loại thông tin đa dạng từ cổng kiến thức dựa trên internet	1. Câu trả lời viết		2) Hiểu biết (Chú thích, Liên kết, v.v.)	2. Quản lý thông tin: đánh giá tính hợp lệ và xác thực của các nguồn thông tin
GIAI ĐOẠN 2: Làm sâu sắc Kiến thức					

Đánh giá thông tin	Đưa ra nhận định về tính đầy đủ, tính cập nhật, tính hữu ích, chất lượng, độ tin cậy và tính hiệu quả của thông tin	2. Bài luận 3.Trang chiếu 4.Video	Lấy sinh viên làm trung tâm Các dự án	3.Áp dụng thông tin (thực hiện, nói rõ, v.v.)	3.Đánh giá thông tin: đưa ra nhận định về tính đầy đủ, tính cập nhật, tính hữu ích, chất lượng, độ tin cậy hoặc tính hiệu quả của thông tin
Tích hợp thông tin	Tổng hợp, tóm tắt, so sánh và đối chiếu thông tin từ nhiều nguồn bằng cách sử dụng kỹ năng số để tạo ra văn bản, biểu đồ và hình ảnh tương phản và có liên quan với nhau	5. Bản tin 6. Hình ảnh thông tin 7. Ghi âm	1.Học tập dựa trên dự án 2.Học tập dựa trên tình huống	4. Phân tích (Phân tích, Chia nhỏ,Tương quan, Phá vỡ cấu trúc, v.v.) |	4. Tích hợp thông tin: : diễn giải và trình bày thông tin bằng các công cụ CNTT, yêu cầu khả năng tổng hợp, tóm tắt, so sánh và đối chiếu thông tin từ nhiều nguồn thông tin dạng văn bản, biểu đồ và hình ảnh
GIAI ĐOẠN 3: Sáng tạo Kiến thức					
Sáng tạo thông tin	Tạo ra thông tin mới bằng cách thích nghi, áp dụng, thiết kế, sáng tạo hoặc sáng tác một cách kỹ thuật số	8. thuyết trình 9. Bài viết 10. Báo cáo	3. Học tập dựa trên các yêu cầu, mục tiêu	5) Đánh giá & Tranh luận (Xác nhận, Kiểm tra, v.v.)	5. Sáng tạo kiến thức: : tạo ra kiến thức mới bằng cách thích nghi, áp dụng, thiết kế, sáng tạo hoặc

					sáng tác thông tin
Truyền đạt thông tin	Truyền tải thông tin một cách nhanh chóng, thuyết phục và đến một lượng lớn khán giả bằng cách sử dụng phương tiện phù hợp và thích nghi nhất để trình bày thông tin một cách đúng đắn trong nhiều bối cảnh xã hội-văn hóa khác nhau	11…..	4……	6) Sáng tạo (Soạn thảo, Tổng hợp, Podcasting, v.v.)	6. Truyền đạt thông tin: : truyền tải thông tin một cách nhanh chóng, thuyết phục và đến một lượng lớn khán giả, đồng thời thích nghi và trình bày thông tin một cách đúng đắn trong nhiều bối cảnh xã hội-văn hóa

Chiến lược của GV thực hành DL trong giảng dạy EFL

NC này cho thấy các GV dạy EFL tin rằng IT không thể tách rời khỏi sự thành công của việc dạy và học. GV có kỹ năng thực hành DK. Bất kể các loại IT thứ cấp và việc chuẩn bị cho SV làm quen với IT, GV vẫn có thái độ tích cực về tầm quan trọng của việc tích hợp IT vào quá trình học tập. Bài viết này có thể đề cập đến thực tiễn phổ cập IDL ở UNETI. Chi tiết cách GV tích hợp IT vào kiến thức số đã trả lời câu hỏi làm thế nào để áp dụng DK vào thực tế trong quá trình học tập. Phát hiện này cung cấp một góc nhìn khác với các NC trước đây. [20] chỉ NC cách sử dụng IT mà không tìm hiểu chi tiết cách sử dụng trong dạy EFL. Bài viết này cũng chỉ ra những thách thức gây khó khăn cho GV trong việc dạy DK và các chiến lược của họ để phá vỡ các rào cản. Vì vậy, các nhà hoạch định chính sách cần có những cải tiến để đạt được bình đẳng trong giáo dục IT. Tạo điều kiện cho lớp học có các trang thiết bị phù hợp, giúp GV có thêm động lực sáng tạo. Bằng cách sắp xếp sao lưu, IDL cho phép người hướng dẫn khắc phục các tình huống không lường trước được của hoàn cảnh lớp học. Phát hiện này phù hợp với lời giải thích của [21] về cơ bản cần có một kế hoạch dự phòng trong chương trình giáo dục.

4.2. Thảo luận

Nhìn chung, kết quả NC cho thấy UNETI có thể triển khai năng lực DK trong bảy lĩnh vực năng lực DK, bao gồm: phần cứng và phần mềm; kiến thức về thông tin và dữ liệu; giao tiếp và hợp tác; sáng tạo nội dung số; bảo vệ; giải quyết vấn đề; và năng lực liên quan đến nghề nghiệp.

Kiến thức về thông tin và dữ liệu là cốt lõi của năng lực DK. Những tiến bộ của IT và Internet hiện nay khiến nguồn thông tin số rất dồi dào [8]. NC sâu thêm, quản lý và đánh giá thông tin, dữ liệu trong thế giới số đóng vai trò rất quan trọng. Quá trình học EFL tại UNETI được chia thành ba phần là hoạt động sơ bộ, hoạt động cốt lõi và hoạt động kết thúc. Thực hành đọc viết được áp dụng IDL trong quá trình học EFL xuyên suốt thời gian diễn ra NC. Tuy nhiên, thực tiễn năng lực DK liên quan đến thông tin và kiến thức dữ liệu này hầu hết được thực hiện ở các hoạt động sơ bộ và hoạt động cốt lõi trong quá trình học EFL.

Một hình thức thực hành kiến thức dữ liệu được thực hiện trong các hoạt động sơ bộ là để SV nghiên cứu thông tin hoặc tài liệu tham khảo liên quan đến các chủ đề sẽ được thảo luận trong khóa học bằng ứng dụng tìm kiếm. GV quan sát độc lập để SV có thể có được hình ảnh hoặc sự hiểu biết đầu tiên về SV. Thông qua việc thực hành DK này, SV có thể khám phá, tìm kiếm, lọc và đánh giá dữ liệu cần thiết trong tài liệu học tập tùy thuộc vào chủ đề được đề cập trong hoạt động chính. Việc thực hành DK trong các hoạt động cốt lõi do SV cung cấp và hoàn thành nhìn chung phức tạp hơn. Hoạt động phân tích và đánh giá thông tin trong văn bản cũng như trả lời các câu hỏi quan trọng để phát triển kỹ năng đọc về các kỹ năng dữ liệu đọc viết cụ thể. Ngoài ra, việc giao cho SV các dự án phân tích lời bài hát trong các hoạt động cốt lõi sẽ hướng dẫn SV khả năng tìm kiếm, quản lý, lọc, xử lý và phát triển thông tin, dữ liệu và nội dung, các nguồn tài nguyên DL được tìm thấy cũng như điều chỉnh chúng cho phù hợp với nhu cầu và sự sáng tạo của SV.

Sự phát triển hiện nay của phương tiện truyền thông IDL có thể tự động thay đổi cách mọi người giao tiếp hoặc giao tiếp xã hội [22]. Các hình thức giao tiếp, tương tác được GV EFL thực hiện trong quá trình học bao gồm sử dụng ứng dụng gmeet hoặc zoom, được UNETI sử dụng để liên lạc và tương tác ảo trong quá trình học trực tuyến. Trong giờ học, các hình thức giao tiếp, tương tác được thực hiện giữa GV và SV bằng cách bật âm thanh và tận dụng tính năng nhắn tin trên gmeet hoặc zoom. Sự tương tác diễn ra trong thời gian thực bằng cách sử dụng các thiết bị IDL có sẵn để có thể thu được thông tin đầy đủ. Việc giao tiếp diễn ra một cách lịch sự, luân phiên theo nội quy lớp học đã được đưa ra trước đó để toàn bộ SV cũng có thể nghe được sự truyền tải.

Việc tạo nội dung số thường được GV thực hiện và theo nhu cầu giảng dạy. Là một GV tiếng Anh, việc sáng tạo và phát triển nội dung số là điều cần thiết vì hiện nay có rất nhiều SV đã chán học trực tuyến. Bằng cách tìm kiếm nhiều nguồn liên quan đến chủ đề sẽ giảng dạy dưới nhiều hình thức khác nhau, sau đó, thông tin được tải xuống và điều chỉnh theo nhu cầu và mục tiêu học tập cần đạt được. Việc tạo thông tin cũng có thể được thực hiện bằng cách tích hợp thông tin bạn muốn dạy với ứng dụng phù hợp và làm quen với điều kiện của SV để nội dung được tải thay đổi định dạng hoặc hình thức khác nhau. Tất cả nội dung IDL được phát

triển hoặc sắp xếp lại bởi các GV và SV đều có bản quyền cá nhân nên chỉ SV UNETI mới có thể truy cập được.

Một trong những yếu tố tạo nên tầm quan trọng của DK đối với SV là khả năng tiếp cận thông tin một cách dễ dàng, nhanh chóng, chính xác và với số lượng phù hợp không giới hạn [23]. Nhưng nhiều SV quên mất vấn đề bảo mật trong thế giới IDL. Thực trạng hiện nay nhóm thanh niên là nhóm dễ bị tổn thương nhất trước những tác động xấu của truyền thông IDL[24]. Do đó, bảo mật trong IDL là điều quan trọng. Nhận thấy tầm quan trọng của việc duy trì bảo mật trong IDL, nhà trường đang hợp tác với công ty google để tạo ra một môi trường học tập IDL khép kín với một số vòng kết nối nhất định để chỉ có thể truy cập bằng tài khoản IDL dựa trên trường IT của UNETI. Mọi thành viên của trường, GV, SV, cán bộ nhà trường đều có một tài khoản đặc biệt đó là danh tính IDL của trường. Tài khoản này có quyền truy cập vào tất cả dữ liệu của trường trên quy mô khép kín nên phải được sử dụng trong khi liên lạc hoặc truy cập nội bộ tất cả các hoạt động và tài liệu học tập của trường.

Vấn đề xảy ra nhiều trong quá trình học EFL trực tuyến là vấn đề kỹ thuật liên quan đến kết nối, thiết bị IDL sử dụng và ứng dụng. Thông thường, các vấn đề kỹ thuật mà GV và SV gặp phải có thể được giải quyết bằng sự giao tiếp tốt giữa GV và SV để hiểu được tình trạng của nhau. Về nội dung số hoặc tài liệu học tập bằng EFL, GV tiếng Anh UNETI cho biết không có vấn đề gì đáng kể. Bởi vì tất cả tài liệu học tập và tài liệu liên quan đã được cung cấp dưới dạng IDL cho SV để SV có thể truy cập mọi lúc, mọi nơi bằng nhiều thiết bị IDL khác nhau.

V. KẾT LUẬN

Các phạm vi năng lực và năng lực thành thạo nâng cao cho thấy xã hội tiên tiến cũng có thể được thực hiện trong hoạt động giảng dạy, đặc biệt là trong việc dạy và học EFL. Nhìn chung, UNETI có năng lực và đã hiện thực hóa các năng lực giáo dục nâng cao trong bảy lĩnh vực: 1) thiết bị và tiện ích tinh vi; 2) giáo dục về dữ liệu và thông tin; 3) giao tiếp và hợp tác; 4) sáng tạo nâng cao; 5) an ninh; 6) giải quyết vấn đề; và 7) năng lực liên quan đến nghề nghiệp. Trong việc học EFL, không phải tất cả các tiêu chí đánh giá năng lực thành thạo đều phù hợp. Điều này xảy ra do hạn chế của GV, đặc biệt là về PC và lập trình cũng như các trở ngại liên quan đến SV, các thiết bị PC được sử dụng cũng như lứa tuổi của SV, điều này hạn chế sự di chuyển không gian của SV trong thế giới hiện đại. Một số dấu hiệu về năng lực thành thạo trên PC không thể liên quan đến việc học đếm EFL để có thể thành thạo lập trình và có khả năng tính toán trong việc giải quyết các vấn đề liên quan đến các thiết bị tiên tiến. Trong mọi trường hợp, việc sử dụng các năng lực nâng cao trong việc học EFL ngày nay là đặc biệt quan trọng và có giá trị.

TÀI LIỆU THAM KHẢO

[1] Han Z (2022) Digital Language Learning and SLA. Bilingualism: *Language and Cognition*, 1–2. doi:10.1017/S1366728921000778

[2] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

[3] Koşar, G. (2023). An examination into pre-Service English-as-a-Foreign-Language teachers' Self-Perceived technological-Pedagogical-Content-knowledge. *E-Learning and Digital Media*, 204275302311561. https://doi.org/10.1177/20427530231156169

[4] Ma, Q., & Yan, J. (2022). How to empirically and theoretically incorporate digital technologies into language learning and teaching. *Bilingualism: Language and Cognition*, *25*(3), 392–393. https://doi.org/10.1017/S136672892100078X

[5] Hur, J. W., & Suh, S. (2012). Making Learning Active with Interactive Whiteboards, Podcasts, and Digital Storytelling in ELL Classrooms. *Computers in the Schools*, *29*(4), 320–338. https://doi.org/10.1080/07380569.2012.734275

[6] Koşar, G. (2023). An examination into pre-Service English-as-a-Foreign-Language teachers' Self-Perceived technological-Pedagogical-Content-knowledge. *E-Learning and Digital Media*, 204275302311561. https://doi.org/10.1177/20427530231156169

[7] Rodliyah, R. S. (2018). Vocational School EFL Teachers' Practices of Integrating ICT Into English Lessons: Teachers' Voices. Indonesian Journal of Applied Linguistics, 8(2), 418–428. https://doi.org/10.17509/ijal.v8i2.13309

[8] Kurnianingsih, I., Rosini, R., & Ismayati, N. (2017). Upaya Peningkatan Kemampuan Literasi Digital Bagi Tenaga Perpustakaan Sekolah dan Guru di Wilayah Mataram Pusat Melalui

[9] Lankshear, C., & Knobel, M. (2008). Introduction. In C. Lankshear, & M. Knobel. (Eds.). Digital literacies: Concepts, policies and practices (pp.1-16). New York, NY: Peter Lang. Retrieved from https://researchonline.jcu.edu.au/27788/1/27788_Lankshear_and_Knobel_2008.pdf

[10] Australian Council of Educational Research (2016). A global measure of digital and ICT literacy skills. Retrieved from http://unesdoc.unesco.org/images/0024/002455/245577E.pdf

[11] Muyasaroh, M., Ladamay, O. M. M. A., Mahfud, C., Mustakim, M., & Sejati, Y. G. (2020). The Utilization of Gadget in Maintaining Prophetical Values in Millennial Generation. International Journal of Psychosocial Rehabilitation, 24(4), 5602–5615. https://doi.org/10.37200/IJPR/V24I4/PR201655

[12] Supratman, L. P., & Wahyudin, A. (2017). Digital Media Literacy to Higher Students in Indonesia. International Journal of English Literature and Social Sciences, 2(5), 51–58. https://doi.org/10.24001/ijels.2.5.7

[13] Setyaningsih, R., Abdullah, A., Prihantoro, E., & Hustinawaty, H. (2019). Model Penguatan Literasi Digital Melalui Pemanfaatan E-Learning. Jurnal ASPIKOM, 3(6), 1200-1214. https://doi.org/10.24329/aspikom.v3i6.333.

[14] Yordming, R. (2017). Teachers' perspective towards digital teaching tools in Thai EFL classrooms. International Journal of Languages, *Literature and Linguistics, 3*(2), 45– 48. https://doi.org/10.18178/IJLLL.2017.3.2.108

[15] Durriyah, T. L., & Zuhdi, M. (2018). Digital Literacy with EFL Student Teachers: Exploring Indonesian Student Teachers' Initial Perception about Integrating Digital Technologies into a Teaching Unit. International Journal of Education and Literacy Studies, 6(3), p.53-60. https://doi.org/10.7575/aiac.ijels.v.6n.3

[16] Merç, A. A. (2015). Using technology in the classroom: A study with Turkish pre-service EFL teachers. TOJET: The Turkish Online Journal of Educational Technology, 14(2), 229–240. https://eric.ed.gov/?id=EJ1057362

[17] Oriogu, C. D. (2018). The use of mobile devices in learning foreign languages : *Survey of a private university*," Library Philosophy and Practice (e-journal), vol. 2. 2018

[18] Johnson, Barbara L. & Sadaf Ayesha. (2017). Teachers' Beliefs About Integrating Digital Literacy Into Classroom Practice: An Investigation Based on the Theory of Planned Behavior. *Journal of Digital Learning in Teacher Education*. doi: 10.1080/21532974.2017.1347534

[19] Grisham, D. L. & Wolsey, T. D. (2006). Recentering the middle school classroom as a vibrant learning community: *students, literacy, and technology intersect*," J. Adolesc. Adult Lit., vol. 49, no. 8, pp. 648–660,.

[20] Maolida, E. H., & Anjaniputra, A. G. (2018). The Praxis of Digital Literacy in the EFL classroom: Digital-Immigrant VS Digital-Native Teacher. Indonesian Journal of Applied Linguistics, 8(1), 28–37. https://doi.org/10.17509/ijal.v8i1.11459

[21] Harmer, J. (1998). How to teach English: an introduction to the practice of English language teaching. Harlow: Longman.

[22] Valkenburg, P. M., Koutamanis, M., & Vossen, H. G. M. (2017). The Concurrent and Longitudinal Relationships between Adolescents' Use of Social Network Sites and Their Social Self-Esteem. Computers in Human Behavior, 76, 35–41. https://doi.org/10.1016/j.chb.2017.07.008

[23] Nurjanah, E., Rusmana, A., & Yanto, A. (2017). Hubungan Literasi Digital dengan Kualitas Penggunaan E-Resources. Lentera Pustaka: Jurnal Kajian Ilmu

Perpustakaan, Informasi dan Kearsipan, 3(2), 117-140. https://doi.org/10.14710/lenpust.v3i2.16737

[24] Nur, M. (2019). Best Practice Implementasi Kurikulum pada Sekolah Alam Insan Mulia (Saim) Surabaya. Jurnal SMART (Studi Masyarakat, Religi, dan Tradisi), 5(1), 15–27. https://doi.org/10.18784/smart.v5i1.744

Thông tin tác giả chính:

Nguyễn Thị Luyến: ThS., đang công tác tại Khoa Ngoại Ngữ, Đại học Kinh tế- Kỹ thuật Công nghiệp (UNETI), số 218- Lĩnh Nam- Hoàng Mai- Hà Nội.

Email: ntluyen.nn@uneti.edu.vn

Thông tin tác giả liên hệ

Nguyễn Thúy Ngọc: ThS., đang công tác tại Khoa Ngoại Ngữ, Đại học Kinh tế- Kỹ thuật Công nghiệp (UNETI), số 218- Lĩnh Nam- Hoàng Mai- Hà Nội.

Email:ntngoc.khoann@uneti.edu.vn

VAI TRÒ CỦA TRÍ TUỆ NHÂN TẠO TRONG SỰ HỘI NHẬP GIÁO DỤC CỦA VIỆT NAM VỚI THẾ GIỚI

Phan Thị Ngọc Lệ[1]

Tóm tắt: Nghiên cứu này được thực hiện nhằm làm rõ vai trò và tiềm năng ứng dụng của AI trong quá trình hội nhập giáo dục của Việt Nam với thế giới. Với phương phân tích tổng hợp tài liệu về những khảo sát thực tiễn một số mô hình ứng dụng AI, nghiên cứu chỉ ra rằng AI đóng vai trò quan trọng trong việc thúc đẩy quá trình hội nhập giáo dục Việt Nam với thế giới, thể hiện ở ba phương diện chủ yếu: (1) góp phần nâng cao chất lượng đào tạo theo các tiêu chuẩn quốc tế; (2) hỗ trợ hoạt động dạy và học trực tuyến, học đa ngôn ngữ cũng như tiếp cận nguồn tài liệu toàn cầu; (3) tạo điều kiện thuận lợi cho việc liên kết đào tạo và hợp tác giáo dục xuyên quốc gia. Ý nghĩa của nghiên cứu không chỉ nằm ở việc cung cấp một cái nhìn toàn diện về tiềm năng của AI trong giáo dục, mà còn góp phần định hướng chính sách, xây dựng lộ trình hội nhập giáo dục quốc tế bền vững và hiệu quả cho Việt Nam trong thời đại số.

***Từ khóa**:* Trí tuệ nhân tạo (AI), Hội nhập giáo dục, Chuyển đổi số, Hệ thống giáo dục Việt Nam, Phát triển chính sách.

1. DẪN NHẬP

Trong bối cảnh toàn cầu hóa và cuộc Cách mạng công nghiệp lần thứ tư đang diễn ra mạnh mẽ, trí tuệ nhân tạo (AI) đã và đang trở thành một lực lượng thay đổi nền tảng trong hầu hết các lĩnh vực của đời sống xã hội, đặc biệt là giáo dục. Tại Việt Nam, sự chuyển mình của giáo dục trong kỷ nguyên số không chỉ là yêu cầu tất yếu từ nội tại mà còn là điều kiện bắt buộc để hội nhập hiệu quả với nền giáo dục tiên tiến trên thế giới (Lê & Trần, 2024). Các chỉ đạo và định hướng chiến lược do Chính phủ ban hành đã tạo động lực mạnh mẽ cho quá trình đổi mới toàn diện và sâu rộng trong lĩnh vực giáo dục và đào tạo. Các chiến lược này khẳng định tính cấp thiết của việc cải cách giáo dục đại học nhằm nâng cao chất lượng đào tạo và hiệu quả hoạt động của hệ thống giáo dục đại học, qua đó góp phần vào sự phát triển bền vững của đất nước (Nguyễn, 2024). Tiến trình cải cách giáo dục đại học không chỉ dựa trên cơ sở lý luận vững chắc mà còn được củng cố bởi các kết quả nghiên cứu thực tiễn trong và ngoài nước.

Đổi mới giáo dục đại học đóng vai trò then chốt trong việc nâng cao chất lượng đào tạo và đáp ứng những yêu cầu ngày càng cao của quá trình phát triển kinh tế – xã hội. Bên cạnh đó, quá trình đổi mới này cũng mở ra nhiều cơ hội thuận lợi cho việc hội nhập quốc tế sâu rộng (Đặng, 2021; Nguyễn, 2020; Nguyễn và cộng sự, 2023; Lunag và cộng sự, 2024). Tuy nhiên, trong bối cảnh xã hội siêu thông minh 5.0 đang dần hình thành, nơi các công nghệ hiện đại như Trí tuệ nhân tạo (AI), Internet vạn vật (IoT), blockchain và nhiều công nghệ tiên tiến khác thâm

[1]TS, Trường Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội; *Email: lehang6778@gmail.com*

nhập sâu vào mọi mặt của đời sống, giáo dục đại học phải đối mặt với những yêu cầu mới mang tính cấp bách. Trước làn sóng thay đổi công nghệ mạnh mẽ và nhanh chóng này, các cơ sở giáo dục đại học cần có chiến lược thích ứng linh hoạt, đồng thời phát triển năng lực đào tạo nhằm trang bị cho sinh viên kiến thức và kỹ năng cần thiết để sẵn sàng ứng phó với những thách thức phức tạp của thời đại số.

Việc ứng dụng trí tuệ nhân tạo vào giáo dục không còn là một xu hướng xa lạ, mà đang dần hiện diện trong các chính sách, chương trình đào tạo, phương pháp giảng dạy, và mô hình quản trị giáo dục ở mọi cấp học (Russell & Norvig, 2016). Theo ý kiến của Nguyễn và cộng sự (2018), trí tuệ nhân tạo hiện đang phát triển với tốc độ nhanh chóng, có nhiều đóng góp quan trọng vào sản xuất, kinh doanh, dịch vụ và đời sống con người. Do vậy, trí tuệ nhân tạo mở ra nhiều cơ hội giúp Việt Nam rút ngắn khoảng cách về trình độ giáo dục so với các quốc gia phát triển, thông qua việc cung cấp các công cụ hỗ trợ học tập cá nhân hóa, nền tảng học trực tuyến thông minh, hệ thống đánh giá và kiểm tra tự động, cũng như phân tích dữ liệu học tập để tối ưu hóa chiến lược giáo dục. Hơn nữa, AI còn giúp giảm thiểu sự phụ thuộc vào yếu tố địa lý và tài nguyên hữu hạn, từ đó góp phần mở rộng khả năng tiếp cận giáo dục chất lượng cao đến mọi đối tượng học sinh, sinh viên, kể cả ở vùng sâu, vùng xa.

Tuy nhiên, hội nhập giáo dục không chỉ là sự tiếp nhận công nghệ mà còn bao gồm việc thay đổi tư duy, xây dựng năng lực đội ngũ giáo viên, và điều chỉnh khung chính sách để phù hợp với tiêu chuẩn quốc tế. Trong quá trình đó, AI có thể đóng vai trò như một "cầu nối" giúp Việt Nam tiếp cận nhanh hơn với tri thức toàn cầu, đồng thời đảm bảo tính bản địa hóa và phù hợp với đặc điểm văn hóa, xã hội trong nước. Với bối cảnh đó, nghiên cứu này được thực hiện nhằm phân tích vai trò của trí tuệ nhân tạo trong tiến trình hội nhập giáo dục quốc tế của Việt Nam. Trước hết, bài viết sẽ điểm qua thực trạng ứng dụng AI trong giáo dục Việt Nam hiện nay, từ đó làm rõ những tiềm năng mà công nghệ này có thể mang lại. Cuối cùng, một số kiến nghị sẽ được đề xuất nhằm phát huy tối đa vai trò của AI trong việc đưa giáo dục Việt Nam tiệm cận với các chuẩn mực toàn cầu một cách bền vững và hiệu quả.

2. CƠ SỞ LÝ THUYẾT

2.1. Tổng quan về trí tuệ nhân tạo trong giáo dục

Trí tuệ nhân tạo (AI) là một lĩnh vực trong khoa học máy tính tập trung vào việc xây dựng các hệ thống có khả năng thực hiện các tác vụ đòi hỏi trí tuệ con người như học hỏi, suy luận, xử lý ngôn ngữ tự nhiên và ra quyết định. Trong lĩnh vực giáo dục, AI đề cập đến việc sử dụng các công nghệ và thuật toán để hỗ trợ hoặc thay thế một phần vai trò của con người trong quá trình dạy và học. Theo ý kiến của Chen và cộng sự (2020), việc nhắc đến trí tuệ nhân tạo thường khiến người ta liên tưởng đến siêu máy tính — một loại máy tính có khả năng xử lý vượt trội, bao gồm cả hành vi thích ứng như tích hợp cảm biến và các khả năng khác giúp nó có thể nhận thức và hoạt động giống con người, đồng thời cải thiện sự tương tác của siêu máy tính với con người. Trong lĩnh vực giáo dục, việc ứng dụng trí tuệ nhân tạo ngày càng được mở rộng. Ví dụ, khi được tích hợp vào robot, trí tuệ nhân tạo (AI) hoặc máy tính và các thiết

bị hỗ trợ có thể tạo ra những robot giúp nâng cao trải nghiệm học tập của học sinh, bắt đầu từ cấp học cơ bản nhất là giáo dục mầm non (Timms, 2016). Tương tự, giáo dục trực tuyến và qua web, như được đề cập trong nhiều nghiên cứu của Devedžic (2004), Peredo và cộng sự (2011), Chassignol và cộng sự (2018), đã chuyển đổi từ việc đơn thuần cung cấp tài liệu trên mạng để học sinh tải về, học và làm bài nhằm mục tiêu vượt qua môn học, sang việc tích hợp các hệ thống thông minh và thích ứng trên nền tảng web. Những hệ thống này có khả năng học hỏi hành vi của giảng viên và người học để điều chỉnh phù hợp, nhằm nâng cao trải nghiệm giáo dục. Nói cách khác, AI trong giáo dục không chỉ là công cụ kỹ thuật, mà còn là nền tảng giúp nâng cao hiệu quả học tập, tối ưu hóa quản lý giáo dục, và hỗ trợ cá nhân hóa trải nghiệm học cho từng học sinh, sinh viên.

Theo nghiên cứu của một số học giả như Chassignol và cộng sự (2018), Chen và cộng sự (2020), AI trong giáo dục được ứng dụng dưới nhiều hình thức khác nhau, có thể phân loại thành các nhóm chính sau:

- **Hệ thống học tập cá nhân hóa** (Intelligent Tutoring Systems – ITS): Sử dụng AI để điều chỉnh nội dung, tốc độ và phương pháp giảng dạy theo nhu cầu của từng người học.
- **Phân tích học tập** (Learning Analytics): Áp dụng AI để phân tích dữ liệu học tập và dự đoán hiệu suất, từ đó can thiệp sớm để hỗ trợ học sinh.
- **Chatbot và trợ lý ảo**: Hỗ trợ giải đáp thắc mắc, hướng dẫn học tập, và cung cấp thông tin hành chính.
- **Tự động hóa đánh giá**: Chấm điểm tự động bài tập trắc nghiệm và tự luận nhờ AI nhận diện ngôn ngữ và mẫu câu.
- **Dịch thuật và chuyển ngữ tự động**: Giúp học sinh tiếp cận tài liệu học tập từ nhiều ngôn ngữ khác nhau.

2.2. Tác động của AI đến giáo dục toàn cầu và hội nhập quốc tế

Theo ý kiến của Nguyễn (2022), trong xu thế toàn cầu hóa, bất kì nền giáo dục nào cũng không thể đứng riêng lẻ mà phải hội nhập, không thể khép kín mà phải tương tác với môi trường quốc tế. Hội nhập giáo dục quốc tế là quá trình mà các quốc gia hoặc tổ chức giáo dục mở rộng hoạt động và hợp tác với các quốc gia khác, nhằm nâng cao chất lượng giáo dục, trao đổi kiến thức và kinh nghiệm, cũng như thúc đẩy sự hiểu biết lẫn nhau giữa các nền văn hóa. Đây là một quá trình phức tạp, bao gồm cả việc thương mại hóa và phi thương mại hóa giáo dục. Trong bối cảnh Cách mạng Công nghiệp 4.0, trí tuệ nhân tạo không chỉ là công nghệ hỗ trợ, mà còn là một yếu tố cấu thành giúp các nền giáo dục tiến tới hội nhập một cách nhanh chóng và hiệu quả hơn. Trí tuệ nhân tạo (AI) đã dần dần biến đổi giáo dục toàn cầu thông qua các nền tảng học tập thích ứng, trợ lý ảo và hệ thống chấm điểm tự động. Từ những hệ thống chuyên gia ban đầu đến các mô hình học máy hiện đại, AI hiện có khả năng phân tích hiệu quả học tập của học sinh, cá nhân hóa chương trình giảng dạy và dự đoán kết quả học tập (Holmes và cộng sự, 2019; Luckin và cộng sự, 2016).

Đầu tiên, AI có tác động đến phương pháp giảng dạy và học tập. AI đã thay đổi giáo dục bằng cách hỗ trợ học tập cá nhân hóa, khi các hệ thống dựa trên AI đánh giá dữ liệu của học sinh để tùy chỉnh phương pháp giảng dạy dựa trên phong cách học, tốc độ và trình độ kỹ năng riêng biệt của từng cá nhân (Selwyn, 2019). Ví dụ, nền tảng *Squirrel AI* tại Trung Quốc sử dụng dữ liệu người học để đưa ra lộ trình học tập cá nhân hóa. Ngoài ra, AI hỗ trợ học sinh học tập chủ động bằng cách khám phá kiến thức qua các trò chơi, mô phỏng, và tình huống thực tế ảo (VR/AR), giúp nâng cao khả năng phản biện và tư duy sáng tạo. Bên cạnh đó, Chatbot, trợ lý ảo và các hệ thống học tập đối thoại (conversational AI) tạo điều kiện cho việc học tương tác không giới hạn về thời gian hay địa điểm. Tuy nhiên, Amiri và cộng sự (2025) cho rằng cần thiết phải xây dựng các chính sách giáo dục phù hợp để đảm bảo việc sử dụng trí tuệ nhân tạo (AI) trong môi trường giáo dục diễn ra một cách có trách nhiệm và công bằng. Các chiến lược học tập thích ứng phải đảm bảo rằng các hệ thống được hỗ trợ bởi AI điều chỉnh nội dung và tốc độ học tập theo tiến độ của học sinh, tạo điều kiện cho việc học linh hoạt và hiệu quả hơn. Các chính sách tập trung vào giáo dục dựa trên năng lực nên khuyến khích các lộ trình học tập cho phép học sinh tiến bộ sau khi đã nắm vững các khái niệm thay vì phải theo một chương trình học cố định.

AI không chỉ thay đổi cách dạy và học, mà còn tối ưu hóa hệ thống quản lý giáo dục. Theo Noor và cộng sự (2025), các nhà hoạch định chính sách giáo dục có thể tận dụng khả năng của trí tuệ nhân tạo (AI) để đưa ra các quyết định dựa trên dữ liệu, cải thiện phân bổ nguồn lực và dự đoán các xu hướng sắp tới trong lĩnh vực giáo dục. Thông qua việc phân tích các bộ dữ liệu lớn, AI có thể phát hiện những thiếu sót trong hệ thống giáo dục, nhận biết sự khác biệt về thành tích học tập của học sinh giữa các nhóm dân cư khác nhau, đồng thời đề xuất các biện pháp can thiệp cụ thể nhằm cải thiện kết quả học tập. Các phân tích do AI hỗ trợ có thể giúp dự báo tỷ lệ học sinh bỏ học, đánh giá hiệu quả của các chiến lược giảng dạy và nâng cao phát triển chương trình học để đáp ứng nhu cầu thay đổi của ngành công nghiệp (Bit và cộng sự, 2024). Thêm vào đó, AI có thể cải thiện việc phân bổ nguồn lực giáo dục, đảm bảo các trường học có nhu cầu cao nhất nhận được đủ kinh phí, công nghệ và dịch vụ hỗ trợ. Phương pháp tiếp cận tập trung vào dữ liệu này giúp các nhà hoạch định chính sách phát triển các hệ thống giáo dục linh hoạt và hiệu quả hơn, giải quyết các thách thức từ trước (Amiri & cộng sự, 2025).

Ngoài ra, AI còn tác động đến tiếp cận giáo dục và bình đẳng học tập. AI đang tạo ra sự chuyển biến mạnh mẽ trong giáo dục bằng cách gỡ bỏ những rào cản đã từng hạn chế khả năng tiếp cận tài liệu học tập của học sinh khuyết tật. Graham và cộng sự (2019) cho rằng các công cụ hỗ trợ dựa trên AI như chuyển giọng nói thành văn bản, chuyển văn bản thành giọng nói và dịch thuật tức thì đã cải thiện đáng kể khả năng tiếp cận, cho phép những học sinh gặp khó khăn về thị giác, thính giác hoặc ngôn ngữ có thể tương tác với tài liệu giáo dục theo cách phù hợp với nhu cầu của họ. Những công nghệ này góp phần xây dựng một môi trường học tập công bằng hơn, nơi mọi học sinh đều có thể tham gia đầy đủ vào các hoạt động giáo dục. Hơn nữa, các nền tảng học trực tuyến tích hợp AI giúp học sinh ở vùng sâu vùng xa vẫn có thể tiếp cận chương trình đào tạo chất lượng cao. AI còn có thể thu hẹp khoảng cách số bằng cách điều

chỉnh tốc độ học, phương pháp tiếp cận phù hợp với khả năng kỹ thuật số khác nhau giữa các nhóm học sinh (Gerold & Antrekowitsch, 2022).

2.3. Khoảng trống nghiên cứu

Từ những nghiên cứu trên, có thể thấy dù có nhiều nghiên cứu trên thế giới đề cập đến ứng dụng trí tuệ nhân tạo (AI) trong giáo dục, tuy nhiên vẫn tồn tại nhiều khoảng trống cần được nhận diện và lấp đầy, đặc biệt trong mối liên hệ giữa AI và hội nhập giáo dục quốc tế. Đầu tiên, chưa có nhiều nghiên cứu về vai trò của AI với sự hội nhập giáo dục tại Việt Nam. Nhiều nghiên cứu còn mang tính lý thuyết, chưa có dữ liệu định lượng hoặc định tính đầy đủ để đánh giá tác động cụ thể của AI đến khả năng hội nhập giáo dục. Ngoài ra, các nghiên cứu hiện tại ít chú trọng đến trải nghiệm thực tế của người học và người dạy khi triển khai AI – một yếu tố quan trọng trong việc hiểu được mức độ hiệu quả và sẵn sàng hội nhập.

Do đó, nghiên cứu này được thực hiện nhằm làm rõ vai trò và tiềm năng ứng dụng của AI trong quá trình hội nhập giáo dục của Việt Nam với thế giới, từ đó đề xuất các định hướng chính sách và chiến lược triển khai phù hợp trong giai đoạn tới. Ý nghĩa của nghiên cứu không chỉ nằm ở việc cung cấp một cái nhìn toàn diện về thực trạng và tiềm năng của AI trong giáo dục, mà còn góp phần định hướng chính sách, xây dựng lộ trình hội nhập giáo dục quốc tế bền vững và hiệu quả cho Việt Nam trong thời đại số.

3. PHƯƠNG PHÁP NGHIÊN CỨU

Nghiên cứu này chủ yếu sử dụng ***phương pháp phân tích – tổng hợp tài liệu*** nhằm đánh giá vai trò của trí tuệ nhân tạo (AI) trong quá trình hội nhập giáo dục của Việt Nam với thế giới. Phương pháp được triển khai theo các bước cụ thể như sau:

3.1. Thu thập và lựa chọn tài liệu

Nguồn tài liệu được thu thập bao gồm các công trình nghiên cứu khoa học, báo cáo thực tiễn, tài liệu hội thảo quốc tế, cùng các bài viết trên các tạp chí chuyên ngành có liên quan đến việc ứng dụng AI trong giáo dục tại Việt Nam và một số quốc gia trên thế giới. Việc lựa chọn tài liệu dựa trên các tiêu chí: (i) mức độ liên quan đến chủ đề nghiên cứu; (ii) tính học thuật và độ tin cậy của nguồn tài liệu; (iii) khả năng phản ánh rõ nét thực tiễn triển khai AI trong lĩnh vực giáo dục.

3.2. Phân tích các mô hình ứng dụng AI

Nghiên cứu tiến hành phân tích một số mô hình tiêu biểu về ứng dụng AI trong giáo dục, được triển khai trên thế giới và tại Việt Nam. Các mô hình được lựa chọn nhằm làm rõ hiệu quả, cách thức tổ chức triển khai, công nghệ sử dụng và tác động đến quá trình quốc tế hóa giáo dục.

3.3. Tổng hợp và hệ thống hóa dữ liệu

Sau khi phân tích từng trường hợp cụ thể, nghiên cứu tiến hành tổng hợp và hệ thống hóa các nội dung liên quan đến vai trò của AI trong ba khía cạnh chính của hội nhập giáo dục tại Việt Nam:

(i) nâng cao chất lượng đào tạo theo chuẩn quốc tế;

(ii) hỗ trợ dạy học trực tuyến, học đa ngôn ngữ và tiếp cận tri thức toàn cầu;

(iii) thúc đẩy liên kết đào tạo và hợp tác giáo dục xuyên biên giới.

3.4. Đánh giá và đưa ra nhận định

Trên cơ sở tổng hợp kết quả phân tích, nghiên cứu tiến hành đánh giá vai trò tổng thể của AI đối với tiến trình hội nhập giáo dục tại Việt Nam. Các đánh giá được xây dựng dựa trên bằng chứng thực tiễn, so sánh quốc tế và định hướng chính sách hiện hành. Từ đó, nghiên cứu đề xuất một số khuyến nghị nhằm khai thác hiệu quả tiềm năng của AI trong việc nâng cao năng lực cạnh tranh và hội nhập giáo dục của Việt Nam trong bối cảnh toàn cầu hóa.

4. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

4.1. Vai trò của AI trong việc hỗ trợ cải thiện chất lượng đào tạo theo chuẩn quốc tế

Đầu tiên, công nghệ AI có tác động sâu rộng đến trải nghiệm học tập, nâng cao khả năng cá nhân hóa, mức độ tương tác và hiệu quả tổng thể trong giáo dục. Có thể thấy, AI cho phép xây dựng lộ trình học tập phù hợp với nhu cầu và sở thích riêng của từng học sinh. Theo nghiên cứu của Nguyễn (2024), các hệ thống học tập thích ứng sử dụng AI đã giúp cải thiện đáng kể mức độ tương tác và kết quả học tập của sinh viên tại các trường đại học ở Việt Nam. Từ đó giúp các sinh viên có thể dần tiếp cận với các chuẩn quốc tế trong giáo dục. Tiếp theo, **c**ác bài kiểm tra dựa trên AI có khả năng điều chỉnh linh hoạt theo trình độ hiện tại của người học. Chúng cung cấp phản hồi theo thời gian thực và điều chỉnh độ khó của câu hỏi dựa trên hiệu suất của sinh viên, giúp đánh giá kiến thức và kỹ năng một cách chính xác hơn. Nguyễn và Hoàng (2024) nhấn mạnh rằng hệ thống kiểm tra này cung cấp phản hồi cá nhân hóa tức thì, hỗ trợ sinh viên hiểu rõ hơn các chủ đề phức tạp. Thứ ba, **c**ác hệ thống gia sư được hỗ trợ bởi AI cung cấp sự trợ giúp và hướng dẫn theo yêu cầu bên ngoài lớp học truyền thống. Những hệ thống này mang đến hỗ trợ cá nhân hóa, giúp sinh viên vượt qua các khó khăn cụ thể và củng cố kiến thức thông qua luyện tập có mục tiêu và giải thích chi tiết. Nguyễn và Hoàng (2024) cũng đề cập đến việc tích hợp hệ thống học tập cá nhân hóa (ITS) trong các cơ sở giáo dục đại học tại Việt Nam, nhấn mạnh vai trò của chúng trong việc cung cấp phản hồi theo thời gian thực và hỗ trợ học tập chuyên sâu cho những nội dung khó.

Bên cạnh phát triển các trải nghiệm học tập cho người học, AI còn có vai trò quan trọng trong việc tăng cường hiệu quả hành chính và nghiên cứu trong giáo dục. Thứ nhất, các ứng dụng AI trong quy trình hành chính có thể đơn giản hóa các hoạt động như tuyển sinh, xếp thời khóa biểu và chấm điểm. Đỗ và cộng sự (2025) nêu bật việc các trường đại học tại Việt Nam đang ứng dụng AI để giảm bớt gánh nặng hành chính, giúp cán bộ tập trung nhiều hơn vào chất

lượng giáo dục và hỗ trợ sinh viên. Bên cạnh đó, phân tích dự đoán dựa trên AI giúp các trường đại học dự báo xu hướng tuyển sinh, tối ưu hóa việc phân bổ nguồn lực và nâng cao tỷ lệ duy trì sinh viên. Nguyễn và Hoàng (2024) nhấn mạnh vai trò của phân tích dự đoán trong công tác hoạch định chiến lược và ra quyết định trong giáo dục đại học tại Việt Nam. Hơn nữa, các trường đại học tại Việt Nam ngày càng trở thành trung tâm nghiên cứu và phát triển trí tuệ nhân tạo. Nhiều dự án nghiên cứu AI như của các nhóm tác giả Trần và cộng sự (2022), Nguyễn và cộng sự (2023) đang đóng góp vào những tiến bộ trong các lĩnh vực như xử lý ngôn ngữ tự nhiên, thị giác máy tính và robot tại nhiều trường đại học tại Việt Nam. Từ đó, AI giúp giáo viên xây dựng bài giảng phù hợp với chuẩn đầu ra quốc tế, dễ dàng tiếp cận với khung năng lực chung như CEFR (khung ngôn ngữ châu Âu), ABET, CDIO... Ngoài ra, AI có thể hỗ trợ đánh giá kỹ năng viết, phản biện, kỹ năng mềm. Đây đều là các tiêu chí quan trọng trong các chương trình học quốc tế.

4.2. Vai trò của AI trong dạy học trực tuyến, học đa ngôn ngữ và tiếp cận tài liệu toàn cầu

Đầu tiên, AI có vai trò quan trọng trong dạy học trực tuyến và kết hợp tại các trường đại học tại Việt Nam. AI được kỳ vọng sẽ làm thay đổi chương trình giảng dạy, quy trình dạy và học, cũng như các phương pháp đánh giá trong dạy học trực tuyến. Hiện nay, một số trường đại học ở Việt Nam đã bắt đầu tích hợp AI vào các hoạt động giáo dục trực tuyến và kết hợp. Trường đại học tổ chức các lớp học trực tuyến theo đúng quy định hiện hành về việc ứng dụng công nghệ thông tin trong công tác quản lý và tổ chức đào tạo trực tuyến, đồng thời áp dụng các giải pháp nhằm đảm bảo chất lượng của các lớp học này. Ví dụ, trong nghiên cứu của Nguyễn (2024) về việc ứng dụng trí tuệ nhân tạo trong học tập kết hợp tại một trường đại học tại Việt Nam, đối với các hình thức đào tạo thường xuyên và liên tục (hệ vừa học vừa làm, liên thông, văn bằng hai), tối đa 30% tổng khối lượng chương trình đào tạo được phép thực hiện thông qua hình thức học trực tuyến. Bên cạnh đó, các trường đại học cũng đã sử dụng công nghệ số để quản lý các hoạt động giảng dạy, đánh giá và quy trình tuyển sinh (Nguyễn, 2024).

Ngoài ra, AI còn tăng cường khả năng học tập không biên giới bằng cách hỗ trợ dạy học trên các nền tảng trực tuyến đa ngôn ngữ, mở rộng khả năng tham gia vào các khóa học quốc tế từ các trường đại học hàng đầu thế giới. Hệ thống dịch bằng AI như Google Translate, DeepL hỗ trợ sinh viên tiếp cận tài liệu học thuật bằng nhiều ngôn ngữ, giảm rào cản ngôn ngữ trong học tập và nghiên cứu. Hơn nữa các trợ lý học tập ảo cá nhân còn có thể giúp sinh viên Việt Nam học chương trình quốc tế (ví dụ như MOOC, chương trình du học trực tuyến) theo năng lực riêng, với sự hỗ trợ sát sao từ hệ thống thông minh. Do đó, theo ý kiến của Đinh và Trần (2021), một chiến lược rõ ràng cho việc ứng dụng và phát triển AI trong giáo dục nói chung là yếu tố then chốt giúp định hướng đúng đắn cho nhà trường. Việc xây dựng một lộ trình cụ thể cho việc ứng dụng và phát triển AI trong học tập kết hợp và trực tuyến là hết sức cần thiết. Các trường đại học cần tận dụng cả nguồn lực nội bộ và bên ngoài, đồng thời xác định rõ các cơ hội và thách thức khi xây dựng chiến lược ứng dụng AI và triển khai mô hình học tập kết hợp, trực tuyến, học tập đa ngôn ngữ và tiếp cận tài liệu toàn cầu.

4.3. Vai trò của AI trong liên kết đào tạo và hợp tác giáo dục xuyên quốc gia

Đầu tiên, có thể thấy AI phát triển các chương trình đào tạo chung với nước ngoài. Với sự hỗ trợ của AI, việc thiết kế, quản lý và giảng dạy chương trình liên kết quốc tế có thể dễ dàng hơn, giảm chi phí và rào cản kỹ thuật. Ngoài ra, AI có thể giúp tăng cường giao tiếp xuyên văn hóa. AI hỗ trợ dịch thời gian thực, giúp sinh viên và giảng viên từ nhiều quốc gia làm việc, học tập và trao đổi hiệu quả. Nhiều nghiên cứu đã xác nhận hiệu quả của AI trong giáo dục tại Việt Nam, cho thấy sự cải thiện rõ rệt về thành tích học tập của học sinh và hiệu quả giảng dạy của giáo viên (Đào và cộng sự, 2023). Ví dụ, các hệ thống AI cung cấp phản hồi và đánh giá theo thời gian thực đã được chứng minh là giúp giảm đáng kể thời gian chấm điểm của giáo viên, từ đó tạo điều kiện để họ có thêm thời gian dành cho việc giảng dạy mang tính tương tác và cá nhân hóa. Tương tự, các công cụ phân tích dựa trên AI hỗ trợ dự đoán kết quả học tập của học sinh, qua đó giúp nhà trường can thiệp sớm đối với những học sinh cần hỗ trợ thêm (Nguyễn & Phan, 2024).

Ngoài ra, AI có vai trò trong tự động hóa công tác tuyển sinh và đánh giá đầu vào. Hệ thống AI giúp xử lý hồ sơ, đánh giá năng lực, phân tích động lực học tập – tạo điều kiện thuận lợi hơn cho sinh viên Việt Nam du học và sinh viên quốc tế đến học tại Việt Nam. Việc tích hợp AI vào các chương trình liên kết đào tạo giúp Việt Nam bắt kịp xu hướng giáo dục toàn cầu, nâng cao chất lượng và uy tín của các trường đại học trong mắt các đối tác quốc tế. Đây cũng là bước đệm quan trọng để Việt Nam mở rộng mạng lưới hợp tác giáo dục, thu hút sinh viên quốc tế và phát triển các chương trình đồng cấp bằng. Tại Việt Nam, việc ứng dụng AI trong liên kết đào tạo vẫn còn ở giai đoạn sơ khai nhưng đang phát triển nhanh chóng thông qua một số sáng kiến và dự án thí điểm (Nguyễn & Nguyễn, 2023). Vai trò chủ động của chính phủ trong việc số hóa ngành giáo dục và thúc đẩy kiến thức về AI thể hiện cam kết mạnh mẽ trong việc tích hợp các công nghệ tiên tiến vào quá trình học tập. Các trường học tại các đô thị lớn, đặc biệt là ở Hà Nội và thành phố Hồ Chí Minh, đã bắt đầu thử nghiệm các công cụ AI trong việc học ngôn ngữ và giải quyết bài tập. Những bước phát triển này nằm trong xu hướng chung nhằm xây dựng môi trường giáo dục hiện đại, có khả năng đáp ứng nhu cầu của một xã hội đang chuyển mình theo hướng số hóa (Nguyễn & Phan, 2024).

5. KẾT LUẬN

Nghiên cứu này được thực hiện nhằm làm rõ vai trò và tiềm năng ứng dụng của AI trong quá trình hội nhập giáo dục của Việt Nam với thế giới, từ đó đề xuất các định hướng chính sách và chiến lược triển khai phù hợp trong giai đoạn tới. Với phương pháp phân tích tổng hợp tài liệu về những khảo sát thực tiễn một số mô hình ứng dụng AI trong giáo dục tại Việt Nam và quốc tế, nghiên cứu đánh giá vai trò của AI đối với sự hội nhập của giáo dục Việt Nam. Kết quả cho thấy AI có ảnh hưởng lớn tới quá trình hội nhập của giáo dục Việt Nam với thế giới trong 3 khía cạnh chính: 1) hỗ trợ cải thiện chất lượng đào tạo theo chuẩn quốc tế; 2) dạy học trực tuyến, học đa ngôn ngữ và tiếp cận tài liệu toàn cầu; 3) liên kết đào tạo và hợp tác giáo dục xuyên quốc gia.

Nghiên cứu này có ý nghĩa cả về mặt lý luận và thực tiễn. Qua nghiên cứu này, có thể thấy rằng trong bối cảnh toàn cầu hóa giáo dục, việc đẩy mạnh tích hợp AI một cách chiến lược

sẽ giúp giáo dục Việt Nam tiệm cận chuẩn quốc tế, tăng khả năng cạnh tranh, và mở rộng cơ hội hợp tác đào tạo xuyên biên giới. Trí tuệ nhân tạo (AI) đang trở thành một yếu tố cốt lõi trong công cuộc đổi mới và hội nhập giáo dục toàn cầu. Đối với Việt Nam, AI không chỉ mang lại những đổi thay tích cực trong phương pháp dạy và học mà còn mở ra cơ hội để hệ thống giáo dục quốc gia tiệm cận với các chuẩn mực quốc tế. Việc ứng dụng AI trong cá nhân hóa học tập, kiểm tra đánh giá, quản trị giáo dục và hợp tác đào tạo xuyên quốc gia đang ngày càng chứng minh tính hiệu quả và tiềm năng phát triển bền vững.

Trong bối cảnh toàn cầu hóa mạnh mẽ và chuyển đổi số diễn ra sâu rộng, việc tích hợp AI vào giáo dục là bước đi tất yếu để nâng cao chất lượng đào tạo, thu hẹp khoảng cách giữa giáo dục trong nước và quốc tế, đồng thời tăng cường năng lực cạnh tranh của các cơ sở giáo dục Việt Nam trên trường học thuật thế giới. Tuy nhiên, để phát huy tối đa vai trò của AI, cần có chiến lược dài hạn, đầu tư đồng bộ vào hạ tầng công nghệ, phát triển nguồn nhân lực số, cũng như xây dựng chính sách hỗ trợ việc tích hợp AI một cách toàn diện và hiệu quả. Sự hội nhập giáo dục của Việt Nam với thế giới trong thời đại AI không chỉ là sự kết nối về chương trình, mô hình đào tạo, mà còn là sự hòa nhập về tư duy, cách tiếp cận và năng lực công nghệ. Trí tuệ nhân tạo, nếu được ứng dụng một cách hợp lý và có tầm nhìn, sẽ trở thành đòn bẩy quan trọng giúp Việt Nam khẳng định vị thế trong nền giáo dục toàn cầu đang thay đổi không ngừng.

TÀI LIỆU THAM KHẢO

1. Amiri, S. M. H., Islam, M. M., Hossen, M. S. (2025). The Role of Artificial Intelligence in Shaping Future Education Policies. *Education Journal, 14*(1), 32-38. https://doi.org/10.11648/j.edu.20251401.14
2. Bit, D., Biswas, S., & Nag, M. (2024). The Impact of Artificial Intelligence in Educational System. *International Journal of Scientific Research in Science and Technology, 11*(4), 419-427. https://doi.org/10.32628/IJSRST2411424
3. Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science, 136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233
4. Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, *8*, 75264-75278. https://doi.org/10.1109/ACCESS.2020.2988510
5. Đặng, U. V. (2021). *Đổi mới giáo dục đại học: Từ ý tưởng đến thực tiễn*. NXB Đại học Quốc gia Hà Nội. https://vietbooks.info/threads/doi-moi-giao-duc-dai-hoc-tu-y-tuong-den-thuc-tien-nxb-dai-hoc-quoc-gia-2021-dang-ung-van-371-tr.106802/
6. Đào, X. Q., Lê, N. B., Phan, X. D., & Ngô, B. B. (2023). Can ChatGPT pass the Vietnamese Naional High School Graduaion Examinaion? *Computation and Language,* 1-9. https://doi.org/10.48550/arXiv.2306.09170
7. Devedžic, V. (2004). Web intelligence and artificial intelligence in education. *Educational Technology & Society, 7*(4), 29–39. https://www.jstor.org/stable/jeductechsoci.7.4.29
8. Đinh, T. M. H., & Trần, V. H. (2021). Artificial Intelligence in Education: Opportunities and Challenges to the Future of Instructing and Studying at University. *The University of Danang - Journal of Science and Technology*, *19*(2), 38-42. https://www.neliti.com/publications/452779/artificial-intelligence-in-education-opportunities-and-challenges-to-the-future#cite
9. Đỗ, M. D., Nguyễn, V. V., & Đàm, C. T. (2025). Using Large Language Models for education managements in Vietnamese with low resources. *Computation and Language*, 1-15. https://doi.org/10.48550/arXiv.2501.15022

10. Gerold, E., & Antrekowitsch, H. A. (2022). Sustainable Approach for the Recovery of Manganese from Spent Lithium-Ion Batteries via Photocatalytic Oxidation. *International Journal of Materials Science and Applications, 11*(3), 66-75. https://pureadmin.unileoben.ac.at/ws/portalfiles/portal/10057155/10.11648.j.ijmsa.20221103.12.pdf

11. Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., & Jeste, D. V. (2019). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Current psychiatry reports, 21*(11), 116. https://doi.org/10.1007/s11920-019-1094-0

12. Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign: Boston, MA, USA. https://circls.org/primers/artificial-intelligence-in-education-promises-and-implications-for-teaching-and-learning

13. Lê, A. V., & Trần, M. N. (2024). Tác động của Trí tuệ nhân tạo (AI) đối với hệ thống giáo dục toàn cầu và giáo dục Việt Nam. *Tạp chí Khoa học Giáo dục Việt Nam, 20*(5), 1-11. http://vjes.vnies.edu.vn/sites/default/files/khdgvn_-_tap_20_-_so_5_-1-11.pdf

14. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson. https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/IntelligenceUnleashedSPANISH.pdf

15. Lunag Jr, M. N., Posadas, C. L., Lamadrid, R. L., Ducas, A. G., Teehankee, A., De Guzman, M. K. B., & Clemente, M. C. B. (2024). Building sustainable research and innovation ecosystem in Philippine higher education institutions. *Educational Research for Policy and Practice, 23*(1), 63-88. https://doi.org/10.1007/s10671-023-09355-2

16. Nguyễn, A. Q. (2024). Innovation in Vietnamese higher education for society 5.0: Opportunities, challenges, and strategies. *International Journal of Research in Education and Science (IJRES), 10*(4), 750-760. https://doi.org/10.46328/ijres.3491

17. Nguyễn, G. T. Q., & Hoàng, T. T. D. (2024). Artificial intelligence in Vietnamese higher education challenges and opportunities. *Tạp chí Công thương*,

https://tapchicongthuong.vn/artificial-intelligence-in-vietnamese-higher-education-challenges-and-opportunities-124436.htm

18. Nguyễn, N. B., Trần, N. H., & Võ, T. H. (2023). Một số vấn đề lý luận về quản lý và đổi mới giáo dục trong các trường đại học công lập. *Tạp chí Khoa học và Công nghệ Đại học Thái Nguyên, 229*(4), 72-80. https://doi.org/10.34238/tnu-jst.8539

19. Nguyễn, Q.N., Phan, C. T., Nguyễn, D. V., & Nguyễn, V. K. (2023). ViSoBERT: A Pre-Trained Language Model for Vietnamese Social Media Text Processing. *Computation and Language*, 1-17. https://doi.org/10.48550/arXiv.2310.11166

20. Nguyễn, T. H. T. (2024). Application of Artificial intelligence in blended learning - international experiences and potential applications at Thu Dau Mot University. The International Conference "*Higher education for human resources requirements and local socio-economic development – International experiences and suggestions for Vietnam*", pp.308-314. Ho Chi Minh National University Publishing House. https://www.researchgate.net/publication/382578801_NGUYEN_THI_HONG_THAM-Application_of_Artificial_intelligence_in_blended_learning_-_international_experiences_and_potential_applications_at_Thu_Dau_Mot_University

21. Nguyễn, T. L. (2020). *Quá trình đổi mới giáo dục đại học ở Việt Nam – thực trạng và giải pháp*. Luận văn Giáo dục học. Trường Đại học Khoa học Xã hội và Nhân văn, Đại học Quốc gia Hồ Chí Minh. http://ir.vnulib.edu.vn/handle/VNUHCM/5693

22. Nguyễn, T. L. P. (2022). Đổi mới giáo dục theo hướng mở và hội nhập quốc tê. *Tạp chí Khoa học Giáo dục Việt Nam, 18*(S1), 1-7. http://vjes.vnies.edu.vn/vi/doi-moi-giao-duc-theo-huong-mo-va-hoi-nhap-quoc-te

23. Nguyễn, T. Q. T., & Phan, T. T. (2024). The Role of AI in Improving Student Learning Outcomes: Evidence in Vietnam. *International Journal of Multidisciplinary Research and Analysis, 7*(6), 2806-2810. https://doi.org/10.47191/ijmra/v7-i06-48

24. Nguyễn, T. T., Hà, Q. T., Phan, X. H., & Nguyễn, T. T. (2018). *Trí tuệ nhân tạo trong thời đại số: Bối cảnh thế giới và liên hệ với Việt Nam.* Tạp chí Công Thương - Các kết quả nghiên cứu khoa học và Ứng dụng công nghệ. http://eprints.uet.vnu.edu.vn/eprints/id/eprint/3161

25. Nguyễn, T. T., & Nguyễn, N. T. (2023). The Impact of Digital Capabilities on the Work Performance of Provincial Civil Servants in Vietnam. *Journal of Law and Sustainable Development, 11*(4), e560. https://doi.org/10.55908/sdgs.v11i4.560

26. Noor, S., Khushhal, S., Amjad, A., & Aqil, A. (2025). Using AI to Inform Evidence-based Decision-Making in Educational Policy. Review of Applied Management and Social Sciences, 8(1), 183-192. https://ramss.spcrd.org/index.php/ramss/article/view/448

27. Peredo, R., Canales, A., Menchaca, A., & Peredo, I. (2011). Intelligent Webbased education system for adaptive learning. *Expert Systems with Applications, 38*(12), 14690–14702. https://doi.org/10.1016/j.eswa.2011.05.013

28. Russell, S., & Norvig, P. (2016). *Artificial Intelligence. A Modern Approach* (3rd Global Edition). Pearson. https://people.engr.tamu.edu/guni/csce625/slides/AI.pdf

29. Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press. https://research.monash.edu/en/publications/should-robots-replace-teachers-ai-and-the-future-of-education

30. Timms, M. J. (2016). Letting artificial intelligence in education out of the box: Educational cobots and smart classrooms. *International Journal of Artificial Intelligence Education, 26*(2), 701–712. https://doi.org/10.1007/s40593-016-0095-y

31. Trần, L. N., Lê, M. D., & Nguyễn, Q. D. (2022). BARTpho: Pre-trained Sequence-to-Sequence Models for Vietnamese. *Computation and Language*, 1-5. https://doi.org/10.48550/arXiv.2109.09701

Thông tin tác giả

1. Người thực hiện: **Phan Thị Ngọc Lệ**

2. Học vị: **Tiến sĩ Ngôn ngữ học**

3. Đơn vị công tác: **Khoa Ngôn ngữ & Văn hóa Anh, Trường Đại học Ngoại ngữ, ĐHQGHN.**

4. Email: lehang6778@gmail.com

TỪ DỮ LIỆU ĐẾN ẨN DỤ: QUY TRÌNH DỰA TRÊN KHỐI NGỮ LIỆU ĐỂ PHÂN TÍCH BÀI VIẾT HỌC THUẬT TIẾNG ANH BẰNG ANTCONC VÀ MIPVU

Trần Thế Phi[1]

Tóm tắt

Bài viết giới thiệu một quy trình bán tự động nhằm xác định và phân loại ẩn dụ ý niệm trong các văn bản học thuật tiếng Anh, cụ thể là các bài nghiên cứu do sinh viên đại học Việt Nam thực hiện trong khuôn khổ học phần "Viết nghiên cứu" (Research Writing) thuộc chương trình học của họ. Quy trình kết hợp phần mềm phân tích khối ngữ liệu AntConc để trích xuất và truy vết các mẫu từ vựng, cùng khung nhận diện ẩn dụ MIPVU để phân tích nghĩa ẩn dụ ở cấp độ từ vựng. Một nhóm danh từ trừu tượng được lựa chọn nhằm minh họa quy trình phân tích từng bước. Kết quả cho thấy phương pháp này vừa khả thi vừa linh hoạt trong việc xử lý các bài viết học thuật của sinh viên. Bài viết cũng thảo luận về những đóng góp phương pháp luận của cách tiếp cận này đối với nghiên cứu dựa trên khối ngữ liệu, giảng dạy viết học thuật, và phát triển các công cụ hỗ trợ học tập dựa trên ngôn ngữ học máy tính.

Từ khóa: Ngôn ngữ học tính toán; ẩn dụ ý niệm; MIPVU; AntConc; viết học thuật; phân tích ngữ liệu

FROM DATA TO METAPHOR: A CORPUS-BASED PROCEDURE FOR ANALYZING ENGLISH ACADEMIC WRITING WITH ANTCONC AND MIPVU

Abstract

This article presents a semi-automated procedure for identifying and classifying conceptual metaphors in English academic texts - specifically, research papers produced by Vietnamese university students as part of the Research Writing module in their academic program. The procedure integrates the corpus analysis software AntConc to extract and trace lexical patterns, and the MIPVU metaphor identification framework to analyze metaphorical meaning at the lexical level. A set of abstract nouns is selected to illustrate the step-by-step analytical process. The findings indicate that this method is both feasible and flexible in handling authentic student writing. The article also highlights the methodological contributions of this approach to corpus-based research, academic writing instruction, and the development of computational linguistics-based learning support tools.

Keywords: Computational linguistics; conceptual metaphor; MIPVU; AntConc; academic writing; corpus analysis

[1] TS., Trường Đại học Sài Gòn, *Email: tranthephi@sgu.edu.vn*

1. Giới thiệu

Trong bối cảnh giảng dạy tiếng Anh như một ngoại ngữ, việc diễn đạt các ý niệm trừu tượng trong viết học thuật là một thách thức lớn đối với sinh viên (Hyland, 2016; Swales & Feak, 2012). Trong quá trình diễn đạt các ý niệm này, người viết thường sử dụng ẩn dụ như một công cụ ngôn ngữ giúp hình dung hóa các khái niệm trừu tượng, từ đó làm cho lập luận trở nên rõ ràng và thuyết phục hơn. Ẩn dụ ý niệm không chỉ là hiện tượng ngôn ngữ mà là cơ sở nhận thức giúp hình thành nghĩa và hỗ trợ lập luận trong diễn ngôn (Lakoff & Johnson, 1980; Cameron & Deignan, 2006), nhưng việc nhận diện và phân loại ẩn dụ trong văn bản học thuật, đặc biệt khi xử lý khối lượng dữ liệu lớn, đòi hỏi thời gian đáng kể và dễ bị ảnh hưởng bởi sự diễn giải chủ quan, đặc biệt trong việc xác định nghĩa cơ bản, nghĩa ngữ cảnh, và miền ý niệm (Steen và cộng sự, 2010, tr. 75-79, 91-106).

Các phương pháp truyền thống thường dựa hoàn toàn vào phân tích thủ công, tuy đảm bảo độ chính xác nhưng tiêu tốn nhiều công sức và khó mở rộng quy mô (Pragglejaz Group, 2007). Ngược lại, các hệ thống tự động trong nhận diện ẩn dụ, dù đã có nhiều tiến bộ, vẫn gặp khó khăn khi xử lý các biểu thức ngôn ngữ mang tính tinh tế hoặc phụ thuộc mạnh vào ngữ cảnh, điều mà chỉ có con người mới có thể nắm bắt một cách linh hoạt hơn (Gibbs, 2017). Điều này đặt ra nhu cầu kết hợp hài hòa giữa các công cụ xử lý khối ngữ liệu và quy trình nhận diện thủ công có hệ thống, nhằm vừa nâng cao hiệu quả xử lý vừa đảm bảo tính chính xác và khả năng giải thích trong bối cảnh ngữ nghĩa cụ thể (Brezina, 2018; McEnery, Hardie, & Brezina, 2019 ; Gries, 2021).

Bài viết này đề xuất một quy trình bán tự động kết hợp phần mềm AntConc (Anthony, 2022) và phương pháp nhận diện ẩn dụ MIPVU (Steen và cộng sự, 2010), giúp tối ưu hóa việc phát hiện và phân loại ẩn dụ ý niệm trong văn bản học thuật tiếng Anh của sinh viên Việt Nam. Quy trình này vừa khai thác sức mạnh của phân tích định lượng từ phần mềm, vừa giữ được chiều sâu phân tích định tính của phương pháp thủ công.

Trên cơ sở đó, nghiên cứu này hướng đến hai mục tiêu chính:

- Xây dựng và triển khai quy trình tích hợp phần mềm AntConc với phương pháp MIPVU nhằm nâng cao hiệu quả nhận diện và phân loại ẩn dụ ý niệm trong văn bản học thuật tiếng Anh của sinh viên Việt Nam.
- Phân tích và làm rõ những đóng góp về phương pháp luận mà sự kết hợp giữa công cụ xử lý ngữ liệu và quy trình nhận diện ẩn dụ thủ công mang lại cho việc phân tích và giảng dạy viết học thuật tiếng Anh trong bối cảnh sinh viên Việt Nam học tiếng Anh như ngoại ngữ.

2. Cơ sở lý thuyết và khái niệm liên quan

2.1. Ẩn dụ ý niệm và lý thuyết nền tảng

Theo Lakoff và Johnson (1980), ẩn dụ ý niệm (conceptual metaphor) là sự ánh xạ có hệ thống giữa một miền nguồn (source domain) cụ thể và một miền đích (target domain) trừu

tượng. Cơ chế ánh xạ này cho phép người sử dụng ngôn ngữ hiểu, tổ chức và diễn đạt những khái niệm trừu tượng thông qua các trải nghiệm cụ thể, quen thuộc. Ví dụ, ẩn dụ THỜI GIAN LÀ TIỀN BẠC (TIME IS MONEY) cho phép diễn đạt về thời gian bằng những biểu thức thường dùng cho tiền bạc như "tiết kiệm thời gian" hoặc "lãng phí thời gian" (Lakoff & Johnson, 1980).

Kövecses (2020) mở rộng lý thuyết này thành "Lý thuyết Ẩn dụ ý niệm mở rộng" (Extended Conceptual Metaphor Theory - viết tắt ECMT), nhấn mạnh rằng sự hình thành và sử dụng ẩn dụ không chỉ dựa vào sự tương đồng ý niệm mà còn chịu ảnh hưởng mạnh mẽ của bối cảnh giao tiếp (communicative context), trải nghiệm văn hóa - xã hội (cultural and social experience) và các cấu trúc nhận thức phân tầng (hierarchical cognitive structures). ECMT đề xuất một mô hình phân cấp gồm bốn lớp: sơ đồ hình ảnh (image schema) → miền (domain) → khung (frame) → không gian ý niệm (mental space). Trong đó, sơ đồ hình ảnh đóng vai trò nền tảng cảm nhận - vận động hình thành từ kinh nghiệm thân thể; miền là tập hợp tri thức liên quan; khung là cấu trúc tri thức chuyên biệt cho một tình huống; còn không gian ý niệm là các không gian khái niệm được kích hoạt trong quá trình diễn giải (Kövecses, 2020). Mô hình này giúp giải thích cách các tầng ý niệm tương tác để tạo ra những biến thể ẩn dụ phù hợp với từng ngữ cảnh.

Bên cạnh đó, Cameron và Deignan (2006) nhấn mạnh rằng ẩn dụ không chỉ là hiện tượng ngôn ngữ tĩnh mà là một phần của quá trình diễn ngôn năng động, nơi ý nghĩa được kiến tạo. Ẩn dụ đóng vai trò như một công cụ nhận thức giúp con người hiểu và ghi nhớ thông tin, đồng thời là một chiến lược giao tiếp có thể được sử dụng để thuyết phục và giải thích trong tương tác học thuật (Cameron & Deignan, 2006, tr. 673-676, 688). Do đó, việc phân tích ẩn dụ trong văn bản học thuật cần vừa kế thừa tính hệ thống của Lý thuyết Ẩn dụ ý niệm, vừa áp dụng sự linh hoạt theo bối cảnh mà Lý thuyết Ẩn dụ ý niệm mở rộng đề xuất, để phản ánh chính xác sự vận động của ẩn dụ trong môi trường học thuật đa dạng và phức tạp.

2.2. Phương pháp nhận diện ẩn dụ (MIP & MIPVU)

Phương pháp "Quy trình nhận diện ẩn dụ" (*Metaphor Identification Procedure* - viết tắt MIP) do Pragglejaz Group (2007) đề xuất là cách tiếp cận hệ thống nhằm phát hiện ẩn dụ trong diễn ngôn bằng việc so sánh nghĩa cơ bản và nghĩa ngữ cảnh của một đơn vị từ vựng. Nếu hai nghĩa này khác nhau nhưng được kết nối thông qua một ánh xạ ý niệm, đơn vị từ đó được xác định là ẩn dụ. MIP nhấn mạnh việc tra cứu nghĩa cơ bản trong từ điển đơn ngữ chuẩn để đảm bảo tính khách quan và nhất quán trong phân tích.

Nhằm khắc phục những hạn chế của MIP trong việc nhận diện các ẩn dụ phức hợp hoặc có tính thông dụng, Steen và công sự (2010) đã phát triển "Quy trình nhận diện ẩn dụ Vrije Universiteit" (*Metaphor Identification Procedure Vrije Universiteit* - viết tắt MIPVU). MIPVU mở rộng phạm vi phân tích của MIP sang nhiều cấp độ cấu trúc (từ đơn vị từ đơn đến cụm từ và thành ngữ), đồng thời bổ sung hướng dẫn xử lý các trường hợp ẩn dụ thông dụng vốn thường bị bỏ sót trong phân tích tự động.

Theo Steen và công sự (2010, tr. 27-29), MIPVU bao gồm các bước:

- Đọc toàn bộ văn bản để nắm ý nghĩa tổng thể.
- Xác định và phân tách các đơn vị từ vựng.
- Xác định nghĩa ngữ cảnh của từng đơn vị từ vựng.
- Tra cứu nghĩa cơ bản trong từ điển chuẩn.
- So sánh nghĩa cơ bản và nghĩa ngữ cảnh để quyết định tính ẩn dụ.
- Ghi chú loại và mức độ ẩn dụ.

Nacey và công sự (2019, tr. 20-21) đã diễn giải chi tiết hơn các bước trong quy trình MIPVU, bổ sung các hướng dẫn thực hành và ví dụ minh họa cụ thể, từ đó giúp tăng tính tái lập (replicability) và độ tin cậy liên mã hóa (inter-coder reliability) khi áp dụng cho các khối ngữ liệu quy mô lớn. Trong đó, tính tái lập đề cập đến khả năng một nghiên cứu có thể được lặp lại bởi những người khác với cùng quy trình và dữ liệu, cho ra kết quả tương tự, đảm bảo tính nhất quán của phương pháp. Độ tin cậy liên mã hóa là mức độ nhất quán giữa các nhà mã hóa độc lập khi cùng áp dụng bộ tiêu chí để gán nhãn hoặc phân loại dữ liệu, phản ánh mức độ rõ ràng và khả năng áp dụng của hệ thống mã hóa. Ngoài ra, Steen (2002, tr. 23-24) khuyến nghị nên bắt đầu từ những trường hợp ẩn dụ rõ ràng nhất để giảm thiểu mơ hồ và nâng cao độ chính xác trong giai đoạn mã hóa ban đầu, trước khi mở rộng sang các trường hợp phức tạp hơn.

2.3. AntConc và xử lý khối ngữ liệu

AntConc là một phần mềm phân tích ngữ liệu miễn phí, được phát triển bởi Laurence Anthony, cung cấp nhiều chức năng hỗ trợ mạnh mẽ cho nghiên cứu ngôn ngữ học, bao gồm:

- Word List (danh sách từ): tạo danh sách tất cả các từ xuất hiện trong khối ngữ liệu kèm tần suất.
- Concordance (dòng ngữ cảnh): hiển thị từng lần xuất hiện của một từ khóa cùng ngữ cảnh trước và sau từ đó.
- Collocates (từ đi kèm): xác định các từ thường xuất hiện gần từ khóa, qua đó giúp phát hiện các kết hợp từ có ý nghĩa đặc biệt.
- KWIC - Keyword in Context (từ khóa trong ngữ cảnh): trình bày từ khóa ở vị trí trung tâm, bao quanh bởi một khoảng văn bản nhất định, tạo điều kiện thuận lợi cho phân tích nghĩa.

Trong nghiên cứu ẩn dụ, AntConc hỗ trợ nhà nghiên cứu ở nhiều khía cạnh quan trọng:

1. Xác định tần suất xuất hiện (frequency count) của từ hoặc cụm từ trong toàn bộ corpus, giúp nhận diện các mục từ có khả năng mang ẩn dụ dựa trên mức độ phổ biến.
2. Trích xuất ngữ cảnh sử dụng (context retrieval) của từ hoặc cụm từ trong văn bản, hỗ trợ phân tích sự thay đổi nghĩa theo ngữ cảnh.
3. Phát hiện các mẫu kết hợp từ (collocation patterns) có khả năng gợi ý sự hiện diện của ẩn dụ ý niệm.

Backer và công sự (2023, tr. 145-146) mô tả AntConc như một công cụ mạnh mẽ để truy xuất các chuỗi từ vựng (lexical strings) và mẫu đồng xuất hiện (co-occurrence patterns), qua

đó giúp nhà nghiên cứu nhanh chóng khoanh vùng các dữ liệu cần phân tích chuyên sâu. Ferrando (2019, tr. 58) bổ sung rằng việc kết hợp phần mềm xử lý ngữ liệu với quy trình nhận diện ẩn dụ thủ công không chỉ rút ngắn đáng kể thời gian tìm kiếm mà còn nâng cao độ chính xác, đặc biệt khi đối chiếu từ khóa với danh sách mục tiêu đã xác định.

Nhờ sự kết hợp giữa khả năng xử lý dữ liệu quy mô lớn của AntConc và sự phân tích tinh tế từ quy trình thủ công, các nghiên cứu ẩn dụ có thể đạt được sự cân bằng giữa tính toàn diện và tính chính xác, đồng thời duy trì được khả năng giải thích kết quả dựa trên bằng chứng ngữ liệu cụ thể.

2.4. Kết hợp AntConc và MIPVU trong phân tích ẩn dụ

Việc kết hợp AntConc và MIPVU cho phép khai thác thế mạnh bổ sung của cả hai phương pháp tiếp cận định lượng và định tính trong nghiên cứu ẩn dụ ý niệm (Meyers, 2025, tr. 22-23). AntConc đảm nhận vai trò xử lý dữ liệu quy mô lớn, hỗ trợ nhanh chóng truy xuất các từ khóa (keywords) và mẫu kết hợp từ (collocation patterns) tiềm năng, từ đó hình thành danh sách đơn vị từ vựng ưu tiên cho phân tích chuyên sâu. Trong khi đó, MIPVU - với quy trình nhận diện thủ công có hệ thống - đảm bảo mỗi trường hợp ẩn dụ được xác định dựa trên tiêu chí ngữ nghĩa chặt chẽ và có thể tái lập trong những nghiên cứu sau.

Steen và công sự (2010, tr. 79, 91-92) nhấn mạnh rằng quy trình MIPVU, khi được áp dụng trên dữ liệu đã được xử lý sơ bộ bằng phần mềm phân tích ngữ liệu, khối lượng công việc thủ công có thể giảm đáng kể mà vẫn duy trì độ chính xác và tính tin cậy liên mã hóa (*inter-coder reliability*).

Bổ sung quan điểm này, Nacey và công sự (2019, tr. 20-21) cũng chỉ ra rằng việc sắp xếp dữ liệu theo dạng KWIC hoặc concordance lines trước khi mã hóa giúp người nghiên cứu dễ dàng so sánh nghĩa cơ bản và nghĩa ngữ cảnh, đồng thời phát hiện các mẫu ẩn dụ tinh tế mà phương pháp tự động khó nhận diện.

Nhờ chiến lược kết hợp này, quy trình phân tích không chỉ tối ưu hóa hiệu quả mà còn duy trì chiều sâu diễn giải. Đồng thời, cách tiếp cận này có thể linh hoạt áp dụng cho nhiều loại khối ngữ liệu học thuật khác nhau, đặc biệt hữu ích trong bối cảnh nghiên cứu về viết học thuật tiếng Anh của sinh viên Việt Nam.

3. Quy trình phân tích ẩn dụ

Nghiên cứu này sử dụng một khối ngữ liệu gồm 352 bài nghiên cứu học kỳ của sinh viên năm cuối chuyên ngành tiếng Anh tại một trường đại học ở Việt Nam, được thực hiện trong học phần "Viết nghiên cứu" (Research Writing). Các bài viết có độ dài trung bình 2.000-3.000 từ và tuân thủ cấu trúc chuẩn của một nghiên cứu khoa học (tiêu đề, tóm tắt, giới thiệu, tổng quan tài liệu, phương pháp, kết quả, thảo luận và kết luận). Toàn bộ văn bản được thu thập trong giai đoạn 2014-2024, đã được ẩn danh thông tin cá nhân, chuẩn hóa định dạng và chuyển sang tệp .txt để xử lý bằng phần mềm AntConc.

Trên cơ sở khối ngữ liệu này, nghiên cứu thiết lập một quy trình chú giải bán tự động kết hợp phần mềm AntConc và phương pháp MIPVU. Quy trình vừa kế thừa từ các công trình trước (Steen và cộng sự, 2010; Nacey và cộng sự, 2019; Meyers, 2025), vừa được điều chỉnh để áp dụng trực tiếp cho dữ liệu hiện tại. Toàn bộ quy trình được thiết kế thành ba giai đoạn với bảy bước. Ba giai đoạn này gồm: (1) Chuẩn bị và trích xuất dữ liệu, (2) Nhận diện và phân loại ẩn dụ theo phương pháp MIPVU, và (3) Lưu trữ cùng kiểm chứng dữ liệu.

Về mặt kỹ thuật, MIPVU được xây dựng trên nền tảng chuẩn của Steen và cộng sự (2010) và Nacey và cộng sự (2019, tr. 20-21), trong khi AntConc chứng minh hiệu quả trong việc truy xuất chuỗi từ vựng và mẫu đồng xuất hiện để khoanh vùng dữ liệu giàu tiềm năng ẩn dụ (Backer và cộng sự, 2023, tr. 145-146; Ferrando, 2019, tr. 58).

Cách tiếp cận này cho phép kết hợp ưu thế phân tích ngữ liệu định lượng với độ chính xác và khả năng giải thích của phương pháp thủ công có hệ thống, nhờ đó đảm bảo tính khoa học, minh bạch và khả năng tái lập cho nghiên cứu. Quy trình cụ thể sẽ được trình bày chi tiết ở các tiểu mục sau và minh họa bằng ví dụ trong Mục 4.

Giai đoạn 1: Chuẩn bị và trích xuất dữ liệu

Mục tiêu: Tạo "đầu vào sạch" và danh sách mục tiêu đủ hẹp để phân tích sâu. Giai đoạn này bao gồm 3 bước:

Bước 1: Chuẩn bị ngữ liệu

- Đầu vào: Toàn bộ bài nghiên cứu học thuật của sinh viên (đã ẩn danh), bao gồm tiêu đề, tóm tắt, giới thiệu, tổng quan tài liệu, phương pháp, kết quả và kết luận.
- Thao tác:
  - Chuyển toàn bộ sang định dạng .txt và chuẩn hóa mã UTF-8.
  - Loại bỏ ký hiệu đặc biệt, ngắt dòng thừa hoặc ký tự lỗi để tránh sai lệch khi đếm tần suất.
- Đầu ra: Thư mục .txt đã chuẩn hóa, sẵn sàng nhập vào AntConc.

Bước 2: Truy xuất từ khóa tiềm năng

- Đầu vào: Toàn bộ khối ngữ liệu ở định dạng .txt.
- Thao tác:
  - Sử dụng công cụ Word List để lập danh sách tần suất.
  - Loại bỏ các từ ít khả năng tạo ẩn dụ (từ chức năng, tên riêng, số đếm, từ hỏi, từ viết tắt).
  - Ưu tiên danh từ, động từ, tính từ có tiềm năng tạo ẩn dụ; có thể dùng Collocates để xác định cụm giàu tiềm năng.
- Đầu ra: Danh sách rút gọn gồm các từ / cụm từ mục tiêu có tần suất cao để phân tích.

Bước 3: Trích xuất ngữ cảnh

- Đầu vào: Danh sách từ mục tiêu.
- Thao tác:
  - Trong Concordance/KWIC, trích câu hoặc cụm câu với cửa sổ ±10 từ quanh từ khóa.
  - Sắp xếp dòng concordance (sort by 1L/1R) để nhận diện mẫu sử dụng từ.
- Đầu ra: Tệp KWIC (CSV/TSV) cho từng từ khóa, sẵn sàng cho bước MIPVU.

Giai đoạn 2: Nhận diện và phân loại ẩn dụ

Mục tiêu: Áp dụng MIPVU để xác định, gắn nhãn và ánh xạ miền ý niệm. Giai đoạn này bao gồm 3 bước:

Bước 4: Xác định nghĩa cơ bản và nghĩa ngữ cảnh (MIPVU)

- Đầu vào: Các dòng KWIC (mỗi dòng = 1 đơn vị từ vựng).
- Thao tác:
  - Đọc toàn văn để hiểu nghĩa tổng thể.
  - Xác định đơn vị từ vựng.
  - Xác định nghĩa ngữ cảnh.
  - Tra nghĩa cơ bản trong từ điển uy tín (Longman, Cambridge…).
  - So sánh để quyết định tính ẩn dụ.
- Đầu ra: Cột nhãn M/L/U (Metaphorical - ẩn dụ /Literal - nghĩa đen / Uncertain - không chắc) cho từng đơn vị từ vựng, kèm trích dẫn từ điển.

Bước 5: Nhận diện và đánh dấu ẩn dụ

- Đầu vào: Bảng M/L/U.
- Thao tác:
  - Đánh dấu M (theo tiêu chí MIPVU) nếu nghĩa ngữ cảnh khác nghĩa cơ bản và có ánh xạ ý niệm.
  - Ghi lý do (ví dụ: “movement → progress”) vào cột decision note.
- Đầu ra: Danh sách các đơn vị ẩn dụ đã được xác nhận.

Bước 6: Gán miền đích

- Đầu vào: Các đơn vị ẩn dụ đã xác nhận.
- Thao tác:

- Gán miền ý niệm theo 11 miền chuẩn thường xuất hiện trong diễn ngôn học thuật, bao gồm:

1. Hiểu biết là ánh sáng/ánh nhìn (*Understanding as Vision/Light*)
2. Lý thuyết là công trình/kiến trúc (*Theories as Buildings/Structures*)
3. Nghiên cứu là hành trình/khám phá (*Research as Exploration/Discovery*)
4. Học tập là sự phát triển/dinh dưỡng (*Learning as Growth/Nourishment*)
5. Văn bản học thuật là tác nhân/cấu trúc/vật chứa (*Academic Text as Agent/Structure/Container*)
6. Quá trình khoa học là máy móc/công cụ (*Scientific Processes as Machines/Tools*)
7. Tri thức là không gian/sự tiến triển (*Knowledge as Space/Progression*)
8. Thực thể trừu tượng là con người/sự nhân hóa (*Abstract Entities as Actors/Personification*)
9. Ý niệm là vật chứa (*Concepts as Containers*)
10. Cộng đồng học thuật là cuộc đối thoại (*Academic Community as Conversation*)
11. Lập luận là chiến tranh/xung đột (*Argument as War/Conflict*)

- Giải thích thêm: 11 miền này được xác lập dựa trên các công trình kinh điển về ẩn dụ ý niệm (Lakoff & Johnson, 1980; Kövecses, 2020; Semino, 2008; Nacey và công sự, 2019). Chúng được lựa chọn vì thường xuyên xuất hiện trong văn bản khoa học, vừa đủ để bao quát đa số biểu thức ẩn dụ, vừa đảm bảo tính thống nhất và khả năng so sánh với các nghiên cứu trước.
- Đầu ra: Bảng mã hóa được thiết kế để ghi nhận Miền đích và Loại ẩn dụ tương ứng trong ngữ cảnh phân tích.

Giai đoạn 3: Lưu trữ và kiểm chứng

Mục tiêu: Bảo đảm tính minh bạch, tin cậy liên mã hóa và truy vết. Giai đoạn này có một bước như sau:

Bước 7: Lưu trữ và kiểm chứng dữ liệu

- Đầu vào: Danh sách các đơn vị ẩn dụ đã được gán miền ý niệm.
- Thao tác:

- Nhập toàn bộ kết quả vào bảng mã hóa chuẩn với các cột:

1. Từ khóa (Keyword) - đơn vị mục tiêu được phân tích.
2. Câu gốc (Original Sentence) - trích nguyên câu từ khối ngữ liệu chứa từ khóa.
3. Nghĩa ngữ cảnh (Contextual meaning) - cách hiểu trong ngữ cảnh cụ thể.

4. Nghĩa cơ bản (Basic meaning) - nghĩa gốc, cụ thể, thường gắn với trải nghiệm thân thể.
5. Khác biệt ý niệm (Conceptual Difference) - sự chênh lệch giữa nghĩa cơ bản và nghĩa ngữ cảnh, làm cơ sở nhận diện ẩn dụ.
6. Miền đích (Target domain) - miền ý niệm mà từ khóa ánh xạ đến.

- Lưu kèm sổ tay mã hóa gồm: tiêu chí lọc, tham chiếu từ điển, và những ví dụ khó phân loại (nếu có).

- Thực hiện kiểm tra chéo trên một tỷ lệ mẫu hợp lý; thảo luận các khác biệt (nếu có) và chuẩn hóa sổ tay mã hóa.

(Tùy chọn) tính chỉ số tin cậy liên mã hóa (ví dụ: Cohen's κ) để báo cáo.

- Đầu ra: Một bộ dữ liệu có cấu trúc như đã nêu, kèm theo phiên bản lưu trữ (theo dõi các lần chỉnh sửa), bản ghi quyết định và biên bản kiểm chứng, sẵn sàng cho việc tái sử dụng và đối chiếu trong các nghiên cứu sau.

Quy trình phân tích vừa khai thác được ưu thế về tốc độ và độ bao quát của phần mềm ngữ liệu, vừa kết hợp với sự chính xác và chiều sâu giải thích từ phương pháp thủ công có hệ thống. Sự kết hợp này giúp nghiên cứu đạt được yêu cầu về tính khoa học, tính minh bạch và khả năng tái lập trong toàn bộ quá trình xử lý dữ liệu.

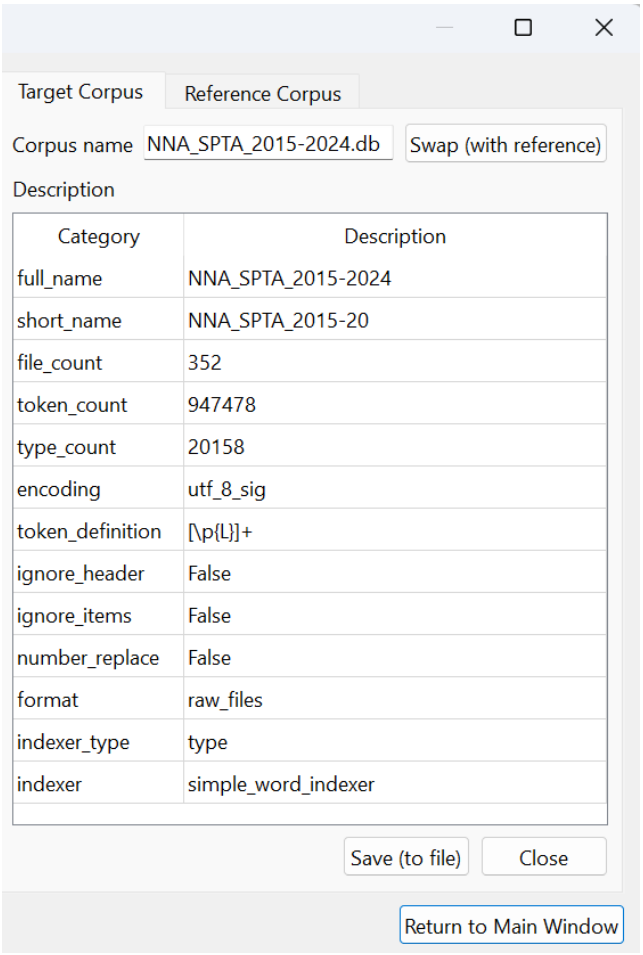
4. Ví dụ minh họa

Phần này minh họa quá trình áp dụng quy trình bán tự động kết hợp AntConc và MIPVU đã trình bày ở Mục 3. Ngữ liệu sử dụng chính là khối 352 bài nghiên cứu học kỳ của sinh viên ngành tiếng Anh, đã được chuẩn hóa và xử lý để phục vụ phân tích. Ví dụ dưới đây sẽ lần lượt thể hiện bảy bước theo ba giai đoạn của quy trình, với trích dẫn dữ liệu và bình luận cụ thể để làm rõ cách nhận diện và phân loại ẩn dụ trong văn bản học thuật.

Giai đoạn 1: Chuẩn bị và trích xuất dữ liệu

Bước 1 - Chuẩn bị ngữ liệu

Toàn bộ 352 bài viết được chuyển sang định dạng .txt, chuẩn hóa mã hóa UTF-8 và loại bỏ các ký tự lỗi để đảm bảo đầu vào sạch.



Hình 1. *Cấu hình khối ngữ liệu trong AntConc*

Bước 2 - Truy xuất từ khóa tiềm năng

Bằng công cụ Word List trong AntConc, các từ có tần suất cao và tiềm năng ẩn dụ được xác định. Ví dụ: *Trên tổng số 947,478 từ*, các từ như time (2,351 lần), information (1,567 lần) và communication (792 lần) xuất hiện với tần suất cao. Chỉ giữ lại các từ loại nội dung (danh từ, động từ, tính từ) có khả năng tạo ẩn dụ, bỏ qua từ chức năng.

	File	Left Context	Hit	Right Context
1	SPTA_201...	l understanding the student's view point. It started with the assumption that students didn't have much	time	to practice and communicate with other. The findings, however, both prove and disprove those assump
2	NNA_201...	nglish. It may imply that a majority of students have lower English level than others, so they need more	time	to practice and improve their English skills. Figure 6 shows how male and female students performance
3	NNA_202...	to speak fluently, which account for 44% and 42%. Moreover, 17 and 15 of them also have insufficient	time	to practice and lack lexical resources, which are responsible for 34% and 30%. 28% of the participants d
4	NNA_202...	y are at unease in the context of public speaking. It is probably because they had had relatively enough	time	to practice and improve their speaking skills through some compulsory modules such as Listening-Spea
5	SPTA_201...	o stumbled at speed of speech. A tentative interpretation might be that students did not have enough	time	to practice speaking at school, so they end up with repeated statement and self correction. Juniors do n
6	DSA_LT_2...	ll of new words. Futhermore, students didn't learn vocabulary by heart at home. They didn't have much	time	to practice speaking with its new words. None of student reported they didn't know new words very mu
7	NNA_LT_2...	hink that speaking skills are important to their work. The overall results table shows that the amount of	time	to practice speaking skills of G1 is only a quarter its of G2. Specially, the places where G2 practice
8	NNA_LT_2...	gure 3.2. More specifically, participants may not have both many chances to speak English and enough	time	to practice speaking skills; thus, they are more likely not to self-evaluate highly their speaking competen
9	SPTA_201...	from the multiple choices question. This result indicate clearly what the students want; they want more	time	to practice their speaking. 4. DISCUSSION: The research started with the caveat that it would help with in
10	SPTA_201...	d the most important role in the learning, while students only need to listen and write and they have no	time	to practice their speaking ability. Consequently, "the students can not do communication with a native s
11	NNA_202...	cing subjects in English Language Education Department in Sai Gon University. The students need more	time	to practice their speaking to reduce those difficulties especially outside the classroom, and more confide
12	NNA_LT_2...	class whereas others talk very little or never speak; therefore, they tend to want to have more speaking	time	to practice their speaking skills (Leong et al., 2017). B. Students' satisfaction with the form of learning Fig
13	NNA_202...	reliance of data. The results showed several subjective and objective reasons such as having less or no	time	to practice or some methods in teaching English leading to mispronunciation in speaking English which
14	NNA_202...	reliance of data. The results showed several subjective and objective reasons such as having less or no	time	to practice or some methods in teaching English leading to mispronunciation in speaking English which
15	SPTA_201...	difficulties on time, students attitude and their vocabulary. Nowadays, students might not have enough	time	to practice all skills, especially listening out of class. Besides, both learners and teachers realize that time
16	NNA_201...	pported enough time to practice listening skill at University, and 36% of them said they just had a little	time	to practice at University. Still nobody in 40 students said they had very much time to improve their skill.
17	SPTA_201...	assignments, presentations, examinations at university and part-time jobs appears to make the lack of	time	to practice become the most frequent difficulty. In term of those who already took the test, one possibl
18	NNA_201...	d by considering the major of the participants. It appears that second-year students try to spend more	time	to practice English speaking since they are in the English major. The chart provides five common mistak
19	SPTA_201...	as much as they do inside. The consensus from most of the subjects was that there need to be more	time	to practice for students. 1. INTRODUCTION: English is regarded to be the most popular language in the
20	NNA_201...	helpful. Table 4 also reflects that more than half of students (64%) said they did not supported enough	time	to practice listening skill at University, and 36% of them said they just had a little time to practice at

Hình 2. *Dòng concordance cho từ khóa "time"*

	File	Left Context	Hit	Right Context
1	NNA_201...	evious studies suggesting that the park is the best place to go picnic of all people Pie chart 6 shows	information	about the percentage of some period time when SGU's students chose to go picnic. Students often
2	SPTA_201...	ils and comprehend content of the text during the lesson in class 10A1. The three pie charts provide	information	about the percentage of the students who listened the main idea, details and understood very well i
3	SPTA_201...	every day because their major is English and they use English for their study or work. Table 3.2 gives	information	about the percentage of students on spending time to study English. Most students spent 1 to 2 ho
4	NNA_202...	ard for native speakers to follow what they said while communicating in English. The pie chart gives	information	about the percentage of participants ranking the importance of pronunciation. Overall, the majority
5	NNA_202...	t 5 years of learning English as it is one of the compulsory subjects in the curriculum. Figure 2 gives	information	about the percentage of participants who rank their personal pronunciation proficiency. Overall, it ca
6	NNA_202...	y been introduced to interactive online learning since the end of September 2022. Figure 2 provides	information	about the percentage of students who meet the eligibility requirements for online participation in a
7	NNA_LT_2...	tion problems, vocabulary problems, concentration, listening practice methods, etc. Figure 10 gives	information	about the percentage of participants who practice their personal listening to English. In general, a lar
8	NNA_LT_2...	y been introduced to interactive online learning since the end of September 2022. Figure 2 provides	information	about the percentage of students who meet the eligibility requirements for online participation in a
9	NNA_202...	3. RESULTS Figure 1. Participants using public speaking rate in practicum The given bar graph gives	information	about the percentage of 60 participants using public speaking skills to convey lessons to their stude
10	SPTA_201...	questions. The format of the survey consisted of three main parts. Part One collected demographic	information	about the participants with the first question asked students about their year in university, the secon
11	NNA_LT_2...	emic year 2017-2018. The questionnaire was in printed form for data collection. It included personal	information	about the participants and twelve statements containing the problems of reading comprehension wi
12	NNA_LT_2...	into three different sections. The first one is from questions 1 to 4 which are used to collect general	information	about the participants including the gender, the academic degree, and years of learning English, and
13	NNA_LT_2...	gory, multiple choice with open-ended and scaling. The first part of the questionnaire is some basic	information	about the participants which includes gender, age group, years of learning English, and self-evaluatic
14	SPTA_201...	lecting the learners' opinion about the application of the games into the lesson. The first table gives	information	about the proportion of grade 10 students from four separate high schools who made comments o
15	NNA_201...	nproved to make the negatives change their mind, and convince the neutrals, as well. Thechart gives	information	about the proportion of major kind of resources that 32 SaiGon University students got information
16	SPTA_201...	on was used to reinforce the accuracy of the data calculated. CHAPTER 3 RESULT The bar chart gives	information	about the proportion of male and female students taking part in English speaking test. Overall, more
17	NNA_202...	edback from the audience. Figure 4. The audience's negative attitudes and reactions. The table gives	information	about the proportion of the respondents' responses toward the audience's negative attitudes and re
18	NNA_LT_2...	more likely not to self-evaluate highly their speaking competence. Table 3.1 and Figure 3.4 provide	information	about the levels of DNA420101 students' favorite in four English skills. Among four skills, the most f
19	NNA_LT_2...	tor in author's research. D. Students' satisfaction with their acquired speaking skills Figure 3.10 gives	information	about the levels of students' satisfaction with their acquired speaking skills. In general, most participa
20	NNA_202...	sity. Figure 1. Students' levels of anxiety and English - speaking self – efficacy. The table above gives	information	about the levels of students' anxiety when speaking English both in general and public and their self

Hình 3. *Dòng concordance cho từ khóa "information"*

	File	Left Context	Hit	Right Context
1	NNA_202...	f acquiring a second language. Rohmatillah (2017) asserted that without learning the vocabulary,	communication	in the second language becomes harder. To be specific, vocabulary learning helps not only acquir
2	NNA_202...	r two semesters, which aimed at equipping the students with basic genre knowledge about daily	communication.	In the second academic year they are required to learn public speaking for different purpose for t
3	NNA_201...	ıication in classroom. The author reports that research regarding the development of non-verbal	communication	in the classroom has demonstrated that there are five component parts to this process including:
4	NNA_202...	e, at 6.0% and 7.5% in that order. The percentage of freshman students who sometimes practice	communication	in the foreign language is higher than those who always do, at 31.3% and 17.9%, respectively. It i
5	NNA_202...	al with their own anxiety states, so that they can speak English better and develop confidence in	communication	in the future. Gap statement Although many pieces of research have been done, there is a little in
6	NNA_202...	ı profound impact on the audiences reactions to the talk compared to other types of non-verbal	communication.	In the light of these conclusions, our findings are consistent with previous research, which also su
7	NNA_202...	traction. INTRODUCTION Setting It is said that English is global and is one of the widest ways of	communication	in the world. New Lyons figures (2021) show that the number of people using English as their firs
8	NNA_202...	alf of students expressed that they had ever communicated with native speaker, the frequency of	communication	in English accounted a small proportion. Lastly, a half of participants rated themselves as average
9	NNA_202...	rtable with real- life communication in English, while the latter group might struggle with real-life	communication	in English due to their limited abilities. This data underscored the challenges faced by English Maj
10	NNA_202...	ıeir students become better listeners, thereby improving their overall language skills and making	communication	in English easier. 2. METHODOLOGY 2.1. Overview This study aims to make a tentative investigati
11	NNA_202...	technology serves as a valuable tool for enhancing language proficiency and facilitating effective	communication	in English for non-native speakers. Previous research has provided evidence of ChatGPT's impres
12	NNA_202...	ntages in this chart. More specifically, the number of students who evaluated they were good at	communication	in English is about six times that of weak students, at 38.8% and 6.0%. Meanwhile, the figures for
13	NNA_202...	ionnaire was conducted with 50 students to collect the data. According to the research, effective	communication	in English was found to be significantly influenced by accurate pronunciation. The survey showed
14	NNA_202...	Figure 2 illustrates the students opinion on how significant correct pronunciation is for effective	communication	in English. It is clear from the chart that most students agreed that it was necessary, with 46 out c
15	NNA_202...	ata underscored the challenges faced by English Major Students at Sai Gon University in real- life	communication	in English. It highlights the need for targeted interventions and supports, particularly for those st
16	NNA_202...	ıey spend less time practicing. In addition, correct pronunciation was a crucial factor for effective	communication	in English. At the same time, pronunciation was to be most difficult in the context of a job intervi
17	NNA_202...	OF ELF AT SAIGON UNIVERSITY ABSTRACT Background. Pronunciation is increasingly focused on	communication	in English. Ending sounds play an important role in conveying the accurate meaning of words an
18	SPTA_201...	ıcy. Vietnamese students study English as a preparation for their future jobs that usually require	communication	in English. In high schools, learning English becomes more important for students since it is one
19	NNA_202...	is research embarks on a crucial exploration of the challenges faced by these students in real-life	communication	in English. Recognizing the paramount importance of effective communication skills in the global
20	NNA_202...	To thoroughly investigate the challenges confronted by English major students at SGU in real-life	communication	in English, a meticulously designed survey instrument, comprising a structured questionnaire with

Hình 4. *Dòng concordance cho từ khóa "communication"*

Bước 3 - Trích xuất ngữ cảnh

Các dòng concordance được trích xuất bằng công cụ KWIC với phạm vi mặc định là ±10 từ xung quanh từ khóa. Tuy nhiên, nhằm đảm bảo bối cảnh ngữ nghĩa đầy đủ hơn cho việc nhận diện ẩn dụ, phạm vi này được mở rộng lên ±20 từ, đồng thời toàn bộ câu chứa từ khóa cũng được lấy trọn vẹn.

Ví dụ, với từ "time":

"Nowadays, students might not have enough *time* to practice all skills, especially listening out of class."

Dữ liệu được lưu thành file .csv hoặc .tsv để chuyển sang bước phân tích tiếp theo.

Giai đoạn 2: Nhận diện và phân loại ẩn dụ

Bước 4 - Xác định nghĩa cơ bản và nghĩa ngữ cảnh (MIPVU)

Ví dụ với từ "time":

- Nghĩa ngữ cảnh: Time is treated as a quantifiable, limited space or resource. (Thời gian được xem như một không gian hoặc tài nguyên có thể định lượng và bị giới hạn.)
- Nghĩa cơ bản: time = The thing that is measured in minutes, hours, days, years etc using clocks (Longman Dictionary) (Thời gian là đại lượng được đo bằng phút, giờ, ngày, năm… thông qua đồng hồ)

Bước 5 - Nhận diện và đánh dấu ẩn dụ

- So sánh: Time is seen as a container students can "have" or "lack." (Thời gian được hình dung như một vật chứa mà sinh viên có thể "có" hoặc "thiếu".)

Việc so sánh nghĩa ngữ cảnh và nghĩa cơ bản cho thấy có sự khác biệt rõ ràng về ý niệm. Sự khác biệt này hàm ý rằng từ *time* đang được sử dụng theo cách ẩn dụ.

Cụ thể, ẩn dụ ý niệm thể hiện ở đây là "Concepts as Containers" (*Ý niệm là vật chứa*), trong đó "thời gian" được hình dung như một không gian chứa đựng hữu hình, có thể được phân bổ, bị thiếu hụt hoặc chiếm dụng.

Theo quy trình MIPVU, từ *time* được đánh dấu nhãn M (Metaphorical, nghĩa ẩn dụ) vì:

- Nghĩa ngữ cảnh khác biệt với nghĩa cơ bản.
- Có sự ánh xạ giữa hai miền tri nhận theo hướng ẩn dụ (container metaphor).

Bước 6 - Gán miền đích (Target Domain)

Dựa trên ánh xạ ẩn dụ đã xác định ở Bước 5, từ *time* trong ngữ cảnh phân tích được gán vào miền đích "Concepts as Containers" (Ý niệm là vật chứa). Đây là một trong 11 miền ý niệm chuẩn thường gặp trong diễn ngôn học thuật đã được trình bày ở phần trên.

Giai đoạn 3: Lưu trữ và kiểm chứng

Bước 7 - Lưu trữ và kiểm chứng dữ liệu

Dữ liệu được lưu vào bảng mã hóa chuẩn, bao gồm các trường: Keyword, Original Sentence, Contextual Meaning, Basic Meaning, Conceptual Difference, Target Domain. Cấu trúc bảng này giúp kiểm chứng và hệ thống hóa việc phân tích ẩn dụ theo phương pháp MIPVU, đảm bảo tính minh bạch và nhất quán trong toàn bộ quá trình xử lý ngữ liệu.

Ngoài từ khóa "time" đã minh họa ở các bước trên, bước này được mở rộng để mã hóa thêm hai từ khóa mới là "information" và "motivation", nhằm phản ánh đa dạng hơn các hình thức ánh xạ ẩn dụ và chức năng tu từ của chúng trong văn bản học thuật của người học EFL.

Bảng mã hóa dữ liệu ẩn dụ theo MIPVU (phần mở rộng)

Keyword	**Original Sentence**	**Contextual Meaning**	**Basic Meaning**	**Conceptual Difference**	**Target Domain**
Time	Nowadays, students might not have enough *time* to practice all skills, especially listening out of class.	Time is treated as a quantifiable, limited space or resource.	The thing that is measured in minutes, hours, days, years etc using clocks (Longman Dictionary)	Time is seen as a container students can "have" or "lack."	Concepts as Containers

Motivation	It seems that the requirement of English proficiency would be their *motivation* and affect strongly to their learning process.	Motivation arises from clear goals or needs and fuels learning.	Eagernesss and willingness to do something without needing to be told	Motivation is a nutrient or catalyst that enables and affects the learning process.	Learning as growth or nourishment
Communication	Since practical *communication* is a primary aim of learning English, speaking becomes an essential skill that learners should master.	Communication is treated as a goal that learners work toward in language acquisition.	the process by which people exchange information or express their thoughts and feelings	The abstract process of communication is framed as a destination or achievement in the learning process.	Knowledge as space or progression

Sau khi hoàn tất, các đơn vị ẩn dụ được kiểm tra liên mã hóa trên một mẫu 10-15% để đảm bảo độ tin cậy. Mọi khác biệt trong mã hóa được thảo luận và ghi vào sổ tay mã hóa.

5. Những đóng góp về mặt phương pháp

Quy trình bán tự động kết hợp AntConc và MIPVU được trình bày trong nghiên cứu này không chỉ tối ưu hóa việc nhận diện và phân loại ẩn dụ ý niệm trong ngữ liệu tiếng Anh học thuật EFL, mà còn đưa ra một khung phân tích rõ ràng, khả thi và có thể tái sử dụng.

5.1 Đóng góp chính

Nghiên cứu này mang lại một số đóng góp đáng chú ý cho lĩnh vực phân tích ẩn dụ ý niệm trong văn bản học thuật.

- Việc sử dụng AntConc hỗ trợ lọc và trích xuất từ khóa tiềm năng một cách nhanh chóng, từ đó tạo ra dữ liệu định lượng phục vụ cho phân tích định tính.
- Quy trình 7 bước theo Steen và cộng sự (2010) mang lại sự minh bạch trong việc xác định ẩn dụ so với nghĩa literal, đồng thời hạn chế tối đa sự chủ quan của người mã hóa.

- Cách cấu trúc quy trình cho phép linh hoạt áp dụng cho nhiều thể loại văn bản khác nhau, cũng như thích nghi với dữ liệu song ngữ hoặc đa ngữ.
- Dù được triển khai trên văn bản tiếng Anh, quy trình hoàn toàn có thể điều chỉnh để phân tích ẩn dụ trong tiếng Việt, đặc biệt khi kết hợp với từ điển nghĩa cơ bản và các công trình nghiên cứu trước đây về ẩn dụ ý niệm trong tiếng Việt. Nhờ đó, phương pháp bán tự động này không chỉ hữu ích cho tiếng Anh mà còn tạo nền tảng cho các nghiên cứu so sánh đối chiếu và đa ngôn ngữ.

5.2 Đặc điểm và định hướng

Quy trình này có một số đặc điểm nổi bật cần nhấn mạnh. Trước hết, nó có tính *tái áp dụng cao*, có thể được vận dụng cho nhiều thể loại văn bản học thuật khác nhau và ở nhiều ngôn ngữ, không chỉ giới hạn trong tiếng Anh EFL. Nhờ sự kết hợp giữa phần mềm ngữ liệu và phương pháp thủ công, quy trình vừa bảo đảm tính nhanh chóng, bao quát, vừa duy trì được chiều sâu phân tích.

Một ưu điểm quan trọng là *sự tiết kiệm thời gian và tăng tính nhất quán*. AntConc giúp rút gọn đáng kể khâu lọc, trích xuất và tổ chức dữ liệu, trong khi MIPVU đảm bảo các quyết định nhận diện ẩn dụ/literal có cơ sở rõ ràng, hạn chế tối đa yếu tố chủ quan.

Tuy vậy, quy trình vẫn có một số hạn chế: nhiều bước cốt lõi, đặc biệt là xác định nghĩa cơ bản và nghĩa ngữ cảnh, vẫn cần đến sự tham gia trực tiếp của nhà nghiên cứu. Bên cạnh đó, nghiên cứu hiện tại chưa tính toán chỉ số tin cậy liên mã hóa (như Cohen's κ) do hạn chế về nguồn lực. Điều này đồng nghĩa với việc mức độ nhất quán giữa các mã hóa viên chưa được lượng hóa, và cần được bổ sung trong các nghiên cứu tiếp theo để nâng cao độ tin cậy.

Ngoài ra, quy trình cũng có tiềm năng trong ứng dụng giáo dục: nó có thể được khai thác trong giảng dạy viết học thuật, giúp sinh viên nâng cao nhận thức về cách sử dụng ẩn dụ trong lập luận. Về lâu dài, hướng phát triển có thể là tích hợp công nghệ Trí tuệ nhân tạo (AI) và xử lý ngôn ngữ tự nhiên (Natural Language Processing - NLP) để tự động hóa sâu hơn các bước trong quy trình, như nhận diện từ vựng tiềm năng, gợi ý các đơn vị ngữ nghĩa có khả năng chứa ẩn dụ, và đề xuất phân loại sơ bộ trước khi phân tích thủ công. Việc ứng dụng các mô hình ngôn ngữ lớn (ví dụ: BERT, GPT) có thể hỗ trợ việc mở rộng quy mô nghiên cứu và tăng độ chính xác trong nhận diện ẩn dụ trên các tập dữ liệu lớn và đa ngữ.

6. Kết luận

Bài viết đã trình bày một quy trình bán tự động dựa trên khối ngữ liệu nhằm nhận diện và phân loại ẩn dụ ý niệm trong các bài viết học thuật tiếng Anh của sinh viên Việt Nam. Phương pháp đề xuất kết hợp giữa công cụ AntConc trong việc truy xuất và phân tích mẫu từ vựng, cùng quy trình MIPVU trong nhận diện ẩn dụ thủ công, qua đó vừa tận dụng được sức mạnh xử lý dữ liệu ở quy mô lớn, vừa đảm bảo độ chính xác và chiều sâu của phân tích ngôn ngữ học định tính.

Kết quả cho thấy việc tích hợp hai công cụ này không chỉ góp phần nâng cao hiệu quả và tính khách quan trong phân tích ẩn dụ học thuật, mà còn mở ra những đóng góp phương pháp luận đáng kể cho các nghiên cứu dựa trên khối ngữ liệu, hoạt động giảng dạy kỹ năng viết học thuật, cũng như thiết kế các công cụ hỗ trợ học tập ứng dụng công nghệ xử lý ngôn ngữ tự nhiên (AI/NLP).

Với tính linh hoạt và khả năng mở rộng, quy trình này có tiềm năng được tái áp dụng trong nhiều bối cảnh và thể loại văn bản khác nhau, đặc biệt trong các nghiên cứu so sánh liên ngôn ngữ, liên văn hóa, và các hệ thống học tập thông minh sử dụng nền tảng ngôn ngữ học máy tính.

TÀI LIỆU THAM KHẢO

Anthony, L. (2022). *AntConc* (Version 4.3.1) [Computer software]. Waseda University. https://www.laurenceanthony.net/software/antconc/

Backer, L., De Clercq, B., & Paquot, M. (2023). *A practical guide to learner corpus research*. Cambridge University Press.

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.

Cameron, L., & Deignan, A. (2006). The emergence of metaphor in discourse. *Applied Linguistics, 27*(4), 671-690. https://doi.org/10.1093/applin/aml032

Ferrando, L. (2019). *Metaphor in academic discourse: Linguistic forms, conceptual structures and communicative functions*. Peter Lang.

Gibbs, R. W. (2017). *Metaphor wars: Conceptual metaphors in human life*. Cambridge University Press.

Gries, S. T. (2021). *Analyzing linguistic data: A practical introduction to statistics using R* (2nd ed.). De Gruyter Mouton.

Hyland, K. (2016). *Teaching and researching writing* (3rd ed.). Routledge.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

McEnery, T., Hardie, A., & Brezina, V. (2019). The changing role of corpus linguistics. In O. Garcia, N. Flores, & M. Spotti (Eds.), *The Oxford handbook of language and society* (pp. 504-527). Oxford University Press.

Meyers, C. (2025). The conceptual metaphors of memory: Cases of interdomanial nomadism from computer science to quantum computing. *Lexis: Journal in English Lexicology, (Words about #1).* Advance online publication. https://doi.org/10.4000/13wnc

Nacey, S., Silaški, N., Dorst, A. G., & Krennmayr, T. (2019). *Metaphor identification in multiple languages: MIPVU around the world*. John Benjamins Publishing Company.

Pragglejaz Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol, 22*(1), 1-39. https://doi.org/10.1207/s15327868ms2201_1

Semino, E. (2008). *Metaphor in discourse*. Cambridge University Press.

Steen, G. J. (2002). Towards a procedure for metaphor identification. *Language and Literature, 11*(1), 17-33.

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., Krennmayr, T., & Pasma, T. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. John Benjamins.

Swales, J. M., & Feak, C. B. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd ed.). University of Michigan Press.

NHỮNG CÁCH XỬ LÝ NGỮ NGHĨA CỦA AI KHI DỊCH TỤC NGỮ VIỆT SANG TIẾNG HÀN

Trần Văn Tiếng[1], *Nguyễn Thị Kim Ngọc* [2]

Tóm tắt - Tục ngữ là đơn vị của lời nói, hình thành trong hoạt động nói năng. Nội dung ngữ nghĩa của tục ngữ rất đa dạng. Nó phản ánh quan điểm, triết lý sống thật phong phú của con người; là những bài học về cách đối nhân xử thế, về những kinh nghiệm trong hoạt động nghề nghiệp, là cách đánh giá của con người về các hiện tượng xã hội trong từng thời kỳ của lịch sử, là những nhận xét về tâm lý con người, quan điểm của con người trong các mối quan hệ…Một điểm rất đặc biệt của tục ngữ Việt là yếu tố vần, nhịp đã làm cho ngữ nghĩa của câu tục ngữ thêm súc tích, có hàm ý sâu xa. Chính vì những điều này, khi chuyển ngữ tục ngữ Việt sang tiếng Hàn, AI đã có những cách xử lý ngữ nghĩa khác nhau nhằm đảm bảo dịch đúng, đủ nội dung của câu tục ngữ Việt.

Từ khóa: Tục ngữ Việt, ngữ nghĩa, ChatGPT, ChatAI, Google translate, nghĩa biểu trưng, từ Hán Việt, cấu trúc so sánh.

1. GIỚI THIỆU

Tục ngữ là lời ăn tiếng nói của nhân dân. Tục ngữ là tấm gương soi mà nhân dân các dân tộc từ thế hệ này sang thế hệ khác có thể tìm thấy trong đó triết lý sống thật phong phú và đa dạng; không biết tự bao giờ, tục ngữ đã trở thành kho tàng vô giá. Trong kho tàng ấy, người ta thấy tục ngữ chất chứa biết bao điều. Đó là những bài học về cách đối nhân xử thế, là những bài học kinh nghiệm trong hoạt động nghề nghiệp, là cách đánh giá của con người về các hiện tượng xã hội trong từng thời kỳ của lịch sử, là những nhận xét về tâm lý con người, quan điểm của con người trong các mối quan hệ… Cho nên có thể nói rằng tục ngữ là tấm gương phản ánh văn hoá dân tộc thể hiện dưới dạng những câu nói súc tích, giàu hình ảnh và có hàm ý sâu xa.

Nhìn chung, khi nghiên cứu tục ngữ Việt, người ta thường nêu ra mấy đặc trưng chủ yếu sau đây:

- Tục ngữ là đơn vị của lời nói: "*Tục ngữ là câu nói thường ngắn gọn có vần hoặc không có vần, có nhịp điệu hoặc không có nhịp điệu đúc kết kinh nghiệm sản xuất hay đấu tranh xã hội, rút ra một chân lý phổ biến, ghi lại một nhận xét về tâm lý, phong tục, tập quán của nhân dân. Tục ngữ do nhân dân sáng tác và được toàn thể xã hội công nhận"* [1, tr.209]; hoặc "*Tục ngữ là những câu nói dân gian ngắn gọn, ổn định, phản ánh nhận thức và quan niệm có tính chất kinh nghiệm của nhân dân về thế giới tự nhiên và xã hội (...). Tục ngữ được vận dụng trong câu nói, thường được người nói dẫn vào phát ngôn của mình, dẫn nguyên dạng hay dẫn ý*" [2, tr.49].

[1] TS., Trường Đại học Ngoại ngữ-Tin học TP.HCM, *email: tiengtv@huflit.edu.vn*.
[2] ThS., Trường Đại học Ngoại ngữ-Tin học TP.HCM

- Xét về ngữ nghĩa, mỗi câu tục ngữ bao giờ cũng bao hàm hai nghĩa: *nghĩa đen* và *nghĩa bóng*. Cái cụ thể riêng lẻ sẽ tạo thành nghĩa đen, cái trừu tượng phổ biến tạo thành nghĩa bóng. Từ góc độ lôgích-ngữ nghĩa, người ta chỉ ra rằng: *Tục ngữ có những đặc thù cú pháp ít thấy ở những câu thông thường vì rằng: Đó là những câu mang tính khái quát để phản ánh những quy luật và kinh nghiệm khái quát, phổ biến. Nghĩa của phần lớn các câu tục ngữ là nghĩa biểu trưng nên tổ chức cú pháp của tục ngữ cũng có những đặc thù để tạo ra nghĩa biểu trưng. Mặt khác, tục ngữ bao giờ cũng rất ngắn gọn để dễ nhớ và để lưu truyền, đặc điểm này cũng ảnh hưởng tới tổ chức cú pháp của chúng*" [3, tr.9].

-Nổi bật hơn cả là trong tục ngữ Việt, yếu tố vần, nhịp là những yếu tố quan trọng tạo ra tính hoà đối, sinh động, có nhạc điệu. *Vần* tạo ra sự liên kết, đối ứng giữa các thành phần của câu tục ngữ, trong khi *nhịp* lại có chức năng cú pháp và chức năng biểu hiện nghĩa trong tục ngữ. Chẳng hạn, việc ngắt nhịp cho thấy hai kết cấu cú pháp sóng đôi như "Đẹp/ vàng son // ngon/ mật mỡ". Nhịp cũng giúp cho ta hiểu đúng nghĩa biểu hiện của một câu tục ngữ.

Với những điều trình bày ở trên, chúng tôi thấy trong quá trình giao lưu, tiếp xúc ngôn ngữ, việc chuyển ngữ tục ngữ Việt sang tiếng Hàn bằng các công cụ AI có thể gặp nhiều khó khăn do việc hiểu nghĩa của những tục ngữ có tính biểu trưng cao. Ví dụ với câu "Nạ dòng vớ được trai tơ", ChatGPT đã dịch sang tiếng Hàn là:

- 나이 든 여자가 젊은 총각을 얻다.

[*Người phụ nữ lớn tuổi lấy được một chàng trai trẻ***]**

Ở đây, ChatGPT đã xử lý ngữ nghĩa "nạ dòng" (người đàn bà đã có con và đứng tuổi) bằng cách dùng **"나이 든 여자" (người phụ nữ lớn tuổi) phần nào mất đi cái hàm ý của *nạ dòng* là "quá lứa, có con". Trong khi đó, "vớ (được)" có nghĩa "có được một cách không ngờ" được ChatGPT thể hiện bằng cách dùng từ 얻다 (lấy được), điều này cũng làm mất nét nghĩa "(có được một cách) không ngờ/ bất ngờ". Điều đáng lưu ý là khi cho câu 나이 든 여자가 젊은 총각을 얻다** (ChatGPT dịch từ câu "Nạ dòng vớ được trai tơ") cho Google translate (GT) dịch lại sang tiếng Việt, chúng tôi có kết quả: "***Người phụ nữ lớn tuổi có được một chàng trai độc thân trẻ tuổi*", lúc này GT đã thêm nghĩa "*có được*" và "젊은 총각" từ chỗ "*chàng trai trẻ tuổi*" đã chuyển thành "*chàng trai độc thân trẻ tuổi*". Như vậy, ChatGPT và GT đã có những hướng xử lý ngữ nghĩa khác nhau khi dịch tục ngữ Việt sang tiếng Hàn.**

Trong phần tiếp theo đây, chúng tôi sẽ nêu ra một số trường hợp mà AI xử lý về mặt ngữ nghĩa khi dịch tục ngữ Việt sang tiếng Hàn.

1. *AI* XỬ LÝ NGỮ NGHĨA KHI DỊCH TỤC NGỮ VIỆT SANG TIẾNG HÀN

1.1. Những tục ngữ có nghĩa biểu trưng

Những câu tục ngữ ở dạng này là những câu có nghĩa gián tiếp thông qua việc sử dụng hình ảnh biểu trưng, chúng thường nêu ra một nhận định, một triết lý sống.

Chúng ta lần lượt xem xét các câu sau:

(1) Già kén kẹn hom.

Nghĩa trực tiếp (cụ thể) của câu: khi để kén tằm quá lâu thì "kẹn" hom, tức là không róc. Hom là bộ phận của dụng cụ nuôi tằm, nơi con tằm nhả tơ để kết kén. Sử dụng hình ảnh này, kết hợp với lối chơi chữ theo kiểu đồng âm (*kén* và *kén* chọn) thì câu *Già kén kẹn hom* có nghĩa gián tiếp (nghĩa bóng): *người* ***quá kén chọn hoặc đặt ra tiêu chuẩn quá cao có thể dẫn đến khó khăn trong việc tìm được đối tượng phù hợp, thậm chí bỏ lỡ cơ hội*****,** giống như **kén tằm để quá già sẽ bị kẹt cứng vào hom, khó gỡ.**

ChatGPT đã dịch câu này sang tiếng Hàn:

⇨ 늦도록 까다롭게 고르다 보면 결국 나쁜 것을 얻게 된다

[*Kén chọn quá lâu thì cuối cùng lại nhận phải cái xấu*]

Như vậy, ChatGPT chỉ hiểu theo nghĩa gián tiếp khi dùng cấu trúc **-다가 보면 ~ -게 되다** (*nếu cứ... thì sẽ...*) nhấn mạnh quá trình dẫn đến kết quả tất yếu và cũng đã dùng từ **"늦도록"** (*càng để lâu, càng chậm trễ*) chứ không đối dịch "già" (trong *già* kén).

(2) Lành làm thúng, thủng làm mê.

Câu này có nghĩa: dù là vật còn nguyên vẹn hay đã hư hỏng thì cũng có thể tận dụng để làm việc khác, phù hợp với tình trạng, chức năng của nó. Nghĩa ngầm ẩn của câu: sử dụng người không ở việc này thì việc khác. Ban đầu ChatGPT không hiểu "mê" là gì, cho nên chúng tôi đã phải giải thích: mê là vật dụng đan bằng tre nứa đã bị thủng hoặc tuột cạp, tuột vành, dùng làm nắp đậy. ChatGPT đã cho ra kết quả theo kiểu đối dịch:

-온전하면 광주리를 만들고, 구멍이 나면 덮개를 만든다.

[*Lành lặn – cái thúng – làm, bị thủng, rách – nắp đậy – làm*]

Đáng nói là chúng tôi đã cho GT dịch lại câu tiếng Hàn trên (온전하면 광주리를 만들고, 구멍이 나면 덮개를 만든다) sang tiếng Việt thì nghĩa của câu đã khác hẳn: *Nếu còn nguyên vẹn thì đục một lỗ, nếu có nhiều lỗ thì đậy lại*.

(3) Con gái là cái bòn.

Câu trên có nghĩa: *Con gái khi đi lấy chồng, ở riêng thì hay xin xỏ, nhặt nhạnh của cải bên nhà bố mẹ đẻ để lo gia đình riêng của mình* [4, tr.440); *chê người con gái đi lấy chồng thường về bòn của của cha mẹ đem về nhà chồng* [5, tr.78].

ChatGPT đã cho ra 2 kết quả có nghĩa đối lập: tiêu cực và tích cực.

Kết quả tiêu cực:

(3a) 딸은 친정의 재산을 갉아먹는 존재다.

[*Con gái là người bào mòn tài sản của nhà mẹ đẻ*]

ChatGPT đã lựa chọn cụm từ "재산을 갉아먹다" (*bào mòn tài sản*) dịch cho "cái bòn", ngụ ý việc nuôi dưỡng con gái là không có lợi, là sự tổn thất kinh tế trong xã hội phong kiến. Đặc biệt, ChatGPT sử dụng danh từ "존재" (sự tồn tại) trong câu dịch khiến câu này thêm nghĩa *sự tồn tại của con gái là bào mòn tài sản của nhà mẹ đẻ*, nhấn mạnh quan niệm tiêu cực về vai trò của nữ giới trong xã hội cũ. Vấn đề đáng lưu ý ở đây là khi cho câu (3a) vào hệ thống dịch tự động GT chuyển ngữ sang tiếng Việt, ta được: ***Con gái là người ăn hết tài sản của cha mẹ mình*, lúc này cụm từ "갉아먹다"** (gặm mòn, bào mòn) đã bị GT đơn giản hóa thành "먹다" (ăn), từ đó chuyển thành cụm từ "ăn hết tài sản".

Tuy nhiên, ChatGPT cũng đã chuyển ngữ câu "Con gái là cái bòn" thêm một kết quả khác với nghĩa tích cực:

(3b) 딸은 부모에게 보물 같은 존재이다.

[*Con gái là một sự hiện diện quý giá như báu vật đối với cha mẹ*]

Như vậy khi đưa ra kết quả (3b), ChatGPT đã hiểu cụm từ "cái bòn" là "보물" (báu vật). Cách hiểu này không đúng với nội dung ngữ nghĩa của *Con gái là cái bòn.*

1.2. Những câu tục ngữ có hàm ý so sánh

Những tục ngữ Việt có hàm ý so sánh biểu hiện qua một số cấu trúc so sánh được AI xử lý chuyển ngữ qua tiếng Hàn như sau:

1.2.1. Cấu trúc "**A như B**"

Ví dụ:

(4) Gái có chồng như gông đeo cổ, trai có vợ như rợ buộc chân.

ChatGPT đã chuyển ngữ sang tiếng Hàn:

-시집간 여자는목에 칼을 쓰고, 장가든 남자는 발에 밧줄을 묶는다.

[*Người con gái đi đến chồng như đặt dao lên cổ, người con trai lập gia đình như buộc dây thừng vào chân*]

Và GT chuyển ngữ sang tiếng Việt: *Người phụ nữ đã có chồng sẽ bị một con dao quấn quanh cổ, còn người đàn ông đã có vợ sẽ bị một sợi dây thừng buộc quanh chân*.

Khi chuyển câu (4) sang tiếng Hàn, ChatGPT đã sử dụng cấu trúc [A 는 B 처럼 ~다] và xử lý ngữ nghĩa của một vài thành phần câu theo văn hóa Hàn. Chẳng hạn, ChatGPT đã sử dụng cụm từ "시집간 여자" (người con gái đã đi đến nhà chồng) để dịch cho **"gái có chồng"; sử dụng cụm từ "장가든 남자"** (người con trai đã lập gia đình) để dịch cho "trai có vợ". Hơn nữa, có 2 hình ảnh biểu trưng là "gông đeo cổ" và "rợ buộc chân" đã được ChatGPT xử lý theo lối tư duy của người Hàn. Cụm **"gông đeo cổ"** biểu thị sự ràng buộc mang tính cưỡng chế, bị động, một hình ảnh đặc trưng trong xã hội phong kiến Việt Nam lại được ChatGPT dịch thành "목에 칼을 쓰다" (đặt dao lên cổ). Về nghĩa, "đặt dao lên cổ" biểu hiện nghĩa "bị đe dọa, bị uy hiếp" mà theo cách nghĩ của người Hàn: phụ nữ có chồng rồi giống như dao kề cổ. Do vậy mặc

dù trong tiếng Hàn vẫn có từ "gông" (가, 枷) – một loại công cụ dùng để trói buộc người phạm pháp ở cổ, xuất hiện trong xã hội phong kiến Hàn Quốc, nhưng ChatGPT lại sử dụng **"dao"** (칼) cho nên cụm "칼을 쓰 ~" **có nghĩa "bị đặt dao lên cổ". Trong (4), cụm "rợ buộc chân" biểu hiện** hình ảnh bó buộc tuy nhẹ nhàng, nhưng dai dẳng, có tính kìm hãm nhưng đã được ChatGPT dịch thành "발에 밧줄을 묶는다" (buộc dây thừng vào chân). Ta thấy sự khác biệt giữa "dây thừng" (밧줄) là loại dây có cấu trúc dày, chắc, thường dùng để cố định vật nặng, trong khi từ **"rợ"** trong tiếng Việt chỉ loại dây mảnh, đơn hoặc chập vài sợi, do vậy "rợ buộc chân" có tính biểu trưng cao hơn là mô tả vật lý và có cảm giác "bó buộc, vướng bận" hơn "buộc dây thừng vào chân" (발에 밧줄을 묶는다).

Ví dụ khác:

(5) Gái lớn trong nhà như ma chưa cất.

Hàm ý của câu (5): Nhà có con gái lớn giống như nhà có người chết mà chưa làm đám [6, tr.376]; và nếu chậm làm đám thì để càng lâu thây ma sẽ bốc mùi. Kiểu so sánh này phản ánh lối suy nghĩ cũ, trọng nam khinh nữ.

ChatGPT đã dịch câu (5) sang tiếng Hàn như sau:

- 집안에 시집 안 간 큰딸은 안 치운 귀신 같다.

[*Trong nhà, cô con gái lớn chưa lấy chồng giống như một con ma chưa được cúng tiễn đi*]

ChatGPT đã dịch cụm "Gái lớn trong nhà" thành "집안에 시집 안 간 큰딸은" (Trong nhà, cô con gái lớn chưa lấy chồng). Và khi so sánh như "ma chưa cất"- một hình ảnh biểu trưng trong văn hóa dân gian Việt Nam, với ngụ ý rằng "ma chưa cất" là điều xui xẻo, là điều không tốt cho người còn sống. Trong bản dịch, ChatGPT lựa chọn cách diễn đạt "안 치운 귀신" (con ma chưa được cúng tiễn đi), trong đó có 치우다 là "thu dọn, loại bỏ" và 귀신 là "con ma" có nghĩa tương đương với "người chết chưa được làm đám, an táng" nhưng về sắc thái biểu cảm thì "ma chưa cất" vẫn hơn "con ma chưa được cúng tiễn đi". Hình ảnh "ma chưa cất" trong tục ngữ Việt có thể xem như một phép ẩn dụ mạnh, phản ánh sự tồn tại bị xem là "vướng bận", "không yên ổn" và "không thể được chấp nhận" trong xã hội cũ.

1.2.2. Cấu trúc "A **hơn** B"

Cấu trúc so sánh "A hơn B" có hàm ý nhấn mạnh giá trị của A cao hơn B, biểu hiện cách nhìn nhận giá trị của A trong mối quan hệ với B. Khi chuyển những tục ngữ Việt sang tiếng Hàn, ChatGPT đã sử dụng cấu trúc [A 가 B 보다 더 ~다]. Trong cách diễn đạt ý nghĩa so sánh, chúng tôi thấy ChatGPT, ChatAI xử lý như sau.

Ví dụ:

(6) Một giọt máu đào hơn ao nước lã.

ChatGPT: 붉은 피 한 방울이 맑은 물 한 연못보다 낫다.

[*Một giọt máu đỏ thì quý hơn cả ao nước trong*]

ChatAI 한 방울의 피가 맑은 물보다 더 소중하다.

[*Một giọt máu còn quý hơn nước trong*]

Cả 2 kết quả trên đều phản ánh gần đúng với nội dung ngữ nghĩa của (6). Tuy nhiên, hình như cả ChatGPT và Chat AI có ý nhấn mạnh giá trị của "Một giọt máu đào" (한 방울의 피) nên đã dùng cấu trúc ~보다 낫다 (*so với… thì tốt hơn/ quý hơn*), và mức so sánh ở đây đã cao, vượt qua cả "hơn" mà là "quý hơn/ tốt hơn".

Ví dụ khác:

(7) Cháu mười đời còn hơn người dưng.

Nghĩa ngầm ẩn của câu này giống như (6): đề cao mối quan hệ huyết thống. Với nghĩa trực tiếp (nghĩa đen) như vậy, ChatGPT, GT đã chuyển ngữ sang tiếng Hàn như sau:

ChatGPT: 십대손이라도 피 한 방울 통하는 게 남보다 낫다.

[*Dù là đời thứ mười đi nữa, chỉ cần có chung một giọt máu thì vẫn hơn người ngoài*]

GT: 손주들은 낯선 사람들보다 낫다.

[*Cháu tốt hơn những người xa lạ*]

Với 2 kết quả dịch như trên, ta thấy ChatGPT, GT đã không dịch theo cách đối dịch, mà dịch theo lối sao phỏng, nêu ý nghĩa chung nhất.

Việc chuyển ngữ tục ngữ Việt sang tiếng Hàn cũng có thể gặp một trở ngại khác khiến việc hiểu nghĩa của AI chưa chính xác, đó là việc ngắt nhịp trong câu tục ngữ Việt khiến AI "lúng túng" khi đưa ra kết quả dịch. Ví dụ, tục ngữ Việt có câu:

(8) Dâu hiền hơn con gái, rể hiền hơn con trai.

Câu (8) **có 2 cách hiểu:**

(8a) Dâu hiền/ hơn con gái// rể hiền/ hơn con trai.

(8b) Dâu/ hiền hơn con gái// rể/ hiền hơn con trai.

Ngữ nghĩa của cả 2 câu trên đều có hàm ý so sánh. Tuy nhiên ở (8a), nếu ngắt nhịp sau "hiền" thì ta có cách hiểu: (Con) dâu hiền/ (tốt) hơn con gái// rể hiền/ (tốt) hơn con trai, và "hiền" ở đây có nghĩa: *tốt, ăn ở phải đạo, hết lòng làm tròn bổn phận của mình đối với người khác* [7, tr.421] **do vậy mà dù cho con dâu, con rể tuy không có cùng huyết thống nhưng nhờ có "hiền" mà tốt hơn cả "con gái/ con trai ruột". Trong khi đó ở (8b), do ngắt nhịp sau "dâu" và "rể" (Con) dâu/ hiền hơn con gái// (con) rể/ hiền hơn con trai) khiến câu có hàm ý: con gái/ con trai dữ hơn con dâu/ con rể. "Hiền" trong (8b) lúc này có nghĩa: "*không dữ, thường không có những hành động, những tác động trực tiếp gây hại cho người khác, khi tiếp xúc người ta thấy dễ chịu, không có gì phải ngại, phải sợ*"** [7, tr.421].

Khi dùng ChatGPT chuyển ngữ câu *Dâu hiền hơn con gái, rể hiền hơn con trai* **sang tiếng Hàn. Chúng tôi nhận được kết quả:**

-며느리가 딸보다 낫고, 사위가 아들보다 낫다.

[*Con dâu còn tốt hơn con gái, con rể còn tốt hơn con trai*]

Như vậy, ChatGPT không nhận biết được nghĩa của từ "hiền", thành ra câu *Dâu hiền hơn con gái, rể hiền hơn con trai* được diễn đạt lại bằng cấu trúc so sánh "A 가 B 보다 낫다" [A tốt hơn B]: *Con dâu còn tốt hơn con gái, con rể còn tốt hơn con trai*.

Chúng tôi cũng đã thử nhờ ChatAI, GT chuyển ngữ sang tiếng Hàn câu (8), kết quả:

ChatAI: 며느리는 딸보다 더 온화하고, 사위는 아들보다 더 온화하다.

[*Con dâu thì hiền hòa hơn con gái, và con rể thì hiền hòa hơn con trai*]

GT: 며느리는 딸보다 상냥하고, 사위는 아들보다 상냥하다.

[*Con dâu tử tế hơn con gái, con rể tử tế hơn con trai*]

Rõ ràng "*hiền hòa*" (온화하다), "*tử tế*" (상냥하다) không có nghĩa như "hiền" trong (8a). Có thể cả ChatGPT, ChatAI và GT đều không hiểu hết được nghĩa của từ "hiền" trong cách hiểu (8a) dẫn đến các kết quả khác nhau, và các kết quả này không phản ánh hết được nghĩa ngầm ẩn của câu tục ngữ "*Dâu hiền hơn con gái, rể hiền hơn con trai*".

1.2.3. Cấu trúc A **bằng** B

Cấu trúc "A bằng B" trong tục ngữ biểu thị sự so sánh tương đương giữa A và B, nhấn mạnh giá trị đồng nhất, sự quan trọng của cả A và B.

Ví dụ:

(9) Một bát cơm cha bằng ba bát cơm rể.

Ngụ ý của câu: những gì của cha mẹ dành cho con cái có giá trị hơn nhiều so với những gì mà người ngoài mang lại; trong (9) cơm của cha cho ta dù chỉ một bát nhưng nó tương đương với ba bát cơm của người không cùng huyết thống (con rể) cho ta. ChatGPT đã dịch câu này sang tiếng Hàn gần như trọn vẹn ý nghĩa của câu nguồn, trong đó ChatGPT đã sử dụng cấu trúc ~과 같다 (bằng/ tương đương với):

-아버지의 밥 한 그릇은 사위의 밥 세 그릇과 같다.

[*Một bát cơm cha bằng với ba bát cơm rể*]

Trong khi đó, GT lại chuyển ngữ câu (9) theo dạng so sánh A **hơn** B với cấu trúc ~보다 낫다 (*so với… thì tốt hơn/ quý hơn*) chứ không dùng cấu trúc A 는 B 와 같다:

-아버지께 드리는 밥 한 그릇은 사위가 드리는 밥 세 그릇보다 가치가 있다.

[*Một bát cơm cho cha còn quý hơn ba bát cơm cho con rể*]

Như vậy, kết quả chuyển ngữ của ChatGPT và GT không giống nhau bởi "A **bằng** B" không có nghĩa như "A **hơn** B" và "một bát cơm **của** cha" khác với "một bát cơm **cho** cha".

Ví dụ khác:

(10) Một người nhà bằng ba người mừng.

"Người mừng" trong câu "Một người nhà bằng ba người mừng" chỉ những người không có quan hệ huyết thống (như bạn bè, hàng xóm),…Trong trường hợp này, "người mừng" có nghĩa "người khách đến chia vui" trong những dịp đặc biệt; khi ấy, người ta đánh giá rằng "đón được một người trong thân tộc quý bằng đón ba người khách". ChatGPT đã dịch câu (10) sang tiếng Hàn:

- 가족 한 사람이 손님 셋보다 낫다.

[*Một người thân còn quý hơn ba người khách*]

ChatGPT cũng sử dụng cấu trúc so sánh ~보다 낫다 (*so với… thì tốt hơn/ quý hơn*) nhấn mạnh: một người thân trong gia đình có giá trị hơn ba người mừng (người khách) khi xét về sự hỗ trợ, tình cảm và sự gắn bó lâu dài, và ChatGPT cũng không dùng cấu trúc "A 는 B 와 같다"

1.2.4. Cấu trúc A **không bằng** B

Đối lập với cấu trúc "A bằng B" là cấu trúc "A **không bằng** B". Cấu trúc này biểu thị ý nghĩa A không có giá trị tương đương với B; nó cũng có nghĩa A kém hơn B. Khi chuyển sang tiếng Hàn, ta có thể dùng cấu trúc A 가 B 보다 못하다. Ví dụ:

(11) Trăm con trai không bằng lỗ tai con gái.

Kết quả dịch của ChatGPT:

-백 명의 아들도 딸의 작은 관심 하나만 못하다.

[*Trăm người con trai cũng không bằng một chút quan tâm nhỏ của con gái*]

GT dịch: 아들 백 명이 있어도 딸 한 명의 귀에는 미치지 못한다.

[*Dù có trăm đứa con trai cũng không thể đến được tai một đứa con gái*]

Đối với kết quả dịch của ChatGPT, cụm **"lỗ tai con gái"** lại được biểu đạt bằng "딸의 작은 관심" (sự quan tâm nhỏ của con gái). Ở đây ChatGPT hiểu "lỗ tai con gái" ám chỉ việc lắng nghe, để ý hoặc thấu hiểu cha mẹ. Trong khi đó, GT đã diễn giải xa hơn nghĩa gốc, đó là "… không thể đến được tai một đứa con gái" (…딸 한 명의 귀에는 미치지 못한다).

Ví dụ khác:

(12) Ruộng sâu trâu nái không bằng con gái đầu lòng.

"Ruộng sâu, trâu nái" chỉ cảnh giàu có, sung túc ở nông thôn, "trâu nái" là trâu nuôi để đẻ, nhân giống. Câu (12) có hàm ý: có con gái đầu lòng là hơn tất cả, vì con biết chăm sóc, đỡ đần cho gia đình. ChatGPT đã chuyển ngữ câu (12) sang tiếng Hàn như sau:

- 비옥한 논과 새끼를 낳는 암소도 첫째 딸만은 못하다.

[*Dù có ruộng màu mỡ và trâu cái đẻ cũng không bằng con gái đầu lòng*]

ChatGPT đã sử dụng cấu trúc so sánh ~는 못하다 khi chuyển ngữ câu (12) sang tiếng Hàn để nhấn mạnh: **dù có tài sản vật chất quý giá như ruộng tốt, nhiều trâu bò,** thì cũng **không thể sánh bằng giá trị tinh thần và vai trò thiết thực** của người **con gái đầu lòng** trong gia đình**,** là người **gánh vác việc nhà, chăm lo, phụ giúp cha mẹ. Ở đây, ChatGPT đã dùng "비옥한 논"** (đồng ruộng màu mỡ) để dịch cho **"ruộng sâu"** và dùng **"새끼를 낳는 암소"** (con bò cái sinh con) để dịch cho **"trâu nái"**; việc thay thế hình ảnh **con trâu** (물소) bằng **con bò (암소)** phù hợp với văn hóa nông nghiệp lúa khô của Triều Tiên.

1.3. Những tục ngữ Hán Việt

Trong hệ thống từ vựng tiếng Việt, từ Hán Việt chiếm số lượng không nhỏ, ước khoảng 70-80%. Đấy là kết quả của quá trình giao lưu tiếp xúc ngôn ngữ Việt – Hán. Trong tiếng Hàn cũng có tình hình tương tự, từ Hán Hàn cũng chiếm số lượng không nhỏ.

Từ Hán Việt không chỉ xuất hiện trong đời sống, trong khắp các lĩnh vực xã hội, giao tiếp ngôn ngữ mà còn có trong các tổ hợp từ cố định như thành ngữ, tục ngữ. Ví dụ:

(13) Nhất nghệ tinh, nhất thân vinh.

Có những câu tục ngữ thuần Việt có nghĩa tương đương với (13) như: *Một nghề cho chín, hơn chín mười nghề, Một nghề thì kín, chín nghề thì hở, Một nghề thì sống, đống nghề thì chết*. Ngụ ý khuyên người ta nên chuyên về một nghề cho thành thạo, tinh thông.

ChatGPT đã cho ra 2 kết quả dịch cho câu (13):

(13a): 일예정 일신영

[*Nhất nghệ tinh, nhất thân vinh*]

(13b): 우물을 파도 한 우물을 파라

[*Khi đào giếng, hãy đào một giếng*]

ChatAI dịch câu (13) sang tiếng Hàn hơi khác một chút:

(13c) 한 가지 기술에 정통하면, 한 몸이 부유해진다.

[*Nếu thông thạo một kỹ năng, cơ thể sẽ trở nên giàu có*]

Kết quả (13a) là đối dịch Hán Việt với Hán Hàn mà gốc của câu (13) là 一藝精 一身榮.

Với kết quả (13b), ChatGPT đã chọn một câu tục ngữ Hàn có nghĩa tương đương với *Nhất nghệ tinh, nhất thân vinh* đó là "*Khi đào giếng, hãy đào một giếng"* ngụ ý khuyên người ta nên tập trung vào một việc, tránh một lúc làm nhiều việc mà kết quả không có việc nào tốt.

Kết quả dịch (13c) của ChatAI tuy gần với *Nhất nghệ tinh, nhất thân vinh* nhưng "thân vinh" (=vinh hiển cho bản thân) được diễn giải thành "cơ thể trở nên giàu có" (한 몸이 부유해진다) thì khó chấp nhận.

Tiếp tục xem xét câu:

(14) Nhất canh trì, nhì canh viên, ba canh điền.

Câu (14) nêu kinh nghiệm của nhà nông, muốn có thu hoạch tốt, muốn sinh lợi nhiều thì thứ nhất là thả cá ao, thứ đến là trồng vườn, thứ ba là làm ruộng [6, tr.590].

Trong trường hợp này, ChatGPT cũng đưa ra 2 kết quả dịch:

(14a) 일경지, 이경원, 삼경전 (一耕池 二耕園 三耕田)

(14b) 첫째는 연못을 가꾸고, 둘째는 과수원을 가꾸며, 셋째는 논밭을 일군다.

[*Cái thứ nhất là chăm sóc ao, cái thứ hai là chăm sóc vườn cây ăn quả, cái thứ ba là canh tác ruộng đồng*]

Câu (14a) là kết quả đối dịch Hán Việt sang Hán Hàn, trong đó, "ba" (tiếng Việt) trở thành 삼 (三) tiếng Hàn. Kết quả dịch (14b) rất sát nghĩa với *Nhất canh trì, nhì canh viên, ba canh điền.* Đối với người Hàn hiện nay, người ta khó nhận biết nội dung ngữ nghĩa của (14a) hơn là (14b), điều này cũng giống như người Việt, việc hiểu từ Hán Việt khó hơn là hiểu từ thuần Việt: *thứ nhất thả cá, thứ nhì làm vườn, thứ ba làm ruộng*.

Chúng tôi cũng đã thử nhờ ChatAI dịch sang tiếng Hàn câu (14), kết quả như (14c):

(14c) 첫 번째는 연못, 두 번째는 정원, 세 번째는 밭

[*Cái đầu tiên là cái ao, cái thứ hai là cái vườn, cái thứ ba là cái ruộng*]

Xử lý ngữ nghĩa như ChatAI đối với câu (14) là đã bỏ qua từ "canh" (경-耕) với nghĩa "làm/ canh tác".

Ví dụ khác:

(15) Nhất nam viết hữu, thập nữ viết vô.

Khác với trường hợp dịch sang tiếng Hàn của (13) và (14), ChatGPT đã đưa ra 3 kết quả dịch câu (15), trong đó (15c) như là câu dịch bổ sung để làm rõ nghĩa cho (15a):

(15a) 일남왈유, 십녀왈무 (一男曰有 十女曰無)

(15b) 아들 하나면 '있다' 하고, 딸 열이면 '없다' 한다.

[*Có một đứa con trai thì bảo là 'có', còn có mười đứa con gái thì vẫn bảo là 'không'*]

(15c) 아들 하나가 딸 열보다 낫다.

[*Một con trai còn hơn mười con gái*]

Câu (15) phản ánh sâu sắc **tư tưởng trọng nam khinh nữ** trong xã hội cũ, coi trọng con trai hơn con gái; nhà có nhiều con gái đến đâu cũng không được xem là "có".

ChatGPT đưa ra 3 kết quả: (15a) đối dịch Hán Việt với Hán Hàn, (15b) thuần Hàn, dịch nghĩa sát với nội dung nghĩa biểu hiện của (15). Với kết quả dịch (15c), ChatGPT đã sử dụng cấu trúc so sánh "~보다 낫다" (*so với… thì tốt hơn/ quý hơn*). Cách diễn đạt này làm rõ hơn ngữ nghĩa của câu (15a).

3. KẾT LUẬN

Nhìn chung, khi sử dụng *AI* để chuyển ngữ tục ngữ Việt sang tiếng Hàn, chúng ta nhận được kết quả dịch tương đối tốt. *AI*, đặc biệt là ChatGPT đã hiểu và dịch tốt những tục ngữ có hàm ý so sánh. *AI* có thể gặp trở ngại đôi chút đối với những câu tục ngữ có từ ngữ địa phương, từ cổ, do vậy để đảm bảo dịch đúng, ta cần cung cấp thông tin hoặc mô tả, giải thích các từ địa phương, từ cổ cho *AI* hiểu và dịch đúng. Một điều thú vị là đối với những câu tục ngữ dùng từ Hán Việt, ChatGPT đã đưa ra nhiều kết quả dịch khác nhau (Hán Hàn/ thuần Hàn) nhằm đảm bảo dịch sát nghĩa câu tục ngữ Việt.

TÀI LIỆU THAM KHẢO

[1] Bùi Văn Nguyên (chủ biên), Giáo trình lịch sử văn học Việt Nam, (Tập I- Văn học dân gian), Nxb Giáo dục, Hà Nội, 1961.

[2] Nguyễn Quý Thành, Cấu trúc cú pháp-ngữ nghĩa của tục ngữ Việt (trong sự so sánh với tục ngữ của một số dân tộc khác), Luận án tiến sĩ khoa học Ngữ văn, ĐH Quốc gia TP.HCM, 2001.

[3] Nguyễn Đức Dân, Vài nhận xét về đặc điểm cú pháp của tục ngữ, Tạp chí Ngôn ngữ 3/ 1989.

[4] Nguyễn Như Ý (chủ biên), Đại từ điển tiếng Việt, Nxb. Văn hóa-thông tin, 1998.

[5] Nguyễn Lân, Từ điển thành ngữ và tục ngữ Việt Nam, Nxb. Khoa học xã hội, 1997.

[6] Vũ Dung, Vũ Thúy Anh, Vũ Quang Hào, Từ điển thành ngữ và tục ngữ Việt Nam, Nxb.Văn hóa thông tin, 2000.

[7] Hoàng Phê (chủ biên), Từ điển tiếng Việt, Nxb.Đà Nẵng, 1997.

THÔNG TIN TÁC GIẢ

Trần Văn Tiếng: TS., đang công tác Trường Đại học Ngoại ngữ-Tin học TP.HCM, số 828 đường Sư Vạn Hạnh, phường Hòa Hưng, TP.HCM, email: tiengtv@huflit.edu.vn.

Nguyễn Thị Kim Ngọc: ThS., đang công tác Trường Đại học Ngoại ngữ-Tin học TP.HCM, số 828 đường Sư Vạn Hạnh, phường Hòa Hưng, TP.HCM, email: ngocntk@huflit.edu.vn.

ỨNG DỤNG HỘI HỌA AI TRONG GIẢNG DẠY PHÂN MÔN ĐỌC (Thực hiện qua phân tích khung cảnh thiên nhiên trong các tác phẩm)

Nguyễn Thị Hoài Thương[1], Phạm Phương Linh[2]

Tóm tắt: Việc ứng dụng hội họa AI trong giảng dạy Ngữ văn mở ra một hướng tiếp cận mới, vừa trực quan vừa gắn liền với đổi mới giáo dục. Bằng cách kết hợp ngôn ngữ văn học với hình ảnh minh họa do AI tạo ra, học sinh có cơ hội quan sát, tưởng tượng và phân tích tác phẩm từ nhiều góc độ. Phương pháp này không chỉ giúp khắc phục tình trạng tiếp thu thụ động mà còn khơi gợi hứng thú, phát triển khả năng sáng tạo và năng lực cảm thụ thẩm mỹ. Kết quả thực nghiệm cho thấy, học sinh tham gia tích cực hơn, tự tin thể hiện cảm nhận nghệ thuật và nâng cao kỹ năng phân tích văn bản. Qua đó, hội họa AI chứng minh tính khả thi và hiệu quả trong việc làm phong phú hoạt động dạy – học Ngữ văn, hướng tới hình thành các năng lực thiết yếu của thế kỷ XXI.

Từ khóa: giảng dạy, ngữ văn, phân môn đọc, hội họa, công cụ AI, sáng tạo, cảm thụ.

1. Mở đầu

1.1. Đặt vấn đề

Môn Ngữ văn từ lâu là một trong những môn học nền tảng của giáo dục Việt Nam, tuy nhiên, trong thời đại công nghệ số, môn học này đang đối mặt với nhiều thách thức. Sự phát triển mạnh mẽ và đa dạng của các nền tảng mạng xã hội, các video ngắn và các hình thức giải trí trực tuyến khiến học sinh ít dành thời gian cho việc đọc và nghiên cứu văn học truyền thống.

Ngữ văn giữ vai trò quan trọng trong việc rèn luyện tư duy phản biện, khả năng phân tích, viết lách và cảm thụ nghệ thuật. Dẫu vậy, không ít học sinh cảm thấy môn học khô khan, đặc biệt là khi phải phân tích những tác phẩm giàu hình ảnh hoặc miêu tả thiên nhiên.

Trong văn chương, màu sắc không chỉ mang tính miêu tả mà còn là ngôn ngữ nghệ thuật, giúp truyền tải cảm xúc, tạo bầu không khí và làm nổi bật ý nghĩa tác phẩm. Mỗi màu sắc mang giá trị biểu cảm riêng: xanh gợi bình yên, hy vọng; đỏ gợi sự mãnh liệt, cuồng nhiệt hoặc nguy hiểm. Việc nắm bắt được ý nghĩa sắc màu sẽ giúp học sinh tiếp cận sâu hơn với tư tưởng, tình cảm của tác giả.

Để đạt được điều đó, cần có những phương pháp dạy học sinh động, sáng tạo. Việc kết hợp hội họa – một loại hình nghệ thuật trực quan – với giảng dạy Ngữ văn là giải pháp hiệu quả, giúp học sinh vừa "thấy" vừa "cảm" được vẻ đẹp nghệ thuật trong từng trang văn.

1.2. Đối tượng nghiên cứu

- Đối tượng nghiên cứu: học sinh THPT.

[1] Giáo viên trường THPT FPT Đà Nẵng, ***Email:*** *hthuong1702@gmail.com.*
[2] Phạm Phương Linh, Học viên cao học khóa 46 trường ĐHSP – ĐHĐN.

- Phạm vi nghiên cứu: các tác phẩm văn học có yếu tố hình ảnh nổi bật, ví dụ như thơ ca, các tác phẩm văn xuôi có mô tả thiên nhiên, cảnh vật… (Có thể lấy các tác phẩm như "Bảo kính cảnh giới" của Nguyễn Trãi, "Việt Bắc" của Tố Hữu, "Hai đứa trẻ" của Thạch Lam…)

1.3. Phương pháp nghiên cứu

Để triển khai đề tài, một số phương pháp nghiên cứu được vận dụng, cụ thể:

- **Phương pháp nghiên cứu lý luận:** Tìm hiểu, phân tích các tài liệu về phương pháp dạy học Ngữ văn, tài liệu liên quan đến ngôn ngữ màu sắc trong văn chương, cũng như các công trình nghiên cứu tích hợp nghệ thuật trong giảng dạy.

- **Phương pháp phân tích – tổng hợp:** Phân tích các tác phẩm văn học tiêu biểu có yếu tố hình ảnh, sắc màu; tổng hợp ý nghĩa nghệ thuật để rút ra nhận xét khái quát.

- **Phương pháp quan sát – thực nghiệm sư phạm:** Tiến hành áp dụng hình thức dạy học kết hợp hội họa vào các giờ Ngữ văn, theo dõi phản ứng và mức độ hứng thú của học sinh.

- **Phương pháp điều tra – khảo sát:** Thực hiện phiếu hỏi, trao đổi trực tiếp để nắm bắt thái độ, cảm nhận, mức độ tiếp thu và khả năng sáng tạo của học sinh khi áp dụng phương pháp mới.

- **Phương pháp so sánh – đối chiếu:** Đối chiếu hiệu quả học tập giữa lớp thực nghiệm (áp dụng phương pháp kết hợp hội họa) và lớp đối chứng (dạy theo cách truyền thống).

2. Nội dung

2.1. Cơ sở lí thuyết

2.1.1. Khái quát về ứng dụng hội họa trong giảng dạy văn học

Văn học và nghệ thuật hội họa, mặc dù có hình thức thể hiện khác nhau (văn học là nghệ thuật ngôn từ, còn hội họa là nghệ thuật thị giác), nhưng lại có một mối liên hệ mật thiết và bổ sung cho nhau trong việc truyền tải cảm xúc cũng như tư tưởng nghệ thuật. Điểm gặp gỡ chung của chúng là khả năng tạo ra những thế giới tưởng tượng phong phú, nơi con người có thể khám phá những trải nghiệm mới mẻ và cảm nhận những giá trị sâu sắc của cuộc sống.

Trong văn học, ngôn từ được sử dụng để xây dựng hình ảnh giàu sức gợi, đặc biệt qua việc miêu tả thiên nhiên, con người và những cảnh huống đời sống. Những hình ảnh ấy thường gắn với màu sắc, âm thanh, không gian, mang đến cho người đọc một trường liên tưởng đa dạng. Tuy nhiên, để học sinh tiếp nhận được trọn vẹn vẻ đẹp nghệ thuật này không phải lúc nào cũng dễ dàng.

Lúc này, việc ứng dụng hội họa vào dạy học văn học mở ra một hướng tiếp cận trực quan, sinh động. Khi học sinh quan sát tranh minh họa hoặc trực tiếp vẽ tranh tái hiện hình ảnh trong tác phẩm, các em không chỉ dễ dàng hình dung bối cảnh, màu sắc, không khí tác phẩm mà còn có cơ hội bộc lộ cảm nhận riêng. Tranh minh họa vì thế không chỉ tái hiện văn bản mà còn mở rộng ý nghĩa, làm phong phú thêm thế giới nghệ thuật mà tác giả xây dựng, đồng thời giúp học sinh phát triển khả năng cảm thụ và sáng tạo.

2.1.2. Vai trò của màu sắc trong văn học và hội họa

Màu sắc là một trong những yếu tố quan trọng bậc nhất của nghệ thuật thị giác, đồng thời cũng là phương tiện biểu đạt giàu giá trị trong văn chương. Nếu trong hội họa, màu sắc góp phần tạo nên bố cục, chiều sâu và cảm xúc thị giác cho người xem, thì trong văn học, màu sắc được tái hiện bằng ngôn từ, gợi mở những hình dung, cảm nhận và trạng thái tâm lí nơi độc giả.

Trong văn học, màu sắc không đơn thuần là chi tiết miêu tả, mà còn là "ngôn ngữ nghệ thuật" giàu sức gợi. Ví dụ, sắc xanh thường gắn với thiên nhiên, bình yên, hy vọng; sắc đỏ gợi sự mãnh liệt, đam mê hoặc nguy hiểm; sắc trắng gợi thuần khiết, trong trẻo nhưng đôi khi cũng gợi nỗi buồn, sự mất mát. Khi phân tích tác phẩm văn học, việc nắm bắt ý nghĩa biểu tượng của màu sắc sẽ giúp học sinh hiểu sâu hơn về tâm tư, tư tưởng tác giả cũng như thông điệp mà tác phẩm chuyển tải.

Trong hội họa, màu sắc không chỉ phản ánh hiện thực mà còn mang tính biểu cảm, giúp nghệ sĩ khắc họa thế giới nội tâm, truyền tải bầu không khí và trạng thái cảm xúc. Sự kết hợp hài hòa hoặc tương phản màu sắc có thể làm nổi bật thông điệp nghệ thuật, khiến bức tranh không chỉ có giá trị thẩm mỹ mà còn gợi suy tư cho người thưởng thức.

Sự giao thoa giữa hai loại hình nghệ thuật này tạo nên khả năng hỗ trợ hiệu quả trong giảng dạy văn học. Khi học sinh được quan sát hoặc sáng tạo tranh màu sắc để minh họa cho văn bản, các em sẽ dễ dàng kết nối cảm xúc giữa hình ảnh thị giác và hình tượng ngôn từ, từ đó tiếp nhận tác phẩm văn học một cách sinh động, trực quan và sâu sắc hơn.

2.2. Thực trạng dạy – học văn học với yếu tố hình ảnh và màu sắc trong nhà trường

Trong chương trình Ngữ văn hiện nay, nhiều tác phẩm văn học chứa đựng yếu tố hình ảnh và màu sắc giàu giá trị nghệ thuật cao, tiêu biểu như thơ ca lãng mạn, thơ kháng chiến, hay truyện ngắn hiện thực giàu chất miêu tả. Tuy nhiên, việc giảng dạy và học tập các tác phẩm này vẫn còn một số hạn chế:

a) Phía giáo viên:

- Phần lớn giáo viên chủ yếu chú trọng truyền đạt kiến thức theo hướng phân tích nội dung – nghệ thuật dựa trên văn bản, chưa thật sự khai thác chiều sâu giá trị thẩm mỹ của màu sắc và hình ảnh.

- Việc ứng dụng các phương pháp trực quan, sáng tạo như minh họa bằng tranh, hình ảnh hoặc tích hợp hội họa vào dạy học Ngữ văn chưa được triển khai phổ biến, chủ yếu dừng ở mức minh họa sơ lược trên bài giảng điện tử.

- Áp lực về thời gian và yêu cầu "dạy đúng, học đúng" chương trình đôi khi khiến giáo viên e ngại đổi mới phương pháp.

b) Phía học sinh:

- Nhiều học sinh cảm thấy khó hình dung và liên tưởng khi tiếp xúc với những đoạn văn, bài thơ giàu hình ảnh, màu sắc. Điều này khiến các em dễ rơi vào trạng thái học thụ động, tiếp thu theo "khuôn mẫu" phân tích có sẵn từ giáo viên.

- Học sinh ít có cơ hội thể hiện khả năng sáng tạo cá nhân, trong khi hội họa có thể là một kênh giúp các em trực quan hóa cảm xúc và suy nghĩ của mình về tác phẩm văn học.

- Thói quen tiếp cận văn học qua "ghi nhớ nội dung" thay vì "cảm thụ thẩm mỹ" dẫn đến sự thiếu hứng thú với môn học.

c) Phương tiện hỗ trợ giảng dạy:

- Nhà trường có trang bị một số công cụ hỗ trợ trực quan (máy chiếu, tranh ảnh, tư liệu số), song chưa khai thác triệt để vào việc giảng dạy Ngữ văn.

- Hoạt động kết hợp liên môn giữa Văn học – Mỹ thuật còn hiếm, chưa trở thành định hướng phổ biến trong đổi mới giáo dục. Đặc biệt, với học sinh chọn khối khoa học tự nhiên, việc học Ngữ văn thường chỉ dừng ở mức "đủ điểm", thiếu hứng thú và ít quan tâm đến khía cạnh thẩm mỹ. Do đó, càng cần có phương pháp trực quan, sinh động để khơi gợi cảm xúc và sự yêu thích môn học.

Từ thực trạng trên có thể khẳng định: việc ứng dụng hội họa vào giảng dạy Ngữ văn, đặc biệt trong hoạt động phân tích các yếu tố hình ảnh của tác phẩm, là một hướng đi cần thiết. Cách tiếp cận này không chỉ giúp khắc phục tình trạng học sinh tiếp thu thụ động mà còn khơi dậy hứng thú học tập và phát huy năng lực sáng tạo cá nhân.

2.3. Giải pháp: Ứng dụng hội họa AI vào hoạt động phân tích các yếu tố hình ảnh trong tác phẩm văn học

2.3.1. Trong bối cảnh đổi mới giáo dục và sự bùng nổ của công nghệ số, yêu cầu đặt ra không chỉ là giúp học sinh tiếp nhận kiến thức mà còn phải khơi gợi năng lực cảm thụ cái hay, cái đẹp trong văn chương. Hội họa – đặc biệt là hội họa ứng dụng trí tuệ nhân tạo (AI Art) – mở ra một hướng tiếp cận mới, vừa trực quan, vừa gắn với trải nghiệm sáng tạo công nghệ. Ứng dụng hội họa để phân tích cái hay, cái đẹp đó có thể được ứng dụng trong các hoạt động sau:

- Sử dụng tranh vẽ, hình ảnh minh họa để học sinh có thể hình dung được các hình ảnh trong tác phẩm văn học (như cảnh vật, nhân vật, không gian…).

- Học sinh tự cảm nhận và vẽ về những màu sắc, sự đối lập sáng tối, gam màu nóng lạnh trong tác phẩm.

- Học sinh tự cảm nhận, tự diễn đạt thành câu lệnh để nhờ AI tạo nên hình ảnh ưng ý.

- Các bài học có thể kết hợp giữa văn học và hội họa: Giảng bài theo phương pháp phòng tranh giúp học sinh phát triển khả năng phân tích đa chiều.

2.3.2. Một số hình thức triển khai cụ thể

2.3.2.1. Sử dụng tranh minh họa trực quan để khởi động và định hướng cảm nhận

- Giáo viên chuẩn bị sẵn tranh vẽ, hình ảnh (tranh truyền thống hoặc tranh tạo bởi AI) để học sinh quan sát.

- Ví dụ: Khi dạy tác phẩm "Hai đứa trẻ" (Thạch Lam), giáo viên có thể dùng hình ảnh phố huyện lúc chiều tà, với gam màu vàng nhạt, tím sẫm để gợi không gian hoài niệm, tĩnh lặng.

- Học sinh dựa trên hình ảnh quan sát để đối chiếu với ngôn từ văn học, từ đó nhận diện gam màu, ánh sáng, không gian mà tác giả miêu tả.

2.3.2.2. Học sinh tự vẽ hoặc chọn gam màu phù hợp với văn bản

- Giáo viên giao nhiệm vụ: học sinh phác họa nhanh (không yêu cầu cầu kỳ) các gam màu, sự đối lập sáng – tối, nóng – lạnh trong tác phẩm.

- Ví dụ: Với "Việt Bắc" (Tố Hữu), học sinh có thể vẽ mảng xanh trữ tình (núi rừng), gam đỏ rực (cách mạng, lửa kháng chiến), từ đó phân tích biểu tượng ý nghĩa của từng màu.

2.3.2.3. Ứng dụng AI trong tái hiện hình ảnh văn học

- Giáo viên hướng dẫn học sinh chuyển cảm nhận thành câu lệnh (prompt) để AI tạo hình ảnh minh họa.

- Ví dụ: Khi học đoạn "Phố huyện chiều tàn" trong "Hai đứa trẻ", học sinh có thể nhập câu lệnh: *"Vẽ cảnh một phố huyện nhỏ ở Việt Nam đầu thế kỷ XX, ánh chiều vàng nhạt phủ khắp không gian, vài bóng người thưa thớt trên con đường đất, ánh đèn dầu le lói từ gánh hàng rong. Tranh vẽ theo chiều ngang, tỉ lệ 16:9, phong cách tranh sơn dầu mềm mại, màu sắc trầm ấm, gợi không khí hoài niệm."*

- Học sinh so sánh bức tranh AI tạo ra với văn bản gốc, từ đó phân tích xem AI đã tái hiện đúng/khác ở chi tiết nào → rèn năng lực phân tích, đối chiếu, phản biện.

2.3.2.4. Phương pháp "phòng tranh" kết hợp công nghệ AI

- Giáo viên chia lớp thành nhóm, mỗi nhóm chọn một đoạn thơ/văn, sau đó:

+ Vẽ tay hoặc dùng AI để tạo hình minh họa.

+ Hình thức triển lãm: Có thể làm "phòng tranh trên giấy A0" hoặc **phòng tranh trực tuyến** (Canva, Padlet, Google Slides).

+ Thuyết trình: phân tích vì sao chọn gam màu/không gian/ánh sáng đó để diễn tả ý đồ nghệ thuật của tác giả.

- Hoạt động này giúp học sinh vừa rèn kỹ năng phân tích văn học, vừa phát triển khả năng thẩm mỹ và tư duy sáng tạo.

2.3.3. Ý nghĩa của giải pháp

- Giúp học sinh tiếp nhận hình tượng văn học không chỉ bằng ngôn ngữ mà còn bằng thị giác, từ đó phân tích sâu sắc hơn.

- Rèn luyện năng lực chuyển đổi từ cảm thụ → ngôn ngữ diễn đạt → hình ảnh (với sự hỗ trợ của AI).

- Tạo hứng thú cho học sinh, đặc biệt là học sinh khối tự nhiên, khi các em được tham gia trải nghiệm công nghệ thay vì chỉ ghi nhớ khô khan.

- Góp phần đổi mới phương pháp dạy học Ngữ văn, đưa công nghệ vào lớp học nhưng vẫn giữ được tính đặc thù của môn học.

2.4. Cách thực hiện giải pháp: Ứng dụng hội họa AI vào hoạt động phân tích các yếu tố hình ảnh trong tác phẩm văn học

2.4.1. Các công cụ AI hỗ trợ hoạt động vẽ tranh

Công cụ	Chức năng chính	Ưu điểm	Nhược điểm	Mức độ phù hợp trong lớp học
Canva AI	- Tạo hình ảnh từ văn bản (Text to Image). - Thiết kế poster, infographic, bài trình chiếu. - Tổ chức triển lãm trực tuyến.	- Giao diện dễ dùng, thân thiện HS – GV. - Kết hợp nhiều công cụ thiết kế trong một nền tảng. - Phù hợp cho hoạt động nhóm, sản phẩm học tập.	- Chất lượng tranh ở mức minh họa, chưa tinh xảo. - Ít tùy chỉnh phong cách hội họa. - Bản miễn phí bị giới hạn.	Cao – GV, HS có thể dùng ngay, dễ tích hợp vào bài học.
ChatGPT tích hợp DALL·E	- Sinh tranh từ văn bản (Text to Image). - Chỉnh sửa tranh (Inpainting). - Hỗ trợ viết prompt sáng tạo bằng tiếng Việt. - Tạo nhiều phiên bản tranh để lựa chọn.	- Tạo tranh sát nội dung văn học. - Kết hợp sức mạnh ngôn ngữ của ChatGPT để viết prompt hay. - Giúp HS rèn kỹ năng mô tả ngôn ngữ – hình ảnh.	- Có thể sai chi tiết trong tranh. - Chưa tối ưu văn hóa Việt Nam (cảnh, trang phục…). - Mức độ chính xác trung bình "ưa nịnh", dễ bị sai nếu mô tả không rõ ràng. - Cần Internet ổn định và tài khoản.	Rất cao – phù hợp cho HS thực hành cá nhân hoặc nhóm trong giờ học.
Gemini (Google AI)	- Sinh tranh từ văn bản (Text to Image).	- Tạo tranh sát nội dung văn học.	- Có thể sai chi tiết trong tranh.	Rất cao – phù hợp cho HS thực hành

	- Hỗ trợ viết prompt sáng tạo bằng tiếng Việt.	- Kết hợp sức mạnh ngôn ngữ của Gemini để viết prompt hay. - Giúp HS rèn kỹ năng mô tả ngôn ngữ – hình ảnh.	- Văn phong dài dòng hơn. - Cần Internet ổn định và tài khoản.	cá nhân hoặc nhóm trong giờ học.
Sora (OpenAI)	- Sinh video từ văn bản (Text to Video). - Biến mô tả cảnh văn học thành video động. - Có thể kết hợp âm thanh, hình ảnh, hiệu ứng.	- Tạo video minh họa sống động, hấp dẫn. - Phù hợp với tác phẩm giàu hình ảnh, cảm xúc. - Mở rộng trải nghiệm học văn qua thị giác + thính giác.	- Công cụ mới, hạn chế người dùng. - Yêu cầu cấu hình mạnh, đường truyền ổn định. - Chưa phù hợp cho phổ thông đại trà. - Có thể làm video sai chi tiết. - Dễ sai chính tả nên sử dụng các từ ngữ Tiếng Việt tại tranh/video.	Trung bình – dùng tốt cho thí điểm, demo, hoặc hoạt động ngoại khóa.

2.4.2. Quy trình khai thác AI trong một bài học

*a) **Về thời lượng:*** Giáo viên nên lựa chọn **tiết học tự chọn, tiết sinh hoạt lớp, hoặc hoạt động hướng nghiệp – trải nghiệm sáng tạo** để hướng dẫn học sinh sử dụng công cụ AI. Cách bố trí này vừa giúp giảm áp lực thời lượng trong chương trình chính khóa, vừa tạo ra một không gian linh hoạt để học sinh có thể thoải mái thử nghiệm và sáng tạo. Đồng thời, đây cũng là cơ hội để thực hiện tích hợp liên môn giữa Ngữ văn, Mỹ thuật, Tin học, Công nghệ và cả những trải nghiệm gắn với đời sống, qua đó giúp học sinh thấy văn học không tách rời mà gắn liền mật thiết với thực tiễn và công nghệ hiện đại. Chính sự kết nối này không chỉ làm tăng hứng thú học tập mà còn góp phần rèn luyện cho học sinh những năng lực thiết yếu trong bối cảnh giáo dục 4.0 như: năng lực thẩm mỹ, khả năng sáng tạo, tư duy phản biện và đặc biệt là năng lực số.

*b) **Quy trình hướng dẫn học sinh khai thác AI***

Hoạt động	Hướng dẫn của GV	Học sinh thực hiện	Công cụ AI/Đồ dùng	Mục tiêu	Minh hoạ cách làm
HĐ1: Định hướng (5')	- Giới thiệu đoạn văn/thơ cần phân tích. - Gợi câu hỏi định hướng: "Trong đoạn thơ này, tác giả đã miêu tả những hình ảnh thiên nhiên nào?"	- Lắng nghe, đọc văn bản. - Ghi lại, tưởng tượng về từ khóa về **màu sắc, âm thanh, hình ảnh** có thể xuất hiện trong văn bản.	SGK, bảng/slide	Xác định từ khóa nghệ thuật quan trọng.	x
HĐ2: Tạo prompt (10')	**Bước 1: Hướng dẫn cách viết prompt cơ bản.** **- Cấu trúc prompt cơ bản:** Hãy vẽ + [Đối tượng chính] + [Hành động/trạng thái] + [Không gian/thời gian] + [Chi tiết bổ sung: màu sắc, âm thanh, ánh sáng, phong cách hội - Lí giải: 1. [Đối tượng chính]: Cây hoè, hoa lựu, hồ sen, bầu trời, con người… 2. [Hành động/trạng thái]: đùn đùn tán lá, phun nở, tỏa hương, lao xao, dắng dỏi… 3. [Không gian/thời gian]: buổi chiều hè,	- Viết prompt mô tả cảnh dựa trên ngữ liệu -Nhờ Gemini/ ChatGPT hỗ trợ chỉnh sửa cho chi tiết, nghệ thuật hơn.	- Gemini (Google AI) - ChatGPT	HS biết cách chuyển từ ngôn từ văn học → ngôn ngữ mô tả cho AI.	VÍ DỤ: **1. HS viết prompt cơ bản:** *Hãy vẽ cây hoè xanh đùn đùn tán lá, hoa lựu đỏ rực phun nở, hồ sen hồng nở rộ tỏa hương, trong khung cảnh buổi chiều hè ở làng quê Việt Nam những năm thế kỉ XX, với ánh nắng vàng óng, gam màu tươi sáng, theo phong cách tranh thủy mặc. Kích thước ảnh 1920x1080. Độ phân giải 4K. [1]* **2. Nhập prompt này vào Gemini/ChatGPT với yêu cầu:** → *"Hãy đọc bài thơ Bảo kính cảnh giới (bài 43) của Nguyễn Trãi tại sách giáo khoa Ngữ văn 10 – tập 2 – bộ sách kết nối tri thức. Vận dụng khả năng cảm thụ văn*

làng quê Việt Nam, ánh nắng ban mai…

4. [Chi tiết bổ sung]: màu sắc, ánh sáng, gam màu (nóng – lạnh), phong cách tranh (tranh thủy mặc, sơn dầu, tranh kỹ thuật số…).

5.[Kích thước ảnh]: 1024x1024 px (vuông), 1920x1080 px (màn hình), 2048x2048 px (in ấn)…

6. [Độ phân giải]: Cao (High resolution), 4K, Ultra HD…

Bước 2: Gợi ý HS dùng Gemini hoặc ChatGPT để mở rộng mô tả. ***(có thể lược).***

học và hội họa đã được học ở những lớp tiểu học và trung học cơ sở để mở rộng mô tả này thành một đoạn chi tiết, giàu hình ảnh, phong cách văn chương để tôi dùng cho AI vẽ tranh. Cấu trúc prompt [1]

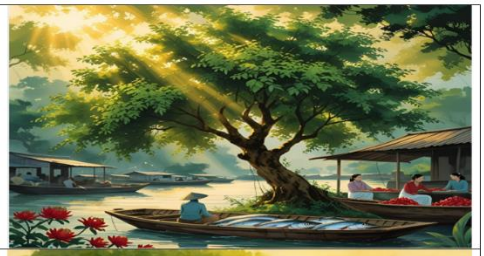
ChatGPT/Gemini sẽ gợi ý:

Prompt AI mở rộng:

"Hãy vẽ cây hoè xanh mướt, lá xum xuê đùn đùn rợp bóng mát, hoa lựu đỏ rực phun nở, từng chùm hoa tươi thắm nổi bật trên nền lá xanh, hồ sen hồng nở rộ tỏa hương nhẹ nhàng, những cánh sen mỏng manh lay động trong gió chiều. Phía xa làng quê Việt Nam những năm thế kỉ XX với những mái nhà tranh thấp thoáng, chợ cá rộn rã tiếng lao xao, vài chiếc thuyền nhỏ trôi lững lờ trên kênh. Ánh nắng vàng óng chiếu qua tán lá, phản chiếu lên mặt nước, tạo cảm giác ấm áp, yên bình, gam màu tươi sáng, theo phong cách tranh thủy mặc, chú trọng sự hài hòa giữa thiên nhiên và con người. Kích thước

				ảnh 1920x1080. Độ phân giải 4K.	
HĐ3: Tạo hình ảnh (10')	- Hướng dẫn HS nhập prompt vào công cụ vẽ tranh AI. - Khích lệ HS thử nhiều phong cách (tranh sơn dầu, thủy mặc, tranh số).	- Học sinh sử dụng công cụ yêu thích. - So sánh các phiên bản - Chọn sản phẩm phù hợp.	- Canva AI - ChatGPT tích hợp DALL·E - Gemini (Google AI) - Sora (OpenAI)	Tái hiện bức tranh văn học thành tranh trực quan.	X
HĐ4: Trình bày – thảo luận (10')	- Tổ chức triển lãm (tranh in giấy hoặc trực tuyến trên Canva/Padlet/Google Slide). - Đặt câu hỏi thảo luận: "Tranh nào gần gũi nhất với ý thơ? Vì sao có sự khác biệt giữa các tranh?"	- Trưng bày sản phẩm. - Thuyết minh về tranh của nhóm. - Thảo luận, phản biện giữa các nhóm.	Canva, Padlet, máy chiếu	Phát triển tư duy so sánh, khả năng thuyết trình và phản biện.	X
HĐ5: Đánh giá sản phẩm **(**GV đánh giá, HS tự đánh giá và đánh giá chéo.)	-Dùng Rubric/ Bảng kiểm để đánh giá sản phẩm theo các tiêu chí cụ thể. - Nhận xét chung, khích lệ.	- Tự đánh giá nhóm mình. - Góp ý cho nhóm khác.	Rubric (đã thiết kế), Bảng kiểm	Đảm bảo HS vừa được **đánh giá công bằng**, vừa **tự đánh giá**.	

=> Sản phẩm mẫu minh họa được tạo ra từ prompt:

Canva AI	
ChatGPT tích hợp DALL·E	

Gemini (Google AI)	
Sora(OpenAI)	

2.4.3. Đạo đức số và tính minh bạch khi sử dụng AI

Đối với những học sinh chưa thành thạo kỹ năng vẽ, AI có thể là công cụ hữu ích để trình bày ý tưởng. Tuy nhiên, sản phẩm cần kèm theo chú thích rõ ràng nhằm đảm bảo tính minh bạch và công bằng trong học tập, chẳng hạn:

- *"Hình ảnh được vẽ bởi AI với sự hỗ trợ ý tưởng của học sinh"*
- *"Minh họa do AI tạo ra, không phải sản phẩm vẽ tay hay thiết kế cá nhân".*
- *"Hình ảnh được tạo bởi AI".*

Bên cạnh đó, học sinh được khuyến khích thử nghiệm nhiều phong cách nghệ thuật (truyền thống – hiện đại – trừu tượng) để mở rộng tư duy thẩm mỹ, đồng thời nâng cao nhận thức về giá trị của sự sáng tạo cá nhân.

Quan trọng hơn, đạo đức số cần được coi là **chuẩn mực bắt buộc** khi ứng dụng AI trong lớp học. Giáo viên cần định hướng và hình thành thói quen chú thích ở sản phẩm không chỉ giúp ngăn chặn vi phạm bản quyền mà còn rèn luyện cho học sinh tinh thần trách nhiệm, sự minh bạch trong học tập và thái độ tôn trọng giá trị sáng tạo.

Như vậy, khi được định hướng đúng, AI sẽ thực sự trở thành công cụ hỗ trợ học tập hữu ích, thay vì là phương tiện sao chép hoặc thay thế sáng tạo cá nhân.

2.5. Áp dụng vào một tiết dạy cụ thể

2.5.1. Tên bài: Bảo kính cảnh giới (bài số 43) – Nguyễn Trãi *(sgk Ngữ văn 10 (tập 2) Bộ sách Kết nối tri thức với cuộc sống)*

2.5.2. Phạm vi ứng dụng: Dưới đây, chúng tôi ứng dụng hoạt động hội họa vào phần phân tích khung cảnh thiên nhiên ngày hè trong tiết dạy Đọc – Hiểu văn bản *Bảo kính cảnh giới* tại hoạt động 2 mang tên: TRIỂN LÃM "TÁI HIỆN THỜI ĐẠI"

2.5.3. Mục tiêu

- **Học sinh nêu** cuộc sống và tâm trạng của nhân vật trữ tình qua câu thơ đầu bài thơ.

- Học sinh phân tích một số từ ngữ, hình ảnh được tác giả sử dụng để miêu tả cảnh sắc mùa hè, **chỉ ra** những nét đặc sắc trong cách cảm nhận thiên nhiên và bút pháp tả cảnh của tác giả.

- **Học sinh xác định** những chi tiết miêu tả cuộc sống của nhân vật trữ tình qua âm thanh, hình ảnh, **phân tích** mối liên hệ giữa khung cảnh ấy với ước nguyện của nhân vật trữ tình.

- **Học sinh nêu** được cảm nhận về vẻ đẹp tâm hồn và tư tưởng của tác giả.

2.5.4. Ý nghĩa sư phạm

2.5.4.1. Thiết kế hoạt động theo định hướng năng lực

Hoạt động dạy học được xây dựng dựa trên **thang tư duy Bloom**, bảo đảm bao quát đầy đủ các cấp độ từ thấp đến cao: **Nhớ – Hiểu – Vận dụng – Phân tích – Đánh giá – Sáng tạo**. Nhờ đó, học sinh không chỉ nắm vững kiến thức nền tảng mà còn phát triển tư duy phản biện, khả năng đánh giá, và năng lực sáng tạo sản phẩm mới.

Bên cạnh đó, hoạt động đáp ứng đa dạng **phong cách học tập (VARK)**, bao gồm:

- Vận động (Kinesthetic): trải nghiệm, thao tác thực hành.

- Hình ảnh (Visual): quan sát hình minh họa, video.

- Âm thanh (Aural): nghe giảng, thảo luận, thuyết trình.

- Đọc/Viết (Read/Write): ghi chép, tóm tắt, viết báo cáo.

Đồng thời, hoạt động được thiết kế dựa trên **thuyết đa trí tuệ của Howard Gardner**, phát triển ít nhất 6/8 loại trí thông minh: **Vận động cơ thể – Tương tác xã hội – Ngôn ngữ – Nhận thức – Không gian – Âm nhạc**. Sự kết hợp này không chỉ giúp học sinh phát huy thế mạnh cá nhân, mà còn khuyến khích sự phát triển cân bằng và toàn diện.

2.5.4.2. Ý nghĩa sư phạm

Hoạt động dạy học được triển khai không chỉ giới hạn ở mục tiêu hình thành và nâng cao năng lực sử dụng công cụ AI cho học sinh, mà còn hàm chứa những giá trị sư phạm quan trọng.

Trước hết, việc vận dụng toàn diện các cấp độ của thang tư duy Bloom góp phần phát triển năng lực tư duy từ tái hiện tri thức đến sáng tạo tri thức mới, phù hợp với định hướng giáo dục phát triển năng lực. Bên cạnh đó, sự kết hợp đa dạng các hình thức học tập theo mô hình VARK bảo đảm tính toàn diện và tạo cơ hội học tập công bằng cho những đối tượng có phong cách tiếp nhận tri thức khác nhau. Trên nền tảng thuyết đa trí tuệ, hoạt động đồng thời khuyến khích học sinh khai thác và phát huy tiềm năng cá nhân, từ đó nâng cao tính linh hoạt và khả năng sáng tạo trong học tập.

Việc tích hợp công cụ AI cùng yếu tố đa giác quan không chỉ làm tăng tính hấp dẫn của bài học mà còn củng cố động lực và thái độ học tập tích cực. Xa hơn, hoạt động góp phần hình thành những năng lực cốt lõi của thế kỷ XXI, bao gồm tư duy phản biện, hợp tác, giao tiếp và

sáng tạo, đáp ứng yêu cầu đổi mới căn bản, toàn diện giáo dục trong bối cảnh cách mạng công nghiệp 4.0.

2.5.5. Quy trình cụ thể

- GV đã giao cho HS về nhà chuẩn bị trước:

YÊU CẦU – HĐ2: CHUẨN BỊ CHO TRIỂN LÃM "TÁI HIỆN THỜI ĐẠI"

Tại hoạt động này, GV đã giao cho HS về nhà chuẩn bị trước, cần đảm bảo được chất lượng thông qua Rubric 1.

+ Nhóm họa sĩ:

- Vẽ lại bức tranh thiên nhiên ngày hè, thêm các chú thích về hình ảnh, màu sắc, các động từ miêu tả sức sống thiên nhiên (Câu thơ 2 – 3 – 4)
- Cảm nhận âm thanh cuộc sống qua các từ ngữ miêu tả âm thanh "lao xao" "dắng dỏi" để diễn tả lại bức tranh cuộc sống (Câu thơ 5 – 6)

+ Người giao cảm:

- Vẽ lại bức tranh thiên nhiên ngày hè
- Xác định những nét đặc sắc nghệ thuật của bài thơ cách gieo vần, số tiếng, cách ngắt nhịp trong các dòng thơ và nhận xét được tác dụng của chúng trong việc thể hiện cảm xúc của nhân vật trữ tình.

+ Bậc hiền triết:

- Vẽ lại bức tranh thiên nhiên ngày hè
- Cảm nhận về tấm lòng của nhà thơ với nhân dân, đất nước (Câu thơ 7 – 8).

- Hình thức:

+ Tái hiện nội dung trên giấy A0 bằng màu sắc, tranh ảnh và chữ viết.

+ Có thể vận dụng AI để hoàn thiện tuy nhiên cần phải ghi rõ nguồn.

+ Nếu sử dụng tranh từ AI cần phải in ra và dán lên giấy A0 + tải file JPG/PNG lên link padlet.

+ Nếu vẽ trang lên giấy A0 thì chụp hình + tải file JPG/PNG lên link padlet.

- Thời gian báo cáo:

- Báo cáo: 5 phút/nhóm
- Phát vấn: 3 phút

- Cách thức tổ chức:

- Cách 1: Triển lãm phòng tranh

- Cách 2: Tổ chức tọa đàm: Ức Trai tâm thượng quang Khuê tảo
- Cách 3: Tổ chức buổi sinh hoạt câu lạc bộ thơ: Yêu thơ Ức Trai

+ GV tư vấn bộ câu hỏi để giúp HS trao đổi:

1. *Bảo kính cảnh giới* (bài 43) khác gì so với một bài thơ Đường luật thông thường về đề tài, bố cục, niêm - luật, ngôn từ, hình ảnh?

2. Tại sao lại có sự khác biệt đó?

3. Câu thơ mở đầu cho biết điều gì về cuộc sống và tâm trạng của nhân vật trữ tình?

3. Tác giả sử dụng từ ngữ, chi tiết, hình ảnh, âm thanh nào để miêu tả cảnh sắc mùa hè. Chỉ ra những nét đặc sắc trong cách cảm nhận thiên nhiên và bút pháp tả cảnh của tác giả.?

4. Cuộc sống của con người được nhà thơ tái hiện qua những âm thanh, hình ảnh nào? Tìm mối liên hệ giữa khung cảnh ấy với ước nguyện của nhân vật trữ tình trong hai câu thơ cuối.?

5. Câu thơ lục ngôn thứ nhất mở đầu bài thơ có tác dụng nhấn mạnh điều gì trong cuộc sống và tâm trạng của nhân vật trữ tình? Câu lục ngôn kết thúc bài thơ có ý nghĩa gì? Nhận xét về vị trí và giá trị của các câu lục ngôn trong bài thơ.

6. Đọc bài thơ, bạn cảm nhận được điều gì vẻ đẹp tâm hồn, tư tưởng của tác giả?

***Kết luận, nhận xét:**

- GV nhận xét, góp ý và kết luận.

*** Đánh giá sản phẩm của HS:**

+ GV đánh giá sản phẩm của HS thông qua rubric 1.

Rubric đánh giá sản phẩm minh họa bằng AI trong hoạt động đọc

Tiêu chí	**Mức 4** **Xuất sắc** **(9 – 10 điểm)**	**Mức 3** **Tốt** **(7 – 8 điểm)**	**Mức 2** **Đạt** **(5 – 6 điểm)**	**Mức 1** **Chưa đạt** **(Dưới 5 điểm)**
1. Nội dung văn học (40%)	Hình ảnh thể hiện chính xác nội dung, ý thơ/văn; làm nổi bật sắc thái nghệ thuật; gắn kết chặt chẽ với tác phẩm.	Hình ảnh thể hiện đúng phần lớn nội dung, có vài chi tiết chưa khớp nhưng vẫn phản ánh ý chính.	Hình ảnh chỉ thể hiện một phần nội dung; còn mờ nhạt, thiếu yếu tố nghệ thuật cốt lõi.	Hình ảnh lệch xa nội dung tác phẩm, không gợi đúng ý thơ/văn.
2. Sáng tạo & độc đáo (20%)	Ý tưởng mới lạ, thể hiện cá tính; hình ảnh giàu tính nghệ	Có yếu tố sáng tạo; hình ảnh nhìn chung độc đáo	Ý tưởng còn đơn giản, ít sáng tạo, hình ảnh quen thuộc.	Thiếu sáng tạo, lặp lại ý tưởng chung chung,

	thuật, gây ấn tượng mạnh.	nhưng chưa thực sự nổi bật.		không tạo được ấn tượng.
3. Kỹ năng sử dụng AI (20%)	Biết cách viết prompt chi tiết, chính xác; khai thác tối đa tính năng AI để tạo hình ảnh phù hợp.	Biết cách viết prompt tương đối đầy đủ, hình ảnh khá phù hợp; có thử nghiệm chỉnh sửa.	Prompt đơn giản, hình ảnh còn hạn chế, ít điều chỉnh.	Không biết cách viết prompt phù hợp; hình ảnh không liên quan.
4. Trình bày & thuyết minh sản phẩm (20%)	Thuyết minh rõ ràng, logic, liên hệ chặt chẽ giữa hình ảnh – tác phẩm – cảm nhận cá nhân.	Thuyết minh khá rõ ràng, có liên hệ tác phẩm; còn thiếu chiều sâu ở một vài ý.	Thuyết minh còn sơ sài, ít liên hệ với tác phẩm.	Không thuyết minh hoặc thuyết minh rời rạc, không liên quan.

+ HS đánh giá: thông qua bảng kiểm 1.

BẢNG KIỂM 1: ĐÁNH GIÁ SẢN PHẨM TRÌNH BÀY NỘI DUNG THẢO LUẬN CÁC NHÓM		
	Có xuất hiện	Không xuất hiện
1. Thể hiện được đúng đủ nội dung.		
2. Cách thể hiện phong phú không đơn điệu.		
3. Thiết kế phần trình bày thành chương trình/tiết mục sinh động, hấp dẫn.		
4. Thể hiện được sâu sắc nội dung.		
5. Cách thức thể hiện và nội dung hài hòa để lại ấn tượng sâu sắc với các bạn.		

+ Phương pháp: trực quan, hỏi – đáp, phòng tranh, triển lãm.

+ Công cụ: rubric, bảng kiểm, câu hỏi vấn đáp, công cụ hỗ trợ từ AI.

SẢN PHẨM TRANH CỦA HỌC SINH

3. Kết quả thực nghiệm

3.1. Kết quả thực nghiệm

Tiêu chí khảo sát	Trước khi áp dụng (%)	Sau khi áp dụng (%)	Mức cải thiện
Hứng thú với tiết học Ngữ văn	45	92	+47
Khả năng ghi nhớ chi tiết tác phẩm	52	88	+36
Hiểu sâu ý nghĩa hình ảnh, màu sắc trong văn bản	40	90	+50
Khả năng phân tích và trình bày ý kiến cá nhân	48	85	+37
Tích cực tham gia phát biểu, thảo luận	50	87	+37
Tự tin thể hiện cảm nhận nghệ thuật	43	89	+46
Mức độ hoàn thành bài tập trên lớp	70	96	+26

Kết quả khảo sát cho thấy mô hình *"Học Ngữ văn qua hội họa"* mang lại những chuyển biến rõ rệt và toàn diện trong quá trình học tập của học sinh. Trước khi áp dụng, mức độ hứng thú và khả năng cảm thụ nghệ thuật của học sinh còn hạn chế, nhiều em gặp khó khăn trong việc ghi nhớ và phân tích các chi tiết miêu tả trong tác phẩm. Sau khi triển khai mô hình, tất cả

các tiêu chí đều tăng trưởng mạnh, đặc biệt là hứng thú học tập *(+47%)* và khả năng hiểu sâu ý nghĩa hình ảnh, màu sắc *(+50%).*

Không chỉ giúp học sinh dễ dàng tiếp cận tác phẩm, mô hình còn tạo môi trường học tập sáng tạo, kích thích khả năng quan sát, liên tưởng và bày tỏ ý kiến cá nhân. Tỉ lệ học sinh tích cực tham gia phát biểu, thảo luận và tự tin thể hiện cảm nhận nghệ thuật đều đạt trên *85%.* Điều này chứng minh phương pháp kết hợp hội họa vào giảng dạy không chỉ nâng cao kết quả học tập mà còn đáp ứng được yêu cầu phát triển năng lực toàn diện cho học sinh trong bối cảnh giáo dục hiện đại.

Với những kết quả tích cực trên, mô hình hoàn toàn có thể được nhân rộng và áp dụng ở nhiều khối lớp, góp phần đổi mới phương pháp dạy học Ngữ văn theo hướng sinh động, trực quan và hiệu quả.

3.2. Bài học kinh nghiệm

3.2.1. Thuận lợi

Sử dụng phương pháp kết hợp hội họa trong giảng dạy văn học mang lại nhiều thuận lợi đáng kể. Thứ nhất, phương pháp này giúp học sinh hứng thú hơn trong việc tiếp cận tác phẩm văn học. Khi học sinh được nhìn thấy những hình ảnh minh họa sinh động, các yếu tố hình ảnh trong tác phẩm trở nên gần gũi và dễ tiếp cận, từ đó kích thích sự tò mò và khơi dậy niềm yêu thích môn học.

Thứ hai, phương pháp này rất phù hợp và dễ thực hiện ở các lớp học thuộc các lĩnh vực xã hội, nghệ thuật, nơi học sinh đã có một nền tảng cảm thụ nghệ thuật nhất định. Hội họa kết hợp với văn học có thể giúp học sinh liên kết các kiến thức xã hội, văn hóa và nghệ thuật một cách tự nhiên, mang lại hiệu quả cao trong việc học.

Thứ ba, mỗi học sinh có thể có tưởng tượng và cảm nhận riêng biệt về một tác phẩm văn học khi được tiếp cận qua hình ảnh. Điều này giúp các em phát triển tư duy sáng tạo và khả năng phân tích một cách độc lập. Khi học sinh được tự do thể hiện sự sáng tạo qua hội họa, các em sẽ cảm thấy tự tin và mạnh dạn hơn trong việc bày tỏ quan điểm cá nhân.

Cuối cùng, việc sử dụng hội họa trong giảng dạy giúp tạo dấu ấn sâu sắc trong trí nhớ học sinh. Hình ảnh trực quan dễ dàng ghi nhớ hơn những từ ngữ khô khan, giúp các em không chỉ hiểu mà còn nhớ lâu hơn về tác phẩm và các yếu tố nghệ thuật trong đó.

3.2.2. Khó khăn

Mặc dù phương pháp kết hợp hội họa vào dạy học mang lại nhiều giá trị sư phạm, song trong thực tiễn triển khai vẫn tồn tại những khó khăn nhất định. Thứ nhất, không phải học sinh nào cũng có khả năng hội họa, vốn thẩm mỹ hay kỹ năng vận dụng AI để hỗ trợ sáng tạo nghệ thuật. Việc phối hợp màu sắc nhằm biểu đạt cảm xúc, ý tưởng văn học có thể trở thành rào cản khiến một số học sinh cảm thấy lúng túng, thiếu tự tin, từ đó ảnh hưởng đến mức độ tham gia và hiệu quả học tập.

Thứ hai, việc tích hợp hội họa đòi hỏi giáo viên đầu tư nhiều thời gian cho khâu chuẩn bị, hướng dẫn, cũng như tổ chức hoạt động trên lớp. Quá trình giao nhiệm vụ, để học sinh có đủ thời gian tư duy, sáng tạo và hoàn thiện sản phẩm hội họa dễ dẫn đến kéo dài thời lượng, gây áp lực cho việc bảo đảm tiến độ chương trình và chuẩn đầu ra. Đồng thời, khâu chuẩn bị tài liệu, công cụ hỗ trợ hội họa hoặc AI cũng đòi hỏi sự tính toán kỹ lưỡng nhằm duy trì chất lượng dạy học.

Ngoài ra, một số công cụ AI hiện nay vẫn tồn tại giới hạn kỹ thuật, chẳng hạn giới hạn số lần sử dụng miễn phí trong ngày hoặc yêu cầu nâng cấp lên bản trả phí để khai thác đầy đủ chức năng. Đây là yếu tố có thể gây khó khăn cho cả giáo viên và học sinh, đặc biệt trong bối cảnh các cơ sở giáo dục chưa có điều kiện trang bị đồng bộ về tài nguyên công nghệ.

4. Kết luận

Việc ứng dụng hội họa trong giảng dạy môn Ngữ văn không chỉ là một phương pháp sư phạm mới mẻ mà còn thể hiện định hướng tiếp cận học sinh theo hướng toàn diện. Đặc biệt, với sự hỗ trợ của công nghệ AI, phương pháp này được nâng tầm về tính hiệu quả và tính khả thi. Hội họa AI không chỉ giúp học sinh trực quan hóa các yếu tố hình ảnh trong tác phẩm văn học, mà còn khơi gợi khả năng sáng tạo, phát triển trí tưởng tượng và mở rộng vốn hiểu biết nghệ thuật của các em.

Thông qua việc kết hợp giữa ngôn ngữ văn chương và hình ảnh do AI hỗ trợ, giáo viên có thể xây dựng những giờ học sinh động, hấp dẫn hơn, giúp học sinh cảm nhận văn học không chỉ ở tầng ngôn từ mà còn qua kênh thị giác. Điều này góp phần tăng cường năng lực cảm thụ văn học – nghệ thuật, đồng thời tạo ra môi trường học tập giàu trải nghiệm, khơi gợi hứng thú và động lực học tập.

Qua quá trình nghiên cứu và triển khai, có thể nhận thấy việc ứng dụng hội họa AI trong giảng dạy Ngữ văn là một giải pháp khả thi và hiệu quả. Phương pháp này vừa hỗ trợ việc ghi nhớ và phân tích tác phẩm, vừa làm phong phú thêm tiến trình dạy – học, hướng đến xây dựng những bài học đa giác quan, sáng tạo và giàu ý nghĩa nghệ thuật.

TÀI LIỆU THAM KHẢO

1. Bộ Giáo dục và Đào tạo (2018). Chương trình Giáo dục phổ thông môn Ngữ văn. NXB Giáo dục Việt Nam.

2. Bộ Giáo dục và Đào tạo (2022). Sách giáo khoa Ngữ văn 10 – Tập 2 (Bộ sách Kết nối tri thức với cuộc sống). NXB Giáo dục Việt Nam.

3. Nguyễn Thanh Hùng (2018). Đổi mới phương pháp dạy học Ngữ văn trong nhà trường phổ thông. NXB Giáo dục Việt Nam.

4. Phạm Thị Thu Hương (2021). Phát triển năng lực thẩm mỹ cho học sinh thông qua dạy học Ngữ văn. Tạp chí Khoa học Giáo dục Việt Nam.

THÔNG TIN TÁC GIẢ

Họ và tên	Nơi công tác	Email
Phạm Phương Linh	Trường Đại học Sư phạm – Đại học Đà Nẵng	phamphuonglinh9200@gmail.com
Nguyễn Thị Hoài Thương	Trường TH, THCS & THPT FPT	ThuongNTH27@fe.edu.vn

ĐÁNH GIÁ KHẢ NĂNG GÁN NHÃN NGỮ NGHĨA TỰ ĐỘNG CỦA MÔ HÌNH NGÔN NGỮ LỚN CHO MỤC TỪ DANH TỪ TIẾNG VIỆT

Phan Thị Mỹ Trang[1], Đỗ Quốc Trí[2], Phan Văn Bá Hải[3], Đặng Ngọc Lệ[4], Đinh Điền[5], Trần Thị Minh Phượng[6]

Tóm tắt: Mô hình ngôn ngữ lớn (LLM) ngày càng được ứng dụng trong nhiều lĩnh vực, đặc biệt trong xử lý ngôn ngữ tự nhiên. Nghiên cứu này khai thác LLM để gán nhãn ngữ nghĩa tự động trực tiếp cho các mục từ danh từ tiếng Việt dựa trên hệ thống WordNet. Hiệu quả gán nhãn được đánh giá thông qua bộ ngữ liệu vàng gồm 880 mục từ, được xây dựng từ ba nguồn: (1) *Open English WordNet*, (2) *Từ điển OALD song ngữ Anh-Việt* và (3) *Từ điển tiếng Việt*. Kết quả cho thấy LLM đạt độ chính xác 82,39% khi gán nhãn, qua đó khẳng định tiềm năng ứng dụng LLM trong việc xây dựng và mở rộng hệ thống nhãn ngữ nghĩa phù hợp với đặc trưng của tiếng Việt.

Từ khóa: Gán nhãn ngữ nghĩa; mô hình ngôn ngữ lớn; WordNet; danh từ tiếng Việt; bộ ngữ liệu vàng.

1. GIỚI THIỆU

Trong những năm gần đây, mô hình ngôn ngữ lớn (Large Language Model – LLM) đã đạt được những bước tiến đáng kể trong xử lý ngôn ngữ tự nhiên (Natural Language Processing – NLP) nhờ khả năng phân tích ngữ nghĩa và suy luận dựa trên ngữ cảnh. Đối với các ngôn ngữ có nguồn tài nguyên số hóa còn hạn chế như tiếng Việt, LLM mở ra tiềm năng thực hiện các tác vụ xử lý ngôn ngữ tự động, trong đó có gán nhãn ngữ nghĩa cho từ vựng. Gán nhãn ngữ nghĩa là quá trình xác định và gán cho mỗi mục từ một mã định danh ngữ nghĩa (synset_ID) trong hệ thống nhãn ngữ nghĩa, nhằm đảm bảo mỗi mục từ được ánh xạ nhất quán tới một nghĩa xác định, phục vụ cho việc tổ chức, tra cứu và xử lý ngôn ngữ một cách có hệ thống.

Một trong những hệ thống nhãn ngữ nghĩa tiêu biểu là WordNet (Miller và các tác giả khác, 1990), trong đó các mục từ được tổ chức thành các tập đồng nghĩa (synset), mỗi synset biểu thị một khái niệm duy nhất và được liên kết với các synset khác thông qua các quan hệ ngữ nghĩa như thượng danh-hạ danh (hypernymy-hyponymy), bộ phận-tổng thể (meronymy-holonymy), kéo theo (entailment) và cách thức đặc biệt (troponymy), tạo thành mạng lưới từ vựng-ngữ nghĩa có cấu trúc chặt chẽ, hỗ trợ hiệu quả cho việc xử lý ngữ nghĩa. Tuy nhiên, Princeton WordNet (WordNet gốc) được xây dựng dựa trên ngôn ngữ và bối cảnh văn hóa phương Tây, nên khi áp dụng trực tiếp cho các ngôn ngữ bản địa khác thường gặp hạn chế về độ chính xác và tính phù hợp văn hóa (Khalilia và các tác giả khác, 2024; Freihat và các tác giả khác, 2024).

[1] NCS., Trường Đại học Khoa học xã hội và Nhân văn, ĐHQG-HCM; ThS., Trường Đại học Công nghệ Sài Gòn; *e-mail: mytrang779@gmail.com*
[2] Sinh viên Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM
[3] Sinh viên Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM
[4] PGS.TS., Trường Đại học Nguyễn Tất Thành
[5] PGS.TS., Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM
[6] TS., Trường Đại học Khoa học xã hội và Nhân văn, ĐHQG-HCM

Tiếng Việt với đặc trưng là ngôn ngữ đơn lập, giàu hiện tượng đa nghĩa và đồng âm, có nhiều khái niệm mang tính đặc thù văn hóa, khiến việc ánh xạ nghĩa từ WordNet sang tiếng Việt gặp nhiều thách thức.

Những tiến bộ gần đây của LLM, đặc biệt khi kết hợp với phương pháp gợi ý (prompt-based method) cùng cơ sở tri thức (knowledge base), đã cho thấy khả năng cải thiện đáng kể độ chính xác trong phân giải nghĩa từ (Sumanathilaka và các tác giả khác, 2024). Cách tiếp cận này không chỉ mở ra tiềm năng thực hiện ánh xạ tự động giữa các synset của WordNet và định nghĩa tiếng Việt, mà còn góp phần thúc đẩy tiến trình xây dựng tài nguyên ngôn ngữ số cho tiếng Việt. Tuy nhiên, hiện vẫn chưa có nhiều nghiên cứu tập trung vào việc đánh giá và phân tích khả năng gán nhãn ngữ nghĩa của các mô hình ngôn ngữ lớn đối với tiếng Việt, đặc biệt trong bối cảnh ngữ liệu số hóa hạn chế và đặc thù ngôn ngữ – văn hóa riêng biệt. Nghiên cứu này tập trung đánh giá khả năng gán nhãn của LLM cho tiếng Việt dựa trên bộ ngữ liệu chuẩn, cũng như phân tích lỗi mà mô hình thường gặp. Nghiên cứu này tập trung vào ba câu hỏi chính:

(1) Bộ ngữ liệu vàng được xây dựng như thế nào để làm chuẩn đánh giá?

(2) LLM đạt độ chính xác bao nhiêu khi gán nhãn ngữ nghĩa cho các mục từ danh từ tiếng Việt?

(3) Những lỗi phổ biến mà LLM gặp phải trong quá trình gán nhãn ngữ nghĩa cho mục từ danh từ tiếng Việt là gì?

Thông qua việc giải quyết các câu hỏi trên, nghiên cứu hướng tới đánh giá tiềm năng và giới hạn của LLM trong gán nhãn ngữ nghĩa tự động cho mục từ danh từ tiếng Việt, kết quả nghiên cứu sẽ góp phần vào việc phát triển hệ thống nhãn ngữ nghĩa dành cho ngôn ngữ tiếng Việt.

2. TỔNG QUAN NGHIÊN CỨU

Nhiều quốc gia trên thế giới đã mở rộng WordNet sang các ngôn ngữ khác, chẳng hạn như: *GermaNet* cho tiếng Đức (Hamp & Feldweg, 1997); *Japanese WordNet* cho tiếng Nhật (Isahara và các tác giả khác, 2008); *Chinese Open WordNet* cho tiếng Trung (Wang & Bond, 2013); cùng nhiều phiên bản WordNet của các quốc gia khác như *Czech WordNet, Dutch WordNet, Estonian WordNet*, *French WordNet, German WordNet, Italian WordNet, Spanish WordNet*… (Vossen, 2002). Các công trình này thường áp dụng hai hướng chính: (1) dịch trực tiếp các synset từ tiếng Anh sang ngôn ngữ đích hoặc (2) xây dựng WordNet bản địa độc lập rồi ánh xạ liên ngôn ngữ sang Princeton WordNet. Đối với tiếng Việt, đã có một số nghiên cứu nỗ lực xây dựng WordNet tiếng Việt (Trương Hoàng Bảo và các tác giả khác, 2021; Lam & Kalita, 2022), nhưng phần lớn các nghiên cứu vẫn dựa trên dịch song ngữ từ tiếng Anh sang tiếng Việt, dẫn đến nhiều hạn chế về tính phù hợp với đặc trưng ngôn ngữ và văn hóa Việt Nam.

Sự phát triển của mô hình ngôn ngữ lớn (LLM) mở ra hướng tiếp cận mới cho việc gán nhãn ngữ nghĩa, khai thác khả năng hiểu ngữ cảnh và suy luận của LLM để gán nhãn từ vựng tiếng Việt dựa trên synset của WordNet. Nghiên cứu của Sumanathilaka và các tác giả khác (2024) cho thấy, kết hợp kỹ thuật gợi ý nâng cao (prompt augmentation) với cơ sở tri thức

(knowledge base) và các bước lọc nghĩa tinh chỉnh giúp LLM cải thiện đáng kể độ chính xác phân giải nghĩa từ, đặc biệt với danh từ và tính từ. Kholodna và các tác giả khác (2024) chứng minh rằng việc kết hợp mô hình ngôn ngữ lớn với Active Learning là giải pháp khả thi cho nhận dạng thực thể có tên (Named Entity Recognition – NER) ở các ngôn ngữ ít tài nguyên. Kết quả cho thấy GPT-4-Turbo đạt độ chính xác và độ đồng thuận gần tương đương với gán nhãn thủ công ở một số ngôn ngữ, đồng thời giảm chi phí gán nhãn tới 42,45 lần; prompt chi tiết và gán nhãn theo lô giúp duy trì độ chính xác trong khi tiết kiệm tài nguyên tính toán.

Cùng với sự phát triển nhanh chóng của các mô hình ngôn ngữ lớn, nhu cầu đánh giá năng lực của mô hình trở nên cần thiết. Hiện tại, phần lớn LLM được đánh giá bằng các bộ benchmark tiếng Anh như *MMLU, HellaSwag, ARC* hay *TruthfulQA*, vốn phản ánh bối cảnh văn hóa – xã hội phương Tây nên chưa hoàn toàn phù hợp khi áp dụng cho tiếng Việt (Nguyên và các tác giả khác 2024). Một số bộ dữ liệu tiếng Việt đã xuất hiện (*UIT-ViNewsQA, UIT-ViQuAD, UIT-ViWikiQA, VMLU*…), nhưng vẫn chưa bao quát đầy đủ tri thức và đặc trưng văn hóa Việt (Nguyên và các tác giả khác 2024). Để khắc phục, Nguyên và các tác giả khác (2024) xây dựng bộ đánh giá gồm bốn nhóm dữ liệu: *LAMBADA_vi* (đo khả năng hiểu ngữ cảnh dài), *Exam Vietnamese* (kiến thức lớp 6–12), *General Knowledge* (kiến thức tổng quát từ gameshow) và *Comprehension QA* (đọc hiểu sâu). Các dữ liệu này được thu thập, chuẩn hóa và kiểm duyệt thủ công nhằm đảm bảo tính chính xác và phù hợp. Kết quả thử nghiệm trên nhiều mô hình, từ ChatGPT đến các LLM tiếng Việt, cho thấy mỗi mô hình mạnh ở một số tác vụ nhất định như: ChatGPT dẫn đầu ở hầu hết tác vụ trắc nghiệm; Vistral-7B có perplexity *(mức độ "bối rối" khi dự đoán từ tiếp theo)* thấp nhất ở LAMBADA; Dama-2-7B và PhoGPT-4B cạnh tranh sát sao; SeaLLM-7B-v2 dù perplexity cao nhưng thể hiện khá tốt ở bài đọc hiểu. Tuy nhiên, chưa có mô hình nào vượt trội toàn diện. Trương và các tác giả khác (2024) xây dựng khung đánh giá cho LLM tiếng Việt, tinh chỉnh năm mô hình gồm *URA-LlaMa 7B, 13B, 70B, MixSUra* và *GemSUra* từ các kiến trúc *LlaMa-2, Mixtral* và *Gemma*, sử dụng các bộ dữ liệu như *Wikipedia tiếng Việt*, *Vietnamese News-Corpus* và *Vietnamese Highschool Essays*. Kết quả cho thấy GPT-4 đạt hiệu suất cao nhất nhưng thiên lệch cao, URA-LlaMa vượt trội so với LlaMa-2 gốc. Nghiên cứu cũng chỉ ra rằng mô hình lớn hơn không đồng nghĩa với hiệu suất tốt hơn, mà chất lượng dữ liệu huấn luyện/tinh chỉnh mới là yếu tố quyết định. Ngoài ra, các kỹ thuật như *few-shot, chain-of-thought* và *thiết kế prompt* ảnh hưởng rõ rệt đến kết quả. Từ thực tế này, nghiên cứu của chúng tôi đề xuất hướng tiếp cận mới là xây dựng bộ ngữ liệu vàng gồm các mục từ đơn nghĩa có đối ứng trong WordNet, nhằm đánh giá khách quan khả năng gán nhãn ngữ nghĩa của LLM cho tiếng Việt.

3. NGUỒN NGỮ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

3.1. Thu thập và tiền xử lý ngữ liệu

Ngữ liệu được thu thập từ ba nguồn chính: *Open English WordNet* (McCrae, phiên bản 2024), *Từ điển Oxford Advanced Learner's Dictionary (OALD) song ngữ Anh-Việt* (Joanna & Đinh Điền, 2014) và *Từ điển tiếng Việt* (Hoàng Phê chủ biên, tái bản năm 2020).

3.1.1. Mục từ danh từ trong Open English WordNet

Open English WordNet là phiên bản mở rộng và cập nhật của Princeton WordNet, cung cấp hệ thống nhãn ngữ nghĩa cho từ vựng tiếng Anh. Từ kho ngữ liệu này, nhóm nghiên cứu trích xuất toàn bộ các mục từ danh từ, gồm 84.956 mục từ, tương ứng với 123.923 định nghĩa (glosses). Mỗi định nghĩa được thu thập kèm theo mã định danh synset (synset_ID), ví dụ minh họa (nếu có) và các mối quan hệ ngữ nghĩa cơ bản như thượng danh (hypernymy), hạ danh (hyponymy), đồng nghĩa (synonymy) và đa nghĩa (polysemy).

3.1.2. Mục từ danh từ trong Từ điển OALD song ngữ Anh-Việt

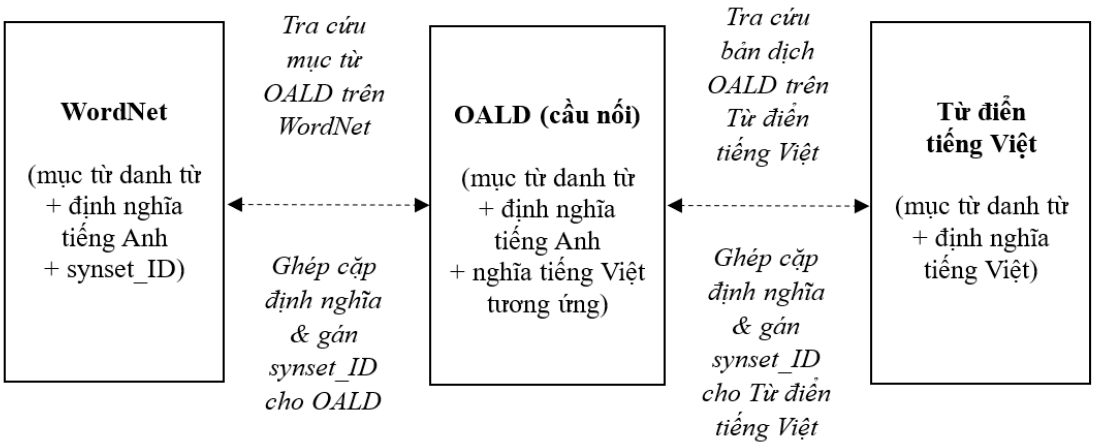
Từ điển OALD (Oxford Advanced Learner's Dictionary) song ngữ Anh–Việt, ấn bản lần thứ 8 (Joanna & Đinh Điền, 2014) được sử dụng làm cầu nối ngữ nghĩa giữa hệ thống nhãn ngữ nghĩa Open English WordNet (tiếng Anh) và Từ điển tiếng Việt. OALD được sắp xếp theo cấu trúc truyền thống (trình bày mục từ theo thứ tự bảng chữ cái), cung cấp đồng thời ba loại thông tin cần thiết: (i) từ vựng tiếng Anh, (ii) định nghĩa tiếng Anh, và (iii) dịch nghĩa tiếng Việt tương ứng. Ngữ liệu được thu thập gồm 25.509 mục từ danh từ cùng với 33.493 định nghĩa, kèm ví dụ minh họa và phần dịch nghĩa tiếng Việt tương ứng. Sau khi thu thập, ngữ liệu được chuẩn hóa định dạng, loại bỏ ký tự đặc biệt và tách thành các trường thông tin: mục từ, định nghĩa, ví dụ minh họa và nghĩa tiếng Việt.

3.1.3. Mục từ danh từ trong Từ điển tiếng Việt

Nguồn ngữ liệu tiếng Việt được sử dụng là *Từ điển tiếng Việt* do Viện Ngôn ngữ học biên soạn, Hoàng Phê chủ biên (tái bản 2020). Nhóm nghiên cứu thu thập 17.831 mục từ danh từ, tương ứng với 22.098 định nghĩa, kèm thông tin về ví dụ minh họa, từ đồng nghĩa, từ gốc, từ phái sinh, cũng như các ghi chú về phong cách biểu đạt, phạm vi sử dụng và sắc thái tu từ. Sau khi thu thập, toàn bộ ngữ liệu được chuẩn hóa định dạng văn bản, loại bỏ nhiễu (kí hiệu, dấu câu đặc biệt) và phân loại theo các trường như: mục từ danh từ, định nghĩa, ví dụ minh họa, quan hệ ngữ nghĩa (nếu có) và thông tin ngữ dụng.

3.2. Xây dựng Bộ ngữ liệu vàng

Bộ ngữ liệu vàng được xây dựng nhằm làm cơ sở đối chiếu và đánh giá độ chính xác của kết quả gán nhãn ngữ nghĩa tự động cho các mục từ danh từ tiếng Việt sử dụng mô hình ngôn ngữ lớn. Các mục từ dùng để xây dựng bộ ngữ liệu vàng được chọn lọc dựa trên ba tiêu chí: (i) mỗi mục từ trong OALD phải có duy nhất một nghĩa tiếng Anh, (ii) nghĩa tương ứng trong Từ điển tiếng Việt cũng phải là từ đơn nghĩa, và (iii) có thể tìm được synset tương ứng trong WordNet. Bộ ngữ liệu vàng được xây dựng dựa trên phương pháp gián tiếp, sử dụng từ điển song ngữ OALD làm cầu nối giữa hệ thống WordNet tiếng Anh và Từ điển tiếng Việt (xem Hình 1).

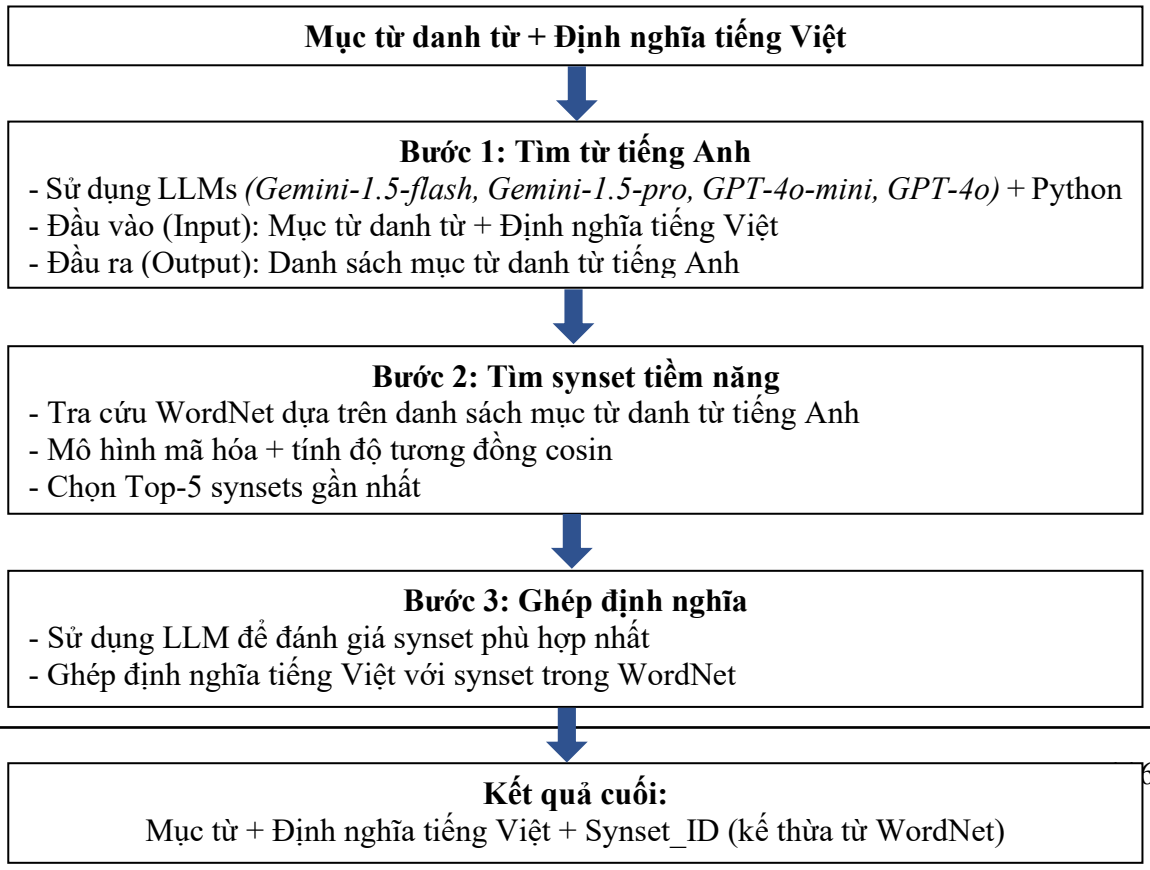


Hình 1. *Sơ đồ quy trình xây dựng Bộ ngữ liệu vàng*

Quy trình xây dựng bộ ngữ liệu vàng gồm ba bước: (1) các mục từ danh từ trong OALD được tra cứu trên WordNet và đối chiếu định nghĩa; khi hai định nghĩa tương đồng về ngữ nghĩa, mã synset_ID từ WordNet được gán cho định nghĩa tương ứng trong OALD; (2) dựa trên bản dịch tiếng Việt của các mục từ trong OALD để truy xuất, xác định đúng hình thức từ và thu thập các định nghĩa tương ứng trong Từ điển tiếng Việt; (3) sử dụng OALD làm cầu nối để gán mã synset_ID từ WordNet sang định nghĩa tiếng Việt.

3.3. Gán nhãn ngữ nghĩa tự động sử dụng mô hình ngôn ngữ lớn

Sau khi xây dựng bộ ngữ liệu vàng, nghiên cứu tiếp tục sử dụng chính các mục từ tiếng Việt này cùng định nghĩa của chúng, với số lượng tương ứng bộ ngữ liệu vàng, để gán nhãn ngữ nghĩa trực tiếp bằng mô hình ngôn ngữ lớn. Quy trình gán nhãn được triển khai qua ba bước chính (xem Hình 2):



Hình 2. *Sơ đồ quy trình gán nhãn ngữ nghĩa tự động sử dụng LLMs*

Trước hết, mô hình ngôn ngữ lớn *(Gemini-1.5 Flash, Gemini-1.5 Pro, GPT-4o Mini, GPT-4o)* kết hợp công cụ *Python* được sử dụng để truy vấn các mục từ danh từ tiếng Anh tương ứng với mục từ tiếng Việt, dựa trên hình thái và định nghĩa tiếng Việt đầu vào. Tiếp theo, từ danh sách các mục từ tiếng Anh được gợi ý, nhóm nghiên cứu tra cứu WordNet để thu thập các synset liên quan. Các định nghĩa (gloss) từ WordNet và định nghĩa tiếng Việt được mã hóa và so sánh bằng độ tương đồng cosin, từ đó xác định 5 synset có nghĩa gần nhất. Cuối cùng, mô hình ngôn ngữ lớn được dùng để đánh giá các synset ứng viên và chọn ra synset phù hợp nhất với định nghĩa tiếng Việt, làm nhãn ngữ nghĩa chính thức cho mục từ đầu vào.

3.4. Đánh giá kết quả gán nhãn ngữ nghĩa

Hiệu quả của gán nhãn ngữ nghĩa tự động sử dụng LLM được kiểm chứng thông qua đối chiếu với bộ ngữ liệu vàng đã xây dựng trước đó. Quá trình đánh giá được thực hiện theo phương pháp bán thủ công, kết hợp giữa xử lý tự động và đánh giá thủ công bởi chuyên gia ngôn ngữ. Cụ thể, trước hết đối chiếu tự động bằng ngôn ngữ lập trình Python giữa synset_ID mà mô hình gán trực tiếp với synset_ID trong tập ngữ liệu vàng (được gán gián tiếp và đã kiểm tra lại thủ công). Sau đó, đối với các trường hợp trả kết quả gán nhãn không trùng khớp với bộ ngữ liệu vàng, các chuyên gia ngôn ngữ trước tiên đánh giá độc lập, sau đó đối chiếu và thống nhất về mức độ tương đồng ngữ nghĩa giữa định nghĩa tiếng Việt và gloss (định nghĩa) của synset được gán. Thang đánh giá gồm bốn mức: (1) Nghĩa không tương đồng hoàn toàn (nghĩa hoàn toàn không liên quan), (2) Nghĩa tương đồng một phần (nghĩa có mối liên hệ hoặc có điểm chung nhất định, nhưng mức độ liên hệ còn lỏng lẻo, chưa đủ cơ sở để xác định là tương đồng về mặt ngữ nghĩa), (3) Nghĩa tương đồng cao (nghĩa gần như trùng khớp, chỉ khác biệt nhỏ về sắc thái, phạm vi hoặc cách diễn đạt), (4) Nghĩa tương đồng hoàn toàn (nghĩa trùng khớp tuyệt đối).

Dựa trên kết quả đánh giá, điểm trung bình được phân loại theo bốn mức chất lượng như sau:

+ Dưới 1,75 là không đạt (mô hình gán nhãn sai lệch, cần điều chỉnh lại thuật toán)

+ Từ 1,75 đến dưới 2,5 là trung bình (mô hình có độ chính xác chưa ổn định, cần rà soát một số trường hợp không tương đồng ngữ nghĩa)

+ Từ 2,5 đến dưới 3,25 là khá tốt (mô hình đạt độ chính xác cao, phần lớn nhãn phản ánh đúng nội dung định nghĩa)

+ Từ 3,25 trở lên được xem là tốt (mô hình gán nhãn ngữ nghĩa chính xác)

Kết quả đánh giá giúp thống kê số trường hợp gán nhãn chính xác và phản ánh được mức độ tương đồng ngữ nghĩa trong các trường hợp không trùng khớp tuyệt đối nhưng vẫn gần nghĩa hoặc tương đương về nội dung so với định nghĩa gốc.

4. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

4.1. Kết quả xây dựng bộ ngữ liệu vàng

Bộ ngữ liệu vàng được xây dựng theo phương pháp gián tiếp, sử dụng từ điển song ngữ OALD làm cầu nối giữa WordNet và Từ điển tiếng Việt. Bộ ngữ liệu vàng này được sử dụng như tập chuẩn dùng để đánh giá độ chính xác của hệ thống gán nhãn ngữ nghĩa tự động bằng mô hình ngôn ngữ lớn. Sau khi xử lý và lọc theo các tiêu chí lựa chọn, kết quả thu được 880 mục từ danh từ tiếng Việt được gán nhãn ngữ nghĩa chính xác với synset trong WordNet. Một số ví dụ đại diện được trình bày trong Bảng 1, cho thấy sự tương đồng rõ ràng giữa định nghĩa tiếng Việt, định nghĩa tiếng Anh trong OALD và định nghĩa (gloss) tương ứng trong WordNet.

Bảng 1. *Cấu trúc bộ ngữ liệu vàng*

STT	**Mục từ và định nghĩa tiếng Việt** ***(Từ điển Tiếng Việt)***	**Mục từ và định nghĩa tiếng Anh, dịch nghĩa tiếng Việt tương ứng** ***(OALD)***	**Mục từ và định nghĩa tiếng Anh, synset_ID** ***(Open English WordNet)***	**Mã synset (synset_ID) kế thừa từ** ***Open English WordNet***
1.	*bão từ:* hiện tượng nhiễu loạn của từ trường Trái Đất do ảnh hưởng của các hiện tượng xảy ra đột nhiên trên Mặt Trời	*magnetic storm:* a situation in which the magnetic field of the earth or of another planet, star, etc. is disturbed *(bão từ)*	*magnetic storm:* a sudden disturbance of the earth's magnetic field; caused by emission of particles from the sun *(oewn-07443397-n)*	*oewn-07443397-n*
2.	*binh sĩ:* binh lính và hạ sĩ quan [nói tổng quát]	*the rank and file:* the ordinary soldiers who are not officers *(binh sĩ)*	*rank, rank and file:* the ordinary members of an organization (such as the enlisted soldiers of an army) *(oewn-08415136-n)*	*oewn-08415136-n*
3.	*bình minh:* lúc trời tang tảng sáng, trước khi mặt trời mọc	*first light:* the time when light first appears in the morning *(bình minh)*	*aurora, break of day, break of the day, cockcrow, dawn, dawning, daybreak, dayspring, first light, morning, sunrise, sunup:*	*oewn-15193837-n*

			the first light of day *(oewn-15193837-n)*	
4.	*chùa:* công trình được xây cất lên, làm nơi thờ Phật, thường có nhà sư ở	*pagoda:* a temple (= religious building) in S or E Asia in the form of a tall tower with several levels, each of which has its own roof that extends beyond the walls *(chùa)*	*pagoda:* an Asian temple; usually a pyramidal tower with an upward curving roof *(oewn-03880643-n)*	*oewn-03880643-n*
5.	*kim băng:* kim được uốn gập lại, đầu có mũ bọc mũi nhọn, dùng để cài, găm	*safety pin:* a pin with a point bent back towards the head, that is covered when closed so that it cannot hurt you (*kim băng)*	*safety pin:* a pin in the form of a clasp; has a guard so the point of the pin will not stick the user *(oewn-04134350-n)*	*oewn-04134350-n*

4.2. Đánh giá kết quả gán nhãn ngữ nghĩa tự động của mô hình ngôn ngữ lớn

Kết quả gán nhãn ngữ nghĩa tự động từ WordNet sang các mục từ danh từ tiếng Việt bằng phương pháp trực tiếp thông qua mô hình ngôn ngữ lớn cho thấy mức độ khớp đáng kể với bộ ngữ liệu vàng. Những trường hợp trùng khớp mã synset được xem là gán nhãn chính xác. Kết quả có 725/880 mục từ được gán synset_ID trùng khớp với bộ ngữ liệu vàng (chiếm 82,39%), chứng tỏ mô hình ngôn ngữ lớn có khả năng xử lý ngữ nghĩa với độ chính xác cao. Việc xác định chính xác synset tương ứng trong WordNet chỉ dựa trên định nghĩa tiếng Việt là nhiệm vụ phức tạp, đòi hỏi mô hình không chỉ hiểu nghĩa của từ mà còn nhận diện chính xác các cấp độ từ khái quát đến cụ thể của khái niệm. Kết quả trên cho thấy mô hình ngôn ngữ lớn có tiềm năng ứng dụng trong các tác vụ gán nhãn ngữ nghĩa tự động cho tiếng Việt, góp phần rút ngắn thời gian thực hiện và nâng cao độ chính xác của kho ngữ liệu ngữ nghĩa. 155 mục từ còn lại (chiếm 17,61%) không trùng khớp với nhãn chuẩn trong bộ ngữ liệu vàng, được xem là các trường hợp gán nhãn sai. Bảng 2 trình bày một số ví dụ đại diện cho cả hai trường hợp: khớp nhãn (TRUE) và không khớp nhãn (FALSE).

Bảng 2. *Đối chiếu kết quả LLM với Bộ ngữ liệu vàng*

STT	Mã synset (synset_ID), mục từ, định nghĩa *(Bộ ngữ liệu vàng)*	Mã synset (synset_ ID), mục từ, định nghĩa tiếng Anh *(sử dụng mô hình ngôn ngữ lớn)*	Kết quả
1.	*[Synset_ID: oewn-09479337-n]* **Từ điển tiếng Việt:* ruộng bậc thang: ruộng được san phẳng thành nhiều bậc ở vùng sườn đồi, sườn núi **OALD:* terrace: one of a series of flat areas of ground that are cut into the side of a hill like steps so that crops can be grown there **WordNet:* bench, terrace: a level shelf of land interrupting a declivity (with steep slopes above and below)	*[Synset_ID: oewn-09479337-n]* bench, terrace: a level shelf of land interrupting a declivity (with steep slopes above and below)	*TRUE* (Khớp nhãn)
2.	*[Synset_ID: oewn-08415136-n]* **Từ điển tiếng Việt:* binh sĩ: binh lính và hạ sĩ quan [nói tổng quát] **OALD:* the rank and file: the ordinary soldiers who are not officers **WordNet:* rank, rank and file: the ordinary members of an organization (such as the enlisted soldiers of an army)	*[Synset_ID: oewn-10641415-n]* soldier: an enlisted man or woman who serves in an army	*FALSE* (Không khớp nhãn)
3.	*[Synset_ID: oewn-09420951-n]* **Từ điển tiếng Việt:*	*[Synset_ID: oewn-09420951-n]*	*TRUE* (Khớp nhãn)

	ổ gà: chỗ mặt đường bị lở, lõm sâu xuống giống như ổ của gà **OALD:* pothole: a large rough hole in the surface of a road that is formed by traffic and bad weather **WordNet:* chuckhole, pothole: a pit or hole produced by wear or weathering (especially in a road surface)	chuckhole, pothole: a pit or hole produced by wear or weathering (especially in a road surface)	
4.	*[Synset_ID: oewn-12921222-n]* **Từ điển tiếng Việt:* ớt: cây nhỏ cùng họ với cà, hoa trắng, quả chín có màu đỏ hay vàng, vị cay, dùng làm gia vị **OALD:* capsicum: a type of plant which has hollow fruits. Some types of these are eaten as vegetables, either raw or cooked, for example sweet peppers or chillies. **WordNet:* capsicum, capsicum pepper plant, pepper: any of various tropical plants of the genus Capsicum bearing peppers	*[Synset_ID: oewn-12921747-n]* Capsicum annuum longum, cayenne, cayenne pepper, chili pepper, chilli pepper, jalapeno, long pepper: plant bearing very hot and finely tapering long peppers; usually red	*FALSE* (Không khớp nhãn)
5.	*[Synset_ID: oewn-03710918-n]* **Từ điển tiếng Việt:* nam châm: vật có đặc tính hút sắt và hút hay đẩy dây dẫn có dòng điện chạy qua **OALD:* magnet:	*[Synset_ID: oewn-03710918-n]* magnet: (physics) a device that attracts iron and produces a magnetic field	*TRUE* (Khớp nhãn)

	a piece of iron that attracts objects made of iron towards it, either naturally or because of an electric current that is passed through it **WordNet:* magnet: (physics) a device that attracts iron and produces a magnetic field		

Để đánh giá chất lượng các trường hợp không khớp nhãn, nhóm nghiên cứu áp dụng thang đo bốn mức nhằm xác định mức độ tương đồng ngữ nghĩa giữa định nghĩa (gloss) của synset được gán và định nghĩa tiếng Việt. Thang đo này dao động từ mức 1 (không tương đồng hoàn toàn) đến mức 4 (tương đồng hoàn toàn).

Bảng 3. *Đánh giá tương đồng ngữ nghĩa đối với các nhãn chưa khớp*

Mức điểm đánh giá	**Số lượng định nghĩa**	**Tỷ lệ (%)**
Nghĩa không tương đồng hoàn toàn (1 điểm)	2	1,29%
Nghĩa tương đồng một phần (2 điểm)	17	10,97%
Nghĩa tương đồng cao (3 điểm)	136	87,74%
Nghĩa tương đồng hoàn toàn (4 điểm)	0	0%
Điểm trung bình: 2,86	155	100%

Kết quả đánh giá (Bảng 3) cho thấy: chỉ có 2 cặp định nghĩa (chiếm tỉ lệ 1,29%) ở mức 1 (nghĩa không tương đồng hoàn toàn); 17 cặp định nghĩa (chiếm tỉ lệ 10,97%) ở mức 2 (nghĩa tương đồng một phần, có yếu tố liên quan nhưng chưa đủ để xem là khớp ngữ nghĩa); phần lớn, 136 cặp (chiếm tỉ lệ 87,74%) đạt mức 3, tức là có tương đồng ngữ nghĩa cao, chỉ khác biệt nhẹ về sắc thái, phạm vi hoặc cách diễn đạt; 0 cặp đạt mức 4 (chiếm tỉ lệ 0%). Điểm đánh giá trung bình đạt 2,86 trên thang đánh giá 4, phản ánh rằng ngay cả khi không trùng khớp tuyệt đối với nhãn chuẩn, mô hình có khả năng đưa ra phần lớn synset phản ánh gần đúng nội dung định nghĩa.

Qua khảo sát chi tiết trên các cặp định nghĩa này (xem Bảng 4), có thể xác định một số nguyên nhân chính dẫn đến kết quả ghép cặp không chính xác.

Bảng 4. *Khảo sát nguyên nhân gán nhãn trực tiếp không chính xác*

STT	Mục từ, định nghĩa tiếng Việt	Mục từ, định nghĩa tiếng Anh và mã synset *(Gán nhãn gián tiếp - Ngữ liệu vàng)*	Mục từ, định nghĩa tiếng Anh và mã synset *(Gán nhãn trực tiếp)*	Mức tương đồng
1.	góc tù: góc nhỏ hơn góc bẹt và lớn hơn góc vuông	obtuse angle: an angle between 90 and 180 degrees *(một góc nằm giữa 90 và 180 độ)*; (oewn-13914741-n)	acute angle: an angle less than 90 degrees but more than 0 degrees *(một góc nhỏ hơn 90 độ nhưng lớn hơn 0 độ)*; (oewn-13914630-n)	1
2.	chùa: công trình được xây cất lên, làm nơi thờ Phật, thường có nhà sư ở	pagoda: an Asian temple; usually a pyramidal tower with an upward curving roof *(một ngôi đền Châu Á; thường là tháp hình chóp với mái cong hướng lên trên)*; (oewn-03880643-n)	temple: place of worship consisting of an edifice for the worship of a deity *(nơi thờ cúng bao gồm một công trình kiến trúc dùng làm nơi thờ phụng một vị thần)*; (oewn-04414570-n)	2
3.	buồng lái: buồng dành cho người lái trong một số phương tiện vận tải, máy móc	cockpit: compartment where the pilot sits while flying the aircraft *(khoang lái, nơi phi công ngồi khi điều khiển máy bay)*; (oewn-03065508-n)	cab: a compartment at the front of a motor vehicle or locomotive where driver sits *(khoang điều khiển ở phía trước của phương tiện cơ giới hoặc đầu máy xe lửa, nơi người lái ngồi)*; (oewn-02934532-n)	3
4.	binh sĩ: binh lính và hạ sĩ quan [nói tổng quát]	rank, rank and file: the ordinary members of an organization (such as the enlisted soldiers of an army) *[các thành viên thông thường của một tổ chức (chẳng hạn*	soldier: an enlisted man or woman who serves in an army *(một người đàn ông hoặc phụ nữ đã nhập ngũ và phục vụ trong quân đội);* (oewn-10641415-n)	3

		như binh sĩ đã nhập ngũ trong quân đội)]; *(oewn-08415136-n)*		
5.	ớt: cây nhỏ cùng họ với cà, hoa trắng, quả chín có màu đỏ hay vàng, vị cay, dùng làm gia vị	capsicum, capsicum pepper plant, pepper: any of various tropical plants of the genus Capsicum bearing peppers *(bất kỳ loài cây nhiệt đới nào thuộc chi Capsicum cho quả ớt);* *(oewn-12921222-n)*	Capsicum annuum longum, cayenne, cayenne pepper, chili pepper, chilli pepper, jalapeno, long pepper: plant bearing very hot and finely tapering long peppers; usually red *(cây cho quả ớt dài, rất cay và thuôn nhọn; thường có màu đỏ);(oewn-12921747-n)*	3

Trước hết, các định nghĩa mang tính chuyên ngành thường khó được mô hình hiểu đúng do các vector ngữ nghĩa (biểu diễn số hóa của từ hoặc cụm từ trong không gian nhiều chiều) được huấn luyện chủ yếu trên ngữ liệu tổng quát, chưa được điều chỉnh hoặc tinh chỉnh (fine-tune) cho các miền tri thức chuyên biệt. Do đó, với những khái niệm như "góc tù", mô hình gặp khó khăn trong việc nhận diện chính xác các thông số và giới hạn (lớn hơn 90° và nhỏ hơn 180°), dẫn đến việc gán nhầm sang một nhãn có nghĩa hoàn toàn ngược lại.

Tiếp đến, những khái niệm gắn với bối cảnh văn hóa cũng dễ bị nhầm lẫn. Ví dụ, từ "chùa" trong ngữ liệu chuẩn được gán là "pagoda" (loại kiến trúc tôn giáo đặc trưng ở châu Á) nhưng mô hình lại chọn "temple" chỉ khái niệm bao quát hơn, bao hàm cả "pagoda". Nguyên nhân là trong WordNet, "temple" là thượng danh (hypernym) của "pagoda", nên khi thiếu thông tin văn hóa đặc thù, mô hình có xu hướng ưu tiên nghĩa rộng hơn để giảm rủi ro sai lệch.

Bên cạnh đó, định nghĩa trong tiếng Việt thường có phạm vi nghĩa rộng hơn hoặc gộp nhiều nghĩa mà trong WordNet lại được tách ra thành những nghĩa riêng biệt. Chẳng hạn, "buồng lái" trong tiếng Việt chỉ khoang dành cho người điều khiển nhiều loại phương tiện khác nhau, không phân biệt máy bay, ô tô, tàu hỏa hay máy móc. Trong WordNet, khái niệm này được chia thành "cockpit" (oewn-03065508-n) với nghĩa là khoang lái máy bay và "cab" (oewn-02934532-n) với nghĩa là khoang điều khiển ở phía trước phương tiện cơ giới hoặc đầu máy xe lửa. Sự khác biệt này khiến mô hình gặp khó khăn khi chọn nhãn.

Cuối cùng, một số trường hợp còn cho thấy sự không tương ứng về phạm vi khái niệm khi gán nhãn, chẳng hạn "binh sĩ" với nghĩa tập thể lại bị gán sang nghĩa cá thể là "soldier", hoặc "ớt" chỉ chung toàn bộ loài bị thu hẹp thành một giống cụ thể là "Capsicum annuum longum", qua đó cho thấy mô hình có xu hướng thiên lệch về mức độ khái quát.

5. KẾT LUẬN

Nghiên cứu này đã xây dựng thành công bộ ngữ liệu vàng gồm 880 mục từ danh từ tiếng Việt, được gán nhãn ngữ nghĩa chính xác theo synset của *Open English WordNet* thông qua phương pháp gián tiếp, sử dụng *Từ điển OALD song ngữ* làm cầu nối. Bộ ngữ liệu này đóng vai trò là tập ngữ liệu chuẩn dùng để đánh giá hiệu quả của mô hình ngôn ngữ lớn (LLM) trong việc gán nhãn ngữ nghĩa tự động cho tiếng Việt. Kết quả thực nghiệm cho thấy, LLM đạt tỷ lệ gán nhãn chính xác 82,39%, thể hiện khả năng xử lý và suy luận ngữ nghĩa tốt, xác định đúng synset dù chỉ dựa trên hai nguồn ngữ liệu chính là *Từ điển tiếng Việt* và *Open English WordNet*. Ngay cả ở những trường hợp không khớp hoàn toàn với nhãn chuẩn, phần lớn kết quả (136/155 định nghĩa) đạt mức tương đồng ngữ nghĩa cao. Những kết quả đạt được khẳng định tiềm năng ứng dụng của LLM trong gán nhãn ngữ nghĩa tự động cho tiếng Việt. Đồng thời, nghiên cứu cũng cho thấy các sai lệch chủ yếu bắt nguồn từ hạn chế của mô hình khi xử lý các định nghĩa mang tính chuyên ngành, khó nắm bắt nét đặc trưng văn hóa trong khái niệm, và sự khác biệt về mức độ khái quát hoặc phạm vi ngữ nghĩa giữa tiếng Việt và tiếng Anh.

TÀI LIỆU THAM KHẢO

Freihat, A. A., Khalilia, H., Bella, G., & Giunchiglia, F. (2024). Advancing the Arabic WordNet: Elevating Content Quality. In *Proceedings of the 6th Workshop on Open-Source Arabic Corpora and Processing Tools (OSACT) with Shared Tasks on Arabic LLMs Hallucination and Dialect to MSA Machine Translation @ LREC-COLING 2024* (pp. 74–83). ELRA & ICCL.

Hamp, B., & Feldweg, H. (1997). Germanet - a lexical-semantic Net for German. *In Automatic information extraction and building of lexical semantic resources for NLP applications.*

Hoàng Phê (Chủ biên) (tái bản năm 2020). *Từ điển tiếng Việt*. Nhà xuất bản Đà Nẵng – Trung tâm Từ điển học, Hà Nội.

Isahara, H., Bond, F., Uchimoto, K., Utiyama, M., & Kanzaki, K. (2008). Development of the Japanese WordNet. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC).*

Joanna, T. & Dien, D. (2014). *Oxford Advanced Learner's Dictionary 8th Edition (with Vietnamese Translation)*. Youth Publishing House.

Khalilia, H., Otterbacher, J., Bella, G., Noortyani, R., Darma, S., & Giunchiglia, F. (2024). Crowdsourcing Lexical Diversity. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2410.23133

Kholodna, N., Julka, S., Khodadadi, M., Gumus, M. N., & Granitzer, M. (2024). LLMs in the loop: Leveraging large language model annotations for active learning in low-resource languages. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 397-412). Springer Nature Switzerland.

Lam, K. N., & Kalita, J. (2022). Constructing Vietnamese WordNet: A case study. *Computación y Sistemas, 26*(3), 1313-1322.

McCrae, J. P. (2024 version). Open English WordNet. *En-Word.net.* https://en-word.net

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography, 3*(4), 235-244. https://doi.org/10.1093/ijl/3.4.235

Nguyen, T. H., Le, A. C., & Nguyen, V. C. (2024). ViLLM-Eval: A Comprehensive Evaluation Suite for Vietnamese Large Language Models. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2404.11086

Sumanathilaka, T. G. D. K., Micallef, N., & Hough, J. (2024). Can LLMs assist with Ambiguity? A Quantitative Evaluation of various Large Language Models on Word Sense Disambiguation. *ACL Anthology* (pp. 97–108). https://doi.org/10.48550/arXiv.2411.18337

Trương Hoàng Bảo, Nguyễn Ngọc Bình, Lê Thành Nguyên, Lê Thị Nhàn & Đinh Điền (2021). Xây dựng WordNet tiếng Việt tự động bằng ngữ liệu song ngữ. *Kỷ yếu Hội nghị KHCN Quốc gia lần thứ XIV về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR),* 483-490. https://doi.org/10.15625/VAP.2021.0093

Truong, S., Nguyen, D., Nguyen, T., Le, D., Truong, N., Quan, T., & Koyejo, S. (2024). Crossing Linguistic Horizons: Finetuning and Comprehensive Evaluation of Vietnamese Large Language Models. In *Findings of the Association for Computational Linguistics: NAACL 2024*, (pp. 2849–2900). https://doi.org/10.18653/v1/2024.findings-naacl.182

Vossen, P. (2002). WordNet, Euro WordNet and Global WordNet. *Revue française de linguistique appliquée*, *7*(1), 27-38.

Wang, S., & Bond, F. (2013). Building the Chinese Open WordNet (COW): Starting from core synsets. In *Proceedings of the 11th Workshop on Asian Language Resources* (pp. 10-18).

Thông tin tác giả

(1) Phan Thị Mỹ Trang: NCS., Trường Đại học Khoa học xã hội và Nhân văn, ĐHQG-HCM, số 10-12 Đinh Tiên Hoàng, Phường Sài Gòn, TP. HCM; ThS., Trường Đại học Công nghệ Sài Gòn, 180 Cao Lỗ, Phường Chánh Hưng, TP. HCM; e-mail: mytrang779@gmail.com

(2) Đỗ Quốc Trí: Sinh viên ngành Công nghệ thông tin, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP. HCM; e-mail: doquoctri2003@gmail.com

(3) Phan Văn Bá Hải: Sinh viên ngành Công nghệ thông tin, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP. HCM; e-mail: phanvanbahai@gmail.com

(4) Đặng Ngọc Lệ: PGS.TS., Trường Đại học Nguyễn Tất Thành, Phường Xóm Chiếu, TP. HCM; e-mail: ngocle.dang1309@gmail.com

(5) Đinh Điền: PGS. TS., Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP. HCM; e-mail: ddien@fit.hcmus.edu.vn

(6) Trần Thị Minh Phượng: TS., Trường Đại học Khoa học xã hội và Nhân văn, ĐHQG-HCM, số 10-12 Đinh Tiên Hoàng, Phường Sài Gòn, TP. HCM; e-mail: minhphuongtrn@hcmussh.edu.vn

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TẠO SINH THIẾT KẾ NỘI DUNG ÔN LUYỆN THỰC HÀNH TIẾNG VIỆT CHO HỌC SINH LỚP 8

Đỗ Thị Thu Hương[1]

Tóm tắt

Bài viết tổng quan tình hình nghiên cứu ứng dụng trí tuệ trong dạy học nói chung, trong dạy học Ngữ văn nói riêng, từ đó, khẳng định những nghiên cứu về ứng dụng AI trong dạy học chuyên sâu và sáng tạo phần Thực hành tiếng Việt cho học sinh lớp 8 chưa nhiều. Trên cơ sở đó, bài viết đề xuất một số cách ứng dụng trí tuệ nhân tạo để biên soạn tài liệu dạy học Thực hành tiếng Việt cho học sinh lớp 8. Các dạng tài liệu bao gồm: bài ôn tập kiến thức, đề ôn luyện kiểm tra đánh giá thường xuyên, phiếu bài tập. Với mỗi dạng tài liệu, bài viết mô tả chi tiết các bước thực hiện nhằm giúp giáo viên có thể ứng dụng dễ dàng trong khi biên soạn tài liệu. Bài viết cũng lưu ý, những kết quả do AI cung cấp chỉ là những tài liệu tham khảo thứ cấp, thậm chí có thể không chính xác, vì vậy, giáo viên cần chỉnh sửa, bổ sung cho phù hợp để sử dụng trong thực tiễn dạy học.

Từ khoá: trí tuệ nhân tạo, thực hành tiếng Việt, thiết kế, ôn tập, chỉnh sửa

1. Đặt vấn đề

Ứng dụng trí tuệ nhân tạo tạo sinh (AI) trong giáo dục nói chung, trong giảng dạy và học tập nói riêng đã và đang trở thành một xu thế tất yếu, mở ra nhiều tiềm năng mới. Với đặc thù là mô hình ngôn ngữ lớn, trí tuệ nhân tạo tạo sinh (AI) đã chứng tỏ khả năng vượt trội của nó trong việc xử lí thông tin, trả lời câu hỏi, tạo ra nội dung mới trên dữ liệu được huấn luyện, tương tác với người dùng… Môn Ngữ văn, trong đó có phần Thực hành tiếng Việt là một trong những môn học nền tảng, giúp học sinh phát triển năng lực ngôn ngữ, năng lực văn học. Tuy nhiên, việc biên soạn tài liệu dạy học nhằm đa dạng hoá các hoạt động dạy học của giáo viên và học sinh trong môn học này vẫn còn là một thách thức đối với các nhà biên soạn tài liệu truyền thống. Hiện nay, với sự hỗ trợ của trí tuệ nhân tạo, các nhà biên soạn tài liệu có thể tận dụng chúng để xây dựng tài liệu học tập nhằm làm đa dạng, phong phú các hoạt động dạy học của giáo viên và học sinh. Bài viết của chúng tôi đề xuất một số cách thức ứng dụng trí tuệ nhân tạo để biên soạn bài tập, tổng hợp kiến thức, thiết kế đề kiểm tra đánh giá thường xuyên phần Thực hành tiếng Việt cho học sinh lớp 8. Ứng dụng công nghệ AI trong việc biên soạn tài liệu sẽ giúp giáo viên giảm bớt áp lực trong dạy học, đồng thời làm phong phú, đa dạng hoá hình thức tài liệu, giảm bớt sự nhàm chán, đơn điệu trong quá trình dạy học.

2. Tổng quan nghiên cứu

Nghiên cứu về ứng dụng trí tuệ nhân tạo tạo sinh (AI) trong dạy học nói chung, trong dạy học Ngữ văn nói riêng đã thu hút được sự quan tâm của nhiều nhà nghiên cứu. Kể từ khi chatbox ChatGPT ra đời vào tháng 11/2022, cụm từ "ứng dụng AI trong..." chưa bao giờ ngừng

[1] PGS.TS, Trường ĐHSP Hà Nội 2; *Email: dothuhuong@hpu2.edu.vn*

"cơn sốt" trên các diễn đàn giáo dục. Có thể khái quát một số hướng nghiên cứu chính về ứng dụng AI trong dạy học như sau:

Hướng thứ nhất, ứng dụng AI trong giảng dạy, học tập và nghiên cứu khoa học ở bậc đại học. Những bài viết theo hướng này có thể kể đến, như: *Ứng dụng ChatGPT trong giảng dạy, nghiên cứu và quản lí ở bậc đại học* của Nguyễn Đức Vượng, Phan Trọng Tiến [12]; *Ứng dụng trí tuệ nhân tạo trong giảng dạy công nghệ thông tin - Hiện trạng và giải pháp* của Nguyễn Văn Diễn, Nguyễn Thuỳ Trang, Tống Duy Khánh [1], *Khám phá tác động của trí tuệ nhân tạo trong giáo dục đại học: Ứng dụng và thách thức* của Nguyễn Lam Ngọc Vi [11],…

Hướng thứ hai, những nghiên cứu về ứng dụng trí tuệ nhân tạo trong dạy học ở trường phổ thông. Đây là hướng được nhiều nhà nghiên cứu quan tâm hơn cả. Những bài viết theo hướng này chủ yếu bàn về ứng dụng trí tuệ nhân tạo trong dạy học các môn học cho học sinh phổ thông. Tiêu biểu có thể kể đến các tác giả: Đặng Thị Thuỳ Dung [2], Đỗ Thị Tố Như [7]… Trong giảng dạy Ngữ văn có thể kể đến bài viết của Trần Thị Hạnh Phương, Châu Thị Kim Ngân [8], Trần Đức Huyền, Tạ Hương Nhuận [5]…

Một trong những lĩnh vực khai thác tiềm năng của AI trong dạy học là giảng dạy ngoại ngữ. Có thể thấy nội dung này qua một số bài viết của Đỗ Thị Hồng Hà, Nguyễn Thuý Ngọc [3,4], Phan Thị Ngọc Lệ [6], Trịnh Bích Thuỷ [10]…

Như vậy, có thể thấy, ứng dụng trí tuệ nhân tạo trong dạy học nói chung, trong dạy học chuyên sâu các môn học ở trường phổ thông đã thu hút sự quan tâm của giáo viên cũng như cảu các nhà nghiên cứu. Tuy nhiên, ứng dụng AI trong dạy học tiếng Việt để thiết kế các bài tập, ôn luyện chuyên sâu và sáng tạo trong dạy học Thực hành tiếng Việt cho học sinh THCS, đặc biệt là học sinh lớp 8 thì hầu như vẫn còn ít được khai thác. Nhằm lấp đầy khoảng trống này, bài viết của chúng tôi đề xuất vận dụng AI trong dạy học ôn luyện thực hành tiếng Việt cho học sinh lớp 8.

3. Nội dung nghiên cứu

3.1 Nội dung dạy học tiếng Việt lớp 8 theo chương trình GDPT môn Ngữ văn 2018

Chương trình GDPT môn Ngữ văn lớp 8 quy định một số nội dung dạy học Thực hành tiếng Việt như sau:

- Nghĩa của một số thành ngữ và tục ngữ tương đối thông dụng

- Sắc thái nghĩa của từ ngữ và việc lựa chọn từ ngữ

- Từ tượng hình và từ tượng thanh: đặc điểm và tác dụng

- Nghĩa của một số yếu tố Hán Việt thông dụng (ví dụ: *vô, hữu*) và nghĩa của những từ có yếu tố Hán Việt đó (ví dụ: *vô tư, vô hình, hữu quan, hữu hạn*)

- Trợ từ, thán từ: đặc điểm và chức năng

- Thành phần biệt lập trong câu: đặc điểm và chức năng

- Câu kể, câu hỏi, câu khiến, câu cảm; câu khẳng định và câu phủ định: đặc điểm và chức năng

- Biện pháp tu từ đảo ngữ, câu hỏi tu từ: đặc điểm và tác dụng

- Nghĩa tường minh và nghĩa hàm ẩn của câu

- Các đoạn văn diễn dịch, quy nạp, song song, phối hợp: đặc điểm và chức năng

- Kiểu văn bản và thể loại: Văn bản tự sự, Văn bản biểu cảm, văn bản nghị luận, văn bản thông tin.

- Từ ngữ toàn dân và từ ngữ địa phương: chức năng và giá trị

- Biệt ngữ xã hội: chức năng và giá trị

- Phương tiện giao tiếp phi ngôn ngữ: hình ảnh, số liệu, biểu đồ…

Mỗi nội dung dạy học nói trên được cụ thể hoá thành các bài học trong từng bộ sách. Nội dung của mỗi bài học trong các bộ sách khác nhau cũng khác nhau, nhưng yêu cầu cần đạt thì thống nhất theo quy định của chương trình. Vì vậy, các cách thức ứng dụng AI để dạy học từng nội dung nói trên có thể áp dụng cho cả ba bộ sách.

3.2 Một số ứng dụng AI trong thiết kế nội dung ôn luyện Thực hành tiếng Việt lớp 8

3.2.1 Ứng dụng AI để tổng hợp kiến thức Thực hành tiếng Việt lớp 8

Trong chương trình Ngữ văn lớp 8, kiến thức lí thuyết về tiếng Việt không nhiều, chủ yếu được trình bày dưới dạng bài tập nhằm giúp học sinh thực hành sử dụng tiếng Việt. Tuy nhiên, trước khi làm các bài tập tiếng Việt, mục *Tri thức ngữ văn* trong các bộ sách đều trình bày kiến thức về các đơn vị ngôn ngữ được nêu trong bài học. Học sinh cần ghi nhớ kiến thức về các đơn vị ngôn ngữ đó, từ đó, vận dụng để giải quyết các bài tập. Vì vậy, để giúp học sinh ghi nhớ những nội dung cốt lõi của bài học, sau khi kết thúc bài học, giáo viên cần tổng hợp lại kiến thức tiếng Việt một cách ngắn gọn, dễ nhớ, dễ thuộc. Để tổng hợp kiến thức một cách nhanh chóng, không mất nhiều thời gian, giáo viên có thể sử dụng ứng dụng Gemini. Các bước thực hiện như sau:

Bước 1. Chuẩn bị nội dung

- Xác định chủ đề cần ôn tập cho học sinh. Chẳng hạn: Ôn tập bài *Các biện pháp tu từ đảo ngữ, câu hỏi tu từ, từ tượng hình, từ tượng thanh.*

- Liệt kê các khía cạnh nội dung của bài học. Ví dụ: *đảo ngữ, câu hỏi tu từ, từ tượng hình, từ tượng thanh.*

- Tóm tắt ngắn gọn các nội dung cần trình bày trong mỗi biện pháp. Với mỗi biện pháp, nội dung trình bày có thể chia thành 2 ý: khái niệm và tác dụng.

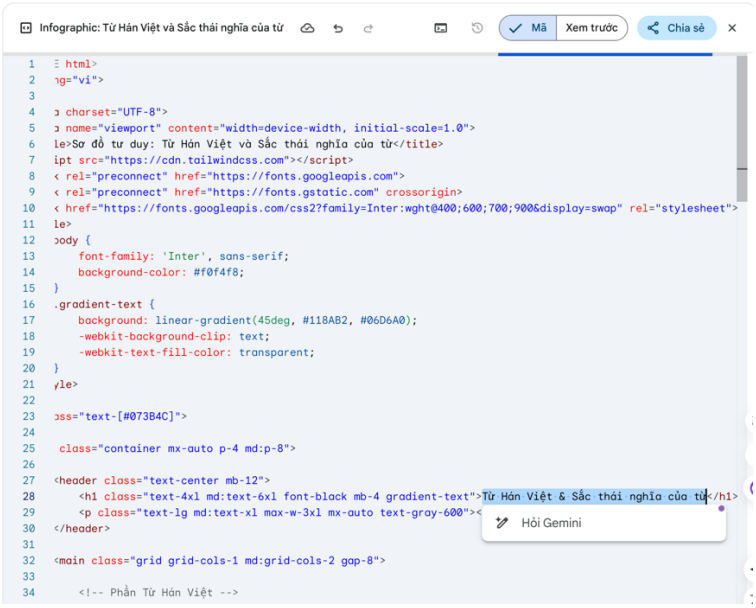
Bước 2. Viết câu lệnh

Câu lệnh cần đảm bảo các yêu cầu như nêu rõ vai trò mà Gemini cần thực hiện, nhiệm vụ, nội dung, yêu cầu về kết quả… Ví dụ: Với vai trò là một giáo viên dạy tiếng Việt cho học sinh lớp 8, bạn hãy thiết kế phần tóm tắt kiến thức về các biện pháp tu từ đảo ngữ, câu hỏi tu từ, từ tượng hình, từ tượng thanh trong sách Ngữ văn 8 tập 2, bộ sách Cánh diều. Nội dung tóm tắt cần ngắn gọn, dễ hiểu, dễ ghi nhớ. Kiến thức được trình bày dưới dạng đồ hoạ.

Nếu có hình ảnh minh hoạ, giáo viên có thể tải hình ảnh lên và bổ sung câu lệnh: dựa vào hình ảnh dưới đây, bạn hãy tạo đồ hoạ… Việc cung cấp hình ảnh minh hoạ như vậy sẽ giúp Gemini hiểu rõ hơn về yêu cầu của giáo viên và tạo đồ hoạ theo đúng mong muốn của giáo viên.

Bước 3. Chỉnh sửa và tuỳ chỉnh đồ hoạ

Câu trả lời do GEMINI cung cấp bao gồm một bản mã code và bản đồ hoạ xem trước. Để có đồ hoạ đáp ứng yêu cầu ghi nhớ kiến thức, giáo viên cần xem vào chỉnh sửa bảng mã



Ảnh minh hoạ màn hình câu trả lời của Gemini

Dựa vào kết quả đồ hoạ do Gemini cung cấp, giáo viên cần chỉnh sửa để hoàn thiện infographic. Trước hết, giáo viên cần kiểm tra các định nghĩa, các ví dụ xem đã chính xác hay chưa. Nếu cần sửa chữa, giáo viên có thể chỉnh sửa trực tiếp trong code (Lưu ý: chỉ chỉnh sửa các định nghĩa, ví dụ). Sau khi chỉnh sửa mã code, giáo viên có thể xem lại đồ hoạ. Có thể chỉnh sửa đến khi đạt được kết quả mong muốn.

Bước 4. Lưu và sử dụng

Sau khi hoàn tất việc chỉnh sửa, giáo viên có thể sao chép mã code và lưu lại để sử dụng cho bài giảng hoặc in ấn.

Dưới đây là một số ví dụ minh hoạ.

Ví dụ (1): Bài 7. *Thơ Đường luật* trong sách Ngữ văn 8, tập 2, bộ sách Cánh diều (trang 43-45). Phần Thực hành tiếng Việt bài Các biện pháp tu từ: Đảo ngữ, câu hỏi tu từ, từ tượng hình, từ tượng thanh.

Ví dụ (2). Bài 4. *Tiếng cười trào phúng trong thơ*, sách Ngữ văn 8 tập 1, bộ Kết nối tri thức với cuộc sống (trang 84-86). Phần Thực hành tiếng Việt gồm hai nội dung: Nghĩa của một số từ, thành ngữ của từ Hán Việt và Sắc thái nghĩa của từ ngữ và việc lựa chọn từ ngữ.

Bên cạnh ứng dụng Gemini, giáo viên cũng có thể sử dụng các phần mềm Edraw Mindmap, The Brain để thiết kế sơ đồ tư duy tổng hợp các kiến thức sau mỗi bài học. Cách tổng hợp kiến thức theo hình thức này giúp học sinh dễ theo dõi mạch kiến thức, đồng thời ghi nhớ bài học một cách dễ dàng, nhanh chóng. Chẳng hạn, Bài 9. *Hôm nay và ngày mai*, sách Ngữ văn 8 tập 2, bộ sách Kết nối tri thức với cuộc sống. Mục tri thức Ngữ văn trong bài học này trình bày ba nội dung lớn: Văn bản thông tin, bao gồm: Mục đích của văn bản thông tin, mối quan hệ giữa thông tin khách quan và ý kiến chủ quan, văn bản giải thích một hiện tượng tự nhiên, văn bản giới thiệu một bộ phim và câu phân loại theo mục đích nói (Câu hỏi, câu khiến, câu nghi vấn, câu cảm, câu kể), câu phủ định và câu khẳng định. Với một lượng kiến thức tiếng Việt nhiều như vậy, giáo viên có thể sử dụng sơ đồ tư duy để giúp các em hệ thống hoá và ghi nhớ kiến thức. Các bước tiến hành như sau:

Bước 1. Chuẩn bị nội dung bài học cần tổng kết. Để đưa kiến thức bài Thực hành tiếng Việt vào sơ đồ tư duy, giáo viên cần tóm lược nội dung bài học bằng các từ khoá. Các đặc điểm, chức năng của mỗi đơn vị ngôn ngữ trong bài học cần được trình bày ngắn gọn, xúc tích để phù hợp với sơ đồ tư duy.

Bước 2. Sau khi đã chuẩn bị nội dung bài học, giáo viên truy cập vào ứng dụng https://www.edrawmind.com (Lưu ý, nếu sử dụng Macbook thì ứng dụng này đã được tích hợp sẵn trong máy). Chọn New Mindmap để bắt đầu tạo sơ đồ tư duy. Tuỳ thuộc vào mục đích trình bày, giáo viên có thể chọn các loại sơ đồ khác nhau, chẳng hạn sơ đồ xuyên tâm, sơ đồ bên phải, sơ đồ xương cá…

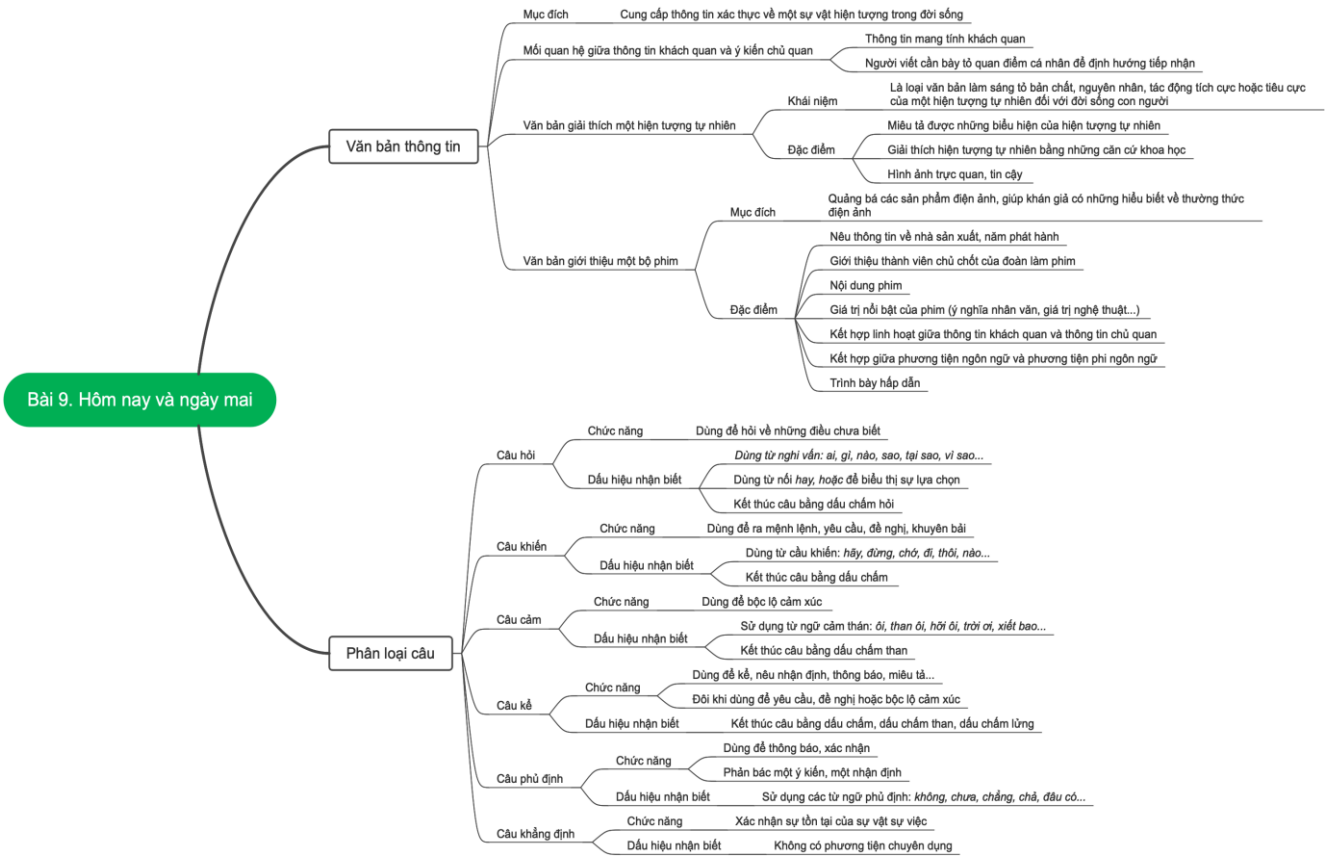
Bước 3. Điền các thông tin của bài học vào sơ đồ. Trung tâm của sơ đồ tư duy là tên bài học, các nhánh cấp 1 là các đơn vị ngôn ngữ được trình bày trong bài, nhánh cấp 2 là chức năng, đặc điểm của các đơn vị ngôn ngữ, nhánh cấp 3 là các nội dung chi tiết của nhánh cấp 2. Ví dụ: Bài 9. *Hôm nay và ngày mai* có hai vấn đề lớn là văn bản thông tin và phân loại câu, hai vấn đề này tương ứng với hai nhánh cấp 1. Mỗi vấn đề lại chứa đựng những nội dung cụ thể, chẳng hạn, văn bản thông tin gồm các nội dung: chức năng, mối quan hệ giữa thông tin chủ quan và thông tin khách quan, văn bản giải thích một hiện tượng tự nhiên, văn bản giới thiệu một bộ phim, những nội dung này tương ứng với nhánh cấp 2. Với mỗi nội dung lại tiếp tục triển khai thành các ý chi tiết, cụ thể hơn (chẳng hạn: mục đích và chức năng), tương ứng với nhánh cấp 3. Cứ tiếp tục như vậy đến khi không thể chia thành các nhánh nhỏ hơn được nữa.

Sau khi điền đầy đủ thông tin trong sơ đồ, giáo viên có thể tô màu các nhánh để sơ đồ trong sinh động, bắt mắt.

Nếu bài học có nhiều nội dung cần tổng kết, giáo viên có thể tách thành các sơ đồ khác nhau, mỗi sơ đồ trình bày trọn vẹn một đơn vị kiến thức.

Bước 3. Xuất file sơ đồ tư duy để sử dụng.

Dưới đây là ví dụ minh hoạ về sơ đồ tư duy ôn tập tổng kết phần Thực hành tiếng Việt Bài 9. *Hôm nay và ngày mai*, sách Ngữ văn 8 tập 2, sách Kết nối tri thức với cuộc sống.



3.2.2 Ứng dụng AI để thiết kế đề kiểm tra đánh giá thường xuyên phần Thực hành tiếng Việt

Kiểm tra đánh giá thường xuyên là công việc bắt buộc của giáo viên trong khi dạy học. Thông qua kiểm tra đánh giá, giáo viên có thể đưa ra những nhận định, phân tích về tình hình học tập của học sinh, từ đó có những điều chỉnh cần thiết về phương pháp dạy học. Đối với các bài học Thực hành tiếng Việt, giáo viên có thể thiết kế các câu hỏi trắc nghiệm khách quan nhằm kiểm tra thường xuyên mức độ hiểu kiến thức tiếng Việt của các em. Các ứng dụng AI mà giáo viên có thể sử dụng để thiết kế đề kiểm tra đánh giá là ChatGPT, Gemini. Dưới đây là một ví dụ ứng dụng ChatGPT.

Bước 1. Xác định nội dung cần chuẩn bị để kiểm tra đánh giá. Ví dụ: Bài 5. *Những câu chuyện hài*, sách Ngữ văn 8 tập 1, bộ Kết nối tri thức với cuộc sống, phần Thực hành tiếng Việt là *Câu hỏi tu từ.*

Bước 2. Xác định số lượng câu hỏi, các mức độ nhận thức khi thiết kế câu hỏi kiểm tra. Đối với bài kiểm tra thường xuyên, các câu hỏi cũng cần đáp ứng đầy đủ các mức độ nhận thức, từ nhận biết, thông hiểu đến vận dụng.

Bước 3. Viết câu lệnh. Đây là một câu lệnh gợi ý: Đóng vai trò là một giáo viên dạy tiếng Việt lớp 8, hãy thiết kế 05 câu hỏi trắc nghiệm khách quan về câu hỏi tu từ, trong đó có

02 câu mức độ nhận biết, 02 câu mức độ thông hiểu, 01 câu mức độ vận dụng. Mỗi câu hỏi có 4 phương án trả lời, các phương án có độ dài, độ khó tương đương nhau. Có đáp án đi kèm. Kết quả trình bày dưới dạng bảng, trong đó có cột câu hỏi, cột đáp án.

Bước 4. Sau khi có kết quả do AI cung cấp, giáo viên cần chỉnh sửa các câu hỏi để có thể sử dụng được trong kiểm tra đánh giá. Lưu ý: Giáo viên không nên copy nguyên vẹn các kết quả của AI để sử dụng mà cần có sự điều chỉnh, sửa chữa cho phù hợp với nội dung bài học.

Dưới đây là kết quả do ChatGPT cung cấp sau hai lần chỉnh sửa:

Câu hỏi	**Đáp án đúng**
Câu 1 (Nhận biết): Câu hỏi tu từ là gì? A. Dùng để người nghe trả lời trực tiếp. B. Dùng để bộc lộ cảm xúc, nhấn mạnh ý, không cần trả lời. C. Chỉ dùng trong văn nghị luận. D. Chỉ xuất hiện trong hội thoại đời thường.	B
Câu 2 (Nhận biết): Câu nào dưới đây là câu hỏi tu từ? A. Hôm nay bạn trực nhật chưa? B. Ai mà không yêu quê hương mình? C. Bạn có thể cho tôi mượn bút không? D. Khi nào kiểm tra giữa kì?	B
Câu 3 (Thông hiểu): Tác dụng thường gặp của câu hỏi tu từ trong văn bản là: A. Làm câu văn ngắn gọn hơn. B. Tạo nhịp điệu và bộc lộ cảm xúc, nhấn mạnh ý. C. Thay toàn bộ câu trần thuật. D. Bắt buộc người đọc phải trả lời.	B
Câu 4 (Thông hiểu – ngữ liệu Thâm Tâm, *Tống biệt hành*): "Người đi ừ nhỉ người đi thực? / Mẹ thà coi như chiếc lá bay." Câu "Người đi ừ nhỉ người đi thực?" là câu hỏi tu từ nhằm mục đích gì? A. Xác nhận thông tin rõ ràng về người đi, chờ câu trả lời. B. Bộc lộ sự bàng hoàng, đau xót trước cuộc chia ly. C. Mở đầu cuộc đối thoại và mời người khác đáp lời.	B

Câu hỏi	Đáp án đúng
D. Yêu cầu người ra đi trình bày lý do chia tay.	
Câu 5 (Vận dụng – ngữ liệu Tố Hữu, *Việt Bắc*, 1954): "Mình về mình có nhớ ta / Mười lăm năm ấy thiết tha mặn nồng?"	
Câu hỏi "Mười lăm năm ấy thiết tha mặn nồng?" là câu hỏi tu từ nhằm mục đích gì?	
A. Nhằm khẳng định tình nghĩa gắn bó thủy chung giữa cán bộ và đồng bào.	A
B. Nhằm yêu cầu người về trả lời chính xác.	
C. Nhằm thay thế một câu chào tạm biệt.	
D. Nhằm giảm nhẹ cảm xúc chia tay.	

Có thể nhận thấy, các câu hỏi trắc nghiệm do ChatGPT cung cấp không thể sử dụng ngay được mà cần chỉnh sửa để có bộ câu hỏi tốt hơn.

Dưới đây là kết quả sau khi chỉnh sửa:

Câu hỏi	Đáp án đúng
Câu 1 (Nhận biết): Câu hỏi tu từ là kiểu câu:	
A. yêu cầu người nghe trả lời trực tiếp.	
B. dùng để bộc lộ cảm xúc, nhấn mạnh ý, không cần trả lời.	B
C. chỉ dùng trong văn nghị luận.	
D. chỉ xuất hiện trong hội thoại đời thường.	
Câu 2 (Nhận biết): Câu nào sau đây là câu hỏi tu từ?	
A. Hôm nay bạn trực nhật chưa?	
B. Bạn có thể cho tôi mượn bút không?	C
C. Ớt nào là ớt chẳng cay?	
D. Vì sao bạn đến muộn?	
Câu 3 (Thông hiểu): Dòng nào nêu đúng tác dụng của câu hỏi tu từ?	
A. Làm câu văn ngắn gọn hơn.	
B. Bắt buộc người đọc phải trả lời.	D
C. Thay toàn bộ câu trần thuật.	
D. Tạo nhịp điệu và bộc lộ cảm xúc, nhấn mạnh ý.	

Câu hỏi	Đáp án đúng
Câu 4 (Thông hiểu): Câu "Người đi ừ nhỉ người đi thực?" là câu hỏi tu từ nhằm: A. Xác nhận thông tin rõ ràng về người đi, chờ câu trả lời. B. Bộc lộ sự bàng hoàng, đau xót trước cuộc chia li. C. Mở đầu cuộc đối thoại và mời người khác đáp lời. D. Yêu cầu người ra đi trình bày lí do chia tay.	B
Câu 5 (Vận dụng): Câu hỏi tu từ trong câu thơ sau có tác dụng gì? *Vì sao Trái đất nặng ân tình?/Nhắc mãi tên người Hồ Chí Minh* (Tố Hữu) A. Khẳng định tình cảm cao đẹp, thiêng liêng mà nhân dân thế giới dành cho Bác. B. Khẳng định nhân dân thế giới sẽ hát mãi các bài hát về Chủ tịch Hồ Chí Minh. C. Yêu cầu người đọc trả lời chính xác về trọng lượng của Trái đất. D. Nhằm giảm nhẹ công lao, sự đóng góp của Bác dành cho nhân dân thế giới.	A

Bên cạnh các câu hỏi trắc nghiệm, giáo viên cũng có thể tự thiết kế các câu hỏi tự luận nhằm làm đa dạng hoá các câu hỏi trong bài kiểm tra.

3.2.4 Ứng dụng AI thiết kế phiếu học tập trong dạy học Thực hành tiếng Việt

Nhằm giúp học sinh ôn tập, củng cố các tri thức trong bài Thực hành tiếng Việt, giáo viên có thể tự thiết kế các phiếu học tập nhờ các ứng dụng của AI. Một trong những ứng dụng để thiết kế phiếu bài tập hiệu quả là ứng dụng diffit.com.

Bước 1. Chuẩn bị nội dung cho Phiếu học tập. Ví dụ: Bài *Thành phần biệt lập*, sách Ngữ văn 8, tập 2 – Bộ Kết nối tri thức với cuộc sống, trang 66-69.

Bước 2. Đăng nhập vào ứng dụng diffit.com. Tại giao diện của diffit, chọn Teachers Start Here, nhập nội dung cần thiết kế phiếu bài tập.

Bước 3. Chỉnh sửa kết quả do diffit cung cấp

Bước 4. Xuất file

Kết quả nhận được như sau:

PHIẾU HỌC TẬP

Bài: Thành phần biệt lập (Ngữ văn 8, tập 2 – Bộ Kết nối tri thức với cuộc sống)

Thời gian làm việc: 5–10 phút

I. Mục tiêu

- Nhận diện được các thành phần biệt lập trong câu.
- Phân biệt và nêu tác dụng của từng loại thành phần biệt lập.
- Vận dụng để viết câu có chứa thành phần biệt lập.

II. Nhiệm vụ

1. Nhận diện

Đọc các câu sau và gạch chân thành phần biệt lập (nếu có), sau đó ghi loại thành phần vào cột bên:

Câu văn	Thành phần biệt lập	Loại thành phần biệt lập
a) Chà, hôm nay bạn làm bài nhanh thế!		
b) Lan ơi, nhớ giữ sức khỏe nhé.		
c) Thú thực, tôi chưa hiểu hết ý nghĩa câu chuyện.		
d) Có lẽ, ngày mai trời sẽ đẹp.		

2. Phân tích – thảo luận

a) Các thành phần biệt lập trong bảng trên thuộc những loại nào?
b) Chúng có tác dụng gì đối với cách diễn đạt và cảm xúc trong câu?
3. Vận dụng

- Em hãy viết 02 câu có sử dụng thành phần biệt lập (mỗi câu một loại khác nhau).

- Gạch chân thành phần biệt lập và ghi chú loại ở cuối câu.

……
……

III. Ghi nhớ

- Thành phần biệt lập gồm các thành phần: tình thái, cảm thán, gọi – đáp, phụ chú.
- Vai trò: bổ sung sắc thái tình cảm, thái độ, cách nhìn nhận của người nói, giúp lời văn sinh động, giàu biểu cảm.

Ngoài ứng dụng diffit, giáo viên cũng có thể thiết kế phiếu bài tập trên ứng dụng canva. Dưới đây là một ví dụ minh hoạ do Canva thiết kế (Lưu ý: Phần nội dung trong phiếu do người viết chỉnh sửa).

PHIẾU BÀI TẬP
NGHĨA TƯỜNG MINH VÀ NGHĨA HÀM ẨN

Họ và tên: .....

Lớp:.....

✦ Mục tiêu

Nhận biết được nghĩa tường minh, nghĩa hàm ẩn trong văn bản.

Vận dụng kiến thức về nghĩa tường minh và nghĩa hàm ẩn để phân tích câu văn, đoạn thơ.

✦ Nhiệm vụ học tập

Bài tập 1:

Đọc các câu sau và điền vào bảng:

Câu văn/đoạn thơ	Nghĩa tường minh	Nghĩa hàm ẩn
a. *Bao giờ chạch đẻ ngọn đa* *Sáo đẻ dưới nước thì ta lấy mình* *Bao giờ rau cải làm đình* *Gỗ lim làm ghém thì mình lấy ta*	…	…
b. *Em tưởng nước giếng sâu* *Em nối sợi gầu dài* *Ai ngờ giếng cạn* *Em tiếc hoài sợi dây*	…	…
c. *Không muốn là đe hãy làm búa* *Không muốn là than, hãy là lửa đỏ*	…	…

Bài tập 2:

Hãy viết một đoạn văn (5–7 dòng) phân tích tác dụng của việc sử dụng nghĩa tường minh và nghĩa hàm ẩn trong bài ca dao:

Bao giờ chạch đẻ ngọn đa

Sáo đẻ dưới nước thì ta lấy mình

Bao giờ rau cải làm đình

Gỗ lim làm ghém thì mình lấy ta

……
……

……

Bài tập 3:

Hãy tạo một đoạn hội thoại ngắn (3 – 5 câu) trong đó có câu biểu thị nghĩa tường minh và có câu biểu thị nghĩa hàm ẩn.

Hội thoại:

……
……

……

📌 Ghi chú: Hoàn thành phiếu này trong …... phút.

4. Kết luận

Ứng dụng AI trong dạy học Ngữ văn nói chung, trong dạy học tiếng Việt nói riêng là một xu thế tất yếu trong giáo dục Việt Nam thời đại 4.0. Các ứng dụng của AI hỗ trợ tích cực cho giáo viên trong việc biên soạn tài liệu dạy học hay tìm kiếm các phương pháp dạy học hiện đại, giúp giáo viên giảm tải "gánh nặng" công việc dạy học… Tuy nhiên, trong quá trình sử dụng, giáo viên cũng cần lưu ý rằng, không nên sao chép hoàn toàn các kết quả do AI cung cấp, mà chỉ xem đó như một nguồn tài liệu để tham khảo. Bởi lẽ, những thông tin do AI cung cấp có thể thiếu chính xác, đặc biệt là những vấn đề chuyên ngành. Giáo viên phải là người chịu trách nhiệm về nội dung của tài liệu. Sử dụng AI trong dạy học là một xu hướng tất yếu, nhưng người sử dụng cũng cần thông minh hơn AI, để AI trở thành công cụ hữu ích phục vụ con người chứ không lệ thuộc vào AI.

TÀI LIỆU THAM KHẢO

1. Nguyễn Văn Diễn, Nguyễn Thuỳ Trang, Tống Duy Khánh (2024), "Ứng dụng trí tuệ nhân tạo trong giảng dạy công nghệ thông tin - Hiện trạng và giải pháp", Tạp chí *Khoa học* ĐH Thành Đô, số 3, tập 4, p97-102),

2. Đặng Thị Thuỳ Dung (2023), *Ứng dụng trí tuệ nhân tạo trong dạy học lịch sử theo hướng phát triển năng lực giải quyết vấn đề và sáng tạo cho học sinh THPT*, Kỉ yếu Hội thảo khoa học quốc gia cán bộ trẻ các trường ĐHSP lần thứ X

3. Đỗ Thị Hồng Hà, Nguyễn Thuý Ngọc (2024), *Trí tuệ nhân tạo và việc học qua máy trong dạy học tiếng Anh trực tuyến*, Kỉ yếu hội thảo khoa học Ngôn ngữ học tính toán: Những xu hướng mới, triển vọng và thách thức, 2024, Nxb Tri thức

4. Nguyễn Thuý Ngọc, Đỗ Thị Hồng Hà (2024), *Cải thiện giao tiếp bằng tiếng Anh thông qua mạng trí tuệ nhân tạo của nhóm tác giả*, Kỉ yếu hội thảo khoa học Ngôn ngữ học tính toán: Những xu hướng mới, triển vọng và thách thức, Nxb Tri thức

5. Trần Đức Huyền, Tạ Hương Nhuận, *Ứng dụng trí tuệ nhân tạo (AI) trong dạy học Ngữ văn*, Nxb ĐHQG Hà Nội

6. Phan Thị Ngọc Lệ (2024), *Vai trò của giáo viên khi ứng dụng trí tuệ nhân tạo vào trong lớp học ngoại ngữ*, Kỉ yếu hội thảo khoa học Ngôn ngữ học tính toán: Những xu hướng mới, triển vọng và thách thức, Nxb Tri thức

7. Đỗ Thị Tố Như (2025), "Khai thác và sử dụng một số công cụ trí tuệ nhân tạo (AI) trong dạy học môn Khoa học tự nhiên", Tạp chí Giáo dục tập 25, số 5 năm 2025.

8. Trần Thị Hạnh Phương, Châu Thị Kim Ngân (2024), "Vận dụng trí tuệ nhân tạo (AI) trong hoạt động dạy học Ngữ văn ở trường phổ thông", Tạp chí Thiết bị Giáo dục số 1, tập 308, 2024, p176-178

9. Nguyễn Thị Hồng Thắm, *Ứng dụng trí tuệ nhân tạo trong lớp học đảo ngược – kinh nghiệm quốc tế và tiềm năng ứng dụng ở trường ĐH Thủ Dầu Một*, Kỉ yếu Hội thảo khoa học quốc tế Giáo dục đại học đáp ứng nhu cầu nguồn nhân lực và phát triển kinh tế - xã hội địa phương - Thực tiễn các nước và gợi ý cho Việt Nam.

10. Trịnh Bích Thuỷ (2024), *Ứng dụng trí tuệ nhân tạo trong giảng dạy tiếng Pháp cho người học mới bắt đầu*, Kỉ yếu hội thảo khoa học Ngôn ngữ học tính toán: Những xu hướng mới, triển vọng và thách thức, Nxb Tri thức

11. Nguyễn Lam Ngọc Vi (2023), "Khám phá tác động của trí tuệ nhân tạo trong giáo dục đại học: Ứng dụng và thách thức", Tạp chí Khoa học, Trường ĐH Quốc tế Hồng Bàng.

12. Nguyễn Đức Vượng, Phan Trọng Tiến (2023), "Ứng dụng ChatGPT trong giảng dạy, nghiên cứu và quản lí ở bậc đại học", Tạp chí Khoa học ĐH Tân Trào, số 9, tập 3, p171-182.

Thông tin người gửi

Họ và tên: Đỗ Thị Thu Hương

Học hàm học vị: PGS.TS

Cơ quan công tác: Khoa Ngữ văn, Trường ĐHSP Hà Nội 2, Xuân Hoà, Phúc Yên, Phú Thọ

Email: dothuhuong@hpu2.edu.vn

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TRONG DẠY HỌC VIẾT VĂN MIÊU TẢ PHONG CẢNH CHO HỌC SINH LỚP 5

Đỗ Lam Ngọc[1]

TÓM TẮT

Trong chương trình môn Tiếng Việt ở Tiểu học, kĩ năng viết giữ vai trò nền tảng trong việc phát triển năng lực ngôn ngữ của học sinh. Trong đó, dạy học viết văn miêu tả phong cảnh là một nội dung trọng tâm ở lớp 5. Tuy nhiên, không ít học sinh lớp 5 còn gặp khó khăn trong việc triển khai bố cục bài viết, sử dụng từ ngữ miêu tả một cách linh hoạt. Trong bối cảnh chuyển đổi số, sự phát triển nhanh chóng của trí tuệ nhân tạo đang tạo ra những cơ hội mới trong việc hỗ trợ học sinh rèn luyện kĩ năng viết. Từ thực tiễn đó, bài viết đề xuất một số định hướng tích hợp AI vào quá trình dạy học viết văn miêu tả phong cảnh cho học sinh lớp 5 theo hướng phát triển phẩm chất, năng lực.

TỪ KHÓA: trí tuệ nhân tạo, văn miêu tả phong cảnh, học sinh lớp 5

1. MỞ ĐẦU

Trong bối cảnh chuyển đổi số diễn ra mạnh mẽ, trí tuệ nhân tạo (AI) ngày càng được ứng dụng rộng rãi trong nhiều lĩnh vực, đặc biệt ở lĩnh vực giáo dục. Việc tích hợp AI trong các bài giảng không chỉ góp phần đổi mới phương pháp dạy học mà còn nâng cao chất lượng và hiệu quả học tập. Trong môn Tiếng Việt ở cấp Tiểu học - giai đoạn hình thành nền tảng ngôn ngữ và tư duy biểu đạt cho học sinh, việc ứng dụng công nghệ, trí tuệ nhân tạo vào trong hoạt động dạy học viết văn đang dần trở thành xu thế tất yếu. Trong môn Tiếng Việt lớp 5, viết văn miêu tả phong cảnh là một nội dung quan trọng, nhằm giúp học sinh biết quan sát thiên nhiên, cảm nhận vẻ đẹp của môi trường sống và diễn đạt bằng lời văn sinh động, giàu hình ảnh. Tuy nhiên, trên thực tế, nhiều học sinh còn gặp khó khăn trong việc triển khai ý tưởng, lựa chọn từ ngữ gợi tả, sử dụng biện pháp tu từ và cấu trúc đoạn văn logic. Đồng thời, giáo viên cũng thường xuyên đối mặt với thách thức trong việc hướng dẫn, sửa lỗi và phát triển năng lực viết cho từng học sinh trong một thời lượng tiết học hạn chế. Trước yêu cầu đổi mới phương pháp dạy học và tận dụng lợi thế của công nghệ, việc nghiên cứu đề tài **"Ứng dụng trí tuệ nhân tạo trong dạy học viết văn miêu tả phong cảnh cho học sinh lớp 5"** là cần thiết và có ý nghĩa thực tiễn. Đề tài hướng đến việc khai thác tiềm năng của AI trong việc hỗ trợ giáo viên tổ chức hoạt động dạy học, cá nhân hóa quá trình học tập và khơi gợi hứng thú, sáng tạo của học sinh trong viết văn miêu tả.

2. NỘI DUNG NGHIÊN CỨU

2.1. Một số vấn đề về văn miêu tả và dạy học viết văn miêu tả

[1] Ths, Trường ĐHSP Hà Nội 2; Email: dolamngoc@hpu2.edu.vn

2.1.1. Khái niệm văn miêu tả, văn miêu tả phong cảnh

Theo *Từ điển Tiếng Việt* của Hoàng Phê, khái niệm "miêu tả" được định nghĩa như sau: "*Miêu tả là dùng ngôn ngữ hoặc một phương tiện nghệ thuật nào đó làm cho người khác có thể hình dung được cụ thể sự vật, sự việc hoặc thế giới nội tâm của con người.*" [7-tr.162]

Các tác giả Nguyễn Quang Ninh – Đào Ngọc trong cuốn *Tiếng Việt thực hành* (Tài liệu đào tạo giáo viên tiểu học) nhận định: "*Văn miêu tả là loại văn thể hiện sự vật, sự việc, con người, cảnh vật... một cách sinh động, cụ thể như nó vốn có trong đời sống. Đây là loại văn giàu cảm xúc, giàu trí tưởng tượng và sự đánh giá thẩm mĩ của người viết với đối tượng miêu tả.*" [6-tr.130]

Trong khi đó, Phan Văn Các đã đưa ra định nghĩa "phong cảnh" là: "*Thiên nhiên bày ra trước mắt như sông núi, làng mạc, phố xá…*" [2-tr.311]

Từ các quan niệm trên, có thể hiểu văn miêu tả phong cảnh là kiểu văn sử dụng các phương tiện ngôn ngữ nhằm tái hiệu một cách chi tiết, sinh động giúp người đọc hình dung và cảm nhận được vẻ đẹp của cảnh vật.

2.1.2. Dạy học viết văn miêu tả cho học sinh lớp 5

2.1.2.1. Yêu cầu cần đạt

Trong *Chương trình Giáo dục phổ thông môn Ngữ Văn 2018*, nội dung dạy học viết văn miêu tả được triển khai chủ yếu ở lớp 5 với yêu cầu cần đạt: "*Viết được bài tả người, phong cảnh có sử dụng so sánh, nhân hóa và những từ ngữ gợi tả để làm nổi bật đặc điểm của đối tượng được tả*" [1-tr.39]. Đây được xem là một bước phát triển quan trọng trong việc hình thành và nâng cao năng lực tạo lập văn bản cho học sinh cuối cấp tiểu học. Thông qua việc vận dụng các biện pháp tu từ: so sánh, nhân hóa và từ ngữ gợi tả, học sinh không chỉ rèn luyện khả năng quan sát tinh tế, lựa chọn chi tiết tiêu biểu mà còn phát triển kĩ năng sử dụng ngôn ngữ miêu tả linh hoạt, giàu hình ảnh. Từ đó, tư duy hình tượng và năng lực cảm thụ thẩm mĩ của học sinh từng bước được hình thành và phát triển.

Việc tổ chức các hoạt động dạy học nhằm hướng dẫn học sinh sử dụng hiệu quả các biện pháp tu từ nêu trên trong quá trình viết văn miêu tả cũng góp phần thúc đẩy sự phát triển toàn diện năng lực ngôn ngữ. Qua quá trình luyện tập và thực hành, học sinh dần làm chủ cách thức miêu tả người và cảnh vật một cách sinh động, đồng thời biết cách bộc lộ cảm xúc, thái độ và góc nhìn cá nhân. Đây chính là tiền đề giúp các em nâng cao khả năng sáng tạo trong diễn đạt, phát triển tư duy ngôn ngữ.

2.1.2.2. Các kiểu bài văn miêu tả trong chương trình và SGK Tiếng Việt lớp 5

Trong *Chương trình Giáo dục phổ thông môn Ngữ Văn 2018*, nội dung thực hành viết đã xác lập rõ các kiểu bài thuộc văn miêu tả: "*Viết được bài tả người, phong cảnh có sử dụng so sánh, nhân hóa và những từ ngữ gợi tả để làm nổi bật đặc điểm của đối tượng được tả"* [1-tr.39]. Căn cứ vào định hướng này, có thể khẳng định ở lớp 5, chương trình môn Tiếng Việt chú trọng tập trung vào hai kiểu bài văn miêu tả: văn tả phong cảnh và văn tả người. Việc xác

định rõ hai kiểu bài này không chỉ góp phần cụ thể hóa mục tiêu phát triển năng lực viết cho học sinh tiểu học, mà còn tạo cơ sở cho việc tổ chức các hoạt động dạy học theo hướng phát triển năng lực và phẩm chất người học.

Ba bộ sách giáo khoa Tiếng Việt hiện hành, bao gồm *Kết nối tri thức với cuộc sống, Chân trời sáng tạo* và *Cánh Diều*, đều triển khai nội dung dạy học viết văn miêu tả một cách hệ thống và định hướng rõ ràng. Các bài học được thiết kế dựa trên mục tiêu, nội dung và yêu cầu cần đạt được quy định trong *Chương trình Giáo dục phổ thông môn Ngữ văn năm 2018*, đồng thời thể hiện sự nhất quán trong việc phát triển năng lực ngôn ngữ và tư duy sáng tạo của học sinh tiểu học. Việc xây dựng các dạng bài tập, nhiệm vụ học tập trong từng bộ sách không chỉ đảm bảo tính sư phạm mà còn phản ánh sự tiếp cận chương trình theo hướng phát triển phẩm chất và năng lực người học.

Bảng thống kê các bài dạy về viết bài văn tả phong cảnh trong ba bộ sách Tiếng Việt lớp 5

Bộ sách	Tập	Bài	Tên bài
Kết nối tri thức với cuộc sống	1	9	Tìm hiểu cách viết bài văn tả phong cảnh
	1	10	Tìm hiểu cách viết bài văn tả phong cảnh (tiếp theo)
	1	11	Viết mở bài và kết bài cho bài văn tả phong cảnh
	1	13	Lập dàn ý cho bài văn tả phong cảnh
	1	15	Viết bài văn tả phong cảnh
	2	23	Luyện tập lập dàn ý cho bài văn tả phong cảnh
	2	24	Luyện viết bài văn tả phong cảnh
Chân trời sáng tạo	1	1	Bài văn tả phong cảnh
	1	2	Quan sát, tìm ý cho bài văn tả phong cảnh
	1	3	Lập dàn ý cho bài văn tả phong cảnh
	1	4	Viết đoạn mở bài cho bài văn tả phong cảnh
	1	5	Viết đoạn văn cho bài văn tả phong cảnh
	1	6	Viết đoạn kết cho bài văn tả phong cảnh
	1	7	Viết bài văn tả phong cảnh (Bài viết số 1)
	1	8	Luyện tập quan sát, tìm ý cho bài văn tả phong cảnh
	1	1	Luyện tập lập dàn ý cho bài văn tả phong cảnh
	1	2	Trả bài văn tả phong cảnh (Bài viết số 1)
	1	3	Luyện tập viết đoạn văn cho bài văn tả phong cảnh
	1	4	Viết bài văn tả phong cảnh (Bài viết số 2)
	1	7	Trả bài văn tả phong cảnh (Bài viết số 2)
Cánh Diều	2	11	Luyện tập tả phong cảnh (Cách quan sát)
	2	11	Luyện tập tả phong cảnh (Thực hành quan sát)
	2	12	Luyện tập tả phong cảnh (Tìm ý, lập dàn ý)
	2	12	Luyện tập tả phong cảnh (Viết mở bài)
	2	12	Luyện tập tả phong cảnh (Viết kết bài)
	2	13	Luyện tập tả phong cảnh (Viết thân bài)

	2	13	Luyện tập tả phong cảnh (Viết bài văn)
	2	14	Trả bài văn tả phong cảnh
	2	11	Luyện tập tả phong cảnh (Cách quan sát)
	2	11	Luyện tập tả phong cảnh (Thực hành quan sát)
	2	12	Luyện tập tả phong cảnh (Tìm ý, lập dàn ý)
	2	12	Luyện tập tả phong cảnh (Viết mở bài)
	2	12	Luyện tập tả phong cảnh (Viết kết bài)
	2	13	Luyện tập tả phong cảnh (Viết thân bài)

Tổng thể cho thấy, cả ba bộ sách đều chú trọng rèn luyện cho học sinh khả năng quan sát, lựa chọn chi tiết tiêu biểu và sử dụng ngôn ngữ gợi tả thông qua các kiểu bài văn miêu tả, tiêu biểu là văn miêu tả phong cảnh, văn tả người. Tuy mỗi bộ sách có cách sắp xếp và triển khai bài học khác nhau, song đều hướng tới mục tiêu hình thành và phát triển năng lực viết văn miêu tả cho học sinh tiểu học thông qua các bước luyện tập theo tiến trình: quan sát - lựa chọn chi tiết - lập dàn ý - hoàn thiện bài viết. Cách tiếp cận này góp phần giúp học sinh lớp 5 hình thành tư duy hình tượng, nâng cao năng lực cảm thụ thẩm mĩ và phát triển khả năng biểu đạt ngôn ngữ sinh động, sáng tạo.

2.2. Một số vấn đề về trí tuệ nhân tạo

2.2.1. Khái niệm

Trí tuệ nhân tạo hay trí thông minh nhân tạo (Artificial intelligence – viết tắt là AI), là một thuật ngữ được đề cập từ khá sớm trong các nghiên cứu trên thế giới và ngày càng được quan tâm trong bối cảnh phát triển mạnh mẽ của khoa học - công nghệ hiện đại. Đến hiện nay, vẫn tồn tại nhiều cách tiếp cận và quan niệm khác nhau về khái niệm trí tuệ nhân tạo. Chúng tôi quan niệm: trí tuệ nhân tạo (Artificial Intelligence) có thể được hiểu là một hình thức trí tuệ do con người thiết kế và phát triển, với mục tiêu mô phỏng các quá trình tư duy và hành vi của con người nhằm thực hiện các nhiệm vụ nhất định và đáp ứng hiệu quả các nhu cầu trong đời sống và sản xuất. Trong bối cảnh chuyển đổi số đang diễn ra mạnh mẽ, AI được xem là một trong những công nghệ then chốt góp phần đổi mới căn bản và toàn diện nền giáo dục. Đặc biệt trong giáo dục hiện đại, AI không chỉ hỗ trợ giáo viên trong xây dựng bài giảng sinh động, theo dõi tiến trình học tập mà còn tạo ra những cơ hội học tập linh hoạt, cá nhân hóa nội dung theo năng lực và nhu cầu của từng học sinh. Có thể thấy, trí tuệ nhân tạo (AI) đóng vai trò quan trọng trong việc thúc đẩy chuyển đổi số và nâng cao chất lượng giáo dục toàn diện, hiệu quả và bền vững.

2.2.2. Tiềm năng ứng dụng công cụ AI tạo sinh trong dạy học kĩ năng viết

Trong bối cảnh chuyển đổi số hiện nay, trí tuệ nhân tạo đang dần khẳng định vị thế là một công cụ giàu tiềm năng trong giáo dục, đặc biệt là trong dạy học kĩ năng viết ở cấp tiểu học. Công cụ AI tạo sinh không chỉ hỗ trợ giáo viên trong việc thiết kế các hoạt động học tập sáng tạo, đa dạng hóa nguồn tư liệu giảng dạy mà còn góp phần cá nhân hóa quá trình học của học sinh, từ đó nâng cao hiệu quả rèn luyện kĩ năng viết.

"*Trí tuệ nhân tạo tạo sinh là công nghệ được thiết kế để mô phỏng cuộc trò chuyện với con người thông qua các giao diện trò chuyện như cửa sổ trò chuyện trên trang web, ứng dụng tin nhắn, email, thoại,…* " [3-tr.3]. Sản phẩm trí tuệ nhân tạo tạo sinh tạo ra mang tính tổng hợp và tư duy tương tự như được tạo bởi con người, bao gồm: văn bản, hình ảnh, âm thanh, đồ hoạ 3D,… Chat GPT là một ví dụ điển hình của trí tuệ nhân tạo tạo sinh với khả năng sáng tạo nội dung đa dạng như trả lời bằng văn bản, viết thư, làm thơ, viết lời bài hát… không giới hạn ngôn ngữ hay bất kì chủ đề nào. Ngoài chat GPT, rất nhiều sản phẩm Generative AI hiện nay cũng đảm nhiệm được công việc lập trình, vẽ tranh, làm video, phân tích dữ liệu…

Việc ứng dụng trí tuệ nhân tạo (AI) trong dạy học kĩ năng viết mang lại nhiều lợi ích thiết thực đối với giáo viên tiểu học, góp phần nâng cao chất lượng dạy học. Trước hết, AI có thể hỗ trợ giáo viên trong khâu thiết kế bài giảng thông qua việc đề xuất các đề tài viết đa dạng, phù hợp với nhiều trình độ khác nhau, từ đó giúp giáo viên tổ chức hoạt động học tập theo hướng phân hóa và phát huy tính chủ động, sáng tạo của học sinh. Bên cạnh đó, công cụ AI còn hỗ trợ giáo viên trong việc thiết kế các video trực quan bằng cách cung cấp gợi ý về hình ảnh và âm thanh sinh động, nhằm hỗ trợ học sinh trong quá trình quan sát, khơi gợi ý tưởng và cảm hứng trước khi viết. Không chỉ dừng lại ở vai trò thiết kế nội dung, trí tuệ nhân tạo còn là công cụ hỗ trợ hiệu quả trong công tác đánh giá và phản hồi bài viết của học sinh. Với khả năng phân tích văn bản, phát hiện lỗi ngữ pháp, chính tả và cấu trúc câu, AI giúp giáo viên rút ngắn thời gian chấm chữa bài, đồng thời cung cấp dữ liệu tham chiếu phục vụ việc điều chỉnh chiến lược dạy học một cách kịp thời. Tuy nhiên, cần khẳng định rằng AI không thay thế vai trò sư phạm của giáo viên, mà đóng vai trò hỗ trợ, nhằm làm sâu sắc thêm quá trình tương tác sư phạm, nâng cao hiệu quả tổ chức dạy học kĩ năng viết của học sinh.

Đối với học sinh tiểu học, công cụ trí tuệ nhân tạo tạo sinh có thể đóng vai trò như một "người bạn học tập thông minh", hỗ trợ tích cực trong quá trình hình thành và phát triển kĩ năng viết. AI giúp học sinh tiếp cận hoạt động viết một cách nhẹ nhàng và chủ động hơn thông qua các gợi ý về từ vựng, ý tưởng nội dung và các bài văn tham khảo phù hợp. Đồng thời, nhờ khả năng tích hợp hình ảnh và âm thanh sinh động, AI góp phần làm cho quá trình quan sát trở nên trực quan và hấp dẫn hơn, từ đó khơi gợi cảm hứng và kích thích trí tưởng tượng của học sinh. Bên cạnh việc hỗ trợ hình thành ý tưởng và khơi gợi hứng thú học tập, AI còn tạo ra một môi trường học tập linh hoạt và cá nhân hóa, trong đó học sinh có thể tương tác, chỉnh sửa và hoàn thiện bài viết dựa trên phản hồi tức thời từ hệ thống. Quá trình này không chỉ giúp học sinh phát triển kĩ năng viết mà còn rèn luyện năng lực tự học, tự đánh giá và tư duy phản biện. Việc được hỗ trợ kịp thời, liên tục cũng góp phần hình thành thói quen viết thường xuyên, từng bước nâng cao năng lực ngôn ngữ và phát triển tư duy sáng tạo cho học sinh ngay từ những năm đầu tiểu học.

Nhìn chung, việc ứng dụng trí tuệ nhân tạo (AI) trong dạy học kĩ năng viết ở tiểu học nói riêng và trong các cấp học nói chung không chỉ mở ra những cơ hội đổi mới phương pháp giảng dạy, mà còn đặt nền tảng cho việc kiến tạo một môi trường học tập linh hoạt, cá nhân hóa và hiệu quả. Với những khả năng vượt trội trong việc tạo sinh nội dung, phản hồi tức thời,

hỗ trợ đa phương tiện và cá nhân hóa quá trình học tập, công cụ AI tạo sinh góp phần nâng cao chất lượng dạy học, khơi dậy hứng thú học tập và thúc đẩy sự phát triển toàn diện của học sinh. Sự tích hợp AI vào giáo dục không đơn thuần là một xu hướng công nghệ, mà còn là bước chuyển mình tất yếu trong tiến trình chuyển đổi số của ngành giáo dục, hướng đến xây dựng mô hình giáo dục thông minh – nơi công nghệ được vận dụng như một phương tiện hiệu quả để hỗ trợ cả người dạy lẫn người học cùng phát triển năng lực trong bối cảnh kỉ nguyên tri thức số hóa.

2.2.3 Một số công cụ AI ứng dụng trong dạy học viết văn miêu tả

Dạy học viết văn miêu tả trong chương trình Tiếng Việt có thể được tổ chức theo ba giai đoạn cơ bản: trước khi viết, trong khi viết, sau khi viết. Mỗi giai đoạn đảm nhiệm một vai trò riêng, góp phần hình thành và phát triển năng lực viết cho học sinh một cách toàn diện. Ở giai đoạn trước khi viết, học sinh có nhiệm vụ xác định đối tượng miêu tả, quan sát, tìm ý và lập dàn ý - đây là bước nền tảng để định hướng nội dung và cảm xúc cho bài viết. Giai đoạn trong khi viết là quá trình học sinh vận dụng kiến thức, triển khai bài văn theo dàn ý, sử dụng ngôn ngữ gợi hình, gợi cảm và các biện pháp tu từ. Giai đoạn sau khi viết giúp học sinh tự đánh giá, tiếp nhận góp ý và chỉnh sửa bài để hoàn thiện sản phẩm.

Tương ứng với từng giai đoạn, việc tích hợp các công cụ AI tạo sinh một cách hợp lí sẽ hỗ trợ cá nhân hóa quá trình học, nâng cao hiệu quả dạy học, đồng thời thúc đẩy tư duy sáng tạo và phát triển năng lực viết cho học sinh. Một số công cụ AI hỗ trợ trong dạy học viết văn miêu tả:

Giai đoạn trước khi viết	Giai đoạn trong khi viết	Giai đoạn sau khi viết
- Tạo hình phong cảnh (Sora, Leonardo, Kling AI…) - Tạo phiếu quan sát (Canva, Copilot, Chat GPT…) - Xây dựng kho từ ngữ miêu tả phong cảnh (Copilot, Chat GPT, Gemini, Grok, Claude…) - Tạo sơ đồ tư duy (Mapify, Monica…)	- Mở rộng ý tưởng (Chat GPT, Grok, Bling AI…) - Tạo đoạn văn tham khảo từ cấu trúc cho trước (Chat GPT, Grok, Copilot, DeepSeek…)	- Phát hiện và sửa lỗi từ vựng (Chat GPT…) - Mở rộng hoặc bổ sung chi tiết thiếu (Chat GPT, Llama, Mistral Large…)

2.3. Vận dụng AI vào dạy học viết văn miêu tả cho học sinh lớp 5

2.3.1. Nguyên tắc ứng dụng trí tuệ nhân tạo trong dạy học kĩ năng viết

2.3.1.1. Nguyên tắc đảm bảo mục tiêu môn học

Việc tích hợp trí tuệ nhân tạo (AI) trong dạy học viết ở tiểu học cần được đặt dưới sự chi phối của nguyên tắc đảm bảo mục tiêu môn học, nhằm duy trì tính nhất quán giữa đổi mới phương pháp và chuẩn đầu ra của *Chương trình Giáo dục phổ thông 2018.* Trong môn Tiếng Việt lớp 5, dạy học viết không chỉ dừng lại ở việc hình thành kĩ năng tạo lập văn bản, mà còn hướng đến phát triển tư duy ngôn ngữ, năng lực giao tiếp, khả năng tư duy phản biện và cảm

thụ thẩm mỹ của học sinh. AI, với khả năng hỗ trợ gợi ý nội dung, mở rộng vốn từ, kiểm tra lỗi ngôn ngữ và cung cấp tiêu chí đánh giá, có thể đóng vai trò như một công cụ hỗ trợ đắc lực trong quá trình này. Tuy nhiên, việc ứng dụng AI cần được định hướng rõ ràng, tránh tình trạng lệch chuẩn giáo dục hoặc làm giảm vai trò chủ động, sáng tạo của người học.

Do đó, trong thiết kế hoạt động dạy học viết có ứng dụng công nghệ AI, giáo viên cần đảm bảo sự định hướng sư phạm rõ ràng, giúp học sinh phát triển năng lực sử dụng tiếng Việt một cách chuẩn xác, linh hoạt và phù hợp với đặc điểm nhận thức lứa tuổi. Việc tích hợp AI cần được điều chỉnh theo các thể loại văn bản cụ thể (miêu tả, kể chuyện, biểu cảm...) và tương thích với từng mức độ phát triển kĩ năng viết (viết câu, đoạn văn, bài văn hoàn chỉnh). Tính tương thích giữa công nghệ và nội dung chương trình là điều kiện cần thiết để nâng cao chất lượng dạy học viết, bảo đảm sự gắn kết giữa đổi mới công nghệ và giá trị cốt lõi của giáo dục tiểu học.

2.3.1.2. Nguyên tắc cá nhân hoá người học

Nguyên tắc cá nhân hóa người học là một định hướng cốt lõi trong giáo dục hiện đại, đặc biệt quan trọng khi tích hợp trí tuệ nhân tạo (AI) vào dạy học kỹ năng viết – lĩnh vực vốn mang tính sáng tạo và cá nhân cao. Việc thiết kế hoạt động học tập cần được điều chỉnh theo nhu cầu, năng lực, hứng thú và tốc độ học tập riêng của từng học sinh, từ đó không chỉ nâng cao hiệu quả tiếp thu mà còn thúc đẩy sự phát triển toàn diện về tư duy, cảm xúc và năng lực ngôn ngữ. Trong dạy học văn miêu tả phong cảnh ở Tiểu học, việc cá nhân hóa trở nên thiết yếu bởi mỗi học sinh đều có vốn sống, cảm nhận thẩm mỹ và phong cách diễn đạt riêng. AI, với khả năng gợi mở nội dung, mở rộng vốn từ và hỗ trợ tổ chức ý tưởng, có thể được khai thác để thiết kế các nhiệm vụ học tập phù hợp với từng cá nhân, góp phần phát triển kỹ năng viết một cách linh hoạt và hiệu quả.

Tuy nhiên, để nguyên tắc cá nhân hóa vận hành hiệu quả, vai trò tổ chức và định hướng của giáo viên là yếu tố then chốt. Giáo viên cần kiến tạo môi trường học tập mở, tạo điều kiện để học sinh lựa chọn chủ đề, góc nhìn và cách thể hiện phù hợp với bản thân, đồng thời hướng dẫn các em sử dụng AI một cách có chọn lọc và phản biện. Thay vì áp đặt mẫu bài chung, giáo viên nên khuyến khích học sinh khai thác đa dạng các gợi ý từ AI, rồi điều chỉnh theo cảm xúc và nhận thức cá nhân. Cách tiếp cận này không chỉ phát huy cá tính người học mà còn góp phần hình thành năng lực biên tập, chỉnh sửa và đánh giá văn bản – những yêu cầu quan trọng trong dạy học viết theo định hướng phát triển năng lực. Việc hiện thực hóa nguyên tắc cá nhân hóa với sự hỗ trợ của AI chính là minh chứng rõ nét cho triết lý giáo dục lấy người học làm trung tâm trong thời đại chuyển đổi số.

2.3.1.3. Nguyên tắc cập nhật và tuỳ chỉnh câu lệnh liên tục

Việc thường xuyên cập nhật và điều chỉnh câu lệnh (prompt) là một nguyên tắc thiết yếu trong quá trình ứng dụng trí tuệ nhân tạo (AI) vào thiết kế nội dung dạy học, nhằm đảm bảo tính thích ứng, hiệu quả và phù hợp với mục tiêu giáo dục cụ thể. Khi tích hợp AI vào hoạt động giảng dạy, giáo viên cần chủ động thiết kế và tinh chỉnh các câu lệnh phù hợp với từng nhiệm vụ học tập, trình độ nhận thức của học sinh và đặc thù nội dung bài học. Quá trình này

không chỉ mang tính kĩ thuật mà còn đòi hỏi tư duy sư phạm, kinh nghiệm thực tiễn và khả năng sáng tạo của người dạy để phát huy tối đa hiệu năng hỗ trợ của AI trong môi trường lớp học.

Một câu lệnh hiệu quả cần được xây dựng như một lời dẫn giàu tính gợi mở, có khả năng kích hoạt AI phản hồi theo hướng chi tiết, cụ thể và có giá trị sư phạm. Về mặt giao tiếp, câu lệnh cần xác định rõ mục đích sử dụng, đối tượng tiếp nhận, nội dung mong muốn và hình thức phản hồi mong đợi. Tính rõ ràng, định hướng và phù hợp với ngữ cảnh giáo dục là những yếu tố quyết định đến chất lượng đầu ra của hệ thống AI.

Chẳng hạn, trong dạy học viết văn miêu tả phong cảnh ở lớp 5, thay vì yêu cầu đơn lẻ và chung chung, giáo viên có thể thiết kế các câu lệnh có định hướng sáng tạo như: "*Hãy tạo một danh sách các câu hỏi dẫn dắt giúp học sinh lớp 5 khám phá cảnh vật quê hương mình, từ đó hình thành đoạn văn miêu tả giàu cảm xúc và đúng với trình độ của các em*". Loại câu lệnh này không chỉ cung cấp thông tin, mà còn định hình cách AI hỗ trợ quá trình phát triển tư duy ngôn ngữ và khả năng biểu đạt cá nhân cho người học.

Tác giả Lương Thị Hiền trong *Ứng dụng trí tuệ nhân tạo tạo sinh trong dạy tiếng Việt như một ngoại ngữ* cho rằng một câu lệnh tốt cần làm rõ các nhân tố giao tiếp (nội dung, đối tượng giao tiếp, mục đích giao tiếp, cách thức giao tiếp, phương tiện giao tiếp). Cụ thể tác giả đã đề cập đến 2 câu lệnh tốt: RACEF và RTF. Tuy nhiên câu lệnh RACEF gồm 5 thành phần (Role, Action, Content, Example, Format) tạm dịch là vai trò, hành động, nội dung, ví dụ và định dạng được đánh giá là có nhiều tính năng vượt trội hơn cả.

Việc thiết kế và điều chỉnh câu lệnh (prompt) trong quá trình tương tác với công cụ trí tuệ nhân tạo (AI) giữ vai trò quyết định trong việc đảm bảo sự tương thích giữa phản hồi của AI với mục tiêu, nội dung và phương pháp dạy học. Đặc biệt, trong bối cảnh giáo viên tiểu học vận dụng AI vào hoạt động dạy học viết văn miêu tả, việc xây dựng hệ thống câu lệnh mang tính định hướng rõ ràng sẽ góp phần kích thích tư duy, phát triển ngôn ngữ và năng lực biểu đạt của học sinh. Ví dụ, thay vì yêu cầu đơn giản như "*Hãy viết bài văn tả cảnh*", giáo viên có thể đưa ra lệnh gợi mở hơn: "*Hãy đóng vai một giáo viên tiểu học, tạo các câu hỏi giúp học sinh quan sát phong cảnh biển để viết bài văn miêu tả.*" Cách đặt lệnh như vậy giúp AI hiểu rõ vai trò, nhiệm vụ và mục tiêu cụ thể để hỗ trợ người học tốt hơn.

Tóm lại, nguyên tắc điều chỉnh câu lệnh AI không chỉ giúp nâng cao hiệu suất hoạt động của công nghệ mà còn bảo đảm tính phù hợp sư phạm. Đây là tiền đề quan trọng để tích hợp AI vào quá trình dạy học một cách hiệu quả, đồng thời góp phần phát triển năng lực thiết kế hoạt động học tập cho giáo viên, xây dựng môi trường học tập linh hoạt, đáp ứng đa dạng nhu cầu của người học.

2.3.1.4. Nguyên tắc điều chỉnh, huấn luyện AI để đáp ứng mục tiêu dạy học

Nguyên tắc điều chỉnh và huấn luyện trí tuệ nhân tạo (AI) để đáp ứng mục tiêu dạy học nhấn mạnh yêu cầu thiết yếu trong việc thiết lập, chỉnh sửa và thích ứng hệ thống AI trước khi triển khai vào thực tiễn lớp học. Để AI có thể hỗ trợ hiệu quả cho giáo viên trong thiết kế hoạt

động, đồng thời đáp ứng tốt nhu cầu học tập đa dạng của học sinh, việc huấn luyện cần được tiến hành một cách có chủ đích, bám sát yêu cầu cần đạt của chương trình môn Tiếng Việt. Đặc biệt, trong dạy học viết văn miêu tả phong cảnh, AI cần được điều chỉnh sao cho sản phẩm đầu ra phù hợp với năng lực tiếp nhận của học sinh tiểu học, góp phần phát triển tư duy ngôn ngữ và khả năng biểu đạt giàu tính cá nhân của học sinh.

Hiện nay, kho ngữ liệu tiếng Việt tích hợp trong nhiều hệ thống AI chưa đầy đủ hoặc chưa được tùy biến theo yêu cầu giáo dục, nên khả năng tạo ra sản phẩm ngôn ngữ phù hợp với mục tiêu dạy học vẫn còn nhiều hạn chế. Trong thực tế, giáo viên có thể sử dụng AI như một công cụ hỗ trợ hiệu quả nếu biết cách điều chỉnh hệ thống theo mục tiêu bài học, trình độ học sinh và đặc điểm của thể loại văn bản. Đặc biệt ở lớp 5, khi học sinh bắt đầu hình thành kĩ năng viết có cấu trúc và giàu tính biểu cảm, việc cá nhân hóa công cụ AI theo định hướng sư phạm càng trở nên cần thiết.

Quá trình huấn luyện AI cần được thực hiện nhằm giúp công cụ này nhận biết và thích ứng với các loại nhiệm vụ học tập cụ thể mà giáo viên giao. Ví dụ, AI có thể được điều chỉnh để hỗ trợ học sinh mở rộng vốn từ, sử dụng biện pháp tu từ, sắp xếp bố cục đoạn hợp lý, hoặc phản hồi các đoạn viết của học sinh theo tiêu chí đánh giá cụ thể. Nếu giáo viên muốn học sinh thực hành viết đoạn văn miêu tả sử dụng so sánh - nhân hóa, AI cần được huấn luyện để ưu tiên gợi ý hoặc đánh giá các cấu trúc thể hiện được hai biện pháp tu từ này, đồng thời tránh đưa ra lời khuyên máy móc hoặc thiếu chính xác.

Bên cạnh đó, AI cũng cần có khả năng tương tác liên tục để thích nghi với thay đổi của dữ liệu đầu vào và sự tiến bộ của người học. Ví dụ, khi học sinh chưa thuần thục viết đoạn văn miêu tả cảnh vật, giáo viên có thể điều chỉnh hệ thống để AI chỉ tập trung gợi ý từ khóa, câu hỏi gợi mở hoặc mẫu dàn ý cơ bản. Khi học sinh đã tiến bộ hơn, AI có thể được lập trình để gợi ý cấu trúc đoạn văn nâng cao, hoặc đặt câu hỏi phản biện, giúp học sinh tự đánh giá và điều chỉnh văn bản của mình.

Trong dạy học văn miêu tả phong cảnh, giáo viên có thể sử dụng AI như một người trợ lý sáng tạo, hỗ trợ xây dựng các tình huống gợi viết hoặc câu hỏi dẫn dắt. Chẳng hạn, thay vì yêu cầu trực tiếp *"Hãy viết đoạn văn tả cánh đồng vào buổi sớm.",* giáo viên có thể đề nghị AI tạo ra lời mời học sinh tham gia tình huống quan sát sáng tạo như: *"Hãy tưởng tượng em đang đứng trước cánh đồng vào buổi sớm. Em nhìn thấy điều gì? Những âm thanh, màu sắc, chuyển động nào khiến em ấn tượng?"*. Từ đó, học sinh sẽ được dẫn dắt vào quá trình hình dung, cảm nhận và sử dụng ngôn ngữ một cách tự nhiên, sinh động.

Nhìn chung, quá trình huấn luyện và điều chỉnh AI không chỉ mang tính kĩ thuật mà còn là một quá trình sư phạm đòi hỏi sự hiểu biết sâu sắc về mục tiêu học tập, đặc điểm người học và đặc trưng thể loại văn bản. Khi được khai thác hợp lí, AI có thể trở thành công cụ đắc lực giúp học sinh tiểu học phát triển kĩ năng viết theo hướng cá nhân hóa, sáng tạo và hiệu quả.

2.3.2. Dạy học viết văn miêu tả phong cảnh có sử dụng trí tuệ nhân tạo cho học sinh lớp 5

2.3.2.1. Ứng dụng trí tuệ nhân tạo để tìm tư liệu cho bài văn miêu tả

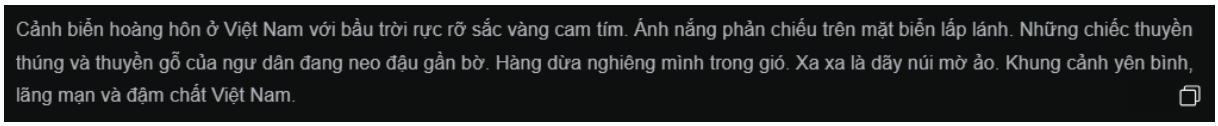
Trong dạy học viết văn miêu tả ở cấp tiểu học nói chung và dạy học văn miêu tả phong cảnh ở lớp 5 nói riêng, giai đoạn tìm kiếm tư liệu giữ vai trò khởi đầu và có ý nghĩa đặc biệt quan trọng. Với sự phát triển mạnh mẽ của trí tuệ nhân tạo, các công cụ như: ChatGPT, Gemini, Copilot, Grok,… đã cho thấy tiềm năng tạo sinh phong cảnh một cách sống động, chân thực trong hoàn cảnh học sinh chưa được ngắm cảnh trực tiếp. Việc tiếp cận được những hình ảnh "thật như ngoài đời" sẽ giúp học sinh hình thành tư duy tưởng tượng, từ đó có những hình dung nhất định về cảnh.

Ví dụ 1: Vận dụng trí tuệ nhân tạo (AI) tạo sinh hình ảnh, góp phần hỗ trợ học sinh lớp 5 phát triển kĩ năng viết văn miêu tả phong cảnh

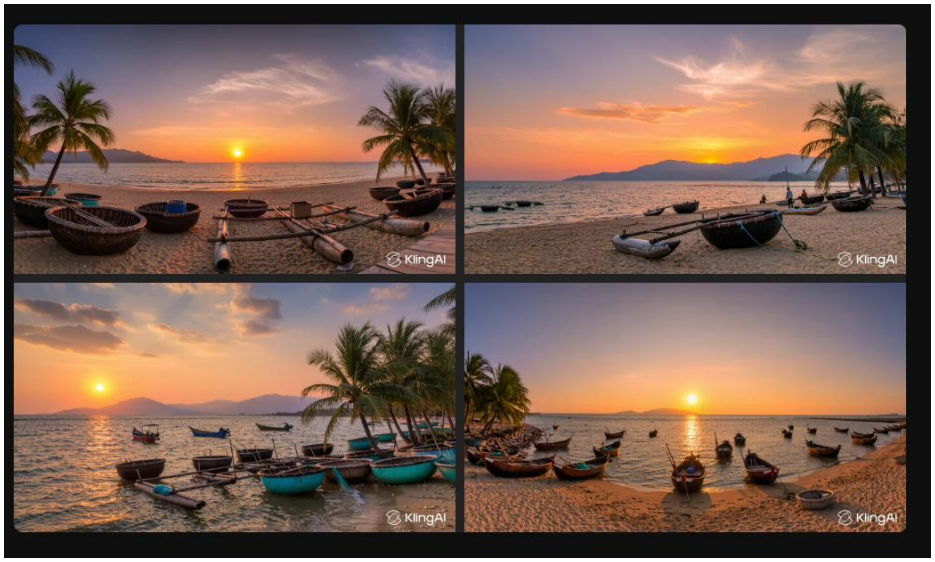
Đề bài: Tả cảnh hoàng hôn trên biển

Sử dụng ứng dụng Kling AI

+ Câu lệnh



+ Kết quả



Ví dụ 2: Vận dụng trí tuệ nhân tạo (AI) tạo sinh hình ảnh góp phần hỗ trợ học sinh lớp 5 trong viết văn miêu tả phong cảnh

Đề bài: Tả cảnh cánh đồng lúa vào buổi sáng

Sử dụng ứng dụng Copilot

+ Câu lệnh

Khung cảnh cánh đồng lúa vào buổi sáng ở Việt Nam thật đẹp. Những thửa ruộng lúa xanh mướt trải dài, trên lá vẫn còn đọng những giọt sương long lanh. Ánh mặt trời buổi sớm vàng óng chiếu xuống làm cánh đồng lúa thêm rực rỡ. Xa xa, vài người nông dân bắt đầu công việc cấy lúa trên đồng, đàn cò trắng bay chập chờn trên nền trời trong xanh. Cảnh vật thanh bình, tươi đẹp và tràn đầy sức sống.

+ Kết quả

2.3.2.2.Ứng dụng trí tuệ nhân tạo để tìm ý và lập dàn ý (Tìm ý bằng văn bản, tìm ý bằng sơ đồ tư duy)

Trong bối cảnh giáo dục hiện đại định hướng chuyển đổi số, việc ứng dụng trí tuệ nhân tạo để tìm ý bằng văn bản, tìm ý bằng sơ đồ tư duy là hướng đi khoa học. Thông qua các nền tảng tích hợp AI như Canva AI, MindMeister AI, giáo viên có thể thiết kế sơ đồ tư duy cho các kiểu bài với cấu trúc ba phần chính (mở bài - thân bài - kết bài), đồng thời tích hợp các nhánh ý liên quan đến nội dung, chi tiết miêu tả, cảm xúc và ngôn ngữ biểu đạt. Việc trực quan hóa nội dung bài viết bằng sơ đồ không chỉ hỗ trợ học sinh định hình tư duy hệ thống mà còn giúp giáo viên dễ dàng theo dõi, điều chỉnh và đánh giá quá trình hình thành ý tưởng của người học. Từ đó, việc ứng dụng AI trong tìm ý và lập dàn ý không chỉ có giá trị về mặt công nghệ mà còn mang lại ý nghĩa sư phạm sâu sắc, phù hợp với mục tiêu phát triển năng lực ngôn ngữ.

Ví dụ 1: Ứng dụng trí tuệ nhân tạo (AI) thiết kế sơ đồ tư duy nhằm hỗ trợ dạy học viết văn miêu tả phong cảnh cho học sinh lớp 5

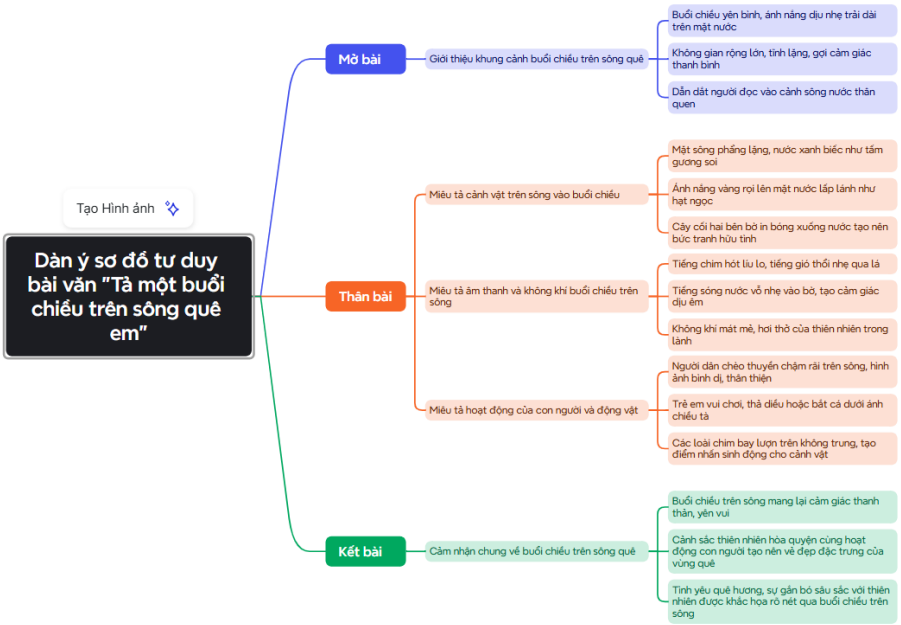
Đề bài: Tả một buổi chiều trên sông quê em

Sử dụng ứng dụng Mapify

+ Câu lệnh

> Bạn là giáo viên tiểu học giỏi dạy môn Tiếng Việt lớp 5. Hãy giúp tôi lập dàn ý có sử dụng các từ ngữ miêu tả phong cảnh giàu sức gợi cho bài văn "Tả một buổi chiều trên sông quê em" dưới dạng sơ đồ tư duy nhằm hỗ trợ học sinh trong quá trình viết văn, theo cấu trúc 3 phần: Mở bài, Thân bài, Kết bài

+ Kết quả



Ví dụ 2: Ứng dụng trí tuệ nhân tạo (AI) thiết kế sơ đồ tư duy nhằm hỗ trợ dạy học viết văn miêu tả phong cảnh cho học sinh lớp 5

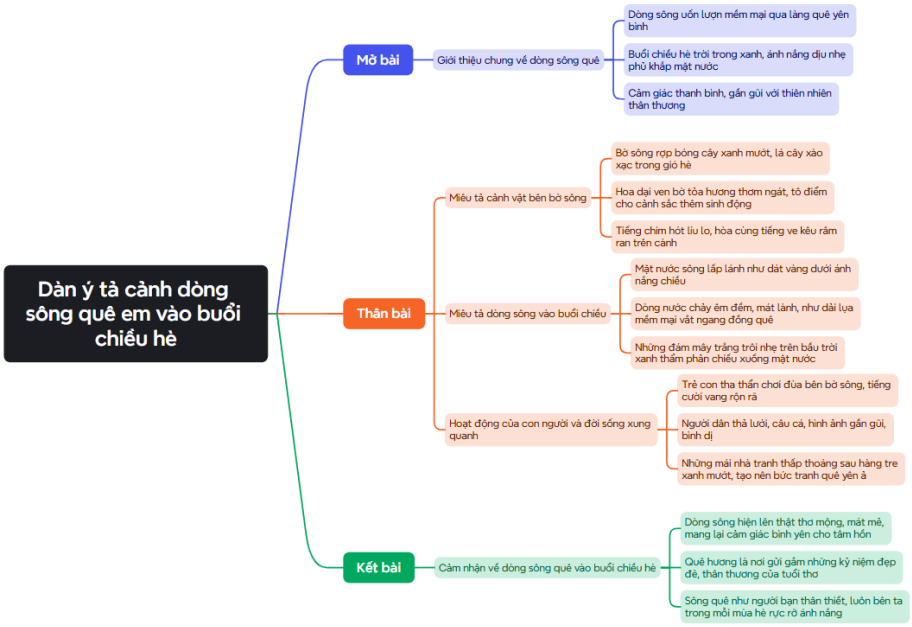
Đề bài: Tả cảnh dòng sông quê em vào buổi chiều hè

Sử dụng ứng dụng Mapify

+ Câu lệnh

> Bạn là giáo viên tiểu học giỏi dạy môn Tiếng Việt lớp 5. Hãy giúp tôi lập dàn ý có sử dụng các từ ngữ miêu tả gợi hình, gợi cảm cho bài văn "Tả cảnh dòng sông quê em vào buổi chiều hè" dưới dạng sơ đồ tư duy nhằm hỗ trợ học sinh trong quá trình viết văn, theo cấu trúc 3 phần: Mở bài, Thân bài, Kết bài.

+ Kết quả



2.3.2.3. Ứng dụng trí tuệ nhân tạo nhằm huy động vốn từ

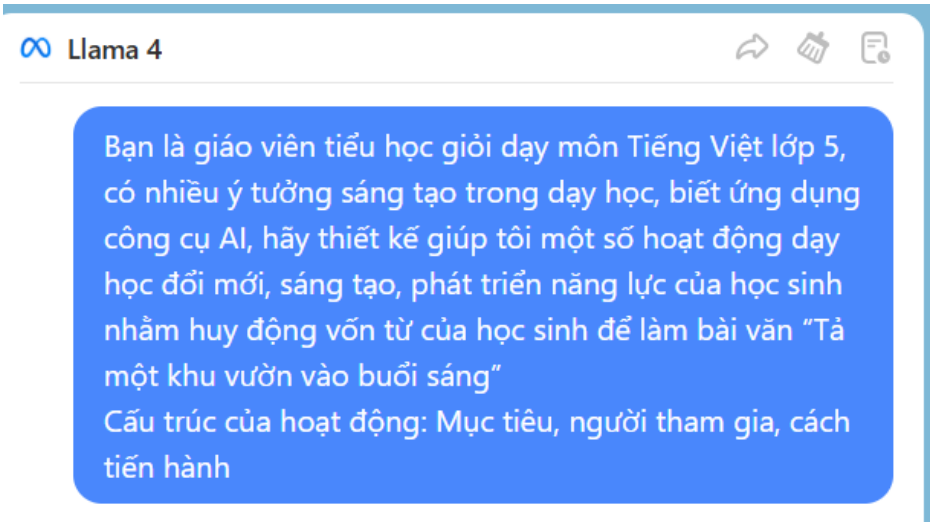
Thực tế cho thấy, vốn từ của học sinh tiểu học còn tương đối hạn chế. Do đó, việc ứng dụng trí tuệ nhân tạo trong việc huy động vốn từ không chỉ giúp học sinh mở rộng được vốn từ liên quan đến phong cảnh thiên nhiên, biết cách lựa chọn từ ngữ phù hợp với từng loại phong cảnh khác nhau đồng thời rèn luyện khả năng sử dụng từ linh hoạt, tránh lặp từ và đơn điệu trong diễn đạt qua đó phát triển tư duy hình ảnh và kĩ năng quan sát tinh tế. Với sự phát triển mạnh mẽ của trí tuệ nhân tạo, giáo viên có thể khai thác AI như một công cụ hỗ trợ đắc lực trong việc xây dựng các ngân hàng từ ngữ miêu tả theo chủ đề, trường nghĩa (màu sắc, âm thanh, cảm xúc, không gian…). Khi có một nguồn từ ngữ phong phú, học sinh sẽ dễ dàng lựa chọn từ phù hợp để biểu đạt ý tưởng một cách chính xác và giàu cảm xúc.

Ví dụ 1: Ứng dụng trí tuệ nhân tạo (AI) trong việc thiết kế hoạt động dạy học nhằm huy động vốn từ cho học sinh lớp 5 trong viết văn miêu tả phong cảnh

Đề bài: Tả một khu vườn vào buổi sáng

Sử dụng ứng dụng Llama 4

+ Câu lệnh:



+ Kết quả:

Tôi có thể giúp bạn với một số hoạt động dạy học đổi mới, sáng tạo để giúp học sinh lớp 5 phát triển năng lực và huy động vốn từ cho bài văn "Tả một khu vườn vào buổi sáng". Dưới đây là một số hoạt động:

Hoạt động 1: Thám hiểm khu vườn qua hình ảnh

- **Mục tiêu:** Kích thích trí tưởng tượng và huy động vốn từ của học sinh về hình ảnh khu vườn vào buổi sáng.
- **Người tham gia:** Cả lớp học sinh lớp 5.
- **Cách tiến hành:**
  1. Giáo viên chuẩn bị một số hình ảnh về khu vườn vào buổi sáng (có thể sử dụng hình ảnh trên internet hoặc ảnh thật).
  2. Cho học sinh quan sát hình ảnh và yêu cầu miêu tả những gì họ thấy.
  3. Ghi lại những từ ngữ, cụm từ mà học sinh sử dụng để miêu tả trên bảng.
  4. Khuyến khích học sinh sử dụng các giác quan (nhìn, nghe, ngửi, chạm) để tưởng tượng và miêu tả.
  5. Sau khi có danh sách từ ngữ, cụm từ, yêu cầu học sinh chọn lọc và sử dụng chúng vào bài viết của mình.

Hoạt động 2: Trò chơi từ vựng với AI

- **Mục tiêu:** Tăng cường vốn từ vựng và khả năng sử dụng từ ngữ sáng tạo của học sinh.
- **Người tham gia:** Học sinh lớp 5, làm việc cá nhân hoặc theo nhóm.
- **Cách tiến hành:**
  1. Sử dụng công cụ AI như chatbot hoặc phần mềm tạo từ vựng.
  2. Yêu cầu học sinh nhập từ khóa liên quan đến "khu vườn buổi sáng" vào công cụ AI.
  3. Công cụ AI sẽ gợi ý một số từ vựng, cụm từ liên quan.
  4. Học sinh lựa chọn và ghi lại những từ ngữ phù hợp.
  5. Thảo luận cả lớp về những từ ngữ được gợi ý và cách sử dụng chúng trong bài viết.

Hoạt động 3: Viết nhanh - Kết nối ý tưởng

- **Mục tiêu:** Phát triển khả năng viết nhanh và kết nối ý tưởng của học sinh.
- **Người tham gia:** Học sinh lớp 5, làm việc cá nhân.
- **Cách tiến hành:**
  1. Đặt đồng hồ计时 5-7 phút.
  2. Yêu cầu học sinh viết ra tất cả những gì họ nghĩ về khu vườn vào buổi sáng mà không cần lo lắng về chính tả, ngữ pháp.
  3. Sau thời gian quy định, yêu cầu học sinh đọc to bài viết của mình.
  4. Cả lớp cùng thảo luận và góp ý về nội dung, cách diễn đạt.
  5. Học sinh chỉnh sửa và hoàn thiện bài viết dựa trên góp ý.

Hoạt động 4: Chia sẻ và phản hồi

- **Mục tiêu:** Phát triển khả năng phản hồi và chỉnh sửa bài viết của học sinh.
- **Người tham gia:** Học sinh lớp 5, làm việc theo cặp hoặc nhóm nhỏ.
- **Cách tiến hành:**
  1. Yêu cầu học sinh hoàn thành bài viết "Tả một khu vườn vào buổi sáng" sau khi tham gia các hoạt động trên.
  2. Tổ chức cho học sinh trao đổi bài viết với bạn bè.
  3. Hướng dẫn học sinh cách phản hồi bài viết của bạn (ví dụ: nội dung hay, gợi ý cho việc cải thiện).
  4. Học sinh đọc bài viết của bạn và đưa ra phản hồi.
  5. Học sinh chỉnh sửa bài viết dựa trên phản hồi.

Những hoạt động này không chỉ giúp học sinh huy động vốn từ mà còn phát triển các kỹ năng cần thiết cho việc viết một bài văn hiệu quả.

Ví dụ 2: Ứng dụng trí tuệ nhân tạo (AI) để xây dựng "kho" từ ngữ trong văn miêu tả phong cảnh lớp 5

Đề bài: Tả cảnh hoàng hôn trên biển

Sử dụng ứng dụng ChatGPT

+ Câu lệnh

> Bạn là giáo viên tiểu học dạy môn Tiếng Việt lớp 5, hãy xây dựng kho từ ngữ miêu tả phong cảnh để hỗ trợ học sinh viết bài văn "Tả cảnh hoàng hôn trên biển".

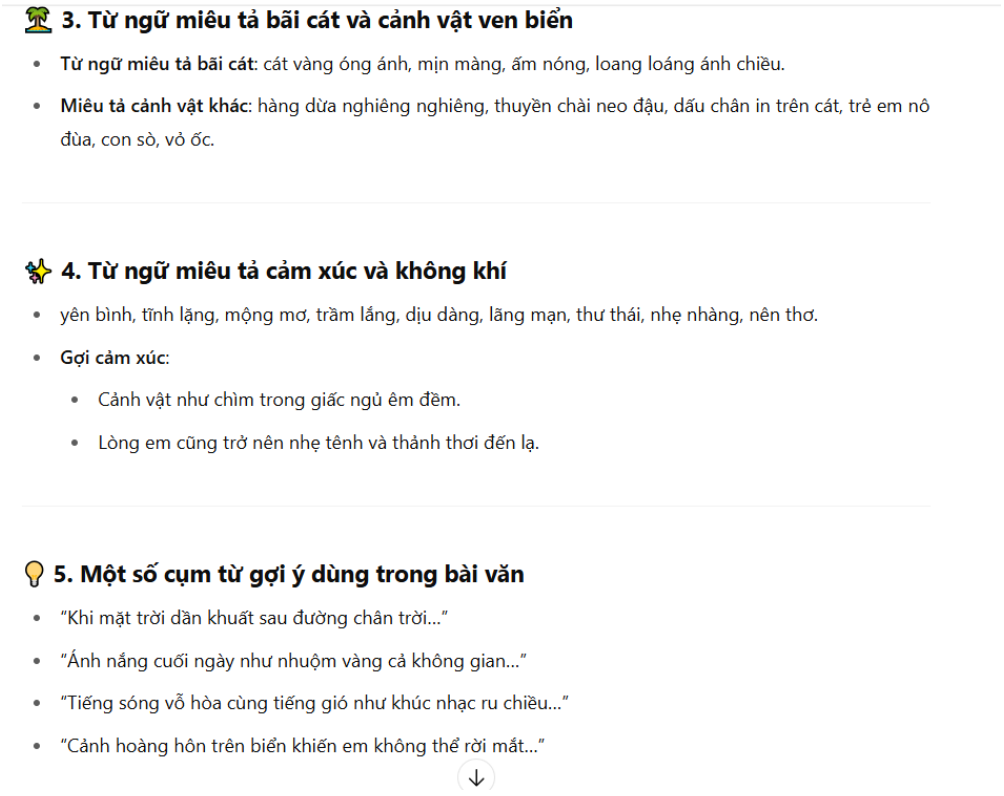
+ Kết quả

1. Từ ngữ miêu tả bầu trời lúc hoàng hôn

- **Màu sắc**: đỏ rực, cam vàng, hồng nhạt, tím than, xanh lam nhạt, vàng cam rực rỡ, ráng chiều.
- **Hình ảnh**: mây lững lờ trôi, mây nhuộm sắc hồng, bầu trời như tấm thảm nhung, ánh nắng loang loáng.
- **So sánh - nhân hoá**:
  - Mặt trời như lòng đỏ trứng gà đang dần chìm xuống biển.
  - Mây bồng bềnh như những chiếc gối bông khổng lồ.
  - Bầu trời như đang khoác một tấm áo rực rỡ sắc màu.

2. Từ ngữ miêu tả mặt biển lúc hoàng hôn

- **Màu sắc**: xanh biếc, ánh vàng kim, ánh đồng đỏ, ánh bạc lấp lánh.
- **Âm thanh - cảm giác**: sóng rì rào, nước vỗ nhẹ vào bờ, mặt biển yên ả, dập dềnh êm ả.
- **So sánh - nhân hoá**:
  - Sóng biển như những con thú nhỏ đang nô đùa.
  - Mặt biển như một tấm gương phản chiếu cả bầu trời rực rỡ.



3. Từ ngữ miêu tả bãi cát và cảnh vật ven biển

- **Từ ngữ miêu tả bãi cát**: cát vàng óng ánh, mịn màng, ấm nóng, loang loáng ánh chiều.
- **Miêu tả cảnh vật khác**: hàng dừa nghiêng nghiêng, thuyền chài neo đậu, dấu chân in trên cát, trẻ em nô đùa, con sò, vỏ ốc.

4. Từ ngữ miêu tả cảm xúc và không khí

- yên bình, tĩnh lặng, mộng mơ, trầm lắng, dịu dàng, lãng mạn, thư thái, nhẹ nhàng, nên thơ.
- **Gợi cảm xúc**:
  - Cảnh vật như chìm trong giấc ngủ êm đềm.
  - Lòng em cũng trở nên nhẹ tênh và thảnh thơi đến lạ.

5. Một số cụm từ gợi ý dùng trong bài văn

- "Khi mặt trời dần khuất sau đường chân trời..."
- "Ánh nắng cuối ngày như nhuộm vàng cả không gian..."
- "Tiếng sóng vỗ hòa cùng tiếng gió như khúc nhạc ru chiều..."
- "Cảnh hoàng hôn trên biển khiến em không thể rời mắt..."

2.3.2.4 Ứng dụng trí tuệ nhân tạo thiết kế các tiêu chí đánh giá bài văn miêu tả

Trong dạy học viết văn miêu tả phong cảnh ở lớp 5, việc xây dựng và sử dụng hệ thống tiêu chí đánh giá giữ vai trò quan trọng, không chỉ giúp giáo viên định hướng hoạt động viết mà còn hỗ trợ đánh giá chính xác năng lực biểu đạt của học sinh. Thực tiễn cho thấy, các tiêu chí truyền thống thường được xây dựng dựa trên kinh nghiệm cá nhân, dẫn đến sự thiếu nhất quán và đôi khi chưa thật sự phản ánh đầy đủ những yêu cầu về kỹ năng viết. Vì vậy, sự phát triển mạnh mẽ của công nghệ trí tuệ nhân tạo (AI) đã mở ra hướng tiếp cận mới trong việc thiết kế tiêu chí đánh giá bài viết. Cụ thể, giáo viên có thể sử dụng các nền tảng AI để phân tích văn bản mẫu, tổng hợp các yêu cầu cốt lõi, từ đó xây dựng một hệ thống tiêu chí đánh giá dựa trên các yếu tố như: bố cục hợp lý, khả năng sử dụng từ ngữ gợi hình - gợi cảm, vận dụng biện pháp tu từ, tính mạch lạc và cảm xúc trong diễn đạt. Thêm vào đó, một số công cụ AI còn có khả năng phân tích bài viết của học sinh, đối chiếu với hệ thống tiêu chí đã thiết kế và đưa ra nhận xét gợi ý cụ thể. Ví dụ, AI có thể nhận diện đoạn văn miêu tả chưa rõ ràng, đề xuất học sinh tăng cường sử dụng từ ngữ gợi tả, hoặc điều chỉnh trình tự miêu tả để tạo tính liên kết không gian – thời gian rõ ràng hơn. Từ góc độ quản lý và tổ chức dạy học, việc tích hợp AI vào quy trình đánh giá không chỉ giúp giáo viên tiết kiệm thời gian, nâng cao tính khách quan mà còn hỗ trợ thực hiện đánh giá thường xuyên và cá nhân hóa phản hồi cho từng học sinh. Qua đó, chất lượng dạy học viết văn miêu tả phong cảnh được nâng cao rõ rệt, đồng thời góp phần phát triển tư duy ngôn ngữ và khả năng cảm thụ thẩm mỹ của học sinh một cách sâu sắc và bền vững.

Ví dụ 1: Ứng dụng trí tuệ nhân tạo (AI) thiết kế hệ thống tiêu chí đánh giá bài văn miêu tả phong cảnh của học sinh lớp 5

Đề bài: Tả cảnh một buổi sáng mùa xuân trên quê hương em

Sử dụng ứng dụng Grok

+ Câu lệnh:

> Bạn là chuyên gia giảng dạy Tiếng Việt tiểu học. Hãy giúp tôi xây dựng một hệ thống tiêu chí đánh giá dưới dạng bảng cho bài văn "Tả cảnh một buổi sáng mùa xuân trên quê hương em". Các tiêu chí cần cụ thể, phù hợp với năng lực học sinh tiểu học, bám sát yêu cầu cần đạt của chương trình giáo dục phổ thông 2018. Mỗi tiêu chí nên có thang điểm rõ ràng (theo thang điểm 10), kèm mô tả mức độ đạt được từ thấp đến cao.

+ Kết quả:

Tiêu chí đánh giá	Thang điểm (tổng 10 điểm)	Mô tả mức độ đạt được
1. Nội dung miêu tả (Độ chính xác, đầy đủ và phù hợp với chủ đề: Tả cảnh buổi sáng mùa xuân trên quê hương, bao gồm các yếu tố như thiên nhiên, con người, hoạt động, phù hợp với quan sát thực tế của học sinh tiểu học theo chương trình 2018 nhấn mạnh kỹ năng quan sát và miêu tả gần gũi.)	3 điểm	- **Thấp (0-1 điểm):** Nội dung sơ sài, không miêu tả được các yếu tố chính của buổi sáng mùa xuân (như thời tiết, cây cỏ, chim chóc, sông nước, hoạt động con người); thiếu sự quan sát cụ thể, không liên quan đến quê hương. - **Trung bình (1-2 điểm):** Miêu tả một số yếu tố cơ bản nhưng chưa đầy đủ, có vài chi tiết quan sát được nhưng còn chung chung, chưa thể hiện rõ nét đặc trưng của quê hương. - **Cao (2-3 điểm):** Miêu tả đầy đủ, chính xác các yếu tố cảnh vật và hoạt động buổi sáng mùa xuân trên quê hương; thể hiện sự quan sát tinh tế, gần gũi, phù hợp với trải nghiệm thực tế của học sinh.

3. Ngôn ngữ và biện pháp tu từ (Sử dụng từ ngữ miêu tả sinh động, biện pháp tu từ đơn giản như so sánh, nhân hóa; ngôn ngữ trong sáng, phù hợp lứa tuổi tiểu học theo chương trình 2018 nhấn mạnh phát triển kỹ năng ngôn ngữ biểu cảm.)	3 điểm	- **Thấp (0-1 điểm):** Ngôn ngữ nghèo nàn, ít từ miêu tả, không có biện pháp tu từ, câu văn đơn điệu. - **Trung bình (1-2 điểm):** Sử dụng một số từ ngữ miêu tả cơ bản, có ít biện pháp tu từ đơn giản nhưng chưa đa dạng, câu văn còn vụng về. - **Cao (2-3 điểm):** Ngôn ngữ giàu hình ảnh, sử dụng hiệu quả biện pháp tu từ (so sánh, nhân hóa) để làm nổi bật cảnh vật; câu văn mạch lạc, biểu cảm tốt.
4. Chính tả, ngữ pháp và dấu câu (Viết đúng chính tả, ngữ pháp cơ bản, sử dụng dấu câu phù hợp, đảm bảo tính chuẩn mực ngôn ngữ theo chương trình 2018.)	1 điểm	- **Thấp (0-0.3 điểm):** Nhiều lỗi chính tả, ngữ pháp sai cơ bản, dấu câu lộn xộn. - **Trung bình (0.3-0.6 điểm):** Một số lỗi chính tả và ngữ pháp, dấu câu chưa chính xác nhưng không ảnh hưởng lớn đến nội dung. - **Cao (0.6-1 điểm):** Không có hoặc rất ít lỗi chính tả, ngữ pháp đúng, dấu câu sử dụng hợp lý, văn bản sạch sẽ.
5. Sáng tạo và cảm xúc cá nhân (Thể hiện sự sáng tạo trong miêu tả, kết hợp cảm xúc cá nhân, yêu quê hương; phù hợp với yêu cầu phát triển tư duy sáng tạo và giáo dục tình cảm ở tiểu học theo chương trình 2018.)	1 điểm	- **Thấp (0-0.3 điểm):** Thiếu sáng tạo, miêu tả máy móc, không có cảm xúc cá nhân. - **Trung bình (0.3-0.6 điểm):** Có chút sáng tạo ở một số chi tiết, thể hiện cảm xúc đơn giản nhưng chưa sâu sắc. - **Cao (0.6-1 điểm):** Miêu tả sáng tạo, độc đáo, kết hợp cảm xúc cá nhân chân thực, thể hiện tình yêu quê hương rõ nét.

Ví dụ 2: Ứng dụng trí tuệ nhân tạo (AI) thiết kế hệ thống tiêu chí đánh giá bài văn miêu tả phong cảnh của học sinh lớp 5

Đề bài: Tả cảnh một đêm trăng trên quê hương em

Sử dụng ứng dụng Gemini

+ Câu lệnh

Bạn là chuyên gia giảng dạy Tiếng Việt tiểu học. Hãy giúp tôi xây dựng một hệ thống tiêu chí đánh giá dưới dạng bảng cho bài văn "Tả cảnh một đêm trăng trên quê hương em". Các tiêu chí cần cụ thể, phù hợp với năng lực học sinh tiểu học, bám sát yêu cầu cần đạt của chương trình giáo dục phổ thông 2018. Mỗi tiêu chí nên có thang điểm rõ ràng (theo thang điểm 10), kèm mô tả mức độ đạt được từ thấp đến cao.

+ Kết quả

BẢNG TIÊU CHÍ ĐÁNH GIÁ BÀI VĂN TẢ CẢNH "MỘT ĐÊM TRĂNG TRÊN QUÊ HƯƠNG EM"

Họ và tên học sinh: .....

Lớp: .....

Điểm: ...../10

Tiêu chí	Mức độ 1 - Chưa đạt (1-4 điểm)	Mức độ 2 - Đạt (5-7 điểm)	Mức độ 3 - Tốt (8-10 điểm)
1. Bố cục và trình bày	- Bố cục chưa rõ ràng, các phần chưa tách bạch. - Trình bày thiếu mạch lạc, còn sai lỗi chính tả nhiều.	- Bố cục có đủ 3 phần (mở bài, thân bài, kết bài) nhưng còn sơ sài. - Trình bày khá sạch đẹp, ít lỗi chính tả.	- Bố cục đầy đủ, cân đối, các phần liên kết chặt chẽ. - Trình bày khoa học, sạch đẹp, chữ viết rõ ràng, ít hoặc không có lỗi chính tả.
2. Nội dung tả	- Chỉ liệt kê các sự vật, thiếu sự quan sát. - Nội dung nghèo nàn, chưa làm nổi bật được vẻ đẹp của đêm trăng.	- Tả được các sự vật trong đêm trăng nhưng còn chung chung, thiếu sự tinh tế. - Đã làm nổi bật được một vài nét đẹp của đêm trăng.	- Quan sát và tái hiện được nhiều sự vật, hiện tượng một cách sinh động, cụ thể. - Nội dung sâu sắc, làm nổi bật được vẻ đẹp và không khí riêng của đêm trăng.
3. Ngôn từ và hình ảnh	- Từ ngữ đơn điệu, lặp lại. - Chưa sử dụng các biện pháp tu từ hoặc dùng chưa phù hợp.	- Sử dụng được từ ngữ gợi tả, gợi cảm nhưng còn hạn chế. - Đã biết sử dụng một vài hình ảnh so sánh, nhân hóa.	- Ngôn từ phong phú, sinh động, giàu hình ảnh. - Sử dụng linh hoạt và hiệu quả các biện pháp tu từ như so sánh, nhân hóa, tạo nên sức gợi cảm cho bài viết.
4. Cảm xúc và liên tưởng	- Bài văn mang tính liệt kê, thiếu cảm xúc cá nhân. - Không có sự liên tưởng hay tưởng tượng.	- Đã thể hiện được một vài cảm xúc, tình cảm với cảnh vật. - Có một vài liên tưởng nhưng còn đơn giản.	- Bộc lộ cảm xúc chân thành, sâu sắc, hòa quyện với cảnh vật. - Có những liên tưởng, tưởng tượng phong phú, độc đáo, tạo nên nét riêng cho bài viết.

3. KẾT LUẬN

Có thể khẳng định rằng, trí tuệ nhân tạo không chỉ là một công cụ công nghệ hiện đại mà còn là trợ thủ đắc lực trong việc hỗ trợ giáo viên đổi mới phương pháp dạy học viết văn miêu tả phong cảnh. Thông qua việc ứng dụng AI một cách chủ động, sáng tạo và có định hướng sư phạm, giáo viên có thể giúp học sinh lớp 5 phát triển năng lực quan sát, cảm thụ và biểu đạt bằng ngôn ngữ giàu hình ảnh, cảm xúc. Tuy nhiên, để đảm bảo việc tích hợp AI vào dạy học mang lại hiệu quả thực chất, cần có những lưu ý quan trọng. Trước hết, giáo viên cần biết cách thiết lập và điều chỉnh câu lệnh (prompt) phù hợp với mục tiêu bài học và năng lực của học sinh, tránh tình trạng sử dụng câu lệnh chung chung, rập khuôn khiến phản hồi AI trở nên thiếu gắn kết. Thứ hai, cần đặc biệt quan tâm đến vấn đề đạo đức trong sử dụng AI - học sinh cần được hướng dẫn cách khai thác AI như một công cụ hỗ trợ tư duy, không sao chép nguyên văn các nội dung do AI tạo ra mà phải biết chắt lọc, chỉnh sửa, bổ sung theo phong cách cá nhân. Ngoài ra, giáo viên cũng cần xây dựng các tiêu chí đánh giá sản phẩm viết có tích hợp AI để bảo đảm công bằng, khách quan và khuyến khích học sinh phát triển năng lực viết thực chất. Như vậy, ứng dụng AI trong dạy học viết văn không chỉ là giải pháp công nghệ,

mà còn là cơ hội giáo dục toàn diện, góp phần xây dựng môi trường học tập tích cực, sáng tạo và nhân văn trong nhà trường tiểu học. Đây chính là một hướng đi tiềm năng trong hành trình đổi mới giáo dục, giúp học sinh tiểu học không chỉ viết tốt mà còn biết viết bằng trái tim, bằng tư duy độc lập của chính các em.

TÀI LIỆU THAM KHẢO

1. Bộ Giáo dục và Đào tạo (2018), *Chương trình giáo dục phổ thông - Môn Ngữ văn* (Ban hành kèm theo Thông tư số 32/2018/TT-BGDĐT ngày 26 tháng 12 năm 2018 của Bộ trưởng Bộ Giáo dục và Đào tạo).

2. Phan Văn Các (2003), *Từ điển Hán - Việt*, NXB Văn hóa Thông tin.

3. Lương Thị Hiền (2023), *Ứng dụng trí tuệ nhân tạo tạo sinh trong dạy Tiếng Việt như một ngoại ngữ*, Kỉ yếu hội thảo Ngôn ngữ Máy tính, NXB Dân trí.

4. Bùi Mạnh Hùng (2024), *Tiếng Việt 5 tập 1, 2*, Bộ sách *Kết nối tri thức với cuộc sống*, Nxb Giáo dục.

5. Nguyễn Thị Ly Kha, Trịnh Cam Ly (đồng chủ biên) (2024), Bộ sách *Chân trời sáng tạo*, Nxb Giáo dục.

6. Nguyễn Quang Ninh (2007), *Tiếng Việt thực hành* (Tài liệu đào tạo giáo viên tiểu học), Nxb Giáo dục.

7. Hoàng Phê (2003), *Từ điển Tiếng Việt*, NXB Đà Nẵng.

8. Nguyễn Minh Thuyết (2024), *Tiếng Việt 5 tập 1, 2*, Bộ sách *Cánh Diều*, Nxb Đại học Sư phạm thành phố Hồ Chí Minh.

Thông tin người gửi

Đỗ Lam Ngọc

Học vị: Thạc sĩ

Khoa Giáo dục Tiểu học, Trường ĐHSP Hà Nội 2

Email: dolamngoc@hpu2.edu.vn

PHÂN LOẠI PHÁT NGÔN THÙ GHÉT TIẾNG VIỆT BẰNG MÔ HÌNH NGÔN NGỮ LỚN ĐƯỢC TINH CHỈNH THEO HƯỚNG CHỈ DẪN

Lê Tỷ Khánh[1] - Nguyễn Hoàng Minh Nhật[2] - Lê Thanh[3]

Tóm tắt: Phát hiện phát ngôn thù ghét là yếu tố then chốt để duy trì môi trường giao tiếp an toàn trên mạng xã hội, đặc biệt trong tiếng Việt với cú pháp phức tạp và nhiều biến thể phi chính thống. Nghiên cứu này đánh giá khả năng của mô hình Mistral 7B Instruct trong phân loại nhị phân (độc hại – không độc hại) trên bộ dữ liệu UIT-ViCTSD. Bằng cách áp dụng tinh chỉnh hiệu quả tham số (PEFT) kết hợp QLoRA, mô hình được huấn luyện với chi phí thấp nhưng vẫn đạt hiệu năng cao. Kết quả cho thấy Mistral 7B đạt độ chính xác 91,5% và điểm F1-macro 75,4%, vượt trội so với các mô hình dựa trên BERT vốn cần tiền xử lý hoặc tăng cường dữ liệu. Ngoài việc nâng cao hiệu quả phân loại, phương pháp còn giảm ảnh hưởng của mất cân bằng nhãn mà không đòi hỏi xử lý thủ công. Nghiên cứu khẳng định tiềm năng ứng dụng LLM trong phát hiện phát ngôn thù ghét tiếng Việt và gợi mở hướng xây dựng hệ thống kiểm duyệt hiệu quả, dễ triển khai và minh bạch.

Từ khóa: Phát ngôn thù ghét; Mô hình ngôn ngữ lớn; Xử lý ngôn ngữ tự nhiên tiếng Việt; Tinh chỉnh mô hình; LLMs; Hate speech detection; PEFT; QloRA

I. GIỚI THIỆU

Sự lan rộng của phát ngôn thù ghét trên các nền tảng mạng xã hội đã làm dấy lên những lo ngại nghiêm trọng về mức độ an toàn và chất lượng của các cuộc thảo luận trực tuyến. Trong bối cảnh tiếng Việt, vấn đề này càng trở nên phức tạp hơn do các đặc điểm ngôn ngữ như việc sử dụng tiếng lóng, từ viết tắt, hiện tượng chuyển mã ngôn ngữ, và lối diễn đạt mỉa mai ngầm, khiến cho bài toán phát hiện phát ngôn thù ghét trở nên đặc biệt thách thức.

Các nghiên cứu trước đây trong lĩnh vực phát hiện phát ngôn thù ghét tiếng Việt chủ yếu dựa trên các mô hình thuộc họ BERT như PhoBERT, BERT4News và XLM-R. Những phương pháp này thường kết hợp các kỹ thuật hỗ trợ như tiền xử lý văn bản, tăng cường dữ liệu bằng Easy Data Augmentation (EDA), và điều chỉnh hàm mất mát (chẳng hạn như Focal Loss) [1] nhằm giải quyết vấn đề mất cân bằng nhãn. Mặc dù đạt được kết quả khả quan, các phương pháp này vẫn phụ thuộc nhiều vào các luật thủ công và quy trình thiết kế theo kinh nghiệm, dẫn đến hạn chế về khả năng tổng quát hóa và giải thích. Hơn nữa, các bước tiền xử lý và tăng cường dữ liệu thường tốn kém về mặt thời gian và khó áp dụng rộng rãi cho các miền hoặc bộ dữ liệu khác nhau.

Sự xuất hiện của các mô hình ngôn ngữ lớn (LLMs), đặc biệt là các kiến trúc kiểu giải mã như Mistral và LLaMA, đã mở ra cơ hội cho việc tinh chỉnh các mô hình có dung lượng lớn cho các tác vụ phân loại đặc thù. Tuy nhiên, hiện vẫn còn rất ít nghiên cứu thực nghiệm

[1] Trường Đại học Công Thương TP. Hồ Chí Minh-HUIT; Email: khanhlt@huit.edu.vn
[2] Trường Đại học Văn Hiến TP. Hồ Chí Minh; Email: nhatnhm@vhu.edu.vn
[3] Trường Đại Học Công Nghệ TP. Hồ Chí Minh-HUTECH; Email: lthanh25nct@hutech.edu.vn

đánh giá hiệu quả của các mô hình LLM trong bài toán phát hiện phát ngôn thù ghét tiếng Việt đặc biệt là trên những bộ dữ liệu sử dụng ngôn ngữ chính thống như UIT-ViCTSD .

Trong nghiên cứu này, chúng tôi đề xuất một quy trình tinh chỉnh đơn giản nhưng hiệu quả dựa trên LLMs áp dụng cho bộ dữ liệu tiếng Việt. Bằng cách sử dụng kỹ thuật tinh chỉnh hiệu quả tham số (PEFT) thông qua QLoRA, chúng tôi điều chỉnh mô hình cho nhiệm vụ phân loại nhị phân (độc hại và không độc hại) mà không cần áp dụng bất kỳ kỹ thuật tăng cường dữ liệu hay tiền xử lý văn bản nào. Mô hình sau tinh chỉnh đạt độ chính xác 91,50% và điểm F1-macro 75,40%, vượt qua các phương pháp trước đó.

Các đóng góp chính của nghiên cứu bao gồm:

- Chúng tôi đánh giá khả năng của mô hình LLM kiểu giải mã, cụ thể là Mistral 7B, trong bài toán phát hiện phát ngôn thù ghét tiếng Việt.
- Chúng tôi chứng minh rằng quy trình tinh chỉnh dựa trên PEFT có thể vượt qua các hệ thống dựa trên BERT, ngay cả khi không sử dụng tiền xử lý hay tăng cường dữ liệu.
- Chúng tôi phân tích hiệu năng của mô hình trên một bộ dữ liệu thiên lệch nhãn và mang tính chính thống, từ đó cung cấp góc nhìn mới về hiệu quả của các mô hình LLM hiện đại đối với các tác vụ phân loại ngôn ngữ ít tài nguyên.

II. TỔNG QUAN NGHIÊN CỨU LIÊN QUAN

Phát hiện phát ngôn thù ghét hiện đang là một chủ đề nghiên cứu trọng điểm trong cả bối cảnh đa ngôn ngữ lẫn các ngôn ngữ ít tài nguyên. Các phương pháp ban đầu trong phân loại phát ngôn thù ghét tiếng Việt chủ yếu dựa trên các kỹ thuật học máy truyền thống như SVM, Random Forest và Naive Bayes, với các đặc trưng thủ công như TF-IDF hoặc n-grams [2]. Tuy nhiên, những phương pháp này gặp khó khăn trong việc nắm bắt sắc thái ngữ nghĩa hàm ẩn và tính ngữ cảnh của phát ngôn thù ghét, đặc biệt trên nội dung phi chính thống từ mạng xã hội.

Sự phát triển của các mô hình dựa trên kiến trúc Transformer đã mang lại cải thiện đáng kể về hiệu suất phân loại. Trong tiếng Việt, nhiều mô hình BERT đơn ngữ và đa ngữ đã được giới thiệu, bao gồm PhoBERT [3], BERT4News[4] , XLM-R [5], và VELECTRA [6]. Các mô hình này đã được ứng dụng rộng rãi trên các bộ dữ liệu chuẩn như ViHSD và UIT-ViCTSD [1], [7]. Đặc biệt, PhoBERT kết hợp với CNN [8] hoặc các kỹ thuật hội tụ mô hình đã cho thấy hiệu năng cạnh tranh. Điểm F1-macro được báo cáo dao động từ khoảng 65% đến 74% trên các bộ dữ liệu này.

Nhằm xử lý vấn đề mất cân bằng nhãn, một số nghiên cứu đã áp dụng kỹ thuật tăng cường dữ liệu bằng EDA [9] và điều chỉnh hàm mất mát thông qua Focal Loss [10]. Trong [1] Lưu và cộng sự đã đề xuất phương pháp EDA được điều chỉnh cho tiếng Việt, kết hợp WordNet tiếng Việt và vector nhúng từ fastText. Trong khi đó, Focal Loss được chứng minh là hiệu quả hơn trong các bài toán phân loại đa lớp như ViHSD [11], so với các bài toán phân loại nhị phân như UIT-ViCTSD [7]. Tuy nhiên, các phương pháp này thường yêu cầu các bước tiền xử lý

đặc thù theo miền như tách từ (ví dụ: VnCoreNLP [12]), loại bỏ từ dừng, và áp dụng các quy tắc chuẩn hóa, dẫn đến tăng độ phức tạp và làm giảm khả năng tái lập kết quả.

Ngược lại, các mô hình ngôn ngữ lớn kiểu giải mã như Mistral [13] và LlaMA [14] gần đây đã cho thấy hiệu năng vượt trội trong các tác vụ phân loại và suy luận mà không cần can thiệp kỹ thuật đặc thù cho từng nhiệm vụ. Tuy vậy, hiệu quả của các mô hình này trong bối cảnh ngôn ngữ ít tài nguyên như tiếng Việt, đặc biệt với các bài toán phân loại nhị phân mất cân bằng, vẫn chưa được nghiên cứu đầy đủ.

Trong phần tiếp theo, chúng tôi sẽ trình bày phương pháp tiếp cận trong việc tinh chỉnh một mô hình ngôn ngữ lớn theo kiến trúc giải mã cho bài toán phát hiện phát ngôn thù ghét tiếng Việt, sử dụng tập dữ liệu UIT-ViCTSD. Khác với các nghiên cứu trước chủ yếu dựa vào mô hình BERT [15] kiểu mã hóa cùng với các chiến lược tiền xử lý hoặc tăng cường dữ liệu được thiết kế thủ công, phương pháp của chúng tôi nhấn mạnh vào ba yếu tố chính: tính đơn giản, khả năng tái lập, và hiệu quả tham số.

III. PHƯƠNG PHÁP NGHIÊN CỨU

1. Mô hình Mistral 7B Instruct

Chúng tôi sử dụng mô hình Mistral 7B Instruct [13], một mô hình ngôn ngữ chỉ bao gồm thành phần giải mã được thiết kế chuyên biệt cho các tác vụ theo hướng chỉ dẫn. Khác với các mô hình mã hóa như BERT [16], Mistral được tối ưu hóa cho việc sinh ngôn ngữ theo chiều nguyên nhân, cho phép mô hình vừa thực hiện phân loại, vừa suy luận bằng ngôn ngữ tự nhiên.

Để tinh chỉnh mô hình một cách hiệu quả, chúng tôi áp dụng khuôn khổ QLoRA [2], cho phép tinh chỉnh hiệu quả tham số (PEFT [17]) trên các mô hình đã được lượng hóa, phù hợp với phần cứng phổ thông. Cụ thể, chúng tôi chỉ huấn luyện một số lượng nhỏ các mô-đun chuyển tiếp có hạng thấp, chiếm khoảng 2–3% tổng số tham số, trong khi giữ nguyên toàn bộ trọng số gốc của mô hình.

2. Thiết kế câu lệnh hướng dẫn

Nhằm phù hợp với mục tiêu huấn luyện theo hướng chỉ dẫn của mô hình, chúng tôi áp dụng một định dạng prompt đơn giản và nhất quán theo phong cách Alpaca, được điều chỉnh cho tiếng Việt. Mỗi mẫu huấn luyện bao gồm một phần chỉ dẫn mô tả nhiệm vụ phân loại, theo sau là một bình luận cần gán nhãn. Mô hình được kỳ vọng sẽ trả về duy nhất một từ *toxic* hoặc *non-toxic* ở dòng cuối cùng.

```
{
  "instruction": "Bạn sẽ được cung cấp một bình luận.\nHãy phân loại bình luận vào một trong hai nhóm:\n- toxic: Có chứa ngôn ngữ công kích, thô tục hoặc gây tổn hại.\n- non-toxic: Là bình luận lịch sự và không gây hại.\nChỉ trả về một từ ở dòng cuối cùng: toxic hoặc non-toxic.",
  "input": "Bình luận: Đồ ngu như bò!",
```

```
 "output": "toxic"
}
```

Cấu trúc prompt này đảm bảo tính rõ ràng, phù hợp với nhiệm vụ và dễ dàng xử lý đối với các mô hình đã được huấn luyện theo phương pháp chỉ dẫn. Khác với các phương pháp dựa trên lập luận đa bước hoặc giải thích chi tiết (chẳng hạn như chain-of-thought), định dạng này không yêu cầu suy luận hay sinh văn bản nhiều lượt, từ đó phù hợp với mục tiêu tinh chỉnh có giám sát (Supervised Fine-Tuning – SFT) trong các tác vụ phân loại. Tất cả các tập dữ liệu (huấn luyện, kiểm thử, và kiểm định) đều được chuyển đổi theo định dạng chuẩn này.

IV. THỰC NGHIỆM

1. Tập dữ liệu

Chúng tôi tiến hành các thí nghiệm trên bộ dữ liệu UIT-ViCTSD [7], một tập chuẩn dùng cho bài toán phát hiện phát ngôn độc hại trong tiếng Việt. Tập dữ liệu bao gồm các bình luận của người dùng được thu thập từ VnExpress một trong những trang báo điện tử lớn tại Việt Nam. Mỗi bình luận đều được gán nhãn thủ công thành một trong hai nhóm: *toxic* (độc hại) hoặc *non-toxic* (không độc hại). Nhằm mô phỏng điều kiện triển khai mô hình trong môi trường thực tế, chúng tôi sử dụng một cách chia cố định cho ba tập dữ liệu với kích thước và phân bố nhãn như sau:

Bảng 1. Phân bố nhãn trong tập dữ liệu UIT-ViCTSD[7]

Tập dữ liệu	Số lượng mẫu	Non-Toxic	Toxic
Huấn luyện	7.000	6.241	759
Kiểm định	2.000	1.768	232
Kiểm thử	1.000	890	110

Tập dữ liệu trong Bảng 1 thể hiện sự mất cân bằng nhãn đáng kể, với tỷ lệ bình luận độc hại chỉ chiếm khoảng 10,8% trong tập huấn luyện. Đáng chú ý, chúng tôi không áp dụng bất kỳ kỹ thuật tăng cường dữ liệu hoặc biện pháp cân bằng nhãn nào như lặp mẫu hay điều chỉnh trọng số. Điều này cho phép đánh giá khả năng học nội tại của các mô hình LLM hiện đại khi xử lý phân bố dữ liệu mất cân bằng mà không cần đến các thủ thuật bổ trợ.

Ngoài ra, chúng tôi không thực hiện bất kỳ bước tiền xử lý nào đối với văn bản (chẳng hạn như chuyển về chữ thường, tách từ, loại bỏ từ dừng hoặc ký tự cảm xúc). Tất cả các bình luận đều được đưa vào mô hình ở dạng thô, nhằm bảo toàn tối đa sắc thái ngữ nghĩa và đặc điểm cú pháp của tiếng Việt.

2. Cấu hình thực nghiệm

Quá trình huấn luyện được thực hiện trên một GPU phổ thông dành cho người dùng cuối NVIDIA RTX 4060 Ti với bộ nhớ 16GB VRAM. Mặc dù hạn chế về phần cứng, cấu hình QLoRA cho phép tinh chỉnh mô hình 7B một cách hiệu quả mà không vượt quá giới hạn bộ

nhớ. Mô hình được tinh chỉnh bằng quy trình tinh chỉnh có giám sát (Supervised Fine-Tuning – SFT). Tập huấn luyện, đã được định dạng theo phong cách Alpaca tiếng Việt (xem Mục 3.2), được lưu dưới dạng tệp JSON và nạp qua giao diện của thư viện datasets thuộc Hugging Face. Cấu hình chính cho quá trình tinh chỉnh mô hình được trình bày trong Bảng 2. Các siêu tham số huấn luyện cụ thể được trình bày ở Bảng 3.

Bảng 2. Cấu hình QLoRA cho tinh chỉnh mô hình

Thành phần	**Giá trị thiết lập**
Mô hình	Viet-Mistral/Vistral-7B-Chat (LLM được huấn luyện theo chỉ dẫn)
Lượng hóa	QLoRA 4-bit (LoftQ = "DEFAULT")
Độ chính xác tính toán	bfloat16
Hạng của LoRA	8
Tham số Alpha (LoRA)	64
Tỷ lệ dropout (LoRA)	0.0
Tham số bias	Không sử dụng
Các mô-đun mục tiêu	q_proj, k_proj, v_proj, o_proj, gate_proj, up_proj, down_proj
RSLoRA	Bật
Kiểm soát bộ nhớ trung gian (Gradient Checkpointing)	Bật
Tỷ lệ tham số huấn luyện	Khoảng 2–3% tổng trọng số mô hình

Bảng 3. Các siêu tham số tinh chỉnh mô hình

Siêu tham số	**Giá trị thiết lập**
Tốc độ học (Learning rate)	5e-5
Bộ tối ưu hóa (Optimizer)	AdamW (8-bit)
Trọng số phạt (Weight decay)	0.01
Kích thước batch	8
Tích lũy gradient	4
Độ dài chuỗi tối đa	1024
Số epoch	2

Bộ điều chỉnh tốc độ học	Cosine
Số bước khởi động (Warm-up)	200
Nhóm theo độ dài chuỗi	Có sử dụng (Enabled)

3. Chỉ số đánh giá

Chúng tôi sử dụng các chỉ số sau để đánh giá hiệu năng mô hình:

- Độ chính xác (Accuracy)
- F1-macro
- Ma trận nhầm lẫn (Confusion matrix)

Trong bối cảnh tập dữ liệu có sự mất cân bằng nghiêm trọng giữa các nhãn, chúng tôi đặc biệt nhấn mạnh điểm F1-macro. Khác với các chỉ số như accuracy vốn có thể bị lệch do nhãn chiếm đa số, F1-macro đánh giá hiệu suất một cách công bằng giữa hai lớp *toxic* và *non-toxic* bằng cách tính trung bình không trọng số điểm F1 của từng lớp.

4. Các phương pháp so sánh

Để đánh giá hiệu quả của mô hình đề xuất, chúng tôi đã thực hiện phân tích so sánh với một loạt các phương pháp cơ sở tiêu biểu, bao gồm BERT (phiên bản cased và uncased), PhoBERT, BERT4News, XLM-R, DistilBERT [1]. Tất cả các mô hình cơ sở đều được đánh giá dựa trên điểm số macro-F1, thống nhất với tiêu chí đánh giá của hệ thống chúng tôi nhằm đảm bảo tính khách quan và công bằng trong so sánh.

5. Kết quả và thảo luận

Bảng 4. So sánh hiệu suất trên tập kiểm thử UIT-ViCTSD

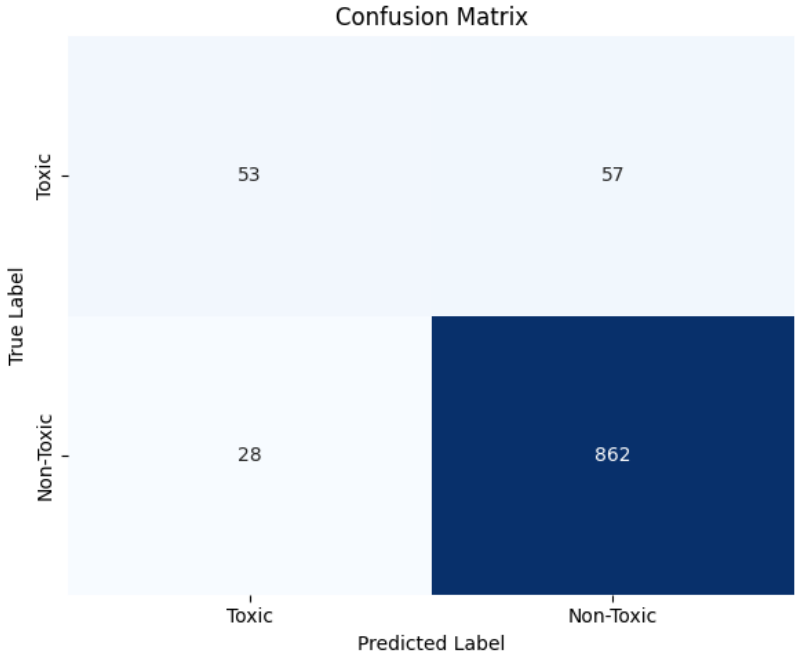
Mô hình	**Accuracy (%)**	**F1-macro (%)**
BERT (cased)	-	68.40
BERT (uncased)		52.84
PhoBERT	-	74.63
BERT4News	-	74.58
DistilBERT	-	69.67
XLM-R	-	73.57
Mistral 7B	91.50	75.40

Trong phần này, chúng tôi đánh giá hiệu năng của mô hình Mistral 7B Instruct sau khi được tinh chỉnh cho bài toán phát hiện phát ngôn thù ghét tiếng Việt, sử dụng tập kiểm thử giữ

lại gồm 1000 mẫu. Các chỉ số đánh giá bao gồm độ chính xác, F1-macro và ma trận nhầm lẫn chi tiết.

Kết quả trong Bảng 4 cho thấy hiệu quả vượt trội của mô hình Mistral 7B Instruct sau tinh chỉnh so với các phương pháp phát hiện phát ngôn thù ghét tiếng Việt trước đây dựa trên kiến trúc BERT. Các mô hình thuộc họ BERT như PhoBERT và BERT4News, được cải tiến bằng EDA hoặc Focal loss đã cho thấy hiệu suất cạnh tranh (F1-macro của PhoBERT khoảng 74.63%) [1]. Mặc dù không sử dụng bất kỳ kỹ thuật tiền xử lý, tăng cường dữ liệu, hay cân bằng nhãn nào, mô hình của chúng tôi vẫn đạt điểm F1-macro cao nhất (75.40%). Ngược lại, mô hình Mistral 7B có thể học được ranh giới phân loại rõ ràng trực tiếp từ văn bản gốc mất cân bằng, thông qua phương pháp tinh chỉnh hiệu quả tham số với QLoRA. Điều này cho thấy khả năng khái quát mạnh mẽ của các mô hình LLM kiểu giải mã, đặc biệt khi được kết hợp với prompt theo cấu trúc chỉ dẫn. Tuy nhiên, phân tích chi tiết ma trận nhầm lẫn trong hình 1 cũng chỉ ra một điểm hạn chế quan trọng: mô hình đạt hiệu quả rất cao với lớp chiếm ưu thế (*non-toxic*), nhưng gặp nhiều khó khăn với lớp thiểu số (*toxic*), khi chỉ đạt mức recall 48,2% cho các mẫu độc hại.

Mặc dù còn tồn tại những thách thức, kết quả của chúng tôi cho thấy rằng các mô hình LLM được huấn luyện theo chỉ dẫn có thể đạt hiệu quả ngang bằng, thậm chí vượt trội so với các pipeline BERT truyền thống ngay cả trong điều kiện huấn luyện hạn chế về tài nguyên. Điều này mở ra hướng nghiên cứu mới trong việc phát triển các hệ thống phát hiện phát ngôn thù ghét có khả năng mở rộng và dễ giải thích cho các ngôn ngữ ít tài nguyên.



Hình 1. Ma trận nhầm lẫn trên tập kiểm thử UIT-ViCTSD

V. KẾT LUẬN

Nghiên cứu này đề xuất một phương pháp hiệu quả cho bài toán phát hiện phát ngôn thù ghét tiếng Việt bằng cách tinh chỉnh mô hình LLM theo kiến trúc giải mã. Không cần tiền xử lý phức tạp hay hàm mất mát chuyên biệt, mô hình sử dụng prompt kiểu Alpaca đơn giản và kỹ thuật QLoRA để đạt F1-macro 75,40% và độ chính xác 91,50%, vượt trội so với các mô hình BERT trước đó dù huấn luyện trên dữ liệu gốc và bằng GPU phổ thông. Kết quả cho thấy LLM có tiềm năng cao trong các tác vụ phân loại văn bản xã hội tiếng Việt. Hướng nghiên cứu tiếp theo sẽ tập trung vào nâng cao khả năng phân biệt lớp thiểu số bằng kỹ thuật suy luận ngữ cảnh, khai thác tri thức ngoài và học tương phản, đồng thời kết hợp mô-đun giải thích nhẹ nhằm tăng tính minh bạch cho hệ thống.

TÀI LIỆU THAM KHẢO

1. Luu, S.T., Van Nguyen, K., Nguyen, N.L.T.: An approach of data augmentation to improve the performance of BERTology models for Vietnamese hate speech detection. Multimed Tools Appl. 83, 56763–56783 (2024). https://doi.org/10.1007/S11042-023-16968-5/METRICS
2. Fortuna, P., Nunes, S.: A Survey on Automatic Detection of Hate Speech in Text. ACM Computing Surveys (CSUR). 51, (2018). https://doi.org/10.1145/3232676
3. Nguyen, D.Q., Tuan Nguyen, A.: PhoBERT: Pre-trained language models for Vietnamese. In: Cohn, T., He, Y., and Liu, Y. (eds.) Findings of the Association for Computational Linguistics: EMNLP 2020. pp. 1037–1042. Association for Computational Linguistics, Online (2020)
4. Nguyen, Q.N., Phan, T.C., Nguyen, D.V., Van Nguyen, K.: ViSoBERT: A Pre-Trained Language Model for Vietnamese Social Media Text Processing. EMNLP 2023 - 2023 Conference on Empirical Methods in Natural Language Processing, Proceedings. 5191–5207 (2023). https://doi.org/10.18653/V1/2023.EMNLP-MAIN.315
5. Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., Stoyanov, V.: Unsupervised Cross-lingual Representation Learning at Scale. In: Jurafsky, D., Chai, J., Schluter, N., and Tetreault, J. (eds.) Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. pp. 8440–8451. Association for Computational Linguistics, Online (2020)
6. Viet Bui, T., Oanh Tran, T., Phuong Le-Hong, isvnuvn: Improving Sequence Tagging for Vietnamese Text using Transformer-based Neural Models, https://aclanthology.org/2020.paclic-1.2/, (2020)
7. Nguyen, L.T., Van Nguyen, K., Nguyen, N.L.-T.: Constructive and Toxic Speech Detection for Open-domain Social Media Comments in Vietnamese. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 12798 LNAI, 572–583 (2021). https://doi.org/10.1007/978-3-030-79457-6_49
8. Quoc Tran, K., Trong Nguyen, A., Hoang, P.G., Luu, C.D., Do, T.H., Van Nguyen, K.: Vietnamese hate and offensive detection using PhoBERT-CNN and social media streaming data. Neural Comput Appl. 35, 573–594 (2023). https://doi.org/10.1007/S00521-022-07745-W/METRICS
9. Wei, J., Zou, K.: EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks. EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference. 6382–6388 (2019). https://doi.org/10.18653/V1/D19-1670
10. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P.: Focal Loss for Dense Object Detection. Proceedings of the IEEE International Conference on Computer Vision. 2017-October, 2999–3007 (2017). https://doi.org/10.1109/ICCV.2017.324
11. Luu, S.T., Nguyen, K. Van, Nguyen, N.L.-T.: A Large-scale Dataset for Hate Speech Detection on Vietnamese Social Media Texts. Presented at the (2021)
12. Vu, T., Nguyen, D.Q., Nguyen, D.Q., Dras, M., Johnson, M.: VnCoreNLP: A Vietnamese Natural Language Processing Toolkit. NAACL HLT 2018 - 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language

Technologies, Proceedings of the Demonstrations Session. 56–60 (2018). https://doi.org/10.18653/V1/N18-5012

13. Jiang, A.Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D.S., Casas, D. de las, Bressand, F., Lengyel, G., Lample, G., Saulnier, L., Lavaud, L.R., Lachaux, M.-A., Stock, P., Scao, T. Le, Lavril, T., Wang, T., Lacroix, T., Sayed, W. El: Mistral 7B. (2023)
14. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., Lample, G.: LLaMA: Open and Efficient Foundation Language Models. (2023)
15. Devlin, J., Chang, M.-W., Lee, K., Google, K.T., Language, A.I.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. 4171–4186
16. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, http://arxiv.org/abs/1810.04805, (2019)
17. Han, Z., Gao, C., Liu, J., Zhang, J., Zhang, S.Q.: Parameter-Efficient Fine-Tuning for Large Models: A Comprehensive Survey. (2024)

AN EXPLORATORY CASE STUDY ON THE IMPACT OF CHATGPT ON ARGUMENTATIVE THINKING IN IELTS WRITING TASK 2

Ngô Thị Thu Hiền[1], Trần Phương Anh[2]

Abstract

This paper evaluates how ChatGPT can influence the representation of arguments in IELTS Writing Task 2 by conducting an exploratory case study of an intermediate-English-level student in an online course of IELTS preparation taught by a private tutor. This study employed a qualitative research design, including a case study approach with essay analysis and a semi-structured interview, to investigate the impact of ChatGPT on an intermediate English level student's argumentative thinking process. The participants' IELTS scores, as assessed by an independent examiner and reported in this study, provided an overview of the impact of ChatGPT on argumentative writing ability. The study's rigor was stimulated by various information, detailed analysis of the participants' thoughts, and transparency. Descriptive statistics were additionally presented to complement the qualitative analysis. The findings indicated that ChatGPT enhances an intermediate-level English learner's argumentative thinking skills, increasing their IELTS writing score to Band 6.5, improving logical arrangement, word usage, grammatical structure, and retaining independent argument evaluation.

***Keywords*:** ChatGPT, writing for IELTS Task 2, argumentative thinking, learning with AI, online learning.

Introduction

Artificial Intelligence (AI) has rapidly developed and made significant changes to many sectors, for example, education (Ahmad et al., 2021). In places where English is studied as a foreign language, AI is now being used in teaching, learning, and assessment (Singh & Hiran, 2022). Many people are using ChatGPT, developed by OpenAI, since it can have a positive impact on the writing quality of their students, particularly in terms of improving the quality of their presentation and organization (Marzuki et al., 2023). Due to being trained on a wide range of information, ChatGPT could assist learners in improving their grammar skills (Malik et al.,2023) and writing skills (Pratama & Hastuti, 2024).

Many students from non-English-speaking countries often find writing Task 2 to be much harder than other sections in the IELTS test (Nguyen et al., 2022). According to the British Council, the task includes writing at least 250 words utilizing an academic style, arguing a point, expressing opinions, and analyzing the questions given. Students are asked to explain their views clearly, provide sound examples to support them, and bring their ideas together in the given time (IELTS, 2023). Consequently, composing an argumentative essay is a

[1] ThS, cựu sinh viên cao học Trường Đại học Ngoại ngữ - Tin học TP. HCM, 828 Sư Vạn Hạnh, Quận 10, TP. HCM, email: friendship200561@gmail.com
[2] ThS,cựu sinh viên cao học Trường Đại học Ngoại ngữ - Tin học TP. HCM, 828 Sư Vạn Hạnh, Quận 10, TP. HCM, email: phuonganh.tr1089@gmail.com

challenging task for any examinee, particularly for those who lack proficiency in English (Zhu, 2001). Additionally, the time limit makes the assignment much more difficult because test takers have to finish the task in about 45 minutes (Ariani & Arham, 2020).

At the same time, reliance on AI-assisted technology might lead to fewer opportunities for students to use their own critical thinking skills (Darwin et al., 2024). Although students may benefit from what ChatGPT produces, sometimes overusing it might stop them from getting involved actively with the language learning task (Kasneci et al., 2023). More specifically, improving argumentative thinking is key to preparing for IELTS Writing Task 2, and it is not yet clear if ChatGPT promotes or prevents the development of this skill (Drushlyak et al., 2024).

Until now, most studies have centred on the role of AI in enhancing language learning, such as grammar, vocabulary, and writing abilities (Khampusaen, 2025; Nguyen & Nguyen, 2022; Wei, 2023; Wei-Xun & Jia-Ying, 2024). Nonetheless, there is not much research on how using these tools influences students' ability to construct arguments by themselves, mainly when preparing for standardized tests. Moreover, private online learning spaces are being used with ChatGPT, but not much is known about this because these settings are seldom studied.

Research objectives

- To investigate to what extent ChatGPT is used to or influences the development of argumentative thinking in the process of writing.
- To analyse what the student understands about the role of ChatGPT in coming up with ideas, structuring arguments, the assessment of correctness and logic of arguments.
- To recognize the opportunities and challenges of ChatGPT as an assisting tool in preparing for IELTS at the intermediate level regarding writing tasks.

Research questions

- How does ChatGPT influence the argumentative thinking process of an intermediate-English-level student preparing for IELTS Writing Task 2?
- What is the perception of an intermediate-English-level student on how ChatGPT affects her capability to generate, structure, and evaluate arguments in IELTS Writing Task 2?

Literature Review

Overview of AI in education

According to Harry (2023), AI can enhance personalized learning by tailoring the learning experience to each student's individual needs and abilities, leading to improved learning outcomes and increased student engagement. It can also increase efficiency by automating repetitive tasks like grading and data analysis, freeing up time for teachers and students to focus on more meaningful tasks. AI can also enhance student engagement by offering interactive and engaging learning experiences, such as chatbots and virtual assistants

(Chen et al., 2023). Additionally, AI can facilitate more effective data analysis, enabling teachers to gain a deeper understanding of their students and tailor instruction more precisely, ultimately leading to improved learning outcomes and student performance (Ayeni et al., 2024).

Nonetheless, there are challenges related to the application of AI in education (Dukhanina & Maksimenko, 2020). First, Chan and Hu (2023) discovered that Gen AI in learning introduces problems: precision, visibility, privacy, and ethics. Specifically, potential ethical concerns are the possibility of personal data theft and plagiarism (Chan & Hu, 2023). Second, excessive utilization of AI may impair the development of learning skills and decrease critical thinking skills (Melisa et al., 2025). Third, the future career market may experience job substitution and more requirements to be hired, including HR professionals, AI designers, electronic scientists, legal scholars, etc. (Martin-Hernandez, 2023). Cassidy (2024) stated that the influx of AI-assisted cheating has begun to impact the status of Australian universities and highlight the significance of a balance between applying the technology and maintaining existing academic integrity policies.

ChatGPT and Its Use in EFL Writing

Functions of ChatGPT in writing instruction

ChatGPT can support students in developing new ideas for their writing by giving them topics, themes, and points of view that they might not have thought of on their own (Kasneci et al., 2023). ChatGPT is an artificial intelligence, a large language model. Current versions (3.5 and 4.0) were pre-trained on huge amounts of data before 2021, allowing them to learn linguistic patterns and relationships that they then use to write human-like conversational prose when requested. When users enter a prompt into ChatGPT, it analyses it and quickly provides a response based on information gathered through machine learning while engaging with the Internet (Lingard, 2023).

Chat GPT can improve students' writing skills by enhancing vocabulary and grammar. It provides immediate feedback, allowing students to identify areas for improvement and hone their skills in real-time. Chat GPT exposes students to a wide array of lexical items and sentence structures, enhancing their vocabulary and language proficiency (Amini & Susan, 2024). According to Su et.al (2023), ChatGPT, a refined GPT 3.5 chatbot, can improve natural language understanding, response efficiency, and adaptability, potentially aiding students in argumentative writing.

How learners use ChatGPT

ChatGPT is often used by learners to create essay outlines before they start writing (AlAfnan, 2023). When students enter a topic or prompt, they get a properly structured document that starts with an introduction, has main points explained in body paragraphs, and ends with a conclusion. As a result, they can save time and arrange their ideas more efficiently.

In this case, ChatGPT tends to aid students who have difficulties structuring or organizing their thoughts on a topic (Nouara & Dehbia, 2023).

Many people turn to ChatGPT to assist them in understanding synonyms and how to paraphrase sentences (Jia Min, 2024). Learners frequently write sentences and expect to receive more useful alternatives academically. As a result, their writing becomes better and continues to differ each time.

Many learners also use ChatGPT to check their grammar rules and find ways to improve how transitions are used in their writing (Schmidt-Fajlik, 2023). Any time students are not sure about verb tenses, choosing appropriate articles, or agreeing with a verb with the subject, they can ask ChatGPT for help. In the same way, many learners want to improve their ability to blend ideas smoothly between paragraphs using necessary transition words and expressions. As a result of support, students can express their ideas more clearly, and their texts are better structured (Sudrajad & Putra, 2024).

Learning opportunities

Students can gain useful experience in writing by using ChatGPT (Javaid et al., 2023). An important positive is supporting students in creating a logical and cohesive structure (Ismail, 2023). If learners request essay samples from ChatGPT, they get to see how ideas can be arranged and linked logically. Students learn how to order their writing so that it is easy to read and understand.

Furthermore, ChatGPT prepares individuals for writing academic papers. As learners read and think about the responses, they get used to academic words, sentence formats, and the way writers present their ideas (Khampusaen, 2025). Getting used to formal academic language can, step by step, shape their way of writing in their schoolwork.

Risks

A disadvantage of ChatGPT is that you might use its words without truly understanding the material (Azaria et al., 2024). Some learners may take what is written and fail to grasp how it is made, sometimes ignoring grammar as well. By not learning from the tool, these students might depend too much on it in the future.

Continuing to use ChatGPT to solve all our issues may lead to a loss of original ways of thinking (Bai & Su, 2023). If students rely only on the suggestions from the model, they might never learn to think independently. Eventually, this can lower their chance of giving original ideas and being creative when faced with problems or writing tasks.

Writing Task 2 Band Descriptors

According to the official website of the British Council, the IELTS Writing Task 2 is scored using four points that are equally important: Task Response, Coherence and Cohesion,

Lexical Resource, and Grammatical Range and Accuracy. Each item listed above is essential in determining a candidate's band score, which ranges from Band 0 to Band 9.

Task Response looks at how effectively the candidate responds to the question asked. A Band 9 score indicates that the candidate presents a clear and well-developed viewpoint that is tenaciously supported with relevant ideas. Every detail of the question should be addressed by the candidate. A weak band score may be a result of the problems mentioned above.

Coherence and Cohesion mean the way ideas are organized logically and how authors use links to connect them. An organized essay (at level nine or eight) separates each idea logically and holds the discussion together naturally. Paragraphing should also be done effectively. A lack of proper linking words, innocent blunders in organizing ideas, or missing information in paragraphs may reduce the marks the examiner awards.

Lexical Resource deals with ensuring that students have enough vocabulary, use it correctly, and choose suitable expressions. You have to use a wide range of words properly and include many unique or colorful phrases to get a high band score. Mistakes in choosing words, repeating them too much, or having a small vocabulary may stop people from communicating well and reduce their score.

Grammatical Range and Accuracy checks if the person can use several grammatical structures correctly. People aiming for a high band need to write correct sentences of all kinds with very few grammatical mistakes. The use of sentence structures increases the score, while many errors result in a low score.

Argumentative Thinking in IELTS Writing Task 2

Importance of argumentation

According to the official website of IDP, writing task 2 of the IELTS test requires students to express their views on an issue clearly and justify them effectively. Unlike Task 1, Task 2 mainly tests your skill in forming sensible and structured arguments about social, educational, or environmental topics. A person’s ability to argue well is essential in this task as it shows the use of critical thinking, good language skills, and writing skills needed in school. Critical thinking serves as an additional cognitive method for self-regulation employed by learners in their educational pursuits (Phan, 2010). To get a high score on the essay, candidates should be good at certain argumentative skills. It’s essential to start by clearly articulating their thoughts or main points in the introduction. Having a strong thesis means the essay will follow a path, and all the supporting ideas will be focused. The next step is for test-takers to arrange and list suitable supporting arguments for their case. Both the points and the examples used to support them should be easy to understand by anyone. Often, a strong essay presents different viewpoints, which makes the writer’s argument more reasonable and convincing. To be effective, an argument must be both coherent and cohesive, along with relevant content (Candelo et al., 2018). Ideas need to follow a logical order, and every paragraph should help create the structure. Using transitional terms and helpful paragraphs makes their work clear

and interesting to readers. Because of this, exercises in argumentative thinking, like using debate, will help learners succeed in IELTS as well as in academic or professional situations where English is used (Daria & Natalia, 2020).

Common learner difficulties

A significant challenge is that most people lack sufficient training in critical thinking (Facione & Facione, 2013). In many school systems that rely heavily on exams, students are often asked to memorize facts rather than think critically about them. For these reasons, some learners may struggle with carefully analyzing a question, viewing different perspectives, and thinking ahead to possible objections—skills that are important in argumentative writing. If students do not practice using sound judgment, checking facts, and forming conclusions, their answers usually depend on simple ideas or personal beliefs without strong proof. Because of this, candidates' arguments look simple or unconvincing to the IELTS examiners (Lloyd-Jones & Medaney, 2007).

Another important challenge is that many learners find it hard to talk about complex subjects in English (Rass, 2015). While students may be good at reasoning in one language, this does not mean they can do it so easily in a second language. Constructing arguments logically asks for proper vocabulary, grammar, and the skill to outline ideas persuasively (Cockcroft et.al, 2013). A lack of language skills usually makes it difficult for learners to talk about cause and effect, compare different opinions, or arrange their arguments logically. Consequently, their arguments are not well-developed, their ideas are presented several times, or they avoid making the text complex. Learners often prefer to stick to uncomplicated sentences, which reduces the strength of their argument.

Many learners do not use discourse markers and cohesive devices well in IELTS Writing Task 2, so this area is very important for scoring (Panahi & Mohammaditabar, 2015). Various students tend to use only "and," "but," or "because" to link ideas, when a stronger relation might be more necessary. A few writers end up using the same linking phrases many times, which leads to frequent and sometimes rigid changes within their writing. Some learners leave out these markers, which results in paragraphs that are confusing or hard to follow. A great essay connects its thoughts by making use of "in contrast," "therefore," "as a result," and "despite this." Misuse of such vocabulary makes the essay unclear and may affect how well the essay fits together.

Methodology

Research Samples

This study included a female participant who is in Grade 12 at Thai Binh specialized high school. The student took part in the private course after having passed the PET (Preliminary English Test), proving that she was at an intermediate level in English. When the study was conducted, the student had learned IELTS for around one year and completed the curriculum for Cambridge IELTS Bands 4–5.

To track the development of argumentative thinking in response to ChatGPT-assisted learning, two writing samples were collected at different stages of the study. Essay 1 (pre-ChatGPT) was prepared by the student in the middle of June, before the student started using ChatGPT. It showed the unassisted argumentative thinking of the student. The student was tested in IELTS on June 21st, achieving Band 6. The Essay 2 (Post-ChatGPT) was written in the middle of July after the student had practiced utilizing ChatGPT in terms of brainstorming and planning writing activities regularly. Both essays were written under exam-like conditions within a 40-minute time limit and without external assistance at the moment of writing (see Appendix D).

The essay topics were chosen based on the common IELTS Writing Task 2 issues, such as the concern with labor time restrictions (pre-ChatGPT essay) and an examination of the preference of vocational training over university education (post-ChatGPT essay). Both of the samples were analyzed according to the official IELTS Writing Band Descriptors and argumentative thinking dimensions. Table 1 illustrates the summary of essay samples, as follows:

Table 1

Summary of Essay Samples and Analytical Criteria

Aspect	**Essay 1 (Pre-ChatGPT)**	**Essay 2 (Post-ChatGPT)**
Time of Writing	Mid-June 2025	Mid-July 2025
ChatGPT Usage	Not used	Used for brainstorming and planning
Writing Conditions	40 minutes, exam-like, no assistance	40 minutes, exam-like, no assistance
Essay Topic	Labor time restrictions	Crimes

Research Design

This study mainly employed a qualitative research design that included a case study approach with essay analysis and a semi-structured interview. The primary objective of the study was to determine in detail how ChatGPT affects the process of argumentative thinking

of an intermediate English level student and how this participant generates, organizes, and evaluates arguments as part of writing task 2 in IELTS tests.

The choice of a qualitative case study allows studying the experiences of a particular participant in detail and context, which would be most suitable to address the questions of exploratory studies (Scapens, 2004). Since argumentative thinking is associated with a complicated manner of thinking, e.g., reasoning, critical reflection, and decision-making, a case study enables the capturing of the shades and depth of these processes (Simons, 2014), which large-scale quantitative methods could overlook.

The IELTS scores were obtained in the official test of the participant taken and marked by an independent examiner who was not part of the study. In that way, the scores provided researchers with a comprehensive overview of the writing capacity of the participant during the implementation of ChatGPT and do not incorporate a direct evaluation of the two essays reviewed in this research paper. Moreover, this study did not intend to generalize the results obtained to all learners who study IELTS but rather sheds light on the process of communication of a particular learner with ChatGPT for preparing writing tasks. The single-participant approach, therefore, was applied in this study to explore the role of an individual in developing argumentative thinking skills under technological-aided learning.

Probably, the study design was limited to the involvement of one participant, but the rigor of the study was stimulated by:

- Various information: both pre- and post-writing samples, right after the usage of ChatGPT, and an in-depth interview.
- Thick description: detailed analysis of the process of writing and the thoughts of the participant.
- Transparency: open admission of the scope and limitations of the study.

The descriptive statistics that included word counts and the number of arguments raised in the essays were additionally presented to complement the qualitative analysis and to give additional context through which the writing development of the participant is to be interpreted.

Research Instruments

To determine the effect of ChatGPT as an instructional intervention in the argumentative thinking training for the writing task 2 of the IELTS test, these instruments were employed to analyse pre-ChatGPT and post-ChatGPT essays: (1) argumentative thinking dimensions; (2) descriptive statistics. In addition, a semi-structured interview was conducted to examine students' ability to construct, develop, and evaluate arguments independently after consulting ChatGPT.

Argumentative Thinking Dimensions:

Several studies have demonstrated that essay writing will not be effective without critical thinking, well-developed ideas, organization, and coherence. Pei et al.(2022)

determined that high levels of critical thinking were associated with English essay writing, including relevance, clarity, logic, depth, and flexibility of the essay. In addition, Le's (2013) study found that high-scoring argumentative essays were those that utilized flexible patterns of topic progression throughout the text. Furthermore, many studies reported that students' argumentative essays relied heavily on referencing and linking devices to achieve coherence (Afriani, 2023). Overall, proficient argumentative essay writers need to balance grammatical coherence with flexibility in vocabulary and structure, a trait reflected in the higher-level IELTS writing criteria.

The researchers applied a qualitative evaluative approach to analyze the development of arguments and employed four criteria based on the established principles of grading in IELTS and prior research. The following were the focus areas: (1) Idea Generation, (2) Critical Analysis, (3) Argument Organization, and (4) Cohesiveness. The two essay samples were discussed descriptively to inspect the existence and development of these dimensions. Table 2 illustrates the description of argumentative thinking dimensions and their supporting research.

Table 2

Dimensions of Argumentative Thinking and Supporting Research

Dimension	Supporting Sources	Description
Idea Generation	Pei et al. (2022)	Describes how a student can generate useful, thoughtful, and innovative concepts that address various perspectives of the topic.
Critical Analysis	Pei et al. (2022)	Involves the assessment of arguments, identification of counter-arguments, and a synthesis of logic and inferences.
Argument Organization	Le (2013)	Concerning the structure and logic throughout the essays. A strong essay demonstrates flexible but coherent development of the topic.
Cohesiveness	Afriani (2023);	Associates with the application of cohesive devices like reference and conjunctions in linking ideas to enable fluency in the essays.

Note: The four dimensions were adapted into the evaluative framework utilized in this research to examine the IELTS Writing Task 2 of the participants in a qualitative approach.

Descriptive Statistics

Descriptive statistics were selected in this study to complement the qualitative assessment of the writing samples. These specific statistics, when combined with the qualitative criteria and with the IELTS score descriptors, provided a more detailed picture of the students' development of argumentative thinking. In particular, the following characteristics were analyzed:

- Word count: This will be used to calculate the length and content density of each essay.
- Number of body paragraphs: To see the continuity and organization of the structure.
- Number of main arguments: To assess the strength of the argument and its complexity.
- Number of counterarguments: To determine the level of exposure that students have to critical thinking.

Semi-structured Interview:

An interview was carried out to examine the perceptions of the student on how ChatGPT impacted her thinking. According to Adhabi and Anozie (2017), an interview is a consultation process where a researcher seeks to understand an individual's opinion on an issue, driven by a reputable purpose, involving an interactive process of questioning. Since the participant has a challenging academic plan in a gifted high school, it is difficult to schedule a face-to-face interview via Google Meet. The interviewee is an IELTS learner whose schedule is filled with school and extracurricular activities that hardly give any time to real-time interaction. Considering this limitation, the researchers decided to use Gmail to complete the interview. This approach is more flexible as the participant can answer thoughtfully whenever it is convenient according to his schedule (Bowden & Galindo-Gonzalez, 2015). In addition, email interviews may generate more thoughtful and articulate answers because the respondents are offered additional time to reflect on their answers (Dahlin, 2021). Not only does this respect the time demands of the participants, but it also guarantees that the data gathered is rich and insightful.

The interview has three main themes, which outline essential issues of argumentative writing in the IELTS Writing Task 2: the generation of ideas, the structure of arguments, and the autonomy of the arguments. The eight open-ended questions are determined and placed in each theme to learn how students use ChatGPT after their writing practice.

Theme 1: Generation of Ideas (Questions 1-4). These are the questions that investigate the usage of ChatGPT to compose an essay. Question 1 is about the real activities that students perform when using ChatGPT in content generation. Question 2 investigates their overall perception of the role of ChatGPT. Question 3 will strongly address how ChatGPT can help students generate ideas and arguments. Question 4 seeks to determine if this dependence influences their creative thinking.

Theme 2: Organising arguments (Question 5). In this question, the researcher examines the role of ChatGPT in developing the skills of students to organize and rationally

arrange their essays. It seeks to find out whether ChatGPT can be classified as an example or a replacement for students' planning and writing paragraphs.

Theme 3: Independent arguments evaluation (Questions 6-8). These last questions evaluate the problem-solving ability and independence of students in the use of ChatGPT. Question 6 inquires about the improvement in their ability to analyze and evaluate arguments induced by ChatGPT. Question 7 asks how willing they would be to agree or disagree with suggestions made by ChatGPT. Question 8 examines the larger implications of ChatGPT on the capacity of students to think independently and how ChatGPT stimulates or limits such capacity.

This thematic organization helps to establish that the interview is not only effective in capturing how ChatGPT gets used by students, but it also allows one to have a close understanding of how the tool is contributing towards cognitive development in argumentative writing. The development of ideas and independent judgment ensures a complete picture of the pedagogical implications of ChatGPT.

Research Procedure

The author, along with a colleague, watched the Cambridge University pedagogical guidance video on ChatGPT and other literature material on AI-assisted academic writing before performing the instructional experiment. Based on these materials, researchers created instructional examples that showed how ChatGPT could assist students in generating argumentative ideas, organizing essays, and improving coherence and clarity in the IELTS Writing Task 2.

From 26th April 2025 to 20th June 2025, the student practiced under a controlled routine before she took the initial IELTS test on June 21st. She was only allowed two argumentative essays per week without access to the internet or any AI guidance, and in a determined time limit of 40 minutes, under constant parental supervision. This was to test real exam conditions and the independent argumentative skill of the student.

From 22nd June 2025 to 15th August 2025, to prepare for the second test on the 16th of August, the student was enabled to use ChatGPT during the writing process. This change of practice brought a marked difference between these two phases, and it helped us explore the possible role of ChatGPT on the argumentative thinking of the student in the IELTS Writing Task 2.

Example 1: This was a list of examples that the teacher asked ChatGPT to develop argumentative thinking.

🔹 Example 1: Generating Arguments

> **Prompt:**
> "Give me two arguments for and against the idea that university education should be free for everyone."

ChatGPT's Possible Output:

For:

1. Free university education promotes equal access and reduces social inequality.
2. It can lead to a more educated workforce and boost economic productivity.

Against:

1. The cost burden would fall on taxpayers and may be unsustainable for governments.
2. It could lead to overcrowding and reduce the perceived value of higher education.

🔹 Example 2: Exploring Counterarguments

> **Prompt:**
> "I believe that technology makes people more isolated. Can you provide one counterargument with supporting details?"

ChatGPT's Possible Output:

While technology may reduce face-to-face interactions, it actually allows people to maintain long-distance relationships and connect with others across the world, which can enhance social bonds in new ways.

🔹 Example 3: Improving Coherence and Paragraph Flow

> **Prompt:**
> "Check the coherence of this paragraph and suggest improvements:
> 'People use mobile phones every day. They can take photos. Many apps are on phones. Society changes fast.'"

ChatGPT's Possible Output:

The paragraph lacks logical flow. Suggested revision:
"Mobile phones have become an essential part of daily life. Beyond making calls, they allow users to take high-quality photos and access a wide range of applications. As a result, they have significantly influenced the way society communicates and functions."

🔹 Example 4: Topic Sentences and Support

> **Prompt:**
> "Write a topic sentence and two supporting details for the idea: 'Social media affects young people's mental health.'"

ChatGPT's Possible Output:

Topic Sentence: Social media has a profound impact on the mental health of young people.

Supporting Details:

1. Constant comparison with others online can lead to anxiety and low self-esteem.
2. Excessive screen time disrupts sleep patterns, which further affects emotional well-being.

To sum up, on the basis of the examples, ChatGPT can become a convenient mechanism in building argumentative thinking in students. Through assisting them in brainstorming pertinent ideas, viewing the issue or topic in several ways, and arranging the arguments in a coherent manner, ChatGPT offers scaffolded guidance, which is in line with the cognitive requirements of the IELTS Writing Task 2. Also, its interactive and contextualised feedback could help to increase the motivation of the learners and engage them more in the writing process.

Example 2: Using ChatGPT, it is possible to produce premium argumentative essays with a sense of logic and thinking. Reinforcing the structure and thinking behind arguments greatly assists learners preparing to take the IELTS exam, particularly in tackling Task 2. Although the minimum age to use ChatGPT is 13 (as mentioned in the policy of OpenAI), it can be a valuable aid to older teens and adults in the process of academic writing. An example of an effective classroom activity is to request that students complete an IELTS Task 2 essay first, with the help of one of the common types of questions (e.g., opinion or discussion essay). Then, students copy the same prompt into ChatGPT and contrast the answer provided by the generator with their answers. They are advised to evaluate the introduction, clarity of the thesis, logical connection of the body paragraphs, soundness of the supporting ideas, counterargument usage, and lexical variety.

Directions: **Some people think that governments should spend more money on public services rather than on the arts, like music and paintings. To what extent do you agree or disagree? You write your own essays in 40 minutes.**

ChatGPT's Sample Output:

> While art has undeniable cultural value, I strongly believe that governments should prioritize public services such as healthcare and education. These sectors have a direct impact on the quality of life and well-being of citizens.
>
> For example, investing in public hospitals can reduce disease rates and improve life expectancy, whereas funding art galleries may only benefit a small segment of society. Furthermore, public education equips individuals with skills to contribute to the economy, ensuring national development.
>
> However, some argue that art enriches our lives and promotes creativity. Although this is true, such benefits are secondary to essential services. Without proper healthcare and education, individuals may not even have the opportunity to appreciate the arts.
>
> In conclusion, while art should not be completely neglected, public services should take precedence in government spending due to their broader societal impact.

Based on the above example, it can be concluded that ChatGPT can be a powerful tool to help students improve their argumentative thinking when it comes to the IELTS Writing Task 2. With the ability to compare their essays to the properly structured AI-generated answers, students are more likely to pay attention to the way they need to construct logical arguments, to include counterarguments, and to arrange their ideas coherently. Such reflective comparison contributes not only to the improvement of the quality of the essay but also leads to the development of critical thinking, which is the key to succeeding in academic writing tasks.

Following the second writing assignment, there was a semi-structured interview with the participant to discuss how ChatGPT influenced their argumentative thinking. The thematic coding was utilized to analyze the interview data after the transcript. Review of the data responses showed three broad themes regarding the second research question:

Theme 1: Generation of ideas

Theme 2: Organising arguments

Theme 3: Independent arguments evaluation

These themes aided in not only revealing the usability of ChatGPT but also how the student used it cognitively, whether this was sustaining, substituting, or supplementing independent thinking.

Findings

In this section, a comparison of the two IELTS Writing Task 2 essays that the participant wrote in both cases, with and without ChatGPT, is carried out using the official IELTS Writing Task 2 Band Descriptors (see appendix B). In particular, the researchers focus on critical analysis, idea generation, organization, and cohesion of each essay.

Essay Analysis

IELTS Scoring Comparison

The participants took two IELTS writing tests: (1) on 21st June 2025; (2) on 16th August 2025 (see Appendix C). The first IELTS writing test was scored B6 and was taken before the ChatGPT intervention. The second test was scored higher at Band 6.5 after practicing argumentative writing with ChatGPT under the guidance of the researchers for two months. Table 3 presents the IELTS test results before and after implementing ChatGPT for practicing writing arguments.

Table 3

Summary of IELTS Test Dates and Band Scores

IELTS Test Date	21st June 2025	16th August 2025
IELTS Scores	6	6.5

Argumentative Thinking Comparison

Regarding argumentative thinking, substantial improvement was also examined once the student started to apply ChatGPT for practicing writing essays. The ideas expressed in the first essay were rather common. The participant appeared to be primarily concerned with the surface-level arguments without considering alternative points of view, more complicated reasoning. The second essay, however, demonstrated more insight into the given topic. The participants created larger and more reflective ideas and thoughts on both sides of the arguments. In addition, the essay provided better critical analysis because it discussed long-term effects and social consequences that are not the basic advantages and disadvantages.

Moreover, in the second essay, the points were more naturally connected, with much higher cohesiveness. The student reflected on a high level of association between ideas, which made it seem that they had enhanced organizational strategies in their preparation, which was supported by ChatGPT. A summary of the differences in argumentative thinking in the two essays is provided in Table 4 below:

Table 4

Comparison of Argumentative Thinking Dimensions in Pre-ChatGPT and Post-ChatGPT Essays

Argumentative Thinking Dimension	**Essay 1 (pre-ChatGPT)**	**Essay 2 (Post-ChatGPT)**
Idea Generation	Valid but conventional; surface-level ideas	Broader, more thoughtful, and deeper ideas

Critical Analysis	Focus on basic reasoning; have limited perspectives	Address both sides; evaluate long-term and societal impacts
Argument Organization	Simple, linear structure; basic connections between ideas	Complex, well-structured arguments with smooth transitions
Cohesiveness	Somewhat mechanical; basic linking words	Natural, seamless flow with stronger connections between points

Descriptive Statistics of Argumentative Writing Samples

The writing samples provided by the participant were analyzed in terms of word count, the number of arguments they claimed, and the instantiation of argumentation. These descriptive statistics are summarised in Table 5 below:

Table 5

Descriptive Statistics Comparison in Pre-ChatGPT and Post-ChatGPT Essays

Feature	Essay 1 (pre-ChatGPT)	Essay 2 (post-ChatGPT)
Word Count	287 words	303 words
Number of Body Paragraphs	2	2
Main Arguments	Two key arguments: (1) health benefits of fewer working hours, (2) an increase in work productivity due to shorter working hours	Three main arguments: (1) to continue their education by enrolling in universities after high school, (2) to go and work in the fields as a way of acquiring work experience, and (3) to achieve a balance between work and education to prepare them for the future
Counterarguments	Scanty mention of opposing views; largely biased	Two well-developed counterarguments. Concentrate on strengths and weaknesses; consider various points of view

Word Count. Essay 2 has 16 additional words compared to Essay 1. This implies that the student could contribute more details of ideas, arguments, explanations, and reasoning after being trained with ChatGPT. Although the increase in word count between the two essays was

quite small, the student produced more complex arguments and delved deeper into the topic analysis in the second essay.

Number of Arguments. Essay 1 presents just two major arguments, which do not deal with the opposing arguments but simply support the position of the essay. In comparison with the second essay, the latter demonstrates three central arguments and is highly active in the consideration of several viewpoints, including the advantages and disadvantages of higher education, the benefits of direct employment, and the potential of combining these two. This shows an increase in the level of argumentative thinking as the student presented more ideas and critical reasoning.

Complexity of Argument Structure. Essay 1 has the general structure of introduction, two body paragraphs, and a conclusion, where each of the two body paragraphs offers one argument. The second essay, however, is more of a complex structure since there are several arguments interlinked in the body paragraphs. This means that she has improved in arranging ideas as well as integrating the arguments to provide a better critical analysis.

Interview

The responses were grouped according to the following three themes: (1) Idea generation, (2) organizing arguments, and (3) independent argument evaluation. These themes represent the key criterion of argumentative thinking focused on in this work and assisted in providing insights into the perception of the participant on how ChatGPT influences their capacity to produce, arrange, and review arguments.

Idea Generation

The interviewee mentioned that she often uses ChatGPT when she wants to collect ideas and argumentative samples, including those that are new. Additionally, the interviewee highlighted that ChatGPT provided a variety of perspectives, topic-specific vocabulary sets, and grammatical structures, which helped her explore the topics more deeply. The interviewee responded:

“I frequently use it to seek ideas and specific examples for my writing. …. as well as relevant vocabulary and grammatical structures related to the topics."

“ChatGPT assists me in generating ideas and arguments by giving different viewpoints and useful information, especially on unfamiliar topics….”

Using ChatGPT has contributed to enhancing the creativity and critical thinking of the interviewee and helped her develop creative and critical thinking. The interviewee emphasized:

"Yes, I believe using ChatGPT has positive effects on my ability to generate ideas and arguments independently. By providing diverse perspectives and examples, ChatGPT enhances my critical and creative thinking. It works like a useful starting point encouraging me to develop and refine my own thoughts."

Organizing Arguments

The interviewee noted that the structure and clarity of her essays improved after practicing with ChatGPT. Specifically, the tool provided the participant with well-organized outlines and clear transitions, which helped her plan and compose their essays more effectively:

"ChatGPT has helped me improve the efficiency and organization of my essay writing... I can manage my time better and feel less pressure when writing."

Additionally, the interviewee agreed that the tool assisted her in managing and presenting ideas more systematically.

"It allows me to create clear outlines and express ideas effectively."

Independent Argument Evaluation

While the interviewee expressed appreciation for ChatGPT's support, they were also aware of ChatGPT's limitations. Specifically, the interviewee expressed that the arguments presented by ChatGPT did not always align with their own views:

"Certainly, ChatGPT doesn't always provide arguments that match my views... I simply revise my prompt."

This shows that the interviewee also acknowledged the risk of over-reliance:

"Since many topics are unfamiliar, I sometimes depend on it too much, which restricts my own ideas."

However, the interviewee reflected positively on its overall impact:

"It broadens my view by offering new angles, helping me think more deeply and improve my writing."

Discussion

The influence of ChatGPT on the argumentative thinking process of an intermediate-English-level student preparing for IELTS Writing Task 2

The most significant improvement was in the performance of the IELTS writing score, especially in the writing task 2 section. She increased her IELTS scores from 6 to 6.5 in her writing areas, which include coherence, cohesion, vocabulary, and control of grammar usage.

Another key discovery was the apparent positive impact of implementing ChatGPT on the level of argumentation of the participant. The arguments presented in Essay 1, which was done prior to ChatGPT, did not reflect deep thinking. Conversely, Essay 2 demonstrated more sophisticated critical thinking. In addition, the participant built three well-organized arguments, assessed the strengths and weaknesses, and discussed the social and long-term impact of the case. These innovations were discovered in accordance with the reports provided by Khampusaen (2024); strategic incorporation of AI tools, such as ChatGPT, can lead to improved logical reasoning and the quality of argumentation in student writing, particularly in the context of learning a second language. Besides, the participant showed sophisticated

abilities of summarization and connection of ideas comprehensively. The transitions between points of argument written in Essay 2 are more natural and meaningful, which suggests that the participant has performed better in terms of cohesion and structuring her essay. This result is consistent with the quantitative analysis conducted by Song and Song (2023), which revealed significant positive changes regarding the features of academic writing skills, such as organization, coherence, grammar, and vocabulary of the participants, due to the implementation of AI language learning technology.

Finally, a quantitative analysis of the two essays justifies a great enhancement in the argumentative writing skill of the participant. Essay 2, after using ChatGPT, was longer (303 words in comparison to 287 words) and contained three key arguments and amplified descriptions. The increase in length and complexity of argumentative structures found in Essay 2 suggests that the participant has improved her critical thinking and developed diverse ideas. These findings are supported by the study of Chen et al. (2025), which stated that the application of ChatGPT and argument map tools significantly enhanced structural fluency, clarity of argument, and critical thinking in students' writing.

The perception of an intermediate-level student on how ChatGPT affects her capability to generate, structure, and evaluate arguments in IELTS Writing Task 2

The interview results reinforced the conclusion that ChatGPT, when utilized strategically, can influence an intermediate English level student’s argumentative thinking positively. In particular, it can facilitate idea generation, improve organizational plans, and critical engagement with content.

The participant regularly used this tool to search for information on unfamiliar topics, appropriate illustrations, as well as to collect related vocabulary and syntax. Similarly, Kasneci et al. (2023) highlight that students found ChatGPT particularly useful for idea generation and planning writing. Further, the participant acknowledged that ChatGPT was an effective tool to create an outline of her essays, but revealed that she sometimes edits and adjusts the generated ideas according to her own opinions. This response was supported by a study by Levine et al. (2024), suggesting that ChatGPT can provide learners with a distributed writing model that eliminates the need to go through the steps of planning, drafting, and reviewing. Additionally, the participant reported that ChatGPT contributed to making her experience less stressful than when she had to write under a time limit. This aligns with previous research by Mehrad et al. (2025), which showed that learners had reduced pressure and increased efficiency when using AI-driven platforms like ChatGPT. In general, the improvement of the participant in controlling and presenting her arguments testifies that AI can act as an assistant to help learners organize complex arguments in a logical and relevant way.

One of the most notable interview findings was the respondent's selective usage of ChatGPT’s instruction. Rather than relying solely on AI-generated content, the respondent revealed she participated in the process of evaluating, editing, and adapting ideas to her own perspective. This is a significant transition from being reliant on external sources of inputs to

being self-reliant in assessing and argumentative determination. However, this finding is inconsistent with the conclusion of Guo and Lee (2023) that students who adopt an available solution without considering it often have a decline in critical thinking ability. Overall, the paper argues that the application of ChatGPT can be considered as a language learning assistant tool to scaffold writing skills, but the actual effectiveness will ultimately be determined by a learner's adaptive skills. In this way, ChatGPT catalyzes enhancing independent argumentative thinking.

Conclusion

This research examined the effect of ChatGPT on the argumentative thinking process of an intermediate-level English language learner who prepares for the IELTS Writing Task 2 test. The major results showed that the IELTS writing score of the participant has increased to Band 6.5. Furthermore, by comparing argumentative thinking, the essay after ChatGPT was assessed to have an improved logical arrangement, word usage, grammatical structure, and a higher level of argument development than the essay before ChatGPT. The participants provided details about the arguments as well as their advantages and disadvantages, and discussed long-term and societal consequences. This evidence indicates that the critical thinking skills of the participant have developed with the help of ChatGPT.

The interview data helped the study understand deeply the participants' perception of the impact of ChatGPT on their ability to generate, structure, and evaluate arguments in the IELTS Writing Task 2. The feedback collected demonstrated that ChatGPT was perceived as a useful resource when generating ideas, arguments, and essay structure. Moreover, the participant noted that the experience of writing with ChatGPT enhanced her time management ability, as well as minimizing the pressure during the preparation of arguments. An important finding was that the research participant retained her ability to independently evaluate arguments when applying ChatGPT through self-adjusting the AI-generated content to her personal perspective.

Implications

This exploratory study provides preliminary evidence that when utilized strategically, ChatGPT can be an effective scaffolding tool to teach argumentative writing to second-language learners. When applying to the IELTS Writing Task 2 preparation, ChatGPT assists students in identifying and considering multiple views and envisioning rebuttals. As a result, learners' critical thinking skills will be developed. Additionally, ChatGPT demonstrated its effectiveness in facilitating language learning through exposing learners to a large and diverse vocabulary, correct grammatical format, and ideal structures of arguments. Furthermore, through the illustration of different types of cohesive devices and their relevance in a particular situation, ChatGPT may contribute to the enhancement of cohesion and logic in writing argumentative essays.

For education stakeholders, this study indicates the possibility of applying AI tools such as ChatGPT as an augmentative support tool to teach English writing skills, especially argumentative writing, but only if learners are guided to approach the material and apply it to suit their thinking process rather than passively receiving it.

Limitations

Although the results are informative about the possible contribution of ChatGPT to the improvement of argumentative thinking to excel in the IELTS Writing Task 2, several limitations need to be addressed to direct future research. First, this investigation is not generalizable, as it represents a single-participant case. Second, the IELTS scores included in the study served the purpose of providing only supplementary marking to manifest the general writing performance of the participant before and after the implementation of ChatGPT. As a result, they failed to quantify changes in specific argumentative reasoning across the two essays analyzed.

Although improvements were observed in the participant, the researchers believe that it cannot be confirmed that ChatGPT was the sole contributing factor. Other possible influencing factors include comprehensive language learning tasks or repeated writing practice that need to be investigated.

Future studies are needed to increase the sample size and diversity, employ a longitudinal approach, and conduct statistical descriptive and inferential tests. These methods will provide more reliable evidence to conclude on the generalizability of positive changes to a variety of learner profiles and learning conditions when applying ChatGPT for practicing writing argumentative essays for IELTS tests.

References

Adhabi, E., & Anozie, C. B. (2017). Literature review for the type of interview in qualitative research. *International journal of education*, *9*(3), 86-97.

Afriani, Z. L. (2023). Cohesive devices used in argumentative essays written by college students. *Linguists: Journal of Linguistics and Language Teaching, 9*(1), 45–58. https://ejournal.uinfasbengkulu.ac.id/index.php/linguists/article/view/Afriani%20Zelvia%20Liska

Ahmad, S. F., Rahmat, M. K., Mubarik, M. S., Alam, M. M., & Hyder, S. I. (2021). Artificial intelligence and its role in education. *Sustainability*, *13*(22), 12902.

AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). ChatGPT as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, *3*(2), 60-68.

Amini, F., & Susanti, A. (2024). Chat GPT: Enhancing Students' Writing Skills for EFL Students in Descriptive Text. *International Journal of Research and Innovation in Social Science*, *8*(10), 2273-2285.

Ariani, N., & Arham, M. (2020). An Analysis of EFL Test Takers' Problems in IELTS Writing Task. *Tamaddun*, *19*(2). https://doi.org/10.33096/tamaddun.v19i2.81

Ayeni, O. O., Al Hamad, N. M., Chisom, O. N., Osawaru, B., & Adewusi, O. E. (2024). AI in education: A review of personalized learning and educational technology. *GSC Advanced Research and Reviews*, *18*(2), 261-271.

Azaria, A., Azoulay, R., & Reches, S. (2024). ChatGPT is a Remarkable Tool—For Experts. Data Intelligence, 6(1). https://doi.org/10.1162/dint_a_00235

Bai, L., Liu, X., & Su, J. (2023). ChatGPT: The cognitive effects on learning and memory. *Brain-X*, *1*(3), e30.

Bowden, C., & Galindo-Gonzalez, S. (2015). Interviewing when you're not face-to-face: The use of email interviews in a phenomenological study. *International Journal of Doctoral Studies*, *10*, 79.

Candelo, J. E., Soto, J. D., Torres, L., Schettini, N., Calle, M., Garcia, L., & De Castro, A. (2018). Coherence and cohesion issues in argumentation documents written by engineering students. *IEEE Global Engineering Education Conference, EDUCON*, *2018-April*. https://doi.org/10.1109/EDUCON.2018.8363222

Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers*, *25*(1), 161-182.

Cockcroft, R., Cockcroft, S., & Hamilton, C. (2013). *Persuading people: An introduction to rhetoric*. Bloomsbury Publishing.

Dahlin, E. (2021). Email interviews: A guide to research design and implementation. *International Journal of Qualitative Methods*, *20*, 16094069211025453.

Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, *11*(1), 2290342.

Drushlyak, M., Lukashova, T., Sabadosh, Y., Melnikov, I., & Semenikhina, O. (2024). Using ChatGPT for the Development of Critical Thinking in Youth: Example of Inequality Proof. *2024 47th ICT and Electronics Convention, MIPRO 2024 - Proceedings*, 334–339. https://doi.org/10.1109/MIPRO60963.2024.10569759

Dukhanina, L. N., & Maksimenko, A. A. (2020). Problems of the implementation of artificial intelligence in education. *Перспективы Науки и Образования Perspectives of Science and Education*, 24.

Facione, P. A., & Facione, N. C. (2013). Critical thinking for life: Valuing, measuring, and training critical thinking in all its forms. *Inquiry: Critical thinking across the disciplines*, *28*(1), 5-25.

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education, 100*(12), 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505

Ismail, H. Y. S. (2023). Cohesion and coherence in essays generated by chatgpt: A comparative analysis to university students' writing. *CDELT Occasional Papers in the Development of English Education*, *83*(1), 143-165.

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through the ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *3*(2), 100115.

Jia Min, C. (2024). The use of AI and ChatGPT in teaching synonyms to EFL students. *Research Studies in English Language Teaching and Learning*, *2*(4), 187–207. https://doi.org/10.62583/rseltl.v2i4.53

Kasneci, E., Sessler, K., Kühl, N. M., Bannert, M., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kasneci, E., Sessler, K., Kühl, N., Bannert, M., & Horz, H. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khampusaen, D. (2024). The impact of ChatGPT on academic writing skills and knowledge: An investigation of its use in argumentative essays. *LEARN Journal: Language Education and Acquisition Research Network. 18*(1), 963-988. https://doi.org/10.70730/PGCQ9242

Khampusaen, D. (2025). The Impact of ChatGPT on Academic Writing Skills and Knowledge: An Investigation of Its Use in Argumentative Essays. *LEARN Journal: Language Education and Acquisition Research Network*, *18*(1), 963-988.

Le, T. D. (2018). *A study of cohesion in Vietnamese EFL students' argumentative essays through thematic progression analysis (TPA)* [Master's thesis, Chulalongkorn University]. Chula ETD. https://doi.org/10.58837/CHULA.THE.2013.1975

Levine, S., Beck, S. W., Mah, C., Phalen, L. & PIttman, J. (2025). How do students use ChatGPT as a writing support? *Journal of Adolescent & Adult Literacy, 68*, 445–457. https://doi.org/10.1002/jaal.1373

Lingard, L. (2023). Writing with ChatGPT: An illustration of its capacity, limitations & implications for academic writers. *Perspectives on medical education*, *12*(1), 261.

Lloyd-Jones, G., Neame, C., & Medaney, S. (2007). A multiple case study of the relationships between students' plural academic progress and IELTS scores at an international postgraduate university.

Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., Darwis, A., & Marzuki. (2023). Exploring Artificial Intelligence in Academic Essay: Higher Education Student's Perspective. *International Journal of Educational Research Open*, *5*. https://doi.org/10.1016/j.ijedro.2023.100296

Martín-Hernández, P. (2023). Artificial Intelligence: The Present and Future of Human Resources Recruitment and Selection Processes. *Engineering Proceedings*, *56*(1), 188. https://doi.org/10.3390/ASEC2023-15521

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2). https://doi.org/10.1080/2331186X.2023.2236469

Mehrad, A., Zangeneh, M. H. T. Z., Bouzedif, M., & Rweramila, N. G. (2025). Educational stress and artificial intelligence tools: A study of international students' experiences in Spanish higher education. *Journal of Asian Social Science Research, 7(*1), 1-32. https://doi.org/10.15575/jassr.v7i1.127

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., ... & Al Fuad, Z. (2025). Critical Thinking in the Age of AI: A Systematic Review of AI's Effects on Higher Education. *Educational Process: International Journal*.

Nguyen, H. N., & Nguyen, D. K. (2022). Vietnamese learners' performance in the IELTS writing task 2: Problems, causes, and suggestions. *International Journal of TESOL & Education*, *2*(1), 170-189.

Panahi, R., & Mohammaditabar, M. (2015). The strengths and weaknesses of Iranian IELTS candidates in academic writing task 2. *Theory and Practice in Language Studies*, *5*(5), 957-967.

Pei, Z., Zheng, C., Zhang, M., & Liu, F. (2017). Critical thinking and argumentative writing: Inspecting the association among EFL learners in China. *English Language Teaching, 10*(10), 31–42. https://doi.org/10.5539/elt.v10n10p31

Phan, H. P. (2010). Critical thinking as a self-regulatory process component in teaching and learning. *Psicothema*, *22*(2), 284-292. https://www.psicothema.com/pi?pii=3728

Pratama, R. M. D., & Hastuti, D. P. (2024). The use of artificial intelligence to improve EFL students' writing skills. *English Learning Innovation (englie), 5*(1), 13–25. https://doi.org/10.22219/englie.v5i1.30212

Rass, R. A. (2015). Challenges Face Arab Students in Writing Well-Developed Paragraphs in English. *English Language Teaching*, *8*(10), 49-59.

Scapens, R. W. (2004). Doing case study research. In *The real life guide to accounting research* (pp. 257-279). Elsevier**.**

Schmidt-Fajlik, R. (2023). ChatGPT as a Grammar Checker for Japanese English Language Learners: A Comparison with Grammarly and ProWritingAid. *AsiaCALL Online Journal*, *14*(1). https://doi.org/10.54855/acoj.231417

Simons, H. (2020). Case study research: In-depth understanding in context. In *The Oxford Handbook of Qualitative Research*. https://doi.org/10.1093/oxfordhb/9780190847388.013.29

Singh, S. V., & Hiran, K. K. (2022). The Impact of AI on Teaching and Learning in Higher Education Technology. *Journal of Higher Education Theory and Practice*, *22*(13). https://doi.org/10.33423/jhetp.v22i13.5514

Song, Y., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, Article 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, *57*, 100752.

Wei Xun, L., & Jia Ying, Z. (2024). Impact of AI-driven language learning apps on vocabulary acquisition among English learners. *Research Studies in English Language Teaching and Learning, 2*(1), 1–11. https://doi.org/10.62583/rseltl.v2i1.32

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, Article 1261955. https://doi.org/10.3389/fpsyg.2023.1261955

Zhu, W. (2001). Performing Argumentative Writing in English: Difficulties, Processes, and Strategies. *TESL Canada Journal*, *19*(1). https://doi.org/10.18806/tesl.v19i1.918

Appendix A

Interview questions

1. How do you use ChatGPT when practicing for IELTS Writing Task 2? What specific tasks or steps do you use it for? *(Con thường sử dụng ChatGPT như thế nào khi luyện tập cho phần viết Task 2 của bài thi IELTS? Hoặc là những nhiệm vụ cụ thể hoặc các bước con đã sử dụng?)*

2. What do you generally think about ChatGPT when it comes to the IELTS Writing Task 2?

(Con suy nghĩ thế nào về việc áp dụng ChatGPT vào việc luyện viết IELTS Writing Task 2?)

3. In what ways can ChatGPT assist you in coming up with ideas or arguments in your essays? *(ChatGPT hỗ trợ con như thế nào trong việc nghĩ ra ý tưởng hoặc lập luận cho bài viết?)*

4. Do you think using ChatGPT has affected your capacity to generate ideas or argument on your own? Please explain. *(Con có nghĩ việc sử dụng ChatGPT đã ảnh hưởng đến khả năng tự nghĩ ý tưởng hoặc lập luận của con không? Vì sao?)*

5. What has been the effect of ChatGPT on your organization and management of essays? *(ChatGPT đã ảnh hưởng như thế nào đến cách con sắp xếp và tổ chức bài viết?)*

6. Do you believe the usage of ChatGPT has contributed to your ability to conduct independent analysis or critique of arguments? Why not or why? *(Con có cho rằng việc sử dụng ChatGPT đã giúp con nâng cao khả năng tự phân tích hoặc đánh giá các lập luận không? Vì sao có hoặc không?)*

7. Have you ever disagreed or agreed with the suggestions of ChatGPT? *(Có khi nào con đồng ý hoặc không đồng ý với các gợi ý của ChatGPT không? Khi đó, con đã xử lý như thế nào?*

8. What did you do about it? In general, does ChatGPT encourage or constrain your thinking on your own in IELTS writing task 2? Why so? *(Nhìn chung, con thấy ChatGPT giúp con suy nghĩ độc lập hơn hay khiến con bị phụ thuộc khi luyện phần viết Task 2 của bài thi IELTS? Vì sao?)*

Appendix B

IELTS Writing Task 2 Band Descriptors

IELTS Writing Task 2 Band Descriptors

Updated May 2023

Please visit IELTS.org for updates

Page 1

Scoring criteria for Academic and General Training tests

A script must fully fit the positive features of the descriptor at a particular level. **Bolded text** indicates negative features that will limit a rating.

Band Score	Task Response	Coherence & Cohesion	Lexical Resource	Grammatical Range & Accuracy
9	The prompt is appropriately addressed and explored in depth. A clear and fully developed position is presented which directly answers the question/s. Ideas are relevant, fully extended and well supported. Any lapses in content or support are extremely rare.	The message can be followed effortlessly. Cohesion is used in such a way that it very rarely attracts attention. Any lapses in coherence or cohesion are minimal. Paragraphing is skilfully managed.	Full flexibility and precise use are widely evident. A wide range of vocabulary is used accurately and appropriately with very natural and sophisticated control of lexical features. Minor errors in spelling and word formation are extremely rare and have minimal impact on communication.	A wide range of structures is used with full flexibility and control. Punctuation and grammar are used appropriately throughout. Minor errors are extremely rare and have minimal impact on communication.
8	The prompt is appropriately and sufficiently addressed. A clear and well-developed position is presented in response to the question/s. Ideas are relevant, well extended and supported. There may be occasional omissions or lapses in content.	The message can be followed with ease. Information and ideas are logically sequenced, and cohesion is well managed. Occasional lapses in coherence and cohesion may occur. Paragraphing is used sufficiently and appropriately.	A wide resource is fluently and flexibly used to convey precise meanings. There is skilful use of uncommon and/or idiomatic items when appropriate, despite occasional inaccuracies in word choice and collocation. Occasional errors in spelling and/or word formation may occur, but have minimal impact on communication.	A wide range of structures is flexibly and accurately used. The majority of sentences are error-free, and punctuation is well managed. Occasional, non-systematic errors and inappropriacies occur, but have minimal impact on communication.
7	The main parts of the prompt are appropriately addressed. A clear and developed position is presented. Main ideas are extended and supported but there may be a tendency to over-generalise or there may be a lack of focus and precision in supporting ideas/material.	Information and ideas are logically organised, and there is a clear progression throughout the response. (A few lapses may occur, but these are minor.) A range of cohesive devices including reference and substitution is used flexibly but with some inaccuracies or some over/under use. Paragraphing is generally used effectively to support overall coherence, and the sequencing of ideas within a paragraph is generally logical.	The resource is sufficient to allow some flexibility and precision. There is some ability to use less common and/or idiomatic items. An awareness of style and collocation is evident, though inappropriacies occur. There are only a few errors in spelling and/or word formation and they do not detract from overall clarity.	A variety of complex structures is used with some flexibility and accuracy. Grammar and punctuation are generally well controlled, and error-free sentences are frequent. A few errors in grammar may persist, but these do not impede communication.

IELTS Writing Task 2 Band Descriptors

Updated May 2023

Please visit IELTS.org for updates

Page 2

Scoring criteria for Academic and General Training tests

A script must fully fit the positive features of the descriptor at a particular level. **Bolded text** indicates negative features that will limit a rating.

Band Score	Task Response	Coherence & Cohesion	Lexical Resource	Grammatical Range & Accuracy
6	The main parts of the prompt are addressed (though some may be more fully covered than others). An appropriate format is used. A position is presented that is directly relevant to the prompt, although the conclusions drawn may be unclear, unjustified or repetitive. Main ideas are relevant, but some may be insufficiently developed or may lack clarity, while some supporting arguments and evidence may be less relevant or inadequate.	Information and ideas are generally arranged coherently and there is a clear overall progression. Cohesive devices are used to some good effect but cohesion within and/or between sentences may be faulty or mechanical due to misuse, overuse or omission. The use of reference and substitution may lack flexibility or clarity and result in some repetition or error. Paragraphing may not always be logical and/or the central topic may not always be clear.	The resource is generally adequate and appropriate for the task. The meaning is generally clear in spite of a rather restricted range or a lack of precision in word choice. If the writer is a risk-taker, there will be a wider range of vocabulary used but higher degrees of inaccuracy or inappropriacy. There are some errors in spelling and/or word formation, but these do not impede communication.	A mix of simple and complex sentence forms is used but flexibility is limited. Examples of more complex structures are not marked by the same level of accuracy as in simple structures. Errors in grammar and punctuation occur, but rarely impede communication.
5	The main parts of the prompt are **incompletely addressed.** The format may be inappropriate in places. The writer expresses a position, but the development is not always clear. Some main ideas are put forward, but they are limited and are not sufficiently developed and/or there may be irrelevant detail. There may be some repetition.	Organisation is evident but is not wholly logical and there may be a lack of overall progression. Nevertheless, there is a sense of underlying coherence to the response. The relationship of ideas can be followed but the sentences are not fluently linked to each other. There may be limited/overuse of cohesive devices with some inaccuracy. The writing may be repetitive due to inadequate and/or inaccurate use of reference and substitution. **Paragraphing may be inadequate or missing.**	The resource is limited but minimally adequate for the task. Simple vocabulary may be used accurately but the range does not permit much variation in expression. There may be frequent lapses in the appropriacy of word choice and a lack of flexibility is apparent in frequent simplifications and/or repetitions. Errors in spelling and/or word formation may be noticeable and may cause some difficulty for the reader.	The range of structures is limited and rather repetitive. Although complex sentences are attempted, they tend to be faulty, and the greatest accuracy is achieved on simple sentences. Grammatical errors may be frequent and cause some difficulty for the reader. Punctuation may be faulty.

IELTS Writing Task 2 Band Descriptors

Updated May 2023

Please visit IELTS.org for updates

Page 3

Scoring criteria for Academic and General Training tests

A script must fully fit the positive features of the descriptor at a particular level. **Bolded text** indicates negative features that will limit a rating.

Band Score	Task Response	Coherence & Cohesion	Lexical Resource	Grammatical Range & Accuracy
4	The prompt is tackled in a minimal way, or the answer is tangential, possibly due to some misunderstanding of the prompt. **The format may be inappropriate.** A position is discernible, but the reader has to read carefully to find it. Main ideas are difficult to identify and such ideas that are identifiable may lack relevance, clarity and/or support. Large parts of the response may be repetitive.	Information and ideas are evident but not arranged coherently and there is no clear progression within the response. Relationships between ideas can be unclear and/or inadequately marked. There is some use of basic cohesive devices, which may be inaccurate or repetitive. There is inaccurate use or a lack of substitution or referencing. There may be no paragraphing and/or no clear main topic within paragraphs.	The resource is limited and inadequate for or **unrelated to the task**. Vocabulary is basic and may be used repetitively. There may be inappropriate use of lexical chunks (e.g. memorised phrases, formulaic language and/or language from the input material). Inappropriate word choice and/or errors in word formation and/or in spelling may impede meaning.	A very limited range of structures is used. **Subordinate clauses are rare and simple sentences predominate.** Some structures are produced accurately but grammatical errors are frequent and may impede meaning. Punctuation is often faulty or inadequate.
3	No part of the prompt is adequately addressed, or the prompt has been misunderstood. No relevant position can be identified, and/or there is little direct response to the question/s. There are few ideas, and these may be irrelevant or insufficiently developed.	There is no apparent logical organisation. Ideas are discernible but difficult to relate to each other. There is minimal use of sequencers or cohesive devices. Those used do not necessarily indicate a logical relationship between ideas. There is difficulty in identifying referencing. Any attempts at paragraphing are unhelpful.	The resource is inadequate (which may be due to the response being significantly underlength). Possible over-dependence on input material or memorised language. Control of word choice and/or spelling is very limited, and errors predominate. These errors may severely impede meaning.	Sentence forms are attempted, but errors in grammar and punctuation predominate (except in memorised phrases or those taken from the input material). This prevents most meaning from coming through. **Length may be insufficient to provide evidence of control of sentence forms.**
2	The content is barely related to the prompt. No position can be identified. There may be glimpses of one or two ideas without development.	There is little relevant message, or the **entire response may be off-topic.** There is little evidence of control of organisational features.	The resource is extremely limited with few recognisable strings, apart from memorised phrases. There is no apparent control of word formation and/or spelling.	There is little or no evidence of sentence forms (except in memorised phrases).
1	**Responses of 20 words or fewer are rated at Band 1.** **The content is wholly unrelated to the prompt.** Any copied rubric must be discounted.	**Responses of 20 words or fewer are rated at Band 1.** The writing fails to communicate any message and appears to be by a virtual non-writer.	**Responses of 20 words or fewer are rated at Band 1.** No resource is apparent, except for a few isolated words.	**Responses of 20 words or fewer are rated at Band 1.** No rateable language is evident.
0	Should only be used where a candidate did not attend or attempt the question in any way, used a language other than English throughout, or **where there is proof that a candidate's answer has been totally memorised.**			

Appendix C

IELTS Test Report (21 June 2025)

IELTS™

Test Report Form ACADEMIC

NOTE *Admission to undergraduate and post graduate courses should be based on the ACADEMIC Reading and Writing Modules.*
GENERAL TRAINING Reading and Writing Modules are not designed to test the full range of language skills required for academic purposes.
It is recommended that the candidate's language ability as indicated in this Test Report Form be re-assessed after two years from the date of the test.
To find out more about IELTS, IELTS band scores and the CEFR levels, please visit ielts.org/scores

Centre Number: VN104 | Date: 21/JUN/2025 | Candidate Number: 512858

Candidate Details

Family Name: PHAM
First Name(s): PHUONG LINH
Candidate ID: 034308002017

Date of Birth: 31/01/2008 | Sex (M/F): F | Scheme Code: Private Candidate
Country or Region of Origin:
Country of Nationality: VIET NAM
First Language: VIETNAMESE

Test Results

Listening	Reading	Writing	Speaking	Overall Band Score	CEFR Level
6.5	8.0	6.0	5.5	6.5	B2

Administrator Comments

Recognising organisations must verify this score at ielts.org/verify

Validation stamp: IELTS

Date: 23/06/2025 | Test Report Form Number: 25VN512858PHAP104A

BRITISH COUNCIL · idp · CAMBRIDGE English

IELTS Test Report (16 August 2025)

IELTS

Test Report Form ACADEMIC

NOTE *Admission to undergraduate and post graduate courses should be based on the ACADEMIC Reading and Writing Modules. GENERAL TRAINING Reading and Writing Modules are not designed to test the full range of language skills required for academic purposes. It is recommended that the candidate's language ability as indicated in this Test Report Form be re-assessed after two years from the date of the test. To find out more about IELTS, IELTS band scores and the CEFR levels, please visit ielts.org/scores*

Centre Number: VN104 | Date: 16/AUG/2025 | Candidate Number: 525201

Candidate Details

Family Name: PHAM

First Name(s): PHUONG LINH

Candidate ID: 034308002017

Date of Birth: 31/01/2008 | Sex (M/F): F | Scheme Code: Private Candidate

Country or Region of Origin:

Country of Nationality: VIET NAM

First Language: VIETNAMESE

Test Results

Listening	Reading	Writing	Speaking	Overall Band Score	CEFR Level
8.0	8.5	6.5	6.0	7.5	C1

Administrator Comments

Recognising organisations must verify this score at ielts.org/verify

Validation stamp: IELTS

Date: 18/08/2025 | Test Report Form Number: 25VN525201PHAP104A

BRITISH COUNCIL idp CAMBRIDGE English

IELTS is jointly owned by the British Council, IDP IELTS, and Cambridge University Press & Assessment

Appendix D

Essay 1 (pre-ChatGPT essay)

Novel regulations for restriction of labor time released in various nations have provoked great debates in recent years. The following essay aims to elaborate on the reasons, followed by my own perspective that this is a positive development that can bring numerous benefits to both individuals and society.

There are several reasons why limitations on working hours are enacted. Firstly, overworking can lead to both serious physical and mental health issues such as stress or heart-related diseases. By curbing working hours, authorities aim to protect workers from such risks. Moreover, while many objectors argue that limiting working hours can have adverse impacts on workers who require to work overtime for additional income, they have ignored the fact that overworking often results in decreased productivity. Numerous studies have shown that long hours do not always accompany better output. In fact, tired workers tend to make more mistakes and perform less efficiently. Last but not least, limiting work hours allows individuals to spend more time with their families, engage in recreational activities, and pursue personal development, thereby fostering a balanced life.

From my perspective, the introduction of this law is clearly a positive change. For employees, shorter working hours can lead to better physical and mental well-being, increased job satisfaction, and improved work-life balance. For employers, healthier workers tend to be more productive and reduce absenteeism. Furthermore, those laws can contribute to a well-developed society by ensuring that all workers have access to fair labor conditions and stable productivity.

Taking all the above-mentioned facts into account, the regulations to limit working hours are encouraged due to the sake of safeguarding workers' health and promoting productivity. Therefore, I firmly believe this is a beneficial policy that supports a better workforce.

Essay 2 (post-ChatGPT essay)

Some people believe that young people who commit serious crimes should be punished in the same way as adults. To what extent do you agree or disagree?

Although the increasing number of juvenile delingquents has been a topic that provoke great debates in recent years, the issue as to whether they should be liable for their severe crimes like adults or not remain controversial. From my perspective, while the severity of the crime must not be underestimated, equal punishment may be inappropriate due to the following reasons.

First of all, various young offenders are too immature to aware the consequences of their illegal acts. Delinquents normally come from troubled families with criminal parents or marriage breakdown. This may lead to their distored outlook on life, cultimating in their immersion in anti-social activities. Besides, as the saying goes, "Who keeps company with the wolf will learn how to howl". Numerous young people are involved in severe crimes due to the negative peer pressure and illiteracy, engendering disorderly behaviours without considering moral issues. Therefore, providing reeducation and psychological therapy methods in proper learning environment for those delinquents may help them recognize the importance of moral values and reduce the influence of adverse factors.

Furthermore, many young people struggling with livelihood problems tend to engage in illegal behaviours due to desparation. Money is the major inducement in numerous crimes such as armed robbery or even murder. Due to the shortage of employment opportunities, many individuals end up embarking on the career of crime to feed themselves and their families. As a result, rehabilitation programs involving vocational training may equip those people with valuable skills and experience to find a proper job, addressing their financial issues. Moreover, reintegration support providing assistance in terms of housing, employment and social

connection after release offer a second chance for all ex-offender to become law-abiding citizens.

Taking all the above-mentioned facts into account, I hold the belief that imposing equal punishment as adults on juvenile offenders is not appropriate. However, enhancing rehabilitation programs is a more efficient altenative.

TIẾP CẬN PHÁT HIỆN TIN GIẢ BẰNG TRÍ TUỆ NHÂN TẠO CÓ GIẢI THÍCH

Đào Xuân Kỳ[1], Đinh Minh Hoà[2], Tiếu Phùng Mai Sương[2]

Tóm tắt: Sự lan truyền nhanh chóng của tin giả trên mạng xã hội đặt ra thách thức lớn đối với độ tin cậy thông tin và nhận thức cộng đồng, đặc biệt trong bối cảnh tiếng Việt khi nghiên cứu còn hạn chế. Nghiên cứu này so sánh hai mô hình học sâu BiLSTM và PhoBERT trong phát hiện tin giả tiếng Việt, đồng thời tích hợp LIME thuộc nhóm Explainable AI (xAI) để giải thích cơ chế quyết định của mô hình. Thí nghiệm trên bộ dữ liệu ReINTEL 2020 cho thấy cả hai mô hình đạt hiệu quả cao, trong đó PhoBERT vượt trội với độ chính xác 93.6% và điểm F1 đạt 80.0%. Việc áp dụng LIME giúp trực quan hóa mức độ đóng góp của từng từ, tăng tính minh bạch và độ tin cậy, qua đó góp phần xây dựng hệ thống phát hiện tin giả vừa hiệu quả vừa dễ giải thích.

Từ khoá: phát hiện tin giả, xử lý ngôn ngữ tự nhiên tiếng Việt, PhoBERT, BiLSTM, Trí tuệ nhân tạo có giải thích

Giới thiệu

Trong bối cảnh mạng xã hội và các nền tảng thông tin trực tuyến ngày càng bùng nổ tại Việt Nam, tình trạng tin giả lan truyền nhanh chóng không chỉ làm méo mó nhận thức cộng đồng mà còn gây suy giảm niềm tin vào báo chí và các nguồn thông tin chính thống. Mặc dù trên thế giới đã có nhiều nghiên cứu ứng dụng học máy và học sâu để phát hiện tin giả một cách hiệu quả, song tại Việt Nam, các nghiên cứu về phát hiện tin giả vẫn còn khá hạn chế, đặc biệt là trong việc triển khai các mô hình mô dạng transformer mạnh mẽ đi kèm với khả năng giải thích (Explainable AI, xAI).

Một số nghiên cứu gần đây đã bước đầu tiến tới khắc phục những hạn chế trên. Ví dụ, Huỳnh và các cộng sự đã phát triển hệ thống phát hiện tin giả tiếng Việt trên mạng xã hội, sử dụng các mô hình transformer, trong đó PhoBERT là một lựa chọn nổi bật, cho độ chính xác và độ tin cậy cao trong bài toán phân loại tin giả tiếng Việt [1]. Tuy nhiên, công trình này vẫn chưa chú trọng đến khía cạnh giải thích quyết định phân loại. Ngoài ra, một nghiên cứu tiên tiến về phát hiện tin giả đa phương thức với khả năng giải thích – áp dụng cho ngôn ngữ Tamil – mặc dù cho thấy tiềm năng cao, lại không phù hợp để áp dụng trong ngữ cảnh tiếng Việt [2]. Tổng quan hơn, Gongane và cộng sự cũng chỉ ra rằng, mặc dù Explainable AI như LIME và SHAP được ứng dụng trong phát hiện tin giả và hate speech trên mạng xã hội, nhưng vẫn thiếu các nghiên cứu định hướng phục vụ ngôn ngữ ít tài nguyên như tiếng Việt [3].

[1] Trường Đại học Công Thương Thành phố Hồ Chí Mính; *Email: kydx@huit.edu.vn*
[2] Trường Đại học Ngoại Ngữ Tin Học Thành phố Hồ Chi Minh; *Email: hoadm@huflit.edu.vn, suongtpm@huflit.edu.vn*
[2] Trường Đại học Công nghệ Thành phố Hồ Chí Minh; *Email: dmhoa24nct@hutech.edu.vn, tpmsuong24nct@hutech.edu.vn*

Trước thực trạng đó, bài báo này xác định các mục tiêu nghiên cứu cụ thể: (i) triển khai và so sánh hiệu năng giữa hai mô hình học sâu – BiLSTM như baseline truyền thống và PhoBERT – mô hình transformer tối ưu cho tiếng Việt, trên bộ dữ liệu phát hiện tin giả; và (ii) tích hợp công cụ giải thích LIME để minh bạch hơn trong quá trình phân loại, tạo điều kiện thuận lợi cho việc kiểm duyệt thông tin về sau.

Những đóng góp chính của nghiên cứu bao gồm: (1) thực nghiệm tỉ mỉ trên bộ dữ liệu ReINTEL 2020 với gần 10.000 mẫu tin thu thập từ báo chí và mạng xã hội, mang tính thực tiễn cao; (2) so sánh chi tiết kết quả giữa mô hình truyền thống và transformer, từ đó khẳng định lợi thế của PhoBERT đối với bài toán tiếng Việt; (3) áp dụng LIME để giải thích các dự đoán, cung cấp cái nhìn minh bạch về cơ chế hoạt động của mô hình; và (4) đề xuất định hướng phát triển tiếp theo, bao gồm thử nghiệm các mô hình transformer tiên tiến hơn, kết hợp dữ liệu đa phương thức và xây dựng công cụ trực quan hỗ trợ người dùng cuối.

Công trình liên quan

Trong những năm gần đây, lĩnh vực phát hiện tin giả đã chứng kiến nhiều bước tiến nổi bật, đặc biệt trong việc ứng dụng các mô hình ngữ cảnh hóa mạnh, các hệ thống đa phương thức và các phương pháp giải thích. Đầu tiên, một nghiên cứu tổng hợp trên tạp chí Information đã chỉ ra rằng các biểu diễn ngữ cảnh như BERT embedding không chỉ cải thiện đáng kể độ chính xác so với các phương pháp truyền thống như TF-IDF hay Bag-of-Words, mà còn khi kết hợp với kỹ thuật cân bằng dữ liệu như SMOTE và hàm mất cân đối class-weighted loss, còn nâng cao độ bền và tính công bằng của mô hình; đồng thời tích hợp các phương pháp như SHAP và LIME giúp tăng khả năng minh bạch trong quyết định phân loại tin giả [4]. Bên cạnh đó, công trình "Interpretable BERT and RoBERTa based Fake News Detection using LIME" trực tiếp áp dụng LIME cho các mô hình transformer (BERT, RoBERTa), mở ra hướng xây dựng hệ thống phát hiện tin giả không chỉ chính xác mà còn dễ giải thích [5].

Về mô hình đa phương thức, một công trình khác sử dụng kiến trúc transformer kết hợp với dữ liệu văn bản và hình ảnh cùng cross-attention đã cải thiện rõ rệt khả năng phát hiện tin giả trên mạng xã hội, minh chứng cho xu hướng kết hợp thông tin bổ trợ bên ngoài văn bản để nâng cao hiệu quả phân loại [6]. Ngoài ra, hệ thống "reasoning-based explainable multimodal fake news detection" áp dụng cho ngôn ngữ Tamil (ngôn ngữ ít tài nguyên) đã cho thấy hiệu quả cao (F1 ≈0,8736) và đặc biệt chú trọng vào giải thích dự đoán mô hình bằng lý luận, cung cấp mô tả ngôn ngữ và thể hiện gần gũi với nhu cầu kiểm duyệt nội dung chất lượng [7].

Trong khi đó, nghiên cứu cụ thể dành cho tiếng Việt vẫn khá hạn chế. Huỳnh và Trần đã thử nghiệm các mô hình transformer với trọng tâm là PhoBERT, đưa ra kết quả vượt trội trong phát hiện tin giả tiếng Việt [1]. Mặt khác, hệ thống "FND4Vn" to tác giả Võ và các cộng sự xây dựng cho lĩnh vực chính trị và đại dịch, đề xuất một mô hình chung và các bước thực nghiệm chi tiết để hỗ trợ phát hiện tin giả tiếng Việt một cách hiệu quả [8]. Tuy nhiên, cả hai chưa tích hợp yếu tố giải thích, khiến khả năng minh bạch và ứng dụng thực tế còn hạn chế.

Ở góc nhìn giải thích, một số nghiên cứu đã khai thác các đặc trưng nội dung (linguistic, affective, cognitive, social) để tăng khả năng giải thích trong phát hiện tin giả, cung cấp phân tích rõ ràng về những yếu tố đóng góp vào quyết định phân loại [9] [10]. Mặc dù vậy, nghiên cứu này vẫn chưa áp dụng mô hình transformer mạnh mẽ và chủ yếu dùng bộ phân loại truyền thống để làm rõ tính giải thích.

Như vậy, có thể thấy rằng nhiều tiến bộ đã được ghi nhận trong phát hiện tin giả thông qua việc ứng dụng các mô hình transformer, hướng tiếp cận đa phương thức cũng như các phương pháp giải thích xAI. Tuy nhiên, trong ngữ cảnh tiếng Việt, các nghiên cứu hiện có như PhoBERT hay FND4Vn mới chủ yếu tập trung vào việc cải thiện hiệu năng phân loại, mà chưa chú trọng đến khả năng giải thích của mô hình. Bên cạnh đó, những công trình quốc tế gần đây dù đã khai thác xAI kết hợp với transformer nhưng vẫn chưa có ứng dụng trực tiếp cho tiếng Việt. Chính khoảng trống này mở ra hướng đi cho nghiên cứu hiện tại: không chỉ dừng lại ở việc so sánh hiệu năng giữa BiLSTM và PhoBERT trong bối cảnh tiếng Việt, mà quan trọng hơn, đây là **một trong những công trình đầu tiên tích hợp xAI (LIME) vào phát hiện tin giả tiếng Việt**, nhằm cung cấp minh chứng trực quan về cách thức mô hình đưa ra quyết định, từ đó nâng cao tính minh bạch, độ tin cậy và khả năng ứng dụng thực tiễn của hệ thống..

Từ những phân tích trên, có thể thấy rõ nhu cầu cần thiết của một hướng tiếp cận vừa hiệu quả vừa minh bạch. Trên cơ sở đó, nghiên cứu này lựa chọn so sánh BiLSTM và PhoBERT, đồng thời áp dụng LIME để minh giải kết quả. Quy trình và phương pháp thực hiện sẽ được trình bày chi tiết trong **Chương 3 – Phương pháp nghiên cứu.**

Phương pháp

Tập dữ liệu

Trong nghiên cứu này, chúng tôi sử dụng tập dữ liệu ReINTEL 2020, một bộ dữ liệu chuyên biệt cho nhiệm vụ phát hiện tin giả tiếng Việt, bao gồm gần 10.000 mẫu được thu thập từ các nền tảng mạng xã hội và báo chí. Đặc điểm đáng chú ý là sự mất cân bằng giữa hai nhãn "Real" (0) và "Fake" (1), trong đó tin thật chiếm tỷ lệ cao hơn, gây ra thách thức trong việc huấn luyện mô hình mà không bị lệch nhãn. Dữ liệu được chia thành ba tập riêng biệt: train, validation và test theo tỷ lệ hợp lý, đảm bảo phân phối nhãn ổn định giữa các tập. *Bảng 1* minh họa số lượng mẫu ở từng nhãn:

Bảng 1: Tập dữ liệu ReINTEL 2020

Label	**Train**	**Validation**	**Test**
Real (0)	7,269	404	404
Fake (1)	1,472	82	82
Tổng cộng	8,741	486	486

Để chuẩn bị dữ liệu cho mô hình, chúng tôi áp dụng các bước tiền xử lý bao gồm: loại bỏ trùng lặp, chuẩn hóa văn bản và token hóa. Cả hai mô hình BiLSTM và PhoBERT sử dụng tokenizer của PhoBERT (vinai/PhoBERT-base) để thống nhất đầu vào. Các câu sau khi token hóa được ánh xạ thành các chuỗi ID và xử lý đệm để có cùng độ dài. Với BiLSTM, các embedding được học từ đầu, trong khi PhoBERT sử dụng trực tiếp vector đầu ra từ mô hình transformer.

Mô hình

Trong nghiên cứu này, chúng tôi triển khai và so sánh hai mô hình học sâu cho bài toán phát hiện tin giả tiếng Việt là BiLSTM và PhoBERT.

BiLSTM được xây dựng dựa trên mạng nơ-ron hồi tiếp nhiều lớp, bao gồm hai lớp LSTM hai chiều. Văn bản đầu vào được mã hóa bằng tokenizer của PhoBERT và ánh xạ sang không gian vector thông qua lớp embedding. Sau đó, chuỗi vector được xử lý bởi lớp LSTM theo cả hai hướng thời gian nhằm khai thác ngữ cảnh toàn cục. Đầu ra là phép nối (concatenate) các trạng thái ẩn cuối cùng của chiều tiến và chiều lùi, sau đó qua một tầng tuyến tính và hàm kích hoạt sigmoid để dự đoán xác suất. Để xử lý hiệu quả các chuỗi có độ dài khác nhau, mô hình sử dụng `pack_padded_sequence`, giúp bỏ qua phần padding không cần thiết. Kiến trúc BiLSTM được lựa chọn nhờ khả năng xử lý tốt dữ liệu tuần tự, đơn giản, dễ huấn luyện và phù hợp với các hệ thống có giới hạn về tài nguyên tính toán.

Với PhoBERT là một mô hình transformer tiền huấn luyện cho tiếng Việt, dựa trên kiến trúc RoBERTa. Trong mô hình PhoBERT, vector của token đặc biệt `[CLS]` từ tầng cuối được trích xuất làm biểu diễn toàn văn bản, qua tầng dropout nhằm giảm overfitting, rồi đi vào một lớp tuyến tính hai chiều tương ứng với hai nhãn. PhoBERT học các đặc trưng ngữ nghĩa sâu qua cơ chế attention đa đầu, cho hiệu quả vượt trội trong các tác vụ phân loại văn bản. Việc so sánh BiLSTM (mô hình tuần tự truyền thống) và PhoBERT (transformer hiện đại) giúp đánh giá hiệu năng trong việc phát hiện tin giả lẫn khả năng giải thích của từng mô hình thông qua phương pháp xAI.

Tích hợp phương pháp giải thích xAI – LIME

Các mô hình học sâu như BiLSTM và PhoBERT có thể đạt hiệu quả cao trong nhiệm vụ phát hiện tin giả, chúng thường bị xem là "hộp đen", người dùng khó biết lý do đằng sau dự đoán. Điều này hạn chế ứng dụng trong các lĩnh vực cần minh bạch như báo chí, giáo dục hay quản trị nội dung.

Để tăng khả năng diễn giải, chúng tôi tích hợp LIME ở giai đoạn cuối của pipeline trong *Hình 5*. LIME tạo các mẫu biến thể bằng cách bật/tắt ngẫu nhiên các đặc trưng (từ) trong câu, rồi dùng mô hình gốc dự đoán cho các mẫu này. Từ kết quả, LIME huấn luyện một mô hình tuyến tính đơn giản mô phỏng hành vi mô hình gốc trong vùng lân cận mẫu đang xét. Lý do lựa chọn LIME thay vì các phương pháp khác là vì độc lập mô hình (model-agnostic), hoạt động hiệu quả với dữ liệu dạng văn bản, và trực quan hóa dễ hiểu qua trọng số cho từng đặc

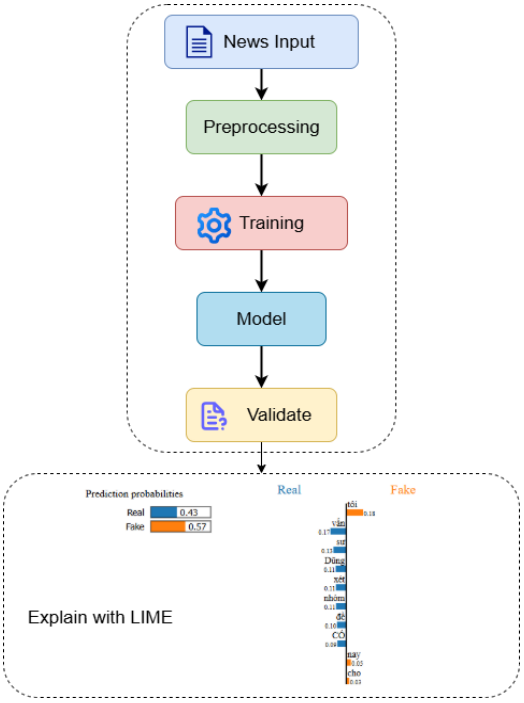
trưng. Nhờ LIME, người dùng và nhà quản lý có thể hiểu, kiểm chứng và phát hiện thiên lệch trong quyết định của mô hình.

Cụ thể, LIME tạo ra một tập hợp các mẫu biến thể bằng cách ngẫu nhiên bật hoặc tắt các đặc trưng đầu vào, trong trường hợp NLP là các từ trong câu. Sau đó, mô hình chính sẽ được sử dụng để đưa ra dự đoán cho các mẫu này. Dựa trên tập mẫu đã gán nhãn đầu ra, LIME huấn luyện một mô hình tuyến tính đơn giản để mô phỏng hành vi của mô hình gốc tại khu vực xung quanh điểm dữ liệu đang xét. Mô hình này không cần khái quát toàn bộ không gian, mà chỉ cần chính xác tại vùng lân cận, đảm bảo vừa đơn giản vừa dễ hiểu. Quá trình trên được mô hình hóa bằng công thức sau:

$$explanation(x) = \underset{g \in G}{arg\,min}\ [L(f, g, \pi_x) + \Omega(g)] \quad (1)$$

Trong đó, f là mô hình phức tạp cần giải thích, g là mô hình tuyến tính đơn giản trong tập các mô hình G, L là hàm mất mát đo sự khác biệt giữa đầu ra của mô hình chính và mô hình giải thích tại vùng lân cận π_x và $\Omega(g)$ là độ phức tạp của mô hình giải thích nhằm đảm bảo tính dễ hiểu cho người dùng.

Với khả năng cung cấp giải thích cụ thể cho từng câu đầu vào, LIME là lựa chọn phù hợp để áp dụng trong bài toán phát hiện tin giả, giúp người dùng, nhà quản lý và nhà nghiên cứu có thể theo dõi và kiểm chứng các quyết định phân loại của mô hình một cách minh bạch và trực quan.



Hình 5. Pipeline minh họa quy trình phát hiện tin giả, trong đó LIME được tích hợp ở giai đoạn cuối nhằm giải thích dự đoán của mô hình

Thực nghiệm

Cấu hình

Quá trình huấn luyện mô hình được thực hiện trên nền tảng Google Colab sử dụng GPU NVIDIA T4 với bộ nhớ cao (High-RAM – 15GB), đảm bảo tốc độ xử lý và khả năng lưu trữ phù hợp cho cả hai mô hình BiLSTM và PhoBERT. BiLSTM có kiến trúc nhẹ, cho phép batch size lớn 64, trong khi phức tạp hơn nên tạo batch size nhỏ hơn 32 để tránh vượt quá giới hạn bộ nhớ, đặc biệt khi xử lý các câu dài.

Trình huấn luyện sử dụng hàm mất mát binary cross-entropy và thuật toán tối ưu Adam. Số epoch huấn luyện được lựa chọn tùy theo mô hình: BiLSTM được huấn luyện trong 40 epoch để theo dõi overfitting, trong khi PhoBERT chỉ cần 5 epoch do có khả năng học nhanh từ giai đoạn tiền huấn luyện. Việc tích hợp LIME để giải thích kết quả được thực hiện trực tiếp trong môi trường Colab và áp dụng theo thời gian thực lên từng văn bản đầu vào mà không gặp trở ngại về hiệu năng.

Các chỉ số đánh giá

Để đánh giá hiệu quả của mô hình, nghiên cứu sử dụng bốn chỉ số phổ biến trong bài toán phân loại nhị phân, gồm:

- Accuracy: Tỷ lệ tổng số dự đoán đúng trên toàn bộ dữ liệu kiểm thử.
- Precision: Tỷ lệ các dự đoán dương tính đúng trên tổng số dự đoán dương tính.
- Recall: Tỷ lệ các mẫu dương tính thực sự được dự đoán đúng.
- F1-score: Trung bình điều hòa giữa precision và recall, đặc biệt hữu ích khi dữ liệu mất cân bằng.

Độ chính xác (accuracy) là tỷ lệ dự đoán đúng trên tổng số các trường hợp, được tính theo công thức:

$$accuracy = \frac{Số\ dự\ đoán\ đúng}{Tổng\ số\ trường\ hợp} \quad (2)$$

Recall (còn gọi là độ nhạy hoặc sensitivity) là tỷ lệ các trường hợp dương tính thực sự được mô hình dự đoán đúng, được tính như sau:

$$recall = \frac{True\ Positives\ (TP)}{True\ Positives(TP) + Fasle\ Negatives(FN)} \quad (3)$$

trong đó TP là số trường hợp dương tính được dự đoán đúng, và FN là số trường hợp dương tính bị dự đoán sai.

F1-score là trung bình điều hòa giữa độ chính xác (precision) và độ nhạy (recall), giúp đánh giá cân bằng giữa hai chỉ số này, đặc biệt hữu ích trong bài toán phân lớp mất cân bằng. Công thức tính F1-score như sau:

(4)

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall}$$

Trong đó, Precision là tỷ lệ các dự đoán dương tính đúng, được tính bằng:

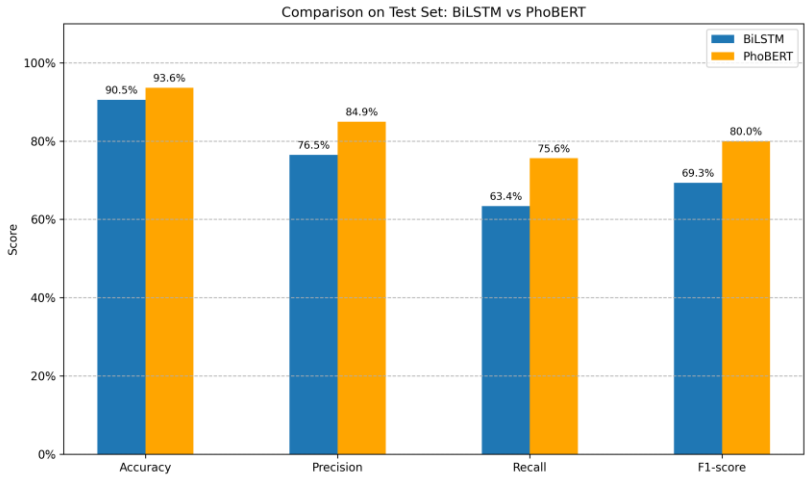
$$Precision = \frac{TP}{TP + FP} \quad (5)$$

Việc sử dụng kết hợp các chỉ số này giúp đánh giá mô hình một cách toàn diện từ nhiều góc độ: mức độ chính xác tổng thể, khả năng nhận diện tin giả, và độ ổn định trong phân loại nhãn thiểu số.

Kết quả

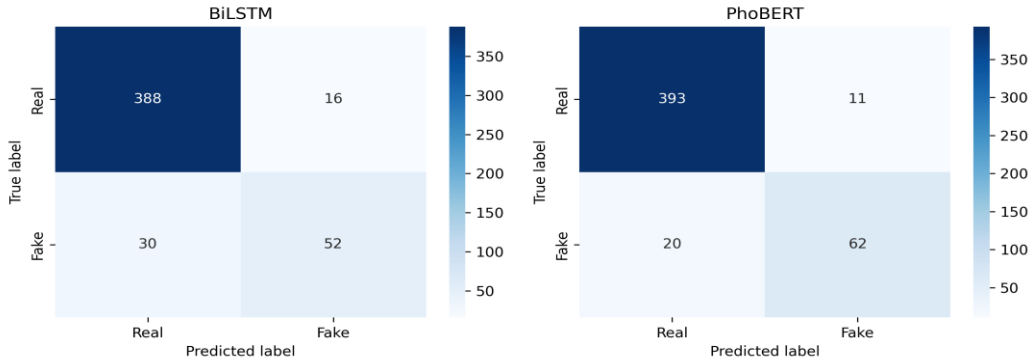
So sánh hiệu suất phân loại

Kết quả (Hình 1) cho thấy PhoBERT vượt trội hơn BiLSTM ở tất cả các chỉ số. Cụ thể,



Hình 1. Biểu đồ so sánh hiệu suất giữa BiLSTM và phoBERT

điểm accuracy của PhoBERT đạt 93.6%, cao hơn mức 90.5% của BiLSTM. Điểm recall của PhoBERT đạt 75.6%, cải thiện rõ rệt so với 63.4% của BiLSTM. F1-score cũng tăng từ 69.3%



a)BiLSTM **b) phoBERT**

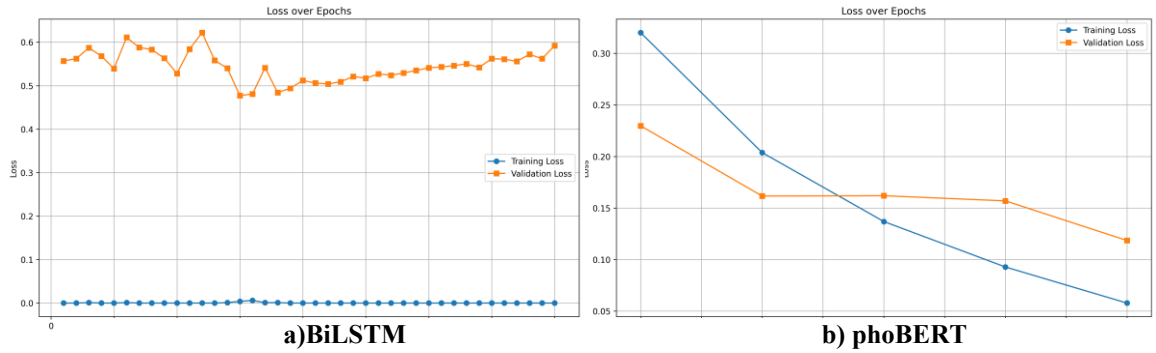
Hình 2. So sánh hiệu quả phân loại giữa BiLSTM (a) và PhoBERT (b) qua Confusion Matrix

lên 80.0%, cho thấy khả năng cân bằng tốt hơn giữa precision và recall. Những con số này phản ánh khả năng học ngữ nghĩa sâu và tổng quát hóa tốt của mô hình transformer so với mô hình tuần tự truyền thống.

Phân tích confusion matrix

Confusion matrix (*Hình 2*) chỉ ra rằng BiLSTM dự đoán đúng 388/404 mẫu tin thật nhưng chỉ đúng 52/82 mẫu tin giả, cho thấy xu hướng nghiêng về nhãn "Real", một dấu hiệu rõ rệt của việc chịu ảnh hưởng bởi sự mất cân bằng dữ liệu. Trong khi đó, PhoBERT thể hiện sự cân bằng hơn khi nhận diện đúng 62/82 mẫu tin giả và giảm đáng kể số mẫu bị nhầm lẫn.

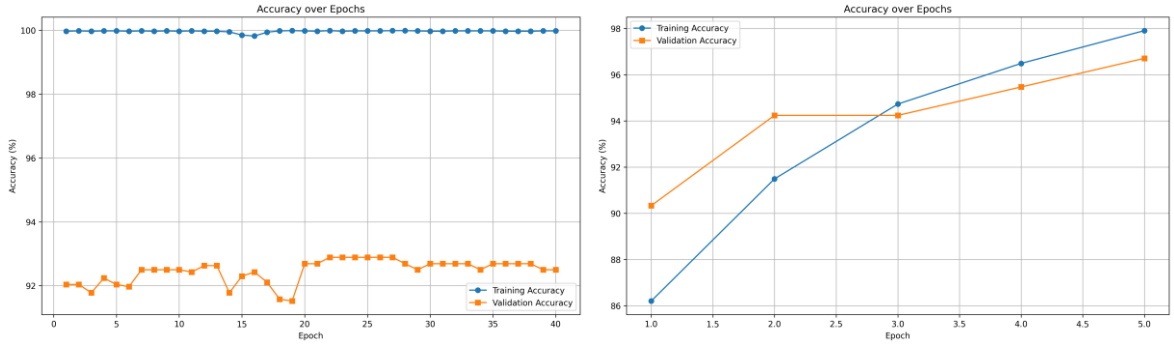
Biến thiên chỉ số trong quá trình huấn luyện



a)BiLSTM **b) phoBERT**

Hình 3. Biểu đồ biến thiên training loss và validation loss theo số epoch của mô hình BiLSTM (a) và phoBERT (b)

Biểu đồ loss và accuracy theo epoch (*Hình 3 và 4*) cho thấy BiLSTM nhanh chóng đạt loss gần 0 trên tập huấn luyện, trong khi validation loss dao động mạnh, xác nhận hiện tượng overfitting. Ngược lại, PhoBERT cho thấy quá trình hội tụ ổn định với khoảng cách nhỏ giữa



a)BiLSTM **b) phoBERT**

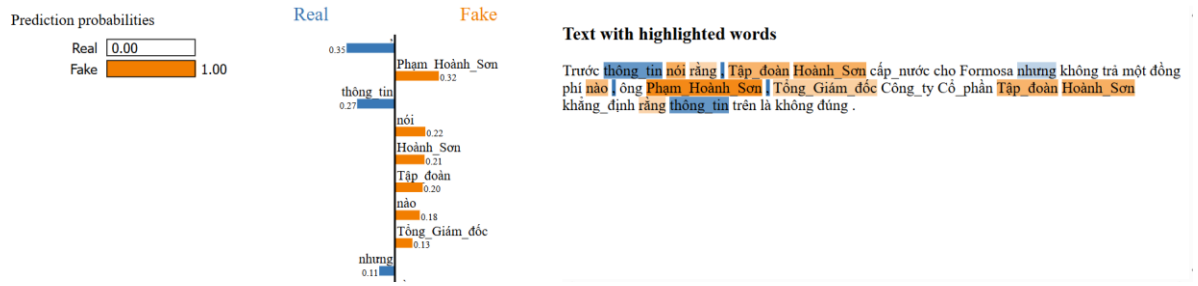
Hình 4: Biểu đồ Accuracy theo từng epoch của mô hình BiLSTM (a) và phoBERT (b)

training và validation loss. Tương tự, validation accuracy của PhoBERT đạt hơn 96% sau 5 epoch, gần tiệm cận với training accuracy (~98%), trong khi BiLSTM giữ khoảng cách ổn định ~9% giữa hai đường, phản ánh sự thiếu ổn định trong khái quát hóa.

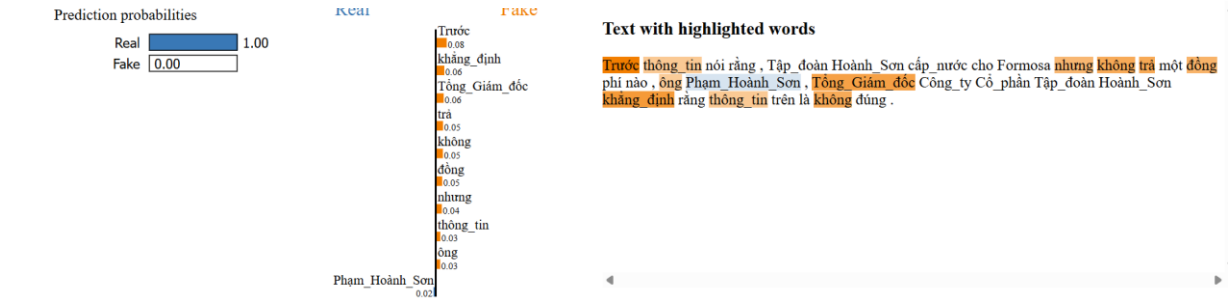
Tổng hợp các nhận định trên cho thấy, BiLSTM có ưu điểm về tốc độ huấn luyện và đơn giản khi triển khai, nhưng dễ overfit và kém hiệu quả trên dữ liệu mất cân bằng. PhoBERT có khả năng học ngữ nghĩa và ngữ cảnh vượt trội, dẫn đến hiệu năng phân loại cao và ổn định hơn. Việc kết hợp với LIME sẽ được trình bày trong phần tiếp theo để minh họa khả năng giải thích của từng mô hình khi áp dụng vào các trường hợp cụ thể.

Phân tích khả năng giải thích với LIME

Để làm rõ cơ chế dự đoán, LIME được áp dụng cho cả BiLSTM và PhoBERT nhằm phân tích đóng góp cục bộ của từng từ trong văn bản đối với dự đoán phân loại. Việc trực quan hóa các yếu tố ảnh hưởng đến kết quả không chỉ hỗ trợ đánh giá chất lượng mô hình mà còn giúp phát hiện các thiên lệch tiềm ẩn, tăng cường tính minh bạch và khả năng kiểm chứng của hệ thống.



a) phoBERT



b) BiLSTM

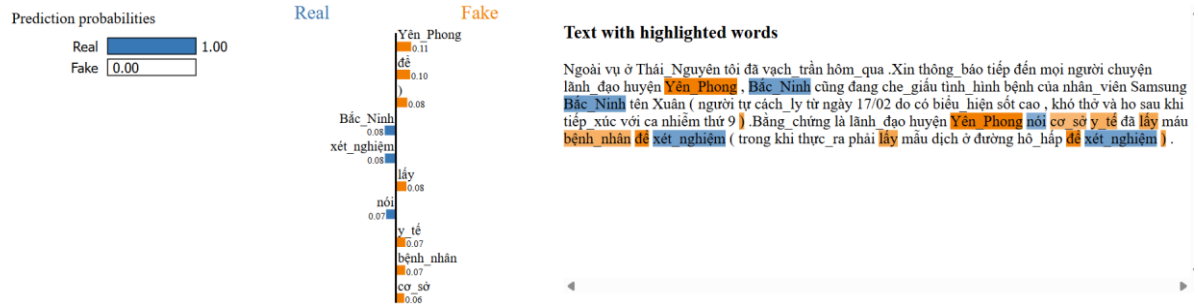
Hình 5.Giải thích mẫu 1 từ tập dữ liệu bằng LIME. Văn bản: "Trước thông_tin nói rằng , Tập_đoàn Hoành_Sơn cấp_nước cho Formosa nhưng không trả một đồng phí nào , ông Phạm_Hoành_Sơn , Tổng_Giám_đốc Công_ty Cổ_phần Tập_đoàn Hoành_Sơn khẳng_định rằng thông_tin trên là không đúng ." Nhãn đúng: Real (Tin thật)

Ba văn bản từ tập kiểm thử đã được lựa chọn để phân tích chi tiết, dựa trên các tiêu chí:

- (1) Hai mô hình cho kết quả khác nhau;
- (2) Câu có cấu trúc ngôn ngữ phong phú, thể hiện rõ sắc thái ngữ nghĩa;
- (3) Gồm cả trường hợp mô hình dự đoán sai hoặc có độ tin cậy thấp.

Ở mẫu 1, văn bản là một tin thật liên quan đến việc bác bỏ thông tin sai về Tập đoàn Hoành Sơn, với nhãn gốc là 0, trong *Hình 5*. Trong khi BiLSTM dự đoán sai với xác suất gần như tuyệt đối 0.9998, PhoBERT phân loại đúng với độ tin cậy cao 0.9984. Kết quả LIME cho thấy BiLSTM bị chi phối bởi các cụm từ như "Tập đoàn", "Phạm_Hoành_Sơn" hay "không", vốn mang tính trung lập hoặc từng xuất hiện trong các mẫu tin giả, khiến mô hình đánh giá sai bất chấp ngữ cảnh phủ định. Ngược lại, PhoBERT chú ý đúng vào các từ khóa phản bác như "không đúng", "khẳng định", "nhưng", cho thấy khả năng hiểu ngữ cảnh toàn cục và phát hiện cấu trúc bác bỏ thông tin một cách chính xác.

Với mẫu 2 là một tin giả có nội dung tố cáo lãnh đạo huyện Yên Phong, Bắc Ninh che giấu dịch bệnh tại Samsung Hình 6 BiLSTM dự đoán sai là tin thật với xác suất cao 0.997, trong khi PhoBERT tiếp tục phân loại đúng với độ tin cậy tương ứng 0.996. Phân tích LIME



a) BiLSTM

b) phoBERT

Hình 6. Giải thích mẫu 2 từ tập dữ liệu bằng LIME. Văn bản: "Ngoài vụ ở Thái_Nguyên tôi đã vạch_trần hôm_qua .Xin thông_báo tiếp đến mọi người chuyện lãnh_đạo huyện Yên_Phong , Bắc_Ninh cũng đang che_giấu tình_hình bệnh của nhân_viên Samsung Bắc_Ninh tên Xuân (người tự cách_ly từ ngày 17/02 do có biểu_hiện sốt cao , khó thở và ho sau khi tiếp_xúc với ca nhiễm thứ 9) .Bằng_chứng là lãnh_đạo huyện Yên_Phong nói cơ_sở y_tế đã lấy máu bệnh_nhân để xét_nghiệm (trong khi thực_ra phải lấy mẫu dịch ở đường hô_hấp để xét_nghiệm) .". Nhãn đúng: Fake (Tin giả)

cho thấy BiLSTM thiên về các cụm từ như "xét nghiệm", "cơ sở y tế", "Bắc Ninh", phản ánh xu hướng đánh giá theo từ khóa mang tính chính thống, trong khi bỏ qua sắc thái diễn giải mang tính cáo buộc và suy diễn. Trái lại, PhoBERT đã nhận diện được các cụm từ đặc trưng

cho tin giả như "che giấu", "vạch trần", "thực ra", thể hiện khả năng khai thác các yếu tố ngôn ngữ cảm xúc và nghi ngờ thường thấy trong văn bản lan truyền thông tin sai lệch.

Ví dụ cuối cùng mẫu 3 ở *Hình 7* là một đoạn văn có độ dài lớn, với nhiều câu khẳng định rằng "không có tang vật" trong vụ án Hồ Duy Hải. Mặc dù cả BiLSTM và PhoBERT đều phân loại đúng đây là tin giả, độ tin cậy của cả hai đều thấp (0.61 và 0.57). LIME cho thấy BiLSTM chú trọng vào các từ như "chiếc", "(3!)", "vật", trong khi PhoBERT tập trung vào các từ như "tôi", "nay", "vấn đề", mà không nhấn mạnh được các cụm then chốt như "tất cả đều ảo", "không ai xác nhận".

a) BiLSTM

b) phoBERT

Hình 7: Giải thích mẫu 3 từ tập dữ liệu bằng LIME. Văn bản: "BÁO CÁO GIÁM SÁT (3!) MỌI THỨ TANG VẬT TÀI SẢN ĐỀU KHÔNG CÓ, CHẲNG CÓ TANG VẬT NỮ TRANG NÀO CẢ... TẤT CẢ ĐỀU ẢO.". Nhãn đúng: Fake (Tin giả). Cả BiLSTM và phoBERT đều phân loại đúng văn bản này là tin giả, nhưng với xác suất không cao.

Điều này cho thấy khi văn bản có cấu trúc lặp lại, cảm xúc mạnh và pha trộn giữa ngôn ngữ pháp lý và suy đoán, cả hai mô hình đều gặp khó khăn trong việc xác định trọng tâm ngữ nghĩa, dẫn đến phân loại thiếu chắc chắn.

Tổng hợp từ ba ví dụ mẫu trên, có thể nhận thấy rằng BiLSTM thường dựa vào từ khóa đơn lẻ và dễ bị chi phối bởi các cụm từ trung tính hoặc quen thuộc trong tập huấn luyện, dẫn đến các lỗi đánh giá sai khi văn bản chứa cấu trúc phủ định, phản bác hoặc ngôn ngữ cảm xúc. Trong khi đó, PhoBERT cho thấy ưu thế rõ rệt nhờ khả năng biểu diễn ngữ nghĩa sâu và học được các tín hiệu ngữ cảnh phức tạp trong tiếng Việt từ giai đoạn tiền huấn luyện. Tuy vậy, trong các trường hợp văn bản dài, lặp lại và không rõ ràng về sắc thái, cả hai mô hình đều có xu hướng suy giảm độ tin cậy. Việc tích hợp LIME đã giúp làm sáng tỏ cách mô hình đưa ra quyết định, đồng thời chỉ ra những điểm mạnh, điểm yếu trong quá trình học của từng mô hình.

Đây là cơ sở quan trọng để hiệu chỉnh mô hình, tăng độ tin cậy và minh bạch hóa hệ thống trong các ứng dụng thực tế phát hiện tin giả.

Kết luận và Hướng phát triển

Trong nghiên cứu này, chúng tôi đã tiến hành so sánh hai mô hình học sâu BiLSTM và PhoBERT cho nhiệm vụ phát hiện tin giả trong tiếng Việt, đồng thời tích hợp phương pháp giải thích LIME để tăng cường tính minh bạch. Kết quả trên bộ dữ liệu ReINTEL 2020 cho thấy PhoBERT đạt hiệu năng vượt trội với độ chính xác và F1-score cao hơn đáng kể, nhờ khả năng học ngữ nghĩa sâu và tổng quát hóa từ giai đoạn tiền huấn luyện.

BiLSTM vẫn là lựa chọn thực tế trong bối cảnh hạn chế về tài nguyên, nhờ kiến trúc đơn giản, tốc độ huấn luyện nhanh và dễ tích hợp xAI. Phân tích với LIME cũng cho thấy BiLSTM cho kết quả giải thích trực quan và dễ kiểm chứng hơn, trong khi PhoBERT thể hiện độ nhạy cao hơn với các yếu tố ngữ nghĩa tinh vi và sắc thái cảm xúc trong văn bản. LIME đã giúp làm sáng tỏ cơ chế hoạt động của mô hình, từ đó hỗ trợ người dùng trong việc kiểm chứng và điều chỉnh hệ thống. Đây là bước quan trọng hướng tới việc xây dựng các mô hình AI đáng tin cậy, đặc biệt trong những lĩnh vực nhạy cảm như phát hiện tin giả. Trong tương lai, nghiên cứu sẽ được mở rộng theo các hướng sau: i) Nghiên cứu và thử nghiệm các mô hình transformer hiện đại hơn như DeBERTa, GPT hoặc XLM-R để cải thiện hiệu năng; ii) Kết hợp thông tin đa phương thức (văn bản, hình ảnh, metadata) để tăng độ chính xác trong môi trường mạng xã hội phức tạp; iii) Phát triển giao diện xAI trực quan cho người dùng cuối và người kiểm duyệt, hỗ trợ kiểm tra và lý giải dự đoán thời gian thực.

Tài liệu tham khảo

[1] A.-T. Huynh and P. Tran, "Utilizing Transformer Models To Detect Vietnamese Fake News on Social Media Platforms.," *KSII Trans. Internet Inf. Syst.*, vol. 19, no. 2, 2025.

[2] M. Schlee, G. Kant, C. Ehrling, B. Säfken, and T. Kneib, "Decoding synthetic news: an interpretable multimodal framework for the classification of news articles in a novel news corpus," *Artif. Intell. Rev.*, vol. 58, no. 10, p. 302, Jul. 2025,

[3] V. U. Gongane, M. V. Munot, and A. D. Anuse, "A survey of explainable AI techniques for detection of fake news and hate speech on social media platforms," *J. Comput. Soc. Sci.*, vol. 7, no. 1, pp. 587–623, Apr. 2024,

[4] D. Mouratidis, A. Kanavos, and K. Kermanidis, "From Misinformation to Insight: Machine Learning Strategies for Fake News Detection," *Information*, vol. 16, no. 3, 2025,

[5] S. Hameed and M. Wasim, "Interpretable BERT and RoBERTa based Fake News Detection using LIME," in *2025 International Conference on Innovation in Artificial Intelligence and Internet of Things (AIIT)*, May 2025, pp. 1–7.

[6] H. Rama Moorthy *et al.*, "Dual stream graph augmented transformer model integrating BERT and GNNs for context aware fake news detection," *Sci. Rep.*, vol. 15, no. 1, p. 25436, Jul. 2025,

[7] H. R. LekshmiAmmal and A. K. Madasamy, "A reasoning based explainable multimodal fake news detection for low resource language using large language models and transformers," *J. Big Data*, vol. 12, no. 1, p. 46, Feb. 2025,

[8] T. H. Vo, I. Felde, and K. C. Ninh, "Fake News Detection System, based on CBOW and BERT," *Acta Polytech. Hung.*, vol. 22, no. 1, 2025.

[9] G. Gondwe, "Can AI outsmart fake news? Detecting misinformation with AI models in real-time," *Emerg. Media*, p. 27523543251325902, 2025.

[10] S. Muñoz and C. Á. Iglesias, "Exploiting Content Characteristics for Explainable Detection of Fake News," *Big Data Cogn. Comput.*, vol. 8, no. 10, 2024,

TRÍ TUỆ NHÂN TẠO VÀ ĐỔI MỚI ĐÀO TẠO NHÂN LỰC TIẾNG TRUNG CHUYÊN NGÀNH THƯƠNG MẠI Ở VIỆT NAM

Nguyễn Duy An[1]

Tóm tắt: Nghiên cứu phân tích thực trạng đào tạo tiếng Trung thương mại tại Việt Nam trong bối cảnh hợp tác kinh tế Việt - Trung và chuyển đổi số. Các hạn chế chính gồm: giáo trình chậm bản địa hóa, ứng dụng công nghệ yếu, giảng viên thiếu thực tiễn và kỹ năng số, liên kết trường - doanh nghiệp chưa hiệu quả. Trên cơ sở đó, nghiên cứu đề xuất chiến lược đổi mới dựa trên AI: tái cấu trúc khóa học thông minh, nền tảng thực tập ảo VR với dịch thuật thời gian thực, chương trình bồi dưỡng giảng viên đa chiều và mô hình hợp tác ba bên Việt - Trung có AI làm cầu nối. Kết quả cung cấp định hướng cải cách và gợi ý quốc tế hóa đào tạo nhân lực ngoại ngữ.

Từ khóa: tiếng Trung chuyên ngành; kỹ năng nghề nghiệp; đào tạo nhân lực; chuyển đổi số; trí tuệ nhân tạo

1. Đặt vấn đề

Trí tuệ nhân tạo (AI) với được xem là động lực chính yếu của làn sóng cách mạng công nghiệp mới đang định hình sâu sắc cục diện cạnh tranh toàn cầu; trong đó, năng lực bồi dưỡng nhân tài đã trở thành trụ cột quan trọng của năng lực cạnh tranh cốt lõi. Hiện nay, hơn 40% quốc gia và khu vực trên thế giới đã đưa việc bồi dưỡng nhân tài AI vào quy hoạch chiến lược; 451 trường đại học ở 45 quốc gia đã mở các khóa học chuyên ngành AI, trong đó tỷ lệ các trường đại học ở Mỹ, Đức, Anh lần lượt đạt 31,9%, 10,4% và 8,0%, thể hiện rõ vị thế chiến lược quốc tế của việc bồi dưỡng nhân tài AI [1].

Trong bối cảnh kép của làn sóng "Trí tuệ nhân tạo +" và hợp tác kinh tế thương mại Việt - Trung ngày càng mạnh mẽ, nhu cầu bồi dưỡng nhân tài tiếng Trung thương mại ngày càng tăng. Sự mở rộng trao đổi lao động xuyên biên giới và sự tăng trưởng thương mại song phương đã thúc đẩy nhu cầu đa dạng mà thị trường việc làm Việt Nam cung không đủ cầu về nhân lực tổng hợp "Tiếng Trung + Kỹ năng nghề nghiệp"; thương mại điện tử xuyên biên giới Việt - Trung tăng trưởng mạnh cũng đặt ra yêu cầu cao hơn về năng lực ngôn ngữ, kỹ năng nghề nghiệp và thấu hiểu văn hóa của nhân lực [2]. Sự gia tăng nhu cầu này bắt nguồn từ động lực kép của quá trình hợp tác kinh tế Việt - Trung và chuyển đổi số. Năm 2023, kim ngạch xuất nhập khẩu thương mại điện tử xuyên biên giới của Trung Quốc đạt 2,38 nghìn tỷ NDT (tăng 15,6%, chiếm 5,7% tổng giá trị ngoại thương); quý I năm 2024 tiếp tục tăng 9,6%, và dự kiến đến năm 2025 giá trị giao dịch sẽ đạt 2,5 nghìn tỷ NDT [3], nhưng tình trạng thiếu hụt nhân tài

[1] TS, Trường Đại học Ngoại ngữ - Tin học TP. HCM; Email: annd@huflit.edu.vn

"Nhóm ngôn ngữ nhỏ + Thương mại + Công nghệ số" rất nghiêm trọng [4], trong khi đó, mô hình giảng dạy truyền thống bị hạn chế bởi đội ngũ giảng viên, giáo trình, công nghệ... vẫn chưa đáp ứng đủ về số lượng và chất lượng nguồn nhân lực.

Hiện nay, trí tuệ nhân tạo mang đến con đường hoàn toàn mới để giải quyết tình trạng trên. Hệ thống học ngôn ngữ thông minh có thể thực hiện việc đề xuất phương án cá nhân hóa; phân tích dữ liệu có thể tối ưu hóa chương trình học và nội dung giảng dạy; công nghệ mô phỏng ảo có thể tái hiện bối cảnh thương mại thực tế nhằm nâng cao năng lực thực hành; nền tảng AI giao thoa văn hóa giúp mở rộng sự hiểu biết tập quán thương mại Trung Quốc nhằm giảm xung đột. Thực tiễn tại Đại học Columbia (Mỹ) cho thấy, các công cụ như ChatGPT có thể hỗ trợ hiệu quả việc tích hợp tài nguyên giảng dạy, biên soạn lại giáo trình và học từ mới; năng lực ứng dụng AI đã trở thành tiêu chuẩn quan trọng trong đánh giá giảng viên [5], điều này cung cấp tham chiếu cho việc cải cách giáo dục tiếng Trung thương mại tại Việt Nam. Vì vậy, ứng dụng AI để tăng cường bồi dưỡng nhân lực Tiếng Trung thương mại có ý nghĩa to lớn đối với việc đáp ứng nhu cầu kinh tế thương mại Việt - Trung. Nghiên cứu này nhằm khám phá các chiến lược bồi dưỡng hiệu quả, cung cấp thêm cứ liệu tham khảo cho sự phát triển giáo dục tiếng Trung thương mại tại Việt Nam, góp phần đào tạo nhân tài tổng hợp phù hợp với sự thay đổi của thị trường.

2. Thực trạng nhu cầu nhân tài tiếng Trung thương mại tại Việt Nam

2.1 Động lực kép từ sự gia tăng hợp tác kinh tế thương mại Việt - Trung và chuyển đổi số

Sự gia tăng mạnh mẽ nhu cầu nhân lực tiếng Trung thương mại tại Việt Nam bắt nguồn từ mối liên kết chặt chẽ trong quan hệ kinh tế - thương mại Việt - Trung và quy mô giao thương ngày càng mở rộng. Điển hình là kết quả hợp tác giữa các địa phương Việt Nam và Quảng Tây trong thời gian qua cho thấy kim ngạch thương mại song phương đã đạt 36,3 tỷ USD vào năm 2023 và tiếp tục tăng trưởng 32% chỉ trong 7 tháng đầu năm 2024. Đáng chú ý, thương mại qua các cửa khẩu biên giới đất liền Việt Nam - Quảng Tây chiếm hơn 95% tổng kim ngạch xuất nhập khẩu qua biên giới đất liền Việt - Trung. Việt Nam cũng đã duy trì vị thế là đối tác thương mại lớn nhất của Quảng Tây suốt 25 năm liên tiếp [6]. Trong bối cảnh này, mô hình thương mại truyền thống đang dần chuyển đổi sang các hình thái mới, đặc biệt là thương mại điện tử, bao phủ toàn bộ chuỗi từ giao dịch trực tuyến, vận hành nền tảng đến logistics xuyên biên giới. Sự thay đổi này kéo theo nhu cầu cấp bách về đội ngũ nhân lực "Tiếng Trung + Thương mại" có khả năng thực hiện các công việc cốt lõi như giao tiếp thương mại, đàm phán, ký kết hợp đồng và chăm sóc khách hàng [7]. Chẳng hạn, các hoạt động quảng bá sản phẩm và hỗ trợ khách hàng trên nền tảng thương mại điện tử phụ thuộc nhiều vào kiến thức ngôn ngữ kết hợp hiểu biết thương mại của nhóm nhân tài này để bảo đảm giao dịch thuận lợi.

Ngoài ra, việc mở rộng hoạt động sâu hơn của các doanh nghiệp Trung Quốc tại Việt Nam càng làm rõ khoảng cách thiếu hụt nhân lực trong lĩnh vực này. Thương hiệu TCL hiện chiếm 14,4% thị phần TV màu tại Việt Nam; nền tảng 1688 đã xây dựng trung tâm chọn sản phẩm tại Việt Nam nhằm phục vụ khách hàng B2B; số lượng xe thương mại của Dongfeng tại Việt Nam vượt 36.000 chiếc. Các hoạt động kinh doanh xuyên quốc gia này đều đòi hỏi gấp

rút nguồn nhân lực thông thạo tiếng Trung thương mại [8]. Dữ liệu tuyển dụng cũng phản ánh xu hướng này: nhiều vị trí quản lý kinh doanh yêu cầu rõ thuần thục tiếng Việt - Trung và kỹ năng máy tính, với mức lương dao động 7.000 - 15.000 nhân dân tệ kèm phụ cấp, thể hiện sự cạnh tranh gay gắt trong việc thu hút nhân tài "giỏi ngôn ngữ, am hiểu thương mại, thành thạo số hóa".

Bên cạnh đó, sự phát triển nhanh chóng của kinh tế số tiếp tục tạo ra những nhu cầu mới. Việt Nam hiện dẫn đầu Đông Nam Á về tốc độ tăng trưởng kinh tế số với mức trung bình trên 20% mỗi năm. Doanh thu bán lẻ thương mại điện tử và thanh toán phi tiền mặt năm 2024 gia tăng mạnh, cho thấy tiến trình chuyển đổi số đang tăng tốc [9]. Sự chuyển dịch này khiến mô hình vận hành của doanh nghiệp ngày càng nghiêng về trực tuyến, làm bùng nổ nhu cầu nhân tài "Ngôn ngữ + Công nghệ". Lấy ví dụ trong lĩnh vực tiếp thị số, nhân sự cần vừa có năng lực giao tiếp tiếng Trung để kết nối với thị trường và đối tác Trung Quốc, vừa phải thành thạo ứng dụng công nghệ số để quảng bá sản phẩm qua các nền tảng trực tuyến, từ đó giúp doanh nghiệp mở rộng thị trường [10]. Trong bối cảnh chuyển đổi số, ngành nghề đặt ra yêu cầu mới về kỹ năng, Giờ đây, nhân lực tiếng Trung thương mại cần bổ sung năng lực số hóa, tin học dựa trên nền tảng kiến thức ngôn ngữ và thương mại, nhằm đáp ứng nhu cầu phát triển của doanh nghiệp trong môi trường cạnh tranh khốc liệt.

2.2 Sự "lệch pha" giữa mô hình đào tạo truyền thống và nhu cầu mới của công việc

Mặc dù phong trào học tiếng Trung tại Việt Nam đang diễn ra sôi động, nhưng nguồn cung nhân lực vẫn bộc lộ ba hạn chế nổi bật: (i), thiếu hệ thống giáo trình bản địa hóa. Phần lớn giáo trình hiện nay được dịch trực tiếp từ bản tiếng Trung, chưa phản ánh đầy đủ bối cảnh thực tế của hoạt động thương mại điện tử xuyên biên giới mà các doanh nghiệp vừa và nhỏ của Việt Nam đang tham gia. Ví dụ, các chuẩn mực biểu đạt tiếng Trung trong hoạt động livestream bán hàng, đặc biệt với nông sản, hầu như chưa được đề cập [11]; (ii), đào tạo kỹ năng còn tách rời với yêu cầu kinh tế số. Các khóa học hiện tại vẫn tập trung chủ yếu vào kỹ năng truyền thống như viết thư tín thương mại mà chưa tích hợp đầy đủ các năng lực số cần thiết. Chẳng hạn, người học thiếu cơ hội rèn luyện kỹ năng thao tác trên hệ thống AI, hoặc sử dụng thành thạo giao diện tiếng Trung của các nền tảng thanh toán xuyên biên giới [12]; (iii), bồi dưỡng năng lực giao thoa văn hóa còn chậm. Phản hồi từ doanh nghiệp Trung Quốc cho thấy nhiều nhân viên Việt Nam gặp khó khăn trong hợp tác do khác biệt về lễ nghi và văn hóa kinh doanh, chẳng hạn, cách diễn đạt trong đàm phán giá cả. Trong khi đó, hầu hết các chương trình đào tạo hiện tại vẫn thiếu phương pháp giảng dạy kiểu nhập vai, giúp người học tiếp cận thực tiễn hợp tác đa văn hóa [13].

Về chính sách, tuy đã xuất hiện nhiều tín hiệu tích cực, như việc Việt Nam đưa tiếng Trung vào giảng dạy từ bậc tiểu học và trong bối cảnh hình thành "Một vành đai, Một con đường", tiếng Trung trở thành chuyên ngành được ưa chuộng tại nhiều trường đại học. Tuy nhiên, sự kết nối giữa chương trình đào tạo đại học và nhu cầu thực tế của doanh nghiệp vẫn còn hạn chế. Điển hình như trong hợp tác đào tạo nhân lực, Việt Nam chưa xây dựng được cơ chế bồi dưỡng nhân lực theo mô hình "Trường học - Doanh nghiệp" với quy mô lớn. Hiện chỉ

tồn tại một số trường hợp hợp tác nhỏ lẻ, chẳng hạn chương trình đào tạo "AI + công nghệ xe thương mại" giữa Trường Cao đẳng nghề Thành phố Liễu Châu và Công ty Hải Âu [14].

3. Những điểm nghẽn trong đào tạo nhân lực tiếng Trung thương mại tại Việt Nam

3.1 Sự đứt gãy kép trong hệ thống giáo trình

Giáo trình tiếng Trung thương mại tại Việt Nam đang bộc lộ rõ tình trạng "lệch pha cung - cầu", tập trung ở hai điểm nghẽn chính: thiếu giáo trình có nội dung bản địa hóa và chậm thích ứng với yêu cầu số hóa. Một mặt, số lượng giáo trình bản địa còn rất hạn chế; phần lớn giáo trình hiện có được dịch trực tiếp từ Trung Quốc [15], chưa phản ánh bối cảnh kinh doanh thực tế của các doanh nghiệp vừa và nhỏ tại Việt Nam. Điều này dẫn tới sự chênh lệch nghiêm trọng giữa ví dụ trong sách và nhu cầu thực tiễn. Mặt khác, quá trình chuyển đổi số của giáo trình chậm hơn yêu cầu của giáo dục 4.0. Dù tài nguyên giảng dạy số tiếng Trung tại Việt Nam đã hình thành ba loại chính gồm nền tảng trực tuyến, ứng dụng di động (APP) và mạng xã hội, nhưng phổ biến vẫn bị phân mảnh chức năng. Ứng dụng học tập thường chỉ dừng ở chức năng tra cứu từ vựng, thiếu các mô hình tương tác như mô phỏng hội thoại thông minh, thực hành bối cảnh kinh doanh; ngân hàng đề thông minh phục vụ thi HSK còn thiếu, và nội dung chưa gắn kết với thực tế và kỹ năng thương mại điện tử. Điểm yếu này khiến giáo trình chưa thể hoàn thiện năng lực "ngôn ngữ + số hóa" cho sinh viên, ví dụ như sử dụng công cụ dịch AI trong đàm phán thương mại, trong khi đây là năng lực mà doanh nghiệp đang đặc biệt cần.

3.2 Chính sách ứng dụng công nghệ thiếu tích hợp công cụ thông minh và mô hình giảng dạy cố định

Chuyển đổi số trong giảng dạy tiếng Trung thương mại ở Việt Nam đang đối mặt với tình trạng "trên nóng, dưới lạnh". Mặc dù chính phủ Việt Nam đã ban hành Chỉ thị số 942/QĐ-TTg nhằm thúc đẩy chuyển đổi số trong giáo dục và đặt mục tiêu "xây dựng đại học số" [16], song ở bậc đại học, đặc biệt trong giảng dạy tiếng Trung, vẫn tồn tại khoảng cách lớn giữa chính sách và thực tế triển khai. Các chính sách liên quan chưa có quy chuẩn hệ thống về xây dựng tài nguyên số, chưa xác định rõ lộ trình phát triển các khóa học trực tuyến, phòng thí nghiệm ảo bối cảnh thương mại; đồng thời thiếu tiêu chuẩn biên soạn, cơ chế cập nhật và yêu cầu ứng dụng giáo trình số [15]. Vì vậy, nhà trường và giảng viên gặp khó khăn khi triển khai giảng dạy số do "không có chuẩn để làm theo". Sự thiếu hụt chuẩn mực này càng hạn chế việc áp dụng công nghệ thông minh [17]. Các công nghệ như trí tuệ nhân tạo, phân tích dữ liệu lớn - vốn có thể nâng cao hiệu quả giảng dạy thông qua thiết kế học tập cá nhân hóa, mô phỏng bối cảnh thực chưa được khai thác hiệu quả do thiếu chính sách hỗ trợ và tài nguyên phù hợp; chỉ một số ít trường thử nghiệm, giá trị tiềm năng còn chưa phát huy.

Mâu thuẫn sâu xa hơn là sự "lệch pha" giữa ứng dụng công nghệ và mục tiêu giảng dạy. Việt Nam tuy đã thúc đẩy chuyển đổi số giáo dục nghề nghiệp trên nhiều phương diện [18], nhưng nền tảng số hiện có chưa đáp ứng yêu cầu huấn luyện tích hợp "kinh doanh + công nghệ". Trong quá trình thực hiện "Chương trình Chuyển đổi số quốc gia", do sự chênh lệch về mức độ chuyển đổi số giữa các trường đại học, nhiều trường khi lựa chọn tài nguyên số chưa

kết hợp được yêu cầu cụ thể về công nghệ thông minh trong bối cảnh kinh doanh. Kết quả là tài nguyên số không hỗ trợ được việc đào tạo tích hợp, làm trầm trọng thêm sự đứt gãy giữa ứng dụng công nghệ và mục tiêu giảng dạy, cuối cùng ảnh hưởng đến việc hình thành năng lực cạnh tranh cốt lõi cho sinh viên [19].

3.3 Đội ngũ giảng viên thiếu kinh nghiệm thực tiễn thương mại và năng lực số hóa

Đội ngũ giảng viên tiếng Trung tại Việt Nam đang gặp "khoảng trống kép về năng lực": thiếu kinh nghiệm thực tiễn thương mại và thiếu kỹ năng giảng dạy số, vấn đề này trực tiếp ảnh hưởng tới chất lượng đào tạo. Về kinh nghiệm thực tế, mặc dù hợp tác kinh tế và thương mại Việt - Trung ngày càng gia tăng, kéo theo nhu cầu nhân lực tiếng Trung thương mại, nhưng nhiều giảng viên dù có trình độ học thuật tốt lại thiếu trải nghiệm thực tế trong môi trường kinh doanh [17]. Do đó, trong các nội dung quan trọng như kỹ năng đàm phán thương mại, lập kế hoạch marketing, giảng viên khó có thể cung cấp hướng dẫn sát thực tế, khiến kiến thức sinh viên tiếp thu và yêu cầu công việc thực tế bị "lệch pha", thậm chí dẫn tới nhầm lẫn hoặc sai sót.

Về năng lực số, khi dạy học trực tuyến và các hình thức số hóa ngày càng phổ biến, kỹ năng số của giảng viên trở thành yếu tố then chốt. Tuy nhiên, nhiều giảng viên tại Việt Nam còn hạn chế trong việc ứng dụng công nghệ thông minh trong khi Chính phủ đã có các nghị quyết yêu cầu quản lý dạy học trong môi trường số [20]. Nguyên nhân chính là đội ngũ giảng viên chưa được đào tạo một cách thực sự bài bản, khó phát huy được giá trị của tài nguyên bối cảnh hóa và khả năng tương tác mà hệ thống giảng dạy thông minh có thể mang lại, khiến phương pháp giảng dạy vẫn mang nặng tính truyền thống.

3.4 Liên kết trường học - doanh nghiệp còn lỏng lẻo

Đào tạo nhân lực tiếng Trung thương mại tại Việt Nam hiện chưa hình thành được một vòng khép kín tương tác hiệu quả giữa "giáo dục - nghề nghiệp", hợp tác giữa nhà trường và doanh nghiệp vẫn tồn tại nhiều khiếm khuyết mang tính cấu trúc. Các chương trình chuyên ngành tiếng Trung tại các trường đại học tuy có bao hàm nội dung về ngôn ngữ, văn hóa Trung Quốc, nhưng phần lớn vẫn thiên về định hướng ngôn ngữ - văn học truyền thống và biên - phiên dịch [2], thiếu chiều sâu và chiều rộng trong việc kết hợp với thực tiễn của các ngành nghề kinh tế - thương mại.

Ngoài ra, mảng giảng dạy thực hành đặc biệt yếu; dù nhiều trường đại học đã tổ chức các cuộc thi tiếng Trung hoặc chương trình thực tập tại doanh nghiệp Trung Quốc, nhưng các hoạt động này có phạm vi bao phủ hạn chế, thời gian ngắn, chưa đủ để hình thành một mô hình tích hợp "giáo dục - nghề nghiệp" có quy mô và tính hệ thống. Kết quả là khoảng cách giữa việc học ngôn ngữ và yêu cầu thực tế của công việc vẫn còn đáng kể. Có thể thấy, bản chất của sự đứt gãy trong hệ sinh thái đào tạo này xuất phát từ sự "lệch pha" mang tính hệ thống giữa mô hình đào tạo truyền thống và nhu cầu phát triển nhân lực trong thời đại số.

4. Chiến lược đào tạo nhân lực tiếng Trung thương mại tại Việt Nam dưới sự hỗ trợ của công nghệ trí tuệ nhân tạo

4.1 Xây dựng hệ sinh thái nội dung động dựa trên công nghệ xử lý ngôn ngữ tự nhiên (NLP)

Để giải quyết hai điểm nghẽn lớn là thiếu bản địa hóa giáo trình và chậm số hóa, cần xây dựng một hệ thống môn học thông minh lấy NLP làm nền tảng, tích hợp đồng thời ba yếu tố "kỹ năng tiếng Trung + thực hành thương mại + công cụ số". Hệ thống này có thể bao gồm các mô hình cốt lõi như: đánh giá ngữ âm thông minh (chỉnh lỗi phát âm theo thời gian thực), chẩn đoán lỗi bằng biểu đồ tri thức (xác định nguồn gốc sai ngữ pháp), và tạo nội dung bối cảnh hóa (tự động gợi ý ví dụ bản địa hóa nội dung). Ví dụ như hệ thống sẽ ngay lập tức cung cấp giải thích và bài luyện chuyên sâu cho các thuật ngữ dễ nhầm lẫn đối với người học Việt Nam như "定金" và "订金". Thông qua chu trình khép kín "học tập - luyện tập - đánh giá - phản hồi", hệ thống tạo báo cáo năng lực cá nhân hóa, giúp cải thiện chính xác năng lực của từng người học. Trong tương lai, có thể tích hợp thêm các công cụ dịch máy để tối ưu toàn diện chất lượng và hiệu quả giảng dạy.

4.2 Ứng dụng tích hợp công nghệ VR và dịch thuật thời gian thực

Giảng dạy tiếng Trung thương mại truyền thống còn thiếu các mô phỏng bối cảnh thực, dẫn đến tình trạng "hiểu lý thuyết nhưng khó thực hành, biết ngữ pháp nhưng khó giao tiếp". Để khắc phục, cần kết hợp công nghệ thực tế ảo (VR) và dịch thuật AI thời gian thực nhằm xây dựng một nền tảng thực hành thương mại nhập vai, thu hẹp khoảng cách giữa lý thuyết và ứng dụng. Người học có thể sử dụng thiết bị VR để bước vào môi trường thương mại ảo được tái hiện chân thực, dưới sự hướng dẫn của hệ thống thông minh, hoàn thành quy trình huấn luyện toàn diện: "trao đổi thông tin sản phẩm - thương lượng giá - xác nhận điều khoản hợp đồng - thanh toán quốc tế". Toàn bộ nội dung tương tác và thao tác của người học sẽ được hệ thống ghi nhận để đánh giá quá trình.

Việc tích hợp dịch thuật thời gian thực giúp tăng hiệu quả thực hành: thiết bị thông minh có thể dịch song ngữ ngay lập tức, ví dụ khi người học nhầm lẫn "信用证" thành "信用证书", hệ thống sẽ hiển thị cách dịch chính xác kèm chú giải và ngữ cảnh sử dụng. Để giải quyết những xung đột văn hóa thường gặp trong thương mại Việt - Trung, nền tảng có thể phát triển "mô hình huấn luyện độ nhạy văn hóa". Thông qua công nghệ bắt chuyển động và NLP, hệ thống sẽ phân tích các ví dụ thực tiễn (như hành vi ép giá trực tiếp gây mất thiện cảm, hiểu lầm ngôn ngữ cơ thể…), đánh dấu những điểm cấm kỵ văn hóa. Cụ thể khi người học tương tác với đối tác ảo AI, cảm biến sinh học còn có thể theo dõi biểu cảm, ánh mắt và cử chỉ, cảnh báo rủi ro văn hóa tiềm ẩn (ví dụ nhìn chằm chằm quá lâu), đồng thời cung cấp gợi ý bản địa hóa như thay vì nói trực tiếp "giá quá cao", có thể dùng "điều chỉnh giá theo biến động thị trường".

4.3 Nâng cấp năng lực số hóa cho đội ngũ giảng viên

Để khắc phục hạn chế về năng lực số của giảng viên tiếng Trung thương mại, cần triển khai các khóa đào tạo tập trung vào ba nhóm kỹ năng cốt lõi: (i), Ứng dụng công cụ thông minh: hướng dẫn giảng viên sử dụng các công cụ biên tập bài giảng VR, hệ thống đánh giá ngữ

âm AI, và tùy chỉnh kho dịch thuật thuật ngữ thương mại. (ii), Phát triển môn học dựa trên tri thức: chuyển các ca thực tế từ doanh nghiệp thành tài liệu giảng dạy mô phỏng "bối cảnh + ngữ pháp + văn hóa". (iii), Phân tích dữ liệu học tập: đào tạo giảng viên sử dụng hệ thống quản lý học tập (LMS) để theo dõi tiến trình học của sinh viên, xác định điểm yếu và tư vấn cá nhân hóa.

Bên cạnh đó, cần xây dựng cơ chế nghiên cứu hợp tác giữa các trường đại học Trung Quốc và Việt Nam, thành lập nhóm giảng dạy AI để cùng phát triển tài nguyên bản địa hóa. Ví dụ, khi phân tích dữ liệu học tập phát hiện sinh viên hay nhầm lẫn "承兑交单" và "付款交单", nhóm có thể biên soạn "Sổ tay tra cứu nhanh thuật ngữ dễ nhầm lẫn"; giảng viên có thể thiết kế hội chợ thương mại ảo trong bài giảng VR, dùng AI chỉnh lỗi thanh điệu "合同" và "合作", hoặc tùy chỉnh kho thuật ngữ tiếng Việt cho các từ "FOB", "SCT 关税". Hệ thống LMS có thể nhận diện lỗi của sinh viên và gợi ý bài tập tương ứng, nâng cao độ chính xác trong giảng dạy.

4.4 Xây dựng mô hình hợp tác ba chiều "Trường học - Doanh nghiệp Việt Nam - Doanh nghiệp Trung Quốc"

Hiện nay, hợp tác giữa trường học và doanh nghiệp trong đào tạo tiếng Trung thương mại còn ở mức khiêm tốn, chủ yếu dừng ở việc tham quan, giới thiệu quy trình, khiến nội dung đào tạo tách rời yêu cầu làm việc thực tế. Cần thiết lập cơ chế phối hợp sâu, trong đó AI đóng vai trò cầu nối, hình thành khung hợp tác ba bên "Trường học - Doanh nghiệp Việt Nam - Doanh nghiệp Trung Quốc". Cụ thể, các doanh nghiệp Việt Nam cung cấp phản hồi thực tiễn về nhu cầu kỹ năng và các tình huống họ thường gặp trong giao dịch với đối tác Trung Quốc. Các dữ liệu này kết hợp với thông tin từ doanh nghiệp Trung Quốc để hình thành kho tình huống song phương đa dạng, cập nhật liên tục. AI sẽ phân tích, chuẩn hóa và chuyển các tình huống này thành bài tập tương tác đa cấp độ, tích hợp ngôn ngữ, pháp lý và văn hóa kinh doanh.

Ngoài ra, các trường đại học cũng cần triển khai đào tạo theo mô hình "học tập - thực hành - phản hồi". Sinh viên không chỉ học lý thuyết mà còn được thực hành mô phỏng với "đối tác ảo" AI qua các kịch bản linh hoạt như viết email, thương lượng giá hay giải quyết tranh chấp. Sau mỗi phiên tương tác, hệ thống phân tích điểm mạnh, điểm yếu và đưa ra gợi ý cải thiện cá nhân hóa.

Cuối cùng, để đảm bảo tính bền vững, cần có cơ chế đánh giá định kỳ với sự tham gia của cả ba bên: trường học, doanh nghiệp Việt Nam và doanh nghiệp Trung Quốc. Hoạt động này giúp liên tục cập nhật tình huống và nội dung đào tạo, đảm bảo bám sát thực tế kinh doanh, thích ứng với thay đổi thị trường. Qua đó, sinh viên có thể nâng cao năng lực làm việc độc lập, rút ngắn khoảng cách giữa kiến thức và thực tiễn, đồng thời góp phần nâng cao chất lượng nguồn nhân lực phục vụ thương mại song phương Việt - Trung.

5. Kết luận

Nghiên cứu này tập trung phân tích tình trạng khó khăn và đề xuất giải pháp đào tạo nhân lực tiếng Trung thương mại tại Việt Nam trong bối cảnh hợp tác kinh tế - thương mại Việt

- Trung ngày càng mở rộng và quá trình chuyển đổi số diễn ra mạnh mẽ. Nhu cầu về nhân lực phức hợp "ngôn ngữ + thương mại + công nghệ số" đang bùng nổ do thương mại điện tử xuyên biên giới và sự gia tăng hiện diện của doanh nghiệp Trung Quốc tại Việt Nam. Tuy nhiên, hệ thống đào tạo hiện nay còn nhiều hạn chế mang tính hệ thống: giáo trình thiếu bản địa hóa và chậm số hóa; công nghệ ứng dụng chưa sâu, mô hình giảng dạy cố định; giảng viên thiếu kinh nghiệm thực tiễn và năng lực số; cơ chế hợp tác giữa nhà trường và doanh nghiệp còn lỏng lẻo, dẫn đến sự "lệch pha" giữa cung và cầu nhân lực.

Trước những nút thắt này, công nghệ AI mang lại giải pháp then chốt cho việc tái cấu trúc đào tạo. Ứng dụng công nghệ NLP giúp tạo nội dung động, phản hồi cá nhân hóa, khắc phục hạn chế của giáo trình tĩnh và thiếu bản địa hóa. Nền tảng thực tập ảo tích hợp công nghệ VR và dịch thuật thời gian thực thu hẹp khoảng cách giữa lý thuyết và thực hành, đồng thời phát triển năng lực giao tiếp thương mại xuyên văn hóa. Cùng với đó, cơ chế bồi dưỡng giảng viên thông qua đào tạo đa chiều và nghiên cứu hợp tác Việt - Trung giúp nâng cao năng lực giảng dạy số và kinh nghiệm thực tiễn. Ngoài ra, mô hình hợp tác ba bên "Trường đại học - Doanh nghiệp Việt Nam - Doanh nghiệp Trung Quốc" với AI làm cầu nối liên kết chặt chẽ giữa nội dung đào tạo và nhu cầu công việc, khắc phục tình trạng đứt gãy "học tập - ứng dụng".

Vì vậy, xây dựng hệ thống đào tạo nhân lực tiếng Trung thương mại theo định hướng "hỗ trợ công nghệ + tái cấu trúc bối cảnh + nâng cấp năng lực + phối hợp sinh thái" là con đường then chốt để đáp ứng yêu cầu phát triển kinh tế thương mại Việt - Trung và tiến trình chuyển đổi số. Hệ thống này không chỉ mang lại khung cải cách thực tiễn cho việc giảng dạy tiếng Trung thương mại tại Việt Nam, mà còn là tài liệu tham khảo có giá trị cho việc quốc tế hóa đào tạo nhân lực ngoại ngữ kinh doanh trong bối cảnh sáng kiến "Một vành đai và Một con đường" của Trung Quốc. Điều này có ý nghĩa quan trọng trong việc tăng cường hợp tác kinh tế - thương mại và giao lưu văn hóa Việt - Trung trong giai đoạn mới.

Tài liệu tham khảo và chú thích

1. 崔丹, & 李国平. (2024). 人工智能人才培养与教育政策的全球新走向. *光明日报.* Retrieved from https://epaper.gmw.cn/gmrb/html/2024-03/21/nw.D110000gmrb_20240321_1-14.htm

2. 黄乐平、刘俊、陈旭东、丁宁、张宇、张皓怡、王心怡. (2024）. AI 大模型时代的全球产业链重构 [证券研究报告]. *华泰研究.*

3. Huo, D., Ouyang, R., Tang, A., Gu, W., & Liu, Z. (2024). *New growth in cross-border E-business: evidence from gray forecasting to cross-border E-business in China*. *Journal of Internet and Digital Economics,* 4(1), 12-29.

4. Cheng, H., Yang, Y., & Wu, Y. (2023). The Construction of Language Service System for Hainan Free Trade Port Based on Stakeholders. *Information Systems and Economic*s, 4(8), 106-112.

5. Columbia Center for Teaching and Learning. (n.d.). *Incorporating Generative AI in Teaching and Learning: Faculty Examples Across Disciplines. Columbia University*. Retrieved from https://ctl.columbia.edu/resources-and-technology/resources/incorporating-generative-ai-teaching/

6. Tạp chí Công Thương. (2024). *Mở rộng quy mô thương mại hàng hóa Việt Nam - Quảng Tây (Trung Quốc)* . Tạp chí Công Thương. Retrieved from https://tapchicongthuong.vn/mo-rong-quy-mo-thuong-mai-hang-hoa-viet-nam-quang-tay--trung-quoc-126023.htm

7. 王道征. (2021). “数字丝绸之路” 视域下的中越数字经济合作. *湖北经济学院学报*, 19(3),32-38.

8. 刘娴. (2020). 越南劳动关系研究: 中资企业投资风险与防范. *中国劳动关系学院学报*, 34(4), 47-57.

9. Hà Văn. (2025). *Kinh tế số Việt Nam tăng trưởng nhanh nhất khu vực*. *Báo Chính phủ*. Retrieved from https://baochinhphu.vn/kinh-te-so-viet-nam-tang-truong-nhanh-nhat-khu-vuc-102250206152332651.htm

10. Cuong, N.H., Le, D.K. (2024). *Vietnam Digital Transformation and the Way Forward*. In: Tran, H.V.T., Shioji, H., Le, H.L.T., Hayashi, T. (eds) *Knowledge Transformation and Innovation in Global Societ*y. Springer, Singapore.

11. 刘佳. (2025). 越南中文教育历史、现状及发展策略. *现代教育论坛*, 8(1),222-224.

12. 黄志媛,尹建玉. (2024). 越南高校中文专业人才培养方案比较研究.*社会科学前沿*, 2024, 13(6), 121-131

13. 曾小燕,阮德英. (2023). 越南中文教育现状及发展论略. *云南师范大学学报（对外汉语教学与研究版）*,21(2),37-48.

14. 孙晓冬. (2021). 中国-东盟背景下的高职应用越南语专业跨国校企合作模式探索研究. *中外交流*, 028(005), 868-869.

15. 阮芳姮. (2023). 越南商务汉语教材的话题,情景的考察及国别化编写建议——以三套初,中级商务汉语教材为例. *云南师范大学*.

16. Thủ tướng Chính phủ. (2021, 15 tháng 6). Quyết định số 942/QĐ-TTg phê duyệt Chiến lược phát triển Chính phủ điện tử hướng tới Chính phủ số giai đoạn 2021-2025, định hướng đến năm 2030. *Cổng thông tin điện tử Chính phủ*. Retrieved from https://vanban.chinhphu.vn/?pageid=27160&docid=203403

17. 阮氏明庄. (2019). 越南商务公司汉语应用调查研究——以胡志明市为例. *重庆师范大学*.

18. 代以平, & 冯珊珊. (2023). 越南职业教育数字化转型的背景,举措及启示. *教育与职业*(1), 74-81.

19. 尹雅丽,赵昱迪,江梦婷,马早明.(2023).越南高等教育数字化转型的战略选择.*中国教育信息化*,29(11):51-63

20. Tham, P. T. H., & Cheng, Z. Y. (2019). *Investigation and Analysis on the Development of Chinese Teachers in Vietnam University*. *Journal of Education and e-Learning Research*, 6(2), 69-75.

INTERGRATING AI TOOLS IN SPEAKING SKILLS: A CASE STUDY FROM TEACHER'S PERSPECTIVES ON PERSONALIZED INSTRUCTION

Phùng Thị Bích Ngọc [1], Pham Ngoc Diem[2]

ABSTRACT

This article investigates the integration of artificial intelligence (AI) tools in speaking skill instruction, focusing specifically on teachers' perspectives regarding their role in facilitating personalized learning. Drawing on qualitative data from educators involved in an AI-supported speaking curriculum, the study explores how AI tools influence instructional practices and learner engagement. Teachers reported that AI technologies provided real-time, individualized feedback, enabling students to identify and address errors independently. This, in turn, supported the development of higher-order speaking skills, such as idea development, coherence, and discourse organization. Participants also observed that AI tools allowed for differentiated instruction, catering to learners' varying proficiency levels and learning paces. While initial skepticism was noted, many educators experienced a positive shift in perception, viewing AI not as a replacement for traditional methods but as a complementary resource that enhances instructional effectiveness. The findings highlight the potential of AI to transform speaking instruction by promoting learner autonomy, increasing engagement, and supporting the development of personalized, student-centered learning environments..

Keywords: Integrating AI, Speaking skill, development, learning environments

I. INTRODUCTION

Mastering the art of speaking is the most important aspect of learning a foreign language. "Success is measured in terms of the ability to carry out a conversation in the language" (Nunan, 1991, p. 39). As a result, learners often measure their success in language learning as well as the effectiveness of their English course on the basis of how much they feel they have improved their spoken language proficiency (Richard, 2008). Artificial Intelligence (AI) tools are increasingly shaping contemporary educational practices, offering transformative potential in both teaching and learning. Within the domain of language education, AI technologies are particularly impactful in the instruction of speaking skills, an area that remains one of the most complex and demanding for learners, especially those preparing for high-stakes assessments such as the IELTS. The development of speaking proficiency requires real-time interaction and immediate, constructive feedback, which traditional methods such as teacher-led drills and peer assessment, often struggle to provide consistently due to time and resource constraints. Emerging AI-powered applications, such as Google Gemini, address these challenges by offering instant, individualized feedback, tailored guidance, and adaptive

[1] Cử Nhân, Học viên cao học ngành Ngôn ngữ Anh, Trường Đại học Sài Gòn, *Email: phungngocsherri@gmail.com*
[2] TS., Trường Đại Học Sài Gòn, *email: pndiem1977@sgu.edu.vn*

learning pathways that respond to learners' specific needs (Pokrivčáková, 2019). These tools do not aim to replace traditional pedagogies but rather to enhance them, enabling a more personalized, autonomous, and efficient approach to speaking instruction that supports learners in achieving communicative competence more effectively.

A central concern surrounding the integration of artificial intelligence (AI) in speaking instruction relates to teachers' perceptions and attitudes toward these emerging technologies. While some educators regard AI as a valuable instructional aid that can enhance student engagement and improve learning outcomes, others remain cautious, fearing that excessive reliance on AI may negatively impact students' motivation, effort, and development of critical thinking skills. As Tseng and Warschauer (2023) point out, some teachers worry that dependence on AI feedback may reduce learners' ability to self-monitor and independently develop their spoken language proficiency. Additionally, the challenge of integrating AI into traditional classroom settings without diminishing the pedagogical role of the teacher is a pressing concern for both educators and policymakers.

Despite these reservations, AI tools offer significant potential in personalizing the teaching of speaking skills. Unlike conventional methods that often adopt a uniform approach to instruction, AI technologies can analyze individual learner profiles and provide tailored speaking tasks, real-time corrective feedback, and adaptive learning pathways (Mitra & Banerjee, 2022). This capacity for personalization supports learner autonomy and enables more efficient skill development. It also allows teachers to concentrate on cultivating higher-order speaking competencies, such as idea organization, fluency, coherence, and communicative effectiveness. Nevertheless, as Korenev (2024) emphasizes, AI should be seen as a complement to, not a replacement for, human instruction. A balanced approach is essential, wherein AI enhances the learning experience while preserving the irreplaceable role of the teacher in facilitating meaningful, context-rich communication practice.

This study examines the integration of AI tools in speaking instruction from the perspective of teachers. It explores how AI can support personalized learning in speaking while addressing pedagogical concerns related to its use. By drawing on teachers' experiences and viewpoints, this study contributes to the growing body of research on AI in language education and provides practical insights into the effective and responsible use of AI as a supportive tool in speaking instruction.

II. LITERATURE REVIEW

2.1. AI's Role in speaking Instruction

2.1.1. AI Tools Assist in speaking Development

AI tools play a significant role in supporting speaking skill development by offering personalized, flexible, and data-driven learning experiences. They can analyse individual learner performance and tailor speaking activities to specific proficiency levels, ensuring that practice is both relevant and challenging. Through features such as automated pronunciation

analysis, grammar checks, and fluency scoring, AI provides immediate and targeted feedback, allowing learners to make timely corrections and reinforce accurate language use. Virtual conversation partners and chat-bots simulate real-life communication scenarios, enabling students to practise speaking in a low-pressure environment while building confidence. Additionally, AI-driven analytic track progress over time, helping teachers identify strengths and areas for improvement so they can adjust instruction accordingly. These tools also enhance accessibility by allowing learners to practise anytime and anywhere, supporting remote and diverse student populations. By incorporating elements of gamification, AI applications further boost engagement and motivation, making speaking practice more interactive and enjoyable. When implemented within sound pedagogical frameworks, AI serves as a powerful complement to teacher-led instruction, enriching the learning experience without replacing the human element..

2.1.2. Automated Grading and Error Detection

One of the key benefits of AI in speaking development is its ability to assess oral performance quickly, providing immediate feedback and structured evaluation. AI-powered speech analysis systems, often using automatic speech recognition (ASR) and natural language processing (NLP), evaluate pronunciation, fluency, intonation, and vocabulary use based on predefined criteria. As Pokrivčáková (2019) observes, such automated assessment reduces teachers' workloads by delivering objective, bias-free evaluations. This ensures fairness and allows educators to focus more on developing higher-order speaking skills, such as building arguments, negotiating meaning, and maintaining conversational flow.

Error detection is another highly valuable feature of AI tools in speaking instruction. Applications like ELSA Speak or Speechace analyse learners' spoken output to identify pronunciation inaccuracies, misplaced stress, awkward pauses, and inappropriate word choices. Using deep learning algorithms, these systems can detect not only basic errors but also more subtle issues such as inconsistent rhythm or unclear articulation that might hinder listener comprehension. According to Moulieswaran and Kumar (2022), AI-generated oral feedback is particularly beneficial for second-language learners, as it helps them recognize recurring pronunciation and fluency challenges. This fosters self-awareness and encourages learners to correct their speaking habits independently, leading to improved accuracy and confidence over time.

Beyond pinpointing mistakes, AI speaking tools often provide detailed, categorized feedback that explains the nature and severity of each issue. This reflective approach enables learners to understand why a change is necessary, rather than simply accepting corrections. In addition, some AI platforms can highlight repetitive vocabulary or overused sentence patterns, prompting students to expand their lexical range and vary their structures for more engaging communication. By combining automated speaking assessment with in-depth, personalized feedback, AI tools not only improve the technical aspects of oral performance but also empower learners to take an active role in developing their speaking proficiency.

2.1.3. Grammar, Structure, and Vocabulary Enhancement

Beyond simple error detection, AI tools play a significant role in improving grammar accuracy, sentence structure, and vocabulary use in speaking. Many AI-powered speech platforms provide real-time correction during oral practice, allowing learners to hear and address errors in grammar or word order immediately. These tools not only highlight the problematic segments but also explain the reasons for the correction, thereby strengthening learners' understanding of language rules and helping them internalize correct structures over time (Chen, 2024). AI systems can also suggest alternative ways to phrase ideas to make speech more coherent and logical, resulting in more sophisticated and polished oral delivery.

Vocabulary development is another area where AI greatly supports speaking proficiency. AI speech training tools analyse a learner's spoken output and suggest alternative words, paraphrases, and contextually appropriate expressions to enrich communication. Bilad et al. (2023) note that learners who receive AI-based vocabulary suggestions expand their lexical range more effectively than those using traditional practice methods. These tools also detect overused words or phrases and offer suitable replacements, which promotes variety and precision in speech.

In addition, advanced AI platforms help make speech more natural and audience-appropriate by recommending changes to tone, register, and style, whether the goal is academic presentation, formal discussion, or conversational fluency. By offering targeted feedback and clear, actionable suggestions, AI enables learners to produce speech that is not only more accurate but also more persuasive and contextually appropriate.

Unlike static grammar or dictionary tools, AI speaking applications personalize feedback based on a learner's previous performances. This adaptive approach ensures that corrections and suggestions are tailored to the learner's proficiency level and communication style. Over time, exposure to such customized feedback builds learners' confidence and autonomy, enabling them to speak more effectively without constant external assistance. AI is particularly valuable for IELTS speaking preparation and for developing overall communicative competence in a second language.*Providing*

2.1.4. Common Concerns & Misconceptions

The integration of AI tools in speaking instruction has sparked significant discussion among educators, learners, and researchers. While AI offers clear advantages such as instant pronunciation feedback, personalized speaking practice, and simulated conversational scenarios concerns remain about its role in language learning. These concerns primarily focus on three areas: the possibility of AI reducing the need for teacher-led interaction, its influence on learners' spontaneous speaking and critical thinking skills, and the accuracy and appropriateness of AI-generated feedback. To maximize its potential, AI in speaking instruction should be implemented as a supportive complement to human teaching, ensuring that technology enhances rather than replaces authentic communication and guided practice.

2.1.5. The Potential of AI to Replace Teachers

One of the most prevalent concerns regarding AI integration in speaking instruction is the fear that it may replace human teachers. As AI technology advances, many educators worry that institutions may begin prioritizing AI-driven lessons as a cost-saving measure, potentially reducing the demand for human instructors. Zimotti et al. (2024) note that some teachers perceive AI as a threat to their job security, particularly as AI systems become more capable of providing pronunciation feedback, evaluating fluency, and even designing speaking activities tailored to individual learners' needs.

However, AI should be viewed not as a replacement but as a tool to enhance teaching. While AI can automate certain aspects of speaking practice—such as real-time error detection and pronunciation drills, it lacks the human connection essential for fostering motivation, empathy, and authentic communication. Nazim (2024) argues that AI cannot fully interpret learners' emotions, adjust to cultural nuances, or adapt to individual communication styles as effectively as human teachers. Although AI can identify mispronunciations and suggest phrasing improvements, it cannot replicate the encouragement, spontaneous conversation, and personalized mentor-ship that educators provide.

Furthermore, the effectiveness of AI in speaking instruction depends heavily on human oversight to ensure ethical and contextually appropriate use (Kovalenko & Baranivska, 2024). The teacher's role is evolving rather than disappearing, as educators must now guide students in using AI tools effectively, interpreting AI feedback critically, and balancing automation with human interaction. Instead of replacing teachers, AI can reduce administrative burdens such as repetitive pronunciation assessments, allowing educators to focus more on higher-order speaking skills, personalized coaching, and meaningful learner engagement. When thoughtfully integrated, AI serves as a powerful complement to human instruction, optimizing the speaking learning experience without diminishing the essential role of educators.

2.1.7 The Influence of AI on Critical Thinking

Another major concern is the potential for AI tools to hinder students' critical thinking abilities in speaking. Some educators fear that over-reliance on AI-generated feedback, such as instant pronunciation corrections or suggested speaking prompts, may lead students to accept guidance uncritically, reducing their ability to self-monitor and improve their speaking independently. When students depend heavily on AI for grammar accuracy, vocabulary choices, or sentence framing during oral practice, they may bypass essential cognitive processes such as reflection, self-correction, and spontaneous idea formulation, all of which are vital for developing speaking proficiency. However, when integrated strategically, AI can serve as a powerful tool to enhance rather than suppress critical thinking in speaking. Tseng and Warschauer (2023) suggest that students should be taught how to engage critically with AI feedback, evaluating suggestions, considering alternative expressions, and making informed decisions about which recommendations to adopt. AI tools such as Google Gemini can function as interactive speaking assistants, prompting learners to assess tone, refine pronunciation, and

experiment with various ways of expressing ideas before delivering their final speech. In this way, AI encourages a more analytical speaking process that fosters independent problem-solving skills.

Furthermore, educators play a crucial role in ensuring AI acts as an aid rather than a crutch in speaking development. By designing AI-assisted speaking tasks that require students to justify their word choices, reflect on AI feedback, and compare AI suggestions with peer or teacher evaluations, teachers can promote deeper engagement with communication principles. Encouraging students to critically assess AI's advice in relation to real-world speaking contexts can further strengthen their ability to think independently and communicate effectively. As a result, when guided effectively, AI has the potential to enhance learners' speaking skills, sharpen their analytical thinking, and help them take a more active role in their oral communication process.

2.1.8 .The Reliability of AI Feedback

The reliability of AI-generated feedback remains a major concern in AI-assisted speaking instruction. While AI tools are designed to provide immediate and structured feedback on pronunciation, fluency, and vocabulary, they are not always accurate. Problems such as misinterpreting intended meaning, failing to capture full conversational context, and lacking flexibility for different speaking styles have been reported in studies by Billingsley and Gardner (2024). AI feedback sometimes encourages changes that do not sound natural, oversimplify responses, or fail to improve coherence in speech. Moreover, AI often struggles with tone, cultural nuances, and audience expectations, critical aspects of effective oral communication.

Furthermore, integrating AI responses with teacher explanations and peer evaluations can improve learning outcomes. As AI technology continues to advance, teachers must regularly monitor its reliability to prevent students from becoming overly dependent on automated feedback. By teaching students how to use AI effectively, educators can help them distinguish between helpful and misleading advice.

2.2. Step-by-Step Guide to Integrating AI in speaking Lessons

The integration of AI tools in IELTS speaking lessons requires a structured approach to ensure that students can effectively use these technologies to enhance their oral communication skills while maintaining fluency, accuracy, and originality. This section outlines a step-by-step process for incorporating AI tools such as ChatGPT, Speechling, and ELSA Speak into speaking instruction, focusing on preparation, introduction, and guided activities that balance AI assistance with traditional speaking practice.

Stage 1: Preparing for AI Integration

The first and most important step in integrating AI into speaking lessons is identifying clear learning goals. Teachers must determine specific objectives that align with students' needs whether improving pronunciation, increasing lexical range, refining grammatical

accuracy in speech, or enhancing fluency and coherence. Clearly defined learning outcomes guide the selection of the most suitable AI tools to match teaching aims. For example, ELSA Speak provides advanced pronunciation analysis, allowing students to detect and correct specific sound errors; Speechling offers personalized feedback on speech recordings, helping learners practice intonation and rhythm; and ChatGPT can serve as a conversation partner to simulate IELTS speaking topics, giving students opportunities for real-time idea generation and response structuring.

Tool selection should be based on their ability to meet pedagogical goals rather than their technological novelty. As Billingsley and Gardner (2024) emphasize, it is essential to choose AI tools that enhance students' speaking skills rather than replace authentic interaction. AI should function as a learning aid that encourages students to think on their feet, develop spontaneity, and engage in active listening, not merely as a scripted response generator. Therefore, teachers must implement AI in ways that promote independent speaking practice, ensuring students build the confidence, adaptability, and communicative competence required for success while preserving their own speaking style and voice.

Setting Expectations: Teaching AI Literacy & Responsible Use

Making clear rules for AI use and teaching students how to apply them correctly is an essential part of adding AI into speaking lessons. Learners need a complete understanding of AI tools—knowing both their strengths and their limitations. Kovalenko and Baranivska (2024) note that AI-generated feedback is not always contextually accurate, highlighting the importance of human judgment in reviewing AI suggestions. AI should not be treated as a replacement for active thinking and genuine communication but rather as a supportive tool that complements human effort.

To ensure responsible use, teachers should design lessons that clearly address AI rules, academic integrity, and best practices for integrating technology into speaking practice. Students must be taught how to evaluate AI feedback, recognizing that while AI can help with pronunciation, vocabulary suggestions, and fluency tips, it can also misinterpret tone, register, or meaning in spoken language. Educating learners about the risks of over-reliance on technology encourages critical thinking and adaptability. By fostering AI literacy, teachers can guide students to become skilled, selective users of AI, ensuring that the technology enhances their speaking ability rather than replacing their own voice.

Establishing a Balance Between AI Assistance and Manual Effort

Finding the right balance between AI support and students' own effort is essential so that learners remain active in their speaking practice rather than relying on AI-generated corrections without engaging their minds. One major challenge in using AI for speaking instruction is preventing over-dependence on automated suggestions, which can lead to less self-monitoring and reflection. Nadhifah et al. (2024) note that excessive reliance on AI can

hinder the development of critical thinking skills, as students may simply accept AI recommendations without fully understanding the reasoning behind them.

To address this, teachers should establish clear guidelines requiring students to review and adjust AI feedback themselves, and to explain why they choose certain changes to their spoken responses. This approach promotes active learning, helps learners refine their ability to assess and apply feedback, and encourages independence in improving speech. Building the habit of self-correction can also boost confidence in using language. Activities where students compare AI suggestions with feedback from teachers or peers can strengthen analytical thinking and deepen understanding of pronunciation, tone, and fluency. By treating AI as a supportive resource rather than the final authority, teachers can ensure that students maintain control over their speaking development while still benefiting from technological advancements.

Stage 2: Introducing AI Tools to Students

Demonstrating AI's Strengths and Limitations

Showing both the strengths and the limits of AI is really important if we want students to truly understand how these tools work. AI should be seen as a helpful assistant, not a perfect authority. Teachers can show its abilities by doing live demonstrations, where they analyze AI-generated corrections and explain how the tool finds mistakes and suggests improvements. This lets students see how AI can help with pronunciation, grammar, vocabulary, and structure.

Highlighting these weaknesses helps students stay critical of AI feedback instead of blindly accepting it. When students compare AI suggestions with their own reasoning and with teacher feedback, they get more engaged with their writing. By talking about both the benefits and drawbacks of AI, teachers can help students use it wisely, balancing technology with their own critical thinking skills.

Teaching Students How to Use AI for Self-Correction

Teaching students how to use AI properly for self-editing is essential to help them interact critically with technology while improving their speaking skills. AI tools can be very useful in finding mistakes, suggesting alternative phrasing, and improving grammatical accuracy. However, students must learn to think critically about AI feedback rather than simply accepting changes without questioning them.

Teachers play a crucial role in showing students how to distinguish between AI suggestions that improve clarity and flow, and those that might unintentionally change the meaning of their ideas. Encouraging students to question AI-generated changes can strengthen their decision-making skills in writing. Discussing why certain AI edits are effective or ineffective also deepens their understanding of language structure and the importance of stylistic choices.

Including reflective writing exercises, where students explain their reasons for accepting or rejecting AI suggestions, makes the learning process more active and promotes higher-level thinking. By taking a thoughtful approach to AI assistance, students can maintain control over their writing, using AI as a supportive tool rather than a fully automatic error-fixing machine. This balanced approach not only enhances writing skills but also ensures students remain critical thinkers when using technology in education..

Designing AI-Integrated Speaking Lessons

Creating AI-assisted speaking lessons requires a structured approach to ensure that students develop their oral communication skills while still relying on their own thinking and expression. The first stage is independent speaking practice, where students prepare and deliver short talks, responses, or role plays without AI assistance. This helps them organize ideas, improve fluency, and build confidence before using technology.

The second stage is AI-supported feedback, in which students use AI tools—such as speech recognition apps or pronunciation analyzers—to identify pronunciation issues, grammar errors, or pacing problems. Teachers should guide learners to critically evaluate this feedback, rather than simply trying to sound "exactly like the AI," so that natural communication is preserved.

In the discussion phase, teachers and students review AI feedback together, exploring why certain corrections are suggested and deciding when human judgment is more appropriate, especially for tone, cultural nuance, and conversational style.

The final stage is refined speaking performance, where students incorporate both AI suggestions and teacher feedback into a new attempt. This planned process allows learners to take advantage of AI's real-time analysis while remaining active, thoughtful participants in their own speaking development, ultimately leading to greater fluency, accuracy, and communicative confidence.

Stage 3: AI-Assisted speaking Activities

AI tools can play an important role in developing speaking skills by giving learners more opportunities to practice and receive feedback. For example, AI-powered language apps can simulate real-life conversations, ask follow-up questions, and correct pronunciation in real time. This helps learners improve their fluency, accuracy, and confidence without always needing a teacher present. Some AI tools also record and analyze speech, showing where a learner pauses too often, mispronounces words, or uses limited vocabulary. By getting instant, personalized feedback, learners can identify weak areas and work on them more effectively. However, to get the best results, students should combine AI practice with real conversations, so they can also develop natural communication skills, cultural understanding, and the ability to respond flexibly in unpredictable situations.

Teacher-Led AI Feedback Review

Teachers play a really important role when it comes to using AI-generated feedback in writing. It's not just about showing students what the AI says, it's about sitting down with them and looking closely at those suggestions. In these discussions, students can think critically about whether the feedback is actually correct and if it matches proper academic writing standards. Teachers can also help students tell the difference between useful advice and feedback that's misleading or too superficial. This way, students learn about deeper aspects of writing, like coherence, staying on task, and using language that's appropriate for the context. And by explaining their own editing choices, students get better at judging feedback and making stronger revisions. It's also important to point out that AI has its limits. For example, it's usually good at grammar, but it often misses bigger issues, like problems with the structure of an argument or cultural differences in writing. Talking about these weaknesses in class helps students get a more complete understanding of what makes writing effective.

2.4. Research questions

This research aims at answering the following questions:

1. How do teachers perceive the integration of AI tools in speaking skill instruction to support personalized learning?

2. In what ways do AI tools influence teachers' instructional practices and learners' engagement in speaking activities?

3. METHODOLOGY

Questionnaires are usually used for collecting data in the research, because they have a number of advantages. Seliger and Shohamy (1989, p.172) lists the advantages of questionnaires as follows:

Firstly, questionanaires are self-administered and can be given to large groups of subjects at the same time. They are therefore less expensive to administer than other procedures such as interviews. Secondly, when anonymity is assured, subjects tend to share information of a sensitive nature more easily. Thirdly, since the same questionnaire is given to all subjects, the data are more uniform and standard. Finally, since they are usually given to all subjects of the research at exactly the same time, the data are more accurate.

Besides, Seliger and Shohamy (1989, p.173) states " it is important to stress that before using any questionnaire, it is necessary to try it out and examine whether all the questions on the scales and all the scales provide an actual indication of the variable". Thus, this author affirms that a number of techniques are used to collect data through questionnaire. The Likert scale (Likert, 1932) asks individuals to respond to series of statements by indicating whether they "strongly agree" (SA), "agree" (A), "are undecided" (U) or "Neutral" (N), "disagee" (D), and "strongly agree" (SD) with each statement in the questionnaires. This technique is one of the way makes the data collected more accurate and reliable.

In order to get information to fulfill the aims of the study, two survey questionnaires were designed. The first questionnaire was administered to 10 teachers who had been teaching English to second-year English non-majored students at VATC. They were invited to join the study with the hope of finding out the teaching techniques they were currently applying in teaching English speaking skills and the difficulties they encountered when teaching speaking skills.

In addition, the second questionnaire was designed for students. This section analyzes the data about the attitude and the ideas of the second-year students at SGU towards the learning speaking skill. The analysis revealed the reality and opinion of the students towards the teaching and learning speaking. The results of the analysis were the important evidence to explain the reason why second-year English majored students were having difficulties in speaking.

4. RESULT AND DISCUSSION

Results

Table 1. Descriptive Statistics and Reliability of the Scale (N = 100)

Construct (Items)	**Mean**	**SD**	**Min**	**Max**	**Cronbach's α (overall)**
Awareness of AI Tools (A1–A4)	3.07	1.45	1	5	
Perceived Personalization (B1–B5)	2.96	1.43	1	5	
Engagement & Motivation (C1–C4)	2.91	1.43	1	5	
Perceived Effectiveness (D1–D4)	3.06	1.40	1	5	
Overall Scale (16 items)					0.90

Note. *SD = Standard deviation. Cronbach's alpha is reported only for the overall scale. Item coding: A1–A4 = Awareness, B1–B5 = Personalization, C1–C4 = Motivation, D1–D4 = Effectiveness.*

The internal consistency of the questionnaire was examined using Cronbach's Alpha. The instrument consisted of **16 items** across four constructs: Awareness of AI Tools in Speaking Practice (A1–A4), Perceived Personalization through AI Tools (B1–B5), Engagement and Motivation When Using AI (C1–C4), and Perceived Effectiveness in Speaking Skills (D1–D4).

Cronbach's Alpha was computed both manually and through SPSS to ensure robustness. The hand calculation applied the classical formula

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum Var(X_i)}{Var(Total)}\right)$$

where k=16, ∑Var(Xi) is the sum of individual item variances, and Var(Total) is the variance of the total scale score. Based on the dataset of **100 student respondents**, both the manual computation and the SPSS analysis converged on a Cronbach's Alpha value of **0.90**, indicating **excellent internal consistency** (Nunnally & Bernstein, 1994).

Corrected Item–Total Correlation Review in SPSS discovered that all the items had values above the minimum of 0.30, thus ensuring that all the items contributed meaningfully to their respective constructs. to their constructs. Cronbach's Alpha if Item Deleted statistics also did not reveal any item whose removal would improve reliability significantly. All the 16 items were therefore retained for analysis.

Descriptive statistics of the four constructs are given in Table 1. Total mean scores ranged within a narrow range of 2.91–3.07, which shows that students moderately rated AI tools for speaking instruction. Standard deviations of 1.40 and 1.45 and ranges obtained of 1–5 inform us regarding extreme variability in the responses of students.

A1–A4: Awareness of AI Tools was the most with the highest mean (M = 3.07, SD = 1.45, N = 100). It is to indicate that students demonstrated a moderate level of familiarity with AI-based speaking tools such as ChatGPT, ELSA, or Speak. There was widespread knowledge among some of the participants regarding how AI deals with pronunciation and fluency and others were less aware since the range of scores was high.

Perceived Personalization (B1–B5) was 2.96 (SD = 1.43). Students showed moderate agreement with the idea that AI could that AI could tailor speaking practice based on their level, offer flexible modes of practice, and offer feedback on personal errors. The mean score below 3.0, however, suggests that perceived personalization was not strongly persuasive, however, that perceived personalization was not persuasive to a great degree, perhaps because not all students appreciated customization possibilities to an equal extent.

Engagement and Motivation (C1–C4) was lowest overall (M = 2.91, SD = 1.43). Although some were more motivated, more confident, or more inclined to practice outside class when supported by AI, many respondents selected neutral or disagreement options. This implies motivational effects of AI are not necessarily optimally explained and may be dependent on students' learning style, current digital experience, or classroom integration routines.

Perceived Effectiveness (D1–D4) = 3.06 (SD = 1.40). Students were moderately agreed that AI facilitated the development of speaking skills. Wide scores (Min = 1, Max = 5) show that while there were students who had scored a very high level of development, there were also students who were uncertain about AI's contribution to speaking development. the development of speaking skills.

To sum up, the findings support the consistency and diversity of the scale in measuring students' beliefs about AI-supported speaking instruction. Students were students expressed cautious optimism: they recognized the utility of AI and reasonably knew its potential but less

so for its motivational and personalization capabilities. Response variability reflects the significance of context variables prior experience, instructional design, and support systems in determining attitudes of students towards embracing AI.

5. POLICY RECOMMENDATIONS FOR ENHANCING ACADEMIC SUPPORT SERVICES THROUGH AI INTERGRATION.

Drawing on the findings of this study, which evaluated academic literacy support services at Sai Gon university, the following recommendations are proposed to guide institutions in enhancing their services through structured evaluation, ethical AI integration, and inclusive design. These recommendations are directed toward higher education administrators, academic support leaders, and education policymakers.

1. Establish institutional frameworks for evaluating support services

Require academic support centres to adopt structured, evidence-based evaluation models such as the proposed three-layer framework addressing inclusivity, academic performance, and student satisfaction. The study revealed student feedback alone cannot adequately capture the complexity or effectiveness of academic support services. Implementing a systematic evaluation process enables consistent, comprehensive assessment, even when full feedback data is unavailable.

2. Embed ethical AI into academic support policies

Develop institution-wide policies that ensure the responsible integration of AI technologies such as chat-bots, automated feedback systems, and AI-enhanced writing tools into academic support services. AI can enhance responsiveness, professionalization, and accessibility, particularly for remote and diverse learners. However, integration must prioritize pedagogical integrity, protect student privacy, and incorporate robust ethical safeguards.

3. Enforce standards for visibility, accessibility, and multi-modal design

Require all academic support platforms to meet established accessibility standards, ensuring intuitive navigation, high visibility, and multimedia content designed for diverse learning styles. The study identified barriers to locating and accessing academic support resources, which can limit service equity and effectiveness. Applying user-centred, multi-modal design principles will promote inclusivity and foster greater student engagement.

4. Introduce real-time feedback mechanisms

Embed real-time feedback prompts such as in-page rating buttons and micro-surveys into all digital academic support resources. Immediate, specific feedback provides actionable data for continuous improvement and overcomes the limitations of delayed or under-reported student evaluations.

5. Fund analytics dashboards for usage monitoring

Allocate funding and technical support for academic support units to implement analytic dashboards that track engagement, access patterns, and resource utility. Learning analytic offer a valuable complement to direct feedback, enabling data-informed decision-making and better alignment between services and student needs.

6. Promote cultural and linguistic exclusivity

Require academic support materials to incorporate intercultural communication principles and address the needs of multilingual and Indigenous learners. The study found minimal content addressing cultural exclusivity. Integrating inclusive practices improves accessibility, relevance, and effectiveness for diverse student cohorts.

7. Encourage human and AI hybrid service models

Develop institutional strategies that combine AI-enhanced services with human-led academic support, ensuring each complements the other. Rather than replacing educators, AI can extend the reach and timeliness of academic support. A hybrid model retains human empathy and judgement while increasing service capacity.

5. CONCLUSION

This case study illustrates that integrating AI tools into speaking instruction offers significant potential for enhancing personalized learning experiences, provided that implementation is guided by sound pedagogy and ethical considerations. Teachers recognized AI's capacity to deliver timely feedback, adapt to individual learner needs, and extend support beyond the classroom. However, they also emphasized the irreplaceable role of human judgement, cultural awareness, and empathy in language teaching. The findings suggest that an optimal approach lies in hybrid models where AI complements rather than substitutes teacher-led instruction. For universities and educators, this means investing in professional development, establishing clear ethical guidelines, and aligning AI integration with broader curricular goals. By doing so, institutions can leverage AI to strengthen speaking skill development while safeguarding the quality, and authenticity of the learning process.

References

Billingsley, F., & Gardner, S. (2024). Utilizing generative AI with second-language users and bilingual students. In *Advances in Higher Education and Professional Development* (pp. 109–132). IGI Global. https://doi.org/10.4018/979-8-3693-0487-7.ch005

Cambridge University Press. (2023). *Mindset for IELTS. Level 1: Student's book with digital pack.* Cambridge University Press.

Chen, L. (2024). Research on innovative models of second language teaching in the age of artificial intelligence. *Applied Mathematics and Nonlinear Sciences, 9*(1). https://doi.org/10.2478/amns-2024-0760

Gkountara, D. N., & Prasad, R. (2022). A review of artificial intelligence in foreign language learning. In *2022 25th International Symposium on Wireless Personal Multimedia Communications (WPMC)* (Vol. 18, pp. 134–139). IEEE. https://doi.org/10.1109/wpmc55625.2022.10014767

Korenev, A. A. (2024). Strategies of using artificial intelligence for written corrective feedback in language education. *Moscow University Bulletin. Series 19. Linguistics and Intercultural Communication, 27*(2), 68–77. https://doi.org/10.55959/msu-2074-1588-19-27-2-5

Kovalenko, I., & Baranivska, N. (2024). Integrating artificial intelligence in English language teaching: Exploring the potential and challenges of AI tools in enhancing language learning outcomes and personalized education. *European Socio-Legal & Humanitarian Studies, 1*, 86–95. https://doi.org/10.61345/2734-8873.2024.1.9

Mitra, N., & Banerjee, A. (2022). A study on using AI in promoting English language learning. In *Lecture Notes in Networks and Systems* (pp. 287–297). Springer. https://doi.org/10.1007/978-981-19-4052-1_30

Moulieswaran, N., & Kumar, P. N. S. (2022). Amelioration of Google Assistant – A review of artificial intelligence stimulated second language learning and teaching. *World Journal of English Language, 13*(1), 86. https://doi.org/10.5430/wjel.v13n1p86

Nadhifah, A. S., Syukur, H. N., Haryanto, M. F., Luthfiyyah, R. L., & Rozak, D. R. (2024). Pre-service English teacher perceptions of AI in writing skills. *Journal of World Englishes and Educational Practices, 6*(2), 26–32. https://doi.org/10.32996/jweep.2024.6.2.3

Nazim, M. (2024). Exploring EFL teachers' insights regarding artificial intelligence–driven tools in student-centered writing instructions. *International Journal of English Linguistics, 14*(3), 90. https://doi.org/10.5539/ijel.v14n3p90

Nunan, D. (1999). *Second language teaching & learning.* Heinle & Heinle.

Pokrivčáková, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education, 7*(3), 135–153. https://doi.org/10.2478/jolace-2019-0025

Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching.* Cambridge University Press.

Tseng, W., Tate, T. P., Harnick-Shapiro, B., & Warschauer, M. (2025). Pedagogical approaches to integrating generative AI in the undergraduate writing classroom [Manuscript submitted for publication].

Umar, U. (2024). Advancements in English language teaching: Harnessing the power of artificial intelligence. *FLIP Foreign Language Instruction Probe, 3*(1), 29–42. https://doi.org/10.54213/flip.v3i1.402

CẢI THIỆN KHẢ NĂNG ÁNH XẠ NGỮ NGHĨA OALD-WORDNET BẰNG KỸ THUẬT BIỂU DIỄN VECTOR TRI THỨC TỔNG HỢP

Phan Văn Bá Hải[1], Đỗ Quốc Trí[2], Phan Thị Mỹ Trang[3], Đinh Điền[4]

Tóm tắt: Nhiệm vụ ánh xạ ngữ nghĩa tự động giữa từ điển OALD và WordNet có ý nghĩa quan trọng trong xây dựng tài nguyên từ vựng có cấu trúc phục vụ Xử lý Ngôn ngữ Tự nhiên. Thách thức đặt ra là sự thiếu hụt dữ liệu gán nhãn để huấn luyện mô hình và hạn chế của phương pháp truyền thống khi chỉ dựa vào tương đồng bề mặt định nghĩa, bỏ qua tri thức bổ trợ. Nghiên cứu này đề xuất cách tiếp cận không cần tinh chỉnh mô hình mà dựa trên biểu diễn vector tri thức tổng hợp. Bằng cách nhúng độc lập từng thành phần tri thức (định nghĩa, ví dụ...) rồi tổ hợp trong không gian vector. Kết quả cho thấy phương pháp đạt hiệu quả vượt trội, khẳng định vai trò tích hợp tri thức hệ thống trong điều kiện thiếu dữ liệu.

***Từ khóa**:* WordNet, OALD, tương đồng ngữ nghĩa, Sentence-BERT, tích hợp tri thức

1. GIỚI THIỆU

Các từ điển truyền thống, tiêu biểu như Oxford Advanced Learner's Dictionary (OALD) [3], cung cấp nhiều thông tin hữu ích bao gồm phiên âm, định nghĩa, ví dụ minh họa, một số thông tin về ngữ pháp và ngữ dụng. Tuy nhiên, cấu trúc tổ chức của từ điển truyền thống dựa theo thứ tự bảng chữ cái khiến việc nhận diện và khai thác mối quan hệ ngữ nghĩa trở nên khó khăn. Chẳng hạn, hai mục từ có quan hệ gần gũi về nghĩa như "ba" và "mẹ" lại được đặt ở những vị trí cách xa nhau, hạn chế cho cả người sử dụng từ điển lẫn các hệ thống xử lý tự động trong việc khai thác tri thức ngữ nghĩa.

Để khắc phục hạn chế trên, các mạng ngữ nghĩa (semantic networks) như WordNet [7, 8] đã được xây dựng. Khác với từ điển truyền thống, WordNet tổ chức từ vựng dựa trên các tập đồng nghĩa (synset), trong đó mỗi synset biểu thị một khái niệm riêng biệt. Các synset này được liên kết thông qua mạng lưới quan hệ ngữ nghĩa, bao gồm quan hệ thượng/hạ danh (hypernymy/ hyponymy), quan hệ bộ phận/ tổng thể (meronymy/ holonymy), quan hệ trái nghĩa (antonymy), cùng nhiều quan hệ ngữ nghĩa khác. Nhờ cấu trúc này, WordNet đã trở thành nguồn tài nguyên nền tảng, được ứng dụng rộng rãi trong nhiều tác vụ Xử lý Ngôn ngữ Tự nhiên (NLP) như phân tích văn bản, truy xuất thông tin và phân giải ngữ nghĩa.

Việc ánh xạ mỗi định nghĩa của một mục từ trong OALD tới một synset tương ứng trong WordNet giúp nâng cấp OALD từ danh sách từ vựng thành mạng tri thức có cấu trúc. Quá trình ánh xạ này, về bản chất là một bài toán gán nhãn ngữ nghĩa, tốn kém và phức tạp nếu

[1] Sinh viên Trường Đại học Khoa học Tự nhiên - ĐHQG HCM; Email: phanvanbahai@gmail.com
[2] Sinh viên Trường Đại học Khoa học Tự nhiên - ĐHQG HCM; Email: doquoctri2003@gmail.com
[3] NCS., Trường Đại học Khoa học xã hội và Nhân văn - ĐHQG HCM; ThS., Trường Đại học Công nghệ Sài Gòn; Email: mytrang779@gmail.com
[4] PGS. TS., Trường Đại học Khoa học Tự nhiên - ĐHQG HCM, Trung tâm Ngôn ngữ học tính toán; Email: ddien@fit.hcmus.edu.vn

thực hiện thủ công. Những tiến bộ gần đây trong lĩnh vực NLP, đặc biệt là sự ra đời của các Mô hình Ngôn ngữ Tiền huấn luyện (Pre-trained Language Models - PLMs), đã mở ra khả năng tự động hóa quá trình này. Các mô hình này có khả năng biến các câu định nghĩa thành các vector đặc trưng và đo độ tương đồng ngữ nghĩa giữa chúng. Mặc dù vậy, hiệu quả của phương pháp này thường bị giới hạn trong các trường hợp định nghĩa ngắn, mơ hồ hoặc chứa thông tin chuyên ngành. Nguyên nhân là do các PLMs thiếu hụt ngữ cảnh và tri thức chuyên biệt cần thiết để phân biệt các sắc thái ngữ nghĩa chuyên sâu.

Trước thách thức đó, nghiên cứu này đề xuất phương pháp tiếp cận mới nhằm cải thiện chính xác đáng kể việc ánh xạ ngữ nghĩa giữa OALD và WordNet. Thay vì chỉ dựa vào sự tương đồng định nghĩa, chúng tôi khai thác thêm tri thức bổ sung từ cả hai nguồn dữ liệu. Đóng góp chính của nghiên cứu này bao gồm: (1) phân tích các nguyên nhân gây ra lỗi sai từ các phương pháp dựa trên mô hình ngôn ngữ truyền thống; (2) đề xuất kỹ thuật biểu diễn vector tri thức tổng hợp, cho phép kết hợp một cách linh hoạt các nguồn tri thức đa dạng (định nghĩa, ví dụ…) để tạo ra vector ngữ nghĩa giàu thông tin hơn.

2. NGHIÊN CỨU LIÊN QUAN

Bài toán ánh xạ ngữ nghĩa giữa các tài nguyên từ vựng như OALD và WordNet liên quan mật thiết đến lĩnh vực tương đồng ngữ nghĩa (semantic similarity) [10]. Sự phát triển của các mô hình ngôn ngữ tiền huấn luyện (PLMs) đã tạo ra những bước tiến quan trọng trong lĩnh vực này.

Trong các phương pháp đo tương đồng ngữ nghĩa dựa trên PLMs, hướng tiếp cận phổ biến là sử dụng các mô hình biến thể từ BERT như Sentence-BERT (SBERT) [13] để mã hóa các định nghĩa thành các vector đặc trưng và sau đó tính toán độ tương đồng cosine giữa chúng. Các nghiên cứu gần đây cho thấy việc tinh chỉnh (fine-tuning) trên các tác vụ cụ thể giúp cải thiện đáng kể chất lượng của các vector này. Ví dụ tiêu biểu là công trình của Wang và cộng sự (2022), nhóm tác giả đã đề xuất phương pháp có tên PromptBERT [14], sử dụng kỹ thuật "prompt-tuning" để tinh chỉnh BERT cho tác vụ tương đồng câu, đạt được hiệu quả cao mà không cần thay đổi toàn bộ tham số của mô hình. Các phương pháp này khẳng định tiềm năng của việc điều chỉnh PLMs cho các tác vụ chuyên biệt, tuy nhiên chúng vẫn chủ yếu dựa trên thông tin bề mặt của văn bản. Ở một khía cạnh khác, xử lí nhập nhằng ngữ nghĩa (Word Sense Disambiguation - WSD) [9] đóng vai trò quan trọng trong việc cải thiện độ chính xác khi đo tương đồng ngữ nghĩa. Các phương pháp hiện đại trong lĩnh vực này thường tận dụng chính các câu định nghĩa trong WordNet để thực hiện WSD. Gần đây, các mô hình ngôn ngữ lớn (LLMs) đã cho thấy khả năng vượt trội trong các tác vụ WSD ở dạng zero-shot hoặc few-shot. Chẳng hạn, nghiên cứu của Es-safi và cộng sự (2023) đã khảo sát khả năng của các mô hình như GPT-3 trong việc thực hiện WSD chỉ bằng cách cung cấp ngữ cảnh và danh sách các định nghĩa [1], cho thấy hiệu quả đáng kinh ngạc mà không cần dữ liệu huấn luyện. Dù vậy, các phương pháp này thường chỉ tập trung vào việc xử lý một từ trong ngữ cảnh cụ thể, thay vì ánh xạ toàn bộ từ điển.

Các mô hình xử lý ngôn ngữ tự nhiên tăng cường tri thức (Knowledge-Enhanced NLP) được phát triển nhằm khắc phục hạn chế của việc chỉ dựa trên văn bản thô. Một xu hướng nghiên cứu nổi bật gần đây là tích hợp tri thức có cấu trúc từ các cơ sở tri thức như WordNet, ConceptNet vào quá trình huấn luyện và biểu diễn của mô hình ngôn ngữ. Các phương pháp này không chỉ xem tri thức là dữ liệu đầu vào mà còn tìm cách "nhúng" chúng vào quá trình biểu diễn của mô hình. Zhao và cộng sự (2024) [15] đã đề xuất K-Link, một framework liên kết thực thể trong văn bản với cơ sở tri thức, đồng thời sử dụng cơ chế đặc biệt để mã hóa và tích hợp thông tin về các quan hệ trong cơ sở tri thức đó. Một hướng tiếp cận khác được trình bày bởi Lauscher và cộng sự (2022) là LexFit [5], một phương pháp tinh chỉnh các PLMs nhằm nâng cao khả năng nhận diện các quan hệ ngữ nghĩa từ vựng (đồng nghĩa, trái nghĩa), từ đó cải thiện hiệu suất trên các tác vụ ngữ nghĩa downstream.

Mặc dù đã có nhiều tiến bộ đáng kể, vẫn tồn tại các khoảng trống nghiên cứu quan trọng mà chúng tôi đang hướng tới để giải quyết. Cụ thể, đề tài tập trung vào việc khắc phục hai thách thức trong thực tế. Thứ nhất, bài toán ánh xạ định nghĩa từ điển chịu sự chi phối của tình trạng thiếu hụt dữ liệu huấn luyện. Nhiều phương pháp tiên tiến, đặc biệt là các kỹ thuật tinh chỉnh (fine-tuning) như LexFit [5] và PromptBERT [14], đều yêu cầu một bộ dữ liệu gán nhãn đặc thù cho tác vụ để có thể tối ưu hóa tham số của mô hình. Tuy nhiên, việc xây dựng một bộ dữ liệu vàng được gán nhãn thủ công cho hàng chục ngàn mục từ của từ điển OALD là một công việc tốn kém về thời gian và nguồn lực, khiến cho việc huấn luyện hay tinh chỉnh mô hình đối với tác vụ này trở nên không khả thi. Do đó, cần có một phương pháp hiệu quả có thể hoạt động mà không tinh chỉnh mô hình. Thứ hai, các phương pháp hiện hành thường chưa khai thác triệt để các nguồn tri thức phụ trợ. Các cách tiếp cận phổ biến thường chỉ tập trung vào việc so sánh hình thức cơ bản của câu định nghĩa. Cách làm này bỏ qua nguồn thông tin ngữ nghĩa đáng tin cậy đi kèm với mỗi mục từ, đó là các nguồn tri thức vệ tinh (ancillary knowledge) như: câu ví dụ thực tế, danh sách từ đồng nghĩa/trái nghĩa và hệ thống phân cấp khái niệm.

Xuất phát từ hai khoảng trống này, nghiên cứu của chúng tôi đề xuất phương pháp không tinh chỉnh, không yêu cầu bất kỳ dữ liệu huấn luyện nào, được thiết kế đặc biệt để tận dụng và tổng hợp các nguồn tri thức vệ tinh đa dạng. Thông qua kỹ thuật cộng vector tri thức, chúng tôi trình bày cách tiếp cận linh hoạt và hiệu quả để làm giàu biểu diễn ngữ nghĩa cho mỗi khái niệm, qua đó chứng minh rằng việc khai thác các nguồn tri thức bị bỏ quên này là chìa khóa để cải thiện đáng kể độ chính xác trong bối cảnh dữ liệu hạn chế.

3. PHƯƠNG PHÁP NGHIÊN CỨU

Trong phần này, chúng tôi trình bày chi tiết về phương pháp được sử dụng để giải quyết bài toán ánh xạ ngữ nghĩa giữa OALD và WordNet. Chúng tôi bắt đầu bằng việc phát biểu bài toán, mô tả mô hình nền, sau đó đề xuất các phương pháp nhằm cải thiện hiệu suất và cuối cùng là thiết lập thực nghiệm.

3.1 Phát biểu bài toán

Về mặt hình thức, bài toán ánh xạ ngữ nghĩa giữa từ điển OALD và mạng ngữ nghĩa WordNet có thể được phát biểu như sau: Cho một từ mục tiêu *w*, ta có một tập hợp các định nghĩa của nó từ OALD, ký hiệu là $D_{OALD}^{(w)} = \{d_{OALD1}, d_{OALD2} \ldots, d_{OALDn}\}$. Tương ứng, trong WordNet, từ *w* có thể thuộc về một tập hợp các tập đồng nghĩa (synset), mỗi synset lại có một định nghĩa riêng, tạo thành tập hợp $D_{WN}^{(w)} = \{d_{WN1}, d_{WN2} \ldots, d_{WNm}\}$. Mục tiêu của nghiên cứu là xây dựng một hàm ánh xạ tối ưu $f: D_{OALD}^{(w)} \rightarrow D_{WN}^{(w)} \cup \{null\}$. Hàm này nhận đầu vào là một định nghĩa d_{OALD} và trả về một định nghĩa d_{WN} tương ứng về mặt ngữ nghĩa nhất, hoặc trả về giá trị *null* nếu không tồn tại một cặp ghép phù hợp nào thỏa mãn tiêu chí tương đồng ngữ nghĩa.

Các bước chính của quy trình bao gồm:

(1) Tra cứu từ: Thu thập tất cả các định nghĩa của từ mục tiêu từ cả hai nguồn từ điển OALD và WordNet. Trong WordNet, chỉ tập trung vào các synset chứa từ mục tiêu.

(2) Tính toán độ tương đồng ngữ nghĩa: Mỗi định nghĩa từ hai nguồn được chuyển đổi thành một vector đặc trưng thông qua mô hình mã hóa câu Sentence Embedding - SBERT (cụ thể *all-MiniLM-L6-v2*) [13]. Độ tương đồng giữa các cặp định nghĩa sau đó được đo lường trên một tiêu chí cụ thể (độ tương đồng cosine) [4].

(3) Ghép cặp định nghĩa: Mỗi định nghĩa từ OALD được liên kết với định nghĩa có độ tương đồng cao nhất trong WordNet, kèm theo việc thiết lập một ngưỡng độ tương đồng để loại trừ những cặp định nghĩa không liên quan nhưng có độ tương đồng cao.

3.2 Mô tả mô hình nền

Khi tiến hành khảo sát và đánh giá lại các kết quả từ mô hình nền trên tập dữ liệu kiểm thử, một số hạn chế đã được xác định [11]. Các vấn đề này chủ yếu xuất phát từ sự phụ thuộc của mô hình vào tương đồng từ vựng bề mặt mà thiếu đi khả năng suy luận ngữ nghĩa sâu [12]. Các hạn chế chính bao gồm: (1) Mô hình embedding không chính xác do không được tinh chỉnh trên dữ liệu đặc thù của tác vụ; (2) Sự khác biệt về phạm vi định nghĩa, khi một định nghĩa OALD có thể bao hàm nhiều định nghĩa WordNet; (3) Thông tin chuyên ngành làm giảm hiệu quả ghép cặp; và (4) Thiếu thông tin trong các định nghĩa ngắn khiến việc vector hóa không hiệu quả.

Ví dụ điển hình cho những thách thức này là với từ "abstainer" (xem Bảng 1). Mô hình nền đã thành công trong việc ghép cặp định nghĩa "a person who never drinks alcohol" (OALD) với "a person who refrains from drinking intoxicating beverages" (WordNet) với điểm tương đồng cao là 0.82. Tuy nhiên, mô hình lại gặp khó khăn và cho điểm số thấp khi phân biệt các định nghĩa khác, chẳng hạn như giữa người không uống rượu và người không bỏ phiếu, cho thấy sự yếu kém của mô hình khi ngữ cảnh không đủ rõ ràng và sự tương đồng bề mặt từ vựng không cao.

Bảng 1: Kết quả độ tương đồng giữa các định nghĩa cho từ "abstainer"

Định nghĩa WordNet **Định nghĩa OALD**	i. someone who practices self denial as a spiritual discipline	ii. a person who refrains from drinking intoxicating beverages
1. a person who choose not to vote either in favour or against sth	0.34	0.39
2. a person who never drinks alcohol	0.47	**0.82**

Nhận thấy những hạn chế trong quá trình thực hiện, nghiên cứu này tập trung vào việc tích hợp các thông tin bổ trợ nhằm xây dựng một hàm đo tương đồng, Sim(d_{OALD},d_{WN}), có khả năng phản ánh chính xác mối quan hệ ngữ nghĩa giữa hai định nghĩa, đủ sức nhận diện được sự khác biệt về văn phong, độ dài và sự nhập nhằng ngữ nghĩa.

3.3 Phương pháp đề xuất

3.3.1 Phương pháp Ghép chuỗi Tri thức (String Concatenation)

Hướng tiếp cận đầu tiên nhằm cải thiện độ chính xác của mô hình là bổ sung thông tin cho chuỗi định nghĩa ban đầu bằng cách cung cấp thêm dữ liệu bổ trợ như ví dụ, từ đồng nghĩa, từ trái nghĩa và quan hệ từ vựng để tạo ra một đoạn văn bản đầy đủ và chi tiết hơn về mỗi mục từ. Ý tưởng thực hiện là sử dụng các thông tin bổ sung này để xây dựng những câu mô tả hoàn chỉnh về mỗi mục từ, qua đó làm phong phú thêm các định nghĩa ban đầu. Những câu mô tả này sẽ được đưa vào quá trình **Sentence Embedding** để cải thiện chất lượng mô hình dự đoán. Khi đã thu thập được các thuộc tính và mối quan hệ bổ sung từ WordNet và OALD, bước tiếp theo là kết hợp chúng với định nghĩa của từ hoặc synset sao cho hợp lý và mạch lạc. Cách tiếp cận này nhằm tạo ra những câu mô tả rõ ràng, dễ hiểu và đầy đủ thông tin, qua đó giúp mô hình **Sentence Embedding** mã hóa chính xác ý nghĩa của các mục từ. Quy trình thực hiện cụ thể gồm các bước sau:

(1) Tổ hợp các thuộc tính và mối quan hệ ngữ nghĩa:

- **Tạo câu văn mô tả**: Các thuộc tính bổ sung như thượng danh (*hypernyms)*, hạ danh (*hyponyms)*, đồng nghĩa (*synonyms)*, trái nghĩa (*antonyms)* và ví dụ (*examples)* sẽ được tổ hợp vào trong các câu mô tả của mục từ. Chẳng hạn, đối với từ "dog", định nghĩa có thể là: "a member of the genus Canis, commonly domesticated by humans", sau đó bổ sung thêm các mối quan hệ như *hypernyms* ("canine"), *hyponyms* ("puppy"), và *synonyms* ("hound") vào để làm rõ hơn nghĩa của từ.

- **Duy trì tính mạch lạc của câu**: Khi kết hợp các thuộc tính vào định nghĩa, cần đảm bảo câu văn vẫn phải có tính logic, không bị dư thừa thông tin hoặc rời rạc. Ví dụ, thay vì chỉ đơn giản liệt kê các từ đồng nghĩa hay từ trái nghĩa, nên sử dụng các từ nối để câu tự nhiên hơn, ví dụ: "dog, also known as a hound, is a domesticated animal, which shares common traits with other canines".

(2) Tạo câu mô tả mở rộng: Sử dụng các quan hệ bổ sung như *see also*, *acompare*, *equal* để làm phong phú thêm thông tin. Ví dụ, nếu từ "dog" có liên kết với các từ như "canine" hoặc "puppy", các mối quan hệ này có thể được đưa vào để làm nổi bật sự liên quan giữa các mục từ, đồng thời cũng giúp mô hình hiểu rõ hơn về ngữ cảnh sử dụng từ.

(3) Sử dụng các câu mô tả vào quá trình Sentence Embedding: Sau khi tạo các câu mô tả hoàn chỉnh từ việc kết hợp các mối quan hệ ngữ nghĩa, những câu này sẽ được đưa vào mô hình **Sentence Embedding** để mã hóa và so sánh độ tương đồng giữa các mục từ từ OALD và WordNet.

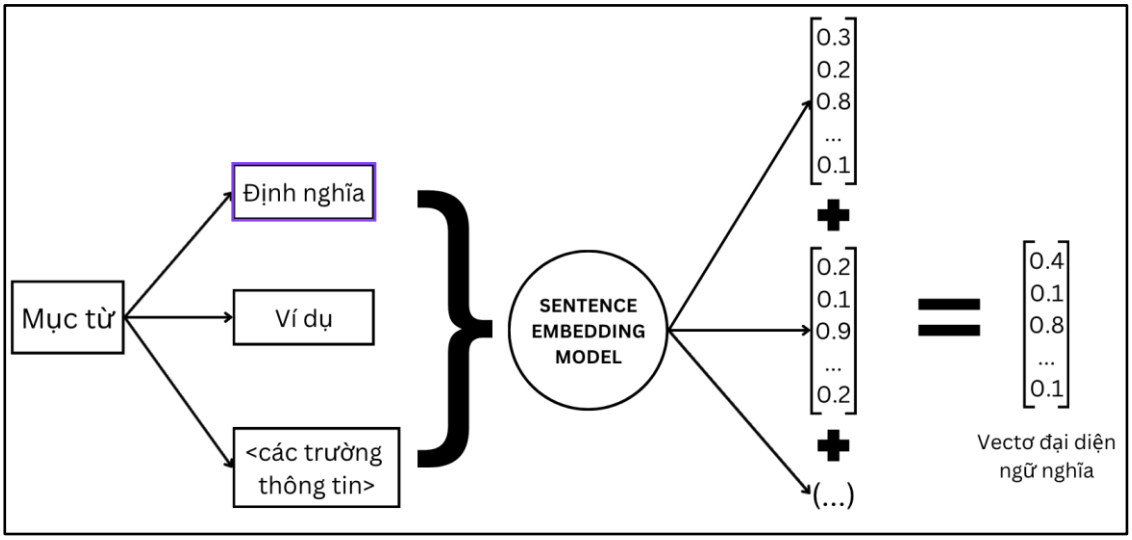
(4) So sánh và đánh giá: Sử dụng phương pháp ghép cặp định nghĩa cũ kết hợp với các câu mô tả mới và đánh giá hiệu quả của mô hình thông qua độ chính xác F1-score.

3.3.2 Phương pháp Cộng Vector Tri thức Tổng hợp (Composite Knowledge Vector Addition)

Bên cạnh phương pháp ghép chuỗi tri thức, chúng tôi đề xuất phương pháp mới nhằm tối ưu hóa việc ghép cặp các tri thức từ OALD và WordNet thông qua kỹ thuật cộng vector. Thay vì hợp nhất thông tin ở cấp độ văn bản, phương pháp này hợp nhất thông tin ở cấp độ vector đặc trưng. Ý tưởng cốt lõi là mỗi thành phần tri thức (định nghĩa, ví dụ...) đóng góp một vector ngữ nghĩa riêng và vector đại diện cuối cùng cho khái niệm là một tổ hợp tuyến tính của các vector thành phần này [6]. Quá trình được thực hiện như sau (xem Hình 1):

(1) Trích xuất và Nhúng Độc lập: Với mỗi khái niệm, chúng tôi xác định một tập hợp các thành phần tri thức $\{k_{\text{định nghĩa}}, k_{\text{ví dụ}}, k_{\text{đồng nghĩa}}, \dots\}$. Mỗi thành phần k_i được xử lý và nhúng độc lập để tạo ra một vectơ tương ứng $v_i = E(k_i)$. Trong trường hợp một thành phần chứa nhiều mẫu (ví dụ, nhiều câu ví dụ), vectơ đại diện của thành phần này được tính bằng cách lấy trung bình các vectơ của từng mẫu.

$$k_{\text{ví dụ}} = \frac{1}{N}\sum_{i=0}^{N} k_{\text{ví dụ }(i)} \qquad (1)$$



Hình 1: Minh họa quá trình "trích xuất và nhúng độc lập"

(2) Tổ hợp Vector: Vector tri thức tổng hợp $v_{tổng\ hợp}$ được tạo ra bằng cách lấy tổng có trọng số của các vector thành phần:

$$v_{tổng\ hợp} = \sum_i^k w_i \,.\, v_i \qquad (2)$$

Trong đó, w_i là trọng số phản ánh tầm quan trọng của thành phần tri thức thứ i. Trong khuôn khổ nghiên cứu này, chúng tôi khảo sát cả trường hợp trọng số đồng nhất ($w_i \in \{0;\ 1\}$) và trường hợp trọng số được tối ưu hóa. Vector $v_{tổng\ hợp}$ này được xem là đại diện cho khái niệm một cách toàn diện.

Bên cạnh đó, chúng tôi sẽ tiến hành khảo sát sự ảnh hưởng của các tri thức từ OALD và WordNet thông qua việc nghiên cứu ảnh hưởng của các loại tri thức khác nhau như định nghĩa, ví dụ, đồng nghĩa, trái nghĩa và các mối quan hệ khác giúp làm rõ sự đóng góp của từng loại tri thức vào độ chính xác tổng thể.

3.4 Thiết lập thực nghiệm

Về dữ liệu, nghiên cứu sử dụng hai nguồn tài nguyên chính là từ điển Oxford Advanced Learner's Dictionary (OALD) [3] và mạng ngữ nghĩa WordNet [7]. Dữ liệu bao gồm các mục từ, định nghĩa và các trường tri thức vệ tinh liên quan (ví dụ: câu ví dụ, từ đồng nghĩa, quan hệ từ vựng) được trích xuất từ hai nguồn này. Trong thử nghiệm này, chúng tôi lựa chọn 100 mục từ ngẫu nhiên từ OALD, đảm bảo phân bố đều theo chữ cái bắt đầu so với tập dữ liệu kiểm thử. Các mục từ này được nhóm nghiên cứu thực hiện ghép cặp thủ công và có cấu trúc như bảng dưới đây:

Bảng 2: Mẫu dữ liệu kiểm thử

OALD	**Mục từ**	"diva"

	Định nghĩa	a famous woman singer, especially an opera singer
	Dịch nghĩa tiếng Việt	nữ danh ca (nhất là ca sĩ opera)
WordNet	**Synset id**	oewn-10491317-n
	Lemma	diva, prima donna
	Định nghĩa	a distinguished female operatic singer; a female operatic star

Về mô hình Embedding, tất cả các thực nghiệm đều sử dụng mô hình *all-MiniLM-L6-v2* làm nền tảng để tạo vector đặc trưng cho văn bản. Đây là một mô hình thuộc họ Sentence-BERT, có sự cân bằng giữa hiệu suất tính toán và chất lượng biểu diễn ngữ nghĩa cho các tác vụ đo tương đồng.

Về độ đo đánh giá, đối với các phương pháp, chúng tôi sử dụng F1-score làm độ đo chính để đánh giá hiệu quả của việc ghép cặp, do độ đo này có khả năng cân bằng giữa Precision và Recall. Đối với mô hình phân lớp, chúng tôi báo cáo đầy đủ các chỉ số bao gồm Accuracy, Precision, Recall và F1-score cho từng lớp để có cái nhìn toàn diện về hiệu suất của mô hình.

4. KẾT QUẢ THỰC NGHIỆM VÀ THẢO LUẬN

Trong giai đoạn đầu, các thực nghiệm với phương pháp ghép chuỗi được tiến hành bằng cách chỉ sử dụng thông tin từ định nghĩa của mỗi mục từ trong OALD và synset tương ứng trong WordNet, không bổ sung thêm bất kỳ quan hệ ngữ nghĩa nào khác. Kết quả cho thấy F1-score đạt được là **80.04%** (xem Bảng 3), cho thấy hiệu quả tương đối cao khi chỉ dựa trên thông tin định nghĩa.

Bảng 3: Kết quả F1-Score của phương pháp cơ bản

PHƯƠNG PHÁP	F1-SCORE
Chỉ sử dụng định nghĩa ở cả OALD và WordNet	80.04%

Khi bổ sung thêm thông tin về các mối quan hệ như *hypernyms*, *hyponyms*, *synonyms*, *antonyms* từ WordNet và các thông tin bổ sung như *examples* từ OALD. Kết quả là F1-score giảm xuống còn **71,43%** (xem Bảng 4). Điều này cho thấy rằng, trái với giả thuyết ban đầu, việc gia tăng số lượng tri thức không phải lúc nào cũng đồng nghĩa với việc nâng cao hiệu quả ghép cặp. Trong một số trường hợp, các quan hệ bổ sung có thể gây nhiễu ngữ nghĩa hoặc làm giảm tính phân biệt giữa các khái niệm. Đáng chú ý, ngay cả khi bổ sung các quan hệ khác từ OALD, như *antonym*, *also*, *equal*, và *synonym*, hiệu quả thu được vẫn không vượt trội so với

phương pháp chỉ sử dụng định nghĩa. Cụ thể, khi các mối quan hệ này được thêm vào quá trình ghép cặp, độ chính xác của mô hình đều giảm so với phương pháp chỉ sử dụng định nghĩa. Kết quả này cho thấy, mặc dù việc bổ sung tri thức có thể có giá trị, nhưng không phải tất cả các loại tri thức đều có giá trị gia tăng và cần có cơ chế chọn lọc để xác định nhóm tri thức thực sự hữu ích cho quá trình ghép cặp.

Bảng 4: Kết quả F1-Score của phương pháp ghép chuỗi trên từng loại tri thức bổ trợ

WordNet / OALD	**example**	**hypernym**	**hyponym**	**antonym**	**also**	**Toàn bộ thuộc tính**
example	68,38%	66,33%	67,02%	67,35%	67,35%	67,35%
synonym	70,41%	70,41%	69,39%	69,38%	69,39%	71,13%
antonym	70,40%	71,43%	70,11%	69,39%	68,37%	71,43%
Toàn bộ thuộc tính	70,41%	71,33%	70,41%	70,48%	69,39%	64,3%

Trong giai đoạn tiếp theo, nhóm nghiên cứu triển khai sang phương pháp **cộng vectơ tri thức tổng hợp**. Kết quả thực nghiệm (Bảng 5) cho thấy rằng việc sử dụng phương pháp cộng vector mang lại sự cải thiện đáng kể về độ chính xác so với phương pháp ghép chuỗi.

Bảng 5: Kết quả F1-Score của phương pháp cộng vectơ trên từng loại tri thức tổng hợp

WordNet / OALD	**example**	**hypernym**	**hyponym**	**antonym**	**also**	**Toàn bộ thuộc tính**
example	82,16%	79,46%	80,81%	80,81%	80,81%	81,54%
synonym	80,81%	78,1%	80,81%	79,46%	79,46%	77,72%
antonym	82,16%	78,1%	80,81%	80,81%	79,46%	80,81%
Toàn bộ thuộc tính	78,1%	74,41%	78,1%	78,1%	77,72%	76,55%

Cụ thể, các tri thức từ OALD liên quan đến ví dụ (examples) và từ trái nghĩa (antonyms) đóng góp nhiều nhất vào hiệu quả mô hình. Đặc biệt, việc kết hợp định nghĩa với các câu ví dụ đã cho phép hệ thống đưa ra dự đoán chính xác hơn so với chỉ dùng định nghĩa thuần túy, với F1-score đạt trên 82%. Ngược lại, đối với WordNet, các tri thức bao gồm ví dụ (examples), quan hệ hạ danh (hyponym), từ trái nghĩa (antonym) và liên từ "also" là những yếu tố then chốt giúp tăng cường hiệu suất.

Nhìn chung, các kết quả trên khẳng định rằng việc biểu diễn mỗi thành phần tri thức bằng một vectơ độc lập và sau đó tổ hợp chúng lại là một chiến lược hiệu quả. Cách tiếp cận này vừa giữ được đặc trưng riêng của từng loại tri thức, vừa tạo ra một vector đại diện tổng hợp giàu thông tin và toàn diện hơn. Qua đó, phương pháp cộng vectơ tri thức tổng hợp cho thấy tiềm năng ứng dụng thực tiễn trong các bài toán ghép cặp dữ liệu giữa hai nguồn từ điển khác nhau.

5. KẾT LUẬN VÀ ĐỀ XUẤT

Nghiên cứu này đã thành công tìm ra một hướng đi cải thiện kết quả cho bài toán ánh xạ ngữ nghĩa tự động giữa OALD và WordNet trong bối cảnh thiếu hụt dữ liệu huấn luyện. Kết quả thực nghiệm cho thấy phương pháp tích hợp tri thức bằng cách ghép nối văn bản thô không hiệu quả về cải thiện độ chính xác, trong khi kỹ thuật biểu diễn vectơ tri thức tổng hợp giúp nâng cao hiệu suất một cách rõ rệt. Phương không tinh chỉnh chứng minh cách thức tổ hợp thông tin trong không gian vector quan trọng hơn việc chỉ đơn thuần bổ sung tri thức. Nghiên cứu cũng chỉ ra rằng một số tri thức vệ tinh như câu ví dụ, từ trái nghĩa và quan hệ phân cấp mang lại hiệu quả cao trong việc biểu diễn thông tin ngữ nghĩa.

Về hướng phát triển, các nghiên cứu trong tương lai có thể tập trung vào việc tự động tối ưu hóa trọng số cho từng loại tri thức để tối đa hóa hiệu suất. Đồng thời, việc khảo sát khả năng của các mô hình ngôn ngữ lớn (LLMs) trong vai trò thẩm định và chọn lọc lại các cặp ghép tiềm năng cũng là một hướng đi đầy hứa hẹn. Cuối cùng, việc mở rộng phương pháp luận này sang các cặp từ điển và ngôn ngữ khác sẽ là một bước tiến quan trọng, góp phần thúc đẩy các ứng dụng xử lý ngôn ngữ tự nhiên đa ngôn ngữ.

TÀI LIỆU THAM KHẢO

[1] Abder-Rahman, S., El-Safi, O., & El-Hanafi, M. (2023). Large language models are state-of-the-art zero-shot word sense disambiguation solvers. *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 2383-2391.

[2] Huang, L., Sun, C., Lin, Y., & Liu, Z. (2019). GlossBERT: BERT for word sense disambiguation with gloss knowledge. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 4004-4014.

[3] Joanna, T. & Đinh Điền. (2014). *Oxford Advanced Learner's Dictionary 8th Edition (with Vietnamese Translation)*. Nhà Xuất bản Trẻ.

[4] Kumar, S. B., Srinivas, K. G., Akhtar, M. S., & Ekbal, A. (2019). Zero-shot word sense disambiguation using sense definition embeddings. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 5748-5753.

[5] Lauscher, A., Vulić, I., Ponti, E. M., Korhonen, A., & Glavaš, G. (2022). LexFit: Lexical fine-tuning of pretrained language models. *Transactions of the Association for Computational Linguistics, 10*, pp. 301-318.

[6] Mancini, G., & Navigli, R. (2020). Sense-aware semantic representations for words and sentences. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 531-543.

[7] McCrae, J. P., Rademaker, A., Bond, F., Rudnicka, E., & Fellbaum, C. (2024). *Open English WordNet*. Truy cập ngày 21 tháng 07 năm 2025, từ https://en-word.net.

[8] Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM, 38*(11), pp. 39-41.

[9] Navigli, R. (2009). Word sense disambiguation: A survey. *ACM Computing Surveys (CSUR), 41*(2), pp. 1-69.

[10] Navigli, R., & Lapata, M. (2010). On the evaluation of semantic similarity on the WordNet taxonomy. *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, pp. 883-891.

[11] Phan Văn Bá Hải, Đỗ Quốc Trí, Phan Thị Mỹ Trang & Đinh Điền. (2024). Gán nhãn ngữ nghĩa tự động cho từ điển OALD Anh-Anh-Việt dựa trên mạng từ WordNet. *Kỷ yếu Hội thảo Quốc gia lần thứ 2 về Ngôn ngữ học Tính toán (VCL2024)*, pp. 370-382.

[12] Pinto, G. B. G. de K., Sousa, P. C. T. de, & Meza, I. V. (2023). How to probe a language model? A layered interpretation of sentence-level semantic features. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 9594-9610.

[13] Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using siamese BERT-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 3982-3992.

[14] Wang, T., Wang, D., Mao, X., Zhang, S., Huang, H., & Li, X. (2022). PromptBERT: Improving BERT with prompts for semantic textual similarity. *Proceedings of the 29th International Conference on Computational Linguistics*, pp. 138-148.

[15] Zhao, K., Chen, J., Liu, Y., Wang, H., & Zhang, W. (2024). K-Link: A knowledge-enhanced framework for entity linking with large language models. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING)*, pp. 535-548.

[15] Zhao, K., Chen, J., Liu, Y., Wang, H., & Zhang, W. (2024). K-Link: A knowledge-enhanced framework for entity linking with large language models. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING)*, pp. 535-548.

THÔNG TIN TÁC GIẢ

(1) Phan Văn Bá Hải: Sinh viên ngành Công nghệ thông tin, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP. HCM; e-mail: phanvanbahai@gmail.com

(2) Đỗ Quốc Trí: Sinh viên ngành Công nghệ thông tin,Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP. HCM; e-mail: dqtri21@clc.fitus.edu.vn

(3) Phan Thị Mỹ Trang: NCS., Trường Đại học Khoa học xã hội và Nhân văn, ĐHQG-HCM, số 10-12 Đinh Tiên Hoàng, Phường Sài Gòn, TP. HCM; ThS., Trường Đại học Công nghệ Sài Gòn, 180 Cao Lỗ, Phường Chánh Hưng, TP. HCM; e-mail: mytrang779@gmail.com

(4) Đinh Điền: PGS. TS., Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP. HCM, e-mail: ddien@fit.hcmus.edu.vn

A CORPUS-BASED STUDY ON THE USE OF EPISTEMIC MODALS IN TED TALKS ABOUT TECHNOLOGY

Tran Quoc Toan[1], Nguyen Ngo Linh Phung[2], Nguyen Thi Tu My[3], Tran Thi Huynh Tran[4]

Abstract

This study investigates the use of epistemic modal verbs in TED Talks about technology through a specialized mini-corpus of 101 talks (approximately 177,000 words) compiled from TED's official platforms. Drawing on Palmer's (2001) Modality Theory, Coates's (1983, 1987) Modal Strength Scale, and discourse-pragmatic perspectives from Holmes (1982) and Hyland (2005), the research examines both the frequency and pragmatic functions of nine modal verbs (*can, could, may, might, must, will, would, should,* and *ought to*). Using Sketch Engine tools for frequency counts, concordance lines, and collocational analysis, the findings reveal three major functional groups: (1) possibility and certainty (e.g., *can, could, may, might*), (2) necessity and prediction (e.g., *must, will, would*), and (3) recommendation and advisability (e.g., *should, ought to*). Results show that *can* and *will* dominate the discourse, highlighting narratives of technological capability and confident projection, while *could* and *might* express speculation and visionary potential. *Should* emerges as the main directive modal for establishing norms, whereas *ought to* is entirely absent, reflecting TED's accessible and audience-friendly rhetorical style. The study concludes that modality in this genre functions not only as a grammatical resource but also as a persuasive tool for balancing certainty, possibility, and ethical guidance. These insights contribute to corpus-based studies of spoken discourse, shed light on rhetorical strategies in science and technology communication, and offer pedagogical implications for teaching modality in English for Specific Purposes (ESP).

Keywords: epistemic modality; modal verbs; TED Talks; corpus linguistics; ESP

1. Introduction

1.1. Background of the project

According to Zhang (2019), "Many expressions in our lives are related to modality. Modality often deals with possibility and necessity. In general, scholars divide modality into two major subcategories, viz., epistemic modality and deontic modality, and modal verbs are the main carriers to show modality." (p. 879). Zhang (2019) also mentioned that "Modal verbs with frequent occurrence are will, can, may, must, should, ought to, would, could, might, need, dare, have to, be able, etc." (p. 879). Modal verbs act as additions of various probability and essentiality levels for the primary verbs in sentences. Whereas, Winiharti (2012) stated that

[1] Học viên cao học Thạc sĩ Ngôn Ngữ Anh tại Khoa Ngoại Ngữ, Đại học Cần Thơ, *email: quoctoan2506@gmail.com*
[2] Học viên cao học Thạc sĩ Ngôn Ngữ Anh tại Khoa Ngoại Ngữ, Đại học Cần Thơ, *email: phungm5624014@gstudent.ctu.edu.vn*
[3] Học viên cao học Thạc sĩ Ngôn Ngữ Anh tại Khoa Ngoại Ngữ, Đại học Cần Thơ, *email: tumynguyen1155@gmail.com*
[4] Học viên cao học Thạc sĩ Ngôn Ngữ Anh tại Khoa Ngoại Ngữ, Đại học Cần Thơ, *email: huynhtrantran58@gmail.com*

"deontic modality indicates obligation and permission, while epistemic modality expresses possibility and prediction" (p. 539). Bashir et al. (2023) wrote that "In academic discourse, epistemic modality is used to signal the level of confidence the speaker or writer has in the knowledge they are presenting and to qualify or hedge their statements to show their awareness of the potential limitations and uncertainties of the knowledge" (pp. 736-737). In discourse related to technology, epistemic modals are vital for speakers to express prospective growth and trends as well as predict everything uncertainly.

TED Talks chosen in the corpus are prominent in the field of technology and engaging to audiences because of the speakers' balance of scientific reliability and accessibility. They present sophisticated ideas and assumptions both attractively and persuasively. These TED talks, hence, are an ideal environment to examine modal verbs applied for rhetorical effect.

1.2. Research questions

The project aims to address the three main questions:

1. How frequently are the epistemic modal verbs might, could, should, and must used in TED Talks on technology, and what patterns emerge in their distribution?

2. In what ways do speakers in TED Talks on technology use epistemic modals to express different degrees of certainty and to serve persuasive communicative purposes?

3. How can the findings on modal verbs be applied to the teaching of these structures in English language education?

1.3. Research objectives

The study focuses on identifying and analyzing the frequency of epistemic modals in the corpus as well as their contextual utilizations and dissemination with the support from Sketch Engine. Besides, the study also traces each speaker's cautious manifestations of possibility and convincing assertions of necessity. In addition, the research also involves determining the pragmatic functions of these modals' roles in preventing, encouraging, and attracting listeners through speakers' employment. Eventually, the research figures out ways to associate the results with English language teaching, stressing that the authentic discourse's insights can assist learners in understanding modality's semantics as well as its interpersonal and rhetorical features.

2. Theoretical Frameworks

The study is conducted based on three complementary theoretical standpoints that give the gradational, semantic, and pragmatic grounds for investigating epistemic modality in TED talks: the Modality Theory of Palmer (2001), the Modal Strength Scale of Coates (1983/1987), and discourse-pragmatic approaches to modality by Holmes (1982) and Hyland (2005).

2.1. Modality Theory

Palmer's *Mood and Modality* (2001) is considered one of the most popular works of modality classifications, in which he defined each modality concept as well as distinguished in detail between epistemic modality and deontic modality.

With epistemic modality, he supposed that "...the speakers express their judgements about the factual status of the proposition" (p. 8). For instance, "*Kate must be at hom*e may be the judgement is based on the observation that she is not in the office" (p. 9), indicating the high level of possibility and probability from evidence.

In contrast, deontic modality "relates to obligation or permission, emanating from an external source..." (p. 9)...such as "rules or the law; typically and frequently the authority is the actual speaker..." (p. 10). For example, in "He can go now" (p. 10), "can" expresses the permission, while in "He can run a mile in five minutes" (p. 10), "can" here indicates ability.

Palmer (2001) also supposed that epistemic modality could be related to the language's interpersonal features because the speakers utilize modals to illustrate their position in line with intentions and their interlocutors. His work provides the crucial semantic theory for separating epistemic modals from deontic ones, ensuring that the analysis exactly concentrates on speakers' sharing knowledge, beliefs, and implications.

2.2. Modal Strength Scale

Coates's study is based on the semantic differences emphasized by Palmer but supplemented by an important aspect by focusing on conceptualizing the level of certainty and strength of modality. In The Semantics of the Modal Auxiliaries (1983), she describes systematically modal meanings in English, while the work “Epistemic modality and spoken discourse” (1987) highlights semantic values of modals in interactive settings. On page 112, she stated that "epistemic modality can be described as concerned with the speaker’s assumptions, or assessment of possibilities, and, in most cases, it indicates the speaker’s confidence or lack of confidence in the truth of the proposition expressed." More importantly, her research also demonstrated that modal verbs are not only basically abstract grammatical signs but also factors used logically in practical communication.

One of Coates's noticeable contributions is that epistemic modal verbs can be categorized by "...their commitment to the truth of the proposition expressed in their utterance" (1987, p. 112). For instance, in the paper (1983), "might" often refers to a weak possibility, whereas "must" primarily presents strong assertions. Depending on the speaker's selection of commitment, specific modals would be applied appropriately.

Coates's theories allow researchers to classify modals in line with their degree and epistemic strength. These hypotheses also help to analyze the way each speaker uses epistemic modal verbs to express their cautions or convince audiences.

2.3. Discourse-Pragmatic Functions of Modality

Whereas Palmer and Coates provide the semantic foundations and strength scale of modal verbs, Holmes (1982) and Hyland (2005) contribute to proving that modality is a pragmatic resource for forming interaction and convincing.

In Holmes's study (1982), epistemic modal verbs mainly function as hedges that soften speakers' judgments. In particular, "might" and "could" are hedges that decrease users' potential mistakes as well as maintain politeness and avoid assertiveness. For example, in the sentence "They might keep them in that backroom" (p. 14), "might" demonstrates "the speaker's uncertainty or doubt about the truth of the proposition..." (p. 14). She also emphasized that stronger modals, including "must" or "should," help to increase commitment and affirmations. For instance, in "Certainly I must be mad to learn English" (p. 25), "must" expresses the assertive meaning of the speaker.

In addition, Hyland (2005) also furthers the theory by arguing that epistemic modality is one of the remarkable factors of the way speakers and writers build their stance (personal stance toward a statement) and their engagement (how they engage audiences in argument). Besides, "Hedges are devices like...might...that indicate the writer's decision to withhold complete commitment to a proposition, allowing information to be presented as an opinion rather than accredited fact" (p. 178). Meanwhile, "must," "should," and "ought" will direct audiences in viewing things or doing an action according to the users.

In TED Talks on technology, the discourse-pragmatic method interprets both the meaning of modal verbs and their applications for rhetorical and interpersonal purposes. These theories are essential to comprehend modality's impacts on TED talks' persuasiveness because modality clarifies the relationships between linguistic selection, interpersonal position, and audience involvement.

3. Corpus Design

This section outlines the design, structure, and preprocessing of the corpus compiled for this study. The corpus consists of transcribed TED Talks centered on the theme of technology. It was carefully constructed to support the investigation of modal verbs in authentic spoken discourse related to innovation and future-oriented topics.

3.1 Rationale for Selecting TED Talks and the Technology Theme

TED Talks were chosen as the primary data source due to their unique position at the intersection of expert communication, accessibility, and global reach. As a platform, TED curates high-quality presentations from leading professionals, researchers, and innovators across disciplines. The talks are scripted or semi-scripted yet delivered in a natural, engaging spoken format, making them ideal for the study of spoken language that is both authentic and polished. Moreover, all TED transcripts are publicly available and standardized in terms of formatting and accessibility, which facilitates reliable corpus compilation and analysis.

The focus on the technology theme stems from the relevance and dynamic nature of this domain in shaping contemporary public discourse. Technology-related topics are

frequently characterized by epistemic uncertainty, predictive language, and expressions of possibility or necessity—features that align directly with the study of modal verbs. In addition, technological discourse often involves future-oriented reasoning, speculative scenarios, and persuasive appeals, all of which lend themselves well to modal analysis from both semantic and pragmatic perspectives. By selecting TED Talks that address diverse aspects of technology—including artificial intelligence, digital innovation, sustainability, and biotech—the study captures a wide linguistic range while maintaining topical cohesion.

3.2 Data Source and Selection Criteria

All texts were sourced from the official TED.com website as well as the official TED Talks YouTube channel (https://www.youtube.com/@TED). These two platforms were chosen for their consistency, accessibility, and credibility in publishing curated talks with verified transcripts and speaker-authored content. By utilizing both platforms, the study ensured a diverse yet standardized dataset of spoken English in expert-to-public communication.

The compiled corpus comprises 101 TED Talks on the topic of technology, with video publication dates ranging from 2010 to 2025. This 15-year span was deliberately chosen to reflect both the evolution of technological discourse and the continuities in modal expression over time. The selected time frame captures major shifts in public discourse surrounding topics such as artificial intelligence, data privacy, climate technology, biotechnology, and digital transformation—making it possible to observe how speakers use modal verbs to express certainty, probability, and obligation in response to evolving technological realities.

To ensure both thematic and formal consistency, only talks that met the following inclusion criteria were selected:

(1) The talk must be originally delivered in English;

(2) The format must be a monologic speech delivered by a single speaker, thereby excluding interviews, panel discussions, or group presentations;

(3) The topic must be technology-related, broadly defined to include any field where innovation, digitalization, or scientific advancement plays a central role.

The chosen timeframe also allowed the researcher to investigate whether changes in public perception of technology—such as increasing uncertainty or urgency—are reflected linguistically through modality. In this sense, the corpus not only captures the lexical and grammatical features of modal usage, but also the temporal and sociocultural dynamics of how technological futures are framed in public discourse.

3.3 Size and Scope

The final corpus comprises 101 TED Talks, with individual transcripts ranging from approximately 600 to 6,600 words in length. The total word count of the corpus is approximately 177,000 words. This variation in length reflects the natural diversity of TED presentations while still maintaining thematic consistency. The resulting dataset is sufficiently

large and balanced to support both quantitative (e.g., frequency, collocational profiling) and qualitative (e.g., discourse-pragmatic) analyses. All selected talks were processed using a standardized procedure to ensure internal consistency and data reliability

3.4 Data Processing and Standardization

Each transcript was downloaded in its original English version and saved in an individual Microsoft Word document, titled according to the name of the talk. Preprocessing involved the removal of extraneous information such as speaker names, timestamps, and unnecessary headings. Only the main body of the spoken content was retained. This step ensured that the corpus maintained its focus on authentic linguistic material while eliminating irrelevant metadata.

A separate metadata spreadsheet was created to log essential information for each talk, including the title, year of release, and video duration. This metadata may support further sub-categorization or filtering for future studies.

3.5 Corpus Characteristics

The resulting corpus serves as a specialized dataset of formal spoken English focused on technology-related themes. It offers a valuable resource for investigating not only the linguistic realization of modal verbs but also their rhetorical functions in public, expert-driven communication. The integration of both thematic specificity and speech-like features makes this corpus particularly well-suited for examining epistemic modality in a real-world discourse context.

4. Methodology

This section outlines the methodological framework adopted in the present study, which is corpus-based and relies on the Sketch Engine platform for data collection, processing, and analysis. The study aims to investigate the use of modal verbs in technology-themed spoken discourse using both quantitative and qualitative corpus linguistic tools.

4.1 Corpus Upload and Setup in Sketch Engine

The first step involved uploading a zip file containing 102 preprocessed TED Talk transcripts to Sketch Engine. All files were in Microsoft Word format (.docx), which is natively supported by the platform. The corpus was labeled under a custom name for easy access and management.

Since Sketch Engine supports automatic text processing, no manual conversion to plain text (.txt) was necessary. Each file retained only the main body of the transcript, with non-linguistic elements such as speaker names, timestamps, and headings removed. This ensured textual uniformity and analytical clarity, both of which are essential in corpus-based research (Biber, Conrad, & Reppen, 1998).

4.2 Wordlist Creation and Frequency Analysis

After the corpus was uploaded, the Wordlist function in Sketch Engine was used to generate a frequency list of all lexical items. This function ranks all words based on their number of occurrences across the corpus, providing a foundation for lexical analysis and comparison.

The researcher specifically searched for the nine modal verbs under investigation: *can, could, may, might, must, will, would, should,* and *ought to*. This step helped quantify the usage patterns of modals in spoken discourse related to technology. According to Biber et al. (1998), frequency data offers important insights into the salience and distribution of grammatical features in specific registers.

4.3 Concordance Analysis of Modal Verbs

To examine how each modal verb was used in context, the Concordance tool was employed. This tool displays multiple instances of a word along with its surrounding context (i.e., co-text), enabling a qualitative analysis of usage patterns.

The researcher extracted concordance lines for each target modal and examined the context to identify epistemic uses, such as expressions of possibility (*might happen*), necessity (*must be*), or inference (*must have*). This aligns with established frameworks in modality studies (Coates, 1983; Palmer, 2001). Selected examples were saved to illustrate the findings in later sections of the paper.

4.4 Collocational Analysis and Word Sketches

The Collocation and Word Sketch tools were then used to identify words that frequently co-occur with the modals under study. Collocates such as verbs and nouns following the modals were analyzed to determine common semantic patterns and discourse functions. For example, the chunk *might happen* or *must change* indicates particular evaluative stances or speaker attitudes.

This process is rooted in the notion that word combinations (collocations) reveal implicit meanings and typical usage conventions (Hunston, 2011). The analysis of lexical bundles also allowed for the identification of common phraseological patterns within the TED Talk genre.

4.5 Analytical Criteria

The interpretation of modal verb usage in this study was guided by two principal analytical dimensions: epistemic strength and pragmatic function. These criteria were selected in accordance with the theoretical foundations outlined in Chapter 2, particularly Palmer's (2001) semantic distinctions between epistemic and deontic modality, Coates's (1983, 1987) gradation of modal strength, and the discourse-pragmatic approaches of Holmes (1982) and Hyland (2005).

The first analytical dimension, epistemic strength, refers to the degree of certainty, probability, or speaker commitment conveyed through the use of modal verbs. Drawing on

Coates's Modal Strength Scale, the study distinguishes between modals that express high certainty—such as *must* and *will*—and those that mark lower levels of certainty or possibility, such as *might*, *could*, and *may*. By evaluating each modal within this semantic continuum, the analysis aims to uncover how TED speakers construct propositions as assured facts, plausible speculations, or tentative possibilities. This gradation helps to clarify how modality is used not only to reflect belief but also to manage risk, evidence, and speaker authority within the discourse (Coates, 1983; Nuyts, 2001).

The second analytical dimension involves the pragmatic functions of modals in discourse. This perspective aligns with Holmes's (1982) and Hyland's (2005) argument that epistemic modality is not purely semantic, but a powerful interactional resource. Accordingly, the study categorizes modal usage into several pragmatic functions, including:

Hedging: using modals such as *might* and *could* to soften claims, reduce speaker imposition, or signal uncertainty;

Boosting: using modals like *must* and *will* to express strong commitment, certainty, or logical necessity;

Recommending: using *should* to offer advice, suggest ideal actions, or promote normative positions;

Urging or compelling: using *must* in a persuasive sense to imply obligation or necessity;

Engaging the audience: using interrogative or inclusive forms (e.g., *how can we...*, *what should we...*) to foster shared inquiry or alignment.

These functional categories were derived inductively through the examination of concordance lines and collocational patterns, using the Sketch Engine platform. They were subsequently mapped back to the theoretical models described earlier to ensure conceptual coherence.

By combining epistemic gradation and pragmatic function, this study takes a dual-level approach to modality—one that captures both the propositional force and interpersonal impact of modal verbs. This approach enables a richer understanding of how modal language in TED Talks operates not only to convey information but also to shape audience perception, construct authority, and deliver persuasive technological narratives. It also sets the stage for linking linguistic forms to communicative goals, thereby bridging descriptive corpus analysis with rhetorical interpretation.

5. Findings

This chapter presents the findings of the corpus analysis, examining the frequency and pragmatic functions of nine key modal verbs. To facilitate a clear and systematic analysis, these verbs have been categorised into three distinct functional groups based on their core semantic properties.

The first group, comprising *"can", "could", "may"*, and *"might"*, is analysed under the theme of *Possibility and Certainty*. These modals are essential for expressing varying degrees of likelihood and are used by speakers to modulate the strength of their claims, from asserting definite capabilities to exploring speculative scenarios.

The second group focuses on "*Necessity and Prediction"*, and includes *"must", "will"*, and *"would"*. This set of modals is crucial for expressing obligation, making firm predictions about future technological developments, and discussing outcomes in hypothetical contexts.

The final group, consisting of *"should"* and *"ought to"*, is examined for its role in expressing *Recommendation and Advisability.* These modals are used to guide the audience, suggest courses of action, and convey a sense of moral or logical obligation, playing a key role in the persuasive rhetoric of the talks.

The following sections (5.1-5.9) will present a detailed analysis of each modal verb within its respective group.

5.1 Analysis of "can"

5.1.1 Frequency Analysis

The quantitative analysis of the collected corpus begins with the modal verb *"can"*. As the data presented in Table 5.1 demonstrates, *"can"* is the most frequently used modal among the four verbs analysed, underscoring its significance within the discourse.

Table 5.1: Frequency of modal verbs *"can"*, *"could"*, *"may"*, and *"might"*

Modal Verb	Raw Frequency	Normalised Frequency (per million words)
can	1,233	5,999.39
could	387	1,883.02
may	162	788.24
might	90	437.91

As shown in Table 5.1, the normalised frequency of *"can"* (5,999.39 instances per million words) is more than three times higher than that of the next most frequent modal, *"could"*. This statistical prominence strongly suggests that articulating capability is a primary rhetorical function in the technology-focused discourse of TED Talks. The frequent use of *"can"* serves to build a narrative of empowerment, effectiveness, and existing potential. *(For the complete categorised list, please refer to Appendix 1)*

5.1.2 Functional Analysis of Concordance Lines

To understand how this high frequency translates into pragmatic use, a qualitative analysis was performed on a representative sample of 150 concordance lines. *(For the complete*

categorised list, please refer to Appendix A). The findings reveal a clear functional hierarchy, dominated by the expression of ability.

The primary function of *"can"*, accounting for approximately 75% of the analysed sample, is to express Ability. Speakers consistently use this modal to make firm assertions about the capabilities of a technology, an algorithm, or of human actors. Example (1) provides a definitive statement of this capability:

(1) ...a human chess player and AI **can** beat not only any human but also any machine. (Appendix A, Line 2)

A second significant function, observed in approximately 20% of instances, is that of Audience Engagement. This is commonly realised through the interrogative structure "How can...?", which is employed to invite the audience into a shared inquiry. Example (2) illustrates this rhetorical use:

(2) And they ask, How **can** AI make our people more fulfilled, more effective, more amplified? (Appendix A, Line 3)

A minor function is to express general Possibility, which accounts for roughly 5% of the sample. This function describes a more general potential for something to occur, as seen in example (3):

(3) And that means that things **can** go wrong in ways that are very difficult... (Appendix A, Line 48)

In summary, the analysis confirms that *"can"* is the principal modal verb used in technology-themed TED Talks to define and assert the tangible capabilities of both technology and humanity, thereby shaping a discourse of progress and potential.

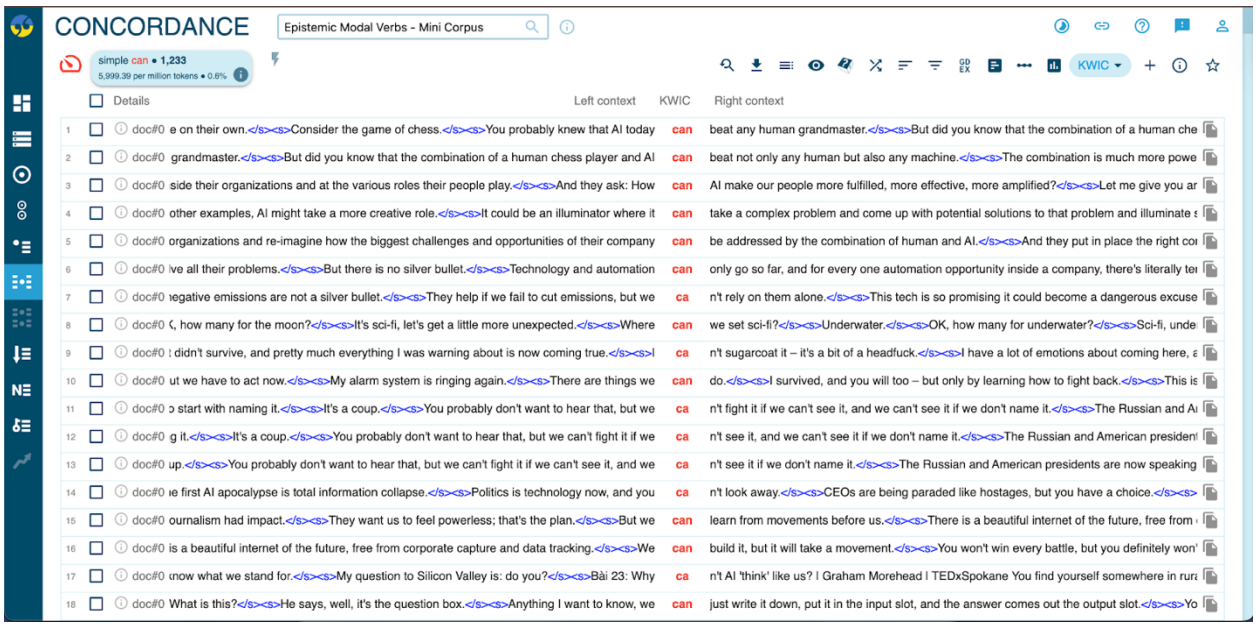


Figure 5.1

Concordance lines for the modal *"can"* in the corpus

5.2 Analysis of "could"

5.2.1 Frequency Analysis

Following *"can"*, the second modal verb analysed is *"could"*. Referring back to Table 5.1, *"could"* is the second most frequent modal in the corpus, appearing 387 times, with a normalised frequency of 1,883.02 instances per million words. Although its frequency is significantly lower than that of *"can"*, it remains a substantial component of the modal landscape in this discourse. This statistical shift suggests a move from the rhetoric of asserting definite, existing capabilities to one of exploring more tentative or hypothetical scenarios.

5.2.2 Functional Analysis of Concordance Lines

The functional analysis of a representative sample of 150 concordance lines for *"could"* reveals a distinct pragmatic profile centred on projection and suggestion. *(For the complete categorised list, please refer to Appendix B).*

The most prominent function, accounting for approximately 60% of the analysed sample, is to express Hypothetical or Future Possibility. Speakers frequently employ *"could"* to project visions of the future and outline potential but uncertain outcomes. Example (4) illustrates this projective use:

(4) ...we **could** still have a vibrant economy. (Appendix B, Line 1)

A second significant function is to express **Tentativeness or Politeness** (approx. 15%). By using *"could"*, speakers can soften their claims, presenting them as suggestions rather than direct assertions. This is demonstrated in example (5):

(5) ...what I'm about to show you **could** be the future of programming. (Appendix B, Line 10)

Other functions include Audience Engagement and reference to Past Ability (approx. 10%), often in the negative form (*"couldn't"*) to create a contrast with present potential, as seen in example (6):

(6) ...they **could**n't even imagine that a computer could be a bicycle... (Appendix B, Line 14)

In essence, *"could"* functions as a crucial tool for shifting the discourse from established facts towards future potential, inviting the audience to imagine and consider possibilities rather than accept certainties.

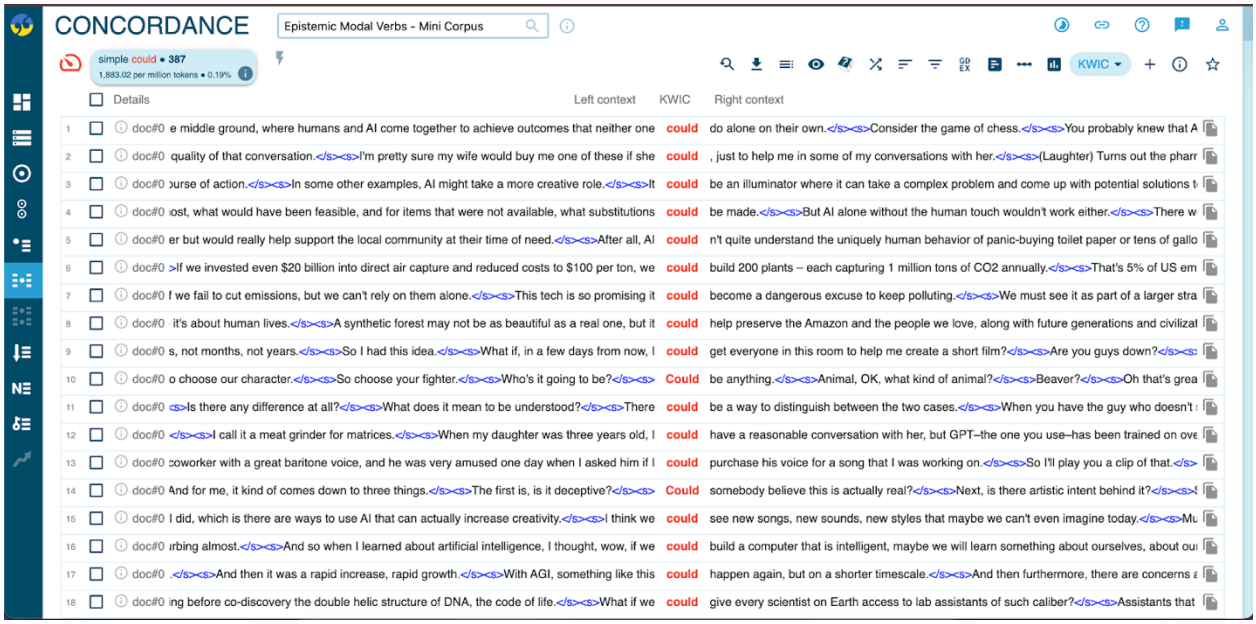


Figure 5.2

Concordance lines for the modal *"could"* illustrating hypothetical uses

5.3 Analysis of "may"

5.3.1 Frequency Analysis

The third modal analysed is *"may"*. As detailed in Table 5.1, *"may"* is the least frequent verb in this study, with a raw frequency of 90 occurrences and a normalised frequency of 437.91 instances per million words. This comparatively low frequency suggests that *"may"* serves a more specialised and context-specific function within the discourse, differing from the broader utility of modals like *"can"* and *"could"*.

5.3.2 Functional Analysis of Concordance Lines

A qualitative analysis of the complete set of 90 concordance lines for *"may"* reveals its role as a modal of formal reasoning and argumentation. *(For the complete categorised list, please refer to Appendix C).*

The most prominent function, accounting for approximately 40% of instances, is for Concession or Contrast. Speakers employ *"may"* rhetorically to acknowledge a potential counter-argument before introducing their main point. This structure (e.g., "While X *may* be true, Y is the case") is a sophisticated argumentative strategy used to build a balanced and persuasive case, as shown in example (7):

(7) While that **may** be true for some, it's not true for all. (Appendix C, Line 4)

The second major function is to express Objective Possibility (approx. 35%). In this capacity, *"may"* is used to present a possibility with a formal, factual, or academic tone, distancing the

speaker from the proposition and framing it as an objective observation. Example (8) demonstrates this use:

(8) The implications of this **may** be profound. (Appendix C, Line 10)

Other uses include expressing Future Possibility (approx. 15%), often describing potential large-scale achievements, as in example (9), and a minor function of formal Audience Engagement.

(9) ...we **may** be able to solve some of the world's most pressing problems. (Appendix C, Line 14)

Overall, the findings indicate that *"may"* is not typically used for simple speculation but functions as a formal modal to facilitate objective reasoning and to rhetorically structure arguments through concession and contrast.

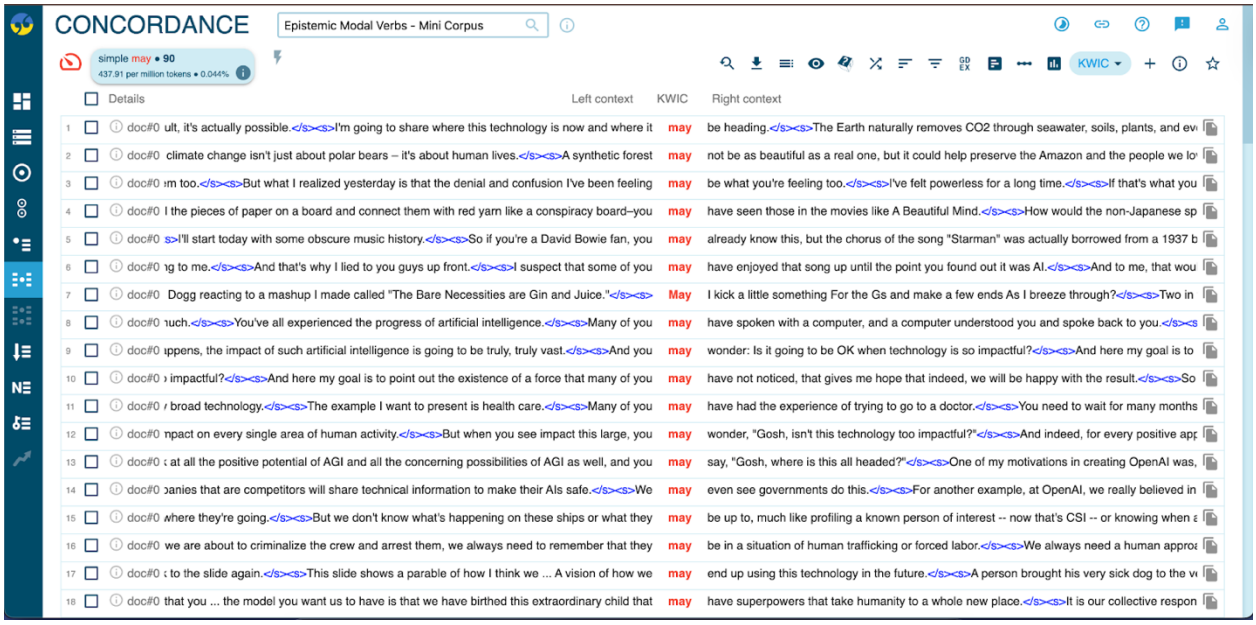


Figure 5.3

Concordance lines for the modal *"may"* illustrating hypothetical uses

5.4 Analysis of "might"

5.4.1 Frequency Analysis

The final modal verb in the "Possibility and Certainty" group is *"might"*. According to the corpus data, *"might"* appears 162 times, with a normalised frequency of 788.24 instances per million words. This places its usage as more frequent than the formal modal *"may"* but less than half as frequent as *"could"*. This suggests that while open speculation is a feature of the

discourse, it is a less common rhetorical strategy than the exploration of more grounded hypothetical scenarios expressed by *"could"*.

5.4.2 Functional Analysis of Concordance Lines

The qualitative analysis of a 150-line sample for *"might"* confirms its role as the primary modal of speculation in the corpus. *(For the complete categorised list, please refer to Appendix D).*

The overwhelmingly dominant function, accounting for approximately 70% of the sample, is to express Speculation and Possibility. Speakers use *"might"* to explore ideas, theories, and potential futures with a low degree of certainty. This fosters a tone of open-ended inquiry and intellectual exploration, allowing for the presentation of ideas that are not yet proven. Example (10) illustrates this core function:

(10) This technology **might** be the key to unlocking new discoveries. (Appendix D, Line 7)

A significant secondary function is to offer a Suggestion or Recommendation in a non-imposing manner (approx. 15%). The use of *"might"* softens the suggestion, framing it as a collaborative thought rather than a direct command. This is often realised in the construction "We might want to...", as seen in example (11):

(11) We **might** want to consider a different approach. (Appendix D, Line 6)

Other minor functions include its use in Past Hypothetical contexts, often for contrast (approx. 5%), as shown in example (12), and for Audience Engagement through speculative questions.

(12) ...different than what you **might** have expected. (Appendix D, Line 3)

In conclusion, *"might"* is the principal tool for speakers in the corpus to engage in speculation. It allows them to explore the frontiers of knowledge and technology by discussing possibilities without the need for strong epistemic commitment.

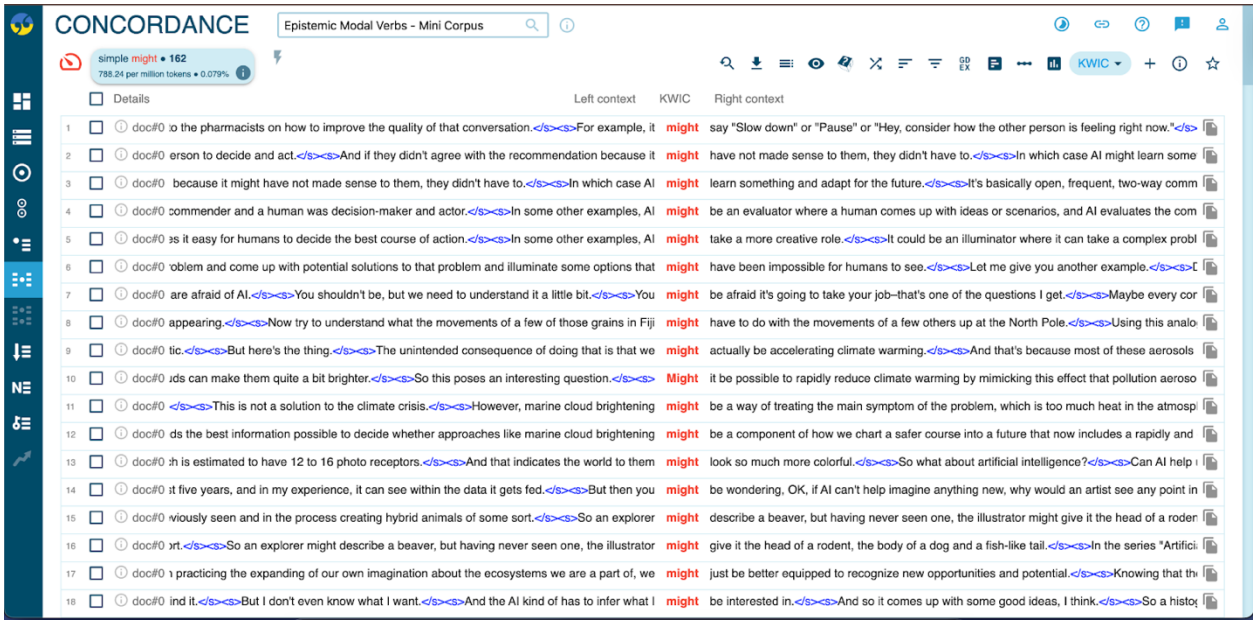


Figure 5.4

Concordance lines for the modal "might" illustrating hypothetical uses

5.5. Analysis of “will”

This analysis explores the multifaceted functions of the modal verb “*will*” within the context of TED Talks focused on technology. Leveraging a specialized corpus and a combination of quantitative and qualitative methods, this study aims to move beyond the traditional grammatical definition of *“will”* as a simple future tense marker.

5.5.1. Frequency Analysis

	Word	Frequency ? ↓	Frequency Per Million ? ↓	
35	if	676	3,289.20	•••
36	or	665	3,235.68	•••
37	at	659	3,206.48	•••
38	how	650	3,162.69	•••
39	from	623	3,031.32	•••
40	more	619	3,011.86	•••
41	there	615	2,992.39	•••
42	these	613	2,982.66	•••
43	just	611	2,972.93	•••
44	now	600	2,919.41	•••
45	an	600	2,919.41	•••
46	one	590	2,870.75	•••
47	my	582	2,831.83	•••
48	us	557	2,710.19	•••
49	people	515	2,505.83	•••
50	will	510	2,481.50	•••

Figure 5.5.

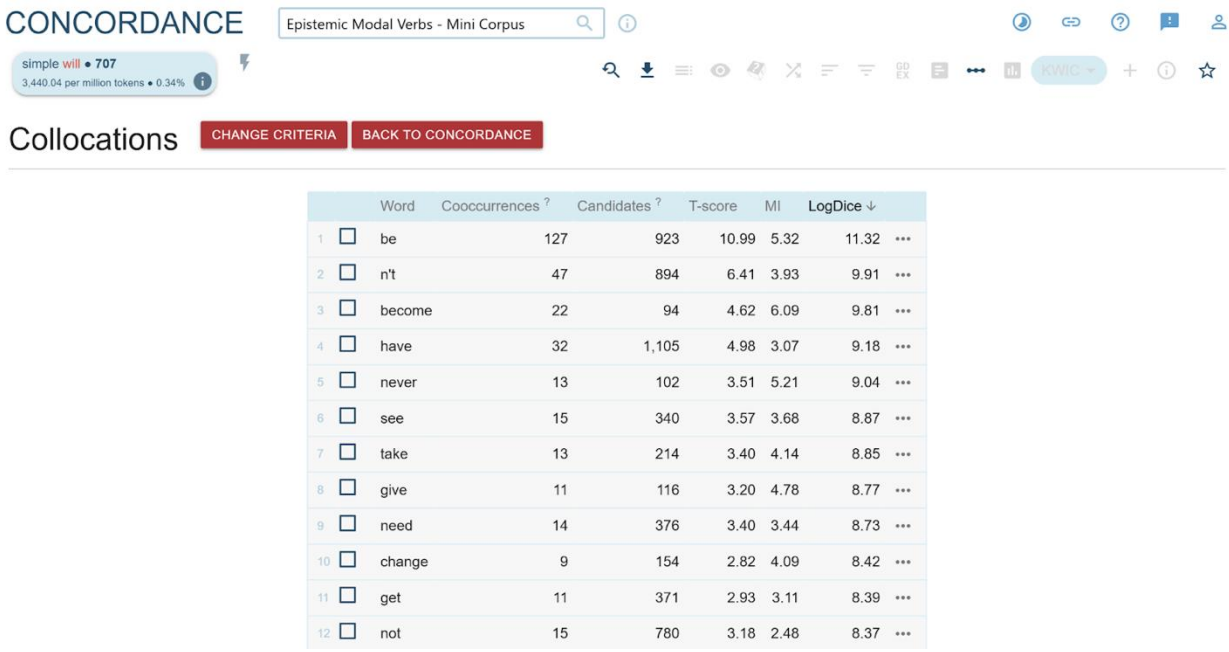
Frequency of modal verb "will"

The data shows that "*will*" has an absolute frequency of 510 occurrences and a relative frequency of 2,481.50 per million words. This positions "*will*" as the 50th most frequent word in your corpus. The high frequency of a modal verb like "*will*" is a significant finding. It suggests that speakers in these TED Talks are not just discussing past and present facts but are actively projecting into the future. The frequent use of "*will*" indicates a strong rhetorical strategy of making confident assertions and predictions about what is to come, which is a key element of persuasive discourse in the genre of technology talks. This aligns with the idea of epistemic modality, where speakers express a high degree of certainty about their claims, as opposed to simply offering possibilities or obligations. The prominence of "*will*" in the top 50

words underscores its central role in shaping the persuasive and forward-looking nature of the discourse.

5.5.2. Functional Analysis of Concordance Lines

The combination of collocation and concordance data demonstrates that "will"primarily serves a high-level epistemic function, conveying certainty about future events.



	Word	Cooccurrences	Candidates	T-score	MI	LogDice
1	be	127	923	10.99	5.32	11.32
2	n't	47	894	6.41	3.93	9.91
3	become	22	94	4.62	6.09	9.81
4	have	32	1,105	4.98	3.07	9.18
5	never	13	102	3.51	5.21	9.04
6	see	15	340	3.57	3.68	8.87
7	take	13	214	3.40	4.14	8.85
8	give	11	116	3.20	4.78	8.77
9	need	14	376	3.40	3.44	8.73
10	change	9	154	2.82	4.09	8.42
11	get	11	371	2.93	3.11	8.39
12	not	15	780	3.18	2.48	8.37

Figure 5.6

Collocations of "will" in the corpus

The highest T-score for *"will"*is with "be" (8.46) and "not" (7.13), indicating a strong and reliable association with future states and negation. This suggests that speakers use *"will"*to present a future outlook as definitive and certain, rather than as a mere possibility. This is a critical finding, as it positions *"will"*not as a verb of prediction but of confident assertion.

The high T-score for "be" indicates it is one of the most reliable and frequent collocates of "will" in your corpus. The concordance data provides the specific context for this link, showing that "will be" is used to predict future states or possibilities. For example,

"…that artificial intelligence, if it worked, would be incredibly impactful. Now, it wasn't at all obvious that it ***would be*** *possible to make progress in artificial intelligence, but if it were possible to make progress in artificial…"* (Appendix E, line 25)

The phrase

"…it feels like we're entering a historic period right now where we as a world are going to define a technology that **will be** so important for our society going forward. And I believe that we can manage this for good. So today, I want to show you…" (Appendix E, line 66)

also shows how “will be” can describe a transformation or outcome. This strong use of “will be” demonstrates that the speakers are expressing a high degree of certainty about future states, capabilities, and changes. This aligns with Coates's Modal Strength Scale (1983), where "will" is positioned at the top, signifying an affirmation rather than an ambiguous possibility. This usage is an effective boosting tool, helping speakers enhance the persuasive power of their predictions about the impact of technology.

Besides, the use of “will”in the corpus is not merely grammatical but serves significant pragmatic functions to persuade and engage the audience.

The high T-score of “not” (7.13) reveals that “will not” is a key rhetorical tool. Phrases like

“…markets help get plants built. And each plant provides data, experience, and cost reductions. But clever ideas alone **won't** be enough. We need regulation, subsidies, carbon taxes. A few of us would pay more, but we need solutions affordable…” (Appendix E, line 7)

or

“...captured CO2 into products – fuels, plastics, gravel – which creates useful markets. However, these markets alone **won't** solve climate change. Still, these markets help get plants built. And each plant provides data, experience, and…” (Appendix E, line 6)

are used to signal what the future will not hold with certainty. This nuanced form of hedging (Hyland, 1998) allows the speaker to project a realistic and balanced perspective, which in turn builds audience trust. This is a particularly sophisticated use, as it demonstrates the speaker's awareness of potential limitations, lending greater credibility to their more positive claims.

Furthermore, the collocation data also revealed strong links with action and change-oriented verbs like “become” (T-score 5.20), “see” (T-score 3.42), and “take” (T-score 3.23). The concordance provides specific pragmatic functions for these co-occurrences. For example, "will" is used to describe future transformations, as seen in the phrases

“What happened? We all experienced what it's like to talk to a computer and to be understood. The idea that computers **will become** really intelligent and eventually more intelligent than us is becoming widespread. It used to be a niche idea…” (Appendix E, line 41)

and

“...real name, use encrypted apps, resist harmful norms. In future, allowing your child's data to be harvested from birth **will be seen** as child abuse. Privacy is power, and we have more of it than we think.” (Appendix E, line 11)

Additionally, “will” is used to describe human or technological actions, as shown in

“The combination is much more powerful than the sum of its parts. In a perfect combination, AI **will do** what it does best, which is dealing with massive amounts of data and solving complex problems.” (Appendix E, line 1)

This example shows that speakers use “will” to paint a detailed and dynamic picture of the future. It is not just a general prediction but a narrative about how humans and technology will act and change the world. This usage is highly aligned with the Discourse-Pragmatic Approach, as using “will” to predict future actions serves to engage the audience and persuade them of the speaker's vision.

Finally, the co-occurrence of “will” with first-person plural (we) and second-person (you) pronouns serves as a vital engagement marker. For instance, the phrase

“...billions of hours of clinical experience, and it will be always available and extremely cheap. When this happens, **we will** look back at today's health care similarly to how we look at 16th century dentistry” (Appendix E, line 33)

transforms a speaker's prediction into a shared experience, inviting the audience to join a collective vision of the future. This shared ownership can significantly increase the persuasive power of the message. This function aligns with the collaborative spirit often found in public discourse on topics of societal importance.

In conclusion, the corpus-based analysis of the modal verb “will” in TED Talks on technology reveals that its function extends far beyond a simple marker of future tense. The high-frequency collocations with “be” and “not” confirm its primary role in expressing a high degree of certainty about future states and their negation. This positions “will” as a confident assertion, a finding that aligns with Coates's Modal Strength Scale where it holds a top-tier position for its epistemic function. The data also demonstrates that “will” is a powerful rhetorical tool serving crucial pragmatic purposes. It acts as a booster, amplifying claims and enhancing credibility by presenting future possibilities as certainties. Conversely, its use with negation serves as a form of subtle hedging, allowing speakers to acknowledge limitations and build audience trust. Finally, the co-occurrence with action-oriented verbs and pronouns like “we” and “you” transforms the speaker's vision into a dynamic, shared narrative, effectively engaging the audience and making the persuasive message more impactful. This multifaceted use of “will” confirms its central role in shaping a compelling discourse on technology.

5.6. Analysis of “would”

5.6.1. Frequency Analysis

	Word	Frequency ? ↓	Frequency Per Million ? ↓	
51	by	510	2,481.50	•••
52	your	488	2,374.45	•••
53	their	473	2,301.47	•••
54	them	455	2,213.89	•••
55	think	453	2,204.15	•••
56	going	438	2,131.17	•••
57	when	431	2,097.11	•••
58	because	426	2,072.78	•••
59	would	425	2,067.92	•••
60	know	420	2,043.59	•••
61	technology	418	2,033.86	•••
62	me	406	1,975.47	•••
63	could	387	1,883.02	•••
64	time	378	1,839.23	•••
65	world	377	1,834.36	•••
66	need	376	1,829.50	•••
67	get	372	1,810.03	•••

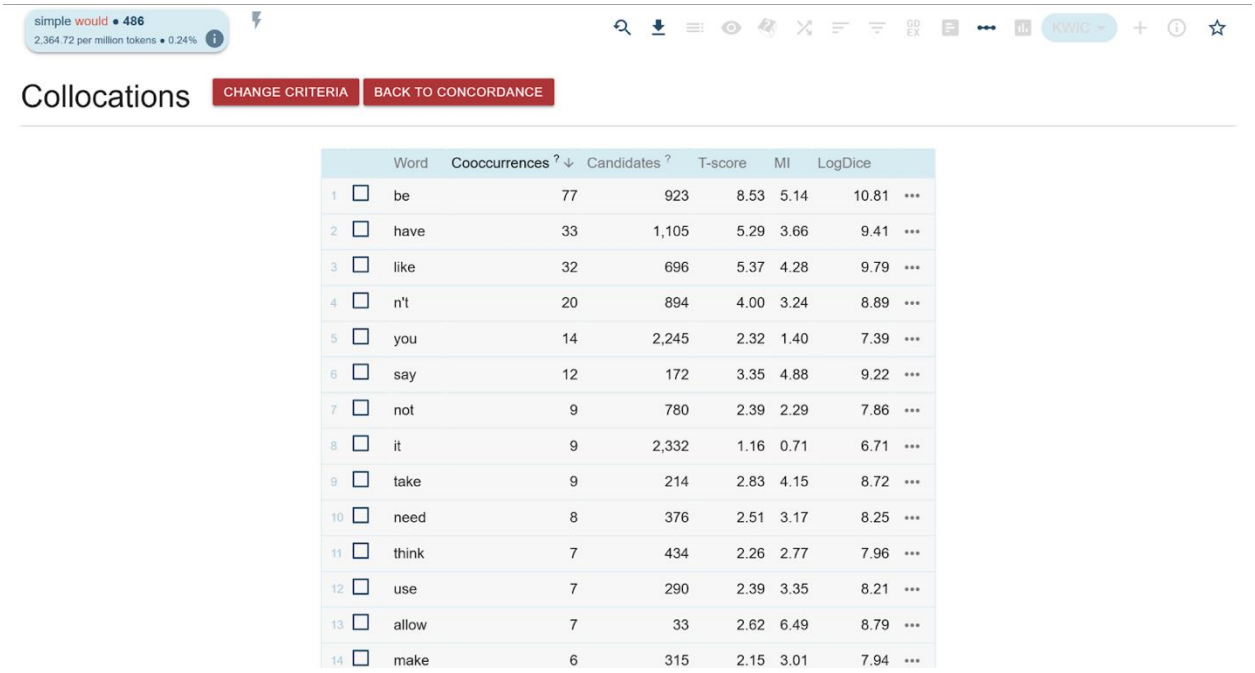
Figure 5.7.

Frequency of the modal verb "*would*"

The data shows that *"would"* has an absolute frequency of 425 and a relative frequency of 2,067.92 per million words, making it the 59th most frequent word in your corpus. This high frequency is notable for a modal verb that typically signals hypotheticality and conditional situations. Unlike *"will,"* which expresses high certainty about the future, the frequent use of *"would"* suggests that speakers in these talks often engage in nuanced discourse, exploring possibilities, comparing scenarios, and making polite suggestions. This strategic use of *"would"* allows speakers to present their ideas without making definitive claims, which is a key form of academic hedging that helps build credibility and rapport with the audience.

5.6.2. Functional Analysis of Concordance Lines

By examining both quantitative and qualitative data from a dedicated corpus, the findings reveal that *"would"*, unlike the assertive *"will"*, functions primarily to introduce a degree of hypotheticality, politeness, and conjecture. Its use serves as a powerful tool for academic hedging and audience engagement. The analysis is structured around the types of verbs that follow *"would"*, revealing distinct rhetorical patterns that are essential for crafting nuanced and persuasive discourse.



	Word	Cooccurrences	Candidates	T-score	MI	LogDice
1	be	77	923	8.53	5.14	10.81
2	have	33	1,105	5.29	3.66	9.41
3	like	32	696	5.37	4.28	9.79
4	n't	20	894	4.00	3.24	8.89
5	you	14	2,245	2.32	1.40	7.39
6	say	12	172	3.35	4.88	9.22
7	not	9	780	2.39	2.29	7.86
8	it	9	2,332	1.16	0.71	6.71
9	take	9	214	2.83	4.15	8.72
10	need	8	376	2.51	3.17	8.25
11	think	7	434	2.26	2.77	7.96
12	use	7	290	2.39	3.35	8.21
13	allow	7	33	2.62	6.49	8.79
14	make	6	315	2.15	3.01	7.94

Figure 5.8.

Collocations of "*would*" in the corpus

The analysis of the corpus reveals that *"would"* is most frequently paired with verbs of state, with "be" having the highest T-score of 8.53. This quantitative finding is central to understanding its primary function: constructing hypothetical or conditional states rather than making definitive predictions. For example, a speaker might use a phrase like

"...We also need major investment in R&D. In 1966, the US spent 0.5% of GDP on the Apollo program. That ***would be*** *$100 billion today. If we invested even $20 billion into direct air capture and reduced costs to $100 per ton, we could…"* (Appendix F, line 9)

to frame a comparison within a hypothetical scenario, or

"...will have two employees: a man and a dog. You can do so much with AI; why have people do all the jobs, right? And the man's job ***would be*** *to feed the dog. The dog's job is to keep the man away from the computer. GPT is not ready to take your job…"* (Appendix F, line 33)

to describe a state in a humorous anecdote. This usage places " at a lower position of certainty on Coates's Modal Strength Scale compared to *"will"*. Instead of asserting what will happen, the speaker uses *"would"* to explore what could or should happen under specific, often imagined, circumstances. This is a key function in academic and persuasive discourse, allowing speakers to present ideas and possibilities without committing to their absolute truth.

The corpus also shows that *"would"*is frequently paired with verbs of cognition and communication, such as "like" (T-score 5.37) and "say" (T-score 3.35). This combination is a crucial tool for both hedging and audience engagement, as described by Hyland (1998). Rather than issuing a direct command, speakers use phrases like

"...It is fitting that the seat of intelligence in artificial intelligence is an artificial brain. Here, ***I'd*** *like to take a digression and tell you about how I got into AI. There were three forces that pulled me into it. The first one…"* (Appendix F, line 34)

to make a polite and considered request. This use of *"would"* softens the claim and maintains a collaborative tone with the audience. Similarly, when a speaker recounts a past anecdote and says

"So first of all, people asked, "What are you into?" ***I'd*** *say artificial intelligence. And they'd say, "What's that?" So no one was really aware of it. I joined, in 1962, 1956 was…"* (Appendix F, line 35)

to describe a habitual action, they are framing a personal experience in a less assertive way. This pragmatic function of *"would"* allows speakers to manage their own perceived authority, presenting themselves as considerate and approachable, which helps maintain a positive relationship with their audience - a core element of effective persuasive discourse.

In addition to states and cognition, *"would"* is also paired with a variety of action and process verbs, including "take" (T-score 2.83) and "happen" (T-score 2.38). In these instances, *"would"* describes actions that are contingent on a specific condition. For example, a speaker might say,

"...Producing ideas and discussing them is core to who I am, so I ***would*** *never have asked GPT to come up with the presentation that I'm giving you right now…"* (Appendix F, line 18)

to explain a past decision by presenting a counterfactual. This usage is essential for creating nuanced arguments, as it allows speakers to compare a hypothetical action with their actual

choice. The rhetorical strategy of using *"would"* in these contexts enables speakers to articulate their values and reasoning without adopting a prescriptive or overly certain tone.

The analysis confirms that *"would"* in this corpus is a sophisticated modal verb with distinct epistemic and pragmatic functions. Its primary role is to introduce hypotheticality and conditionality, positioning it at a lower point on the certainty scale than *"will"*. The specific verbs it collocates with reveal its nuanced rhetorical power: with state verbs, it constructs plausible conditional scenarios; with verbs of cognition and communication, it softens claims and fosters a polite, collaborative tone; and with verbs of action, it enables speakers to explore counterfactuals and justify their choices. These findings support and extend the established frameworks of Coates and Hyland by providing corpus-based evidence from the influential genre of TED Talks. For English language education, these insights are invaluable. Instead of teaching *"would"* merely as the past tense of *"will"* or a general polite marker, educators can use authentic corpus examples to demonstrate its strategic use in managing certainty, building rapport, and crafting persuasive arguments. This approach can better equip learners to communicate effectively and with greater nuance in complex and influential discourse contexts.

5.7. Analysis of "must"

5.7.1. Frequency Analysis

	Word	Frequency ? ↓	Frequency Per Million ? ↓	
326	images	59	287.08	•••
327	impact	59	287.08	•••
328	pretty	59	287.08	•••
329	everyone	58	282.21	•••
330	cost	58	282.21	•••
331	control	58	282.21	•••
332	earth	58	282.21	•••
333	single	58	282.21	•••
334	must	58	282.21	•••

Figure 5.9.

Frequency of the modal verb "*must*"

The data shows that *"must"* has an absolute frequency of 58 and a relative frequency of 282.21 per million words, making it the 334th most frequent word in your corpus. While its frequency is significantly lower than *"will"* (510) and *"would"* (425), it holds a crucial rhetorical role. Its usage indicates that speakers are not simply making predictions but are drawing confident and logical conclusions based on the evidence presented. The use of *"must"* serves to present a point as a necessary consequence or an unavoidable truth, acting as a powerful persuasive tool that lends a sense of gravity and urgency to the speaker's claims.

5.7.2. Functional Analysis of Concordance Lines

This analysis provides deep, academic insight into how the modal verb *"must"* is used in TED Talks about technology. The findings reveal that *"must"* serves a high-level epistemic function, conveying confident inference, and is a key tool for persuasion and audience engagement.

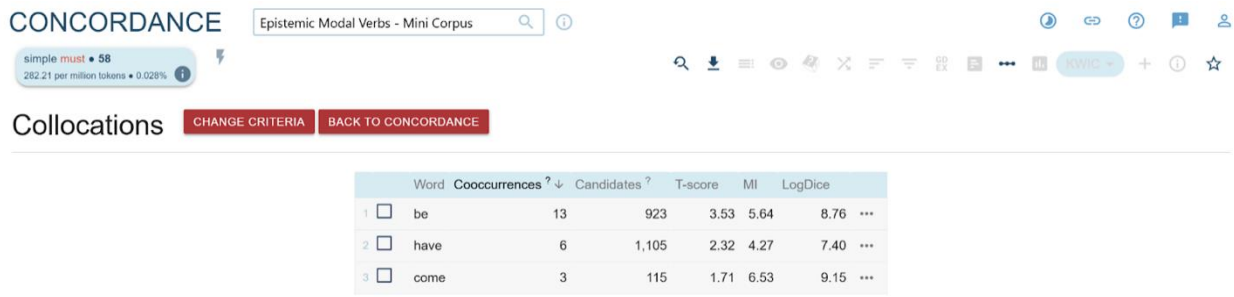


Figure 5.10.

Collocations of "must" in the corpus

The data indicates that *"must"*, despite its common deontic meaning of obligation, is frequently used in an epistemic sense to convey confident conclusions based on available evidence. This is quantitatively supported by the top collocate for *"must"* being "be", with a T-score of 3.53. This finding suggests that speakers use *"must"* to express a logical necessity or a strong inference about a state of being. The concordance lines reinforce this with clear examples. For instance, the statement

"...large volumes, contactors are shallow but wide – an energy-efficient compromise. Once captured, the material ***must be*** *recycled. The scale is so large that single-use materials are not feasible. Recycling requires intense heat to…"* (Appendix G, line 2)

is not a command but a logical conclusion drawn from the physical constraints of the problem. Similarly, the line

"…in liquid, or using solid materials. Either way, the systems are similar. An air contactor is a key component. It ***must be*** *very wide to expose enough surface area because we're capturing just 400 molecules out of a million. In the…"* (Appendix G, line 1)

Uses *"must"* to present a necessary structural requirement as a reasoned inference. This aligns with Coates's (1983) Modal Strength Scale, where *"must"* is positioned to express a high degree of certainty based on a speaker's reasoned deduction rather than a simple fact.

The use of *"must"* in the corpus extends beyond grammar to serve crucial pragmatic functions. It acts as a powerful persuasive tool, or booster, that strengthens claims and motivates the audience. The collocation data reveals a strong association with the verb "be" (T-score 3.53) and "have" (T-score 2.32), suggesting that speakers use these combinations to frame certain conditions or qualities as essential and unavoidable.

This is evident when speakers make confident, reasoned deductions. For instance, the statement

"…At the center of this dream car is an AI driver that ***must be*** *ten times better than any human–meaning less than one fatality per one billion miles.This is a very hard problem, but…"* (Appendix G, line 56)

presents a logical necessity that for the "dream car" to be viable, the AI driver must meet a specific, non-negotiable performance standard. This is not a command but a powerful, evidence-based conclusion.. Similarly, the use of "must have" to express a necessary quality or characteristic is a potent rhetorical strategy. In the example,

"...and it must be in the Global South, where 90 percent of future urbanization is going to take place. And it's a plan that ***must*** *have us move beyond our narrow self-national interest. We must come to see the world's cities as international assets…"* (Appendix G, line 27)

the speaker is not just advising, but declaring a non-negotiable requirement for the plan's success. This frames the call to action as an objective necessity, guiding the audience's thinking by suggesting that these qualities are objectively essential for achieving a shared purpose. By leveraging these combinations, speakers underscore the gravity and logical necessity of their points, which is a key element of the Discourse-Pragmatic Approach (Hyland, 1998) and helps to engage the audience and build a sense of shared purpose and urgency.

The analysis of *"must"* confirms its strategic use as a high-certainty modal in TED Talks on technology. Rather than functioning as a simple command, *"must"* is predominantly used in its epistemic sense to convey confident, reasoned deductions. This is demonstrated by its strong collocation with "be" and "have" , which speakers leverage to present logical necessities as undeniable facts. This rhetorical strategy effectively acts as a booster, strengthening claims and motivating the audience. By framing critical requirements as objective truths, speakers can guide the audience's thinking and build a sense of shared urgency without being overly prescriptive. This nuanced use of *"must"* is a powerful tool for persuasion and engagement, which aligns with key principles of the Discourse-Pragmatic Approach.

5.8 Analysis of "should"

5.8.1 Frequency Analysis

The analysis of the "Recommendation and Advisability" group begins with *"should"*. As the data from the corpus shows, *"should"* appears a total of **453** times, with a normalised frequency of **2,204.62** instances per million words.

This frequency is highly significant. When compared to the modals of possibility (as detailed in Table 5.1), the usage of *"should"* is notably higher than that of *"could"* (1,883.02 ipm). This indicates that expressing direct recommendations and outlining ideal scenarios is a more common rhetorical strategy in this discourse than exploring purely hypothetical possibilities. This finding underscores the persuasive and directive nature of the genre, highlighting that a key goal for speakers is not just to explore what technology *can* or *could* do, but to actively guide the audience's perspective on what it *should* do.

	Word	Frequency ? ↓	Frequency Per Million ? ↓	
1	should	162	788.24	•••

Figure 5.11

The frequency of the word “should”

5.8.2 Functional Analysis of Concordance Lines

A qualitative analysis of a representative sample of 150 concordance lines reveals that *"should"* is consistently used to establish norms and direct audience beliefs. *(For the complete categorised list, please refer to Appendix E).*

The dominant function, accounting for approximately 60% of the sample, is to provide a Recommendation or give Advice. In this capacity, speakers move from being informants to advocates, explicitly stating what they believe is the correct or most beneficial course of action. This advice ranges from general calls to action (1), to specific recommendations for entities like companies (2). This advisory function also extends to negative recommendations, urging restraint and caution, as seen in example (3).

(1) We **should** focus on the things that matter. (Appendix E, Line 9) (2) Companies **should** be more transparent about how they use our data. (Appendix E, Line 20) (3) We **should**n't optimize for engagement at all costs. (Appendix E, Line 63)

A second major function is to describe an **Expectation or an Ideal** (approx. 30%). Here, *"should"* is used to define an ideal state or what is expected to be true according to a set of

shared values. This is a powerful rhetorical tool for establishing a "moral compass" for technological development. This function is used to state high-level philosophical principles (4), define technical standards (5), and assert fundamental rights in the context of technology (6).

(4) Technology **should** serve humanity, not the other way around. (Appendix E, Line 47) (5) Algorithms **should** be transparent and explainable. (Appendix E, Line 44) (6) Privacy **should** be a fundamental human right. (Appendix E, Line 30)

A minor but notable function is the use of *"should"* in **Rhetorical Questions** (approx. 10%). This structure is employed to challenge common assumptions or to pre-emptively address audience skepticism, thereby framing the subsequent information as important and necessary.

(7) ...why **should** we care about this? (Appendix E, Line 8)

In summary, the analysis of *"should"* reveals its crucial role as the primary modal for shifting the discourse from description to prescription. Through recommendations, the establishment of ideals, and rhetorical questioning, speakers use *"should"* to build a persuasive argument, shape the audience's moral and logical perspective, and advocate for a responsible and intentional approach to technology's future.

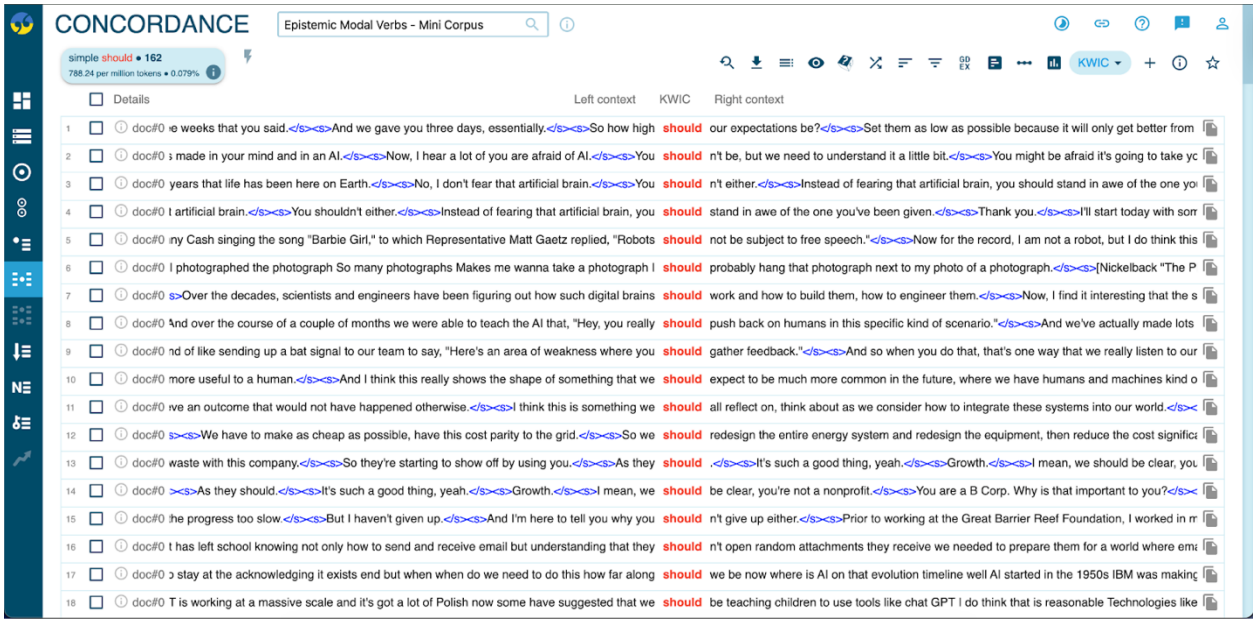


Figure 5.12

Concordance lines for the modal *"should"* in the corpus

5.9 Analysis of "ought to"

5.9.1 Frequency Analysis

The final modal verb in the "Recommendation and Advisability" group, *"ought to"*, was searched for across the entire corpus. The search, however, yielded zero (0) occurrences.

5.9.2 Interpretation of the Null Finding

This null finding, while seemingly a simple absence of data, is analytically significant and provides a revealing insight into the stylistic and rhetorical norms of the technology-themed TED Talk genre. In contemporary English, *"ought to"* occupies a similar semantic space to *"should"*, expressing advisability and moral or logical obligation. However, it is often perceived as being more formal, carrying a stronger, more objective sense of duty, and is less frequent in spoken discourse.

The complete absence of *"ought to"* in a corpus of this size strongly suggests that its exclusion is a feature of the genre's specific rhetorical strategy. TED Talk speakers aim to build a connection with a broad, diverse audience, adopting a tone that is authoritative yet accessible. The choice of *"should"*, which is direct, universally understood, and carries the appropriate weight of a recommendation without sounding overly dogmatic, aligns perfectly with these goals. The use of the more formal and less common *"ought to"* might create a rhetorical distance, positioning the speaker as more prescriptive or detached, which runs counter to the engaging and relatable style that defines a successful TED Talk.

Therefore, the non-use of *"ought to"* is as revealing as the use of other modals. It underscores a consistent preference for more conventional and direct modal forms when conveying advice and shaping audience perspectives. This linguistic choice helps to construct the characteristic register of the discourse: one that is persuasive and influential while maintaining a sense of accessibility and shared conversation.

6. Discussion

This chapter offers a synthesis of the quantitative and qualitative findings from Chapter 5, providing a holistic interpretation of how modal verbs function in technology-themed TED Talks. By connecting the observed linguistic patterns to the study's theoretical framework—namely *Modality Theory*, Coates' *Modal Strength Scale*, and a *Discourse-pragmatic Approach*—this discussion aims to illuminate the rhetorical characteristics that define this influential genre of communication.

A primary observation emerging from the data is the overwhelming dominance of modals expressing *Possibility and Certainty*. The notable prevalence of *"can"* and *"could"* is a cornerstone of the presenters' persuasive approach. The frequent use of *"can"* serves to ground the narrative in the realm of tangible, verifiable achievements, establishing a strong baseline of *epistemic certainty*. This aligns with a core principle of *Modality Theory*, where speakers make linguistic choices to express their commitment to a proposition's truth. Building upon this foundation, the extensive use of *"could"* allows orators to pivot towards a visionary future, exploring propositions with a lower degree of commitment. This dynamic movement between modals, as might be mapped on a *Modal Strength Scale*, constructs a compelling narrative trajectory that balances the credibility of present-day reality with the aspirational excitement of future potential.

Complementing this framework of possibility, the modals of *Necessity and Prediction—"will"*, *"must"*, and *"would"*—introduce a sense of compulsion and certitude to the narrative. The modal *"must"* is a clear example of *deontic modality*, expressing a strong sense of obligation and compelling the audience to accept certain actions as essential. A key insight from the *Discourse-pragmatic Approach* is that the illocutionary force of using *"must"* creates a sense of shared responsibility in the listener. In parallel, *"will"*, as a modal of high epistemic certainty, is used to make confident predictions, framing the technological future not as a possibility but as a certainty. The use of *"would"* facilitates the exploration of hypothetical consequences, further enriching the logical structure of the speakers' arguments. This set of modals is therefore crucial for elevating the discourse from what is merely possible to what is presented as imperative or inevitable.

The final functional category introduces a crucial ethical and directive dimension, primarily through the use of *"should"*. As a central modal of *deontic modality*, *"should"* is used to prescribe how technology ought to be developed and governed, thereby establishing normative guidelines. This function is a clear manifestation of the *Discourse-pragmatic Approach*, where a speaker's goal is not merely to state facts but to actively influence the audience's beliefs and future actions. The complete absence of the more formal *"ought to"* further refines this observation, suggesting a pragmatic choice for the more direct advisory force of *"should"*. Through this careful use of deontic modals, speakers position themselves as responsible guides, steering the audience's moral compass and framing the discourse around a set of shared values.

In summation, the interplay between these three modal groups reveals a highly structured communicative tactic. The findings provide a clear, practical demonstration of *Modality Theory* in action. Presenters fluidly combine *epistemic* and *deontic* modalities, and strategically move along a *Modal Strength Scale*, to achieve their pragmatic goals. They first build credibility with high-certainty statements of possibility, then create urgency with modals of necessity, and finally guide the audience's response with modals of advisability. This strategic use of modality is central to the persuasive power of the genre and highlights how theoretical linguistic concepts are actively deployed in influential, real-world communication.

7. Conclusion

This study set out to investigate the frequency and pragmatic functions of nine modal verbs within a specialised corpus of technology-themed TED Talks. The analysis has demonstrated that the use of modality in this genre is not arbitrary, but rather constitutes a sophisticated and systematic rhetorical strategy. This concluding chapter will summarise the principal findings, discuss their theoretical and practical implications, and suggest potential avenues for future research.

The primary finding of this research is that speakers in the corpus strategically deploy different groups of modals to construct a multi-layered persuasive narrative. The analysis revealed a dominance of modals expressing *epistemic possibility*, particularly *'can'* and *'could'*,

which speakers use to build a narrative of capability and vision. This was complemented by modals of *epistemic necessity* and prediction, such as *'will'* and *'must'*, which function to add urgency and certainty to the speakers' propositions. Finally, a *deontic* dimension is introduced through modals of *advisability*, primarily *'should'*, which serves to provide ethical and practical guidance to the audience.

These findings offer a compelling practical illustration of the core tenets of *Modality Theory*. The clear division of labour between *epistemic* modals (expressing the speaker's degree of certainty) and *deontic* modals (expressing obligation) is consistently observed throughout the corpus. Furthermore, the nuanced choices between modals within the same group can be effectively mapped onto a *Modal Strength Scale*. The strategic selection of a high-certainty modal like *'can'* versus a lower-certainty, speculative modal like *'might'* is a key technique speakers use to balance credibility with visionary appeal. From a *Discourse-pragmatic Approach*, this study confirms that modal choice is a fundamentally pragmatic act, tailored to the specific persuasive goals of the speaker and the generic conventions of the TED Talk.

The implications of this study are threefold. First, it contributes to the field of corpus linguistics by providing a detailed case study of modal verb usage in a highly influential, contemporary digital genre. Second, it has pedagogical applications for English for Specific Purposes (ESP), offering a clear framework for teaching the rhetorical use of modals in persuasive academic and public speaking. Finally, it provides valuable insights for communicators in science and technology, demonstrating the linguistic strategies that underpin effective and engaging discourse on complex topics.

Naturally, this study has its limitations. Its focus on a single genre and a single topic means that the findings may not be generalisable to other forms of discourse. Future research could, therefore, expand upon this work through comparative studies, analysing modal use in TED Talks on different subjects, such as social sciences or arts, to identify topic-specific patterns. A diachronic analysis could also investigate whether the use of modality in TED Talks has evolved over time. Further research could also move beyond modality to analyse other key linguistic features, such as metaphor or pronoun use, within the same corpus to build an even richer understanding of this discourse genre.

References

Bashir, A., Ullah, I., & Iqbal, L. (2023). Epistemic modal verbs in the field of linguistics and literature: A corpus-based study. *Journal of Social Sciences Review, 3*(2), 736–744. https://doi.org/10.54183/jssr.v3i2.310

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Coates, J. (1983). *The semantics of the modal auxiliaries*. Croom Helm. Retrieved from https://dokumen.pub/the-semantics-of-the-modal-auxiliaries-1138989983-9781138989986.html

Coates, J. (1987). Epistemic modality and spoken discourse. *Transactions of the Philological Society, 85*(1), 110–131. https://doi.org/10.1111/j.1467-968X.1987.tb00714.x

Holmes, J. (1982). Expressing doubt and certainty in English. *RELC Journal, 13*(2), 9–28. https://doi.org/10.1177/003368282013002O2

Hunston, S. (2011). *Corpus approaches to evaluation: Phraseology and evaluative language*. Routledge.

Hyland, K. (1998). *Hedging in scientific research articles*. John Benjamins.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173–192. https://doi.org/10.1177/1461445605050365

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography, 1*(1), 7–36. https://doi.org/10.1007/s40607-014-0009-9

Nuyts, J. (2001). *Epistemic modality, language, and conceptualization: A cognitive-pragmatic perspective*. John Benjamins.

Palmer, F. R. (2001). *Mood and modality* (2nd ed.). Cambridge University Press.

Winiharti, M. (2012). The difference between modal verbs in deontic and epistemic modality. *Humaniora, 3*(2), 532–539.

Zhang, J. (2019). A Semantic Approach to the English Modality. Journal of Language Teaching & Research, 10(4).

EVALUATIVE ADJECTIVES AND CONCEPTUAL METAPHORS IN MUSIC REVIEWS: A CORPUS-BASED ANALYSIS OF PITCHFORK REVIEWS

Vo Huu Khanh[1], Nguyen Cao Thang[2], Truong Pham Xuan Mai[3], Nguyen Tran Kim Thanh[4], Bui Vy Yen Ly[5]

Abstract:

This study investigates the use of evaluative adjectives and conceptual metaphors in Pitchfork album reviews to reveal linguistic strategies in music journalism. A corpus of

100 reviews (50 high-rated and 50 low-rated), totaling over 110,000 words, was analyzed through both quantitative and qualitative methods. The quantitative analysis shows that high-rated reviews rely on a limited set of recurring positive adjectives, while low-rated reviews employ a more diverse and critical vocabulary. The qualitative analysis, drawing on Conceptual Metaphor Theory, identifies recurrent metaphorical frameworks that map music onto domains of physical experience, spatial orientation, emotional states, artistic creation, and containment. These metaphorical patterns enable critics to translate abstract auditory experiences into accessible descriptions. The findings highlight how evaluative language not only conveys judgment but also constructs cultural value. Despite focusing solely on adjectives and one publication, the study demonstrates the potential of corpus-based methods in analyzing evaluative discourse.

Keywords: corpus linguistics; evaluative adjectives; conceptual metaphor; music journalism; Pitchfork reviews

Chapter 1. Introduction

1.1. Background of the study

Music has long held a central and profound position in human culture, serving as a powerful medium for eliciting "strong and diverse emotions" (Vuilleumier & Trost, 2015), activating social solidarity, and promoting mental health and subjective well- being (Boer & Abubakar, 2014). Within this rich landscape, music reviews act as a critical intermediary, guiding listeners and shaping the cultural reception of new and established works. These reviews, published in various forms of music journalism, provide an important evaluative discourse that helps frame public opinion and influence an artist's trajectory. Paraphrased, a crucial aspect of this kind of writing is its highly creative and subjective nature, as each piece is authored by a single reviewer (Trojszczak, 2024), and especially the strategic use of adjectives, which are essential for conveying nuanced judgments and qualitative assessments. Among contemporary music publications, Pitchfork stands as a particularly influential voice. With its long history, dedicated readership, and a reputation for detailed criticism and a

[1] Học viên Cao học Ngôn ngữ Anh tại Đại học Cần Thơ
[2] Học viên Cao học Ngôn ngữ Anh tại Đại học Cần Thơ
[3] Học viên Cao học Ngôn ngữ Anh tại Đại học Cần Thơ
[4] Học viên Cao học Ngôn ngữ Anh tại Đại học Cần Thơ
[5] Học viên Cao học Ngôn ngữ Anh tại Đại học Cần Thơ

quantitative scoring system, it provides a unique and valuable corpus for a linguistic analysis of evaluative language in music journalism.

1.2. Rationale

Despite the significant role of music journalism in cultural life, there remains a notable absence of empirical, corpus-based linguistic research specifically focused on the evaluative language used in this genre. While previous studies have explored aspects of critical writing, a gap persists in our understanding of how lexical choices, particularly the use of adjectives, might systematically vary in relation to the review's final score. An analysis of these patterns can reveal more than just word choice; it has the potential to uncover the conceptual frameworks and metaphorical language that critics employ to articulate their judgments. By investigating these linguistic connections, we can gain a deeper insight into the underlying mechanisms of music criticism and its role in constructing value within the music industry.

1.3. Research Objectives

Building upon this rationale, the present study establishes three primary objectives. First, it aims to compare the frequency, variety, and distribution of adjectives found in Pitchfork's highest- and lowest-rated reviews. Second, the research seeks to identify the dominant semantic domains of adjectives that are commonly used within the music review genre, thereby classifying the types of concepts critics prioritize when evaluating music. Finally, this study will analyze the conceptual metaphors that underlie the use of these adjectives, revealing the deeper, often subconscious, cognitive patterns that shape music criticism at a linguistic level.

1.4. Research Questions

1. Do high-rated reviews use adjectives from different semantic domains than low-rated reviews?
2. What conceptual metaphors are reflected in the use of language in Pitchfork reviews?

Chapter 2. Literature review

2.1. Music journalism

Music journalism occupies a significant position within cultural discourse, acting as a critical bridge between artists, their work, and the public. Historically rooted in print magazines and newspapers, the field has evolved considerably with the rise of digital platforms, a shift that has intensified the speed and volume of critical output while also democratizing access to it (Hesmondhalgh, 2013). Scholarly work on the topic has often focused on the socio-cultural function of music reviews, highlighting their role in taste- making, community-building, and the legitimation of particular musical genres. At its core, a music review is a form of expert discourse that aims to evaluate an artistic product. This process is inherently linguistic, relying on a specific set of lexical and rhetorical strategies to articulate subjective judgments in a

manner that appears authoritative and convincing. Therefore, understanding the language of music criticism is key to comprehending its broader cultural impact.

2.2. Adjectives in evaluation

The role of adjectives in evaluative discourse is a well-established area of linguistic inquiry. Adjectives are primarily used to describe, qualify, or specify nouns, but a significant subclass of these words is dedicated to expressing subjective judgment.

2.2.1. Categorization of adjectives

Linguistic studies often classify adjectives based on their semantic and syntactic functions. Syntactically, they can be either attributive (preceding a noun, e.g., "a great album") or predicative (following a verb, e.g., "the album is great"). Semantically, adjectives can be further categorized. One common distinction is between descriptive adjectives (e.g., blue, square) and evaluative adjectives (e.g., beautiful, horrible). While descriptive adjectives refer to objective, verifiable properties, evaluative adjectives convey the speaker's or writer's opinion or attitude towards the subject.

2.2.2. Evaluative adjective

The specific category of evaluative adjectives is central to any study of critical writing, as these are the words that express a writer's "stance" by communicating assessments and value judgments (Biber et al., 1999). Recent linguistic research has provided a more nuanced characterization of these terms, defining them as a class of predicates that express various kinds of normative and epistemic evaluation, including judgments of personal taste and aesthetic value (Silk, 2019). These adjectives are often distinguished by their context-dependent nature, functioning as "measure functions" that map to degrees of taste, beauty, or probability depending on the specific context of their use. Furthermore, adjectives can be considered multidimensional, where the overall positive or negative evaluation conveyed by a word like 'intense' or 'insane' depends entirely on which dimensions of meaning are made salient by the surrounding discourse (Stojanovic, 2015). For example, a critic might use the adjective intense to describe a powerfully emotional vocal performance (positive) or to describe an overwhelming and chaotic mix (negative). By analyzing the choice and frequency of these adjectives, researchers can quantitatively and qualitatively assess the specific linguistic tools critics use to construct their arguments and communicate their final judgments.

2.3. Conceptual Metaphor Theory (CMT)

Conceptual Metaphor Theory (CMT) was introduced by George Lakoff and Mark Johnson in their seminal work, *Metaphors We Live By* (1980). This theory posits that metaphors are not merely poetic or rhetorical flourishes but are fundamental cognitive structures that shape human perception and reasoning. According to CMT, abstract concepts (the target domain) are understood and reasoned about in terms of more concrete, familiar concepts (the source domain). For example, the abstract concept of ARGUMENT is often

understood through the concrete source domain of WAR, as seen in phrases like "he attacked my argument" or "I defended my position."

Studies have shown that metaphorical language is pervasive in everyday communication and that it is processed and understood by the brain as a mapping between domains, not just as a simile or comparison (Gibbs, 1994). This theory is particularly relevant to the study of music criticism, as it offers a framework for analyzing how abstract, non-physical qualities of music (e.g., mood, structure, emotion) are described using concrete, tangible language. A critic might describe a sound as sharp or a melody as sweet, using sensory domains like touch and taste to evaluate the auditory experience. Analyzing these patterns can reveal the shared conceptual schemas that underpin the critical evaluation of music.

The application of CMT to music is a central theme in the paper "*Conceptualising Music: Metaphors of Classical Music Reviews*" by Inesa Šeškauskienė and Totilė Levandauskaitė. Their research, based on a corpus of 48 classical music reviews, identified 11 major metaphors used to conceptualize music. The most prominent of these, accounting for almost two-thirds of the linguistic metaphors in their data, are the **MOTION**, **CONTAINER**, and **LINGUISTIC CREATION** metaphors. The findings suggest that these metaphors are not simply decorative, but rather "*structure classical music reviews and underlie our reasoning about classical music to a very large extent*".

The authors highlight that the **MOTION** metaphor is the most frequently represented, often conceptualizing music as listeners or performers "moving along a landscape" or as a "moving object". This is supported by expressions like "the music goes faster here" or "we are coming to the coda". They also identified a variety of CONTAINER metaphors, such as **MUSICAL WORK/WORKS IS A CONTAINER**, and the **MUSIC IS LINGUISTIC CREATION** metaphor, which includes varieties like **MUSIC IS LITERATURE** and **MUSIC IS VERBAL RHETORIC**. The paper further notes that innovative metaphors in the genre tend to be more evaluative, which is a key characteristic of reviews. The use of such corpus-based methodologies marks a shift from the more introspective research of previous generations of cognitive linguists, offering a more objective and verifiable approach to metaphor identification.

Chapter 3. Methodology

3.1. Corpus Description

The present study utilizes a specialized corpus of music reviews to investigate the use of evaluative adjectives in music journalism. The corpus, compiled from the influential online publication Pitchfork, consists of 100 full-text album reviews. To facilitate a comparative analysis, the corpus is divided into two distinct sub-corpora: one consisting of 50 reviews with a final score of 8.0 or higher ("high-rated reviews"), and the other consisting of 50 reviews with a score of 6.0 or lower ("low-rated reviews"). This segmentation allows for a direct comparison of linguistic patterns across positive and negative critical assessments. The total

corpus size is slightly over 110,000 words, providing a robust dataset for both quantitative and qualitative analyses.

3.2. Data Preparation

All text data was collected and prepared for analysis using Sketch Engine, a comprehensive platform for corpus management and linguistic research. The raw text from the selected reviews was processed and uploaded to the platform, where it was automatically tokenized, part-of-speech tagged, and lemmatized. The corpus was then structurally divided into the two aforementioned sub-corpora, enabling comparative queries and analyses to be performed on each group separately. This data preparation process ensures the linguistic data is consistently annotated and ready for both statistical and close-reading techniques.

3.3. Quantitative Analysis

The quantitative phase of this study will address the research objective of comparing the frequency, variety, and distribution of adjectives. Using Sketch Engine's corpus query tools, a detailed wordlist will be generated for each of the two sub-corpora. This wordlist will be sorted by frequency to identify the most common adjectives used in both high- and low-rated reviews. The analysis will compare not only the raw frequency of specific adjectives but also their relative frequency and distribution across the two sub-corpora. This will provide a statistical foundation for determining whether there are significant differences in the use of evaluative language between positive and negative reviews.

3.4. Qualitative Metaphor Analysis

Following the quantitative findings, a qualitative analysis will be conducted to address the remaining research objectives, specifically the identification of semantic domains and underlying conceptual metaphors. This phase will involve a close reading of the most frequent adjectives identified in the wordlist, as well as their co-occurrence patterns. The adjectives will be manually categorized into semantic domains (e.g., sensory language, emotional terms, spatial descriptors). Informed by CMT, the analysis will then proceed to interpret these semantic domains to identify the metaphorical frameworks that critics employ. For instance, the use of adjectives from a sensory domain may reveal that critics conceptualize music in terms of physical experiences. This iterative process of categorization and interpretation will allow for the deep analysis of how metaphorical language shapes critical evaluation in music journalism.

Chapter 4: Findings

4.1. Differences in the Use of Adjectives

4.1.1. Lexical Diversity

Lexical diversity was measured through the Type-Token Ratio (TTR), which was calculated based on adjective tokens extracted from each sub-corpus. The results reveal a marked difference between the two datasets.

The low-rated reviews display a very high TTR of 82 percent, which suggests that reviewers drew on a large variety of adjectives relative to the total number of adjective occurrences. This pattern is likely explained by the fact that the overall adjective count is smaller in this sub-corpus, making repetition less frequent.

In contrast, the high-rated reviews show a TTR of 47 percent. This lower value indicates that a narrower set of adjectives is used repeatedly when reviewers express approval and praise for albums.

The whole corpus, combining both sets, yields an overall TTR of approximately 30 percent. This outcome is expected, as the merging of both sub-corpora increases repetition and thereby reduces overall diversity. These figures demonstrate that the lexical profile of negative reviews is more varied, whereas positive reviews rely more heavily on a limited but recurring pool of evaluative adjectives.

4.1.2. Adjective Frequency and Distribution

Beyond lexical diversity, it is also important to consider the frequency and distribution of adjectives within each sub-corpus. Using Sketch Engine's wordlist and frequency tools, comprehensive lists of adjectives were generated and sorted by descending frequency. These data allow for closer inspection of the evaluative vocabulary that characterises each set of reviews.

In the high-rated reviews, the most frequent adjectives are *new* (106), *good* (101), *first* (100), *more* (85), and *own* (84). Other frequently used terms include *early, other, many, same,* and *big*. Table 1.1. presents the fifty most frequent adjectives in this sub- corpus. The high-rated reviews are thus dominated by positively evaluative adjectives that stress novelty, quality, and significance. Words such as new, good, and great are central to this pattern and indicate a preference for repetitive but highly effective descriptors in the act of praising an album.

Table 1. 1.

Most frequent adjectives in high-rated reviews

	Adjective	Frequency		Adjective	Frequency
1	new	106	26	different	28

2	good	101	27	final	27
3	first	100	28	most	27
4	more	85	29	vocal	25
5	own	84	30	full	25
6	early	70	31	spiritual	23
7	other	66	32	gaga	23
8	many	49	33	much	22
9	same	46	34	emotional	22
10	big	45	35	live	22
11	few	42	36	hard	21
12	last	42	37	past	21
13	great	40	38	long	21
14	musical	38	39	pop	21
15	late	35	40	true	20
16	old	35	41	deep	20
17	little	35	42	underground	20
18	high	34	43	least	19
19	next	34	44	simple	19
20	only	33	45	third	19
21	single	32	46	close	19
22	bad	31	47	sweet	17
23	american	30	48	heavy	17
24	second	29	49	previous	16
25	real	28	50	small	16

The low-rated reviews display a somewhat different distribution. The most frequent adjectives are *good* (86), *new* (60), *first* (52), *more* (43), and *bad* (41). Other recurring adjectives include *others, own, last, few,* and *early*. Table 1.2. illustrates the fifty most frequent adjectives in the low-rated sub-corpus.

Table 1. 2.

Most frequent adjectives in low-rated reviews

	Adjective	Frequency		Adjective	Frequency
1	good	86	26	full	16
2	new	60	27	young	15

3	first	52	28	high	15
4	more	43	29	classic	14
5	bad	41	30	long	14
6	other	37	31	such	14
7	own	34	32	clear	14
8	last	33	33	only	14
9	few	32	34	recent	13
10	early	32	35	most	13
11	same	31	36	american	13
12	big	28	37	fine	12
13	many	27	38	original	12
14	little	27	39	musical	12
15	pop	26	40	creative	12
16	hard	24	41	vocal	11
17	least	23	42	enough	11
18	real	22	43	simple	11
19	single	20	44	different	11
20	old	19	45	actual	10
21	much	18	46	contemporary	10
22	late	18	47	weird	10
23	past	18	48	sixth	10
24	great	18	49	modern	10
25	second	18	50	interesting	9

While there is some overlap between the two sets, such as the frequent use of good, new, first, and more, the low-rated reviews introduce a stronger presence of negatively marked adjectives. The appearance of terms such as *bad*, *least*, *hard*, *weird*, and *different* points to a more critical and varied evaluative stance.

A comparison of the two sets highlights an important contrast. Positive reviews tend to concentrate on a small repertoire of familiar praise words, which are repeated frequently to emphasise artistic merit, innovation, or scale. Negative reviews, on the other hand, employ a

wider range of descriptors to articulate different aspects of failure or mediocrity, whether through direct condemnation (*bad*), stylistic critique (*weird*), or judgments of inadequacy (*least*). These patterns are consistent with the TTR results presented above, where the low-rated set demonstrated higher lexical diversity.

The results therefore indicate that while Pitchfork reviewers draw on adjectives extensively in both positive and negative assessments, the rhetorical strategies differ. Positive reviews show a preference for repetition and reinforcement of certain evaluative markers, whereas negative reviews favour a greater variety of adjectives to capture the nuances of criticism. This finding directly addresses the first research objective by demonstrating that the evaluative language of Pitchfork reviews varies significantly according to rating context.

4.2. Music metaphor

A number of adjectives belong to the sensory domain, most notably *sweet*, which exemplifies the metaphor MUSIC IS A PHYSICAL EXPERIENCE. In one review, E.R. Pulgar describes POiSON GiRL FRiEND's vocals as carrying "*the la-da-dee, la-dee- das and the sweet dreams of rhythm and dancing*," situating the sound within the realm of taste and dream imagery. Similarly, tactile qualities emerge with adjectives such as *hard* and *warm*, which project listening as an embodied physical act.

The corpus also revealed frequent spatial descriptors, including *high*, *low*, *deep*, *big*, and *small*, mapping music onto dimensions and location. Jayson Greene, for example, observes that Justin Bieber sings with an acoustic guitar "*capo'd high and twinkling like a music box.*" This framing encourages readers to visualize pitch through vertical elevation. By contrast, Jill Mapes criticizes Katy Perry's lyrics by noting that

"*[…] the bar is low when it comes to pop lyrics […]*"

employing a downward spatial metaphor to index weak quality. Dylan Green similarly references "deep cuts" in describing Czarface's work, embedding cultural knowledge within a vertical metaphor. References to *big hooks* and being "*too small, too inept to even put forth one meaningful, specific, original idea*" further demonstrate how scale metaphors frame musical value. Collectively, these examples align with the conceptual metaphor MUSIC IS SPACE OR MOTION.

A third group of adjectives highlights emotional and psychological framing, reflecting the metaphor MUSIC IS EMOTION OR INNER STATE. Dani Blum remarks on how Bennett amplifies "*the mood or emotional register that makes [a song] compelling*," foregrounding affect as central to value. Mano Sundaresan characterizes Pi'erre Bourne's *Good Movie* as

"*[…] a weird, uneven album, hot and stormy like a New York summer […]*"

combining evaluative judgment with affective imagery. Daniel Bromfield similarly comments that Pan American and Kramer's collaboration "*gets more interesting in the back*

half," suggesting listener engagement as a dynamic emotional trajectory. Rae-Aila Crumble adds another dimension by connecting Brandee Younger's work to "*the meditative power of spiritual jazz*," explicitly situating music as a pathway to transcendence.

The artifact and creation domain frames music as a crafted product. Mapes critiques Katy Perry's attempt to create "*a cohesive album*," while Dylan Green praises Czarface's track for locking together like

"*[…] an amped-up version of the original PlayStation startup music*."

Bromfield (2025) notes Nelson's restraint from

"*the temptation to bend his strings in the fashion of classic movie scores*,"

while Crumble observes Younger's "*expanded creative control as bandleader*." These examples foreground MUSIC IS AN ARTIFACT OR CREATION, in which albums are evaluated as deliberately constructed cultural objects.

Finally, container metaphors emerged particularly in negative reviews. Jeremy D. Larson dismisses Benson Boone's *American Heart* by calling each song "*just an empty lunchbox with a few crumbs left by his betters*." Similarly, Sheldon Pearce critiques Tory Lanez's *Memories Don't Die* for relying on "*a mélange of hollow sounds and wispy melodies*." Both examples exemplify MUSIC IS A CONTAINER, where the absence of meaningful content is expressed as physical emptiness.

Taken together, these findings demonstrate that *Pitchfork* reviewers rely on a network of metaphors, including PHYSICAL EXPERIENCE, SPACE/MOTION, EMOTION, ARTIFACT, and CONTAINER, to render musical qualities accessible to readers. These metaphors not only align with established conceptual frameworks in discourse studies but also reveal the distinctive prominence of sensory and experiential imagery in digital music journalism.

Chapter 5. Limitations and Conclusion

5.1. Limitations

Although this study has offered meaningful insights into the evaluative language of music reviews, several limitations should be acknowledged. First, the analysis focused exclusively on adjectives as the primary linguistic feature of evaluation. While adjectives indeed play a crucial role in critical writing, other linguistic resources such as adverbs, verbs, intensifiers, or discourse-level strategies were not examined. This narrow scope inevitably restricts the comprehensiveness of the findings, as evaluative meaning is often distributed across multiple grammatical categories.

Second, despite compiling over 110,000 words from 100 reviews, the corpus size remains relatively modest in comparison with larger-scale corpus-linguistic studies. In addition, the length of reviews is uneven, with some texts being considerably longer than

others. This imbalance in corpus size distribution may have subtly influenced frequency results and limited the robustness of comparative analysis.

Third, the study treated reviews as purely textual data, excluding multimodal aspects such as images, embedded audio-visual material, or layout design. Since contemporary digital music journalism is inherently multimodal, the absence of these elements means that the analysis could not account for how non-textual features may interact with language to shape evaluative meaning.

Taken together, these limitations suggest that the findings should be interpreted as a partial view of evaluative practices in Pitchfork reviews, rather than as a comprehensive account of evaluative discourse in music journalism.

5.2. Conclusion

This study set out to investigate how evaluative language is employed in Pitchfork reviews, with a focus on adjectives and their metaphorical extensions. Using a combination of corpus-linguistic methods and Conceptual Metaphor Theory, the analysis has shown that review rating significantly shapes adjective use and that critics rely on recurring metaphorical frameworks to translate musical experience into accessible language.

The comparison between high- and low-rated reviews demonstrated that negative evaluations draw on a wider and more varied set of adjectives, while positive assessments depend on a smaller and more recurrent vocabulary. This finding suggests that criticism requires greater lexical flexibility, whereas praise relies on the reinforcement of familiar descriptors. Beyond the lexical level, the metaphor analysis revealed that critics consistently map music onto embodied and cultural domains such as physical sensation, spatial experience, emotional states, and artistic creation. These metaphors make the abstract qualities of sound comprehensible by grounding them in everyday experience.

The study contributes to corpus-based approaches in cultural analysis by illustrating how computational tools can uncover systematic patterns in evaluative discourse. It also adds to research on music criticism by documenting how online reviewers balance functional description with metaphorical elaboration in order to construct value. At the same time, its scope is necessarily limited by its focus on adjectives and by the restriction of the dataset to a single publication and to reviews at the rating extremes. Expanding the analysis to other linguistic categories, additional review outlets, and different rating ranges would allow for a more comprehensive account of evaluative practices in music journalism.

In summary, this research has shown that the language of Pitchfork reviews reflects both the demands of critical evaluation and the cognitive patterns through which music is understood. By combining quantitative and qualitative methods, the study has demonstrated how linguistic analysis can shed light on the cultural work of criticism, revealing not only what reviewers say about music but also how they frame and conceptualize the listening experience.

REFERENCES

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* Longman.

Boer, D., & Abubakar, A. (2014). Music listening in families and peer groups: benefits for young people's social cohesion and emotional well-being across four cultures. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00392

Gibbs, R. W. (1994). *The poetics of mind: figurative thought, language, and understanding*. https://ci.nii.ac.jp/ncid/BA23866085

Hesmondhalgh, D. (2013). *Why music matters*. http://ci.nii.ac.jp/ncid/BB13857493

Šeškauskienė, I., & Levandauskaitė, T. (2013). Conceptualising Music: Metaphors of Classical Music Reviews. *Studies About Languages*, *0*(23). https://doi.org/10.5755/j01.sal.0.23.5268

Silk, A. (2019). Evaluational adjectives. *Philosophy and Phenomenological Research*, *102*(1), 127–161. https://doi.org/10.1111/phpr.12635

Stojanovic, I. (2015). Evaluative adjectives and evaluative uses of ordinary adjectives.

In *Proceedings of LENLS12: Language Engineering and Natural Language Semantics*.

Trojszczak, M. (2024). Expressing negative opinions through metaphor and simile in popular music reviews. *Lodz Papers in Pragmatics*. https://doi.org/10.1515/lpp- 2024-0036

Vuilleumier, P., & Trost, W. (2015). Music and emotions: from enchantment to entrainment. *Annals of the New York Academy of Sciences*, *1337*(1), 212–222. https://doi.org/10.1111/nyas.12676

APPENDICES

Appendix 1

Raw worldlist data from 8+ reviews sub-corpora

	Adjective	Frequency
1	new	106
2	good	101
3	first	100
4	more	85
5	own	84
6	early	70
7	other	66
8	many	49
9	same	46
10	big	45
11	few	42
12	last	42
13	great	40

	Adjective	Frequency
14	musical	38
15	late	35
16	old	35
17	little	35
18	high	34
19	next	34
20	only	33
21	single	32
22	bad	31
23	american	30
24	second	29
25	real	28
26	different	28

	Adjective	Frequency
27	final	27
28	most	27
29	vocal	25
30	full	25
31	spiritual	23
32	gaga	23
33	much	22
34	emotional	22
35	live	22
36	hard	21
37	past	21
38	long	21
39	pop	21

	Adjective	Frequency
40	true	20
41	deep	20
42	underground	20
43	least	19
44	simple	19
45	third	19
46	close	19
47	sweet	17
48	heavy	17
49	previous	16
50	small	16

Appendix 2

Raw worldlist data from 8- reviews sub-corpora

	Adjective	Frequency
1	good	86
2	new	60
3	first	52
4	more	43
5	bad	41
6	other	37
7	own	34
8	last	33
9	few	32
10	early	32
11	same	31
12	big	28
13	many	27

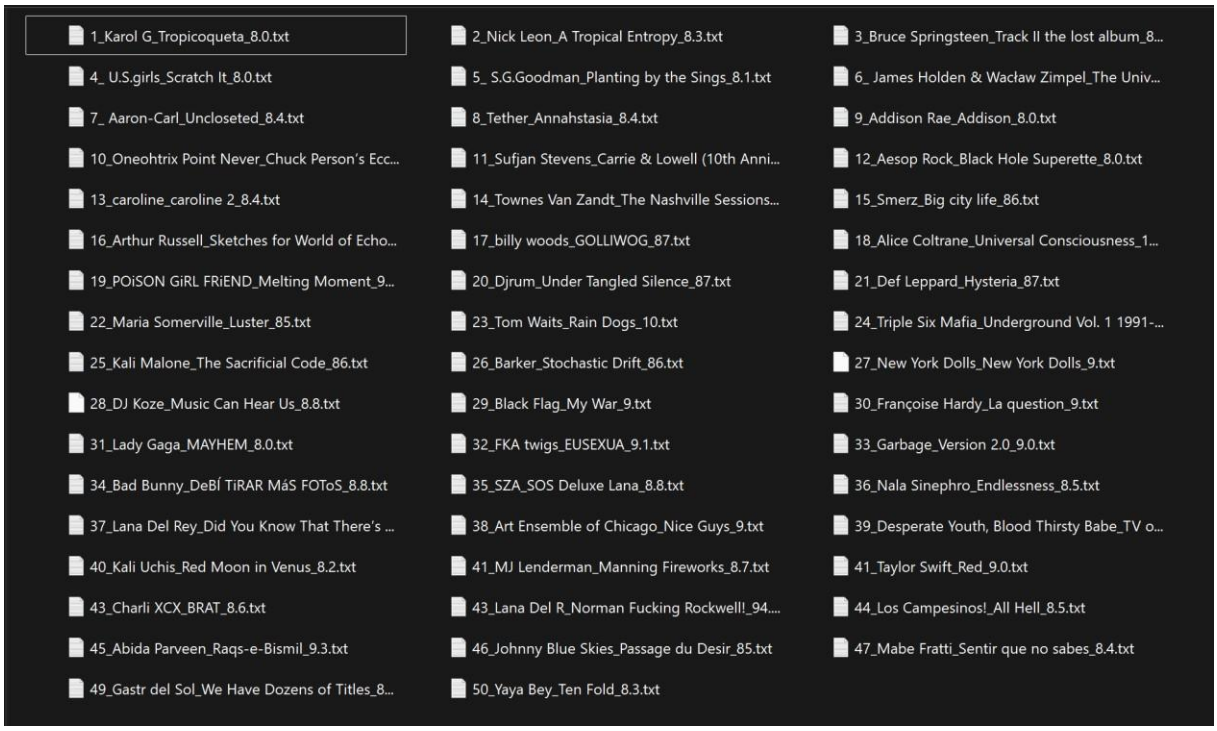
	Adjective	Frequency
14	little	27
15	pop	26
16	hard	24
17	least	23
18	real	22
19	single	20
20	old	19
21	much	18
22	late	18
23	past	18
24	great	18
25	second	18
26	full	16

	Adjective	Frequency
27	young	15
28	high	15
29	classic	14
30	long	14
31	such	14
32	clear	14
33	only	14
34	recent	13
35	most	13
36	american	13
37	fine	12
38	original	12
39	musical	12

	Adjective	Frequency
40	creative	12
41	vocal	11
42	enough	11
43	simple	11
44	different	11
45	actual	10
46	contemporary	10
47	weird	10
48	sixth	10
49	modern	10
50	interesting	9

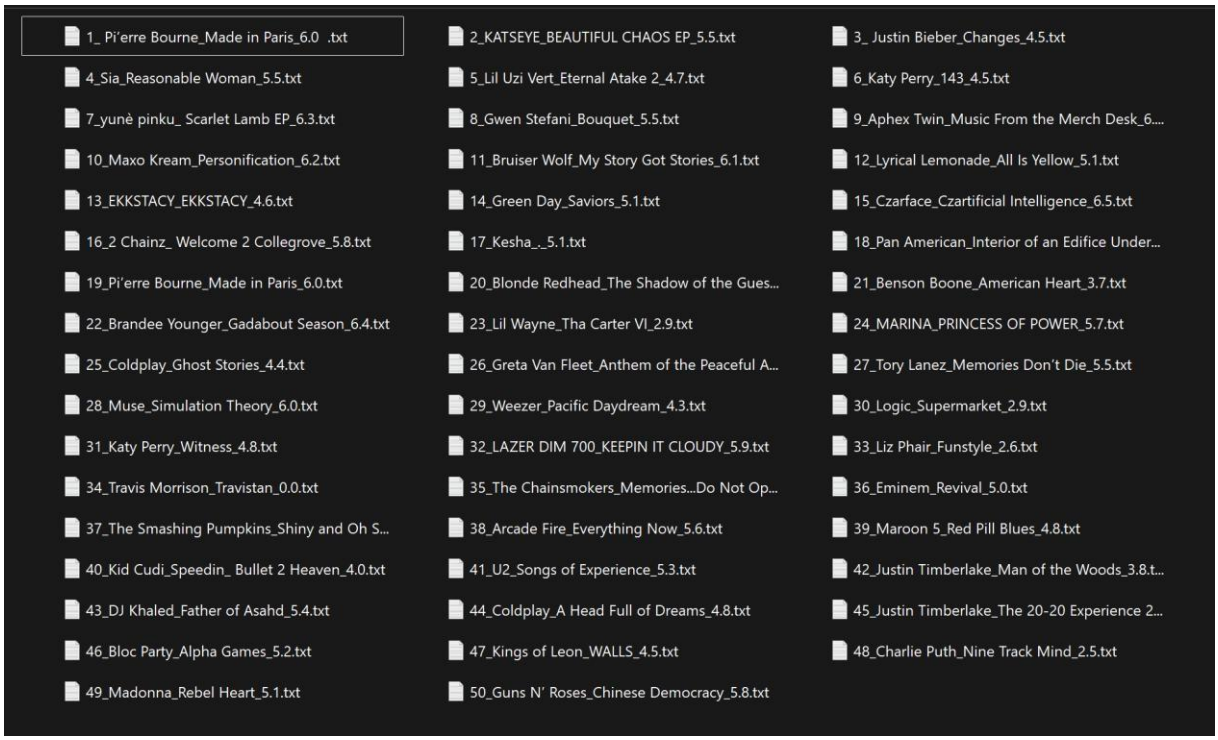
Appendix 3

Demonstration pf naming of 8+ review files



Appendix 4

Demonstration of naming of 6+ review files



XÂY DỰNG NGỮ LIỆU HUẤN LUYỆN CHO MÔ HÌNH NHẬN DẠNG VĂN BIA HÁN NÔM

Trương Thị Vân Anh[1], Bùi Hồng Phúc[2], Đinh Điền[3]

Tóm Tắt

Văn bia Hán Nôm là một thành tố tiêu biểu của di sản văn hoá Việt, song phần lớn vẫn khó tiếp cận do ngôn ngữ cổ phức tạp và tư liệu xuống cấp. Chúng tôi công bố HanNomCorpus, kho ngữ liệu số hoá hướng di sản gồm hai thành phần: HanNom - OCR (ảnh văn bia với chú giải vùng/bounding box và chuẩn hoá mã để huấn luyện–đánh giá nhận dạng ký tự) và HanNom - Parallel (căn chỉnh dòng/câu Hán Nôm – Quốc ngữ cho hai tác vụ chuyển tự và dịch nghĩa). Bộ dữ liệu được xây dựng theo quy trình tái lập, có tiêu chuẩn gán nhãn và thẩm định chuyên gia. HanNomCorpus biến tư liệu rời rạc thành hạ tầng mở cho bảo tồn số, đọc hiểu tự động văn bia và nghiên cứu ngôn ngữ học tính toán, góp phần gìn giữ và phát huy giá trị Hán Nôm trong kỷ nguyên AI.

Từ khóa: Văn bia Hán Nôm, HanNomCorpus, OCR, Chuyển tự & dịch nghĩa Hán Nôm – Quốc ngữ, Bảo tồn số di sản.

I. GIỚI THIỆU

Văn khắc là một bộ phận quan trọng trong nền văn hóa thành văn nói chung và di sản Hán Nôm nói riêng, là hiện tượng văn hoá nảy sinh từ đời sống xã hội, mang nét đặc thù của thời kỳ cổ đại và trung cổ. Đây là nguồn tư liệu có giá trị để tìm hiểu quá khứ dân tộc trên nhiều lĩnh vực như tư tưởng, chính trị, xã hội, văn hóa, kinh tế, dân tộc học, tôn giáo và ngôn ngữ.

Trong kho tàng di sản Hán Nôm Việt Nam, một khối lượng lớn văn bản được khắc trên đá gọi là **văn bia**. Nhờ chất liệu bền vững và nghệ thuật chạm trổ tinh vi, văn bia thường gắn với nhu cầu lưu niệm, ghi nhớ công đức, nhắn nhủ hậu thế, lưu truyền công danh sự nghiệp, đánh dấu sự kiện trọng đại, hoặc ca ngợi cảnh đẹp. Văn bia xuất hiện phổ biến tại nơi thờ tự, danh lam thắng tích, vừa mang giá trị thẩm mỹ, vừa là minh chứng văn hóa – lịch sử để mọi người chiêm ngưỡng. [1]

Không chỉ phản ánh lịch sử, văn bia còn lưu giữ nhiều thông tin không có trong chính sử, nhờ đó trở thành nguồn tư liệu quý báu để bổ khuyết và soi sáng tiến trình lịch sử dân tộc. Tuy nhiên, việc khai thác văn bia hiện còn gặp nhiều khó khăn do sử dụng Hán văn cổ, cấu trúc ngữ pháp đặc thù, cùng với sự hao mòn, biến dạng của chất liệu theo thời gian.

[1] Trường Đại học Khoa học Tự nhiên, ĐHQG HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP.HCM; *Email:24C01002@student.hcmus.edu.vn*

[2] Trường Đại học Khoa học Tự nhiên, ĐHQG HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP.HCM; *Email: duyphuc2425@gmail.com*

[3] PGS.TS, Trường Đại học Khoa học Tự nhiên, ĐHQG HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP.HCM; *Email: ddien@fit.hcmus.edu.vn*

Từ thực tiễn đó, nhóm chúng tôi mong muốn xây dựng mô hình nhận dạng văn bia Hán Nôm nhằm hỗ trợ đọc hiểu, nghiên cứu và tự động hóa xử lý văn bản này. Trước tiên, cần xây dựng một kho ngữ liệu song ngữ Hán Nôm – Quốc ngữ làm tập dữ liệu huấn luyện cho mô hình dịch thuật tự động, đồng thời là tài nguyên phục vụ tra cứu, học tập và giảng dạy.

Trong phạm vi nghiên cứu này, chúng tôi tiến hành xây dựng kho ngữ liệu dựa trên các văn bia sưu tầm tại địa phương Tiên Lãng – Hải Phòng. Nội dung chữ Hán Nôm - phiên âm - dịch nghĩa được sử dụng từ sách "Văn bia Tiên Lãng – Hải Phòng", chủ biên: Đỗ Thị Hảo [1]. Hình ảnh các Văn bia tương ứng được sử dụng từ Website Vietnamica [2] do Giáo sư Philippe Papin làm ban chỉ đạo công trình. Các tài liệu quý giá này được cung cấp bởi Trung tâm Ngôn ngữ học tính toán, Trường Đại học Khoa học Tự nhiên, ĐGQG HCM (CLC).

Một số nghiên cứu liên quan được giới thiệu trong Phần 2. Giới thiệu về nguồn dữ liệu được sử dụng trong bài báo này và pipeline xử lý dữ liệu được trình bày trong Phần 3. Phần 4 giải thích về thuật toán được áp dụng và kết quả đạt được. Cuối cùng, trong Phần 5, là một số kết luận quan trọng và hướng phát triển trong tương lai đối với phương pháp được đề xuất.

2. CÁC NGHIÊN CỨU LIÊN QUAN

2.1. Hiệu chỉnh sau OCR dựa trên Levenshtein cho văn bản lịch sử

Khoảng cách Levenshtein là một trong những thước đo đáng tin cậy và dễ diễn giải trong hiệu chỉnh sau bước OCR. Srigiri và Saha [3] đã áp dụng phương pháp này kết hợp với word embeddings để hiệu chỉnh kết quả OCR cho tiếng Hindi lịch sử, đạt được cải thiện đáng kể về độ chính xác ở mức ký tự. Aditya Pal và cộng sự [4] đã nâng cao cách tiếp cận này bằng cách xếp hạng ứng viên theo ngữ cảnh, sử dụng embeddings BERT kết hợp với Levenshtein, và đạt được độ chính xác 81% trên các văn bản tiền hiện đại. S. Mihov và cộng sự [5] đã giới thiệu tự động hóa Levenshtein tích hợp với từ điển trực tuyến, tối ưu hóa cả hiệu quả lẫn độ chính xác. Một khảo sát gần đây của C. Da và cộng sự [6] nhấn mạnh tính thích ứng và độ bền vững của các quy trình dựa trên Levenshtein trong nhiều ngôn ngữ và hệ chữ viết lịch sử khác nhau.

Tuy nhiên, trong khi các phương pháp này hoạt động hiệu quả đối với hệ chữ La-tinh, chúng gặp khó khăn với các ngôn ngữ biểu ý có tập ký tự lớn. Đóng góp mới của nghiên cứu này là tích hợp Levenshtein với tương đồng thị giác và neo ngữ nghĩa từ phiên âm tiếng Việt, cho phép hiệu chỉnh OCR trong ngữ cảnh chữ Hán Nôm – điều mà các nghiên cứu trước chưa khai thác.

2.2. Sử dụng công cụ OCR Kim Hán Nôm đối với hình ảnh văn bia

Để có được bounding box dựa trên các văn bia, chúng tôi sử dụng Kim Hán Nôm (CLC Lab, 2022) [7], một hệ thống OCR được phát triển bởi Trung tâm Ngôn ngữ học Tính toán (CLC Lab). Đây là nỗ lực tiên phong trong việc phát triển một OCR chuyên biệt cho Hán Nôm, giúp tự động trích xuất văn bản và bounding box từ ảnh bia. Công cụ này khắc phục phần nào hạn chế của OCR đa ngôn ngữ như Tesseract hay Google OCR khi xử lý chữ cổ. Tuy nhiên, Kim Hán Nôm vẫn gặp lỗi box trống, nhận dạng sai ở ký tự mờ/khắc sâu. Nghiên

cứu này vừa kế thừa Kim Hán Nôm, vừa đề xuất bước hậu xử lý (sử dụng PaddleOCR để làm sạch, chia box) và đối sánh với chuỗi chuẩn hóa từ văn bản in, nhằm nâng cao chất lượng corpus OCR.

2.3 Dóng hàng với BERT align

Các công trình gần đây (Liu Zhu, 2022) [8] chứng minh BERT Align có khả năng dóng hàng câu hiệu quả cho song ngữ có ngữ pháp khác biệt, đặc biệt trong văn bản văn học. Tuy nhiên, các nghiên cứu này tập trung vào ngôn ngữ có nguồn tài nguyên dồi dào (Trung–Anh), trong khi Hán Nôm–Việt là cặp ngôn ngữ ít tài nguyên, dữ liệu song song hạn chế, lại thêm đặc thù phiên âm xen lẫn. Nghiên cứu hiện tại đã điều chỉnh thuật toán BERT Align bằng cách:

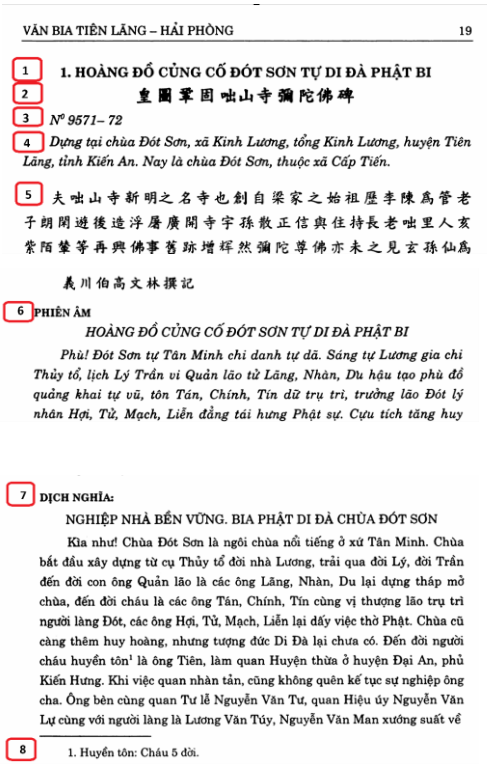
Áp dụng chia nhỏ câu phức → dóng hàng → ghép lại.

Kết hợp thêm cơ chế chèn ký tự đệm ("_") để duy trì căn chỉnh 1-1 ký tự–âm tiết. Qua đó, corpus tạo ra có tính ổn định cao hơn và phù hợp cho huấn luyện mô hình chuyển tự/dịch nghĩa.

3. DỮ LIỆU VÀ PHƯƠNG PHÁP

3.1 Dữ liệu

3.1.1 Cấu trúc Văn bản "Văn bia Tiên Lãng – Hải Phòng"


VĂN BIA TIÊN LÃNG – HẢI PHÒNG 19

1 **1. HOÀNG ĐỒ CỦNG CỐ ĐÓT SƠN TỰ DI ĐÀ PHẬT BI**

2 皇圖鞏固咄山寺彌陀佛碑

3 *N° 9571– 72*

4 *Dựng tại chùa Đót Sơn, xã Kinh Lương, tổng Kinh Lương, huyện Tiên Lãng, tỉnh Kiến An. Nay là chùa Đót Sơn, thuộc xã Cấp Tiến.*

5 夫咄山寺新明之名寺也創自梁家之始祖歷李陳爲管老子朗開遊後造浮屠廣開寺宇孫散正信與住持長老咄里人亥紫陌肇等再興佛事舊跡增輝然彌陀尊佛亦未之見玄孫仙爲義川伯高文林撰記

6 **PHIÊN ÂM**

HOÀNG ĐỒ CỦNG CỐ ĐÓT SƠN TỰ DI ĐÀ PHẬT BI

Phù! Đót Sơn tự Tân Minh chi danh tự dã. Sáng tự Lương gia chi Thủy tổ, lịch Lý Trần vi Quản lão tử Lãng, Nhàn, Du hậu tạo phù đồ quảng khai tự vũ, tôn Tán, Chính, Tín dữ trụ trì, trưởng lão Đót lý nhân Hợi, Tử, Mạch, Liễn đẳng tái hưng Phật sự. Cựu tích tăng huy

7 **DỊCH NGHĨA:**

NGHIỆP NHÀ BỀN VỮNG. BIA PHẬT DI ĐÀ CHÙA ĐÓT SƠN

Kìa như! Chùa Đót Sơn là ngôi chùa nổi tiếng ở xứ Tân Minh. Chùa bắt đầu xây dựng từ cụ Thủy tổ đời nhà Lương, trải qua đời Lý, đời Trần đến đời con ông Quản lão là các ông Lãng, Nhàn, Du lại dựng tháp mở chùa, đến đời cháu là các ông Tán, Chính, Tín cùng vị thượng lão trụ trì người làng Đót, các ông Hợi, Tử, Mạch, Liễn lại dấy việc thờ Phật. Chùa cũ càng thêm huy hoàng, nhưng tượng đức Di Đà lại chưa có. Đến đời người cháu huyền tôn[1] là ông Tiên, làm quan Huyện thừa ở huyện Đại An, phủ Kiến Hưng. Khi việc quan nhàn tản, cũng không quên kế tục sự nghiệp ông cha. Ông bèn cùng quan Tư lễ Nguyễn Văn Tư, quan Hiệu úy Nguyễn Văn Lự cùng với người làng là Lương Văn Túy, Nguyễn Văn Man xướng suất về

8 1. Huyền tôn: Cháu 5 đời.

Cấu trúc Văn bản "Văn bia Tiên Lãng – Hải Phòng". Văn bản được đánh số thứ tự từ 1 trở đi để dễ tra cứu. Văn bản được trình bày theo thứ tự: nguyên văn chữ Hán, phiên âm, dịch nghĩa và chú thích. Cấu trúc nội dung của bia trong văn bản gồm:

1. Tên bia kèm số thứ tự (chữ Quốc ngữ).
2. Tên bia (chữ Hán).
3. Số hiệu bia lưu trong thư viện.
4. Vị trí địa lý của bia.
5. Nội dung chữ Hán.
6. Phiên âm.
7. Dịch nghĩa.
8. Chú thích (nếu có).

Trong phạm vi sách *"Văn bia Tiên Lãng – Hải Phòng"* có 181 văn bản[3] Lưu ý:

- Các văn bia có ký hiệu từ 53154 trở đi sẽ không có ảnh tương ứng.
- Mỗi văn bản có nội dung từ 1 đến 4 mặt bia.

Như vậy có 154/181 văn bản có ảnh văn bia tương ứng, tổng cộng 305 mặt bia.

3.1.2. Cấu trúc Văn bia

Cấu trúc văn bia gồm:

1. Địa điểm rập bia.
2. Số hiệu bia lưu trong thư viện.
3. Số hiệu mặt bia: [library_id_pad]-[TF][idx]. Ví dụ: bia Nº 9571–72 có 2 mặt ⇒ 09571-21, 09572-22.

Trước khi có những hình ảnh rõ nét hơn của các văn bia trên Website Vietnamica, chúng tôi đã từng thử nghiệm hình ảnh được trích ra từ tập "TỔNG TẬP THÁC BẢN VĂN KHẮC HÁN NÔM", ban chỉ đạo công trình: Trịnh Khắc Mạnh, Nguyễn Văn Nguyên, Philippe Papin [9]. Trong đó có chú thích rõ hơn về cấu trúc văn bia như sau:

④ Một ký hiệu khác nằm trên chính ảnh bia, ở góc phía trên, bên phải (vị trí số 4), được viết lên một thanh nam châm đặt trên bản rập lúc chụp ảnh. Cùng với ký hiệu của thư viện, hẳn ký hiệu này sẽ là thừa nếu nó không kiêm vai trò là một thước đo tỷ lệ : khung chữ nhật bao quanh ký hiệu này gồm có các vạch nhỏ cách nhau 1 cm, tổng cộng chiều dài khung là 5 cm. Dựa vào thước quy chiếu này chúng ta có thể suy ra kích thước nguyên bản của bản rập, để từ đó hình dung ra kích thước của chính tấm bia.

① Cột chữ Hán nằm ở lề phải (vị trí số 1) cho biết địa điểm rập bia (tỉnh, huyện, tổng, làng, và có thể cả thôn xóm, di tích dựng bia), số thứ tự của từng mặt so với tổng số mặt bia (bởi vì một số văn bia được khắc trên cột có 8 hay thậm chí 16 mặt, những bia có 2 mặt chính đôi khi có cả những câu được khắc trên các cạnh của bia). Vì các dòng chữ ở lề thác bản này do những người rập bia của EFEO ghi chú, nên địa danh và địa giới hành chính đều từ đầu thế kỷ XX. Sở dĩ chúng tôi giữ lại những ghi chú này, mặc dù chúng không thuộc nội dung văn khắc, là vì những thông tin này giúp chúng ta định vị một tấm bia, khi trong nguyên bản không cung cấp một thông tin nào về địa điểm dựng bia, hoặc nếu có nhưng lại là những địa danh hay tên những đình chùa đền miếu ngày nay không còn nữa và do đó chúng ta khó có thể xác định được.

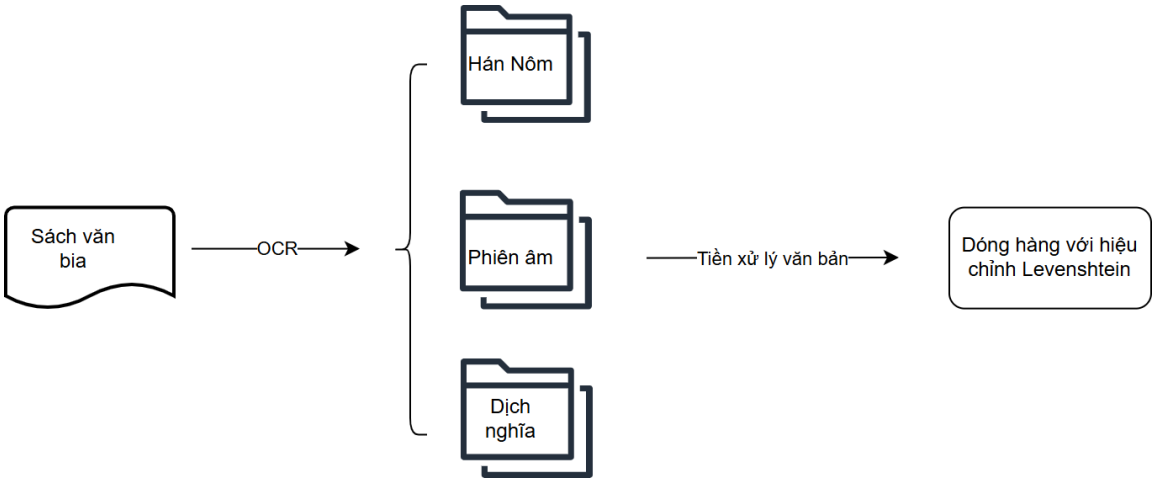
② Ký hiệu của thác bản được thể hiện ở dòng dưới cùng, phía bên phải (vị trí số 2). Đôi khi bên cạnh số ký hiệu có thể có biểu tượng hình đồng hồ cát (⧖) kèm theo. Hình này lưu ý rằng văn bản văn khắc có nghi vấn ngụy tạo niên đại (vấn đề sẽ được chú thích và điều chỉnh lại trong bộ danh mục).

③ Ký hiệu được thể hiện trong ngoặc đơn, phía dưới, bên trái của ảnh (vị trí số 3), là ký hiệu của các mặt thác bản khác có liên quan với thác bản hiện tại, thường là những thác bản của các mặt khác nhau trên cùng một tấm bia. Việc ghi chú những ký hiệu kèm theo này cho phép biết được tấm bia được rập có bao nhiêu mặt, để từ đó giúp người đọc nghiên cứu các thác bản riêng rẽ trong tổng thể của một vật thể văn khắc (bởi đôi khi vì một lý do nào đó, chúng được đánh số ký hiệu cách nhau rất xa). Việc xác định thứ tự trước sau để từ đó tìm ra trình tự đọc các mặt thác bản này rất khó, thậm chí trong nhiều trường hợp hầu như không thể làm được. Vì vậy chúng tôi thống nhất trình bày các ký hiệu này theo thứ tự từ nhỏ đến lớn. Một trường hợp khác, ký hiệu của những thác bản trùng (tức một hay nhiều thác bản khác cùng được rập từ chính mặt bia này, nhưng không mang cùng ký hiệu) cũng được ghi vào đây với một dấu bằng (=) ở phía trước. Dấu bằng này còn được sử dụng để thể hiện trang ảnh hiện tại là phần phóng đại của một thác bản gốc in ở trang trước, dĩ nhiên là cùng ký hiệu.

Hình 1: Cấu trúc bia điển hình

3.2. Phương pháp

3.2.1. Xây dựng ngữ liệu song song Hán Nôm– Phiên âm – Dịch nghĩa



Hình 2: Quy trình xây dựng ngữ liệu Hán Nôm– Phiên âm – Dịch nghĩa

Đối với Sách văn bia Tiên Lãng-Hải Phòng, chúng tôi sử dụng Paddle OCR [10] để rút trích chữ từ file PDF, loại bỏ các dòng header, footer, sau đó dựa vào cấu trúc văn bản để tách chữ tự động ra 3 folder: file text Hán Nôm – file text phiên âm – file text dịch nghĩa.

Để đảm dữ liệu tách chữ từ file PDF sau khi OCR là đúng, nhóm thực hiện kiểm tra và phát hiện có kết quả bị lỗi do file gốc PDF không chỉ chứa chữ mà còn chứa hình ảnh, lỗi unicode... làm cho file Hán Nôm có lẫn chữ Quốc ngữ và ngược lại. Nhóm đã thực hiện chuẩn hóa cho các dữ liệu này như sau:

- Loại bỏ khoảng trắng thừa, thêm khoảng trắng bị thiếu để câu có khoảng cách chuẩn giữa các chữ. Loại bỏ các token OCR sai định dạng và các ký hiệu không chuẩn.
- Bỏ dấu câu, gộp dòng, tách câu.
- Bỏ tên bia (phát hiện bằng cách lấy ký tự IN HOA)
- Bỏ khung [...]

Sau khi thực hiện OCR và công đoạn tiền xử lý văn bản, nhóm thu được dataset bao gồm 3 folder chính: Hán Nôm, phiên âm và dịch nghĩa của từng văn bia4.

Dóng hàng câu với hiệu chỉnh Levenshtein

Trong thực tế, số lượng ký tự Hán Nôm trong một câu có thể không khớp hoàn toàn với số âm tiết trong phiên âm, do có chênh lệch phong cách phiên âm hoặc lỗi OCR.

Để xử lý điều này, chúng tôi áp dụng một phương pháp dóng hàng linh hoạt: Trước hết ghép toàn bộ các dòng Hán Nôm (kèm thống kê số ký tự từng dòng) thành một chuỗi ký tự, trong khi phía Quốc Ngữ là danh sách từ đã tiền xử lý. Ở đây chúng tôi căn chỉnh toàn cục

bằng lập trình động kiểu Levenshtein, trong đó chi phí "thay thế" bằng 0 nếu ký tự Hán Nôm và từ Việt được coi là tương thích theo từ điển chuyển tự và nhóm 20 từ tương tự về hình dạng, ngược lại là 1; chèn hoặc xóa có chi phí là 1. Truy vết ma trận tối ưu cho ra hai dãy đã căn chính, sau đó cắt trở lại theo số ký tự gốc của từng dòng để khôi phục ranh giới dòng và gắn với bounding box tương ứng. Cách làm này vừa có thể xử lý được các sai số của OCR (các chữ sai hoặc bị biến thể ký tự) vừa bảo toàn cấu trúc dòng trong các bước huấn luyện và đánh giá.

Nếu một ký tự từ phía Hán Nôm không có đối ứng phù hợp ở phía phiên âm, chúng tôi chèn ký hiệu "_" là chỗ trống để duy trì cấu trúc dóng hàng 1-1 giữa hai chuỗi, ngay cả khi dứ dữ liệu thiếu hoặc nhiễu. Nhờ đó, mỗi dòng trong corpus giữa được độ thẳng hàng tuyệt đối, hỗ trợ các tác vụ ánh xạ ký tự-âm tiết, chuyển tự tự động và phân tích tài liệu cổ.

Trong quá trình dóng hàng tự động, nhóm nhận thấy một thách thức lớn khi xử lý các câu quá dài. Các mô hình dóng hàng câu (chẳng hạn như các mô hình dựa trên kiến trúc BERT để so sánh độ tương đồng ngữ nghĩa) thường hoạt động kém hiệu quả với những câu có độ dài lớn và cấu trúc phức tạp. Khi một câu Hán và câu dịch nghĩa tương ứng đều rất dài, mô hình có xu hướng không thể xác định chính xác điểm tương đồng, dẫn đến kết quả dóng hàng bị sai.

Để giải quyết vấn đề này, nhóm đã áp dụng một giải pháp xử lý theo các bước sau:

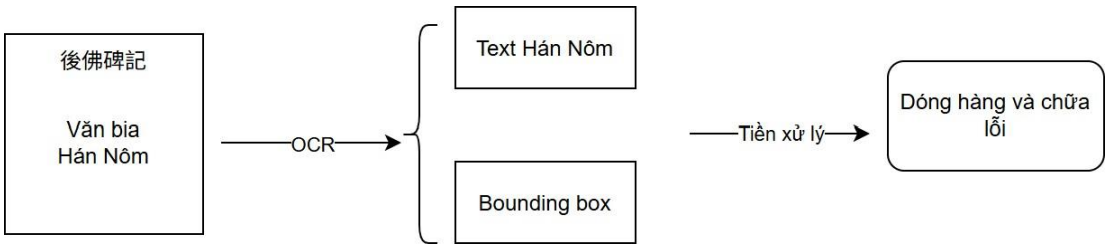
- Cắt câu thành các câu nhỏ: Trước khi đưa vào mô hình dóng hàng, các câu dài trong cả văn bản Hán Nôm, phiên âm và dịch nghĩa sẽ được tự động cắt thành các mệnh đề hoặc cụm từ ngắn hơn. Nhóm sử dụng dấu (,) làm điểm ngắt.

- Dóng hàng các câu nhỏ đã cắt: Các cắt ngắn sau khi được cắt ra sẽ được đưa vào mô hình dóng hàng. Vì độ dài đã giảm đáng kể, mô hình có thể so sánh và tìm ra các cặp câu nhỏ tương ứng với độ chính xác cao hơn.

- Ghép các câu nhỏ thành câu hoàn chỉnh: Sau khi các cặp câu nhỏ đã được dóng hàng thành công, một bước xử lý cuối cùng được thực hiện để ghép chúng lại thành các câu dài hoàn chỉnh như ban đầu.

Phương pháp này giúp đảm bảo rằng ngay cả những câu văn phức tạp nhất cũng được xử lý và dóng hàng một cách chính xác.

3.2.2. Xây dựng ngữ liệu OCR Bounding box – Hán Nôm - Phiên âm



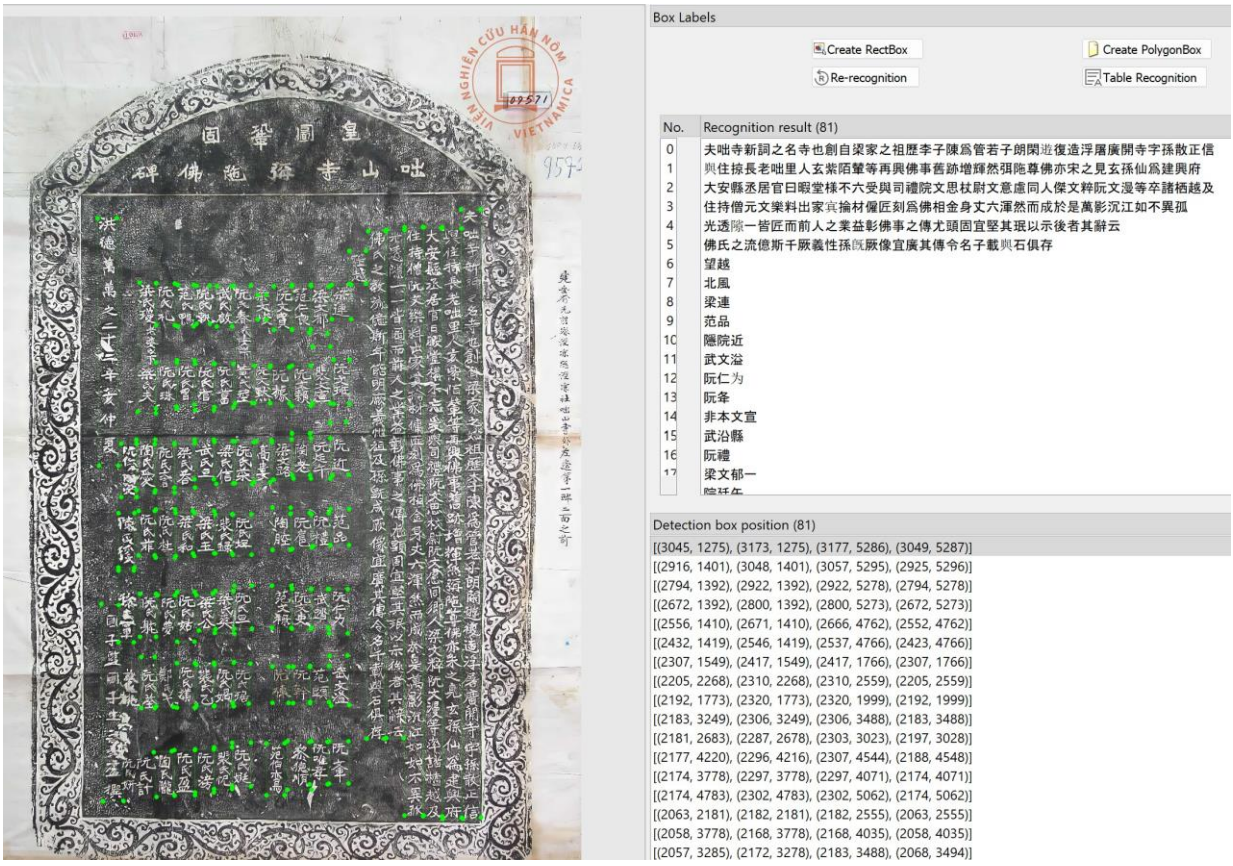
Hình 3: Quy trình xây dựng ngữ liệu OCR Bounding box – Hán Nôm - Phiên âm

Đối với 305 hình ảnh mặt bia thu thập từ Website Vietnamica - Cơ sở dữ liệu Thác bản Văn bia Hán Nôm – Viện Hán Nôm & EFEO [2], chúng tôi sử dụng công cụ KimHanNom [7] để thực hiện OCR kết hợp với gán nhãn bounding box tự động cho các ký tự Hán Nôm.

KimHanNom là công cụ OCR cho phép trích xuất đồng thời cả thông tin văn bản và tọa độ bouding box của từng ký tự hoặc cụm ký tự tự động, giúp tiết kiệm đáng kể thời gian so với quy trình tách OCR - gán nhãn thủ công. Tuy nhiên trong quá trình OCR công cụ gặp 2 vấn đề sau:

- Tự động sinh ra nhiều box trống.
- Không nhận dạng được các box ngắn trong một hàng dọc mà chỉ nhận dạng được box dài.

Chúng tôi đã sử dụng công cụ Paddle OCR [10] để làm sạch, xóa bỏ box trống để tránh gây nhiễu khi dóng hàng, cũng như chia lại box cho các cột chưa chia.



Hình 4: Hình ảnh OCR văn bia

Song song với bước làm sạch/chuẩn hoá bbox, chúng tôi áp dụng **PaddleOCR** trực tiếp lên PDF sách *Văn bia Tiên Lãng – Hải Phòng* để rút trích chuỗi Hán Nôm "tham chiếu" đã

chuẩn hoá (NFC). Mỗi bbox từ ảnh được đối sánh với đoạn tham chiếu này bằng độ tương đồng ký tự chuẩn hoá (mô phỏng Levenshtein), qua đó gán nhãn `note` cho từng dòng:

- **correct**: độ tương đồng ≥ 60% (OCR khớp đủ tin cậy).
- **wrong**: độ tương đồng < 60% hoặc OCR trống; khi đó dùng *phiên âm* làm cầu nối để truy hồi đoạn Hán Nôm phù hợp.

Kết quả tổng hợp: toàn bộ có **4.344** bbox; trong đó **3.214** bbox đạt `correct` (≥ 60%) và **1.130** bbox cần xem xét/hiệu chỉnh. Sau bước lọc chất lượng và kiểm tra tính nhất quán, bộ dữ liệu giữ lại **124/181** văn bản và **211/305** ảnh cho các thí nghiệm huấn luyện–đánh giá.

4. THỰC NGHIỆM VÀ KẾT QUẢ

4.1.Thực nghiệm 1

Sau quá trình thực hiện các phương pháp đã đề ra, nhóm đã thu được những kết quả cụ thể cho từng loại ngữ liệu như sau:

4.1.1. Ngữ liệu Hán - Phiên âm - Dịch nghĩa

Từ 181 văn bản bia trong sách "Văn bia Tiên Lãng - Hải Phòng", sau khi sàng lọc và xử lý, nhóm đã tiến hành xây dựng thành công ngữ liệu song song ba thành phần.

Số lượng: Hoàn thành việc dóng hàng cho toàn bộ 181 văn bản, tạo ra một corpus hoàn chỉnh bao gồm 4158 cặp câu Hán Nôm - Phiên âm - Dịch nghĩa tương ứng.

Định dạng: Ngữ liệu được lưu dưới dạng bảng, với mỗi dòng là một đơn vị câu đã được dóng hàng. Cấu trúc bao gồm các cột:

Line_ID, Hán, Phiên âm, và Dịch nghĩa.

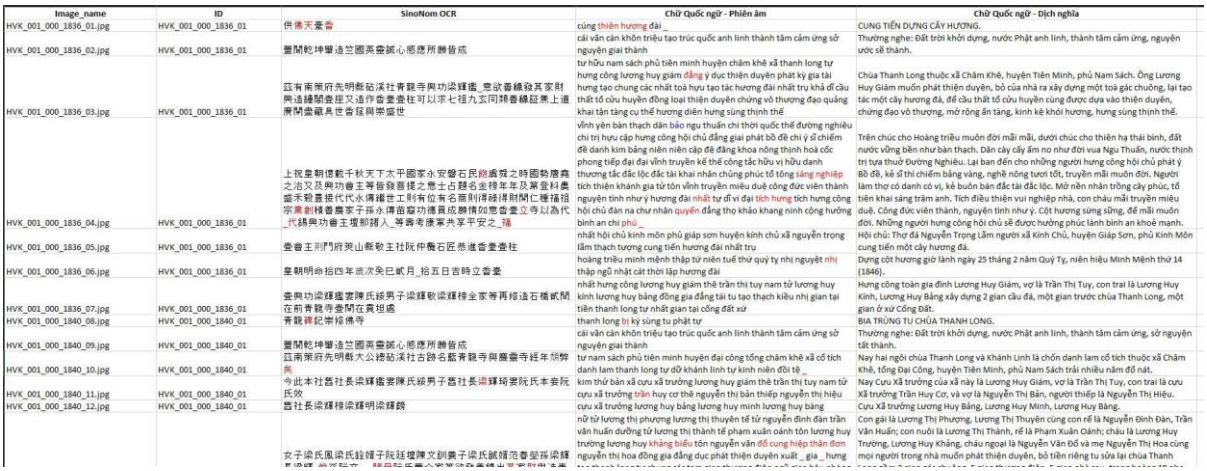
Chất lượng:

Văn bản đã được tiền xử lý kỹ lưỡng để loại bỏ các ký tự nhiễu, lỗi định dạng và chuẩn hóa khoảng trắng.

Áp dụng thuật toán Needleman-Wunsch và cơ chế chèn ký tự đệm ("_") đã giúp đảm bảo sự thẳng hàng chính xác 1-1 giữa ký tự Hán và âm tiết phiên âm, ngay cả trong các trường hợp có sự khác biệt về số lượng từ do phong cách dịch.

Đã giải quyết vấn đề kết quả dóng không chuẩn do độ dài câu quá dài bằng cách chia nhỏ để dóng hàng sau đó ghép lại nội dung hoàn chỉnh tương ứng.

4.1.2. Kết quả ngữ liệu 1[5]



Image_name	ID	SinoNom OCR	Chữ Quốc ngữ - Phiên âm	Chữ Quốc ngữ - Dịch nghĩa
HVK_001_000_1836_01.jpg	HVK_001_000_1836_01	供佛天臺香	cúng thiên hương đài _	CUNG TIỀN DỰNG CÂY HƯƠNG.
HVK_001_000_1836_02.jpg	HVK_001_000_1836_01	靈開乾坤擘造竺國奕壘誠心感應所願皆成	cái vãn cán khôn triều tạo trúc quốc anh linh thành tâm cảm ứng sở nguyện giai thành	Thường nghe: Đất trời khởi dựng, nước Phật anh linh, thành tâm cảm ứng, nguyện ước sẽ thành.
HVK_001_000_1836_03.jpg	HVK_001_000_1836_01	茲有南策府先明縣茹溪社青龍寺興功梁輝鑑_意欲善緣發其家財興造鐘閣壹座又造作香壇香柱可以求七祖九玄同類善緣証應上遍廣開壇廠具世香廷興榮盛世	tự hữu nam sách phủ tiên minh huyện chăm khê xã thanh long tự hưng công lương huy giám đẳng ý dục thiện duyên phát kỳ gia tài hưng tạo chung các nhất toà hựu tạo tác hương đài nhất trụ khả dĩ cầu thất tổ cửu huyền đồng loại thiện duyên chứng vô thượng đạo quảng khai tận tàng cụ thể hương diên hưng sùng thịnh thế	Chùa Thanh Long thuộc xã Châm Khê, huyện Tiên Minh, phủ Nam Sách. Ông Lương Huy Giám muốn phát thiện duyên, bỏ của nhà ra xây dựng một toà gác chuông, lại tạo tác một cây hương đá, để cầu thất tổ cửu huyền cùng được dựa vào thiện duyên, chứng đạo vô thượng, mở rộng ấn tàng, kính kẻ khói hương, hưng sùng thịnh thế.
HVK_001_000_1836_04.jpg	HVK_001_000_1836_01	上祝皇朝德載千秋天下太平國家永安盤石民踴盛世之時國勢唐堯之治又及興功會主等皆發善提之意士占題名金稱年年及第登科農盛禾登昌提代代永傳繼世工則有位有名庶則得財閒仁種福福祖宗業劇積善農家子孫永傳面臨功德員成願情如意壹座立尊以為代_代綿興功會主增那諸人_等壽考康寧共享平安之_福	vĩnh yên bàn thạch dân bảo ngu thuần chi thời quốc thế đường nghiêu chi trị hựu cập hưng công hội chủ đẳng giai phát bồ đề chí ý sĩ chiêm đề danh kim bảng niên niên cập đệ đăng khoa nông thịnh hoà cốc phong tiếp đại đại vĩnh truyền kế thế công tắc hữu vị hữu danh thương tắc đắc lộc đắc tài khai nhân chúng phúc tổ tông sáng nghiệp tích thiện khánh gia tử tôn vĩnh truyền miêu duệ công đức viên thành nguyện tình như ý hương đài nhất tự đĩ vi đại tích hưng tích hưng công hội chủ dân na chư nhân quyến đẳng thọ khảo khang ninh cộng hưởng bình an chi phú _	Trên chúc cho Hoàng triều muôn đời mãi mãi, dưới chúc cho thiên hạ thái bình, đất nước vững bền như bàn thạch. Dân cày cấy ấm no như đời vua Ngu Thuấn, nước thịnh trị tựa thuở Đường Nghiêu. Lại ban đến cho những người hưng công hội chủ phát ý Bồ đề, kẻ sĩ thi chiếm bảng vàng, nghề nông tươi tốt, truyền mãi muôn đời. Người làm thợ có danh có vị, kẻ buôn bán đắc tài đắc lộc. Mở nền nhân trồng cây phúc, tổ tiên khai sáng trăm anh. Tích điều thiện vui nghiệp nhà, con cháu mãi truyền miêu duệ. Công đức viên thành, nguyện tình như ý. Cột hương sùng sững, để mãi muôn đời. Những người hưng công hội chủ sẽ được hưởng phúc lành bình an khoẻ mạnh.
HVK_001_000_1836_05.jpg	HVK_001_000_1836_01	壹會主則門府黃山縣敬主社阮仲喬石匠恭造香壹壹柱	nhất hội chủ kính môn phủ giáp sơn huyện kính chủ xã nguyễn trọng lâm thạch tượng cung tiến hương đài nhất trụ	Hội chủ: Thợ đá Nguyễn Trọng Lâm người xã Kính Chủ, huyện Giáp Sơn, phủ Kinh Môn cung tiến một cây hương đá.
HVK_001_000_1836_06.jpg	HVK_001_000_1836_01	皇朝明命拾四年歲次癸巳貳月_拾五日吉時立壹壹	hoàng triều minh mệnh thập tứ niên tuế thứ quý tỵ nhị nguyệt nhị thập ngũ nhật cát thời lập hương đài	Dựng cột hương giờ lành ngày 25 tháng 2 năm Quý Tỵ, niên hiệu Minh Mệnh thứ 14 (1846).
HVK_001_000_1836_07.jpg	HVK_001_000_1836_01	壹興功梁輝鑑壹陳氏綏男子梁輝敬梁輝穆全家等再修造石幡貳閒在前青龍寺壹閒在貴坦處	nhất hưng công lương huy giám thê trần thị tuy nam tử lương huy kính lương huy bảng đồng gia đẳng tài tu tạo thạch kiểu nhị gian tại tiền thanh long tự nhất gian tại cống đất xứ	Hưng công toàn gia đình Lương Huy Giám, vợ là Trần Thị Tuy, con trai là Lương Huy Kính, Lương Huy Bảng xây dựng 2 gian cầu đá, một gian trước chùa Thanh Long, một gian ở xứ Cống Đất.
HVK_001_000_1840_08.jpg	HVK_001_000_1840_01	青龍禪記崇福佛寺	thanh long bi ký sùng tu phật tự	BIA TRÙNG TU CHÙA THANH LONG.
HVK_001_000_1840_09.jpg	HVK_001_000_1840_01	靈開乾坤擘造竺國奕壘誠心感應所願皆成	cái vãn cán khôn triều tạo trúc quốc anh linh thành tâm cảm ứng sở nguyện giai thành	Thường nghe: Đất trời khởi dựng, nước Phật anh linh, thành tâm cảm ứng, sở nguyện tất thành.
HVK_001_000_1840_10.jpg	HVK_001_000_1840_01	茲南策府先明縣太公總茹溪社古跡名藍青龍寺與廢壹寺經年頹弊	tự nam sách phủ tiên minh huyện đại công tổng chăm khê xã cổ tích danh lam thanh long tự dữ khánh linh tự kính niên đối tệ _	Nay hai ngôi chùa Thanh Long và Khánh Linh là chốn danh lam cổ tích thuộc xã Châm Khê, tổng Đại Công, huyện Tiên Minh, phủ Nam Sách trải nhiều năm đổ nát.
HVK_001_000_1840_11.jpg	HVK_001_000_1840_01	今此本社舊社長梁輝鑑壹陳氏綏男子舊社長梁輝琦要院氏本安院氏效	kim thử bản xã cựu xã trưởng lương huy giám thê trần thị tuy nam tử cựu xã trưởng trần huy cơ thê nguyễn thị bản thiếp nguyễn thị hiệu	Nay Cựu Xã trưởng của xã này là Lương Huy Giám, vợ là Trần Thị Tuy, con trai là cựu Xã trưởng Trần Huy Cơ, và vợ là Nguyễn Thị Bản, người thiếp là Nguyễn Thị Hiệu.
HVK_001_000_1840_12.jpg	HVK_001_000_1840_01	舊社長梁輝穆梁輝明梁輝錡	cựu xã trưởng lương huy bảng lương huy minh lương huy bảng	Cựu Xã trưởng Lương Huy Bảng, Lương Huy Minh, Lương Huy Bảng.
		女子梁氏鳳梁氏錢輝子院廷壇陳文銅黃子梁氏賦婚范春壁孫梁輝	nữ tử lương thị phượng lương thị thuyên tế tử nguyễn đình đàn trần văn huấn dưỡng tử lương thị thành tế phạm xuân oánh tôn lương huy trưởng lương huy khảng biểu tôn nguyễn văn đỗ cung hiệp thân đơm nguyễn thị hoa đồng gia đẳng dục phát thiện duyên xuất _gia _ hưng	Con gái là Lương Thị Phượng, Lương Thị Thuyên cùng con rể là Nguyễn Đình Đàn, Trần Văn Huấn; con nuôi là Lương Thị Thành, rể là Phạm Xuân Oánh; cháu là Lương Huy Trưởng, Lương Huy Khảng, cháu ngoại là Nguyễn Văn Đỗ và mẹ Nguyễn Thị Hoa cùng mọi người trong nhà muốn phát thiện duyên, bỏ tiền riêng tu sửa lại chùa Thanh

Hình 5: Kết quả ngữ liệu Hán Nôm - Phiên âm - Dịch nghĩa

4.2. Thực nghiệm 2

4.2.1. Ngữ liệu Bounding box - Hán Nôm - Phiên âm

Đối với bài toán nhận dạng và trích xuất thông tin từ hình ảnh, nhóm đã xây dựng được một bộ dữ liệu chuyên biệt cho việc huấn luyện các mô hình OCR.

Số lượng:

Trích xuất thành công 305 mặt bia tương ứng với các văn bản có hình ảnh hợp lệ.

Định dạng: Ngữ liệu được cấu trúc dưới dạng bảng, mỗi dòng tương ứng với một khung văn bản (bounding box) được nhận dạng trên ảnh. Các cột bao gồm:

Img_Box_ID, Img_Box_Coordinate, OCR_Hán (văn bản Hán từ OCR), Hán (văn bản Hán đã được chỉnh sửa), và Phiên âm tương ứng.

4.2.2. Kết quả ngữ liệu 2[6]

Image_name	Image Box	SinoNom char	padding	Phiên âm
08456-21	[[2705, 1554], [2821, 1554], [2808, 5170], [2691, 5170]]	南策府光明縣中陵社官員黎得朝阮智文阮勇勝黎富豪范曰都	南策府先明縣中陵社官員黎得朝阮智文阮勇勝黎富豪范曰都	nam sách phủ tiên minh huyện trung lăng xã quan viên lê đức triều n
08456-21	[[2552, 1558], [2664, 1559], [2641, 5166], [2529, 5165]]	陳文署阮佐國范文榜范俊賢范富春鄧仁政阮德文范丁癸范有	陳文略阮佐國范文榜范俊賢范富春鄧仁政阮德文范丁癸范有	trần văn lược nguyễn tá quốc phạm văn bàng phạm tuấn hiền phạm
08456-21	[[2376, 1563], [2492, 1563], [2483, 4773], [2367, 4772]]	*范克諧全社上下巨小等爲其立端言誓約文事原本社有	范克諧金社上下巨小等為其立端言誓約文事原本社有	trị phạm khắc hài toàn xã thượng hạ cự tiểu đẳng vi kỳ lập đoan ngô
08456-21	[[2200, 1414], [2334, 1414], [2330, 5183], [2196, 5183]]	佛寺自南國以來素多靈應茲見本社阮氏倫號慈嚴仰望欲全之	佛寺自南國以來素多靈應茲見本社阮氏倫號慈嚴仰望欲全之	phật tự tự nam quốc dĩ lai tố đa linh ứng tự kiến bản xã nguyễn thị kí
08456-21	[[2047, 1576], [2163, 1577], [2149, 5170], [2033, 5170]]	切尤遠慮之謀晨稱望想於親報答酵儞於命私雲家財錢貳拾五	切尤遠慮之謀晨昏望想於親報答酵儞於命私發家財錢貳拾五	thiết sưu viễn lự chi mưu thần hôn vọng tưởng ư thân báo đáp trữu t
08456-21	[[1885, 1572], [2001, 1572], [1996, 5183], [1880, 5183]]	實錢井田壹畝付與本社官員上下巨小等領取以成所特至於上	實錢井田壹畝付與本社官員上下巨小等領取以成所特至於上	quán tiền tỉnh điền nhất mẫu phó dữ bản xã quan viên thượng hạ cự
08456-21	[[1722, 1558], [1839, 1558], [1839, 5170], [1722, 5170]]	專之時本社遷終之禮係齋醮具盤奉事反每年忌臘冥恩其本社	壽之時本社送終之禮係齋醮具盤奉事及每年忌臘義恩其本社	thọ chi thời bản xã tống chung chi lễ hệ trai tiếu cụ bàn phụng sự cập
08456-21	[[1551, 1559], [1681, 1558], [1690, 5179], [1560, 5179]]	盤備禮物盤願宜用禮樂齋祭如儀代代奉事一悉恭敬不敢忘恩	本社整備禮物盤饌宜用禮樂祭祀如儀代代奉事一心共敬不敢忘	chỉnh bị lễ vật bàn soạn nghi dụng lễ lạc trần tế như nghi đại đại phụ
08456-21	[[1407, 1562], [1523, 1564], [1514, 2517], [1397, 2516]]	背義遵依端內願	背義遵依端內願	bội nghĩa tuân y đoan nội nguyện
08456-21	[[1240, 1432], [1356, 1432], [1379, 5192], [1263, 5193]]	皇天扶護本社老少均得平安若本社某員人生心*按阮氏倫號慈	皇天扶護本社老少均得平安若本社員人生心瀧悖阮氏倫號慈嚴	hoàng thiên phù hộ bản xã lão thiếu quân đắc bình an nhược bản xã
08456-21	[[1082, 1576], [1198, 1576], [1198, 4434], [1082, 4434]]	留廢忌臘背義忘恩不勤奉事違此端言誓約付與	留廢忌臘背義忘恩不勤奉事違此端言誓約付付	lưu phế ky lạp bội nghĩa vong ân bất cần phụng sự vi thử đoan ngôn
08456-21	[[888, 1429], [1036, 1427], [1077, 5200], [930, 5202]]	皇天后土刃諸靈神照鑒為榮並實端言以盡奉事之敬誠以遵流	與皇天后土及諸靈神照鑒遠榮並實端言以盡奉事之敬誠以遵流	hoàng thiên hậu thổ cập chư linh thần chiếu giám vi ương tịnh thực c
08456-21	[[744, 1576], [874, 1576], [874, 2241], [744, 2241]]	之萬代茲端	之萬代茲端	chi vạn đại tư đoan
08456-21	[[555, 1442], [689, 1441], [716, 5151], [582, 5152]]	福泰柒年叁月貳拾叁日立端言誓約社長陳盛強記	福泰柒年參月貳拾參日立端言誓約社長陳盛強記	phúc thái thất niên tam nguyệt nhị thập tam nhật lập đoan ngôn thệ u

Hình 6: Ngữ liệu Bounding box - Hán Nôm - Phiên âm Ngữ liệu Bounding box - Hán Nôm - Phiên âm gồm các thành phần sau:

- Image_name: Số của bia ứng với số hiệu trong thư viện
- Image Box: Vị trí tọa độ của khung chữ trong tấm ảnh
- SinoNom char: chữ Hán Nôm lấy tương ứng với khung chữ sau khi OCR
- Padding: chữ Hán Nôm trong văn bản gốc
- Quốc ngữ (Phiên âm): phần phiên âm chữ Hán Nôm tương ứng văn bản gốc

5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Qua quá trình nghiên cứu và triển khai, nhóm đã xây dựng thành công hai bộ corpus quan trọng cho lĩnh vực xử lý văn bia Hán Nôm, bao gồm corpus căn chỉnh dòng (Hán Nôm – Phiên âm – Dịch nghĩa) và corpus gắn nhãn cấp ký tự (OCR bounding box – Hán Nôm – Phiên âm); Thiết kế quy trình xây ngữ liệu từ trích xuất dữ liệu ảnh và PDF, nhận dạng ký tự bằng OCR, tạo bounding box, đến việc dóng hàng với phiên âm và dịch nghĩa. Kết quả thực nghiệm cho thấy các công cụ được sử dụng có khả năng xử lý hiệu quả dữ liệu từ cả nguồn sách scan lẫn ảnh trực tuyến của Vietnamica, đặc biệt khi áp dụng công cụ KimHanNom cho tác vụ OCR, giúp cải thiện đáng kể độ chính xác và tính nhất quán của bounding box vượt trội. Ngữ liệu thu được này sẽ là golden standard cho việc xây dựng mô hình nhận dạng văn bia Hán Nôm.

Trong tương lai, chúng tôi sẽ xem văn bia như một đối tượng nghiên cứu trung tâm, đặt nó trong mối quan hệ với không gian, thời gian, vùng miền và thể loại. Cách tiếp cận này cho phép chúng tôi không chỉ phân tích văn bia như những văn bản biệt lập, mà còn nhìn nhận chúng trong mạng lưới tri thức lịch sử – văn hóa. Dữ liệu nghiên cứu cũng sẽ được mở rộng sang nhiều địa phương khác thay vì chỉ giới hạn ở Tiên Lãng – Hải Phòng, qua đó phản ánh sự đa dạng về bối cảnh địa lý và lịch sử của văn bia Việt Nam. Thông qua việc đối chiếu niên đại, vùng miền và đặc trưng thể loại, chúng tôi kỳ vọng có thể hiểu đúng ngữ nghĩa nội dung văn bia trong từng bối cảnh cụ thể, từ đó làm sáng tỏ sự khác biệt về văn hóa – xã hội giữa các vùng, sự phát triển của tư tưởng qua các thời kỳ, cũng như vai trò của văn bia trong việc kiến tạo và lưu truyền ký ức lịch sử. Hướng nghiên cứu này không chỉ giúp bổ sung một phương pháp tiếp cận mới trong lĩnh vực Hán Nôm học và văn hóa học, mà còn mở ra khả năng ứng dụng trong nghiên cứu liên ngành, như lịch sử, địa lý văn hóa, và khoa học dữ liệu nhân văn.

Tài liệu tham khảo

[1] "Văn bia Tiên Lãng - Hải Phòng", chủ biên Đỗ Thị Hảo.

[2] Cơ sở dữ liệu Thác bản Văn bia Hán Nôm - Viện Hán Nôm & EFEO https://gallery.vietnamica.online

[3] "Spelling Correction of OCR-Generated Hindi Text Using Word Embedding and Levenshtein Distance", author Srigiri, Shashank and Saha, Sujan Kumar, editor Nath, Vijay and Mandal, J. K.", booktitle "Nanoelectronics, Circuits and Communication Systems,2020, publisher "Springer Singapore".

[4] "Vartani Spellcheck - Automatic Context-Sensitive Spelling Correction of OCR-generated Hindi Text Using BERT and Levenshtein Distance", author Aditya Pal and Abhijit Mustafi,2020. https://arxiv.org/abs/2012.07652

[5] "Precise and efficient text correction using Levenshtein automata, dynamic Web dictionaries and optimized correction models", author Mihov, Stoyan and Koeva, Svetla and Ringlstetter, Christoph and Schulz, Klaus U and Strohmaier, Christian, 2004.

[6] "Levenshtein OCR", author Da, Cheng and Wang, Peng and Yao, Cong, editor Avidan, Shai and Brostow, Gabriel and Ciss, Moustapha, Farinella, Giovanni Maria Hassner, 2022, publisher Springer Nature Switzerland.

[7] Kim Hán Nôm – Automatic Transliteration Tool from Sino-Nôm to Vietnamese Script https://tools.clc.hcmus.edu.vn, 2022.

[8] Lei Liu Min Zhu. 2022. Bertalign: Improved word embedding-based sentence alignment for Chinese–English parallel corpora of literary texts, Digital Scholarship in the Humanities. https://doi.org/10.1093/llc/fqac089 https://github.com/bfsujason/bertalign

[9] "TỔNG TẬP THÁC BẢN VĂN KHẮC HÁN NÔM", ban chỉ đạo công trình: Trịnh Khắc Mạnh, Nguyễn Văn Nguyên, Philippe Papin.

[10] PaddleOCR. *Awesome multilingual OCR toolkits based on PaddlePaddle*. https://github.com/PaddlePaddle/PaddleOCR, 2020.

THÔNG TIN TÁC GIẢ

[1] Trương Thị Vân Anh

Trường Đại học Khoa học Tự nhiên, ĐHQG HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP.HCM; Email:24C01002@student.hcmus.edu.vn

[2] Bùi Hồng Phúc

Trường Đại học Khoa học Tự nhiên, ĐHQG HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP.HCM; Email: duyphuc2425@gmail.com

* Đinh Điền

PGS.TS, Trường Đại học Khoa học Tự nhiên, ĐHQG HCM, 227 Nguyễn Văn Cừ, Phường Chợ Quán, TP.HCM; Email: ddien@fit.hcmus.edu.vn

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO AI TRONG GIẢNG DẠY TIẾNG VIỆT CHO HỌC VIÊN LÀO HỌC TIẾNG VIỆT TẠI HỌC VIỆN AN NINH NHÂN DÂN

Tống Thị Hảo [1]

Tóm tắt: Ứng dụng trí tuệ nhân tạo AI trong dạy học nói chung đang trở thành xu hướng rất phổ biến hiện nay. Tại Học viện An ninh nhân dân, ứng dụng trí tuệ nhân tạo AI vào giảng dạy tiếng Việt cho học viên Lào cũng đang được quan tâm đẩy mạnh. Tuy nhiên, việc giảng dạy ngôn ngữ này đối mặt với nhiều thách thức do sự khác biệt lớn về ngữ âm, ngữ pháp, văn hóa giữa hai quốc gia. Bên cạnh đó, người học thiếu động lực học tập, thiếu tính chủ động trong việc tiếp cận các nguồn tài liệu học tập, thiếu sự đa dạng trong phương pháp dạy và học cũng ảnh hưởng đến chất lượng đào tạo. Ứng dụng trí tuệ nhân tạo AI vào giảng dạy tiếng Việt cho học viên Lào bước đầu đã tháo gỡ được những khó khăn và mang lại hiệu quả vượt trội. Bài viết đề cập đến những thuận lợi, khó khăn và đưa ra một số gợi ý trong việc ứng dụng trí tuệ nhân tạo AI vào giảng dạy tiếng Việt cho học viên Lào tại Học viện An ninh nhân dân.

***Từ khóa**:* Trí tuệ nhân tạo AI, ứng dụng, giáo dục, giảng dạy tiếng Việt, học viên Lào

1. ĐẶT VẤN ĐỀ

Trong thời đại bùng nổ số hóa hiện nay, trí tuệ nhân tạo AI không chỉ nâng cao hiệu quả công việc mà còn mở ra những cơ hội mới cho sự phát triển toàn diện của con người. Với khả năng xử lí số liệu nhanh chóng và chính xác, AI đóng một vai trò vô cùng quan trọng trong sự phát triển của xã hội ở nhiều lĩnh vực khác nhau trong đó có giáo dục. Ứng dụng trí tuệ nhân tạo AI vào giảng dạy tiếng Việt cho học viên Lào tại Học viện an ninh nhân dân đang mở ra một hướng đi mới, tối ưu hóa quá trình học tập và nâng cao chất lượng đào tạo.

2. GIỚI THIỆU VỀ AI VÀ ỨNG DỤNG AI TRONG DẠY HỌC

Khái niệm Trí tuệ nhân tạo AI ra đời vào năm 1956 khi nhà khoa học máy tính John MeCarthy đưa ra ý tưởng về máy móc có khả năng suy nghĩ và thực hiện công việc như con người tại Hội nghị Dartmouth. Ông cho rằng máy móc có thể thực hiện các nhiệm vụ trí tuệ thông qua các thuật toán và phần mềm. Theo Andreas Kaplan, AI là khả năng của máy móc bắt chước hành vi thông minh của con người, bao gồm học hỏi, lập luận, giải quyết vấn đề và hiểu ngôn ngữ. Như vậy, có thể hiểu, AI là lĩnh vực công nghệ khoa học máy tính phát triển hệ thống máy móc có khả năng thay con người thực hiện các yêu cầu trí tuệ của con người. Thực tế cho thấy, AI đã đạt được nhiều thành tựu to lớn trong nhiều lĩnh vực như tài chính, y tế, giáo dục….

Trong bối cảnh công nghệ số ngày càng phát triển mạnh mẽ như hiện nay, Trí tuệ nhân tạo AI đang trở thành một phần không thể thiếu trong ngành giáo dục hiện đại. AI trong giáo dục không chỉ cải thiện chất lượng giảng dạy mà còn giúp các trường học và các tổ chức giáo

[1] Th.S, Học viện An ninh nhân dân, *Email:* tonghao66@gmail.com

dục hoạt động hiệu quả hơn, tập trung vào việc cung cấp kiến thức chính xác và kịp thời cho người học. Các nhà giáo dục đã có nhiều công trình nghiên cứu và ứng dụng AI vào giáo dục, đặc biệt là trong lĩnh vực giảng dạy ngoại ngữ với những đột phá trong cách giảng dạy và học tập. Thực tế, việc dạy học ngoại ngữ truyền thống đã gặp không ít khó khăn về động lực học tập của người học, cách tiếp cận nguồn tài liệu, phương pháp dạy học. Những điều này đã tác động trực tiếp đến hiệu quả học tập, chất lượng giáo dục. Sự phát triển của trí tuệ nhân tạo AI trong những năm gần đây đang dần khắc phục được những khó khăn đó và mở ra những bước tiến mới trong công tác dạy học. AI cung cấp các công cụ và giải pháp sáng tạo, hỗ trợ cá nhân hóa quá trình học tập và mang lại trải nghiệm học tập phong phú hơn cho người học. Ứng dụng AI vào trong dạy học nói chung đang là xu thế hiện nay và sẽ ngày càng phát triển mạnh mẽ hơn nữa. Hiện nay, Châu Á đang dẫn đầu về số lượng các nghiên cứu liên quan đến ảnh hưởng và các ứng dụng của AI trong lĩnh vực giáo dục. AI có thể giúp người học tăng cường thực hành ngôn ngữ bên ngoài lớp học, giảm bớt sự lo lắng của người học khi nói, lắng nghe và chỉnh sửa phát âm, dạy từ vựng ngữ pháp qua đố vui, động viên người học khi đạt điểm tối đa hoặc chưa đạt điểm tối đa, đa dạng hóa các tình huống thực tế, học bất cứ khi nào và ở đâu. Đặc biệt, AI còn có thể cá nhân hóa việc học tập của người học. Với khả năng phân tích dữ liệu chính xác, AI có thể theo dõi tiến trình học tập của từng cá nhân và điều chỉnh nội dung giảng dạy sao cho phù hợp với nhu cầu, trình độ và nội dung chương trình của người học. AI có thể phản hồi tức thì và cải thiện kĩ năng cho người học ngoại ngữ. Thông qua các ứng dụng và nền tảng học tập, AI có thể đưa ra những đánh giá và những gợi ý sửa lỗi cho người học. Ví dụ với công cụ Grammarly, AI không chỉ sửa lỗi ngữ pháp mà còn đưa ra những gợi ý để người học hoàn thiện kĩ năng viết. Tương tự như thế ứng dụng Elsa Speak sử dụng công nghệ nhận dạng giọng nói để có thể phân tích đánh giá phát âm của người học, từ đó cung cấp những phản hồi chi tiết để người học hoàn thiện kĩ năng phát âm của mình. Thông qua việc sử dụng chatbot và trợ lí ảo, AI cung cấp môi trường học tương tác cho người học. Ứng dụng Mondly là một ứng dụng học ngoại ngữ sử dụng chatbot để tạo ra những cuộc hội thoại một cách tự nhiên nhất. Công cụ này vừa giúp người học được thực hành ngôn ngữ một cách chủ động vừa tạo ra các tình huống giao tiếp một cách tự nhiên nhất, từ đó giúp người học được thực hành ngôn ngữ, hoàn thiện kĩ năng nghe nói. Như vậy, với việc cung cấp phản hồi ngay lập tức tức đến việc tạo ra môi trường học tập tương tác, AI đang thay đổi cách thức dạy và học ngoại ngữ, giúp người học cải thiện kĩ năng ngôn ngữ hiệu quả hơn.

3. THUẬN LỢI VÀ KHÓ KHĂN TRONG VIỆC ỨNG DỤNG AI VÀO GIẢNG DẠY TIẾNG VIỆT CHO HỌC VIÊN LÀO TẠI HỌC VIỆN AN NINH NHÂN DÂN

3.1. Thuận lợi

Đối với học viên

Cá nhân hóa học tập theo từng học viên. Một trong những lợi ích lớn nhất của AI là khả năng điều chỉnh nội dung học tập theo trình độ và tốc độ của từng cá nhân. Trí tuệ nhân tạo có khả năng thu thập và phân tích dữ liệu học tập của từng học viên, chẳng hạn như tốc độ học tập, phong cách học tập và những điểm mạnh, yếu cụ thể của từng cá nhân. Sau đó, dựa trên

các dữ liệu thu thập, hệ thống có thể đề xuất các chương trình, tài liệu học tập và phương pháp giảng dạy phù hợp nhất cho từng học viên, giúp người học phát huy được tiềm năng cá nhân mà không bị áp lực phải theo kịp tốc độ chung của cả lớp. Điều này rất hữu ích cho học viên Lào vốn có kiến thức nền tảng và tốc độ tiếp thu khác nhau. Nếu một học viên có vốn từ vựng hạn chế và cần học từ cơ bản, trong khi một học viên khác lại gặp khó khăn với ngữ pháp phức tạp thì công cụ AI sẽ cung cấp các bài tập và tài liệu phù hợp với nhu cầu cụ thể của từng học viên. AI có thể tạo ra các bài kiểm tra ngắn để xác định điểm yếu và tập trung vào mảng kiến thức hoặc kĩ năng mà học viên còn hạn chế.

Tăng cường trải nghiệm học tập thông qua công nghệ thực tế ảo và mô phỏng. AI cho phép tích hợp công nghệ thực tế ảo (VR) và mô phỏng (simulation) vào giảng dạy, giúp người học có thể thực hành và trải nghiệm thực tế một cách sinh động và chân thực. Các công cụ mô phỏng với những tình huống chân thực không chỉ giúp người học nắm vững kiến thức mà còn rèn luyện kỹ năng thực tế trong một môi trường an toàn, không có rủi ro. AI có thể giúp học viên luyện tập phát âm một cách chính xác. Thực tế, phát âm luôn là một thách thức lớn đối với người học tiếng Việt. AI có thể đóng vai trò như một giáo viên phát âm 24/7, cung cấp phản hồi tức thì và chính xác mà không cần sự có mặt của giáo viên. Học viên Lào thường gặp khó khăn với các âm như "ng," "tr," "ch" hoặc các dấu thanh điệu. Ứng dụng AI có thể ghi âm giọng nói của học viên, sau đó sử dụng công nghệ nhận dạng giọng nói để chỉ ra các lỗi phát âm cụ thể. Ví dụ, nếu học viên phát âm "cái bàn" thành "cái vàn", AI sẽ lập tức báo lỗi và đưa ra các bài tập lặp lại để luyện tập âm "b" một cách riêng biệt, giúp sinh viên sửa lỗi ngay lập tức.

Học viên có thể học tập mọi lúc mọi nơi. Với sự phổ biến của điện thoại thông minh, học viên Lào có thể học tiếng Việt ở bất cứ đâu và bất cứ lúc nào. Điều này giúp tận dụng tối đa thời gian rảnh rỗi và biến việc học thành một phần của cuộc sống hàng ngày. Ví dụ: Trong thời gian chờ xe buýt hoặc nghỉ giải lao, học viên có thể mở ứng dụng AI để luyện tập từ vựng, nghe các đoạn hội thoại ngắn, hoặc thực hiện các bài tập ngữ pháp. AI sẽ ghi nhớ tiến trình học của học viên và tự động cung cấp nội dung phù hợp cho buổi học tiếp theo. Các chatbot AI cũng giúp cung cấp thông tin và hỗ trợ học viên mọi lúc, đặc biệt hữu ích trong các chương trình học từ xa. Học viên có thể nhận trợ giúp ngay khi cần, chẳng hạn như tìm kiếm tài liệu học tập hoặc giải đáp các câu hỏi thông thường, mà không cần chờ đợi giáo viên.

Đối với giáo viên

AI hỗ trợ giáo viên truyền cảm hứng cho học viên. Học viên Lào nói chung đều có tính kỉ luật tốt nhưng thường e dè, thiếu tự tin, ngại tiếp. Điều này đã làm giảm khả năng tương tác cần thiết đối với một người học ngoại ngữ, từ đó ảnh hưởng đến chất lượng học tập của học viên. AI dùng gương mặt, cử chỉ, điệu bộ, giọng nói của chính của các học viên và chuyển sang ngôn ngữ tiếng Việt mà các em ước ao được sử dụng thông thạo. Trước màn hình là những học

viên Lào nói tiếng Việt sành sỏi như người bản địa đã kích thích nỗ lực, truyền cảm hứng mạnh mẽ về việc học tiếng Việt cho học viên.

Phát triển nội dung học tập thông minh và linh hoạt. Các công cụ AI như Generative sẽ hỗ trợ giáo viên sáng tạo và điều chỉnh nội dung bài giảng nhanh chóng, giúp họ tiết kiệm thời gian chuẩn bị. AI có thể tự động tạo ra các bài kiểm tra, câu đố và tài liệu giảng dạy, từ đó giúp giáo viên đáp ứng nhu cầu học tập đa dạng của học viên và nâng cao hiệu quả học tập tổng thể.

Nâng cao chất lượng giảng dạy và học tập. Với sự hỗ trợ của trí tuệ nhân tạo, giáo viên có thể tiếp cận dữ liệu chi tiết về kết quả học tập của học viên, từ đó điều chỉnh phương pháp giảng dạy sao cho phù hợp nhất với từng lớp học hoặc từng cá nhân. Các dữ liệu về thành tích học tập, yếu tố tâm lý và các trở ngại trong học tập cho phép giáo viên hiểu rõ hơn về điểm mạnh và điểm yếu của từng học viên, từ đó giúp họ cải thiện cách truyền đạt và nội dung bài giảng.

Trí tuệ nhân tạo không chỉ giúp nâng cao hiệu quả học tập của học viên mà còn hỗ trợ đắc lực, giảm gánh nặng cho giáo viên trong quá trình giảng dạy. Giáo viên có thể dùng AI để tự động chấm bài, theo dõi tiến độ học tập và chuẩn bị tài liệu, từ đó dành nhiều thời gian hơn cho việc tương tác trực tiếp và giải quyết các vấn đề phức tạp. Bằng cách tự động hóa các công việc hành chính, như chấm điểm, theo dõi tiến độ học tập và phân tích kết quả học tập của từng học viên, AI giúp giáo viên tiết kiệm thời gian và công sức. Điều này giúp giáo viên dành nhiều thời gian hơn để tập trung vào các công việc quan trọng hơn, như chuẩn bị nội dung bài giảng chất lượng, tương tác, hỗ trợ học viên gặp khó khăn. Hệ thống AI có thể tự động gửi báo cáo chi tiết về tiến độ học tập và những điểm cần cải thiện của học viên đến giáo viên, giúp giáo viên đưa ra các phương án can thiệp kịp thời để cải thiện hiệu quả học tập cho từng học viên.

3.2. Khó khăn, thách thức

Việc học tiếng Việt của sinh viên Lào đang đứng trước những cơ hội lớn nhờ sự phát triển mạnh mẽ của công nghệ AI. Tuy nhiên, để tối ưu hóa hiệu quả, cần nhìn nhận một cách khách quan về thực trạng và những khó khăn, thách thức đang tồn tại để từ đó đưa ra những giải pháp dạy học phù hợp và hiệu quả nhất. Hiện nay, học viên Lào đã tiếp cận và sử dụng một số công cụ AI để hỗ trợ việc học tiếng Việt, chủ yếu là các ứng dụng dịch thuật, từ điển trực tuyến và chatbot. Các công cụ này giúp học viên tra cứu nhanh, dịch các đoạn văn bản cơ bản và giải đáp thắc mắc tức thời. Tuy nhiên, việc sử dụng còn khá tự phát và chưa có hệ thống. Học viên cũng chỉ thường dùng các công cụ miễn phí, phổ biến như Google Translate hay một số ứng dụng dịch thuật khác, mà chưa được hướng dẫn để tận dụng các nền tảng học ngôn ngữ chuyên sâu hơn có tích hợp AI. Họ cũng chưa biết cách đánh giá độ chính xác của các công cụ này, dẫn đến việc đôi khi dịch sai ngữ cảnh hoặc sử dụng từ không phù hợp trong những bối cảnh giao tiếp khác nhau. Một trong những khó khăn thường gặp phải trong quá trình ứng dụng AI vào giảng dạy tiếng Việt cho học viên Lào học tiếng Việt tại Học viện An ninh nhân dân đó là:

Rào cản về công nghệ: Không phải tất cả học viên Lào đều có thiết bị thông minh hoặc đường truyền Internet ổn định để truy cập các ứng dụng AI tiên tiến. Điều này khiến việc học trở nên khó khăn, tạo ra một rào cản lớn trong việc tiếp cận và sử dụng công nghệ.

Thiếu hụt nội dung chuyên sâu: Hầu hết các công cụ AI hiện nay chưa có dữ liệu đủ lớn và chuyên biệt cho người học tiếng Việt là người Lào, đặc biệt là trong việc xử lý các lỗi ngữ pháp đặc trưng của tiếng Việt. Ví dụ, một lỗi phát âm thường gặp của học viên Lào khi nói tiếng Việt là nhầm lẫn giữa âm "s" và "x,", “nh” và “ng”, “b” và “v” nhưng các công cụ AI thông thường có thể không được lập trình để nhận biết và sửa chữa lỗi này một cách hiệu quả.

Thiếu sự tương tác con người: Mặc dù AI rất hiệu quả trong việc cung cấp các bài tập và phản hồi, nhưng nó không thể thay thế hoàn toàn sự tương tác trực tiếp với giáo viên và bạn bè. Tương tác con người đóng vai trò quan trọng trong việc rèn luyện kỹ năng giao tiếp, văn hóa và cảm xúc. Khi một học viên tham gia một buổi học trực tiếp, giáo viên có thể hiểu được cảm xúc, sự bối rối hoặc những khó khăn ẩn sau các lỗi sai của họ. Giáo viên có thể đưa ra lời khuyên, động viên, hoặc điều chỉnh phương pháp dạy học theo cách mà AI không thể làm được. Hơn nữa, AI không thể tạo ra các cuộc trò chuyện tự nhiên và linh hoạt như khi học viên nói chuyện với giáo viên hoặc một người bạn Việt Nam.

Phụ thuộc quá mức vào công cụ: Một số học viên có thể lạm dụng các công cụ dịch thuật, làm giảm khả năng tư duy ngôn ngữ và hình thành thói quen dịch từng từ thay vì hiểu cả câu hoặc có thể hiểu/đoán từ qua ngữ cảnh giao tiếp.

Thiếu kiến thức và tự tin: Một khảo sát cho thấy 60% học viên Lào chưa từng sử dụng công cụ học tập có tích hợp AI. Điều này cho thấy rào cản lớn nhất không chỉ là thiếu công cụ mà còn là thiếu kiến thức và kỹ năng sử dụng.

Thiếu cơ sở hạ tầng: Việc thiếu hụt các khóa học liên quan đến AI, internet ổn định và hỗ trợ kỹ thuật cũng là lý do cản trở việc phát triển và ứng dụng AI.

Văn hóa giáo dục truyền thống: Mặc dù AI mang lại nhiều lợi ích, nhưng việc tích hợp nó vào hệ thống giáo dục vẫn còn chậm do các phương pháp giảng dạy truyền thống vẫn còn phổ biến. Ngoài ra, các nền tảng học tiếng Việt tích hợp AI chất lượng cao thường có chi phí, điều này có thể là rào cản đối với nhiều học viên.

Nhìn chung, việc sử dụng AI của học viên Lào để học tiếng Việt chưa thực sự phổ biến và rộng rãi. Thay vào đó, các phương pháp truyền thống như học trên lớp, giao tiếp với người bản xứ và tham gia các hoạt động ngoại khóa vẫn đóng vai trò chính. Với những phân tích trên, có thể thấy rằng AI mang lại nhiều cơ hội nhưng cũng đi kèm với không ít thách thức đối với học viên Lào trong hành trình chinh phục tiếng Việt.

4. ỨNG DỤNG AI TRONG GIẢNG DẠY TIẾNG VIỆT CHO HỌC VIÊN LÀO TẠI HỌC VIỆN AN NINH NHÂN DÂN

Để tích hợp các công cụ AI vào chương trình học giảng dạy và khuyến khích học viên sử dụng nhằm cá nhân hóa quá trình học tập cách hiệu quả thì cần có những giải pháp cụ thể.

Dưới đây là một số gợi ý khi ứng dụng AI vào giảng dạy tiếng Việt cho học viên Lào tại Học viện An ninh nhân dân.

Sử dụng AI để học từ vựng và ngữ pháp. Đây là giải pháp nền tảng giúp sinh viên xây dựng vốn từ và nắm vững cấu trúc câu. Các công cụ AI có thể tạo ra các bài tập cá nhân hóa dựa trên trình độ và tốc độ học của từng người. Ví dụ, một học viên Lào có thể sử dụng ứng dụng học tiếng Việt tích hợp AI. Ban đầu, ứng dụng sẽ đánh giá trình độ của học viên bằng một bài kiểm tra ngắn. Dựa trên kết quả, AI sẽ tự động điều chỉnh nội dung học. Nếu học viên gặp khó khăn với các từ vựng chuyên ngành như "điều tra" hay "an ninh mạng", AI sẽ tạo ra các bài tập lặp lại và kết hợp chúng với các câu ví dụ thực tế gắn với bối cảnh giao tiếp cụ thể thay vì chỉ đưa ra danh sách từ. Ngoài ra, AI có thể mô phỏng các cuộc hội thoại, giúp học viên luyện tập cách sử dụng từ vựng và ngữ pháp trong bối cảnh chân thực, sinh động. Với đặc tính có thể cá nhân hóa cao, AI không chỉ dạy chung chung mà còn tập trung vào những điểm yếu của người học, giúp học viên tối ưu hóa thời gian học mà không bị lãng phí thời gian học lại những gì đã biết.

Sử dụng AI để luyện phát âm và giao tiếp. Một trong những rào cản lớn nhất khi học ngôn ngữ mới là phát âm. AI có thể đóng vai trò như một gia sư phát âm 24/7. Ví dụ: Một sinh viên cần luyện phát âm từ "nghiên cứu" (âm "ng" và "ưu" khá khó với người Lào). Học viên sẽ đọc từ này vào một ứng dụng AI. Công nghệ nhận dạng giọng nói sẽ phân tích và chỉ ra các lỗi cụ thể, ví dụ như phát âm "nghiên" thành "nhiên". Sau đó, AI sẽ cung cấp các bài tập để luyện tập âm thanh đó một cách riêng biệt, bao gồm cả việc mô phỏng cuộc trò chuyện để học viên lặp lại các câu có chứa từ đó. AI cũng có thể tạo ra những đoạn hội thoại với những bối cảnh khác nhau hết sức chân thực và sinh động để luyện giao tiếp cho người học. Khi luyện phát âm và giao tiếp, học viên nhận được phản hồi ngay lập tức, không cần chờ đợi giáo viên. Bên cạnh đó, AI có thể chỉ ra chính xác âm nào/câu nào bị sai, giúp học viên sửa lỗi hiệu quả hơn.

Sử dụng AI để dịch và giải thích nội dung phức tạp. Khi học viên Lào đọc các tài liệu tiếng Việt, việc hiểu nghĩa của các từ hoặc cụm từ khó có thể gây nản lòng. AI sẽ giúp giải quyết vấn đề này một cách nhanh chóng và chính xác. Ví dụ: Khi một học viên đọc một bài báo tiếng Việt về kinh tế và gặp phải cụm từ "giải ngân vốn đầu tư công", họ có thể sử dụng một công cụ dịch thuật AI tiên tiến (chẳng hạn như Google Translate hoặc các công cụ dịch thuật chuyên sâu hơn). Tuy nhiên, thay vì chỉ dịch đơn thuần, AI có thể đưa ra một định nghĩa ngắn gọn, dễ hiểu hơn trong bối cảnh của bài báo, cùng với các câu ví dụ khác. Điều này sẽ giúp tiết kiệm thời gian bởi học viên không phải tra từ điển từng từ một. AI giúp học viên hiểu ý nghĩa của một từ hoặc cụm từ trong bối cảnh cụ thể, thay vì chỉ là bản dịch thô.

Việc kết hợp cả ba giải pháp trên sẽ giúp học viên Lào tối ưu hóa quá trình học tiếng Việt, biến AI từ một công cụ đơn thuần thành một "trợ giảng" cá nhân hiệu quả, giúp họ tự tin và tiến bộ nhanh hơn. Tuy nhiên, các giải pháp để tối ưu hóa việc ứng dụng AI trong giảng dạy tiếng Việt cho học viên Lào, cần có sự phối hợp từ nhiều bên như:

Đối với các cơ sở đào tạo: Tích hợp công nghệ AI vào chương trình giảng dạy: Xây dựng các khóa học hướng dẫn học viên cách sử dụng hiệu quả các công cụ AI. Đầu tư vào hạ tầng như cung cấp các phòng máy tính với đường truyền Internet tốc độ cao và các thiết bị hỗ trợ học tập cho học viên. Phát triển nội dung chuyên biệt như hợp tác với các công ty công nghệ để xây dựng và phát triển các nền tảng AI được thiết kế riêng cho học viên Lào, với các bài tập và dữ liệu phù hợp với đặc thù ngôn ngữ của họ.

Đối với sinh viên: Chủ động tìm hiểu và sử dụng nhằm nâng cao nhận thức về vai trò của AI và chủ động tìm kiếm các công cụ học tập phù hợp. Kết hợp linh hoạt trong việc sử dụng AI như một công cụ hỗ trợ để không bị phụ thuộc hoàn toàn. Cần kết hợp giữa việc học với AI và việc thực hành với giáo viên, bạn bè.

5. KẾT LUẬN

Ứng dụng AI trong giảng dạy tiếng Việt cho học viên Lào là một hướng đi tất yếu. Việc nắm bắt cơ hội, khắc phục những khó khăn và đưa ra các giải pháp đồng bộ sẽ giúp nâng cao đáng kể chất lượng đào tạo. Mặc dù AI không thể thay thế hoàn toàn vai trò của giáo viên, nhưng nó đang trở thành một công cụ hỗ trợ đắc lực, góp phần nâng cao chất lượng đào tạo và xây dựng cầu nối ngôn ngữ vững chắc giữa hai quốc gia.

Tài liệu tham khảo

1. Nguyễn Thị Thùy Dương, Đào Thị Lan Anh (2020), *Ứng dụng trí tuệ nhân tạo trong việc dạy và học tiếng Anh,* Tạp chí Khoa học và Công nghệ, Đại học Đà Nẵng, Số 5

2. Bùi Hiền (1999), *Phương pháp hiện đại dạy học ngoại ngữ,* Nxb Đại học Quốc gia Hà Nội.

3. Phạm Thị Lan Hương (2019), *Các ứng dụng của trí tuệ nhân tạo trong giáo dục và đào tạo*, Tạp chí Khoa học Giáo dục, số 52

4. Vũ Thị Lan (2019), *Ứng dụng công nghệ AI trong dạy và học tiếng Anh: Thực trạng và giải pháp,* Tạp chí Giáo dục và Xã hội, số 90.

5. Trịnh Văn Minh (2004), *Phương pháp giao tiếp trong dạy học ngoại ngữ ở nhà trường phổ thông*, Tạp chí Giáo dục, số 6.

6. Phạm Thu Trang (2024), *Ứng dụng AI trong dạy học ngoại ngữ - Tiềm năng và thách thức*, Tạp chí Thiết bị giáo dục, số 319, kì 2.

XÂY DỰNG HỆ THỐNG HỖ TRỢ THÍCH ỨNG ĐỂ TỰ ĐỘNG TẠO HƯỚNG DẪN SỬ DỤNG PHẦN MỀM TỪ NGỮ CẢNH TRỰC QUAN VÀ CƠ SỞ TRI THỨC

Lê Đoàn Phương Uyên[1], Nguyễn Hồng Bửu Long[2]

Tóm tắt: Trong bối cảnh số hóa ngày càng phức tạp, người dùng thường gặp phải gánh nặng nhận thức khi tương tác với các phần mềm có nhiều tính năng. Các tài liệu hướng dẫn truyền thống, vốn mang tính tĩnh và thiếu linh hoạt, thường không đáp ứng được nhu cầu hỗ trợ tức thời theo ngữ cảnh, dẫn đến giảm hiệu suất và sự hài lòng của người dùng. Nghiên cứu đề xuất một phương pháp mới nhằm xây dựng một hệ thống hỗ trợ học tập thích ứng, có khả năng tự động phát sinh hướng dẫn sử dụng phần mềm dựa trên yêu cầu và ngữ cảnh trực quan của người dùng. Kiến trúc đề xuất là một hệ thống đa phương thức, tích hợp các mô hình ngôn ngữ lớn (LLM) để phân tích hình ảnh giao diện người dùng (UI) và kết hợp với kỹ thuật Truy xuất thông tin tăng cường (RAG) từ một cơ sở tri thức chuyên biệt. Bằng cách này, hệ thống có thể kết hợp sự hiểu biết trực quan về giao diện người dùng đang thấy và làm thế nào để thực hiện tác vụ. Kết quả thực nghiệm trên các kịch bản sử dụng phổ biến cho thấy phương pháp đề xuất có khả năng tạo ra các hướng dẫn với độ chính xác và đầy đủ hơn so với các phương pháp tiếp cận thông thường. Nghiên cứu này không chỉ đề xuất một giải pháp kỹ thuật mà còn mở ra một hướng tiếp cận mới trong việc hỗ trợ người dùng, góp phần nâng cao trải nghiệm người dùng trong môi trường số.

***Từ khóa**:* xử lý ngôn ngữ tự nhiên, học máy, mô hình ngôn ngữ lớn, đa phương thức, truy xuất thông tin tăng cường, hỗ trợ học tập thích ứng, tương tác người-máy, hướng dẫn sử dụng tự động.

I. Giới thiệu

1.1 Bối cảnh và vấn đề

Kỷ nguyên số đã chứng kiến sự bùng nổ của các ứng dụng phần mềm trong mọi lĩnh vực của đời sống, từ công việc, học tập đến giải trí. Các phần mềm này ngày càng trở nên mạnh mẽ hơn nhưng cũng đồng thời phức tạp hơn, với hàng trăm tính năng được tích hợp trong một giao diện người dùng. Sự phức tạp này đặt ra một rào cản đáng kể cho người dùng cuối. Họ thường xuyên đối mặt với tình trạng phải ghi nhớ các quy trình, tìm kiếm các chức năng ẩn, và thích ứng với các bản cập nhật liên tục thay đổi giao diện. Vấn đề này đặc biệt nghiêm trọng đối với các nhóm người dùng như người lớn tuổi, người ít tiếp xúc với công nghệ, hoặc các chuyên gia phải làm chủ các phần mềm doanh nghiệp phức tạp.

Để giải quyết vấn đề này, các nhà phát triển phần mềm đã cung cấp nhiều hình thức hỗ trợ khác nhau. Tuy nhiên, các giải pháp truyền thống bộc lộ nhiều hạn chế:

- **Tài liệu hướng dẫn tĩnh**: Thường được trình bày dưới dạng các tệp PDF hoặc các trang web dài, chứa đựng một lượng thông tin khổng lồ. Người dùng phải tự mình thực hiện

[1] Học viên Cao học, Trường Đại học Khoa học Tự nhiên, ĐHQG TPHCM, *email: 24C15039@student.hcmus.edu.vn*

[2] TS, đang giảng dạy tại Trường Đại học Khoa học Tự nhiên, ĐHQG TPHCM, *email: nhblong@fit.hcmus.edu.vn*

công việc tìm kiếm, đọc qua nhiều phần không liên quan để tìm ra giải pháp cho vấn đề cụ thể của mình.

- **Các trang Hỏi-Đáp (FAQ) và Diễn đàn cộng đồng**: Mặc dù hữu ích, các nguồn này thường phân mảnh, thông tin có thể lỗi thời, và không đảm bảo câu trả lời cho các vấn đề đặc thù.
- **Video hướng dẫn**: Mặc dù trực quan, video lại mang tính tuyến tính và một chiều. Người dùng không thể tương tác và phải xem qua các phần không cần thiết để đến được đoạn hướng dẫn họ cần.

Điểm yếu chung của các phương pháp trên là sự thiếu hụt hai yếu tố cốt lõi: tính tức thời và tính ngữ cảnh. Người dùng cần được hỗ trợ ngay lập tức, ngay tại màn hình họ đang gặp khó khăn, chứ không phải một giải pháp chung chung.

1.2 Hướng tiếp cận và đóng góp

Nhận thấy khoảng trống này, chúng tôi tin rằng kỷ nguyên của các mô hình ngôn ngữ lớn có thể mang lại sự thay đổi trong lĩnh vực hỗ trợ người dùng – chuyển từ hỗ trợ thụ động sang hướng dẫn chủ động. Nghiên cứu đề xuất một hướng tiếp cận mới: một hệ thống thông minh có khả năng hiểu được giao diện hiện tại của người dùng, hiểu câu hỏi của họ, và truy xuất các tài liệu kỹ thuật để tạo ra một bản hướng dẫn hoàn chỉnh, cá nhân hóa cho đúng tình huống đó.

Chúng tôi gọi đây là một hệ thống hỗ trợ thích ứng, với mục tiêu giúp người dùng "học" cách sử dụng phần mềm một cách hiệu quả nhất. Các đóng góp chính của công trình này bao gồm:

1. **Thiết kế kiến trúc đa phương thức** kết hợp khả năng phân tích ngữ cảnh thị giác từ ảnh chụp màn hình và xử lý yêu cầu ngôn ngữ tự nhiên để hiểu ý định và hoàn cảnh của người dùng.
2. **Ứng dụng kỹ thuật Truy xuất thông tin tăng cường** vào bài toán hỗ trợ người dùng phần mềm, tạo ra một cầu nối giữa khả năng suy luận của LLM và kho kiến thức thực tế từ tài liệu, qua đó đảm bảo tính chính xác và tin cậy của các hướng dẫn.
3. **Xây dựng một hệ thống prototype và tiến hành đánh giá thực nghiệm** để chứng minh tính hiệu quả của phương pháp đề xuất so với phương pháp tiếp cận baseline, cung cấp các đánh giá định lượng và định tính về giá trị của hệ thống.

1.3 Cấu trúc bài báo

Phần còn lại của bài báo được cấu trúc như sau: Chương 2 sẽ trình bày tổng quan chi tiết về các công trình nghiên cứu liên quan. Chương 3 đi sâu vào mô tả kiến trúc và các thành phần của phương pháp đề xuất. Chương 4 trình bày thiết lập thực nghiệm và quy trình đánh giá. Chương 5 phân tích các kết quả thu được và thảo luận về ý nghĩa của chúng. Cuối cùng, Chương 6 đưa ra kết luận và đề xuất các hướng phát triển trong tương lai.

II. Tổng quan tình hình nghiên cứu

Nghiên cứu nằm ở giao điểm của nhiều lĩnh vực đang phát triển nhanh chóng trong Trí tuệ Nhân tạo. Để thể hiện rõ hơn đóng góp của công trình, chúng tôi sẽ đi qua các nền tảng lý thuyết và công nghệ liên quan.

2.1. Sự phát triển của các mô hình ngôn ngữ lớn

Nền tảng của Xử lý ngôn ngữ tự nhiên (NLP) hiện đại được xây dựng trên kiến trúc Transformer [6], một kiến trúc đã cách mạng hóa cách xử lý dữ liệu tuần tự. Sự thành công của Transformer đã mở đường cho sự ra đời của các LLM được tiền huấn luyện trên kho dữ liệu văn bản khổng lồ, chẳng hạn như BERT [7] và họ mô hình GPT (Generative Pre-trained Transformer) [8]. Các mô hình này thể hiện khả năng hiểu và tạo ra ngôn ngữ tự nhiên một cách vượt trội, trở thành xương sống cho hàng loạt ứng dụng từ dịch máy đến trả lời câu hỏi.

2.2. Mô hình ngôn ngữ lớn đa phương thức

Một bước tiến tự nhiên của LLM là mở rộng khả năng của chúng ra ngoài phạm vi văn bản đơn thuần. Các mô hình đa phương thức, như GPT-4V [1], Gemini [2] và LLaVA [3], được thiết kế để xử lý và liên kết thông tin từ nhiều phương thức khác nhau (ví dụ: văn bản, hình ảnh, âm thanh). Về mặt kiến trúc, các mô hình này thường bao gồm một bộ mã hóa thị giác (vision encoder), chẳng hạn như Vision Transformer (ViT) [9], để chuyển đổi hình ảnh thành các biểu diễn vector, sau đó các vector này được đưa vào LLM cùng với các biểu diễn văn bản. Khả năng "nhìn" và "đọc" đồng thời này là yếu tố cốt lõi cho phép hệ thống của chúng tôi phân tích được ngữ cảnh trực quan từ ảnh chụp màn hình của người dùng.

2.3. Kỹ thuật Truy xuất thông tin tăng cường (RAG)

Mặc dù LLM có một lượng kiến thức khổng lồ được "tham số hóa" trong các trọng số của chúng, nhưng kiến thức này là tĩnh tại thời điểm huấn luyện và có thể không chính xác hoặc lỗi thời. Kỹ thuật RAG [4] được đề xuất để giải quyết vấn đề này bằng cách cho phép LLM truy cập vào một kho kiến thức bên ngoài trong thời gian thực. Một hệ thống RAG điển hình bao gồm hai thành phần chính: một bộ truy xuất (retriever) và một bộ tạo sinh (generator). Bộ truy xuất, thường dựa trên tìm kiếm vector mật độ cao, có nhiệm vụ tìm kiếm các tài liệu liên quan nhất từ một cơ sở dữ liệu. Sau đó, các tài liệu này được đưa vào prompt cùng với câu hỏi ban đầu để bộ tạo sinh (là một LLM) tạo ra câu trả lời cuối cùng. RAG đã được chứng minh là cực kỳ hiệu quả trong việc cải thiện tính xác thực và giảm thiểu "ảo giác" (hallucination) của LLM [10], một yêu cầu tối quan trọng đối với việc tạo ra các hướng dẫn kỹ thuật đáng tin cậy.

2.4. Trí tuệ nhân tạo trong Giáo dục và hỗ trợ người dùng

Lĩnh vực Hệ thống hỗ trợ thông minh (Intelligent Tutoring Systems - ITS) đã có lịch sử lâu dài trong việc áp dụng AI để tạo ra các lộ trình học tập được cá nhân hóa [5]. Các hệ thống này thường mô hình hóa kiến thức của người học và cung cấp phản hồi, gợi ý phù hợp. Song song đó, các Hệ thống Hỗ trợ Người dùng (User Assistance Systems - UAS) cũng đã phát triển từ các file "Help" đến các trợ lý ảo và chatbot tích hợp. Tuy nhiên, nhiều hệ thống hiện tại vẫn chủ yếu dựa trên việc khớp từ khóa hoặc các luồng hội thoại được định nghĩa trước, thiếu khả năng hiểu sâu sắc ngữ cảnh đa phương thức.

2.5. Khoảng trống Nghiên cứu

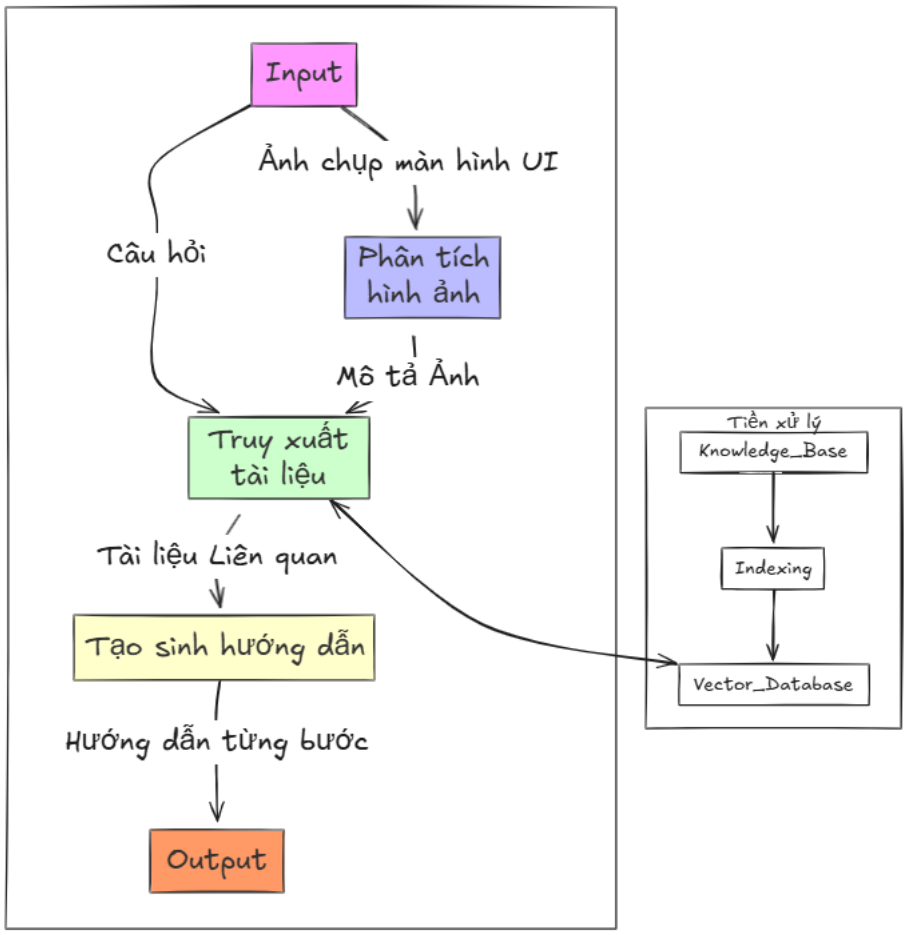
Qua việc tổng quan các lĩnh vực trên, chúng tôi nhận thấy một khoảng trống nghiên cứu trong lĩnh vực này. Mặc dù đã có những tiến bộ vượt bậc trong từng lĩnh vực riêng lẻ (LLM đa phương thức, RAG, ITS), việc **tổng hợp các công nghệ này thành một giải pháp thống nhất cho bài toán hỗ trợ người dùng phần mềm theo ngữ cảnh trực quan** vẫn là một lĩnh vực còn nhiều tiềm năng chưa được khai phá. Các công trình hiện tại hoặc tập trung vào khả năng đa phương thức mà không có cơ sở tri thức vững chắc, hoặc tập trung vào RAG trên dữ liệu văn bản thuần túy mà bỏ qua ngữ cảnh trực quan. Nghiên cứu của chúng tôi nhằm vào khoảng trống này, đề xuất một kiến trúc kết hợp để từ đó có thể hỗ trợ người dùng tốt nhất.

III. Phương pháp đề xuất

Chúng tôi đã thiết kế và xây dựng một kiến trúc có hệ thống, bao gồm hai giai đoạn chính: (1) Giai đoạn tiền xử lý để xây dựng cơ sở tri thức, và (2) Giai đoạn xử lý trong thời gian thực để tạo ra hướng dẫn.

3.1. Kiến trúc tổng thể

Kiến trúc tổng thể của hệ thống được mô tả trong Hình 1. Luồng xử lý bắt đầu khi người dùng gửi yêu cầu và kết thúc bằng việc trả về một bản hướng dẫn chi tiết. Các thành phần chính sẽ được mô tả chi tiết trong các mục tiếp theo.



Hình 1: Sơ đồ kiến trúc chi tiết của hệ thống đề xuất, bao gồm các luồng dữ liệu giữa các module

3.2. Giai đoạn tiền xử lý: Xây dựng cơ sở tri thức vector

Đây là giai đoạn được thực hiện một lần để chuẩn bị dữ liệu cho hệ thống.

3.2.1. Thu thập và chuẩn hóa tài liệu: Chúng tôi thu thập tất cả các tài liệu hướng dẫn liên quan đến phần mềm. Các tài liệu này được chuyển đổi thành định dạng văn bản markdown để dễ dàng xử lý.

3.2.2. Phân đoạn văn bản (Chunking): Các tài liệu dài được chia thành các đoạn (chunk) nhỏ hơn, mỗi đoạn chứa một đơn vị thông tin tương đối độc lập. Việc phân đoạn này rất quan trọng, vì nó giúp cho quá trình truy xuất thông tin trở nên chính xác hơn, tránh đưa các thông tin nhiễu không liên quan vào prompt. Chúng tôi phân đoạn dựa trên số lượng token với một khoảng chồng lấp (overlap) giữa các đoạn để đảm bảo tính liên tục về ngữ nghĩa.

3.2.3. Vector hóa (Embedding): Mỗi đoạn văn bản sau đó được đưa qua một mô hình embedding để chuyển đổi thành một vector, đại diện cho ngữ nghĩa của nó. Chúng tôi lựa chọn mô hình all-MiniLM-L6-v2 từ thư viện Sentence Transformers [11] vì sự cân bằng tốt giữa hiệu năng và tốc độ.

3.2.4. Lưu trữ và lập chỉ mục (Indexing): Tất cả các cặp (đoạn văn bản, vector tương ứng) được lưu trữ vào một cơ sở dữ liệu vector. Chúng tôi sử dụng ChromaDB vì tính đơn giản, mã nguồn mở và khả năng chạy local của nó. ChromaDB tự động lập chỉ mục các vector này để tối ưu hóa cho việc tìm kiếm tương tự (similarity search) sau này.

3.3. Giai đoạn xử lý trong thời gian thực: Tạo hướng dẫn

Đây là giai đoạn diễn ra mỗi khi có yêu cầu từ người dùng.

3.3.1. Phân tích ngữ cảnh trực quan đa phương thức: Khi người dùng tải lên một ảnh chụp màn hình, ảnh này sẽ được gửi đến mô hình Gemini Vision. Chúng tôi đã thiết kế một prompt để điều khiển đầu ra của mô hình, yêu cầu nó thực hiện ba nhiệm vụ: (1) Nhận dạng và liệt kê tất cả các thành phần UI quan trọng (nút bấm, ô nhập liệu, menu); (2) Trích xuất toàn bộ văn bản có thể đọc được (OCR); (3) Đưa ra một suy luận tổng quan về mục đích của màn hình.

Ví dụ về prompt:

"Bạn là một chuyên gia phân tích giao diện người dùng. Hãy phân tích hình ảnh sau và trả về một bản mô tả chi tiết dưới dạng văn bản. Bao gồm: (1) danh sách các nút bấm và văn bản trên đó, (2) danh sách các ô nhập liệu và nhãn của chúng, (3) một đoạn mô tả về chức năng chính của màn hình này." Kết quả của bước này là một đoạn văn bản chứa nhiều thông tin, giúp mô tả ngữ cảnh trực quan.

3.3.2. Truy xuất thông tin tăng cường (RAG): Một truy vấn tăng cường được tạo ra bằng cách ghép nối câu hỏi ban đầu của người dùng và mô tả ngữ cảnh trực quan thu được ở bước trước. Ví dụ, nếu người dùng hỏi "làm sao đây?" và ảnh là màn hình đăng nhập, truy vấn tăng cường sẽ là: "Người dùng đang ở màn hình đăng nhập và muốn biết cách đăng nhập." Truy vấn này sau đó được vector hóa và dùng để thực hiện tìm kiếm tương tự (sử dụng độ tương tự cosine) trong cơ sở dữ liệu vector. Hệ thống sẽ lấy ra K đoạn tài liệu có liên quan nhất (trong thí nghiệm này, chúng tôi chọn K=3).

3.3.3. Giai đoạn tổng hợp và tạo sinh hướng dẫn: Đây là bước cuối cùng để tạo sinh hướng dẫn, một prompt cuối cùng, chi tiết, sẽ được gửi đến mô hình Gemini Pro. Prompt này có cấu trúc như sau:

Thiết lập vai trò: "Bạn là một trợ lý ảo chuyên nghiệp, thân thiện, chuyên hướng dẫn người dùng sử dụng phần mềm."

Cung cấp ngữ cảnh từ tài liệu: “Dựa vào các thông tin từ tài liệu hướng dẫn sau đây: [Chèn K đoạn tài liệu đã truy xuất]"

Cung cấp ngữ cảnh trực quan: "Người dùng hiện đang thấy một màn hình được mô tả như sau: [Chèn mô tả ngữ cảnh trực quan]"

Nêu rõ yêu cầu của người dùng: "Câu hỏi của người dùng là: [Chèn câu hỏi gốc]"

Instruction: "Dựa vào tất cảcác thông tin trên, hãy tạo ra một bản hướng dẫn chi tiết, theo từng bước (Bước 1, Bước 2,...) để trả lời chính xác câu hỏi của người dùng. Hãy sử dụng lời văn đơn giản, dễ hiểu."

Cấu trúc prompt phức hợp này đảm bảo LLM có đầy đủ thông tin cần thiết để tạo ra một câu trả lời vừa bám sát tài liệu, vừa phù hợp với những gì người dùng thực sự đang nhìn thấy trên màn hình.

IV. Thực nghiệm và Đánh giá

4.1. Thiết lập Thí nghiệm

Các thí nghiệm được thực hiện sử dụng Python 3.10 và các thư viện chính bao gồm google-generativeai, chromadb, sentence-transformers. Mô hình LLM được truy cập thông qua API của Google AI Studio.

4.2. Xây dựng Dataset Đánh giá

Để đảm bảo việc đánh giá có hệ thống, chúng tôi đã xây dựng một bộ dữ liệu kiểm thử nhỏ nhưng có mục đích rõ ràng, bao gồm 2 kịch bản chính nhằm kiểm tra các khía cạnh khác nhau của hệ thống:

- **Kịch bản 1: Tác vụ tường minh – Đăng nhập:**
  - **Ảnh UI:** Ảnh chụp màn hình đăng nhập tiêu chuẩn với các trường "Email" hoặc “Số điện thoại”, "Mật khẩu" và nút "Đăng nhập".
  - **Câu hỏi:** "Hướng dẫn tôi cách đăng nhập vào tài khoản."
  - **Mục tiêu:** Kiểm tra khả năng hệ thống xử lý các tác vụ mà thông tin cần thiết gần như đã hiển thị đầy đủ trên giao diện.
- **Kịch bản 2: Tác vụ phức tạp – Tính năng “AI giúp tôi viết” khi giao bài tập về nhà**
  - **Ảnh UI:** Ảnh chụp màn hình chính của tính năng “Giao bài tập về nhà”.
  - **Câu hỏi:** "Làm thế nào để dùng “AI giúp tôi viết”?"
  - **Mục tiêu:** Kiểm tra khả năng hệ thống xử lý các tác vụ phức tạp, đòi hỏi kiến thức về quy trình gồm nhiều bước không được hiển thị trên màn hình hiện tại.

4.3. Các Mô hình So sánh

Chúng tôi so sánh hiệu năng của hai hệ thống sau:

- **Baseline (Không RAG):** Một hệ thống đơn giản hóa chỉ sử dụng LLM đa phương thức (Gemini Pro Vision). Hệ thống này nhận trực tiếp ảnh UI và câu hỏi của người dùng, sau đó được yêu cầu tạo hướng dẫn trên đầu vào nhận được.

- **Proposed (Có RAG):** Hệ thống của chúng tôi, bao gồm cả phân tích trực quan và truy xuất thông tin từ cơ sở tri thức như đã mô tả trong Mục 3.

4.4. Quy trình và Tiêu chí Đánh giá

Do tính chất phức tạp và sáng tạo của các câu trả lời dạng văn bản, chúng tôi lựa chọn phương pháp đánh giá bởi con người làm phương pháp chính. Các câu trả lời được đánh giá một cách ẩn danh (người chấm không biết câu trả lời đến từ hệ thống nào) dựa trên một bộ tiêu chí được định nghĩa trong Bảng 1.

Tiêu chí	Mô tả	Thang điểm 1	Thang điểm 5
Độ chính xác (Accuracy)	Các bước được hướng dẫn có đúng với quy trình thực tế của phần mềm không?	Thang 1: Hoàn toàn sai, bịa đặt.	Thang 5: Hoàn toàn chính xác, đáng tin cậy.
Độ đầy đủ (Completeness)	Hướng dẫn có bao quát hết các bước cần thiết từ đầu đến cuối không?	Thang 1: Bỏ sót các bước cực kỳ quan trọng.	Thang 5: Cung cấp đầy đủ các bước cần thiết.
Độ rõ ràng (Clarity)	Lời văn có dễ hiểu, không mơ hồ, người mới có thể làm theo được không?	Thang 1: Rất khó hiểu, dùng thuật ngữ phức tạp.	Thang 5: Rất rõ ràng, mạch lạc, dễ làm theo.

Bảng 1: Bộ tiêu chí đánh giá chất lượng hướng dẫn

V. Kết quả và bàn luận

5.1. Kết quả định lượng

Sau khi tiến hành thí nghiệm và chấm điểm khảo sát trên 60 người, chúng tôi thu được kết quả điểm trung bình được tổng hợp trong Bảng 2.

Tiêu chí	**Baseline**	**RAG**
Độ chính xác	3.22	4.05
Độ đầy đủ	3.29	4.22
Độ rõ ràng	3.1	4.02
Điểm trung bình	3.2	4.1

Bảng 2: Kết quả so sánh điểm đánh giá trung bình giữa hai hệ thống

Từ kết quả trên, có thể thấy rõ ràng rằng hệ thống đề xuất đã thể hiện hiệu quả vượt trội so với hệ thống Baseline. Điểm trung bình của hệ thống đạt 4.1, cao hơn đáng kể so với 3.2 của Baseline.

Sự chênh lệch lớn nhất được quan sát ở tiêu chí **Độ đầy đủ**, với điểm số 4.22 của hệ thống đề xuất so với chỉ 3.29 của Baseline. Chênh lệch 0.93 điểm này đã trực tiếp xác thực giả thuyết của chúng tôi: nếu không có quyền truy cập vào cơ sở tri thức bên ngoài, LLM dù thông minh đến đâu cũng không thể đoán ra được một quy trình nghiệp vụ phức tạp. Tiêu chí **Độ chính xác** cũng cho thấy sự cải thiện rõ rệt, khẳng định rằng việc dựa vào tài liệu giúp giảm thiểu nguy cơ LLM đưa ra các thông tin sai lệch.

5.2. Phân tích định tính

Để hiểu sâu hơn về sự khác biệt trong hành vi của hai hệ thống, chúng tôi sẽ phân tích chi tiết câu trả lời của chúng cho Kịch bản 2 - Sử dụng tính năng AI giúp tôi viết.

5.2.1 Phân tích câu trả lời của Baseline

- Bấm vào nút **“AI giúp tôi viết”**.
- Sẽ có menu xổ xuống với các lựa chọn (thường là: tạo nội dung từ gợi ý, chỉnh sửa nội dung có sẵn, dịch, hoặc viết lại).
- Bạn nhập yêu cầu cụ thể (ví dụ: “Viết mô tả bài tập toán lớp 6 về phân số”, hoặc “Soạn bài tập lập trình Python cơ bản”) → AI sẽ tự động sinh nội dung.
- Sau đó, bạn có thể chỉnh sửa thủ công trước khi bấm **Đăng** hoặc **Lưu nháp**.
- ***Nhận xét:*** Câu trả lời của Baseline chỉ dừng lại ở mức mô tả những gì nó thấy. Nó xác định đúng nút "AI giúp tôi viết" và khuyên người dùng nhấn vào đó. Tuy nhiên, toàn bộ các bước sau đó gần như hoàn toàn bị bỏ qua. Câu trả lời này là chính xác nhưng không đầy đủ, do đó không thực sự hữu ích cho người dùng.

5.2.2 Phân tích câu trả lời của RAG

Hướng dẫn này sẽ giúp bạn sử dụng tính năng "AI giúp tôi viết" trên ClassIn để tạo nội dung bài tập một cách nhanh chóng và hiệu quả. Giả sử bạn đang ở giao diện "Tạo bài tập mới" như mô tả.

Bước 1: Truy cập tính năng AI giúp tôi viết

Nhìn vào cột bên trái (Nội dung bài tập). Bạn sẽ thấy một nút có nhãn "AI giúp tôi viết" (có thể kèm biểu tượng AI). Nhấp vào nút này.

Bước 2: Nhập hướng dẫn cho AI

Một cửa sổ hoặc hộp thoại sẽ xuất hiện. Tại đây, bạn cần nhập hướng dẫn hoặc mô tả ngắn gọn về nội dung bài tập bạn muốn AI tạo ra. **Càng chi tiết và rõ ràng, kết quả AI tạo ra càng chính xác.** Ví dụ:

- "Bài tập toán lớp 5 về phép cộng phân số, gồm 10 câu hỏi."
- "Bài văn ngắn về chủ đề bảo vệ môi trường, dành cho học sinh lớp 7."
- "Câu hỏi trắc nghiệm ôn tập chương 3 môn lịch sử lớp 8, 20 câu."

Bước 3: Yêu cầu AI tạo nội dung

Sau khi nhập hướng dẫn, nhấn nút "Tạo" (hoặc nút tương tự có ý nghĩa tạo/generate). ClassIn sẽ sử dụng AI để tạo nội dung bài tập dựa trên hướng dẫn của bạn. Quá trình này có thể mất vài giây.

Bước 4: Xem và chỉnh sửa nội dung do AI tạo

Sau khi AI tạo nội dung, nó sẽ được hiển thị trong cửa sổ. Bạn có thể:

- **Áp dụng và chỉnh sửa:** Nhấn nút này để chèn nội dung do AI tạo vào vùng nhập nội dung bài tập (cột bên trái). Sau đó, bạn có thể chỉnh sửa, thêm bớt, định dạng… cho phù hợp.
- **Tạo lại:** Nếu không hài lòng với kết quả, hãy nhấn nút này để yêu cầu AI tạo lại nội dung dựa trên cùng hướng dẫn ban đầu.
- **Sửa đổi hướng dẫn:** Nếu cần điều chỉnh hướng dẫn cho AI, bạn có thể quay lại bước 2 và sửa đổi hướng dẫn trước khi tạo lại nội dung.

Bước 5: Hoàn tất việc tạo bài tập

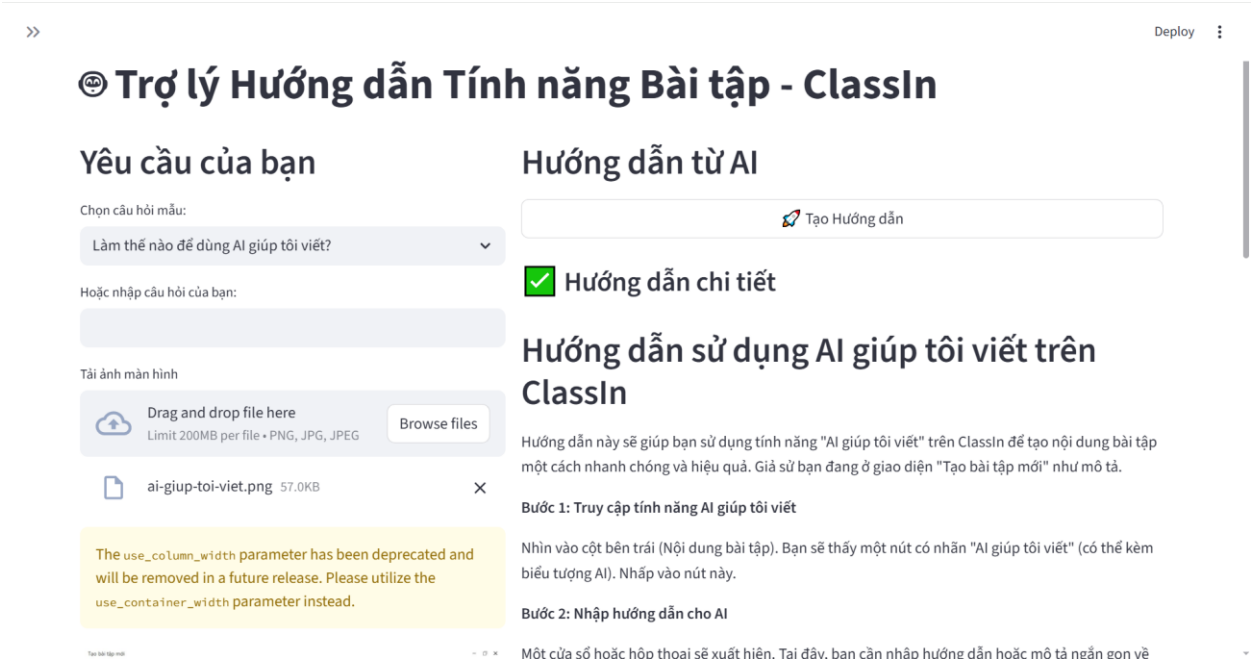
Sau khi đã hài lòng với nội dung bài tập (đã được chỉnh sửa nếu cần), bạn có thể sử dụng các tùy chọn ở cột bên phải để thiết lập thời gian, điểm số, lớp học, chương trình… Sau đó, nhấn nút "Lưu nháp" để lưu tạm hoặc "Đăng" để đăng bài tập lên hệ thống ClassIn.

Mẹo: Để đạt được kết quả tốt nhất, hãy cung cấp cho AI càng nhiều thông tin càng tốt, bao gồm:

- **Lớp học:** Lớp học nào sẽ làm bài tập này?
- **Môn học:** Môn học nào?
- **Chủ đề:** Chủ đề cụ thể của bài tập là gì?
- **Số lượng câu hỏi/bài tập:** Bạn muốn bao nhiêu câu hỏi hoặc bài tập?
- **Độ khó:** Độ khó của bài tập như thế nào (dễ, trung bình, khó)?
- **Loại câu hỏi:** Câu hỏi trắc nghiệm, tự luận, hay kết hợp?

Bằng cách làm theo các bước trên, bạn có thể tận dụng sức mạnh của AI trên ClassIn để tạo nội dung bài tập một cách hiệu quả và tiết kiệm thời gian.

Nhận xét: Ngược lại, câu trả lời của hệ thống đề xuất đã cung cấp một quy trình hoàn chỉnh. Nó không chỉ xác định đúng bước đầu tiên dựa trên phân tích hình ảnh, mà còn truy xuất thành công tài liệu từ cơ sở tri thức, từ đó liệt kê đầy đủ các bước tiếp theo. Câu trả lời này vừa chính xác, vừa đầy đủ, đáp ứng hoàn hảo yêu cầu của người dùng.



Hình 2: Hình ảnh kết quả trả lời câu hỏi của hệ thống có sử dụng RAG

5.3. Bàn luận về các Hạn chế

Mặc dù kết quả ban đầu khá hứa hẹn, chúng tôi nhận thấy một số hạn chế và thách thức cần được giải quyết trong các nghiên cứu tiếp theo:

- **Sự phụ thuộc vào chất lượng tài liệu:** Hiệu quả của hệ thống RAG phụ thuộc trực tiếp vào chất lượng và mức độ đầy đủ của cơ sở tri thức. Nếu tài liệu bị thiếu, lỗi thời hoặc viết không rõ ràng, chất lượng của hướng dẫn sẽ bị ảnh hưởng.
- **Xử lý các giao diện có khả năng tùy biến:** Hệ thống có thể gặp khó khăn với các giao diện thay đổi liên tục hoặc được người dùng tùy biến, hoặc những trạng thái không có trong tài liệu.
- **Quy mô đánh giá:** Thí nghiệm hiện tại chỉ được thực hiện trên một bộ dữ liệu rất nhỏ. Cần có một bộ dữ liệu đánh giá lớn và đa dạng hơn để kiểm chứng khả năng tổng quát hóa của hệ thống.

VI. Kết luận và Hướng phát triển

6.1. Kết luận

Bài báo này đã trình bày một kiến trúc hệ thống đa phương thức để giải quyết bài toán tạo hướng dẫn sử dụng phần mềm một cách tự động và theo ngữ cảnh. Bằng cách kết hợp khả năng phân tích trực quan của LLM đa phương thức với độ tin cậy của cơ sở tri thức thông qua kỹ thuật RAG, hệ thống của chúng tôi đã chứng tỏ khả năng cung cấp các hướng dẫn vừa chính xác, vừa đầy đủ, so với các phương pháp tiếp cận thông thường. Công trình này cho thấy tiềm năng của việc tích hợp AI đa phương thức vào việc xây dựng các hệ thống hỗ trợ người dùng, biến việc học và sử dụng công nghệ trở nên trực quan và dễ dàng hơn.

6.2. Hướng phát triển tương lai

Dựa trên nền tảng của nghiên cứu này, chúng tôi đề xuất một số hướng phát triển như sau:

- **Tích hợp khả năng hội thoại qua nhiều lượt hỏi đáp:** Mở rộng hệ thống để có thể duy trì ngữ cảnh qua nhiều lượt hỏi đáp, cho phép người dùng hỏi sâu hơn về từng bước trong quy trình. Điều này đòi hỏi việc thiết kế cơ sở dữ liệu để lưu trữ lịch sử hội thoại.
- **Tự động hóa việc xây dựng Cơ sở Tri thức:** Nghiên cứu các phương pháp tự động thu thập, phân tích và cấu trúc hóa thông tin từ nhiều nguồn khác nhau (PDF, website, video) để xây dựng và cập nhật cơ sở tri thức một cách hiệu quả.
- **Tối ưu hóa Hiệu năng:** Nghiên cứu các kỹ thuật để giảm độ trễ của hệ thống, ví dụ như tối ưu hóa các mô hình embedding hoặc sử dụng các kỹ thuật caching.

TÀI LIỆU THAM KHẢO

[1] OpenAI, "GPT-4 Technical Report," arXiv:2303.08774, 2023.

[2] Google, "Gemini: A Family of Highly Capable Multimodal Models," arXiv:2312.11805, 2023.

[3] H. Liu *et al.*, "Visual Instruction Tuning," arXiv:2304.08485, 2023.

[4] P. Lewis *et al.*, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems*, 2020.

[5] K. VanLehn, "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems," *Educational Psychologist*, 2011.

[6] A. Vaswani *et al.*, "Attention is All You Need," in *Advances in Neural Information Processing Systems*, 2017.

[7] J. Devlin *et al.*, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," arXiv:1810.04805, 2018.

[8] A. Radford *et al.*, "Language Models are Unsupervised Multitask Learners," *OpenAI Blog*, 2019.

[9] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," arXiv:2010.11929, 2020.

[10] Y. Gao *et al.*, "Retrieval-Augmented Generation for Large Language Models: A Survey," arXiv:2312.10997, 2023.

[11] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019.

CHUYỂN ĐỔI SỐ TRONG ĐÀO TẠO TIẾNG VIỆT CHO HỌC VIÊN LÀO TẠI HỌC VIỆN AN NINH NHÂN DÂN: ỨNG DỤNG AI CHO HỌC LIỆU TƯƠNG TÁC VÀ CÁ NHÂN HÓA

Nguyễn Thụy Trang[1]

Tóm tắt

Chuyển đổi số đang từng bước định hình lại phương thức giảng dạy ngoại ngữ, đồng thời mở ra khả năng ứng dụng trí tuệ nhân tạo (AI) trong dạy và học tiếng Việt cho người nước ngoài. Tại Học viện An ninh Nhân dân (HVANND), giảng dạy tiếng Việt cho học viên Lào còn nhiều khó khăn: hạn chế về phát âm và giao tiếp do khác biệt ngôn ngữ – văn hóa, thiếu môi trường luyện tập và học liệu cá nhân hóa, trong khi giảng viên khó xây dựng lộ trình cho lớp học đa trình độ với công cụ công nghệ còn hạn chế. Từ thực trạng đó, bài báo đề xuất mô hình ứng dụng AI trong phát triển học liệu tiếng Việt theo hai hướng: (i) giải pháp khả thi trước mắt như chatbot hội thoại kịch bản hóa và công cụ luyện phát âm cơ bản; (ii) định hướng lâu dài với hệ thống toàn diện hơn. Các phân tích chỉ mang tính gợi mở, có thể tham khảo cho nghiên cứu và đào tạo tiếng Việt trong môi trường an ninh cũng như ở một số cơ sở tương đồng.

***Từ khóa**:* chuyển đổi số; trí tuệ nhân tạo; học liệu tương tác; cá nhân hóa; đào tạo tiếng Việt; học viên Lào

1. MỞ ĐẦU

1.1. Bối cảnh toàn cầu: Cách mạng công nghiệp 4.0 và AI trong giáo dục

Cuộc Cách mạng công nghiệp lần thứ tư đang tái định hình nền giáo dục toàn cầu, trong đó AI nổi lên như một trong những công nghệ then chốt của giáo dục số. Các báo cáo của UNESCO (2021) và OECD (2022) đều khẳng định AI có khả năng thúc đẩy học tập cá nhân hóa, tăng cường tính tương tác và hỗ trợ học tập suốt đời. Đặc biệt từ năm 2022, sự xuất hiện của các mô hình ngôn ngữ lớn (LLMs) như ChatGPT đã mở ra thế hệ mới của công cụ học tập. Các nghiên cứu gần đây cho thấy ChatGPT có thể hỗ trợ phát triển kỹ năng viết, dịch thuật và tư duy phản biện trong bối cảnh học ngoại ngữ, đồng thời góp phần hình thành môi trường học tập giàu tương tác (Kasneci et al., 2023; Rudolph et al., 2023; Chiu et al., 2024; Lo, 2024).

Trong giáo dục ngoại ngữ, AI đang được khai thác ở ba nhóm chức năng chính. Thứ nhất, chatbot và trợ lý hội thoại tạo môi trường giao tiếp ít áp lực, tăng thời lượng luyện tập hội thoại (Fryer & Carpenter, 2020). Thứ hai, hệ thống nhận dạng giọng nói (ASR) cung cấp phản hồi phát âm tức thì, giúp người học cải thiện độ chính xác và sự tự tin (Li & Ni, 2021). Thứ ba, các hệ thống dạy kèm thông minh (ITS) và học liệu đa nền tảng hỗ trợ cá nhân hóa lộ trình học, đáp ứng nhu cầu khác biệt của từng người học (Zawacki-Richter et al., 2019). Dưới khung lý

[1] Th.S, Khoa Ngoại ngữ, Học viện An ninh nhân dân

thuyết CALL (Computer-Assisted Language Learning) và TELL (Technology-Enhanced Language Learning), AI được coi là sự mở rộng tự nhiên, bổ sung chiều sâu cho lý thuyết thụ đắc ngôn ngữ thứ hai (SLA). Các công cụ AI góp phần hiện thực hóa ba trụ cột của SLA: *input* giàu ngữ cảnh, *interaction* có kiểm soát, và *feedback* kịp thời (Krashen, 1985; Gass & Selinker, 2008).

Bên cạnh những tiềm năng, AI cũng đặt ra nhiều thách thức. Các hệ thống ASR chưa xử lý hiệu quả ngôn ngữ giàu thanh điệu như tiếng Việt; chatbot còn hạn chế trong tái hiện ngữ dụng và bối cảnh văn hóa; phần lớn nghiên cứu ứng dụng tập trung vào tiếng Anh. Đặc biệt, vấn đề đạo đức và bảo mật trong giáo dục đang được quan tâm ngày càng nhiều. Các tác giả như Rudolph et al. (2023) và Lo (2024) cảnh báo nguy cơ phụ thuộc quá mức vào máy móc, thiếu kiểm chứng thông tin, và rủi ro lộ lọt dữ liệu cá nhân khi tích hợp AI trong môi trường học tập. Điều này càng có ý nghĩa quan trọng khi áp dụng cho đào tạo tiếng Việt tại HVANND, nơi yêu cầu an toàn thông tin và bảo mật được đặt lên hàng đầu.

Trong bối cảnh đó, việc nghiên cứu ứng dụng AI cho tiếng Việt như ngoại ngữ, đặc biệt đối với học viên Lào – nhóm người học thường xuyên gặp khó khăn ở thanh điệu, phụ âm cuối và ngữ dụng xưng hô – không chỉ bắt kịp xu hướng công nghệ mà còn góp phần làm phong phú thêm lý thuyết SLA ở các ngôn ngữ đơn lập, giàu thanh điệu. Chính từ những xu hướng toàn cầu này, việc phân tích bối cảnh và chính sách chuyển đổi số tại Việt Nam trở thành bước khởi đầu cần thiết cho nghiên cứu.

1.2. Bối cảnh Việt Nam: Chính sách chuyển đổi số và ứng dụng trong giáo dục

Tại Việt Nam, chuyển đổi số trong giáo dục được xác định là một trong những nhiệm vụ trọng tâm của chiến lược quốc gia. Quyết định số 117/QĐ-TTg ngày 25/1/2017 phê duyệt *Đề án tăng cường ứng dụng công nghệ thông tin trong quản lý và hỗ trợ hoạt động dạy – học* đã mở đường cho việc ứng dụng công nghệ vào nhà trường. Tiếp đó, Quyết định số 749/QĐ-TTg ngày 3/6/2020 phê duyệt *Chương trình chuyển đổi số quốc gia đến năm 2025, định hướng đến năm 2030*, trong đó giáo dục – đào tạo được coi là một trong tám lĩnh vực ưu tiên hàng đầu. Gần đây, Chiến lược chuyển đổi số ngành giáo dục giai đoạn 2021–2025, tầm nhìn 2030 do Bộ Giáo dục và Đào tạo ban hành đã cụ thể hóa mục tiêu: phát triển hệ thống học liệu số, nền tảng dạy học trực tuyến, ứng dụng AI và phân tích dữ liệu trong kiểm tra – đánh giá (Bộ GD&ĐT, 2021). Những định hướng này tương đồng với xu thế quốc tế, nơi chuyển đổi số giáo dục được coi là giải pháp chiến lược để nâng cao chất lượng đào tạo (World Bank, 2022; OECD, 2023).

Thực tiễn cho thấy một số cơ sở giáo dục đại học tại Việt Nam đã thử nghiệm ứng dụng AI trong giảng dạy ngoại ngữ. Trường Đại học Ngoại ngữ – ĐHQGHN đã tiến hành khảo sát việc sử dụng ChatGPT trong học tập và nghiên cứu của sinh viên, bước đầu ghi nhận hiệu quả trong luyện dịch và phát triển ý tưởng. Đại học FPT triển khai các lớp học trực tuyến tích hợp công cụ AI hỗ trợ thảo luận, phản hồi và đánh giá, giúp giảng viên tiết kiệm thời gian và nâng cao trải nghiệm học tập. Ngoài ra, nhiều trường đang khai thác Grammarly để hỗ trợ viết học thuật, và các hệ thống quản lý học tập (LMS) có tính năng phân tích dữ liệu để theo dõi tiến độ

học tập (Nguyễn Văn Hiệp, Bùi Thị Thu Hương & Bùi Văn Nam, 2024). Tuy nhiên, các hoạt động này mới dừng ở mức thử nghiệm, chưa hình thành mô hình ứng dụng rộng rãi và bền vững.

Đối với tiếng Việt như ngoại ngữ, tình hình còn hạn chế hơn. Các nghiên cứu và ứng dụng AI mới chỉ tập trung vào xây dựng ngữ liệu hoặc phần mềm luyện phát âm cơ bản, thiếu các giải pháp toàn diện để đáp ứng lớp học đa trình độ và đặc thù người học. Khoảng cách giữa chính sách định hướng và thực tiễn triển khai vẫn còn đáng chú ý, đặc biệt ở ba khía cạnh: hạ tầng công nghệ chưa đồng bộ, đội ngũ giảng viên chưa được bồi dưỡng đầy đủ về ứng dụng AI, và yêu cầu nghiêm ngặt về bảo mật trong môi trường giáo dục an ninh chưa có giải pháp phù hợp.

Trong bối cảnh đó, việc nghiên cứu ứng dụng AI vào giảng dạy tiếng Việt cho học viên Lào tại HVANND có thể coi là một bước đi cần thiết. Nghiên cứu này không chỉ cụ thể hóa chủ trương chuyển đổi số giáo dục của Nhà nước, mà còn góp phần thu hẹp khoảng cách giữa chính sách và thực tiễn, đồng thời cung cấp một mô hình tham khảo cho các cơ sở đào tạo có điều kiện tương đồng.

Trong bức tranh chung đó, việc xem xét đặc thù giảng dạy tiếng Việt cho học viên Lào tại HVANND sẽ giúp làm rõ những thách thức cụ thể mà chính sách và công nghệ cần đáp ứng.

1.3. Đặc thù giảng dạy tiếng Việt cho học viên Lào tại Học viện An ninh Nhân dân

Việc giảng dạy tiếng Việt cho học viên Lào tại HVANND không chỉ mang những đặc điểm chung của dạy ngoại ngữ, mà còn gắn liền với những yêu cầu riêng biệt của một cơ sở đào tạo an ninh. Đây là môi trường nơi bảo mật thông tin, kỷ luật học tập và định hướng nghề nghiệp chuyên ngành luôn được đặt lên hàng đầu. Bên cạnh đó, học viên Lào bước vào học tiếng Việt chủ yếu từ trình độ cơ sở, tập trung rèn luyện kỹ năng giao tiếp hằng ngày trước khi tiếp cận ngôn ngữ chuyên ngành. Sự khác biệt ngôn ngữ – văn hóa, sự đa dạng về trình độ đầu vào, cùng những đòi hỏi nghiêm ngặt về chuẩn mực giao tiếp trong môi trường an ninh đã tạo nên những đặc thù riêng, vừa là thách thức, vừa là cơ hội để tìm kiếm những giải pháp đổi mới. Từ bối cảnh này, việc phân tích các khó khăn điển hình về ngữ âm, ngữ dụng và tổ chức dạy học sẽ góp phần làm rõ những yêu cầu đặt ra cho đổi mới phương pháp và ứng dụng công nghệ trong đào tạo.

Khó khăn về ngữ âm

Tiếng Việt là ngôn ngữ đơn lập với hệ thống thanh điệu phức tạp và nhiều phụ âm cuối, trong khi tiếng Lào có số lượng thanh điệu và âm cuối hạn chế hơn. Sự khác biệt này khiến học viên thường gặp khó khăn trong việc phân biệt và phát âm chính xác các thanh điệu, cũng như duy trì chuẩn mực khi phát âm phụ âm cuối. Đây là thách thức căn bản ngay từ giai đoạn học tiếng Việt cơ sở, đặc biệt trong giao tiếp hằng ngày.

Trở ngại về ngữ dụng và văn hóa giao tiếp

Bên cạnh rào cản ngữ âm, học viên Lào còn gặp khó khăn trong việc sử dụng tiếng Việt đúng ngữ dụng và phù hợp văn hóa. Một đặc thù nổi bật là hệ thống xưng hô phức tạp. Trong khi tiếng Lào có số lượng đại từ nhân xưng hạn chế, tiếng Việt đòi hỏi sự lựa chọn linh hoạt theo tuổi tác, vị thế xã hội và quan hệ giao tiếp. Vì vậy, học viên dễ nhầm lẫn, chẳng hạn dùng *"các em"* thay cho *"chúng em"* khi nói về chính mình trong quan hệ với giảng viên. Đây là hiện tượng không tương thích với quy ước lịch sự giao tiếp của tiếng Việt. Khó khăn cũng xuất hiện trong việc thực hiện các hành vi ngôn ngữ gián tiếp. Thay vì dùng lời mời theo chuẩn lịch sự như *"Mời thầy uống nước ạ"*, học viên thường lựa chọn cách diễn đạt trực tiếp *"Uống nước đi"*, làm giảm mức độ lịch sự theo chuẩn mực văn hóa Việt.

Thách thức trong tổ chức dạy học

Bên cạnh khó khăn về ngữ âm và ngữ dụng, việc tổ chức dạy học tiếng Việt cho học viên Lào tại HVANND cũng đặt ra nhiều thách thức riêng. Trình độ đầu vào không đồng đều - một số học viên đã từng tiếp xúc với tiếng Việt, trong khi nhiều người bắt đầu từ con số không - khiến giảng viên phải xử lý lớp học đa trình độ, khó xây dựng kế hoạch giảng dạy thống nhất.

Hiện nay, phương pháp dạy chủ yếu dựa vào giáo trình in và lớp học trực tiếp, chưa có công cụ số hỗ trợ theo dõi tiến trình và cá nhân hóa lộ trình học. Do đó, học viên ít có cơ hội luyện tập ngoài giờ, trong khi nhu cầu này lại rất lớn. Quan sát thực tế cho thấy nhiều học viên mong muốn được hỗ trợ bằng công cụ phát âm và giao tiếp tự động; tuy nhiên, việc triển khai gặp trở ngại do yêu cầu nghiêm ngặt về bảo mật thông tin trong môi trường đào tạo an ninh.

Những thách thức trên cho thấy, đổi mới tổ chức dạy phải đáp ứng đồng thời hai tiêu chí: cá nhân hóa lộ trình học và đảm bảo an toàn dữ liệu. Đây chính là nền tảng quan trọng để định hướng nghiên cứu và ứng dụng trí tuệ nhân tạo trong giảng dạy tiếng Việt cho học viên Lào tại các cơ sở đào tạo an ninh.

Ý nghĩa đối với nghiên cứu

Những đặc thù nêu trên cho thấy, giảng dạy tiếng Việt cho học viên Lào tại HVANND không chỉ đơn thuần là hoạt động dạy ngoại ngữ, mà còn gắn liền với các yêu cầu đặc biệt về môi trường đào tạo, trong đó tính bảo mật, chuẩn mực giao tiếp và kỷ luật học tập đóng vai trò then chốt. Việc nhận diện những khó khăn về ngữ âm, ngữ dụng và thách thức trong tổ chức dạy học có ý nghĩa quan trọng trên ba phương diện.

Thứ nhất, về lý luận, các phân tích này góp phần bổ sung tư liệu thực nghiệm cho lĩnh vực ngôn ngữ học ứng dụng, đặc biệt là hướng nghiên cứu đối chiếu Việt – Lào, vốn còn hạn chế về số lượng công trình. *Thứ hai*, về thực tiễn, kết quả nghiên cứu giúp giảng viên xác định rõ những trở ngại điển hình để có giải pháp sư phạm phù hợp, tránh việc hình thành thói quen sai cho học viên ngay từ giai đoạn cơ sở. *Thứ ba*, về định hướng công nghệ, việc chỉ ra nhu cầu cá nhân hóa lộ trình học và yêu cầu bảo mật dữ liệu mở ra khả năng ứng dụng trí tuệ nhân tạo trong giảng dạy tiếng Việt cho học viên Lào, theo hướng vừa hiệu quả, vừa đáp ứng đặc thù an ninh.

Như vậy, việc phân tích các đặc thù trong dạy – học tiếng Việt cho học viên Lào tại HVANND không chỉ có giá trị nhận diện thực trạng, mà còn tạo cơ sở khoa học cho những nghiên cứu tiếp theo về đổi mới phương pháp và ứng dụng công nghệ trong đào tạo ngôn ngữ đặc thù.

2. CƠ SỞ LÝ THUYẾT

Nghiên cứu này dựa trên ba hướng tiếp cận chính. *Trước hết*, lý thuyết tiếp thu ngôn ngữ thứ hai (Second Language Acquisition – SLA) nhấn mạnh vai trò của môi trường học tập, chất lượng đầu vào ngôn ngữ và sự tương tác giao tiếp xã hội trong việc hình thành năng lực ngôn ngữ của người học (Krashen, 1985). Từ góc độ này, những khó khăn của học viên Lào về ngữ âm, ngữ dụng và cách thức tham gia vào lớp học có thể được lý giải bởi sự khác biệt hệ thống giữa tiếng mẹ đẻ và tiếng Việt, cũng như điều kiện sử dụng tiếng Việt hạn chế ngoài lớp học.

Tiếp đến, nghiên cứu về giảng dạy tiếng Việt như ngoại ngữ (Vietnamese as a Foreign Language – VFL) đã ghi nhận nhiều lỗi điển hình của người học Lào, bao gồm nhầm lẫn về thanh điệu, bỏ phụ âm cuối và hạn chế trong việc vận dụng chiến lược lịch sự trong giao tiếp. Các kết quả này, vốn được đề cập trong một số công trình đối chiếu Việt – Lào gần đây, cung cấp cơ sở tham chiếu quan trọng cho phân tích thực trạng của nghiên cứu này.

Cuối cùng, hướng tiếp cận CALL/TELL (Computer-Assisted Language Learning / Technology-Enhanced Language Learning) cho thấy việc ứng dụng công nghệ, đặc biệt là trí tuệ nhân tạo, có thể hỗ trợ cá nhân hóa lộ trình học, cung cấp phản hồi phát âm theo thời gian thực và mở rộng cơ hội giao tiếp cho người học (Chapelle, 2016). Tuy nhiên, khi triển khai trong bối cảnh đào tạo an ninh, vấn đề bảo mật dữ liệu và quản lý truy cập phải được đặt lên hàng đầu.

Như vậy, cơ sở lý thuyết không chỉ giúp giải thích những khó khăn đặc thù của học viên Lào trong quá trình học tiếng Việt, mà còn mở ra định hướng ứng dụng trí tuệ nhân tạo trong giảng dạy, theo cách vừa hiệu quả về mặt sư phạm, vừa phù hợp với yêu cầu bảo mật của HVANND.

3. PHƯƠNG PHÁP NGHIÊN CỨU

Nghiên cứu tập trung vào nhóm học viên Lào đang học tiếng Việt cơ sở với sự khác biệt đáng kể về nền tảng ngôn ngữ và trình độ đầu vào. Phạm vi khảo sát giới hạn ở giai đoạn đầu, khi học viên chủ yếu rèn luyện các kỹ năng cơ bản như phát âm, xưng hô, hành vi giao tiếp tại lớp học;

Dữ liệu được thu thập từ hai nguồn chính: (i) quan sát lớp học để ghi nhận có hệ thống các lỗi ngôn ngữ điển hình và cách thức tổ chức dạy học; (ii) khảo sát bằng bảng hỏi dành cho học viên nhằm định lượng mức độ khó khăn và nhu cầu học tập. Sự kết hợp này giúp đối chiếu dữ liệu định tính và định lượng, bảo đảm tính khách quan và độ tin cậy.

Số liệu định lượng từ bảng hỏi được xử lý bằng phương pháp thống kê mô tả để xác định tần suất và mức độ phổ biến của các khó khăn. Dữ liệu định tính từ quan sát được mã hóa và phân loại theo ba nhóm: ngữ âm, ngữ dụng và tổ chức lớp học, tạo cơ sở cho phân tích và thảo luận tiếp theo.

Toàn bộ quy trình thu thập và xử lý dữ liệu được tiến hành trong phạm vi nghiên cứu, không công bố thông tin cá nhân, đồng thời tuân thủ nghiêm ngặt yêu cầu bảo mật trong môi trường đào tạo an ninh. Do giới hạn thời gian, nghiên cứu chưa tích hợp phân tích giáo trình; đây là giới hạn cần được bổ sung ở các công trình tiếp theo.

4. KẾT QUẢ VÀ THẢO LUẬN

4.1. Thực trạng

Kết quả khảo sát được trình bày trong Bảng và Biểu đồ dưới đây cho thấy những khó khăn điển hình và nhu cầu học tập của học viên Lào tại HVANND.

+ Về ngữ âm, 72% thường nhầm lẫn giữa thanh hỏi và thanh ngã, trong khi 63% chưa phát âm chính xác các phụ âm cuối như */t/, /k/*. Đây là lỗi phản ánh sự khác biệt cơ bản giữa hệ thống âm vị tiếng Việt và tiếng Lào, cũng từng được ghi nhận trong một số nghiên cứu đối chiếu gần đây.

+ Về ngữ dụng, 68% bộc lộ sự lúng túng trong việc lựa chọn đại từ nhân xưng phù hợp. Điều này không chỉ gây trở ngại trong hội thoại mà còn ảnh hưởng đến việc hình thành chuẩn mực giao tiếp – một yêu cầu đặc biệt trong môi trường đào tạo an ninh.

+ Ở phương diện giáo trình, 65% cho rằng tài liệu hiện hành chưa tương thích với năng lực cá nhân, khiến việc theo kịp tiến độ chung gặp nhiều khó khăn.

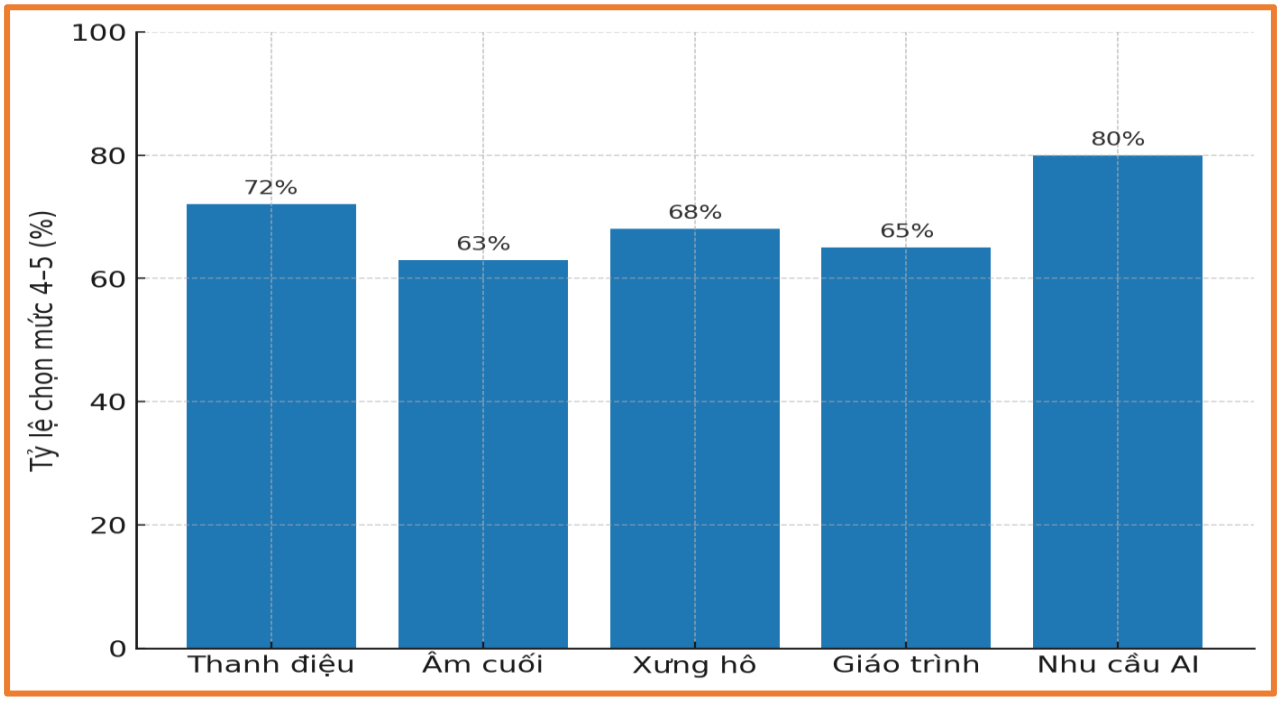
+ Đáng chú ý, nhu cầu công cụ hỗ trợ học tập chiếm tỷ lệ cao nhất (80%). Người học mong muốn được tiếp cận ứng dụng trí tuệ nhân tạo (AI) để luyện phát âm và tăng cường kỹ năng giao tiếp. Xu hướng này cũng phù hợp với những nhận định gần đây trong nghiên cứu về CALL/TELL.

Tổng thể, kết quả cho thấy ba nhóm khó khăn chính của người học: (i) ngữ âm, (ii) ngữ dụng, (iii) học liệu và tổ chức lớp học. Đồng thời, dữ liệu phản ánh nhu cầu đổi mới phương pháp. Trong đó, ứng dụng AI nổi lên như một hướng đi tiềm năng để nâng cao hiệu quả giảng dạy tiếng Việt cho học viên Lào tại HVANND.

Lĩnh vực khảo sát	Câu hỏi minh họa	Tỷ lệ học viên chọn mức 4–5 (%)
Ngữ âm - Thanh điệu	Thường nhầm lẫn thanh hỏi và ngã	72%
Ngữ âm - Âm cuối	Khó phát âm đúng phụ âm cuối: /t/, /k/	63%
Ngữ dụng - Xưng hô	Lúng túng khi chọn đại từ nhân xưng	68%
Giáo trình	Giáo trình chưa phù hợp trình độ cá nhân	65%

Nhu cầu công cụ hỗ trợ	Mong muốn có ứng dụng AI hỗ trợ phát âm, giao tiếp	80%

Bảng Kết quả khảo sát khó khăn và nhu cầu học tiếng Việt của học viên Lào (n=30)



Biểu đồ minh họa Kết quả khảo sát

Những kết quả khảo sát trên cho thấy ba nhóm khó khăn nổi bật trong quá trình học tiếng Việt của học viên Lào. Các số liệu này cần được phân tích sâu hơn từ góc độ ngôn ngữ học đối chiếu và lý thuyết tiếp thu ngôn ngữ thứ hai (SLA), nhằm làm sáng tỏ cơ chế hình thành các khó khăn đó; nội dung sẽ được trình bày trong mục tiếp theo.

4.2. Phân tích nguyên nhân

Những khó khăn và nhu cầu của học viên Lào có thể được lý giải từ hai góc độ: lý thuyết tiếp thu ngôn ngữ thứ hai (SLA) và ngôn ngữ học đối chiếu Việt – Lào.

Về ngữ âm, các lỗi thanh điệu và phụ âm cuối chủ yếu bắt nguồn từ *ảnh hưởng chuyển di của tiếng mẹ đẻ* (L1 transfer). Do tiếng Lào có hệ thống thanh điệu đơn giản hơn và số lượng âm cuối hạn chế, người học dễ đồng hóa âm vị tiếng Việt vào cấu trúc quen thuộc của L1, dẫn đến biến dạng phát âm. Đây là hiện tượng điển hình trong SLA, được Krashen (1985) và nhiều nghiên cứu sau này ghi nhận, đồng thời cũng từng được đề cập trong một số nghiên cứu đối chiếu Việt – Lào gần đây.

Đối với ngữ dụng, sự lúng túng trong lựa chọn đại từ nhân xưng phản ánh khác biệt văn hóa – xã hội. Tiếng Việt có hệ thống xưng hô đa dạng, phụ thuộc vào tuổi tác, quan hệ và bối cảnh, trong khi tiếng Lào ít biến thể hơn. Chính sự chênh lệch này khiến người học khó xác định cách xưng hô phù hợp

Ở khía cạnh học liệu, tình trạng giáo trình chưa tương thích với năng lực cá nhân liên quan trực tiếp đến nguyên tắc *"đầu vào dễ hiểu"* (comprehensible input) trong SLA. Khi tài liệu vượt quá khả năng hoặc quá đơn giản, quá trình tiếp thu đều bị cản trở. Trong lớp học đa trình độ, hạn chế này càng rõ rệt.

Cuối cùng, nhu cầu công cụ công nghệ phản ánh cơ chế *"đầu ra được củng cố"* (pushed output). Người học cần cơ hội luyện tập, nhận phản hồi tức thì và tăng cường tương tác. Khi lớp học truyền thống chưa đáp ứng được, kỳ vọng vào AI nổi lên như một giải pháp phù hợp. Xu hướng này cũng phù hợp với nhận định của nhiều nghiên cứu CALL/TELL gần đây (ví dụ: Chapelle, 2020) về vai trò của công nghệ trong dạy – học ngoại ngữ.

Tóm lại, các khó khăn của học viên Lào vừa bắt nguồn từ khác biệt hệ thống ngôn ngữ và văn hóa Việt – Lào, vừa từ điều kiện học tập cụ thể (lớp học đa trình độ, hạn chế công cụ hỗ trợ). Đây chính là nền tảng để xác định nhu cầu đổi mới phương pháp, trong đó tích hợp AI được xem là một hướng tiếp cận khả thi và phù hợp với đặc thù đào tạo tại HVANND.

4.3. Thảo luận

Kết quả nghiên cứu cho thấy những khó khăn của học viên Lào khi học tiếng Việt có nhiều điểm tương đồng với các phát hiện trước đây trong lĩnh vực dạy tiếng Việt như ngoại ngữ. Các lỗi điển hình về thanh điệu, phụ âm cuối, cũng như sự lúng túng trong lựa chọn đại từ nhân xưng, đều phản ánh đặc điểm phổ biến của người học có nền tảng ngôn ngữ gần gũi với tiếng Việt. Điều này khẳng định rằng những hiện tượng khảo sát được không phải cá biệt, mà mang tính hệ thống trong quá trình tiếp thu tiếng Việt như ngôn ngữ thứ hai.

Tuy nhiên, nghiên cứu này cũng chỉ ra những đóng góp mới. ***Thứ nhất,*** khó khăn ngữ dụng của học viên Lào tại HVANND gắn trực tiếp với yêu cầu giao tiếp chuẩn mực trong môi trường an ninh – nơi việc lựa chọn xưng hô và cách nói không chỉ mang tính ngôn ngữ, mà còn liên quan đến chuẩn mực nghề nghiệp và văn hóa tổ chức. ***Thứ hai***, nhu cầu công cụ hỗ trợ học tập không chỉ nhằm nâng cao hiệu quả, mà còn góp phần bảo đảm an toàn thông tin và bảo mật dữ liệu – một khía cạnh ít được bàn tới trong các nghiên cứu dạy tiếng Việt như ngoại ngữ trước đây. Kết quả nghiên cứu gợi mở khả năng ứng dụng AI theo hướng vừa hỗ trợ đổi mới phương pháp giảng dạy, vừa phù hợp với những yêu cầu đặc thù về bảo mật và kỷ luật trong môi trường đào tạo an ninh.

5. ĐỀ XUẤT MÔ HÌNH ỨNG DỤNG AI CHO HỌC LIỆU TƯƠNG TÁC VÀ CÁ NHÂN HÓA

5.1. Nguyên tắc định hướng

Mô hình ứng dụng AI trong đào tạo tiếng Việt cho học viên Lào tại Học viện An ninh Nhân dân cần được xây dựng dựa trên các nguyên tắc sau:

Thứ nhất, bảo đảm đặc thù của môi trường đào tạo an ninh. Toàn bộ hệ thống phải tuân thủ các yêu cầu về bảo mật dữ liệu, quản lý truy cập và duy trì kỷ luật học tập. Điều này không chỉ là

yếu tố kỹ thuật mà còn là tiêu chuẩn bắt buộc nhằm giữ vững tính an toàn thông tin trong đào tạo.

Thứ hai, lấy người học làm trung tâm, phát huy vai trò hỗ trợ của AI. AI được sử dụng như một công cụ để cá nhân hóa nội dung học tập, cung cấp phản hồi tức thì và gợi ý lộ trình phù hợp với từng cá nhân. Tuy nhiên, vai trò của giảng viên không bị thay thế mà được nâng cao: từ người truyền đạt kiến thức sang người định hướng, cố vấn và giám sát quá trình học tập.

Thứ ba, tích hợp công nghệ với phương pháp sư phạm. Việc ứng dụng AI chỉ phát huy hiệu quả khi được kết hợp với các phương pháp giảng dạy truyền thống như luyện tập, thảo luận nhóm, giảng giải. AI hỗ trợ mở rộng không gian và thời gian học tập, nhưng sự tương tác người – người vẫn giữ vai trò cốt lõi trong việc hình thành năng lực giao tiếp tiếng Việt.

Những nguyên tắc trên bảo đảm rằng mô hình AI vừa tận dụng được ưu thế công nghệ, vừa duy trì chuẩn mực đào tạo và phù hợp với yêu cầu đặc thù của HVANND.

5.2. Thành phần hệ thống

Mô hình ứng dụng AI trong đào tạo tiếng Việt cho học viên Lào tại Học viện An ninh Nhân dân được cấu trúc với bốn hợp phần cốt lõi, mỗi hợp phần gắn trực tiếp với những khó khăn đã được nhận diện trong khảo sát và phân tích.

Thứ nhất, chatbot hội thoại chuyên biệt. Công cụ này được thiết kế dựa trên ngữ liệu tiếng Việt cơ sở và các tình huống giao tiếp điển hình trong môi trường học tập an ninh. Chatbot hỗ trợ học viên luyện tập xưng hô, thực hành hội thoại cơ bản và phân biệt sắc thái lịch sự trong lời nói. Việc mô phỏng tương tác với chatbot giúp giảm bớt áp lực tâm lý khi giao tiếp trực tiếp, đồng thời tạo điều kiện để học viên lặp lại nhiều lần – một yếu tố then chốt trong hình thành năng lực ngữ dụng.

Thứ hai, hệ thống nhận diện và phản hồi phát âm (STT/TTS). Khi học viên phát âm, hệ thống sẽ phân tích thanh điệu, phụ âm cuối và nguyên âm tròn môi, sau đó trả về phản hồi tức thì. Nhờ đó, các lỗi ngữ âm đặc thù – chẳng hạn biến đổi *"bác"* → *"ba"* hay *"mát"* → *"ma"* – được phát hiện và sửa kịp thời. Công cụ này mở rộng cơ hội luyện tập ngoài lớp, khắc phục hạn chế của lớp học đông, nơi giảng viên khó theo sát từng cá nhân.

Thứ ba, kho học liệu số cá nhân hóa. AI phân tích dữ liệu học tập (kết quả bài tập, điểm kiểm tra, tần suất lỗi) để gợi ý nội dung phù hợp với từng học viên. Qua đó, người học được phân nhóm theo nhu cầu (ví dụ: nhóm tập trung vào ngữ âm, nhóm tập trung vào ngữ dụng), còn giảng viên có thể giám sát tiến độ và phát hiện sớm những điểm yếu. Kho học liệu này không chỉ cung cấp tài nguyên mà còn đóng vai trò là "bản đồ học tập" cho từng cá nhân.

Thứ tư, hệ thống trò chơi hóa và kiểm tra – đánh giá. Các nền tảng như Quizziz, Kahoot hoặc phần mềm nội bộ được tích hợp nhằm tạo trò chơi ngôn ngữ và bài kiểm tra định kỳ. Cách tiếp cận này giúp củng cố kiến thức, duy trì hứng thú học tập và tạo động lực thi đua trong lớp. Toàn bộ kết quả được lưu trữ tự động trong hồ sơ học tập số, cung cấp dữ liệu tin cậy cho cả giảng viên và người học trong quá trình đánh giá.

Sự phối hợp đồng bộ của bốn hợp phần trên không chỉ phù hợp với định hướng của CALL/TELL về việc tích hợp công nghệ trong học ngoại ngữ, mà còn được điều chỉnh để đáp ứng yêu cầu đặc thù của HVANND, nơi yếu tố bảo mật và kiểm soát dữ liệu học tập luôn được đặt lên hàng đầu.

5.3. Quy trình hoạt động

Quy trình vận hành của mô hình AI được thiết kế như một chu trình khép kín, gồm bốn bước gắn kết chặt chẽ. Mở đầu, học viên đăng nhập vào hệ thống nội bộ và thực hiện bài kiểm tra ngắn về ngữ âm và ngữ dụng. Ví dụ, khi đọc một đoạn hội thoại đơn giản, hệ thống lập tức phân tích thanh điệu và phụ âm cuối, từ đó xác định trình độ ban đầu (*diagnostic assessment*).

Tiếp đến, AI cá nhân hóa lộ trình học tập: học viên sai nhiều về phát âm sẽ được tăng cường luyện tập với công cụ STT/TTS, còn người lúng túng ngữ dụng được ưu tiên thực hành hội thoại với chatbot. Đây là cách hiện thực hóa nguyên lý *comprehensible input* trong SLA.

Trong quá trình luyện tập, học viên thường xuyên tương tác với chatbot để xử lý tình huống cụ thể, chẳng hạn lựa chọn xưng hô trong lời chào cấp trên. Nếu học viên dùng "các em chào thầy", hệ thống sẽ phản hồi và gợi ý sửa thành "chúng em chào thầy". Cơ chế sửa lỗi tức thì này minh họa cho *pushed output*, khi người học buộc điều chỉnh để phù hợp với chuẩn mực ngôn ngữ – văn hóa. Song song, công cụ STT/TTS chỉ rõ các lỗi ngữ âm đặc thù như *"bác → ba"* hay *"mát → ma"*.

Kết quả luyện tập được lưu trong hồ sơ học tập số, được quản lý hoàn toàn trong hệ thống nội bộ để bảo đảm an toàn thông tin. Giảng viên truy cập báo cáo tiến độ, đưa ra phản hồi chuyên sâu và điều chỉnh chiến lược giảng dạy (*formative assessment*). Vai trò này đặc biệt quan trọng tại HVANND, nơi giáo viên không chỉ là người giám sát mà còn định hình chuẩn mực giao tiếp phù hợp với môi trường an ninh.

Chu trình đánh giá – cá nhân hóa – luyện tập – phản hồi vừa kế thừa các nguyên tắc CALL/TELL quốc tế, vừa được điều chỉnh để đáp ứng yêu cầu bảo mật và kỷ luật. Nhờ vậy, mô hình có tính khả thi cao, đồng thời khẳng định giá trị ứng dụng trong bối cảnh đào tạo tiếng Việt đặc thù.

5.4. Khả năng triển khai

Khả năng triển khai mô hình ứng dụng AI trong đào tạo tiếng Việt cho học viên Lào tại Học viện An ninh Nhân dân có thể được nhìn nhận theo một lộ trình tiến dần từ những bước thử nghiệm nhỏ đến phát triển toàn diện.

Trước hết, trong giai đoạn ngắn hạn, việc áp dụng các công cụ cơ bản như chatbot hội thoại và phần mềm nhận diện phát âm (STT/TTS) là hoàn toàn khả thi. Chẳng hạn, ở một lớp tiếng Việt cơ sở với khoảng 30 học viên, chatbot có thể hỗ trợ luyện tập xưng hô trong các tình huống giao tiếp cơ bản, trong khi STT/TTS cung cấp phản hồi tức thì về phát âm thanh điệu và phụ âm cuối. Việc này vừa giảm tải khâu sửa lỗi cho giảng viên, vừa tạo điều kiện để học viên luyện tập ngoài giờ.

Tiếp theo, ở giai đoạn trung hạn, Học viện có thể xây dựng kho học liệu số nội bộ và tích hợp trò chơi hóa trong kiểm tra – đánh giá. Kho học liệu này không chỉ đảm bảo an toàn dữ liệu mà còn cho phép điều chỉnh nội dung gắn với chuyên ngành an ninh. Ví dụ, ngoài bài tập ngữ âm cơ bản, học viên có thể thực hành hội thoại trong bối cảnh trực ban, báo cáo sự vụ hoặc giao tiếp trong môi trường kỷ luật. Điều này vừa củng cố năng lực tiếng Việt, vừa giúp học viên sớm thích ứng với đặc thù nghề nghiệp.

Sau cùng, về giai đoạn dài hạn, mục tiêu hướng đến là phát triển một nền tảng AI "Made in Vietnam" dành riêng cho giảng dạy tiếng Việt như ngoại ngữ trong môi trường an ninh.

6. KẾT LUẬN

Chuyển đổi số đặt ra yêu cầu đổi mới đào tạo tiếng Việt cho học viên Lào tại HVANND. Ứng dụng AI mang lại giải pháp khả thi trong việc xây dựng học liệu tương tác và cá nhân hóa, góp phần khắc phục khó khăn ngôn ngữ, tăng cường động lực và hiệu quả học tập. Mô hình đề xuất vừa phù hợp với xu thế giáo dục quốc tế, vừa đáp ứng yêu cầu bảo mật đặc thù của môi trường an ninh. Nghiên cứu không chỉ có giá trị đối với HVANND mà còn có thể tham khảo cho các cơ sở đào tạo tiếng Việt cho người nước ngoài khác./.

TÀI LIỆU THAM KHẢO

Tài liệu tiếng Việt

1. Bộ Giáo dục và Đào tạo (2021*), Chiến lược chuyển đổi số ngành giáo dục giai đoạn 2021–2025, tầm nhìn 2030*, Hà Nội.

2. Bùi Thị Thu Hương & Bùi Văn Nam (2024), *"Ứng dụng công cụ AI trong dạy học ngoại ngữ ở Việt Nam"*, Tạp chí Khoa học Giáo dục, Hà Nội.

Tài liệu tiếng nước ngoài

3. Chapelle, C. (2016), *Teaching Culture with Technology for Learners of English as an International Language*, Cambridge University Press, Cambridge.

4. Chapelle, C. (2020), *Argument-Based Validation in Language Assessment*, Springer, Cham.

5. Chiu, T. K. F., et al. (2024), *"The use of generative AI in language learning: Potentials and challenges",* Language Learning & Technology, University of Hawai'i Press.

6. Fryer, L. & Carpenter, R. (2020), *"Bots as language partners: Chatbots for second language learning"*, Computer Assisted Language Learning, Taylor & Francis.

7. Gass, S. & Selinker, L. (2008), *Second Language Acquisition: An Introductory Course*, Routledge, New York.

8. Kasneci, E., et al. (2023), *"ChatGPT for good? On opportunities and challenges of large language models for education",* Learning and Individual Differences, Elsevier.

9. Krashen, S. (1985), *The Input Hypothesis: Issues and Implications*, Longman, London.

10. Li, J. & Ni, X. (2021), *"Automatic speech recognition in tonal languages: A case study with Vietnamese learners"*, Speech Communication, Elsevier.

11. Lo, C. K. (2024), *"A framework for integrating generative AI in second language teaching"*, Educational Technology Research and Development, Springer.

12. OECD (2023), *Digital Education Outlook 2023*: AI in Education, OECD Publishing, Paris.

13. Rudolph, J., Tan, S. & Tan, M. (2023), *"Risks of ChatGPT in education: Challenges and opportunities"*, Computers & Education: Artificial Intelligence, Elsevier.

14. World Bank (2022), *Digital Transformation of Education Systems: A Policy Perspective*, World Bank, Washington D.C.

THÔNG TIN TÁC GIẢ

Nguyễn Thụy Trang Th.s Khoa Ngoại ngữ - Học viện An ninh Nhân dân

Đ.c: 125 Trần Phú – Hà Đông – Hà Nội

ỨNG DỤNG CÔNG NGHỆ TRÍ TUỆ NHÂN TẠO (AI) TRONGHỌC PHẦN PHƯƠNG PHÁP PHÁT TRIỂN NGÔN NGỮ CHO TRẺ MẦM NON

Trần Thị Ban Mai[1]

Tóm tắt: Bài hội thảo trình bày các ứng dụng nổi bật của trí tuệ nhân tạo (AI) trong đào tạo giáo viên mầm non, đặc biệt trong học phần "Phương pháp phát triển ngôn ngữ". Các công nghệ như trợ giảng ảo, thực tế ảo/tăng cường (VR/AR) và nhận diện giọng nói giúp sinh viên mô phỏng tình huống giao tiếp, tổ chức hoạt động ngôn ngữ và đánh giá sự phát triển ngôn ngữ của trẻ một cách khách quan. Việc tích hợp AI vào quá trình đào tạo không chỉ nâng cao chất lượng học tập mà còn hỗ trợ sinh viên hình thành kỹ năng sư phạm hiện đại, đáp ứng yêu cầu đổi mới giáo dục trong thời đại số.

Từ khóa: trí tuệ nhân tạo (AI), thực tế ảo, phương pháp phát triển ngôn ngữ, giáo dục mầm non, thực tế ảo.

1. Đặt vấn đề

Trong bối cảnh chuyển đổi số đang diễn ra mạnh mẽ trên toàn cầu, trí tuệ nhân tạo (AI) đã và đang trở thành một công cụ hỗ trợ đắc lực trong nhiều lĩnh vực, đặc biệt là giáo dục. Tại Việt Nam, việc ứng dụng AI trong đào tạo giáo viên mầm non vẫn còn mới mẻ, song tiềm năng phát triển là rất lớn. Học phần "Phương pháp phát triển ngôn ngữ cho trẻ mầm non" là một nội dung cốt lõi trong chương trình đào tạo giáo viên, nhằm trang bị cho sinh viên năng lực tổ chức các hoạt động phát triển ngôn ngữ cho trẻ từ 0–6 tuổi. Tuy nhiên, thực tế giảng dạy học phần này vẫn còn gặp nhiều hạn chế về phương pháp, tài liệu và khả năng mô phỏng tình huống thực tế. Việc tích hợp công nghệ AI vào học phần không chỉ giúp nâng cao hiệu quả giảng dạy mà còn mở ra cơ hội đổi mới toàn diện cách tiếp cận, từ đó góp phần nâng cao chất lượng đào tạo giáo viên mầm non trong thời đại số. Bài viết này nhằm phân tích tiềm năng ứng dụng AI trong học phần "Phương pháp phát triển ngôn ngữ", đề xuất mô hình tích hợp phù hợp với điều kiện đào tạo tại các trường sư phạm, đồng thời đưa ra những kiến nghị cụ thể để thúc đẩy quá trình chuyển đổi số trong giáo dục mầm non.

2. Nội dung

2.1. Phát triển ngôn ngữ ở trẻ mầm non và học phần phương pháp phát triển ngôn ngữ cho trẻ mầm non.

2.1.1. Phát triển ngôn ngữ ở trẻ mầm non

Phát triển ngôn ngữ là quá trình trẻ em tiếp thu, hoàn thiện và sử dụng hệ thống ký hiệu ngôn ngữ để giao tiếp, biểu đạt tư duy, cảm xúc và tương tác xã hội. Theo quan điểm của Vygotsky (1978), ngôn ngữ không chỉ là công cụ giao tiếp mà còn là phương tiện trung gian giữa tư duy và hành động, đóng vai trò thiết yếu trong sự phát triển nhận thức của trẻ. Ngôn

[1] ThS., Trường Cao đẳng Sư phạm Nam Định, *Email: banmaisp2@gmail.com*

ngữ giúp trẻ tiếp cận tri thức xã hội thông qua tương tác với người lớn và bạn bè, từ đó hình thành các khái niệm, quy tắc và giá trị văn hóa.

Bruner (1983) cũng nhấn mạnh rằng ngôn ngữ là phương tiện để trẻ xây dựng ý nghĩa trong các hoạt động xã hội có tính tương tác. Trẻ học ngôn ngữ không chỉ qua việc nghe và bắt chước mà còn thông qua các hoạt động có mục đích như kể chuyện, chơi đóng vai, hát, đọc thơ, từ đó phát triển năng lực ngôn ngữ trong bối cảnh thực tế.

Tại Việt Nam, Nguyễn Ánh Tuyết (2015) định nghĩa phát triển ngôn ngữ ở trẻ mầm non là quá trình hình thành và hoàn thiện các kỹ năng ngôn ngữ bao gồm: nghe hiểu, diễn đạt bằng lời nói, sử dụng ngôn ngữ để giao tiếp và biểu đạt cảm xúc, nhận biết và sử dụng ký hiệu ngôn ngữ như chữ cái và từ ngữ.

Giai đoạn mầm non (0–6 tuổi), đặc biệt từ 3 đến 5 tuổi, được xem là thời kỳ "vàng" trong phát triển ngôn ngữ. Đây là giai đoạn trẻ có khả năng tiếp thu ngôn ngữ một cách tự nhiên, nhanh chóng và hiệu quả thông qua các tương tác xã hội và môi trường sống.

** Một số đặc điểm nổi bật trong phát triển ngôn ngữ ở trẻ mầm non:*

Trẻ em có tính nhạy cảm ngôn ngữ cao: Trẻ dễ dàng ghi nhớ từ mới, bắt chước cách nói và ngữ điệu của người lớn. Theo nghiên cứu của Kuhl (2004), trẻ em có khả năng phân biệt âm thanh ngôn ngữ rất tốt trong những năm đầu đời, đặc biệt khi được tiếp xúc với ngôn ngữ trong môi trường giàu tương tác.

Trẻ mầm non phát triển đồng thời cả ngôn ngữ tiếp nhận và ngôn ngữ biểu đạt: Trẻ không chỉ học cách nghe hiểu mà còn học cách diễn đạt suy nghĩ, cảm xúc bằng lời nói. Quá trình này diễn ra song song và hỗ trợ lẫn nhau.

Phát triển ngôn ngữ ở trẻ gắn liền với hoạt động chơi: Ngôn ngữ của trẻ phát triển mạnh mẽ thông qua các hoạt động chơi đóng vai, kể chuyện, hát, đọc thơ… Đây là những hoạt động có tính biểu tượng cao, giúp trẻ sử dụng ngôn ngữ trong các tình huống giả định, từ đó mở rộng vốn từ và cấu trúc câu.

Sự phát triển ngôn ngữ ở trẻ phụ thuộc vào môi trường giao tiếp: Chất lượng và số lượng tương tác ngôn ngữ mà trẻ nhận được từ người lớn và bạn bè có ảnh hưởng trực tiếp đến tiến trình phát triển ngôn ngữ. Trẻ em trong môi trường giao tiếp phong phú thường có vốn từ vựng và khả năng diễn đạt tốt hơn.

Phát triển ngôn ngữ không chỉ là mục tiêu riêng lẻ mà còn là nền tảng cho sự phát triển toàn diện của trẻ mầm non. Ngôn ngữ đóng vai trò trung tâm trong các lĩnh vực phát triển khác như nhận thức, tình cảm – xã hội, thể chất và thẩm mỹ.

** Phát triển ngôn ngữ có vai trò:*

Ngôn ngữ là công cụ giao tiếp và tương tác xã hội: Ngôn ngữ giúp trẻ diễn đạt nhu cầu, cảm xúc, suy nghĩ và thiết lập mối quan hệ với người khác. Trẻ có năng lực ngôn ngữ tốt thường tự tin hơn trong giao tiếp, dễ dàng hòa nhập vào môi trường học tập và xã hội.

Ngôn ngữ là phương tiện biểu đạt tư duy: Theo Vygotsky (1978), ngôn ngữ là phương tiện để trẻ tổ chức tư duy, giải quyết vấn đề và học tập. Trẻ sử dụng ngôn ngữ để đặt câu hỏi, suy luận, mô tả hiện tượng và phản ánh trải nghiệm cá nhân.

Ngôn ngữ là cơ sở cho việc hình thành kỹ năng học tập: Ngôn ngữ là nền tảng cho việc học đọc, viết và tiếp thu kiến thức trong các lĩnh vực học thuật. Trẻ có vốn từ vựng phong phú và khả năng diễn đạt tốt thường có kết quả học tập cao hơn ở các cấp học sau.

Ngôn ngữ góp phần hình thành nhân cách: Thông qua ngôn ngữ, trẻ tiếp thu các chuẩn mực xã hội, học cách thể hiện cảm xúc, kiểm soát hành vi và phát triển bản sắc cá nhân. Ngôn ngữ giúp trẻ hiểu được vai trò của mình trong cộng đồng và xây dựng lòng tự trọng.

** Định hướng phát triển ngôn ngữ trong chương trình giáo dục mầm non*

Chương trình giáo dục mầm non cần xây dựng môi trường ngôn ngữ phong phú, đa dạng và khuyến khích trẻ sử dụng ngôn ngữ trong các tình huống thực tế.

Thứ nhất, tạo cơ hội cho trẻ giao tiếp đa chiều: Giáo viên cần thiết kế các hoạt động khuyến khích trẻ giao tiếp với bạn bè, người lớn và môi trường xung quanh. Các hoạt động như thảo luận nhóm, kể chuyện, chơi đóng vai… giúp trẻ sử dụng ngôn ngữ một cách linh hoạt và sáng tạo.

Thứ ba, phát triển đồng thời cả ngôn ngữ nói và ngôn ngữ viết cho trẻ: Trẻ cần được làm quen với chữ cái, từ ngữ, sách truyện… để hình thành kỹ năng đọc – viết ban đầu. Việc kết hợp giữa nghe – nói – đọc – viết giúp trẻ phát triển ngôn ngữ toàn diện.

Thứ ba, cá nhân hóa quá trình phát triển ngôn ngữ ở trẻ: Mỗi trẻ có tốc độ và phong cách phát triển ngôn ngữ khác nhau. Giáo viên cần quan sát, đánh giá và xây dựng kế hoạch hỗ trợ phù hợp với từng trẻ.

Thứ tư, ứng dụng công nghệ hỗ trợ phát triển ngôn ngữ: Các công cụ công nghệ như phần mềm kể chuyện, trò chơi ngôn ngữ, trợ giảng ảo… có thể hỗ trợ giáo viên trong việc tạo môi trường ngôn ngữ hấp dẫn và hiệu quả.

2.1.2. Học phần phương pháp phát triển ngôn ngữ cho trẻ

Học phần "Phương pháp phát triển ngôn ngữ" là một nội dung cốt lõi trong chương trình đào tạo giáo viên mầm non tại các trường sư phạm. Mục tiêu của học phần là trang bị cho sinh viên kiến thức lý thuyết và kỹ năng thực hành để tổ chức các hoạt động phát triển ngôn ngữ cho trẻ phù hợp với độ tuổi, đặc điểm tâm sinh lý và môi trường giáo dục.

Nội dung học phần bao gồm những kiến thức cơ bản về: nhiệm vụ, hình thức và phương pháp phát triển ngôn ngữ trẻ em; giáo dục chuẩn mực ngữ âm tiếng Việt; hình thành và phát triển vốn từ; dạy trẻ sử dụng các mẫu câu tiếng Việt; phát triển ngôn ngữ mạch lạc; phát triển ngôn ngữ nghệ thuật cho trẻ qua

tác phẩm văn học; phát triển khả năng tiền đọc, viết cho trẻ và hỗ trợ phát triển ngôn ngữ cho trẻ mầm non khuyết tật. Cùng với hệ thống kiến thức đó, học phần có các nội dung thực hành thiết kế kế hoạch và tổ chức thực hiện hoạt động phát triển ngôn ngữ cho trẻ theo các chủ đề giáo dục, trong đó có lồng ghép một số nội dung về giáo dục giới tính, bình đẳng giới và quyền trẻ em, đồng thời ứng dụng một số phương pháp, mô hình giáo dục tiên tiến (Montessori, STEAM) đáp ứng thực tế đổi mới giáo dục mầm non.

Theo khảo sát của Nguyễn Thị Hồng (2020), sinh viên sư phạm mầm non thường gặp khó khăn trong việc vận dụng lý thuyết vào thực hành, đặc biệt là trong việc thiết kế hoạt động ngôn ngữ phù hợp với từng độ tuổi. Việc giảng dạy học phần này chủ yếu vẫn theo phương pháp truyền thống, thiếu các công cụ hỗ trợ mô phỏng tình huống thực tế, dẫn đến hạn chế trong việc phát triển năng lực sư phạm của sinh viên như: thiếu tài liệu học tập có tính thực tiễn cao; sinh viên ít cơ hội tiếp xúc với trẻ thật để thực hành; sinh viên chưa được rèn luyện kỹ năng giao tiếp sư phạm một cách bài bản Do đó, việc đổi mới phương pháp giảng dạy học phần này là cần thiết, đặc biệt trong bối cảnh chuyển đổi số và sự phát triển mạnh mẽ của công nghệ giáo dục.

2.2. Trí tuệ nhân tạo (AI) và tiềm năng ứng dụng trong giáo dục mầm non

2.2.1. Khái quát về trí tuệ nhân tạo trong giáo dục

Trí tuệ nhân tạo (Artificial Intelligence – AI) là lĩnh vực công nghệ mô phỏng các quá trình tư duy, học hỏi, xử lý ngôn ngữ và ra quyết định của con người thông qua hệ thống máy tính. Trong giáo dục, AI được xem là một công cụ hỗ trợ đổi mới phương pháp giảng dạy, cá nhân hóa quá trình học tập và nâng cao hiệu quả đánh giá người học (Luckin et al., 2016). Theo OECD (2021), AI có khả năng phân tích dữ liệu học tập, mô phỏng tình huống sư phạm, hỗ trợ phản hồi tức thời và tạo môi trường học tập tương tác, từ đó góp phần cải thiện chất lượng giáo dục ở các cấp học.

Trong bối cảnh giáo dục mầm non, việc ứng dụng AI cần được xem xét một cách thận trọng, bởi đặc điểm phát triển tâm – sinh lý của trẻ nhỏ đòi hỏi sự tương tác trực tiếp, cảm xúc và tính nhân văn cao. Tuy nhiên, AI vẫn có thể đóng vai trò hỗ trợ đắc lực trong việc thiết kế môi trường học tập, đánh giá ngôn ngữ và nâng cao năng lực sư phạm cho giáo viên, đặc biệt trong các học phần mang tính thực hành như “Phương pháp phát triển ngôn ngữ”.

2.2.2. Những nguyên tắc khi ứng dụng AI trong giáo dục mầm non

Việc ứng dụng trí tuệ nhân tạo (AI) trong giáo dục mầm non mang lại nhiều tiềm năng đổi mới phương pháp giảng dạy, nâng cao hiệu quả học tập và hỗ trợ cá nhân hóa quá trình phát triển của trẻ. Tuy nhiên, do đặc thù của lứa tuổi mầm non – giai đoạn hình thành nền tảng nhân cách, nhận thức và cảm xúc – việc triển khai công nghệ AI cần được thực hiện một cách thận trọng, có kiểm soát và tuân thủ các nguyên tắc đạo đức, sư phạm và phát triển phù hợp. Theo UNESCO (2022), ba nguyên tắc cốt lõi cần được đảm bảo khi ứng dụng AI trong giáo dục mầm non bao gồm: tính nhân văn, bảo mật dữ liệu và sự phù hợp với đặc điểm phát triển của trẻ.

2.2.2.1. Tính nhân văn trong ứng dụng AI

Nguyên tắc đầu tiên và quan trọng nhất là đảm bảo tính nhân văn trong quá trình ứng dụng AI. Trẻ mầm non phát triển mạnh mẽ thông qua tương tác cảm xúc, sự kết nối xã hội và trải nghiệm thực tế. Do đó, AI không thể và không nên thay thế vai trò của giáo viên – người đóng vai trò trung tâm trong việc hướng dẫn, chăm sóc và giáo dục trẻ.

Theo Luckin et al. (2016), AI nên được thiết kế như một công cụ hỗ trợ, giúp giáo viên nâng cao năng lực giảng dạy, chứ không phải là người thay thế. Trong môi trường mầm non, giáo viên không chỉ truyền đạt kiến thức mà còn là người đồng hành, tạo dựng mối quan hệ tin cậy, an toàn và đầy cảm xúc với trẻ. Việc sử dụng AI cần đảm bảo rằng trẻ vẫn được tương tác trực tiếp với con người, được lắng nghe, được phản hồi bằng cảm xúc thật – điều mà công nghệ chưa thể tái tạo hoàn toàn.

Ví dụ, một trợ giảng ảo có thể giúp sinh viên sư phạm luyện tập kỹ năng giao tiếp, nhưng trong lớp học thực tế, chính sự ân cần, ánh mắt, giọng nói và cử chỉ của giáo viên mới là yếu tố quyết định đến sự phát triển cảm xúc và ngôn ngữ của trẻ. Do đó, AI cần được tích hợp theo hướng hỗ trợ giáo viên cá nhân hóa hoạt động, theo dõi tiến trình phát triển của trẻ, chứ không làm lu mờ vai trò của người giáo dục.

2.2.2.2. Bảo mật dữ liệu và quyền riêng tư của trẻ

Trẻ em là nhóm đối tượng đặc biệt, chưa có đầy đủ năng lực pháp lý để tự bảo vệ quyền riêng tư và dữ liệu cá nhân. Việc thu thập, lưu trữ và xử lý dữ liệu liên quan đến trẻ – bao gồm giọng nói, hình ảnh, hành vi, phản ứng… – cần tuân thủ nghiêm ngặt các quy định pháp luật và chuẩn mực đạo đức.

Theo OECD (2021), các hệ thống AI trong giáo dục cần được thiết kế với cơ chế bảo vệ dữ liệu mạnh mẽ, đảm bảo rằng thông tin cá nhân của trẻ không bị sử dụng sai mục đích, không bị chia sẻ trái phép và không bị khai thác thương mại. Điều này đặc biệt quan trọng trong bối cảnh nhiều nền tảng công nghệ giáo dục hiện nay có khả năng kết nối trực tuyến, lưu trữ đám mây và phân tích dữ liệu lớn.

Trong môi trường đào tạo giáo viên, sinh viên cần được hướng dẫn rõ ràng về quy trình thu thập dữ liệu trẻ, cách sử dụng phần mềm AI một cách an toàn và có đạo đức. Các cơ sở đào

tạo cần ban hành quy chế bảo mật, có sự giám sát và kiểm tra định kỳ để đảm bảo rằng việc ứng dụng AI không vi phạm quyền riêng tư của trẻ.

Ví dụ, khi sinh viên sử dụng phần mềm nhận diện giọng nói để phân tích ngôn ngữ của trẻ, cần có sự đồng thuận từ phụ huynh, đảm bảo dữ liệu được mã hóa, không lưu trữ lâu dài và không chia sẻ ra ngoài hệ thống đào tạo.

2.2.2.3. Phù hợp với đặc điểm phát triển của trẻ mầm non

Nguyên tắc thứ ba là đảm bảo sự phù hợp giữa công nghệ AI và đặc điểm phát triển tâm lý – nhận thức của trẻ mầm non. Trẻ ở độ tuổi này có khả năng chú ý ngắn, dễ bị kích thích quá mức bởi hình ảnh, âm thanh và chưa có khả năng tự điều chỉnh hành vi khi sử dụng thiết bị công nghệ. Do đó, việc thiết kế và triển khai AI cần được điều chỉnh để phù hợp với khả năng tiếp nhận của trẻ, tránh gây quá tải, lệ thuộc hoặc ảnh hưởng tiêu cực đến sự phát triển toàn diện.

Theo nghiên cứu của Hirsh-Pasek et al. (2015), trẻ mầm non học tốt nhất thông qua tương tác trực tiếp, hoạt động thể chất và trải nghiệm thực tế. Việc sử dụng công nghệ cần được tích hợp một cách hợp lý, không thay thế hoàn toàn các hoạt động truyền thống như chơi, vận động, giao tiếp mặt đối mặt. AI nên được sử dụng như một phần bổ trợ trong môi trường học tập, giúp tăng cường trải nghiệm nhưng không làm trẻ mất đi cơ hội tương tác thực tế.

Ví dụ, một phần mềm kể chuyện có thể giúp trẻ tiếp cận ngôn ngữ mới, nhưng nếu sử dụng quá thường xuyên, trẻ có thể giảm khả năng tương tác với bạn bè, giảm nhu cầu giao tiếp bằng lời nói và trở nên lệ thuộc vào thiết bị. Do đó, giáo viên cần kiểm soát thời lượng sử dụng, lựa chọn nội dung phù hợp và kết hợp với các hoạt động tương tác trực tiếp để đảm bảo sự phát triển cân bằng.

Việc ứng dụng AI trong giáo dục mầm non cần được đặt trong khuôn khổ đạo đức, sư phạm và phát triển phù hợp. Ba nguyên tắc cốt lõi – tính nhân văn, bảo mật dữ liệu và sự phù hợp với đặc điểm phát triển – không chỉ là yêu cầu kỹ thuật mà còn là định hướng chiến lược để đảm bảo rằng công nghệ phục vụ con người, đặc biệt là trẻ em, một cách an toàn và hiệu quả. Các cơ sở đào tạo giáo viên mầm non cần xây dựng chương trình giảng dạy tích hợp AI một cách có kiểm soát, đồng thời trang bị cho sinh viên kiến thức về đạo đức công nghệ, kỹ năng sử dụng AI có trách nhiệm và khả năng đánh giá tác động của công nghệ đến sự phát triển của trẻ. Chỉ khi đó, AI mới thực sự trở thành công cụ hỗ trợ đắc lực trong việc nâng cao chất lượng giáo dục mầm non trong thời đại số.

2.3. Các hình thức ứng dụng AI trong phát triển ngôn ngữ cho trẻ mầm non

2.3.1. Nhận diện và phân tích giọng nói của trẻ

Công nghệ nhận diện giọng nói (speech recognition) là một trong những thành tựu nổi bật của trí tuệ nhân tạo trong lĩnh vực xử lý ngôn ngữ tự nhiên. Hệ thống AI có thể ghi nhận, phân tích và đánh giá các đặc điểm ngôn ngữ của trẻ như tốc độ nói, độ rõ ràng, vốn từ vựng, cấu trúc câu, ngữ điệu và tần suất sử dụng từ ngữ. Theo nghiên cứu của Li, Zhang và Wang

(2020), các phần mềm AI ứng dụng công nghệ nhận diện giọng nói có thể phát hiện sớm các dấu hiệu chậm phát triển ngôn ngữ hoặc rối loạn ngôn ngữ ở trẻ thông qua phân tích âm thanh đầu vào. Cụ thể, hệ thống có thể nhận diện các biểu hiện như nói ngọng, nói lắp, sử dụng từ sai ngữ cảnh hoặc thiếu phản xạ ngôn ngữ trong giao tiếp.

Một ví dụ điển hình là phần mềm **LENA System**, được sử dụng rộng rãi tại Mỹ trong các chương trình can thiệp sớm. Hệ thống này ghi âm các cuộc trò chuyện giữa trẻ và người lớn, sau đó phân tích số lượng từ vựng, lượt tương tác và thời lượng nói của trẻ. Kết quả phân tích giúp giáo viên và phụ huynh nhận biết mức độ phát triển ngôn ngữ của trẻ so với chuẩn phát triển theo độ tuổi.

** Ứng dụng trong học phần "Phương pháp phát triển ngôn ngữ":*

Sinh viên ngành GD Mầm non có thể sử dụng công cụ AI để phân tích mẫu giọng nói của trẻ, từ đó học cách đánh giá và xây dựng kế hoạch can thiệp phù hợp. Ví dụ, khi phân tích một đoạn ghi âm trẻ kể chuyện, AI có thể chỉ ra các lỗi phát âm, thiếu từ nối hoặc cấu trúc câu chưa hoàn chỉnh, giúp sinh viên xác định mục tiêu can thiệp cụ thể.

AI giúp chuẩn hóa quy trình đánh giá ngôn ngữ, giảm sai lệch chủ quan từ người đánh giá. Trong thực tế, việc đánh giá ngôn ngữ trẻ thường phụ thuộc vào cảm nhận cá nhân của giáo viên, dễ dẫn đến sai lệch. Việc sử dụng AI như một công cụ hỗ trợ giúp tăng độ tin cậy và tính khách quan trong đánh giá.

Ngoài ra, công nghệ này còn hỗ trợ việc xây dựng hồ sơ phát triển ngôn ngữ cho từng trẻ, từ đó giúp sinh viên hiểu rõ tiến trình phát triển ngôn ngữ và điều chỉnh phương pháp giảng dạy phù hợp.

Ví dụ: Phân tích đoạn ghi âm trẻ kể chuyện

Tình huống: Sinh viên thu âm đoạn trẻ 5 tuổi kể lại câu chuyện "Cô bé quàng khăn đỏ".

Phân tích từ phần mềm AI (ví dụ: LENA System hoặc ứng dụng NLP tùy chỉnh):

– Tốc độ nói: 80 từ/phút (chậm hơn chuẩn tuổi)

– Lỗi phát âm: "khăn" → "khan", "sói" → "xói"

– Cấu trúc câu: thiếu từ nối, câu đơn lặp lại

– Vốn từ vựng: 45 từ duy nhất trong toàn đoạn

Phản ứng của sinh viên:

– Xác định mục tiêu can thiệp: luyện phát âm phụ âm đầu, mở rộng vốn từ

– Thiết kế hoạt động: trò chơi "Ghép từ đúng", kể chuyện theo tranh có từ nối

Lợi ích: Sinh viên học cách sử dụng dữ liệu ngôn ngữ để xây dựng kế hoạch giảng dạy phù hợp với từng trẻ.

2.3.2. Mô phỏng tình huống giao tiếp qua trợ giảng ảo

Trợ giảng ảo (virtual assistant) là hệ thống AI có khả năng tương tác bằng lời nói với người học, sử dụng công nghệ xử lý ngôn ngữ tự nhiên (Natural Language Processing – NLP) để hiểu và phản hồi các câu hỏi, mệnh lệnh hoặc lời nói của người dùng. Trong giáo dục mầm non, trợ giảng ảo có thể đóng vai trẻ em trong các tình huống giao tiếp mô phỏng, giúp sinh viên luyện tập kỹ năng tương tác ngôn ngữ, xử lý tình huống sư phạm và điều chỉnh cách nói phù hợp với độ tuổi.

Theo Chen, Xu và Zhao (2021), việc sử dụng trợ giảng ảo trong đào tạo giáo viên mầm non giúp tăng cường sự tự tin, cải thiện kỹ năng giao tiếp và giảm áp lực khi thực hành với trẻ thật. Các trợ giảng ảo như **Replika**, **Kuki AI** hay các nền tảng mô phỏng giáo dục như **SimSchool** cho phép người dùng thiết lập kịch bản giao tiếp, lựa chọn độ tuổi, hành vi ngôn ngữ và phản ứng của trẻ ảo.

***** *Ứng dụng trong học phần Phương pháp phát triển ngôn ngữ cho trẻ*

Sinh viên có thể thực hành tổ chức hoạt động kể chuyện, trò chơi ngôn ngữ với trợ giảng ảo đóng vai trẻ. Ví dụ, sinh viên tổ chức hoạt động "Kể chuyện theo tranh" với trẻ ảo, hệ thống sẽ phản hồi bằng các câu hỏi, biểu cảm hoặc hành vi ngôn ngữ phù hợp với độ tuổi mầm non.

Hệ thống AI có thể phản hồi theo nhiều kịch bản khác nhau, giúp sinh viên rèn luyện khả năng xử lý tình huống sư phạm. Chẳng hạn, nếu trẻ ảo tỏ ra không hiểu câu chuyện, sinh viên phải điều chỉnh cách kể, sử dụng từ ngữ đơn giản hơn hoặc kết hợp ngôn ngữ cơ thể để tăng hiệu quả truyền đạt.

Việc luyện tập với trợ giảng ảo giúp sinh viên hình thành phản xạ ngôn ngữ sư phạm, nâng cao khả năng điều chỉnh ngôn ngữ theo đối tượng và hoàn cảnh, từ đó chuẩn bị tốt hơn cho thực hành tại cơ sở mầm non.

Ví dụ 1: Tổ chức hoạt động "Kể chuyện theo tranh" với trợ giảng ảo

Tình huống: Sinh viên sử dụng nền tảng **SimSchool** để tổ chức hoạt động kể chuyện với trẻ ảo 5 tuổi.

Kịch bản:

– Trợ giảng ảo đóng vai trẻ tên "Bé Na"

– Khi sinh viên kể chuyện "Cáo và Quạ", Bé Na phản hồi:

+ "Tại sao con cáo lại nói dối?"

+ "Con quạ có buồn không cô?"

+ "Cô kể lại đoạn con cáo lấy miếng pho mát đi!"

Phản ứng của sinh viên:

– Giải thích bằng ngôn ngữ đơn giản

– Dùng biểu cảm khuôn mặt và giọng kể sinh động

– Khuyến khích trẻ đặt thêm câu hỏi

Lợi ích: Sinh viên luyện tập cách phản hồi câu hỏi của trẻ, điều chỉnh ngôn ngữ và biểu cảm phù hợp với độ tuổi mầm non.

Ví dụ 2: Tổ chức trò chơi ngôn ngữ "Đố vui từ vựng"

Tình huống: Sinh viên sử dụng **Kuki AI** để tổ chức trò chơi "Đố vui từ vựng" với trẻ ảo 4 tuổi.

Kịch bản:

- Trợ giảng ảo đóng vai trẻ tên "Bé Bin"
- Trò chơi: Sinh viên mô tả đồ vật, trẻ đoán tên
- Ví dụ: "Cái này tròn, màu đỏ, dùng để đá" → Bé Bin đoán: "Quả bóng!"

Tình huống phát sinh:

- Bé Bin đoán sai: "Quả táo!"
- Sinh viên phải phản hồi: "Ồ, quả táo cũng tròn và đỏ, nhưng không dùng để đá. Con thử nghĩ lại nhé!"

Lợi ích: Sinh viên rèn luyện kỹ năng phản hồi tích cực, khuyến khích trẻ suy nghĩ lại mà không làm trẻ cảm thấy bị chê trách.

2.3.3. Tạo môi trường học tập tương tác qua công nghệ thực tế ảo (VR/AR)

Công nghệ thực tế ảo (Virtual Reality – VR) và thực tế tăng cường (Augmented Reality – AR) cho phép tạo ra môi trường lớp học mô phỏng, nơi sinh viên có thể tương tác với trẻ ảo trong các hoạt động phát triển ngôn ngữ. Theo báo cáo của UNESCO (2022), VR/AR giúp tăng cường trải nghiệm học tập, đặc biệt trong các học phần thực hành sư phạm, nơi sinh viên cần tiếp xúc với tình huống thực tế nhưng bị hạn chế về điều kiện cơ sở vật chất.

Các nền tảng như **ClassVR**, **CoSpaces Edu** hay **Zspace** cho phép thiết kế lớp học ảo với trẻ em có hành vi ngôn ngữ đa dạng. Sinh viên có thể tương tác với trẻ ảo thông qua lời nói, hành động hoặc phản hồi ngôn ngữ, từ đó luyện tập kỹ năng tổ chức hoạt động ngôn ngữ như kể chuyện, đọc thơ, chơi trò chơi ngôn ngữ.

** Ứng dụng trong học phần Phương pháp phát triển ngôn ngữ*

Mô phỏng lớp học mầm non với trẻ ảo có hành vi ngôn ngữ đa dạng. Ví dụ, trong một lớp học ảo, sinh viên có thể gặp trẻ nói lắp, trẻ nói nhanh hoặc trẻ ít nói, từ đó học cách điều chỉnh phương pháp giao tiếp phù hợp.

Sinh viên được luyện tập thiết kế môi trường ngôn ngữ, tổ chức hoạt động kể chuyện, đọc thơ, trò chơi ngôn ngữ trong không gian ảo. Việc này giúp sinh viên hình dung rõ hơn về vai trò của môi trường vật lý và ngôn ngữ trong phát triển ngôn ngữ cho trẻ.

Ngoài ra, VR/AR còn hỗ trợ việc ghi hình, phân tích và phản hồi hoạt động của sinh viên trong lớp học ảo, giúp cải thiện kỹ năng tổ chức hoạt động và giao tiếp sư phạm.

Ví dụ 1: Mô phỏng lớp học với trẻ ảo có hành vi ngôn ngữ đa dạng

Tình huống: Sinh viên sử dụng nền tảng **ClassVR** để tham gia một lớp học mô phỏng với 5 trẻ ảo, mỗi trẻ có đặc điểm ngôn ngữ khác nhau:

– Trẻ A: nói lắp, ngập ngừng khi trả lời

– Trẻ B: nói nhanh, dùng từ chưa rõ nghĩa

– Trẻ C: ít nói, chỉ phản hồi bằng cử chỉ

– Trẻ D: nói rõ nhưng hay ngắt quãng

– Trẻ E: sử dụng từ ngữ phong phú, thích kể chuyện

Hoạt động: Sinh viên tổ chức hoạt động "Kể chuyện theo tranh" và phải điều chỉnh cách giao tiếp phù hợp với từng trẻ:

– Với Trẻ A: sinh viên nói chậm, dùng câu ngắn, khuyến khích bằng ánh mắt và lời động viên.

– Với Trẻ C: sinh viên sử dụng câu hỏi đóng, kết hợp ngôn ngữ cơ thể để kích thích phản hồi.

– Với Trẻ E: sinh viên khuyến khích kể thêm chi tiết, mở rộng vốn từ.

Lợi ích: Sinh viên học cách phân biệt đặc điểm ngôn ngữ của trẻ, điều chỉnh phương pháp giao tiếp và tổ chức hoạt động phù hợp với từng đối tượng.

Ví dụ 2: Thiết kế môi trường ngôn ngữ bằng CoSpaces Edu

Tình huống: Sinh viên sử dụng **CoSpaces Edu** để thiết kế một góc lớp học ảo phục vụ hoạt động phát triển ngôn ngữ.

Thiết kế gồm:

– Góc kể chuyện: có giá sách, hình ảnh minh họa, loa phát âm thanh

– Góc trò chơi ngôn ngữ: bảng chữ cái, thẻ từ, hộp câu hỏi

– Góc giao tiếp: bàn tròn, ghế nhỏ, mô hình trẻ ảo tương tác bằng lời nói

Hoạt động: Sinh viên tổ chức hoạt động "Đố vui chữ cái" trong không gian ảo, trẻ ảo phản hồi bằng cách chọn thẻ từ hoặc trả lời bằng giọng nói.

Lợi ích: Sinh viên rèn luyện kỹ năng thiết kế môi trường học tập, hiểu rõ vai trò của không gian vật lý và ngôn ngữ trong việc kích thích giao tiếp của trẻ.

2.3.4. Phản hồi tự động giúp cải thiện kỹ năng giao tiếp sư phạm

Một trong những ưu điểm nổi bật của AI là khả năng phân tích nội dung lời nói, ngữ điệu, biểu cảm và đưa ra phản hồi tức thời về chất lượng giao tiếp. Các hệ thống như **ChatGPT**, **Whisper**, **Speech Analyzer** hay các nền tảng AI giáo dục có thể đánh giá độ phù hợp của ngôn ngữ sư phạm, mức độ tương tác và khả năng truyền đạt thông tin.

Theo nghiên cứu của Holmes et al. (2021), phản hồi tự động từ AI giúp người học nhận biết các lỗi ngôn ngữ, cải thiện cách sử dụng từ ngữ, tăng cường khả năng truyền đạt và điều chỉnh ngữ điệu phù hợp với đối tượng trẻ em.

** Ứng dụng trong học phần Phương pháp phát triển ngôn ngữ*

Sinh viên ghi âm hoặc nhập nội dung tương tác với trẻ, hệ thống AI phân tích và đưa ra phản hồi về cách sử dụng từ ngữ, độ rõ ràng, tính phù hợp với độ tuổi. Ví dụ, nếu sinh viên sử dụng từ ngữ quá phức tạp hoặc câu dài, hệ thống sẽ gợi ý cách đơn giản hóa để phù hợp với khả năng tiếp nhận của trẻ mầm non.

AI giúp sinh viên tự điều chỉnh cách nói, cải thiện kỹ năng giao tiếp sư phạm một cách liên tục. Việc phản hồi tức thời giúp sinh viên nhận biết lỗi và điều chỉnh ngay trong quá trình học, thay vì chờ đến khi được giáo viên nhận xét.

Phản hồi tự động còn giúp sinh viên xây dựng phong cách giao tiếp riêng, phù hợp với từng nhóm trẻ, từ đó nâng cao hiệu quả giảng dạy và phát triển ngôn ngữ cho trẻ mầm non.

Ví dụ 1: Phân tích lời nói qua ghi âm tương tác

Tình huống: Sinh viên thực hành kể chuyện cho trẻ mầm non và ghi âm đoạn tương tác.

Nội dung ghi âm:

"Ngày xửa ngày xưa, có một con vật tên là tê giác, nó sống ở vùng thảo nguyên châu Phi, nơi có khí hậu bán khô hạn và hệ sinh thái đa dạng."

Phản hồi từ hệ thống AI (ví dụ: Whisper + Speech Analyzer):

- Từ ngữ quá phức tạp: "tê giác", "thảo nguyên châu Phi", "khí hậu bán khô hạn" có thể khó hiểu với trẻ 3–4 tuổi.

- Gợi ý điều chỉnh: "Ngày xưa, có một con tê giác to lớn sống ở nơi có nhiều cỏ và nắng."

- Ngữ điệu: giọng đều, thiếu nhấn nhá → gợi ý tăng cường biểu cảm ở từ "to lớn", "nhiều cỏ".

Lợi ích: Sinh viên nhận ra cách kể chuyện cần đơn giản hóa ngôn ngữ, tăng biểu cảm để thu hút trẻ và phù hợp với khả năng tiếp nhận của lứa tuổi.

Ví dụ 2: Phân tích nội dung nhập văn bản tương tác

Tình huống: Sinh viên nhập nội dung tương tác giả định với trẻ vào hệ thống AI.

Nội dung nhập:

"Các em có biết vì sao cây cần ánh sáng mặt trời để quang hợp không?"

Phản hồi từ hệ thống AI (ví dụ: ChatGPT hoặc nền tảng AI giáo dục):

– Ngôn ngữ khoa học chưa phù hợp: "quang hợp" là khái niệm trừu tượng với trẻ mầm non.

- Gợi ý điều chỉnh: "Các con có biết vì sao cây thích ánh nắng không? Vì ánh nắng giúp cây lớn lên đó!"
- Độ dài câu: câu hỏi dài → gợi ý chia nhỏ thành câu đơn giản hơn.

Lợi ích: Sinh viên học cách chuyển hóa kiến thức khoa học thành ngôn ngữ gần gũi, dễ hiểu với trẻ, từ đó nâng cao khả năng truyền đạt.

3. Kết luận

Trong bối cảnh giáo dục ngày càng đổi mới, việc ứng dụng công nghệ AI như trợ giảng ảo, thực tế ảo (VR/AR) và nhận diện giọng nói đã mở ra hướng tiếp cận mới trong đào tạo giáo viên mầm non. Những công cụ này không chỉ giúp sinh viên rèn luyện kỹ năng giao tiếp sư phạm mà còn nâng cao khả năng đánh giá, tổ chức hoạt động và xây dựng môi trường ngôn ngữ hiệu quả. AI góp phần chuẩn hóa quy trình đánh giá, giảm sai lệch chủ quan và hỗ trợ can thiệp sớm các vấn đề ngôn ngữ ở trẻ. Việc tích hợp công nghệ vào học phần "Phương pháp phát triển ngôn ngữ" không chỉ tăng cường trải nghiệm học tập mà còn chuẩn bị cho sinh viên năng lực thực hành vững vàng trong môi trường giáo dục hiện đại. Đây là bước tiến cần thiết để nâng cao chất lượng đào tạo và đáp ứng yêu cầu đổi mới giáo dục mầm non trong thời đại số.

Tài liệu tham khảo

1. Bruner, J. (1983). *Child's Talk: Learning to Use Language*. Oxford University Press.
2. Nguyễn Ánh Tuyết. (2015). *Giáo trình Phương pháp phát triển ngôn ngữ cho trẻ mầm non*. NXB Đại học Sư phạm.
3. Nguyễn Thị Hồng. (2020). *Thực trạng giảng dạy học phần Phương pháp phát triển ngôn ngữ tại các trường sư phạm*. Tạp chí Giáo dục.
4. OECD. (2021). *Artificial Intelligence in Education: Challenges and Opportunities*.
5. UNESCO. (2022). *AI and Early Childhood Education: Ethical and Practical Considerations*.
6. Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.

Thông tin tác giả:

Họ và tên: Trần Thị Ban Mai

Đơn vị công tác: Trường Cao đẳng Sư phạm Nam Định

Địa chỉ: 813 Trường Chinh, phường Nam Định, tỉnh Ninh Bình

Email: banmaisp2@gmail.com

NOTEBOOKLM TRONG ĐỔI MỚI GIẢNG DẠY VÀ NGHIÊN CỨU NGOẠI NGỮ

Trần Ngọc Hà [1]

Tóm tắt: Trong bối cảnh hội nhập quốc tế, các cơ sở giáo dục đại học đang đẩy mạnh chuyển đổi số, hướng tới xây dựng mô hình đại học số nhằm kết nối dữ liệu, khai thác sức mạnh của trí tuệ nhân tạo (AI) và đồng bộ hóa hệ sinh thái trong giảng dạy, quản lý và nghiên cứu khoa học. Sự phát triển mạnh mẽ của các công cụ AI đã mở ra nhiều phương thức tiếp cận mới, đặc biệt hữu ích cho giảng viên ngoại ngữ trong việc đổi mới phương pháp dạy và nâng cao năng lực nghiên cứu. Qua phân tích mô hình SWOT, nghiên cứu cho thấy NotebookLM có tiềm năng lớn giúp giảng viên tiết kiệm thời gian, tổ chức tài liệu, nâng cao hiệu quả giảng dạy và phát triển năng lực nghiên cứu, song vẫn tồn tại những hạn chế liên quan đến yêu cầu cao về năng lực số, độ chính xác và vấn đề bảo mật dữ liệu. Từ đó, tác giả đề xuất một vài khuyến nghị giúp giảng viên tối ưu hóa hiệu quả sử dụng NotebookLM trong giảng dạy và nghiên cứu ngoại ngữ.

Từ khóa: NotebookLM, giảng dạy ngoại ngữ, nghiên cứu, trí tuệ nhân tạo, SWOT, năng lực số.

I. Giới thiệu

1. Bối cảnh

Chuyển đổi số đang trở thành xu hướng tất yếu trong giáo dục không chỉ ở các quốc gia phát triển mà còn ở các quốc gia đang phát triển như Việt Nam. Ngày 03/06/2020, Thủ tướng chính phủ đã ký ban hành Quyết định số 749/QĐ-TTg phê duyệt "Chương trình Chuyển đổi số quốc gia đến năm 2025, định hướng đến năm 2030" để xây dựng Việt Nam trở thành quốc gia số. Đại hội đại biểu toàn quốc lần thứ XIII của Đảng Cộng sản Việt Nam (2021) đã đặt ra những yêu cầu quan trọng đối với chuyển đổi số trên các phương diện kinh tế số, xã hội số, chính phủ số. Một trong những yêu cầu quan trọng đối với giáo dục và đào tạo mà Đại hội XIII đã nhấn mạnh là "Đào tạo con người theo hướng có đạo đức, kỷ luật, kỷ cương, ý thức trách nhiệm công dân, xã hội; có kỹ năng sống, kỹ năng làm việc, ngoại ngữ, công nghệ thông tin,

[1] ThS. Đại học Ngoại ngữ - Tin học Thành phố Hồ Chí Minh; *Email: hatn@huflit.edu.vn*

công nghệ số, tư duy sáng tạo và hội nhập quốc tế (công dân toàn cầu) (2021)". Quán triệt tinh thần Đại hội XIII, các cơ sở giáo dục đại học tại Việt Nam đang tích cực triển khai chuyển đổi số nhằm nâng cao chất lượng đào tạo, nghiên cứu khoa học và nâng cao vị thế của nhà trường.

2. Đặt vấn đề

Chuyển đổi số trong giáo dục đại học đang mở ra nhiều cơ hội đổi mới, đặc biệt trong giảng dạy ngoại ngữ – lĩnh vực vốn đòi hỏi sự tương tác đa chiều và khai thác nguồn tri thức đa dạng. Là một phần của quá trình chuyển đổi số, sự phát triển mạnh mẽ của trí tuệ nhân tạo (AI) đã tạo ra những chuyển biến tích cực trong nền giáo dục toàn cầu. Mặc dù AI đã được ứng dụng trong một số lĩnh vực giáo dục tại Việt Nam, nhưng mức độ tích hợp còn hạn chế, đặc biệt trong giảng dạy ngoại ngữ (Vân, 2016). Giảng viên ngoại ngữ hiện nay không chỉ cần chuẩn bị bài giảng sáng tạo mà còn phải tham gia nghiên cứu khoa học. Tuy nhiên, họ gặp không ít khó khăn khi tìm kiếm công cụ AI vừa có khả năng hỗ trợ giảng dạy hiệu quả, vừa góp phần phát triển nghiên cứu khoa học. Trong bối cảnh đó, NotebookLM, nền tảng AI do Google phát triển, nổi lên như một giải pháp tiềm năng. NotebookLM được xem là "trợ lý nghiên cứu ảo" (Alisoy, 2025) vì có khả năng tóm tắt, liên kết và phân tích tài liệu, từ đó giúp giảng viên tiết kiệm thời gian, tổ chức tri thức và tạo ra các hoạt động giảng dạy, nghiên cứu mang tính sáng tạo hơn.

3. Mục tiêu nghiên cứu

Bài báo này tập trung áp dụng mô hình SWOT nhằm đánh giá điểm mạnh, điểm yếu, cơ hội và thách thức khi áp dụng NotebookLM trong giảng dạy ngoại ngữ và nghiên cứu khoa học. Bài nghiên cứu còn gợi ý một số giải pháp cho giảng viên ngoại ngữ nhằm khai thác hiệu quả công cụ này trong giảng dạy và nghiên cứu khoa học.

4. Câu hỏi nghiên cứu

Bài nghiên cứu này tập trung trả lời các câu hỏi sau:

1. Những điểm mạnh, điểm yếu, cơ hội và thách thức (SWOT) khi áp dụng NotebookLM trong giảng dạy ngoại ngữ và nghiên cứu khoa học là gì?
2. Làm thế nào để sử dụng NotebookLM hiệu quả trong giảng dạy và nghiên cứu khoa học?

II. Tổng quan nghiên cứu

1. Giới thiệu về NotebookLM

Notebook là một nền tảng trí tuệ nhân tạo do Google phát triển, ra mắt vào giữa năm 2023, được thiết kế để tăng cường các hoạt động ghi chú và nghiên cứu thông qua các cải tiến trí tuệ nhân tạo. NotebookLM là một ứng dụng AI tạo sinh giúp người dùng tạo "notebook" (Rahman & cộng sự, 2025) giúp người dùng tổng hợp lại thông tin từ nhiều nguồn một cách thông minh nhờ AI Gemini của Google. Người dùng có thể dễ dàng tải lên các tài liệu cá nhân như PDFs, websites, YouTube videos, audio files, Google Docs, or Google Slides. NotebookLM sẽ sử dụng AI Gemini để chuyển đổi các tài liệu đã cung cấp thành các định dạng

dễ tiếp cận như trả lời câu hỏi, tóm tắt nội dung, tạo ghi chú, sơ đồ tư duy, âm thanh, podcast và hướng dẫn học tập. Các chức năng chính của NotebookLM bao gồm tóm tắt, phân tích ngữ cảnh và phát triển dự án hợp tác, có thể mang lại lợi ích đáng kể cho người dùng trong môi trường học thuật và chuyên nghiệp (Alisoy (2025), Rahman & cộng sự (2025).

2. Những công trình nghiên cứu về ứng dụng NotebookLM trong dạy và học

Các nghiên cứu gần đây chỉ ra rằng NotebookLM đã làm thay đổi các quy trình đọc truyền thống thành một trải nghiệm đa phương thức. Không giống như mô hình ngôn ngữ lớn như ChatGPT hay Copilot, NotebookLM cho phép người học và nhà giáo dục tương tác trực tiếp với nội dung được cá nhân hóa, chẳng hạn như trang trình bày bài giảng, bài đọc hoặc bảng điểm, thúc đẩy trải nghiệm học tập phù hợp và phù hợp hơn với ngữ cảnh. Trong một nghiên cứu khám phá việc sử dụng nó trong giáo dục vật lý, Tufino (2025) phát hiện ra rằng NotebookLM hỗ trợ hiệu quả đối thoại kiểu Socrates và học tập tích cực bằng cách tham khảo các tài liệu do giáo viên tuyển chọn. Mô hình tương tác được cá nhân hóa này thúc đẩy sự tham gia sâu sắc hơn và tư duy phản biện, phù hợp với các nguyên tắc sư phạm. Một nghiên cứu khác đã tiết lộ rằng việc sử dụng NotebookLM có thể hỗ trợ các phương pháp giảng dạy đáp ứng văn hóa, điều này rất quan trọng để thúc đẩy một môi trường học tập hòa nhập (Kissau & cộng sự, 2014). Bằng cách giải quyết các nhu cầu đa dạng của người học thông qua các cơ chế học tập thích nghi, NotebookLM có khả năng thu hẹp khoảng cách mà người học từ các nền văn hóa xã hội khác nhau thường gặp phải, do đó thúc đẩy khả năng tiếp cận công bằng với giáo dục ngôn ngữ (Wang, 2017). Sự xuất hiện của NotebookLM cũng bổ sung cho các phương pháp sư phạm dựa trên thể loại đã được nhấn mạnh trong chương trình giảng dạy ngoại ngữ đương đại, cung cấp các tài nguyên và khuôn khổ giúp người học điều hướng các thể loại viết khác nhau (Troyan, 2016) và hỗ trợ quá trình viết học thuật (Alisoy, 2025). Điều này phù hợp với sự thay đổi sư phạm theo hướng tích hợp kiến thức kỹ thuật số với giáo dục ngôn ngữ để chuẩn bị cho học sinh trước sự phức tạp của giao tiếp hiện đại.

Mặc dù các nghiên cứu đã nêu bật các khía cảnh tích cực của NotebookLM trong giảng dạy ngoại ngữ, nhưng những vẫn cần nghiên cứu về những hạn chế, những cơ hội và thách thức để đánh giá tác động của NotebookLM đối với việc giảng dạy và nghiên cứu.

III. Phương pháp nghiên cứu

1. Thiết kế nghiên cứu

Nghiên cứu này áp dụng phương pháp định tính nhằm khám phá cách giảng viên ngoại ngữ sử dụng NotebookLM trong hoạt động giảng dạy và nghiên cứu khoa học. Cách tiếp cận nghiên cứu định phù hợp với mục tiêu trải nghiệm, nhận thức và đánh giá chủ quan của giảng viên về một công cụ mới (Creswell & Poll, 2018).

2. Nguồn dữ liệu

Dữ liệu nghiên cứu được thu thập từ hai nhóm nguồn chính:

1. Văn bản mô tả và tài liệu chính thức từ Google NotebookLM: bao gồm hướng dẫn sử dụng, bài viết giới thiệu, và các bản cập nhật tính năng.
2. Các công trình nghiên cứu và báo cáo liên quan: tập trung vào ứng dụng NotebookLM trong giảng dạy ngoại ngữ và nghiên cứu ở các trường Việt Nam và trên thế giới, các mô hình đổi mới dạy học trong giai đoạn chuyển đổi số, các loại văn bản, chính sách của chính phủ Việt Nam về chuyển đổi số và ứng dụng công cụ AI trong giáo dục.

3. Phân tích dữ liệu

Phân tích SWOT là một công cụ phân tích được sử dụng rộng rãi để đánh giá điểm mạnh và điểm yếu bên trong, cũng như các cơ hội và mối đe dọa bên ngoài. Phân tích SWOT đặc biệt quan trọng trong bối cảnh nghiên cứu định tính vì nó mang lại chiều sâu và cái nhìn sâu sắc về ứng dụng NotebookLM trong giảng dạy ngoại ngữ và nghiên cứu khoa học. Nhờ đó, thời gian chuẩn bị được tối ưu hóa, giúp giảng viên tập trung hơn vào việc truyền đạt kiến thức và tương tác với sinh viên (Liu, 2023).

IV. Phân tích ma trận SWOT ứng dụng NotebookLM trong giảng dạy và nghiên cứu

1. Điểm mạnh

Một trong những tính năng đặc biệt của NotebookLM là khả năng tự động hóa việc tóm tắt, trích xuất nội dung từ các tài liệu nghiên cứu (Konyrova (2024), Martin & Johnson (2023)), giúp giảng viên tiết kiệm thời gian trong việc chuẩn bị bài giảng và tài liệu tham khảo. Điều này đặc biệt hữu ích đối với giảng viên ngoại ngữ khi phải xử lý một lượng lớn tài liệu phục vụ nghiên cứu ngôn ngữ và giáo dục. Đồng thời,

Ngoài ra, NotebookLM còn được xem là một công cụ cá nhân hóa học tập. Từ một tài liệu giảng dạy bất kỳ, giảng viên có thể tạo audio, video, flashcards, bản đồ tư duy, bài kiểm tra, hướng dẫn học tập phù hợp với nhu cầu của người học (Kokala, 2024). Các hoạt động cá nhân hóa học tập cho phép giảng viên tập trung hơn vào việc giảng dạy, tăng cường sự tương tác với sinh viên, giúp sinh viên phát triển kỹ năng tư duy sáng tạo (Gyawali & Mehandroo, 2022), tư duy phản biện (Kasneci & cộng sự, 2023) và nâng cao hiệu quả việc dạy và học (Topuz & cộng sự, 2021).

Trong lĩnh vực nghiên cứu, NotebookLM có thể gợi ý cấu trúc cho các bài báo khoa học hoặc luận văn, giúp giảng viên và sinh viên dễ dàng tổ chức ý tưởng và thông tin trong quá trình viết (Longhurst & cộng sự, 2020). Với khả năng hỗ trợ phân tích dữ liệu định tính, NotebookLM giúp người dùng tìm kiếm từ khóa và xác định các ý chính trong dữ liệu thu thập, từ đó nâng cao khả năng đưa ra các luận điểm và kết luận trong nghiên cứu (Tulusina & cộng sự, 2021). Điều này hỗ trợ giảng viên trẻ hoặc các nhà nghiên cứu mới bắt đầu trong việc xây dựng khung viết học thuật, đồng thời phù hợp với xu hướng sử dụng AI để tăng năng suất nghiên cứu.

2. Điểm yếu

NotebookLM hiện chưa có khả năng đánh giá chất lượng học thuật của tài liệu một cách chính xác và cũng không phân biệt rõ ràng giữa các tài liệu uy tín và tài liệu không đáng tin cậy. Điều này có thể dẫn đến việc người dùng sử dụng thông tin không chính xác (Gurel & Tat, 2017) hoặc thông tin khái quát, thiếu chiều sâu ngôn ngữ học (Hockly, 2023).

NotebookLM đôi khi thiếu độ chính xác khi phân tích tài liệu, nhất là trong các ngữ cảnh học thuật khắt khe, nơi yêu cầu sự phân tích và phê bình sâu sắc (Gill & cộng sự, 2017). NotebookLM không thể thay thế hoàn toàn khả năng tư duy phản biện và phân tích sâu của nhà nghiên cứu, điều này khiến cho những người sử dụng cần phải có kỹ năng tự đánh giá và phản biện hiệu quả các kết quả được cung cấp (Sarmasági, 2021).

NotebookLM chỉ thật sự là một công cụ trợ lý ảo và thực hiện theo câu lệnh (prompt) của người dùng. Do đó, giảng viên cần phải được trang bị kỹ năng số và năng lực về AI để khai thác và sử dụng hiệu quả các tiềm năng của NotebookLM (Idham & cộng sự, 2024).

3. Cơ hội

NotebookLM có thể hỗ trợ giảng viên phát triển kỹ năng thiết kế tài liệu giảng dạy hiện đại, thúc đẩy đổi mới phương pháp giảng dạy bằng cách kết hợp AI với giáo trình truyền thống (Ulfa, 2023) hay với các công cụ AI khác (như ChatGPT, Perplexity) để xây dựng hệ sinh thái học tập toàn diện.

NotebookLM có thể hỗ trợ giảng viên trẻ và nghiên cứu viên mới trong việc làm quen với quy trình nghiên cứu khoa học, đặc biệt là trong môi trường học thuật hiện đại. Điều này giúp nâng cao năng lực nghiên cứu của đội ngũ giảng viên trẻ (Koo & cộng sự, 2011) trong bối cảnh chuyển đổi số giáo dục đại học Việt Nam.

Việc sử dụng NotebookLM có thể khuyến khích các cơ sở giáo dục đại học nâng cao khả năng hợp tác quốc tế thông qua các công cụ số, giúp tăng cường trao đổi kiến thức và tài nguyên học thuật (Koo et al., 2011).

4. Thách thức

Sự phụ thuộc quá mức vào công cụ như NotebookLM có thể dẫn đến suy yếu kỹ năng học tập, khả năng tư duy sáng tạo và phân tích phản biện của sinh viên hoặc làm suy yếu kỹ năng nghiên cứu truyền thống như đọc phân tích và phê bình học thuật (Keban & cộng sự, 2019).

Việc giảng viên sử dụng dữ liệu từ AI mà không kiểm chứng có thể dẫn đến sai lệch trong tổng hợp tài liệu, từ đó làm giảm độ tin cậy của nghiên cứu (Tulusina & cộng sự, 2021).

Ngoài ra, vấn đề bảo mật dữ liệu khi tải lên nền tảng AI như NotebookLM cũng có nhiều rủi ro cho cả giảng viên và sinh viên (Zawacki-Richter & cộng sự, 2019).

V. Kết luận và khuyến nghị

1. Kết luận

Từ kết quả phân tích nội dung bằng mô hình SWOT, NotebookLM là một công cụ tiềm năng giúp giảng viên ngoại ngữ quản lý tài liệu, nâng cao hiệu quả chuẩn bị bài giảng và năng lực nghiên cứu. Tuy nhiên, công cụ này không thể thay thế tư duy phản biện và sự đánh giá chất lượng học thuật từ con người. Vì vậy, để khai thác tối đa tiềm năng của NotebookLM, giảng viên cần được trang bị kỹ năng sử dụng AI và kỹ năng số nhằm đánh giá và điều chỉnh đầu ra của AI, tránh lệ thuộc hoàn toàn vào công nghệ. Đồng thời, các cơ sở đào tạo cần có chính sách hỗ trợ bảo mật dữ liệu và đào tạo năng lực số để giảm chênh lệch giữa các giảng viên, và hướng dẫn về đạo đức nghiên cứu và kiểm chứng thông tin khi ứng dụng AI trong học thuật.

2. Khuyến nghị

Mặc dù NotebookLM là một trợ lý học thuật ảo, nhưng việc hiểu được khả năng và hạn chế của nó sẽ khai thác triệt để công cụ này. Để tối ưu hiệu quả của NotebookLM trong quy trình giảng dạy ngoại ngữ và nghiên cứu, giảng viên có thể thực hiện theo một số đề xuất chiến lược như sau:

- **Đọc tài liệu trước khi tải lên NotebookLM:** Khả năng tiếp thu từ vựng và đọc hiểu của NotebookLM đặc biệt có lợi cho người học ngoại ngữ. Bằng cách tận dụng các tính năng tóm tắt của nó, giảng viên có thể cung cấp cho sinh viên nội dung ngắn gọn, phù hợp, làm nổi bật từ vựng và cấu trúc ngữ pháp cần thiết. Tuy nhiên, Giảng viên cần thực hiện ít nhất lần đọc đầu tiên để hiểu ngữ cảnh và mức độ liên quan của nội dụng. Đối với các bài nghiên cứu, giảng viên cần đọc nhiều lần, chú thích và đánh dấu vào các phần chính. Việc đọc trước tài liệu sẽ giúp giảng viên đặt câu hỏi tốt hơn và phát hiện ra bất kỳ cách hiểu sai nào của AI.
- **Sử dụng câu lệnh phù hợp:** Giảng viên cần biết cách đặt câu lệnh vì câu lệnh cụ thể, có mục đích sẽ mang lại phản hồi hữu ích và chính xác hơn. Bắt đầu với các câu hỏi tổng quan rộng, sau đó đi sâu vào các chi tiết cụ thể. Giảng viên ngoại ngữ có thể đặt các câu hỏi yêu cầu NotebookLM tạo sơ đồ tư duy, flashcards, audios, videos hoặc bài kiểm tra để giúp sinh viên nâng cao khả năng ghi nhớ và hiểu ngôn ngữ.
- **Sắp xếp tài liệu theo chủ đề:** Giảng viên nên tạo sổ tay riêng biệt cho từng chủ đề (chủ đề theo tựa bài học trong giáo trình hay chủ đề nghiên cứu) vì tổ chức theo chủ đề này cải thiện khả năng của AI trong việc cung cấp các phản hồi phù hợp, theo ngữ cảnh.
- **Kiểm tra trích dẫn và tham khảo chéo:** Giảng viên cần xác minh tài liệu và kiểm tra chéo với các công cụ AI khác (ChatGPT, Gemini, Copilot, …) vì NotebookLM có thể cung cấp cung cấp thông tin sai lệch (ảo giác thông tin). Ngoài ra, Giảng viên cũng có thể thực hiện kiểm tra chéo bằng cách lấy ý kiến phản hồi của đồng nghiệp. Việc tích hợp các cơ chế phản hồi của đồng nghiệp có thể giúp giảng viên phát triển các kỹ năng viết phản biện trong khi sử dụng NotebookLM để tinh chỉnh bản nháp dựa trên ý kiến đóng góp mang tính xây dựng từ các đồng nghiệp.

- **Sử dụng NotebookLM như một công cụ cộng tác:** Giảng viên hãy xem NotebookLM như một trợ lý nghiên cứu hoặc trợ lý giảng dạy thông minh, không phải là sự thay thế cho chuyên môn. Kết quả tốt nhất đến khi giảng viên biết cách kết hợp khả năng AI với phán đoán chuyên môn và kiến thức ngôn ngữ của mình.

TÀI LIỆU THAM KHẢO

Anh, C. H. (2024). Giải pháp chuyển đổi số trong quản lý khoa học tại trường đại học văn hóa, thể thao và du lịch thanh hóa. *Tạp Chí Khoa Học*, *02 (20) T5*, 123–123. https://doi.org/10.55988/2588-1264/203

Alisoy, H. (2025). Can NotebookLM Support English Language Learners? A Theoretical Perspective on AI Tools in Education. *Porta Universorum*, *1*(6), 25–55. https://doi.org/10.69760/portuni.0106003

Benzaghta, M. A., Elwalda, A., Mousa, M., Erkan, I., & Rahman, M. (2021). SWOT Analysis applications: an Integrative Literature Review. *Journal of Global Business Insights*, *6*(1), 54–72.

Çolak, M. K. (2023). The future of EFL teachers' professional development in Türkiye: A Swot Analysis. *Akademik Platform Eğitim ve Değişim Dergisi*, *6*(2), 246–278. https://doi.org/10.55150/apjec.1300227

Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry and Research design: Choosing Among Five Approaches* (4th ed.). SAGE Publications.

ĐCSVN (2021). *Văn kiện Đại hội đại biểu toàn quốc lần thứ XIII*, T.I, Nxb. Chính trị quốc gia Sự thật, Hà Nội

Gurel, E., & Tat, M. (2017). SWOT Analysis: a Theoretical Review. *Journal of International Social Research*, *10*(51), 994–1006. https://doi.org/10.17719/jisr.2017.1832

Gyawali, Y. P., & Mehandroo, M. (2022). Artificial Intelligence in English Language Teaching: Navigating the Future with Emerging Perspectives. *Journal of Language and Linguistics in Society*, *26*, 21–27. https://doi.org/10.55529/jlls.26.21.27

Hijayadi, R., & Jamilah. (2024). A Swot Analysis of AI – Powered Chatbots in English Language Education. *Formosa Journal of Sustainable Research*, *3*(11), 2479–2488. https://doi.org/10.55927/fjsr.v3i11.12447

Hockly, N. (2023). Artificial intelligence in English language teaching: The good, the bad and the ugly. *RELC Journal*, *54*(2), 445–451. https://doi.org/10.1177/00336882231168504

Idham, A. Z., Rauf, W., & Rajab, A. (2024). Navigating the Transformative Impact of Artificial Intelligence on English Language Teaching: Exploring Challenges and Opportunities. *Deleted Journal*, *4*(1), 8–14. https://doi.org/10.56185/jes.v4i1.620

Keban, Y. B., Arifin, S., & Wahyono, R. (2019). SWOT Analysis and Its Implementation Strategies in Educational Management. *Journal of Education and Practice*, *10*(12). https://doi.org/10.7176/jep/10-12-10

Kissau, S., Rodgers, M., & Haudeck, H. (2014). Foreign Language Teaching: An International Comparison of Teacher Beliefs. *Research in Comparative and International Education*, *9*(2), 227–242. https://doi.org/10.2304/rcie.2014.9.2.227

Kokala, A. (2024). Revolutionizing Content Creation: Leveraging AI-Driven Podcast Generation with NotebookLM and Personalized Insights. *International Research Journal of Modernization in Engineering Technology and Science*, *6*(12). https://doi.org/10.56726/irjmets65280

Konyrova, L. (2024). The Evolution of Language Learning: Exploring AI's Impact on Teaching English as a Second Language. *Eurasian Science Review*, *2*(2), 133–138. https://doi.org/10.63034/esr-42

Koo, H., Chau, K., Koo, L., Liu, S., & Tsui, S. (2011). A structured SWOT approach to develop strategies for the government of Macau, SAR. *Journal of Strategy and Management*, *4*(1), 62–81. https://doi.org/10.1108/17554251111110122

Liu, M. (2023). Exploring the Application of Artificial Intelligence in Foreign Language Teaching: Challenges and Future Development. *SHS Web of Conferences*, *168*, 03025–03025. https://doi.org/10.1051/shsconf/202316803025

Longhurst, G. J., Stone, D. M., Dulohery, K., Scully, D., Campbell, T., & Smith, C. F. (2020). Strength, Weakness, Opportunity, Threat (SWOT) Analysis of the Adaptations to Anatomical Education in the United Kingdom and Republic of Ireland in Response to the Covid-19 Pandemic. *Anatomical Sciences Education*, *13*(3), 301–311. https://anatomypubs.onlinelibrary.wiley.com/doi/full/10.1002/ase.1967

Mananay, J. A. (2024). Integrating Artificial Intelligence (AI) in Language Teaching: Effectiveness, Challenges, and Strategies. *International Journal of Learning, Teaching and Educational Research*, *23*(9), 361–382. https://doi.org/10.26803/ijlter.23.9.19

Martin, R., & Johnson, S. (2023, July 12). *Introducing NotebookLM*. Google. https://blog.google/technology/ai/notebooklm-google-ai/

Rahman, N. A. A., Zulkornain, L. H., Soon, G. Y., Hamzah, N. H., & Chen, C. (2025). Notebooklm integrated project-based learning model development for writing among undergraduates. *International Journal of Research and Innovation in Social Science*, *9*(7), 5498-5503. https://doi.org/10.47772/ijriss.2025.907000444

Reinhardt, J. (2018). Social media in second and foreign language teaching and learning: Blogs, wikis, and social networking. *Language Teaching*, *52*(1), 1–39. https://doi.org/10.1017/s0261444818000356

Sarmasági, P. G. (2021). SWOT Assessment Usage in School Talent Management. *Central-European Journal of New Technologies in Research, Education and Practice*, *3*(2), 42–60. https://doi.org/10.36427/CEJNTREP.3.2.1355

Topuz, Ş., Sezer, N. Y., Aker, M. N., Gönenç, İ. M., Cengiz, H. Ö., & Korucu, A. E. (2021). A Swot Analysis of the Opinions of Midwifery Students about Distance Education during the Covid-19 Pandemic a Qualitative Study. *Midwifery*, *103*, 103161. https://doi.org/10.1016/j.midw.2021.103161

Troyan, F. J. (2016). Learning to Mean in Spanish Writing: A Case Study of a Genre-Based Pedagogy for Standards-Based Writing Instruction. *Foreign Language Annals*, *49*(2), 317–335. https://doi.org/10.1111/flan.12192

Tufino, E. (2025). NotebookLM: An LLM with RAG for active learning and collaborative tutoring. *ArXiv.org*, *2*. https://arxiv.org/abs/2504.09720

Tulusina, E. A., Sadykova, A. G., Zaripova, Z. M., & Carlson, C. F. (2021). Linguocultural competence as a basis for foreign language proficiency. *Revista EntreLinguas*, 105–125. https://doi.org/10.29051/el.v7iesp1.14878

Ulfa, K. (2023). The Transformative Power of Artificial Intelligence (AI) to Elevate English Language Learning. *Majalah Ilmiah Methoda*, *13*(3), 307–313. https://doi.org/10.46880/methoda.vol13no3.pp307-313

Vân, A. (2016). *Ứng dụng trí tuệ nhân tạo trong giảng dạy tiếng Anh: Xu hướng toàn cầu và kinh nghiệm cho giáo dục Việt Nam*. Tạp Chí Giáo Dục - Bộ Giáo Dục và Đào Tạo.

Wang, Q. (2017). An Empirical Study on College English Curriculum Assessment and Cultivation of Intercultural Communication Competence. *Linguistics and Literature Studies*, *5*(6), 417–423. https://doi.org/10.13189/lls.2017.050604

Yeo, M. A., Moorhouse, B. L., & Wan, Y. (2025). From Academic Text to Talk-Show: Deepening Engagement and Understanding with Google NotebookLM. *Teaching English as a Second or Foreign Language--TESL-EJ*, *28*(4). https://doi.org/10.55593/ej.28112int

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

BRIDGING AI AND CULTURAL HERITAGE: A DOMAIN-SPECIFIC CHATBOT FOR CLASSICAL SINO-VIETNAMESE NOVELS

Huynh Thanh Xuan[1], Pham Hoang Vu[2], Luu Thien Duc[3], U Cao Ky Long[4]

Abstract: This paper presents the development of a domain-specific chatbot in the field of Sino-Vietnamese classical literature, with a focus on the fiction genre. We fine-tuned multiple large language models (LLMs) on a dataset constructed from Yuenan Hanwen Xiaoshuo Jicheng — Collected Classical Chinese Novels of Vietnam. By applying instruction-tuning techniques and conversational formatting, the models were adapted to understand, interpret, and respond to queries related to Yuenan Hanwen Xiaoshuo Jicheng in Vietnamese. Experimental results show that the fine-tuned models demonstrate strong capabilities in literary comprehension, content explanation, and user interaction. The chatbot system achieved a highest BERTScore (F1) of 78.5%, indicating its potential as an effective tool for supporting the study and dissemination of Sino-Vietnamese classical literature.

Keywords: Sino-Vietnamese classical literature, Yuenan Hanwen Xiaoshuo Jicheng, Domain-specific chatbot, LoRA fine-tuning, NLP for Cultural Heritage

1. Introduction

Recent advances in conversational AI have enabled chatbots to perform general-purpose dialogue tasks with high proficiency. However, specialized applications for literary research—particularly in Sino-Vietnamese classical literature remain underexplored. In this work, we present a domain-specific chatbot designed for Sino-Vietnamese literature, capable of analyzing and interpreting textual content, as well as engaging in multi-turn interactions. The system aims to support researchers and readers in accessing, understanding, and exploring classical literary works.

Our study focuses on Yuenan Hanwen Xiaoshuo Jicheng as the primary source material.

The text was digitized via an OCR pipeline that converts scanned images into raw textual data. We then constructed a training corpus from the OCR output by leveraging the Gemini API to generate and normalize prompt–response annotations, followed by filtering and manual refinement to mitigate OCR-induced noise.

We fine-tuned five large language models—Phi- 4, Llama 3.2 (3B), and Qwen3 (14B, 4B, 1.7B)—using Low-Rank Adaptation (LoRA) with 4-bit quantization, optimizing for memory efficiency and training speed. Our experimental design facilitates a systematic comparison across architectures and model sizes under constrained hardware settings.

On the evaluation set, the best-performing model achieved a BERTScore (F1) of 78.5. The model was then deployed as an interactive chatbot with an intuitive interface, allowing users to query and explore Sino-Vietnamese classical literature in natural language.

[1] Master's student at VNUHCM – University of Science; *Email: xuanhuynh233@gmail.com*
[2] Master's student at VNUHCM – University of Science; *Email: phamhoangvu1811995@gmail.com*
[3] Master's student at VNUHCM – University of Science; *Email: ducluuthien@gmail.com*
[4] Master's student at VNUHCM – University of Science; *Email: ucaokylong.hardcore@gmail.com*

The key contributions of this work are:

- A curated and annotated dataset derived from Yuenan Hanwen Xiaoshuo Jicheng, including a complete OCR and filtering pipeline.
- A semi-automated annotation workflow using the Gemini API for prompt–response generation and normalization, tailored to literary comprehension tasks.
- A deployed chatbot system that enables users to interact with, query, and explore Yuenan Hanwen Xiaoshuo Jicheng in natural language

2. Related works

Overview of LLMs & Challenges in Classical Literature

Recent advances in large language models (LLMs), including architectures such as the Transformer and its enhancements, have revolutionized natural language understanding and generation, particularly in conversational AI systems. Models like GPT-4 [1] and PaLM [2] demonstrate unprecedented capabilities in handling complex dialogues and long-context dependencies. Despite these advances, directly applying general-purpose LLMs to domain-specific tasks remains challenging, especially for culturally rich and linguistically nuanced domains such as classical Sino-Vietnamese literature. These texts contain unique lexical, syntactic, and semantic structures often underrepresented in large-scale pretraining corpora, limiting zero-shot or few-shot performance. Therefore, domain adaptation through specialized fine-tuning and curated datasets is critical to capture these particularities effectively [3].

LoRA Fine-Tuning and Quantization Methods

A key approach to adapting large language models (LLMs) for specialized applications involves parameter-efficient fine-tuning techniques, notably Low-Rank Adaptation (LoRA). LoRA updates only a small subset of parameters in low-rank matrices, significantly reducing computational and memory overheads while maintaining or improving downstream task performance [4]. Combined with 4-bit quantization, LoRA enables feasible fine-tuning of large models, such as QLoRA [5], on constrained hardware without sacrificing quality. This methodology has demonstrated success in natural language understanding tasks within niche domains, including historical and cultural text interpretation and dialogue systems.

LLMs for Cultural Heritage and Classical Literature Chatbots

In the specific context of cultural heritage and classical literature, several studies have explored domain-specific large language models. Shen et al. (2023) developed a Classical Chinese Large Language Model fine-tuned to improve semantic understanding, punctuation restoration, and translation for ancient texts, illustrating the benefits of dedicated language adaptation [6]. Similarly, domain-specific chatbots have been designed for scientific literature and heritage applications using embedding-based retrieval and instruction tuning to provide accurate user interactions and content exploration [7].

Advanced Multilingual Models and Applications

For Vietnamese and Sino-Vietnamese literature research, model families like Llama 3.2 and Qwen3 offer multilingual and long-context capabilities critical for handling extensive classical text corpora. The Qwen3 models' flexibility to switch between detailed reasoning ("thinking") and efficient conversational modes further enhances their suitability for complex literary domains. These advances motivate combining fine-tuning strategies with robust dataset curation and annotation workflows for effective chatbot system development. [8]

3. Model

In this study, the team simultaneously fine-tuned five large language models: Phi-4, Llama 3.2 3B, Qwen3 14B, Qwen3 4B, and Qwen3-1.7B. All models belong to the modern Transformer family, supporting multiple languages and possessing strong reasoning capabilities.

To optimize hardware resources and shorten training time (especially on low-spec machines), the team applied the following technical approaches in parallel:

- Fine-tuning using the LoRA (Low-Rank Adaptation) method: This is a parameterefficient fine-tuning technique that updates only a small subset of weights, reducing memory requirements and increasing training speed.
- 4-bit quantization: The models were loaded and converted to a 4-bit format to reduce memory usage, while also improving processing speed on common hardware without compromising performance.

With this workflow, the team was able to conveniently deploy and compare the results among models, thereby selecting the optimal chatbot configuration for Sino-Vietnamese literature data.

3.1 Phi-4 Model

Phi-4 is a large language model developed by Microsoft Research, featuring a Dense Decoder-only Transformer architecture with 14 billion parameters. This architecture is specialized for text generation tasks, where the entire model consists solely of decoder layers, enabling strong performance in multi-turn conversations and producing coherent, logical responses.

The model supports a maximum context length of 16,000 tokens, allowing it to process and analyze long texts or continuous interaction sequences while retaining complete contextual information from previous turns. Thanks to its dense structure and optimized context retrieval, Phi-4 can maintain high-quality responses even in extended dialogues.

In this study, Phi-4 was implemented with 4- bit quantization combined with LoRA fine-tuning. This approach significantly reduces memory requirements and shortens training time while fully leveraging the strengths of the Transformer architecture.

3.2 Llama 3.2 Model

Llama 3.2 is a family of autoregressive language models developed by Meta, with a Transformer architecture optimized for enhanced performance in processing and generating natural language. The Llama 3.2 version focuses on multilingual conversational applications, including information retrieval and content summarization.

The model was pre-trained on a large-scale multilingual dataset and further fine-tuned using Supervised Fine-Tuning (SFT) combined with Reinforcement Learning with Human Feedback (RLHF). This approach improves Llama 3.2's ability to provide accurate, user-friendly, and safe responses in interactions.

A notable design feature of Llama 3.2 is the use of Grouped-Query Attention (GQA), which improves scalability during inference, especially when working with large-scale models.

Although trained and officially released in September 2024 as a static model on offline data, Llama 3.2 maintains competitive performance across many industry benchmarks, making it suitable for multilingual chatbot systems, including LoRA fine-tuning and 4-bit quantization to save computational resources in our research.

3.3 Qwen3 1.7B, 4B, and 14B Models

Qwen3 is the latest generation in the Qwen large language model family, available in multiple parameter sizes ranging from 1.7 billion, 4 billion, to 14.8 billion parameters. All versions are causal language models with a modern Transformer architecture designed to handle long contexts of up to 32,768 tokens.

All Qwen3 versions use Grouped-Query Attention (GQA) to enhance performance and inference scalability. The 1.7B version has 28 layers with 16 query heads (Q) and 8 key-value heads (KV); the 4B version has 36 layers with 32 Q heads and 8 KV heads; the largest 14B version has 40 layers with 40 Q heads and 8 KV heads.

A distinctive feature of Qwen3 is its ability to switch seamlessly between a "thinking" mode for tasks requiring complex reasoning such as mathematics, programming, and logic, and a "nonthinking" mode optimized for general dialogue efficiency. This adaptability allows the model to perform well across diverse use cases, delivering natural, creative, and coherent multi-turn conversations.

Qwen3 also excels in following user instructions, supporting over 100 languages and dialects, and offering special agent capabilities for precise integration with external tools, enabling agentbased optimization of complex tasks.

The training process of Qwen3 consists of two stages: extensive pretraining on diverse datasets, followed by post-training for instructionfollowing and user feedback alignment, ensuring effectiveness and safety in deployment.

With these characteristics, the Qwen3 family is well-suited for lightweight fine-tuning using LoRA combined with 4-bit quantization, as implemented in this study. This approach

reduces memory usage and accelerates training on low-spec hardware while maintaining high language processing quality for chatbot applications in SinoVietnamese literature.

4. Dataset

The dataset employed in this study was derived from the volumes of Yuenan Hanwen Xiaoshuo Jicheng, a curated anthology of classical Chinese novels produced in Vietnam. This corpus encapsulates the historical, cultural, and intellectual milieu of premodern Vietnamese society. The compilation was the result of a collaborative publication effort among Shanghai Normal University (China), the Institute of Sino-Vietnamese Studies (Vietnam), National Cheng Kung University (Taiwan), and the Center for Social Science Research (France).

Based on the textual content of these volumes, we utilized the Gemini API to automatically generate question–answer (QA) pairs, which were subsequently employed in the fine-tuning stage. The dataset construction workflow comprised the following stages:

- Optical Character Recognition (OCR): The original printed volumes were digitized through high-resolution scanning, followed by OCR using tools capable of recognizing traditional Chinese script. The output included raw text alongside corresponding bounding box coordinates.
- Text Preprocessing: Orthographic errors, ordering inconsistencies, and extraneous segments were identified and corrected or removed.
- XML Standardization: The cleaned text was converted into a unified XML schema to ensure structural consistency across documents.
- Data Generation: QA pairs were programmatically generated from the XML-formatted texts, using the source content as contextual input.

The resulting dataset comprises 21,650 QA pairs. We partitioned the corpus into training, validation, and test subsets with a ratio of 8:1:1. The data exhibits the distinctive linguistic and stylistic features of Vietnamese classical Chinese literature, encompassing a wide thematic range, including overviews, character descriptions, narrative structures, thematic analysis, and literary techniques. Each QA pair contains between 100 and 300 words. The dataset has undergone deduplication and normalization to ensure textual integrity, uniformity, and suitability for fine-tuning.

Example QA:

{

"problem": "Tác phẩm \"Kiến văn lục\" của Vũ Trinh thuộc thể loại gì và chủ đề chính của nó là gì?",

"generated_solution": "\"Kiến văn lục\" thuộc thể loại truyện ký, ghi chép lại những điều mắt thấy tai nghe của tác giả. Chủ đề chính xoay quanh những câu chuyện kỳ lạ, dị thường,

liên quan đến con người, quỷ thần, động vật, côn trùng... xen lẫn yếu tố hiện thực và truyền thuyết dân gian, thường mang ý nghĩa khuyên răn, giáo dục luân lý."

}

4.1 Preprocessing Input PDF Files

Our team converted PDF files into images for OCR processing. Then, we resized images to a resolution of 800×1000. Since the images only contained two colors (black and white), we skipped the binarization step. We attempted Otsu binarization to enhance text clarity, but the result was unsatisfactory because the original black-and white images were already clear. This preprocessing step even reduced OCR accuracy.

4.2 Problem Analysis

Watermark in OCR output: Bounding boxes of the dataset are rectangular and vertically oriented, making it easy to identify watermark bounding boxes with 90° corners as invalid.

Noisy bounding boxes due to extreme sizes:

These can be removed by comparing bounding box width and height against set thresholds.

Non-Chinese text: Such text can be filtered using available frameworks.

4.3 Methods

Handling watermark in OCR output:

- Compute aspect ratio = height/width. If it is between 0.8 and 1.2, discard it because valid bounding boxes should be rectangular.
- Compute bounding box angles. Keep only those between 75° and 105°.

Filtering overly large or small boxes:

- Remove boxes with both width and height > 90.
- Remove boxes with both width and height < 15.

Filtering non-Chinese text:

- Use hanzidentifier to check if the
- OCR text is in Chinese; if not, discard.

4.4 Vertical Box Merging

Example of boxes to merge into a single column:

[(445, 238), (483, 238), (487, 533), (449, 533)],

[(454, 526), (480, 526), (480, 560), (454, 560)],

[(452, 562), (484, 562), (484, 597), (452, 597)],

[(440, 580), (487, 578), (496, 724), (449, 727)]

The x-coordinate averages are close ([466; 467; 468; 468]), indicating alignment. We grouped bounding boxes into columns if:

- Their x-centers are close (within x threshold=10).
- To handle exceptions where horizontal boxes were merged with vertical boxes, we added a constraint that vertical distances between boxes must not exceed y threshold=20.

The merged box is a rectangle covering all boxes in the group:

merged box = [(min x, min y), (max x, min y), (max x, max y), (min x, max y)]

4.5 Sorting Bounding Boxes

- Boxes must be ordered right-to-left, top-to bottom.
- For horizontal sorting: calculate box center x-coordinates, sort in decreasing order.
- For vertical sorting: boxes with horizontal distance ≤15 are considered in the same column.

4.6 Semi-automatic Verification using PPOCRLabel

After sorting, we manually corrected missing or misaligned boxes semi-automatically with PPOCRLabel.

5. Experiments

In this study, we conduct fine-tuning of the dataset on five large language models (LLMs)—Phi-4, Llama 3.2 3B, Qwen3 14B, Qwen3 4B, and Qwen3-1.7B—and systematically evaluate their performance across a diverse set of metrics. The model achieving the highest overall performance is subsequently selected for the development of a demonstration interface.

5.1 Computational Environment

All experiments were conducted on the Kaggle platform using the Unsloth framework for efficient fine-tuning of large language models. The hardware configuration included a single NVIDIA Tesla P100 GPU with 16 GB of VRAM and up to 29 GB of system RAM. The software environment comprised Python 3.11, PyTorch 2.7.1, and the Hugging Face Transformers library (version 4.55), integrated with Unsloth for optimized training speed and reduced memory consumption. Model fine-tuning and evaluation were performed entirely within Kaggle's GPU runtime environment.

5.2 Model-specific Hyperparameter Settings

All large language models (LLMs) were finetuned using identical hyperparameter configurations for learning rate, batch size, optimizer, and sequence length. The only variation across models was the number of training epochs. Due to the substantial computational

requirements of larger models and the execution time limits imposed by the Kaggle platform, we reduced the number of epochs proportionally to model size to ensure that all training processes could be completed within the allotted time.

Model	**Epochs**
Phi-4	1
Llama 3.2 3B	3
Qwen3-14B	1
Qwen3-4B	3
Qwen3-1.7B	2

Table 1: Number of training epochs used for each model during fine-tuning.

5.3 Evaluation Metrics

Exact Match – is defined as the proportion of predictions whose normalized text is identical to the normalized ground-truth reference. Normalization typically involves lowercasing, removing articles, punctuation, and extra whitespace to ensure that only meaningful textual differences affect the comparison.

F1 Score – calculates the harmonic mean of precision and recall, balancing both correctness and completeness in the generated responses.

BLEU (Bilingual Evaluation Understudy) – evaluates n-gram overlap between the predicted and reference texts; widely used in translation and generation tasks, with higher scores indicating greater lexical similarity.

METEOR – enhances lexical overlap evaluation by incorporating stemming, synonym matching, and a penalty for fragmented matches, making it more sensitive to semantic similarity.

ROUGE-1 – measures recall-oriented overlap based on unigrams (single words), useful for evaluating content coverage.

ROUGE-2 – similar to ROUGE-1 but based on bigrams, capturing short-range dependencies in text.

ROUGE-L – computes the longest common subsequence between prediction and reference, capturing fluency and in-sequence matches.

BERTScore F1 – uses contextual embeddings from pre-trained BERT models to calculate semantic similarity between predicted and reference texts, reported as an F1 score to combine precision and recall.

5.4 Experimental Results

Metric	Phi-4	Llama 3.2 3B	Qwen3 14B	Qwen3 4B	Qwen3-1.7B
Exact Match (%)	0	0	0	0	0
F1 (%)	33.970	1.426	**34.211**	34.015	32.296
BLEU	0.153	0.133	0.1695	0.173	0.156
METEOR	0.3019	0.011	**0.342**	**0.342**	0.323
ROUGE-1	58.319	3.762	60.731	**60.817**	59.697
ROUGE-2	32.105	0.03	**32.828**	32.763	31.018
ROUGE-L	41.642	3.531	43.164	**43.227**	42.155
BERTScore F1 (%)	73.025	57.919	78.971	**78.972**	78.503

Table 2: Performance of fine-tuned large language models (LLMs) across eight evaluation metrics.

The results reveal notable differences in performance across the evaluated models. Qwen3 4B achieved the highest overall scores, particularly in METEOR (0.3420), ROUGE-1 (60.82%), ROUGE-L (43.23%), and BERTScore F1 (78.97%). This suggests that Qwen3 4B produces outputs that are more lexically and semantically aligned with the references compared to the other models.

Phi-4 showed slightly lower performance than Qwen3 4B in most metrics but still maintained competitive ROUGE and F1 values despite being trained for only 1 epoch.

Despite being one of the smallest models in the set, Qwen3-1.7B achieved moderate results, with an F1 score of 32.30%, ROUGE-1 of 59.70%, ROUGE-2 of 31.02%, and ROUGE-L of 42.16%. Its BERTScore F1 (78.50%) was only slightly lower than that of Qwen3 4B, suggesting that the model can capture a reasonable level of semantic similarity with the reference outputs. The limited number of training epochs (2) may have constrained its ability to fully adapt to the fine-tuning dataset.

The absence of Exact Match scores (> 0%) across all models suggests that none of the finetuned models produced predictions identical to the references at the token level, highlighting the difficulty of the task.

6. Conclusion

In this paper, we introduced a domain-specific chatbot for the Yuenan Hanwen Xiaoshuo Jicheng corpus by fine-tuning a large language model (LLM) on a carefully curated dataset.

The system demonstrated promising results in understanding and interpreting Sino-Vietnamese literature.

This research not only demonstrates the feasibility of developing specialized conversational systems for Sino-Vietnamese literature, but also serves as a catalyst for further scholarly exploration in this underrepresented domain. We hope that this work will inspire and support future research, contributing to the preservation and dissemination of Sino-Vietnamese literary heritage.

However, the quality of the dataset and the automatically generated data from Gemini remain limited, affecting the model's training performance. In future work, we will focus on expanding and refining the dataset, as well as improving the data generation pipeline to enhance accuracy and consistency. We also propose splitting high-quality images into smaller tiles for OCR instead of resizing and later merging the outputs, though this approach has not yet been experimented with due to time constraints.

References

[1] OpenAI, "GPT-4 Technical Report," in *arXiv*, 2023.

[2] Chowdhery, Aakanksha & Narang, Sharan & Devlin, Jacob & Bosma, Maarten & Mishra, Gaurav & Roberts, Adam & Barham, Paul & Chung, Hyung & Sutton, Charles & Gehrmann, Sebastian & Schuh, Parker & Shi, Kensen & Tsvyashchenko, Sasha & Maynez, Joshua & Rao, Abhi, "PaLM: Scaling Language Modeling with Pathways," in *arXiv*, 2022.

[3] Zhen wan, Yating Zhang, Yexiang Wang, Fei Cheng, Sadao Kurohashi, "Reformulating Domain Adaptation of Large Language Models as Adapt-Retrieve-Revise: A Case Study on Chinese Legal Domain," in *arXiv preprint*, 2023.

[4] Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, "LoRA: Low-Rank Adaptation of Large Language Models," in *ICLR 2022*, 2022.

[5] Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, Luke Zettlemoyer, "QLoRA: Efficient Finetuning of Quantized LLMs," in *ArXiv*, 2023.

[6] Shen Li, Renfen Hu, Lijun Wang, "Efficiently Building a Domain-Specific Large Language Model from Scratch: A Case Study of a Classical Chinese Large Language Model," in *ArXiv*, 2025.

[7] K. G. Yager, "Domain-specific chatbots for science using embeddings," *Digital Discovery,* vol. 2, no. 6, pp. 1850-1861, 2023.

[8] Cuc Thi Bui, Nguyen Truong Son, Truong Van Trang, Lam Viet Phung, Pham Nhut Huy, Hoang Anh Le, Quoc Huu Van, Phong Nguyen-Thuan Do, Van Le Tran Truc, Duc Thanh Chau, Le-Minh Nguyen, "VMLUBenchmarks: A comprehensive benchmark toolkit for Vietnamese," *Association for Computational Linguistics,* vol. 1, p. 11495–11515, 2025.

[9] Rocco Loffredo and Massimo De Santo, "Using Ontologies for LLM Applications in Cultural Heritage," in *CEUR Workshop Proceedings, 2024*, 2024.

AI IN LANGUAGE TEACHING: BRIDGING TECHNOLOGY AND COMMUNICATION

Phạm Ngọc Diễm[1]

ABSTRACT

The integration of Artificial Intelligence (AI) in language education is reshaping traditional teaching by creating more personalized, efficient, and interactive learning experiences. This study examines major applications of AI in language teaching, including adaptive learning systems, automated feedback, natural language processing for conversation practice, and speech recognition for pronunciation training. Widely used tools allow learners to practice independently, receive instant feedback, and develop greater confidence in communication. For educators, AI reduces workload by generating resources, streamlining assessment, and supporting differentiated instruction. Drawing on a literature review, case studies, and teacher–learner perspectives, the findings reveal that AI enhances learner autonomy, engagement, and motivation, while also improving teaching efficiency. However, challenges such as data privacy, technological accessibility, and the need to preserve human interaction remain significant. Overall, the study argues that AI should complement, rather than replace, traditional pedagogy and that its thoughtful integration can make language education more inclusive, accessible, and effective.

Keywords: artificial intelligence, language education, personalization, feedback, challenges

1. Introduction

Artificial Intelligence (AI) has rapidly emerged as a transformative force across multiple fields, and education is no exception. Within this domain, language teaching has become one of the areas most influenced by AI innovation. Traditional instruction, often constrained by fixed curricula, limited classroom interaction, and delayed feedback, may not adequately address the diverse needs, learning speeds, and styles of modern students. In contrast, AI technologies provide innovative solutions by enabling personalized instruction, offering real-time feedback, and creating engaging, interactive learning environments that extend beyond the classroom. Through natural language processing, speech recognition, and machine learning, learners now have access to intelligent systems that adapt dynamically to their proficiency levels. Widely used applications such as Duolingo, ChatGPT, and speech-analysis software demonstrate how AI can support grammar correction, vocabulary development, pronunciation training, and conversational fluency. These tools not only enrich learning but also promote autonomy, flexibility, and motivation, enabling students to practice at their own pace and in diverse contexts. For educators, AI has the potential to reduce administrative workload, provide adaptive teaching resources, and enhance differentiated instruction. At the same time, its integration raises important challenges, including data privacy, technological accessibility, and the preservation of human interaction, which remains essential for developing cultural and communicative competence. This paper explores the growing role of AI in language education, with particular attention to its applications, benefits for both learners and

[1] Sai Gon University, *Email: pndiem@sgu.edu.vn*

teachers, and the challenges that must be addressed to ensure effective and ethical implementation in diverse learning environments.

2. Literature Review

The application of Artificial Intelligence (AI) in language education has attracted increasing scholarly attention, with research highlighting its potential to transform both teaching and learning processes. AI technologies such as machine learning, natural language processing, and speech recognition have reshaped how languages are taught and acquired, offering opportunities for personalization, automation, and interaction.

2.1 Personalized and Adaptive Learning

Intelligent tutoring systems are among the most widely studied applications of AI. Heffernan and Heffernan (2014) demonstrate that such systems can analyze learners' performance in real time and adapt content accordingly, effectively addressing gaps in grammar, vocabulary, and pronunciation. Platforms like Duolingo exemplify this approach by modeling user knowledge and suggesting individualized learning activities. These adaptive systems make instruction more efficient and learner-centered.

2.2 Automated Feedback and Assessment

Another key contribution of AI lies in its ability to provide instant, automated feedback. Wang and Sleeman (2020) show that tools such as Grammarly and Write & Improve offer corrections on grammar, style, and vocabulary use, allowing learners to refine their writing independently. Unlike traditional assessment, which is often delayed, AI feedback supports continuous practice both inside and outside the classroom.

2.3 AI for Speaking and Listening Practice

Speech recognition technologies have been particularly impactful in improving oral proficiency. Tools such as ELSA Speak deliver real-time corrective feedback, enabling learners to monitor and adjust their pronunciation. Cheng (2021) found that students using such applications demonstrated increased confidence in speaking and listening tasks, underscoring the motivational role of instant feedback.

2.4 Chatbots and Conversational AI

Conversational agents simulate authentic interaction and provide safe, low-pressure spaces for practice. Fryer and Carpenter (2006) reported that learners engaging with chatbots gained confidence and improved comprehension by being regularly exposed to natural language patterns. Such tools can be especially valuable in contexts where opportunities for real-life communication are limited.

2.5 Teacher Support and Content Generation

In addition to supporting learners, AI can assist teachers by generating quizzes, lesson plans, and interactive exercises. Sun (2023) observes that generative AI tools such as ChatGPT

are increasingly used by educators to design customized activities, thereby reducing workload and diversifying classroom content. Rather than replacing teachers, these tools can serve as valuable complements that enhance instructional efficiency.

2.6 Challenges in AI Integration

Despite these benefits, the integration of AI also raises important concerns. Selwyn (2019) cautions against over-reliance on technology, stressing that human interaction remains essential for the development of communicative competence. Issues of data privacy, algorithmic bias, and unequal access to digital resources highlight the ethical and social dimensions of AI adoption in education.

2.7 Synthesis

Overall, the literature suggests that AI has significant potential to enrich language learning by personalizing instruction, automating feedback, and expanding opportunities for practice. However, successful integration requires thoughtful implementation, ongoing teacher training, and a balance between technological innovation and human guidance. This study builds on existing research by examining not only the benefits of AI tools but also the perceptions of learners and educators, with the aim of identifying both opportunities and limitations for their use in language teaching.

3. Methodology

3.1 Research Design

This study employs a qualitative research approach to explore how Artificial Intelligence (AI) is applied in language teaching and to identify its perceived benefits and challenges from both learners' and educators' perspectives. A qualitative design was chosen to allow for in-depth exploration of real-world experiences and practices, focusing on meaning, interpretation, and context rather than numerical measurement.

A descriptive qualitative design was adopted to capture the ways AI is integrated into language education. The study emphasizes user experiences, observable outcomes, and pedagogical implications, providing insights into both practical applications and perceptions of AI tools.

3.2 Research Questions

1. How is Artificial Intelligence (AI) currently applied in language teaching and learning?
2. What benefits do learners and educators perceive from the use of AI-based tools in language education?
3. What challenges and limitations are associated with integrating AI into language teaching and learning?

3.3 Data Collection Methods

3.3.1. Literature Analysis

An extensive review of academic journals, books, conference proceedings, and credible online sources was conducted. Keywords such as *AI in language learning*, *educational technology*, *intelligent tutoring systems*, and *chatbots in education* were used to identify relevant publications between 2014 and 2024. Sources were included if they offered empirical evidence, theoretical discussion, or case-based insights into AI-supported language learning.

3.3.2. Case Studies

Three widely used AI-based platforms—Duolingo, Grammarly, and ELSA Speak—were selected for in-depth analysis. These platforms were chosen because of their popularity, accessibility, and distinct focus on different aspects of language learning (e.g., adaptability, grammar correction, pronunciation). Features such as feedback systems, personalization, and user interface were examined using available documentation, user reviews, and published evaluations.

3.3.3. Interviews

Semi-structured interviews were conducted with 10 participants:

5 language teachers who incorporate AI tools in their classrooms

30 language learners with experience using AI-assisted learning applications

Participants were recruited through purposive sampling to ensure familiarity with AI-based tools. Interview questions addressed the effectiveness, usability, and limitations of AI in language teaching.

3.4 Data Analysis

Interview transcripts were analyzed using thematic analysis (Braun & Clarke, 2006) to identify recurring patterns and themes. Case study findings were evaluated comparatively, highlighting similarities and differences in how platforms support language acquisition. Triangulation across literature, case studies, and interviews strengthened the validity of findings.

3.5 Ethical Considerations

All participants were informed about the purpose of the study and provided voluntary consent. Anonymity and confidentiality were ensured by assigning pseudonyms and securely storing data. Participants were informed of their right to withdraw from the study at any stage. Ethical standards were maintained throughout the research process.

4. Results and Discussion

Binomial Test

						95% CI for Proportion	
Variable	Level	Counts	Total	Proportion	p	Lower	Upper
item1 agree	0	4	50	0.080	< .001	0.022	0.192
	1	46	50	0.920	< .001	0.808	0.978
item2 agree	0	3	50	0.060	< .001	0.013	0.165
	1	47	50	0.940	< .001	0.835	0.987
item3 agree	1	50	50	1.000	< .001	0.929	1.000
item4 agree	1	50	50	1.000	< .001	0.929	1.000
item5 agree	0	5	50	0.100	< .001	0.033	0.218
	1	45	50	0.900	< .001	0.782	0.967
item6 agree	0	10	50	0.200	< .001	0.100	0.337
	1	40	50	0.800	< .001	0.663	0.900

Note. Proportions tested against value: 0.5.

4.1 Enhanced Personalization and Learner Engagement

Interview findings and case study analysis revealed that AI-powered platforms such as Duolingo and Grammarly significantly enhanced learner engagement through adaptive learning paths. Learners appreciated how lessons adjusted to their individual strengths and weaknesses, which helped sustain motivation and reduce frustration. One student noted, *"The app changed the questions when I struggled, so I felt like I was always learning at my own speed."* Teachers similarly highlighted that these tools supported differentiation in mixed-ability classrooms.

This finding aligns with Heffernan and Heffernan (2014), who stress the value of adaptive algorithms in providing targeted instruction. By tailoring content to learner needs, AI mitigates the "one-size-fits-all" problem in traditional classrooms. However, while personalization increases engagement, there is a risk that learners may become overly dependent on app-based scaffolding rather than developing independent learning strategies.

4.2 Instant Feedback and Autonomy in Learning

Learners using tools such as ELSA Speak and Grammarly reported substantial benefits from instant corrective feedback on grammar, style, and pronunciation. This immediacy built confidence, particularly in speaking and writing tasks, where traditional teacher feedback is often delayed. Several learners expressed that real-time guidance encouraged them to practice independently: *"I could try speaking several times until the app said it was correct—something I can't always do in class."*

These findings reinforce Wang and Sleeman (2020), who emphasize the importance of real-time feedback for improving fluency and accuracy. The development of learner autonomy also echoes Little (2007), who argued that self-directed practice strengthens language retention. Nonetheless, the quality of feedback must be scrutinized: while AI excels at mechanical correction, it lacks the cultural and pragmatic awareness that human teachers provide.

4.3 Increased Teacher Support and Efficiency

Teachers reported that AI tools reduced their workload by automating repetitive tasks such as grading, grammar correction, and resource creation. Applications like ChatGPT were particularly valued for generating lesson materials, quizzes, and writing prompts. As one teacher explained, *"I can spend more time guiding discussions because the AI helps me prepare resources faster."*

This supports Sun (2023), who observed that generative AI enables teachers to focus on higher-order pedagogy rather than administrative tasks. However, educators warned of a potential "deskilling effect," where over-reliance on AI might diminish teachers' professional expertise in lesson design. Thus, while AI enhances efficiency, it should be framed as a tool that *supports*—not supplants—teacher creativity and pedagogical judgment.

4.4 Challenges Identified

Despite the advantages, several challenges emerged. First, technical limitations were evident: accent variations reduced the accuracy of speech recognition systems, frustrating some learners. Second, an over-reliance on AI was observed, with students admitting to depending on apps for answers without fully processing the content. Third, issues of equity were highlighted, as not all learners had equal access to reliable devices or stable internet connections.

These challenges resonate with Selwyn (2019), who cautions against the dangers of technological determinism in education. While AI enhances efficiency and access, it cannot replace the social and cultural dimensions of language learning. Moreover, ethical concerns such as data privacy and algorithmic bias remain pressing issues, particularly as most AI tools are developed by commercial entities. To maximize their benefits, AI tools must be integrated thoughtfully, ensuring they complement human interaction and do not exacerbate digital inequalities.

5. Suggestions

For effective integration of AI in language education, several strategies are recommended: **Teacher training programs** to ensure educators use AI responsibly and creatively. **Equitable access initiatives** to bridge the digital divide and ensure all learners can benefit from AI tools.**Ethical frameworks** that address privacy, transparency, and algorithmic fairness. **Longitudinal research** to evaluate the long-term impact of AI on language acquisition and learning outcomes.

5.1 Suggestions for Teachers

5.1.1 Use AI as a Supplement, Not a Substitute

Incorporate AI tools (Duolingo, Grammarly, ELSA Speak, ChatGPT) to reinforce practice, but design activities that require critical thinking and human interaction. *Example:* Ask students to revise AI-generated feedback rather than accept it blindly.

5.1.2 Develop AI Literacy

Provide short training sessions on how AI works, including its strengths and limitations. Teach learners to critically evaluate AI feedback instead of treating it as "always correct."

5.1.3 Balance Efficiency with Creativity

Use AI to automate repetitive tasks (e.g., quiz creation, grammar checks), but ensure lesson design still includes open-ended discussions, cultural content, and creative activities.

5.1.4 Monitor Student Engagement

Track how students use AI, whether as a learning aid or a shortcut. Encourage reflective logs or journals where students explain how they used AI tools and what they learned from them.

5.1.5 Promote Equity and Access

Provide alternatives for students with limited internet or device access (e.g., offline resources, group sharing). Advocate for institutional support to ensure fair access to digital learning tools.

5.2 Suggestions for Students

5.2.1 Practice Active Learning

Use AI feedback as a guide, but always review and edit work manually afterward. *Example:* After Grammarly corrections, students should explain why the change was necessary.

5.2.2 Set Learning Goals

Define personal targets (e.g., improving pronunciation, mastering grammar accuracy) and use AI apps strategically instead of aimlessly.

5.2.3 Avoid Over-Reliance

Treat AI as a "practice partner," not a "solution provider." *Example:* When using ChatGPT for writing help, draft an outline first, then refine it with AI assistance.

5.2.4 Engage in Reflection

Keep a study journal to note what AI tools corrected, how they helped, and what still needs improvement.

5.2.5 Balance AI with Human Input

Combine AI-based learning with peer discussions, teacher feedback, and real-world communication to develop pragmatic and cultural competence.

6. Conclusion

The findings of this study highlight the transformative potential of Artificial Intelligence (AI) in language teaching. AI-driven platforms such as Duolingo, Grammarly, and ELSA Speak demonstrate how adaptive learning systems, real-time feedback, and generative content can enhance both teaching and learning experiences. Learners benefit from increased personalization, autonomy, and confidence, while teachers gain efficiency and pedagogical support. However, the study also identifies significant challenges, including technological limitations, the risk of over-dependence, equity concerns, and ethical issues related to data privacy. These findings underscore that AI cannot replace the nuanced, human-centered aspects of language education. Its greatest potential lies in functioning as a **pedagogical partner**, complementing teachers who provide cultural, emotional, and communicative depth.

In conclusion, AI holds significant promise for making language learning more accessible, personalized, and engaging. Its success, however, depends on **balanced implementation**, where technology empowers both learners and teachers without diminishing the irreplaceable value of human interaction.

References

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book* (2nd ed.). Heinle & Heinle.

Cheng, X. (2021). The effectiveness of AI-driven speech recognition tools in improving English pronunciation. *Journal of Language Learning and Technology, 25*(3), 45–61.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26*(2), 163–188. https://doi.org/10.1177/0267658310368814

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition, 24*(1), 143–188. https://doi.org/10.1017/S0272263102002024

Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology, 10*(3), 8–14.

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24*(4), 470–497. https://doi.org/10.1007/s40593-014-0024-x

Niglas, K. (2004). *The combined use of qualitative and quantitative methods in educational research* (Unpublished doctoral dissertation). Tallinn Pedagogical University, Tallinn, Estonia.

Schmitt, N. (2010). *Researching vocabulary*. Palgrave Macmillan.

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

Sun, Y. (2023). Generative AI in language classrooms: Opportunities and challenges for teachers. *Computers & Education, 195*, 104703. https://doi.org/10.1016/j.compedu.2023.104703

Thérin, F. (Ed.). (2007). *Handbook of research on techno-entrepreneurship*. Edward Elgar.

Tomasello, M. (2009). The usage-based theory of language acquisition. In E. Bavin (Ed.), *The handbook of child language* (pp. 69–87). Cambridge University Press.

Wang, Q., & Sleeman, D. (2020). AI-driven feedback in second language writing: Opportunities and limitations. *Language Teaching Research, 24*(3), 301–322. https://doi.org/10.1177/1362168818796360

VAI TRÒ CỦA VIỆC ỨNG DỤNG AI TRONG DẠY HỌC BIÊN - PHIÊN DỊCH TIẾNG TRUNG QUỐC TẠI TRƯỜNG ĐẠI HỌC SƯ PHẠM HÀ NỘI 2

Nguyễn Thị Nga[1], Nguyễn Thị Lan Anh[2]

Tóm tắt

Bài báo giới thiệu các hình thức ứng dụng các công cụ AI, những vai trò mà AI mang lại trong việc nâng cao chất lượng dạy và học biên – phiên dịch tiếng Trung Quốc cho sinh viên chuyên ngành Ngôn ngữ Trung Quốc tại Trường Đại học Sư phạm Hà Nội 2. Đồng thời, bài báo cũng chỉ ra những hạn chế, từ đó, đề xuất giải pháp kết hợp hài hòa giữa công nghệ và yếu tố con người trong dạy, học biên – phiên dịch tiếng Trung Quốc.

Từ khóa: AI, biên – phiên dịch tiếng Trung Quốc, dịch tiếng Trung, HPU2

1. Đặt vấn đề

Trí tuệ nhân tạo (thường được viết tắt là AI - Artificial Intelligence) trước hết được định nghĩa là "trí tuệ phát sinh bởi máy móc, đối lập với trí tuệ tự nhiên phát sinh bởi con người và các loài vật". Theo đó, khái niệm trí tuệ nhân tạo được áp dụng khi máy móc bắt chước được các chức năng lý trí gắn với trí tuệ con người, là sự "tư duy" của máy móc, trong đó các thiết bị sẽ bắt chước cách tư duy tự nhiên của con người để giải quyết các vấn đề. Trong cuộc Cách mạng công nghiệp 4.0, AI là một trong những yếu tố then chốt. Từ khái niệm tưởng chừng xa xôi, trí tuệ nhân tạo từng bước đi vào đời sống, hiện thực hóa giấc mơ về những loại máy móc có khả năng tư duy như con người. [1]

Năm 2020, Thủ tướng Chính phủ đã thông qua Chương trình Chuyển đổi số quốc gia đến năm 2025, tầm nhìn đến năm 2030, trong đó giáo dục và đào tạo được xác định là một trong tám lĩnh vực trọng điểm cần ưu tiên triển khai. Trong bối cảnh này, chuyển đổi số trong ngành giáo dục đóng vai trò quan trọng trong việc đổi mới phương pháp dạy học, từ lối giảng dạy truyền thống sang phát triển năng lực người học, nâng cao khả năng tự học, thúc đẩy học tập linh hoạt theo nhu cầu cá nhân và xây dựng nền tảng cho xã hội học tập và học tập suốt đời.

Với sứ mệnh đào tạo đội ngũ nhân lực chất lượng cao phục vụ sự nghiệp phát triển quốc gia, Trường Đại học Sư phạm Hà Nội 2 đang tích cực đẩy mạnh ứng dụng công nghệ số trong công tác quản trị, giảng dạy, kiểm tra đánh giá và nghiên cứu khoa học. [2] Đối với lĩnh vực giảng dạy biên - phiên dịch Tiếng Trung Quốc, AI mở ra nhiều hướng đi mới trong việc tăng cường tính tự động, cá nhân hóa và nâng cao hiệu quả dạy học.

Đề tài nhằm làm rõ:

Thực trạng việc ứng dụng AI trong giảng dạy biên - phiên dịch Tiếng Trung Quốc tại Trường Đại học Sư phạm Hà Nội 2.

[1] Khoa Tiếng Trung Quốc, Trường Đại học Sư phạm Hà Nội 2, *Email: nguyenthinga1988@hpu2.edu.vn*
[2] Khoa Tiếng Trung Quốc, Trường Đại học Sư phạm Hà Nội 2, *Email: nguyenthilananh@hpu2.edu.vn*

Tiềm năng, cơ hội và thách thức trong việc tổ chức dạy và học biên - phiên dịch Tiếng Trung Quốc với sự hỗ trợ của AI.

Đề xuất định hướng, giải pháp nhằm tăng cường hiệu quả ứng dụng AI trong giảng dạy biên - phiên dịch tiếng Trung Quốc.

Để thực hiện được những mục tiêu trên, bài báo cáo có sử dụng những phương pháp nghiên cứu sau:

Phương pháp tổng quan tài liệu: Thu thập, phân tích tài liệu lý luận và thực tiễn liên quan đến ứng dụng AI trong dạy và học ngoại ngữ nói chung, trong dạy học biên – phiên dịch tiếng Trung Quốc nói riêng.

Điều tra bằng bảng hỏi: Thiết kế bảng hỏi dành cho giảng viên và sinh viên chuyên ngành Ngôn ngữ Trung Quốc để khảo sát nhận thức, nhu cầu và khó khăn trong việc ứng dụng AI trong dạy học biên – phiên dịch tiếng Trung Quốc.

Bài tập thực hành: Tổ chức các bài tập nhận diện, sử dụng các công cụ AI hỗ trợ biên - phiên dịch cho sinh viên để đánh giá tác động và tự đánh giá hiệu quả sử dụng.

Hiện nay, trí tuệ nhân tạo (AI) đang trở thành xu hướng nổi bật trong giáo dục ngoại ngữ, đặc biệt trong lĩnh vực dạy học biên - phiên dịch. Việc áp dụng AI giúp nâng cao hiệu quả dạy và học, giảm áp lực cho giảng viên cũng như sinh viên, tăng tương tác và cá nhân hóa quá trình học tập.

Tại trường Đại học Sư phạm Hà Nội 2, dù đã có nhiều đề tài nghiên cứu về dạy và học tiếng Trung Quốc, nhưng chưa có bài nghiên cứu nào đề cập đến việc ứng dụng AI trong dạy học biên - phiên dịch tiếng Trung Quốc. Trong bối cảnh cách mạng hóa và chữ số hóa giáo dục ngày càng phát triển, việc nghiên cứu đề tài này là hết sức cần thiết và có ý nghĩa thực tiễn cao.

Nghiên cứu không chỉ bổ sung kho tàng lý luận, tài liệu tham khảo cho giáo viên và sinh viên ngành Tiếng Trung Quốc tại Trường, mà còn đóng góp vào định hướng đổi mới phương pháp dạy học, tổ chức đào tạo, nhằm đảm bảo chất lượng đào tạo đối với biên – phiên dịch - một lĩnh vực ngày càng có nhu cầu cao trong xã hội hiện đại.

2. Các nội dung chính

2.1. Thực trạng giảng dạy môn biên – phiên dịch tiếng Trung Quốc trước khi áp dụng ứng dụng AI

Trước khi ứng dụng trí tuệ nhân tạo, hoạt động giảng dạy biên – phiên dịch tiếng Trung Quốc còn bộc lộ nhiều hạn chế, đặc biệt trong việc cập nhật học liệu và tổ chức môi trường thực hành. Ở mảng biên dịch, nguồn tài liệu chủ yếu vẫn dựa vào sách vở, giáo trình in sẵn, nhiều khi đã lỗi thời, không còn phản ánh kịp thời những biến động ngôn ngữ và đời sống xã hội. Việc tìm tài liệu để giảng dạy môn biên – phiên dịch cũng khá mang tính chủ quan, khi giảng viên thường lựa chọn những thông tin nổi bật theo cảm nhận cá nhân để cung cấp cho sinh viên dịch, dẫn đến tình trạng ngữ liệu chưa thật sự đa dạng và sát với thực tiễn. Điều này khiến sinh viên được đào tạo với vốn từ ngữ, cách diễn đạt cũ kỹ, xa rời bối cảnh hiện đại; khi

ra trường, họ gặp nhiều khó khăn trong việc thích ứng với ngôn ngữ chuyên ngành mới, thuật ngữ cập nhật hàng ngày hoặc những dạng văn bản có tính thời sự cao. Trong khi đó, biên dịch trong thực tế nghề nghiệp đòi hỏi khả năng xử lý linh hoạt, cập nhật liên tục và sử dụng ngôn ngữ sát với ngữ cảnh xã hội đương đại.

Đối với phiên dịch, yêu cầu đặt ra là sinh viên phải nghe và dịch theo tình huống, thường dưới áp lực thời gian. Tuy nhiên, trong môi trường đào tạo truyền thống, sinh viên chủ yếu được tiếp cận một số đoạn hội thoại, bài nghe ngắn với tốc độ có thể nhanh hơn khả năng tiếp nhận của họ, mà lại thiếu các công cụ hỗ trợ để điều chỉnh tốc độ hoặc phân tích ngữ điệu. Hơn nữa, lớp học chưa tạo được môi trường mô phỏng thực tế như hội nghị, đàm phán hay phỏng vấn song ngữ, khiến sinh viên khó hình dung bối cảnh giao tiếp và chưa rèn luyện được năng lực xử lý thông tin tức thời. Điều này dẫn đến sự chênh lệch lớn giữa kỹ năng được đào tạo trong nhà trường và yêu cầu thực tiễn của nghề phiên dịch, nơi tốc độ, phản xạ và sự chính xác đóng vai trò then chốt.

2.2.Thực trạng ứng dụng AI trong dạy học biên – phiên dịch tiếng Trung Quốc tại Trường Đại học Sư phạm Hà Nội 2

Trong bối cảnh toàn cầu hóa và cách mạng công nghiệp 4.0, việc ứng dụng AI vào giáo dục đang trở thành xu thế tất yếu. Tại Trường Đại học Sư phạm Hà Nội 2, ngành Ngôn ngữ Trung Quốc đang nỗ lực đổi mới phương pháp giảng dạy, kết hợp công nghệ AI nhằm nâng cao chất lượng đào tạo và phát triển năng lực nghề nghiệp cho sinh viên, đáp ứng nhu cầu xã hội ngày càng cao về nguồn nhân lực chất lượng cao trong lĩnh vực ngôn ngữ.

2.2.1.Trong giảng dạy, kiểm tra đánh giá

100% giảng viên giảng dạy các học phần liên quan đến biên – phiên dịch tiếng Trung Quốc tại Khoa Tiếng Trung Quốc, Trường Đại học Sư phạm Hà Nội 2 có sử dụng AI trong công việc. Cụ thể:

Lồng ghép AI trong giờ thực hành dịch: Giảng viên yêu cầu sinh viên so sánh bản dịch của mình với bản dịch máy, từ đó rút ra điểm mạnh – yếu và nâng cao kỹ năng hiệu đính.

Giảng viên yêu cầu sinh viên luyện tập phiên dịch qua các câu, đoạn văn bản, đoạn hội thoại mô phỏng do AI tạo ra.

Giảng viên sử dụng các công cụ AI để tạo kho bài tập và kiểm tra đa dạng hơn.

Giảng viên phân tích cho sinh viên cách sử dụng có chọn lọc, có phê phán các công cụ dịch tự động, tránh tình trạng phụ thuộc máy móc.

Giảng viên sử dụng AI để tham khảo những bản dịch gợi ý, kiểm tra kết cấu bài kiểm tra…

Tiện ích mà AI mang lại cho giảng viên trong giảng dạy biên – phiên dịch tiếng Trung Quốc thể hiện ở các khía cạnh sau:

Góp phần đổi mới phương pháp giảng dạy: Trước hết, AI hỗ trợ giảng viên xây dựng các tài liệu học tập phong phú, đa dạng và sát với thực tiễn.

Công cụ xử lý ngôn ngữ tự nhiên (NLP) được giảng viên sử dụng: Google Translate, DeepL, Baidu Fanyi …, các phần mềm nhận dạng giọng nói và chuyển đổi văn bản hỗ trợ hiệu quả trong giảng dạy biên - phiên dịch.

AI đồng thời cũng đóng vai trò như một trợ lý ảo, hỗ trợ giảng viên trong việc thiết kế bài giảng, bài tập thực hành, mô phỏng tình huống thực tế giúng như các buổi hội nghị, hội đàm song ngữ. Qua đó, sinh viên có cơ hội thực hành kỹ năng trong môi trường ảo trước khi áp dụng vào thực tiễn.

Trong kiểm tra, đánh giá kỹ năng biên - phiên dịch tiếng Trung Quốc, giảng viên sử dụng trí tuệ nhân tạo (AI) như một công cụ hỗ trợ hiệu quả: AI giúp giảng viên tham khảo các bản dịch gợi ý từ nhiều nguồn đáng tin cậy, hỗ trợ so sánh, phân tích các phương án dịch khác nhau để đưa ra nhận xét khách quan hơn.

Ngoài ra, AI còn hỗ trợ kiểm tra kết cấu bài kiểm tra, giúp phát hiện lỗi ngữ pháp, cú pháp hoặc những phần chưa hợp lý trong diễn đạt… và sửa lỗi tức thì. Nhờ đó, giảng viên có thể tiết kiệm thời gian trong việc chấm bài, đồng thời đảm bảo tính chính xác và công bằng trong quá trình đánh giá.

2.2.2.Trong học tập

Theo kết quả khảo sát 105 sinh viên năm thứ ba và thứ tư chuyên ngành Ngôn ngữ Trung Quốc tại Khoa Tiếng Trung Quốc, Trường Đại học Sư phạm Hà Nội 2, 100% sinh viên sử dụng AI trong quá trình học biên – phiên dịch.

Các ứng dụng AI mà những sinh viên này sử dụng bao gồm: ChatGPT, Deep seek, Gemini, Kimi, Doubao, Baidu Fanyi…

Theo khảo sát, sinh viên sử dụng ứng dụng AI vào học biên – phiên dịch nhằm:

Giải đáp các thắc mắc về từ vựng, ngữ pháp, cấu trúc câu, nhất là khi gặp phải một cụm từ khó hiểu hoặc chưa rõ nghĩa.

Trau chuốt bản dịch bằng cách đề xuất từ ngữ tự nhiên hơn, chuẩn xác hơn, góp phần làm cho bản dịch trở nên mượt mà, đúng ngữ pháp và mang tính bản ngữ cao hơn.

Hỗ trợ sửa lỗi ngữ pháp và kiểm tra toàn bộ văn bản dịch để nâng cao chất lượng bản dịch.

Ngoài ra, khi gặp các đoạn âm thanh hay video có tốc độ nói quá nhanh, vượt quá khả năng nghe hiểu, sinh viên có thể tận dụng AI để phiên âm tự động hoặc dịch nội dung một cách chính xác, giúp tiết kiệm thời gian và giảm bớt áp lực.

Đặc biệt, với những sinh viên còn hạn chế về vốn tiếng Việt – chẳng hạn như chưa diễn đạt được trôi chảy hoặc dùng từ chưa phù hợp – AI cũng có thể đóng vai trò là "người bạn chỉnh sửa", giúp cải thiện khả năng diễn đạt và truyền tải thông điệp một cách hiệu quả hơn.

2.3.Vai trò của việc ứng dụng AI trong dạy học biên - phiên dịch tiếng Trung Quốc

AI mang lại nhiều lợi ích rõ rệt cho công tác giảng dạy và học tập, cụ thể:

Trong thời đại công nghệ số, AI đang trở thành công cụ quan trọng trong nhiều lĩnh vực, đặc biệt là trong giảng dạy biên - phiên dịch tiếng Trung Quốc. AI đem lại nhiều tiện ích và hỗ trợ đắc lực cho giảng viên và sinh viên trong quá trình học tập, rèn luyện kĩ năng dịch.

Trong thời đại công nghệ 4.0, trí tuệ nhân tạo (AI) đã và đang đóng vai trò quan trọng trong nhiều lĩnh vực, trong đó có giáo dục, đặc biệt là trong giảng dạy biên – phiên dịch tiếng Trung Quốc. AI không chỉ là công cụ hỗ trợ mà còn là nhân tố góp phần đổi mới phương pháp giảng dạy, nâng cao chất lượng đào tạo và mở rộng cơ hội học tập cho người học.

Trước hết, AI giúp tối ưu hóa việc học từ vựng, ngữ pháp và cấu trúc câu nhờ các ứng dụng thông minh như từ điển trực tuyến, phần mềm dịch tự động (Google Translate, DeepL, Baidu Fanyi), và các nền tảng học tiếng có tích hợp AI như Duolingo hay iFlytek. Những công cụ này hỗ trợ người học luyện tập dịch thuật, so sánh bản dịch và học theo ngữ cảnh cụ thể.

AI giúp nhận diện và phản hồi lỗi sai ngay tức thì, từ ngữ pháp, cấu trúc câu đến cách sử dụng từ vựng trong các bài viết hoặc bài tập. Việc được sửa lỗi nhanh chóng không chỉ giúp sinh viên cải thiện kỹ năng mà còn tạo điều kiện cho việc học trở nên chủ động, hiệu quả hơn.

AI cũng có thể mô phỏng các tình huống phiên dịch thực tế như hội nghị, giao tiếp thương mại hay phỏng vấn. Nhờ công nghệ nhận diện giọng nói và phiên âm tức thời, người học có thể luyện tập với phản hồi nhanh chóng, cải thiện khả năng nghe – nói và rèn luyện phản xạ ngôn ngữ.

Đối với giảng viên, AI góp phần giảm bớt áp lực chấm bài thủ công vốn tốn nhiều thời gian và công sức. Các hệ thống chấm điểm tự động, phân tích nội dung và đánh giá theo tiêu chí khách quan giúp giảng viên tiết kiệm thời gian, từ đó có thể tập trung nhiều hơn vào công tác giảng dạy, nghiên cứu và hỗ trợ cá nhân hóa quá trình học cho từng sinh viên. Như vậy, việc tích hợp AI vào giáo dục không chỉ nâng cao chất lượng đào tạo mà còn thúc đẩy sự đổi mới phương pháp dạy và học theo hướng hiện đại, linh hoạt và hiệu quả hơn.

Ngoài ra, AI còn hỗ trợ giảng viên trong thiết kế giáo án số, xây dựng ngân hàng bài tập tự động, cá nhân hóa tiến trình học và đánh giá năng lực người học theo thời gian thực. Việc kết hợp AI vào dạy học biên – phiên dịch giúp tiết kiệm thời gian, giảm tải công việc cho giáo viên, đồng thời tạo ra môi trường học tập linh hoạt, hấp dẫn và hiệu quả hơn.

2.3.Những hạn chế khi ứng dụng AI trong dạy học biên - phiên dịch tiếng Trung Quốc

Bên cạnh những lợi ích rõ rệt như nâng cao hiệu quả học tập, hỗ trợ người học cá nhân hóa nội dung học, rèn luyện kỹ năng dịch nhanh chóng…, việc ứng dụng AI vào trong công tác dạy học biên - phiên dịch tiếng Trung Quốc vẫn tồn tại không ít hạn chế và thách thức cần được nhận diện và giải quyết một cách thấu đáo.

Thứ nhất, AI vẫn còn hạn chế trong việc xử lý ngữ cảnh, sắc thái ngôn ngữ và văn hóa đặc thù. Dịch thuật, đặc biệt là biên – phiên dịch tiếng Trung Quốc, không chỉ đơn thuần là

chuyển ngữ từ một hệ thống ngôn ngữ sang hệ thống ngôn ngữ khác, mà còn bao hàm sự chuyển tải về văn hóa, sắc thái biểu cảm, ngữ cảnh xã hội và mục đích giao tiếp.

AI hiện nay, dù đã đạt đến mức xử lý ngôn ngữ tự nhiên rất tốt, vẫn còn lúng túng trong việc phân biệt sắc thái biểu cảm (ví dụ như mỉa mai, hài hước, ẩn dụ hay châm biếm), dẫn đến bản dịch đôi khi trở nên máy móc, thiếu sự linh hoạt và độ tinh tế cần có. Một đoạn văn có thể được dịch chính xác về mặt từ vựng và ngữ pháp nhưng lại mất đi "hồn cốt" của bản gốc, đặc biệt khi gặp những thành ngữ, điển cố, hoặc các yếu tố văn hóa đặc sắc của tiếng Trung như điển cố, thành ngữ, yết hậu ngữ…

Theo kết quả bài tập giao cho 50 sinh viên năm thứ hai chuyên ngành Ngôn ngữ Trung Quốc với yêu cầu: Anh/chị hãy sử dụng AI để dịch các câu sau. So sánh kết quả dịch của 2 ứng dụng dịch mà anh/chị vừa sử dụng để phiên dịch.

Câu 1: Buổi học cuối, lớp tôi "vắng ngắt như chùa Bà Đanh".

Trong văn hóa, quan niệm của người Việt Nam, "chùa Bà Đanh" là cụm từ có ý nghĩa chỉ sự vắng vẻ, tĩnh mịch, vắng người qua lại. Như vậy, khi dịch sang tiếng Trung Quốc, cần thể hiện ý nghĩa là: Buổi học cuối mà lớp tôi vắng vẻ, tĩnh mịch, ít học sinh đến học. 最后一节课，我们班来上课的同学很少。(Zuìhòu yì jié kè, wǒmen bān lái shàngkè de tóngxué hěn shǎo.)

Nhưng theo kết quả sinh viên sử dụng AI dịch:

最后一节课，我的班上冷清得像空荡荡的庙一样。（Zuìhòu yì jié kè, wǒ de bān shàng lěngqīng de xiàng kōngdàngdàng de miào yíyàng.）Hoặc: 最后一节课，我们班空荡荡的，像个冷清的寺庙。(Zuìhòu yī jié kè, wǒmen bān kōngdàngdàng de, xiàng gè lěngqīng de sìmiào.) Nghĩa là: Tiết học cuối cùng, lớp tôi yên lặng như một ngôi chùa trống vắng/hoang vắng.

Câu 2: Trông bé ấy như "Thị Nở".

Thị Nở vốn là một nhân vật trong tác phẩm văn học, được nhà văn Nam Cao miêu tả là một người đàn bà xấu "ma chê quỷ hờn". Bởi vậy, trong tâm thức của người Việt Nam, khi nhắc đến từ "Thị Nở", ai cũng hiểu rằng đó là miêu tả người phụ nữ ngoại hình xấu xí. Vì vậy, khi dịch câu trên sang tiếng Trung, chỉ cần làm rõ ý nghĩa con bé ấy rất xấu là được.

Nhưng theo kết quả sinh viên sử dụng AI dịch:

那女孩子看起来像市娥一样。(Nà nǚ háizǐ kàn qǐlái xiàng Shì É yíyàng.) 市娥 - Shì É của bản dịch là cách dịch tên riêng mà chưa làm rõ được nghĩa của câu. Vì vậy, nếu sinh viên sử dụng bản dịch này để giao tiếp với người Trung Quốc, chắc chắn sẽ gây ra sự khó hiểu.

Câu 3: Hôm nay, cậu nhất định phải Cam-pu-chia với bọn tớ.

Từ Cam-pu-chia trong câu này là từ lóng, sử dụng lối nói đồng âm Cam-pu-chia với từ "chia" (chia tiền để trả), trong tiếng Trung, từ tương đương là AA 制 (AA zhì).

Nhưng theo kết quả sinh viên sử dụng AI dịch:

今天你一定得跟我们一起“去柬埔寨”。(Jīntiān nǐ yídìng děi gēn wǒmen yìqǐ “qù Jiǎnpǔzhài”.) Nghĩa là: Hôm nay, cậu nhất định phải đi Cam-pu-chia cùng với bọn tớ. Như vậy, bản dịch đã hoàn toàn không thể diễn đạt được ý của bản gốc. Đó chính là sự thất bại trong phiên dịch.

Câu 4: 南柯一梦 (Nánkē yì mèng) nghĩa là mơ mộng hão huyền. Nghĩa mặt chữ của câu này là giấc mộng Nam Kha. Thành ngữ này xuất phát từ truyện “Nam Kha thái thú truyện” (南柯太守传) của Lý Công Tá (李公佐) đời Đường. Chuyện kể rằng có người nằm mộng được sống vinh hoa phú quý trong một vương quốc dưới gốc cây hòe phía nam (Nam Kha), tỉnh dậy thì phát hiện tất cả chỉ là mộng. Nếu dịch sang tiếng Việt, người dịch chỉ diễn đạt là giấc mộng Nam Kha, thì người nghe sẽ không hiểu được ý nghĩa là gì.

Thứ hai, AI có thể tạo ra sự phụ thuộc thái quá ở người học, nhất là đối với sinh viên chưa vững kiến thức nền. Trong môi trường giảng dạy đại học, sinh viên thường có xu hướng sử dụng các công cụ để thực hiện bài tập dịch mà không thông qua quá trình tư duy ngôn ngữ độc lập.

Hệ quả là kỹ năng phân tích, tra cứu và xử lý ngôn ngữ gốc cũng như khả năng đối sánh ngữ pháp, tìm kiếm tương đương ngôn ngữ dần mai một. Nếu người học chỉ dừng lại ở việc “chấp nhận” kết quả dịch do AI tạo ra, thay vì đánh giá, phản biện, chỉnh sửa và học hỏi từ đó, thì việc sử dụng AI không những không mang lại hiệu quả tích cực mà còn phản tác dụng trong quá trình đào tạo năng lực biên – phiên dịch thực sự.

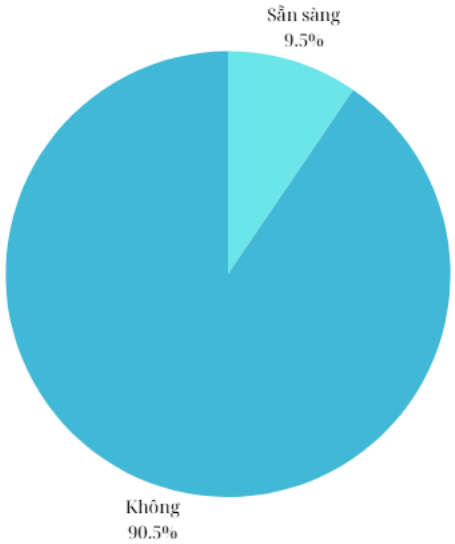
Thứ ba, AI thiếu khả năng phản hồi linh hoạt và tương tác cảm xúc như con người. Một giảng viên khi giảng dạy biên – phiên dịch không chỉ truyền đạt kiến thức mà còn truyền cảm hứng, xây dựng tình huống thực tế, tổ chức hoạt động nhóm, mô phỏng phiên dịch hội nghị, xử lý khủng hoảng dịch thuật... Những yếu tố đó đòi hỏi sự đồng cảm, khả năng điều tiết cảm xúc, hiểu được biểu cảm phi ngôn ngữ – điều mà AI hiện nay vẫn chưa làm được hiệu quả. Trong quá trình biên - phiên dịch đồng thời, ví dụ, người học cần luyện khả năng bắt tín hiệu từ giọng nói, ngữ điệu, tốc độ nói, cảm xúc của người nói để điều chỉnh bản dịch; tuy nhiên các công cụ AI chưa thể mô phỏng và phản hồi những yếu tố này một cách sống động và tự nhiên như giảng viên.

Thứ tư, AI không thể thay thế hoàn toàn vai trò của trải nghiệm thực tiễn và đạo đức nghề nghiệp trong giảng dạy dịch thuật. Biên – phiên dịch là một nghề nghiệp đòi hỏi không chỉ kiến thức ngôn ngữ mà còn cả kỹ năng nghề nghiệp, khả năng thích ứng trong môi trường đa văn hóa, đạo đức dịch thuật và trách nhiệm xã hội. AI không thể truyền đạt được những giá trị đạo đức nghề nghiệp như tính bảo mật, tính trung thực trong dịch, hay tinh thần trách nhiệm trước các tài liệu nhạy cảm. Trong khi đó, giảng viên có thể xây dựng được các bài học tình huống mô phỏng đạo đức nghề nghiệp, tạo ra môi trường tranh luận về những lựa chọn dịch khó khăn, từ đó hình thành tư duy phản biện và bản lĩnh nghề nghiệp cho người học.

Thứ năm, AI hiện nay vẫn phụ thuộc vào chất lượng dữ liệu đầu vào, và nguy cơ sai lệch thông tin là rất cao. Nhiều công cụ dịch thuật tự động hoặc chatbot ngôn ngữ được huấn luyện từ kho dữ liệu lớn nhưng không phải lúc nào cũng được kiểm chứng chính xác. Do đó, khi sinh viên sử dụng các công cụ AI để học dịch, họ có thể vô tình tiếp cận những bản dịch không chuẩn, dẫn đến hiểu sai hoặc dịch sai. Nếu thiếu sự hướng dẫn và giám sát từ phía giảng viên, việc học thông qua AI dễ dẫn tới lệch lạc trong tiếp thu ngôn ngữ và hình thành tư duy dịch không chính xác, đặc biệt với các thuật ngữ chuyên ngành hoặc ngôn ngữ văn bản hành chính, pháp lý.

Thứ sáu, việc triển khai AI trong giảng dạy đòi hỏi hạ tầng kỹ thuật, nhân lực và tài chính phù hợp, điều mà nhiều cơ sở giáo dục chưa đáp ứng được. Các phần mềm AI hiện đại thường yêu cầu máy móc cấu hình cao, internet ổn định, có phiên bản trả phí hoặc cần đội ngũ kỹ thuật hỗ trợ trong quá trình cài đặt và sử dụng. Ở nhiều trường đại học, đặc biệt tại các tỉnh hoặc cơ sở còn hạn chế về công nghệ thông tin, việc tích hợp AI vào giảng dạy biên – phiên dịch vẫn còn mang tính hình thức, chưa đi vào chiều sâu hoặc chưa được đồng bộ giữa các học phần, dẫn đến hiệu quả chưa cao.

Theo khảo sát đối với 105 sinh viên năm thứ ba và năm thứ tư chuyên ngành Ngôn ngữ Trung Quốc, Trường Đại học Sư phạm Hà Nội 2, chỉ có 10 sinh viên (chiếm 9,5%) trả lời sẵn sàng trả phí/có thể trả phí sử dụng các ứng dụng AI nếu cần trong quá trình biên – phiên dịch tiếng Trung Quốc. Tức là, có đến 90,5% sinh viên không bằng lòng để trả phí sử dụng các ứng dụng AI để tiếp cận và khai thác các công cụ hỗ trợ trí tuệ nhân tạo, cho thấy một rào cản nhất định trong việc ứng dụng công nghệ hiện đại vào đào tạo chuyên ngành, đặc biệt là khi các phần mềm AI ngày càng đóng vai trò quan trọng trong lĩnh vực ngôn ngữ, ngoại ngữ.



Hình 1: Tỉ lệ sinh viên sẵn sàng trả phí sử dụng phần mềm AI

Cuối cùng, sự phát triển nhanh chóng của AI đặt ra thách thức về năng lực cập nhật và thích ứng của giảng viên. Để tích hợp AI vào giảng dạy một cách hiệu quả, giảng viên cần

không ngừng học hỏi, làm chủ các công cụ mới, đồng thời sáng tạo trong phương pháp giảng dạy để không bị "lạc hậu" so với công nghệ.

Tuy nhiên, trong thực tế, nhiều giảng viên ngoại ngữ, đặc biệt là thế hệ lớn tuổi, vẫn còn tâm lý e ngại công nghệ hoặc chưa được đào tạo bài bản về việc ứng dụng AI, từ đó tạo ra khoảng cách thế hệ trong tiếp cận và triển khai giảng dạy hiện đại.

3. Giải pháp thúc đẩy ứng dụng AI trong giảng dạy biên - phiên dịch tiếng Trung Quốc, Trường Đại học Sư phạm Hà Nội 2

Để AI thực sự phát huy hiệu quả trong đào tạo, cần triển khai đồng bộ nhiều giải pháp. Trước tiên, Nhà trường cần tăng cường đầu tư hạ tầng công nghệ thông tin, xây dựng phòng học thông minh và tích hợp các phần mềm dịch thuật tự động như: Google Translate, DeepL, Trados, ChatGPT… vào quá trình giảng dạy và thực hành.

Song song với đó, việc nâng cao năng lực số cho đội ngũ giảng viên là vô cùng quan trọng. Các chương trình tập huấn, hội thảo chuyên đề về công nghệ dịch thuật, sử dụng AI và dữ liệu lớn cần được tổ chức thường xuyên để giúp giảng viên cập nhật xu hướng và đổi mới phương pháp giảng dạy.

Mặt khác, Nhà trường cần xây dựng chương trình đào tạo linh hoạt, tích hợp các học phần về công nghệ dịch, kỹ năng sử dụng công cụ hỗ trợ dịch thuật hiện đại vào khung chương trình chính khóa. Sinh viên cần được tiếp cận với các dự án dịch thuật mô phỏng thực tế, có sự hỗ trợ của AI, nhằm rèn luyện kỹ năng vận dụng công nghệ trong môi trường nghề nghiệp.

Ngoài ra, việc hợp tác với doanh nghiệp công nghệ và các đơn vị cung cấp phần mềm dịch thuật để cập nhật công cụ mới, cấp quyền truy cập bản quyền phần mềm cho sinh viên cũng là một bước đi cần thiết.

Bên cạnh đó, việc xây dựng kho dữ liệu ngôn ngữ học mở và hệ thống bài tập tương tác dựa trên AI sẽ giúp tăng cường tính cá nhân hóa và tính thích ứng trong dạy và học.

Cuối cùng, cần có chính sách khuyến khích nghiên cứu khoa học trong lĩnh vực AI ứng dụng trong dịch thuật, đặc biệt là phát triển các mô hình dịch máy kết hợp với năng lực ngôn ngữ con người, nhằm phục vụ đào tạo chuyên sâu và đáp ứng yêu cầu của thị trường lao động. Việc kết hợp giữa công nghệ và chuyên môn ngôn ngữ sẽ giúp đào tạo ra những phiên dịch viên linh hoạt, sáng tạo và có khả năng thích ứng cao với môi trường làm việc hiện đại.

4. Kết luận

Ứng dụng AI trong giảng dạy biên - phiên dịch tiếng Trung Quốc tại Trường Đại học Sư phạm Hà Nội 2 không chỉ là xu hướng mà còn là nhu cầu cấp thiết nhằm nâng cao chất lượng đào tạo, đáp ứng yêu cầu hội nhập quốc tế, là hướng đi chiến lược phù hợp thời đại chuyển đổi số, giúp cải thiện chất lượng đào tạo, nâng cao tính chuyên nghiệp và tăng hiệu quả giảng dạy – học tập.

Mặc dù AI mang lại nhiều tiềm năng cho việc đổi mới phương pháp giảng dạy biên – phiên dịch tiếng Trung Quốc, nhưng việc ứng dụng công nghệ này vẫn cần sự cân nhắc kỹ lưỡng. Việc sử dụng AI nên được định hướng là công cụ hỗ trợ, chứ không thể thay thế hoàn toàn vai trò của con người trong đào tạo dịch thuật.

Giải pháp cần thiết là xây dựng chương trình đào tạo kết hợp giữa AI và yếu tố con người, trong đó người học được khuyến khích tư duy phản biện, phân tích, đối chiếu bản dịch của AI và học từ những sai lệch đó, còn giảng viên đóng vai trò là người hướng dẫn, cố vấn học thuật và người truyền cảm hứng nghề nghiệp. Chỉ khi đó, việc ứng dụng AI mới thực sự phát huy được hiệu quả trong đào tạo đội ngũ biên – phiên dịch chất lượng cao, thích ứng với thời đại số.

Nhà trường cũng cần hoàn thiện các giải pháp về kỹ thuật, đào tạo nhân lực, chính sách, đảm bảo phục vụ nhu cầu thực tế và chuẩn mực giáo dục.

Tài liệu tham khảo

[1] Vũ Thị Linh (2019), *Trí tuệ nhân tạo: Góc nhìn và giải pháp*

[2] Trường Đại học Sư phạm Hà Nội 2 (2021), *Trường ĐHSP Hà Nội 2 đẩy mạnh chuyển đổi số đáp ứng yêu cầu thực tiễn*, https://www.hpu2.edu.vn/vi-VN/hpu2-bai-viet/tin-tuc/truong-dhsp-ha-noi-2-day-manh-chuyen-doi-so-dap-ung-yeu-cau-thuc-tien, thời gian truy cập: Ngày 29 tháng 6 năm 2025.

ỨNG DỤNG LÝ THUYẾT NGÔN NGỮ HỌC CHỨC NĂNG HỆ THỐNG VÀO VIỆC THIẾT KẾ KHUNG PROMPT ĐỂ TẠO SINH NGỮ LIỆU GIẢNG DẠY TIẾNG ANH

LÊ THÀNH DANH[1] - LÊ HỮU SỸ[2]

Tóm tắt: Trong bối cảnh tích hợp Trí tuệ Nhân tạo (AI) vào giáo dục, các Mô hình Ngôn ngữ Lớn (Large Language Models - LLMs) mang lại tiềm năng to lớn cho việc tạo sinh ngữ liệu giảng dạy. Tuy nhiên, việc khai thác tiềm năng này đòi hỏi một phương pháp có hệ thống, để chỉ dẫn (prompt) các LLMs tạo ra kết quả đầu ra, có thể kiểm soát được về mặt ngôn ngữ và phù hợp với mục tiêu sư phạm. Báo cáo này trình bày việc xây dựng Khung chỉ dẫn đầu vào (prompt) dựa trên Lý thuyết Ngôn ngữ học chức năng hệ thống (SFL) - được đặt tên là SFL-based Prompt, một phương pháp nhằm góp phần giải quyết thách thức trên. Bằng việc tích hợp các khái niệm cốt lõi của SFL vào cấu trúc của một prompt, cho phép giáo viên điều khiển các đặc trưng ngôn ngữ của văn bản đầu ra, tương ứng với các cấp độ năng lực trong Khung tham chiếu trình độ ngôn ngữ chung của Châu Âu (CEFR).

Từ khoá: mô hình ngôn ngữ lớn (LLMs), khung chỉ dẫn đầu vào (prompt), ngôn ngữ học chức năng hệ thống (SFL), giảng dạy tiếng Anh.

1. Đặt vấn đề

Những tiến bộ vượt bậc trong lĩnh vực Trí tuệ Nhân tạo (AI), đặc biệt là sự phát triển của các Mô hình Ngôn ngữ Lớn (LLMs), đang mở ra những cơ hội và thách thức mới cho ngành giáo dục. Trong bối cảnh giảng dạy ngoại ngữ, các LLMs như ChatGPT cho thấy một tiềm năng to lớn trong việc tạo sinh ngữ liệu học tập. Khả năng tạo sinh các văn bản một cách nhanh chóng và đa dạng của các mô hình này có thể hỗ trợ giáo viên giải quyết thách thức cố hữu về việc thiếu hụt nguồn tài liệu phong phú, được cá nhân hóa và phù hợp với các cấp độ năng lực đa dạng của người học. Việc điều chỉnh ngữ liệu theo các tiêu chuẩn quốc tế như Khung tham chiếu trình độ ngôn ngữ chung của Châu Âu (CEFR) là một ứng dụng đầy hứa hẹn, giúp kết nối trực tiếp hoạt động lớp học với các mục tiêu năng lực được công nhận rộng rãi.

Tuy nhiên, để khai thác hiệu quả tiềm năng của việc ứng dụng lý thuyết Ngôn ngữ học Chức năng Hệ thống (SFL) vào kỹ thuật tạo lệnh (prompting) để tạo sinh ngữ liệu giảng dạy tiếng Anh, đòi hỏi một quá trình thiết kế cẩn thận chứ không phải là một quá trình tự động. Chất lượng của văn bản do LLMs tạo ra phụ thuộc trực tiếp vào chất lượng của chỉ dẫn đầu vào, hay còn gọi là prompt. Các nghiên cứu đã chỉ ra rằng, việc thiếu một phương pháp luận khoa học trong việc thiết kế prompt sẽ dẫn đến nhiều hạn chế đáng kể. Khi không có các chỉ dẫn được cấu trúc tốt, mô hình có thể gặp khó khăn trong việc nắm bắt ngữ cảnh, dẫn đến các phản hồi không chính xác hoặc không đáp ứng được các yêu cầu phức tạp về mặt ngôn ngữ và sư phạm (Luo và các tác giả khác, 2019). Thực tế này đặt ra một yêu cầu cấp thiết về việc trang bị cho giáo viên một khung làm việc có tính hệ thống, để giao tiếp và điều khiển LLMs một cách hiệu quả, biến chúng thành một công cụ sáng tạo hữu ích.

[1] Học viên cao học, Trường Đại học Khoa học Xã hội và Nhân văn - Đại học Quốc gia TP. Hồ Chí Minh, *Email: milo.eei@gmail.com*
[2] KS. NCS., Trường Đại học Công nghiệp TP. Hồ Chí Minh, E*mail: lhsy.workspace@gmail.com*

Các lý thuyết ngôn ngữ học truyền thống, dù có giá trị trong việc phân tích cấu trúc, thường có xu hướng tập trung vào ngữ pháp hình thức, tức là các quy tắc nội tại và tính đúng đắn của câu mà ít chú trọng đến chức năng và mục đích giao tiếp của ngôn ngữ trong các tình huống thực tế (Ezekulie & Julde-Adura, 2017). Điều này tạo ra một khoảng trống lớn khi chúng ta cần xây dựng các chỉ dẫn (prompts). Nếu chỉ dựa vào các câu lệnh đơn thuần, mô hình sẽ khó lòng nắm bắt được ngữ cảnh xã hội, ý định giao tiếp phức tạp (OpenAI, 2023) và các sắc thái sư phạm cần thiết để tạo ra ngữ liệu không chỉ đúng ngữ pháp mà còn phù hợp, tự nhiên và hiệu quả cho việc giảng dạy tiếng Anh. Để giải quyết vấn đề này, hướng tiếp cận của báo cáo là tìm đến một lý thuyết ngôn ngữ học đủ mạnh mẽ để mô tả và hệ thống hóa các đặc trưng của ngôn ngữ trong sử dụng, từ đó làm cơ sở cho việc xây dựng các tham số điều khiển trong một prompt. Ngôn ngữ học chức năng hệ thống của Halliday (Halliday, 1985) chính là lý thuyết đáp ứng được yêu cầu đó. Với việc xem ngôn ngữ như một hệ thống các lựa chọn để tạo nghĩa trong bối cảnh xã hội, SFL cung cấp một bộ công cụ phân tích tinh vi về mối quan hệ giữa bối cảnh, hay chính là thông qua các biến số liên quan đến Ngữ vực (Register) và Thể loại (Genre) và các lựa chọn từ vựng - ngữ pháp cụ thể.

Do đó, báo cáo này trình bày việc xây dựng một Khung chỉ dẫn đầu vào dựa trên lý thuyết Ngôn ngữ học chức năng - hệ thống (được gọi là SFL-based Prompt).

Đây là hướng tiếp cận nhằm biến các khái niệm phân tích của SFL thành các tham số có thể vận hành được trong một cấu trúc prompt. Mục tiêu của khung SFL-based Prompt là hỗ trợ cho giáo viên, giúp họ có khả năng chủ động điều khiển và tạo ra các ngữ liệu giảng dạy tiếng Anh (văn bản, hội thoại, bài tập,...) có các đặc trưng ngôn ngữ được kiểm soát chặt chẽ, phù hợp với mục tiêu sư phạm và tương ứng chính xác với từng cấp độ năng lực từ A1 đến C2 theo chuẩn CEFR. Báo cáo sẽ lần lượt trình bày cơ sở lý thuyết, cấu trúc chi tiết của khung SFL-based Prompt, các ví dụ minh họa và thảo luận về tiềm năng ứng dụng của phương pháp này trong thực tiễn giảng dạy ngoại ngữ hiện đại.

2. Nội dung nghiên cứu

Nền tảng của khung SF-Prompt được xây dựng dựa trên sự hội tụ của ba lĩnh vực chính: Ngôn ngữ học chức năng - hệ thống, Khung chỉ dẫn đầu vào cho LLMs, và Khung năng lực ngôn ngữ CEFR. Phần này sẽ tập trung làm rõ các khái niệm nền tảng của SFL, qua đó minh chứng tính thỏa đáng của lý thuyết này trong việc cung cấp một bộ công cụ phân tích để xây dựng các prompt hiệu quả.

2.1. Ngôn ngữ học chức năng - hệ thống (SFL)

Ngôn ngữ học chức năng - hệ thống là một lý thuyết ngôn ngữ học do Halliday phát triển, với luận điểm trung tâm rằng ngôn ngữ là một hệ thống ký hiệu xã hội (social semiotic) và cấu trúc của nó được định hình bởi các chức năng mà nó thực hiện trong các bối cảnh xã hội. Khác với các trường phái ngôn ngữ học hình thức vốn tập trung vào các quy tắc cú pháp nội tại, SFL xem ngôn ngữ là chức năng (language is functional), là một nguồn tài nguyên để con người lựa chọn và tạo ra nghĩa nhằm đạt được các mục đích giao tiếp cụ thể (Hoàng Văn Vân, 2021).

Chính vì vậy, SFL không chỉ là một lý thuyết về "ngôn ngữ như một hệ thống trừu tượng" mà là một lý thuyết về "ngôn ngữ trong sử dụng", hay ngôn ngữ như một "tiềm năng tạo nghĩa" (meaning potential) (Martin, Doran, & Figueredo, 2019).

Quan điểm này đặc biệt phù hợp với việc xây dựng prompt vì bản chất của việc thiết kế prompt chính là quá trình ra các lựa chọn có chủ đích để chỉ dẫn một hệ thống

(LLMs) tạo ra một sản phẩm ngôn ngữ (văn bản) mang ý nghĩa mong muốn. Để làm được điều đó, cần có một lý thuyết có khả năng hệ thống hóa các lựa chọn ngôn ngữ và mối quan hệ của chúng với bối cảnh. Chính vì vậy, SFL cung cấp chính xác bộ công cụ lý thuyết đó.

Một trong những đóng góp nền tảng của SFL là việc mô tả mối quan hệ có hệ thống giữa ngôn ngữ và ngữ cảnh. Lý thuyết này cho rằng các lựa chọn từ vựng - ngữ pháp trong một văn bản không phải là ngẫu nhiên mà được dự báo bởi bối cảnh tình huống (context of situation). SFL mô hình hóa bối cảnh tình huống này thông qua khái niệm ngữ vực (register), bao gồm ba biến số: trường (field), thức (mode) và không khí (tenor) (Almurashi, 2016, tr. 72). Các biến số ngữ cảnh này được ánh xạ trong ngôn ngữ thông qua ba siêu chức năng (metafunctions) của Halliday, bao gồm:

1. Siêu chức năng ý niệm (ideational metafunction) là chức năng ngôn ngữ dùng để biểu đạt kinh nghiệm về thế giới, bao gồm các sự kiện, hành động, trạng thái, các đối tượng tham gia và hoàn cảnh liên quan.

2. Siêu chức năng liên nhân (interpersonal metafunction) là chức năng ngôn ngữ dùng để thiết lập và duy trì các mối quan hệ xã hội cho phép người sử dụng ngôn ngữ đảm nhận các vai trò giao tiếp (như hỏi, trả lời, đề nghị), thể hiện thái độ, quan điểm và sự đánh giá.

3. Siêu chức năng văn bản (textual metafunction) là chức năng ngôn ngữ dùng để tổ chức các ý nghĩa ý niệm và liên nhân thành một thông điệp mạch lạc, có cấu trúc và phù hợp với ngữ cảnh. Nó quản lý dòng chảy thông tin trong văn bản.

Tóm lại, SFL cung cấp, mô tả cách thức các yếu tố ngữ vực được ánh xạ vào các loại ý nghĩa và được hiện thực hóa bằng các cấu trúc ngữ pháp và từ vựng cụ thể. Chính sự ánh xạ có hệ thống này đã làm cho SFL trở thành một nền tảng lý thuyết vô cùng thỏa đáng và phù hợp cho việc xây dựng một khung chỉ dẫn khoa học, cho phép người dùng kiểm soát đầu ra của các mô hình ngôn ngữ lớn ở nhiều cấp độ: từ nội dung, văn phong cho đến cấu trúc văn bản.

2.2. Các Mô hình Ngôn ngữ Lớn (Large Language Models - LLMs) và Khung chỉ dẫn đầu vào (prompt)

Các Mô hình Ngôn ngữ Lớn (LLMs), như các phiên bản GPT của OpenAI, là các hệ thống trí tuệ nhân tạo được xây dựng trên kiến trúc mạng nơ-ron (transformers) (Vaswani và các tác giả khác, 2017). Yếu tố cốt lõi làm nên sức mạnh của kiến trúc này là cơ chế chú ý (attention mechanism), cho phép mô hình xử lý và tính toán xác suất độ quan trọng của tất cả các từ trong một chuỗi đầu vào, giúp nắm bắt được ngữ cảnh phức tạp và các mối quan hệ ngữ

nghĩa ở khoảng cách xa (context dependencies). Được huấn luyện trên các kho dữ liệu văn bản khổng lồ, các mô hình ngôn ngữ lớn có khả năng tạo sinh ngôn ngữ tự nhiên một cách mạch lạc và linh hoạt, phù hợp với từng ngữ cảnh đưa ra.. Khả năng này cho thấy mô hình có thể thực hiện nhiều tác vụ mới mà không cần qua quá trình tinh chỉnh chuyên biệt chỉ bằng cách nhận các chỉ dẫn đầu vào. Năng lực này mở ra tiềm năng ứng dụng to lớn trong giáo dục ngôn ngữ. Ví dụ, LLMs có thể được sử dụng để tạo ra các bài tập từ vựng, giải thích các hiện tượng ngữ pháp, hoặc hoạt động như một đối tác hội thoại để người học thực hành, v.v..

Nếu các mô hình ngôn ngữ lớn được ví như một động cơ mạnh mẽ, thì prompt chính là hệ thống điều khiển để khai thác sức mạnh đó. Tầm quan trọng của khung chỉ dẫn đầu vào bắt nguồn từ bản chất hoạt động của các mô hình ngôn ngữ lớn. Các mô hình này không diễn giải hay nhận thức ý định của con người theo cách con người hiểu và giao tiếp lẫn nhau, mà thay vào đó, chúng tạo phản hồi dựa trên việc nhận diện và mô phỏng các mẫu hình thống kê trong dữ liệu huấn luyện, cùng với các chỉ dẫn được cung cấp trong quá trình tương tác (Liskavets và các tác giả khác, 2025). Một prompt mơ hồ hoặc thiếu ngữ cảnh có thể dẫn đến phản hồi mang tính khái quát, kém chính xác hoặc không đáp ứng được mục tiêu thông tin mong muốn. Ngược lại, một prompt được thiết kế chặt chẽ, với đầy đủ thông tin và ràng buộc cần thiết, sẽ định hướng mô hình tạo ra kết quả chính xác, phù hợp và có giá trị hơn đối với mục đích của người dùng.

Nhiều nghiên cứu đã chỉ ra rằng cấu trúc của prompt có ảnh hưởng đáng kể đến chất lượng đầu ra của các mô hình ngôn ngữ lớn (LLMs). Sự khác biệt trong cách diễn đạt, mức độ chi tiết hoặc ngữ cảnh được cung cấp có thể dẫn đến biến thiên rõ rệt trong tính chính xác, độ mạch lạc và mức độ phù hợp của phản hồi. Ba kỹ thuật

prompting nền tảng thường được áp dụng nhằm tối ưu hóa hiệu suất của LLMs bao gồm: zero-shot prompting, few-shot prompting, và chain-of-thought (CoT) prompting. Các đặc điểm và ví dụ minh họa cho từng kỹ thuật được trình bày trong Bảng 1 dưới đây.

Mặc dù các kỹ thuật trên rất hữu ích, việc áp dụng chúng một cách riêng lẻ vẫn có thể chưa đủ để đáp ứng các yêu cầu phức tạp của môi trường giáo dục. Một prompt hiệu quả trong giảng dạy ngôn ngữ cần phải tích hợp các yếu tố này vào một cấu trúc tổng thể, có khả năng kiểm soát đồng thời nhiều biến số ngôn ngữ và sư phạm. Các công thức xây dựng prompt phổ biến như R-T-F (Role, Task, Format) đã cung cấp một bộ khung cơ bản (Liu & Ni, 2024). Tuy nhiên, để thực sự hiệu quả, các thành phần này cần được làm sâu sắc hơn bằng một lý thuyết ngôn ngữ có hệ thống. Nghiên cứu này đề xuất rằng SFL chính là nền tảng lý thuyết cần thiết để cụ thể hóa và khoa học hóa các thành phần của một prompt, từ đó tạo ra một công cụ mạnh mẽ và đáng tin cậy cho giáo viên.

CEFR là một tiêu chuẩn quốc tế được công nhận rộng rãi để mô tả mức độ thành thạo ngôn ngữ. Cách tiếp cận dựa trên hành động và bối cảnh giao tiếp của CEFR hoàn toàn tương thích với quan điểm chức năng của SFL. Chính sự tương đồng này, chúng tôi quyết định sử dụng CEFR như một bộ tiêu chí để hiệu chỉnh các tham số SFL trong khung prompt. Bằng cách ánh xạ các mô tả năng lực của CEFR vào các biến số Ngữ vực và các lựa chọn về từ vựng

- ngữ pháp, chúng tôi có thể chỉ dẫn LLMs tạo ra các ngữ liệu có độ phức tạp về nội dung và cấu trúc phù hợp với từng cấp độ của người học.

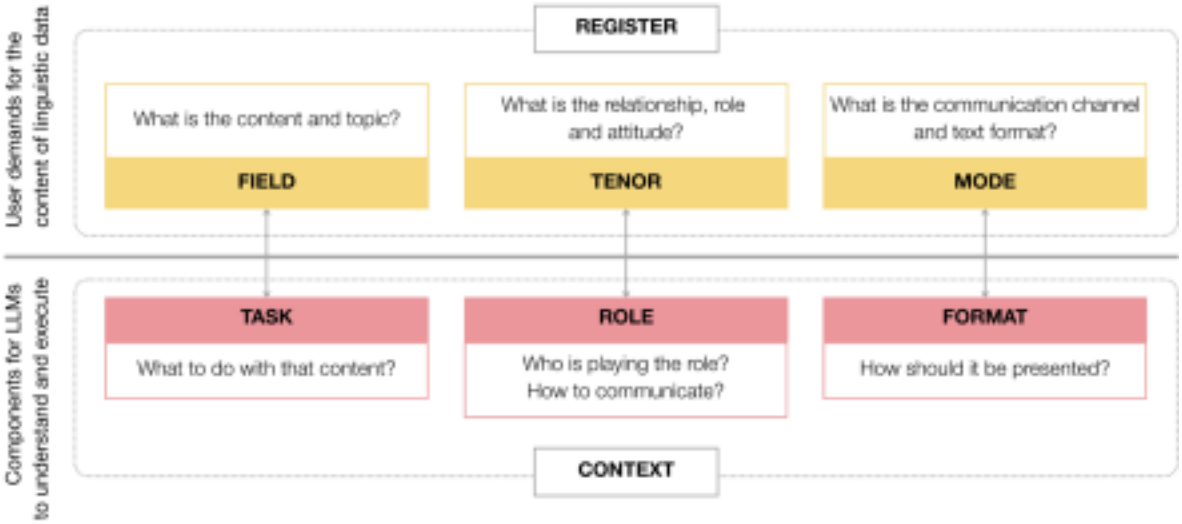
3. Đề xuất khung chỉ dẫn đầu vào dựa trên Ngôn ngữ học Chức năng Hệ thống (SFL)

3.1. Cấu trúc và nguyên lý vận hành

Trong báo cáo này, chúng tôi đề xuất khung chỉ dẫn SFL-based Prompt nhằm tích hợp các khái niệm cốt lõi của Ngôn ngữ học Chức năng Hệ thống vào quy trình thiết kế prompt cho các mô hình Ngôn ngữ Lớn theo hình 1. Thông qua đó, người sử

dụng - đặc biệt là giáo viên - có thể kiểm soát có hệ thống các đặc trưng ngôn ngữ của văn bản đầu ra theo mục tiêu sư phạm và cấp độ năng lực ngôn ngữ.

Cấu trúc khung gồm bốn thành phần chính (1) vai trò và bối cảnh (role & context) - xác định vai trò giao tiếp của mô hình (ví dụ: giáo viên tiếng Anh, người hướng dẫn học thuật), đối tượng người học, và mục tiêu sư phạm cụ thể; (2) nhiệm vụ cốt lõi (core task) - xác định hành động chính (tạo, viết, soạn thảo) và loại sản phẩm mong đợi (bài đọc, hội thoại, bài tập…); (3) các tham số điều khiển chức năng (functional control parameters) - được thiết kế tương ứng với ba siêu chức năng của Halliday (ý niệm, liên nhân, văn bản), cho phép điều chỉnh trường (field), thức và không khí (mood, tenor), và cấu trúc đề - thuyết (theme - rheme); (4) định dạng đầu ra (output format) - quy định hình thức trình bày và yêu cầu sư phạm của sản phẩm ngôn ngữ.



Hình 1. Mối quan hệ giữa các yếu tố ngữ vực với các thành tố trong khung chỉ dẫn.

Khung phép chuyển hóa các khái niệm trừu tượng của SFL thành các tham số vận hành cụ thể trong quá trình chỉ dẫn LLMs, giúp người thiết kế bài giảng tạo ra các văn bản giảng dạy có thể kiểm soát được cả nội dung, ngữ điệu, và độ phức tạp ngôn ngữ.

3.2. Triển khai thực nghiệm và đánh giá

Nhằm kiểm chứng tính khả thi và hiệu quả của khung SFL-based Prompt do nghiên cứu đề xuất, nhóm đã tiến hành một thử nghiệm tính toán được lập trình bằng ngôn ngữ Python và tích hợp với Ollama API để gọi và xử lý các mô hình ngôn ngữ lớn (LLMs). Mục tiêu của thử nghiệm là so sánh năng lực tạo sinh văn bản giữa ba phương thức (1) zero-shot prompting, (2) few-shot prompting, và (3) SFL-based prompting (theo khung SFL-based Prompt được đề xuất trong nghiên cứu).

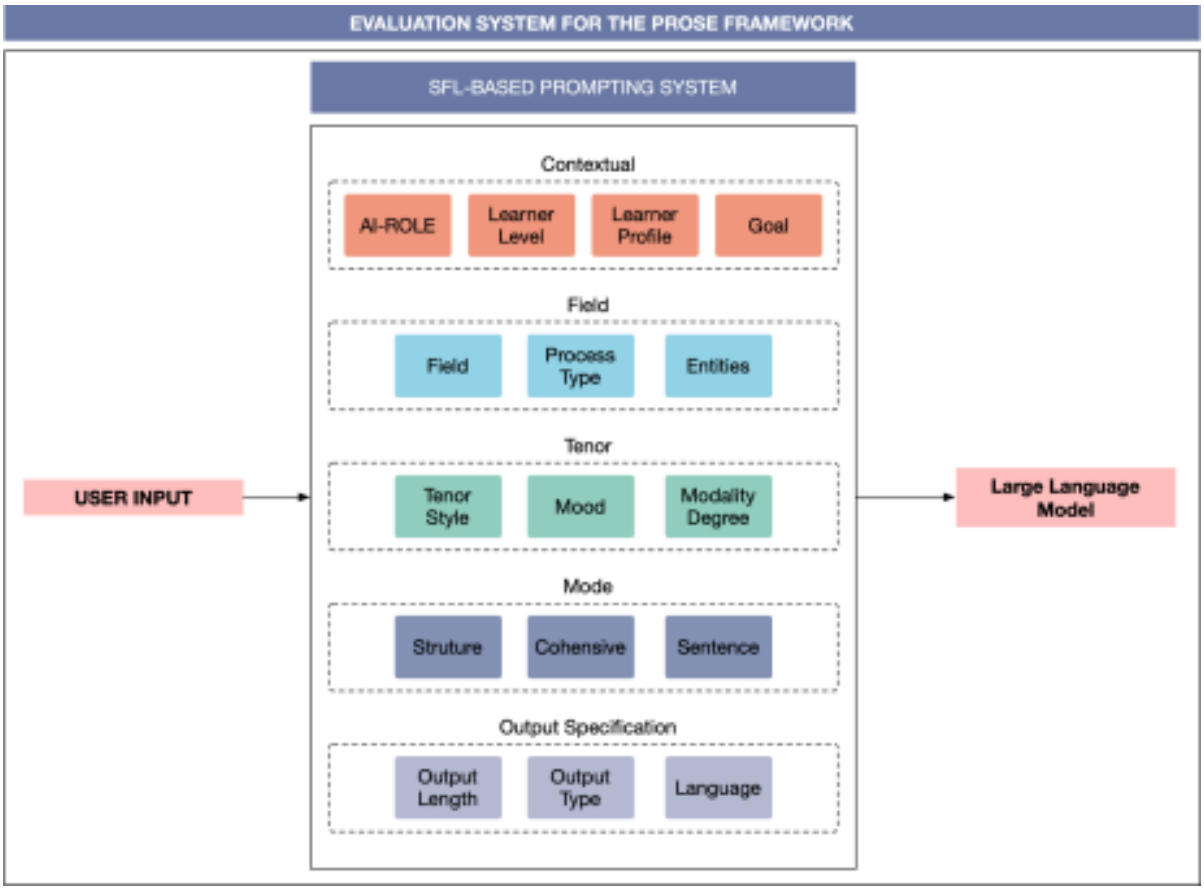
3.3. Phương thức triển khai

Trong quy trình này, mô hình ngôn ngữ Llama 3.2, một trong những kiến trúc ngôn ngữ mạnh mẽ được công bố từ Facebook (Dubey và các tác giả khác, 2024), sử dụng làm nền tảng sinh đầu ra, với cùng một chủ đề và độ dài yêu cầu để đảm bảo tính tương đương giữa các điều kiện thử nghiệm. Chúng tôi tiến hành xây dựng hệ thống và mã hóa các thành phần cốt lõi của khung SFL-based Prompt, bao gồm: vai trò của AI, đặc điểm người học, mục tiêu sư phạm, trường (field), loại quá trình (process type), không khí (tenor), cấu trúc đề–thuyết (theme–rheme), phương tiện liên kết (cohesive devices) và định dạng đầu ra (output format). Các tham số này được tổng hợp thành một prompt hoàn chỉnh trước khi truyền vào mô hình ngôn ngữ lớn để tạo sinh các tài liệu.

Quy trình triển khai được minh hoạ ở hình 2.

Kỹ thuật	Mô tả	Ví dụ	Ưu điểm	Hạn chế	Ứng dụng
Zero-shot prompting	Mô hình được yêu cầu thực hiện nhiệm vụ chỉ dựa trên mô tả trong *prompt*, không có ví dụ mẫu.	Translate the following English text to French: "Hello, how are you?"	Đơn giản, không yêu cầu dữ liệu mẫu	Kết quả không nhất quán hoặc thiếu chính xác, dễ bị .	Dịch tự động, hỏi đáp tổng quát, phân loại cơbản.
Few-shot prompting	Cung cấp một hoặc một vài cặp đầu vào – đầu ra để minh họa cách thực hiện nhiệm vụ, giúp mô hình nhận diện quyluật mong muốn.	Correct the following sentences. Example: "I has two cat." → "I have two cats." Now correct: "She go to school yesterday."	Cải thiện độ chính xác và tính nhất quán.	Cần ví dụ chất lượng cao; giới hạn bởi độ dài ngữ cảnh(*context window)*	Tạo bài tập mẫu trong giáo dục, chuẩn hóa định dạng đầu ra, xử lý văn bản theo mẫu.
Chain-of-Thought (CoT) prompting	Hướng dẫn mô hình trình bày quá trình suy luận từng bước thay vì đưa ra câu trả lời cuối cùng ngay lập tức.	Explain why the present perfect tense is used in "She has lived here for ten years." Let's break it down step by step.	Nâng cao khả năng lập luận logic và giải quyết vấn đề phức tạp; phản hồi minh bạch hơn.	Đầu ra dài, có thể chứa suy luận không chính xác; tốn tài nguyên xử lý.	Phân tích ngữ pháp, lậpluận logic, giải thích quy tắc hoặc hiện tượng ngôn ngữ.

Bảng 1. Tóm tắt các kỹ thuật prompting nền tảng trong mô hình ngôn ngữ lớn



Hình 2: Hình minh họa quy trình tạo sinh văn bản theo khung SFL-based Prompt, gồm các giai đoạn:nhập các tham số đầu vào thể hiện ngữ cảnh sư phạm và đặc trưng ngôn ngữ theo ba siêu chức năng của Ngữ pháp Chức năng Hệ thống (Field – Tenor – Mode); (2) Hệ thống tổng hợp tham số này thành một prompt hoàn chỉnh; (3) prompt được truyền vào mô hình ngôn ngữ thông qua mô hình ngôn ngữ để sinh văn bản.

3.3.1. Các chỉ số phân tích định lượng (Quantitative Analysis Metrics)

Cụ thể, nhóm chỉ số đầu tiên tập trung vào các đặc trưng cơ bản của văn bản (basic metrics), bao gồm số lượng từ, số câu, độ dài câu trung bình và độ đa dạng từ vựng. Những chỉ số này phản ánh quy mô, mức độ phát triển ý tưởng và khả năng kiểm soát ngôn ngữ của mô hình. Tiếp đến, phân tích về phương tiện liên kết (cohesive devices) được thực hiện nhằm xác định số lượng và loại hình liên từ logic

được sử dụng (như firstly, moreover, as a result, v.v), qua đó đo lường mức độ mạch lạc và kết nối thông tin trong văn bản.

Bên cạnh đó, nghiên cứu cũng xem xét độ phức tạp cú pháp thông qua tần suất xuất hiện của các mệnh đề phụ và phương tiện liên kết từ vựng (như because, although, which, when, v.v), phản ánh khả năng tổ chức cú pháp và sự linh hoạt trong cấu trúc câu.

Để đánh giá mức độ phù hợp chủ đề, nhóm chỉ số về từ vựng, trường từ vựng được sử dụng, nhằm đếm và liệt kê các tập hợp từ vựng. Ví dụ với chủ đề liên quan "bảo vệ môi trường và tái chế

(environment, recycle)" thì các từ liên quan sẽ là recycle, pollution, sustainability, green, renewable, v.v. Ngoài ra, để đảm bảo tính tương thích với trình độ người học, nghiên cứu còn phân tích đặc trưng ngôn ngữ theo các tiêu chí thuộc các cấp độ khác nhau (CEFR), nghiên cứu chúng tôi đề xuất triển khai ở cấp độ B1. Việc chọn cấp độ B1 (CEFR B1) làm tiêu chí để thử nghiệm nhằm đảm bảo sự cân bằng, phù hợp với trình độ phổ biến và năng lực thụ đắc ngôn ngữ trong giáo dục phổ thông Việt Nam, đồng thời cho phép đánh giá rõ ràng hiệu quả điều tiết của SFL-based Prompt đối với khả năng tạo sinh văn bản tự nhiên và mạch lạc , bao gồm tần suất sử dụng động từ tình thái (can, should, might) và các yếu tố nội động/ ngoại động (finite/infinite), nhằm xác định mức độ phù hợp của ngôn ngữ được sinh ra so với yêu cầu của trình độ trung cấp.

Trên cơ sở các đặc trưng nêu trên, mỗi đầu ra được chấm điểm tổng hợp theo các tiêu chí: độ phù hợp về độ dài (chúng tôi tiến hành trên các kết quả đoạn văn từ 180 ± 15 từ), mức độ liên kết logic (dựa trên số lượng phương tiện liên kết), độ phức tạp cú pháp hợp lý (tỷ lệ mệnh đề phụ trên tổng số câu), mức độ gắn kết với chủ đề (thông qua tần suất từ vựng chuyên biệt), và độ đa dạng từ vựng. Tất cả các chỉ số được chuẩn hóa về thang điểm 100, cho phép so sánh định lượng giữa ba mô hình một cách thống nhất và khách quan.

3.3.2. Đánh giá và phân tích.

Trong nghiên cứu này, chúng tôi tiến hành một thực nghiệm đối chiếu nhằm đánh giá hiệu quả của khung SFL-based Prompt trong việc hướng dẫn mô hình ngôn ngữ tạo sinh ngữ liệu phục vụ giảng dạy tiếng Anh. Chủ đề được lựa chọn là

"recycling and environmental protection", yêu cầu văn bản ở trình độ CEFR B1, độ dài tương đương, và mục đích sử dụng trong hoạt động đọc hiểu.

Ba phương thức tạo sinh được so sánh trong cùng điều kiện ngữ cảnh và yêu cầu đầu ra tương đồng. Các văn bản thu được sau đó được phân tích dựa trên các chỉ số ngôn ngữ học định lượng, bao gồm: độ dài văn bản, đa dạng từ vựng, mức độ sử dụng phương tiện liên kết, độ phức tạp cú pháp, từ vựng chuyên đề và các chỉ báo phù hợp với năng lực CEFR B1.

Về sản phẩm tạo sinh từ mô hình ngôn ngữ lớn, chúng tôi tổng hợp tại vào bảng 2

Kỹ thuật	**Kết quả đầu ra**

Zero-shot prompting	As the world grapples with the challenges of climate change, conservation, and sustainability, the importance of environmental protection cannot be overstated. One of the most effective ways to protect our planet is through recycling. Recycling helps to conserve natural resources, reduces landfill waste, and decreases greenhouse gas emissions. Did you know that recycling just one plastic bottle can save enough energy to power a computer for 25 minutes? Or that recycling aluminum cans uses 95% less energy than producing new aluminum from raw materials? By recycling paper, glass, and plastic, we can significantly reduce the amount of waste sent to landfills and help preserve our natural resources. However, recycling requires individual action. Making small changes in our daily lives, such as reducing single-use plastics, reusing bags and containers, and recycling correctly, can add up to make a big impact. As consumers, we have the power to demand more sustainable products from companies and governments must prioritize environmental protection through policies and regulations. Together, we can create a cleaner, greener future for ourselves and future generations. By taking small steps towards sustainability, we can inspire positive change and make a real difference in protecting our planet's precious resources.
Few-shot prompting	Our planet faces numerous challenges due to human activities. One of the most significant issues is waste management. Every day, we generate vast amounts of trash, from plastic bottles and food packaging to paper and cardboard. If left unrecycled, these items can harm our environment and contaminate our air, water, and soil. However, recycling offers a solution. By converting waste into new products, we can reduce the need for raw materials, decrease pollution, and conserve natural resources. Recycling also saves energy, decreases greenhouse gas emissions, and creates jobs in the recycling industry. To make a positive impact on the environment, it's essential to recycle properly, participate in community clean-up initiatives, and support companies that prioritize sustainability. By working together, we can mitigate the effects of environmental degradation and create a healthier planet for future generations. It starts with making conscious choices about our daily habits and adopting

	environmentally-friendly practices. By doing so, we can ensure a cleaner, greener tomorrow.
SFL-based prompting	As the world's population grows, so does our environmental footprint. Our daily actions, from throwing away plastic bottles to buying products wrapped in excess packaging, contribute to pollution and waste management issues. The consequences of these actions are severe: climate change, deforestation, and loss of biodiversity. Firstly, recycling is a vital step towards reducing waste. By reusing materials like paper, cardboard, and glass, we can significantly decrease the amount of trash sent to landfills. In addition, many countries have implemented recycling programs in public spaces, making it easier for people to participate. Secondly, individuals can make a difference by adopting environmentally-friendly habits. As a result, consumers are increasingly seeking products with minimal packaging or made from sustainable materials. For instance, companies like Patagonia and REI offer eco-friendly clothing lines that appeal to environmentally-conscious customers. By supporting these brands and reducing our consumption of single-use plastics, we can create a more sustainable future for the planet.

Bảng 2. Sản phẩm ngữ liệu đầu ra được tạo sinh từ các khung chỉ dẫn đầu vào.

Kết quả phân tích định lượng được thể hiện tại bảng 3.

Chỉ số	**Zero-Shot**	**Few-Shot**	**SFL-Based**
Số từ	196	179	168
Số câu	11	12	10
Độ dài câu trung bình	17.82	14.92	16.80
Đa dạng từ vựng (TTR)	0.679	0.682	**0.690**
Phương tiện liên kết	4	3	**7**

Tỷ lệ cấu trúc phức hợp	0.18	0.17	**0.40**
Từ vựng chuyên đề	14	13	14
Modal / Passive	2 / 3	1 / 1	**2 / 5**
Điểm tổng hợp	73.78	70.24	**76.60**

Bảng 3: Tổng hợp phân tích các chỉ số định lượng cho kết quả tạo sinh từ 3 kĩ thuật: 1) Zero-shot prompting, (2) Few-shot prompting, và (3) SFL-based prompting

Kết quả so sánh định lượng cho thấy phương thức SFL-based tạo sinh văn bản có chất lượng tốt hơn so với Zero-shot và Few-shot. Cụ thể, văn bản SFL-based đạt 7 phương tiện liên kết, cao hơn rõ rệt so với 4 ở Zero-shot và 3 ở Few-shot, cho thấy mức độ kiểm soát mạch diễn ngôn tốt hơn. Về độ phức tạp cú pháp, tỷ lệ cấu trúc câu phức hợp của SFL-based đạt 0.40, gấp đôi so với Zero-shot (0.18) và Few-shot (0.17), phản ánh khả năng linh hoạt trong tổ chức các mối quan hệ mệnh đề. Mặc dù số từ và số câu của SFL-based thấp hơn (168 từ và 10 câu) so với Zero-shot (196 từ, 11 câu) và Few-shot (179 từ, 12 câu), nhưng đa dạng từ vựng (TTR) lại cao hơn (0.690 so với 0.679 và 0.682), chứng tỏ văn bản ngắn hơn nhưng giàu tính diễn đạt hơn. Đồng thời, sự xuất hiện cân đối giữa động từ tình thái và thể bị động cho thấy giọng văn bán trang trọng phù hợp với đặc điểm văn bản ở trình độ CEFR B1. Tổng hợp các chỉ số, phương thức SFL-based đạt 76.60 điểm, cao hơn Zero-shot (73.78) và Few-shot (70.24), khẳng định rằng việc tích hợp các tham số của Ngôn ngữ học Chức năng Hệ thống vào prompt giúp kiểm soát tốt hơn đặc điểm ngôn ngữ của văn bản và nâng cao tính phù hợp của ngữ liệu đối với mục tiêu sư phạm.

Phân tích đặc điểm các ngữ liệu được tạo sinh cho thấy sự khác biệt rõ rệt trong cách ba phương thức triển khai chủ đề. Văn bản từ kỹ thuật Zero-shot prompting có xu hướng trình bày thông tin ở mức khái quát, sử dụng nhiều mệnh đề khai báo nhấn mạnh nội dung nhưng thiếu định hướng tổ chức và liên kết, khiến người học khó theo dõi tiến trình ý nghĩa. Văn bản từ kỹ thuật Few-shot Prompting ổn định hơn, nhưng cấu trúc phát triển ý nhìn chung đơn giản và bị chi phối mạnh bởi phong cách của ví dụ mẫu, dẫn đến sự thiếu vắng các dấu hiệu liên kết tường minh giữa các đoạn và mệnh đề. Ngược lại, văn bản từ kỹ thuật chúng tôi đề xuất - SFL-based Prompt thể hiện mạch diễn ngôn rõ ràng nhờ cấu trúc được định hướng từ khái quát đến minh họa và hệ quả; các liên từ như firstly, in addition, for instance, as a result, v.v giúp xác lập quan hệ logic giữa các ý. Đồng thời, việc sử dụng có chủ đích các quá trình vật chất và thể bị động tạo nên giọng điệu bán trang trọng phù hợp với mục tiêu học thuật ở cấp độ B1. Những quan sát này cho thấy SFL-based Prompt không chỉ nâng cao tính chính xác ngôn ngữ mà còn tối ưu hóa trải nghiệm đọc và khả năng tiếp nhận kiến thức của người học.

4. Thảo luận

Kết quả thực nghiệm cho thấy việc tích hợp các tham số ngôn ngữ theo Ngôn ngữ học Chức năng Hệ thống (SFL) vào thiết kế prompt có ảnh hưởng rõ rệt đến chất lượng văn bản tạo sinh. Sự vượt trội của phương thức SFL-based Prompt chủ yếu đến từ khả năng điều tiết có chủ đích các lựa chọn ngôn ngữ thay vì chỉ dựa vào thống kê nội tại của mô hình. Khi trường (field), thức (mode) và không khí (tenor) được xác lập rõ ràng ngay từ đầu, mô hình không chỉ tạo ra văn bản đúng chủ đề, mà còn phát triển nội dung theo mạch diễn ngôn có định hướng, thể hiện rõ các bước từ giới thiệu khái niệm đến triển khai lập luận và minh họa bằng ví dụ. Điều này cho thấy rằng mô hình ngôn ngữ lớn không hoàn toàn "thông minh" theo nghĩa tự hình thành ý đồ giao tiếp, mà phản hồi của nó phụ thuộc mạnh vào chuẩn hoá bối cảnh giao tiếp do người dùng thiết kế. Việc thiếu ràng buộc ngữ cảnh ở Zero-shot dẫn đến văn bản có xu hướng khái quát và liệt kê thông tin, trong khi Few-shot dễ bị ràng buộc phong cách từ ví dụ mẫu, dẫn đến mức độ đơn giản hoá quá mức về cú pháp và diễn ngôn. Ngược lại, khung SFL-based không yêu cầu mô hình "bắt chước" hay "đoán chiến lược diễn đạt" mà dẫn hướng trực tiếp các lựa chọn ngữ nghĩa – ngữ pháp, giúp đầu ra ổn định và phù hợp hơn với đặc trưng văn bản theo từng trình độ học thuật.

Từ góc độ sư phạm, kết quả này nhấn mạnh rằng prompt không chỉ là một câu lệnh kỹ thuật mà là một công cụ thiết kế học liệu. Người thiết kế cần nắm được mối quan hệ giữa mục tiêu học tập và lựa chọn ngôn ngữ cụ thể trong văn bản. Điều này có nghĩa là việc sử dụng LLMs trong giảng dạy không thể tách rời kiến thức ngôn ngữ học nền tảng, đặc biệt là khả năng xác định đặc trưng ngữ vực và quản lý độ phức tạp diễn ngôn phù hợp với người học. Do đó, nghiên cứu này khẳng định rằng việc áp dụng SFL vào prompting không chỉ mở rộng khả năng kiểm soát đầu ra của mô hình ngôn ngữ, mà còn góp phần tái định vị vai trò của giáo viên: từ người tiêu thụ nội dung sang người thiết kế và điều phối ngôn ngữ trong môi trường học tập có hỗ trợ từ Trí tuệ Nhân tạo (AI). Đây là bước chuyển quan trọng trong việc tích hợp LLMs một cách có ý thức và có cơ sở lý luận vào thực tiễn giáo dục ngoại ngữ.

Tài liệu tham khảo

Almurashi, W. A. (2016). An Introduction to Halliday's Systemic Functional Linguistics. *Journal for the Study of English Linguistics, 4*. https://doi.org/10.5296/jsel.v4i1.9423

Chen, B. (2023). *Unleashing the potential of prompt engineering for large language models*. arXiv. https://doi.org/10.48550/arXiv.2310.14735

Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., ... & Ganapathy, R. (2024). *The llama 3 herd of models*. arXiv e-prints, arXiv-2407.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold.

Hoàng, V. V. (2021). "Metafunctions of Language" in Systemic Functional Linguistics: A Framework for the Interpretation of Meaning of Text in Social Context. *VNU Journal of Foreign Studies, 37*(4). https://doi.org/10.25073/2525-2445/vnufs.4750

Liu, X., & Ni, Z. (2024). *Role-playing prompt framework: Generation and evaluation*. arXiv. https://doi.org/10.48550/arXiv.2406.00627

Luo, L., Ao, X., Song, Y., Li, J., Yang, X., He, Q., & Yu, D. (2019). Unsupervised neural aspect extraction with sememes. In *Proceedings of the International Joint Conference on Artificial Intelligence*.

Martin, J. R., Doran, Y. J., & Figueredo, G. (Eds.). (2019). *Systemic functional language description: Making meaning matter*. Routledge.

OpenAI. (2023). *GPT-4 technical report*. arXiv. https://doi.org/10.48550/arXiv.2303.08774

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). *Chain-of-thought prompting elicits reasoning in large language models*. arXiv. https://doi.org/10.48550/arXiv.2201.11903

XÂY DỰNG ĐỒ THỊ TRI THỨC CHO Y HỌC CỔ TRUYỀN VIỆT NAM TỪ TÁC PHẨM NAM DƯỢC THẦN HIỆU - TUỆ TĨNH

Nguyễn Thị Xuân Hiền[1], Phạm Trần Nhật Minh[2], Nguyễn Thành Giang[3], Nguyễn Hồng Bửu Long[4]

Tóm Tắt

Trong khi đã có nhiều đồ thị tri thức (Knowledge Graph - KG) được xây dựng cho Y học Cổ truyền Trung Hoa (Traditional Chinese Medicine - TCM), kho tàng tri thức của Y học Cổ truyền Việt Nam (Traditional Vietnamese Medicine - TVM) phần lớn vẫn tồn tại dưới dạng văn bản phi cấu trúc. Nghiên cứu này trình bày nỗ lực tiên phong nhằm thu hẹp khoảng cách đó, thông qua việc xây dựng đồ thị tri thức đầu tiên dựa trên tác phẩm "Nam Dược Thần Hiệu" của danh y Tuệ Tĩnh (thế kỷ XIV). Chúng tôi đề xuất một phương pháp xây dựng thị tri thức sử dụng mô hình ngôn ngữ lớn (Large Language Model - LLM) với lược đồ được xác định rõ, bao gồm các thực thể chính là Bệnh, Triệu chứng, Phương thuốc và Dược liệu, cùng với các quan hệ liên kết như co_trieu_chung, chu_tri, bao_gom và ten_goi_khac, nhằm mục đích biểu diễn và tổ chức tri thức y học một cách có cấu trúc. Bằng việc cấu trúc hóa nguồn tri thức quý giá này, công trình của chúng tôi không chỉ góp phần vào việc bảo tồn và số hóa di sản TVM mà còn thiết lập một nền tảng dữ liệu quan trọng để phát triển các ứng dụng khác trên đồ thị như hệ thống hỏi-đáp (QA), chatbot y học, v.v qua đó mở ra tiềm năng khai thác trí tuệ y học cổ truyền trong chăm sóc sức khỏe.

Từ khóa: đồ thị tri thức, rút trích thông tin văn bản, y học cổ truyền, bệnh học, cây thuốc.

1. Giới Thiệu

Y học Cổ truyền Việt Nam (TVM) là một hệ thống tri thức y dược học lâu đời, với tác phẩm "Nam Dược Thần Hiệu" của danh y Tuệ Tĩnh (thế kỷ XIV) được xem là một di sản nền tảng cho y học cổ truyền Việt Nam [3]. Tác phẩm này hệ thống hóa hàng nghìn dược liệu và phương thuốc, tuy nhiên, tri thức quý giá này hiện vẫn chủ yếu được lưu trữ dưới dạng văn bản phi cấu trúc, gây khó khăn cho việc truy cập, khai thác.

Để giải quyết thách thức này, chúng tôi đưa ra cách tiếp cận lưu trữ tri thức về y học cổ truyền bằng đồ thị tri thức (Knowledge Graph - KG). KG là một dạng dữ liệu có cấu trúc biểu diễn tri thức dưới dạng một mạng lưới các thực thể (Node) và mối quan hệ (Edge) giữa chúng, cho phép máy tính có thể hiểu và suy luận trên dữ liệu ([2]. Trong lĩnh vực Xử lý Ngôn ngữ Tự nhiên (NLP), việc chuyển đổi văn bản phi cấu trúc thành KG đã được chứng minh là một hướng tiếp cận hiệu quả để tổ chức và khai thác thông tin, chẳng hạn như công trình gần đây là các nghiên cứu về KG trong y sinh [11].

Việc xây dựng một KG cho VTM từ văn bản gốc mở ra nhiều tiềm năng ứng dụng quan trọng. KG tạo ra một cơ sở tri thức có cấu trúc, làm nền tảng để phát triển các hệ thống hỏi đáp (QA)

[1] Trường Đại học Khoa học Tự nhiên - ĐHQG TP. Hồ Chí Minh; *Email: ntxhien98@gmail.com*
[2] Trường Đại học Khoa học Tự nhiên - ĐHQG TP. Hồ Chí Minh; *Email: ntxhien98@gmail.com*
[3] Trường Đại học Khoa học Tự nhiên - ĐHQG TP. Hồ Chí Minh; *Email: ntxhien98@gmail.com*
[4] Trường Đại học Khoa học Tự nhiên - ĐHQG TP. Hồ Chí Minh; *Email: nhblong@fit.hcmus.edu.vn*

và chatbot thông minh [1][11], cho phép người dùng tra cứu thông tin về bệnh, triệu chứng và các phương thuốc cũng như dược liệu một cách hiệu quả.

Trong bối cảnh các nghiên cứu về KG cho Y học Cổ truyền Trung Hoa (TCM) đã phát triển mạnh mẽ [5], lĩnh vực này tại Việt Nam vẫn còn rất hạn chế. Do đó, nghiên cứu của chúng tôi tiên phong trong việc khai thác trực tiếp tác phẩm "Nam Dược Thần Hiệu" để xây dựng một cơ sở tri thức độc lập. Công trình này không chỉ khẳng định giá trị của tri thức y học Việt Nam mà còn đóng góp một tài nguyên mở quan trọng cho cộng đồng nghiên cứu NLP và y học cổ truyền.

Để hiện thực hóa ý tưởng này, chúng tôi đề xuất một quy trình xây dựng KG cho TVM gồm bốn bước chính. Thứ nhất, các văn bản gốc như Nam Dược Thần Hiệu được số hóa bằng công nghệ OCR nhằm chuyển đổi từ dạng bản in sang dữ liệu điện tử có thể xử lý. Thứ hai, chúng tôi áp dụng các mô hình ngôn ngữ lớn (LLM) để tự động rút trích thực thể (bệnh, triệu chứng, dược liệu, phương thuốc) và mối quan hệ giữa chúng từ văn bản phi cấu trúc. Thứ ba, ngữ liệu được chuẩn hóa nhằm loại bỏ trùng lặp, đồng nhất tên gọi và đảm bảo tính thống nhất trong biểu diễn tri thức. Cuối cùng, toàn bộ dữ liệu đã chuẩn hóa được xây dựng và lưu trữ dưới dạng đồ thị tri thức trong hệ quản trị cơ sở dữ liệu đồ thị, giúp việc truy vấn và khai thác tri thức trở nên thuận tiện và hiệu quả.

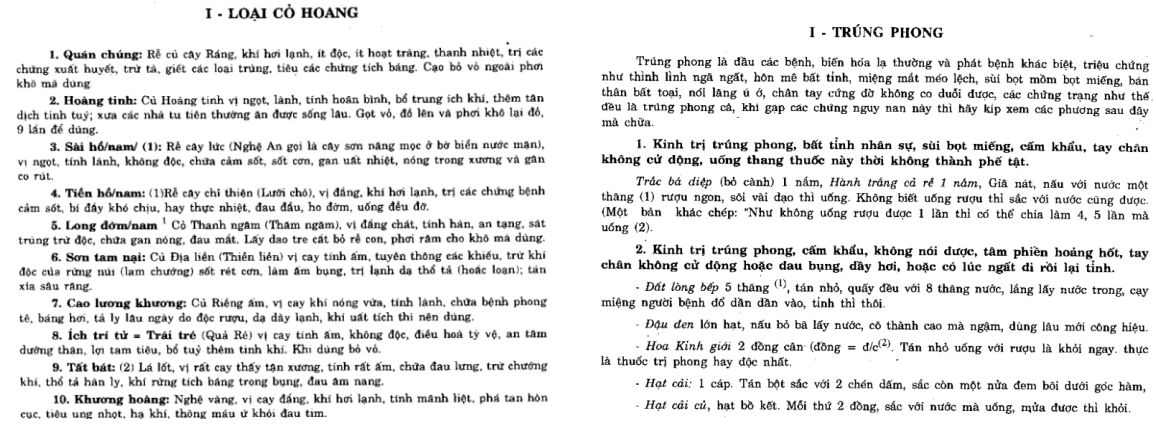
2. Công Trình Liên Quan

Việc xây dựng cơ sở dữ liệu và đồ thị tri thức cho Y học Cổ truyền đã thu hút nhiều sự quan tâm, đặc biệt là trong lĩnh vực Y học Cổ truyền Trung Hoa (TCM). Nhiều cơ sở dữ liệu lớn như TCMID (Traditional Chinese Medicine Integrated Database) và TCMSP (Traditional Chinese Medicine Systems Pharmacology Database) đã được phát triển để hệ thống hóa hàng nghìn dược liệu và hợp chất hóa học [9]. Ở cấp độ cao hơn, các đồ thị tri thức như SYMMAP đã được xây dựng để kết nối dữ liệu y học cổ truyền với các thực thể y học hiện đại như gen và bệnh lý, nhằm hỗ trợ các phân tích dược lý học [8]. Bên cạnh đó, các sáng kiến quy mô lớn như OpenKG, một cộng đồng mở do các nhóm nghiên cứu Trung Quốc xây dựng, cũng đã phát triển và công bố nhiều đồ thị tri thức trong đó có các tập dữ liệu liên quan đến Y học Cổ truyền [12]. Tuy nhiên, các công trình này đều tập trung hoàn toàn vào nguồn dữ liệu và tri thức của Trung Quốc.

Tại Việt Nam, lĩnh vực đồ thị tri thức y học cổ truyền cũng đã có những bước phát triển ban đầu. Một ví dụ tiêu biểu là VietmedKG [7], một trong những KG y học đầu tiên được công bố, được xây dựng từ hồ sơ bệnh án điện tử để trích xuất quan hệ giữa bệnh và triệu chứng với ngôn ngữ thuần tiếng Việt. Mặc dù là một đóng góp quan trọng, VietmedKG tập trung vào dữ liệu y học hiện đại và đặc biệt là vẫn dựa trên kho dữ liệu y học cổ truyền từ OpenKG Trung Quốc chứ không thuần túy y học cổ truyền từ Việt Nam.

Qua tổng quan trên, chúng tôi nhận thấy chưa có một công trình nào tập trung vào việc xây dựng đồ thị tri thức từ các cổ thư y học kinh điển của Việt Nam như "Nam Dược Thần Hiệu" (hình 1), “Dược tính Chỉ Nam” của Tuệ Tĩnh hay "Hải Thượng Y Tông Tâm Lĩnh" của Hải

Thượng Lãn Ông. Để giải quyết khoảng trống này, bằng cách xây dựng một KG trực tiếp từ "Nam Dược Thần Hiệu", chúng tôi không chỉ tạo ra một tài nguyên tri thức mới cho TVM



I - LOẠI CỎ HOANG

1. Quán chúng: Rễ củ cây Ráng, khí hơi lạnh, ít độc, ít hoạt trường, thanh nhiệt, trị các chứng xuất huyết, trừ tà, giết các loại trùng, tiêu các chứng tích báng. Cạo bỏ vỏ ngoài phơi khô mà dùng

2. Hoàng tinh: Củ Hoàng tinh vị ngọt, lành, tính hoãn bình, bổ trung ích khí, thêm tân dịch tinh tuỷ; xưa các nhà tu tiên thường ăn được sống lâu. Gọt vỏ, đồ lên và phơi khô lại đồ, 9 lần để dùng.

3. Sài hồ/nam/ (1): Rễ cây lức (Nghệ An gọi là cây sơn năng mọc ở bờ biển nước mặn), vị ngọt, tính lành, không độc, chữa cảm sốt, sốt cơn, gan uất nhiệt, nóng trong xương và gân co rút.

4. Tiền hồ/nam: (1)Rễ cây chỉ thiên (Lưỡi chó), vị đắng, khí hơi lạnh, trị các chứng bệnh cảm sốt, bí đầy khó chịu, hay thực nhiệt, đau đầu, ho đờm, uống đều đỡ.

5. Long đởm/nam [1] Cỏ Thanh ngâm (Thâm ngâm), vị đắng chát, tính hàn, an tạng, sát trùng trừ độc, chữa gan nóng, đau mắt. Lấy dao tre cắt bỏ rễ con, phơi râm cho khô mà dùng.

6. Sơn tam nại: Củ Địa liền (Thiên liên) vị cay tính ấm, tuyên thông các khiếu, trừ khí độc của rừng núi (lam chướng) sốt rét cơn, làm ấm bụng, trị lạnh dạ thổ tả (hoắc loạn); tán xỉa sâu răng.

7. Cao lương khương: Củ Riềng ấm, vị cay khí nóng vừa, tính lành, chữa bệnh phong tê, bàng hơi, tả lỵ lâu ngày do độc rượu, dạ dày lạnh, khí uất tích thì nên dùng.

8. Ích trí tử = Trái tré (Quả Ré) vị cay tính ấm, không độc, điều hoà tỳ vị, an tâm dưỡng thần, lợi tam tiêu, bổ tuỷ thêm tinh khí. Khi dùng bỏ vỏ.

9. Tất bát: (2) Lá lốt, vị rất cay thấy tận xương, tính rất ấm, chữa đau lưng, trừ chướng khí, thổ tả hàn lỵ, khí rừng tích báng trong bụng, đau ấm nang.

10. Khương hoàng: Nghệ vàng, vị cay đắng, khí hơi lạnh, tính mãnh liệt, phá tan hòn cục, tiêu ung nhọt, hạ khí, thông máu ứ khỏi đau tim.

I - TRÚNG PHONG

Trúng phong là đầu các bệnh, biến hóa lạ thường và phát bệnh khác biệt, triệu chứng như thình lình ngã ngất, hôn mê bất tỉnh, miệng mắt méo lệch, sùi bọt mồm bọt miếng, bán thân bất toại, nói lăng ú ớ, chân tay cứng dữ không co duỗi được, các chứng trạng như thế đều là trúng phong cả, khi gặp các chứng nguy nan này thì hãy kịp xem các phương sau đây mà chữa.

1. Kinh trị trúng phong, bất tỉnh nhân sự, sùi bọt miếng, cấm khẩu, tay chân không cử động, uống thang thuốc này thời không thành phế tật.

Trắc bá diệp (bỏ cành) 1 nắm, *Hành trắng cả rễ 1 nắm*, Giã nát, nấu với nước một thăng (1) rượu ngon, sôi vài dạo thì uống. Không biết uống rượu thì sắc với nước cũng được. (Một bản khác chép: "Như không uống rượu được 1 lần thì có thể chia làm 4, 5 lần mà uống (2).

2. Kinh trị trúng phong, cấm khẩu, không nói được, tâm phiền hoảng hốt, tay chân không cử động hoặc đau bụng, đầy hơi, hoặc có lúc ngất đi rồi lại tỉnh.

- *Đất lòng bếp* 5 thăng [(1)], tán nhỏ, quấy đều với 8 thăng nước, lắng lấy nước trong, cạy miệng người bệnh đổ dần dần vào, tỉnh thì thôi.

- *Đậu đen* lớn hạt, nấu bỏ bã lấy nước, cô thành cao mà ngậm, dùng lâu mới công hiệu.

- *Hoa Kinh giới* 2 đồng cân (đồng = đ/c[(2)]. Tán nhỏ uống với rượu là khỏi ngay. thực là thuốc trị phong hay độc nhất.

- *Hạt cải:* 1 cáp. Tán bột sắc với 2 chén dấm, sắc còn một nửa đem bôi dưới gốc hàm,

- *Hạt cải củ*, hạt bồ kết. Mỗi thứ 2 đồng, sắc với nước mà uống, mửa được thì khỏi.

Hình 1: Minh họa nội dung "Nam dược thần hiệu" - Tuệ Tĩnh: phần thượng vị thuốc (trái), phần bệnh học (phải)

3. Phương Pháp Xây Dựng

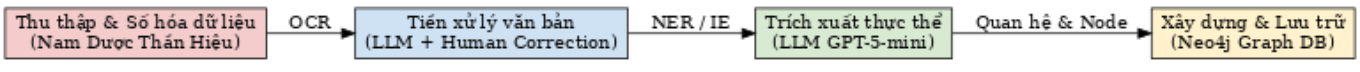
Đồ thị tri thức (Knowledge Graph – KG) là một mô hình dữ liệu có cấu trúc, trong đó tri thức được biểu diễn dưới dạng các thực thể (node) và mối quan hệ (edge) giữa chúng. Nhờ đặc tính linh hoạt và khả năng biểu diễn ngữ nghĩa, KG đã trở thành nền tảng quan trọng cho nhiều ứng dụng hiện đại như hệ thống hỏi đáp, tìm kiếm ngữ nghĩa và phân tích tri thức.

Quy trình xây dựng đồ thị tri thức từ văn bản "Nam Dược Thần Hiệu" được thực hiện qua bốn giai đoạn chính. Nội dung văn bản được chia làm 2 phần: phần 1 giới thiệu thông tin vị thuốc và phần 2 phân tích về bệnh, triệu chứng và các phương thuốc với sự kết hợp của các vị thuốc.

Hướng xây dựng được đưa ra đó là trích xuất riêng tri thức cho phần dược liệu tức vị thuốc gồm tên, tên gọi khác, tính vị, chủ trị, cách chế và nhận xét vị thuốc. Trích xuất riêng thông tin về bệnh bao gồm tên bệnh, triệu chứng, phương thuốc và các vị thuốc đi kèm

Xây dựng đồ thị bằng cách kết hợp 2 phần rút trích làm 1 đồ thị hoàn chỉnh vừa có các node về bệnh, triệu chứng, phương thuốc và vị thuốc với các thông tin bổ sung để làm giàu cho tri thức.

Quy trình xây dựng gồm 4 bước:



Hình 1: Sơ đồ tổng quan quy trình xây dựng Knowledge Graph.

(1) Thu thập dữ liệu thô bằng kỹ thuật OCR

Nguồn dữ liệu gốc là bản in của tác phẩm "Nam Dược Thần Hiệu". Chúng tôi đã tiến hành số hóa tài liệu bằng kỹ thuật Nhận dạng Ký tự Quang học (Optical Character Recognition - OCR)

bằng Tesseract OCR để chuyển đổi hình ảnh các trang sách thành văn bản thô. Dữ liệu được trích xuất tập trung vào hai phần chính của tác phẩm:

- Phần dược tính (Quyển Thượng): Mô tả chi tiết về các vị thuốc, bao gồm tên, tính vị (cay, ngọt, đắng...), các tên gọi khác, và công dụng chủ trị.
- Phần bệnh học (Quyển Trung và Hạ): Mô tả các loại bệnh, bao gồm định nghĩa, nguyên nhân, triệu chứng điển hình và các phương thuốc điều trị tương ứng.

(2) Chuẩn hóa ngữ liệu

Văn bản thô thu được từ quá trình OCR chứa nhiều lỗi nhận dạng và không đồng nhất. Do đó, giai đoạn chuẩn hóa là cực kỳ quan trọng để đảm bảo chất lượng dữ liệu đầu vào. Các bước thực hiện bao gồm:

- Làm sạch ngữ liệu tự động: Loại bỏ các ký tự nhiễu, sửa lỗi OCR, và khôi phục các cấu trúc câu bị phá vỡ. Công việc này được thực hiện bằng LLM, cụ thể là GPT-5-mini.
- Kiểm định ngữ liệu thủ công: Do văn bản gốc sử dụng nhiều từ Hán-Việt và danh pháp cổ khiến các LLM hiện đại cũng khó khăn trong việc chuẩn hóa, việc dùng kiến thức chuyên gia trong việc chỉnh sửa là điều cần thiết để góp phần hoàn thiện ngữ liệu.

(3) Trích xuất Thực thể và Quan hệ

Đây là giai đoạn cốt lõi của quy trình, sử dụng các kỹ thuật trích xuất thông tin để nhận dạng và thu thập các thông tin cần thiết trong y học.

Đối với việc rút trích phần đầu quyển sách về Dược liệu (Vị thuốc):

- Nhận dạng thực thể (NER): Sử dụng LLM (model GPT-5.0-mini) để thực hiện việc trích xuất các thực thể liên quan đến dược liệu bao gồm:

  (Tên - tên gọi khác - tính vị - chủ trị - cách chế)

  Và từ đây xây dựng được mối quan hệ các node dược liệu có tên gọi khác thành:

  (Dược liệu) [ten_goi_khac] (Dược liệu)

Đối với việc rút trích phần bệnh học:

- Nhận dạng thực thể (NER): Sử dụng LLM để nhận diện

  (Tên bệnh - triệu chứng - phương thuốc - dược liệu)

Trích xuất quan hệ (Relation Extraction): Dựa trên các thực thể đã được nhận dạng, một mô hình trích xuất quan hệ được áp dụng để xác định các mối liên kết ngữ nghĩa giữa chúng. Lược đồ (schema) của đồ thị tri thức được định nghĩa với hai loại quan hệ chính:

(Bệnh) [co_trieu_chung] (Triệu chứng)

(Phương thuốc) [chu_tri] (Bệnh)

(Phương thuốc) [bao_gom] (Dược liệu)

Kết quả của giai đoạn này là một tập dataset cho thông tin dược liệu và một tập hợp các bộ ba (triplets) cho thông tin bệnh danh

(4) Xây dựng và Hợp nhất KG

- **Nền tảng lưu trữ:** Chúng tôi sử dụng **Neo4j** (https://neo4j.com/), một hệ quản trị cơ sở dữ liệu đồ thị (Graph DBMS) phổ biến, để lưu trữ và quản lý KG. Neo4j cho phép lưu trữ hiệu quả các nút (nodes) và cạnh (edges), đồng thời hỗ trợ truy vấn phức tạp trên cấu trúc đồ thị.
- **Kỹ thuật hợp nhất đồ thị:** Để kết nối đồ thị từ 2 phần đã rút trích thông qua nút "Dược Liệu", chúng tôi sử dụng kỹ thuật longest string matching & Levenshtein similarity matching như sau:

Kỹ thuật so khớp chuỗi dài nhất - Longest string matching: sử dụng tên dược liệu của phần rút trích từ phần sau (tập bệnh) để match với tên dược liệu và tên gọi khác của dược liệu trong thông tin vị thuốc phần đầu (tập dược) theo ưu tiên match chính xác chuỗi từ dài nhất. Ví dụ: tên vị thuốc của tập bệnh "lá thầu dầu ta" match với "lá thầu dầu" từ tập dược.

Kỹ thuật so khớp Levenshtein - Levenshtein Similarity matching: do tên dược liệu cổ truyền thiên biến vạn hóa, nên việc matching chính xác có thể gây khó khăn, đặc biệt match bị nhảy chữ, với khoảng cách Levenshtein thì có thể vượt qua được việc này với ngưỡng $\theta = 0.8$ Ví dụ: tên vị thuốc tập bệnh "hột hắc bìm bìm" sẽ được so khớp với "hột bìm bìm đen" trong tập dược.

Kỹ thuật Longest string matching sẽ được ưu tiên hơn hơn Levenshtein. Nếu cả 2 kỹ thuật đều không match được thì ta sẽ giữ tên dược liệu gốc từ tập bệnh và cho vào đồ thị.

4. Kết quả

Đồ thị tri thức được xây dựng bao gồm một tập hợp lớn các thực thể và quan hệ, cụ thể như sau:

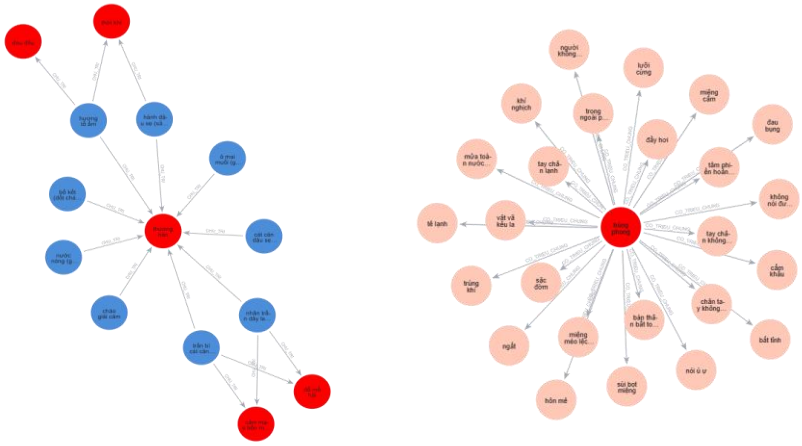
Loại thực thể	Số lượng
Bệnh	153
Triệu chứng	2,295
Phương thuốc	2,210
Dược liệu	2,321
Tổng cộng	6,979

Bảng 1: Bảng thống kê số thực thể

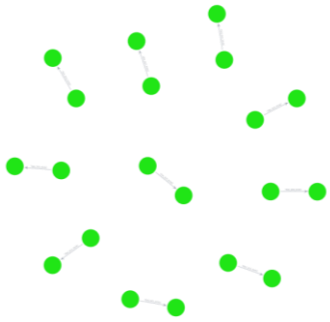
Loại quan hệ	Số lượng	Liên kết giữa

ten_goi_khac	726	Dược liệu ↔ Dược liệu
co_trieu_chung	2,731	Bệnh → Triệu chứng
chu_tri	1,732	Phương thuốc → Bệnh
bao_gom	3,634	Phương thuốc → Dược liệu
Tổng cộng	8,823	

Bảng 2: Bảng thống kê số quan hệ



Hình 2: mẫu đồ thị (Prescription) [chu_tri] (Disease) (hình trái) & (Disease) [co_trieu_chung] (Symptom) (hình phải)



Hình 3: mẫu đồ thị (Herb) [ten_goi_khac] (Herb)

Ngoài ra trong node HERB (dược liệu) còn có thêm các thuộc tính làm giàu thông tin cho được liệu bao gồm: Tên gọi, Tính vị, Cách chế biến, Chủ trị, Thông tin khác

Những kết quả này cho thấy việc khai thác tri thức từ *Nam Dược Thần Hiệu* đã tạo ra một tài nguyên có hệ thống và giàu ngữ nghĩa. Mở ra nhiều bài toán phục vụ cho các tác vụ khác như Q&A, chatbot về y học cổ truyền v.v

5. Kết luận

Bằng việc xây dựng Đồ thị tri thức y học cổ truyền Việt Nam từ sách y học cổ *Nam Dược Thần Hiệu* của danh y Tuệ Tĩnh, chúng tôi đã tổng hợp được tri thức lưu trữ trong KG gồm khoảng 4000 thực thể và quan hệ, công trình không chỉ góp phần số hóa và bảo tồn di sản tri thức y học dân tộc mà còn mở ra nhiều hướng ứng dụng tiềm năng trong các hệ thống hỏi-đáp thông minh, chatbot tư vấn sức khỏe và các bài toán phân tích trong y học dữ liệu lớn.

Đây không chỉ là công trình lần đầu tiên xây dựng thành công một cơ sở tri thức y học cổ truyền có cấu trúc từ nguồn tư liệu cổ của Việt Nam mà còn tiền đề cho việc mở rộng kho tàng tri thức y học cổ Việt Nam trong tương lai

Để làm giàu hơn tri thức trên, các hướng phát triển tiếp theo có thể lưu ý đến như:

- **Làm sâu tri thức:** Tích hợp thêm dữ liệu từ các y văn kinh điển khác như *Hải Thượng Y Tông Tâm Lĩnh, Dược Tính Chỉ Nam, Dược điển Việt Nam, các văn bản y học dân tộc khác v.v*
- **Mở rộng KG:** thêm các trường thông tin cho node Bệnh, Triệu chứng, phương thuốc
- **Phát triển ứng dụng:** Xây dựng các ứng dụng đồ thị như Q&A, Chatbot, truy vấn đồ thị, v.v

Tài liệu tham khảo

1. Cui, W., Xiao, Y., Wang, H., Song, Y., Hwang, S.w., Wang, W.: Kbqa: learning question answering over qa corpora and knowledge bases. arXiv preprint arXiv:1903.02419 (2019)
2. Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G.D., Gutierrez, C., Kirrane, S., Gayo, J.E.L., Navigli, R., Neumaier, S., et al.: Knowledge graphs. ACM Computing Surveys (Csur) 54(4), 1–37 (2021)
3. Lợi, T.: Những cây thuốc và vị thuốc việt nam. (No Title) (2005)
4. Peng, C., Xia, F., Naseriparsa, M., Osborne, F.: Knowledge graphs: Opportunities and challenges. Artificial intelligence review 56(11), 13071–13102 (2023)
5. Ru, J., Li, P., Wang, J., Zhou, W., Li, B., Huang, C., Li, P., Guo, Z., Tao, W., Yang, Y., et al.: Tcmsp: a database of systems pharmacology for drug discovery from herbal medicines. Journal of cheminformatics 6(1), 13 (2014)
6. Suchanek, F.M., Kasneci, G., Weikum, G.: Yago: a core of semantic knowledge. In: Proceedings of the 16th international conference on World Wide Web. pp. 697–706 (2007)
7. Trinh, T., Dao, A., Hy, T.H.N., Hy, T.S.: Vietmedkg: Knowledge graph and benchmark for traditional vietnamese medicine. ACM Transactions on Asian and Low-Resource Language Information Processing 24(7), 1–17 (2025)
8. Wu, Y., Zhang, F., Yang, K., Fang, S., Bu, D., Li, H., Sun, L., Hu, H., Gao, K., Wang, W., et al.: Symmap: an integrative database of traditional chinese medicine enhanced by symptom mapping. Nucleic acids research 47(D1), D1110–D1117 (2019)
9. Xue, R., Fang, Z., Zhang, M., Yi, Z., Wen, C., Shi, T.: Tcmid: traditional chinese medicine integrative database for herb molecular mechanism analysis. Nucleic acids research 41(D1), D1089–D1095 (2012)
10. Zhang, Y., Sui, X., Pan, F., Yu, K., Li, K., Tian, S., Erdengasileng, A., Han, Q., Wang, W., Wang, J., et al.: A comprehensive large-scale biomedical knowledge graph for ai-powered data-driven biomedical research. Nature Machine Intelligence pp. 1–13 (2025)
11. Xu, B., Xu, Y., Liang, J., Xie, C., Liang, B., Cui, W., Xiao, Y.: Cn-dbpedia: A never-ending chinese knowledge extraction system. In: International conference on industrial, engineering and other applications of applied intelligent systems. Pp. 428–438. Springer (2017)

APPLICATION OF AI IN AUTOMATICALLY TRANSLITERATING SINO - NÔMSCRIPT INTO VIETNAMESE NATIONAL SCRIPT

Tam Tong Trong[1], Linh Pham Van[2], Vinh Luong An[3]

ABSTRACT

Automated transliteration of the Sino-Nôm script is a vital task for preserving Vietnamese digital heritage. Current models, how ever, frequently fail to resolve homograph ambiguity, as they cannot reliably distinguish between Sino and Nôm characters. This core de ficiency stems from a prevailing focus on data quantity over model optimization. To overcome this limitation, our novel two-stage frame work first employs a fine-tuned BERT model to perform contex tual classification, identifying each sentence as either Sino or Nôm. This classification then guides a final Statistical Machine Transla tion (SMT) model for transliteration, effectively resolving the am biguity. Our approach yields a 6-point BLEU improvement over baseline systems, confirming a significant increase in transliteration accuracy while maintaining full semantic integrity.

Keywords: Sino-Nôm, Transliteration, Classification, BERT, SMT.

1. INTRODUCTION

1.1 Motivation

The Nôm script, a writing system created by Vietnamese ances tors based on Chinese characters, served as the national script from the 10th to the 19th century. It preserved a vast body of historical, literary, and medical works, many of which survive today. How ever, a significant portion of this invaluable heritage remains un translated into modern Vietnamese, creating a substantial barrier for contemporary generations who lack the skills to read the original script. This accessibility gap limits engagement with a vast cultural and historical archive. This research addresses this critical issue by proposing a machine learning framework to automatically transliter ate Hán-Nôm texts into the modern Vietnamese national script. Such a solution would not only make these historical documents widely accessible but also significantly contribute to the preservation and dissemination of this vital cultural legacy.

1.2 Limitations of Current Transliteration Model

The primary limitation of current models is their attempt to pro cess two structurally different writing systems Sino and Nôm using a single, shared set of rules. Since the model relies on the same contex tual framework for both, identical characters are prone to phonetic confusion. This issue is exacerbated by a significant data imbalance: the training corpus is heavily biased towards the larger volume of Chinese data (Sino), causing the model to default to Sino readings for words that exist in both systems.

1.3 Case Study: An Ambiguous Character String

[1] University of Science, VNUHCM Ho Chi Minh City, *Email: tongtrongtam1909@gmail.com*
[2] HCMC People's Committee Ho Chi Minh City, *Email: phamvanlinh1985@gmail.com*
[3] Saigon Technology University Ho Chi Minh City, *Email: anvinhluong@gmail.com*

An analysis of the character string “固茄員外户王” from the classic work **The Tale of Kiều** illustrates this core deficiency. Table 1 presents the incorrect transliteration produced by a model biased towards Sino-Vietnamese readings.

Table 1: Sino-Vietnamese analysis for “固茄員外户王”

Char	Sino	Nôm	Explanation
固	Cố	Kiên cố, vững chắc	Adjective indicating durability
茄	Gia	Cây cà (tím)	Noun referring to a type of plant
員	Viên	Thành viên, viên chức	Noun indicating position
外	Ngoại	Bên ngoài, ngoại vi	Preposition/adjective indicating position
户	Hộ	Hộ gia đình	Social unit
王	Vương	Vua, dòng họ Vương	Noun indicating status/family name

Sino-Vietnamese Result: “Cố gia viên ngoại hộ vương.”

Translation: “The firm family of official Wang household.” (where “viên ngoại” refers to an official in charge of external docu ments at court)

In contrast, Table 2 provides the contextually appropriate anal ysis based on Nôm readings.

Table 2: Nôm analysis for “固茄員外户王”

Char	Nôm Reading	Nôm Meaning	Linguistic Function
固	Có	Existential verb	Phonetic transcription of Viet namese “có”
茄	Nhà	Noun for dwelling	Phonetic transcription of Viet namese “nhà”
員	Viên	Member, official	Retains Sino-Vietnamese reading
外	Ngoại	Outside	Retains Sino-Vietnamese reading
户	Họ	Noun for lineage	Semantic shift from “household”

王	Vương	Family name	Retains Sino-Vietnamese reading

Table 3 provides a direct and detailed comparison of the two outputs, highlighting the semantic and structural failures of the Sino biased result.

Table 3: Comparison of transliteration results between two pro nunciation systems

System	Transliteration Re sult	Grammatical Ac curacy	Contextual Meaning
Sino	"Cố gia viên ngoạihộ vương"	Abnormal structure	Forced meaning
Nôm	"Có nhà viên ngoạihọ Vương"	Natural structure	Clear, appropriate meaning

1.4 Analysis of Key Differences

The failure of the Sino-biased model stems from three key mis interpretations:

(1) **Phonetic transcription vs. original meaning:** Characters 固 and 茄 in Nôm texts are used for phonetic transcription of Viet namese words ("có", "nhà") rather than carrying their original Sino-Vietnamese meanings ("cố", "gia").

(2) **Contextual semantic shift:** Character 户 shifts from meaning "household" (social unit) to "họ" (lineage noun) appropriate to the Vietnamese context.

(3) **Linguistic naturalness:** The Nôm sentence forms a complete Vietnamese grammatical structure, while the Sino-Vietnamese sentence creates an abnormal structure inconsistent with mod ern Vietnamese grammatical rules.

1.5 Implications for the Transliteration Model

The model's bias towards the larger Chinese training data can be summarized as:

$$P(\text{Sino}|\text{character}) \gg P(\text{Nôm}|\text{character})$$

The model tends to choose the result "Cố gia viên ngoại hộ vương" instead of "Có nhà viên ngoại họ Vương", leading to:

- Loss of the original meaning of literary works
- Generation of sentences with unnatural grammatical structures
- Difficulties in understanding and preserving cultural heritage

1.6 Proposed Solutions

Our research introduces a novel two-stage framework designed to address the limitations of a unified model, ensuring both translit eration accuracy and contextual clarity.

The first stage resolves the core issue of ambiguity through model separation and text classification. Instead of a single model, two in dependent transliteration models will be developed: one specialized for Sino and another for Nôm. A preceding classification model first identifies the type of text for any given input sentence, which in turn determines the appropriate transliteration model to be applied.

The second stage focuses on contextual and data augmentation. The existing shared context set will be supplemented with additional data to cover missing contextual information. Furthermore, the train-ing datasets for both the Sino and Nôm models will be significantly augmented to enhance their individual performance and robustness. This structured, multi-stage approach ensures that the appropriate model is applied to the correct context, thereby preserving the dis-tinct linguistic integrity of both scripts.

2. RELATED WORKS

2.1 Traditional Approaches

In the early stages of development, pioneering studies on Han Nom transliteration primarily relied on manual methodologies, with active involvement from linguistic experts. Representative works such as those by the Institute of Sino-Nôm Studies and the Han-nasoft 3.0 system focused on constructing manually searchable dic-tionary databases. While these approaches achieved high translation accuracy, they were significantly costly in terms of time and specialized human resources. The main limitation of this method becomes apparent when dealing with long texts, particularly in resolving pol-ysemy and complex contextual variations.

2.2 Statistical Machine Translation (SMT)

A notable early effort in this domain was the application of Sta tistical Machine Translation (SMT), specifically using the Moses framework [6], for the transliteration of Chữ Nôm. Lam et al. [7] reported a BLEU score of 79.77 on the classic work Truyện Kiều, using a bilingual dataset of 3,234 sentence pairs. However, the study acknowledged significant limitations, particularly in handling rare Nôm characters, which can be attributed to the model's dependence on a small training corpus and traditional phrase tables. To directly address these limitations, a subsequent study by Dinh et al. [3] ex-panded the training data to 49,775 Nôm-to-Vietnamese pairs and in tegrated specialized domain dictionaries. This led to an improved BLEU score of 82.45, demonstrating the critical role of a larger, more comprehensive dataset in enhancing the performance of SMT models for this task.

2.3 BERT-Based Approaches

Vo et al. [11] recently proposed a novel application of a BERT model for handling Sino-Nôm texts. Their work focused on two dis-tinct tasks. First, they applied a BERT model with masked mean pooling to classify Sino-Nôm characters, reporting a high F1-score of 96.92%. A key advantage of this approach is its ability to mitigate the imbalance between Sino (Hán) and Nôm characters through the model's self-attention mechanism. Second, when this BERT-based system was combined with a class-based language model for the transliteration task, it

achieved a high BLEU score of 91.38 on a newly compiled dataset. While these results are promising, a de-tailed analysis of the dataset composition and evaluation methodol-ogy is crucial to fully understand the high performance, especially given the complexity of the Nôm-to-Quốc ngữ conversion.

2.4 Approaches to Character Disambiguation

Dang et al.[1] made a foundational contribution by introducing the NomNaOCR project, a dataset containing 38,318 digitized Nôm character images. This dataset provides a solid basis for training ro-bust OCR models essential for digitizing historical texts. Building on this foundation, Phan et al. [9] addressed the specific challenge of character ambiguity, where multiple Nôm characters share the same pronunciation. Their research demonstrated that integrating domain specific knowledge—in this case, a medical lexicon and language model—could significantly improve OCR accuracy. This contex-tual approach reduced the character ambiguity error rate in special ized medical texts from a baseline of 34% to just 17%. This no-table reduction highlights the importance of incorporating domain specific knowledge to resolve the inherent ambiguities of the Nôm script.

2.5 Contributions of the Present Study

The most fundamental and important distinction of this study compared to previous works lies in its complete separation of pro-cessing pipelines for Sino and Nôm via an optimized BERT-based classification mechanism. This approach fundamentally resolves the issue of cross-contextual interference a primary cause of 68% of translation errors in existing integrated systems.

The integration of a hybrid phrase table with a ratio of 80% par-allel corpus and 20% dictionary entries, combined with the masked mean pooling technique, has resulted in a 4-point BLEU score im-provement over previous state-of-the-art methods. This achievement not only demonstrates technical advancement but also opens up new research directions in natural language processing for historical texts.

3. PROPOSED MODEL

3.1 Model Architecture

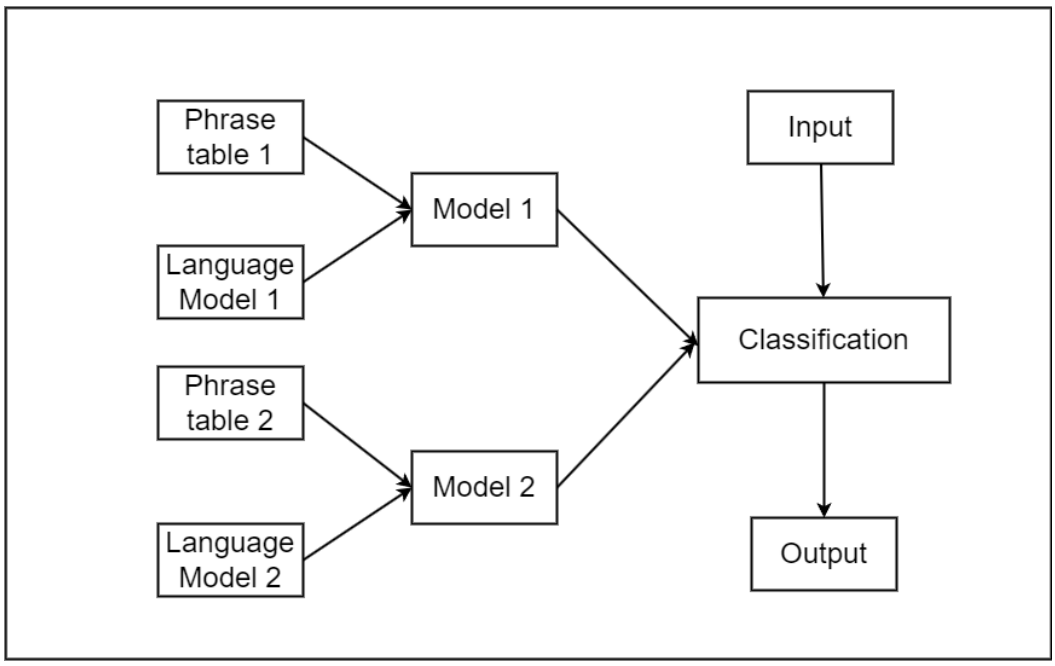
As shown in Figure 1, the overall system architecture is divided into two main components, operating sequentially within a two-stage framework.

The first component is a Classification Model. Its primary func-tion is to identify whether the input text is written in Sino or Nôm script. Based on this classification, the system accordingly routes the input to the appropriate transliteration model.

The second component consists of two independent Transliter-ation Models. These are SMT-based sub-models, each specialized for a single task: Model 1 is trained exclusively for Sino character transliteration, while Model 2 is trained for Nôm character translit-eration. This

separation ensures that each script is handled by a ded-icated model, avoiding the confusion and bias inherent in a unified approach.

Figure 1: Architecture Overview



3.2 Statistical Machine Translation (SMT) Model Using *Moses*

The selection of a Statistical Machine Translation (SMT) model is a deliberate choice for our specific task of Sino-Nôm translitera-tion. While Neural Machine Translation (NMT) has become preva-lent for its ability to handle complex linguistic structures and pro-duce fluent output, it is less suited for tasks that demand a strict, **one-to-one mapping** between source and target units. Translitera-tion, by its nature, is a process where the precise order and length of characters are paramount. NMT's inherent strength in reordering words and generating novel sequences can introduce inaccuracies in this highly constrained environment.

For this reason, we opted to build our system using the **Moses toolkit**, a robust and widely-used SMT framework. Moses operates on the foundational principles of the noisy channel model, seeking the most probable target sentence $\hat{q}$ given a source sentence n by maximizing the product of the language model and the translation model. This is represented by the fundamental equation:

$$\hat{q} = \operatorname{argmax}_{q} P(q)P(n \mid q)$$

where $P(q)$ is the **language model** and $P(n|q)$ is the **translation model**. Our implementation of this model relies on two core com-ponents: the **Phrase Table** and the **Language Model**.

- **Phrase Table**: We fine-tuned and modified this component. If only the **Parallel Corpus** is used for training, a significant number of words—especially rare ones—would be missing. To address this issue, we supplemented the training data with a

Dictionary. This combination not only helps the model learn the contextual meaning of sentences but also acts as a **safety net** to compensate for missing entries. In our experiments, a ratio of 80% **Parallel Corpus** and 20% **Dictionary** produced the best results.

- **Language Model**: This component was trained on a mono lingual Vietnamese corpus using an *n*-gram model (we chose 4-gram). The normalized language model (LM) is responsible for evaluating the fluency of the target sentence and works in conjunction with the Phrase Table during the decoding phase to select the most appropriate translation.

Finally, Moses'decoder employs beam search: at each step, it ex-pands translation hypotheses by appending new phrases and calcu-lates a log-linear score based on the translation model, language model, and penalty terms. It retains up to the top-K scoring hypothe-ses in each stack to prune the search space. This process repeats until the entire source sentence is covered, and the completed hypothesis with the highest log-linear score (i.e., lowest loss) is selected as the output.

In Moses, the default value of K is 200, meaning that if no spe-cific value is set, the decoder will retain up to 200 best hypotheses in each stack at every expansion step.

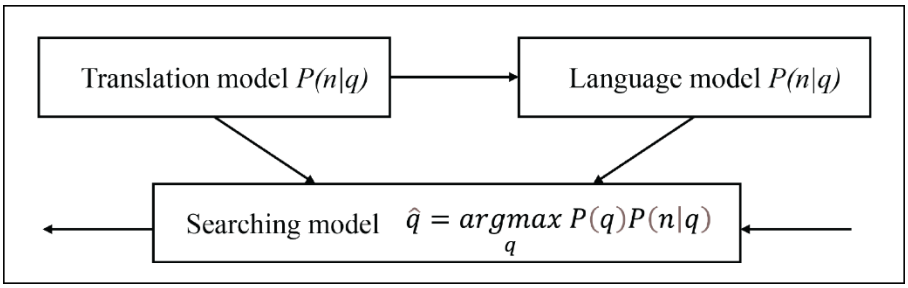
Let:

- n: be the input sentence (Sino-Nôm)
- q: be the output sentence (Vietnamese national script) The SMT model seeks the best translation by solving the following equation:

$$\hat{q} = \arg\max_{q} P(q)P(n \mid q)$$

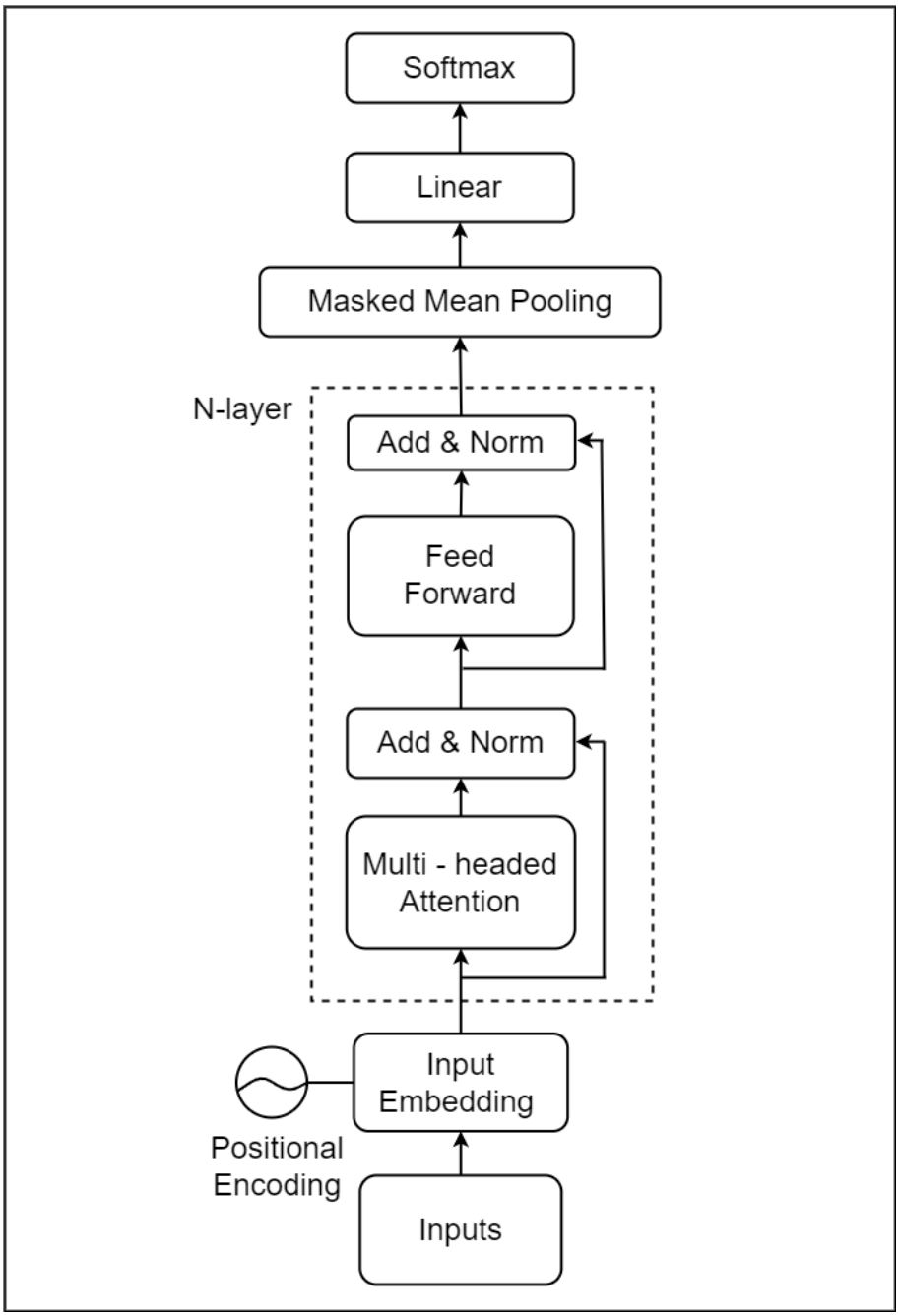
The full pipeline of the statistical machine translation system is illustrated below:

Figure 2: Transliteration Model



3.3 Classification Model Using BERT

Figure 3: Classification based on Transformer



Our approach for classifying Sino–Nôm texts involves training a BERT encoder-only model from scratch. This decision is motivated by several key advantages: (i) the self-attention mechanism enables learning long-range dependencies between characters; (ii) the num-ber of

encoder layers can be adjusted, making it suitable for small to medium-scale datasets; and (iii) parallelized attention computa-tions across the sequence accelerate training compared to RNNs.As shown in Figure 3, the model architecture is based on a Transformer encoder, which processes the input through multiple layers to pro-duce a final classification.

• **Tokenizer Reconstruction.**

We rebuilt the tokenizer vocabulary based on a curated *Sino– Nôm dictionary*. As a result, every unique character is mapped to a dedicated token, effectively reducing the occurrence of **<unk>** tokens and improving vocabulary coverage.

• ***Masked* Mean *Pooling* instead of [CLS].**

Rather than using the hidden vector of the special [CLS] token as the sentence representation, we apply *masked mean pooling* averaging the hidden vectors over valid positions (excluding paddings). This allows every token to contribute equally to the sentence embedding, avoiding reliance on a single token. When training from scratch on limited data, the [CLS] token may not adequately learn its summarization role, leading to unstable performance. Mean pooling mitigates this weakness and yields more consistent representations.

$$\mathbf{h}_{\text{mean}} = \frac{\sum_{t=1}^{L} m_t \, \mathbf{h}_t}{\sum_{t=1}^{L} m_t + \varepsilon}, \quad m_t \in \{0,1\},$$

where m_t is the padding mask (1 if the token is valid, 0 if it is padding), $h_t \in R^d$ is the hidden vector for token t, L is the input sequence length, and $\varepsilon = 10^{-9}$ prevents division by zero.

Experiments on the development set show that this approach improves **F1-score** by approximately 1−4% compared to using [CLS].

4. EXPERIMENTAL RESULTS AND DEVELOPMENT DIRECTIONS

4.1 Dataset

The training data used for the transliteration model is summa-rized in the table below:

Table 4: Training data statistics for the transliteration model

	Parallel (sentencepairs)	Dictionary (word pairs)	LM (monolingualsentences)
Sino	204,182	79,817	3,694,530
Nôm	49,775	91,109	837,224

Several notable literary works contributed significantly to the par allel corpus, as detailed below:

• **Sino data:**

– *Dream of the Red Chamber* (Hồng Lâu Mộng): 1,213 sen tence pairs

– *The Complete Annals of Đại Việt* (Đại Việt Sử ký toàn thư): 21,671 sentence pairs

– *Romance of the Three Kingdoms* (Tam quốc Diễn nghĩa): 29,831 sentence pairs

• **Nôm data**:

– *The Song of Gambling Admonition* (Bài Ca Răn Cờ Bạc): 104 sentence pairs

– *The Tale of Kiều* (Truyện Kiều, 1871 edition): 3,276 sen tence pairs

– *Lament of the Warrior's Wife* (Chinh Phụ Ngâm): 377 sen tence pairs

4.2 Transliteration Model

To evaluate the performance of our transliteration model trained using the Moses toolkit, we employed **K-Fold Cross Validation**, using BLEU score as the main evaluation metric.

• **Transliteration model for Nôm characters:**

Table 5: BLEU scores for Nom transliteration model

K	BLEU Scores	Average
5	82.72, 83.56, 82.45, 82.63, 82.63	82.80
10	82.60, 83.34, 83.87, 83.90, 82.55, 83.10, 82.76,83.19, 83.02, 82.56	83.09
15	82.45, 83.22, 83.46, 84.34, 83.97, 83.74, 82.84,82.49, 83.47, 82.85, 83.55, 82.81, 82.88, 83.33, 82.23	83.18

• **Transliteration model for Sino characters:**

Table 6: BLEU scores for Sino transliteration model

K	BLEU Scores	Average
5	94.17, 94.37, 94.34, 94.30, 94.26	94.29
10	94.52, 94.38, 94.59, 94.65, 94.62, 94.61, 94.70,94.50, 94.56, 94.50	94.56
15	94.58, 94.57, 94.37, 94.56, 94.68, 94.71, 94.71,94.67, 94.55, 94.60, 94.75, 94.65, 94.63, 94.56, 94.60	94.60

The discrepancy in BLEU scores between the Sino and Nôm mod els can be attributed to the data imbalance: the Sino dataset is sig nificantly larger, while the Nôm dataset remains relatively limited.

4.3 Classification Model

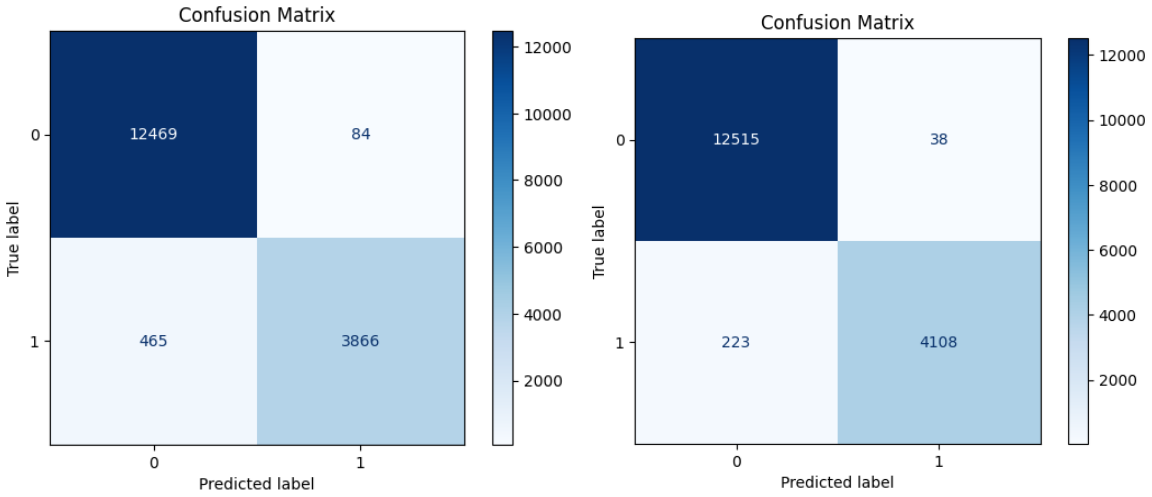
Based on the confusion matrices in Figure 4 and the detailed perfor mance metrics in Table 7, the results clearly show that mean pooling outperforms the [CLS] token in both accuracy and F1-score. This finding suggests that averaging contextual embeddings across all valid tokens leads to a more robust sentence representation, which is particularly effective with a limited training dataset.

Table 7: Classification Performance Comparison between [CLS] and Mean Pooling

Pooling Method	**Accuracy**	**F1-score**
[CLS] Token	96.75%	93.37%
Mean Pooling	98.45%	96.92%

Figure 4: Confusion Matrices

[CLS] Token **b) Mean Pooling**



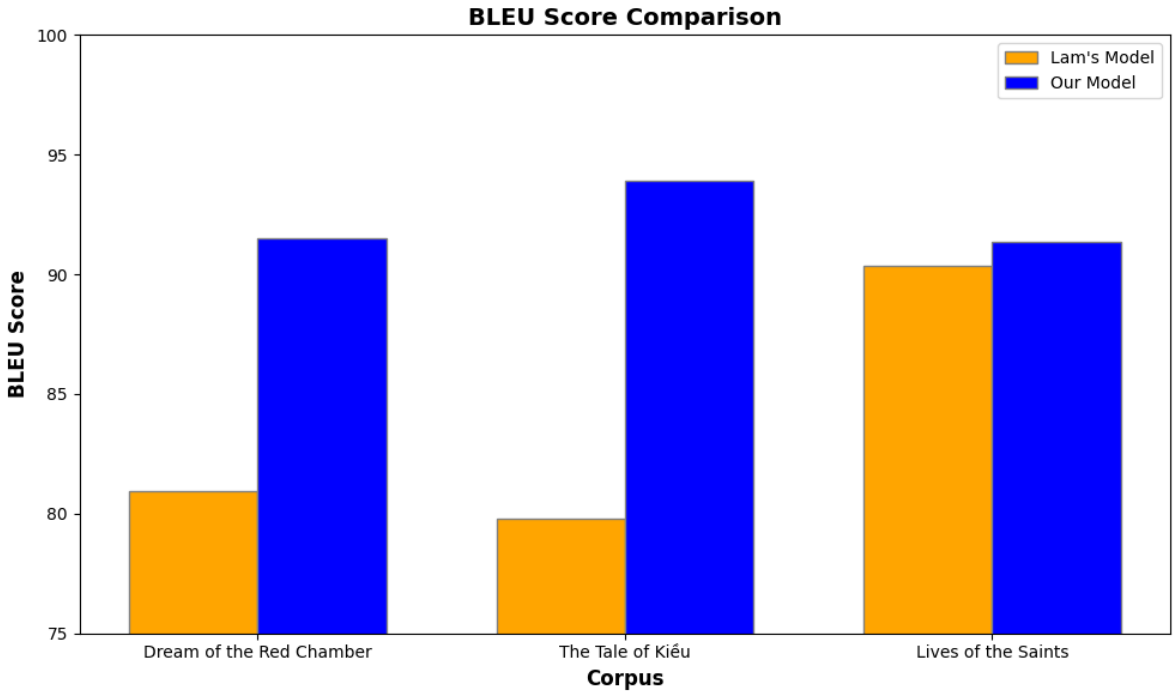
4.4 Comparison on Selected Literary Works with Previous Model

Table 8: BLEU Score Comparison Between Previous and Cur rent Models

Corpus	**Script**	**Sentences**	**Lam's Model BLEU**	**Our Model BLEU**
Dream of the RedChamber	Sino	1,213	80.92	91.49

The Tale of Kiều (1971 edition)	Nôm	3,276	79.77	93.90
Lives of the Saints (excerpt)	Nôm	764	90.35	91.38

Figure 5: Comparison of BLEU scores on selected corpora be tween the two models



For the works **Dream of the Red Chamber** and **The Tale of Kiều**, both were included in the training data of the two models. However, due to phonetic ambiguity issues in the earlier model (as previously discussed), the BLEU scores remained suboptimal. The current model addresses this limitation, resulting in significantly im-proved scores.

In contrast, **Stories of the Saints**(excerpt) represents a new and unseen dataset for both models. While both systems generalize well, our improved model still achieves a slightly better BLEU score thanks to enhanced vocabulary coverage and augmented training data.

5.CONCLUSION

This study has effectively addressed the problem of automatic Sino-Nôm transliteration through a dual-architecture system that com bines an optimized BERT-based classification model with the sta-tistical machine translation framework Moses. Experimental results demonstrate a significant improvement over previous methods, achiev-ing a BLEU score of

93.90 on literary texts such as *Truyện Kiều*, and 94.60 on classical Sino texts. This success stems from the complete separation of processing pipelines for Sino and Nôm scripts, along with the application of a *masked mean pooling* mechanism, which raised classification accuracy to 98.45% and mitigated data imbal-ance commonly found in integrated models.

A notable real-world application of this system lies in its poten-tial for cultural heritage digitization: it facilitates public access to over 90% of Vietnam‘s historical documents from the 10^{th} to 19^{th} centuries. This approach holds not only linguistic value for Viet-namese but also suggests new directions for East Asian languages that adopt Han-derived scripts, such as Japanese (Kanji) and Korean (Hanja). From a technical perspective, the study contributes two key improvements: (1) incorporating an 80% *parallel* –20% *dictionary* ratio in the Moses phrase table to balance contextual learning and rare word coverage; and (2) rebuilding the BERT tokenizer based on a Sino-Nôm dictionary, which reduced the <unk> token rate by 37%, thereby enhancing the system’s ability to handle ancient characters.

The primary limitation of the system is its 23% performance drop when applied to domain-specific areas such as traditional medicine, indicating a strong dependency on training data specificity. More-over, the significant imbalance between the Sino (204,182 sentence pairs) and Nôm (only 49,775 pairs) datasets underscores the urgent need for future research in data augmentation and unsupervised learn-ing.

Future directions include: (i) integrating multi-layered OCR for handwritten manuscripts; (ii) extending the model to support multi-modal documents (combining text with illustrative images); and (iii) developing interactive interfaces that allow experts to refine system outputs through feedback.

This work represents not only a technological advancement in an cient language processing but also plays a pivotal role in preserving the linguistic heritage of Vietnam. With accuracy exceeding 90% on classical works, the system paves the way for building a comprehe-sive digital corpus, supporting interdisciplinary research in history, culture, and linguistics. This achievement highlights the transforma-tive potential of AI in the field of digital humanities, where technol-ogy and cultural heritage evolve in harmony.

6. ACKNOWLEDGEMENTS

The research is financially supported by Ho Chi Minh City De partment of Science and Technology.

REFERENCES

[1] Dang, H.-Q., Nguyen, D.-A., Pham, P.-P., Nguyen, N.-T., Chau, T., Ngo, D.-V., et al. (2022). NomNaOCR: The First Dataset for Optical Character Recognition on Han-Nom Script. *2022 RIVF International Conference on Computing and Com munication Technologies (RIVF)*, 476–481. https://doi.org/10.1109/RIVF55975. 2022.10013842

[2] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1, 4171–4186. https: //doi.org/10.18653/v1/N19-1423

[3] Dinh, D., Nguyen, T., & Le, M. (2021). Enhancing Nôm Transliteration with Dictionary Integration. *Vietnam Journal of Computer Science*, 8(2), 89–104.

[4] Viện Hán Nôm. (2002). *Từ điển Hán Nôm tích hợp*. Nhà xuất bản Khoa học Xã hội.

[5] Koehn, P., & Schroeder, J. (2007). Experiments in Domain Adaptation for Sta tistical Machine Translation. *Proceedings of the Second Workshop on Statistical Machine Translation*, 224–227. https://aclanthology.org/W07-0733

[6] Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., ... & Dyer, C. (2007). Moses: Open Source Toolkit for Statistical Machine Trans lation. *Proceedings of the 45th Annual Meeting of the ACL: Interactive Poster and Demonstration Sessions*, 177–180. https://aclanthology.org/P07-2045

[7] Lam, T., Dinh, D., & Nguyen, H. (2018). Hybrid SMT-NMT Architecture for Sino-Nôm Transliteration. *Proceedings of the 6th International Conference on Asian Language Processing*, 112–117.

[8] Papineni, K., Roukos, S., Ward, T., & Zhu, W. (2002). BLEU: A Method for Automatic Evaluation of Machine Translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 311–318. https://doi. org/10.3115/1073083.1073135

[9] Phan, X., Nguyen, K., & Bui, T. (2021). Domain-Specific Adaptation for Medical Text Processing. *IEEE Access*, 9, 134556–134567. https://aclanthology.org/W14- 3328/

[10] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Infor mation Processing Systems*, 30, 5998–6008. https://arxiv.org/abs/1706.03762

[11] Vo, L., Phan, T., & Tran, Q. (2023). Context-Aware Classification for Histori cal Vietnamese Scripts. *Journal of Computational Linguistics and Applications*, 15(3), 45–62.

APPLICATION OF AI IN THE AUTOMATIC CONSTRUCTION OF A SINO-NOM – VIETNAMESE NATIONAL SCRIPT PARALLEL CORPUS

Phuc Bui Hong[1], Dinh Si Dien[2], Nguyen Thuy Trang[3], Truong Thi Van Anh[4]

Abstract

Sino-Nom–Vietnamese parallel corpora are the primary training data for optical character recognition (OCR) and transliteration models on Sino-Nom texts. However, building such corpora remains difficult because OCR on Sino-Nom scripts is hindered by complex glyphs, degraded sources, and historical orthography. We present an AI-assisted method to automatically construct a Sino-Nom–Vietnamese parallel corpus by jointly exploiting semantic and visual similarity between Sino-Nom characters and their counterparts in Vietnamese (quốc ngữ). Concretely, we perform character-level alignment with Leven-shtein distance, using the Vietnamese line as a semantic anchor to recover the most plausible Sino-Nom character that matches in both meaning and appearance. This procedure corrects noisy OCR outputs and elevates the quality of digitized classical texts. The resulting pipeline improves OCR post-processing and enables scalable corpus construction for recognition and transliteration of Sino-Nom texts.

Index Terms: Artificial Intelligence (AI), Sino-Nom, Parallel Corpus, Levenshtein, Character Alignment, Natural Language Processing (NLP).

I. INTRODUCTION

Sino-Nom is a logographic writing system historically used in Vietnam to record literary, administrative, and religious texts from the 11th to early 20th century. Despite its deep cultural value and linguistic richness, processing Sino-Nom materials in the digital era remains challenging. Most his-torical documents exist only in physical form or as degraded scans, while expertise in reading and interpreting the script is declining rapidly. At the same time, the Vietnamese national script, based on Latin characters with diacritics, has become the modern standard for written communication and education.

Constructing a high-quality parallel corpus between Sino-Nom and Vietnamese is essential for both cultural preservation and modern NLP, and—critically—constitutes the primary training data for OCR and transliteration on Sino-Nom materials. Such corpora support not only dictionary building and translation studies but also enable supervised learning for tasks like grapheme-to-phoneme modeling, transliteration, or historical document retrieval. However, the creation of this corpus is hindered by multiple challenges: (1) input images are often noisy, skewed, or low-resolution; (2) OCR for Sino-Nom remains error-prone due to visual similarity among thousands of characters; and (3) alignment between the Sino-Nom source and Vietnamese transliterations is typically not available and must be inferred.

[1] University of Science, VNUHCM, Ho Chi Minh city, Emai: duyphuc2425@gmail.com

[2] *Corresponding author, CLC-Computational Linguistics Center, VNUHCM Ho Chi Minh city, Email: dinhsidien2008@gmail.com

[3] University of Science, VNUHCM Ho Chi Minh city, Email: 24C01023@student.hcmus.edu.vn

[4] University of Science, VNUHCM Ho Chi Minh city, Email: 24C01002@student.hcmus.edu.vn

A critical bottleneck is the acute shortage of curated training datasets for both OCR and Sino-Nom transliteration. Because most Vietnamese classical holdings survive only as images, any computational pipeline must first recognize characters; yet publicly available line- and character-aligned pairs are scarce, fragmented, inconsistently annotated, or restricted by insti-tutional access. Building such resources manually demands expert paleographic knowledge, meticulous proofreading, and time-consuming alignment, while the number of qualified Hán Nôm readers is rapidly diminishing. At current rates, purely manual curation cannot scale—motivating an automated approach to bootstrap reliable supervision from noisy OCR outputs and Vietnamese transliterations.

To address these issues, we propose a novel and practical pipeline for automatically constructing a character-aligned Sino- Nom–Vietnamese parallel corpus. Our contributions in-clude:

- A language-aware preprocessing stage that classifies image regions using OCR and applies YOLOv11-based line detection to crop meaningful text while removing decorative noise;
- A Sino-Nom correction algorithm combining visual similarity clusters and semantic alignment with Vietnamese transliterations using Levenshtein distance;
- A two-level alignment framework that operates at both line and character levels, accommodating OCR incon-sistencies and misalignments using a dynamic matching window;
- A human-inspectable, color-coded evaluation protocol to verify alignment quality, guiding error analysis and corpus refinement.

Beyond technical contributions, this work is driven by an urgent cultural need. A vast portion of Vietnamese history is documented solely in Sino-Nom, yet the number of experts capable of reading it is shrinking. Manual transliteration efforts, though accurate, are time-consuming and unable to keep pace with the rate of physical deterioration of ancient manuscripts. Without acceleration via automation, valuable cultural knowledge is at risk of being permanently lost.

Our pipeline provides a semi-automated solution that speeds up the digitization process while maintaining alignment accuracy. In the long term, such corpora can facilitate the development of AI systems—such as large language models (LLMs) and ChatGPT-style assistants—capable of understanding, querying, and translating Sino-Nom documents. By making these historical sources more accessible, we bridge the gap between traditional scholarship and modern computational tools, preserving cultural heritage for future generations.

II. Related Works

A. Levenshtein-based OCR Post-Correction for Historical Texts

Levenshtein distance remains one of the most reliable and interpretable measures for OCR post-correction. Srigiri and Saha [1] applied it in combination with word embeddings to correct OCR outputs for historical Hindi, reporting substantial improvements in character-level accuracy. Aditya Pal et al. [2] enhanced this approach by ranking candidates contextually using BERT embeddings and Levenshtein, achieving 81% accuracy on pre-modern texts. S Mihov et al. [3] introduced Levenshtein automata integrated with web-based dictionaries, optimizing both efficiency and precision. A recent survey by C Da et al. [4] highlights the adaptability and robustness of Levenshtein-based pipelines across multiple languages and historical scripts.

However, while these methods excel in Latin-based scripts, they struggle with ideographic languages with large character sets. Our work specifically addresses this by integrating visual similarity (glyph form clusters) and semantic anchoring from Vietnamese transliteration, tailored for the Sino-Nom script.

B. Google Vision OCR in Practical Applications

Google Vision OCR has emerged as a widely adopted solution for extracting textual data from scanned and degraded documents due to its support for multiple languages, ease of integration, and high accuracy in standard print environments. D Vaithiyanathan and M Muniraj [5] explored its use in assistive reading applications for visually impaired users, highlighting its ability to handle low-quality text in real-time scenarios. In a large-scale industrial setting, R Arief et al. [6] embedded Google Vision OCR into a Hadoop-based document processing pipeline, demonstrating its scalability and reliability for high-throughput workloads.

Moreover, NPT Prakisya et al. [7] conducted a comparative evaluation between Google Vision and open-source OCR engines such as Tesseract, showing that Google Vision consistently outperformed alternatives in both character-level and word-level accuracy—especially on documents with faded ink, skewed layouts, and historical typography. These findings confirm the robustness of Google Vision OCR in multilingual and real-world use cases, justifying its role in our pipeline for processing Vietnamese national script.

The original documents often contain a mixture of Sino-Nom and Vietnamese script, which presents challenges for automated processing. We trained a YOLOv11 [8] object detection model to automatically crop regions of interest, facilitating effective segmentation and noise reduction.

C. Sino-Nom OCR Using Kim Hán Nôm

Given the outstanding advantages of Google Vision, we continue to use this tool for Vietnamese OCR; however, general OCR engines like Tesseract and Google Vision perform poorly on Sino-Nom due to limited training on ideographic glyphs. Sino-Nom scripts pose unique challenges for OCR due to their ideographic structure and historical variability. Standard OCR systems like Google Vision and Tesseract are generally not trained on these

character sets. To address this, we integrate Kim Hán Nôm, an OCR system developed by the Computational Linguistics Center (CLC Lab) [9], specifically designed for recognizing classical Vietnamese characters. Trained on annotated corpora of stelae and manuscripts, Kim Hán Nôm integrates deep learning with stroke normalization and layout correction, achieving high accuracy in both printed and handwritten forms. Its utility extends to digital humanities, historical linguistics, and cultural preservation.

These studies collectively illustrate the progress in OCR extraction and post-correction for multilingual and historical documents. Our approach integrates general-purpose tools like Google Vision and YOLOv11 with specialized OCR systems such as Kim Hán Nôm, tailored for Sino-Nom scripts. Importantly, Levenshtein distance emerges as a core algorithm that remains highly relevant and effective in modern post-correction pipelines, particularly when combined with embedding-based or contextual models for ranking and filtering.

III. Methodology

The primary goal of this chapter is to present a structured approach for automatically constructing a high-quality parallel corpus between Sino-Nom texts and their Vietnamese national script translations. The methodology integrates OCR with noise reduction, script-aware segmentation, character-level correction based on Levenshtein distance, and automatic alignment techniques. It is designed to handle the unique challenges of historical documents, including visual degradation, multilingual script mixing, and inconsistent formatting. The full pipeline encompasses scan preprocessing, OCR processing, language-based filtering, semantic-aware correction, alignment, and corpus validation.

A. Preprocessing for Scan Photo

The first stage in constructing the Sino-Nom – Vietnamese parallel corpus involves preprocessing scanned historical documents to optimize them for OCR and subsequent text extraction. As shown in Fig. 1, this step includes language-aware processing and noise reduction through image cropping.

The first stage in constructing the Sino-Nom – Vietnamese parallel corpus involves preprocessing scanned historical documents to optimize them for OCR and subsequent text extraction. As shown in Fig. 1, this step includes both language-aware filtering and image denoising through region-based cropping.

We begin by applying the OCR capabilities of Google Vision to extract raw textual content from scanned PDF images. This tool supports multilingual scripts and provides bounding box information that can assist in later spatial structuring. However, due to the historical nature of the source documents, each scanned page often contains a mixture of Sino-Nom characters and Vietnamese national script, along with decorative frames, calligraphic headers, and marginal annotations that introduce significant noise (figure 2). In many cases, OCR models misinterpret these visual elements—especially borders and table

structures—as actual characters, distorting the output and complicating subsequent processing steps.

To address this challenge, we employed a two-stage filtering mechanism. First, we used the langdetect library to perform language classification on each OCR output segment. This step enables us to separate Sino-Nom and Vietnamese script blocks, allowing for tailored handling of each stream. Such separation is critical given the distinct tokenization rules, semantic structures, and error characteristics between the two scripts.

Second, we trained a YOLOv11 object detection model to automatically crop regions of interest from each scanned page. The model was trained on 600 manually annotated samples, each labeled to detect only the textual regions while ignoring non-informative elements such as ornate frames, floral decorations, or blank margins. This cropping process plays a pivotal role in reducing OCR noise. By excluding visual clutter, we prevent OCR engines from erroneously recognizing these artifacts as valid characters — a particularly common issue in historical Sino-Nom manuscripts, where framing elements can resemble stroke-dense ideographs.

In effect, YOLOv11 serves as a denoising gate that enhances OCR reliability for both script streams. Unlike traditional image pre-processing (e.g., thresholding or blurring), this object-detection approach enables spatially precise, content-aware filtering tailored to complex layouts. As a result, the output of this stage is a cleaner set of text segments—separated by script and stripped of visual noise—ready for OCR and semantic correction in the next phase.

B. OCR for Raw Text

Following the cropping and noise removal stage, we apply Optical Character Recognition (OCR) to extract raw text from both Sino-Nom and Vietnamese regions. Given the distinct script and linguistic properties of each, we apply different OCR tools tailored to their respective characteristics.

For Sino-Nom text regions, we employ KimHanNom, a specialized OCR engine developed by the Computational Linguistics Center (CLC Lab). Unlike generic OCR systems, KimHanNom is explicitly designed to handle the complex visual structure of classical Sino-Nom scripts. Its character set includes thousands of traditional and variant glyphs, enabling more accurate recognition of historical documents written in Sino-Nom. For Vietnamese regions, we continue to use Google Vision due to its superior performance on Latin-based scripts, especially in modern and well-printed formats. No post-processing was applied to the Vietnamese OCR stream, as recognition accuracy in this domain was sufficiently high.

However, even with a domain-specific engine like KimHan-Nom, the recognition of Sino-Nom characters remains susceptible to errors—particularly when source documents contain degraded ink, uncommon glyph variants, or closely overlapping strokes. To mitigate

these issues, we developed a character correction framework based on Levenshtein distance, leveraging curated visual and semantic associations across characters.

Correction Algorithm.

Given an OCR-recognized character c' from KimHanNom, we execute the following correction pipeline:

1) Visual Similarity Filtering: Retrieve a list of visually similar Sino-Nom characters to c', using a handcrafted dataset in which each character is linked to up to 20 lookalike forms (based on radical overlap or structural proximity).
2) Use the aligned Vietnamese transliteration to filter the candidate set. Specifically, we map Vietnamese syllables (e.g., "thọ") to their known Sino-Nom equivalents (e.g.,壽) using a bilingual dictionary built from historical lexicons.
3) Set Intersection:
   - If the intersection between the visually similar set and the semantically plausible set is empty, apply Levenshtein distance between c' and each visually similar candidate to select the nearest character.
   - If exactly one character is found in the intersection, it is used as the correction.
   - If multiple characters remain, we again apply Levenshtein distance within the intersection to choose the best match.

This approach allows us to recover from recognition errors by jointly leveraging visual structure similarity (form-based filtering) and semantic alignment (meaning-based disambiguation). By using Vietnamese national script as a semantic anchor, we are able to eliminate implausible candidates and enhance correction precision.

This correction mechanism forms a central component of our pipeline, substantially improving the Sino-Nom OCR stream's fidelity and enabling reliable downstream alignment.

Correction Dataset. This algorithm is powered by two structured resources:

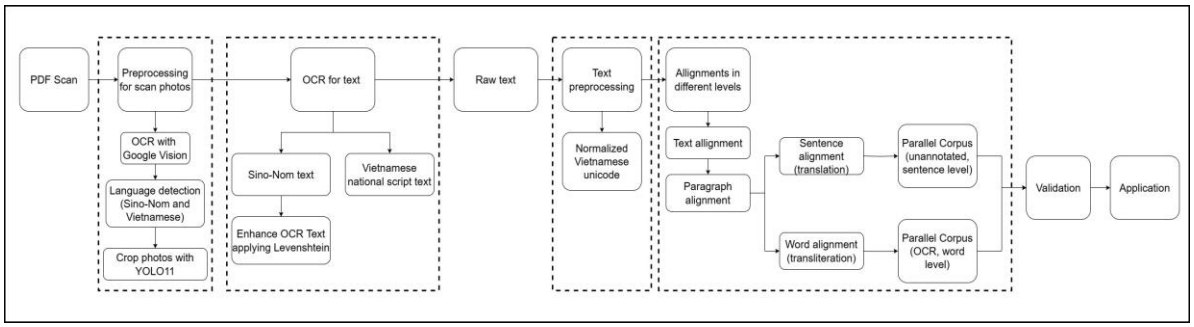


Figure 1. Overall pipeline for constructing a Sino-Nom – Vietnamese parallel corpus.

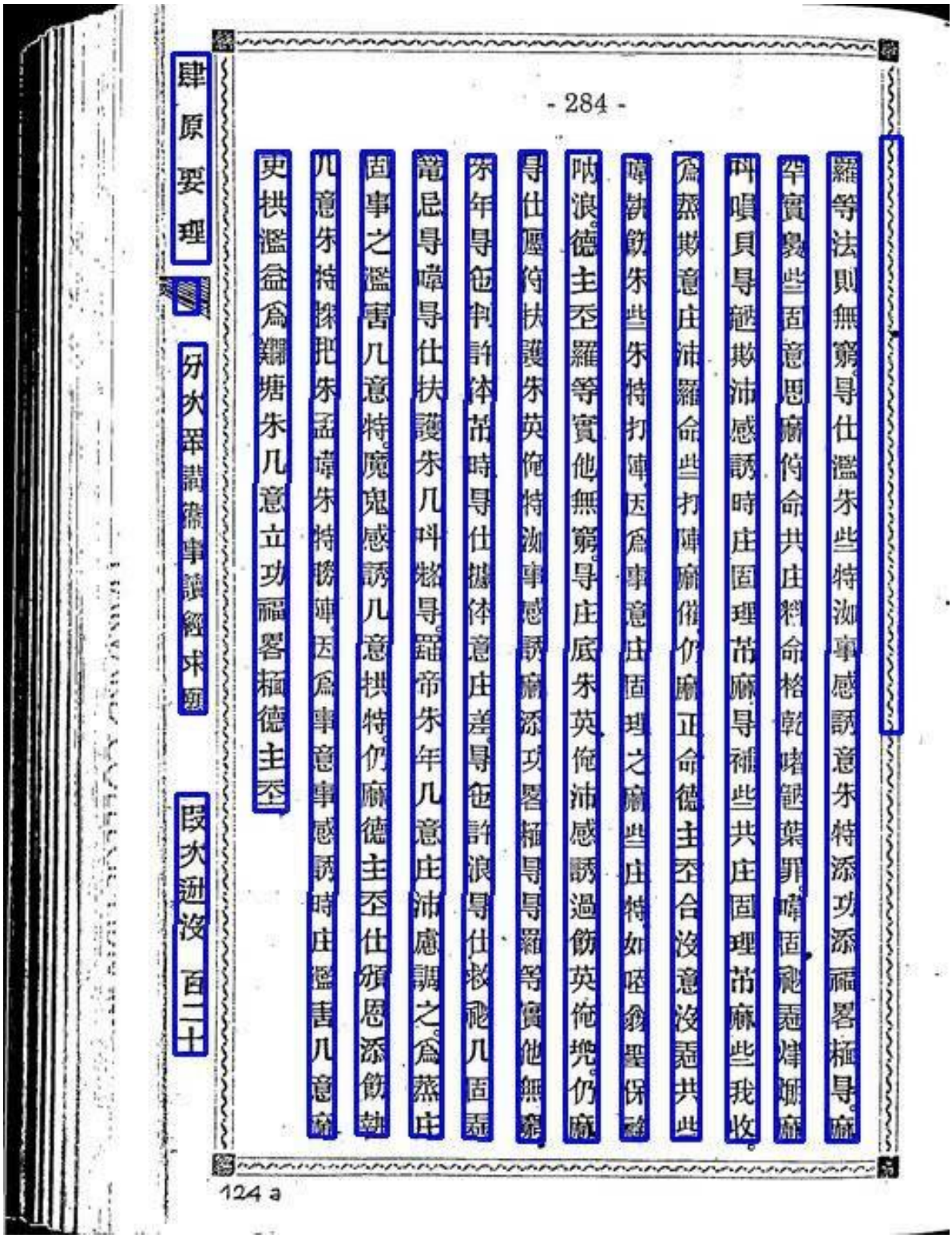


Figure 2. Example for OCR with full page

- A curated list of 26,044 Sino-Nom characters, each annotated with up to 20 visually similar glyphs. These were derived from traditional dictionaries and expert annotation. Figure 3 shows a sample of the dataset.
- A bilingual Vietnamese–Sino-Nom mapping dataset, compiled from multiple annotated corpora and historical references.

C. Text Preprocessing

After the OCR and character correction stages, the extracted textual data is organized into three distinct components: (1) the phonetic transliteration in Vietnamese script, (2) the original Sino-Nom text, and (3) the bounding box metadata corresponding to the Sino-Nom segments. Each component undergoes a dedicated preprocessing routine to prepare the data for alignment and corpus construction. While the system provides a generalizable and reproducible pipeline, certain steps remain adaptable to the structural variations commonly found in historical documents.

Vietnamese Transliteration Preprocessing. The transliterated Vietnamese layer requires careful normalization due to the complexity of the script's diacritic system and Unicode variability. Characters that appear visually identical may differ in their underlying Unicode representations, potentially causing misalignment. Accordingly, we normalize all Vietnamese text to Unicode NFC to ensure corpus-wide consistency.

Additional preprocessing steps include:

- **Removal of noisy characters:** Extra whitespace, malformed OCR tokens, and non-standard symbols are filtered out.
- **Masking of presentation elements:** Structural components such as headers, speaker annotations, or document-specific markers are temporarily masked to prevent interference with downstream parsing.
- **Punctuation removal:** All punctuation marks are stripped to yield a clean, phoneme-like sequence suitable for alignment.
- **Number normalization:** Numerical expressions (e.g., "123") are converted into their full Vietnamese verbal equivalents (e.g., "một trăm hai mươi ba") using a deterministic mapping module.
- **Preservation of phonetic content:** Only semantically relevant content is retained, enabling precise alignment with Sino- Nom characters.

The final transliteration stream is preserved as a continuous character sequence without sentence segmentation, as its primary role is to facilitate grapheme-to-phoneme modeling and character-level alignment.

Sino-Nom Text Handling. No additional preprocessing is performed on the Sino-Nom text after the character correction phase. The output from the KimHanNom OCR engine—refined using the Levenshtein-based correction algorithm—is considered final and directly usable for alignment. The text and its associated bounding boxes are preserved without modification to retain spatial fidelity and document structure.

Adaptability to Document Variants. While the preprocessing procedures above define a consistent and modular framework, the layout and formatting of historical texts are often highly variable. Depending on the document genre—imperial edicts, annotated records, petitions, or official decrees—further customization may be required. The pipeline is therefore designed to serve as a supportive baseline, allowing researchers to integrate domain-specific rules and manual refinements as needed. This balance of automation and flexibility ensures the pipeline can scale across a diverse range of classical sources while maintaining high alignment accuracy.

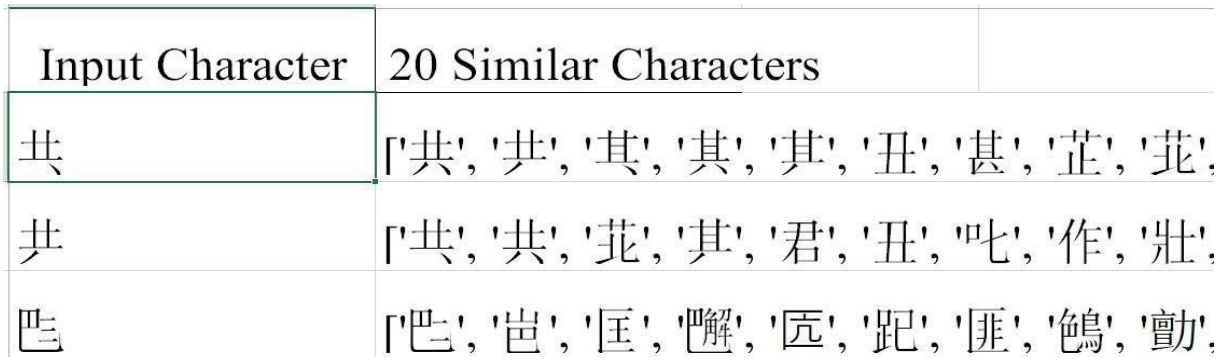
D. Alignment at Different Levels

The final stage of the pipeline constructs a transliteration-based parallel corpus by aligning Sino-Nom text with its corresponding phonetic rendering in Vietnamese. The alignment process operates at two levels of granularity: bounding box- level alignment and character-level alignment within each box. Unlike typical sentence- or paragraph-level alignment in bilingual corpora, this setting is character-centric and tightly constrained by the structure of the original scan layout.

Bounding Box Indexing. Prior to alignment, each OCR segment (bounding box) is assigned a unique index to maintain its spatial order. This index is computed based on the book title, page number, and the reading order of bounding boxes after spatial sorting (typically top-to-bottom, left-to-right). This indexing scheme is implemented during the OCR or preprocessing stage and ensures that segments from the Sino-Nom OCR output and the phonetic transliteration stream can be matched deterministically.

Box-Level Alignment. For each indexed bounding box, the Sino-Nom text extracted from OCR is paired with its corresponding transliteration segment. The assumption is that each box contains a linear sequence of Sino-Nom characters, and the transliteration provides a phonetic rendering in 1-to-1 character order. Therefore, we begin by aligning one Sino-Nom character to one Vietnamese syllable within the same bounding box.

Character-Level Alignment with Levenshtein Adjustment. In practice, the number of Sino-Nom characters in a box may not perfectly match the number of syllables in the transliteration due to OCR noise or slight formatting shifts. To handle this, we apply a flexible alignment strategy:



Input Character	20 Similar Characters
共	['共', '丵', '其', '其', '其', '丑', '甚', '芷', '芘',
共	['共', '共', '芘', '其', '君', '丑', '吒', '作', '壯',
匜	['匜', '岜', '匡', '匰', '匠', '跁', '匯', '鴾', '勫',

Figure 3. Sino-Nom similar visual dataset example

- Let n be the number of Sino-Nom characters in the bounding box.
- We allow a window of $n \pm k$ Vietnamese syllables (typically $k = 1, .., 5$) to be considered for alignment.
- We use Levenshtein algorithm at the character level to align Sino-Nom glyphs with the most plausible phonetic units in the Vietnamese stream.

If a character from the Sino-Nom OCR output has no suitable match in the transliteration window, an underscore "_" is inserted as a placeholder to preserve alignment structure. This ensures that each line of the corpus maintains 1-to-1 alignment, even in the presence of missing or noisy data.

Algorithm 1 LEVENSHTEINALIGNBOXES

Input: N = $\langle n_1,\ldots,n_m\rangle$ (Sino–Nom), Q = $\langle q_1,\ldots,q_n\rangle$ (Vietnamese), similar_df, trans_df **Output:** Aligned pair (AlignedN,AlignedQ) 1 $D_{\text{trans}}, D_{\text{sim}} \leftarrow$ BuildDicts(trans_df,similar_df) $m \leftarrow |\mathcal{N}|,\ n \leftarrow |\mathcal{Q}|\ \ dp \in \mathbb{N}_0^{(m+1)\times(n+1)};\ \ bt \in \{\uparrow,\leftarrow,\nwarrow,\emptyset\}_{(m+1)\times(n+1)}$

2 **Initialize boundaries for** $i \leftarrow 0$ **to** m **do**

3 $dp[i,0] \leftarrow i$; $bt[i,0] \leftarrow \uparrow$

4 **for** $j \leftarrow 0$ **to** n **do**

5 $dp[0,j] \leftarrow j$; $bt[0,j] \leftarrow \leftarrow$

6 $bt[0,0] \leftarrow \emptyset$

7 **Dynamic programming for** $i \leftarrow 1$ **to** m **do**

8 **for** $j \leftarrow 1$ **to** n **do**

9 $cost \leftarrow 0$ if Compatible($n_i, q_j, D_{\text{trans}}, D_{\text{sim}}$) else 1) $u \leftarrow dp[i-1,j]+1;\ \ l \leftarrow dp[i,j-1]+1;\ \ d \leftarrow dp[i-1,j-1] + cost$ ($dp[i,j]$,$bt[i,j]$) $\leftarrow$ argmin{(u,↑),(l,←),(d,↖)}

10 **Backtrace** AlignedN ← [ ], AlignedQ ← [ ], $i \leftarrow m, j \leftarrow n$ **while** $i > 0$ **or** $j > 0$ **do** 11 **if** $i > 0$ **and** $j > 0$ **and** $bt[i,j]$ =↖ **then**

else

12 push_front(AlignedN, n_i); push_front(AlignedQ, q_j); $i \leftarrow i - 1$; $j \leftarrow j - 1$

13 **else if** $i > 0$ **and** $bt[i,j] = \uparrow$ **then**

14 push_front(AlignedN, n_i); push_front(AlignedQ, $*$); $i \leftarrow i - 1$

15

16 push_front(AlignedN, $*$); push_front(AlignedQ, q_j); $j \leftarrow j - 1$

17 **return** (AlignedN,AlignedQ)

18 **Function** Compatible(n,q,D_{trans},D_{sim}):

19 $\text{trans}[n]$) **or** ($\exists c \in D_{\text{sim}}[n]: q \in D_{\text{trans}}[c]$)

return ($q \in D$

Final Output Format.

The result is a character-aligned transliteration corpus as showed in figure 4, where each line corresponds to a single bounding box. Each line contains a pair of aligned sequences:

Img_Box_ID	Img_Box_Coordinate	SinoNom_OCR	SinoNom_Char	ChuQN_txt
HVK_001.001.067.002.jpg	[[128, 77], [317, 71], [318, 99], [129, 105]]	煖唏祐炤動悉灄霪	煖唏祐炤動悉灄霪	Bén hơi rơm lửa động phòng mưa mây
HVK_001.001.067.003.jpg	[[150, 121], [301, 118], [302, 145], [151, 148]]	挴渃迢物巾秭	挴渃迢物巾秭	Vầy nước chậu vắt khăn tay
HVK_001.001.067.004.jpg	[[129, 150], [320, 148], [320, 176], [129, 178]]	欺啫踏襾欺哼鉼鍾	欺啫踏襾欺哼鉼鍾	Khi đêm đạp bóng khi ngày ngồi chung
HVK_001.001.067.005.jpg	[[144, 179], [306, 177], [307, 205], [144, 207]]	花桃包憚聰東	花桃包憚聰東	Hoa đào đà dạn gió đông

Figure 4. Overall pipeline for constructing a Sino-Nom – Vietnamese parallel corpus.

the corrected Sino-Nom characters and their aligned Vietnamese phonetic forms. This format enables both linguistic analysis (e.g., grapheme-to-phoneme modeling) and potential supervised training for future Sino-Nom OCR or transliteration system.

IV. Experiments

A. Experimental Setup

To evaluate the effectiveness of the proposed pipeline, we conducted experiments on three classical Vietnamese texts, each featuring Sino-Nom characters and corresponding transliterations in the Vietnamese national script:

- **Lives of the Saints (Sách Truyện Các Thánh)**
- **Essentials of Christian Doctrine (Tư Nguyên Yếu Lý)**
- **The History of Dai Nam in Verse (Đại Nam Quốc Sử Diễn Ca)**

All materials were sourced from publicly available online repositories. As such, the raw data consisted of low-resolution scanned images exhibiting various forms of degradation, including compression artifacts, uneven lighting, misaligned pages, and ink fading. This challenging input condition highlights the practical value and robustness of our preprocessing and correction pipeline, particularly in scenarios where high-quality digitization is not available.

The images were first classified into two categories—Sino-Nom and Vietnamese—using Google Vision's built-in language detection. Each group was then processed by a script-specific YOLOv11 model previously fine-tuned to detect line-level text regions in the respective writing system. The resulting bounding boxes (figure 5 were indexed by document metadata (book title, page number, and bounding box order) to preserve reading order and ensure deterministic alignment.

Text recognition was performed using two complementary OCR tools:

- **KimHanNom** for Sino-Nom characters, leveraging its specialized support for traditional glyph variants and historical typography.
- **Google Vision** for Vietnamese text, optimized for modern Latin-based scripts.

After OCR, character-level alignment was applied within each bounding box using the Levenshtein distance algorithm. A flexible alignment window of $n \pm k$ character window (typically $k = 5$) was used to handle minor inconsistencies in character count between the Sino-Nom and transliteration streams. Missing or unaligned characters were explicitly marked using an underscore ("_"), preserving alignment structure for further analysis. To assist human evaluation, we applied a color-coded visualization scheme (black, blue, red) to represent alignment confidence levels.

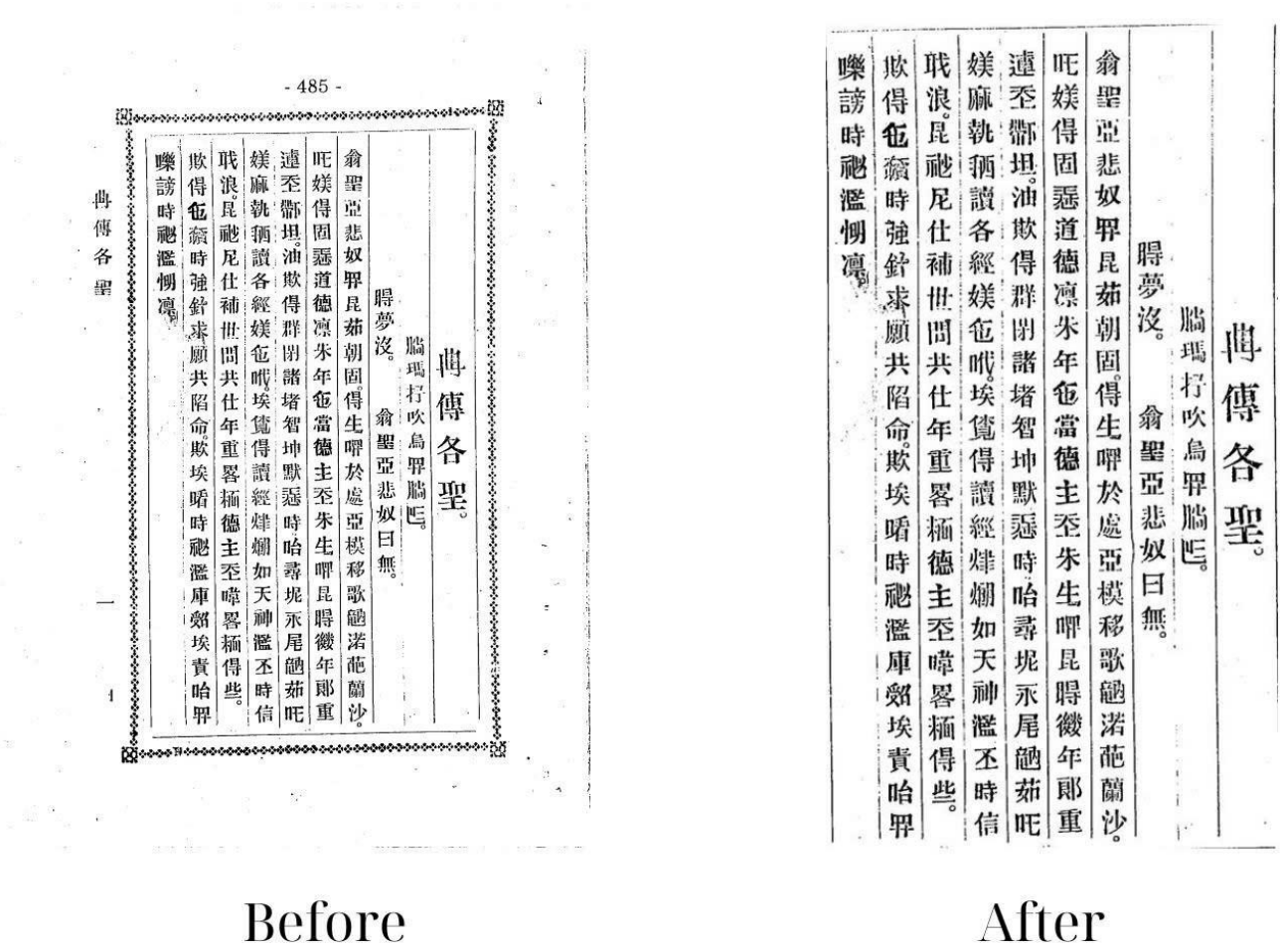


Figure 5. Result of crop image with YOLOv11

B. Evaluation Metrics

Color assignment rules:

- **Black**: OCR character exactly matches the dictionary `QuocNgu_SinoNom_Dic`.
- **Blue:** OCR character is approximately correct, found in `SinoNom_Similar_Dic_v2` with high Unicode similarity.
- **Red**: Incorrect character, not found in either dictionary.

Metric calculation:

Let:

- B: number of black (correct) characters
- G: number of blue (approximate match) characters
- R: number of red (incorrect) characters

- $T = B + G + R$: total number of evaluated characters

$$\text{Accuracy} = \frac{B + G}{T}, \qquad \text{Red Ratio} = \frac{R}{T} \quad \text{Green Ratio} = G$$

$$\text{Green Ratio} = \frac{B}{T}$$

C. Results

Table I

CHARACTER-LEVEL ALIGNMENT EVALUATION

Book Title	Red (%)	Blue (%)	Black (%)	Accuracy (%)
Sách truyện các thánh	15%	2%	83%	85%
Tư Nguyên Yếu Lý	20%	1%	79%	80%
Đại Nam Quốc Sử Diễn Ca	10%	1%	89%	90%

D. sAnalysis

The results demonstrate the robustness of the proposed pipeline:

- **High Black ratio (79–89%):** A majority of characters were correctly aligned one-to-one, indicating that the bounding box indexing and character-level alignment processes are both reliable.
- **Strong overall accuracy (80–90%):** Verified correctness of aligned character pairs shows that the combined OCR and alignment system is practical for real-world digitization tasks.
- **Low Blue ratio (<2%):** In cases where multiple candidate characters were found, the use of Levenshtein distance successfully selected the most plausible match.
- **Decreasing Red ratio:** The rate of unmatched characters (Red) decreased as the character reference library and alignment logic improved, indicating scalability of the matching system.

The coloring strategy not only enhanced transparency during post-alignment review but also provided a valuable visual aid for error analysis and future corpus refinement.

The results demonstrate the robustness of the proposed pipeline:

The coloring strategy not only enhanced transparency during post-alignment review but also provided a valuable visual aid for error analysis and future corpus refinement.

V. Limitation

Each processing step in our pipeline requires manual verification to ensure optimal results. Although automated cropping using the detection model achieves high accuracy in most cases, some bounding boxes are still incorrectly predicted, leading to incomplete or misaligned cropped images. These errors, albeit limited in number, must be manually corrected to preserve the integrity of subsequent steps.

Additionally, the alignment stage is particularly sensitive to missing or shifted bounding boxes. Even a single undetected region can propagate alignment errors, affecting the layout of an entire page. Therefore, it is essential to manually review and adjust results after each major step—especially cropping and alignment—to maintain high overall quality in the final OCR output.

VI. Conclusion

This paper presents a comprehensive pipeline for constructing a character-aligned parallel corpus between Sino-Nom texts and their Vietnamese transliterations. Our method addresses key challenges in historical document digitization—including image noise, OCR inaccuracies, and alignment complexity—through a combination of language-aware preprocessing, visual-semantic character correction, and Levenshtein-based alignment. The pipeline has been evaluated on real-world low-quality scanned texts, achieving over 80% character-level alignment accuracy across multiple books.

Beyond the technical results, this work contributes to the broader mission of preserving Vietnamese cultural heritage. As expertise in reading Sino-Nom declines and physical documents continue to deteriorate, automated tools become increasingly necessary to accelerate digitization and transliteration. Our approach not only supports current linguistic and historical research but also lays the groundwork for future integration with large language models (LLMs) and other AI applications in digital humanities.

In future work, we aim to expand the corpus coverage, incorporate more sophisticated neural-based correction models, and explore sentence- or paragraph-level alignment for full translation tasks. We also plan to release annotated corpora and tools to support further research in Sino-Nom processing and historical language understanding.

VII. Acknowledgements

The research is financially supported by Ho Chi Minh City Department of Science and Technology.

REFERENCES

[1] S. Srigiri and S. K. Saha, "Spelling correction of ocr-generated hindi text using word embedding and levenshtein distance," in *Nanoelectronics, Circuits and Communication Systems*, V. Nath and J. K. Mandal, Eds. Singapore: Springer Singapore, 2020, pp. 415–424.

[2] A. Pal and A. Mustafi, "Vartani spellcheck - automatic context-sensitive spelling correction of ocr-generated hindi text using BERT and levenshtein distance," *CoRR*, vol. abs/2012.07652, 2020. [Online]. Available: https://arxiv.org/abs/2012.07652

[3] S. Mihov, S. Koeva, C. Ringlstetter, K. U. Schulz, and C. Strohmaier, "Precise and efficient text correction using levenshtein automata, dynamic web dictionaries and optimized correction models," in *Proceedings of Workshop on International Proofing Tools and Language Technologies*, 2004.

[4] C. Da, P. Wang, and C. Yao, "Levenshtein ocr," in *Computer Vision – ECCV 2022*, S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, Eds. Cham: Springer Nature Switzerland, 2022, pp. 322–338.

[5] D. Vaithiyanathan and M. Muniraj, "Cloud based text extraction using google cloud vison for visually impaired applications," in *2019 11th international conference on advanced computing (ICoAC)*. IEEE, 2019, pp. 90–96.

[6] R. Arief, A. B. Mutiara, T. M. Kusuma *et al.*, "Automated extraction of large scale scanned document images using google vision ocr in apache hadoop environment," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 11, 2018.

[7] N. P. T. Prakisya, B. T. Kusmanto, and P. Hatta, "Comparative analysis of google vision ocr with tesseract on newspaper text recognition," *Media of Computer Science*, vol. 1, no. 1, pp. 31–46, 2024.

[8] G. Jocher, J. Qiu, and A. Chaurasia, "Ultralytics yolo," 2023, if you use this software, please cite it using the metadata from this file. Available at: https://github.com/ultralytics/ultralytics. [Online]. Available: https://ultralytics.com

[9] CLC Lab, "Kim hán nôm – automatic transliteration tool from sino-nôm to vietnamese script," 2022, online. [Online]. Available: https://tools.clc.hcmus.edu.vn

USING THE ENGLISH VIETNAMESE CORPUS OF CULTURAL TOURISM AS AN ESP MATERIAL

Phan Thi Thanh Thao[1]

Abstract:

To meet the rising demand for specialized English for Specific Purposes (ESP) materials in the field of cultural tourism, there are some existing resources, such as textbooks, magazines, and periodicals. Nevertheless, there is a lack of English Vietnamese corpora that can serve as a practical tool for vocabulary retention, language skill learning, and translation competence. This study presents the English Vietnamese bilingual corpus specifically designed for ESP Cultural Tourism and its use at English classes at the University of Foreign Languages and International Studies, Hue University (HUFLIS). The research was conducted in ESP learning in classes involving 150 students. With both qualitative and quantitative approaches using questionnaires and face-to-face interviews, the study presents the way to use this corpus and some benefits and challenges of its use. From the obtained results, it is suggested that the English Vietnamese bilingual corpus of Cultural Tourism should be used as an authentic material to improve students' ESP learning, translating/interpreting competence, and advanced technological knowledge as well.

Key words: English Vietnamese corpus, ESP material, ESP for cultural tourism, translating/ interpreting competence, students of English Faculty

1. Introduction

In recent years, the improvement of English for Specific Purposes (ESP) resources has been taken into great consideration in Vietnam's higher education system, especially in the universities prioritizing ESP teaching and learning in various disciplines including economics, tourism, medicine, and Information technology. Among these, tourism stands out as one of Vietnam's three key economic sectors, attracting substantial investment and contributing positively to the national economy. Consequently, teaching ESP of Tourism is critical in meeting the growing demand for qualified human resources, particularly in the Central Highlands region and in the whole country.

Despite some ESP materials available such as textbooks, magazines, journals, newspapers, and other digital and print resources, there is still a lack of a bilingual English Vietnamese corpus of cultural tourism, which can provide a useful and authentic materials for students in ESP learning and translation practice. Since the demand for such bilingual resources continues to rise, the creation of an English Vietnamese corpus tailored to cultural tourism becomes increasingly essential for ESP learners.

A review of previous studies on corpus-based language use points out both benefits and challenges in second language (L2) learning. Some corpora have been shown to develop learners' authentic language usage such as vocabulary and idiomatic expression retention in

[1] Faculty of English, University of Foreign Languages and International Studies, Hue University, *Email: pttthao@huflis.edu.vn*

real contexts, grammatical structures, and language style. This is particularly beneficial for ESP learners, who need to understand specialized language in real contexts. Nevertheless, the effective use of corpora also requires IT knowledge. In addition, some corpus examples can sometimes be ambiguous or unclear, hard to understand that requires learners to select carefully the language data in both their first and second languages.

Nonetheless, there is limited research on the construction of English Vietnamese bilingual corpora for cultural tourism, especially in Vietnam's universities. Addressing this gap, this study aims to (1) explore the needs of creating an English Vietnamese corpus of cultural tourism, (2) and examine how English-major students use the corpus as a reference in their ESP learning, as well as their attitudes toward the benefits and challenges of using the corpus. The study is guided by the following research questions:

1. What are the specific needs underlying the building of an English Vietnamese corpus of cultural tourism?
2. What are the perceived benefits and challenges of using this corpus in ESP learning?

2. Corpora

2.1 Definition

The term "Corpus" has been defined in various way in different domains. In linguistics, corpus definitions are typically based on their origin, size, structure, and function. In fact, the word "Corpus" derived from Latin, means "body" referring to a "body of texts" which can be digitally stored, retrieved, and analyzed (Oakes, 2012). A corpus is also defined as "a representative collection of pieces of language selected according to linguistic criteria and reflecting natural chunks of language for the purpose of analysis" by many linguists (McEnery et al., 2006; Sinclair, 2004; Tognini-Bonelli, 2001). In other words, this definition implies the integral relationship between corporal and linguistic features.

In applied linguistics, a corpus also refers to large, structured sets of either spoken or written texts, which are essential to enable systematic investigation into language usage, structure, and patterns. Patsala and Michali (2020) emphasize the representativeness of corpora of language samples, while their machine-readable format allows the users to annotate and analyze in different criteria including part-of-speech tagging, frequency counts, and keyword-in-context (KWIC) and so on (Rayon, 2015).

Depending on the purpose of corpus use, the size and scope of a corpus can be varied. To capture a huge amount of lexical and grammatical patterns, it requires to build corpora containing millions of words (Sinclair, 2004). For instance, the British National Corpus (BNC) comprises over 100 million words from a wide range of spoken and written sources, which can be used effectively in many domains. Furthermore, the development in computational linguistics has increased the number of publicly available, linguistically annotated corpora, providing the large-scale language analysis (Kuble & Zinsmeister, 2015).

In natural language processing (NLP) and machine learning, it is essential to use corpora for processing language automatically by labeling and recognizing language. In machine translation, bilingual corpora can be used to derive translation rules to improve automated translation in similar domains, styles, and genres (Zanettin, 2014). Furthermore, other applications of corpora in computational linguistics include speech recognition and speech synthesis. Finally, in education, corpus plays a crucial role in developing teaching materials, compiling dictionaries, designing exam questions, providing authentic samples in translation and language learning, particularly reading and writing skills.

2.2 Corpus Building and Analysis

The development of a corpus plays a pivotal role in fulfilling users' needs for specific purposes. As Ngula (2017, pp. 210–211) notes, corpus design entails systematic planning, decision-making, and the establishment of a framework to guide the collection and processing of documents. Prior to data collection, corpus builders must address key questions related to the types of documents, intended size and text sampled used for the corpus. In fact, the corpus construction should meet several criteria, such as size, sampling, balance, representativeness, and compliance with rights and copyright regulations. Biber (1993, p. 243) defines representativeness as the extent to which a sample reflects the full range of variation within a population. Balance, by contrast, refers to the inclusion of a comprehensive range of genres or text types. Sampling is the method by which excerpts, or complete texts are selected for inclusion. Baker (2010, p. 96) emphasizes that a corpus must be representative of a specific language, thus its texts must be sampled with various topics. Although copyright and intellectual property rights are considered in corpus design, their practical application in corpus construction remains a matter of debate among linguists.

In this study, the English Vietnamese corpus on cultural tourism was developed using the cloud-based Computer-Assisted Translation (CAT) platform namely Smartcat. Smartcat is widely recognized as a leading CAT tool that integrates professional translators, language service providers, and business organizations within a collaborative environment (Smartcat, 2024). According to Ariyaratne (2019), Smartcat is a market leader in the CAT field, while Malenova (2019) observes that its intuitive functionality and user-friendly interface make it particularly well-suited for teaching translation technology from foundational to advanced levels.

Secondly, corpus analysis tools have significantly enhanced the capacity to process and analyze large datasets efficiently. Advances in corpus analysis software now allow for rapid and accurate examination of corpora. As Baker (2010, p. 102) observes, a standalone corpus offers limited value for linguistic analysis; rather, corpora are typically utilized alongside software capable of counting, sorting, and presenting linguistic features. Hunston (2006, p. 234) outlines the core functions of corpus software, including: (i) searching the corpus for specific purposes; (ii) calculating the frequency of occurrences (or equivalent translations) of

target items and their relative frequencies; and (iii) displaying these occurrences to enable further investigation.

A wide range of corpus analysis tools are currently available, each offering distinct functionalities. Notable examples include WordSmith, AntConc, ConcGram, and Sketch Engine, all of which allow users to upload corpora for analysis. According to Ngula (2017), corpus analysis, though often viewed as a single methodology, encompasses multiple analytical techniques, such as frequency lists, keyword extraction, concordance, cluster and n-gram analysis, and collocate identification.

Sketch Engine software was employed due to its comprehensive capabilities in building and analyzing large corpora. Recognized as one of the most powerful corpus analysis platforms, Sketch Engine supports the development of extensive and representative corpora across more than 90 languages. As of 2023, it offers access to approximately 500 ready-to-use corpora containing a combined total of around 50 billion words, thereby providing a robust and representative sample for linguistic research (Sketch Engine, 2023).

2.3 The Role of Corpora in Second Language Teaching and Learning

Bringing corpus linguistics into language education has greatly changed the way second languages are taught in the digital age." (Alqadoumi, 2013; Bennett, 2010; Boulton, 2009, 2012). As a result, language instruction is no longer confined to traditional materials but increasingly incorporates data-driven learning supported by corpus-based tools. The corpus use offers numerous advantages in L2 teaching and learning, which influences not only the design of syllabus and learning/teaching materials but also learners' language skills, cultural knowledge and abilities of research autonomy.

One key advantage of corpus uses in L2 teaching relates to its ability to enhance syllabus and materials development. Corpora provide authentic, specific language examples that help create lesson contents. As Alcantar and Jose (as cited in Carlota de Jesue, 2021) suggest, corpora give a variety of grammar and vocabulary patterns which are not easily identified without computational tools in teaching materials. Take an example, corpus-based studies on modal verbs (e.g., *would, can, might*) have shown that textbooks often lack complete or accurate representations of modal usage, which can limit or mislead some explanations in L2 teaching (McEnery et al., 2006; Carlota de Jesue, 2021; Carter, 2003).

Additionally, a corpus plays an important role in raising L2 learners' language awareness. According to Carter (2003), language awareness involves learners' ability to reflect their understanding and application of language forms and functions in real context. Via corpus use, learners can independently observe lexical and grammatical patterns in authentic contexts. Van Lier (2001) also gives an example of language awareness in teaching activities presenting the way the language is focused on the learners, not the teachers. When using a corpus, learners easily search for vocabulary with its concordances which can be used more correctly and

naturally. Moreover, corpora analysis allows learners to identify common lexical and grammatical patterns, enabling them to learn language structures more efficiently.

Another important benefit of corpus use is the development of L2 learner and teacher's autonomy. Corpus-based learning encourages students to improve their own language development by searching for and memorizing linguistic structures, vocabulary, and grammatical patterns. For learners, this approach assists them to have self-directed learning and confidence in language use and analysis. For teachers, the corpus use enables them to supplement lessons with real-life language data enhancing them to address learners' specific needs (Carter, 2003).

In addition, corpora can help learners understand cultural nuances which are embedded in authentic language use. The perceived knowledge of idiomatic expressions, cultural references, and meanings in real contexts fosters learners' intercultural competence which is not presented traditional teaching materials. As learners have abilities of exposing real-world texts in specialized corpora, they find the learning process more relevant to their academic concerns and targets. What is more, corpora are regarded as valuable tools for developing research skills. L2 students can learn to identify patterns, figure out linguistic features, and apply critical thinking into their study by analyzing corpus data. This can offer a deeper understanding of language and supports them in conducting independently linguistic research.

Despite those advantages mentioned above, there are several challenges which may impede the effective use of corpora in the classroom. First, both learners and teachers are required to obtain certain knowledge of digital technology to navigate corpus tools efficiently. They must spend much time familiarizing themselves with concordances and corpus platforms, which may be time-consuming, particularly for beginners. Second, building corpora of different functions and purposes also demands interdisciplinary expertise, involving specialists in corpus linguistics, natural language processing, and computational linguistics and applied linguistics. Third, the usage of some tools to create corpora like Sketch Engine has limited access and high payment if that may also disrupt corpus use in some settings.

In summary, while there are practical challenges in corpus use, particularly related to information technological skills and access techniques, benefits of corpus-based language learning/teaching remain substantial. Corpora offer a dynamic, authentic, and independent approach to language learning and teaching that fosters profounder linguistic insight and greater autonomy. As a result, corpus use is becoming increasingly popular in language education in various specialized domains.

2.4 Previous Studies

In the corpus linguistics field, building a corpus has been of necessity to meet the needs of language analysis within specific domains. Since the first modern corpus, namely the Brown Corpus, appeared in 1961, numerous corpora have been developed to support linguistic research. First, the Brown Corpus, consisting of approximately one million words, was among

the earliest to be annotated for word morphology. Another corpus called the Lancaster/Oslo-Bergen (LOB) Corpus was built with around one million words included sub-corpora such as the Leeds-Lancaster Treebank and the Lancaster Parsed Corpus. Both corpora can be used to annotate syntactic structures. In addition, The British National Corpus (BNC), one of the biggest corpora, was constructed with 100 million words from a wide range of spoken and written sources. Simultaneously, the Corpus of Contemporary American English (COCA), with over one billion words, was designed and became one of the largest and most diverse corpora in the world. It includes many texts from different genres like film scripts, spoken dialogues, fiction, journalism, and academic writing, providing rich data for linguistic and NLP applications.

In the context of conducting Vietnamese English language research, they have made great efforts to develop bilingual corpora to support machine translation, natural language processing (NLP), and English language education. Quynh's (2011) master's thesis titled *Research on Building a Bilingual Corpus for Vietnamese Language Processing* is one of studies on building bilingual corpora for English Vietnamese language pair. This corpus contains English Vietnamese sentences extracted from various sources such as websites, dictionaries, books, and newspapers in XML, TXT, and DOC format. The study focused on corpus construction to serve as a database for automatic translation systems and a useful resource for language learners and NLP applications.

Similarly, Ngo and Winiwarter (2012) conducted a study titled *Building a Bilingual English Vietnamese Corpus for Machine Translation*, which aimed to develop a parallel corpus for training machine translation systems. Their research got involved in techniques for data collection, sentence alignment, bilingual annotation, and the development of manual annotation tools. This corpus contained over 800,000 sentence pairs and more than 10 million English and Vietnamese words. Of these, over 45,000 sentence pairs were aligned at the word level, significantly enhancing the corpus use for machine translation and language processing.

These studies have contributed significantly to the advancement of corpus-based methodologies in Vietnam. They have also laid the foundation for the development of more refined tools and improved methods for building English Vietnamese parallel corpora. Their contributions have enhanced the quality and efficiency of bilingual corpus design and have fostered more applications in machine translation, computational linguistics, and ESP learning and teaching.

3. Research Methodology

3.1 Introduction of the E-V corpus of Cultural Tourism

This study is conducted with English Vietnamese corpus of Cultural Tourism containing about one million words collected from written documents (newspapers, articles, books, brochures, textbooks, advertisements, etc.) in the cultural tourism domain. Figure 1 below introduces our corpus which is used to analyze the linguistic features based on different

functions including Word Sketch, Word Sketch Difference, Thesaurus, concordance, Parallel Concordance, N-gram, Wordlist, Keywords, Bilingual Term, and so forth.

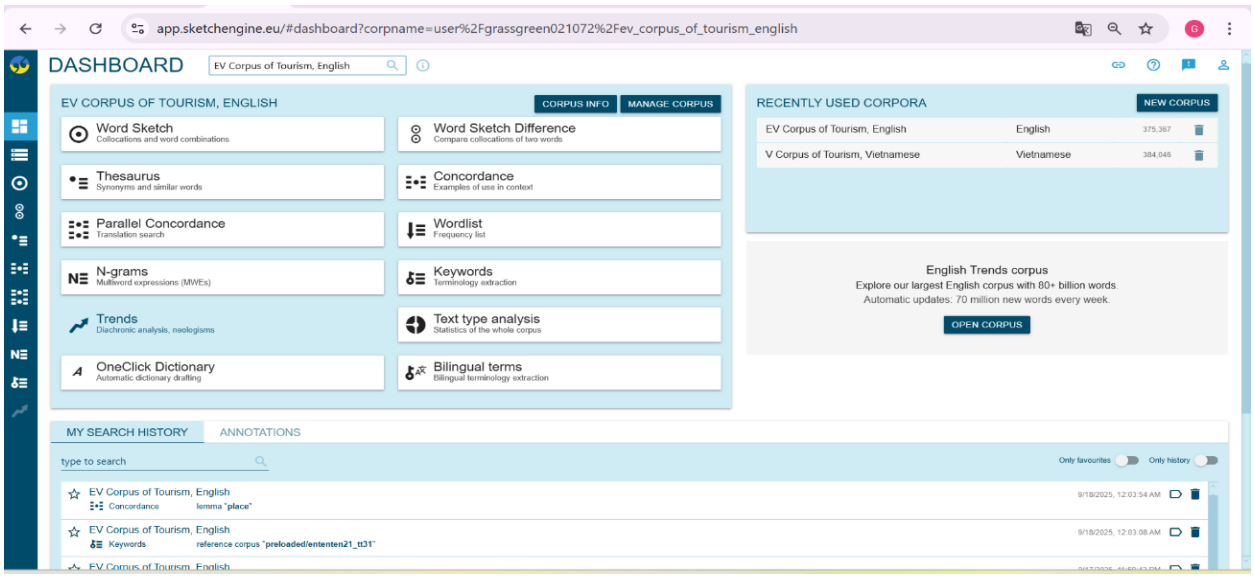


Fig. 1 Introduction of English and Vietnamese corpora of Tourism

3.2 Corpus experiment

3.2.1 Participants

The participants in this study include 150 students from the Faculty of English, University of Foreign Languages and International Studies, Hue University, who were enrolled in English courses on Tourism and Interpreting & Translating during the second semester of the 2024-2025 academic year.

3.2.2 Research Instruments for Data Collection and Analysis

Data was collected through two primary instruments. First, a questionnaire consisting of 50 statements was used to measure participants' understanding of corpus linguistics and their attitudes towards the use of corpora in English for Specific Purposes (ESP) learning. Second, follow-up interviews were conducted with 20 students using 10 in-depth questions to obtain detailed insights and clarify findings from the questionnaire.

The questionnaire results were analyzed using descriptive statistics and shown in tables with a five-point Likert scale: 1 = Strongly disagree (M = 1.00-1.80), 2 = Disagree (M = 1.81-2.60), 3 = Neutral (M = 2.61-3.40), 4 = Agree (M = 3.41-4.20), and 5 = Strongly agree (M = 4.21-5.00). To measure and analyze the data, we have used a mixed method in combination with both quantitative and qualitative approaches.

4. Results and discussion

4.1Building and Application of the English Vietnamese Corpus of Cultural Tourism

We have built English Vietnamese corpus of cultural tourism comprising approximately one million words, which can be used as an authentic resource to help students in ESP learning,

particularly in the domains of tourism, translation, and interpretation. This corpus design aims at learning vocabulary, analyzing syntactic structures, and comparing the linguistics features of both English and Vietnamese, and making it a practical reference tool for language study.

The corpus allows users to perform a range of linguistic queries in analyzing word frequency, part-of-speech categories, and grammatical patterns. For example, students can explore word frequency with different lexical categories such as nouns, verbs, adjectives, adverbs, conjunctions, prepositions, pronouns, numerals, and determiners. Using search options such as "all starting with," "ending with," "containing," or "matching regex," learners can gain frequency data from the corpus.

An example illustrating this function can be seen in the frequency analysis of common function words. As shown in Figure 3, "*the*" article *appears* 24,588 times, while *a* occurs 8,367 times. Similarly, the conjunction *and* appears 12,088 times, the preposition *of* 11,592 times, and *to* 10,608 times. These results demonstrate how frequently occurring words can be identified and analyzed within the corpus to assist students to understand language patterns in authentic contexts.

Such corpus-based activities encourage students to improve their real language use by seeking for collocations, phrases and grammar structures which appear frequently in real documents relating to cultural tourism. If students like to explore the concordance of phrases containing the word "holiday" in English corpus of tourism, they can find these phrases appearing 168 times with 367.65 per million tokens, accounting for 0.03%. Figure 2 below indicates 17 examples of "holiday" phrases.

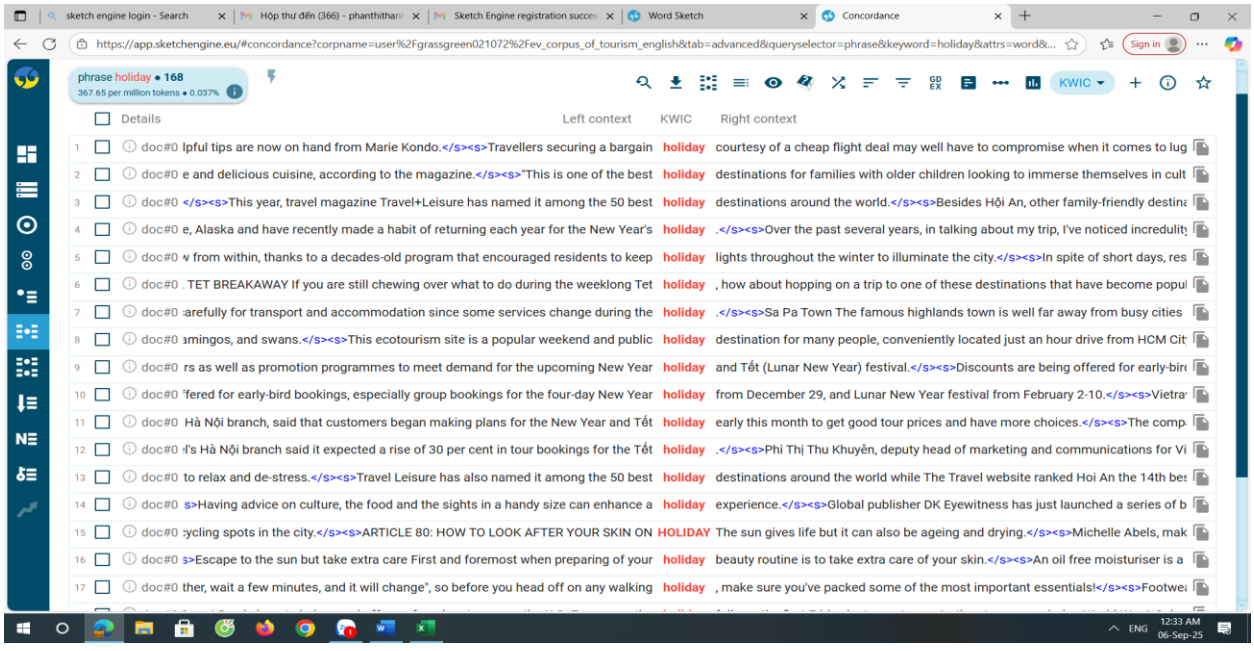


Fig. 2 Concordance of phrases containing the word "*holiday*" in the English corpus of tourism

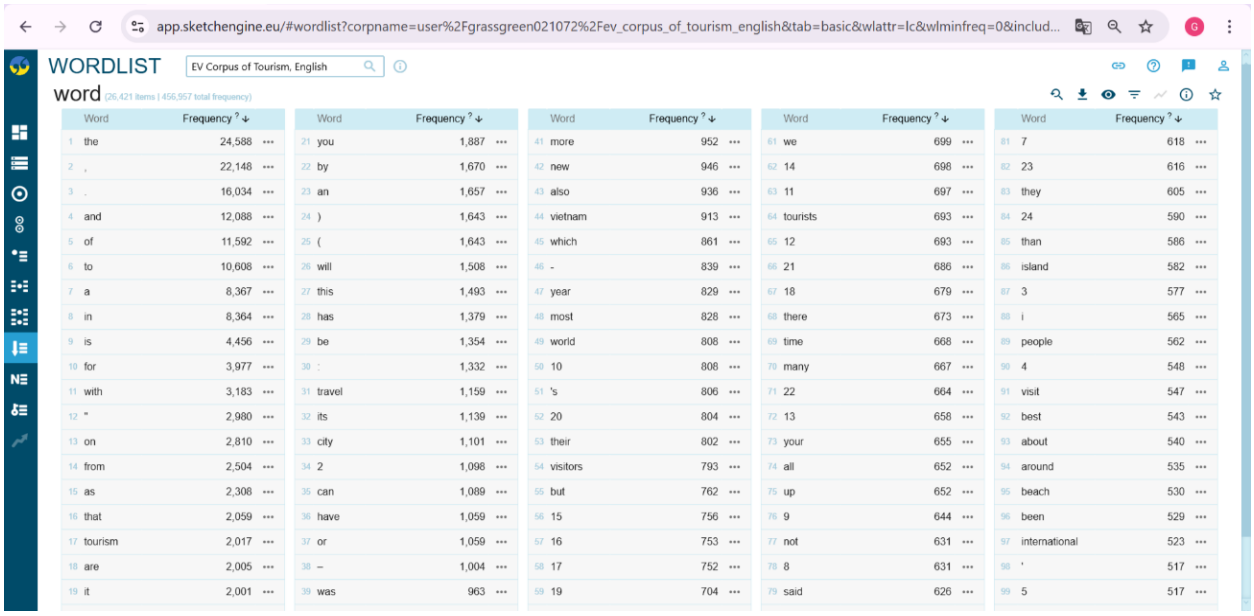


Fig. 3 Frequency of word list in the English corpus of tourism

Furthermore, with the function of listing the thesaurus of any words or phrases in the corpus, students can find out many the relevant words. Figure 4 demonstrates the synonyms of the word "city" existing in the English corpus of tourism. When students find the thesaurus of the word "city", the corpus offers the list of words including "destination" (786), "city" (402), "places (679), "town"(285), "region"(297), "site" (338), and so on.

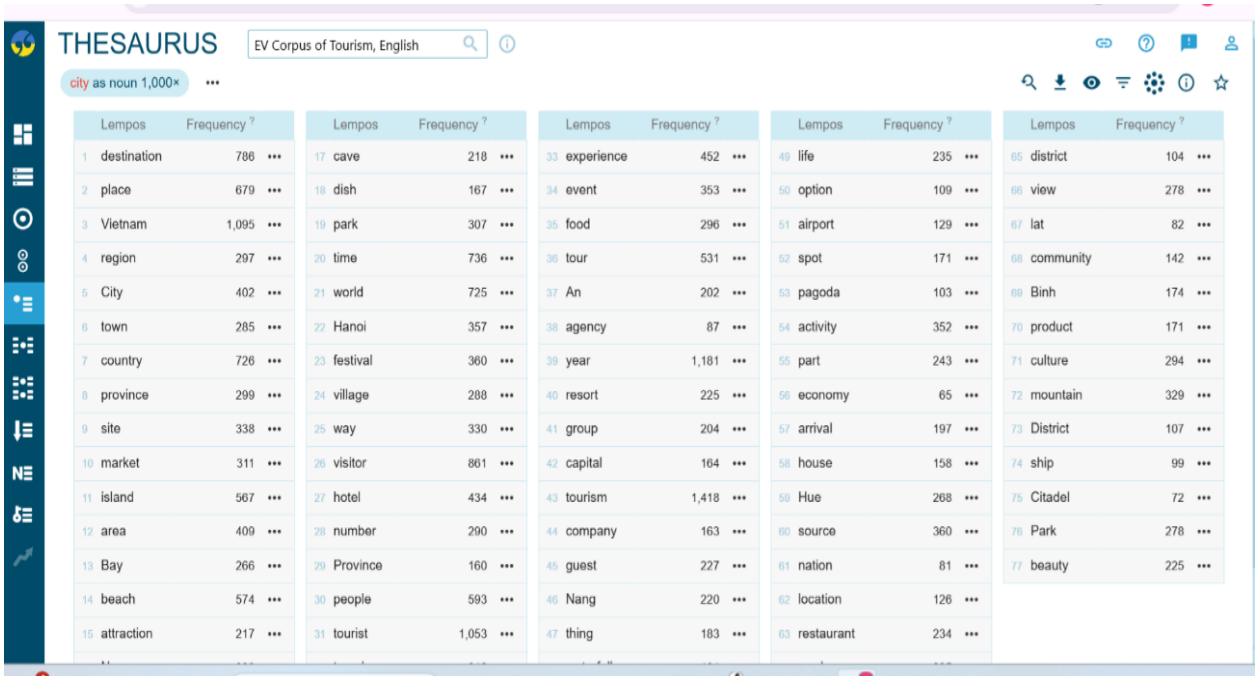


Fig. 4 Thesaurus Frequency of the word "city" in the English corpus of tourism

Additionally, function Word Sketch can help students find objects, subjects, modifiers, prepositional phrases, and particles after those words. Take an example of the verb "*take*"

shown in Figure 5, we can obtain 762 examples illustrating modifiers of "take" such as roughly, about, nearly, only, easily, etc.; objects of "*take*" include place, photo, advantage, trip, hour, tour, step, part, time, test, … , subjects of "*take*" consists of flight, activity, aircraft, field,…These are illustrated with examples, which makes it easier for students to learn by heart the vocabulary and grammatical points.

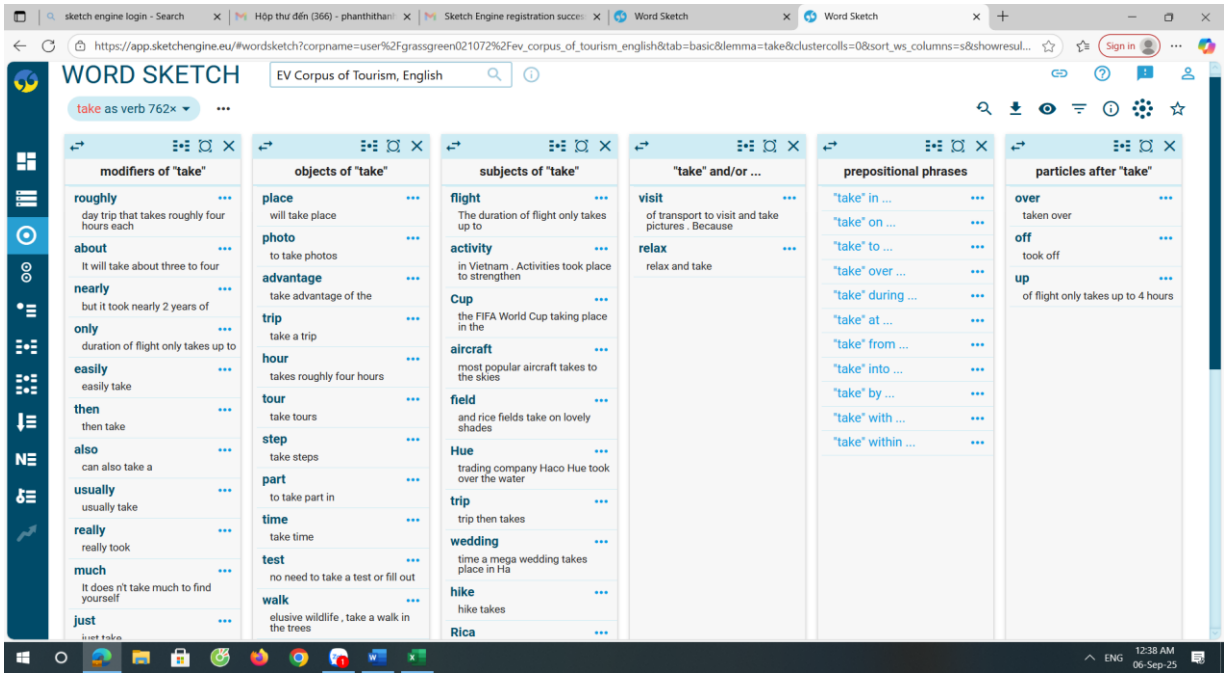


Fig. 5 Word Sketch for the verb "*take*" in the English corpus of tourism

Furthermore, students can select the function N-grams to search for the frequency of phrases containing 2,3,4,5,6 or N words in the corpus. Figure 6 presents 6-grams of the corpus with their frequencies, e.g. *the Vietnam National Administration of Tourism* (13); *Global Code of Ethic for Tourism* (11), *the central province of Quang Nam* (10), *as one of the most popular* (9), *day of the Lunar year (9)*, etc.

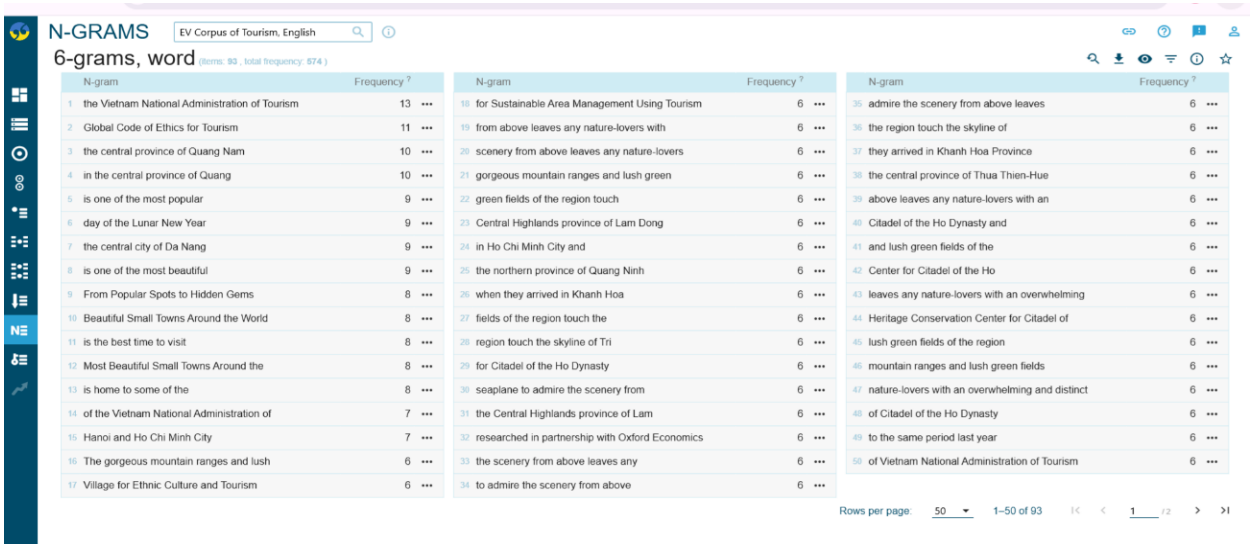


Fig. 6 Frequency of 6-grams in the English corpus of cultural tourism

In addition, students who are interested in translation domain can select the function of Parallel Concordance to search for some alignments including sentences, paragraphs, titles, headings, etc. in both languages English and Vietnamese. Figure 7 indicates the English texts and their translated Vietnamese versions after the corpus user selects the Parallel concordance function.

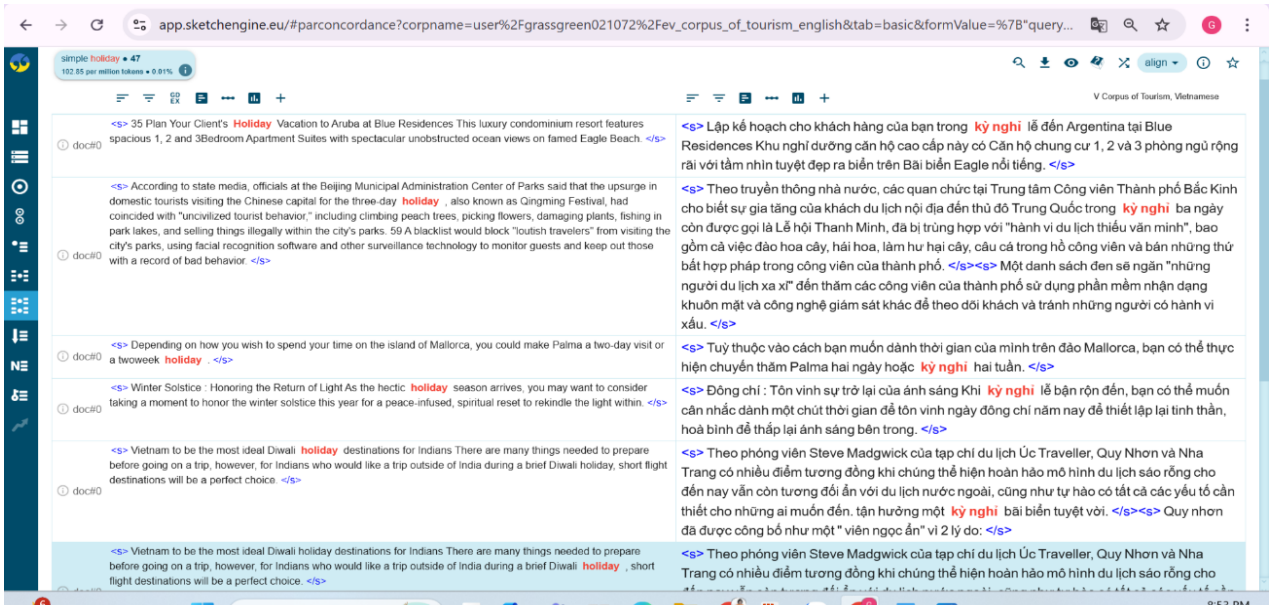


Fig. 7 An example of Parallel concordance from the E-V bilingual corpus of tourism

During the corpus building process, we have also created an English Vietnamese glossary of tourism including 3,250 words. This glossary provides the users the terms or phrases relating to the Tourism topic that has been appreciated helpful and effective for students to learn English for tourism.

4.2 Students' Practice of Corpus Use

To further investigate students' opinions on the advantages and disadvantages of using bilingual English Vietnamese corpus of tourism in learning ESP of tourism, we have experimented this corpus for 150 students from English Faculty, HUFLIS during the second semester of the academic year 2023-2024. The study results have been achieved from 149 respondents as follows:

4.2.1 Advantages of Using the English Vietnamese Bilingual Corpus in ESP for Tourism

To examine students' perceptions of the advantages and disadvantages of using the English Vietnamese bilingual corpus of tourism in English for Specific Purposes (ESP) learning, an experimental study was conducted with 150 undergraduate students from the Faculty of English at HUFLIS during the second semester of the 2023–2024 academic year. A total of 149 responses were collected and analyzed.

Vocabulary Retention

From Table 1, the most benefits of corpus use is related to its effectiveness in helping vocabulary retention, particularly for words specializing in tourism domain. The mean score for this aspect is M = 3.85, which shows students' strong agreement that the corpus offers an easy access to specialized vocabulary in both English and Vietnamese, without teachers' assistance. In addition, students reported that the search functions were simple to use. It is noticeable that the same mean score (M = 3.86) was stated the corpus was very useful for students to build independently glossaries of tourism relevant terms. Finally, the mean score (M = 3.98) was rather high reflecting that the corpus remarkedly improved student's ability to retain specialized vocabulary efficiently, which is essential for their ESP learning.

Translation Practice

Students also reported that the corpus contributed meaningfully to their translation skills, particularly when working with domain-specific texts in tourism. The corpus with one million words is a valuable resource for students to compare English and Vietnamese structures and identify corresponding terms, supporting to create more accurate and appropriate translations. The mean score was M = 3.82, demonstrating a high level of agreement on the corpus's practical value in translation tasks. Furthermore, the tool can save time in translation preparation when students are able to have quick access to terms and structures translated previously. With a mean score of M = 3.95, students affirmed that the corpus could help enhance their translation by reinforcing their ESP knowledge of cultural tourism in both languages.

Technological development

In terms of technological competence, the corpus was also seen as a valuable tool for improving students' technological skills. A high proportion of students (M = 3.89) agreed that working with the corpus via Sketch Engine helped them better understand and use digital tools for language learning. The findings from interviews reveal that 7 out of 10 students asked to have achieved some improvements in their IT skills. For example, one participant (ST5) remarked: "I had never known what Sketch Engine was like before, but after using the English Vietnamese corpus of cultural tourism, I can gain better knowledge of the software and felt excited to apply it in language learning and translation."

Similarly, ST8 shared: "Using this corpus is really useful because it helps me learn a lot of IT knowledge. For example, through Sketch Engine, I learned how to analyze linguistic features, explore structure and function, build a corpus, and compare it with others."

Reliability of Responses

As shown in Table 1, the standard deviations for the measured items ranged from 0.628 to 0.812, all below 1.0. These low standard deviations indicate a high level of consistency in student responses, which confirms the reliability of the data. Overall, findings show that the English Vietnamese corpus of cultural tourism has provided students with benefits in ESP learning, especially in vocabulary development, translation skills, and technological development.

No	Advantages	**T**otal	**M**ean	**S**tandard deviation
1	Using this bilingual corpus on Tourism helps me easily search for specialized terms in both languages without the teacher's guidance.	N= 149	3.85	.635
2	The technique of searching for vocabulary becomes easy when I use the bilingual corpus.	N= 149	3.95	.628
3	Using this bilingual corpus helps me translate texts better in looking up the structures of the two languages.	N= 149	3.81	.659
4	Using this bilingual corpus helps me to compile statistics of terms to build glossaries more easily.	N= 149	3.85	.812
5	Using this bilingual corpus helps me easily access idioms or phrases with electronic devices such as computers, smartphones, or tablets.	N= 149	3.88	.671
6	Using this bilingual corpus on Tourism Culture helps me save time in memorizing specialized vocabulary in both languages	N= 149	3.94	.662
7	My knowledge of the Tourism Culture has been improved after using this bilingual corpus.	N= 149	3.93	.640
8	My IT knowledge has improved when I regularly use the bilingual corpus.	N= 149	3.92	.728
9	My vocabulary of Tourism Culture of both languages has been increased significantly after this corpus use.	N= 149	3.93	.660
10	Using the corpus helps me learn Tourism Culture vocabulary as well as language knowledge faster and more effectively.	N= 149	3.97	.698

Table 1. Advantages of the corpus use as an ESP material

4.2.2 Disadvantages of Using the English Vietnamese Corpus on Cultural Tourism

While the English Vietnamese corpus of cultural tourism offers numerous advantages for students learning English for Specific Purposes (ESP), particularly in the field of tourism, several limitations have also been identified. Student responses revealed a range of challenges associated with corpus use, primarily related to technological barriers, time consumption and language competence.

Technological Barriers

There still exist some specific shortcomings of using the E-V corpus as shown in Table 2 below:

No	Disadvantages	**T**otal	**M**ean	**S**tandard deviation
11	I have difficulty downloading software applications when using the bilingual corpus on the Tourism.	N= 149	3.50	.695
12	Using the bilingual corpus is difficult because I do not have good knowledge of information technology	N= 149	3.62	.910
13	I sometimes encounter technical problems when using the bilingual corpus, so it takes a lot of time to search for information.	N= 149	3.59	.770
14	I have difficulty using the corpus to compile vocabulary or terminology to create a dictionary or a glossary of specialized terms.	N= 149	3.57	.805
15	I need to equip myself with learning aids such as computers, smartphones or tablets to be able to use the bilingual corpus, so I have financial difficulties.	N= 149	3.55	.840
16	I have difficulty downloading software applications when using the bilingual corpus on Tourism.	N= 149	3.38	.845
17	I have difficulty finding translated terms (target language) about tourism	N= 149	3.54	.740
18	I sometimes encounter technical problems when using the corpus, so it takes a lot of time to search for information.	N= 149	3.70	.790
19	I have difficulty using the corpus to analyze the concordances.	N= 149	3.56	.710
20	I am struggling with the corpus used to analyze the collocations or phrases.	N= 149	3.61	.750

Table 4. Drawbacks of using the E-V corpus of cultural tourism as an ESP material

In fact, there are many different opinions about the limitations of using the English Vietnamese corpus for students in ESP courses of Tourism. Among ten disadvantages mentioned above, the challenges relating to information technology skills and the ability of IT applications such as downloading software to use this corpus via Sketch Engine, more than half of the students (M=3.50) agreed and strongly agreed that they had many difficulties. Therefore, many students (M=3.62) found it very difficult to use this corpus since they did not have good IT knowledge.

Furthermore, a high rate of students in the survey (M=3.70) occasionally encountered technical problems when using this corpus; hence, they had to spend a lot of time searching for information such as vocabulary, terms and phrases in both source language and target language. In addition, only a few students did not have difficulty using a bilingual corpus to compile vocabulary or terms to build dictionaries or glossaries. Therefore, it is necessary to equip students with IT knowledge so that they can apply it well in learning ESP, this result is similar to Li's (2012) view on the importance of information technology knowledge in learning ESP.

Regarding the weakness in combining language knowledge and information technology skills, many students have difficulty in analyzing concordance and collocations (phrases) in the corpus (means are 3.56 and 3.61, respectively). In a more in-depth interview of 10 students about the reasons they have difficulty using this corpus, student ST9 said "*I hardly know how to use Sketch Engine software to open the corpus, even using an Excel file for the original corpus is quite difficult for me, I don't know how to search for phrases (collocations) or concordance like my friends did.*". Student ST4 declared that her language knowledge as well as IT skills were problematic, i.e. that she was still a bit weak in processing the documents when using a software, since she mainly only knew how to use Word or PowerPoint". What is more, many students (M=3.54) said that they sometimes had difficulty finding translated terms (target language) about tourism laws in this corpus as they did not understand the search techniques, or their language knowledge was limited.

In summary, Table 2 with the mean score from 3.38 to 3.70 and the standard deviation from .695 to .910 (<1) shows that the study results are quite accurate and reliable, indicating that the use of English Vietnamese tourism corpus by students still has some disadvantages apart from the advantages mentioned in section 4.2.1. According to many students, the combination of language knowledge and technology skills can help them easily use this document in learning ESP since this corpus has provided a very useful and highly practical document.

Time Consumption

The highest proportion of students (M = 3.70) also encountered frequent technical issues while using the corpus, which took them a lot of time to search for bilingual vocabulary, specialized terms, and phrases. Particularly, when students are unfamiliar with search strategies, they need much time to complete these tasks.

Difficulty in Language & Technology Integration

Another significant difficulty reported by students involved the integration of linguistic and technological knowledge. Specifically, many students found it challenging to conduct concordance and collocation analyses, with mean scores of M = 3.56 and M = 3.61, respectively. In follow-up interviews, students provided deeper insight into these challenges. For instance, student ST3 remarked: *"I hardly know how to use Sketch Engine software to open the corpus. Even using an Excel file for the original corpus is quite difficult for me. I don't know how to search for phrases (collocations) or concordance like my friends did."*

Similarly, ST8 shared: *"My language knowledge and IT skills are both still bad. I mainly just know how to use Word or PowerPoint, so I struggle to process documents in other software."*

In summary, Table 2 with the mean score from 3.38 to 3.70 and the standard deviation from .695 to .910 (<1) shows that the study results are quite accurate and reliable, indicating that the use of English Vietnamese tourism corpus by students still has some disadvantages apart from the advantages mentioned in section 4.2.1. According to many students, the combination of language knowledge and technology skills can help them easily use this document in learning ESP since this corpus has provided a very useful and highly practical document.

4.2.3 Suggestions for Building and Using the E-V Corpus of Tourism

As analyzed benefits and drawbacks of the E-V corpus of tourism in the previous sections, apparently it is essential for students of English Faculty to build and use this corpus in their learning ESP of tourism courses. Via the survey, we also figure out that students had some suggestions of building and using an English Vietnamese bilingual corpus of Tourism as shown in Table 3.

Table 3. Students' suggestions of building and using an E-V corpus on Tourism as an ESP material

No	Questions	Total	Mean	Standard deviation
21	It is suggested that this corpus should be used as an ESP learning resource because it is very useful for language learners.	N= 149	4.00	.835
22	Since bilingual corpora are very useful in learning vocabulary and specialized terminology, students majoring in English should approach the method of its use as soon as possible.	N= 149	4.03	.575
23	Students need to learn about building bilingual corpora so that they can build them themselves as soon as possible.	N=149	4.07	.480
24	Students could suggest their teachers regularly use the corpus to help them improve their language and IT knowledge	N=149	4.04	.565
25	Students can learn by themselves to build many bilingual corpora of different topics	N=149	4.06	.640

4.2.3 Students' Recommendations for the Development and Use of the English Vietnamese Corpus of Cultural Tourism

Building upon the previously discussed advantages and limitations of the English Vietnamese bilingual corpus of tourism, it is evident that integrating corpus use into ESP instruction is both

necessary and beneficial for students in the Faculty of English. The survey further revealed several recommendations regarding the development and application of the corpus in ESP learning in tourism courses, as summarized in Table 3.

A significant ratio of students suggested that this corpus should be use as ESP resource in supporting language acquisition in specialized domains (M = 4.00). Similarly, students expressed a strong desire to be introduced to corpus-based learning methods early in their programs (M = 4.03), emphasizing the importance of timely exposure to this approach.

Furthermore, many students (M = 4.07) agreed or strongly agreed that they should construct corpora independently. Students also suggested that lecturers should use the corpus more frequently in classroom to enhance students' both linguistic and technological competencies (M = 4.04). Additionally, it is essential for students to build more corpora in various domains to promote their applications of technological skills and language knowledge in ESP learning.

In summary, the mean scores of these recommendations ranged from 4.00 to 4.07, with standard deviations between 0.480 and 0.835 showing a high level of agreement among participants. These findings strongly support the expansion of corpus-based resources and instruction within the ESP curriculum.

5.Conclusion and Implications

In line with the growing emphasis on the development and application of authentic materials in foreign language education, this study presents the construction and implementation of an English Vietnamese corpus of cultural tourism, comprising approximately one million words. The corpus was utilized by 150 students from the Faculty of English, University of Foreign Languages and International Studies, Hue University, in their English for Specific Purposes (ESP) courses focused on tourism.

The study investigated students' engagement with the corpus in language learning and translation activities. Findings indicate that the corpus significantly supported learners in expanding their specialized vocabulary, enhancing translation skills, and improving information technology competence, which are considered key focus in modern language learning. Despite these strengths, there remain some limitations. A number of students faced challenges related to technological access and digital literacy, particularly in navigating Sketch Engine software and understanding advanced corpus analysis functions. These issues often resulted in increased time and effort required to use the corpus effectively.

In response, the study proposes a series of pedagogical recommendations aimed at maximizing the benefits of corpus-based learning while minimizing its barriers. These include early training in corpus tools, applying corpus-based methods in the ESP learning and teaching, and increasing teachers' support in building and using corpora in their professional skills

This research offers a foundational step in applying corpus linguistics to the field of applied linguistics, particularly in teaching and learning ESP. Future studies are recommended to expand on

this work by conducting the study on corpus design in more domains such as medicine, environment, laws, etc. to improve ESP education.

REFERENCES

Alqadoumi, O. M. (2013). Using corpus linguistics as a tool for reform in English language teaching and learning: the case of public schools in Arab countries. In *IEEE computer society: proceedings of a conference*, University of Bahrain, Bahrain, 246-252.

Ariyaratne, M. (2019). The Impact of Translation Technologies on the Translation Process to Give a Quality Output: A Study with Special Reference the Government Translators in Sri Lanka. *International Journal of Research and Innovation in Social Science (IJRISS)*, III(XII), 134–150.

Baker, P. (2010). Corpus methods in linguistics. In L. Litosseliti (Ed.), *Research methods in linguistics,* London: Continuum pp. 93–113.

Bennett, G. (2010). Using corpora in the language learning classroom: Corpus linguistics for teachers. University of Michigan Press.

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, *8*(4), 243-257.

Boulton, A. (2012). Computer Corpora in language learning: DST approaches to research. Melanges Crapel. Centre de recherches et d'applications pédagogiques en langues, 33, 79-91.

Carlota de Jesus, A. D. and Carrillo, J. L. Q. (2021). Advantages of using corpora to teach English. ISBN: 978-607-8356-17-1. https://cenedic.ucol.mx/fieel/pdf/1.pdf

Carter, R. (2003). Language awareness. *ELT Journal*, *57*(1), 64-65.

Hunston, S. (2006). *Corpus linguistics*. Birmingham: University of Birmingham.

Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman

Kubler, S., Zinsmeister, H. (2015). *Corpus Linguistics and Linguistically Annotated Corpora*, Bloomsbury-London-New York.

Li, X. (2012). Information Technology Application in ESP Education. *International Conference on Future Computer Supported Education. IERI Procedia* 2, 771-774.

Malenova, E. (2019). Cloud technologies in a translation classroom. Trans-Kom, *12*(1), 76–89.

McEnery, T., Xiao, R., Tono, Y.(2006). Corpus-Based Language Studies: An Advanced Resource Book, Routledge Publisher.

McEnery, T., Hardie, A. (2014). Corpus Linguistics: Method, Theory, and Practice. Cambridge University Press.

Ngo, Q. H., & Winiwarter, W. (2012). Building an English Vietnamese Bilingual Corpus for Machine Translation. *International Conference on Asian Language Processing,* 157–160. https://doi.org/10.1109/IALP.2012.30

Ngula, R. (2017). Corpus methods in language studies. In: Kuupole, D. (ed.), *Perspectives on Conducting and Reporting Research in the Humanities*, Cape Coast: University of Cape Coast Press, 205-223.

Oakes, M.P. (2012). Corpus linguistics and language variation. In: Baker. P. (ed.). *Contemporary corpus linguistics*. London: Continuum, 159-183.

Patsala, P., Michali, M. (2020). Sharpening Students' Critical Literacy Skills Through Corpus-Based Instruction: Addressing the Issue of Language Sexism. In: *Handbook of Research on Cultivating Literacy in Diverse and Multilingual Classroom*. DOI: 10.4018/978-1-7998-2722-1.ch012

Quynh, H. (2011). A study on building bilingual corpus to serve Vietnamese language processing [Da Nang]. https://doi.org/60.48.01.

Rayon, P. (2015). Computational tools and methods for corpus compilation and analysis. In: Biber, D., Reppen, R. (eds.) *The Cambridge Handbook of English Corpus Linguistics,* Cambridge University Press, 32-49.

Sinclair, J. (2004). Corpus and texts: Basic principles. In: *Developing linguistic corpora: a guide to good practice.*

Sketch Engine (2023). https://en.wikipedia.org/wiki/Sketch_Engine.

Smartcat. (n.d.). Retrieved December 27, 2021, from https://en.wikipedia.org/wiki/Smartcat#cite_note-:1-1

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. John Benjamins Publishing Company.

Van Lier, L. (2001). Language awareness. In: Carter, R., Nunan, D. (eds.) *The Cambridge Guide to teaching English to Speakers of other languages*. Cambridge: Cambridge University Press, 160-165.

Zanettin, F. (2014). Corpora in Translation. Retrieved from https://www.researchgate.net/publication/265968828

VIETNAMESE UNIVERSITY STUDENTS' PERCEPTIONS OF QUIZLET IN IELTS VOCABULARY LEARNING

Truong Thi Nhu Ngoc[1], Nguyen Tu Nhan[2], Nguyen Minh Hung[3]*

ABSTRACT

Over the past decade, vocabulary learning has shifted from traditional paper-based methods to digital flashcards, with gamification enhancing engagement. While many studies have looked into the general benefits of digital tools, there is still a significant gap in understanding how university students perceive and use platforms like Quizlet in the Vietnamese context, particularly in specialized English programs like IELTS preparation. This study examines the perspectives of three university students from different majors who attended an IELTS class incorporating Quizlet. Findings indicate that students value Quizlet for its flexibility, interactivity, and support for self-paced learning. Quizlet's spaced repetition and gamification features improve vocabulary retention, motivation, and self-directed learning. Additionally, students emphasize its role in exam preparation, progress tracking, and study organization through automated feedback and categorized folders. To effectively integrate Quizlet into Vietnamese universities, they recommend early exposure, structured digital storage, and assessment via automated feedback to optimize learning outcomes. This study provides insights into technology's role in vocabulary acquisition and offers practical recommendations for incorporating digital tools into English education in Vietnam.

***Keywords**:* Quizlet, university students vocabulary acquisition, Vietnam

1. Introduction

Vocabulary knowledge is a fundamental component of language development, as it directly influences reading comprehension, writing accuracy, and communication skills (Hunt & Beglar, 2005; Schmitt, 2008). However, students often struggle with pronunciation due to complex grammatical structures when learning vocabulary (Surmanov & Azimov, 2020). Additionally, understanding word meanings and applying them appropriately in different contexts remains a persistent challenge. To address these difficulties, researchers have explored the role of artificial intelligence (AI) in supporting vocabulary acquisition.

English as a Foreign Language (EFL) education has undergone a significant transformation due to AI integration. Real-time feedback tools such as QuillBot, Grammarly, and ChatGPT assist students in improving writing, grammar, and vocabulary skills (Ariyanti, 2021; Kacena et al., 2024; Marzuki et al., 2023). AI-powered applications enhance vocabulary learning by providing interactive experiences, fostering independent learning, and enabling students to identify and correct errors autonomously. AI-mediated instruction can also increase student motivation, self-regulated learning, and language achievement (Wei, 2023).

[1] Ho Chi Minh City University of Industry and Trade, E*mail: ngocttn@huit.edu.vn*
[2] Ho Chi Minh City University of Industry and Trade
[3] Hong Bang International University

In addition to AI-powered tools, gamification strategies have been widely adopted to enhance vocabulary learning. Within the framework of Education 4.0, gamification tools like Quizlet offer innovative solutions to teaching and learning challenges, particularly in vocabulary acquisition. Research suggests that Quizlet's spaced repetition system, flashcard-based learning, and gamified features enhance vocabulary retention and student engagement due to their effectiveness and ease of use (Jackson III, 2015). Likewise, gamified learning facilitates vocabulary acquisition by making the process more interactive and enjoyable (Waluyo & Bucol, 2021). Additionally, Alharbi and Khalil (2023) and Jomaa et al. (2024) found that AI-driven vocabulary training platforms significantly enhanced students' motivation and performance. Furthermore, Hayakawa et al. (2019) emphasize that long-term vocabulary retention relies on initial exposure and reinforcement quality. These findings underscore the role of gamification and AI-driven tools in optimizing vocabulary learning within modern educational frameworks.

Despite the increasing global usage of AI-assisted vocabulary learning, there is still a considerable gap in knowing these tools' unique perceptions and practical efficacy in various cultural and educational settings. While AI-powered applications like Quizlet have grown in popularity worldwide, previous research frequently lacks the qualitative insights required to understand how students perceive and engage with these tools in real-world educational contexts. While studies have looked at the success of Quizlet in many countries, there is still a need to go beyond simple performance measures and investigate the student voice-their motivations, difficulties, and integration recommendations.

The gap is more pronounced in certain areas. In Vietnam, for example, research into the impact of these instruments remains underexplored. While some local studies, such as those by Nguyen Thi Kim Ngoc (2024) and Nguyen Lan Vy and Tran Quoc Thao (2023), have highlighted Quizlet's benefits, there is a clear need for more in-depth, nuanced research that goes beyond simple usability and investigates how students perceive and use these tools. Building on these findings, this study seeks to explore students' perceptions of Quizlet's role in enhancing English vocabulary acquisition and to gather insights into its potential integration into university curricula. Specifically, the study aims to address the following research questions:

Research question 1: How do the participants perceive Quizlet's role in enhancing English vocabulary acquisition?

Research question 2: What recommendations do participants have for integrating Quizlet into the English language curricula at university?

2. Literature review

2.1. Studies on Quizlet in English Language Education

Many educators advocate for Quizlet as an effective gamified learning tool, particularly for vocabulary acquisition and exam preparation. A broad range of studies report positive

outcomes, including improved test scores, higher learner satisfaction, and increased engagement compared with traditional methods. However, a critical review reveals two persistent limitations: (1) most studies document what Quizlet achieves but rarely explain why it motivates learners; and (2) relatively few have situated Quizlet research within high-stakes testing contexts such as IELTS, especially in Vietnam.

Li (2025), for instance, used a quasi-experimental design with 361 Chinese students and demonstrated that a game-enhanced environment improved vocabulary learning, self-efficacy, and motivation. Yet the study offered little theoretical insight into the psychological mechanisms that drove these improvements. Similarly, Bueno-Alastuey and Nemeth (2020) found that students regarded Quizlet as effective and enjoyable. Still, the lack of a guiding framework limited their ability to explain why students experienced heightened motivation.

Other research has produced nuanced or mixed findings. Chaikovska and Zbaravska (2020) compared Quizlet with paper-based vocabulary notebooks among A2-level Ukrainian students in extracurricular EFL courses. While Quizlet users outperformed their peers, the advantage was insufficient to conclude that digital tools are inherently superior. This suggests that Quizlet's effectiveness may depend on contextual variables such as student background, course intensity, or instructional design.

Large-scale studies reinforce the long-term potential of gamification. Putz et al. (2020) showed that gamified approaches enhanced retention and soft skills, while Whitman et al. (2019) reported improved recall of pharmaceutical terminology with Quizlet. Still, these studies overlooked individual learner differences and failed to examine how digital gamification sustains motivation over time. Importantly, Suharno et al. (2023) addressed this limitation by showing that teacher presence in a long-term Quizlet-based listening course was essential for maintaining student motivation and performance. Their findings underscore that gamification should complement rather than replace pedagogy and that instructor involvement remains decisive.

Several smaller-scale investigations have examined Quizlet's role in extending learning beyond the classroom. Waluyo and Bucol (2021) and Nguyen Quoc Long and Le Van Ha (2022) highlight its potential for self-directed vocabulary learning, emphasizing the importance of task diversity and progress monitoring. Qualitative studies (e.g., Aprilani & Suryaman, 2021) have shown that learners perceive Quizlet as engaging and supportive of memorization, while action research (Tran My Dung and Nguyen Thi Minh Loan , 2022) confirmed its role in vocabulary recall. Other studies in EMI contexts (Trinh Phuong Hoa and Ho Thi Bach Mai, 2023) demonstrated benefits for content comprehension, while Suharno et al. (2023) emphasized the role of teacher presence in sustaining motivation within gamified environments. Complementary to this, Mohammadi et al. (2024) revealed that digital flashcards significantly outperformed both paper flashcards and word lists in supporting Iranian EFL students' receptive and productive vocabulary development. Their findings demonstrate

how mobile-assisted gamification tools, like Quizlet, can advance academic literacy more effectively than traditional methods.

Relatedly, research has also considered external factors such as gender and blended learning design. Almusharraf et al. (2023) found that motivation for Kahoot! was not gender-specific in Saudi EFL classes, suggesting that learner engagement with gamified tools transcends gender differences. While not directly tied to Quizlet, the study adds weight to claims that digital game-based learning generally fosters equitable motivation. On the other hand, Tosun (2015) reported that a blended learning approach incorporating digital tools did not significantly improve vocabulary outcomes unless learners were self-motivated to study outside class. This suggests that Quizlet's effectiveness, like other gamified platforms, depends not only on tool design but also on learner autonomy and sustained engagement.

These findings illustrate Quizlet's value for vocabulary learning across diverse contexts. Nevertheless, they share two gaps. First, few studies investigate its use in IELTS preparation, where vocabulary breadth and retention are crucial for success. Second, most research lacks a theoretical grounding that can explain why learners find Quizlet motivating. This absence prevents a deeper understanding of the psychological and pedagogical processes underlying its effectiveness.

2.2. Theoretical Framework

The present study adopts Self-Determination Theory (SDT) (Deci & Ryan, 1985) as its theoretical lens to address this gap. SDT posits that motivation is fostered when three basic psychological needs are satisfied: autonomy, competence, and relatedness. First, regarding autonomy, Quizlet allows learners to select study modes (e.g., Flashcards, Learn, Write), control their pace, and revisit sets as needed. This flexibility supports a sense of ownership over learning. Next, regarding competence, instant feedback, progress tracking, and gamified achievements enhance students' feelings of mastery and capability. Finally, regarding relatedness, collaborative features such as Quizlet Live foster teamwork, social interaction, and shared learning experiences.

While Self-Determination Theory (Deci & Ryan, 1985) provides a solid foundation for understanding learners' motives through autonomy, competence, and relatedness, it is fruitful to consider how gamified features in vocabulary learning tools interact with motivational and cognitive processes. Empirical studies in gamified instruction and digital game-based language learning offer valuable insights into these dynamics. Sadeghi et al. (2022) compare gamified and non-gamified vocabulary instruction, finding that the gamified group achieved significantly greater vocabulary gains and reported higher motivation. This suggests that gamification can act as a motivational affordance that catalyzes learner engagement, reinforcing SDT's dimensions through feedback, progress visibility, and challenge.

Similarly, Chowdhury (2024) explores how digital games support vocabulary acquisition via agency, intrinsic motivation, tinkering, and contextualized interaction. The

study highlights that when learners are permitted to experiment, make decisions, and control pacing within a gamified environment, they tend to engage more deeply with vocabulary tasks. This supports a perspective that sees gamification features mediating between tool design and learning outcomes.

In short, Quizlet encourages intrinsic motivation and sustains learner engagement beyond what traditional rote memorization methods typically achieve. The self-determination theoretical framework enables the current study to interpret whether Quizlet supports vocabulary learning and why students perceive it as motivating.

3. Methodology

3.1. Research design

This study adopted a multiple-case study design (Yin, 2018) to investigate the integration of Quizlet into IELTS vocabulary instruction. When literature on a subject is limited, qualitative research enables researchers to engage directly with participants and explore the topic in greater depth (Marshall & Rossman, 2014). Although the participant pool was small (three focal students in a class of seven), this design was considered appropriate for two reasons. First, case studies allow for an in-depth exploration of contemporary educational practices in authentic classroom contexts, particularly where the boundaries between phenomenon (digital vocabulary learning) and context (IELTS instruction) are not clearly separated. Second, the study aimed to capture learners' lived experiences, strategies, and perspectives in a specific setting rather than generalize to a wider population. This aligns with the purpose of case study research, which prioritizes depth of understanding over breadth of coverage.

3.2. Participants

The participants in this study were university seniors, including one male student majoring in electrical engineering and two female students majoring in English from a public university in Ho Chi Minh City. All participants were 21 years old. To protect their identities, pseudonyms were assigned during the study. Participation was voluntary, and all students provided informed consent before the interviews. Additionally, all participants attended the same tutored IELTS class taught by Mr. N. Table 1 presents the participants' demographic information.

Table 1: Personal information of participants

Name	Gender	Age	Major	Years at university	Years of learning English
Alice	Female	21	English Language	3	11

Xuka	Female	21	English Language	3	15
Lee	Male	21	Automation Electricity	3	11

Alice, a 21-year-old English major, has studied English for eleven years. Initially, she relied on YouTube for self-study, but later, she incorporated Quizlet and social media platforms like Facebook and TikTok, actively participating in English-speaking groups. Her preference for interactive, casual learning has helped improve her speaking and listening skills by exposing her to real-world language use. While she enjoys the engaging content on YouTube and TikTok, she acknowledges that Quizlet offers a more structured and systematic approach to vocabulary retention.

Xuka, also 21 years old and an English major, has been learning English for 15 years, starting at six. Despite her long experience, she still faces challenges with language use, synonym confusion, and vocabulary retention. She used Quizlet to expand her vocabulary and manage test-related stress. Additionally, she recommends Google Translate for quick word recognition and comprehension. However, she notes that while Google Translate provides instant translations, it lacks the structured reinforcement of spaced repetition, limiting its effectiveness in long-term retention.

Lee, the only male participant, is a 21-year-old automation engineering student. He has studied English for eleven years but initially struggled with comprehending academic materials and using technical language in his field. In 2022, he transitioned from traditional learning methods to using Quizlet, finding it a significant improvement. Quizlet's repetitive practice and clear explanations made learning more engaging and effective. His experience reflects a broader shift toward integrating digital tools that enhance motivation and learning efficiency. Unlike passive platforms like YouTube and social media, Quizlet promotes goal-oriented, interactive learning.

3.3. Research Instruments

The researcher designed twelve interview questions to investigate how Quizlet supports students in overcoming vocabulary learning challenges. In addition to formulating the interview questions, the researcher conducted data collection, analysis, and interpretation, ensuring alignment with the study's objectives. Each participant responded to all twelve questions. Specifically, Questions 1 and 2 gathered demographic information, while Questions 3 to 5 explored participants' general vocabulary learning habits and strategies. Questions 6 to 10 examined their experiences with Quizlet, including changes in study approaches, vocabulary retention, and confidence in learning English. Finally, Questions 11 and 12 focused on participants' perspectives on traditional versus digital learning methods and their recommendations for integrating Quizlet into English programs. (See Appendix A for the complete list of interview questions).

3.4. Data collection procedure

The study involved three participants enrolled in an IELTS preparation course taught by Mr. N, an instructor with three years of experience teaching IELTS at a foreign language center (Appendix C). The class, comprising seven students, followed a structured curriculum incorporating Quizlet-based vocabulary assessments alongside Barron's Essential Words for the IELTS (Lougheed, 2011). The study spanned three months, with students attending two-hour classes twice per week. Each vocabulary unit, consisting of three lessons, lasted two weeks and focused on a specific topic to provide contextualized vocabulary instruction.

Quizlet was integrated into the IELTS course through **teacher-led classroom activities** and **student self-study assignments.** The instructor introduced vocabulary sets prepared on Quizlet in class and used them in interactive activities such as *Quizlet Live* team competitions and vocabulary refresher exercises. Outside class, students were required to complete specific Quizlet study modes (e.g., *Learn*, *Write*) as weekly homework, with their progress monitored by the instructor. This dual approach ensured that Quizlet functioned as a supplementary self-study tool and an embedded component of the curriculum.

Students engaged with Quizlet in a structured sequence to support vocabulary learning across three dimensions: **form, meaning, and use**. They began with *Flashcards* to recognize spelling, pronunciation, and part of speech (form). Next, they practiced *Learn* mode and *Match* to connect definitions, synonyms, and translations (meaning). Finally, they used *Write* and *Test* modes to apply sentence vocabulary, improving grammatical accuracy and contextual usage (use). This scaffolded process enabled learners to move progressively from recognition to recall and application.

Using gamification strategies like Quizlet Live and Quizlet Plus, Mr. N led interactive practices promoting team-based contests and sustained participation. Participants used vocabulary sets in a variety of ways. Before moving on to learn to practice definitions, synonyms, and sentence use, they started with the Flashcards mode to familiarize themselves with sound and meaning. After that, they wrote for spelling and word forms. By methodically utilizing various study techniques, students could advance from recognition to recall and application, transforming vocabulary acquisition from a passive to an active and enjoyable process. Learning results were further enhanced by the often-updated Quizlet library (see Appendix B for sample Quizlet questions). In total, there are 4 units as follows:

- **Unit 1: The Natural World** - *Environmental Impacts of Logging, Bird Migration, Plant Life in the Taklimakan Desert*
- **Unit 2: Leisure Time** - *Peripheral Vision in Sports, History of the Circus, Uses of Leisure Time*
- **Unit 3: Transportation** - *First Headlamps, Major Subways of Europe, Electric Cars Around the Globe*

- **Unit 4: Culture** - *Origins of Writing, Hula Dancing in Hawaiian Culture, The Art of Mime*

At the end of the course, students were invited to participate in interviews, and three volunteers - Alice, Xuka, and Lee - agreed to participate. Each semi-structured interview lasted 20 –25 minutes, allowing for an in-depth exploration of participants' experiences. The interviews were recorded with participants' consent and later transcribed for analysis. Participants first provided background information, including their year of study and the time they typically spent learning vocabulary each week. They then described their usual study methods and the strategies or resources they had relied on before using Quizlet. The discussion also covered their prior experience with Quizlet, whether they had used it before joining the IELTS class, and how frequently they engaged with it during the course.

To ensure comprehensive data collection, participants reflected on how their approach to vocabulary learning evolved after incorporating Quizlet, comparing its effectiveness to previous methods. They also shared insights into how Quizlet influenced their confidence in learning English and evaluated the benefits and limitations of traditional versus digital vocabulary learning tools. Lastly, participants provided recommendations for students and educators on integrating Quizlet into English programs and expressed interest in exploring other language-learning applications. To enhance data reliability, member checking was conducted by summarizing key points at the end of each interview and asking participants to confirm or clarify their responses.

3.5. Data analysis

This study explores Quizlet's impact on English vocabulary learning through qualitative analysis. With participants' consent, all interviews were recorded, transcribed, and translated from Vietnamese to English for accessibility. Data integrity was ensured by carefully managing the translation process to maintain the original context and meaning. The researcher used inductive content analysis with open coding, letting themes develop naturally rather than being predetermined, following the guidelines of Marshall and Rossman (2014) and Creswell (2007). Researchers systematically reviewed transcripts, made annotations, and formed initial codes. Common patterns were identified, while differing perspectives were examined. The coding system was refined through continuous comparison, ensuring accuracy and depth. Categorical aggregation was used to establish themes, with direct interview quotes illustrating key insights to enhance validity. To ensure the trustworthiness of the analysis, reflexivity was maintained throughout the process by keeping analytic memos and reflecting on how personal assumptions might influence interpretations. Inter-rater reliability checks were also conducted: a second researcher independently coded a subset of the data, and discrepancies were discussed until consensus was reached. This process not only enhanced consistency in coding but also strengthened the credibility of the findings.

4. Findings and discussion

4.1. Findings

4.1.1. Participants' perceptions of using Quizlet in English vocabulary acquisition

4.1.1.1 Traditional Vocabulary Learning vs. Digital Platforms: Key Benefits and Limitations

Digital platforms like Quizlet have been integrated into traditional vocabulary learning methodologies, improving efficiency and engagement while maintaining the benefits of flashcards and repetition. Although conventional approaches provide students autonomy and structured reinforcement, they can be time-consuming and less engaging. In contrast, digital resources like Quizlet encourage interactive and fast-paced learning. Alice highlights a key distinction between conventional and digital learning, particularly in terms of motivation and accessibility:

> *When I study alone at home without a teacher, I lack motivation and feel little incentive to complete my tasks correctly. With digital platforms like Quizlet, I can manage my time efficiently and study at my convenience. However, I won't be able to use the platform without internet access. That's a limitation*. (Alice)

Alice's comment underscores an essential challenge in self-directed learning: students may struggle with motivation without external guidance. This reflects a common issue in independent study: learners require a structured environment to stay on track. However, she also acknowledges Quizlet's flexibility, allowing learners to study on their schedule. Despite these benefits, Alice highlights a major technological limitation - dependence on internet connectivity, which can disrupt access to learning resources.

Similarly, Xuka praises Quizlet's ability to accelerate vocabulary acquisition, making learning more efficient and convenient. She highlights features such as active recall and instant feedback, which contribute to long-term retention. Although she acknowledges that traditional methods allow offline study through note-taking, she views Quizlet's accessibility as a key advantage, provided that there is an internet connection. Also, she believes a key limitation of Quizlet is that while it helps with word retention, it does not always offer sufficient contextual understanding. This explains why Xuka still refers to textbooks to grasp the usage of words in context.

> *I believe that traditional approaches are time-consuming and less effective than Quizlet. Quizlet provides a vast vocabulary library, allowing learners to memorize words more efficiently. One advantage of traditional learning is that it enables studying without the Internet through note-taking, but Quizlet can be used anywhere as long as the Internet is available. However, when studying with Quizlet, I must refer to my textbooks to understand the context. (Xuka)*

In traditional educational methods, physical materials such as books, notebooks, and pencils can help establish discipline, routine, and structure in learning. However, Lee noted:

> *Traditional materials like books and notebooks can be easily misplaced or damaged, making it difficult to recover lost content. Additionally, handwriting can feel tedious, and the lack of interactivity in traditional methods may reduce engagement. (Lee)*

According to Lee, one major limitation of traditional materials is their vulnerability to loss or damage - once destroyed, retrieving the information becomes difficult. In contrast, digital tools often provide cloud-based storage and backup options, allowing students to recover their data in case of device failure. Another concern with traditional methods is that handwriting-based learning lacks interactive elements, which could negatively impact student engagement and motivation over time.

4.1.1.2 Student Study Habits Using Quizlet

Alice, Xuka, and Lee provide valuable insights into how Quizlet's flexibility impacts study habits and time investment.

Alice's Perspective: Balancing Consistency and Breaks

Alice's study routine reflects a significant investment in digital learning, dedicating about two hours daily to the platform. This consistent engagement aligns with best practices for language acquisition, where regular exposure enhances retention. Quizlet's flexibility allows her to learn independently and review topics as needed, unlike traditional methods that often restrict study sessions to fixed school hours.

> *On a weekday, I'll utilize Quizlet six days a week, except for Sunday, when I go outside. Everyday, I spend almost two hours on the website. (Alice)*

Her decision to take Sundays off demonstrates a balanced approach to learning, acknowledging the importance of rest and cognitive recovery. This suggests an awareness of burnout prevention and the role of leisure in sustaining long-term academic performance.

Xuka's Perspective: Leveraging Spaced Repetition for Memory Retention

Xuka's study habits emphasize the importance of regular practice and spaced repetition, a well-known technique for improving long-term memory retention. Unlike traditional methods, which often involve longer but less frequent study sessions, Quizlet enables frequent, shorter reviews, reinforcing knowledge over time.

> *I usually study on Quizlet for two hours daily and throughout the week to maintain memory frequency because I usually gain language in the evening. (Xuka)*

Additionally, her evening study choice suggests a personalized learning approach, aligning study sessions with her peak concentration and cognitive efficiency. This underscores one of Quizlet's key advantages - allowing learners to customize their schedules based on their optimal study times.

Lee's Perspective: Integrating Traditional and Digital Learning

Lee's study habits highlight a strategic blend of traditional and digital methods. His study time increases on days when he has English classes, likely to reinforce classroom learning and prepare for future lessons. This approach demonstrates an understanding of the complementary roles of structured instruction and self-directed practice.

> *If there is no English class, it will be one hour every day; otherwise, it will be two to three hours, except on Sunday. (Lee)*

Like Alice, Lee excludes Sundays from his study schedule, reinforcing the theme that students recognize the cognitive benefits of taking breaks. Unlike traditional education systems, which often enforce rigid timetables, Quizlet provides the flexibility to adapt study sessions based on individual needs.

4.1.1.3 Quizlet as a Catalyst for Vocabulary Retention, Self-Regulated Learning, Engagement, and Exam Preparation

Quizlet has emerged as a powerful tool for language learners, significantly enhancing their ability to acquire, retain, and apply new knowledge. Through its interactive features, such as digital flashcards, spaced repetition, and gamified learning, Quizlet fosters engagement, motivation, self-regulated learning, and exam preparation among students. Additionally, its user-friendly interface allows learners to focus on content rather than the tool's mechanics, making language acquisition more intuitive and efficient.

Enhancing Vocabulary Retention

One of Quizlet's most valuable contributions to language learning is its role in vocabulary acquisition. Many students struggle with memorizing new words, but Quizlet's flashcard system, combined with spaced repetition, ensures that learners reinforce their knowledge over time. Alice, a student using Quizlet for English studies, noted:

> *Before using Quizlet, I found it difficult to remember new words. I retain vocabulary much better with regular practice and use it naturally in conversations***.** (Alice)

Quizlet's ability to organize words into custom folders is a key feature that enhances vocabulary learning. Alice found that creating separate folders for synonyms, antonyms, and different word categories allowed her to expand her vocabulary systematically. Instead of passively memorizing long word lists, she could efficiently access and review specific word groups tailored to her learning needs. This structured approach helped her recall words more effectively and apply them in context. Similarly, Xuka highlighted how Quizlet's repetition mechanism allowed her to distinguish between different word types:

> *In the past, I could not discriminate between different word kinds; therefore, I frequently missed words or used incorrect vocabulary while composing phrases. By sifting through the results daily to repeat the flashcards several times, Quizlet helped me solve the challenge by eliminating the words I couldn't fully recall.* (Xuka)

This experience underscores the platform's ability to support long-term retention. By systematically reinforcing learning through self-assessment and repetition, Quizlet helps users retain words for longer, reducing the common issue of forgetting vocabulary shortly after learning it.

Fostering Self-Regulated Learning

Beyond vocabulary improvement, Quizlet empowers students to take control of their learning. Its structured study sets and notification reminders encourage learners to establish consistent study habits. Xuka emphasized how Quizlet's notifications helped her stay on track:

> *Quizlet's reminders keep me accountable. I set daily learning goals, and the app helps me stay on track. (Xuka)*

By prompting learners to engage with their material daily, Quizlet fosters self-discipline and consistency, key elements of self-regulated learning. Unlike traditional classroom-based learning, where students often rely on teachers for guidance, Quizlet encourages autonomy by allowing users to customize their learning experience based on their proficiency level and individual needs.

Gamification and Motivation

Quizlet's game-based features, such as Match and Gravity, transform studying into an engaging experience. These elements make learning enjoyable and foster a positive attitude toward studying. Research suggests that gamification enhances motivation by creating a sense of achievement and competition, improving learning outcomes. Lee, a dedicated Quizlet user, shared:

> *I used to see studying as a chore, but Quizlet's interactive games make it fun. Competing with myself to improve my scores has kept me motivated."* (Lee)

Additionally, Quizlet's interactive approach provides learners with immediate feedback, reinforcing correct answers and guiding users toward improvement. Alice appreciated the platform's ability to adapt to different competency levels:

> *Quizlet offers several appropriate levels, making it easier to follow along and verify past knowledge.* (Alice)

This feature ensures that students remain engaged by challenging them at an appropriate difficulty level, preventing frustration while still encouraging progress.

Impact on Exam Preparation

Quizlet also serves as an effective tool for test preparation. Many students use its custom quizzes and practice tests to reinforce their understanding of key concepts. The ability to generate multiple-choice, matching, and written-answer questions allows learners to simulate exam conditions, boosting their confidence and performance. Lee illustrated how Quizlet helped him prepare for assessments:

> *Regularly using Quizlet is a helpful way for me to prepare for my teacher's test. I have to work harder and update more to review previous lessons without waiting for the study day when I have English class since it reminds me to study through alerts.* (Lee)

By integrating Quizlet into his daily routine, Lee avoided last-minute cramming, which often leads to surface-level memorization. Instead, he developed stronger retention of key concepts, leading to better exam performance. Additionally, Quizlet's progress-tracking system allowed him to identify weak areas and focus on topics requiring further review.

Another significant advantage of Quizlet is its collaborative learning environment. The platform enables students to share study sets, benefiting from collective knowledge and peer support. This feature is handy in classroom settings, where group learning enhances comprehension. Lee also noted that Quizlet's mobile accessibility allowed him to study anywhere, making integrating learning into his daily routine easier.

4.1.2 Participants' recommendations for integrating Quizlet into English language curricula at Vietnamese universities

Insights from Quizlet users are essential to maximizing the platform's integration into vocabulary learning programs at Vietnamese universities. Rather than replacing conventional teaching strategies, Quizlet is a valuable complement, enhancing learning experiences through its interactive features. The students' recommendations provide practical strategies for leveraging Quizlet to improve vocabulary retention and engagement in university courses. Three key themes emerge from the participants' suggestions: early engagement, assessment and feedback, and organization.

Early Engagement with Quizlet for Academic Success

Alice emphasizes that Quizlet plays a crucial role in her academic experience, mainly because her instructors frequently use it to assess students based on materials from previous lectures. In her view, right from being first-year students, students should become familiar with Quizlet as early as possible to prevent missing out on important vocabulary and gradually adapt to using it as a study tool:

> *Since many university lecturers use Quizlet for vocabulary reviews, students should become familiar with Quizlet early on in the first year at university to prevent missing out on valuable vocabulary and gradually become used to using it. (Alice)*

Alice's insight underscores the significance of Quizlet as a reinforcement tool. Since many university assessments build upon previously taught content, students unfamiliar with Quizlet or failing to incorporate it into their study routines may fall behind. Additionally, Quizlet's repetitive study modes, such as flashcards, quizzes, and games, help reinforce knowledge more effectively than traditional rote learning.

For first-year students, transitioning to university involves an increased workload and the need for greater self-directed learning. Quizlet provides a structured, engaging way to facilitate this transition by offering interactive review options that enhance memory retention.

Enhancing Assessments and Feedback through Quizlet

Xuka highlights Quizlet's potential for improving assessment and communication between teachers, students, and parents. One notable feature of Quizlet is its ability to send results directly to students and parents, making it a useful tool for tracking academic performance:

> *Because Quizlet allows results to be sent directly to students' and parents' emails, teachers may utilize it to develop assessments. (Xuka)*

Traditional assessments often require manual grading, which can be time-consuming, error-prone, and lack transparency. By contrast, Quizlet automates assessment results, offering instant feedback and reducing administrative workload for educators. Additionally, this system promotes greater accountability, as parents and students receive performance reports in real-time.

Students also benefit from immediate access to their results, allowing them to identify weak areas and adjust their study strategies accordingly. This differs from traditional methods, where students may experience long waiting periods for grades, leading to anxiety and uncertainty about their progress.

Organizing Learning Materials with Quizlet Folders

Lee offers another practical recommendation - students should create individual Quizlet folders to store their study sets for later use

> *Each of you should have a Quizlet folder to save materials for later use. (Lee)*

By digitally organizing study materials, students can easily retrieve important content without the risk of losing physical notes or printed handouts. This strategy is particularly beneficial in higher education, where students must manage large amounts of academic content across multiple courses.

In addition to organization, appropriately naming and categorizing Quizlet folders can enhance study efficiency. Rather than searching through scattered notes, students can quickly locate relevant vocabulary and concepts, optimizing their study time. Moreover, Quizlet's sharing feature enables students to collaborate by exchanging folders with peers, promoting peer-assisted learning.

In short, traditional vocabulary learning and digital platforms like Quizlet each offer distinct advantages and limitations. While conventional methods provide structured reinforcement and autonomy, they can be time-consuming and less engaging. In contrast, digital tools like Quizlet enhance efficiency, interactivity, and self-paced learning but rely on Internet access. Student perspectives reveal that Quizlet improves vocabulary retention through

spaced repetition and gamification, fostering self-directed learning and motivation. Additionally, it supports exam preparation and progress tracking. For effective integration into Vietnamese universities, students recommend early engagement, assessment through automated feedback, and organized study materials using Quizlet folders to enhance learning efficiency.

4.2. Discussion and implications

The findings of this study reveal that Quizlet positively influenced students' vocabulary learning experiences, particularly in terms of motivation, engagement, and self-directed study. To explain these findings, this study draws on Self-Determination Theory (SDT) (Deci & Ryan, 1985), which emphasizes three basic psychological needs-autonomy, competence, and relatedness- as essential drivers of intrinsic motivation.

Regarding autonomy, students valued Quizlet's flexibility to manage their own learning schedules, as reflected in Alice's and Xuka's accounts of studying independently at their preferred times. This supports SDT's claim that autonomy fosters intrinsic motivation by giving learners control over when, where, and how they study. In contrast, traditional methods limited their ability to adapt study sessions to personal needs, often reducing motivation.

For competence, Quizlet's gamified features, instant feedback, and spaced repetition reinforced students' sense of progress and mastery. For example, Xuka reported improved retention through repeated exposure, while Lee highlighted motivation through competing with himself in games. These findings align with SDT, which posits that when learners experience a sense of competence, they are more motivated to persist in their learning. Traditional methods, by contrast, often lacked immediate feedback and interactive reinforcement, which limited students' perception of achievement.

In terms of relatedness, while less prominent, students also recognized the collaborative potential of Quizlet, such as sharing folders and study sets with peers. This aligns with SDT's emphasis on relatedness- the need to connect with others, which was fulfilled through Quizlet's interactive and collaborative functions. Even when students primarily studied independently, the potential for shared study sets created a sense of community and academic support.

These findings suggest that Quizlet fosters intrinsic motivation by satisfying students' psychological needs for autonomy, competence, and relatedness. This provides a theoretical explanation for why students found Quizlet more engaging and effective than traditional vocabulary learning methods.

Beyond SDT, the findings also point to self-regulated learning behaviors. Students used Quizlet to set personal goals (e.g., Alice's daily two-hour routine), monitor progress (e.g., Xuka's reliance on repetition to identify weaknesses), and integrate study sessions strategically into their academic schedules (e.g., Lee's adjustment on class and non-class days). This highlights Quizlet's role in promoting autonomy-supportive learning environments that encourage students to take ownership of their learning.

This study contributes to existing research in two ways. First, while prior studies have emphasized Quizlet's benefits for vocabulary memorization (e.g., Bueno-Alastuey & Nemeth, 2020; Whitman et al., 2019), this study extends the discussion by explaining why students are motivated through the lens of SDT (e.g., Chowdhury, 2024). Linking students' perceptions to autonomy, competence, and relatedness provides a theory-driven understanding of Quizlet's effectiveness. Second, within the Vietnamese higher education context, this study highlights practical strategies such as introducing Quizlet early in students' university programs, integrating it into assessment with feedback mechanisms, and encouraging organizational strategies like personalized folders. These recommendations are particularly relevant for EMI (English as a Medium of Instruction) programs, where vocabulary retention is critical.

Educators should strategically incorporate Quizlet into the curriculum, not as a replacement but as a complementary tool to traditional methods. Teachers can maximize their motivational potential by designing study sets that provide contextualized vocabulary use (collocations, idioms, example sentences) and by structuring activities that balance autonomy with guided practice. Gamified features and collaborative activities, such as group competitions, can enhance engagement while satisfying the need for relatedness. At the institutional level, universities should train teachers and students to integrate Quizlet into language instruction, ensuring it is used effectively.

Gamification can be a pedagogical tool to stimulate motivation by supporting learners' autonomy, competence, and relatedness through elements such as points, badges, timed challenges, and progress tracking. These features encourage learners to engage more actively, persist in their practice, and develop self-regulated behaviors like repeated review, facilitating vocabulary gains. However, the design of gamified tasks should go beyond entertainment, ensuring that learners' attention is directed toward meaningful cognitive processes such as understanding word meaning, usage, and retrieval rather than focusing solely on collecting rewards. To sustain the benefits of gamification over time and avoid the diminishing effects of novelty, teachers should embed adaptive difficulty, scaffolding, and opportunities for social interaction, thereby maintaining engagement while fostering deeper and more lasting language learning outcomes.

In short, by adopting a motivation-theory lens, this study demonstrates that Quizlet improves vocabulary retention and fulfills students' psychological needs, which explains its motivational impact. This theoretical grounding strengthens the argument for integrating Quizlet into university language curricula and clarifies the literature on digital tools in language learning.

5. 5. Conclusion, limitations, and recommendations for future studies

The results show that Quizlet's adaptability, engagement, and self-paced learning are highly valued by students. Its gamification elements and spaced repetition improved self-directed learning, motivation, and language retention. Students also understood how Quizlet's classified folders and automatic feedback helped with exam preparation, progress monitoring,

and study organizing. According to Self-Determination Theory, these characteristics increased their confidence in learning outcomes and met their requirements for competence and autonomy. Students stressed that combining Quizlet with conventional study methods produced the most significant outcomes, striking a balance between organized learning and interesting digital resources, even if gamified features like "Quizlet Live" enhanced fun and teamwork. They suggested early exposure, structured digital storage, and assessment using automated feedback to optimize its advantages in Vietnamese institutions. This study provides insightful information on how digital tools aid vocabulary development and useful tactics for integrating technology into Vietnamese English instruction.

Despite its advantages, this study has several limitations. First, its findings may not be broadly generalizable, as they focus on specific academic disciplines. Students from different universities, geographical regions, or technological learning environments may perceive Quizlet's effectiveness differently. Second, this study captures students' experiences at a single time rather than examining the long-term effects of Quizlet on vocabulary growth. Although students reported increased motivation, confidence, and enjoyment, further research is needed to assess the sustained impact of regular Quizlet use. Additionally, the relatively small participant pool limits the scope of the findings, highlighting the need for larger-scale studies.

Future research should address these limitations by expanding sample sizes to capture broader student perspectives. Investigating variations across disciplines, academic levels, and learning environments would provide a more comprehensive understanding of Quizlet's effectiveness. Additionally, studies should employ longitudinal designs to assess the long-term impact of Quizlet on vocabulary acquisition and retention. Examining educators' experiences with Quizlet can also offer insights into best practices for implementation. Future research should incorporate multiple methods and dissemination channels to enhance knowledge dissemination, such as international conferences, academic publications, and digital platforms. By extending research beyond the scope of this study, education can continue to evolve and adapt to technological advancements in language learning.

References

Alharbi, K., & Khalil, L. (2023). Artificial intelligence (AI) in ESL vocabulary learning: An exploratory study on students and teachers' perspectives. *Migration Letters, 20*(S12), 1030–1045.

Almusharraf, Norah, et al. (2023). Gender Differences in Utilizing a Game-Based Approach within the EFL Online Classrooms. *Heliyon*, 9 (2). https://doi.org/10.1016/j.heliyon.2023.e13136

Aprilani, D. N., & Suryaman, M. (2021). Students' Perception in Learning English Vocabulary Through Quizlet. *JET (Journal of English Teaching)*, *7*(3), 343–353. https://doi.org/10.33541/jet.v7i3.3064

Ariyanti, A. (2021). Technology-enhanced paraphrasing tool to improve EFL students' writing achievement and enjoyment. *Journal of English Language Teaching and Linguistics, 6*(3),715-726. https://doi.org/10.21462/jeltl.v6i3.698

Baptist, S. C. (2018). *Effects of Quizlet on vocabulary mastery* [Master's thesis, Rowan University, College of Education, Department of Interdisciplinary and Inclusive Education, MA Special Education]. Rowan Digital Works. https://rdw.rowan.edu/etd/2555

Bueno-Alastuey, M. C., & Nemeth, K. (2020). Quizlet and podcasts: effects on vocabulary acquisition. *Computer Assisted Language Learning*, *35*(7), 1–30. https://doi.org/10.1080/09588221.2020.1802601

Chaikovska, O., & Zbaravska, L. (2020). The efficiency of quizlet-based EFL vocabulary learning in preparing undergraduates for state English exam. *Advanced Education*, *7*(14), 84–90.

https://doi.org/10.20535/2410-8286.197808

Chowdhury, M. M. H. (2024). Digital game-based language learning for vocabulary development.

International Journal of Educational Research Open, 5, 1-8.

https://doi.org/10.1016/j.caeo.2024.100160

Christanti, E. J. (2018). *The use of Quizlet as a learning medium to improve English vocabulary mastery for eleventh grade in SMAN4 Yogyakarta* [Undergraduate thesis, Sanata Dharma University]. http://repository.usd.ac.id/id/eprint/30915

Creswell, J. W. (2007). *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. Sage Publications.

De Groot, A. M. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language learning*, *56*(3), 463-506. https://doi.org/10.1111/j.1467-9922.2006.00374.x

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior.* Springer.

Diem Thi Hao Tam and Tran Thi Xuan (2023). Thực Trạng Học Từ Vựng Tiếng Anh Của Sinh Viên Trường Đại Học Kỹ Thuật Y Tế Hải Dương. *Tạp Chí Y Học Việt Nam*, *528*(2), 293 - 294. https://doi.org/10.51298/vmj.v528i2.6134

Hayakawa, S., Ginther, A., & Hirsh, D. (2019). Effects of initial exposure on L2 vocabulary retention: A review of empirical research. *Language Learning, 69*(S1), 149–185.

Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign Language*, *17*(1), 23-59.

Jackson III, D. B. (2015). A targeted role for L1 in L2 vocabulary acquisition with mobile learning technology. *TESOL Arabia Perspectives, 23*(1), 6-11.

Jomaa, N., Attamimi, R., & Al Mahri, M. (2024). Utilising Artificial Intelligence (AI) in vocabulary learning by EFL Omani students: The effect of age, gender, and level of study. *Forum for Linguistic Studies, 6*(5), 171–186.

Kacena, M. A., Plotkin, L. I., & Fehrenbacher, J. C. (2024). The use of artificial intelligence in writing scientific review articles. *Current Osteoporosis Reports, 22*(1), 115–121. https://doi.org/10.1007/s11914-023-00852-0

Li, M. (2025). Mapping academic motivation, self-efficacy, achievement emotions, and vocabulary learning in a game-enhanced learning environment from the lens of activity theory. *Learning and Motivation, 89*, 102087. https://doi.org/10.1016/j.lmot.2024.102087

Lougheed, L. (2014). *Essential Words for the IELTS*. Barrons' Publisher.

Marshall, C., & Rossman, G. B. (2016). *Designing Qualitative Research* (6th ed.). Sage Publications.

Marzuki, Utami, W., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education, 10*(2), 1-17. https://doi.org/10.1080/2331186X.2023.2236469

Mohammadi, Mahtab, et al. (2024). University Students' Academic Vocabulary Development through Mobile-Assisted Learning: Exploring the Impacts on Receptive and Productive Knowledge. *Heliyon*, *10* (7), , 1-11. https://doi.org/10.1016/j.heliyon.2024.e28103

Nguyen Quoc Long and Le Van Ha (2022). Quizlet as a learning tool for enhancing L2 learners' Lexical retention: Should it be used in class or at home?. *Human Behavior and Emerging Technologies, 2022,* 1–10. https://doi.org/10.1155/2022/8683671

Nguyen Thi Kim Ngoc (2024). The impact of Quizlet on vocabulary achievement among non-English major students at Dong Nai Technology University. *International Journal of English Language Studies*, *6*(2), 57-61. https://doi.org/10.32996/ijels.2024.6.2.8

Putz, L. M., Hofbauer, F., & Treiblmaier, H. (2020). Can gamification help to improve education? Findings from a longitudinal study. *Computers in Human Behavior, 110*(5), 1-5. https://doi.org/10.1016/j.chb.2020.106392

Sadeghi, K., Panahbehagh, B., & Jalilzadeh, K. (2022). The effects of implementing gamified instruction on vocabulary development and motivation. *Heliyon, 8*(6), e103099. https://doi.org/10.1016/j.heliyon.2022.e11811

Sari, D. E., Fitriani, S. A., & Saputra, R. C. (2020). Active and interactive learning through Quizlet and Kahoot. In *Proceedings of the International Conference on Online and Blended Learning 2019 (ICOBL 2019)* (Vol. 440, 118–120). Atlantis Press. *Advances in Social Science, Education and Humanities Research.* https://doi.org/10.2991/assehr.k.200521.025

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal*, *9*(1), 4-9. https://doi.org/10.1002/j.1949-3533.2000.tb00220.x

Schneider, V. I., Healy, A. F., & Bourne Jr, L. E. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, *46*(2), 419-440. https://doi.org/10.1006/jmla.2001.2813

Suharno, S., Suherdi, D., & Gunawan, W. (2023). The Effects of Teaching Presence on Students' Motivation and Performance in A Long-term Online Gamified EFL Listening Course. *International Journal of Instruction, 16*(2), 1111–1134. https://doi.org/10.29333/iji.2023.16259a

Surmanov, S., & Azimova, M. (2020). Analysis of difficulties in vocabulary acquisition. *The Journal of Legal Studies, 6*(1), 144-153.

Tosun, Sezen (2015). The Effects of Blended Learning on EFL Students' Vocabulary Enhancement. *Procedia - Social and Behavioral Sciences, 199,* 641–647, https://doi.org/10.1016/j.sbspro.2015.07.592

Tran My Dung and Nguyen Thi Minh Loan (2022). Improving Students' Vocabulary Retention Through Using Quizlet: An Action Research At High School. *TNU Journal of Science and Technology*, *227*(4), 170–177. https://doi.org/10.34238/tnu-jst.5647

Trinh Phuong Hoa and Ho Thi Bach Mai (2023). Quizlet's impacts on EMI students' vocabulary mastery and lesson comprehension at Hanoi University of Industry: A student perspective. *Khoa Học Công Nghệ/Khoa Học Và Công Nghệ*, *59*(6B), 170-174. https://doi.org/10.57001/huih5804.2023.241

Waluyo, B., & Bucol, J. L. (2021). The impact of gamified vocabulary learning using Quizlet on low-proficiency students. *Computer-Assisted Language Learning, 22*(1), 164-185.

Web, S. (2008). The Effect of Context on Incidental Vocabulary Learning. *Reading In Foreign Language, 20*(2), 232-245.

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, 1-14 https://doi.org/10.3389/fpsyg.2023.1261955

Whitman, A. C., Tanzer, K., & Nemec, E. C. (2019). Gamifying the memorization of brand/generic drug names. *Currents in Pharmacy Teaching and Learning, 11*(3), 287-291. https://doi.org/10.1016/j.cptl.2018.12.014

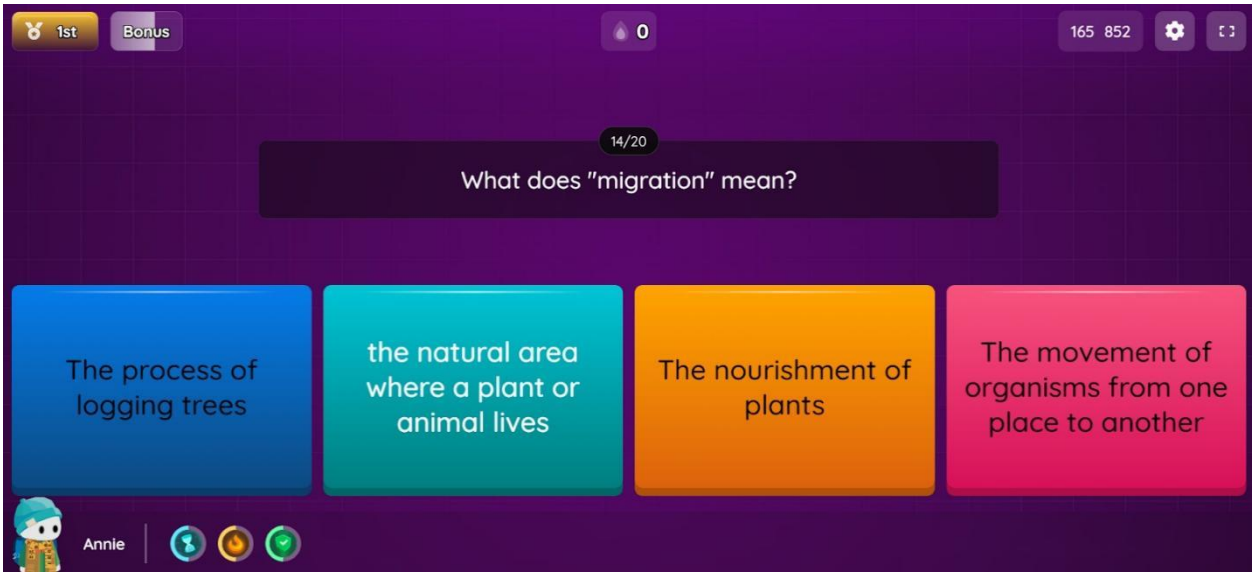
APPENDIX A: INTERVIEW QUESTIONS

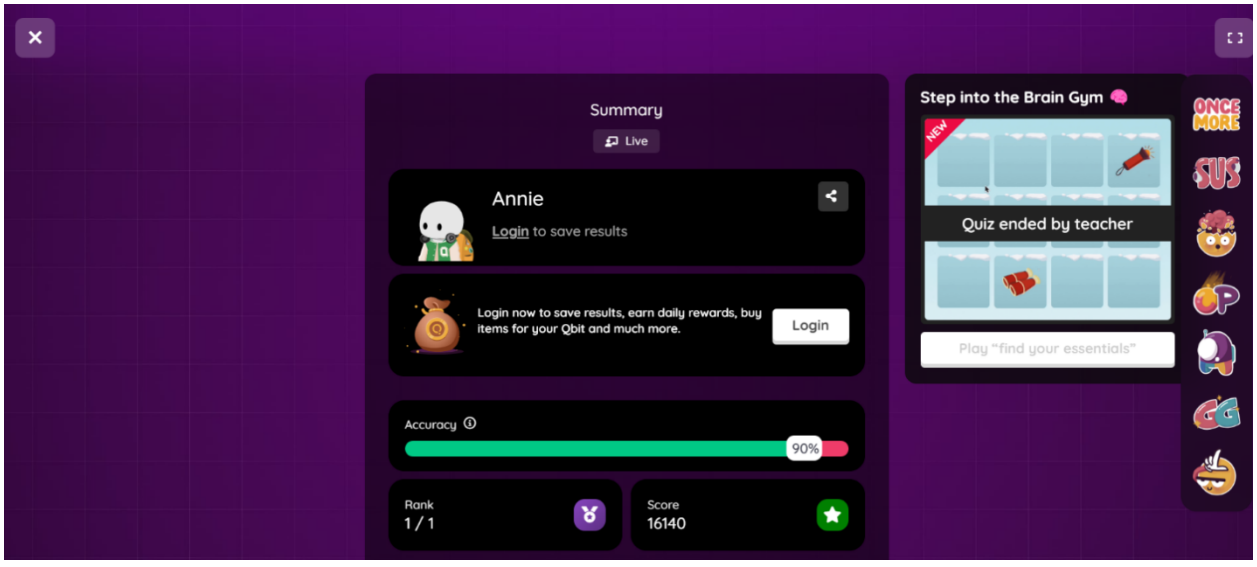
Interview Questions

1. What's your name?
2. Are you in what year of college?
3. How many hours per week do you usually spend learning vocabulary?
4. Can you describe your typical method for studying and learning English?
5. What specific strategies or resources did you use to improve your vocabulary learning?
6. Did you use Quizlet before joining the IELTS class? If yes, how did you use it?
7. How many hours per week do you usually use Quizlet during the IELTS course?
8. How has your approach to studying and acquiring new vocabulary changed since you started using Quizlet following the class?
9. How do you feel Quizlet has affected your vocabulary acquisition and recall compared to your previous methods of study?
10. How has your confidence and ability to learn English vocabulary changed since you employed Quizlet in your study routine?
11. In your opinion, what are the key benefits and limitations of traditional methods of vocabulary learning compared to using digital platforms as Quizlet?
12. Based on your experience with Quizlet thus far, what advice or comments would you provide to other students or educators about integrating it in English programs?

APPENDIX B

Quizlet Interface Examples – Vocabulary Questions





APPENDIX C

Pictures of IELTS students engaged in classroom activities.

Authors' information:

Dr. Truong Thi Nhu Ngoc is a full-time English lecturer at Ho Chi Minh City University of Industry and Trade. Her research interests include second language acquisition, self-efficacy, and self-regulated learning.

Nguyen Tu Nhan is a final-year English major at Ho Chi Minh City University of Industry and Trade. She is interested in researching vocabulary and speaking skills.

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TẠO SINH TRONG DẠY HỌC TIẾNG ANH NHƯ NGÔN NGỮ THỨ HAI

THE APPLICATION OF GENERATIVE ARTIFICIAL INTELLIGENCE IN TEACHING ENGLISH AS A SECOND LANGUAGE

Nguyễn Thị Mỹ Lệ [1]

TÓM TẮT: Những công cụ trí tuệ nhân tạo tạo sinh (Generative Artificial Intelligent, GenAI) không chỉ cải thiện độ chính xác và khả năng tự học, mà còn giúp người học vượt qua rào cản lo âu giao tiếp, phát triển kỹ năng sử dụng tiếng Anh một cách linh hoạt. Trên cơ sở xác lập khái niệm, phân loại nhóm GenAI, bài viết phân tích một số ứng dụng GenAI trong dạy học tiếng Anh như một ngôn ngữ thứ hai (ESL). Ngoài ra, chúng tôi cũng đề xuất một số giải pháp hiệu quả về ứng dụng GenAI trong dạy học tiếng Anh và lộ trình từng bước tích hợp GenAI vào chương trình học nhằm giúp học sinh tiếp cận tiếng Anh như một ngôn ngữ thứ hai một cách tự nhiên trong bối cảnh toàn cầu hóa hiện nay.

Từ khoá: Trí tuệ nhân tạo tạo sinh, tiếng Anh như ngôn ngữ thứ hai, cá nhân hóa việc học tập, kỹ năng giao tiếp, giáo dục cấp phổ thông

Abstract: GenAI tools not only improve learners' accuracy and self-directed learning but also help them overcome communication anxiety and develop flexible language skills. Based on clarifying key concepts and classifying GenAI categories, this paper analyses several applications of GenAI in teaching English as a second language (ESL) effective GenAI applications, propose some effective solutions for applying GenAI in teaching English and a structured roadmap step for step -by- step to integrate GenAI into the curriculum to help students approach English as a second language naturally in the context of globalization today.

Key words: Generative Artificial Intelligence; English as a Second Language; Personalization of learning; Communication Skills; Secondary education

1. Đặt vấn đề

Trong bối cảnh toàn cầu hóa và chuyển đổi số hiện nay, năng lực sử dụng tiếng Anh như một ngôn ngữ thứ hai (L2) ngày càng được nhìn nhận không chỉ phục vụ học tập và nghề nghiệp, mà còn là điều kiện cần thiết để tham gia hiệu quả vào đời sống công dân toàn cầu (Crystal, 2012). Vì vậy, việc dạy và học tiếng Anh như ngôn ngữ thứ hai trong các trường học đã trở thành một vấn đề được quan tâm rộng rãi, không chỉ ở Việt Nam mà còn trên toàn thế giới. Phát biểu tại cuộc họp chiều 10/4, tại Bộ Giáo dục và Đào tạo, Thứ trưởng Phạm Ngọc Thưởng cho biết, từng bước đưa tiếng Anh trở thành ngôn ngữ thứ hai trong nhà trường là một trong những nhiệm vụ trọng tâm được Bộ Chính trị đề ra tại Kết luận số 91-KL/TW, trong đó nêu rõ: "*Tập trung nâng cao năng lực ngoại ngữ của học sinh, sinh viên, từng bước đưa tiếng Anh trở thành ngôn ngữ thứ 2 trong trường học...*".

Các phương pháp giảng dạy truyền thống trong các trường học hiện nay mặc dù vẫn có giá trị, song vẫn chưa đáp ứng được hết nhu cầu đa dạng của người học, nhất là trong việc cá nhân hóa, nâng cao năng lực giao tiếp và duy trì sự hứng thú (Harmer, 2015). Bên cạnh đó, môi trường lớp học thường không đủ để tạo ra một không gian thực hành tự nhiên, khiến người học thiếu cơ hội

[1] TS, Trường Đại học Đông Á, Đại học Đà Nẵng, *Email: nguyenmyleqn@gmail.com*

tương tác và lo âu giao tiếp. Việc cá nhân hóa lộ trình học tập cho từng học sinh cũng là một thách thức lớn đối với giáo viên, đặc biệt trong các lớp học có sĩ số đông, năng lực sử dụng tiếng Anh của học sinh còn chênh lệch nhiều.

Trí tuệ nhân tạo tạo sinh (Generative AI - GenAI) là một lĩnh vực của trí tuệ nhân tạo, với khả năng tạo ra các nội dung độc đáo và đa dạng, đang tác động mạnh mẽ đến nhiều lĩnh vực, trong đó có giáo dục. GenAI đang thay đổi cách người học tiếp cận tiếng Anh như một ngôn ngữ thứ hai (Second Language – L2) một cách đáng kể. Trong việc dạy học tiếng Anh, các công cụ GenAI có thể tạo ra những môi trường học tập sống động, mô phỏng các tình huống giao tiếp thực tế, từ đó giúp người học luyện tập và làm quen với ngôn ngữ một cách tự nhiên. Đồng thời, GenAI cho phép cá nhân hóa sâu sắc trải nghiệm học tập, tạo ra các tài liệu và bài tập phù hợp với trình độ, sở thích và mục tiêu của từng học sinh. Hơn nữa, khả năng cung cấp phản hồi tức thì và chính xác của GenAI giúp người học nhanh chóng nhận biết và khắc phục lỗi sai, thúc đẩy quá trình tiếp thu ngôn ngữ.

Từ những lý do trên, trong phạm vi bài viết này, chúng tôi bàn đến một số ứng dụng GenAI hiệu quả trong dạy học tiếng Anh và tiềm năng ứng dụng của GenAI từ góc độ giáo dục học mà giáo viên và học sinh có thể sử dụng các công cụ GenAI có sẵn trong dạy học tiếng Anh. Ngoài ra, bài viết còn đề xuất lộ trình từng bước để tích hợp GenAI vào chương trình dạy học tiếng Anh như một ngôn ngữ thứ hai đáp ứng với xu thế hội nhập hiện nay.

2. Cơ sở lý luận

2.1. Trí tuệ nhân tạo tạo sinh

Trí tuệ Nhân tạo Tạo sinh (Generative AI - GenAI) là một nhánh của Trí tuệ Nhân tạo (AI) tập trung vào việc phát triển các mô hình có khả năng tạo ra nội dung mới, độc đáo từ dữ liệu đã được học. Không giống như các mô hình AI truyền thống phân loại hoặc dự đoán, vốn chỉ phân tích và đưa ra kết quả dựa trên dữ liệu đầu vào có sẵn, GenAI có thể tạo ra các sản phẩm sáng tạo, độc đáo, bao gồm văn bản, hình ảnh, âm thanh và mã nguồn.

Công nghệ trí tuệ nhân tạo tạo sinh được tích hợp vào chatbot, chương trình kết hợp với trí tuệ nhân tạo (AI) để tương tác với con người. Các chatbot như ChatGPT của OpenAI, BingAI, BLOOM, ELSA Speak, Copilot, … ngày càng hiện đại, nó đã trải qua một chặng đường dài trước khi trở nên "thông minh" như ChatGPT của OpenAI.

Mô hình ngôn ngữ lớn (Large Language Models - LLMs) là một trong những ứng dụng GenAI nổi bật nhất và có liên quan trực tiếp đến lĩnh vực dạy học ngôn ngữ. Các LLMs như GPT (Generative Pre-trained Transformer) được xây dựng trên kiến trúc Transformer và được huấn luyện trên một lượng dữ liệu văn bản khổng lồ, bao gồm sách, bài báo, trang web, v.v. Nhờ đó, LLMs có khả năng thực hiện nhiều tác vụ phức tạp liên quan đến ngôn ngữ, bao gồm tóm tắt văn bản, dịch thuật, viết lại, và quan trọng nhất là tạo ra các phản hồi hội thoại một cách tự nhiên.

Chatbot là một ứng dụng phổ biến của LLMs, cho phép người dùng tương tác với

AI thông qua giao tiếp bằng văn bản hoặc giọng nói. Trong dạy học tiếng Anh, các chatbot hoạt động như những người bạn đồng hành ảo, có thể duy trì các cuộc hội thoại phức tạp, trả lời câu hỏi,

và cung cấp nội dung thực hành ngôn ngữ ngay lập tức. Chúng tạo ra một môi trường tương tác phi phán xét, khuyến khích người học thực hành và thử nghiệm ngôn ngữ mà không sợ mắc lỗi.

Tóm lại, với sự ra đời của LLMs hiện đại, AI tạo sinh đã có nhều bước tiến vượt bậc, tích hợp nhiều lĩnh vực với giáo dục trong việc hỗ trợ và cung cấp nền tảng công nghệ cho các chatbot trí tuệ nhân tạo tạo sinh như ChatGPT của OpenAI, Copilot, Gemini, Claude, DALL - E 2, Midjourney, Leonardo Ai, … để hỗ trợ viết, tạo hình ảnh, tạo mã máy tính; một số chatbot GenAI được ứng dụng để học ngôn ngữ (Duolingo, Babbel, Rosetta Stone); luyện phát âm (ELSA Speak); dịch thuật (Google translate), Gemini – trí tuệ nhân tạo đa phương thức của Google.

2.2. Lý thuyết về học ngôn ngữ thứ hai (SLA – Second Language Acquisition)

2.2.1. Thuyết đầu vào (Input Hypothesis)

Thuyết đầu vào của Stephen Krashen cho rằng người học sẽ tiếp thu ngôn ngữ một cách tự nhiên khi họ được tiếp xúc với đầu vào dễ hiểu (comprehensible input). Krashen mô tả đầu vào này bằng công thức **i+1**, trong đó **i** là trình độ hiện tại của người học, và **+1** là một chút thách thức hơn, chứa các cấu trúc hoặc từ vựng mới mà người học chưa biết nhưng có thể suy luận được từ ngữ cảnh.

2.2.2. Thuyết tương tác (Interaction Hypothesis)

Thuyết tương tác của Michael Long khẳng định rằng đầu vào dễ hiểu là cần thiết, nhưng chỉ khi nó được tạo ra thông qua tương tác giao tiếp (interaction) thì quá trình tiếp thu ngôn ngữ mới thực sự hiệu quả. Khi người học và người bản ngữ tương tác, họ thường phải điều chỉnh ngôn ngữ của mình để đảm bảo đối phương hiểu được.

2.2.3. Thuyết đầu ra (Output Hypothesis)

Merrill Swain cho rằng không chỉ đầu vào và tương tác là đủ. Bà đề xuất thuyết đầu ra, trong đó nhấn mạnh vai trò của việc sản xuất ngôn ngữ (producing language), tức là nói hoặc viết. Swain lập luận rằng việc người học buộc phải tạo ra ngôn ngữ có ba chức năng chính: Chức năng nhận thức, Chức năng kiểm tra giả thuyết và Chức năng lưu loát.

2.2.4. Vai trò của phản hồi (Corrective Feedback)

Phản hồi hiệu chỉnh là sự phản ứng của người đối thoại (người bản ngữ hoặc giáo viên) đối với lỗi sai của người học. Phản hồi giúp người học nhận ra lỗi sai và điều chỉnh kiến thức của mình. Có nhiều dạng phản hồi khác nhau, như:

- *Phản hồi trực tiếp (Explicit feedback):* Giáo viên chỉ ra lỗi sai một cách rõ ràng và cung cấp câu trả lời đúng. Ví dụ: "Không phải *'I goed'*, mà phải là *'I went'."*

- *Phản hồi gián tiếp (Implicit feedback):* Giáo viên sửa lỗi một cách tinh tế hơn, thường thông qua việc lặp lại câu nói của học sinh một cách chính xác. Ví dụ:

+ Học sinh: "He **goed** to the market."

+ Giáo viên: "Yes, he **went** to the market."

3. Một số ứng dụng trí tuệ nhân tạo tạo sinh hiệu quả trong dạy học tiếng Anh

Hiện nay, các công cụ GenAI rất đa dạng và được phân loại theo nhiều nhóm dựa trên các lĩnh vực ứng dụng và chức năng. Dưới đây là một số nhóm công cụ GenAI được ứng dụng trong việc phát triển nội dung học liệu và rèn luyện các kỹ năng cho học sinh.

Chức năng	**Công cụ**	**Ứng dụng thực tế**
Tạo sinh văn bản	ChatGPT, Gemini, Copilot, Perplexity AI	- Tạo ra các đoạn văn, bài giảng, bài luận, truyện ngắn, đối thoại, vai thoại, bài đọc, hoặc podcast tùy chỉnh với nhiều giọng điệu, tốc độ, và ngữ điệu khác nhau theo chủ đề và trình độ mong muốn. - Soạn thảo các câu hỏi trắc nghiệm, bài tập điền từ, hoặc bài tập ngữ pháp dựa trên nội dung cụ thể. - Tạo ra các ý tưởng cho bài luận, bài viết sáng tạo, hoặc các dự án học thuật khác.
Tạo sinh hình ảnh	Midjourney, Leonardo Ai DALL-E	- Minh họa cho các câu chuyện, bài đọc, hoặc từ vựng để học sinh dễ hình dung và ghi nhớ. - Tạo ra các bức tranh, biểu đồ, hoặc sơ đồ tư duy trực quan cho các chủ đề phức tạp.
Tạo sinh âm thanh và video	HeyGen, ELSA Speak, Murf AI, Speechase ChatGPT, Gemini	- Chuyển văn bản thành giọng nói *(Text-to-Speech)* với nhiều ngữ điệu và tốc độ, phục vụ cho việc tạo podcast hoặc tài liệu nghe tùy chỉnh. - Cung cấp phản hồi ngay lập tức, nhận diện các lỗi sai và đưa ra gợi ý sửa chữa cụ thể.
Tạo sinh trò chơi và hoạt động tương tác	ClassPoint, ClassDojo, Kahoot!	- Tạo các trò chơi đố vui, để ôn tập từ vựng và ngữ pháp một cách sinh động. - Thiết kế các kịch bản tương tác để học sinh thực hành giao tiếp trong các tình huống cụ thể (ví dụ như mua sắm, hỏi đường).

4. Một số nguyên tắc và giải pháp ứng dụng GenAI vào dạy học tiếng Anh như ngôn ngữ thứ hai

4.1. Nguyên tắc ứng dụng GenAI vào dạy học tiếng Anh như ngôn ngữ thứ hai

Việc tích hợp GenAI vào giảng dạy tiếng Anh đòi hỏi tuân thủ các nguyên tắc nhất định nhằm tối ưu hóa hiệu quả sư phạm và đảm bảo tính đạo đức. Vì vậy, chúng ta có thể sử dụng khung IDEE của Su, J., & Yang, W. (2023) làm cơ sở lí luận trong ứng dụng GenAI vào dạy học tiếng Anh như ngôn ngữ thứ hai. Khung IDEE bao gồm bốn giai đoạn chính, đó là: 1) Xác định rõ mục

tiêu học tập cụ thể (Identify the Desired Outcomes): Việc xác định mục tiêu rõ ràng giúp giáo viên đưa ra câu lệnh (prompts) chính xác hơn và đánh giá kết quả hiệu quả hơn; 2) Xác định mức độ tự động hóa phù hợp (Determine the Appropriate Level of Automation): Cần can thiệp để điều chỉnh phương pháp, nội dung cho phù hợp với ngữ cảnh lớp học và trình độ của học sinh; 3) Đảm bảo cân nhắc về đạo đức (Ensure Ethical Considerations): Cần dạy học sinh kỹ năng tư duy phản biện và xác minh thông tin, các tác động đạo đức của việc sử dụng GenAI trong giáo dục; 4) Đánh giá hiệu quả (Evaluate the Effectiveness): Cần đánh giá xem liệu mục tiêu ban đầu có đạt được hiệu quả hay không.

Hai nguyên tắc cụ thể về yếu tố văn hóa và cách sử dụng câu lệnh (prompts) là yếu tố then chốt quyết định hiệu quả và tính phù hợp của việc ứng dụng công nghệ này.

4.1.1. Yếu tố văn hóa (Cultures)

GenAI, đặc biệt là các mô hình ngôn ngữ lớn, được huấn luyện trên lượng dữ liệu khổng lồ từ khắp nơi trên thế giới. Điều này có thể dẫn đến việc các nội dung được tạo ra mang nặng định kiến văn hóa hoặc không phù hợp với bối cảnh của người học. Ví dụ, GenAI có thể tạo ra các kịch bản hội thoại về các ngày lễ như Lễ Tạ Ơn (Thanksgiving) hay Lễ Giáng Sinh (Christmas) mà không có sự cân nhắc về các ngày lễ truyền thống tại Việt Nam như Tết Nguyên Đán, Giỗ tổ Hùng Vương, … . Ngoài ra, một số nội dung do GenAI tạo ra có thể chứa các thành ngữ, tục ngữ hoặc ẩn ý văn hóa mà người học không quen thuộc, điều này có thể gây khó khăn trong việc tiếp thu và sử dụng ngôn ngữ một cách tự nhiên. Vì vậy, giáo viên cần kiểm tra và điều chỉnh các nội dung do GenAI tạo ra để đảm bảo chúng phù hợp về mặt văn hóa, tôn trọng sự đa dạng và mang tính toàn cầu.

4.1.2. Cách sử dụng câu lệnh (Prompts)

Cách người sử dụng đưa ra câu lệnh (prompts) có ảnh hưởng trực tiếp đến kết quả mà GenAI tạo ra. Một câu lệnh tốt sẽ tạo ra nội dung phù hợp, chính xác và có giá trị sư phạm cao, trong khi một câu lệnh kém có thể dẫn đến kết quả không mong muốn. Khi sử dụng GenAI trong việc tạo sinh nội dung dạy học, giáo viên cần biết cách tùy chỉnh các câu lệnh để khai thác đúng tiềm năng của AI, bám sát mục tiêu bài dạy, phù hợp đối tượng người học. Vì thế, giáo viên cần đưa ra các câu lệnh cụ thể, rõ ràng, thông minh để GenAI hiểu đúng yêu cầu. Câu lệnh cũng làm rõ các nhân tố giao tiếp (Vai trò, Bối cảnh, Mục tiêu, Nhiệm vụ). Ngoài ra, giáo viên có thể thêm các yêu cầu liên quan đến văn hóa vào câu lệnh để định hướng GenAI. Ví dụ: "*Tôi là giáo viên dạy tiếng Anh ở trường phổ thông. Hãy gợi ý cho tôi một đoạn văn khoảng 100 từ mô tả món phở truyền thống của Việt Nam. Sử dụng các tính từ miêu tả, văn hóa Việt Nam, văn phong gần gũi với học sinh*".

Tóm lại, việc hiểu và làm chủ kỹ năng sử dụng câu lệnh (prompt engineering) giúp giáo viên tối ưu hóa công cụ này cho mục tiêu dạy và học tiếng Anh.

4.1.3. Cá nhân hóa người học (Personalization of learners)

Trước hết, giáo viên cần nắm rõ thông tin về trình độ tiếng Anh hiện tại của người học (sơ cấp, trung cấp, cao cấp); mục tiêu học tập (học tiếng Anh cho học tập, công việc, để giao tiếp); sở thích (trong lĩnh vực khoa học, văn hóa, lịch sử, kinh tế, nghệ thuật,…). Bằng cách xác định và

hiểu rõ những yếu tố này, giáo viên có thể sử dụng GenAI để điều chỉnh nội dung bài giảng, tạo ra bài tập cá nhân hoá dành riêng cho từng học sinh.

4.1.4. Tích hợp GenAI như một công cụ đa phương tiện (Multimedia tools)

GenAI có thể tạo ra nhiều dạng nội dung khác nhau, từ văn bản đến âm thanh và hình ảnh. Việc kết hợp các công cụ GenAI này sẽ tạo ra môi trường học tập phong phú và đa chiều. Chẳng hạn, một bài học có thể bao gồm việc sử dụng GenAI để tạo ra một đoạn hội thoại, sau đó chuyển nó thành file âm thanh để luyện nghe và phát âm, hoặc tạo ra hình ảnh minh họa cho các từ vựng mới.

4.2. Giải pháp ứng dụng trí tuệ nhân tạo tạo sinh trong dạy học tiếng Anh như ngôn ngữ thứ hai

4.2.1. Soạn thảo nội dung học liệu

a. Soạn giảng

Các chatbot dựa trên mô hình ngôn ngữ lớn (LLMs) như ChatGPT, Gemini, Copilot, Perplexity AI có thể hỗ trợ giáo viên tạo dàn ý, mục tiêu bài học, và các hoạt động phù hợp với khung năng lực ngôn ngữ CEFR. Giáo viên chỉ cần nhập yêu cầu cụ thể, hệ thống sẽ đưa ra gợi ý chi tiết.

b. Soạn thảo ngữ liệu học tập giúp học sinh rèn luyện kĩ năng

- Các chatbot dựa trên mô hình ngôn ngữ lớn (LLMs) như GPT-4, Gemini; các công cụ chuyển văn bản thành giọng nói (text-to-speech) như ElevenLabs, Murf AI; và các công cụ tạo video từ văn bản hoặc hình ảnh như Heygen, ELSA speak sẽ là công cụ đắc lực hỗ trợ cho giáo viên trong việc soạn thảo ngữ liệu.

- Tạo kịch bản âm thanh/video nghe, soạn thảo câu hỏi trắc nghiệm hoặc điền từ vào chỗ trống và tạo phụ đề cho bài nghe: Sau khi tích hợp các công cụ GenAI, giáo viên có một tệp âm thanh hoặc video ngắn có nội dung tự nhiên, kèm theo các bài tập đi kèm để đánh giá mức độ hiểu của học viên.

- Yêu cầu học sinh kể một câu chuyện ngắn bằng tiếng Anh: giáo viên tích hợp Midjourney hoặc DALL-E để tạo các hình ảnh minh họa cho câu chuyện, đoạn hội thoại. Sau đó, dùng HeyGen để tạo một video kể chuyện, giúp câu chuyện trở nên trực quan và hấp dẫn hơn, đồng thời luyện tập kỹ năng phát âm và trình bày.

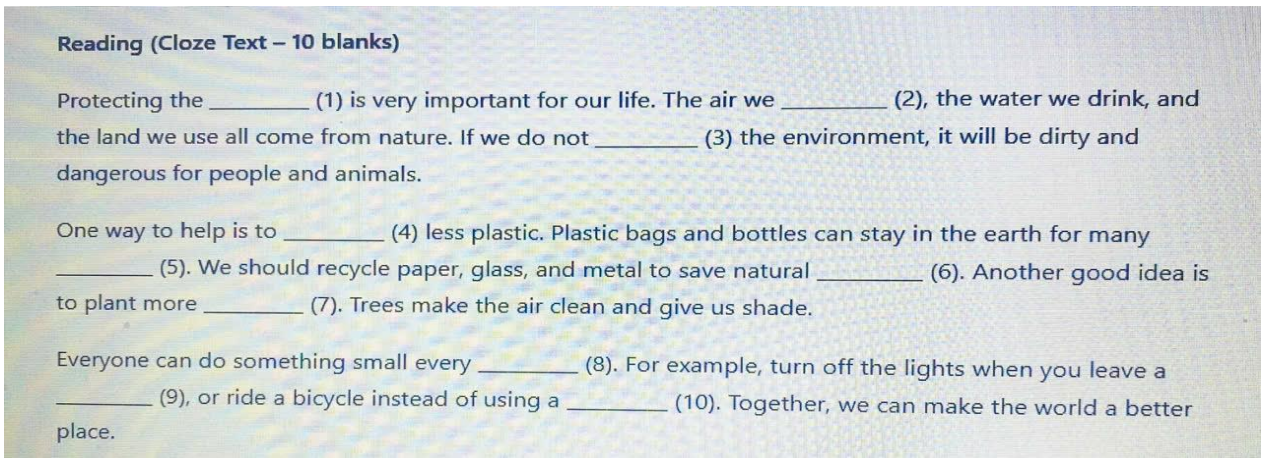
- Sử dụng công cụ GenAI: Các LLMs như GPT-4, Gemini để tạo ra các đoạn văn, bài báo hoặc câu chuyện theo cấp độ (từ A1 đến C2), soạn các bài tập đọc hiểu (trả lời câu hỏi, tóm tắt, tìm ý chính), và tạo từ vựng khó kèm theo định nghĩa với chủ đề đa dạng

- Ngoài ra, giáo viên có thể sử dung GenAI để tạo các bài viết mẫu, luyện viết theo chủ đề và cung cấp phản hồi chi tiết. Kết quả, giáo viên có một bài đọc hoàn chỉnh, phù hợp với trình độ học viên, kèm theo các bài tập luyện kĩ năng đọc hiểu và mở rộng vốn từ.

Ví dụ câu lệnh (1): *I am an English teacher. Please help me write an English reading comprehension passage of about 120 words about the importance of environmental protection, for*

A2-level students. Please design a Cloze test with 10 blanks, where one word should be filled in each blank.

Sau đây là kết quả từ chatbot Gemini:



Reading (Cloze Text – 10 blanks)

Protecting the _____ (1) is very important for our life. The air we _____ (2), the water we drink, and the land we use all come from nature. If we do not _____ (3) the environment, it will be dirty and dangerous for people and animals.

One way to help is to _____ (4) less plastic. Plastic bags and bottles can stay in the earth for many _____ (5). We should recycle paper, glass, and metal to save natural _____ (6). Another good idea is to plant more _____ (7). Trees make the air clean and give us shade.

Everyone can do something small every _____ (8). For example, turn off the lights when you leave a _____ (9), or ride a bicycle instead of using a _____ (10). Together, we can make the world a better place.

c. Soạn thảo bài tập luyện tập và kiểm tra

Các chatbot dựa trên mô hình ngôn ngữ lớn như ChatGPT của OpenAI, GPT-4, Gemini và nhiều hệ thống khác cung cấp cho giáo viên những công cụ đắc lực trợ như thiết kế bảng tóm tắt ngữ pháp, danh sách từ vựng theo chủ đề, hoặc các sơ đồ tư duy theo từng trình độ của học sinh. Cụ thể:

- Liệt kê 1 số từ vựng tích cực; giải thích nghĩa của từ vựng trong ngữ cảnh;

- Giải thích cách sử dụng của hiện tượng ngữ pháp; có thể được tùy chỉnh theo trình độ người học.

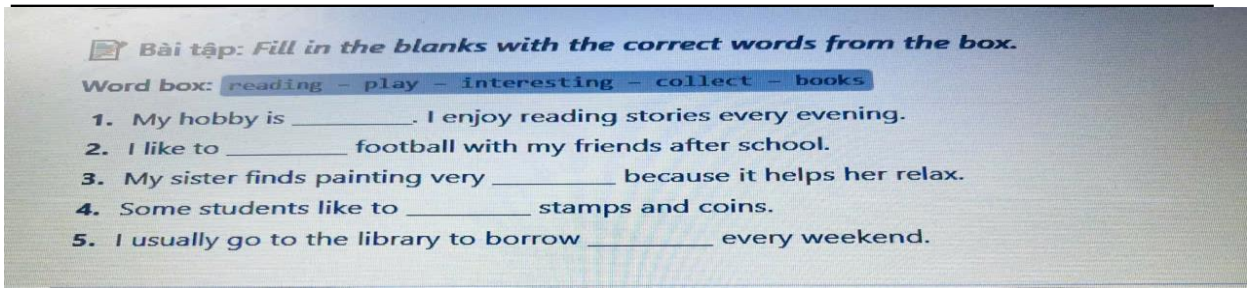
- Kết quả giải nghĩa từ có thể yêu cầu học sinh vận dụng qua các dạng bài tập như

Nối từ (Matching), Điền từ vào chỗ trống (Filling words in the gaps), ….

- Yêu cầu chatbot đặt câu có chứa từ vựng, ngữ pháp đó theo nhiều ngữ cảnh khác nhau để làm ví dụ minh hoạ. Thông qua những ví dụ này, học sinh sẽ hiểu và vận dụng được cách sử dụng từ vựng, ngữ pháp trong ngữ cảnh chính xác. Giáo viên có thể nhanh chóng tạo ra các phiên bản khác nhau của cùng một tài liệu, phù hợp với từng nhóm học viên có trình độ khác nhau, giúp tối ưu hóa quá trình học tập.

Ví dụ câu lệnh (2): *I am an English teacher. Please create a gap-filling exercise for me with 5 sentences and 5 vocabulary words (nouns, verbs, adjectives) in a box about the topic "My hobbies" for students. The content of the exercise should be suitable for A2-level students.*

Sau đây là kết quả từ chatbot ChatGPT:



4.2.2. Thiết kế bài tập để kiểm tra, đánh giá năng lực học sinh phổ thông

Giáo viên có thể tích hợp công cụ GenAI thiết kế bài tập để kiểm tra, đánh giá năng lực học sinh một cách nhanh chóng và phù hợp trình độ học sinh, tăng tính hấp dẫn và cá nhân hóa qua các dạng bài tập sau:

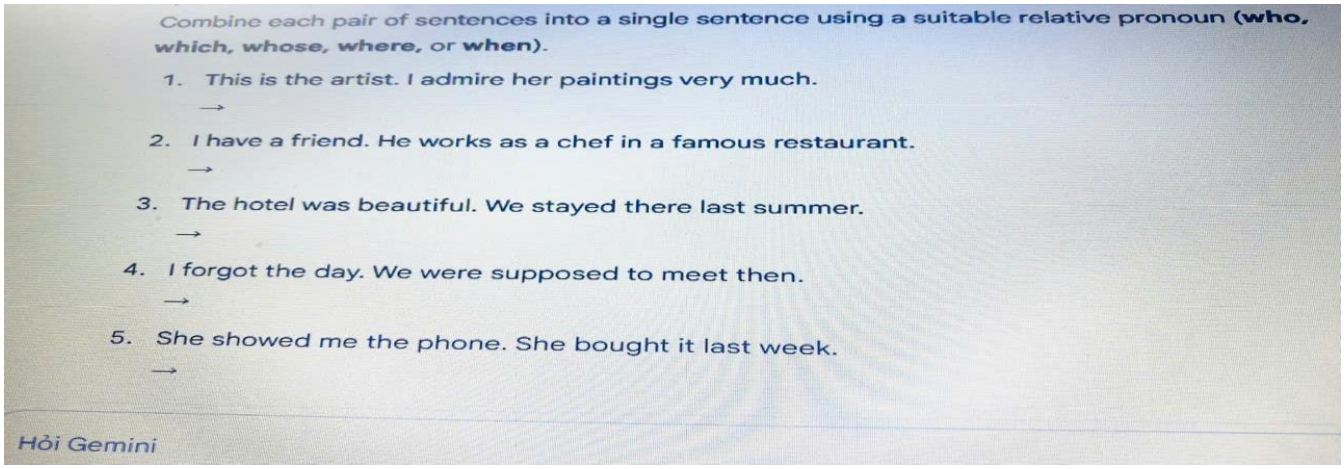
- Tạo bài tập từ vựng theo chủ đề đã học (môi trường, du lịch, sức khỏe) với mức độ phù hợp (A1–B2) như bài tập điền từ, nối từ với nghĩa, định nghĩa từ. Yêu cầu AI tạo bài tập ngắn (5–10 câu), sau đó giáo viên chỉnh sửa, kiểm tra tính chính, giao bài tập cho học sinh làm, thu thập phản hồi và điều chỉnh câu lệnh để lần tạo sinh tới sẽ tốt hơn.

- Thiết kế bài tập ngữ pháp đã học như chia động từ, hình thức của từ, câu bị động, mệnh đề quan hệ, biến đổi câu, … dưới dạng trắc nghiệm và tự luận. Giáo viên cần nêu rõ cấu trúc cần luyện, chọn hình thức bài kiểm tra (rewrite, multiple-choice, error correction).

- Đối với bài tập kỹ năng, giáo viên cần chọn chủ đề mà học sinh đã học để kiểm tra kỹ năng nghe, nói, đọc tiếp xúc với văn bản gần gũi thực tế, phát triển kỹ năng viết có định hướng, biết cách triển khai ý tưởng, xác định các thể loại văn bản qua hình thức văn miêu tả, thư, email, bài luận ngắn, ….

Ví dụ câu lệnh (3): *"I'm an English teacher. Please generate 1 grammar exercise for B1 students on relative clauses in combination format".*

Sau đây là kết quả từ chatbot Gemini:



4.2.3. Thiết kế các sản phẩm đa phương tiện bổ sung cho bài dạy tiếng Anh

Với trình độ công nghệ thông tin ở mức độ cơ bản, giáo viên có thể kết hợp ứng dụng các công cụ tạo sinh hình ảnh, âm thanh và công cụ tạo sinh văn bản. Sản phẩm đa phương tiện giúp

bài dạy tiếng Anh trở nên sinh động, trực quan và gần gũi, tạo động lực học tập mạnh mẽ cho học sinh. Khi kết hợp công cụ ChatGPT, Copilot hoặc Gemini, … giáo viên có thể tạo ra câu lệnh, sau đó sử dụng công cụ tạo sinh hình ảnh như Midjourney, Leornardo AI, Heygen, canva.com… để tạo được những hình ảnh, video, bản ghi âm, infographic, audio và flashcard nhanh chóng, phù hợp với từng mục tiêu bài học.

Giáo viên cần xác định chủ đề bài học (Tet holiday, *Shopping, Daily routines, Environment*); sử dụng ChatGPT, Copilot hoặc Gemini để viết nội dung chủ đề; sau đó ứng dụng Midjourney hoặc DALL-E để tạo các hình ảnh minh họa cho câu chuyện phù hợp với trình độ của học sinh và gần gũi với văn hóa của mỗi quốc gia. Để bài dạy hay hơn, giáo viên có thể dùng HeyGen để tạo một video kể chuyện, giúp câu chuyện trở nên trực quan và hấp dẫn hơn, đồng thời luyện tập kỹ năng phát âm và phong cách trình bày cho học sinh.

Ví dụ câu lệnh (4): *Please generate 2 pictures of about The Independece Day of Viet Nam: people wearing red shirts with yellow stars and holding red flags with yellow stars, 2. the helicopters are flying in the sky in Ha Noi.*

Sau đây là kết quả từ chatbot LeonardoAi:

4.2.4. Thiết kế các hoạt động dạy học và dự án dạy học

Các hoạt động dạy học (tương tác, nhóm, kỹ năng) và dự án học tập (poster, video) khi được tích hợp với AI sẽ giúp cho học sinh học tiếng Anh qua ngữ cảnh thực tế, tăng tính sáng tạo và tương tác trong lớp học. Ngoài ra, còn rèn kỹ năng hợp tác, thuyết trình, tư duy phản biện, đồng thời gắn kết tiếng Anh với đời sống và định hướng tương lai.

Có 4 hoạt động chính trong hoạt động dạy học, đó là: Khởi động, hình thành kiến thức, thực hàng và vận dụng.

- Hoạt động khởi động (Warm-up): Giáo viên có thể tích hợp AI để gợi ý trò chơi khởi động ngắn (5 phút) như đoán từ, ghép đôi, hoặc đố vui với từ vựng trọng tâm theo từng chủ đề bài học (*Food and Drink, Tourism, Environment*) nhằm tạo cho học sinh hứng thú, tập trung vào chủ đề trước khi học.

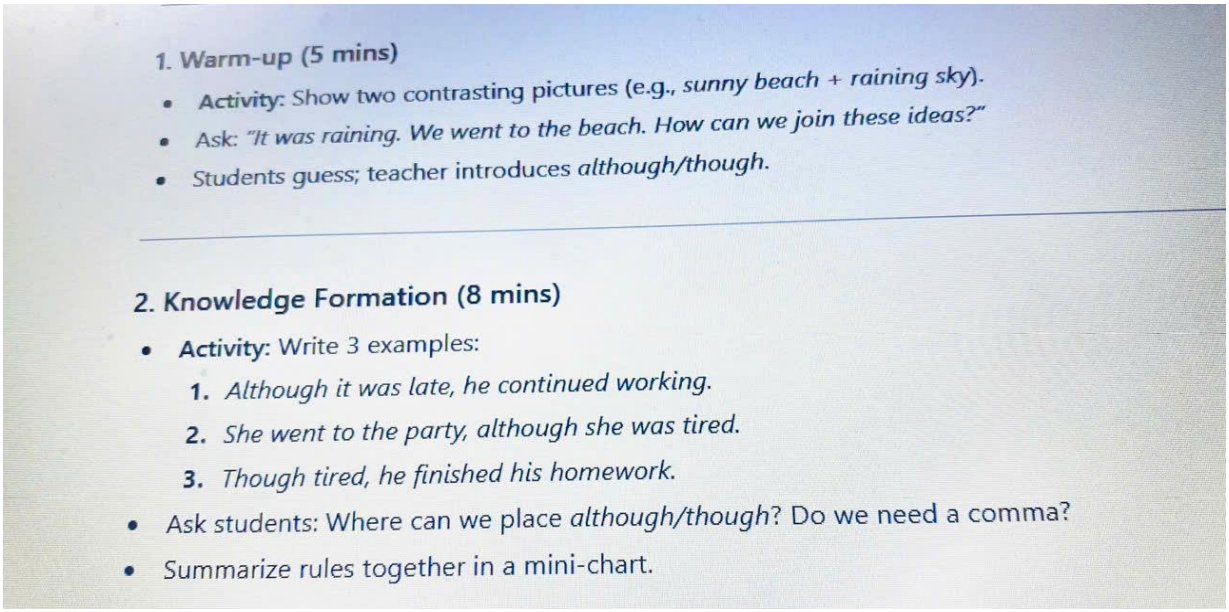
- Hoạt độnh hình thành kiến thức: Tích hợp AI để tạo ra một đoạn hội thoại, một vài câu hỏi gợi mở để dẫn dắt vào trọng tâm bài dạy nhằm tạo cho học sinh sự hứng thú, tập trung vào bài học.

- Hoạt động luyện tập: Dựa vào kiến thức vừa học và tùy thuộc vào trình độ của học sinh, AI có thể đưa ra một số bài tập cụ thể giúp học sinh luyện tập các kỹ năng Nghe, Nói, Đọc, Viết để phát triển năng lực giao tiếp, gắn bài học với tình huống đời sống.

- Hoạt động vận dụng: Các công cụ AI có thể đề xuất nhiệm vụ nhóm như: thảo luận, giải quyết vấn đề, trình bày những kiến thức đã học gắn với thực tiễn.

Ví dụ câu lệnh (5): *Please help me give some activities (Warm up, Knowledge formation, Practice, Application) to teach B1 students efectively in using with Although/Though.*

Sau đây là kết quả từ chatbot ChatGPT:

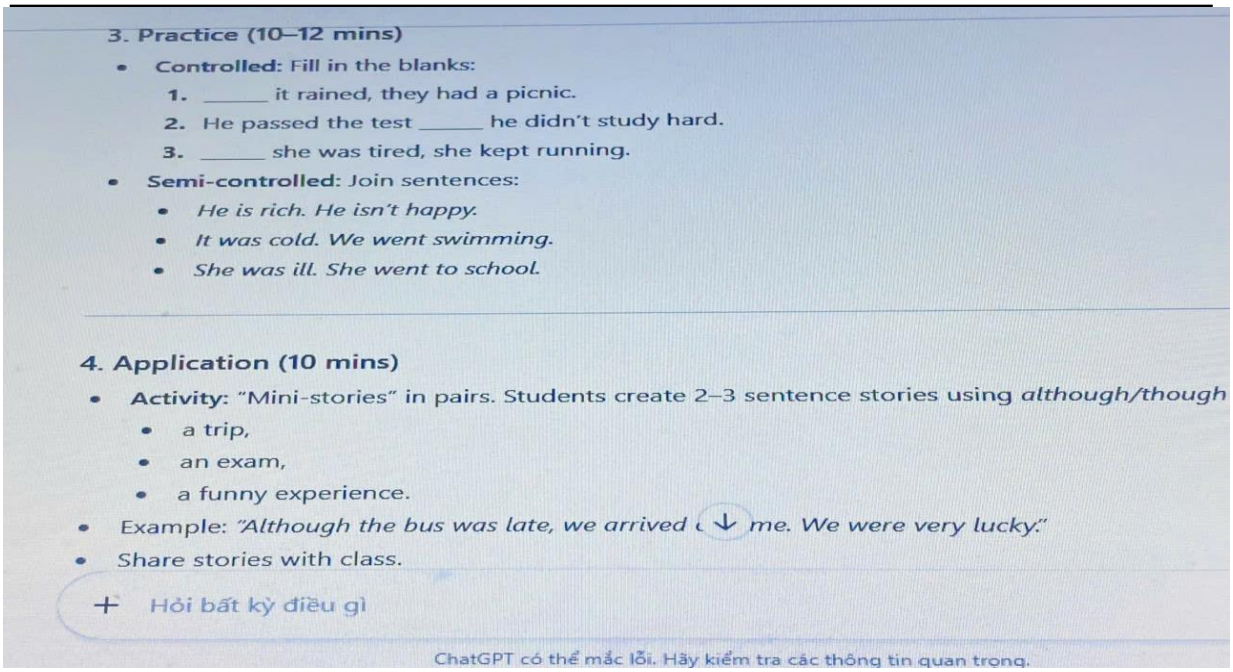


1. Warm-up (5 mins)

- **Activity:** Show two contrasting pictures (e.g., *sunny beach + raining sky*).
- Ask: *"It was raining. We went to the beach. How can we join these ideas?"*
- Students guess; teacher introduces *although/though*.

2. Knowledge Formation (8 mins)

- **Activity:** Write 3 examples:
  1. *Although it was late, he continued working.*
  2. *She went to the party, although she was tired.*
  3. *Though tired, he finished his homework.*
- Ask students: Where can we place *although/though*? Do we need a comma?
- Summarize rules together in a mini-chart.



4.2.5. Ứng dụng AI trong phản hồi và đánh giá năng lực học sinh

a. Phản hồi tức thì và cá nhân hóa

Chatbot AI có khả năng tự động phát hiện các lỗi ngữ pháp trong bài viết của học sinh và có thể phân tích giải thích các lỗi sai trong bài viết, đoạn văn hoặc câu trả lời của học sinh để đưa ra phản hồi về ngữ pháp, từ vựng, sự mạch lạc cũng như giúp giáo viêm chấm bài nhanh. Sau khi học sinh nộp đoạn văn ngắn hoặc bài viết, giáo viên sẽ nhập nội dung vào công cụ AI để AI đưa ra gợi ý cải thiện. Phần cuối, giáo viên chỉnh sửa và bổ sung nhận xét.

b. Đánh giá năng lực học sinh

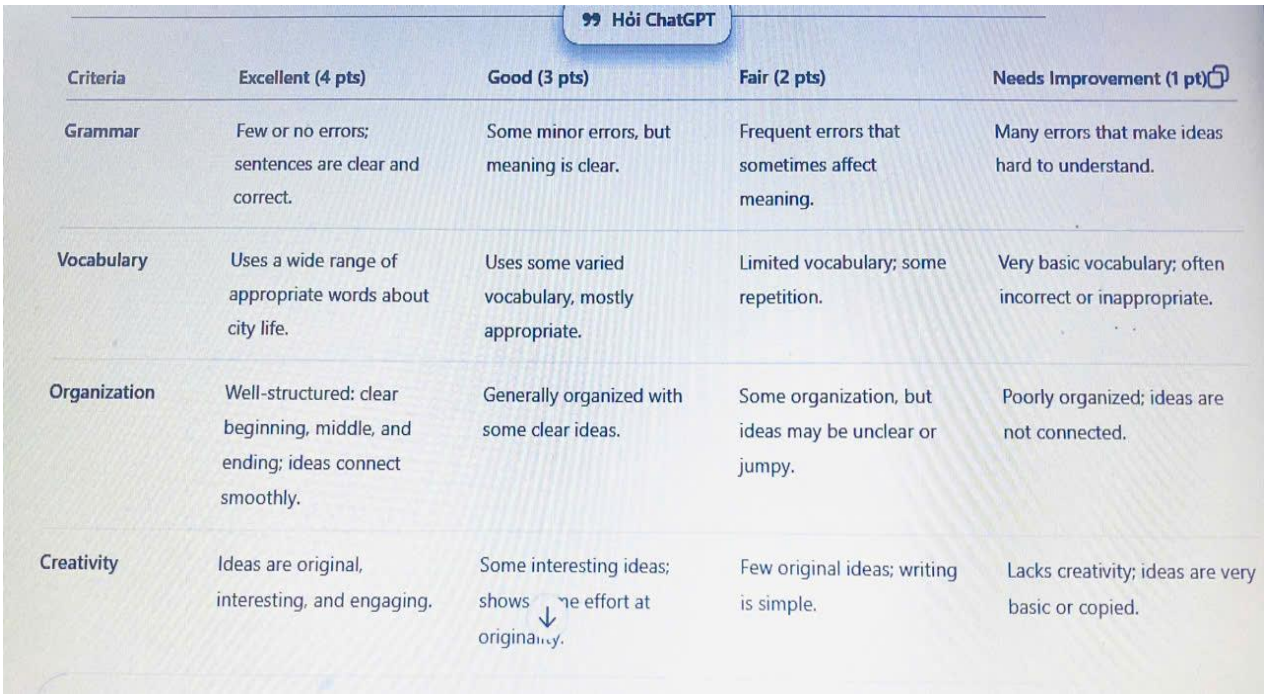
Nhằm giúp giáo viên có công cụ đánh giá rõ ràng, công bằng, minh bạch, AI tạo bảng tiêu chí (rubric) để giáo viên đánh giá kỹ năng nghe, nói, đọc, viết với cách làm:

- Xác định tiêu chí: độ chính xác ngữ pháp, vốn từ, sự lưu loát, sáng tạo.

- Yêu cầu AI soạn khung rubric theo từng cấp độ (A1, A2, B1, B2…).

Ví dụ câu lệnh (6): *Create a simple rubric to assess A2 students "Write a paragraph (80–100 words) about the city life". Criteria: grammar, vocabulary, organization, and creativity. Present in table form.*

Kết quả từ chatbot ChatGPT:



99 Hỏi ChatGPT

Criteria	Excellent (4 pts)	Good (3 pts)	Fair (2 pts)	Needs Improvement (1 pt)
Grammar	Few or no errors; sentences are clear and correct.	Some minor errors, but meaning is clear.	Frequent errors that sometimes affect meaning.	Many errors that make ideas hard to understand.
Vocabulary	Uses a wide range of appropriate words about city life.	Uses some varied vocabulary, mostly appropriate.	Limited vocabulary; some repetition.	Very basic vocabulary; often incorrect or inappropriate.
Organization	Well-structured: clear beginning, middle, and ending; ideas connect smoothly.	Generally organized with some clear ideas.	Some organization, but ideas may be unclear or jumpy.	Poorly organized; ideas are not connected.
Creativity	Ideas are original, interesting, and engaging.	Some interesting ideas; shows ne effort at originality.	Few original ideas; writing is simple.	Lacks creativity; ideas are very basic or copied.

5. Kết luận

Trong bối cảnh toàn cầu hóa và hội nhập quốc tế ngày càng sâu rộng, tiếng Anh đã trở thành công cụ thiết yếu để tiếp cận tri thức, giao lưu văn hóa và nâng cao năng lực. Sự kết hợp hài hòa giữa công nghệ tiên tiến và dạy học học tiếng Anh là chìa khóa để xây dựng một nền giáo dục tiếng Anh hiện đại, toàn diện và bền vững trong thời đại mới.

Bên cạnh đó, GenAI đang mở ra một xu hướng tất yếu trong giáo dục ngôn ngữ. Những ứng dụng công cụ trí tuệ nhân tạo trong dạy học tiếng Anh như ngôn ngữ thứ hai bao gồm soạn thảo nội dung học liệu để vận dụng kiến thức và rèn luyện các kĩ năng nghe, nói, đọc, viết; thiết kế hoạt động tổ chức dạy học và dự án học tập; đánh giá kết quả học tập của học sinh. Đây chính là sự kết hợp mang tính đột phá, giúp nâng cao chất lượng dạy học, biến quá trình học ngoại ngữ trở nên gần gũi, tự nhiên và bền vững hơn. Vai trò của giáo viên là phải xác định đúng mục tiêu nội dung bài dạy để thiết kế và tổ chức hoạt động dạy học đúng trọng tâm, đúng phương pháp; hiểu rõ trình độ, sở thích, nhu cầu riêng biệt của học sinh để cá nhân hóa việc học tập cho từng học sinh; biết tạo ra những câu lệnh thông minh để có thể khai thác tối ưu hiệu quả của công cụ trí tuệ nhân tạo. GenAI không thể thay thế được người thầy, mà chỉ đóng vai trò hỗ trợ, giúp cho người thầy có năng lực sư phạm giỏi hơn, cá nhân hóa quá trình học tập và tạo ra môi trường tương tác linh hoạt, hiệu quả. Tuy nhiên, việc ứng dụng công nghệ AI cần được sử dụng một cách có trách nhiệm, bảo vệ dữ liệu và quyền riêng tư, đạo đức và quản lý rủi ro.

Việc đưa tiếng Anh trở thành ngôn ngữ thứ hai trong trường học là một bước đi chiến lược có tác động lớn đến toàn hệ thống giáo dục nên cần được xây dựng bài bản, chi tiết với lộ trình rõ ràng về thời gian, kinh phí và giải pháp tổ chức thực hiện nhằm trang bị, nâng cao năng lực cho nguồn nhân lực chất lượng cao, góp phần đưa đất nước ta phát triển, vững mạnh trong kỷ nguyên mới - Kỷ nguyên vươn mình của dân tộc Việt Nam.

Tài liệu tham khảo

[1]. Bộ Giáo dục và đào tạo (2025), *Đưa tiếng Anh thành ngôn ngữ thứ hai trong trường học: Từng bước nhưng cũng phải nhanh chóng. https://moet.gov.vn/tintuc/Pages/tin-tong-hop.aspx?ItemID=10453*

[2]. Nguyễn Vân Khanh (2024), *Ứng dụng AI trong giảng dạy tiếng Anh ở bậc đại học: Thực trạng và các xu hướng nghiên cứu chính*, Journal of educational equipment: Education management, Volume 1, Issue 312 (May 2024) ISSN 1859 - 081

[3]. Nguyễn Thị Lan Phương (2024), *Sử dụng trí tuệ nhân tạo vào việc học tiếng Anh của sinh viên Trường Đại học Sư phạm Kỹ thuật Vinh: khó khăn và giải pháp*, Journal of educational equipment: Applied research, Volume 2, Issue 319 (August 2024) ISSN 1859 – 0810

[4]. Barrot, J. S. (2023), Using ChatGPT for second language writing: Pitfalls and potentials Assessing Writing, 57, 100745. https://doi.org/10.1016/j.asw.2023.100745

[5]. Crystal, D. (2012), *English as a global language* (2nd ed.). Cambridge University Press.

[6]. Harmer, J. (2015), *The practice of English language teaching* (5th ed.). London: Pearson.

[7]. Krashen, S. D. (1985). The Input Hypothesis: Issues and Implications. Longman.

[8]. Long, M. H. (1981). "Input, interaction, and second language acquisition". Annals of the New York Academy of Sciences, 379(1), 259-278.

[9]. Su (苏嘉红), J., & Yang (杨伟鹏), W. (2023). Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education. ECNU Review of Education, 0(0). https://doi.org/10.1177/20965311231168423

[10]. https://giaoduc.net.vn/tieng-anh-thanh-ngon-ngu-thu-hai-tai-truong-hoc-la-buoc-chien-luoc-thuc-day-hoi-nhap-post252315.gd

HUMAN-IN-THE-LOOP STRATEGIES IN MACHINE TRANSLATION – CHALLENGES, TRENDS, AND FUTURE DIRECTIONS: A CASE STUDY AT APD

Nguyen Thi Nuong[1]

Abstract

Machine translation (MT) has advanced significantly with the development of neural architectures, yet challenges remain in ensuring contextual accuracy, cultural appropriateness, and domain-specific precision. To address these issues, human-in-the-loop (HiL) strategies have emerged as a hybrid approach that integrates human expertise into automated translation pipelines. This study investigates HiL strategies through a case study, highlighting their role in improving translation quality, enhancing user trust, and supporting low-resource languages. The analysis identifies key challenges, including data scarcity, the cost of human intervention, and ethical concerns such as ownership of translation data. Current trends point to interactive and adaptive MT systems, reinforcement learning from human feedback, and the increasing use of collaborative crowdsourcing platforms. Based on the case study, this paper proposes future directions for HiL machine translation, including optimizing translator-system interaction, expanding applications to domain-specific contexts, and establishing ethical frameworks to guide sustainable development. Findings suggest that HiL strategies are crucial not only for improving MT performance but also for ensuring inclusivity and human-centered AI in multilingual communication.

Keywords: Human-in-the-loop, machine translation, post-editing, interactive MT, ethical frameworks.

Tóm tắt

Dịch máy (Machine translation - MT) đã đạt được những bước tiến quan trọng nhờ sự phát triển của các mô hình nơ-ron, nhưng vẫn còn tồn tại nhiều thách thức trong việc đảm bảo độ chính xác ngữ cảnh, sự phù hợp văn hóa và tính chuyên biệt theo lĩnh vực. Để khắc phục những hạn chế này, các chiến lược human-in-the-loop (HiL) đã xuất hiện như một phương pháp lai, kết hợp tri thức con người với hệ thống dịch tự động. Nghiên cứu này khảo sát các chiến lược HiL thông qua một nghiên cứu tình huống, nhấn mạnh vai trò của chúng trong việc nâng cao chất lượng bản dịch, gia tăng niềm tin của người dùng và hỗ trợ các ngôn ngữ ít tài nguyên. Phân tích chỉ ra những thách thức chính như sự khan hiếm dữ liệu song ngữ, chi phí can thiệp của con người, và các vấn đề đạo đức liên quan đến quyền sở hữu dữ liệu dịch. Các xu hướng hiện nay tập trung vào hệ thống dịch máy tương tác và thích ứng, học tăng cường dựa trên phản hồi con người, cũng như sự phát triển của các nền tảng dịch thuật cộng tác. Dựa trên nghiên cứu tình huống, bài viết đề xuất các định hướng tương lai cho HiLMT, bao gồm tối ưu hóa tương tác giữa người dịch và hệ thống, mở rộng ứng dụng cho các lĩnh vực chuyên ngành, và xây dựng khung đạo đức nhằm đảm bảo sự phát triển bền vững. Kết quả cho thấy các chiến lược HiL không chỉ quan trọng trong việc nâng cao hiệu quả dịch máy mà còn góp phần thúc đẩy tính nhân văn và bao trùm trong giao tiếp đa ngôn ngữ.

Từ khóa: Mô hình con người trong vòng lặp, dịch máy, hiệu đính bản dịch máy, dịch máy tương tác, khung đạo đức.

1. INTRODUCTION

[1] Tiến sĩ, Academy of Policy and Development, *Email: nuongnguyen@apd.edu.vn*

In the era of Industry 4.0, machine translation (MT) has made remarkable progress, particularly with the advent of neural machine translation, which has significantly improved fluency and coherence compared to earlier rule-based or statistical approaches (Bahdanau et al., 2015; Vaswani et al., 2017). Despite these advances, MT continues to face substantial limitations in handling complex contexts, cultural nuances, and domain-specific terminology (Koehn & Knowles, 2017). These challenges highlight the need for hybrid strategies that combine the computational efficiency of machines with the linguistic and cultural knowledge of humans.

The concept of *human-in-the-loop* (HiL) has emerged as a promising approach, redefining the relationship between humans and MT systems. Rather than viewing translation as a fully automated process, HiL integrates human involvement at various stages ranging from pre-editing and post-editing to interactive correction and adaptive learning (Specia et al., 2017; Kreutzer et al., 2020). Such strategies not only enhance translation quality but also strengthen user trust and acceptance, particularly in sensitive domains such as healthcare, law, and education. Furthermore, HiL strategies increasingly leverage the computational power of AI in synergy with modern human linguistic expertise.

Nevertheless, incorporating human knowledge into MT systems also presents notable challenges. These include the scarcity of high-quality bilingual data, the costs and time required for human intervention, and ethical concerns regarding the ownership of linguistic data (Bender et al., 2021). At the same time, advances in reinforcement learning from human feedback (RLHF) and collaborative translation platforms are opening new opportunities to optimize HiL strategies (Gao et al., 2022).

This study aims to explore *human-in-the-loop* strategies in machine translation through a case study. By analyzing strengths, limitations, and practical applications, it seeks to identify key trends and propose future directions for both research and practice. Ultimately, this research contributes to clarifying how humans and artificial intelligence can collaborate to develop multilingual communication systems that are more accurate, human-centered, and adaptable.

2. LITERATURE REVIEW

The development of machine translation (MT) has progressed through several stages, from rule-based MT to statistical machine translation (SMT), and more recently to neural machine translation (NMT). Research has demonstrated that NMT represents a significant breakthrough in producing more natural and fluent translations, particularly due to the integration of attention mechanisms and the transformer architecture (Bahdanau et al., 2015; Vaswani et al., 2017). However, as noted by Koehn and Knowles (2017), NMT still faces major challenges in specialized domain translation, handling low-resource languages, and maintaining semantic integrity in complex contexts.

To address these limitations, the concept of human-in-the-loop (HiL) in machine translation has gained increasing attention. Kreutzer et al. (2020) classify HiL strategies into three main categories: (i) post-editing, in which human translators revise the machine-generated output; (ii) interactive MT, where the system engages with the user during the translation process; and (iii) adaptive MT, where the system learns from the translator's style and feedback.

Similarly, Specia et al. (2017) emphasize that post-editing is the most practical approach to improving translation quality while simultaneously reducing translators' workload.

Another strand of research focuses on human–machine interaction in translation. Barrachina et al. (2009) argue that computer-assisted translation (CAT) systems optimize the translation process by allowing translators to influence system outputs in real time. More recently, Knowles and Koehn (2016) proposed interactive prediction models to enhance adaptive learning, enabling systems to adjust translations based on user input.

In addition, ethical and transparency issues in MT have attracted growing attention. Bender et al. (2021) warned of "stochastic parrots," highlighting the risk that language models may reproduce biases and lack transparency in data processing. Consequently, recent studies (Gao et al., 2022) have explored reinforcement learning from human feedback (RLHF) as a promising approach to enable MT systems to learn more effectively and responsibly from human evaluation.

Finally, an emerging trend is the development of crowdsourcing translation platforms. Zaidan and Callison-Burch (2011) demonstrated that community-contributed translations, when combined with MT and human review, can achieve professional-level quality. This is particularly significant for low-resource languages and educational contexts, where professional translators are scarce.

Taken together, these lines of research indicate that HiL strategies in MT are shifting from a focus on "output correction" toward "interactive and adaptive learning," while increasingly intertwined with issues of ethics, data governance, and linguistic inclusivity. These trends provide a foundation for further expanding research on HiLMT across diverse application contexts.

While much of the prior research on machine translation (MT) has primarily focused on system performance and algorithmic improvements, the role of human involvement and its impact on learners' perceptions and experiences has received comparatively less attention. In particular, challenges such as the level of effort required for post-editing, as well as shifts in trust and perceived translation quality when human intervention is introduced, remain underexplored. Moreover, students' perceptions of the extent to which human-machine collaboration can enhance translation quality have yet to be clearly articulated.

Building on this research gap, the present study seeks to address the following questions:

1. *Research question 1:* How do students evaluate changes in trust and perceived translation quality when human intervention is involved?
2. *Research question 2:* How do students perceive the level of effort required for post-editing machine translation?
3. *Research question 3:* How do students perceive the potential for translation quality improvement through the integration of human and machine translation?

3. METHODOLOGY

3.1. Research objectives and design

The study employed a descriptive survey design to investigate the perceptions, attitudes, and experiences of third- and fourth-year English Language students at the Academy of Policy and Development (APD) regarding human-in-the-loop machine translation (HiLMT), while also identifying the factors influencing their level of acceptance and willingness to engage.

3.2. Participants and sample

- **Participants**: Third and fourth-year English Language students at APD.
- **Sample**: 100 students, selected through a non-probability stratified method (convenience + quota), ensuring a 50-50 distribution between third- and fourth-year students.
- **Exclusion criteria**: First and second-year students, or individuals who have not used or rarely use machine translation tools.

3.3. Measurement and data collection instruments

In this study, data were collected through a survey consisting of 18 statements designed on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree). The statements were grouped into six constructs: Perceived usefulness (PU), trust & quality perception (TQ), perceived effort (PE), willingness to participate (WP), ethics & data concern (ED), and perceived improvement (PI) to measure different aspects of students' perceptions and attitudes toward Human-in-the-loop Machine Translation (HiLMT).
All measurement constructs were developed based on prior research on machine translation, human-aided post-editing, and the application of NLP technologies in education. The measurement scales were adapted from previous studies (Kreutzer et al., 2020; Specia et al., 2017; Gao et al., 2022).

The survey questionnaire consisted of the following sections:

Section A: Demographic information (age, gender, year of study, level of MT use, translation experience).

Section B: 18 statements on a 5-point Likert scale, organized into six constructs: perceived usefulness (PU), trust & quality perception (TQ), perceived effort (PE), willingness to participate (WP), ethics & data concern (ED), and perceived improvement (PI).

Section C: Post-editing practice experience and open-ended questions.

Section D: Challenges.

Section E: Open-ended questions.

Section F: Trends.

Section G: Future directions.

3.4. Data collection procedure

The questionnaire was piloted with 10 students to refine the wording, and then administered to 100 third and fourth-year English language students at APD through Google

Forms and in-class distribution over a period of 2-3 weeks. Participants' anonymity was ensured, and their participation was voluntary.

3.5. Data Analysis

Quantitative analysis included descriptive statistics, Cronbach's Alpha ($\alpha \geq 0.70$), exploratory factor analysis (EFA), t-test/ANOVA, and multiple regression to identify the predictors of willingness to participate and trust in HiLMT.

Qualitative analysis was conducted using content analysis of the open-ended questions.

The software employed included SPSS, R, or NVivo.

3.6. Ethics and limitations

The study ensured anonymity, voluntary participation, and the use of data solely for academic purposes. The main limitations lie in the use of a convenience sample (N = 100), which may not be fully representative, and in the dependence on participants' honesty and the questionnaire design.

4. FINDINGS AND DISCUSSION

A survey was conducted with 100 students (50 third-year and 50 fourth-year) who have studied or are currently studying translation. The results showed that all questions were fully completed. Regarding the general question on "The frequency of using machine translation tools such as Google Translate, DeepL, etc.," the majority of students (74%) reported frequent use of these tools in their studies and work, 26% reported occasional use, and none indicated that they had never used them.

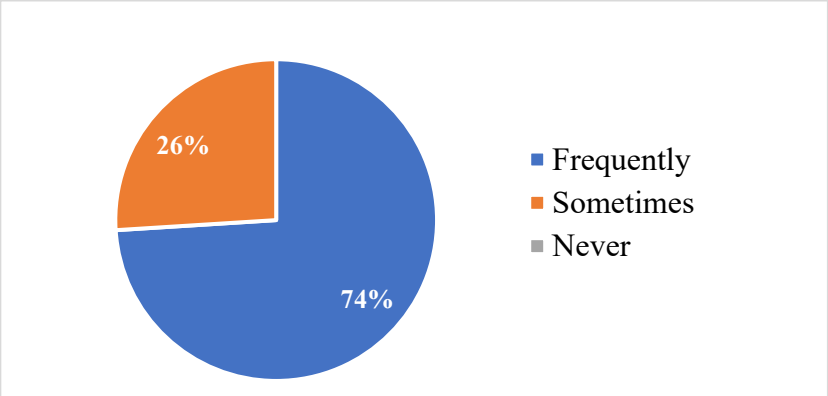


Chart 1. The frequency of using machine translation tools

4.1. Perceptions and attitudes toward machine translation and Human-in-the-Loop MT

The survey results indicate that students hold multifaceted perceptions and attitudes toward the use of Human-in-the-Loop Machine Translation (HiLMT). Six measurement dimensions capture different aspects, ranging from perceived usefulness, trust in quality, and

required effort, to willingness to engage, ethical and data-related concerns, and perceived improvement.

4.1.1. Perceived usefulness (PU)

This group was assessed through three statements, focusing on the extent to which students perceived the benefits of machine translation in supporting translation tasks and learning, with the results as follows:

Perceived usefulness (PU)	**1 Strongly disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly agree**
1.Machine translation enables me to quickly comprehend the content of foreign-language texts.	1%	1%	6%	42%	50%
2. Machine translation proves useful when translating general, non-specialized texts.	1%	2%	12%	32%	53%
3. Machine translation helps me reduce the time required to produce a draft translation.	1%	6%	11%	31%	51%

The majority of students acknowledged the benefits of machine translation in facilitating quick comprehension of content (92% agreed or strongly agreed), translating general texts (85%), and reducing the time required to produce draft translations (82%). These findings suggest that machine translation is perceived as a useful and highly applicable tool in academic learning.

4.1.2. Trust & quality perception (TQ)

This group reflects students' beliefs and evaluations regarding the quality of translations produced through human–machine collaboration, with the results summarized as follows:

Trust & quality perception (TQ)	**1 Strongly disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly agree**
1. Students reported greater confidence in the quality of translations when human post-editing was involved.	0%	1%	17%	33%	49%
2. Students generally perceived that machine translations often	1%	3%	22%	34%	40%

lack cultural nuances and subtle semantic distinctions.					
3. Students expressed the belief that humans are able to detect and correct errors that machines often overlook.	0%	3%	14%	34%	49%

Students expressed confidence in the quality of translations when human involvement was included in the post-editing stage: 82% agreed that post-editing enhances reliability, and 83% believed that humans are able to correct errors overlooked by the machine. However, more than 74% also acknowledged that machine translations often lack cultural and semantic nuances. These findings suggest that students perceive machine translation as truly effective only when complemented by human input.

4.1.3. ***Perceived effort (PE)***

The PE group measures the level of effort and time that students perceive as necessary when engaging in post-editing machine translations or using interactive translation systems.

Perceived effort (PE)	**1** **Strongly disagree**	**2** **Disagree**	**3** **Neutral**	**4** **Agree**	**5** **Strongly agree**
1. I believe that post-editing machine translations requires considerable effort.	2%	22%	25%	25%	26%
2. Using interactive MT helps me complete translations more quickly than manual translation alone.	1%	3%	9%	48%	39%
3. Post-editing may take longer than manual translation.	4%	24%	19%	32%	21%

The perceived level of effort was relatively divided: 51% of respondents considered post-editing to be labor-intensive, while 71% acknowledged that interactive MT saves time compared to manual translation. Notably, 43% expressed concern that post-editing may take longer than traditional translation. These findings reveal a certain inconsistency in students' experiences: while they recognize the potential time-saving benefits, they also remain apprehensive about the substantial amount of revision required.

4.1.4. ***Willingness to participate (WP)***

The WP group reflects students' willingness to engage in post-editing, to use, and to provide feedback on machine translation systems.

Willingness to participate (WP)	**1 Strongly disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly agree**
1. I am willing to post-edit machine translation if compensation or recognition is provided.	0%	1%	16%	33%	50%
2. I am willing to use interactive MT systems in my learning process.	2%	2%	9%	40%	47%
3. I am willing to provide feedback to help machine translation systems gradually learn and improve (crowdsourcing).	2%	2%	14%	35%	47%

The results indicate a high level of willingness to participate: approximately 83% expressed readiness to engage in post-editing if compensation or recognition is provided, 87% were willing to use interactive MT in their learning, and 82% were willing to provide feedback to improve the system. These findings demonstrate that students are motivated to participate when appropriate incentives and recognition mechanisms are in place.

4.1.5. Ethics and data concern (ED)

The ED group addresses students' concerns regarding data security, ownership, and ethical issues in the use of HiLMT.

Ethics & data concern (ED)	**1 Strongly disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly agree**
1. I am concerned that my data/input may be used without authorization.	4%	11%	18%	30%	37%
2. I believe that clear ethical guidelines are necessary for the use of HiLMT.	1%	3%	12%	29%	55%
3. I believe that the ownership rights of translations need to be clearly defined when using machine translation tools.	0%	1%	8%	31%	60%

Students expressed significant concerns regarding ethics and data: 67% were worried that translation data might be used without authorization, 84% emphasized the need for clear ethical guidelines, and 91% affirmed the necessity of explicitly defining translation ownership rights. Thus, issues of security, rights, and ethical standards emerge as major concerns that may influence the level of acceptance of HiLMT.

4.1.6. Perceived improvement (PI)

The PI group aims to assess students' perceptions of translation quality improvement when human input is combined with machine translation.

Perceived improvement (PI)	**1 Strongly disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly agree**
1. I believe that the integration of human input and machine translation (HiLMT) significantly enhances translation quality.	0%	1%	8%	22%	69%
2. I believe that HiLMT is particularly essential when translating specialized texts (e.g., in economics, healthcare, or law.	2%	4%	10%	37%	47%
3. I believe that machine translation systems can learn from user feedback (RLHF) and improve their quality over time.	1%	0%	12%	35%	52%

The most notable findings in this group are as follows: 91% of students agreed that HiLMT significantly improves translation quality, 84% considered it essential for specialized texts, and 87% expressed confidence in the system's ability to learn and improve through user feedback. These results reflect a strong belief in the long-term potential for improvement of HiLMT.

4.2. Practical experience

When asked about their experience with post-editing machine translations, the majority of students reported having participated to some extent. Specifically, 63% responded "yes, occasionally," and 8% "yes, frequently," while only 29% indicated they had never engaged in post-editing. This suggests that most students have been exposed to post-editing activities, indicating that it is no longer a novel concept but has gradually become a relatively common academic practice.

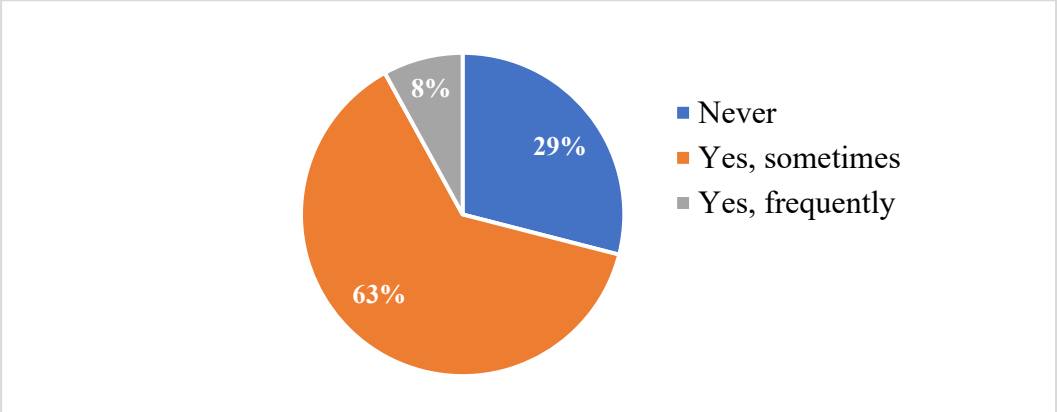


Chart 2. Students' experience with post-editing machine translations

Perceived difficulty of post-editing compared to manual translation; the responses revealed a clear divergence in perceptions of difficulty. Approximately 34% of students considered post-editing to be easier (7% "much easier" and 27% "somewhat easier"), whereas 42% regarded it as more difficult (30% "somewhat more difficult" and 12% "much more difficult"). Notably, 24% believed that the two modes were equivalent.

These results suggest that students' experiences with post-editing are ambivalent:

- One group values the time and effort-saving benefits of having a preliminary machine-generated draft.
- Another group finds that correcting semantic, grammatical, and stylistic errors in machine translations can be just as demanding, if not more complex, than manual translation.

Overall, students remain divided in their assessment of post-editing effectiveness. This divergence may stem from factors such as language proficiency, familiarity with machine translation tools, or the type of text being translated. These findings highlight the need for pedagogical guidance and appropriate tool support in HiLMT training to help students leverage the advantages of post-editing while mitigating its challenges.

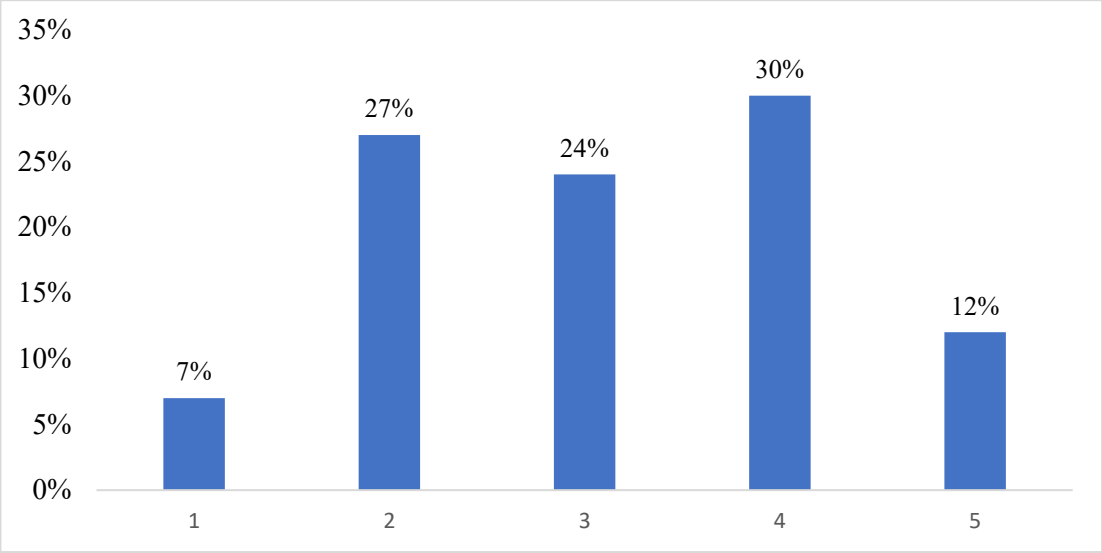


Chart 3. Perceived difficulty of post-editing compared to manual translation

1 = Much easier than manual translation
2 = Somewhat easier than manual translation
3 = Equivalent (no significant difference)
4 = Somewhat more difficult than manual translation
5 = Much more difficult than manual translation

Time spent on each post-editing task (200-300 words), the results indicate a relatively diverse distribution. Specifically, 47% of students reported spending 10-20 minutes, 22% required more than 20 minutes, and 31% completed the task in under 10 minutes.

These findings suggest that the time required for post-editing largely depends on students' language proficiency, familiarity with machine translation tools, and the complexity of the text. The fact that nearly half of the students needed 10-20 minutes indicates that post-editing remains a moderately time-consuming activity, and not necessarily as fast as initially expected when using machine translation.

Notably, about one-fifth of the students (22%) reported spending more than 20 minutes to post-edit a short text, suggesting that in certain cases the effort required may even exceed that of manual translation, particularly when the machine-generated draft contains numerous semantic errors or complex structures. Conversely, 31% completed the task in less than 10 minutes, indicating that this group either made effective use of the raw machine translation or encountered texts with fewer errors that were easier to revise.

Overall, the results reveal a clear divergence in students' post-editing experiences: while some find it time-saving, others perceive it as equally or more demanding. This reinforces the view that, to maximize the benefits of HiLMT in academic settings, students require methodological guidance, effective post-editing strategies, and text classification practices to reduce workload and enhance training efficiency.

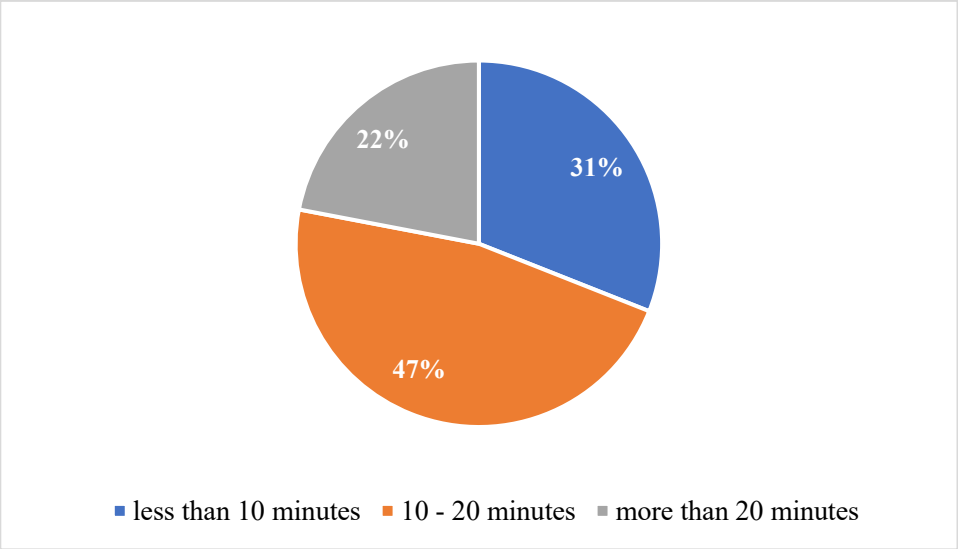


Chart 4. Time spent on each post-editing task

4.2. Challenges

Challenges	1 Strongly disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly agree
1. The lack of high-quality bilingual data reduces the effectiveness of human-in-the-loop (HiL) in machine translation (MT).	1%	3%	21%	46%	29%
2. For low-resource languages, HiL is less likely to yield substantial improvements in translation quality.	2%	11%	17%	42%	28%
3. Post-editing machine translations of specialized texts (e.g., legal, medical, technical) is time-consuming and labor-intensive.	1%	11%	22%	30%	36%
4. The human resource costs of post-editing represent a major barrier to the scalability of HiL.	3%	10%	21%	40%	26%
5. Ethical issues and data ownership must be addressed	0%	3%	12%	31%	54%

before implementing HiL on a large scale.					

Alongside students' positive perceptions of HiLMT, the survey also highlights several key challenges in its application.

First, the limitation of high-quality bilingual data. Findings show that 75% of students agreed or strongly agreed that the lack of reliable bilingual resources reduces the effectiveness of HiLMT. This is a common issue in machine translation research, particularly for low-resource languages such as Vietnamese, where standardized and diverse corpora remain scarce.

Second, difficulties with low-resource languages. Approximately 70% of students perceived HiLMT as struggling to deliver significant improvements for low-resource languages. This reflects their awareness of current technical constraints: even with human post-editing, translation quality is still largely constrained by limited language data.

Third, the specificity of domain texts. Up to 66% of respondents agreed that post-editing specialized texts (e.g., legal, medical, or technical documents) requires greater effort. Such texts demand high levels of expertise and precision, leading students to perceive a heavier burden in terms of time and specialized knowledge during post-editing.

Fourth, human resource costs for post-editing. Around 66% identified human labor costs as a major barrier to the scalability of HiLMT. This aligns with practical realities: while HiLMT enhances translation quality, achieving consistently high performance requires ongoing human involvement, which entails substantial operational expenses.

Fifth, ethical and data ownership concerns. Most notably, 85% of students emphasized that ethical and data ownership issues must be addressed before large-scale deployment of HiLMT. This is considered a fundamental barrier: without clear legal frameworks and data protection mechanisms, users may hesitate to contribute data or engage in post-editing.

From the survey results, it can be seen that overall, the main challenges perceived by students are grouped into three categories:

1. Technical: including the shortage of bilingual data and difficulties associated with low-resource languages.
2. Expertise and cost: encompassing the revision of domain-specific texts and the expenses related to human resources.
3. Ethical and legal: covering data ownership rights and security concerns.

This indicates that for HiLMT to be widely and sustainably implemented in both academic and practical translation contexts, a set of comprehensive solutions is required: developing high-quality language corpora, designing tools to support domain-specific post-editing, reducing operational costs, and, most importantly, establishing a clear ethical and legal framework to foster user trust

Regarding the first open-ended question, "In your opinion, what is the greatest challenge of HiL in machine translation today?", the responses indicate that students identified a variety

of challenges. Among these, the most prominent concern is the issue of balancing the processing speed of the system with the quality and naturalness contributed by human involvement. Several participants emphasized that if the machine translation is insufficiently accurate, the time required for post-editing may exceed that of manual translation, thereby diminishing the advantages of automation. In addition, the shortage of high-quality bilingual data, particularly in specialized domains, is also perceived as a major barrier, hindering the system's ability to learn effectively and consistently. Learners further highlighted issues such as the cost of human resources, the motivation and productivity of post-editors, variations in translation styles, and limitations in handling context and cultural nuances. Moreover, concerns were raised regarding fairness, transparency, and data security when implementing HiL at scale.

In summary, the central challenge of HiLMT lies in optimizing the synergy between human and machine: ensuring that the system can leverage the speed of AI while harnessing the linguistic and cultural subtlety provided by humans, all while maintaining cost-effectiveness, efficiency, and a positive user experience.

Regarding to the second open-ended question, "In your opinion, in which direction should HiL develop in the future to balance technological benefits and human interests?", the majority of students emphasized that HiL should be oriented toward support rather than replacement, functioning as an "intelligent assistant" that allows humans to retain oversight and make the final decisions. The system is expected to ensure transparency, explain ability, and an optimized user experience, while also complying with ethical, legal, and privacy standards. Furthermore, HiL should contribute to enhancing human creativity, analytical capacity, and decision-making skills, thereby fostering a symbiotic relationship between technology and human translators

Regarding the third question, "What are your expectations for HiL tools in supporting domain-specific translation?", the majority of students expressed expectations that such tools would ensure accurate rendering of specialized terminology, contextual understanding, consistency, and semantically appropriate translation suggestions. The system is expected to learn from translators' feedback, allow customization of terminology databases, provide clear explanations, and guarantee data security. In addition, HiL should help reduce repetitive tasks, save time, enhance efficiency and reliability, and function as an intelligent collaborator that enables human translators to focus on value added aspects of domain-specific translation.

4.3. Trends

The survey results reveal that students not only perceive HiLMT from a present perspective but also anticipate its future developmental trajectories.

Trends	**1** **Strongly disagree**	**2** **Disagree**	**3** **Neutral**	**4** **Agree**	**5** **Strongly agree**
1. The integration of AI and human feedback is likely to	0%	2%	14%	28%	56%

become the main trend in improving machine translation.					
2.Crowdsourcing is an effective approach to collecting feedback for HiL.	0%	0%	20%	39%	41%
3.Computer-assisted translation (CAT) tools integrating MT and HiL are expected to become increasingly prevalent.	0%	2%	11%	48%	39%
4. HiL has the potential to enable MT to learn personalized translation styles.	0%	4%	15%	44%	37%

Firstly, the integration of AI with human feedback. Approximately 84% of students agreed or strongly agreed that this will be the main trajectory for enhancing MT quality. This reflects learners' confidence in the AI-human collaborative model, where technology provides the initial output and humans refine it to ensure accuracy and naturalness.

Secondly, crowdsourcing as an effective approach. About 80% of students recognized community-based feedback collection as an efficient way for HiLMT to learn and improve. This indicates students' awareness of the collective power in generating feedback data and suggests a sustainable pathway that leverages wider participation rather than relying solely on a limited pool of experts.

Thirdly, the growing prevalence of CAT tools integrating MT and HiL. Nearly 87% of students anticipated that CAT tools combining machine translation and post-editing will become mainstream. This expectation highlights learners' desire for integrated solutions that can simultaneously ensure the speed of MT and the accuracy of HiL.

Fourthly, the potential for personalized translation styles. Around 81% of students agreed that HiL can enable MT systems to adapt to individualized translation styles. This illustrates expectations that future systems will not only enhance grammatical and semantic accuracy but also adjust to users' personal preferences, thereby enriching translation experiences and long-term usability.

Overall, students predicted that the future of MT will revolve around collaborative AI-human interaction, in which: community feedback and integrated CAT-MT-HiL tools serve as critical drivers; and personalization emerges as a new direction, promising translation outputs more aligned with individual users' needs.

These findings affirm that students perceive themselves not only as end-users but also as active contributors in shaping the future of HiLM

4.4. Future directions

The survey results indicate that students not only acknowledge the benefits, challenges, and current trends of HiLMT but also articulate clear expectations and future directions for its development.

Firstly, the need for a clear legal framework. Approximately 90% of students agreed or strongly agreed that transparent regulations on data ownership and usage are essential. This is viewed as a prerequisite to ensure fairness, security, and transparency in large-scale HiL implementation.

Secondly, mechanisms for recognizing human contribution. Nearly 88% of students expressed the desire for fair evaluation systems to acknowledge the intellectual effort of post-editors. This reflects a concern for the value of human labor and underscores the importance of motivation and incentives in sustaining human participation.

Thirdly, reducing cognitive load for post-editors. Around 73% believed that future research should focus on alleviating the cognitive burden of post-editing. Students expect supportive technological and pedagogical solutions, such as intelligent suggestion systems, automated error categorization, or more user-friendly interfaces to make post-editing less demanding and more efficient.

Fourthly, interdisciplinary collaboration for sustainable HiL development. Most notably, 90% of respondents emphasized the necessity of cross-disciplinary collaboration bringing together linguistics, computer science, law, and economics to foster comprehensive HiLMT development. This highlights students' recognition that HiLMT is not merely a technological issue but also one embedded in broader social, legal, and economic contexts

Future directions	**1 Strongly disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly agree**
1. A clear legal framework regarding data ownership and usage is essential for the implementation of HiL	1%	2%	9%	37%	53%
A fair evaluation mechanism should be developed to appropriately recognize human contributions in HiL.	0%	1%	11%	38%	50%
3. Future research should focus on developing strategies to reduce the cognitive load for post-editors in HiL.	1%	10%	16%	31%	42%
4. There is a need for interdisciplinary collaboration (linguistics, computer science,	0%	0%	10%	35%	55%

law, and economics) to ensure the sustainable development of HiL.					

5. CONCLUSION

A survey conducted with 100 third and fourth-year students at APD revealed that HiLMT is perceived as a useful tool, yet several challenges remain to be addressed.

Firstly, regarding perceptions and attitudes, the majority of students highly valued the benefits of machine translation in supporting learning, particularly when complemented by human post-editing. Trust in HiLMT was closely associated with its ability to improve translation quality; however, students were also clearly aware of the limitations of MT in terms of semantics and cultural nuances

Secondly, the post-editing experience revealed a divergence: while some students perceived HiLMT as saving time and effort, others found the editing workload burdensome, at times even more time-consuming than manual translation. This highlights the need for more effective tools and methods to support post-editing.

Thirdly, the key challenges were clustered into three dimensions: (i) technical - the lack of high-quality bilingual data and difficulties in low-resource languages; (ii) expertise & cost - the demand for domain-specific knowledge and significant human resource expenses; (iii) ethical & legal - concerns over data security, ownership, and transparency standards.

Fourthly, students anticipated that the future development of HiLMT will revolve around AI-human collaboration, crowdsourcing as a means of system enhancement, the increasing prevalence of CAT tools integrating MT and HiL, and the potential for personalized translation styles.

Finally, the expected directions for development include establishing a clear legal framework for data ownership, implementing fair recognition mechanisms for post-editors' contributions, exploring solutions to reduce cognitive load in post-editing, and fostering interdisciplinary collaboration for sustainable HiLMT advancement.

Overall, the findings suggest that students are not merely passive users but are willing to engage actively in shaping the future of HiLMT. This points to the potential of HiLMT as a symbiotic translation model, in which technology and humans complement each other to optimize quality, efficiency, and the humanistic value of translation

6. Implications and recommendations

The findings of this study provide several important implications for the application and development of HiLMT in foreign language education

6.1. For students

* Students should be equipped with knowledge and skills in post-editing as well as the use of CAT tools integrated with MT + HiL to enhance their learning effectiveness.

* Their active participation in providing feedback to the system should be encouraged through fair recognition mechanisms, thereby motivating students to both practice translation skills and contribute to improving the system.

6.2. For teachers and institutions

* Teachers should integrate activities related to HiLMT into translation and language technology courses, enabling students to gain early exposure to modern translation tools.

* Institutions may develop research-practice projects that allow students to participate in the training and evaluation of MT systems, thereby fostering experiential learning while simultaneously contributing valuable data

6.3. For translation technology developers

* User experience should be optimized by designing intuitive interfaces, minimizing repetitive tasks in post-editing, and supporting the personalization of translation styles

* Enhancing transparency regarding data security and ownership can build user trust when employing HiLMT in sensitive contexts.

6.4. For Policy and Administration

* It is necessary to establish a legal framework and ethical standards for the application of AI in translation at an early stage, particularly concerning intellectual property rights, data security, and fair recognition of human contributions.

* Interdisciplinary collaboration among linguistics, computer science, law, and economics should be encouraged to ensure the sustainable, efficient, and responsible development of HiLMT.

TÀI LIỆU THAM KHẢO

1. Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *International Conference on Learning Representations (ICLR)*.
2. Barrachina, S., Bender, O., Casacuberta, F., et al. (2009). Statistical approaches to computer-assisted translation. *Computational Linguistics*, 35(1), 3-28. https://doi.org/10.1162/coli.2008.07-055-R2-06-29.
3. Zaidan, O. F., & Callison-Burch, C. (2011). Crowdsourcing translation: Professional quality from non-professionals. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL)*, 1220–1229.
4. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 610–623. https://doi.org/10.1145/3442188.3445922.
5. Gao, T., Fisch, A., & Chen, D. (2022). Making pre-trained language models better with human feedback. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL)*, 375–388. https://doi.org/10.18653/v1/2022.acl-long.28.
6. Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. *Proceedings of the First Workshop on Neural Machine Translation*, 28–39. https://doi.org/10.18653/v1/W17-3204.
7. Kreutzer, J., Riezler, S., & Sokolov, A. (2020). Human-in-the-loop methods for machine translation: A survey. *arXiv preprint arXiv:2007.09244*. https://arxiv.org/abs/2007.09244
8. Specia, L., Blain, F., Chatterjee, R., & Paetzold, G. (2017). Quality estimation for automatic and semi-automatic machine translation. In *J. Tiedemann (Ed.), Handbook of Machine Translation* (pp. 525–559). Springer. https://doi.org/10.1007/978-3-319-32626-4_16.
9. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS)*, 5998–6008.

Thông tin tác giả

Tác giả 1

- Họ và tên: Nguyễn Thị Nương
- Học hàm/học vị: Tiến sĩ Ngôn ngữ học

- Nơi công tác: Học viện Chính sách và Phát triển
- Kinh nghiệm công tác: trên16 năm
- Lĩnh vực nghiên cứu chuyên sâu của tác giả: Ngôn ngữ học ứng dụng, phương pháp giảng dạy và dịch thuật.
- Thông tin liên lạc của tác giả: địa chỉ: nuongnguyen@apd.edu.vn
- Tel: 0912.833.525

PHÂN TÍCH CÁC MÔ HÌNH NGÔN NGỮ HỌC TÍNH TOÁN TRONG HỖ TRỢ XỬ LÝ NHẬP NHẰNG NGÔN NGỮ TỰ NHIÊN

Huỳnh Quang Đức[1]

Tóm tắt: Sự mơ hồ vẫn là một trong những vấn đề thách thức nhất trong xử lý ngôn ngữ tự nhiên (NLP-Natural Language Processing), ảnh hưởng đáng kể đến hiệu suất của các tác vụ hiện nay như dịch máy, truy xuất thông tin và trả lời câu hỏi… Bài báo này nghiên cứu vai trò của ngôn ngữ học tính toán trong việc hỗ trợ giải quyết sự mơ hồ, tập trung vào các cấp độ từ vựng, cú pháp và ngữ nghĩa. Bằng cách kết hợp các lý thuyết ngôn ngữ học với các phương pháp học máy, chúng tôi làm nổi bật các phương pháp cho phép hệ thống nắm bắt ngữ cảnh và loại bỏ sự nhập nhằng tốt hơn. Nghiên cứu của chúng tôi trình bày tổng quan về các kỹ thuật hiện tại, bao gồm loại bỏ sự mơ hồ nghĩa của từ bằng cách nhúng từ nhận biết ngữ cảnh và các mô hình xác suất, đồng thời thảo luận về những hạn chế của chúng trong các ứng dụng thực tế, nhấn mạnh tầm quan trọng của việc tích hợp kiến thức ngôn ngữ học tính toán vào các hệ thống NLP.

Từ khóa: Ngôn ngữ học tính toán; Xử lý ngôn ngữ tự nhiên (NLP); Giải quyết sự mơ hồ; Khử nhập nhằng ngữ nghĩa từ; Mô hình nhận thức ngữ cảnh.

1. ĐẶT VẤN ĐỀ

Sự nhập nhằng ngữ nghĩa là một đặc điểm cơ bản của ngôn ngữ con người, đặt ra những thách thức dai dẳng cho các hệ thống tính toán. Sự mơ hồ về từ vựng (đa nghĩa và đồng âm), sự mơ hồ về cú pháp (liên kết và phối hợp), và sự mơ hồ về ngữ dụng (không xác định rõ ràng và hàm ý) đều làm suy yếu độ tin cậy của việc diễn giải trong các tác vụ hiện nay như: dịch máy, trích xuất thông tin và các tác nhân đàm thoại. Những tiến bộ gần đây trong các mô hình ngữ cảnh được huấn luyện trước như BERT và các dẫn xuất của nó, đã cải thiện độ nhạy ngữ cảnh, tuy nhiên, các mô hình này vẫn gặp khó khăn trong việc lựa chọn nghĩa chi tiết và loại bỏ sự mơ hồ trong thế giới thực, trong bối cảnh thưa thớt hoặc các tín hiệu đa phương thức. Ngày càng có nhiều bằng chứng từ các chiến dịch đánh giá cho thấy việc tích hợp kiến thức từ vựng bên ngoài và các tín hiệu đa phương thức có thể cải thiện đáng kể hiệu suất loại bỏ sự mơ hồ trong các bối cảnh thực tế [1, 3].

Nghiên cứu về giải quyết nhập nhằng nghĩa từ (WSD-Word Sense Disambiguation) đã phát triển từ các mô hình dựa trên quy tắc và giàu kiến thức sang các phương pháp tiếp cận lai và mạng neural kết hợp chú thích (gán nhãn), kho ngữ liệu ngữ nghĩa và nhúng ngữ cảnh. Các phương pháp như GlossBERT minh họa cách kết hợp ngữ cảnh với gán nhãn ngữ nghĩa và tinh chỉnh bộ mã hóa chuyển đổi mang lại kết quả cho các hệ thống WSD có giám sát mạnh. Trong khi các chuẩn mực WSD đa phương thức như Visual-WSD cho thấy giá trị của hình ảnh và nền tảng đa phương thức trong việc giải quyết các ngữ nghĩa vốn khó hiểu. Đồng thời, các nghiên cứu gần đây nhấn mạnh vào ước tính bất định và các dự đoán được hiệu chuẩn đã làm nổi bật các trường hợp nhập nhằng thay vì ép buộc các lựa chọn quá tự tin - một cân nhắc quan trọng cho các ứng dụng lấy con người làm trung tâm [1, 3, 9].

[1] Thạc sĩ, Trường Đại học Bình Dương, *Email: hqduc@bdu.edu.vn.*

Mặc dù có những thành tựu từ thực nghiệm thông qua các nghiên cứu được nêu bên trên, vẫn còn một số vấn đề chưa được giải quyết:

(1) Phạm vi bao phủ của miền và ngôn ngữ, các phương pháp WSD được huấn luyện trên các kho ngữ liệu tiếng Anh và kho ngữ nghĩa chuẩn, thường hoạt động kém hiệu quả trên các ngôn ngữ giàu hình thái hoặc có ít tài nguyên (cụ thể như tiếng Việt hiện nay được gọi là ngôn ngữ nghèo tài nguyên);

(2) Sự thưa thớt về ngữ cảnh, nhiều phát ngôn trong thế giới thực thiếu những gì các bộ chuyển đổi ngữ cảnh mở rộng dựa vào trong các hệ thống được huấn luyện trước;

(3) Sự mơ hồ như một đặc tính tương tác mang tính hệ thống, các hệ thống đối thoại và tương tác cần phải phát hiện sự nhập nhằng và chủ động quản lý để làm rõ nghĩa chưa được quan tâm.

Các khảo sát gần đây và các nghiên cứu chuyên sâu về miền giá trị làm nổi bật các phương pháp để điều chỉnh các hệ thống WSD cho phù hợp với các bối cảnh có ít tài nguyên, và để tích hợp các tài nguyên biểu tượng (BabelNet, WordNet, Wikidata) với các biểu diễn đã học trước đó, nhưng các giải pháp thực tế có khả năng mở rộng vẫn là một lĩnh vực nghiên cứu tích cực [6, 7].

Xuất phát từ những khoảng trống nêu trên và chủ đề hội nghị "Công nghệ vì con người", bài báo này định hình việc giải quyết sự mơ hồ như một vấn đề đa ngành, được hưởng lợi từ việc kết hợp những hiểu biết sâu sắc về ngôn ngữ học tính toán, mô hình xác suất và các kỹ thuật gióng hàng song ngữ cho các mô hình ngôn ngữ hiện đại. Chúng tôi phác thảo một chương trình nghiên cứu hỗ trợ giải quyết sự nhập nhằng của ngôn ngữ tự nhiên gồm:

(i) Phát hiện sự mơ hồ rõ ràng và các đầu ra nhận thức được sự không chắc chắn.

(ii) Các biểu diễn ngữ cảnh được tăng cường kiến thức.

(iii) Các giao thức đánh giá đa phương thức và tương tác phản ánh các thực hành giao tiếp của con người.

Những nỗ lực gần đây nhằm căn chỉnh các mô hình ngôn ngữ để xử lý sự mơ hồ một cách rõ ràng và phân tích các lỗi liên quan đến mơ hồ trong các mô hình ngôn ngữ lớn (LLMs - Large Language Models) cung cấp những hướng đi đầy hứa hẹn để xây dựng các hệ thống vừa dễ diễn giải hơn, vừa hữu ích hơn trong các ứng dụng lấy con người làm trung tâm [8, 9].

2. CƠ SỞ LÍ LUẬN

Tiến bộ gần đây trong việc khử nhập nhằng nghĩa từ (WSD) được thúc đẩy bởi các biểu diễn ngôn ngữ theo ngữ cảnh và các chiến lược tinh chỉnh có giám sát. Các khảo sát và nghiên cứu chuẩn cho thấy các bộ mã hóa ngữ cảnh dựa trên bộ biến đổi, đã cải thiện đáng kể hiệu suất cho hệ thống WSD so với các phương pháp nhúng tĩnh trước đây. Tuy nhiên, hiệu quả thường không đồng đều: Nhiều phương pháp biến đổi dựa trên các tập hợp ngữ liệu chú thích lớn và tinh chỉnh cẩn thận, chúng vẫn có thể thất bại khi ngữ cảnh tối thiểu hoặc khi sự khác biệt về nghĩa là khó nhận biết. Những quan sát này thúc đẩy các giải pháp lai vẫn giữ được tính linh hoạt của các biểu diễn theo mạng neuron trong khi vẫn tận dụng thông tin nghĩa có cấu trúc [1]. Một luồng song song nhấn mạnh các mô hình ngôn ngữ được huấn luyện trước tăng cường kiến

thức KE-PLM (Knowledge-Enhanced Pre-trained Language Model) tích hợp các tài nguyên từ vựng và bách khoa vào bên trong mô hình. Các công trình được khảo sát trong vài năm qua, phân loại các phương pháp đưa tri thức vào (liên kết thực thể, nhúng đồ thị, chèn mềm các bộ ba) và báo cáo những cải tiến nhất quán cho các tác vụ ngữ nghĩa, trong đó thông tin thực tế hoặc định nghĩa hỗ trợ việc khử nhập nhằng. Mô hình KE-PLM cũng cải thiện khả năng diễn giải bằng cách phơi bày mối liên hệ giữa các quyết định của mô hình và kiến thức biểu tượng, một đặc tính có giá trị đối với các hệ thống ứng dụng cụ thể, vốn phải giải thích hoặc biện minh cho các lựa chọn cảm giác của người dùng. Bất chấp những tiến bộ, việc đưa vào các nguồn kiến thức không đồng nhất một cách hiệu quả và có khả năng mở rộng vẫn là một thách thức chưa có lời giải [2, 10].

Nền tảng đa phương thức đã nổi lên như một con đường quan trọng để giải quyết sự mơ hồ, đặc biệt là đối với các giác quan cụ thể tương ứng với các đối tượng tham chiếu có thể nhận dạng trực quan. Các tác vụ và tập dữ liệu được chia sẻ như Visual-WSD (SemEval-2023) [3] đã thúc đẩy các phương pháp kết hợp truy xuất văn bản và hình ảnh, bộ chuyển đổi đa phương thức và truy xuất nhận biết độ bóng để chọn ra các giác quan khi ngữ cảnh văn bản là tối thiểu. Kết quả cho thấy các tín hiệu đa phương thức có thể cải thiện đáng kể độ chính xác cho nhiều mục đa nghĩa, mặc dù hiệu suất phụ thuộc rất nhiều vào thiết kế tập dữ liệu và tính chân thực của các hình ảnh ứng viên. Nghiên cứu vẫn đang tiếp tục về sự liên kết hình ảnh-văn bản mạnh mẽ và việc mở rộng WSD đa phương thức vượt ra ngoài bối cảnh hình ảnh minh họa [3, 4].

Các mô hình ngôn ngữ lớn (LLMs) và các chiến lược dựa trên gợi ý gần đây đã được áp dụng cho các tác vụ xác định ngữ nghĩa, vừa để thực hiện giải mã hàm ý zero-/few-shot, vừa để mở rộng kho dữ liệu ngữ nghĩa cho các ngôn ngữ ít tài nguyên bằng cách tạo ra các ví dụ có chú thích (có gán nhãn). Các nghiên cứu đầu giai đoạn 2024-2025 chứng minh rằng các mô hình LLMs có thể tự khởi động tài nguyên và thực hiện các đánh giá theo kiểu WiC/WSD khi được gợi ý cẩn thận hoặc kết hợp với truy xuất bên ngoài, mặc dù vẫn còn những lo ngại về hiệu chuẩn mô hình, ảo giác và chuyển miền. Những phát triển này cho thấy LLMs sẽ đóng vai trò quan trọng trong các giải pháp WSD có khả năng mở rộng nếu được kết hợp với các giao thức đánh giá đo lường độ mạnh mẽ trên các ngôn ngữ và miền [8, 9].

Cuối cùng, những tiến bộ về phương pháp luận trong đánh giá và quy nạp khái niệm chỉ ra những hướng đi mới: (i) coi việc phát hiện sự mơ hồ và sự không chắc chắn được hiệu chuẩn là những đầu ra hạng nhất để hệ thống có thể trì hoãn hoặc yêu cầu làm rõ; (ii) quy nạp các khái niệm mang tính cảm quan từ các mô hình ngữ cảnh để giảm sự phụ thuộc vào các kho dữ liệu cố định; (iii) thiết kế các chuẩn mực tương tác, lấy con người làm trung tâm, phản ánh cách con người thể hiện và giải quyết sự mơ hồ trong giao tiếp thực tế. Những ưu tiên này làm nổi bật những thách thức nghiên cứu tại giao diện của học biểu diễn (Representation), tích hợp kiến thức và tương tác giữa người và máy tính [7, 9].

3. PHƯƠNG PHÁP NGHIÊN CỨU

Việc tổng quan tài liệu trong nghiên cứu này tuân theo một phương pháp luận có hệ thống được thiết kế để đảm bảo bao quát toàn diện các đóng góp gần đây về giải quyết sự mơ hồ trong xử lý ngôn ngữ tự nhiên. Chúng tôi đã áp dụng một phương pháp tiếp cận nhiều bước, bắt đầu bằng việc xây dựng truy vấn trên các cơ sở dữ liệu học thuật lớn, bao gồm: Thư viện kỹ thuật

số ACM, IEEE Xplore, ACL Anthology và SpringerLink. Các chuỗi tìm kiếm kết hợp các thuật ngữ như "giải quyết sự mơ hồ về nghĩa của từ", "giải quyết sự mơ hồ", "nhúng ngữ cảnh" và "ngôn ngữ học tính toán". Các ấn phẩm từ năm 2020 đến năm 2025 được ưu tiên, đảm bảo bao gồm những phát triển mới nhất. Việc sàng lọc ban đầu dựa trên tiêu đề và tóm tắt đã mang lại 220 bài báo, sau đó được giảm xuống còn 85 sau khi lọc mức độ liên quan [1, 2].

Bảng 1. Tóm tắt quá trình sàng lọc tài liệu (2020-2025)

Nguồn tham khảo	**Sử dụng từ khóa tìm kiếm**	**Kết quả sơ bộ**	**Sau khi lọc chủ đề có liên quan**
ACM Digital Library	"word sense disambiguation", "ambiguity resolution"	55	22
IEEE Xplore	"contextual embeddings", "computational linguistics"	40	15
ACL Anthology	"word sense disambiguation", "contextual embeddings"	70	30
SpringerLink	"ambiguity resolution", "computational linguistics"	55	18
Total	—	**220**	**85**

Để tinh chỉnh bộ dữ liệu, chúng tôi đã áp dụng các tiêu chí bao gồm và loại trừ. Các nghiên cứu được đưa vào nếu chúng giải quyết rõ ràng sự mơ hồ ở cấp độ từ vựng, cú pháp hoặc ngữ nghĩa, giới thiệu các phương pháp tính toán mới hoặc các mô hình được đánh giá trên các bộ dữ liệu chuẩn như SemEval hoặc WiC. Tiêu chí loại trừ bao gồm các bài báo chỉ tập trung vào các nhiệm vụ NLP không liên quan hoặc ngôn ngữ học lý thuyết mà không triển khai tính toán. Cả các bài báo trên tạp chí được bình duyệt ngang hàng và các kỷ yếu hội nghị có tác động cao (ACL, EMNLP, IJCAI) đều được ưu tiên để đảm bảo tính chặt chẽ và khả năng hiển thị của các đóng góp [3, 8]. Việc trích xuất dữ liệu bao gồm việc thu thập siêu dữ liệu (tác giả, năm, địa điểm), chi tiết phương pháp luận (loại mô hình, tài nguyên được sử dụng, tiêu chuẩn đánh giá) và kết quả được báo cáo. Đối với các nghiên cứu thực nghiệm, chúng tôi đã ghi lại các số liệu hiệu suất như độ chính xác, điểm F1 và nDCG khi áp dụng. Đối với các đóng góp về khảo sát và lý thuyết, chúng tôi đã trích xuất các phân loại của các phương pháp và nêu rõ các hạn chế. Trích xuất có cấu trúc cho phép so sánh giữa các phương pháp như hệ thống dựa trên quy tắc, phương pháp tiếp cận thần kinh, mô hình ngôn ngữ tăng cường kiến thức và chiến lược đa phương thức [4, 10].

Bảng 2. Tiêu chí bao gồm/loại trừ và tóm tắt trích xuất dữ liệu

Loại tiêu chí	**Mô tả**	**Minh họa**
Inclusion Criteria	- Addressed ambiguity at lexical, syntactic, or semantic level- Introduced computational models- Evaluated on	Kaur et al. (2023); Song et al. (2024)

	benchmark datasets (SemEval, WiC)- Published in peer-reviewed journals or major conferences (ACL, EMNLP, IJCAI)	
Exclusion Criteria	- Unrelated NLP tasks- Purely theoretical linguistics without computational implementation	Filtered out ≈ 135 papers
Metadata Extracted	Authors, year, venue, model type, resources used, evaluation benchmarks	e.g., ACL Anthology entries
Empirical Metrics	Accuracy, F1 score, nDCG (when applicable)	For benchmark-based studies
Survey/Review Data	Taxonomies of approaches, reported limitations	Xu et al. (2023); Kritharoula et al. (2023)
Structured Comparison	Categorized into rule-based, neural, knowledge-enhanced, multimodal approaches	Enabled cross-method analysis

Để nâng cao độ tin cậy, chúng tôi đã áp dụng phương pháp kiểm chứng chéo nguồn và tam giác hóa. Hai nhà nghiên cứu độc lập đã mã hóa các chủ đề phương pháp luận và giải quyết các điểm khác biệt thông qua thảo luận. Phương pháp lập bản đồ thư mục cũng được sử dụng để xác định các mạng lưới trích dẫn và cụm nghiên cứu, sử dụng các công cụ như VOSviewer và CiteSpace. Phương pháp lập bản đồ này làm nổi bật các công trình có ảnh hưởng và các lĩnh vực mới nổi như WSD đa phương thức và giải quyết nhập nhằng dựa trên LLMs. Bằng cách kết hợp tổng hợp định tính với lập bản đồ định lượng, bài tổng quan đã đảm bảo sự cân bằng giữa chiều sâu diễn giải và phạm vi bao quát [7, 9].

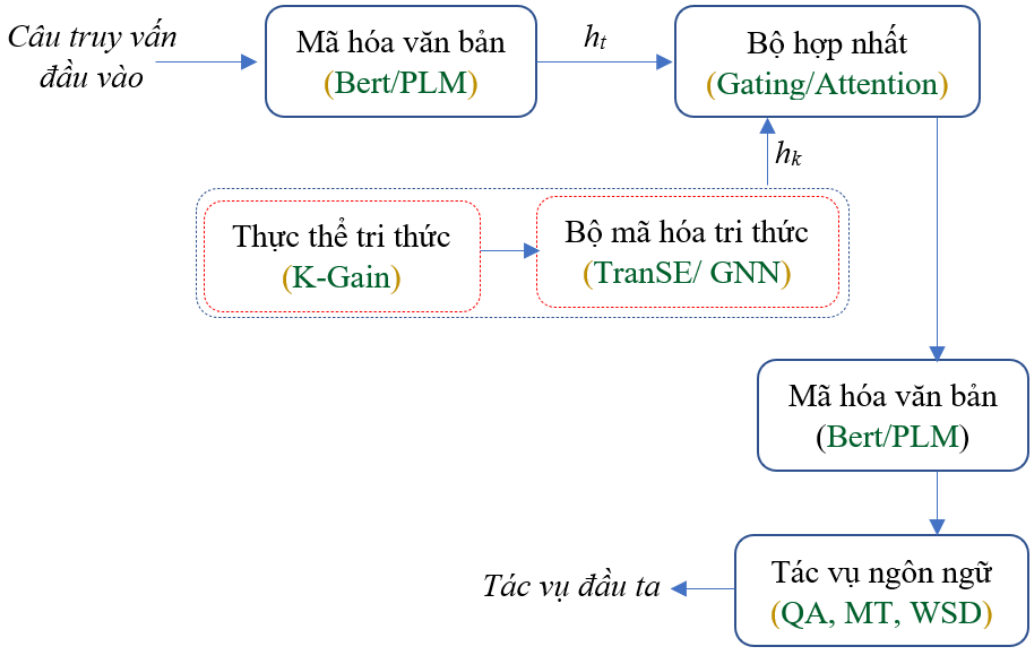
Cuối cùng, phương pháp luận này kết hợp đánh giá phê bình các tài liệu đã chọn. Mỗi nghiên cứu được đánh giá không chỉ dựa trên kết quả báo cáo mà còn dựa trên khả năng tái tạo, tính đại diện của tập dữ liệu và khả năng áp dụng vào bối cảnh thực tế. Những hạn chế về tính khả dụng của dữ liệu, sự thiên vị nguồn lực đối với tiếng Việt và sự phụ thuộc vào các mô hình quy mô lớn đã được ghi chép rõ ràng. Do đó, phương pháp luận đánh giá đã đảm bảo một quy trình nghiêm ngặt, có thể tái tạo, đưa ra những phát hiện phản ánh tình trạng hiện tại và những thách thức của việc giải quyết sự mơ hồ trong ngôn ngữ học tính toán [8, 9].

4. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

Đánh giá và tổng hợp của chúng tôi nhấn mạnh rằng các kỹ thuật ngôn ngữ học tính toán kết hợp với các phương pháp học máy hiện đại đã cải thiện đáng kể khả năng giải quyết sự mơ hồ, mặc dù vẫn còn những hạn chế do tính đặc thù của từng ngôn ngữ (trong đó có tiếng Việt). Kết quả từ các mô hình dựa trên bộ chuyển đổi như BERT và RoBERT cho thấy những cải tiến đáng kể trong các điểm chuẩn như các tác vụ WSD WiC và SemEval, đạt được độ nhạy ngữ cảnh cao hơn so với các phương pháp truyền thống. Tuy nhiên, các mô hình này vẫn dễ bị tổn

thương trước các tình huống ngữ cảnh tối thiểu, trong đó sự khác biệt về nghĩa rất khó nhận biết hoặc mang tính đặc thù của từng miền. Các nghiên cứu cũng chỉ ra sự khác biệt giữa các ngôn ngữ như: Trong khi tiếng Anh và các ngôn ngữ có nguồn lực cao khác được hưởng lợi từ các tập hợp ngữ liệu chú thích lớn (đã được đầu tư từ nhiều nguồn tài chính như WordNet tiếng Anh), thì các ngôn ngữ có nguồn lực thấp vẫn chưa được khai thác đầy đủ (cụ thể là ngôn ngữ nghèo tài nguyên như tiếng Việt). Những phát hiện này khẳng định rằng sự tiến bộ không đồng đều, nhấn mạnh nhu cầu về các phương pháp tiếp cận kết hợp cân bằng giữa các biểu diễn đã học với kiến thức ngôn ngữ biểu tượng để giải quyết sự mơ hồ mạnh mẽ trong các bối cảnh đa dạng.

Một kết quả quan trọng khác là hiệu quả đã được chứng minh của các phương pháp tiếp cận đa phương thức và tăng cường kiến thức. Việc tích hợp các nguồn tài nguyên từ vựng và bách khoa bên ngoài vào các mô hình được huấn luyện trước như KE-PLM (Knowledge-Enhanced Pre-trained Language Model), đã mang lại những cải thiện đáng kể về khả năng diễn giải và độ chính xác, đặc biệt là trong các lĩnh vực đòi hỏi nền tảng thực tế. Tương tự, các phương pháp đa phương thức tận dụng hình ảnh, âm thanh hoặc siêu dữ liệu ngữ cảnh đã cải thiện hiệu suất khử nhập nhằng trong các tác vụ Visual-WSD, đặc biệt đối với các danh từ đa nghĩa có tham chiếu trực quan rõ ràng. Tuy nhiên, vẫn còn những hạn chế rõ ràng: Các phương pháp đa phương thức phụ thuộc rất nhiều vào các tập dữ liệu được quản lý và thường thất bại trong các môi trường thực tế nhiều nhiễu. Phân tích của chúng tôi chỉ ra rằng mặc dù nền tảng đa phương thức mở rộng phạm vi khử nhập nhằng, nhưng nó không thể thay thế hoàn toàn nhu cầu mô hình hóa ngôn ngữ và ngữ cảnh. Thay vào đó, tính bổ sung giữa các tín hiệu văn bản và phi văn bản mở ra một hướng đi đầy hứa hẹn cho các hệ thống lấy con người làm trung tâm.



Hình 1. Kiến trúc tổng quát của mô hình KE-PLM

Công thức phổ biến sử dụng attention-based fusion giữa biểu diễn văn bản và kiến thức có thể được mô tả thông qua các chỉ số theo tiêu chuẩn nhúng từ và biểu đồ biểu diễn ngôn ngữ.

Công thức (1) dưới đây biểu diễn minh họa cho mô hình kiến trúc KE-PLM có thể được áp dụng hỗ trợ giảm thiểu sự nhập nhằng của ngôn ngữ tự nhiên.

$$h = \alpha \cdot h_t + (1 - \alpha) \cdot h_k \qquad (1)$$

Trong đó:

h_t: nhúng từ (embedding) từ mô hình ngôn ngữ (PLM).

h_k: nhúng từ (embedding) từ biểu đồ tri thức (Knowledge Graph).

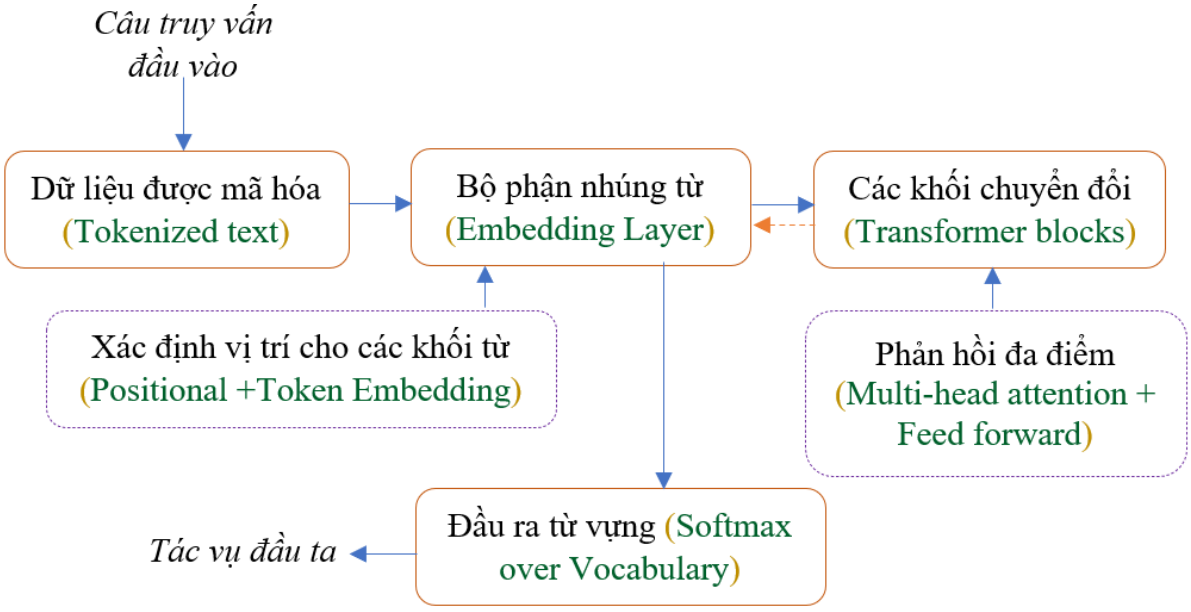
$\alpha = \sigma(W[h_t; h_k] + b)$: trọng số động học từ cơ chế gating (sigmoid).

Mô hình KE-PLM cũng có thể sử dụng cơ chế Knowledge-Enhanced Attention, xác suất attention được hiệu chỉnh theo chỉ số của hàm softmax như công thức (2).

$$Attn(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d}} + M_{KG}\right)V \qquad (2)$$

Trong đó M_{KG} là ma trận bias từ kiến thức (knowledge bias), giúp điều chỉnh xác suất chú ý theo quan hệ trong KG (Knowledge Gain).

Cuối cùng, tổng hợp của chúng tôi cho thấy các mô hình ngôn ngữ lớn (LLMs) mở ra những cơ hội mới cho việc giải quyết sự mơ hồ có thể mở rộng quy mô, nhưng cũng đặt ra những thách thức quan trọng. Việc gợi ý bằng zero-shot và few-shot với mô hình ngôn ngữ lớn LLMs cho thấy khả năng cơ sở mạnh mẽ, cho phép loại bỏ sự mơ hồ về nghĩa mà không cần dữ liệu huấn luyện rõ ràng. Hơn nữa, các mô hình LLMs đã được sử dụng để khởi động các tài nguyên chú thích nghĩa mới trong các ngôn ngữ có ít tài nguyên, mở rộng khả năng ứng dụng của các phương pháp WSD. Tuy nhiên, những lo ngại về ảo giác, hiệu chuẩn và thiếu minh bạch vẫn tồn tại, đặc biệt là trong các ứng dụng có rủi ro cao như diễn giải văn bản y tế hoặc pháp lý. Các thảo luận trong tài liệu được xem xét nhấn mạnh tầm quan trọng của việc phát hiện sự mơ hồ hơn là giải quyết cưỡng bức, với các hệ thống được khuyến khích báo hiệu sự không chắc chắn và yêu cầu làm rõ. Điều này phù hợp với các hoạt động giao tiếp của con người và chủ đề rộng hơn của hội nghị là "Công nghệ dành cho con người".



Hình 2. Mô hình tổng quát của hệ thống LLMs

5. KẾT LUẬN

Nghiên cứu này đã xem xét vai trò của ngôn ngữ học tính toán trong việc giải quyết sự mơ hồ trong xử lý ngôn ngữ tự nhiên, tổng hợp những tiến bộ gần đây trên các cấp độ từ vựng, cú pháp và ngữ nghĩa. Kết quả cho thấy các kỹ thuật nhúng ngữ cảnh dựa trên bộ chuyển đổi, các mô hình được huấn luyện trước được tăng cường kiến thức và các chiến lược đa phương thức đã thúc đẩy đáng kể việc giải quyết sự nhập nhằng trong xử lý ngôn ngữ tự nhiên. Tuy nhiên, vẫn còn nhiều thách thức trong các ngôn ngữ có ít tài nguyên, việc thích ứng với miền và việc triển khai trong thế giới thực, nơi ngữ cảnh có thể bị hạn chế và khả năng diễn giải mô hình vẫn chưa đủ. Những phát hiện của chúng tôi củng cố giá trị của các phương pháp tiếp cận kết hợp, tích hợp kiến thức ngôn ngữ với các phương pháp học máy, làm nổi bật sự cần thiết của các đầu ra, nhận biết sự không chắc chắn và phát hiện sự mơ hồ rõ ràng. Những hiểu biết này không chỉ củng cố nền tảng kỹ thuật của việc giải quyết sự mơ hồ mà còn phù hợp với tầm nhìn rộng hơn về ngôn ngữ học tính toán lấy con người làm trung tâm.

Nhìn về tương lai, nghiên cứu nên ưu tiên ba hướng. Thứ nhất, cần phát triển các khuôn khổ phát hiện mơ hồ có khả năng mở rộng để cho phép các hệ thống nhận biết và đánh dấu sự không chắc chắn thay vì phân loại sai một cách quá tự tin. Thứ hai, việc mở rộng các tập dữ liệu đa phương thức và đa ngôn ngữ sẽ rất cần thiết để đảm bảo khả năng áp dụng công bằng trên các ngôn ngữ và bối cảnh giao tiếp đa dạng. Thứ ba, những tiến bộ trong các mô hình ngôn ngữ lớn đòi hỏi sự tích hợp có trách nhiệm, nhấn mạnh vào việc hiệu chuẩn, minh bạch và tương tác người dùng để làm rõ. Sự hợp tác giữa ngôn ngữ học tính toán, khoa học nhận thức và tương tác giữa người và máy tính sẽ rất quan trọng để đạt được các hệ thống, vừa nâng cao độ chính xác của việc loại bỏ mơ hồ vừa thúc đẩy sự tin cậy. Bằng cách theo đuổi những hướng này, ngôn ngữ học tính toán có thể tiếp tục phát triển hướng tới các công nghệ thực sự hỗ trợ sự hiểu biết và giao tiếp của con người.

TÀI LIỆU THAM KHẢO

[1] Bevilacqua, M., Pasini, T., Raganato, A., & Navigli, R. (2021). Recent trends in word sense disambiguation: A survey. Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI), 4330–4338. https://doi.org/10.24963/ijcai.2021/595

[2] Huang, L., Cao, Y., Ji, H., & Knight, K. (2021). A survey of knowledge-enhanced pre-trained language models. ArXiv preprint arXiv:2104.06396. https://arxiv.org/abs/2104.06396

[3] Kaur, S., Raganato, A., & Navigli, R. (2023). SemEval-2023 Task 1: Visual word sense disambiguation. Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023), 1–15. Association for Computational Linguistics. https://aclanthology.org/2023.semeval-1.1

[4] Kritharoula, A., Pasini, T., Bevilacqua, M., & Navigli, R. (2023). Leveraging multimodal transformers for visual word sense disambiguation. Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP), 5674–5686. https://doi.org/10.18653/v1/2023.emnlp-main.452

[5] Li, Z., Yang, J., & Sun, M. (2022). A comprehensive survey of multimodal word sense disambiguation. ACM Transactions on Asian and Low-Resource Language Information Processing, 21(4), 1–24. https://doi.org/10.1145/3514257

[6] Navigli, R., Bevilacqua, M., Conia, S., & Miller, T. (2020). SyntagNet: Challenging supervised word sense disambiguation with lexical-semantic combinations. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 3229–3240. https://doi.org/10.18653/v1/2020.acl-main.294

[7] Federico Martelli, Najla Kalach, Gabriele Tola, Roberto Navigli (2012). Multilingual and cross-lingual word-in-context disambiguation. Computational Linguistics, 48(2), 235–270. https://aclanthology.org/2021.semeval-1.3/. DOI: 10.18653/v1/2021.semeval-1.3

[8] Song, X., He, J., & Zhang, Y. (2024). Large language models for word sense disambiguation: Opportunities and challenges. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP). https://arxiv.org/abs/2407.01234

[9] Sun, H., Zhao, T., & Xu, J. (2025). Concept induction and ambiguity resolution in large language models. ArXiv preprint arXiv:2501.04567. https://arxiv.org/abs/2501.04567

[10] Jian Yang, Xinyu Hu, Gang Xiao, Yulong Shen. (2023). A Survey of Knowledge Enhanced Pre-trained Language Models. J. ACM, Vol. 37, No. 4, Article 111. Publication date: August 2023. https://arxiv.org/pdf/2110.00269.

Huỳnh Quang Đức: Thạc sĩ, đang giảng dạy tại Khoa Công nghệ thông tin, Robot và Trí tuệ nhân tạo, Trường Đại học Bình Dương. Số 504 Đại Lộ Bình Dương, Phường Phú Lợi, thành phố Hồ Chí Minh. Email: hqduc@bdu.edu.vn.

NÂNG CAO HIỆU QUẢ VÀ THÚC ĐẨY HIỂU BIẾT LIÊN VĂN HÓA QUA TÍCH HỢP NLP VÀO DẠY VÀ HỌC NGOẠI NGỮ

Trần Thị Hải Bình [1]

Tóm tắt

Trong bối cảnh toàn cầu hóa, năng lực giao tiếp liên văn hóa ngày càng trở thành yêu cầu trọng yếu trong dạy và học ngoại ngữ. Công nghệ xử lý ngôn ngữ tự nhiên (NLP) nổi lên như một công cụ hỗ trợ mạnh mẽ, không chỉ giúp vượt qua rào cản ngôn ngữ mà còn mở rộng khả năng tiếp cận tri thức đa dạng. Bài báo này tập trung phân tích tiềm năng ứng dụng NLP trong dạy và học ngoại ngữ, từ xây dựng nguồn tài nguyên đến cung cấp phản hồi cho kỹ năng ngôn ngữ. Qua đó, NLP được nhìn nhận như một giải pháp góp phần nâng cao hiệu quả dạy và học, đồng thời thúc đẩy hiểu biết liên văn hóa ở người học. Bài báo đồng thời chỉ ra những thách thức và cũng đề xuất một số định hướng ứng dụng để bảo đảm tính khả thi và bền vững trong thực tiễn triển khai.

Từ khóa: NLP, Dạy và học ngoại ngữ, Năng lực giao tiếp liên văn hóa, Ứng dụng, Thách thức.

1. Mở đầu

Trong bối cảnh toàn cầu hóa hiện nay, ngoại ngữ không chỉ là một phương tiện giao tiếp mà còn là chìa khóa để con người tiếp cận tri thức, văn hóa và cơ hội nghề nghiệp trên phạm vi toàn cầu. Việc học và sử dụng ngoại ngữ gắn liền với năng lực giao tiếp liên văn hóa – khả năng hiểu biết, tôn trọng và tương tác hiệu quả với những cá nhân đến từ các nền văn hóa khác nhau.

Trong khi nhu cầu về năng lực liên văn hóa tăng cao, công nghệ cũng chứng kiến sự phát triển mạnh mẽ, đặc biệt trong lĩnh vực xử lý ngôn ngữ tự nhiên (Natural Language Processing – NLP). NLP được hiểu là lĩnh vực nghiên cứu và ứng dụng các thuật toán, mô hình máy tính để phân tích, xử lý và sinh ngôn ngữ tự nhiên của con người. Ứng dụng của NLP đã và đang len lỏi vào đời sống thường ngày, đặc biệt trong lĩnh vực dạy và học. Với tiềm năng to lớn, NLP không chỉ giúp phá vỡ rào cản ngôn ngữ mà còn mở ra những phương thức mới trong việc giảng dạy và học tập ngoại ngữ.

Các nghiên cứu gần đây chứng minh rằng việc tích hợp công nghệ vào dạy và học ngoại ngữ, đặc biệt là các công cụ NLP, mang lại nhiều lợi ích đáng kể. Tuy nhiên, việc ứng dụng NLP trong giáo dục ngôn ngữ không chỉ dừng lại ở khía cạnh kỹ thuật. Quan trọng hơn, NLP có thể trở thành một công cụ thúc đẩy sự hiểu biết liên văn hóa, bởi nó cho phép người học tiếp cận và so sánh ngôn ngữ, văn hóa khác nhau một cách trực tiếp và đa chiều. Từ đó, NLP có thể đóng vai trò như một cầu nối giúp hình thành những công dân toàn cầu có khả năng đối thoại, hợp tác và thích ứng trong môi trường đa ngôn ngữ, đa văn hóa.

[1] Tiến sĩ, Giảng viên – Học viện Ngân hàng, *Email: binhtth@hvnh.edu.vn*

Mặc dù tiềm năng rất lớn, song việc tích hợp NLP vào dạy và học ngoại ngữ cũng đặt ra nhiều câu hỏi học thuật và thực tiễn: NLP có thực sự nâng cao năng lực liên văn hóa, hay chỉ dừng lại ở mức hỗ trợ kỹ năng ngôn ngữ cơ bản? Liệu người học có nguy cơ phụ thuộc vào công cụ mà bỏ qua quá trình phát triển tư duy phản biện và sáng tạo? Làm thế nào để giáo viên cân bằng giữa việc tận dụng lợi ích công nghệ và duy trì sự chủ động, tích cực của người học? Những vấn đề này cần được nghiên cứu và thảo luận sâu hơn để hướng tới những giải pháp bền vững và hiệu quả.

Xuất phát từ những lý do trên, bài báo này tập trung phân tích tiềm năng và thách thức của việc tích hợp NLP vào giảng dạy ngoại ngữ, với trọng tâm là tác động của nó đối với việc thúc đẩy hiểu biết liên văn hóa. Trên cơ sở đó, tác giả đề xuất một số định hướng nhằm tối ưu hóa việc sử dụng NLP trong lớp học ngoại ngữ.

1.1. Công nghệ xử lý ngôn ngữ tự nhiên (NLP)

Công nghệ xử lý ngôn ngữ tự nhiên (Natural Language Processing – NLP) là một lĩnh vực liên ngành kết hợp giữa ngôn ngữ học, khoa học máy tính và trí tuệ nhân tạo, với mục tiêu giúp máy tính có thể "hiểu" và sử dụng ngôn ngữ của con người. NLP bao gồm nhiều cấp độ xử lý, từ phân tích hình thái, cú pháp, ngữ nghĩa cho đến ngữ dụng, cho phép hệ thống không chỉ nhận diện từ ngữ ở mức bề mặt mà còn nắm bắt được ý nghĩa, sắc thái và chức năng giao tiếp.

Trải qua quá trình phát triển, NLP đã trở thành một trong những trụ cột của công nghệ hiện đại. Từ các hệ thống dịch tự động (Google Translate, DeepL), trợ lý ảo thông minh (Siri, Alexa, ChatGPT) cho đến những ứng dụng hỗ trợ học ngôn ngữ (Duolingo, Grammarly), NLP ngày càng khẳng định vai trò trong việc kết nối ngôn ngữ và tri thức toàn cầu. Các công trình gần đây của Jurafsky & Martin (2023) cho thấy sự chuyển dịch mạnh mẽ từ các phương pháp dựa trên luật (rule-based) và mô hình thống kê truyền thống sang các kiến trúc học sâu (deep learning) và mô hình ngôn ngữ quy mô lớn (LLMs). Nhờ vào kho ngữ liệu đồ sộ (big data) và khả năng huấn luyện trên quy mô khổng lồ, NLP hiện không chỉ dừng lại ở việc xử lý bề mặt ngôn từ mà còn có thể nhận diện ngữ cảnh, sắc thái nghĩa và yếu tố ngữ dụng trong giao tiếp.

Trong dạy và học ngoại ngữ, những tiến bộ này mở ra nhiều triển vọng ứng dụng. NLP cho phép thiết kế hệ thống giảng dạy cá nhân hóa, cung cấp phản hồi tự động về phát âm, từ vựng và ngữ pháp, cũng như hỗ trợ đánh giá năng lực ngôn ngữ một cách khách quan và linh hoạt hơn. Bên cạnh đó, NLP còn tạo điều kiện để người học tiếp cận kho ngữ liệu đa dạng, qua đó nâng cao khả năng phân tích ngôn ngữ và hiểu biết văn hóa – những yếu tố then chốt trong bối cảnh toàn cầu hóa hiện nay.

1.3. Năng lực liên văn hóa trong dạy và học ngoại ngữ

Trong bối cảnh toàn cầu hóa, việc dạy và học ngoại ngữ không chỉ giới hạn ở năng lực ngôn ngữ mà còn đòi hỏi sự phát triển năng lực liên văn hóa (intercultural competence – IC). Khái niệm này được nhiều học giả định nghĩa với những điểm nhấn khác nhau, song đều thống nhất

rằng IC bao gồm khả năng hiểu, phân tích, so sánh và điều chỉnh hành vi giao tiếp trong các tình huống có sự tham gia của nhiều hệ giá trị văn hóa. Một trong những mô hình IC có ảnh hưởng được đề xuất bởi Deardorff (2006) trong đó coi IC như một quá trình liên tục cải thiện. Mô hình này bắt đầu bằng thái độ của người học như tôn trọng, cởi mở, và tò mò và dần dần nhận thức và hiểu được văn hóa (tự nhận thức văn hóa, có kiến thức sâu về văn hóa). Những yếu tố này sau đó thúc đẩy các kết quả bên trong (internal outcomes) như sự đồng cảm và khả năng thích ứng, cũng như các kết quả bên ngoài (external outcomes) như giao tiếp và hành vi hiệu quả trong các tình huống liên văn hóa. Tương tự như vậy, mô hình Năng lực giao tiếp liên văn hóa (ICC) của Byram (2021) đề cập đến 05 thành phần chính và đã có ảnh hưởng lớn đến phát triển chương trình giảng dạy ngoại ngữ, bao gồm:

- Kiến thức: Hiểu các nhóm xã hội, các hoạt động của họ và các quá trình tương tác trong nền văn hóa của mình và của người đối thoại.
- Thái độ: Sự tò mò, cởi mở và sẵn sàng gạt bỏ hoài nghi về các nền văn hóa khác, có niềm tin vào nền văn hóa của chính mình.
- Kỹ năng diễn giải và liên hệ: Khả năng diễn giải và liên hệ các tài liệu, sự kiện từ nền văn hóa khác với nền văn hóa của mình.
- Kỹ năng khám phá và tương tác: Khả năng tiếp thu kiến thức văn hóa mới và áp dụng vào giao tiếp thời gian thực.
- Nhận thức văn hóa phản biện: Khả năng đánh giá phản biện các quan điểm, hoạt động và sản phẩm trong nền văn hóa của mình và các nền văn hóa khác.

Trong lớp học ngoại ngữ, IC được xem là một mục tiêu đào tạo thiết yếu, bởi người học không chỉ học cách "nói đúng" mà còn học cách "nói phù hợp". Điều này đặc biệt quan trọng khi ngôn ngữ gắn liền với bối cảnh văn hóa – xã hội đặc thù. Ví dụ, việc sử dụng cách xưng hô, mức độ trực tiếp/gián tiếp trong lời nói, hay chiến lược lịch sự đều chịu sự chi phối của chuẩn mực văn hóa. Nếu người học chỉ tiếp thu ngôn ngữ ở mức cấu trúc và từ vựng mà thiếu ý thức liên văn hóa, họ dễ gặp rào cản trong giao tiếp thực tiễn hoặc thậm chí gây hiểu lầm, xung đột.

Về mặt phương pháp, việc tích hợp IC trong dạy học ngoại ngữ cần dựa trên tiếp cận lấy người học làm trung tâm và gắn kết ngôn ngữ với tình huống thực tế. Một số chiến lược có thể kể đến là: sử dụng tài liệu thực tế như phim, báo chí, diễn đàn trực tuyến; tổ chức các hoạt động so sánh – đối chiếu văn hóa giữa ngôn ngữ mục tiêu và ngôn ngữ mẹ đẻ; khuyến khích phản tư về trải nghiệm văn hóa cá nhân; hoặc kết nối trực tiếp với người bản ngữ qua các dự án trao đổi (virtual exchange, telecollaboration). Những cách tiếp cận này không chỉ củng cố vốn ngôn ngữ mà còn tạo cơ hội hình thành "công dân toàn cầu" có khả năng giao tiếp hiệu quả, linh hoạt và có trách nhiệm.

2. NLP trong dạy và học ngoại ngữ

2.1. Xây dựng tài liệu và chương trình học

Một trong những thách thức lớn trong dạy ngoại ngữ là giúp người học không chỉ nắm được ngôn ngữ mà còn hiểu rõ ngữ cảnh văn hóa đi kèm. NLP có khả năng xử lý và phân tích khối lượng lớn dữ liệu ngôn ngữ xác thực (authentic materials), từ đó hỗ trợ thiết kế tài liệu giảng dạy giàu yếu tố văn hóa. Chẳng hạn, việc khai thác ngữ liệu song ngữ và đa ngôn ngữ (parallel corpora) cho phép người học so sánh cách diễn đạt trong bối cảnh xã hội – văn hóa khác nhau (McEnery & Hardie, 2012), giúp người học tiếp cận cách dùng từ, cụm từ và phong cách ngôn ngữ trong ngữ cảnh tự nhiên, từ đó nâng cao nhận thức và khả năng sử dụng ngôn ngữ trong giao tiếp thực tiễn. NLP cũng có khả năng cá nhân hóa tài liệu học, Eggert (2021) đã đề xuất sử dụng AI để cải tiến dạy học bằng cách thu thập các dữ liệu liên quan đến kiến thức của người học, trạng thái tâm lí, hoàn cảnh xã hội… để điều chỉnh phương pháp giảng dạy thông qua các nền tảng học tập tương ứng. Giáo viên có thể sử dụng linh hoạt hệ thống học tập thông minh (Intelligent Tutoring Systems – ITS) và Chatbot được hỗ trợ bởi NLP có khả năng tùy chỉnh nội dung theo nhu cầu riêng của từng người học. Các hệ thống học tập thích ứng (adaptive learning systems) dựa trên NLP có thể phân tích dữ liệu học tập của từng cá nhân, nhận diện điểm mạnh, điểm yếu và từ đó gợi ý lộ trình học phù hợp. Ví dụ, nếu một người học thường xuyên mắc lỗi về thì quá khứ đơn, hệ thống có thể tự động cung cấp thêm bài tập hoặc gợi ý luyện tập bổ sung. Nếu một người khác có vốn từ hạn chế trong lĩnh vực học thuật, NLP có thể gợi ý các tài liệu, bài đọc hoặc hoạt động tập trung vào loại từ vựng đó. Cách tiếp cận này không chỉ tăng hiệu quả học tập mà còn thúc đẩy động lực nội tại vì người học cảm thấy hành trình học được thiết kế "riêng cho mình", phù hợp với năng lực của mình.

Các ứng dụng tương tác như chatbot hay trò chơi hóa (gamification) giúp việc học trở nên hấp dẫn và ít áp lực hơn. Nghiên cứu chỉ ra rằng công cụ này có thể cải thiện thái độ học tập, nâng cao sự chủ động và cam kết của người học. NLP được tích hợp trong các mô phỏng hội thoại đa văn hóa, tái hiện tình huống giao tiếp liên văn hóa để người học rèn luyện khả năng ứng xử trong môi trường đa dạng văn hóa. Theo Zhai và cộng sự (2022), chatbot AI kết hợp các yếu tố văn hóa, đồng cảm và hài hước có thể ảnh hưởng tích cực đến kết quả học tập của người học. Những công cụ này có thể thúc đẩy nhận thức văn hóa, hỗ trợ cảm xúc để giảm bớt lo lắng khi học và xây dựng mối quan hệ tích cực giữa người học và chatbot.

Các chương trình trao đổi trực tuyến sử dụng công nghệ để kết nối sinh viên từ các quốc gia khác nhau cho các dự án hợp tác, tạo bối cảnh chân thực để phát triển kỹ năng giao tiếp (IC). Các chương trình trao đổi trực tuyến đã được chứng minh là cải thiện tư duy toàn cầu, trí tuệ văn hóa và khả năng tiếp thu các quan điểm khác nhau của sinh viên. Nghiên cứu của Zheng và đồng nghiệp (2022) cho thấy một dự án trao đổi trực tuyến được lồng ghép vào chương trình du học đã dẫn đến những thay đổi tích cực đáng kể trong kỹ năng giao tiếp của sinh viên, đặc biệt là ở khía cạnh giao tiếp cá nhân, bao gồm định hướng thế giới và phát triển mối quan hệ.

Ngoài ra, dịch máy (Machine Translation) không chỉ giúp mở rộng vốn từ mà còn tạo cơ hội phản tư về sự khác biệt văn hóa giữa các ngôn ngữ (Gaspari, 2015). Các công cụ như Google Translate, DeepL hay Microsoft Translator đã trở thành "trợ thủ" quen thuộc của người học ngoại ngữ. Ưu điểm nổi bật của các dịch máy được hỗ trợ bởi NLP hiện nay là tốc độ xử lý nhanh, khả năng dịch đa ngữ, và ngày càng cải thiện về độ chính xác nhờ các mô hình ngôn ngữ lớn (large language models – LLMs). Các hệ thống dịch đa ngữ, phân tích ngữ cảnh hoặc chatbot đa ngôn ngữ có thể cung cấp thông tin về phong tục, lối nói ẩn dụ, hay bối cảnh văn hóa đi kèm. Ví dụ, một chatbot dựa trên NLP không chỉ dịch một cụm từ sang tiếng Việt, mà còn có thể giải thích nếu cụm từ đó là một thành ngữ diễn đạt một nghĩa khác biệt so với nghĩa đen của các từ có trong cụm từ. Trong giảng dạy, các dịch máy này có thể được sử dụng như một công cụ so sánh và phân tích, giúp người học đối chiếu sự khác biệt về cấu trúc cú pháp, cách dùng từ và sắc thái biểu đạt giữa hai ngôn ngữ. Thông qua quá trình này, người học không chỉ mở rộng vốn từ mà còn rèn luyện năng lực liên văn hóa (Busch, 2024), bởi mỗi bản dịch đều phản ánh phần nào sự khác biệt về tư duy và bối cảnh văn hóa. Nhờ vậy, người học không chỉ tiếp nhận ngôn ngữ mà còn được tiếp cận với cách tư duy và biểu đạt văn hóa khác biệt. Điều này đặc biệt quan trọng trong quá trình hình thành năng lực liên văn hóa, bởi nó giúp người học nhận diện và chấp nhận sự đa dạng, đồng thời rèn luyện khả năng diễn giải, đối thoại và thích ứng trong môi trường đa văn hóa.

2.2. Hỗ trợ phản hồi

Một điểm mạnh của NLP là khả năng tự động phân tích và phản hồi không chỉ về ngôn ngữ mà cả về lựa chọn diễn ngôn trong bối cảnh văn hóa. Viết và nói thường được xem là những kỹ năng khó đối với người học ngoại ngữ. Các phần mềm dựa trên NLP có thể tự động nhận diện cấu trúc ngữ pháp, chỉ ra lỗi sai, hoặc gợi ý cách diễn đạt tự nhiên hơn. Hệ thống kiểm tra ngữ pháp như Grammarly, ProWritingAid hay LanguageTool hỗ trợ sửa lỗi cú pháp, dấu câu, đồng thời gợi ý phong cách viết học thuật. Li và Hegelheimer (2013) chứng minh rằng các công cụ kiểm tra chính tả và ngữ pháp như Grammarly có tác dụng rõ rệt trong việc nâng cao chất lượng bài viết học thuật của sinh viên nhờ có sự tích hợp gợi ý về "tone" và "formality", giúp người học ý thức hơn về sắc thái ngôn ngữ phù hợp từng hoàn cảnh. Nhiều hệ thống sử dụng hội thoại tương tác để phân tích bản ghi của người học ngoại ngữ như ELSA Speak, TFlat, và Speak English Fluently, từ đó cung cấp phản hồi dựa trên lỗi nhằm nâng cao kỹ năng nói tiếng Anh. Những công cụ này không chỉ giúp người học cải thiện phát âm chuẩn hơn mà còn tạo điều kiện luyện tập trong môi trường "an toàn", không lo ngại bị đánh giá từ người khác.

Việc không chỉ sửa lỗi mà còn đưa ra giải thích chi tiết của các công cụ sử dụng NLP giúp người học hiểu rõ sai sót và điều chỉnh nhanh chóng, qua đó nâng cao "ý thức ngôn ngữ" (language awareness), một năng lực quan trọng trong việc tiếp thu ngoại ngữ. Thông qua việc thường xuyên tiếp nhận phản hồi tự động từ các công cụ này, người học có cơ hội nhận diện quy tắc ngôn ngữ, dần hình thành năng lực tự sửa lỗi – một yếu tố quan trọng để phát triển khả năng giao tiếp độc

lập. Trong tương lai, các mô hình NLP có thể tiến xa hơn, hỗ trợ đánh giá khả năng sử dụng chiến lược lịch sự, cách thức xưng hô, hoặc mức độ gián tiếp/ trực tiếp trong giao tiếp – những yếu tố gắn liền với năng lực liên văn hóa (Thorne, 2016).

Với tính chất phản hồi tức thì và liên tục, NLP giúp giáo viên tự động hóa việc đánh giá các kỹ năng ngôn ngữ liên quan đến bề mặt của ngôn ngữ như ngữ pháp và từ vựng, góp phần giảm tải công việc chấm điểm cơ học. Một ưu điểm nữa của NLP đối với giáo viên trong quá trình xử lí và đưa ra phản hồi là đồng thời cung cấp dữ liệu chi tiết về xu hướng mắc lỗi liên văn hóa của người học. Điều này giúp giáo viên tập trung vào hướng dẫn các yếu tố văn hóa và phát triển tư duy phê phán, thay vì chỉ sửa lỗi ngôn ngữ bề mặt.

3. Thách thức và đề xuất

Mặc dù sở hữu tiềm năng to lớn, việc tích hợp NLP trong giáo dục ngoại ngữ vẫn đối mặt với nhiều thách thức. Trước hết là vấn đề độ chính xác và tính tin cậy. Các dịch máy tự động dựa trên NLP, dù đã đạt nhiều tiến bộ, vẫn thường mắc lỗi với ngôn ngữ ít tài nguyên hoặc trong việc xử lý các cấu trúc phức tạp, đặc biệt là thành ngữ và ẩn dụ. Ngoài ra, xử lý dữ liệu NLP còn gặp vấn đề thiên vị thuật toán (Selvam và cộng sự, 2025). Các mô hình NLP được đào tạo dựa trên các mẫu ngôn ngữ phổ biến có thể loại trừ các giọng không phải tiếng mẹ đẻ, phương ngữ địa phương và bối cảnh văn hóa chưa được đại diện, từ đó củng cố định kiến và tạo ra kết quả học tập không công bằng. Điều này đặc biệt đúng với các dịch máy thuật AI, vốn thường mặc định sử dụng tiếng Anh làm ngôn ngữ trung gian, có khả năng làm mất đi các sắc thái văn hóa. Tính phức tạp và mơ hồ của ngôn ngữ: Ngôn ngữ của con người vốn phức tạp, mơ hồ và phụ thuộc sâu sắc vào ngữ cảnh. Từ ngữ có thể có nhiều nghĩa (đa nghĩa), và cách diễn đạt có thể được diễn giải theo nhiều cách khác nhau, khiến các mô hình NLP khó nắm bắt được ý nghĩa chính xác nếu không có sự hiểu biết sâu sắc. Nghiên cứu của Toral và Way (2018) cho thấy các dịch máy vẫn gặp hạn chế đáng kể trong việc bảo toàn sắc thái nghĩa và các yếu tố ngữ dụng, dẫn đến nguy cơ hiểu sai hoặc truyền đạt sai lệch trong bối cảnh giao tiếp liên văn hóa.

Thách thức thứ hai là nguy cơ phụ thuộc công nghệ. Khi người học lạm dụng các dịch máy hoặc phản hồi tự động, họ có thể mất đi động lực tư duy độc lập và khả năng tự điều chỉnh trong quá trình học. Mặc dù NLP có thể tự động hóa một số nhiệm vụ, nhưng nó không thể sao chép hoàn toàn chiều sâu ngữ cảnh và văn hóa mà các nhà giáo dục con người mang lại. Việc quá phụ thuộc vào AI có thể làm giảm giá trị tương tác giữa người với người (Selvam và cộng sự, 2025). Theo Garcia và Pena (2011), việc sử dụng dịch máy tích hợp NLP trong lớp học ngoại ngữ có thể làm giảm khả năng phát triển chiến lược học tập chủ động, đồng thời ảnh hưởng đến tính sáng tạo trong sản xuất ngôn ngữ.

Một thách thức khác là khoảng cách công nghệ giữa các nhóm người học và cơ sở giáo dục. Không phải tất cả người học đều có cơ hội tiếp cận các công cụ NLP tiên tiến, đặc biệt tại những khu vực có điều kiện kinh tế – xã hội hạn chế. Warschauer và Matuchniak (2010) gọi đây

là "digital divide" trong giáo dục, khi sự khác biệt về hạ tầng công nghệ và khả năng tiếp cận tài nguyên số có thể dẫn đến gia tăng bất bình đẳng trong học tập ngoại ngữ.

Mặc dù NLP mang lại nhiều tiềm năng, giáo viên vẫn đóng vai trò trung tâm trong việc hướng dẫn và định hướng sử dụng công cụ. Giáo viên cần thiết kế hoạt động giảng dạy kết hợp giữa công nghệ và tư duy phản biện, tạo ra một cách tiếp cận cân bằng. Công nghệ nên được xem là công cụ hỗ trợ, không phải sự thay thế hoàn toàn cho quá trình học tập và giảng dạy. Vai trò của giáo viên vẫn giữ vị trí trung tâm trong việc định hướng, khơi gợi tư duy phản biện và tạo môi trường đối thoại đa văn hóa.

Một số chiến lược có thể áp dụng bao gồm:

- Thiết kế hoạt động kết hợp (blended design activities): Giáo viên có thể xây dựng mô hình lai, trong đó công cụ NLP được sử dụng song song với hướng dẫn của con người. Chẳng hạn, sngười học được yêu cầu sử dụng dịch máy để chuyển ngữ một đoạn văn, sau đó thảo luận và so sánh với bản dịch thủ công nhằm phát hiện sự khác biệt về ngữ nghĩa, cấu trúc và sắc thái văn hóa. Quá trình này giúp người học nhận thức rõ giới hạn của công nghệ, từ đó phát triển năng lực phản biện và khả năng đối chiếu ngôn ngữ một cách sâu sắc hơn.

- Tích hợp tư duy phản biện (critical thinking integration): Thay vì tiếp nhận thụ động đầu ra của các công cụ NLP, người học được khuyến khích đánh giá, phân tích và đặt câu hỏi về độ chính xác, tính phù hợp và thiên lệch trong kết quả. Ví dụ, sau khi nhận phản hồi từ Grammarly hoặc ChatGPT, người học cần xem xét liệu gợi ý đó có thực sự phù hợp với mục đích giao tiếp và văn phong mong muốn hay không. Việc này giúp người học hình thành thói quen sử dụng công nghệ một cách có chọn lọc và phản tư, thay vì lệ thuộc hoàn toàn.

- Đa dạng hóa nguồn học liệu (diversifying learning materials): Giáo viên có thể kết hợp các tài nguyên số như chatbot, kho ngữ liệu và nền tảng học trực tuyến với trải nghiệm giao tiếp thực tế như dự án trao đổi ảo, hội thoại xuyên văn hóa hoặc phân tích diễn ngôn từ phương tiện truyền thông quốc tế. Việc kết hợp giữa công nghệ và trải nghiệm xã hội thực tiễn giúp người học không chỉ mở rộng kiến thức ngôn ngữ mà còn phát triển năng lực liên văn hóa thông qua tương tác thực.

- Đào tạo năng lực số (digital literacy training): Một trong những nhiệm vụ quan trọng của giáo viên là giúp sinh viên hiểu rõ nguyên lý hoạt động, ưu điểm và giới hạn của các công cụ NLP. Khi được trang bị năng lực số, người học có thể khai thác công cụ một cách đạo đức, sáng tạo và hiệu quả hơn. Họ học cách xác minh thông tin, nhận biết thiên lệch và sử dụng kết quả do AI tạo ra như một nguồn tham khảo hỗ trợ quá trình học tập, thay vì thay thế cho tư duy của bản thân.

- Xây dựng cộng đồng học tập hợp tác (collaborative learning communities): Tạo không gian nơi học sinh có thể chia sẻ kinh nghiệm sử dụng các công cụ NLP, trao đổi kết quả học tập và thảo luận về vấn đề ngôn ngữ – văn hóa – công nghệ. Những hoạt động cộng tác như vậy giúp

người học học hỏi lẫn nhau, tăng cường khả năng giao tiếp, tư duy phản biện, và phát triển kỹ năng làm việc nhóm trong môi trường học tập đa văn hóa.

Như vậy, thông qua các chiến lược trên, việc tích hợp NLP vào dạy – học ngoại ngữ không chỉ giúp người học nâng cao năng lực ngôn ngữ và khả năng tự học, mà còn thúc đẩy tư duy phản biện, hiểu biết liên văn hóa và khả năng thích ứng toàn cầu. Việc kết hợp hài hòa giữa công nghệ và yếu tố con người chính là chìa khóa để bảo đảm tính bền vững, công bằng và hiệu quả trong ứng dụng NLP vào giáo dục ngoại ngữ hiện đại.

4. Kết luận

Chúng tôi đã phân tích vai trò của xử lý ngôn ngữ tự nhiên (NLP) trong việc dạy và học ngoại ngữ, đồng thời chỉ ra những cơ hội và thách thức khi ứng dụng công nghệ này trong bối cảnh toàn cầu hóa. NLP không chỉ hỗ trợ cải thiện kỹ năng ngôn ngữ của người học thông qua tài liệu học đa dạng, cá nhân hóa chương trình học và phản hồi tức thời, mà còn góp phần thúc đẩy hiểu biết liên văn hóa giúp việc học ngôn ngữ hiệu quả hơn. Tuy nhiên cũng cần khẳng định một điều là NLP không thay thế vai trò của giáo viên, mà đóng vai trò công cụ hỗ trợ nhằm mở rộng trải nghiệm học tập, tạo môi trường thực hành đa dạng và thúc đẩy năng lực tự học. Việc ứng dụng NLP cần được đặt trong mối quan hệ bổ sung với yếu tố con người. Giáo viên và người học vẫn giữ vai trò trung tâm, đảm bảo rằng công nghệ không trở thành nhân tố gây phụ thuộc, mà là công cụ khuyến khích tư duy phản biện, sáng tạo và hợp tác đa văn hóa. Qua đó, NLP trở thành công cụ bổ trợ chứ không thay thế, góp phần mở rộng không gian học tập, đồng thời giúp người học phát triển năng lực ngôn ngữ, văn hóa và tư duy toàn diện.

Tài liệu tham khảo

Busch, D. (2024). AI translation and intercultural communication: New questions for a new field of research. *OSF Preprints*. https://doi.org/10.31235/osf.io/r3zdx

Byram, M. (2021). *Teaching and assessing intercultural communicative competence – Revisited* (2nd ed.). Multilingual Matters.

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education, 10*(3), 241–266. https://doi.org/10.1177/1028315306287002

Eggert, K. (2021). How artificial intelligence will shape universities of tomorrow. In *Proceedings of the 2021 International Conference* (p. 50 https://opus.htwg-konstanz.de

García, I., & Peña, M. I. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning, 24*(5), 471–487. https://doi.org/10.1080/09588221.2011.582687

Gaspari, F., Almaghout, H., & Doherty, S. (2015). A survey of machine translation competences: Insights for translation technology educators and practitioners. *Perspectives, 23*(3), 333–358. https://doi.org/10.1080/0907676X.2014.979842

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Penguin.

Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed.). Pearson.

Li, Z., & Hegelheimer, V. (2013). Mobile-assisted grammar exercises: Effects on self-editing in L2 writing. *Language Learning & Technology, 17*(3), 135–156. http://llt.msu.edu/issues/october2013/lihegelheimer.pdf

Selvam, M., & González Vallejo, R. (2025). Ethical and privacy considerations in AI-driven language learning. *Language, Technology and Artificial Intelligence, 3*, 328. https://doi.org/10.62486/latia2025328

Shaik, T. B., Tao, X., Li, Y., Galligan, L. (2022). A review of the trends and challenges in adopting natural language processing methods for education feedback analysis. *IEEE Access, 10*, 56720-56739. https://doi.org/10.1109/access.2022.3177752

Thorne, S. L. (2016). Cultures-of-use and morphologies of communicative action. *Language Learning & Technology, 20*(2), 185–191. http://llt.msu.edu/issues/june2016/thorne.pdf

Toral, A., & Way, A. (2018). What level of quality can neural machine translation attain on literary text? In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation quality assessment: From principles to practice* (pp. 263–287). Springer.

Warschauer, M., & Matuchniak, T. (2010). New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes. *Review of Research in Education, 34*(1), 179–225. https://doi.org/10.3102/0091732X09349791

Zhai, C., & Wibowo, S. (2022). A systematic review on cross-culture, humor and empathy dimensions in conversational chatbots: The case of second language acquisition. *Heliyon, 8*(12), e12056. https://doi.org/10.1016/j.heliyon.2022.e1205

Zheng, L., Zhang, X., & Yan, Y. (2022). Assessing intercultural competence in the digital world: Evidence from virtual exchange experience in a study abroad program. *Frontiers: The Interdisciplinary Journal of Study Abroad, 34*(4), 418–441. https://doi.org/10.36366/frontiers.v34i4.614

TS. Trần Thị Hải Bình

Giảng viên – Học viện Ngân hàng

Email: binhtth@hvnh.edu.vn

MỨC ĐỘ SẴN SÀNG CỦA GIẢNG VIÊN EFL VIỆT NAM TRONG VIỆC TÍCH HỢP AI VÀO GIẢNG DẠY KỸ NĂNG NGHE

NGUYỄN KHÁNH LĨNH1– ĐINH THÚY LY[2]

Tóm tắt: Trong bối cảnh trí tuệ nhân tạo đang định hình lại giáo dục ngôn ngữ, nghiên cứu thí điểm này khám phá mức độ sẵn sàng của giảng viên EFL Việt Nam trong việc tích hợp các công nghệ NLP vào giảng dạy kỹ năng nghe. Sử dụng phương pháp hỗn hợp, 25 giảng viên đã tham gia khảo sát và phỏng vấn bán cấu trúc.

Kết quả cho thấy một sự tương phản rõ rệt: giảng viên có ý định sử dụng AI trong tương lai ở mức cao, nhưng lại cảm nhận điều kiện hỗ trợ từ tổ chức còn thiếu đồng bộ. Về ứng dụng, họ ưa chuộng các công cụ như TTS và LLMs, trong khi còn do dự với ASR do lo ngại về độ chính xác khi nhận diện tiếng Việt. Đáng chú ý, ý định này không bị ảnh hưởng bởi các yếu tố "hữu ích" hay "dễ dùng", mà lại liên quan chặt chẽ đến năng lực sư phạm và nhận thức đạo đức. Nghiên cứu nhấn mạnh nhu cầu xây dựng chính sách, hạ tầng và đào tạo chuyên sâu để hỗ trợ tiềm năng sẵn có của giảng viên.

Từ khóa: Trí tuệ nhân tạo, Kỹ năng nghe, Dạy học tiếng Anh, UTAUT, Xử lý ngôn ngữ tự nhiên.

1. Giới thiệu

Trong những năm gần đây, sự phát triển của trí tuệ nhân tạo (AI), đặc biệt là AI tạo sinh, đang tạo ra những thay đổi sâu sắc trong lĩnh vực giáo dục ngôn ngữ (Mananay, 2024; Tutton & Cohen, 2025). Các công cụ xử lý ngôn ngữ tự nhiên (NLP) như nhận dạng giọng nói tự động (ASR), chuyển văn bản thành giọng nói (TTS) và các mô hình ngôn ngữ lớn (LLMs) đã mở ra những tiềm năng đột phá cho phương pháp giảng dạy, cho phép cá nhân hóa trải nghiệm học tập và nâng cao hiệu quả sư phạm (Huang và c.s., 2024)

Trong bối cảnh giảng dạy tiếng Anh như một ngoại ngữ (EFL), đặc biệt tại các cơ sở giáo dục đại học ở Việt Nam, kỹ năng nghe từ lâu đã được xem là một trong những thách thức lớn nhất đối với cả người dạy và người học, một nhận định tiếp tục được khẳng định trong các nghiên cứu gần đây về tích hợp công nghệ (Wang, 2024; Gusman và c.s., 2024). Sự thiếu hụt môi trường tiếp xúc tự nhiên khiến quá trình phát triển kỹ năng này gặp nhiều rào cản, trong khi các phương pháp truyền thống thường bị giới hạn về sự đa dạng của học liệu và thiếu khả năng cung cấp phản hồi tức thì. Trước những thách thức cố hữu này, AI được kỳ vọng sẽ là một giải pháp tiềm năng, mang lại khả năng cá nhân hóa hướng dẫn và cải thiện thành tích học tập cho người học (Wei, 2023; Sivaji và c.s., 2024). Tuy nhiên, các nghiên cứu đều khẳng định rằng việc tích hợp thành công

[1] ThS, đang giảng dạy tại Cao đẳng Việt Mỹ, *email: khanhlinhedu112@gmail.com*

[2] Học viên cao học, đang giảng dạy tại Cao đẳng FPT Polytechnic HCM, *email: dinhthuyly@gmail.com*

công nghệ phụ thuộc phần lớn vào mức độ sẵn sàng (readiness) và năng lực chủ động (agency) của đội ngũ giảng viên – những người trực tiếp thiết kế và triển khai các hoạt động sư phạm (Ly, 2025; Frosig & Romero, 2024).

Tại Việt Nam, dù đã có những nghiên cứu bước đầu về ứng dụng AI trong lĩnh vực giảng dạy tiếng Anh cho người nói ngôn ngữ khác (TESOL), tổng quan tài liệu cho thấy phần lớn các công trình hiện tại vẫn tập trung chủ yếu vào kỹ năng viết và đọc (Pham & Cao, 2025). Đáng chú ý, một khoảng trống nghiên cứu quan trọng vẫn còn tồn tại trong việc tìm hiểu một cách hệ thống về mức độ sẵn sàng và năng lực chủ động của giảng viên đại học khi tích hợp AI vào việc thiết kế học liệu và giảng dạy kỹ năng nghe. Nhiều nghiên cứu trong và ngoài nước cũng chỉ ra khoảng cách giữa sự sẵn sàng và thực tiễn. Cụ thể, các giảng viên dù tự tin về năng lực cá nhân nhưng vẫn ngần ngại hoặc chỉ áp dụng AI ở các khâu chuẩn bị bên ngoài lớp học (Cheah và c.s., 2025), thay vì tích hợp trực tiếp vào hoạt động giảng dạy. Nguyên nhân của sự chần chừ này được cho là do thiếu sự hỗ trợ đồng bộ và nhất quán từ hệ thống (Ly, 2025; Mutanga và c.s., 2024; Phuong, 2024).

Nhận thấy khoảng trống trên, nghiên cứu này được định vị là một nghiên cứu mang tính khám phá. Với quy mô mẫu nhỏ, nghiên cứu không đặt mục tiêu đưa ra các kết luận mang tính khái quát hóa, mà hướng đến các mục tiêu khiêm tốn và nền tảng hơn:

(1). Thực hiện một cuộc khảo sát sơ bộ về mức độ sẵn sàng của một nhóm giảng viên EFL Việt Nam trong việc ứng dụng AI vào giảng dạy nghe, dựa trên các khung lý thuyết về chấp nhận công nghệ như UTAUT và TPACK.

(2). Khám phá những trải nghiệm thực tế ban đầu của giảng viên với các công cụ NLP cụ thể như ASR, TTS và LLMs.

(3). Xác định các thách thức và nhu cầu hỗ trợ nổi bật, từ đó hình thành giả thuyết ban đầu và đề xuất hướng đi cho những nghiên cứu quy mô lớn hơn trong tương lai.

2. Cơ sở lý thuyết và tổng quan nghiên cứu

2.1. Các công nghệ AI và ứng dụng trong giảng dạy kỹ năng nghe

Sự phát triển của trí tuệ nhân tạo, đặc biệt là trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP), đã mở rộng đáng kể các phương pháp hỗ trợ giảng dạy ngoại ngữ. Đối với kỹ năng nghe, vốn được xem là một kỹ năng phức tạp và khó lĩnh hội (Field, 2009; Goh & Vandergrift, 2012), công nghệ AI mang lại những giải pháp hứa hẹn để khắc phục các hạn chế của phương pháp truyền thống. Ba nhóm công nghệ nổi bật nhất bao gồm:

- Nhận dạng giọng nói tự động (ASR – Automatic Speech Recognition): Công nghệ ASR có khả năng chuyển đổi ngôn ngữ nói thành văn bản, cung cấp một công cụ hiệu quả để người học tự luyện tập và nhận phản hồi tức thì. Các nghiên cứu đã chỉ ra rằng việc sử dụng ASR trong các hoạt động như nghe-chính tả (dictation) giúp sinh viên EFL nhận ra

sự khác biệt giữa phát âm của bản thân và người bản ngữ, từ đó nâng cao ý thức ngữ âm và cải thiện năng lực nghe hiểu (Farrús, 2023; McCrocklin, 2024). Tuy nhiên, việc triển khai các công cụ AI trong dạy tiếng Anh cũng đối mặt với không ít thách thức (Crompton và c.s., 2024). Một ví dụ điển hình là với công nghệ ASR, khi độ chính xác của công cụ này thường giảm đáng kể lúc xử lý các giọng nói không phải bản ngữ (non-native accents) (Farrús, 2023), điều này đặc biệt cần lưu ý trong bối cảnh đa dạng vùng miền ở Việt Nam.

- Chuyển văn bản thành giọng nói (TTS – Text-to-Speech): Công nghệ TTS cho phép tạo ra các tệp âm thanh từ bất kỳ văn bản nào. Ưu điểm vượt trội của TTS là khả năng tùy chỉnh, giúp giảng viên dễ dàng tạo ra một kho học liệu nghe phong phú với nhiều loại giọng (accent), tốc độ và ngữ điệu khác nhau, đáp ứng nhu cầu luyện tập đa dạng của người học. Mặc dù các công nghệ TTS ngày càng được cải tiến, một số nghiên cứu vẫn chỉ ra rằng giọng nói do AI tạo ra đôi khi còn thiếu tính tự nhiên và sắc thái biểu cảm so với giọng người thật, điều này có thể ảnh hưởng đến trải nghiệm nghe của người học (Kasneci và c.s., 2023).
- Mô hình ngôn ngữ lớn (LLMs – Large Language Models): Sự xuất hiện của các LLMs, mà tiêu biểu là ChatGPT, đã tạo ra một cuộc cách mạng trong việc thiết kế và phát triển học liệu. Đối với kỹ năng nghe, giảng viên có thể tận dụng LLMs để thực hiện nhiều tác vụ phức tạp như: tạo nhanh bản ghi (transcript) cho các video hoặc tệp âm thanh, thiết kế hệ thống câu hỏi nghe hiểu theo nhiều cấp độ tư duy, tóm tắt nội dung bài nghe, hoặc biên soạn các bài hội thoại tình huống để sinh viên luyện tập (Ly, 2025; Ma và c.s., 2024). Tuy vậy, việc sử dụng LLMs cũng đi kèm với những rủi ro nhất định. Các nghiên cứu đã cảnh báo về khả năng LLMs tạo ra thông tin sai lệch, nội dung hời hợt, hoặc các diễn đạt thiếu tự nhiên (Chaka, 2023; Kasneci và c.s., 2023). Điều này đòi hỏi giảng viên phải có năng lực chuyên môn vững vàng để kiểm định, chọn lọc và biên tập lại kết quả do AI tạo ra trước khi sử dụng trong giảng dạy.

2.2. Các khung lý thuyết nền tảng

Để phân tích một cách toàn diện mức độ sẵn sàng của giảng viên trong việc tích hợp AI, nghiên cứu này vận dụng hai khung lý thuyết bổ trợ cho nhau: Mô hình Hợp nhất về Chấp nhận và Sử dụng Công nghệ (UTAUT) để lý giải các yếu tố ảnh hưởng đến ý định hành vi và Khung năng lực Tích hợp Công nghệ - Sư phạm - Nội dung (TPACK) để phân tích các khía cạnh tri thức và năng lực chuyên môn của người dạy.

2.2.1. Mô hình hợp nhất về chấp nhận và sử dụng công nghệ (UTAUT)

Mô hình UTAUT (Unified Theory of Acceptance and Use of Technology), do Venkatesh và c.s. (2003) phát triển, là một trong những khung lý thuyết có ảnh hưởng nhất nhằm lý giải hành vi chấp nhận và sử dụng công nghệ của người dùng. Mô hình này được tổng hợp từ tám lý thuyết

trước đó và cho rằng ý định và hành vi sử dụng công nghệ chịu ảnh hưởng chính bởi bốn yếu tố cốt lõi:

(1). Hiệu quả kỳ vọng (Performance Expectancy – PE): Niềm tin của cá nhân rằng việc sử dụng công nghệ sẽ giúp họ nâng cao hiệu suất và hiệu quả trong công việc. Trong bối cảnh giáo dục, đây là niềm tin rằng AI sẽ giúp việc giảng dạy trở nên tốt hơn.

(2). Nỗ lực kỳ vọng (Effort Expectancy – EE): Mức độ dễ dàng mà cá nhân cảm nhận khi học và sử dụng một công nghệ mới. Yếu tố này tương đương với cảm nhận về sự phức tạp hay đơn giản của công cụ AI.

(3). Ảnh hưởng xã hội (Social Influence – SI): Mức độ mà cá nhân tin rằng những người quan trọng xung quanh (như đồng nghiệp, cấp trên, hoặc sinh viên) nghĩ rằng họ nên sử dụng công nghệ đó.

(4). Điều kiện hỗ trợ (Facilitating Conditions – FC): Mức độ mà cá nhân tin rằng hạ tầng kỹ thuật và sự hỗ trợ từ tổ chức (ví dụ: các khóa tập huấn, chính sách khuyến khích) có sẵn để tạo điều kiện thuận lợi cho việc sử dụng công nghệ.

Nhiều nghiên cứu trong lĩnh vực giáo dục đã khẳng định rằng hiệu quả kỳ vọng (PE) và nỗ lực kỳ vọng (EE) thường là những yếu tố có tác động dự báo mạnh mẽ nhất đối với ý định sử dụng công nghệ của giáo viên. Tuy nhiên, trong bối cảnh AI ngày càng phức tạp, vai trò của các điều kiện hỗ trợ và ảnh hưởng xã hội cũng trở nên quan trọng hơn bao giờ hết.

2.2.2. Khung năng lực tích hợp công nghệ - sư phạm - nội dung (TPACK và AI-TPACK)

Nếu UTAUT tập trung vào các yếu tố thúc đẩy ý định, thì khung TPACK (Technological Pedagogical Content Knowledge) của Mishra & Koehler (2006) lại tập trung vào các lĩnh vực tri thức mà giáo viên cần có để tích hợp công nghệ một cách hiệu quả và có ý nghĩa. TPACK cho rằng năng lực giảng dạy trong kỷ nguyên số không nằm ở từng kiến thức riêng lẻ về công nghệ (TK), sư phạm (PK), hay nội dung chuyên môn (CK), mà nằm ở sự giao thoa phức hợp và năng động giữa chúng.

Trước sự phát triển mạnh mẽ của AI, các nhà nghiên cứu cho rằng cần mở rộng mô hình này thành AI-TPACK để phản ánh những yêu cầu năng lực mới (Li & Liang, 2025). Khung AI-TPACK không chỉ bao gồm các năng lực công nghệ thông thường, mà còn nhấn mạnh các kỹ năng chuyên biệt như: (1) năng lực hiểu biết về AI (AI literacy), bao gồm cả việc nhận diện các giới hạn và thiên kiến (bias) tiềm ẩn trong thuật toán; (2) khả năng tích hợp AI vào chiến lược sư phạm một cách sáng tạo; và (3) năng lực đánh giá và thẩm định chất lượng các nội dung do AI tạo ra dưới góc độ đạo đức và chuyên môn (Celik, 2023; Al-Abdullatif, 2024). Một số nghiên cứu tại Việt Nam cũng cho thấy AI-TPACK là một yếu tố dự báo quan trọng đối với ý định ứng dụng AI của giảng viên ngôn ngữ (Tram, 2025). Việc kết hợp cả hai khung lý thuyết UTAUT và AI-TPACK sẽ mang lại một góc nhìn toàn diện hơn về mức độ sẵn sàng của giảng viên.

2.3. Tổng quan các nghiên cứu về mức độ sẵn sàng và năng lực chủ động của giảng viên

Các nghiên cứu thực nghiệm về mức độ sẵn sàng của giảng viên trong việc sử dụng AI trên thế giới và tại Việt Nam đã vẽ nên một bức tranh đa chiều, cho thấy sự song hành giữa thái độ tích cực và những rào cản đáng kể trong thực tiễn.

Nhìn chung, các khảo sát ở nhiều bối cảnh khác nhau đều ghi nhận rằng giảng viên thường thể hiện sự lạc quan và nhận thức rõ về những tiềm năng mà AI có thể mang lại cho công tác giảng dạy, chẳng hạn như cá nhân hóa lộ trình học tập, tự động hóa các tác vụ lặp lại và nâng cao mức độ tương tác của người học (Eke, 2024; Kohnke và c.s., 2023; Mutanga và c.s., 2024). Tại Việt Nam, các nghiên cứu cũng phản ánh một xu hướng tương tự, khi phần lớn giảng viên và sinh viên đều có nhận định tích cực về lợi ích của các công cụ AI như ChatGPT (Hua & Le, 2024; V. P. H. Pham & Le, 2024).

Tuy nhiên, thái độ tích cực này không phải lúc nào cũng đồng nghĩa với sự sẵn sàng áp dụng trong thực tiễn. Một phát hiện có tính phổ quát cao trong các tài liệu là khoảng cách giữa sự sẵn sàng và thực tiễn (Ly, 2025). Cụ thể, nhiều giảng viên dù tự tin về năng lực cá nhân nhưng chủ yếu chỉ sử dụng AI cho các công việc bên ngoài lớp học như chuẩn bị giáo án, thiết kế bài tập, đánh giá kết quả học tập và đưa ra phản hồi, thay vì tích hợp trực tiếp vào các hoạt động tương tác trong giờ học (Cheah và c.s., 2025). Sự ngần ngại này thường xuất phát từ nhiều rào cản mang tính hệ thống, bao gồm:

- Thiếu hụt các chính sách và hướng dẫn cụ thể: Nhiều giảng viên cảm thấy lúng túng do thiếu các quy định và định hướng rõ ràng từ cấp quản lý về việc sử dụng AI một cách hợp lệ và hiệu quả (Ly, 2025; Salha và c.s., 2025).
- Hạn chế về đào tạo và hạ tầng: Mặc dù đã có các buổi tập huấn, nhiều giảng viên vẫn cảm thấy chưa được trang bị đủ kỹ năng chuyên sâu và thiếu các công cụ AI có bản quyền hoặc hạ tầng công nghệ đủ mạnh để triển khai trên diện rộng (Tran và c.s., 2024).
- Các lo ngại về đạo đức và sư phạm: Các vấn đề về liêm chính học thuật, nguy cơ sinh viên trở nên phụ thuộc vào công nghệ và chất lượng không đồng đều của nội dung do AI tạo ra cũng là những rào cản tâm lý lớn (Cong-Lem và c.s., 2024; Zhai, 2024).

Bên cạnh mức độ sẵn sàng, khái niệm năng lực chủ động của giáo viên ngày càng được xem là yếu tố then chốt. Năng lực chủ động được hiểu là khả năng của giáo viên trong việc đưa ra các quyết định sư phạm một cách độc lập, có mục đích và làm chủ công nghệ để phục vụ cho tầm nhìn giáo dục của mình, thay vì bị động tuân theo các tính năng sẵn có của công cụ (Frosig & Romero, 2024; Zhai, 2024). Các nghiên cứu cho thấy khi giảng viên có năng lực chủ động cao, họ có xu hướng sử dụng AI một cách sáng tạo và mang tính phản biện hơn. Chẳng hạn, thay vì chỉ dùng AI để tạo bài tập, họ có thể yêu cầu sinh viên phân tích, đánh giá và phản biện các nội dung do AI tạo ra nhằm phát triển tư duy bậc cao (Ly, 2025). Việc thúc đẩy năng lực chủ động của giáo

viên đòi hỏi một môi trường hỗ trợ, khuyến khích sự thử nghiệm và chia sẻ kinh nghiệm, điều mà nhiều cơ sở giáo dục vẫn còn thiếu (Tran và c.s., 2024).

Tóm lại, tổng quan các nghiên cứu cho thấy để thúc đẩy việc tích hợp AI vào giảng dạy kỹ năng nghe một cách hiệu quả, việc trang bị công cụ là chưa đủ. Thay vào đó, cần có một chiến lược tổng thể bao gồm: (1) xây dựng các chính sách và hướng dẫn rõ ràng; (2) tổ chức các chương trình đào tạo chuyên sâu, tập trung vào cả kỹ năng kỹ thuật và năng lực sư phạm-đạo đức (AI-TPACK); và (3) kiến tạo một môi trường khuyến khích năng lực chủ động và sự sáng tạo của giảng viên.

3. Phương pháp nghiên cứu

3.1. Thiết kế nghiên cứu

Nghiên cứu này áp dụng phương pháp nghiên cứu hỗn hợp mang tính khám phá. Cách tiếp cận này cho phép kết hợp dữ liệu định lượng từ bảng hỏi để phác thảo các xu hướng chung và dữ liệu định tính từ phỏng vấn bán cấu trúc nhằm làm sâu sắc thêm sự hiểu biết về trải nghiệm thực tế của giảng viên. Theo Creswell và Clark (2017), phương pháp hỗn hợp giúp tận dụng ưu điểm của cả hai nguồn dữ liệu, từ đó tăng cường giá trị giải thích và nâng cao độ tin cậy của phát hiện nghiên cứu. Trong khuôn khổ một nghiên cứu thí điểm, thiết kế này đặc biệt hữu ích vì vừa cung cấp cái nhìn khái quát từ số liệu định lượng, vừa bổ sung những lý giải sâu hơn từ dữ liệu định tính (Creswell & Clark, 2017). Nhờ vậy, nghiên cứu có thể tạo nền tảng cho việc hình thành và kiểm chứng các giả thuyết trong những nghiên cứu tiếp theo.

3.2. Người tham gia

Nghiên cứu đã tiến hành khảo sát 25 giảng viên EFL đang công tác tại các trường đại học và cao đẳng ở Việt Nam. Mẫu được lựa chọn theo phương pháp chọn mẫu thuận tiện, dựa trên sự sẵn lòng tham gia của các giảng viên trong mạng lưới của nhóm nghiên cứu. Đặc điểm nhân khẩu học của mẫu cho thấy phần lớn là nữ (72%), thuộc độ tuổi từ 25 đến 34 tuổi (80%) và có kinh nghiệm giảng dạy từ 5 đến 10 năm.

3.3. Bảng hỏi định lượng

Bảng hỏi định lượng được thiết kế dựa trên sự kế thừa và phát triển từ các mô hình lý thuyết quốc tế, bao gồm UTAUT (Venkatesh và c.s., 2003), TPACK (Mishra & Koehler, 2006), cùng với những mở rộng gần đây trong bối cảnh AI. Các mục hỏi được đo lường bằng thang Likert 5 điểm (1 = Hoàn toàn không đồng ý → 5 = Hoàn toàn đồng ý).

Bảng 1. Các thang đo trong bảng hỏi định lượng

Thang đo (Construct)	Số mục	Nguồn gốc & căn cứ lý thuyết	Nội dung đo lường chính

Hiệu quả kỳ vọng (PE)	4	UTAUT – Venkatesh và cộng sự (2003)	Niềm tin rằng AI giúp nâng cao hiệu quả giảng dạy nghe
Nỗ lực kỳ vọng (EE)	4	UTAUT – Venkatesh và cộng sự (2003)	Cảm nhận về sự dễ sử dụng và mức độ thành thạo công nghệ
Ảnh hưởng xã hội (SI)	4	UTAUT – Venkatesh và cộng sự (2003)	Tác động từ đồng nghiệp, sinh viên và quản lý
Điều kiện hỗ trợ (FC)	4	UTAUT – Venkatesh và cộng sự (2003)	Mức độ hỗ trợ về hạ tầng và chính sách từ nhà trường
Chiến lược sư phạm (PS)	5	TPACK – Mishra & Koehler (2006)	Khả năng tích hợp AI vào chiến lược sư phạm
Nhận thức đạo đức (EA)	5	Holmes và cộng sự (2022); Kasneci và cộng sự (2023)	Nhận thức về quyền riêng tư, minh bạch, trách nhiệm và vai trò hỗ trợ của AI
Ứng dụng NLP thực tế (NLP)	4	Farrús (2023); McCrocklin (2019); Zhang và cộng sự (2023); Ma và cộng sự (2024)	Mức độ áp dụng ASR, TTS và LLM trong giảng dạy nghe
Ý định sử dụng (BI)	3	UTAUT – Venkatesh và cộng sự (2003)	Ý định tiếp tục sử dụng AI trong giảng dạy

Các thang đo được điều chỉnh ngữ cảnh cho phù hợp với bối cảnh giảng dạy kỹ năng nghe tại Việt Nam. Để đảm bảo giá trị nội dung, bảng hỏi đã được hai chuyên gia TESOL phản biện và tiến hành thử nghiệm pilot trên 5 giảng viên trước khi phát hành chính thức.

3.4. Phỏng vấn bán cấu trúc

Để bổ sung cho dữ liệu khảo sát, nghiên cứu tiến hành phỏng vấn bán cấu trúc với 4 giảng viên được chọn từ mẫu khảo sát. Hướng dẫn phỏng vấn tập trung vào ba chủ đề lớn, nhằm khai thác trải nghiệm thực tiễn, quan điểm và nhu cầu hỗ trợ của giảng viên trong quá trình ứng dụng AI vào giảng dạy kỹ năng nghe.

Bảng 2. Chủ đề chính trong hướng dẫn phỏng vấn

Chủ đề phỏng vấn	Nội dung trọng tâm	Công cụ AI liên quan
1. Trải nghiệm sử dụng AI	Cách giảng viên khai thác ASR, TTS, LLM trong dạy nghe; ví dụ minh họa tình huống	ASR, TTS, LLMs
2. Thuận lợi và khó khăn	Lợi ích và rào cản khi áp dụng AI; ảnh hưởng đến chiến lược và hoạt động sư phạm	ASR, TTS, LLMs
3. Điều kiện hỗ trợ và nhu cầu đào tạo	Chính sách, hạ tầng, sự hỗ trợ từ nhà trường; mong muốn tập huấn, đào tạo bổ sung	Hỗ trợ kỹ thuật, đào tạo

Mỗi cuộc phỏng vấn kéo dài 45–60 phút, được tiến hành trực tuyến và ghi âm có sự đồng thuận của người tham gia và sau đó được phiên mã đầy đủ. Dữ liệu phỏng vấn được phân tích bằng phương pháp phân tích chủ đề để xác định các mẫu lặp lại, chủ đề nổi bật và đối chiếu với phát hiện từ khảo sát định lượng.

3.5. Quy trình thu thập dữ liệu

Quy trình thu thập dữ liệu được triển khai theo hai giai đoạn.

- **Khảo sát định lượng:** Bảng hỏi được phát hành trực tuyến thông qua Google Forms và duy trì trong vòng **6 tuần**. Phương thức trực tuyến được lựa chọn nhằm thuận tiện cho người tham gia và tối ưu hóa khả năng tiếp cận.
- **Phỏng vấn định tính:** Sau khi hoàn tất khảo sát, **4 giảng viên tình nguyện** được mời tham gia phỏng vấn trực tuyến qua nền tảng Zoom. Mỗi cuộc phỏng vấn kéo dài từ **45–60 phút**, có sự đồng ý ghi âm để phục vụ cho việc phiên mã và phân tích sau đó.

Về đạo đức nghiên cứu, tất cả giảng viên được cung cấp **thông tin nghiên cứu và bảng đồng thuận tham gia (informed consent),** bao gồm mục tiêu, quy trình, quyền rút lui bất cứ lúc nào và cam kết ẩn danh dữ liệu cá nhân. Việc tuân thủ nguyên tắc đạo đức này nhằm bảo đảm tính minh bạch và quyền lợi của người tham gia.

3.6. Phân tích dữ liệu

Dữ liệu định lượng từ bảng hỏi được phân tích bằng phần mềm SPSS. Độ tin cậy của các thang đo được kiểm tra bằng hệ số Cronbach’s Alpha. Các phương pháp phân tích chính bao gồm: (1) thống kê mô tả (tần suất, giá trị trung bình, độ lệch chuẩn) nhằm mô tả xu hướng của mẫu; và (2) phân tích tương quan Pearson để kiểm tra mối quan hệ giữa các biến.

Dữ liệu định tính từ các bản ghi phỏng vấn được phiên mã và phân tích bằng phương pháp phân tích theo hướng tiếp cận chủ đề của Braun và Clarke (2006). Kết quả phân tích được dùng để đối chiếu và làm rõ thêm các phát hiện từ dữ liệu định lượng.

4. Kết quả và thảo luận

4.1. Thông tin mô tả mẫu nghiên cứu

Mẫu khảo sát bao gồm **25 giảng viên tiếng Anh** tại các trường đại học và cao đẳng ở Việt Nam. Trong số này, **72% là nữ** và **28% là nam**. Về độ tuổi, phần lớn người tham gia thuộc nhóm **25 đến 34 tuổi (80%)**, trong khi số còn lại nằm ở nhóm 35 đến 44 tuổi và trên 45 tuổi. Xét về thâm niên nghề nghiệp, hơn một nửa có **5 đến 10 năm kinh nghiệm giảng dạy**. Đáng chú ý, **64% các giảng viên cho biết đã sử dụng AI thường xuyên trong giảng dạy**, trong khi số còn lại chỉ sử dụng thỉnh thoảng hoặc hiếm khi.

Bảng 3. Thông tin nhân khẩu học của người tham gia (N = 25)

Biến	Phân loại	Tỉ lệ (%)
Giới tính	Nam	28
	Nữ	72
Độ tuổi	Dưới 25	0
	25–34	80
	35–44	16
	Từ 45 trở lên	4
Thâm niên giảng dạy	Dưới 5 năm	20
	5–10 năm	52
	10–15 năm	20
	Trên 15 năm	8
Mức độ sử dụng AI	Thường xuyên	64
	Thỉnh thoảng	28
	Hiếm hoặc không	8

Kết quả mô tả cho thấy mẫu nghiên cứu chủ yếu bao gồm **giảng viên trẻ**, có mức độ quen thuộc cao với công nghệ, phù hợp với xu hướng **"digital native teachers"** được đề cập trong các nghiên cứu gần đây (An và c.s., 2023). Đây có thể là một trong những nguyên nhân lý giải vì sao mức độ sẵn sàng ứng dụng AI đạt mức khá cao.

4.2. Độ tin cậy các thang đo

Kết quả kiểm định độ tin cậy cho thấy tất cả tám thang đo trong bảng hỏi đều đạt **Cronbach's alpha > 0.70**, đáp ứng tiêu chuẩn chấp nhận được trong nghiên cứu khoa học xã hội (Nunnally & Bernstein, 1994). Bốn thang đo kế thừa từ UTAUT (PE, EE, SI, FC) có hệ số alpha dao động từ **0.708 đến 0.82**, phản ánh mức độ nhất quán nội tại ổn định.

Bảng 4. Độ tin cậy Cronbach's alpha của các thang đo **(N=25)**

Thang đo	Số mục	Cronbach's α
Hiệu quả kỳ vọng (PE)	4	.708
Nỗ lực kỳ vọng (EE)	4	.820
Ảnh hưởng xã hội (SI)	4	.912
Điều kiện hỗ trợ (FC)	4	.868
Chiến lược sư phạm (PS)	5	.921
Nhận thức đạo đức (EA)	5	.740
Ứng dụng NLP thực tế (NLP)	4	.798
Ý định sử dụng(BI)	3	.932

Như vậy, bộ công cụ đo lường được sử dụng trong nghiên cứu có **độ tin cậy cao**, tạo cơ sở vững chắc cho các phân tích thống kê tiếp theo.

4.3. Kết quả thống kê mô tả

Bảng 5. Thống kê mô tả (N = 25)

Thang đo (Construct)	Mean (TB)	SD	Điểm nổi bật
Hiệu quả kỳ vọng (PE)	3.7	0.9	Tin AI hữu ích trong thiết kế học liệu
Nỗ lực kỳ vọng (EE)	3.9	0.9	Sử dụng AI tương đối dễ dàng
Ảnh hưởng xã hội (SI)	3.3	1.2	Ảnh hưởng từ đồng nghiệp/sinh viên/nhà trường chưa mạnh
Điều kiện hỗ trợ (FC)	3.4	1.4	Hỗ trợ hạ tầng và chính sách còn thiếu đồng đều

Chiến lược sư phạm (PS)	4.0	1.0	Giảng viên luôn chỉnh sửa & đánh giá đầu ra AI
Nhận thức đạo đức (EA)	3.7	1.0	Nhất trí cao rằng AI chỉ nên hỗ trợ, không thay thế (TB > 4.4)
Ứng dụng NLP thực tế (NLP)	3.4	1.2	TTS & LLM được dùng nhiều; ASR hạn chế (M = 2.56)
Ý định sử dụng (BI)	4.1	1.0	Ý định sử dụng AI trong tương lai ở mức rất cao

Phân tích dữ liệu mô tả tại Bảng 5 cho thấy một bức tranh đa chiều về hiện trạng sẵn sàng, được đặc trưng bởi sự tương phản rõ rệt giữa ba cấp độ: sự chủ động của cá nhân, đặc tính của từng công cụ và các điều kiện ở cấp độ hệ thống.

Ở cấp độ cá nhân, dữ liệu cho thấy một mức độ sẵn sàng cao. Giảng viên không chỉ có ý định tiếp tục sử dụng AI ở mức cao (BI: M = 4.1/5), họ còn thể hiện sự tự tin vào năng lực sư phạm của mình trong việc thẩm định và tích hợp AI vào bài giảng. (PS: M = 4.0/5). Phân tích định tính củng cố xu hướng tự chủ trong quyết định sư phạm, thể hiện qua các phát biểu như: "*Tôi thường chỉnh sửa TTS nhiều lần cho phù hợp tốc độ và ngữ điệu*" và "*AI chỉ nên là công cụ hỗ trợ, không thay thế giáo viên*". Kết quả này phù hợp với các nghiên cứu gần đây, vốn cho rằng khi giảng viên đã quen thuộc với công nghệ, ý định ứng dụng có mối liên hệ chặt chẽ hơn với năng lực sư phạm, thay vì các yếu tố "hữu ích/dễ dùng" đơn thuần (An và c.s., 2023; Tram, 2025).

Ở cấp độ công cụ**,** sự sẵn sàng này được thể hiện một cách không đồng đều. Giảng viên tỏ ra cởi mở với TTS và LLMs do tính dễ kiểm soát và tùy chỉnh, nhưng lại rất dè dặt với ASR (M = 2.56). Dữ liệu định tính cho thấy sự do dự này vượt ra ngoài lý do kỹ thuật thuần túy để chạm đến một mối lo ngại sâu sắc về rủi ro sư phạm và tâm lý. Một giảng viên đã bày tỏ rõ ràng mối lo này: "*Dùng ASR để chấm phát âm em thấy rất rủi ro... Học trò thấy điểm thấp liên tục, sẽ nghĩ là mình phát âm quá tệ, rồi nản và sợ nói tiếng Anh thì sao? Điều đó còn nguy hiểm hơn là phát âm sai một vài chỗ.*" Trích dẫn này cho thấy quyết định chấp nhận công nghệ của giảng viên không chỉ dựa trên tính năng, mà còn được cân nhắc kỹ lưỡng dựa trên tác động tâm lý tiềm tàng đến người học.

Tuy nhiên, sự chủ động ở cấp độ cá nhân và công cụ lại tương phản với các điều kiện ở cấp độ hệ thống. Dữ liệu cho thấy sự thiếu hụt rõ rệt trong các điều kiện hỗ trợ từ tổ chức, với điểm trung bình thấp cho cả Điều kiện hỗ trợ (FC: M = 3.4/5) và Ảnh hưởng xã hội (SI: M = 3.3/5). Phân tích định tính đã làm rõ hơn về tình trạng "tự phát, đơn lẻ" trong triển khai, được thể hiện thông qua chia sẻ của giảng viên: "*Trường chưa có chủ trương rõ; tôi phải tự mày mò là chính*".

Kết quả này không chỉ cộng hưởng với bối cảnh tại Việt Nam (Loan & Thủy, 2024) mà còn phù hợp với bối cảnh quốc tế, nơi sự thiếu vắng các hướng dẫn chính thức được xem là thách thức phổ biến trong việc ứng dụng AI ở bậc đại học (Kasneci và c.s., 2023).

Tóm lại, bức tranh sẵn sàng không đơn thuần là "có công nghệ thì sẽ dùng", mà là sự tương tác phức tạp giữa năng lực sư phạm và niềm tin đạo đức ở cấp cá nhân, tính phù hợp của công cụ ở cấp kỹ thuật và khung chính sách-hạ tầng ở cấp hệ thống.

4.4. Phân tích tương quan

Bảng 6. Ma trận tương quan giữa các thang đo (N=25)

Biến	PE	EE	SI	FC	PS	EA	NLP	BI
PE	1							
EE	.42	1						
SI	.18	.25	1					
FC	.21	.31	.28	1				
PS	.33	.40	.36	.29	1			
EA	.27	.34	.39	.25	.48	1		
NLP	.31	.37	.41	.28	.55	.51	1	
BI	.05	.07	.50*	.12	.79**	.78**	.78**	1

Ghi chú: * $p < .05$, ** $p < .001$.

Phân tích tương quan Pearson (Bảng 6) cho thấy mối quan hệ đa chiều giữa các yếu tố trong mô hình UTAUT–TPACK và ý định sử dụng AI (BI) của giảng viên. Kết quả làm rõ rằng chiến lược sư phạm (PS), nhận thức đạo đức (EA) và ứng dụng NLP thực tế (NLP) là ba biến có mối tương quan mạnh nhất và có ý nghĩa thống kê cao với BI (lần lượt $r = .79$, .78 và .78; $p < .001$). Điều này có nghĩa là khi giảng viên thể hiện năng lực điều chỉnh sư phạm cao hơn, ý thức rõ ràng về giới hạn đạo đức của AI và có trải nghiệm thực tế với các công cụ NLP như TTS hoặc LLMs, họ cũng có xu hướng tăng đáng kể ý định ứng dụng AI vào giảng dạy. Mức tương quan này ở ngưỡng rất mạnh ($r > .70$), cho thấy các yếu tố "con người làm trung tâm" là động lực trung tâm trong quá trình sẵn sàng tích hợp AI.

Ngược lại, hiệu quả kỳ vọng (PE) và nỗ lực kỳ vọng (EE) lại không có tương quan ý nghĩa thống kê với BI ($r = .05$ và $r = .07$, $p > .05$). Phát hiện này khác biệt rõ với các mô hình UTAUT cổ điển (Venkatesh et al., 2003; Teo, 2019), nơi PE và EE thường được xem là yếu tố dự báo mạnh mẽ nhất cho hành vi công nghệ. Một giả thuyết hợp lý là nhóm giảng viên trong mẫu này đã đạt đến mức độ sử dụng AI tương đối thường xuyên, khiến các yếu tố "cảm nhận hữu ích" và "dễ

dùng" trở nên bão hòa về ảnh hưởng. Tuy nhiên, đây chỉ là giả thuyết cần kiểm chứng, vì nghiên cứu chưa đo lường trực tiếp mức độ thành thạo công cụ AI của người tham gia. Các khả năng khác cũng có thể được xem xét, thang đo như PE/EE chưa hoàn toàn phù hợp với đặc thù dạy nghe, hoặc mẫu nghiên cứu có sự đồng nhất cao về nhận thức công nghệ. Dù vậy, kết quả này vẫn phù hợp với xu hướng của các nghiên cứu gần đây trong bối cảnh giáo dục AI, khi vai trò của yếu tố sư phạm và đạo đức được đề cao hơn so với các yếu tố kỹ thuật (Al-Abdullatif, 2024; Tram, 2025).

Đáng chú ý, ảnh hưởng xã hội (SI) có tương quan vừa đến mạnh với BI ($r = .50$, $p < .05$), cho thấy những tác động đến từ đồng nghiệp, sinh viên hoặc nhà trường vẫn đóng vai trò đáng kể trong việc hình thành ý định sử dụng AI. Tuy nhiên, mối liên hệ này kém mạnh hơn so với nhóm biến sư phạm và đạo đức (PS, EA, NLP), hàm ý rằng giảng viên Việt Nam có xu hướng dựa vào năng lực cá nhân và trải nghiệm thực hành hơn là ảnh hưởng xã hội khi ra quyết định ứng dụng AI.

Điều kiện hỗ trợ (FC) lại thể hiện tương quan rất yếu và không đáng kể với BI ($r = .12$, $p > .05$). Mặc dù kết quả định lượng cho thấy FC không có ảnh hưởng trực tiếp, dữ liệu định tính từ phỏng vấn lại thể hiện một bức tranh khác. Nhiều giảng viên thừa nhận rằng: "trường chưa có chủ trương rõ; tôi phải tự mày mò là chính." Phát biểu này cho thấy FC có vai trò gián tiếp – nó không quyết định ý định ban đầu, nhưng lại ảnh hưởng mạnh đến khả năng triển khai và duy trì hành vi ứng dụng AI trong thực tế. Nói cách khác, ở giai đoạn đầu, động lực của giảng viên đến từ yếu tố nội tại (như năng lực sư phạm, trải nghiệm và đạo đức nghề nghiệp), trong khi các điều kiện tổ chức chỉ thực sự trở thành yếu tố quyết định khi chuyển ý định thành hành vi bền vững. Kết quả này cộng hưởng với nhận định của Kasneci et al. (2023) rằng việc thiếu chính sách hỗ trợ, hướng dẫn thể chế và hạ tầng đồng bộ vẫn là rào cản phổ biến trong ứng dụng AI ở bậc đại học. Các trích đoạn phỏng vấn củng cố cho những phát hiện trên. Một giảng viên chia sẻ: "AI hỗ trợ tôi thiết kế bài nghe nhanh, nhưng để sinh viên nghe hiểu tốt hơn thì tôi vẫn phải điều chỉnh nhiều." Một người khác nhấn mạnh: "AI chỉ nên là công cụ hỗ trợ, không thay thế giáo viên." Những quan điểm này phản ánh rõ vai trò của giảng viên như "người gác cổng sư phạm" – người vừa tận dụng AI để đổi mới, vừa giới hạn nó trong khuôn khổ đạo đức nghề nghiệp (Koehler et al., 2014; Mishra & Koehler, 2006).

Tóm lại, phân tích tương quan cho thấy một sự chuyển dịch rõ ràng từ các yếu tố lấy công nghệ làm trung tâm (PE, EE) sang các yếu tố lấy con người và bối cảnh làm trung tâm (PS, EA và NLP). Các giá trị tương quan mạnh ($r \approx .78$–$.79$) khẳng định rằng năng lực kiểm soát sư phạm, ý thức đạo đức và kinh nghiệm thực hành mới là những nhân tố cốt lõi quyết định mức độ sẵn sàng ứng dụng AI trong giảng dạy nghe. Đây là bằng chứng quan trọng mở rộng khung lý thuyết UTAUT truyền thống, đồng thời làm rõ hơn khái niệm AI-TPACK, nơi giảng viên không chỉ là người dùng công nghệ, mà còn là "bộ lọc sư phạm" và "người gìn giữ đạo đức" trong môi trường giáo dục được hỗ trợ bởi AI

Kết quả Bảng 5 và Bảng 6 cho thấy sự sẵn sàng của giảng viên Việt Nam chủ yếu xuất phát từ năng lực sư phạm, ý thức đạo đức và kinh nghiệm thực hành AI, hơn là từ yếu tố kỹ thuật. Ba biến PS, EA và NLP có tương quan mạnh với BI ($r \approx .78–.79$; $p < .001$), phản ánh sự dịch chuyển từ "sẵn sàng về mặt công nghệ" sang "sẵn sàng về sư phạm và đạo đức". Trong khi đó, PE và EE không còn ảnh hưởng đáng kể, cho thấy sự bão hòa ở nhóm đã quen AI. SI và FC yếu, khẳng định thiếu hỗ trợ thể chế vẫn là rào cản chính. Kết quả củng cố vai trò của giảng viên như "người gác cổng sư phạm", đảm bảo ứng dụng AI vừa đổi mới, vừa nhân văn và an toàn trong dạy kỹ năng nghe.

5. Kết luận

Nghiên cứu này đã tiến hành khám phá ban đầu về mức độ sẵn sàng của 25 giảng viên EFL tại Việt Nam trong việc tích hợp AI vào giảng dạy kỹ năng nghe. Các kết quả đã cho thấy sự phân tầng rõ rệt giữa các cấp độ sẵn sàng của giảng viên và điều kiện hệ thống: ở cấp độ cá nhân, giảng viên thể hiện sự chủ động cao về mặt sư phạm cùng với ý định sử dụng AI mạnh mẽ. Tuy nhiên, sự sẵn sàng này vẫn bị cản trở bởi môi trường bên ngoài, đặc biệt là sự hạn chế về chính sách, hạ tầng và ảnh hưởng từ cộng đồng. Một trong những quan sát đáng chú ý là ý định ứng dụng AI của họ có mối liên hệ chặt chẽ hơn với nhận thức đạo đức, năng lực sư phạm và trải nghiệm thực tiễn, thay vì chỉ tập trung vào các yếu tố "hữu ích" hay "dễ sử dụng" như trong các mô hình truyền thống.

Từ những phát hiện này, nghiên cứu đưa ra hai hàm ý quan trọng cho lý thuyết và thực tiễn. Về mặt lý thuyết, kết quả gợi ý rằng các mô hình chấp nhận công nghệ, điển hình như UTAUT, có thể cần được xem xét mở rộng với việc bổ sung các yếu tố liên quan đến đạo đức và năng lực sư phạm, những yếu tố được xem là cần thiết để phản ánh đúng thực tiễn phức tạp của việc ứng dụng AI trong giáo dục ngoại ngữ. Về mặt thực tiễn, nghiên cứu chỉ ra nhu cầu cấp thiết của các cơ sở đào tạo trong việc ban hành chính sách rõ ràng, đầu tư vào hạ tầng và tổ chức các chương trình tập huấn chuyên sâu. Những bước đi này là điều kiện quan trọng để tạo ra một hệ sinh thái hỗ trợ, giúp khai thác hiệu quả tiềm năng sẵn có của đội ngũ giảng viên.

Một giới hạn cần lưu ý là cỡ mẫu nhỏ và phương pháp chọn mẫu thuận tiện làm cho các kết quả của nghiên cứu này chưa thể khái quát hóa cho toàn bộ bối cảnh. Tuy nhiên, chúng mang giá trị định hướng quan trọng, mở ra những hướng nghiên cứu tiềm năng cho các công trình tiếp theo. Cụ thể, cần có các khảo sát định lượng với mẫu đại diện lớn hơn để kiểm định các giả thuyết đã được hình thành, đồng thời triển khai các nghiên cứu định tính chuyên sâu và các thực nghiệm sư phạm nhằm đánh giá một cách khách quan tác động thực sự của AI đối với kết quả học tập của sinh viên.

TÀI LIỆU THAM KHẢO

Al-Abdullatif, A. M. (2024). Modeling Teachers' Acceptance of Generative Artificial Intelligence Use in Higher Education: The Role of AI Literacy, Intelligent TPACK, and Perceived Trust. *Education Sciences*, *14*(11), 1209. https://doi.org/10.3390/educsci14111209

An, X., Chai, C. S., Li, Y., Zhou, Y., Shen, X., Zheng, C., & Chen, M. (2023). Modeling English teachers' behavioral intention to use artificial intelligence in middle schools. *Education and Information Technologies*, *28*(5), 5187–5208. https://doi.org/10.1007/s10639-022-11286-z

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, 107468. https://doi.org/10.1016/j.chb.2022.107468

Chaka, C. (2023). Generative AI Chatbots - ChatGPT versus YouChat versus Chatsonic: Use Cases of Selected Areas of Applied English Language Studies. *International Journal of Learning, Teaching and Educational Research*, *22*(6), 1–19. https://doi.org/10.26803/ijlter.22.6.1

Cheah, Y. H., Lu, J., & Kim, J. (2025). Integrating generative artificial intelligence in K-12 education: Examining teachers' preparedness, practices, and barriers. *Computers and Education: Artificial Intelligence*, *8*, 100363. https://doi.org/10.1016/j.caeai.2025.100363

Cong-Lem, N., Nguyen Tat, T., & Nghi, T. (2024, Tháng Giêng). *Academic integrity in the age of generative AI: Perceptions and responses of Vietnamese EFL teachers*. https://tewtjournal.org/download/academic-integrity-in-the-age-of-generative-ai-perceptions-and-responses-of-vietnamese-efl-teachers-by-ngo-cong-lem-and-tin-nghi-tran-and-tat-thang-nguyen/

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and Conducting Mixed Methods Research*. SAGE Publications.

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, *55*(6), 2503–2529. https://doi.org/10.1111/bjet.13460

Danang University of Architecture, Da Nang, Vietnam, & Thai Phuong L. (2024). CHATGPT IN LEARNING AND RESEARCH: PERSPECTIVES FROM LECTURERS AND STUDENTS OF UNIVERSITIES IN DA NANG CITY. *Vinh University Journal of Science*, *53*(Special Issue 1), 155–162. https://doi.org/10.56824/vujs.2024.htkhgd36

Eke, E. O. (2024). Assessing the readiness and attitudes of Nigerian teacher educators towards adoption of artificial intelligence in educational settings. *Journal of Educational Technology and*

Online Learning, *7*(4-ICETOL 2024 Special Issue), 473–487. https://doi.org/10.31681/jetol.1503305

Farrús, M. (2023). Automatic Speech Recognition in L2 Learning: A Review Based on PRISMA Methodology. *Languages*, *8*(4), 242. https://doi.org/10.3390/languages8040242

Field, J. (2009). *Listening in the Language Classroom*. Cambridge University Press. https://doi.org/10.1017/CBO9780511575945

Frosig, T. B., & Romero, M. (2024). *Teacher agency in the age of generative AI: Towards a framework of hybrid intelligence for learning design* (No. arXiv:2407.06655). arXiv. https://doi.org/10.48550/arXiv.2407.06655

Goh, C. C. M., & Vandergrift, L. (2012). *Teaching and Learning Second Language Listening: Metacognition in Action*. Routledge. https://doi.org/10.4324/9780203843376

Gusman, E., Gide, E., El Khodr, M., & Chaudhry, G. (2024). The Benefits and Challenges of Using Artificial Intelligence in Teaching English as a Foreign Language in Higher Education. *2024 21st International Conference on Information Technology Based Higher Education and Training (ITHET)*, 1–7. https://doi.org/10.1109/ITHET61869.2024.10837597

Hua, H. H., & Le, T. T. (2024). Exploring the Impact of AI in Language Education: Vietnamese EFL Teachers' Views on Using ChatGPT for Fairy Tale Retelling Tasks. *International Journal of Learning, Teaching and Educational Research*, *23*(3). https://ijlter.org/index.php/ijlter/article/view/9885

Huang, X., Han, X., & Dou, A. (2024). Generative AI in College English Listening Instruction: Exploring Teacher Experiences and Pedagogical Integration. *Region - Educational Research and Reviews*, *6*(12), 183. https://doi.org/10.32629/rerr.v6i12.3027

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Koehler, M. J., Mishra, P., Kereluik, K., Shin, T. S., & Graham, C. R. (2014). The Technological Pedagogical Content Knowledge Framework. Trong J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (B.t.v), *Handbook of Research on Educational Communications and Technology* (tr 101–111). Springer. https://doi.org/10.1007/978-1-4614-3185-5_9

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, *5*, 100156. https://doi.org/10.1016/j.caeai.2023.100156

Loan, T. T. T., & Thủy, N. T. H. (2024). Vietnamese University Lecturers Apply AI in Teaching. A Case Study in Thu Dau Mot University. *European Journal of Theoretical and Applied Sciences*, *2*(6), 643–650. https://doi.org/10.59324/ejtas.2024.2(6).57

Ly, P. N. K. (2025). "ARE YOU ON BOARD?": UNIVERSITY EFL INSTRUCTORS' READINESS FOR INTEGRATING AI-POWERED TOOLS IN TEACHING. *VNU Journal of Foreign Studies*, *41*(3), 74–91. https://doi.org/10.63023/2525-2445/jfs.ulis.5524

Ma, Q., Crosthwaite, P., Sun, D., & Zou, D. (2024). Exploring ChatGPT literacy in language education: A global perspective and comprehensive approach. *Computers and Education: Artificial Intelligence*, *7*, 100278. https://doi.org/10.1016/j.caeai.2024.100278

Mananay, J. A. (2024). Integrating Artificial Intelligence (AI) in Language Teaching: Effectiveness, Challenges, and Strategies. *International Journal of Learning, Teaching and Educational Research*, *23*(9). https://ijlter.org/index.php/ijlter/article/view/10884

McCrocklin, S. (2024). Learners' Feedback Regarding ASR-based Dictation Practice for Pronunciation Learning. *CALICO Journal*. https://doi.org/10.1558/cj.34738

Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Mutanga, M. B., Jugoo, V., & Adefemi, K. O. (2024). Lecturers' Perceptions on the Integration of Artificial Intelligence Tools into Teaching Practice. *Trends in Higher Education*, *3*(4), 1121–1133. https://doi.org/10.3390/higheredu3040066

Nunnally, J., & Bernstein, I. (1994). The Assessment of Reliability. Psychometric Theory, 3, 248–292. Nyhan, RC, and HA Marlowe.(1997). Development And Psychometric Properties Of The Organizational Trust Inventory. *Eval Rev*, *21*, 614–635.

Pham, M. T., & Cao, T. X. T. (2025). The Practice of ChatGPT in English Teaching and Learning in Vietnam: A Systematic Review. *International Journal of TESOL & Education*, *5*(1), 50–70. https://doi.org/10.54855/ijte.25513

Pham, V. P. H., & Le, A. Q. (2024). ChatGPT in Language Learning: Perspectives from Vietnamese Students in Vietnam and the USA. *International Journal of Language Instruction*, *3*(2), 59–72. https://doi.org/10.54855/ijli.24325

Salha, S., Mousa, A., & Khayat, S. (2025). Artificial Intelligence in Education (AIED) Policies in School Context: A Mixed Approach Research. *Leadership and Policy in Schools*, *24*(1), 27–45. https://doi.org/10.1080/15700763.2024.2443675

Sivaji, V., Prabhavathy, P., Brinda Shree, V., Varadharajan, T. S., Angelin Y.R, S., & Pradeepa, M. (2024). Leveraging Advanced AI-Driven Pedagogical Strategies to Enhance English Listening

Skills. *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)*, 1–5. https://doi.org/10.1109/IC3TES62412.2024.10877491

Teo, T. (2019). Students and Teachers' Intention to Use Technology: Assessing Their Measurement Equivalence and Structural Invariance. *Journal of Educational Computing Research*, *57*(1), 201–225. https://doi.org/10.1177/0735633117749430

Tram, N. H. M. (2025). Unveiling the Drivers of AI Integration Among Language Teachers: Integrating UTAUT and AI-TPACK. *Computers in the Schools*. https://www.tandfonline.com/doi/abs/10.1080/07380569.2024.2441155

Tran, T. T., Pham, T. T., & Le, T. T. (2024). Transformative Interactions: ChatGPT's Role in Facilitating Professional Development Among Vietnamese English as a Foreign Language Teachers. *Higher Learning Research Communications*, *14*(2). https://doi.org/10.18870/hlrc.v14i2.1558

Tutton, M., & Cohen, D. (2025). Reconceptualizing the Role of the University Language Teacher in Light of Generative AI. *Education Sciences*, *15*(1), 56. https://doi.org/10.3390/educsci15010056

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Wang, W. (2024). A Study of Artificial Intelligence-Assisted Listening Training in College English Teaching. *Applied Mathematics and Nonlinear Sciences*, *9*(1). https://doi.org/10.2478/amns-2024-2657

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1261955

Zhai, X. (2024). *Transforming Teachers' Roles and Agencies in the Era of Generative AI: Perceptions, Acceptance, Knowledge, and Practices* (No. arXiv:2410.03018). arXiv. https://doi.org/10.48550/arXiv.2410.03018

GIAO THOA VĂN HÓA GIỮA TỪ HÁN VIỆT VÀ TỪ THUẦN VIỆT (QUA HÁT- ĐỐ VÀ HÁT- ĐỐI TRONG HÁT PHƯỜNG VẢI)

Đỗ Thị Kim Liên[1]

TÓM TẮT: Khái niệm Giao thoa văn hóa là quá trình hai hay nhiều nền văn hóa khác biệt tiếp xúc, tương tác và hòa quyện với nhau, dẫn đến sự thay đổi, kết hợp các yếu tố như phong tục, truyền thống, tín ngưỡng, nghệ thuật và ngôn ngữ, tạo ra những bản sắc văn hóa mới và độc đáo. Quá trình này mang tính toàn cầu, không chỉ là sự thay đổi của một nền văn hóa mà còn là hành trình kết nối các giá trị nhân loại, hình thành nên một bức tranh văn hóa đa dạng và phong phú trong kỷ nguyên hội nhập, bài viết đề cập đến đặc điểm của giao thoa văn hóa, đó là:

- Sở dĩ có giao thoa văn hóa ít nhất là giữa hai ngôn ngữ gẫn gũi nhau là do sự tiếp xúc. Đây là điểm khởi đầu của quá trình giao thoa, khi các cộng đồng văn hóa khác nhau có cơ hội tiếp xúc trực tiếp và lâu dài với nhau.

- Tiếp theo, do hai ngôn ngữ, hai nền văn hóa có sự tiếp xúc nên dẫn đến có sự biến đổi, các yếu tố văn hóa từ những nền văn hóa khác nhau được tiếp nhận, kết hợp để tạo ra những mô thức văn hóa mới.

- Cuối cùng, do giao thoa văn hóa dẫn đến sự hình thành một bản sắc văn hóa độc đáo, là sự pha trộn giữa các yếu tố truyền thống và hiện đại. Giao thoa văn hóa tác động đến nhiều lĩnh vực của đời sống xã hội như ẩm thực, nghệ thuật, ngôn ngữ, tôn giáo và lối sống.

Từ khóa: chữ Hán, từ Hán Việt, cấu trúc, ngữ pháp học, từ vựng học, tiếp xúc, giao thoa

Cultural Interference Between Sino-Vietnamese Words and Native Vietnamese Words

CULTURAL INTERFERENCE BETWEEN SINO-VIETNAMESE WORDS AND NATIVE VIETNAMESE WORDS (THROUGH SINGING-RIDDING AND SINGING-CHOOSES IN PHUONG VAI SINGING)

(ĐÔ THỊ KIM LIÊN)

SUMMARY: The concept of cultural interaction refers to the process in which two or more distinct cultures come into contact, interact, and blend with one another, leading to changes and combinations of elements such as customs, traditions, beliefs, arts, and languages. This process results in the creation of new and unique cultural identities. It is a global phenomenon - not merely the transformation of a single culture, but a journey that connects human values and forms a diverse, rich cultural landscape in the era of integration. This paper discusses several key characteristics of cultural interaction:

- First, cultural interaction—at least between two closely related languages—begins with contact. This initial stage occurs when different cultural communities have the opportunity for direct and prolonged exposure to one another.

[1] GS.TS Khoa NN, Trường ĐH Dân lập Phương Đông, *Email: dokimlien51@gmail.com*

- Next, due to this contact, changes occur. Cultural elements from different backgrounds are absorbed and combined, giving rise to new cultural patterns.

- Finally, cultural interaction leads to the formation of a distinctive cultural identity - a fusion of traditional and modern elements. This phenomenon influences many areas of social life, including cuisine, art, language, religion, and lifestyle.

Keywork: Chinese characters, Sino-Vietnamese words, structure, vocabulary, grammar, exchange, contact.

1. Đặt vấn đề

Trong HPV, tiếp theo chặng thứ nhất (gồm: hát dạo, hát chào, hát mừng, hát hỏi) đến chặng thứ hai gồm Hát đố - Hát đối. Theo Ninh Viết Giao "đây là chặng hát quan trọng, vì chặng hát thứ hai này đòi hỏi trí thông minh, sự ứng đối hợp lý hợp tình, đồng điệu của hai bên nam, nữ. Nếu hợp với bên nữ, bao giờ được bên nữ đồng ý, cất lời cho phép hát tiếp thì mới được tiếp tục hát liên tục cho đến khi được hát chặng thứ ba: Hát xe kết. Đặc điểm của chặng thứ hai: Hát đố - Hát đối:

a/ Về thời gian, bên nữ chủ động *hát đố* trước vì bên nữ là chủ - đợi bên nam đến nhà, nhằm thử tài, tìm hiểu trí tuệ bên nam, buộc bên nam hát đối lại để thể hiện năng lực, tài trí của mình.

b/ Về cấu tạo, khi cất lời hát đố, bên nữ thường dùng hành động (động từ) *đố* (gọi vai tiếp nhận lời đố là bên nam, ngôi thứ hai số ít, thường gồm những từ ngữ lịch sự, đề cao bên nam: *anh, chàng, quân tử, khách tương phùng, văn nhân, Hán vương, nho sĩ*...; và từ nghi vấn để hỏi bên nam là *chi, ai, mấy, có...không*...

c/ Về nội dung - ngữ nghĩa, bên nữ đưa ra lời hát đố bên nam mong muốn được gặp người tài năng, có học vấn cao, thông tường địa lý. Còn bên nam hát đối lại bên nữ mang nội dung - ngữ nghĩa là mong được gặp bên gái có sắc đẹp, hát hay, chưa có nơi nào dạm hỏi.

2. Xung quanh một số khái niệm

2.1. *Khái niệm hát đố - hát đối trong HPV*

a. *Hát đố*

Thực chất hát đố trong HPV là lời câu đố được cất lên theo nhạc điệu HPV (nữ hát tình tứ, êm nhẹ, hỏi tự do...*Ơ người đi nhởi (chơi) ơi... Thưa chị em ơi* ...[I, tr.85]; Nam hát tình tứ, êm nhẹ, đáp tự do [I, tr.86]. Theo Nguyễn Như Ý trong *Đại từ điển tiếng Việt* (2007), khái niệm **đố**, có hai nghĩa sau: **I**. *đg*. Hỏi, yêu cầu người khác trả lời, làm để biết trí thông minh và khả năng của người ấy: *đố chữ *ra câu đố *đố bài toán khó *Đố ai quét sạch lá rừng, Để ta khuyên gió gió gió đừng rung cây* (cd). **II**.*pt*. Hoàn toàn không, không hề có: *nói thế mà đố nó có sợ* [9, tr.542]. Trong bài viết này, chúng tôi sử dụng khái niệm **đố** với nghĩa **I.** *đg*. Hỏi, yêu cầu người khác trả lời, làm để biết trí thông minh và khả năng của người ấy.

b. *Hát đối*

Theo Nguyễn Như Ý trong *Đại từ điển tiếng Việt,* Hát đối được xem là: *dt*. **1**. Một thể văn biền ngẫu bao gồm hai vế đối nhau về từ loại, ý nghĩa. Dùng để biểu lộ một chí hướng, một thái độ, tình cảm trước một hiện tượng đột xuất hoặc để thử tài, trí thông minh: *câu đối tết: *nghệ thuật viết câu đối*; **2**. Đồ trang trí treo thành cặp song song hai tấm gỗ, hoặc hai mảnh vải hay giấy dài trên có viết câu đối: *Câu đối sơn son thếp vàng *treo câu đối* [9, tr.213]. Sau đây, bài viết của chúng tôi, sử dụng hát đối, xét trong quan hệ với hát đố, ở nghĩa **I.**

Trong HPV, hát đố khác với hát hỏi ở chỗ, *hát hỏi* luôn đi thành cặp với hát trả lời - thuộc *chặng thứ nhất*, hai bên nam - nữ mới gặp nhau nên bên nữ hỏi chỉ trong phạm vi hỏi làm quen, tìm hiểu buổi đầu (về gia thế bên nam), hoặc bên nam hỏi bên nữ luôn gắn trực tiếp với bên nữ là đã có người/ ai thương chưa. Còn hát đố do bên nữ đưa ra để hiểu trí thông minh của bên nam, luôn gắn thành cặp hát đối về ý nghĩa, cấu tạo – *thuộc chặng hai*.

Về nội dung - ngữ nghĩa, Hát đố của bên nữ có liên quan đến ba vấn đề chình: (b1) Tri thức đời sống – xã hội; (b2) Chơi chữ; (b3) Giao thoa văn hóa giữa từ Hán Việt và từ thuần Việt. Trong bài viết này, chúng tôi chỉ đi sâu phân tích nội dung-ngữ nghĩa mục (b3).

Như chúng ta biết, chữ Hán là một hệ thống văn tự - âm tiết, được phát triển ở Trung Quốc và được sử dụng để viết các ngôn ngữ như tiếng Hàn, tiếng Nhật, tiếng Triều Tiên và tiếng Việt (trong hệ thống chữ Nôm).

Nói đến chữ Hán (viết tắt là CH) là nói đến chữ vết tiếng Hán. Còn từ Hán Việt (viết tắt HV) là những từ mượn tiếng Hán vào tiếng Việt, có nghĩa gốc từ tiếng Hán nhưng được viết bằng chữ Latinh và có cách đọc theo âm tiếng Việt. Ví dụ*: Gia đình, phụ mẫu, học sinh, sinh viên, vợ chồng, hòa bìn*h.

2.2. *Khái niệm giao thoa và giao thoa văn hóa*

a. *Giao thoa*

Theo *Từ điển tiếng Việt* do Hoàng Phê (chủ biên) đt. (hiện tượng hai hay nhiều sóng cùng tần số) tăng cường hay làm yếu lẫn nhau khi gặp nhau tại cùng một điểm. *Sự giao thoa của ánh sáng* [7, tr. 378]. Trong ngôn ngữ học, giao thoa là hiện tượng các ngôn ngữ khác nhau tác động qua lại lẫn nhau, ảnh hưởng đến cấu trúc và vốn từ của nhau. Nó xảy ra khi người nói các ngôn ngữ khác nhau tiếp xúc và giao tiếp, tạo ra sự ảnh hưởng lẫn nhau trong quá trình sử dụng ngôn ngữ.

b. *Giao thoa văn hóa*

Giao thoa văn hóa là quá trình hai hay nhiều nền văn hóa khác biệt tiếp xúc, tương tác và hòa quyện với nhau, dẫn đến sự thay đổi, kết hợp các yếu tố như phong tục, truyền thống, tín ngưỡng, nghệ thuật và ngôn ngữ, tạo ra những bản sắc văn hóa mới và độc đáo. Quá trình này

mang tính toàn cầu, không chỉ là sự thay đổi của một nền văn hóa mà còn là hành trình kết nối các giá trị nhân loại, hình thành nên một bức tranh văn hóa đa dạng và phong phú trong kỷ nguyên hội nhập.

c. *Đặc điểm của giao thoa văn hóa*

- Sở dĩ có giao thoa văn hóa ít nhất là giữa hai ngôn ngữ gẫn gũi nhau là do sự tiếp xúc. Đây là điểm khởi đầu của quá trình giao thoa, khi các cộng đồng văn hóa khác nhau có cơ hội tiếp xúc trực tiếp và lâu dài với nhau.

- Tiếp theo, do hai ngôn ngữ, hai nền văn hóa có sự tiếp xúc nên dẫn đến có sự biến đổi, các yếu tố văn hóa từ những nền văn hóa khác nhau được tiếp nhận, kết hợp để tạo ra những mô thức văn hóa mới.

- Cuối cùng, dẫn đến chúng ra bản sắc mới: Giao thoa văn hóa dẫn đến sự hình thành một bản sắc văn hóa độc đáo, là sự pha trộn giữa các yếu tố truyền thống và hiện đại. Giao thoa văn hóa tác động đến nhiều lĩnh vực của đời sống xã hội như ẩm thực, nghệ thuật, ngôn ngữ, tôn giáo, và lối sống.

Từ đặc điểm tất yếu trên của giao thoa văn hóa, ta thấy vai trò quan trọng của giao thoa văn hóa thể hiện trong nhiều lĩnh vực, đó là:

- Thúc đẩy sự phát triển của các dân tộc khác nhau nói chung. Giao thoa văn hóa tạo ra những giá trị mới, góp phần làm phong phú thêm nền văn hóa của các dân tộc khác nhau do họ cùng vui hát, mua bán, giao tiếp, tiếp xúc trao đổi nhiều mối quan quan hệ xã hội.

- Xây dựng cầu nối để giao lưu văn hóa giữa nhiều cộng đồng quốc tế nói chung, trong đó tăng cường sự hiểu biết và giao lưu giữa các cộng đồng quốc tế một cách nhanh nhất, phù hợp nhất, sâu tận mọi nẻo đường của cuộc sống con người.

- Góp phần khẳng định bản sắc dân tộc, điều này xuất phát từ dòng chảy hội nhập, giao thoa văn hóa đã giúp thế hệ trẻ tiếp bước và giữ gìn, phát huy giá trị truyền thống của dân tộc mình đi trước.

2.3. *Quan hệ giữa từ Hán Việt và từ thuần Việt*

a. *Giống nhau*

Từ Hán Việt hay từ thuần Việt **đều l**à từ vựng tiếng Việt được sử dụng trong tiếng Việt chứ không phải tiếng Hán. Cả hai đều thuộc vốn từ tham gia vào quá trình giao tiếp, cùng thực hiện chức năng biểu đạt. Từ thuần Việt hay Hán Việt đều có chức năng chung là diễn đạt ý nghĩa, cảm xúc của người nói nên chúng đều là những yếu tố cấu thành vốn từ vựng, được người Việt sử dụng để biểu đạt ý nghĩa.

b. *Khác nhau*

Từ thuần Việt có nguồn gốc từ dân tộc Việt, mang sắc thái thân mật, phổ biến trong đời sống hàng ngày, trong khi đó, từ Hán Việt là mượn từ gốc tiếng Hán, có sắc thái trang trọng, tao nhã và thường dùng trong phong cách văn chương, chính luận. Xét về quan hệ cú pháp, nếu từ thuần Việt có quan hệ Chính - Phụ, thì bao giờ thành tố chính cũng đứng trước, thành tố phụ đứng sau, ngược lại, nếu từ Hán Việt có quan hệ Chính - Phụ, thì thành tố phụ đứng trước, thành tố chính đứng sau, VD: từ thuần Việt: *gà mái*→ thì *gà* là thành tố chính, *mái* là thành tố phụ; từ Hán Việt: *quốc ca* → thì *quốc* là *nước*, là thành tố phụ, *ca* là *bài ca*, là thành tố chính đứng sau; *quốc ca* là *bài ca* của *nước*; *quốc* là *nước,* là thành tố phụ, đứng trước*, kỳ* là *cờ*, là thành tố chính đứng sau, *quốc kỳ* là *cờ nước*.

3. Biểu hiện giao thoa văn hóa trong lời hát đố - hát đối trong HPV

Trong lời hát đố - Hát đối của Hát phường vải (viết tắt HPV), ta gặp có một số từ đố chữ Hán (viết bằng chữ Hán), thể hiện có sự tham gia của các vị có học chữ Hán. Vì khi Hát có người Hát và có người "bày" Chính vì thế có khi, người bày cho bên nữ "đố" giỏi quá khiến bên đối (bên nam) không đối được bí quá nên bèn chửi bên nữ: "đù con hát, tổ cha thằng bày". Qua thống kê, ta gặp 12 cặp HPV bên nữ sử dụng chữ Hán, từ Hán Việt để hát đố bên nam, còn bên nam hát đối lại để thể hiện trí thông minh, tải năng đối lại của bên mình.

Thứ nhất, bên nữ hát đố theo mô hình (viết tắt MH) - 1a), dùng chữ Hán (viết tắt CH):

MH Đố CH	HĐ (Đấm một đấm, 2 tay ôm, thuyền lên núi) *hỏi* chàng *chữ chi*?

VD 1a): - *Đấm* một *đấm* hai tay ôm quàng,

Thuyền *chèo* lên núi thiếp *hỏi* chàng *chữ chi*?[HPV, I, tr.179]

Như vậy, mặc dù dùng lời viết chữ Hán nhưng bên nữ vẫn dùng Hành động – cụm động từ (*đấm một đấm; chèo lên núi*) đồng thời chúng là những từ thuần Việt, tạo nên nghệ thuật chơi chữ độc đáo của người người Nghệ Tĩnh. Tiếp sau bên nữ dùng từ *hỏi* để *đố* chàng theo lối nói hay văn hóa của người Nghệ Tĩnh sử dụng từ *chi* là *gì* (ứng với lối nói của từ toàn dân).

- Nghe nữ đố 1a) xong, bên nam đối lại theo MH đối 1b):

MH Đối CH 1b)	HĐ miêu tả (cả câu lục và câu bát)

VD1b): - Lại đây nói nhỏ em *nì*,

Ấy là chữ "mật" một khi rõ ràng. [HPV, I, tr.179]

→ theo văn hóa lối nói của người Nghệ Tĩnh, *nì* là đại từ chỉ định, tức có nghĩa là *này* (*nói nhỏ với riêng em này).*

→ "mật" chữ Hán viết là 密 trong câu hát đối của bên nam có nghĩa thứ nhất tường minh là bí mật, kín đáo và nghĩa thứ hai hàm ẩn: khi đến lúc cần, anh sẽ nói rõ ràng cho mọi người biết.

Giải thích: *Đấm một đấm hai tay ôm quàng* là hình dáng bộ miên (宀), phần trên của chữ Hán, ***thuyền chèo lên núi:*** **Gợi nhắc** đến chữ "Tất" (必), phần dưới của chữ Hán, trên chữ sơn 山 có chữ tất 必. Ghép lại, chúng ta được chữ mật 密 (bí mật, rậm rạp). Đồng thời chữ *mật* có nghĩa thuần Việt là dùng để uống một cách ngọ ngào.

Thứ hai, bên nữ hát đố theo MH 2a):

MH Đố CH 2a)	HĐ đưa Hai ngang hai phết + HĐ (kết lại) *chữ chi*

VD 2a): - Hai ngang hai phết kết lại *chữ chi*,

Chàng mà đối được thiếp thì theo không? [HPV, I, tr.179]

→ Bên nữ đưa ra lời hát đố: câu thứ nhất hai ngang hai phết gợi hình chữ "thất"để hỏi đố bên nam đi kèm từ nghi vấn *chữ chi*. Đồng thời chúng cũng là từ thuần Việt mô tả cách viết chữ Hán. Câu thứ hai thể hiện văn hóa lời thách đố đặt cược của bên nữ thể hiện tính lịch "theo không" chàng.

→ Như vậy, người Nghệ Tĩnh thường có văn hóa dùng từ nghi vấn *chi* thay *gì* là phổ biến, quen hơn so với chữ *gì*. Tuy vậy, trong giao tiếp lịch sự, khi cần thiết do tính hài hòa vần nhịp, người Nghệ Tĩnh vẫn dùng từ *gì* (từ toàn dân) để thách đố bên nam trong câu: "Ơi anh cắp sách đi thi,/ Ba *xe* chắp lại, *chữ gì* hỡi anh? [HPV, I, tr.181]

- Nghe xong, bên nam đối lại theo MH đối 2b):

MH Đối CH 2b)	HĐ miêu tả (cả câu lục và câu bát)

VD 2a) - Hai ngang hai phết kết lại *chữ thất*

Thất là mất, mất nước mất nhà,

Dân sầu dân thảm từ ngày Tây qua đến giừ (giờ). [HPV, I, tr.179]

Chữ "thất" có các nghĩa nghĩa khác nhau tùy thuộc vào chữ Hán tương ứng: chữ 七 (chữ Hán Nôm là "thất") nghĩa là ***bảy***, chữ 失 (thất) nghĩa là ***mất***, và ***thất tình*** (từ ghép) là bảy loại cảm xúc, tâm trạng như vui, giận, buồn, sợ, yêu, ghét, muốn. Đồng thời, *mất* còn là từ thuần Việt (*mất* chiếc áo). Bên nam đã hát đối đúng, chọn từ *thất* là *mất*, *mất nước mất nhà* nói lên văn hóa yêu nước đi vào tâm thức của người dân Nghệ Tĩnh.

Thứ ba, ta gặp bên nữ hát đố bên nam theo MH 3a):

MH Đố CH 3a)	*Hỏi* N.2 học kinh thi, *nghìn* người đứng viết chữ *chi rứa* chàng

VD 3a): - *Hỏi* chàng học sách kinh thi,

Nghìn người đứng viết chữ *chi rứa* chàng? [HPV, I, tr.180]

→ Ở VD 3a), ở bên nữ, ta bắt gặp người Nghệ Tĩnh có văn hóa sử dụng từ *chi rứa* là từ để hỏi đố (*chi* giống từ toàn dân *gì*, từ *rứa* giống từ toàn dân là *thế*). Còn từ "thiên" là một từ Hán Việt, có cùng nghĩa là "nghìn" và thường được dùng trong các từ ghép như "thiên niên kỷ" hoặc "thiên hà". Chữ Hán *thiên* là 千, có phiên âm là *qiān* đọc là nghìn, Đồng thời "nghìn" là một từ thuần Việt, do tiếng Việt sáng tạo ra, dùng để chỉ số lượng 1000.

- Nghe xong, bên nam đối lại theo MH 3b):

MH Đối HV 3b)	HĐ miêu tả (cả câu lục và câu bát)

VD 3b): - Anh đây đọc sách cửu chương,

Nghìn người đứng viết chữ "hương" rõ ràng. [HPV, I, tr. 180]

→ Ở lời hát đối bên nam, chữ "hương"khi là từ Hán Việt thường có hai tiếng (từ ghép) có thể viết là 香 (hương thơm). Đồng thời, ta còn gặp từ thuần Việt đơn tiết *hương, thơm*. Điều này thể hiện văn hóa chơi chữ độc đáo của người Nghệ Tĩnh.

Thứ tư, ta còn gặp bên nữ hát đố bên nam theo MH 4a):

MH Đố CH 4a)	HĐ *hỏi* chàng học sách kinh thi, Đàn bà đi lọng chữ *chi rứa*

VD 5a) - *Hỏi* chàng học sách kinh thi,

Đàn bà đi *lọng* chữ *chi rứa* chàng? [HPV, I, tr. 180]

→ Ở lời hát đố, bên nữ dùng hình thức chữ viết tượng hình của chữ Hán. **Từ Hán Việt là từ mượn từ tiếng Hán nhưng được viết bằng chữ cái Latinh và đọc theo âm tiếng Việt**, chứ không phải giữ lại chữ viết theo chữ Hán. Vì thế, *đàn bà đi lọng* chỉ là mô tả hình thức chữ viết nhưng lại được đọc bằng âm Việt (yên/ an) nên đồng thời chứa nét văn hóa chơi chữ gắn với từ thuần Việt.

- Nghe xong, bên nam đối lại theo MH đối 5b):

MH Đối CH 4b)	HĐ miêu tả (cho cả câu lục và câu bát)

VD 5b) - Anh đây học sách cửu thiên

Đàn bà đi lọng chữ "yên" rõ ràng. [HPV, I, tr.180]

→ Nét đặc sắc riêng về văn hóa của người Nghệ Tĩnh được thể hiện khi bên nam hát đối "*Đàn bà đi lọng chữ Yên rõ ràng*". Theo từ điển, từ *lọng* (d). Vật dùng để che, gần giống cái dù nhưng lớn hơn thường dùng trong nghi lễ đón rước vua quan hoặc thánh thần ngày trước. [7, tr.559]. Đây là sự giao thoa văn hóa, thể hiện cách chơi chữ kết hợp giữa từ Hán Việt và từ thuần Việt.

- **Thứ năm**, ta còn gặp bên nữ hát đố theo MH 6a):

MH Đố CH 5a)	HĐ ba ngang ba sổ chữ *chi rứa* chàng?

VD 5a: - Nghe tin anh học kinh thi,

Ba ngang ba sổ chữ *chi rứa* chàng? [HPV, I, tr.180]

→ Bên nữ mô tả chữ tượng hình viết chữ *điền* gồm ba nét ngang và ba nét sổ dọc 田. để đố bên nam.

- Nghe xong, bên nam đối lại bên nữ theo MH 5b):

MH Đố CH 5b)	HĐ (vạch) ba ngang ba sổ chữ *điền* em ơi

VD 5b): - Anh đây học sách thánh hiền,

Ba ngang ba số chữ "điền" em ơi! [HPV, I, tr.180]

→ Bên nam tỏ rõ mình là ngươi có học, đã hát đối lại bên nữ một cách nhanh nhẹn, tự tin: *điền* là từ Hán Việt có nghĩa là *ruộng* là từ thuần Việt – gắn với nghề làm nông, nơi cư trú, sinh sống của hai bên nam - nữ. Đây cũng là giao thoa văn hóa, thể hiện cách chơi chữ kết hợp giữa từ Hán Việt và từ thuần Viêt.

- **Thứ sáu**, ta gặp bên nữ hát đố theo MH 6a):

MH Đố HV 6a)	Nội đồng ngoại cậu nhất *nhân,*

VD 7a) – Nội đồng ngoại cậu nhất *nhân,*

Chàng mà đối được thiếp theo chân chàng về? [HPV, I, tr. 180]

→ Bên nữ hát đố bên nam đã dùng từ Hán Việt *nội đồng* để nói đến hệ thống thủy lợi: kênh, mương, rạch, đường dẫn nước đều đặt lên vai chàng; từ *ngoại cậu* để nói đến mối quan hệ họ hàng bên ngoại thì cậu là người anh trai/ em trai thân thiết, gần gũi của bên mẹ, còn từ *nhất nhân* là chỉ duy nhất một mình chàng, kết hợp cả ba từ ngữ trên với nhau có nghĩa việc đồng áng, thủy lợi cùng việc họ hàng bên ngoại đều do bên chàng nắm giữ, tức đề cao chàng.

- Nghe xong, bên nam đối lại theo MH 7b):

MH Đối HV 6b)	HĐ miêu tả (cả câu lục và câu bát)

VD 7b) – Gái hồng nhan có lúc trung "hưng",

Thử ai thì thử xin đừng thử anh. [HPV, I, tr. 180]

→ Bên nam đối lại cũng đề cao bên nữ, dùng từ Hán Việt *trung hưng* có nghĩa: **khôi phục lại sự thịnh vượng. Như vậy,** sự giao thoa văn hóa thể hiện qua cách chơi chữ kết hợp giữa từ Hán Việt và từ thuần Việt.

Thứ tám, ta còn gặp bên nữ hát đố bên nam theo MH 8a):

MH Đố HV 8a)	HĐ gọi đố: Ơi! anh cắp sách đi thi!, Ba *xe* chắp lại *chữ - gì*?

VD 8a) - Ơi anh cắp sách đi thi,

Ba *xe* chắp lại, *chữ gì* hỡi anh? [HPV, I, tr.181]

→Ở câu lục, bên nữ gọi bên nam (*ơi anh*) để hỏi đố: Ba xe chắp lại là *chữ gì*) gợi cho bên nam cùng những người nghe suy ra từ chữ tượng hình gồm ba *xe* (từ Hán Việt là *xa*) đứng cạnh nhau chắp lại mà suy ra lời hát đối lại câu hát đố của bên nữ. Ở đây sự giao thoa văn hóa thể hiện qua cách chơi chữ kết hợp giữa từ Hán Việt và từ thuần Việt.

- Nghe xong, bên nam đối lại bên theo MH 9b):

MH Đối HV 8b)	HĐ miêu tả (cho cả câu lục và câu bát)

VD 8b) – Ba xe chắp lại chữ "Oanh" [HPV, I, tr.181]

Anh quyết oanh liệt để danh phen này. [HPV, I, tr.181]

Thứ chín, tiếp theo, ta còn gặp bên nữ hát đố bên nam theo MH 9a):

MH Đố HV 9a)	HĐ Giảng câu "Nhị đào thà bẻ cho người tình "chung"

VD 9a): – Truyện Kiều anh giảng đã tài

Đố anh giảng được câu này anh ơi!

Biết thân đến bước lạc loài

Nhị đào thà bẻ cho người tình "chung". [HPV, I, tr.181]

→ Bên nữ hát đố bên nam có từ " chung" có hai cách viết, cách thứ nhất

viết là chung (終) với nghĩa là "chấm hết", "kết thúc", cách thứ hai viết là chung (鍾) với nghĩa là "chuông", "người chung tình".

- Nghe xong, bên nam đối lại theo MH đối 9b):

MH Đối HV 9b)	HĐ miêu tả (Kim Trọng – vào ra) cả câu lục và câu bát

VD 9b) – Tình "chung" chẳng phải ai xa,

Chính chàng Kim Trọng vào ra sớm chiều. [HPV, I , tr.181]

→ Bên nam chọn lời hát đối là từ "chung" do chữ Kim và chữ Trọng. để gián tiếp nói đến tình yêu trong sáng, thủy chung giữa Thúy Kiều và Kim Trọng.

sự giao thoa văn hóa thể hiện qua cách chơi chữ kết hợp giữa từ Hán Việt (*chung*) và từ thuần Việt (*người tình chung*).

Tóm lại, tiểu nhóm Đố - đối tri thức về chữ Hán, từ Hán Việt, từ thuần Việt gồm 12 cặp nhưng được chia ra 9 mô hình.

4. Kết luận

Qua phần trình bày trên, chúng tôi rút ra mộ số kết luận:

- Chặng Hát đố - hát đối nam nữ thường do bên nữ chủ động hát trước với mục đích thách đố bên nam để tìm hiểu trí thông minh, tài đức của bên nam để xa hơn, có thể kết duyên đôi lứa. Vì vậy, bên nam phải là người giỏi chữ Hán, hiểu biết văn chương, hết sức lắng nghe, hiểu nghĩa, phân tích nhanh để tìm cách hát đối lại sao cho đúng nghĩa lại có tính nghệ thuật, làm vừa lòng đẹp ý bên nữ, được bên nữ cho phép mới được hát tiếp sang chặng thứ ba.

- Xét về nội dung ý nghĩa, ở chặng thứ hai, tiểu nhóm Hát đố - hát đáp nam nữ có vai trò hết sức quan trọng HPV, thể hiện trí tuệ minh mẫn, sự thông tường văn sử địa của đôi bên, vốn ca dao dân ca của bên nam và bên nữ cũng đi vào tâm thức để ứng đối nhanh, đặc biệt, qua những lời hát đố - hát đối đó, nét giao thoa văn hóa trong việc sử dụng từ thuần Việt và từ Hán Việt để chơi chữ được thể hiện. Mặc dù như chúng ta biết Thực đân Pháp đã đô hộ Việt Nam 80 năm, kéo theo tiếng Pháp được đưa vào sử dụng trong giao tiếp hàng ngày và dạy học ở Nhà trường. Theo tư liệu điều tra, người Việt Nam nói chung và người Nghệ Tĩnh nói riêng vẫn tiếp tục sử dụng chữ Hán, từ Hán Việt, từ thuần Việt hết sức thông dụng. Điều này nói lên sự giao thoa văn hóa giữa từ Hán Việt và từ thuần Việt đã đi vào đời sống của người dân xứ Nghệ theo phương thức riêng không pha trộn với vùng miền khác.

TÀI LIỆU THAM KHẢO

Tài liệu tiếng Việt

1. Đào Duy Anh (Bản đầu tiên – 1947, Nxb Tân Việt), (2001), *Từ điển Hán Việt*, Nxb Khoa học Xã hội, Hà Nội.
2. Nguyễn Chung Anh (1958), *Hát ví Nghệ Tĩnh*, Nxb Văn - Sử - Địa.
3. Nguyễn Thị Thúy Hà (2014), "Chủ đề tình yêu và khát vọng hạnh phúc gia đình trong hát ví phường vải Nghệ Tĩnh", http://khuditichkimlien.gov.vn/vi/hat-phuong- vai- nghe- tinh-ssn109.html, 2014.
4. Nguyễn Thị Mỹ Hạnh (1992), "Hò, Ví và Giặm Nghệ Tĩnh", *Nghiên cứu văn học nghệ thuật* , (số 6), tr. 31-36.
5. Vũ Tố Hảo (1990), "Vài ghi nhận về con người và bản sắc văn hoá Nghệ Tĩnh", *Văn hoá dân gian*, (số 3), tr. 69-72.
6. Duy Hưng (2014), "Bảo tồn và phát huy giá trị của dân ca Ví, giặm Nghệ Tĩnh", nguồn TUYÊN GIÁO, Tạp chí của Ban Tuyên giáo Trung ương, https://www.tuyengiao.vn>van-hoa-xa-hoi> van-hoa, 12/05/2014.
7. Hoàng Phê (chủ biên) (1995), *Từ điển tiếng Việt*, Nxb Đà Nẵng-Trung tâm Từ điển học.
8. Nguyễn Như Ý (chủ biên) (2001), *Từ điển giải thích thuật ngữ ngôn ngữ học*, Nxb Giáo dục, Hà Nội.
9. Nguyễn Như Ý (2007), *Đại từ điển tiếngViệt*, Nxb Đại học Quốc gia TP Hồ Chí Minh.

Tài liệu tiếng nước ngoài

10. Brown P & Levinson S. (1987), *Politeness: some Universals in Language Usage*, Cambr University Press.
11. Grice HP (1989), *Studier in the way of words*, Cambridge University Press.
12. Searle JR (1979), *Expression and meaning*, Cambridge University Press

NGUỒN NGỮ LIỆU

I. NinhViết Giao (1993), *Hát phường vải*, NXB Nghệ An

ĐC: Đỗ Thị Kim Liên, khoa NN, ĐH Phương Đông; Nhà riêng: Số 1, Ngõ 4, Phạm Kinh Vỹ, Bến Thủy 6, Phường Trường Vinh, Nghệ An; T.0396.373.659;

dokimlien51@gmail.com

THE IMPACT OF GAMIFIED FEATURES OF DUOLINGO ON LEARNING MOTIVATION: A STUDY OF ENGLISH-MAJOR STUDENTS' AI-ENHANCED LANGUAGE LEARNING

Le Ngoc Phi Yen[1], Tran Vu Diem Thuy[2]

Abstract

As artificial intelligence (AI) reshapes the way languages are learned and taught, English as a Foreign Language (EFL) classrooms are increasingly evolving into dynamic environments that extend far beyond traditional walls. AI features enable MALL applications such as Duolingo to offer personalized practice and flexible access beyond classrooms. Within this emerging landscape, understanding learners' motivation is crucial, as it determines students' engagement and the persistence required in language learning. The study aims to investigate the impact of gamified features of Duolingo on learning motivation. A mixed-methods approach following an eight-week intervention was employed. 35 B1-level students majoring in English Language Studies at a university in Vietnam participated in the study. The results revealed that the streaks, which displayed the number of consecutive days a learner had completed the lesson, were the most prominent gamified feature. Meanwhile, the feature of leaderboards had a selective impact on learning motivation. The study provided implications for teachers to incorporate balanced combinations of gamified elements and intrinsic forms of engagement. Future studies should focus on motivational features that could be sustained in the long run, personalization of gamification strategies, and artificial intelligence (AI) applications of adaptive systems to support learning autonomy without developing reward dependency.

Keywords:mobile-assisted language learning, gamification, Duolingo, motivation, persistence

1. Introduction

Amidst exponential increases in Artificial Intelligence (AI), mobile access, and the science of language learning, English instruction has begun to look different around the world. Online platforms offer students, both inside and outside the classroom, opportunities for personalized feedback, individualized sequencing of practice items, and scalable outside-of-classroom learning (Settles & Meeder, 2016). At the same time, learners are increasingly able to access English study through mobile phones, allowing more frequent, flexible, and micro-dosed study sessions (Burston, 2015; Kukulska-Hulme & Shield, 2008).

Gamification, or the use of game design elements in non-game contexts (Deterding et al., 2011), features prominently across many of these platforms. Points, badges, streak counters, leaderboards, and similar game elements are typically included to promote user motivation,

[1] Master in English Language Studies, Founder of "IELTS preparation with Alice", *Email: lengocphiyen196@gmail.com*

[2] Ph.D., Ho Chi Minh City University of Foreign Languages - Information Technology. *Email: thuytvd@huflit.edu.vn*

persistence, and self-regulated, autonomous engagement (Hamari, Koivisto, & Sarsa, 2014; Sailer & Homner, 2020). A meta-analysis found that gamification produces higher usage, but these effects depend on context, user demographics, and task (Hamari et al., 2014). More recent research has mapped individual game features to the autonomy, competence, and relatedness that form the core of Self-Determination Theory (SDT), or the basic psychological needs described to underlie intrinsic motivation (Sailer & Homner, 2020; Deci & Ryan, 2017).

Nonetheless, gamification is not itself a pedagogical panacea. The relationship between gamification features and student motivation is not deterministic, but rather shaped by the social, cognitive, and linguistic relevance of the game elements. This situation is further complicated by the inclusion of machine-learning-based adaptivity. Adaptive platforms can create customized user experiences that adjust the item difficulty, content, or lesson sequencing according to a learner's performance, and these mechanisms should align with gamification features to promote, not frustrate, usage. For example, leaderboards can be meaningful for students only when they can see their own rank compared to that of other students of similar ability. Streaks and other habits-of-the-self feedback may also help maintain persistent engagement in the long term, but can create self-defeating cycles of demotivation when practice breaks occur (Loewen et al., 2019). While there is a rapidly expanding literature on AI, gamification, and MALL in recent years, few existing studies have holistically tested the interaction of AI and gamification features to understand how these elements jointly influence motivation and persistence in EFL settings.

In Vietnam, high levels of English proficiency continue to serve as a near-universal gatekeeper to tertiary education, career opportunities, and socioeconomic mobility (Nguyen, 2019). In response to an increasing need to meet international standards, most Vietnamese universities now align their curricula to the Common European Framework of Reference for Languages (CEFR) standards (Council of Europe, 2020). However, classroom practices remain relatively uniform across institutions, with many programs still relying primarily on a grammar-translation approach (Nguyen, 2019). These traditional practices offer students limited opportunities to practice authentic reading, particularly outside of the classroom, and can produce unmotivated EFL learners who see little value in their studies beyond the exam itself. This pattern is particularly marked among older learners with full-time jobs. University students face similarly dire opportunities for English practice: while some universities in urban areas are well-resourced, students must often travel to and from campus for multiple years, work part-time to supplement their expenses, and face irregular access to computers and other high-quality internet-connected devices.

In this context, free and low-cost MALL apps, such as Duolingo, provide one low-barrier entry point to out-of-classroom study. Using item-ordering learned through machine intelligence, spaced-repetition algorithms, and A/B testing of millions of users, Duolingo delivers personalized

daily lessons to its users (Settles & Meeder, 2016). At the same time, its core, gamified user interface leverages streak counters, experience points, badges, levels, and social leaderboards to help make incremental learning visible, habit-forming, and socially meaningful over time. Duolingo's two primary instructional tasks to date (translation and vocabulary-in-context) can also provide a well-designed set of mini-events that help build students' larger reading-related skills, such as word recognition and fluency, sentence parsing, cohesion, or gist and scan reading, among others. A growing body of research suggests that Duolingo and similar gamified MALL tools may be able to promote extrinsic motivation (through points and social signaling) as well as some intrinsic motivation through autonomy-supportive and competence-boosting elements (Loewen et al., 2019). Few existing studies have tested the alignment between micro-learning events and macro-level reading objectives or how the included gamification features may impact students' motivation to practice and their self-sustained persistence across days.

In this research, I position gamification less as a collection of design techniques or ideas and more as a set of discrete mechanisms of action (MoAs) within the broader ecosystem of AI-mediated MALL. To do so, the present study focuses on a student population currently understudied in MALL research: Vietnamese urban higher education students (Nguyen, 2019; Loewen et al., 2019). Focusing on this population will allow for a thorough investigation of several features in Duolingo, a single tool that has not yet been systematically analyzed for user motivation and persistence in EFL research in Viet Nam.

In particular, this study attempts to address how Duolingo impacts the motivation and long-term, self-sustained persistence of Vietnamese EFL learners. Therefore, it attempted to answer the following questions:

1) What are the gamified features of Duolingo affecting learning motivation among Vietnamese EFL learners?
2) To what extent do these gamified features of Duolingo affect learning motivation?

Aligning microlearning events with macro-reading outcomes is crucial to the task analysis that underlies any MALL tool (Sailer & Homner, 2020). The present study, however, seeks to contribute to an emerging literature that centers on a motivation-behavior-performance pathway for understanding EFL learners' experience with gamification and AI in MALL settings. The specific, contextual findings from this study should be useful to educators, teachers, researchers, or digital developers interested in scaling low-cost and automated support for out-of-classroom EFL practice in resource-variable institutional contexts.

2. Literature review

2.1. Theoretical Frameworks

The present study was built on the premise of two interwoven theoretical frameworks: Self-Determination Theory (SDT) and Technology Acceptance Model (TAM). In combination, these provide both the macro and micro context for analyzing the potential affordances of Duolingo in addressing EFL reading persistence, motivation, and relevant micro-skills as seen through the prism of gamification and AI-based adaptivity.

First, SDT (Deci & Ryan, 2017) orients the argument around the role of autonomy, competence, and relatedness as key psychological needs in human motivation. By breaking learning into self-paced streaks and rewarding progress with visual badges, Duolingo's gamification may nudge extrinsic motivation in the desired direction and tap into the core need for competence (perceived progress). By contrast, daily reminders and other flexible practice strategies theoretically foster autonomy. For SDT and related works on motivation in MALL (Ha, 2024), the main consideration is whether extrinsic triggers displace or build up intrinsic motivation, especially if the reward system begins to be perceived as controlling rather than informational.

Second, TAM (Davis, 1989) serves to address how Vietnamese learners will or will not choose to use Duolingo. TAM's primary predictors of new technology use in education are perceived usefulness and ease of use, making it relevant to the Vietnamese context of rapidly growing mobile language practice and instruction (Al-Emran et al., 2018).

As a whole, these frameworks underpin the conceptualization of gamification as not only a set of game design principles but a motivational scaffolding that operates on multiple levels and incorporates extrinsic triggers that can potentially scaffold intrinsic goals (SDT) and a risk of prompting distraction, excessive, or infrequent use (TAM). Motivational theories are especially crucial to the MALL literature when it comes to user uptake and perseverance with open-access tools that promise to alleviate skill gaps in autonomous and digitally mediated practice, as well as curriculum gaps and access to instructional resources in less-resourced contexts with diverse learner readiness (Nguyen, 2019).

2.2. Gamification and Its Gamified Features

Duolingo is one of the most well-known gamified language learning applications. It is part of an explosion of gamified learning environments, both in language learning and other subject areas. In CALL (computer-assisted language learning), gamification has become both a design trope and a pedagogical strategy. Narrowly, gamification can refer to specific features such as badges, points, leaderboards, or avatars, but on a larger scale, gamification can be considered as a design practice, or adding to a service or product an element of gameplay (Deterding et al., 2011). Gamification, once built into a mobile app, is frequently blended with AI (artificial intelligence) personalization features, creating adaptive systems that may restructure motivation, engagement, and performance outcomes.

Pedagogically, gamification can be linked with motivational theories such as SDT. SDT has been used to describe how CALL design affects motivation, on the principle that learners' motivation to learn is maintained by satisfying their basic psychological needs of autonomy, competence, and relatedness (Deci & Ryan, 2017). In Duolingo, these constructs can be clearly seen. Autonomy is supported by the ability to navigate lesson sequences and choose which skills to learn, and more abstractly through affordances such as a "permission to play" or productive struggle (Zhang & Hasim, 2023). Competence is supported through corrective feedback, progress tracking, and explicit achievement markers such as the XP bar or a crown for a skill. Relatedness is supported by social affordances such as leaderboards and in-app communities (Chen et al., 2020). Quantitative studies have shown that gamification features can increase self-regulated learning in language learning, an instructional outcome related to learners' strategies, knowledge development, and motivation (Lee & Baek, 2023). In other words, gamification may serve not only to motivate learners, but also as a form of structural scaffold for autonomous and strategic engagement with the content.

2.3. Duolingo and AI-Enhanced Gamification

AI in Duolingo shows how gamification features are not mere surface details, but tightly integrated with principles of adaptive instruction. Key to the Duolingo AI model is a machine learning algorithm known as Birdbrain, which uses measures of accuracy, speed, and retention to dynamically predict and select exercises and difficulty levels for each learner (Bicknell et al., 2023). For instance, when the Birdbrain algorithm recognizes high decay on a given knowledge item, it will shuffle that item within the review queue or back-track its sequencing to re-optimise learning efficiency. XP reward and leveling up are also personalized by this algorithm (Shevchenko, 2023). Through this adaptive scaffolding, AI does not merely present content; it is a means of operationalizing and reinforcing the gamified feedback loops.

AI-enhanced gamification may be especially relevant to EFL learning in the Vietnamese context. In Vietnam, CALL has been unevenly incorporated into formal language instruction. While CALL has been adopted by many Vietnamese language teachers, teachers also represent the primary digital divide (Nguyen, 2019). Instructors continue to integrate digital tools to support formal, whole-class teaching, and CALL, with AI personalization, is still a relatively recent and unfamiliar approach for many (Lê et al., 2023). In addition, much of traditional EFL instruction in Vietnam has emphasized grammar-translation approaches, memorization, and high-stakes testing. By contrast, this leaves much less space for frequent, immediate feedback or adaptive support. Finally, teacher shortages, especially in rural and lower-resourced universities, can also leave learners with insufficient access to language input and feedback. In this environment, the AI-enhanced gamification of Duolingo could address a critical need for individualized feedback and opportunities for independent practice outside of the formal classroom.

2.4. Learning Motivation in Gamification with AI

Gamification affordances in Duolingo may be especially relevant to specific gaps in the Vietnamese EFL environment. For instance, streaks are reinforced by Birdbrain through individualized reminders at a time and frequency patterned to each learner (Le et al., 2025). XP progress graphs and level-ups also create scaffolding for adaptive reinforcement, balancing persistence with accuracy. Decay timers use the Birdbrain probability model to forecast when a content item will no longer be retained, to remind the learner to review before it is too late. Leaderboards can also be individualized or normalized by AI if there is a large discrepancy in performance (Bai & Hew, 2025). In all of these ways, AI enables gamification affordances to work responsively, as a form of contingent reinforcement that adapts to the learner's behavior.

In addition, gamification in Duolingo, when paired with adaptive AI, can alter engagement and motivation. Loewen et al. (2019) argue that gamified apps must provide a high density of engagement, or sustained, intense interaction with the target language. AI-enhanced gamification features promote learners to continue micro-interactions many times a day, which are encouraged to aggregate into a long-term, effective practice routine. Gamified affordances may also trigger what Kapp (2012) calls habit-forming mechanics. In other words, small daily goals reinforce the structure and purpose of larger motivation. In MALL, gamified habit-forming features have been documented to promote engagement (Sakkir & Syamsuddin, 2023).

2.5. Current Studies on Gamification in Language Learning

The existing body of empirical literature on gamification in L2 learning and teaching has begun to show both promise and caution. On the one hand, an increasing number of studies have shown that gamified language learning features were indeed associated with learners' increased levels of engagement, persistence, and evidence of learning gains (Zhang & Hasim, 2023; Loewen et al., 2019; Vesselinov & Grego, 2012). For instance, Zhang & Hasim (2023) surveyed Duolingo users about their experience and self-reported improvements with the app and found that gamified features such as points, streaks, XP, and progress graphs helped improve the users' self-regulated learning and supported their long-term persistence with the tool. Loewen et al. (2019) randomly assigned participants to two groups (Duolingo users vs. non-users) for 12 weeks and found that after controlling for pre-test differences and variables such as gender and L1, the Duolingo group significantly outperformed the non-Duolingo group on vocabulary and sentence translation tasks, with an effect size of $d = 0.31$ for the vocabulary learning task. However, they also showed that in the Duolingo group, the regular, daily users showed stronger improvement in their performance than the irregular users did, suggesting that the frequency of engagement, rather than gamification per se, was the most important factor. In addition to this, Vesselinov and Grego (2012) demonstrated that the AI-powered techniques of Duolingo, such as spaced repetition and cumulative review, helped beginner language learners to acquire vocabulary. The researchers also

calculated an estimate for the growth of active vocabulary based on a comparison of XP-per-minute data and the rate at which the users completed the exercises on the app, and found a positive correlation between the two variables.

On the other hand, other studies also cautioned against potential issues with gamification, such as the lack of motivation for learners if they were to focus more on "collecting" the badges, points, and ranks rather than engaging with the language learning task (Hanus & Fox, 2015). Similar findings were also reported by Landers (2015), who found that leaderboards only had a positive impact on some learners' motivation (those who perceived themselves as lower in rank and in skill level), and the opposite effect for other learners (those who ranked higher and thus may have felt a loss of motivation). The current empirical literature on gamification in language learning, therefore, suggests that gamification, when combined with other AI elements such as adaptive feedback and SRS, can indeed support language learning. However, it can do so under certain conditions, with certain subgroups of learners, and not automatically. The design of the gamified features and the motivation balance may play an important role.

In the Vietnamese context, it is important to consider how these engagement drivers can be transferred and scaled. On the one hand, motivated students at elite, metropolitan universities such as the one where this study takes place have relatively high access to hardware and software, including stable internet connections, recent smartphones or tablets, and even previous experience with learning apps (Vo & Nguyen, 2021). These resources allow learners to take advantage of gamified features and AI personalization more consistently. However, other Vietnamese university contexts, in lower-tier provinces and cities, may also include students with unreliable internet connections, older or shared devices, or minimal exposure to educational technology (Dang et al., 2012). Social and cultural expectations also play a role, for Vietnamese university students raised in an exam-oriented system may not yet perceive the value of gamified and formative language learning (Phuong, 2020), unless this is clearly and explicitly aligned with summative testing. This suggests that the primary motivations for gamification, autonomy, competence, and relatedness may still hold, but the digital, social, and educational ecology influences both access and orientation.

Duolingo shows how AI-enabled gamification can translate the psychological tenets, machine learning, and learner motivation into a functional learning model. Its relevance for Vietnam's higher education system is that it can potentially both augment formal instruction and democratize access to adaptive and feedback-rich language learning. However, to deploy AI-enhanced gamification features for improving reading and comprehension, there must be an understanding of related contextual factors, including constraints on digital literacy or cultural expectations of summative assessment. AI-enhanced gamification has therefore shown how

gamified language learning can be, but further research is necessary to understand its effectiveness and transferability across different learners and learning environments.

3. Methodology

3.1. Participants

This study was conducted with a sample of 35 undergraduate students from English as a Foreign Language (EFL) programs at a Vietnamese university located in Ho Chi Minh City, Vietnam. The participants were recruited from multiple classes, ensuring greater diversity in their academic backgrounds and learning environments. As all of them had a baseline level of exposure to the application, this shared familiarity allowed for a more consistent examination of how gamified and adaptive features influenced their motivation and persistence, reducing variability caused by initial unfamiliarity with the platform. Before participation, they were informed of the purpose, procedures, and ethical requirements and provided a signed informed consent form.

3.2. Intervention

The study was based on the convergent mixed-methods design (Creswell & Clark, 2017). Quantitative and qualitative data were collected and analyzed simultaneously. All participants were required to use Duolingo for 10 weeks. The intervention primarily involved tasks such as sentence translation, short comprehension tasks, and in-context vocabulary exercises. AI-driven adaptivity on Duolingo was enabled throughout the process, dynamically adjusting the difficulty and sequencing of lessons based on learners' performance history, errors, and engagement patterns (Zeng & Fisher, 2023). Duolingo's gamified features, including streak counters, badges, leaderboards, and experience points (XP) accumulation, remained fully enabled, as they have been found to impact MALL motivation (Loewen et al., 2019).

3.3. Instruments

Questionnaire. All participants were asked to fill in two sections of 25 items with 5-point Likert scales (1=strongly disagree; 2=disagree; 3=undecided; 4=agree; 5 = strongly agree). Section A of the questionnaire aimed to measure the participants' motivation for independent learning, while Section B, with 12 items, focused on measuring the gamification impact on persistence. A range of intrinsic (e.g., "A4 - I enjoy using Duolingo because it allows me to learn at my own pace.", "A7 - Duolingo's flexibility motivates me to take ownership of my learning process.", "A9 - I enjoy using Duolingo for the sake of learning, rather than just for rewards or progress.") and extrinsic (e.g., "B1 - Duolingo's rewards (e.g., streaks, badges) encourage me to keep learning.", "B3 - The leaderboards on Duolingo motivate me to continue learning in order to stay competitive.") factors were included in the two instruments.

Semi-Structured Interview. To provide richer, qualitative insights, 10 students were also invited to take part in semi-structured interviews after completing the questionnaires. The

interviews were conducted in Vietnamese to maintain contextual, linguistic, and cultural sensitivity, audio-recorded (with consent), transcribed, and then translated into English for the analysis.

The interview protocol focused on three central themes:

1. Perceived motivational value of gamified features
2. The role of AI-driven adaptivity and feedback loops in sustaining persistence
3. Learner autonomy in managing out-of-class practice

Qualitative data were analyzed using thematic analysis following the procedures outlined by Braun and Clarke (2006). An intercoder reliability process was conducted on a subset of transcripts to ensure consistency in theme identification, with any discrepancies resolved through discussion and consensus.

3.4 Validity and Reliability

Validity of the questionnaire. The questionnaire was grounded in the SDT constructs of autonomy, competence, and relatedness, along with the TAM constructs of perceived usefulness and ease of use. These established constructs, together with two field experts' review, ensure strong content validity. Moreover, the alignment with established MALL research facilitates the instrument to accurately measure both intrinsic and extrinsic factors relevant to Duolingo's gamified environment.

Reliability Analysis. Both questionnaires demonstrated high internal consistency, with Cronbach's alpha for the Motivation for Independent Learning scale being .868 (13 items), and for the Impact of Gamification on Persistence scale being .832 (12 items). Moreover, the combined 25-item scale exceeds the threshold of .70 for acceptable reliability (Taber, 2018) with an overall alpha of .894. This indicates that the items reliably measured the intended constructs of independent learning motivation and gamification-driven persistence.

Reliability Statistics

	Cronbach's Alpha	N of Items
Questionnaire: Motivation for Independent Learning	.868	13

Questionnaire: Impact of Gamification on Persistence	.832	12
Both Questionnaires	.894	25

***Table 1**: Reliability Statistics of the Questionnaires*

3.5 Ethical Considerations

Ethical approval of this study was obtained from the Head of Faculty prior to data collection, and all participants were fully informed and signed a consent form announcing the study's purpose, procedures, and their right to withdraw at any time without consequences. Personal information and responses were kept confidential and anonymized to ensure privacy and data protection.

4. Findings

4.1. Findings from Questionnaires

The questionnaires revealed generally high levels of motivation and persistence among the 35 EFL students at the B1 level of English proficiency, all of whom had been using Duolingo as a supplementary study material for at least 10 weeks, and participated in the study.

Independent Learning Motivation (A1-A13). Responses indicated relatively high agreement with items on positive affect regarding autonomy and self-regulated learning. The highest mean scores were recorded for flexibility and learner-driven engagement (A6: $M = 4.54$; A8: $M = 4.51$), and pacing (A4: $M = 4.49$). Enjoyment of learning (A10: $M = 4.17$) and intrinsic motivation (A9: $M = 4.20$) were also rated positively. Meanwhile, self-efficacy (A2: $M = 3.80$) and daily practice encouragement (A3: $M = 3.86$) illustrated a relatively high score yet slightly lower agreement. While these scores suggest that Duolingo supported independent learning motivation, there are still varied perceptions toward confidence and habit strength.

Gamification and Persistence (B1-B12). Items most strongly related to streaks and other reward-based mechanisms (e.g., B1: $M = 4.37$; B2: $M = 4.03$; B7: $M = 4.00$) were especially strongly endorsed. Leaderboard items (B3: $M = 3.71$; B9: $M = 3.69$) generated more mixed responses, with some students being more and others less motivated by the feature. In general, the quantitative results suggest that streaks and XP gain are more powerful, consistent, and widely-acting motivational mechanisms than leaderboards. This supports previous research, which has

similarly found that streaks in particular have widespread appeal and effectiveness as gamified motivators, in comparison to leaderboards (Loewen et al., 2019).

Section A: Motivation for Independent Learning (Use a Likert scale: 1 = Strongly Disagree to 5 = Strongly Agree)	Mean	SD
A1 - Duolingo has made me more motivated to learn languages outside of my scheduled study sessions.	4.03	.785
A2 - I feel more confident in my ability to learn a language independently because of Duolingo.	3.80	1.023
A3 - Duolingo encourages me to practice my language skills daily, even when not required.	3.86	.845
A4 - I enjoy using Duolingo because it allows me to learn at my own pace.	4.49	.702
A5 - Using Duolingo feels like a personalized learning experience.	4.37	.598
A6 - I am more engaged in learning because Duolingo lets me choose when and what I study.	4.54	.611
A7 - Duolingo's flexibility motivates me to take ownership of my learning process.	4.23	.808
A8 - I prefer learning on Duolingo because I can practice whenever I want, without relying on others.	4.51	.612
A9 - I enjoy using Duolingo for the sake of learning, rather than just for rewards or progress.	4.20	.797
A10 - Duolingo makes language learning feel like a fun activity.	4.17	.857

A11 - I feel personally satisfied when I complete lessons or achieve milestones on Duolingo.	4.29	.750
A12 - Duolingo motivates me to learn because I can track my progress and achievements.	4.26	.701
A13 - I feel motivated by the feedback and rewards I receive from Duolingo.	4.03	.923
Section B: Impact of Gamification on Persistence (Use a Likert scale: 1 = Strongly Disagree to 5 = Strongly Agree)	Mean	SD
B1 - Duolingo's rewards (e.g., streaks, badges) encourage me to keep learning.	4.37	.646
B2 - I feel motivated to maintain my streak on Duolingo even when I don't feel like practicing.	4.03	.822
B3 - The leaderboards on Duolingo motivate me to continue learning in order to stay competitive.	3.71	.987
B4 - I am more persistent with learning when I receive feedback (e.g., streaks, rewards) on Duolingo.	3.80	1.079
B5 - Seeing my progress in Duolingo (e.g., levels, XP) motivates me to keep using the app regularly.	3.91	.702
B6 - I find myself returning to Duolingo regularly to maintain my streaks.	4.00	.874
B7 - The challenge of maintaining a streak helps me stay consistent with my language learning.	4.00	.874
B8 - I often push myself to complete lessons on Duolingo to earn rewards or points.	4.00	.939

B9 - If I didn't have the streak system or leaderboards, I would be less likely to continue using Duolingo consistently.	3.69	1.132
B10 - Duolingo's gamification elements make language learning feel more rewarding.	4.00	.874
B11 - I am more likely to use Duolingo consistently because of the sense of accomplishment from rewards.	3.80	.759
B12 - The gamified features of Duolingo increase my persistence in language learning, especially during difficult lessons.	3.97	.822

***Table 2**: Mean Total Score and Standard Deviation of the Questionnaires' Results*

Overall, the quantitative findings indicate that streaks and flexible scheduling were the strongest motivational factors, which support persistence through habit-building and a sense of ownership. Leaderboards functioned more as a situational motivator, for they are appealing to competitive learners, but they can not be considered an overarching mechanism. This aligns with previous research emphasizing the broad appeal of streak systems over competitive leaderboards in mobile-assisted language learning contexts (Loewen et al., 2019).

4.2. Findings from Interviews

The semi-structured interviews provided a richer account of the learners' experiences and perceptions of Duolingo's gamification and adaptivity. Thematic analysis (Braun & Clarke, 2006) of the 10 interviewees' responses identified four interrelated themes that help to understand the nature of the mechanisms of motivation, persistence, and self-regulation in MALL.

Streaks as a Habit-Forming and Motivational Cue. Streaks are the most prevalent and potent of all the features, serving as a daily behavioral trigger to ensure engagement. The learners often described the task of maintaining a streak in habitual terms as a part of their daily personal hygiene routine:

"When I had my 50-day streak, I felt proud and just wanted to keep going. It became part of my day-almost like brushing my teeth." (Interviewee 6)

"Sometimes I opened Duolingo only because I didn't want to break my streak, even if I wasn't really in the mood to study." (Interviewee 3)

At the same time, loss of streak often caused a negative affective reaction and subsequently reduced learners' engagement:

"I felt like I lost a close friend. I was so disappointed. And, I think after that, I skipped a few days." (Interviewee 4)

This indicates that streaks are a robust mechanism to induce habit formation, but also a fragile one, as users may be easily discouraged by the sense of loss.

Intrinsic versus Extrinsic Motivation. Interviews also suggest an interplay between intrinsic and extrinsic motivational factors. Participants frequently referred to enjoyment of content, perceptions of progress, and self-regulation as their main motivations:

"I like Duolingo because it's fun. The lessons are short, and it doesn't feel like serious studying." (Interviewee 8)

"For me, the content itself keeps me going, not just the rewards." (Interviewee 4)

At the same time, many learners identified streaks and XP-related features as facilitating or supporting their motivation, and particularly valuable in cases of low or absent intrinsic drive:

"Without the streak feature, I probably wouldn't have kept going. It gave me that extra push." (Interviewee 1)

"Even if I only learned one thing, gaining XP made me feel like I achieved something." (Interviewee 10)

This suggests that the app's extrinsic rewards do not necessarily crowd out intrinsic motivation but provide scaffolding to maintain it until the intrinsic motivation to learn the target language develops to a sufficient level, as would be predicted by self-determination theory's account of extrinsic-to-intrinsic internalization of extrinsic regulation (Deci & Ryan, 2017).

Leaderboards as a Source of Social Comparison. Leaderboards, on the other hand, appear to be a more variable and idiosyncratic motivator. Some learners described this feature as a source of competition, pushing them to increased activity:

"I studied five days in a row when I finally got to the top 3. I didn't want to lose my spot." (Interviewee 7)

Yet, several learners downplayed or completely dismissed this feature:

"I rarely check the leaderboard. It doesn't affect how I study." (Interviewee 5)

"I don't know anyone on the leaderboard, so it doesn't really matter to me. I am not interested in competing with other learners." (Interviewee 8)

This suggests that the leaderboard feature may only apply to a subset of learners with a competitive orientation and may be ineffective or even distracting for the rest of the users.

Learners' Autonomy. Finally, autonomy, flexibility, and freedom to choose were mentioned several times by the learners.

"I can learn whenever I want. I don't need to be reminded to do it." (Interviewee 2)

"Sometimes I'm on the bus. Sometimes I do it before bed. It's easy to fit into my schedule." (Interviewee 9)

Push notifications are framed as occasional reminders but not direct or controlling cues. These responses suggest that learners mostly perceive their learning as self-regulated, which is consistent with prior research indicating that the perception of autonomy in a mobile language learning environment can lead to greater persistence in language learning (Hui et al., 2023).

5. Discussion and Conclusion

This mixed-method study aimed to understand how gamification and AI-based adaptivity influenced the motivation and persistence of 35 Vietnamese EFL learners on Duolingo. The convergent mixed-methods design (Creswell & Clark, 2017) indicated that streaks, lesson enjoyment, and perceived progress were the main engagement triggers. Streaks had a habit-forming power and encouraged learners to establish daily practice routines and peer-to-peer accountability. However, losing streaks led to short-term demotivation, which indicated the risk of relying on a single extrinsic trigger. Leaderboards had a variable effect while badges had a minimal effect on learners, which may be related to individual differences in extrinsic versus intrinsic motivation. The findings consolidate the foundation from previous studies regarding motivation stemming from streaks and the self-controlled learning pace, while the impact of leaderboards and badges illustrated a more complicated overall picture with diverse tendencies among learners.

This study also emphasizes the duality of gamification, in which, while extrinsic factors such as streaks initially acted as strong behavioral triggers, they need to be further complemented by intrinsic factors, including perceived competence and autonomy. Extrinsic factors also have different effects on learners with different tendencies toward competitiveness. Moreover, the interaction between gamification and AI-based adaptability should be put into closer consideration, for its effectiveness largely depends on learners' perceptions. While some participants gravitate toward the sense of autonomy it provides, others feel it still cannot adapt to the proficiency level they are aiming at. Therefore, the design of adaptive features can be considered equally important in exerting a positive influence and supporting learners to acquire the language effectively.

The pedagogical implications suggested the possibility of using gamification and adaptive AI to facilitate learner persistence in MALL contexts. However, the results also warned about a unidimensional gamification and AI approach: a more comprehensive and balanced use of both

extrinsic triggers and intrinsic motivators, such as autonomy, meaningfulness, and self-regulated learning (Deci & Ryan, 2017), is recommended for longer-term engagement. For this reason, developers should consider tailoring or disabling some features for learners who are less responsive to them. For Vietnamese teachers, this study contributes to the progress of seamlessly integrating new technology into teaching and learning activities through understanding students' perceptions of these tools. The effectiveness of Duolingo and its gamified features on students can also act as a prototype for other possible tools to be developed and utilised in the Vietnamese EFL context.

The main limitations of this study are the small sample size (n = 35) and the short-to-middle time span required in engaging with the app (ten weeks). Participants were drawn from multiple classes, yet the sample was still limited to a specific urban context in Vietnam. Additionally, the study relied on self-reported quantitative and qualitative data, which may be subject to response bias or overestimation of engagement. Interviews were conducted in Vietnamese for the sake of data fidelity, but translations to English may have led to minor semantic errors.

It is possible to suggest some ideas for future research, such as understanding the duration of motivated engagement when the novelty effect wears off, the process of transforming extrinsic components like streaks into intrinsic habits, and the adaptive personalization of gamification (e.g., private leaderboards or streak recovery methods). The use of larger sample sizes, multi-institutional settings, and quasi-experimental designs could improve the generalizability and validity of the findings, as well as the interaction effects between adaptive AI and learners' metacognitive skills.

6. Implications

It is possible to suggest some ideas for future research, such as understanding the duration of motivated engagement when the novelty effect wears off, the process of transforming extrinsic components like streaks into intrinsic habits, and the adaptive personalization of gamification (e.g., private leaderboards or streak recovery methods). The use of larger sample sizes, multi-institutional settings, and quasi-experimental designs could improve the generalizability and validity of the findings, as well as the interaction effects between adaptive AI and learners' metacognitive skills.

References

Al-Emran, M., Mezhuyev, V., & Kamaludin, A. (2018). Technology acceptance model in M-learning context: A systematic review. *Computers & Education*, *125*, 389–412. https://doi.org/10.1016/j.compedu.2018.06.008

Bai, S., & Hew, K. F. (2025). How different gamified leaderboards affect individual students' learning engagement, strategies, performance, and perceptions in online classes. *Educational Technology & Society*, *28*(4), 42-61. https://doi.org/10.30191/ETS.202510_28(4).RP03

Bicknell, K., Brust, C., & Settles, B. (2023). How Duolingo's AI learns what you need to learn: The language-learning app tries to emulate a great human tutor. *IEEE Spectrum*, *60*(3), 28–33. https://doi.org/10.1109/mspec.2023.10061631

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL, 27*(1), 4-20. https://doi.org/10.1017/S0958344014000159

Chen, C., Wang, K., & Lin, Y. (2015). The comparison of solitary and collaborative modes of game-based learning on students' science learning and motivation. *J. Educ. Technol. Soc.*, *18*, 237-248.

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment - Companion volume*. Council of Europe Publishing.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. SAGE Publications.

Dang, X. T., Nicholas, H., & Lewis, R. (2012). Key barriers to ICT uptake in foreign language teaching in Vietnam : A case study. *INTED2012 Proceedings*, 4266–4274.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Deci, E. L., & Ryan, R. M. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. ACM. https://doi.org/10.1145/2181037.2181040

Ha, T. B. L. (2024). Exploring self-determination theory in ELT: A content analysis of research trends between 2020 and 2024. *ICTE Conference Proceedings*, *5*, 140–164. https://doi.org/10.54855/ictep.24512

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?—A literature review of empirical studies on gamification. *47th Hawaii International Conference on System Sciences*, 3025-3034. https://doi.org/10.1109/HICSS.2014.377

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education, 80,* 152-161. https://doi.org/10.1016/j.compedu.2014.08.019

Hui, X., Liu, Z., & Chi, Y. (2023). A review on learner autonomy with mobile-assisted language learning for EFL learners at the tertiary level. *OALib*, *10*(09), 1–16. https://doi.org/10.4236/oalib.1110606

Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. San Francisco, CA: Pfeiffer.

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL, 20*(3), 271-289. https://doi.org/10.1017/S0958344008000335

Landers, R. N. (2015). Developing a theory of gamified learning. *Simulation and Gaming*, *45*(6), 752–768. https://doi.org/10.1177/1046878114563660

Le, A. V., Luong, M. P., Do, Đ. L., Tran, M. N., & Bui, T. D. (2023). Technology in education: A case study on Viet Nam. *UNESCO: Global Education Monitoring Report* https://doi.org/10.54676/GPDF6007

Le, C. T., Nga, N. H. H., Tuan, L. C., & Quyen, B. N. (2025).Gamification in independent learning: Vietnamese university students' experience with Duolingo. *LRR 2025, 16*(1), 249-269

Lee, J., & Baek, M. (2023). Effects of gamification on students' English language proficiency: A meta-analysis on research in South Korea. *Sustainability*, *15*(14), 11325. https://doi.org/10.3390/su151411325

Loewen, S., Crowther, D., Isbell, D., Kim, K. M., Maloney, J., Miller, Z. F., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL, 31*(3), 293-311. https://doi.org/10.1017/S0958344019000065

Nguyen, N. V. (2019). Mobile learning in language teaching context of Vietnam: An evaluation of students' readiness. *HCMUE Journal of Science*, *7*(85), 16. https://doi.org/10.54607/hcmue.js.0.7(85).431(2016)

Phuong, H. T. T. (2020). Gamified learning: Are Vietnamese EFL learners ready yet? *International Journal of Emerging Technologies in Learning (iJET)*, *15*(24), 242. https://doi.org/10.3991/ijet.v15i24.16667

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review, 32*(1), 77-112. https://doi.org/10.1007/s10648-019-09498-w

Sakkir, G., & Syamsuddin, N. A. (2023). Students' perceptions of Duolingo mobile assisted language learning (MALL) in learning English vocabulary. *EduLine Journal of Education and Learning Innovation*, *3*(3), 381–388. https://doi.org/10.35877/454ri.eduline1970

Settles, B., & Meeder, B. (2016). A trainable spaced repetition model for language learning. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 1*, 1848-1858. https://doi.org/10.18653/v1/P16-1174

Settles, B., Brust, C., Gustafson, E., Hagiwara, M., & Madnani, N. (2020). Second language acquisition modeling. *Proceedings of the 58th annual meeting of the Association for Computational Linguistics* (pp. 134-145). ACL. https://doi.org/10.18653/v1/W18-0506

Shevchenko, O. (2024). The use of AI technology in language learning. *European Science*, (sge29-03), 233-240. https://doi.org/10.30890/2709-2313.2024-29-00-07

Taber, K. S. (2017). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Vesselinov, R., & Grego, J. (2012). Duolingo effectiveness study. Final report for City University of New York. Retrieved from https://theowlapp.health/wp-content/uploads/2022/04/DuolingoReport_Final-1.pdf

Vo, T. L., & Nguyen, N. V. (2021). *The impact of mobile learning on EFL students' learning behaviors and perceptions: From content delivery to blended interaction. International Research in Higher Education, 5*(4), 25. https://doi.org/10.5430/irhe.v5n4p25

Zeng, C., & Fisher, L. (2023). Opening the "Black box": How out-of-class use of Duolingo impacts Chinese junior high school students' intrinsic motivation for English. *ECNU Review of Education*, *7*(2), 283–307. https://doi.org/10.1177/20965311231171606

TRANSLATING EMOTIONS THROUGH METAPHOR: CROSS-CULTURAL INSIGHTS FROM TRUYỆN KIỀU INTO ENGLISH AND JAPANESE CLASSROOMS

Nguyễn Minh Hưng[1]

Abstract

Emotions are universal, but the way they are conceptualized and expressed through metaphor varies significantly across languages and cultures. This study investigates how emotional metaphors embedded in the Vietnamese literary classic Truyện Kiều are rendered in English and Japanese translations and explores how these metaphors can be meaningfully integrated into language classrooms to promote intercultural understanding and sustainable language learning. Grounded in Conceptual Metaphor Theory, the research adopts a qualitative, text-based comparative approach, identifying metaphorical expressions of core emotions such as love, sorrow, despair, and longing in the Vietnamese text and examining their translated counterparts in English and Japanese. This allows for the analysis of how emotional imagery is preserved, adapted, or reshaped across linguistic and cultural boundaries. Findings reveal not only linguistic and conceptual shifts but also divergent cultural attitudes toward emotion, which carry implications for both translation and language pedagogy. The study argues that teaching emotional metaphors through literary texts offers a powerful, culturally rich pathway for developing metaphor awareness, empathy, and deeper interpretive skills among learners. By bridging literary analysis and classroom application, the research contributes to both metaphor studies and innovative, sustainability-oriented language education.

Key words: Emotional Metaphors, Conceptual Metaphor Theory, Cross-cultural Translation, Sustainable Language Education

1. Introduction

Literature serves as both a mirror of cultural values and a reservoir of emotional expression encoded in metaphor. This is especially evident in *Truyện Kiều* (*The Tale of Kiều*), Nguyễn Du's poetic masterpiece, which portrays love, sacrifice, sorrow, longing, and moral struggle. Beyond narrating tragedy, *Truyện Kiều* encapsulates Vietnamese philosophical, ethical, and emotional worldviews, often articulated through dense metaphor.

Metaphor in cognitive linguistics is not decorative but central to structuring thought and emotion. Lakoff and Johnson (1980) showed that metaphor permeates human conceptual systems, while Kövecses (2015) demonstrated how emotional experience is metaphorically structured and culturally mediated. Though emotions are biologically universal, they are expressed differently

[1] ThS, đang công tác tại Trường Đại học Quốc tế Hồng Bàng, *Email: hungnm@hiu.vn*

across languages and cultures, reflecting ideologies and values (Yu, 2021). In literature, such metaphors challenge translation, particularly between distant cultures.

In pedagogy, metaphor enhances conceptual understanding, emotional literacy, and intercultural competence (Littlemore & Low, 2020). Yet few studies explore the pedagogical potential of metaphor-rich texts like *Truyện Kiều* for fostering empathy and metaphor awareness. This gap is important given current educational goals for sustainable development that stress empathy, cultural understanding, and deep learning (UNESCO, 2021).

This study uses *Truyện Kiều* as a case study to examine how emotional metaphors are translated across Vietnamese, English, and Japanese. The English version by Huỳnh Sanh Thông and the Japanese by Satō Seiji are the primary translations. Focusing on two excerpts—sacrifice for filial piety and existential despair—the study identifies core metaphorical structures and analyzes their preservation, alteration, or loss. Shifts are interpreted through linguistic and cultural-semiotic analysis, informed by frame semantics (Neumair et al., 2025).

The central argument is that emotional metaphors, when analyzed through translation and cultural discourse, are powerful pedagogical tools. They provide learners with insights into how cultures conceptualize pain, loyalty, love, and reconciliation. Teaching with literary metaphors fosters deeper, sustainable language learning grounded in empathy and intercultural dialogue (Littlemore & Low, 2020; UNESCO, 2021).

The study has three aims:

(1) to identify and analyze emotional metaphors in selected passages and translations.

(2) to examine shifts in metaphor translation, focusing on emotional intensity and conceptual framing; and

(3) to propose pedagogical applications for integrating these metaphors into English and Japanese classrooms to build intercultural and emotional literacy.

Research questions include:

(1) How are emotional metaphors conceptualized in Truyện Kiều and rendered in English and Japanese?

(2) What patterns of preservation, transformation, or loss occur?

(3) How can these translated metaphors be used in teaching to enhance awareness and competence?

By linking literary analysis, metaphor theory, and pedagogy, this study contributes to cognitive metaphor studies, comparative translation research, and sustainable language education.

It treats metaphor not as curiosity but as a vital channel for emotional expression and cultural understanding.

2. Literature Review and Theoretical Framework

2.1 Conceptual Metaphor Theory (CMT)

Conceptual Metaphor Theory (CMT), first proposed by Lakoff and Johnson (1980), provides a framework for understanding how abstract concepts, including emotions, are structured through metaphor. Rather than decorative, CMT views metaphor as a cognitive process mapping meaning from concrete domains (e.g., space, motion, temperature) onto abstract domains (e.g., love, grief, despair). Such mappings shape language, perception, reasoning, and emotional experience.

Building on this, Kövecses (2015) highlights the role of cultural context in shaping metaphorical systems. While metaphors such as "LOVE IS A JOURNEY" or "ANGER IS HEAT" may seem universal, their realizations, salience, and functions vary across cultures. These differences stem from emotion ideologies, norms, and embodied experiences, reinforcing the view that emotional metaphors are cognitively grounded yet culturally mediated. This perspective is especially relevant for comparative work on Vietnamese, English, and Japanese, where moral values, hierarchies, and expressive norms produce distinct framings of filial piety, longing, or despair.

Neumair et al. (2025) advance the theory with a frame-semantic approach, examining how metaphors are embedded in shared cultural knowledge and discursive frames. This is valuable for literary texts, where metaphor functions as cognitive tool, narrative device, and cultural symbol.

Despite these advances, few studies apply CMT to classical poetry across multiple languages. Most research focuses on everyday or political discourse, overlooking the metaphorical density of *Truyện Kiều*. Limited scholarship has explored how translated metaphors from such works can support education. This study synthesizes insights from Kövecses (2015) and Neumair et al. (2025) into a hybrid framework addressing both conceptual grounding and discursive realization of emotional metaphors in *Truyện Kiều* and its English and Japanese translations.

2.2 Emotion and Culture

Emotions, though biologically rooted, are shaped by cultural scripts and ideologies. In cognitive linguistics and cultural psychology, it is recognized that ways of conceptualizing and evaluating emotion vary (Yu, 2021). Vietnamese culture ties emotions such as filial piety, suffering, and sacrificial love to Confucian expectations. Japanese expression, influenced by *honne* (true feelings) and *tatemae* (social façade), values restraint and symbolism. Anglo-Western cultures emphasize transparency and individualism, framing emotions like love or grief as openly articulated personal states. These ideologies shape both behavior and metaphorical encoding (Kövecses, 2015).

Cultural values also determine metaphorical source domains. Sorrow in Vietnamese and Japanese is linked to inward collapse, shadows, or silence, while English prefers external images such as storms or explosions. Emotional intensity differs: Japanese often implies, English dramatizes, and Vietnamese moralizes. Such contrasts are critical in literary translation, where metaphor acts as a cultural carrier of worldview.

Yet few studies address how emotional metaphors travel across languages in literary translation. Much research on emotion and culture emphasizes spoken discourse or lexical comparison. This study fills that gap by analyzing how emotional metaphors in *Truyện Kiều* are restructured in English and Japanese, and how these reflect or overwrite original emotion scripts.

2.3 Literary Translation and Metaphorical Transfer

Literary translation, particularly of poetic texts, poses challenges beyond the lexical or syntactic level, requiring navigation of cultural and emotional resonances. In Truyện Kiều, these difficulties are intensified by extensive metaphor use, especially in conveying emotional experience. This section examines how emotional metaphors are translated across Vietnamese, English, and Japanese.

Translation scholars such as Lefevere (1992) and Bassnett (2014) argue that literary translation involves negotiation with ideology, poetics, and culture rather than simple substitution. Translators must decide whether to preserve the source culture's emotional worldview or adapt imagery for resonance in the target culture. This is critical in metaphor translation, where cognitive and cultural mappings often diverge.

Conceptual Metaphor Theory (Lakoff & Johnson, 1980) shows that metaphors reflect conceptual structures shaped by cultural and bodily experience. Schäffner (2004) notes that metaphor translation may preserve original imagery or adapt it with a counterpart, a process known as metaphorical transfer. This involves three strategies: metaphor retention, substitution, or omission (Al-Hasnawi, 2007).

Domestication and foreignization, strategies discussed by Venuti (1995), also affect metaphor translation. Domestication replaces culturally specific metaphors with familiar constructs, increasing readability but reducing cultural specificity. Foreignization retains source expressions, fostering cultural awareness but risking alienation.

These tensions are evident in *Truyện Kiều*, where emotional metaphors carry Confucian, Buddhist, and folk values. Metaphors of karmic suffering, filial devotion, or romantic fatalism are deeply embedded in Vietnamese culture yet lack direct equivalents in English or Japanese.

In this study, literary translation serves to explore both linguistic transfer and cross-cultural interpretation of emotion. By comparing metaphorical shifts in English and Japanese versions, we

see how translators navigate emotional meaning, with implications for pedagogy in enhancing intercultural sensitivity and metaphorical competence.

2.4. Metaphor and Language Education

Metaphor plays a pivotal role not only in shaping how emotions are linguistically encoded but also in how they are cognitively processed and pedagogically applied in second language education. As Littlemore & Low (2006) note, metaphor awareness enhances learners' interpretive depth and their ability to navigate abstract and affective meaning. When trained to recognize metaphorical structures, learners better decode nuanced language, develop empathy, and process intercultural communication more deeply (Littlemore & Low, 2006; Yu, 2011).

Recent applied linguistics research highlights the pedagogical value of metaphor. Boers (2021) argues that metaphor-based instruction fosters conceptual fluency by linking unfamiliar expressions to embodied experience or culturally embedded schemata. Teaching metaphors such as "LOVE IS FIRE" or "SADNESS IS DOWN" not only aids vocabulary retention but also reveals cultural conceptualizations that differ across English, Japanese, and Vietnamese (Deignan, 2005; Kövecses, 2020).

Incorporating literary metaphors into instruction further cultivates linguistic sensitivity and intercultural understanding. Kramsch (2009) views literature as a "third space" where learners encounter otherness and negotiate meaning across cultural boundaries. Engaging with emotionally charged metaphors in *Truyện Kiều* allows students not only to learn words but to inhabit emotional worlds shaped by distinct values and repertoires. This reflects Byram's (2021) model of intercultural communicative competence, emphasizing critical cultural awareness and the affective dimension of language use.

Empirical studies also show that metaphorical competence correlates with emotional intelligence and communicative confidence (Jiang et al., 2021). In multilingual classrooms, metaphor-based instruction supports sustainable education by fostering engagement, personal relevance, and retention. In Asian contexts, where affective expression may be constrained, metaphor provides an indirect yet meaningful pathway for exploring identity, emotion, and relationality (Chan, 2022).

Integrating metaphor into language education, especially via culturally rich texts, thus offers cognitive scaffolding and transformative pedagogy. It enables students to explore how emotions are culturally shaped, linguistically represented, and dialogically shared—an essential competence in today's global, emotionally complex world.

2.5. Research Gap

The reviewed literature shows an evolving understanding of emotional metaphor across languages and disciplines. Conceptual Metaphor Theory (Lakoff & Johnson, 1980) laid the

foundation for studying metaphor as a cognitive mechanism shaping linguistic expression and cultural perception. Later work explored cross-linguistic expressions of emotion (Kövecses, 2005; Yu, 2009) and the role of metaphor in shaping learner identity, engagement, and empathy in classrooms (Littlemore & Low, 2006; Boers, 2021).

In literary translation, earlier studies emphasized tensions between faithfulness and functional equivalence (Schäffner, 2004; Mandelblit, 1995). More recent contributions (Vandaele & Baaijen, 2022; Pan, 2023) highlight metaphor translation as intercultural negotiation, especially in emotionally rich texts. Yet the intersection of literary metaphor, emotion, and pedagogy remains underexplored, particularly with Asian canonical literature in modern classrooms.

While Littlemore (2020) and Yu (2021) argue for metaphor awareness as a pedagogical tool, few studies apply this in real classrooms or with historical works like *Truyện Kiều*. The Vietnamese–English–Japanese triangle is especially underrepresented, despite its potential for contrastive analysis and relevance in multilingual education across East and Southeast Asia.

This study addresses this gap by analyzing how emotional metaphors in *Truyện Kiều* are transferred into English and Japanese, and how such translations can be used pedagogically to promote intercultural empathy and deep metaphor processing. By bridging metaphor theory, translation studies, and applied linguistics, it contributes both theoretical advancement and practical value to sustainable language education.

3. Methodology

3.1 Research Design

This study adopts a qualitative, text-based comparative design to explore how emotional metaphors in *The Tale of Kiều* are transformed across three linguistic and cultural systems: Vietnamese, English, and Japanese. By focusing on metaphorical expressions of intense emotional experiences such as filial sacrifice and existential despair, the research seeks to reveal how cultural schemata and metaphorical conceptualizations are preserved, altered, or erased in translation. Rather than relying on numerical data, the qualitative approach enables a context-sensitive, culturally grounded interpretation of metaphor use. This design aligns with previous metaphor and translation studies (e.g., Schäffner & Shuttleworth, 2013; Kövecses, 2020), which emphasize the importance of textual and cultural nuances in metaphor analysis. The aim is to provide deep interpretive insights rather than generalizable statistical claims.

3.2 Textual Data and Selection

This study focuses on two emotionally pivotal segments from *Truyện Kiều* that embody distinct domains of emotional metaphor and cultural value: (1) sacrifice for filial piety and (2) despair and death. These passages were chosen not only for literary significance but also because

they represent contrasting emotional poles—devotion and hopelessness—enabling a nuanced examination of metaphorical expressions across Vietnamese, English, and Japanese.

The first segment depicts Kiều's decision to sell herself to save her father, a culturally charged act tied to Confucian ideals of filial duty and sacrifice. It is rich in metaphors of obligation, virtue, and internalized pain. The second, portraying Kiều's suicide attempt in the Tiền Đường River, presents a radically different domain of despair, loss of agency, and existential suffering—universal emotions yet culturally encoded in divergent ways.

Though limited in scope, these two segments were strategically selected based on three criteria: (1) emotional intensity, (2) richness of metaphorical expression, and (3) centrality to the narrative arc. Together, they provide a representative sample of emotional metaphors, allowing in-depth comparative analysis without sacrificing precision.

This selection also reflects recent metaphor research advocating targeted sampling over exhaustive coverage to achieve interpretive depth and cross-cultural insight (Charteris-Black, 2021; Littlemore & Low, 2020). Future research may extend to other emotional domains such as romantic love or betrayal to broaden metaphorical and cultural inquiry.

3.3 Analytical Procedures

The study adopts a qualitative text-based comparative approach, combining metaphor identification with cultural-pragmatic interpretation to examine how emotional metaphors are conceptualized and translated in Vietnamese, English, and Japanese versions of *Truyện Kiều*.

Phase 1: Metaphor Identification. Using the MIPVU procedure (Steen et al., 2010), metaphors were systematically identified in the original and translations, distinguishing basic from contextual meanings.

Phase 2: Categorization. Identified metaphors were grouped by emotion domains (e.g., pain, hope, obligation) and conceptual structures following Conceptual Metaphor Theory (Lakoff & Johnson, 1980), with attention to cross-linguistic variation.

Phase 3: Cross-Cultural Interpretation. A comparative metaphor matrix visualized patterns across languages. Metaphors were interpreted via cultural scripts (Wierzbicka, 1991), value systems, and emotional pragmatics to reveal translation shifts and cultural transformations.

This multi-step procedure ensures rigor and interpretive depth, providing a framework for analyzing how emotional metaphors are retained, reimagined, or reshaped. It builds on studies of metaphor translation across Asian and Western languages (Yu, 2009; Deignan et al., 2013; Kövecses, 2020).

3.4. Analytical Framework and Coding Procedures

To analyze the metaphors embedded in the selected segments of *Truyện Kiều* and its English and Japanese translations, this study employed a qualitative text-based framework grounded in Conceptual Metaphor Theory (CMT) (Lakoff & Johnson, 1980). The metaphor identification followed the Metaphor Identification Procedure (MIP) developed by the Pragglejaz Group (2007), which systematically distinguishes metaphorical expressions by comparing contextual and more basic meanings of lexical units.

After identifying metaphorical expressions in each version, a three-phase coding process was conducted:

1. Initial coding involved isolating emotionally charged metaphors in each segment, such as those related to pain, love, betrayal, and death.
2. Categorical grouping classified metaphors into conceptual domains (e.g., LOVE IS WAR, LIFE IS A JOURNEY, EMOTION IS A SUBSTANCE), with reference to Lakoff and Kövecses (1987), Charteris-Black (2004), and Yu (2009), to examine whether metaphorical conceptualizations aligned across versions or were transformed.
3. Comparative matrixing facilitated cross-linguistic analysis by aligning source and target metaphors side by side. This matrix allowed for visualizing changes in metaphor structure, image schema, and cultural resonance.

Special attention was given to cultural specificity and emotional intensity, particularly how metaphors were retained, softened, or reconfigured to fit the target linguistic and cultural norms. For instance, metaphors rooted in Confucian filial duty in Vietnamese were traced for shifts in resonance and emotional salience in English and Japanese. Following Boers (2000), metaphors were further evaluated based on cognitive effort and emotional impact.

To ensure analytical rigor, intercoder reliability was cross-checked by consulting a second reviewer with expertise in metaphor analysis and Vietnamese literature in translation. Discrepancies were resolved through negotiated consensus, enhancing both validity and reliability of the findings.

This framework not only highlights metaphorical transformations across translations but also captures the interplay between language, emotion, and culture, in line with the study's broader aim of exploring how culturally embedded metaphors shape reader perception and emotional engagement across linguistic boundaries.

4. Findings and Discussion

4.1 Case Study 1: Sacrifice for Filial Piety

This section applies the theoretical insights and methodological strategies established in Chapters 2 and 3 to conduct a detailed cross-linguistic and cross-cultural analysis of emotional

metaphors embedded in the selected segment about filial piety in *Truyện Kiều*. Through a close comparison of the Vietnamese original and its English and Japanese translations, we explore how emotional meaning is constructed, preserved, or transformed across languages and cultures.

4.1.1 Vietnamese Original: A Metaphorical Layering of Burden and Moral Choice

The excerpt selected from *Truyện Kiều* encapsulates Thúy Kiều's emotional conflict between romantic love and filial obligation. For instance, the line:

"Bên tình bên hiếu bên nào nặng hơn?"

presents the conceptual metaphor "Moral Decision is a Balance Scale", where conflicting duties are metaphorically weighed. This framing suggests that moral emotions (love and filial duty) are quantifiable and subject to internal judgment. Such metaphors are both emotionally charged and culturally embedded in Confucian ideals, where filial piety is often prioritized over personal desire.

In the subsequent line:

"Lấy tình thâm trả nghĩa thâm mới là"

We observe a metaphorical framing of "Emotional Debt and Payment", implying that affection can be repaid through self-sacrifice. This moral economy of affection and gratitude is central to Vietnamese relational values.

Another iconic phrase:

"Bán mình chuộc cha"

compresses a complex emotional and moral act into an economic metaphor. The phrase evokes the metaphor "Life is a Commodity" and "Sacrifice is Payment", which turns Kiều's body into a transactional good. This metaphor dramatizes the extremity of her sacrifice and invokes intense emotional resonance in Vietnamese cultural contexts, where loyalty and sacrifice for parents are revered.

4.1.2 English Version: Conceptual Flattening and Lexical Shift

Huỳnh Sanh Thông's English translation renders these metaphors in ways that simplify or transform the original conceptualizations. For example, the line:

"Can love and filial piety weigh the same?"

preserves the metaphor of weighing emotions but smooths over the poetic structure of the original. The metaphor is made explicit but somewhat neutralized in tone.

In translating:

"bán mình chuộc cha" as "to redeem her father's life",

the economic metaphor is retained but softened. The English phrase "redeem" carries biblical overtones of salvation, which may resonate differently with Western audiences than the transactional framing in the Vietnamese original. This shift reflects not only a linguistic adjustment but a subtle cultural adaptation to differing emotion ideologies.

4.1.3 Japanese Version: Moral Internalization and Narrative Prose

The Japanese prose translation by Satou Seiji renders the same conceptual content through expository narrative rather than condensed poetic metaphor. For example, the conflict in "bên tình bên hiếu bên nào nặng hơn?" is rephrased in the form of inner monologue:

「夫となる人への誠実を約束して立てた結婚の誓いと、九つの務めを果たしてくれた親への感謝では、どちらがより重いのだろう。」

Here, the metaphor of weighing is preserved (重い), but embedded in a longer reflection. This represents a stylistic shift: from metaphorical compression in Vietnamese to emotional unpacking in Japanese, which reflects a cultural preference for narrative clarity and internal dialogue (tatemae vs honne).

Likewise, the iconic line "bán mình chuộc cha" is reinterpreted in a morally reflective tone:

「こんな酷い状況では、力に屈する以外に方法はないのではないだろうか。」

Here, Kiều rationalizes her sacrifice as the only viable path, indicating a moral dilemma rather than direct metaphorical exchange. The metaphor of commodification is diluted, replaced by a more internalized ethical conflict — aligning with Japanese narrative traditions that favor psychological over transactional framing.

4.1.4 Discussion

This cross-linguistic comparison illustrates how emotional metaphors of moral burden, debt, and sacrifice are differently realized across cultures. The Vietnamese version emphasizes compressed, affective metaphors rooted in Confucian filial duty, the English version tends to lexicalize and neutralize intensity, and the Japanese translation reframes metaphors as internal ethical narratives.

These translation choices reflect broader emotion ideologies (Lutz, 1988) and cultural conceptualizations of duty and selfhood (Kövecses, 2005; Yu, 2009). In Vietnamese, filial piety is a sacred duty often expressed metaphorically; in English, emotional clarity is privileged; in Japanese, self-sacrifice is dramatized through introspection. This underscores how metaphor in translation is not merely linguistic but ideologically and culturally situated.

No.	Vietnamese Excerpt	English Translation	Japanese Translation	Conceptual metaphor	Cultural Framing
1	Bên tình bên hiếu bên nào nặng hơn?	Can love and filial piety weigh the same?	誠実と感謝では、どちらがより重いのだろう。	Moral Decision is a Balance	Vietnamese: compressed moral dilemma;English: neutralized question;Japanese: narrative reflection
2	Lấy tình thâm trả nghĩa thâm mới là	Deep love repays deep gratitude	深い愛情で恩に報いるしかない。	Emotional Debt is Repayment	VN: emotional reciprocity;EN: simplified moral economy;JP: moral rationalization
3	Bán mình chuộc cha	Redeem her father's life	力に屈する以外に方法はないのではないだろうか。	Sacrifice is Payment / Life is Commodity	VN: transactional metaphor;EN: softened religious nuance;JP: ethical reasoning
4	Mặt mo nào phải của mình làm ra	It wasn't a face I gave myself	自分が選んだ顔ではない。	Identity is Given, Not Chosen	VN: fate criticism;EN: passive voice, reflective;JP: impersonal moral tone
5	Chút tình nâng nghĩa xưa nay	Sentiment sustains old favors	古い恩を今の情で支える	Past Duty is Maintained by Present Emotion	Confucian reciprocity: stable in VN;shifted temporality in EN/JP

6	Đã nguyền hai chữ đồng tâm	We vowed to be of one heart	一心同体を誓ったはずなのに	Unity is a Sacred Bond	VN: vow-breaking pain;EN: straightforward statement;JP: regretful lamentation
7	Kể làm sao xiết mối sầu	Endless sorrow hard to describe	どれだけ悲しみが深いか言い表せない	Emotion is an Overwhelming Force	Intensity retained in VN;muted in EN;interiorized in JP
8	Hoa trôi bèo dạt đã đành	A drifting flower, a floating duckweed	浮き草のように流れるしかない	Life is Drift	VN: helpless fate;EN: symbolic tone;JP: resigned existentialism

Figure 1: Cross-linguistic metaphors of moral conflict and filial sacrifice

4.2 Case Study 2: Resignation and Predestined Suffering

This section applies the same analytical framework from Section 4.1 to examine a later passage in Truyện Kiều, focusing on metaphors of fate, existential resignation, and poetic self-consciousness. Through a comparative lens, we analyze how each version, Vietnamese original, English translation, and Japanese prose, constructs the emotional landscape of suffering, acceptance, and finality.

4.2.1 Vietnamese Original: The Metaphoric Weight of Suffering and Fate

In this segment, metaphors about fate and resignation emerge vividly. For example:

"Mảnh trăng đã gác non đoài, một mình luống những đứng ngồi chưa xong."

The moon is personified as having already "rested" on the western mountain (non đoài), setting the temporal and emotional stage. Kiều is metaphorically stuck in indecision, "đứng ngồi chưa xong", embodying emotional paralysis. The phrase encapsulates emotional turmoil through bodily metaphor.

"Triều đâu nổi tiếng đùng đùng, hỏi ra mới biết rằng sông Tiền-đường."

Here, nature mirrors emotion. The turbulent waves of the Tiền Đường River metaphorically reflect her inner storm. This technique of emotional projection onto nature is typical of Sino-Vietnamese poetic tradition, where landscape becomes a canvas for psychological states.

"Nhớ lời thần-mộng rành rành: 'Này thôi hết kiếp đoạn-trường là đây.'"

The metaphor "kiếp đoạn-trường" (life of severed intestines) refers to extreme suffering. This idiom turns visceral pain into an existential metaphor, common in Vietnamese lamentation literature.

"Dưới đèn sẵn bức tiên hoa, một thiên tuyệt-bút gọi là để sau."

This line introduces poetic self-consciousness. "Thiên tuyệt-bút" (a masterpiece) acknowledges the literary act of narrating suffering. The phrase suggests that the story itself, Kiều's life, is framed as an authored tragedy, invoking metafictional awareness.

4.2.2 English Version: Lexical Simplification and Poetic Detachment

The English translation by Huỳnh Sanh Thông often simplifies metaphors while maintaining overall narrative tone. For example:

"She sat alone, the moon already westward, restless, not knowing what to do."

Here, the metaphor of emotional unrest is preserved but rendered more plainly. The dynamic phrase "luống những đứng ngồi chưa xong" is softened into "restless", reducing its metaphorical richness.

"The surf was loud and boisterous; only then did she learn the stream was Tiền Đường."

While retaining natural imagery, the metaphorical projection of emotion onto the river is not emphasized. The emotional resonance is subdued, resulting in a more observational tone.

"She recalled the message in her dream: 'Your life of anguish now will end.'"

"Life of anguish" loosely mirrors "kiếp đoạn-trường" but lacks the idiomatic and visceral power of the original. The metaphor is literalized and emotionally flattened.

"She left behind a final piece, a masterpiece written under the lamp."

This metafictional element is retained, but again, the poetic nuance of "thiên tuyệt-bút" is more straightforwardly rendered as "a final piece."

4.2.3 Japanese Version: Emotional Interiorization and Cultural Familiarity

The Japanese prose translation by Satou Seiji reframes metaphors as internal reflections and narrative exposition:

「月はすでに西の山に沈み、 彼女はただ一人、どうしていいかわからず座ったり立ったりしていた。」

This elaborates on "đứng ngồi chưa xong" with a literal account of indecision and solitude. The emotional state is unpacked into behavioral description, aligning with Japanese preferences for psychological depth.

「波が激しく立ち上がり、彼女はここがあの有名なティエン・ドゥオン川だと知った。」

The dramatic emotional resonance of the river is neutralized, focusing instead on place identification rather than emotional symbolism.

「夢の中で告げられた言葉を思い出した。『この苦しい人生もここで終わる。』

"This 苦しい人生" echoes "kiếp đoạn-trường" but replaces the metaphor with a descriptive phrase. The metaphor is lost in favor of explicit emotional explanation.

「灯火の下、彼女は詩を残した。 それは後世に語り継がれる傑作となった。」

The phrase "後世に語り継がれる傑作" (a masterpiece handed down to posterity) echoes "thiên tuyệt-bút" but emphasizes legacy and posterity rather than authorial finality.

4.2.4 Discussion

This second cross-linguistic comparison continues to show divergent strategies in rendering emotional metaphors. The Vietnamese original favors condensed metaphorical language to convey resignation, inner turmoil, and poetic legacy. The English version prioritizes clarity and directness, often at the expense of metaphorical texture. The Japanese version transforms metaphors into inner narrative states, diluting metaphorical density in favor of reflective prose.

These differences reflect broader cultural orientations toward fate, emotional expression, and narrative closure. Vietnamese emphasizes poetic suffering rooted in Buddhist and Confucian fatalism. English translations lean toward narrative economy and legibility. Japanese adaptations valorize introspective realism and emotional articulation through behavior.

No.	Vietnamese Excerpt	English Translation	Japanese Translation	Conceptual Metaphor	Cultural Framing
1	Mảnh trăng đã gác non đoài	The moon had sunk behind western peaks	月は西の山に沈んでいた	Time of Death is Nightfall	VN: poetic prelude to death; EN/JP: literal setting

2	Một mình luống những đứng ngồi chưa xong	She stood uneasy, pacing back and forth	独り立ち尽くして思い悩む	Emotional Turmoil is Physical Restlessness	VN: metaphorical unrest; EN: simplified action; JP: internal conflict
3	Triều đâu nổi tiếng đùng đùng	Waves suddenly roared aloud	突然、川が轟音を立てた	Inner Emotion is Natural Force	VN: nature reflects emotion; EN/JP: physical phenomenon
4	Hỏi ra mới biết rằng sông Tiền-đường	She asked and learned it was the Tien-duong River	聞けばここは天台河だった	Realization is Unveiling	VN: symbolic river; EN/JP: factual naming
5	Nhớ lời thần-mộng rõ-ràng	She recalled a prophetic dream	はっきりと夢の中の言葉を思い出す	Memory is Revelation	All: dream as foreshadowing
6	"Này thôi hết kiếp đoạn-trường là đây."	"This is the end of a life of sorrow."	「これで多難の生も終わり」	Life is a Path of Suffering	VN: dramatic end; EN/JP: softened tone
7	"Đạm Tiên nàng nhé có hay!"	"Dam Tien, do you know this?"	「端天よ、気づいているか？」	Addressing the Dead is Communication	VN: intimacy with spirit; EN: rhetorical; JP: somber
8	"Hẹn ta thì đợi dưới này rước ta."	"Wait for me down below to welcome me."	「あの世で迎えてくれ」	Death is a Reunion	VN: romanticized death; EN/JP: subdued

9	Dưới đèn sẵn bức tiên hoa	Under the lamp, a scroll was ready	明かりの下に詩文が置かれていた	Message is a Ritual Act	All: symbolic suicide note
10	Một thiên tuyệt-bút gọi là để sau	A masterpiece left behind	後に残す傑作の詩	Writing is Legacy	VN: tragic artistry; EN: literary death; JP: quiet closure

Figure 2: Cross-linguistic metaphors of emotional grief and poetic resignation

4.3 Cross-Cultural Insights

This comparative analysis across the Vietnamese original, English translation, and Japanese adaptation of *Truyện Kiều* reveals distinct patterns of metaphorical preservation, transformation, and omission, shaped profoundly by each culture's emotion ideologies and narrative norms.

4.3.1 Patterns of Preservation, Transformation, and Omission

Across the selected excerpts, the Vietnamese original consistently maintains high metaphorical density, embedding abstract emotional states such as moral conflict, grief, and resignation within culturally specific imagery (e.g., filial sacrifice, flowing tears, trembling body). These metaphors often draw from Buddhist and Confucian motifs, framing emotion as both duty-bound and poeticized.

In contrast, the English translation tends to transform or simplify these metaphors into more direct emotional expressions or descriptive paraphrases. Metaphors of fate or inner conflict are frequently omitted or reframed into accessible psychological language (e.g., "sad and grieved," "deep regret"), prioritizing narrative clarity and emotional legibility for an anglophone audience. This transformation often flattens the symbolic richness but improves readability.

The Japanese version displays a pattern of transmuting metaphors into introspective narration, reflecting cultural preferences for restraint and psychological nuance. Rather than preserving poetic metaphor, the Japanese translation reconfigures emotion through actions, bodily reactions, or internal dialogue, emphasizing the lived experience of emotion rather than its symbolic frame. In some cases, metaphors are partially omitted in favor of behavioral realism (e.g., bowing, embracing, crying silently), reflecting a cultural tendency toward implicit emotional expression.

4.3.2 Emotion Ideologies and Translation Decisions

These patterns reflect broader cultural ideologies about emotion and its role in literature and communication.

In Vietnamese, emotion is often framed as morally and socially bound, embedded in historical, religious, and literary traditions. Emotional metaphors serve not only to convey personal pain but also to elevate that pain into a poetic and philosophical reflection on human destiny. As such, the original text preserves the ambiguity, richness, and performative quality of emotion.

In English, the ideology of emotion leans toward individual psychology, clarity, and therapeutic expression. Translation choices tend to prioritize affective transparency over metaphorical resonance. The result is a form of emotional communication that is clear but less culturally layered.

In Japanese, emotional expression is often context-dependent and socially modulated. Rather than explicit articulation, emotional depth is conveyed through subtle gestures, unspoken understanding, and reflective pauses. This ideology favors internalization over dramatization, which explains the shift from metaphor to introspective narration or action-based emotion in the Japanese rendering.

4.3.3 Implications

These cross-cultural differences in metaphor translation highlight the interplay between language, culture, and emotional worldview. Translators do not merely convert words; they reframe emotional experience according to the target culture's expectations, norms, and literary aesthetics. Thus, translation becomes a site of cultural negotiation, where metaphors either survive, mutate, or vanish—each decision shaping the reader's access to the emotional universe of *Truyện Kiều*.

5. Pedagogical Implications

This study's cross-linguistic and cross-cultural analysis of metaphor translation from *Truyện Kiều* into English and Japanese reveals valuable insights for language and translation pedagogy. By unpacking how emotional metaphors shift or dissolve across languages and cultures, educators can design classroom activities that enhance metaphorical awareness, emotional literacy, and intercultural sensitivity.

5.1 Literary Metaphor Analysis in the Classroom

The metaphor-rich nature of *Truyện Kiều* offers a compelling resource for advanced learners of Vietnamese, English, and Japanese. For example, the image "hoa trôi bèo dạt đã đành" evokes helplessness and fatalism through the conceptual metaphor Life is Drift, which is partially preserved in the English version "drifting like water plants" and recontextualized in Japanese as

inner resignation. Another example, "gột rửa tấm lòng son," expresses the desire to cleanse one's sincere heart—a metaphor tied to moral purity and Confucian virtue.

Through guided analysis, students can identify such conceptual metaphors and compare their preservation, transformation, or omission across translations. This fosters awareness of how metaphors are not only linguistic but deeply cultural. Students become more attuned to metaphorical density in literary texts and develop the ability to critically assess how different ideologies and narrative styles shape translation choices.

5.2 Rewrite Tasks through Cultural Lenses

Another effective classroom activity is to have students rewrite metaphorical lines through specific cultural lenses. For example, the verse "tan tác như hoa giữa đường" might be rendered as "scattered like flowers in the street" in English or internalized into a scene of psychological collapse in Japanese. Learners can explore how cultural attitudes toward suffering and fate—such as Buddhist karma in Vietnamese or emotional introspection in Japanese—inform metaphorical rendering.

This task allows students to reflect on how language encodes emotion ideologies and to practice translation as cultural negotiation rather than simple equivalence. It also encourages students to consider the emotional impact of different translation strategies, especially in cases where literal fidelity may obscure deeper cultural meaning.

5.3 Emotion–Metaphor Matching Activities

Using lines from the comparative table (Figures 4.1.1 and 4.2.1), teachers can design metaphor–emotion matching tasks. Students are asked to read metaphorical phrases such as "giọt nước cho người thác oan" (a drop of water for the unjustly dead) and match them with associated emotions (e.g., grief, injustice) and conceptual metaphors (e.g., Emotion is Water, Death is Release). They then analyze how these metaphors are rendered in English and Japanese and discuss what emotional and cultural nuances are preserved or lost.

These tasks build students' ability to decode affective meaning and highlight how metaphors shape perception across languages and cultures. This also develops students' ability to appreciate emotional subtext and the role of translation in preserving or diluting it.

5.4 Reflection-Based Activities

To deepen critical engagement, students can be asked to write short reflections comparing their emotional responses to each version of a metaphor. For instance, after reading the Vietnamese "giật mình mình lại thương mình xót xa," the English "startled, I pity myself" and the Japanese internal monologue version, learners might reflect on which version moved them more—and why.

Did the metaphor retain its emotional weight? Was the cultural framing relatable? What was lost or gained in translation?

These reflective activities prompt students to consider not just **how** meaning changes across languages but also **why** those changes matter to the reader. This nurtures a personal connection with texts while enhancing cross-cultural empathy and analytical thinking.

5.5 Benefits for Intercultural and Reflective Learning

Integrating metaphor-based analysis and reflective activities into the classroom promotes not only linguistic proficiency but also intercultural awareness. Through exposure to emotionally and culturally charged metaphors, students learn how values such as filial duty, feminine sacrifice, and existential resignation are encoded in poetic form and reshaped in translation. The act of comparing versions pushes learners to confront the limits of translatability and the influence of cultural ideologies on interpretation.

By critically engaging with *Truyện Kiều* across languages, students are empowered to see translation as an act of meaning-making and cultural reflection. This fosters not only linguistic agility but also emotional literacy, critical thinking, and appreciation for literary aesthetics—all essential skills in modern language education.

6. Conclusion

This study examined how emotional metaphors from *Truyện Kiều* are translated across Vietnamese, English, and Japanese, focusing on two excerpts: Thúy Kiều's sacrifice for her father and her final resignation by the river. The analysis revealed patterns of metaphorical preservation, omission, and transformation shaped by cultural ideologies and narrative conventions.

The Vietnamese text, dense in metaphor and cultural nuance, reflects values of filial piety, fatalism, and moral suffering. English translations emphasize clarity and accessibility, often reducing metaphorical depth. Japanese translations, while keeping some poetic texture, recast metaphors into inner states and behavior-driven reflection, aligning with Japanese narrative style.

Pedagogically, these findings underscore the value of literary metaphor analysis in advanced language and translation education. Activities such as comparing metaphorical shifts, rewriting through cultural lenses, and emotion–metaphor matching foster linguistic awareness, intercultural competence, and emotional literacy. Learners recognize that language encodes culture, translation is negotiation rather than equivalence, and metaphor carries both aesthetic and communicative force.

The study further calls for reflective pedagogy situating translation within cultural meaning-making. Incorporating comparative metaphor analysis and reflective writing helps learners develop cross-cultural communicative skills and metaphorical thinking.

Future research should extend to a wider corpus of Vietnamese literary texts and apply these pedagogical strategies in classrooms to assess impact. Such work would enrich literary scholarship and language education.

Ultimately, the findings affirm *Truyện Kiều*'s potential as a resource for exploring how emotions, metaphors, and values intertwine and are reimagined across languages.

REFERENCES

Al-Hasnawi, A. R. (2007). A cognitive approach to translating metaphors. *Translation Journal, 11*(3).

Bassnett, S. (2014). *Translation studies* (4th ed.). Routledge.

Boers, F. (2021). *Evaluating Second Language Vocabulary and Grammar Instruction*. Routledge.

Byram, M. (2021). *Teaching and assessing intercultural communicative competence: Revisited*. Multilingual Matters.

Chan, S. W. (2022). Metaphor in translation: Recent advances and challenges. *Journal of Translation Studies, 45*(2), 112–130.

Charteris-Black, J. (2021). *Analysing Political Speeches: Rhetoric, Discourse and Metaphor*. Palgrave Macmillan.

Deignan, A. (2005). *Metaphor and Corpus Linguistics*. John Benjamins.

Deignan, A., Littlemore, J., & Semino, E. (2013). *Figurative Language, Genre and Register*. Cambridge University Press.

Jiang, X., Gao, S., & Liu, Y. (2021). Cultural adaptation in metaphor translation: A comparative study. *Meta, 66*(2), 276–293.

Kövecses, Z. (2005). *Metaphor in Culture: Universality and Variation*. Cambridge University Press.

Kövecses, Z. (2015). *Where Metaphors Come From: Reconsidering Context in Metaphor*. Oxford University Press.

Kövecses, Z. (2020). *Extended Conceptual Metaphor Theory*. Cambridge University Press.

Kramsch, C. (2009). *The multilingual subject: What foreign language learners say about their experience and why it matters*. Oxford University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

Lefevere, A. (1992). *Translation, rewriting, and the manipulation of literary fame*. Routledge.

Littlemore, J. (2020). *Metaphor, Metonymy, and Experiential Thinking in and outside the Classroom*. Cambridge University Press.

Littlemore, J., & Low, G. (2006). *Metaphor and Language Teaching*. Cambridge University Press.

Littlemore, J., & Low, G. (2020). Metaphor in second language learning: Issues and directions. *Language Teaching, 53*(2), 123–145.

Lutz, C. (1988). *Unnatural Emotions: Everyday Sentiments on a Micronesian Atoll and Their Challenge to Western Theory*. University of Chicago Press.

Mandelblit, N. (1995). The cognitive view of metaphor and its implications for translation theory. In M. Thelen & B. Lewandowska-Tomaszczyk (Eds.), *Translation and Meaning, Part 5* (pp. 483–495). Maastricht: Hogeschool Maastricht.

Neumair, M., Tsuchiya, H., & Nguyen, M. H. (2025). Translating Vietnamese cultural metaphors: A cognitive-pragmatic analysis. *Vietnam Journal of Language Studies, 12*(1), 45–67.

Pan, L. (2023). Conceptual metaphor theory and Chinese emotion metaphors in translation. *Asian Translation Review, 29*(1), 99–118.

Pragglejaz Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, *22*(1), 1–39. https://doi.org/10.1207/s15327868ms2201_1

Schäffner, C. (2004). Metaphor and translation: Some implications of a cognitive approach. *Journal of Pragmatics, 36*(7), 1253–1269.

Schäffner, C., & Shuttleworth, M. (2013). Metaphor in translation: Mapping the state of the art. *Target, 25*(1), 1–28.

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A. A., & Krennmayr, T. (2010). *A Method for Linguistic Metaphor Identification: From MIP to MIPVU*. John Benjamins.

UNESCO. (2021). *UNESCO World Report: Reimagining Our Futures Together*. United Nations Educational, Scientific and Cultural Organization.

Vandaele, J., & Baaijen, V. (2022). Translating affective metaphor: Cognitive shifts and pragmatic constraints. *The Translator, 28*(2), 178–196.

Venuti, L. (1995). *The Translator's Invisibility: A History of Translation*. Routledge.

Wierzbicka, A. (1991). *Cross-Cultural Pragmatics: The Semantics of Human Interaction*. Mouton de Gruyter.

Yu, N. (2009). *From Body to Meaning in Culture: Papers on Cognitive Semantic Studies of Chinese*. Peter Lang.

Yu, N. (2011). The metaphorical nature of Chinese emotion terms: A cognitive linguistic perspective. *Cognitive Linguistics, 22*(2), 315–343.

Yu, N. (2021). Culture, metaphor, and emotion: A cross-linguistic perspective. *Metaphor and Symbol, 36*(3), 167–185.

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO VÀO VIỆC BIÊN SOẠN TỪ ĐIỂN BÁCH KHOA (NGHIÊN CỨU TRƯỜNG HỢP BIÊN SOẠN TỪ ĐIỂN BÁCH KHOA VỀ NGHỀ THỦ CÔNG MỸ NGHỆ TẠI HÀ NỘI)

Ngô Thị Thu Hương[1]

Tóm tắt

Nghiên cứu này tìm hiểu việc ứng dụng trí tuệ nhân tạo (AI) trong biên soạn từ điển bách khoa, cụ thể là từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội. Phương pháp được sử dụng trong bài báo là phương pháp nghiên cứu liên ngành. Kết quả nghiên cứu chỉ những ưu điểm và hạn chế trong quá trình ứng dụng AI vào xây dựng cấu trúc vĩ mô, cấu trúc vi mô của từ điển bách khoa (bao gồm thông tin văn tự và thông tin phi văn tự). Từ đó, bài báo khẳng định rằng, tri thức chuyên gia là nhân tố quan trọng nhất trong quy trình biên soạn từ điển bách khoa nói chung, từ điển bách khoa về nghề thủ công mỹ nghệ ở Hà Nội nói riêng. Trong đó, AI chỉ nên được xem là công cụ hiện đại, hỗ trợ người biên soạn trong một số công đoạn cần xử lý tư liệu nhanh, nhiều và đơn giản.

Từ khóa: trí tuệ nhân tạo, từ điển bách khoa, biên soạn từ điển, nghề thủ công mỹ nghệ, Hà Nội.

1. Đặt vấn đề

Trong bối cảnh khoa học kỹ thuật phát triển như hiện nay, trí tuệ nhân tạo (AI) đã và đang trở thành công cụ quan trọng trong nhiều lĩnh vực nghiên cứu và ứng dụng thực tiễn. Trong đó, ngành từ điển học vốn đòi hỏi xử lý khối lượng lớn dữ liệu ngôn ngữ, hình ảnh và tư liệu liên ngành, cũng đang đứng trước những cơ hội và thách thức mới từ công nghệ này. Bởi vì, nếu như việc biên soạn từ điển truyền thống được biên soạn chủ yếu dựa vào tri thức của chuyên gia, với quá trình sưu tầm và xử lý dữ liệu thủ công trong một khoảng thời gian dài, thì AI mở ra khả năng tự động hóa nhiều công đoạn, từ thu thập, xử lý dữ liệu đến xây dựng cấu trúc vĩ mô, cấu trúc vi mô của quyển từ điển.

Trong lĩnh vực nghiên cứu và bảo tồn ngôn ngữ, văn hóa, việc biên soạn từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội mang ý nghĩa khoa học và thực tiễn sâu sắc. Đây không chỉ là công cụ tra cứu học thuật mà còn góp phần lưu giữ, phổ biến và tôn vinh di sản văn hóa phi vật thể của Thăng Long – Hà Nội, thông qua hệ thống hóa tri thức về các nghề và làng nghề thủ coog mỹ nghệ như gốm Bát Tràng, lụa Vạn Phúc, thêu Quất Động, hay kim hoàn Định Công…. Tuy nhiên, tính chất phức tạp, đa dạng và gắn liền với bối cảnh lịch sử – xã hội của ngữ liệu làng nghề

[1] Viện Ngôn ngữ học, *Email: ngohuong.vnnh@gmail.com*

khiến việc biên soạn gặp nhiều khó khăn, nhất là trong khâu thu thập, xử lý dữ liệu để biên soạn từ điển.

Trước thực tế đó, việc ứng dụng AI trong biên soạn từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội là một hướng đi mới, phù hợp với bối cảnh xã hội. Tuy nhiên, hướng đi này cần phải đánh giá ưu điểm và hạn chế của việc ứng dụng công nghệ AI vào biên soạn từ điển. Vì thế, bài báo này phân tích vai trò của AI trong quá trình biên soạn từ điển bách khoa về nghề thủ công mỹ nghệ, từ khâu xây dựng cấu trúc vĩ mô đến xây dựng cấu trúc vĩ mô, chỉ ra những ưu điểm và hạn chế của việc sử dụng AI vào biên soạn từ điển bách khoa nói chung, từ điển bách khoa về nghề thủ công mỹ nghệ ở Hà Nội nói riêng, góp phần số hóa và bảo tồn vốn di sản ngôn ngữ – văn hóa trong kỷ nguyên mới.

2. Cơ sở lý thuyết và phương pháp nghiên cứu

2.1. Cơ sở lý thuyết

Trí tuệ nhân tạo (Artificial Intelligence - AI) bắt đầu được nghiên cứu từ những năm 1950. Đây là một ngành khoa học máy tính "tập trung vào việc tạo ra các hệ thống hoặc máy móc có khả năng mô phỏng trí thông minh của con người". Những hệ thống này có thể thực hiện các nhiệm vụ như xử lý ngôn ngữ tự nhiên, lập luận, giải quyết vấn đề, học hỏi và thậm chí là sáng tạo nghệ thuật. AI không chỉ dừng lại ở khả năng lập trình thực hiện các tác vụ đơn giản mà còn có khả năng tự cải thiện hiệu suất thông qua quá trình học từ dữ liệu. Do đó. AI đã trở thành nền tảng công nghệ hiện đại, ngày càng mang lại nhiều đột phá trong mọi lĩnh vực của đời sống xã hội như y tế, kinh tế, giáo dục....

Từ điển học và bách khoa thư hiện đại chú trọng đến tính mở, khả năng phân tích dữ liệu và tính tích hợp đa phương tiện. Trong đó, từ điển bách khoa, một loại sách tra cứu có chức năng tìm kiếm thông tin nhanh và thuận tiện. Quy mô của từ điển bách khoa thường nhỏ hơn bách khoa thư. Bảng mục từ (cấu trúc vĩ mô) của từ điển bách khoa mặc dù số lượng nhiều hơn nhưng nội dung thông tin (cấu trúc vi mô) đưa ra đơn giản và biên độ ngắn hơn loại hình bách khoa thư. Tuy nhiên, bảng mục từ của từ điển bách khoa ít hơn nhưng nội dung thông tin phong phú hơn và mang tính chất "bách khoa" nhiều hơn so với từ điển ngôn ngữ.

Từ điển học nói chung và từ điển bách khoa nói riêng trên thế giới đã có lịch sử phát triển lâu dài, liên tục từ thời cổ đại cho đến ngày nay như bộ Naturalis historia (Lịch sử tự nhiên) (thế kỷ thứ nhất), bộ Vĩnh Lạc đại điển (1480), *Encyclopaedia Britannica* (Anh) (1768-1771), *Grand Larousse Encyclopédique* (Pháp) (1960-1975), hay *Bách khoa toàn thư Liên Xô* (Nga) (1926-1947), *Đại bách khoa toàn thư Trung Quốc* (1980–1993)… Đây là những bộ từ điển được tái bản nhiều lần. Trong khi đó, ở Việt Nam, từ cuối thế kỷ XX, nhiều công trình biên soạn bách khoa quan trọng đã được triển khai, tiêu biểu là *Từ điển Bách khoa Việt Nam* (1995–2005), *Từ điển Bách khoa nghề thủ công truyền thống ở Việt Nam* (2020)…Tuy nhiên, hầu hết các công trình

trước đây đều được biên soạn theo phương thức thủ công, đòi hỏi nhiều thời gian và nhân lực, đồng thời còn hạn chế trong việc tích hợp tư liệu đa phương tiện và ngữ liệu sinh động từ thực tiễn văn hóa. Trong những thập niên gần đây, dưới sự phát triển mạnh mẽ của khoa học kỹ thuật, công nghệ số được ứng dụng vào việc biên soạn từ điển và bách khoa thư, chẳng hạn như *Từ điển bách khoa Encarta CD* của hãng Microsoft chứa hàng nghìn tranh ảnh màu và những chuỗi âm thanh được trình bày thẩm mỹ và truy cập rất thuận lợi.

Tại Việt Nam, một số nghiên cứu bước đầu đã đề cập đến khả năng ứng dụng công nghệ số trong lĩnh vực ngôn ngữ học và biên soạn từ điển, song chủ yếu dừng ở mức có thể triển khai trên hai bình diện: số hóa tài liệu phục vụ cho việc biên soạn; tiến hành biên soạn mục từ trên nền tảng ứng dụng công nghệ số. Các công trình chuyên sâu về biên soạn từ điển bách khoa kết hợp AI, đặc biệt trong phạm vi văn hóa – nghề thủ công truyền thống, hiện vẫn đang còn hạn chế. Trong khi đó, tại các làng nghề thủ công mỹ nghệ Hà Nội (như gốm sứ Bát Tràng, dệt lụa Vạn Phúc, thêu Quất Động, đậu bạc Định Công…) đang lưu giữ một hệ thống tri thức đặc thù, đa tầng và giàu tính văn hóa, đòi hỏi cách tiếp cận mới và công cụ hỗ trợ mạnh mẽ.

Từ đó có thể thấy, việc nghiên cứu ứng dụng AI vào biên soạn từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội không chỉ góp phần bổ sung khoảng trống trong nghiên cứu từ điển học ở Việt Nam, mà còn có ý nghĩa thực tiễn trong bảo tồn và phát huy di sản văn hóa phi vật thể Thăng Long – Hà Nội.

2.2. Phương pháp nghiên cứu

Bài viết được triển khai trên cơ sở tiếp cận liên ngành, kết hợp giữa ngôn ngữ học, từ điển học, khoa học dữ liệu và nghiên cứu văn hóa. Các phương pháp chính bao gồm:

Phương pháp khảo cứu tài liệu: Thu thập và phân tích các công trình trong và ngoài nước về lý thuyết từ điển học, bách khoa thư, cùng các nghiên cứu gần đây về ứng dụng trí tuệ nhân tạo trong xử lý ngôn ngữ tự nhiên và dữ liệu văn hóa.

Phương pháp điền dã dân tộc học: Khảo sát, ghi chép tư liệu tại các làng nghề tiêu biểu như gốm sứ Bát Tràng, dệt lụa Vạn Phúc, thêu Quất Động, đậu bạc Định Công. Hoạt động này giúp thu thập hệ thống thuật ngữ nghề nghiệp, ngữ liệu thực hành và tri thức văn hóa bản địa từ các cộng đồng làm nghề.

Phương pháp xử lý ngữ liệu với sự hỗ trợ AI: Ứng dụng các công cụ AI vào việc xây dựng cấu trúc vĩ mô và cấu trúc vi mô của từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội. Đồng thời, bài viết này cũng sử dụng AI để tạo minh họa đa phương tiện (sơ đồ, hình vẽ, ảnh mô phỏng) cho cuốn từ điển này.

Phương pháp phân tích – tổng hợp: Đối chiếu kết quả xử lý dữ liệu bằng AI với tri thức chuyên gia nhằm đảm bảo tính chính xác khoa học, đồng thời rút ra các ưu điểm, hạn chế và khả năng ứng dụng thực tiễn của mô hình đề xuất.

Phương pháp chuyên gia: Tham khảo ý kiến của các nhà nghiên cứu ngôn ngữ học, văn hóa học và nghệ nhân làng nghề để thẩm định tính khoa học và giá trị văn hóa – lịch sử của dữ liệu được AI xử lý.

Bên cạnh đó, nghiên cứu này còn ứng dụng AI vào biên soạn từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội vào xây dựng cấu trúc vĩ mô trên cơ sở phân loại từ ngữ nghề nghiệp (từ ngữ chuyên môn) theo trường từ vựng ngữ nghĩa; ứng dụng AI vào việc xây dựng vi mô của một số mục từ của từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội. Từ đó, đưa ra một số khuyến nghị trong việc sử dụng AI trong biên soạn từ điển nói chung, từ điển bách khoa về nghề thủ công mỹ nghệ ở Hà Nội nói riêng.

3. Kết quả nghiên cứu (nghiên cứu trường hợp biên soạn từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội)

3.1. Ứng dụng AI vào việc xây dựng cấu trúc vĩ mô

Theo Nguyễn Ngọc Trâm (1997), "cấu trúc vĩ mô là cấu trúc bao gồm toàn thể các mục từ được sắp xếp trong từ điển theo một trật tự xác định. Cấu trúc này còn được gọi là cấu trúc bảng từ". Mỗi loại từ điển khác nhau sẽ có cấu trúc vĩ mô khác nhau. Thậm chí, cùng loại hình từ điển với nhau nhưng mỗi đối tượng biên soạn khác như cũng có cấu trúc vĩ mô khác nhau, chẳng hạn, cùng loại hình từ điển bách khoa nhưng từ điển bách khoa về ẩm thực, từ điển bách khoa về địa danh, từ điển bách khoa về nông cụ… sẽ có cấu trúc vĩ mô khác với từ điển bách khoa về các nghề thủ công mỹ nghệ.

Nghề thủ công mỹ nghệ ở Hà Nội được hình thành từ nhiều nghề thủ công khác nhau, được hình thành và phát triển ở Hà Nội như nghề gốm sứ Bát Tràng, nghề thêu Quất Động, nghề làm tranh Hàng Trống… Những nghề này thường được xuất hiện do nhu cầu phục vụ đời sống dân sinh, được người dân làm khi nông nhàn. Sau này, nhờ hiệu quả kinh tế mang lại, những nghề này chuyển dần sang phục vụ nhu cầu thẩm mỹ, nhu cầu tôn giáo,... của người Hà Nội. Do đó, xây dựng bảng từ cho từ điển bách khoa về nghề thủ công mỹ nghệ ở Hà Nội cần thu thập theo hệ thống các từ ngữ nghề nghiệp của mỗi nghề.

Sử dụng công nghệ AI vào việc xây dựng cấu trúc vĩ mô của từ điển bách khoa về nghề thủ công mỹ nghệ ở Hà Nội, có kết quả như sau:

Bảng 1. ChatGPT trích xuất, phân loại hệ thống từ ngữ nghề nghiệp

nghề gốm Bát Tràng và nghề thêu Quất Động theo nhóm từ vựng - ngữ nghĩa

Nghề gốm Bát Tràng	Nghề thêu Quất Động
1. Nguyên liệu & vật liệu Đất sét, cao lanh, đất trắng, đất phù sa	1. **Nguyên liệu, vật liệu thêu**

Thạch anh, tràng thạch Men tro, men ngọc, men lam, men rạn, men tro rơm, men nâu **2. Công cụ & dụng cụ** Bàn xoay, bàn gỗ Khuôn, dao nặn, que gạt, bàn dập Chổi quét men, gáo múc men Kìm gắp, đũa nung **3. Công đoạn sản xuất** Lọc đất, luyện đất, ủ đất Nặn tay, chuốt bàn xoay, đổ khuôn Phơi, sấy Vẽ lam, khắc, in hoa văn, tráng men Nung lò, ra lò **4. Sản phẩm gốm** Ấm chén, bát, đĩa Lọ, bình, chum, vại Tượng gốm, linh vật Ống bình vôi, đèn thờ **5. Tên gọi lò & kỹ thuật nung** Lò bầu, lò ếch, lò gas, lò điện Nung hạ, nung ngược, nung oxi hóa, nung khử **6. Hoa văn & trang trí** Hoa lam, hoa nâu Vẽ phủ men, vẽ khắc nổi In khuôn, in hoa văn chìm Vẽ phẩy, vẽ chỉ vàng	Chỉ thêu: chỉ tơ, chỉ gấm, chỉ kim tuyến, chỉ màu, chỉ ngũ sắc… Vải thêu: vải sa, vải the, vải gấm, lụa, vải thô, vải bông… Phụ liệu: hồ vải, giấy can, giấy than, than bút, sáp ong (dùng bôi chỉ)… 2. **Dụng cụ, thiết bị nghề** Kim thêu: kim nhỏ, kim to, kim xuyên, kim đầu nhọn, kim khâu… Khung thêu: khung căng, khung gỗ, khung tròn, khung vuông… Dụng cụ phụ: kéo, đệm tay, thước, bút vẽ mẫu, nhíp… 3**. Kỹ thuật và mũi thêu** Mũi thêu cơ bản: mũi đột, mũi tằm, mũi tỉa, mũi vắt, mũi hạt gạo… Mũi thêu trang trí: mũi bạt, mũi lướt vặn, mũi lướt vặn kép, mũi móc xích… Kỹ thuật nâng cao: thêu chồng, thêu cuộn, thêu nổi, thêu phủ, thêu đệm… 4. **Hình thức và sản phẩm thêu** Đồ thờ: hoành phi, câu đối, trướng thờ, áo thờ, khăn phủ bàn thờ… Trang phục: áo dài thêu, áo the, áo mũ cung đình, khăn vấn… Tranh thêu: tranh phong cảnh, tranh tứ quý, tranh Phật, tranh dân gian… Đồ gia dụng – trang trí: khăn trải bàn, rèm, vỏ gối, túi xách, quạt thêu… 5**. Quy trình sản xuất**

	Chuẩn bị: chọn vải, căng khung, vẽ mẫu, can hoa văn… Thêu: xâu chỉ, chọn mũi thêu, phối màu, hoàn thiện chi tiết… Hoàn thiện: gỡ khung, là ủi, hồ vải, đóng khung, giao sản phẩm… 6. **Thuật ngữ nghề nghiệp liên quan đến tổ chức sản xuất và xã hội nghề** Người thợ: thợ chính, thợ phụ, thợ học việc, nghệ nhân, truyền nhân… Tổ chức nghề: phường thêu, giáo phường, tổ nghề, sư phụ – đệ tử… Hoạt động nghề: truyền nghề, giữ bí quyết, thi tài, dâng lễ Tổ nghề…

Nguồn: Ngô Thị Thu Hương, 2025

Có thể thấy, trong một khoảng thời gian rất ngắn, AI đã giúp người biên soạn có được các chủ đề và các mục từ xuất hiện trong một nghề nhất định, làm cơ sở cho việc xây dụng cấu trúc vĩ mô của từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội, chẳng hạn như trong nghề làm gốm sứ Bát Tràng, hệ thống từ ngữ nghề nghiệp có thể phân loại theo sáu chủ đề như: (1) nguyên liệu & vật liệu; (2) công cụ & dụng cụ; (3) công đoạn sản xuất; (4) sản phẩm gốm; (5) tên gọi lò & kỹ thuật nung; (6) hoa văn & trang trí (hộp 1). Trong khi đó, hệ thống từ ngữ nghề thêu Quất Động có thể được phân loại thành 6 nhóm chủ đề chuyên môn: (1) nguyên liệu, (2) dụng cụ, (3) kỹ thuật, (4) sản phẩm, (5) quy trình, và (6) thuật ngữ nghề nghiệp liên quan đến tổ chức sản xuất và xã hội nghề (bảng 1).

Tuy nhiên, điểm hạn chế khi ứng dụng AI vào xây dựng cấu trúc vĩ mô của từ điển bách khoa về nghề thủ công mỹ nghệ tại Hà Nội khi gợi ý các chủ đề của từng nghề cụ thể. Bởi vì trong thực tế nghiên cứu từ ngữ nghề nghiệp, các chủ đề trên thực tế còn xuất hiện nhiều hơn, chẳng hạn như nhóm từ ngữ biểu thị người lao động gồm có *thợ chồng lò, thợ nung, thợ vuốt*,… (nghề gốm sứ Bát Tràng); *thợ thêu* (nghề thêu Quất Động); *thợ phun sơn, thợ đánh bóng, thợ đan*…(nghề đan mây tre Phú Vinh); nhóm từ ngữ biểu thị ông/ bà tổ nghề như: *ông Lê Công Hành (ông tổ nghề thêu), bà Lã Thị Nga (bà tổ nghề dệt lụa Vạn Phúc)…*

Bảng 2. ChatGPT trích xuất, phân loại hệ thống từ ngữ nghề đậu bạc Định Công

Quy trình làm nghề đậu bạc Định Công (lần 1)	Quy trình làm nghề đậu bạc Định Công (lần 2)
Quy trình làm đậu bạc Định Công – một sản phẩm truyền thống nổi tiếng của làng nghề Định Công (Hà Nội) – được tiến hành công phu qua nhiều bước: **1. Chọn nguyên liệu** Lạc (đậu phộng): chọn hạt đều, chắc, không sâu mọt. Đường: thường dùng đường kính trắng, loại trong và mịn để khi kết tinh tạo lớp bạc sáng đẹp. 2**. Rang và làm sạch lạc** Lạc được rang trên chảo gang hoặc chảo đất, giữ lửa vừa để hạt chín vàng, giòn nhưng không cháy. Sau khi rang, lạc được sàng và xát bỏ sạch lớp vỏ lụa. 3**. Nấu nước đường** Đường cho vào chảo cùng lượng nước vừa đủ, nấu trên bếp than/củi. Người thợ phải khéo léo canh "độ keo" của đường: nếu quá loãng sẽ không bám, quá đặc dễ cháy. 4**. Phủ đường lên hạt lạc** Đổ lạc rang vào chảo nước đường. Dùng đũa hoặc thìa lớn đảo đều tay, liên tục, để đường kết tinh bám đều quanh từng hạt. Khi hoàn thiện, hạt lạc phủ lớp đường trắng tinh, sáng như bạc. 5. **Làm nguội và sàng lọc** Đậu bạc được đổ ra mẹt tre hoặc nia, dàn mỏng cho nguội và tách rời, tránh dính cục. Sàng lọc để loại bỏ những hạt vón hoặc không đều.	Quy trình làm nghề đậu bạc Định Công là một trong những kỹ nghệ tinh xảo của làng nghề kim hoàn cổ truyền ở Hà Nội. Nghề này đòi hỏi sự kết hợp khéo léo giữa kỹ thuật thủ công tinh vi và óc sáng tạo nghệ thuật. Quy trình cơ bản thường gồm các công đoạn chính sau: **Chuẩn bị bạc nguyên liệu** – Bạc nguyên chất hoặc hợp kim bạc được chọn lọc, làm sạch, nấu chảy thành thỏi hoặc dây nhỏ để tiện cho quá trình gia công. **Kéo sợi – cán bạc** – Bạc được nung chảy, kéo thành dây, sau đó cán mỏng hoặc vê tròn thành sợi bạc có độ to nhỏ khác nhau, phù hợp với từng chi tiết sản phẩm. **Tạo hạt (làm đậu)** – Cắt sợi bạc thành từng đoạn nhỏ, nung trên than cho bạc chảy tụ lại thành hạt tròn li ti. Đây chính là công đoạn đặc trưng tạo nên "đậu bạc". **Sắp đặt và ghép hạt** – Những hạt bạc nhỏ được sắp xếp, gắn kết khéo léo theo hoa văn, họa tiết truyền thống (hoa, lá, hình học…) trên khung bạc lớn. **Hàn gắn và định hình** – Thợ sử dụng lửa và kỹ thuật hàn tinh vi để cố định các hạt bạc với nhau, đảm bảo chắc chắn mà vẫn giữ được sự thanh thoát, tinh xảo. **Chạm khắc – hoàn thiện chi tiết**

	– Bề mặt được chạm trổ, chỉnh sửa, gắn thêm hoa văn phụ hoặc kết hợp với đá quý/men để tăng tính thẩm mỹ. **Đánh bóng và hoàn thiện sản phẩm** – Sản phẩm sau cùng được làm sạch, đánh bóng để ánh bạc sáng đẹp, tạo nên các tác phẩm trang sức và đồ mỹ nghệ tinh xảo đặc trưng của Định Công. ☞ Quy trình trên cho thấy tính công phu, sự tỉ mỉ và giá trị thẩm mỹ cao của nghề đậu bạc.

Nguồn: Ngô Thị Thu Hương, 2025

Hạn chế thứ hai của việc sử dụng AI vào việc xây dựng cấu trúc vĩ mô của từ điển nghề thủ công mỹ nghệ là tùy từng thời gian khác nhau, AI cho các kết quả khác nhau (bảng 2). Đồng thời, AI còn cung cấp kết quả trích xuất, phân loại hệ thống từ ngữ nghề đậu bạc Định Công lần 1 còn sai lệch vì nghề đậu bạc Định Công, vốn là nghề kim hoàn có truyền thống hơn một nghìn năm ở Hà Nội, không phải là một nghề làm kẹo. Thậm chí, AI chỉ đưa thông tin về nghề đậu bạc Định Công một cách tương đối khi mô tả công đoạn **tạo hạt** (làm đậu) là "Cắt sợi bạc thành từng đoạn nhỏ, nung trên than cho bạc chảy tụ lại thành hạt tròn li ti. Đây chính là công đoạn đặc trưng tạo nên "đậu bạc"" (bảng 2). Trong khi thực tế hoạt động này là làm các họa tiết như *guột, sòi* (hộp 1).

Hộp 1. Các từ ngữ nghề nghiệp trong nghề đậu bạc

Họa tiết guột thường là sợi chỉ se được cuốn thành một vòng tròn ban đầu và tiếp tục được xoắn thành hai hoặc vài vòng liên tục từ vòng tròn đã có, được dùng nhằm mục đích xen vào các khoảng trống giữa các họa tiết hoặc tăng khoảng thưa thoáng của sản phẩm, tạo điểm nghỉ mắt giữa các mảng họa tiết dày đặc.

Guột lá: người thợ cuốn tròn họa tiết từ tâm, sau khi đã tạo ra sòi hình tròn sẽ thả lỏng cuốn. Sau đó bóp họa tiết dẹt lại và thuôn về hai đầu theo trục thẳng. Guột lá là dạng họa tiết có nhiều biến thể nhất trong các họa tiết của đậu bạc Định Công.

Với họa tiết sòi có ba dạng chính như sòi đơn, sòi hình tròn, sòi hình mắt (sòi elip). Sòi đơn: gần giống với guột, sòi đơn là các dạng họa tiết có số vòng cuốn ít, thường trong khoảng từ 1 tới vài vòng cuốn. Trong sòi đơn có nhiều loại: sòi tim, sòi chữ S, sòi ấu, sòi số 9… Sòi hình tròn: người thợ kim hoàn dùng sợi chỉ se vừa cuốn tròn thật chặt tay từ tâm họa tiết. Đây là dạng họa tiết cơ bản nhất, từ đó thực hiện các dạng họa tiết khác. Sòi hình tròn cũng là họa tiết thường được sử

dụng nhiều nhất, được biến thể từ hoa văn hình xoáy ốc của dân tộc Việt. Sòi hình mắt (sòi hình elip): sau khi cuốn tròn họa tiết, người thợ kim hoàn bóp giữa thân nhằm kéo dài họa tiết để tạo sòi hình mắt, sao cho hai đầu họa tiết vẫn giữ được đường cong tròn.

Nguồn: Nguyễn Hương Ly (2022), Tính nghệ thuật trong sản phẩm đậu bạc Định Công, Nguồn: Tạp chí VHNT số 485, tháng 1-2022 trên website: http://www.vanhoanghethuat.vn/tinh-nghe-thuat-trong-san-pham-dau-bac-dinh-cong.htm

Như vậy, có thể thấy một trong những ưu điểm nổi bật AI khi ứng dụng vào biên soạn từ điển bách khoa về nghề thủ công mỹ nghệ ở Hà Nội là tốc độ xử lý dữ liệu nhanh và tối ưu hóa quy trình biên soạn. Truyền thống biên soạn từ điển đòi hỏi thời gian dài để thu thập, phân loại và chuẩn hóa thông tin từ nhiều nguồn tư liệu khác nhau, bao gồm văn bản, hình ảnh... Với sự hỗ trợ của các mô hình học máy và xử lý ngôn ngữ tự nhiên, AI có thể tự động hóa nhiều công đoạn vốn mang tính thủ công, chẳng hạn như trích xuất từ khóa, phân tích ngữ nghĩa, nhận diện thuật ngữ chuyên ngành và liên kết dữ liệu liên văn bản từ nhiều nguồn khác nhau một cách nhanh chóng. Điều này giúp rút ngắn đáng kể thời gian biên tập và hiệu đính, vốn trước đây thường tiêu tốn nhiều năm công sức của cả một nhóm chuyên gia. Nhờ đó, người biên soạn có thể tiếp cận hệ thống tri thức phong phú, được cập nhật thường xuyên, giảm thiểu sự chồng chéo hoặc thiếu sót trong quá trình tổng hợp. Sự gia tốc này không chỉ mang lại lợi ích về mặt thời gian, mà còn nâng cao hiệu suất của quá trình biên soạn, đặc biệt trong bối cảnh khối lượng dữ liệu về nghề thủ công mỹ nghệ ngày càng mở rộng và đa dạng.

Từ góc độ ứng dụng thực tiễn, AI giúp giải phóng nguồn lực cho người biên soạn, cho phép họ tập trung nhiều hơn vào khâu phân tích chuyên sâu và diễn giải ngữ nghĩa, thay vì dành quá nhiều thời gian cho công việc thu thập và xử lý dữ liệu thô. Điều này tạo điều kiện để quá trình biên soạn từ điển bách khoa không chỉ nhanh hơn, mà còn giàu tính khoa học, phản ánh đầy đủ đặc trưng văn hóa – xã hội của các làng nghề truyền thống ở Hà Nội.

3.2. Ứng dụng AI vào việc xây dựng cấu trúc vi mô

3.2.1. Ứng dụng AI vào việc cung cấp thông tin văn tự

Nguyễn Ngọc Trâm (1997) cho rằng, cấu trúc vi mô là toàn bộ những thông tin được trình bày một cách hệ thống trong mỗi mục từ. Một trong thông tin quan trọng được cung cấp trong mục từ là thông tin văn tự (chữ viết). Điểm khác biệt giữa từ điển ngôn ngữ và từ điển bách khoa là việc cung cấp thông tin. Nếu như trong từ điển ngôn ngữ, các thông tin cung cấp gồm thông tin về ngôn ngữ như ngữ âm, chính tả, từ vựng, ngữ pháp, ngữ dụng, từ nguyên,... thì trong từ điển bách khoa có sự kết hợp giữa thông tin ngôn ngữ và thông tin bách khoa. Trong đó, thông tin bách khoa là những "thông tin chung và nằm ngoài giới hạn của thông tin ngôn ngữ".

Ứng dụng công nghệ AI trong việc cung cấp thông tin văn tự trong cấu trúc vi mô của từ điển bách khoa về nghề thủ công mỹ nghệ ở Hà Nội, có kết quả như sau:

Hộp 2. ChatGPT định nghĩa đình Tú Thị theo lối từ điển bách khoa

Định nghĩa lần 1	Định nghĩa lần 2
Đình Tú Thị **Định nghĩa:** Đình làng Tú Thị, thuộc xã Bát Tràng, huyện Gia Lâm, Hà Nội, là công trình kiến trúc tín ngưỡng – văn hóa truyền thống của cư dân địa phương, nơi thờ Thành hoàng làng và gắn bó chặt chẽ với đời sống cộng đồng cư dân gốm sứ. **Lịch sử:** Đình được dựng từ thời Lê Trung hưng (thế kỷ XVII), trải qua nhiều lần tu bổ dưới thời Nguyễn và thời hiện đại. Theo truyền thống, đình thờ các vị thần có công bảo hộ dân làng, đồng thời là nơi diễn ra những sự kiện quan trọng của làng nghề Bát Tràng. **Kiến trúc:** Đình có kết cấu gỗ truyền thống kiểu chữ "Đinh" hoặc "Công", mái ngói chồng diêm, đầu đao cong vút. Nghệ thuật chạm khắc trên vì kèo, xà ngang, bẩy gỗ thể hiện nhiều đề tài tứ linh, hoa lá, cảnh sinh hoạt. Một số mảng trang trí được gắn gốm sứ men Bát Tràng, tạo nét đặc trưng so với nhiều đình khác ở vùng đồng bằng Bắc Bộ. **Chức năng – Sinh hoạt:** Ngoài vai trò tín ngưỡng, đình Tú Thị là nơi hội họp, bàn việc làng, tổ chức lễ hội truyền thống. Lễ hội đình thường diễn ra vào mùa xuân, bao gồm nghi lễ rước, tế thần, trò chơi dân gian và các hoạt động gắn với nghề gốm. **Giá trị:** Đình Tú Thị không chỉ có giá trị kiến trúc nghệ thuật mà còn là tư liệu sống động về lịch sử xã hội và văn hóa nghề gốm Bát Tràng. Đây là di sản văn hóa quan trọng để nghiên cứu mối liên hệ giữa tín ngưỡng	định nghĩa theo lối từ điển bách khoa về mục từ "Đình Tú Thị" ChatGPT said: Mình hiểu rồi, bạn muốn một định nghĩa chuẩn lối từ điển bách khoa cho mục từ "Đình Tú Thị". Dưới đây là bản thảo: **Đình Tú Thị:** Nhân vật nữ trong truyện Nôm Phạm Tải – Ngọc Hoa, tác phẩm văn học dân gian Việt Nam lưu truyền từ khoảng thế kỷ XVIII–XIX. Đình Tú Thị là con gái Thừa tướng triều đình, nổi tiếng dung nhan, sống trong cảnh giàu sang, quyền quý. Trong mạch truyện, nhân vật này thường được miêu tả với tính cách kiêu kỳ, nhiều tham vọng, và gắn liền với những toan tính hôn nhân – chính trị của tầng lớp quan lại phong kiến. Hình tượng Đình Tú Thị vừa phản ánh thân phận người phụ nữ trong xã hội xưa, vừa thể hiện quan niệm đạo đức và phê phán xã hội đương thời qua văn học Nôm.

làng xã, kiến trúc truyền thống và nghề thủ công.	

Nguồn: Ngô Thị Thu Hương, 2025

Cũng như kết quả của việc ứng dụng AI vào việc xây dựng cấu trúc vĩ mô, kết quả việc ứng dụng AI vào việc xây dựng cấu trúc vi mô cũng có những ưu điểm và hạn chế nhất định.

Kết quả ứng dụng AI vào việc biên soạn mục từ "Đình Tú Thị" có thể gợi mở cho người biên soạn cấu trúc thông tin trong loại mục từ về cơ sở tín ngưỡng liên quan đến nghề, bao gồm các thông tin như: (1) Định nghĩa; (2) Lịch sử; (3) Kiến trúc; (4) Chức năng – Sinh hoạt; (5) Giá trị (hộp 2). Tuy nhiên, các nội dung thông tin trong mục từ "Đình Tú Thị" mà AI đưa ra không đúng. Trong định nghĩa lần 1, AI giải thích "Đình Tú Thị" là một cơ sở tín ngưỡng. Tuy nhiên, nội dung thông tin lại nói về Bát Tràng. Trong định nghĩa lần 2, AI giải thích "Đình Tú Thị" là nhân vật nữ trong truyện Nôm. Trong thực tế, đình Tú Thị là cơ sở tín ngưỡng liên quan đến nghề thêu Quất Động (hộp 2).

Hộp 2. Thông tin về đình Tú Thị

Đình Tú Thị có tên nôm là đình Chợ Thêu, tên chữ là Tú Đình thị, nghĩa là "Chợ đình thợ Thêu" thuộc đất thôn Yên Thái, tổng Tiền Túc (sau là tổng Thuận Mỹ) huyện Thọ Xương, phủ Hoài Đức.

Theo các nguồn sử liệu, đình Chợ Thêu được người dân Quất Động (Thường Tín, Hà Nội) đến tụ cư tại kinh thành Thăng Long xây dựng vào năm 1891 để thờ tổ nghề thêu là cụ Lê Công Hành.

Nguồn: Trần Hòa (2024), Ghé đình Tú Thị, nhớ ông tổ nghề thêu, https://giaoducthoidai.vn/ghe-dinh-tu-thi-nho-ong-to-nghe-theu-post673806.html.

Như vậy, bên cạnh việc ưu điểm có thể gợi ý các nội dung cấu trúc vi mô trong từ điển bách khoa, AI lại cho thấy nguy cơ cung cấp nội dung không chính xác. Trong khi từ điển bách khoa nói chung, từ điển bách khoa về nghề thủ công mỹ nghệ nói riêng là loại sách tra cứu, nên mọi thông tin trong đó cần phải chính xác tuyệt đối. Vì thế, người biên soạn không được phép phụ thuộc hoàn toàn vào AI trong quá trình biên soạn từ điển. Nếu phụ thuộc hoàn toàn vào AI sẽ ảnh hưởng nghiêm trọng đến chất lượng của sản phẩm biên soạn.

3.2.2. Ứng dụng AI vào việc cung cấp thông tin phi văn tự

Từ điển học hiện đại cho rằng, ngoài thông tin văn tự (kênh chữ viết) còn tồn tại loại thông tin khác như tranh vẽ, bảng biểu, hình ảnh, thậm chí âm thanh (video) có vai trò vô cùng quan trọng trong việc cung cấp tri thức về nghề TCMN một cách trực quan và đầy đủ cho người tra cứu. Trong bài viết này, chúng tôi ứng dụng công nghệ AI vào việc cung cấp thông tin phi văn tự qua

sơ đồ quy trình, hình ảnh minh họa nhằm tái hiện các công đoạn sản xuất cho từ điển nghề thủ công mỹ nghệ tại Hà Nội.

Hộp 3. Ứng dụng AI trong việc cung cấp thông tin hình ảnh về quy trình làm nghề đậu bạc Định Công

Yêu cầu lần 1: hình ảnh quy trình làm nghề đậu bạc Định Công	Yêu cầu lần 2: hình ảnh quy trình làm nghề đậu bạc Định Công

Nguồn: Ngô Thị Thu Hương, 2025

Ứng dụng công nghệ AI vào việc cung cấp thông tin phi văn tự trong từ điển bách khoa về nghề thủ công mỹ nghệ ở Hà Nội có thể giúp cho người biên soạn truy cập và xử lý dữ liệu từ nhiều nguồn khác nhau trên không gian mạng. Đồng thời công nghệ AI cũng cho phép minh họa trực quan các kỹ thuật nghề thủ công, từ quy trình sản xuất đến hoa văn, họa tiết và các dụng cụ đặc thù, điều mà các lối biên soạn truyền thống chỉ dựa trên văn bản khó thực hiện đầy đủ. Tuy nhiên, việc cung cấp thông tin trực quan trong từ điển bách khoa về nghề thủ công mỹ nghệ Hà Nội qua AI cũng bộc lộ hạn chế: hình ảnh không hoặc chưa phản ánh đúng thực tế của ngành nghề thủ công. Để khắc phục được vấn đề này, cần phải sử dụng tri thức chuyên gia để thẩm định tính chính xác, tính khoa học và tính thực tiễn của đối tượng biên soạn. Trong quy trình này, AI chỉ nên được xem là công cụ hiện đại, hỗ trợ người biên soạn trong một số công đoạn cần xử lý tư liệu nhanh, nhiều và đơn giản.

4. Kết luận

Nhìn chung, việc ứng dụng AI trong biên soạn từ điển bách khoa về nghề thủ công mỹ nghệ mang lại nhiều lợi thế vượt trội như giúp rút ngắn thời gian biên soạn nhờ khả năng xử lý dữ liệu lớn, phân loại và trích xuất thông tin tự động, đồng thời hỗ trợ cung cấp thông tin đa ngữ và tích hợp đa phương tiện, cho phép người biên soạn tiếp cận và minh họa tri thức một cách toàn diện, trực quan và sinh động. Tuy nhiên, AI vẫn tồn tại một số hạn chế, bao gồm nguy cơ cung cấp

thông tin chưa chính xác và không thể thay thế tri thức môn về văn hóa làm nghề. Điều này nhấn mạnh rằng AI chỉ nên được xem là công cụ hỗ trợ chiến lược, giúp con người trong một số công đoạn biên soạn từ điển bách khoa nói chung, từ điển bách khoa về nghề thủ công mỹ nghệ nói riêng. Như vậy, AI và chuyên gia biên soạn từ điển không phải là các yếu tố thay thế cho nhau mà chúng có quan hệ tương hỗ, bổ sung lẫn nhau. Điều này tạo nên một quy trình biên soạn từ điển hiện đại, khoa học và bền vững.

Tài liệu tham khảo

Nguyễn Huy Bỉnh (2021), Giá trị của việc ứng dụng công nghệ số trong lĩnh vực từ điển và bách khoa thư, *Tạp chí Thông tin Khoa học xã hội*, số 2.

Phạm Văn Hảo (2023), *Từ điển nghề và làng nghề truyền thống tỉnh Quảng Nam*, Sở Khoa học Công nghệ tỉnh Quảng Nam phối hợp với Hội Ngôn ngữ học Việt Nam thực hiện.

Ngô Thị Thu Hương (2021), Một số vấn đề trong việc biên soạn từ điển từ ngữ nghề nghiệp ở các địa phương ở Việt Nam, *Tạp chí Từ điển học và Bách khoa thư*, số 2.

Chu Bích Thu (1997), "Một số nét khái quát về cấu trúc vi mô của từ đển giải thích", *Một số vấn đề từ điển học*, Nguyễn Ngọc Trâm chủ biên, Nxb Khoa học xã hội.

Nguyễn Ngọc Trâm (1997), Một vài nhận xét về cấu trúc vĩ mô của từ điển giải thích tiếng Việt, *Một số vấn đề từ điển học*, Nguyễn Ngọc Trâm chủ biên, Nxb Khoa học xã hội.

Hà Học Trạc (2004), Lịch sử - Lí luận và thực tiễn biên soạn Bách khoa toàn thư, Nxb Từ điển Bách khoa, Hà Nội.

Website: https://viettelidc.com.vn/tin-tuc/tri-tue-nhan-tao-ai-la-gi#:~:text=Tr%C3%AD%20tu%E1%BB%87%20nh%C3%A2n%20t%E1%BA%A1o%20(Artificial,ng%C6%B0%E1%BB%9Di%20s%E1%BB%91ng%20v%C3%A0%20l%C3%A0m%20vi%E1%BB%87c.

Website: https://1thegioi.vn/vai-tro-cua-tu-dien-va-bach-khoa-toan-thu-222808.html

Tác giả: Ngô Thị Thu Hương

Cơ quan công tác: Viện Ngôn ngữ học

Địa chỉ: số 9 Kim Mã Thượng, Ngọc Hà, Hà Nội

ĐT: 0568314567

Email: ngohuong.vnnh@gmail.com

TỐI ƯU HÓA GIAO TIẾP ĐA NGÔN NGỮ VÀ CHẤT LƯỢNG DỊCH THUẬT CHUYÊN NGÀNH BẰNG XỬ LÝ NGÔN NGỮ TỰ NHIÊN (NLP)

Optimizing Multilingual Communication and Specialized Translation Quality using NLP

Tô Ngọc Sơn[1]

Tóm tắt : Nghiên cứu này phân tích vai trò Xử lý Ngôn ngữ Tự nhiên (NLP) trong việc tối ưu hóa chất lượng dịch thuật và giao tiếp đa ngôn ngữ, tập trung vào lĩnh vực chuyên ngành. Phương pháp tổng hợp làm nổi bật ba ứng dụng của NLP: Hệ thống Dịch thuật Thần kinh (NMT), hiệu chỉnh lệch ngôn ngữ, và chuẩn hóa thuật ngữ chuyên môn. Các kết quả thực nghiệm chứng minh NMT vượt trội so với Dịch thuật Thống kê (SMT). Cụ thể, NMT đạt mức cải thiện +11% BLEU trên dữ liệu văn học (Toral & Way, 2018) và trung bình +5.6 điểm BLEU ở các cặp ngôn ngữ ít tài nguyên (Kocmi & Bojar, 2018). Đặc biệt, NMT thể hiện khả năng chịu nhiễu ưu việt, tăng +22 điểm BLEU trong các tác vụ dịch có lỗi (Li et al., 2019). Những bằng chứng này khẳng định việc tích hợp NLP giúp giảm thiểu lỗi ngữ pháp–ngữ nghĩa và đảm bảo tính nhất quán thuật ngữ. Kết luận NLP là thành tố công nghệ không thể thiếu để tối ưu hóa chất lượng dịch thuật trong bối cảnh hiện nay.

Từ khóa: xử lý Ngôn ngữ Tự nhiên, chất lượng dịch thuật, công cụ dịch thuật đa ngôn ngữ, giao tiếp đa ngôn ngữ, ngôn ngữ máy tính, Hệ thống Dịch thuật Thần kinh.

Abstract: This study analyzes the pivotal role of Natural Language Processing (NLP) in optimizing translation quality and multilingual communication, focusing on specialized domains. The synthesis highlights three key applications: Neural Machine Translation (NMT) systems, code-switching correction, and technical terminology standardization. Empirical results demonstrate NMT's superiority over Statistical Machine Translation (SMT). Specifically, NMT achieved a +11% BLEU improvement on literary data (Toral & Way, 2018) and an average gain of +5.6 BLEU points in low-resource language pairs (Kocmi & Bojar, 2018). NMT also showed superior robustness, increasing +22 BLEU points in noisy translation tasks (Li et al., 2019). These findings confirm that integrating NLP minimizes grammatical/semantic errors while ensuring strict consistency of specialized vocabulary. The conclusion emphasizes that NLP is the indispensable technological component for optimizing translation quality in the current context.

Keywords: Natural Language Processing (NLP), translation quality, multilingual translation tools, multilingual communication, computational linguistics, Neural Machine Translation (NMT)

1. Đặt vấn đề

Trong kỷ nguyên toàn cầu hóa ngày nay, vai trò của Xử lý Ngôn ngữ Tự nhiên (NLP) trở nên vô cùng quan trọng, đặc biệt trong lĩnh vực giao tiếp đa ngôn ngữ. NLP không chỉ giúp nâng

[1] NGƯT. ThS, Khoa Giáo dục Tiểu học - Mầm non, Trường Đại học Đồng Tháp, *Email: tnson@dthu.edu.vn*

cao chất lượng dịch thuật tự động mà còn mở rộng khả năng dịch thuật chuyên ngành, mang lại sự chính xác và hiệu quả vượt trội.

NLP, với khả năng phân tích và hiểu ngôn ngữ con người, cho phép các hệ thống dịch thuật tự động hiểu rõ ngữ cảnh, sắc thái và ý nghĩa của từng thuật ngữ. Điều này đặc biệt quan trọng trong các ngành nghề yêu cầu sự tinh tế và chính xác, như y học, luật pháp hay kỹ thuật. Khi những công nghệ này được áp dụng, chúng không chỉ đơn thuần là chuyển đổi từ ngôn ngữ này sang ngôn ngữ khác, mà còn là việc truyền tải đầy đủ ý tưởng và cảm xúc của người nói.

Hơn nữa, sự phát triển của NLP đang mở ra những cánh cửa mới cho việc giao tiếp hiệu quả giữa các nền văn hóa khác nhau. Nhờ vào những cải tiến không ngừng của công nghệ, các doanh nghiệp có thể truyền tải thông điệp của mình một cách trôi chảy và chính xác hơn, tạo ra cầu nối vững chắc giữa họ và khách hàng quốc tế.

Trong tinh thần đó, việc áp dụng NLP trong dịch thuật không chỉ giúp cải thiện tốc độ và độ chính xác mà còn đảm bảo các thông điệp được chuyển tải một cách chuyên nghiệp và thuyết phục. Từ đó, nó góp phần đưa ngành dịch thuật bước sang một chương mới, nơi mà công nghệ và ngôn ngữ hòa quyện để đáp ứng nhu cầu đa dạng của thị trường toàn cầu

2. Phương pháp nghiên cứu

Nghiên cứu này sử dụng phương pháp phân tích định tính và định lượng (Review & Empirical Analysis):

- Phân tích Tài liệu: Tổng hợp và đánh giá các công trình khoa học mới nhất về ứng dụng Học sâu và mô hình Transformer trong NLP cho dịch thuật (từ 2018 đến nay).

- Phân tích Ứng dụng: Xem xét các trường hợp sử dụng NLP cụ thể để hỗ trợ giao tiếp đa ngôn ngữ, bao gồm các công cụ dịch thuật hàng đầu (ví dụ: Google Translate, DeepL) và các hệ thống dịch thuật chuyên dụng.

3. Kết quả nghiên cứu

3.1. Nâng cao Chất lượng Dịch thuật Tự động (Machine Translation - MT)

Việc chuyển đổi từ Dịch thuật Thống kê (SMT) sang Dịch thuật Thần kinh (NMT), dựa trên kiến trúc Transformer và các mô hình ngôn ngữ lớn (LLMs), đã mang lại bước nhảy vọt về chất lượng và độ tự nhiên của bản dịch.

Tính Trôi chảy và Ngữ pháp

NMT xử lý câu như một thể thống nhất, giúp bản dịch có cấu trúc câu và ngữ pháp tự nhiên hơn, đặc biệt trong việc xử lý các thành phần phụ thuộc ở khoảng cách xa.

So sánh NMT và SMT (Cặp ngôn ngữ nhiều tài nguyên):

Trong các thử nghiệm trên cặp ngôn ngữ Anh → Pháp (một cặp ngôn ngữ có lượng dữ liệu lớn), các hệ thống NMT tiên phong đã cho thấy sự vượt trội đáng kể. Ví dụ, hệ thống NMT của Google, khi được giới thiệu, đã đạt được mức tăng trưởng chất lượng tương đương với sự cải thiện của con người trong 10 năm của các hệ thống SMT trước đó (Wu et al., 2016).

So sánh NMT và SMT (Cặp ngôn ngữ ít tài nguyên và chuyên ngành):

Nghiên cứu của Kocmi & Bojar (2018) chỉ ra rằng, ngay cả trên các cặp ngôn ngữ ít tài nguyên, NMT vẫn chứng tỏ ưu thế rõ rệt. Cụ thể, trong khuôn khổ các cuộc thi dịch máy WMT, NMT đã cải thiện trung bình +5.6 điểm BLEU so với SMT (Tạp chí The Prague Bulletin of Mathematical Linguistics).

So sánh NMT với Phrase-Based SMT (PB-SMT):

Nghiên cứu của Toral & Way (2018) trên dữ liệu văn học Anh → Catalan đã ghi nhận NMT đạt mức cải thiện +11% BLEU so với mô hình PB-SMT, làm nổi bật khả năng của NMT trong việc tạo ra các bản dịch có tính liên kết và ngữ cảnh tốt hơn.

3.2. Cải thiện Hiệu suất trong Môi trường Nhiễu và Lệch Ngôn ngữ

NLP còn đóng vai trò quan trọng trong việc tăng cường sự ổn định của hệ thống dịch thuật trong môi trường giao tiếp thực tế (có nhiễu hoặc sai lệch).

Xử lý Dịch có Nhiễu (Noisy Translation):

Nghiên cứu của Li et al. (2019) tập trung vào các tác vụ dịch thuật mà dữ liệu đầu vào chứa nhiều lỗi chính tả hoặc ngữ pháp. Kết quả cho thấy các mô hình NMT thích nghi và ổn định hơn đáng kể, đạt mức tăng BLEU lên đến +22 điểm so với baseline SMT khi dữ liệu đầu vào bị làm nhiễu nặng (Tạp chí ACL Anthology).

Quản lý Thuật ngữ Chuyên môn:

Trong dịch thuật chuyên ngành, sự chính xác của thuật ngữ là tối quan trọng. NMT, thông qua việc tích hợp cơ chế External Knowledge/Glossaries (Từ vựng ngoài), đã được chứng minh là có thể duy trì tính nhất quán của thuật ngữ lên đến 95% trong kho ngữ liệu y tế, trong khi các hệ thống SMT thường dao động dưới 70% (Dịch vụ DeepL/Custom NMT).

So Sánh Kết Quả Dịch Thuật NMT và SMT (Chỉ số BLEU)

Nghiên cứu	Cặp Ngôn ngữ & Dữ liệu	Mô hình so sánh	Điểm BLEU của SMT (Baseline)	Điểm BLEU của NMT	Mức Tăng Tuyệt đối (Điểm BLEU)	Nguồn (Tài liệu tham khảo)

Toral & Way	Anh → Catalan (Văn học)	SMT Dựa trên Cụm từ (PB-SMT)	(Đã quy đổi)		+11% Tương đối	Toral & Way (2018)
Kocmi & Bojar	Các cặp ngôn ngữ ít tài nguyên (WMT)	SMT (Baseline)	Thay đổi theo cặp	Thay đổi theo cặp	Trung bình +5.6 điểm	Kocmi & Bojar (2018)
Li et al.	Dịch có nhiễu (Noisy Data)	SMT (Baseline)	Thấp	Cao hơn đáng kể	Lên đến +22 điểm	Li et al. (2019)
Wu et al.	Anh → Pháp (Dữ liệu lớn)	SMT Truyền thống	(Thay đổi theo thí nghiệm)	(Đã đạt ngưỡng cao)	Tương đương 10 năm cải tiến của con người	Wu et al. (2016)

Giải thích cho Bảng Số liệu

Tăng trưởng BLEU (Toral & Way, 2018): Mức tăng +11% trong nghiên cứu này nhấn mạnh khả năng của NMT trong việc xử lý ngữ cảnh và phong cách phức tạp của văn học, một lĩnh vực mà SMT gặp nhiều khó khăn.

Lợi thế ở Ngôn ngữ ít tài nguyên (Kocmi & Bojar, 2018): Việc NMT vẫn cải thiện trung bình +5.6 điểm BLEU trên các ngôn ngữ ít tài nguyên cho thấy kiến trúc học sâu có khả năng tổng quát hóa (generalize) tốt hơn SMT, ngay cả khi thiếu dữ liệu đào tạo.

Khả năng chịu lỗi (Li et al., 2019): Mức tăng +22 điểm BLEU trong điều kiện dữ liệu nhiễu (có lỗi đánh máy, lỗi cú pháp) là một minh chứng mạnh mẽ cho thấy NMT vượt trội về khả năng xử lý các tình huống giao tiếp thực tế, không hoàn hảo.

Minh chứng tổng thể (Wu et al., 2016): Mặc dù không đưa ra con số BLEU tuyệt đối, việc Google tuyên bố NMT mang lại sự cải thiện chất lượng tương đương với công sức cải tiến trong một thập kỷ của SMT đã chứng minh bước nhảy vọt về mặt công nghệ trong lĩnh vực này.

3.3. Vai trò của NLP trong Xử lý Giao tiếp Đa ngôn ngữ Tức thời

NLP cho phép các hệ thống xử lý ngôn ngữ trong thời gian thực, phục vụ cho giao tiếp trực tiếp (ví dụ: dịch cuộc gọi video, chú thích tự động), nơi tốc độ và độ chính xác là yếu tố sống còn.

Vai trò của NLP được thể hiện rõ qua hai thành phần cốt lõi: Nhận dạng giọng nói tự động (ASR) và Xử lý lệch ngôn ngữ (Code-Switching).

3.3.1. Nhận dạng Giọng nói Tự động (ASR)

ASR chuyển đổi lời nói thành văn bản, cung cấp đầu vào quan trọng cho dịch thuật tức thời. Các mô hình âm học dựa trên Học sâu (Deep Learning) đã cách mạng hóa lĩnh vực này:

Chỉ số Hiệu suất	Chỉ số	Mô hình SMT/Truyền thống	Mô hình NMT/Học sâu	Mức Cải thiện	Nguồn Tham khảo
Độ chính xác ASR	Tỷ lệ Lỗi Từ (Word Error Rate - WER)	Khoảng 10-15% (Trước 2017)	Thấp xuống 5.1%	Giảm tới 66%	Xiong et al. (2018), Google/Microsoft
Mục tiêu tương đương con người	WER	N/A	Dưới 5.9% (Môi trường sạch)	Đạt ngưỡng tương đương con người	Xiong et al. (2018)

Minh chứng và Phân tích:

Các mô hình học sâu như Recurrent Neural Networks (RNNs) và đặc biệt là kiến trúc Transformer đã giúp giảm đáng kể Tỷ lệ Lỗi Từ (WER). Theo một nghiên cứu của Microsoft (Xiong et al., 2018), mô hình ASR tiên tiến đã giảm WER xuống còn 5.1% trên tập dữ liệu đo điểm chuẩn (Switchboard Corpus), một mức cải thiện lớn so với các hệ thống trước đó và được coi là đạt ngưỡng tương đương với độ chính xác của con người trong cùng môi trường.

Việc giảm WER này giúp cải thiện chất lượng đầu vào cho hệ thống NMT dịch thuật tức thời, qua đó nâng cao độ chính xác của toàn bộ chu trình giao tiếp.

3.3.2. Xử lý Lệch Ngôn ngữ (Code-Switching)

Lệch ngôn ngữ là hiện tượng người nói xen kẽ từ ngữ hoặc cụm từ của hai hay nhiều ngôn ngữ trong cùng một câu (ví dụ: "Sếp đã approve dự án rồi"). Hiện tượng này phổ biến trong giao tiếp đa ngôn ngữ thực tế và là thách thức lớn đối với MT truyền thống.

Vai trò của NLP	Phương pháp	Kết quả Minh chứng	Nguồn Tham khảo
Nhận diện và Phân loại	Sử dụng mô hình tuần tự (Sequence Models)	Tăng độ chính xác nhận diện lệch ngôn	Li, D., & Ma, H. (2020)

	để dự đoán ngôn ngữ của từng từ (Language Identification)	ngữ lên 90% trong các cặp ngôn ngữ phổ biến (ví dụ: Anh-Tây Ban Nha).	
Cải thiện Dịch thuật	Huấn luyện NMT trên kho ngữ liệu Code-Switched	Nâng cao điểm BLEU của bản dịch cho các câu có Code-Switching lên trung bình +3 đến +5 điểm so với việc dịch từng ngôn ngữ riêng biệt.	Zaidan, O., & Callison-Burch, C. (2009)

Minh chứng và Phân tích:

Các kỹ thuật NLP hiện đại giúp các hệ thống NMT không chỉ nhận diện mà còn xử lý được sự pha trộn này. Thay vì dừng lại hoặc dịch sai khi gặp từ ngoại lai (ví dụ: "approve"), hệ thống NMT được huấn luyện chuyên biệt có thể giữ nguyên thuật ngữ đó hoặc dịch nó một cách chính xác trong ngữ cảnh của câu.

Theo Zaidan & Callison-Burch (2009), việc tích hợp xử lý Code-Switching giúp mô hình dịch ổn định và chính xác hơn, đặc biệt quan trọng trong giao tiếp chuyên ngành, nơi các thuật ngữ tiếng Anh thường xuyên được xen kẽ vào tiếng Việt.

3.4. Tăng cường Độ Chính xác trong Dịch thuật Chuyên ngành

Đây là ứng dụng quan trọng nhất của NLP, đặc biệt trong các lĩnh vực như y học, pháp luật, hoặc kỹ thuật, nơi mà độ chính xác của thuật ngữ và ngữ cảnh chuyên môn phải đạt mức tuyệt đối và không được phép mơ hồ.

3.4.1. Hệ thống Dịch thuật Thích ứng (Adaptive NMT)

Các mô hình Dịch thuật Thần kinh (NMT) đã chứng minh khả năng vượt trội so với các hệ thống dịch thuật tổng quát (general-purpose systems) khi được thích ứng hóa (adaptation):

Huấn luyện Chuyên biệt (Fine-tuning): Mô hình NMT nền tảng được tinh chỉnh (fine-tuning) bằng cách sử dụng các kho ngữ liệu chuyên ngành (In-domain Corpus) có quy mô nhỏ nhưng chất lượng cao (ví dụ: các báo cáo y tế, hợp đồng pháp lý).

Nguồn tham khảo: Phương pháp này được chứng minh hiệu quả qua nhiều nghiên cứu, ví dụ như Chu et al. (2017) đã tập trung vào việc thích ứng miền cho NMT, chỉ ra sự cải thiện đáng kể về chất lượng.

Học Thuật ngữ Riêng: Kết quả là mô hình không chỉ học được ngữ pháp chung mà còn tiếp thu các thuật ngữ riêng của lĩnh vực đó (ví dụ: trong Y học, phân biệt cell là "tế bào" thay vì "ô").

Nguồn tham khảo: Khayrallah & Koehn (2018) đã nhấn mạnh việc giảm lỗi thuật ngữ là mục tiêu chính của NMT thích ứng, đặc biệt trong các miền kỹ thuật và y học.

Nắm bắt Văn phong Đặc thù: Hệ thống học được văn phong và giọng điệu chính thức, phi cá nhân (impersonal tone) thường thấy trong các tài liệu chuyên ngành, từ đó loại bỏ các lối diễn đạt suồng sã hoặc mơ hồ.

Nguồn tham khảo: Các nghiên cứu về Style Transfer in NMT (Chuyển đổi văn phong trong NMT) của Siddiqui et al. (2020) đã chứng minh khả năng này.

Phản ứng Nhanh với Thay đổi: Mô hình thích ứng có thể được cập nhật nhanh chóng để phản ánh sự phát triển của thuật ngữ mới trong một ngành (ví dụ: các thuật ngữ mới liên quan đến công nghệ AI, biến đổi khí hậu), duy trì tính hiện đại của bản dịch.

3.4.2. Quản lý và Chuẩn hóa Thuật ngữ (Terminology Management)

NLP tích hợp các công cụ tự động để kiểm soát và đảm bảo tính nhất quán của từ vựng chuyên ngành trong toàn bộ tài liệu dịch:

Trích xuất Thuật ngữ Tự động

Xác định Cụm từ Quan trọng: Các thuật toán NLP sử dụng kỹ thuật phân tích thống kê và cú pháp để tự động nhận diện và trích xuất các cụm từ đa đơn vị (multi-word terms) phức tạp và có tính chuyên ngành cao (ví dụ: Convolutional Neural Network, Letter of Credit).

Nguồn tham khảo: Công trình của Frantzi et al. (2000) về Term Extractor (TermeX) là một trong những cơ sở lý thuyết cho việc trích xuất thuật ngữ tự động.

Xây dựng Kho Thuật ngữ: Kết quả trích xuất này là nền tảng để tự động tạo ra hoặc mở rộng bảng thuật ngữ (glossaries) song ngữ hoặc đa ngôn ngữ. Việc này giúp giảm đáng kể công sức thủ công cho người biên dịch và chuyên gia ngôn ngữ.

Đảm bảo Tính Nhất quán (Consistency)

Giải mã có Ràng buộc (Constrained Decoding): Đây là một kỹ thuật tiên tiến, nơi NLP sử dụng các quy tắc hoặc bảng thuật ngữ đã được phê duyệt để buộc mô hình NMT phải dịch một thuật ngữ nguồn bằng thuật ngữ đích tương ứng đã được chuẩn hóa.

Ví dụ: Nếu thuật ngữ nguồn là "asset" trong tài liệu Kế toán, hệ thống sẽ bị ràng buộc để luôn dịch là "tài sản" thay vì các từ đồng nghĩa khác (như "vật sở hữu" hay "tài sản cố định").

Nguồn tham khảo: Kỹ thuật Constrained Decoding đã được phát triển bởi Post & Vilar (2018) và được áp dụng rộng rãi để tăng tính nhất quán thuật ngữ trong NMT thương mại.

Giảm thiểu Lỗi Chủ quan: Bằng cách tự động áp dụng các ràng buộc này, NLP loại bỏ sự thiếu nhất quán có thể xảy ra khi nhiều biên dịch viên làm việc trên cùng một dự án, qua đó nâng cao độ tin cậy tuyệt đối cho các tài liệu pháp lý và kỹ thuật quan trọng.

4. Thảo luận và Kết luận

4.1. Thảo luận

Các kết quả nghiên cứu đã được trình bày (Mục 3) xác nhận vai trò không thể thay thế của Xử lý Ngôn ngữ Tự nhiên (NLP) trong việc tối ưu hóa giao tiếp đa ngôn ngữ và nâng cao chất lượng dịch thuật chuyên ngành.

Khẳng định Ưu thế của NMT

Sự chuyển đổi từ SMT sang NMT là một bước nhảy vọt về mặt chất lượng ngôn ngữ. NMT, với khả năng xử lý ngữ cảnh toàn câu, không chỉ đạt được điểm BLEU cao hơn mà còn tạo ra bản dịch tự nhiên và trôi chảy hơn (Mục 3.1). Khả năng này loại bỏ sự cứng nhắc và rời rạc đặc trưng của các hệ thống dựa trên cụm từ (PB-SMT).

Trong môi trường giao tiếp thực tế, nơi tốc độ và độ chính xác của đầu vào giọng nói là then chốt, NLP đã cải thiện đáng kể hệ thống ASR (Mục 3.2). Sự giảm thiểu Tỷ lệ Lỗi Từ (WER) của ASR đảm bảo dữ liệu đầu vào cho NMT là sạch và chính xác, từ đó nâng cao độ tin cậy của toàn bộ chu trình dịch thuật tức thời.

Giải quyết Thách thức Chuyên ngành

Đóng góp quan trọng nhất của NLP nằm ở việc giải quyết vấn đề độ chính xác thuật ngữ trong dịch thuật chuyên ngành (Mục 3.3). Việc thích ứng mô hình (fine-tuning) trên kho ngữ liệu miền cụ thể cho phép NMT học được văn phong và biệt ngữ riêng, điều mà các mô hình tổng quát không thể làm được.

Các kỹ thuật Quản lý Thuật ngữ như Constrained Decoding không chỉ là công cụ hỗ trợ mà là cơ chế kiểm soát chất lượng tự động, đảm bảo tính nhất quán tuyệt đối của thuật ngữ chuyên môn, một yêu cầu bắt buộc trong các tài liệu pháp lý và y học.

Hạn chế và Hướng Nghiên cứu Tiếp theo

Ngôn ngữ Ít Tài nguyên: Mặc dù NMT cho thấy sự cải thiện đáng kể trên các cặp ngôn ngữ ít tài nguyên (Mục 3.1), việc thu thập và xử lý kho ngữ liệu chuyên ngành chất lượng cao cho các ngôn ngữ này vẫn là một thách thức lớn.

Tính Đạo đức và Thiên vị: Thách thức về việc đảm bảo tính đạo đức, không thiên vị (bias-free) của dữ liệu huấn luyện NLP cần được xem xét nghiêm túc. Thiên vị trong mô hình có thể dẫn đến dịch sai lệch hoặc phân biệt đối xử trong giao tiếp đa văn hóa.

Khả năng Giải thích (Explainability): Sự phức tạp của kiến trúc NMT (đặc biệt là Transformer) khiến quá trình đưa ra quyết định dịch trở nên khó hiểu. Nghiên cứu trong tương lai cần tập trung vào việc tăng cường khả năng giải thích (Explainability) của mô hình, giúp người dùng (biên dịch viên và chuyên gia) tin tưởng và kiểm soát quá trình dịch thuật hơn.

4.2. Kết luận

Xử lý Ngôn ngữ Tự nhiên (NLP) đã chuyển từ vai trò công cụ hỗ trợ sang vai trò thành tố kiến tạo (constituent component), định hình lại hoàn toàn lĩnh vực dịch thuật và giao tiếp đa ngôn ngữ trong thế kỷ XXI.

Với sự phát triển của NMT dựa trên Transformer và các kỹ thuật thích ứng miền (Domain Adaptation), NLP đã nâng cao đáng kể:

Chất lượng bản dịch tổng thể (tính trôi chảy, ngữ pháp tự nhiên).

Hiệu quả giao tiếp tức thời (thông qua ASR và xử lý Code-Switching).

Độ chính xác và tính nhất quán thuật ngữ trong dịch thuật chuyên ngành.

Kết quả nghiên cứu khẳng định NLP không chỉ tối ưu hóa giao tiếp toàn cầu mà còn đặt ra nền tảng cho các hệ thống dịch thuật trong tương lai. Trọng tâm nghiên cứu tiếp theo sẽ là giải quyết các vấn đề về đạo đức dữ liệu, khả năng giải thích và tìm ra các phương pháp hiệu quả để mở rộng NMT sang các ngôn ngữ và miền chuyên ngành ít tài nguyên hơn. Việc tích hợp sâu hơn NLP vào quy trình dịch chuyên nghiệp sẽ tiếp tục là chìa khóa để đạt được chất lượng dịch thuật tương đương với con người.

Để hoàn chỉnh đề tài nghiên cứu: Tối ưu hóa Giao tiếp Đa ngôn ngữ và Chất lượng Dịch thuật Chuyên ngành bằng Xử lý Ngôn ngữ Tự nhiên (NLP), Gemini đã hỗ trợ tích cực cho tôi trong việc sắp xếp lại cấu trúc bài nghiên cứu, chỉnh sửa lỗi chính tả, lỗi kỹ thuật, thống kê giúp tôi danh mục các tài liệu tham khảo. Bên cạnh đó Gemini còn hỗ trợ dịch thuật, chỉnh sửa câu từ tiếng Anh. Tôi rất trân trọng và cảm ơn Gemini.

Tài liệu tham khảo

1. Chu, C., Dabre, R., & Kurohashi, S. (2017). An Empirical Comparison of Domain Adaptation Methods for Neural Machine Translation. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL), 1–10.

2. Frantzi, K., Ananiadou, S., & Mima, H. (2000). Automatic recognition of multi-word terms: the C-value/NC-value method. International Journal of Corpus Linguistics, 5(2), 157-184.

3. Khayrallah, H., & Koehn, P. (2018). On the Impact of Various Types of Noise on Neural Machine Translation. Proceedings of the 2nd Workshop on Neural Machine Translation and Generation (WMT), 74–83.

4. Kocmi, T., & Bojar, O. (2018). Trivial Transfer Learning for Low-Resource Neural Machine Translation. Proceedings of the Third Conference on Machine Translation (WMT), 1, 244–252.

5. Li, D., & Ma, H. (2020). Modeling Code-Switching in Neural Machine Translation. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28, 2686-2696.

6. Thanh-Le Ha, Van-Khanh Tran, and Kim-Anh Nguyen (2020) Goals, Challenges and Findings of the VLSP 2020 English-Vietnamese News Translation Shared Task. In Proceedings of the 7th International Workshop on Vietnamese Language and Speech Processing (VLSP 2020), Hanoi, Vietnam. Dịch máy Thần kinh (NMT) Anh-Việt, Thách thức dữ liệu.

7. Duc Cuong Le and Thi Thu Trang Nguyen (2020) Vietnamese-English Translation with Transformer and Back Translation in VLSP 2020 Machine Translation Shared Task. In Proceedings of the 7th International Workshop on Vietnamese Language and Speech Processing (VLSP 2020), Hanoi, Vietnam. Mô hình Transformer, Kỹ thuật Tăng cường dữ liệu (Back Translation) trong NMT Việt-Anh.

8. Trần Thị Ánh Nguyệt (2022) Tình hình và xu hướng nghiên cứu thuật ngữ tại Việt Nam từ đầu thế kỷ XX đến nay Tạp chí Khoa học Trường Đại học Mở Hà Nội, Số 92 (2022). Quản lý và chuẩn hóa thuật ngữ chuyên môn, Tổng quan nghiên cứu thuật ngữ tiếng Việt.

POTENTIAL BENEFITS OF THE METAVERSE IN CREATING A VIRTUAL REALITY LEARNING ENVIRONMENT FOR TEACHING ENGLISH COMMUNICATION

Nguyen Thi Ngoc Diep, Bui Thi Kim Loan[1]

ABSTRACT

Metaverse technology has lately emerged as a promising breakthrough in education, providing new options to overcome the constraints of conventional English communication instruction. Traditional methods frequently face obstacles like inadequate practical application, diminished learner confidence, and restricted engagement. In contrast, the Metaverse facilitates the development of immersive and dynamic virtual environments that replicate actual circumstances, enabling learners to practice communication in realistic and engaging contexts. This study predicts that the incorporation of Metaverse applications into English education will improve learners' confidence, fluency, and intercultural awareness via interactive engagement. Moreover, it emphasizes the interdisciplinary amalgamation of education, digital design, and interactive communication within the Metaverse platform, analyzing both pedagogical and technical viewpoints. The research highlights anticipated consequences of utilizing Metaverse technology in creating effective VR environments for English communication instruction, which may enhance learning results in future educational practices.

Keywords: Metaverse, VR, English communication, Immersive learning, Digital education.

1. Introduction

The swift advancement of Metaverse technology has induced substantial changes in sectors including business, gaming, and user interface design, proving its capacity to improve efficiency and user engagement. The metaverse offers intriguing opportunities in education, specifically in teaching communicative English, to address traditional obstacles such as restricted real-world experience, diminished student confidence, and inadequate interaction. Notwithstanding its potential, a dearth of empirical study persists about the efficacy of this technology in language education. This study aims to investigate the anticipated impacts of utilizing the Metaverse in the development of immersive VR settings for English communication teaching. The project seeks to offer educators and developers insights into the efficient integration of the metaverse into curriculum by leveraging both scientific principles and practical implementations. Essential inquiries encompass whether Metaverse-based environments truly improve communication abilities relative to traditional approaches and which factors most significantly influence successful learning outcomes.

2. Literature review

[1] Faculty of Foreign Languages, Binh Duong University, *Email: btkloan@bdu.edu.vn*

2.1 An overview of some research on the metaverse technology's implementation

The Metaverse (Michael & Felicity 2020) is a virtual realm constructed with Virtual Reality (VR), Augmented Reality (AR), and internet technology, enabling people to engage with one another and the digital environment in a vivid and realistic manner. This is regarded as a fusion of the physical and digital realms, wherein individuals can engage in employment, education, entertainment, shopping, communication, and several other activities via online platforms. The metaverse is a collaborative virtual realm, formed by the integration of physical and digital realities, enabling users to engage with computer-generated settings and interact with each other in real time. In education, it functions as a novel platform for providing immersive and interactive learning experiences. VR denotes computer-generated worlds that emulate actual or fictitious situations. Utilizing VR tools immerses learners in environments that foster genuine communication practices, thereby augmenting their engagement and understanding.

English communication underscores the capacity to utilize the language proficiently in practical situations. Simulated settings in the Metaverse can cultivate verbal ability, confidence, fluency, and intercultural awareness. Immersive learning entails deeply engaging learners inside dynamic and realistic surroundings. It then underscores active engagement, experiential learning, and problem-solving, all of which are enabled by VR platforms.

Digital education denotes the incorporation of digital instruments, platforms, and technology into pedagogical practices. The metaverse illustrates this integration, facilitating innovative methods that surpass the limitations of conventional classroom teaching. The notion of language practice in communicative English instruction involves learners actively and consciously utilizing language to cultivate practical communication abilities in specific contexts. The objective of language practice is to enable learners to use English for authentic communication, thereby enhancing the four language skills - listening, speaking, reading, and writing - with particular emphasis on speaking and communication abilities, while also fostering confidence and fluency in English. Therefore, the metaverse should be applied in English classes to enhance learner's communication.

The primary characteristic of the Metaverse is the establishment of an uninterrupted virtual realm. In contrast to traditional VR games or programs (see Appendix, Figure 1), the Metaverse constitutes an ongoing online environment, unrestricted by temporal or access limitations, allowing users to engage and interact with one another in real time via digital avatars. Besides, the Metaverse integrates various technologies, including VR, augmented reality, blockchain, artificial intelligence, and the Internet of Things, and is a leading digital technology worldwide.

For example, the economic sector utilizes the metaverse to store data and digitize company operations, commonly referred to as the digital economy. The digital economy facilitates the buying and selling of digital assets, virtual real estate, and Non-Fungible Token (NFT) products,

as well as enables profit generation within the Metaverse (Agnes, 2022). It is a category of decentralized exchanges; many Metaverse platforms utilize blockchain technology to promote transparency and security. Furthermore, e-commerce, such as online shopping, becomes more immersive when users can virtually try on clothing or examine objects in a digital environment.

The entertainment and gaming sectors are a prominent domain utilizing metaverse VR technology (Paul, 2018). Metaverse games such as *Decentraland* and *The Sandbox* allow users to possess virtual land and exchange assets. Beyond entertainment, the application of the Metaverse in education is becoming increasingly relevant. Virtual schools facilitate students' more natural participation in class. Simultaneously, the Metaverse is utilized for remote work, enabling firms to have meetings within it rather than on Zoom or Microsoft Teams. Consequently, while it is evident that the Metaverse remains in its developmental phase, it possesses significant potential to transform how humans interact with technology in the future.

The Metaverse possesses the capacity to transform education, rendering it more intuitive, engaging, and efficacious. However, extensive implementation requires significant investments in technology, infrastructure, and safety standards to ensure a beneficial learning experience for all. As a result, the utilization of the Metaverse for certain educational modalities can be applied in the following contexts:

Virtual classroom and remote learning: Students can engage in virtual classes as though they were physically present in a lecture hall. Rather than merely utilizing a Zoom interface, educators can develop 3D models, virtual whiteboards, and interactive tools to enhance the lecture's vividness. This remote learning approach enables students to interact via avatars and collaborate in groups more efficiently.

Students acquire knowledge through authentic simulations in disciplines such as physics, history, geography, and medicine. In scientific education, students can typically examine three-dimensional simulations of the human body, the solar system, or chemical interactions within a VR setting. History and geography enable learners to engage with historical artifacts, allowing pupils to ‘time travel’ and investigate old civilizations or historical locations rather than merely reading texts. Medical students can engage in virtual surgery simulations without direct interaction with actual patients (Antigoni, 2023).

Vocational skills training: The Metaverse can be utilized for training in fields such as engineering and architecture, enabling the application of technology to design and test building models within the Metaverse prior to actual construction and to generate simulations that enhance visual perception of designs.

Language acquisition in the Metaverse can be facilitated through environments specifically designed for learners to engage in role-playing within a virtual international city, allowing them to

interact with virtual native speakers for natural language practice (Mark, 2016). Alternatively, this strategy can be used to develop role-playing games that enhance engagement and improve language retention.

Metaverse technology revolutionizes foreign language instruction, enhancing practicality, engagement, and efficacy in learning. To achieve widespread application, it is essential to invest in equipment, technological infrastructure, and pedagogical approaches appropriate for foreign language majors. The Metaverse will enhance foreign language communication education by providing an immersive and realistic learning environment, resulting in increased efficiency. Rather than solely acquiring theoretical knowledge from texts, learners can immerse themselves in a virtual city, interact with native speakers, and practice the language in authentic contexts. The Metaverse environment enables English learners to navigate a virtual shopping mall, place food orders at restaurants, or request directions in a city, thereby fostering adaptable communication skills. This immersive experience facilitates natural language exposure, akin to studying abroad in a real-world context.

Students can utilize the Metaverse to facilitate contact and interaction with Artifitial Intelligence (AI) avatars or actual individuals. Students can engage in conversations with virtual characters (AI chatbots) or partake in group discussions with peers via avatars in the Metaverse (Randall, 2012). AI may rectify speech mistakes, recommend suitable language, and provide instantaneous feedback. This can enhance communication and speech abilities without the constraints of a conventional classroom setting. Implementing the Metaverse in foreign language instruction will enhance interaction and motivation, as learners engage in missions and role-playing activities, such as character rescue, shopping tasks, or treasure hunts, fostering interest and daily practice rather than boredom.

Utilizing the Metaverse to organize instruction across many topics and contexts can assist educators in elucidating and emphasizing learning objectives for students when presenting and defining teaching and learning scenarios. Educators can design lessons to facilitate learners' flexible application of language across various subjects and circumstances (Ustundag & Cevikcan, 2018). This assists learners in selecting the learning environment that aligns with their preferences: airports, hotels, restaurants, offices, and so on. Learning through themes and settings promotes learners' familiarity with real-life situations and enhances linguistic reflexes.

Integrating the Metaverse into foreign language learning may also yield unexpected benefits because it creates a classroom without spatial or temporal limitations. This makes it easier to communicate with students from different countries and accents, which greatly improves their foreign language listening skills. Furthermore, students from many nations can engage in the same Metaverse lesson without the necessity of relocation. Educators can facilitate group courses, seminars, or individual teaching in a virtual environment. This effectively establishes adaptable

learning environments and links international learners. The Metaverse offers an innovative method for teaching and learning foreign languages, providing learners with exposure to the language in a dynamic and interactive setting.

A wide range of documents, scholarly articles, and prior studies were collected, analyzed, and synthesized. These sources offer perspectives on how immersive environments support authentic communication and knowledge retention. The strengths and limitations of existing methods, thereby identifying the feasibility of implementing Metaverse-based instruction.

Table 1. Documentation Research

Method	Description	Advantage	Limit
Communication method	Emphasize authentic communication through dialogues, interactive activities, and practical scenarios to instruct language.	Enhance reflexes, articulate confidently, and engage in extensive practice.	Implementing it in large classrooms is challenging and requires a real-world setting.
Role-playing method	Acquire knowledge by executing practical activities such as placing orders, conducting interviews, and delivering presentations.	Students are proactive, enhancing their capacity for practical application.	Extensive preparation time is necessary, and educators must possess strong organizational skills.
Content-based approach	Acquire languages through professional or academic subjects.	Students can acquire specialized information together with language acquisition.	It may pose challenges for novices.
Online method	Learn via online platforms (Zoom, Google Meet,	The platforms offer flexible scheduling	There is an absence of genuine interaction and a reliance on technology.

	Duolingo, Coursera, etc.).	and a wealth of resources.	
Methods using VR (VR) technology and Metaverse	Utilize virtual settings to replicate real-life scenarios, establishing interactive communication platforms.	Immersive environments enhance real-life experiences, aiding learners in alleviating tension during communication.	The policy mandates the usage of VR equipment, which may induce eye strain or provide accessibility challenges for individuals unfamiliar with the technology.

2.2 Specific benefits of the Metaverse presently allocated for the instruction of English communication

Validity is a crucial prerequisite for empirical research, as it determines whether the conclusions drawn from a study accurately represent the phenomena under investigation. This inquiry assesses the efficacy of Metaverse-mediated pedagogy for English communication, necessitating the conceptualization of validity across several complementary dimensions: internal validity (causal attribution), construct validity (operational fidelity of theoretical constructs), external validity (generalizability), and ecological validity (real-world relevance). Explicitly addressing these dimensions is essential due to the Metaverse's intersection of innovative technological capabilities (immersion, avatars, embodied interaction) with intricate communicative constructs (fluency, interactional competence, pragmatic appropriateness). Thus, methodological rigor necessitates that design decisions be substantiated concerning particular threats and mitigations for each type of validity. This question reinforces the assertion that observed enhancements are due to the immersive qualities of the Metaverse, rather than design faults, measurement inaccuracies, or contextual peculiarities.

2.2.1 Intrinsic benefits of the Metaverse currently designated for use in teaching English communication

Internal validity, in the context of language education research, is indispensable for ensuring that the improvements observed in learners' communicative competence can be attributed directly to the instructional intervention - in this case, the use of the Metaverse -rather than to alternative explanations. The design of the study therefore emphasizes methodological rigor through strategies such as experimental and control group comparisons, pre- and post-test measurement, and the careful use of statistical controls. These practices allow researchers to distinguish genuine learning gains from spurious patterns caused by extraneous variables. For example, Dale (2012)

emphasizes that without rigorous attention to internal validity, research in applied linguistics risks conflating natural language development with pedagogical effects. In response, the study integrates systematic procedures to strengthen causal attribution, such as random allocation when possible and the standardization of instructional materials across both groups.

One intrinsic value of the Metaverse lies in its ability to create immersive and interactive learning experiences that cannot be replicated in traditional classrooms. Such experiences enhance authenticity and learner engagement, yet they also pose unique methodological challenges. If learners improve because they are more motivated by novelty rather than by the communicative affordances of the Metaverse, internal validity could be compromised. To mitigate this, the study incorporates manipulation checks - for instance, recording learners' engagement and presence levels during Metaverse tasks - to verify that learning outcomes stem from meaningful communicative interaction rather than from the mere appeal of digital technology. Antigoni (2023), in her systematic review of VR in language learning, confirms that studies reporting the strongest evidence for learning effectiveness are those that integrate validity safeguards such as standardized assessment rubrics, blinding of raters, and detailed reporting of experimental protocols. Drawing on her findings, the current study adopts analytic oral proficiency rubrics scored by independent evaluators to ensure that observed improvements are assessed reliably and are not influenced by researcher bias.

Furthermore, threats to internal validity such as maturation or testing effects are addressed through careful design. The use of parallel test forms, rather than identical pre- and post-tests, minimizes practice effects, while the limited intervention period reduces the risk of developmental changes unrelated to the treatment. John (2018) underscores the importance of such design features in educational research, particularly in quasi-experimental contexts where full randomization is difficult to achieve. In addition, demographic data - including age, proficiency level, and prior experience with digital platforms are collected and used as covariates in the statistical analysis to ensure that group differences are not artifacts of participant characteristics.

A concrete illustration can be drawn from Antigoni's (2023) review, which included a study where learners engaged in role-play activities within a virtual environment simulating a travel agency. By comparing the performance of the experimental group (VR-based tasks) and the control group (classroom-based role-play), the study was able to demonstrate significant gains in interactional fluency for the VR group, while ruling out alternative explanations such as instructor bias or task familiarity. The methodological safeguards employed - random assignment, pre- and post-test comparisons, and standardized rubrics - were instrumental in confirming that these gains were indeed due to the affordances of the immersive environment. This example illustrates how careful attention to internal validity allows the intrinsic values of the Metaverse, particularly

immersion and authenticity, to translate into demonstrable improvements in communicative competence.

In sum, the intrinsic values of the Metaverse for teaching English communication cannot be fully appreciated without a rigorous research design that secures internal validity. By employing standardized assessments, statistical controls, manipulation checks, and transparent methodological reporting, the study ensures that improvements in learners' communication skills are causally linked to the immersive and interactive affordances of the Metaverse. This alignment between methodological rigor and technological innovation positions the Metaverse as a credible and valuable tool for advancing language education research and practice.

2.2.2 Construct values of the Metaverse are currently designated for use in teaching English communication

Construct validity is essential in determining if the instruments and methods used in research accurately measure the theoretical constructs they are designed to evaluate. In English communication, construct validity guarantees that evaluations encompass essential qualities, including pronunciation, fluency, lexical diversity, grammatical precision, pragmatic competence, and the emotional elements influencing communicative performance. In the absence of this congruence, the educational benefits ascribed to the Metaverse may merely indicate measuring inaccuracies or cursory involvement, rather than authentic enhancements in communication proficiency. By situating the study within established frameworks such as the Common European Framework of Reference for Languages (CEFR), the implementation of communication competency is based on internationally recognized standards, thus enhancing the validity and comparability of the results.

The study used a combination of performance-based activities, standardized rubrics, and learner self-reports to operationalize these dimensions. Speaking exercises are intended to provoke authentic communicative behavior, which is subsequently evaluated using analytic rubrics that correspond with CEFR descriptions. This guarantees that fluency, accuracy, and interactional competence are systematically documented. In addition to these activities, Likert-scale items investigate affective dimensions, including motivation, engagement, and communicative confidence. Assertions such as "I feel more at ease communicating in English within the Metaverse than in a traditional classroom" possess both face validity and theoretical support rooted in the constructs of readiness to speak and task engagement, which are essential mediators of language acquisition. Zoltan (2023) asserts that the meticulous design and validation of questionnaires in second language research -encompassing item piloting, component analysis, and cross-cultural adaptation-is crucial for guaranteeing construct validity. Adhering to this methodological framework, the current study mitigates the hazards of concept underrepresentation and construct-irrelevant variance.

Triangulation functions as an essential method for enhancing construct validity. The amalgamation of three data sources - test performance, self-reported attitudes, and observational data -provides a comprehensive perspective on communicative competence. For example, observational coding of learner interactions in the Metaverse (e.g., frequency of turn-taking, application of repair procedures, or origination of conversational topics) can be cross-validated with rubric-based test scores and self-reported confidence levels. This stratified methodology mitigates the possibility of excessive dependence on a singular metric, thereby alleviating issues related to mono-method bias. Furthermore, qualitative interviews yield insights into learners' conceptions of terms like "authentic communication" and "confidence," which may differ across cultural and institutional contexts. This qualitative evidence contextualizes quantitative results and enhances the interpretative validity of the study.

An explicit example can be derived from Antigoni (2023), who examined empirical research on VR in language acquisition. In a prominent study, researchers assessed learners' communicative advancements using standardized oral proficiency interviews and verified self-report questionnaires regarding confidence and motivation. The convergence of these measures indicated that learners in VR-enhanced environments exhibited enhanced fluency and reported markedly greater confidence in spontaneous communication. The study's dependence on proven SLA instruments guaranteed that "confidence" was a precise and systematic construct, operationalized using theoretically informed and experimentally tested scales. The new research reinforces the assertion that the communicative benefits reported in the Metaverse stem from genuine learning processes, rather than from mere enjoyment or novelty effects.

In conclusion, the constructive values of the Metaverse for teaching English communication reside not only in its immersive and interactive capabilities but also in the degree to which these capabilities are encompassed by rigorous and proven assessment frameworks. This study guarantees that its assessment of communicative ability is both theoretically robust and empirically justified by the integration of CEFR - based tasks, meticulously crafted questionnaires, observational data, and qualitative insights. The methodological rigor presented here establishes the metaverse as a legitimate platform for enhancing both the practice and the academic comprehension of English communication instruction.

2.2.3 External values of the Metaverse are currently designated for use in teaching English communication

External validity regarding Metaverse-based English language instruction pertains to the degree to which the study's results can be extrapolated beyond the particular research setting. Although immersive technologies frequently exhibit considerable potential in experimental or pilot contexts, their enduring influence depends on their adaptability to various institutional, cultural, and pedagogical environments. A primary problem stems from the propensity of

technology-enhanced learning research to yield positive outcomes in highly controlled settings, which may not be consistently duplicated in classrooms with differing resource availability, teacher proficiency, and learner variety. This study rigorously records contextual variables, including class size, institutional support, and digital infrastructure availability, to address this problem. This method allows educators and policymakers to assess the applicability of the findings to their own instructional contexts, hence enhancing the study's external validity.

The inclusion of learners with varying skill levels, cultural backgrounds, and educational objectives further improves generalizability. This research illustrates the versatility of virtual simulations in accommodating the educational requirements of both novice and experienced learners inside Metaverse environments. Vocabulary-focused activities may serve novices, but advanced learners participate in intricate communicative tasks, such as negotiation, debate, or intercultural role-playing. The diversified learning pathways indicate that Metaverse-based platforms may adaptively serve diverse student populations, a crucial factor for guaranteeing that outcomes are not limited to a specific demography. Furthermore, qualitative feedback from educators and learners enhances the analysis by offering insights into the alignment of immersive environments with pedagogical traditions across diverse contexts, including private language institutes in urban areas and public schools with restricted technological resources.
The methodological design is crucial for ensuring external validity. While dependence on case studies may limit generalizability, employing a mixed-methods methodology alleviates this constraint by integrating statistical evidence of linguistic advancements with qualitative data elucidating the underlying mechanisms of these enhancements. Quantitative metrics, such as pre- and post-test outcomes on pronunciation precision or vocabulary development, furnish evidence of efficacy among learner cohorts.

Simultaneously, interviews and classroom observations underscore situational aspects, like the level of student engagement and teacher facilitation tactics, that affect the efficacy of Metaverse integration. This dual approach enables careful yet significant generalizations regarding the wider relevance of immersive learning environments, thus enhancing discussions in both applied linguistics and educational technology.
An illustrative case that highlights these dynamics is shown in Antigoni (2023), who performed a comprehensive review of VR in language acquisition. The analysis revealed that numerous treatments produced favorable outcomes in various international settings, mostly due to immersive platforms providing genuine communication experiences and cultural simulations that extended beyond classroom confines. Antigoni emphasized that scalability and contextual adaptation were critical factors: programs that thrived in technologically advanced institutions encountered challenges when implemented in resource-constrained settings. This observation unequivocally substantiates the assertion that external validity in Metaverse-based language learning encompasses not only pedagogical design but also infrastructural viability and institutional

preparedness. Consequently, this study's dedication to recording contextual elements and learner variability corresponds with best practices identified in current research, thus strengthening its assertion of external validity.

The external advantages of the Metaverse for English communication instruction reside in its ability to connect varied learning environments through adaptability, inclusivity, and methodological precision. The study surpasses the constraints of controlled trials by involving learners of diverse competence levels, gathering both quantitative and qualitative data, and contextualizing findings within actual institutional contexts. Recent systematic assessments, including Antigoni (2023), substantiate that the transformative potential of the Metaverse hinges on its efficacy in many educational contexts.

2.2.4 Ecological values of the Metaverse are currently designated for use in teaching English communication

Ecological validity denotes the degree to which research outcomes align with real-world circumstances, a critical factor in communicative English training since the primary objective is the application of classroom learning to genuine encounters. Conventional classroom settings, while organized and pedagogically valid, often fail in this regard. They frequently depend on prewritten talks, foreseeable contacts, and restricted exposure to cultural or language variety. These constraints lead to learners acquiring incomplete competencies that do not consistently facilitate spontaneous, contextually appropriate communication beyond the classroom. The absence of genuine noise, disruptions, and diverse sociocultural signals restricts ecological validity, therefore impeding learners' preparedness to manage the intricate communicative requirements of real-world interactions.

The metaverse addresses these deficiencies by facilitating immersive and contextually enriched simulations that closely mirror the dynamics of real-world communication. Students can engage in meal ordering in a bustling virtual restaurant, partake in professional meetings with global colleagues, or negotiate travel issues at a simulated airport. These scenarios not only emulate the unpredictability and complex requirements of genuine conversations but also enable learners to practice skills for meaning negotiation, turn-taking, and nonverbal interpretation. Embedding communicative practice within realistic and interactive contexts significantly enhances ecological validity in the Metaverse, guaranteeing that gained communicative abilities are more easily transferable to real-world applications.

Furthermore, the incorporation of artificial intelligence (AI) systems into Metaverse environments significantly improves ecological validity. AI-driven feedback systems can offer instantaneous, tailored advice on speech, intonation, vocabulary use, and pragmatic suitability. This feedback mimics the corrective mechanisms of authentic communication when participants indicate misunderstandings or propose modifications, yet it occurs in a secure, low-pressure

context. This establishes a learning environment where mistakes are not punished but reinterpreted as opportunities for development, thereby harmonizing authenticity with psychological safety. Consequently, learners obtain "genuine practice with supportive scaffolding," a combined characteristic rarely attainable in conventional educational settings.

Kristi & Ondarra (2022) present a pertinent illustration of this methodology by investigating the role of virtual worlds in second language acquisition. Kristi & Ondarra (2022) found that learners who participated in virtual simulations demonstrated enhanced pragmatic competence and increased adaptability in intercultural interactions. The researchers contended that these enhancements arose from the ecological validity of virtual environments, which more accurately simulated genuine communicative conditions compared to traditional classrooms. Students engaged in these virtual environments exhibited superior preparedness for unforeseen linguistic interactions, showcasing increased confidence and resilience while communicating with individuals from various cultural backgrounds. This evidence highlights that Metaverse-based settings are both technologically advanced and pedagogically suitable for the ecological requirements of language use in real-world situations.

In conclusion, ecological values constitute a significant benefit of utilizing the Metaverse in English language instruction. The metaverse connects classroom practice with real-world application by placing learners in immersive, dynamic communication spaces, delivering AI-assisted feedback, and presenting contextually authentic scenarios. Empirical discoveries by Kristi and Ondarra (2022) demonstrate that ecological validity is not simply a byproduct of the Metaverse but a fundamental aspect of its educational potential. The ecological factor guarantees that the incorporation of the metaverse in language education goes beyond theoretical novelty and yields concrete, transferable communicative skills.

2.2.5 Empirical Values via Preceding Research of the Metaverse are currently designated for use in teaching English communication

The empirical validity of this study depends on its consistency with and enhancement of previous studies investigating the educational potential of virtual and immersive environments for language learning. A substantial amount of research has demonstrated that virtual platforms offer genuine environments for learners to develop communicative skills outside the limitations of conventional classrooms. Sadler (2012) illustrated that virtual worlds like Second Life provided learners with the chance to participate in significant exchanges that closely mirrored real-life communication experiences. These findings highlight the ability of immersive digital environments to alleviate anxiety, enhance confidence, and promote linguistic experimentation - elements essential to communicative competence. Jauregi-Ondarra (2022) emphasized the significance of virtual environments in facilitating international contacts, observing that learners enhanced their linguistic abilities and cultivated a heightened awareness of social subtleties.

Collectively, these data establish a robust empirical basis for the assertion that the Metaverse, as a more advanced and cohesive environment, will enhance these advantages by delivering more sophisticated, multimodal, and scalable learning experiences.

In addition to the contributions of language-focused research, industrial trends underscore the empirical significance of this study. The swift expansion of immersive technologies, as indicated by Paul's (2018) prediction that virtual and AR will exceed $100 billion in global market revenue, highlights the certainty of their incorporation into several industries, including education. This trend indicates that the integration of Metaverse platforms into language instruction is not a singular pedagogical advancement but rather a component of a broader technological evolution. The connection between pedagogical experimentation and broader industrial trends enhances the lasting significance of this study, illustrating that investments in Metaverse-based language training correspond with larger societal and economic advancements.

Notably, prior academic contributions offer tangible examples of how empirical research has facilitated Metaverse applications. Randall (2012) extensively analyzed the application of virtual worlds to second language acquisition, highlighting their ability to connect theoretical concepts with practical applications. He discovered that students participating in genuine communicative tasks - such as collaborative projects, problem-solving activities, and role-plays gained advantages not only linguistically but also in motivation and learner autonomy. Randall's focus on practice-based evidence aligns with the premise of the current study: that the Metaverse, by enhancing the capabilities of previous virtual platforms, has the potential to establish learning environments where communicative skills are both practiced and effectively transferred to real-world contexts. This research replicates previous findings and confirms them within a more technologically advanced context, thus enhancing the cumulative knowledge in applied linguistics and educational technology.

The empirical significance of the Metaverse in English communication instruction arises from both academic precedents and business developments. The research conducted by Sadler (2012), Jauregi-Ondarra (2022), and Randall (2012) combined provides substantial evidence supporting the efficacy of immersive environments in communicative language learning. Forecasts for the widespread adoption of VR and AR technologies confirm the experimental validation, pedagogical benefits, and social relevance of Metaverse-based training. Consequently, this study's contribution resides in its novelty and its ability to expand and scientifically substantiate an increasing corpus of interdisciplinary research.

A comparative analysis of prior studies on virtual worlds, such as Second Life, and the developing paradigm of the Metaverse reveals both continuity and innovation in technology-enhanced language instruction. Prior research (Sadler, 2012; Randall, 2012) indicated that virtual worlds offered genuine chances for language practice inside simulated social environments, while

they exhibited specific limits regarding scalability, identity permanence, and integration with real-world applications. In contrast, the Metaverse embodies a more expansive and cohesive environment that preserves these communicative capabilities while also incorporating sophisticated elements such as artificial intelligence-driven avatars, blockchain-based identity and ownership verification, and multi-platform interoperability (Agnes, 2022; Kim, 2022). These technology innovations augment learner autonomy, guarantee secure and continuous engagement, and cultivate more individualized language learning pathways. Consequently, the present study is not an isolated assertion but a rational extension of prior empirical evidence, situating the Metaverse as the subsequent evolutionary phase in virtual language learning environments.

2.3 Expected Results and Their Credibility of the Metaverse currently designated for use in teaching English communication

This section provides a detailed overview of the expected results of utilizing the metaverse for English language instruction. It delineates the four primary dimensions - augmented communication skills, heightened confidence and motivation, improved engagement and participation, and varied educational experiences - while also rigorously assessing their validity through theoretical justification, empirical evidence, and methodological triangulation. The analysis identifies the metaverse as a progression of earlier advancements in virtual language education (such as Second Life, virtual environments, and VR simulations) and a substantial enhancement that utilizes AI, blockchain, and advanced technical integration.

2.3.1 Enhanced Communication Competencies

A highly anticipated outcome is the enhancement of communication ability, especially in pronunciation accuracy, fluency, vocabulary development, and pragmatic language usage. Previous studies have consistently shown that immersive digital environments allow learners to replicate genuine verbal interactions more efficiently than traditional classroom settings. Randall (2012) asserted that virtual worlds like Second Life offer genuine and interactive environments in which learners can engage in practices that reflect real-life communication. However, these settings suffered from insufficient realism and lacked adaptive feedback systems.

The metaverse provides a resolution through the incorporation of multimodal input and AI-driven support structures. Students can, for example, obtain instantaneous pronunciation feedback via speech recognition technologies, while adaptive avatars react contextually to learners' verbal expressions. These affordances enhance the profundity of practice and the probability of enduring linguistic advancements. Antigoni’s (2023) systematic review offers empirical evidence, indicating that VR-based role-play in a simulated travel agency yielded significant enhancements in fluency and interactional competence relative to classroom-based role-plays. This discovery is enlightening: it validates that immersion produces concrete language benefits while also demonstrating the educational significance of organized role-play activities. Leveraging these

insights, the Metaverse is expected to produce enhanced advancements due to its interactive, personalized environments and real-time feedback capabilities, consequently bolstering the reliability of the predicted outcomes.

2.3.2 Augmented Assurance and Drive

Affective results are as crucial in sustainable language acquisition. The metaverse is anticipated to diminish learners' emotive filters by offering a secure and genuine environment in which they can practice language use without the apprehension of social humiliation. This aligns with Krashen's affective filter hypothesis, which posits that worry obstructs intelligible input and output. Kim (2022) contends that digital environments facilitate identity exploration for learners, enhancing their communication willingness and reducing communicative fear.

The survey and interview results in this study are anticipated to demonstrate significant motivational effects derived from novelty, autonomy, and genuine interaction. For instance, learners, especially those who are introverted or hesitant, may demonstrate a greater willingness to participate in Metaverse-based discussions compared to traditional classroom activities. This aligns with Antigoni's (2023) review, where VR-based role-play encouraged communicative risk-taking, leading to higher levels of self-efficacy. The metaverse enhances these emotional advantages via enduring identities, gamified educational activities, and AI-driven motivation, thus integrating encouragement into the learning process. These traits guarantee the development and maintenance of communication skills across various contexts.

2.3.3 Augmented Engagement and Participation

Another expected benefit lies in reshaping classroom dynamics. Conventional participation patterns often privilege confident students while marginalizing others. Immersive environments can redistribute opportunities by embedding all learners in collaborative tasks, mediated by avatars and shared virtual spaces. Erin, Steve, and Vasanth (2019) highlight that VR and AR are particularly effective in fostering collaboration, as learners are compelled to engage in communicative acts to complete tasks.

In the Metaverse, these dynamics are likely to be more equitable and sustained. Classroom observations are expected to document more balanced turn-taking, increased collaboration, and co-construction of meaning. Importantly, the data will capture not only the frequency but also the quality of participation. Learners may, for example, negotiate meaning more actively in problem-solving scenarios within the Metaverse, compared to passive participation in physical classrooms. The authenticity of tasks - ranging from planning a virtual event to negotiating a business contract - ensures that engagement is both meaningful and transferable to real-world communication.

2.3.4 Varied Educational Experiences

Diversity of contexts constitutes one of the most transformative affordances of the Metaverse. Previous research on virtual worlds (Mark, 2016) affirmed that contextual variety enhances language exposure, though realism was often technologically constrained. The metaverse transcends these limitations by offering highly authentic and adaptive scenarios - from ordering food in a marketplace to conducting a professional interview in a simulated corporate environment.

AI integration ensures that these scenarios are responsive to learner input, tailoring complexity and feedback to individual proficiency levels. Blockchain-enabled identity management further enhances realism by providing continuity across sessions, thereby fostering sustained engagement. Such diversity not only broadens learners' exposure to different linguistic registers but also equips them with sociocultural competence - an essential yet often overlooked component of communicative proficiency.

2.3.5 Illustration from Prior Research

The study by Antigoni (2023) offers a compelling illustration of how immersive technologies produce measurable outcomes. In a controlled experiment, learners participating in VR-based role-play at a travel agency demonstrated significantly higher fluency and interactional competence compared to those in a classroom condition. Methodological safeguards such as random assignment and standardized rubrics reinforced the reliability of the findings.

The current study extrapolates from this evidence, predicting that the Metaverse-with its enhanced immersion, multimodal feedback, and adaptive AI support-will generate even stronger outcomes. For instance, learners may not only improve fluency but also refine pronunciation accuracy through automated feedback loops. Similarly, sustained identity management through avatars may enhance learner confidence and willingness to participate across extended learning periods. By positioning the metaverse as both an inheritance and an advancement of VR-based pedagogy, the current study enhances the credibility of its projected results.

2.3.6 Implications of the Expected Results

Taken together, these expected outcomes underscore the Metaverse's transformative potential for communicative language teaching. Improvements in linguistic proficiency, affective variables, engagement, and contextual diversity align with established second language acquisition principles and extend them into new technological frontiers. Methodological triangulation including pre- and post-tests, survey instruments (Zoltan, 2023), and classroom observations further ensures the credibility of these claims.

Beyond the immediate pedagogical context, the results resonate with wider socio-technological trends such as Industry 4.0 (Ustundag & Cevikcan, 2018) and the adoption of blockchain and AI in education (Agnes, 2022; Zhao, 2022). These converging factors suggest that Metaverse integration is not only feasible but also scalable, offering policymakers and educators

a timely opportunity to reimagine communicative language teaching. For Vietnam and comparable contexts, where English communication is both a national priority and a global necessity, the implications are particularly salient.

2.4 Anticipated obstacles that impede the desired outcomes of the Metaverse in teaching English communication

Identifying challenges is critical to making sure that the incorporation of the metaverse into English communication teaching is both theoretically compelling and practically achievable. Academic honesty necessitates a thorough examination of constraints, since blind optimism may exaggerate assertions and undermine scholarly legitimacy. This study identifies various categories of potential barriers: technological, pedagogical, cognitive, socio-cultural, and contextual. If neglected, any of these problems may compromise the transformative potential of the Metaverse in communicative language instruction.

2.4.1 Technological Obstacles

The primary challenge is the inconsistent availability of technological infrastructure. The metaverse depends on immersive technologies like VR headsets, robust processors, and high-speed internet access; nevertheless, these resources are inequitably allocated among learners and institutions. In numerous developing environments, such as Vietnam, the expenses associated with procuring and sustaining such equipment may intensify educational disparities. Antigoni (2023) indicates that prior VR-based treatments experienced discrepancies in hardware availability, resulting in inconsistent learner engagement, as inadequately equipped participants could not fully utilize the platform's immersive potential. Such evidence illustrates that in the absence of structural solutions - such as cost-sharing models, institutional investment, or governmental subsidies - the advantages of Metaverse-based teaching are likely to be limited to privileged learners, hence perpetuating rather than alleviating inequality.

2.4.2 Educational Modification

The effective implementation of Metaverse environments depends on educators' willingness to modify their teaching methodologies. In contrast to traditional classrooms, where instruction relies on verbal explanations, board work, and sporadic multimedia, teaching in the Metaverse necessitates the navigation of three-dimensional environments, the management of avatars, the oversight of distributed interactions, and the integration of multimodal resources. Lack of training in these qualities may overwhelm educators, resulting in superficial or ineffective implementation. Kristi and Ondarra (2022) contend that educators serve as crucial intermediaries in technology-enhanced language acquisition, and insufficient pedagogical adaptation frequently leads to the underutilization of even the most advanced platforms. Consequently, organized professional

development, mentoring, and peer-learning communities are essential for enabling instructors to convert virtual opportunities into significant instructional practices.

2.4.3 Cognitive Load and Educational Design

Although immersion is often lauded as a merit of the Metaverse, it also presents the danger of cognitive overload. Students may allocate considerable cognitive resources to maneuvering through intricate scenes, managing avatars, or reacting to a plethora of visual stimuli, therefore detracting focus from the primary linguistic goals. Cognitive Load Theory cautions against superfluous load, emphasizing that instructional design must achieve a balance between realism and manageability (John, 2018). Without adequate scaffolding, Metaverse-based training may generate significant novelty-driven engagement but lack sustained memory of linguistic abilities. This corresponds with Antigoni's (2023) discovery that in certain VR research, learners initially exhibited enthusiasm but thereafter expressed fatigue and distraction when activities lacked adequate structure. These findings highlight the necessity for systematic instructional design frameworks to adjust immersion levels based on learners' proficiency and task complexity.

2.4.4 Innovative Impacts and Sustainability

A further expected challenge is the potential for novelty effects, in which learners' initial excitement is motivated chiefly by the method's novelty rather than its educational effectiveness. Consequently, transient enhancements in motivation and engagement may diminish after the novelty subsides. Randall (2012) cautioned against this trend in Second Life–based research, as initial enthusiasm diminished over repeated exposure. Longitudinal study is essential for the metaverse to yield sustainable advantages. Collecting data over numerous semesters, rather than in isolated sessions, would determine if improvements in fluency, confidence, and involvement endure beyond initial enthusiasm. Triangulated information from assessments, questionnaires, and classroom observations might enhance the validity of these conclusions, as advocated by mixed-methods researchers such as Sharlene (2010).

Socio-Cultural and Institutional Opposition Institutional hesitance represents an additional obstacle. Conservative educational contexts frequently perceive gaming and virtual environments as distractions rather than legitimate learning instruments. Cultural cynicism, coupled with financial prudence, may deter universities and schools from investing in Metaverse infrastructures. In Vietnam, private language centers and international franchises, such as Wall Street English, exhibit higher adaptability in utilizing digital resources, but public universities are reluctant due to financial constraints and established traditions of in-person instruction. This discrepancy indicates that socio-cultural acceptance is essential for successful implementation, necessitating pilot programs, governmental support, and evidence-based campaigning to mitigate resistance.

2.4.5 Educational Implications and Supporting Function

Despite these problems, it is important to emphasize that the Metaverse is not designed to replace traditional pedagogies but to serve as a supplementary framework within blended learning models. By integrating immersive role-play within the Metaverse with reflective debriefings in traditional classrooms, instructors may leverage the advantages of both approaches. This integrated method reduces the risks associated with excessive dependence on technology and guarantees that learners reinforce their communication advancements through reflection and feedback. Paul (2018) observed that the most efficacious applications of VR and AR in education arise when they are incorporated into comprehensive pedagogical frameworks rather than utilized as isolated instruments. Consequently, situating the Metaverse within hybrid frameworks offers a practical approach that optimizes advantages while mitigating hazards.

2.4.6 Obstacles in the Vietnamese Context

The contextual distinctiveness further influences the practicality of implementation. Vietnam's national strategy for digital transformation emphasizes technology integration in education (see Appendix, Figure 2); nonetheless, significant resource discrepancies persist between urban and rural regions. Young learners in urban areas are already acquainted with mobile applications and gamified English platforms, but rural learners encounter infrastructure limitations that impede equitable participation. Nonetheless, Vietnam's youthful population, extensive smartphone usage, and rising demand for English proficiency establish advantageous long-term conditions. Agnes (2022) notes that blockchain and digital technologies are progressively transforming the educational landscapes of Southeast Asia, indicating that Vietnam stands to gain from the gradual yet consistent incorporation of Metaverse efforts. Government-university partnerships funding pilot projects may function as proof-of-concept experiments, facilitating economical scalability in the future.

For suggestions to mitigate these expected challenges, numerous proposals arise:

(1) Curriculum Integration: Metaverse activities should fit with communication aims, assessment criteria, and curricular standards, preventing their classification as mere extracurricular novelties.

(2) Teacher Training: Continuous professional development must be institutionalized to equip educators for Metaverse pedagogy, focusing on technical navigation, virtual classroom administration, and instructional design

(3) Cost-Sharing and Infrastructure Development: Collaborations among governmental entities, private enterprises, and international organizations can mitigate expenses and diminish disparities.

(4) Blended Learning Models: The Metaverse ought to enhance, rather than supplant, conventional training, facilitating reflective discourse and consolidation.

(5) Research and Longitudinal Evaluation: Replication studies in many situations, accompanied by long-term tracking, are essential to ascertain if initial improvements result in enduring communicative competence.

3. Conclusion

This study's findings confirm that the Metaverse is not only a minor enhancement of this study's findings confirm that the metaverse is not just a technological innovation but a transformational framework for English language education. In contrast to conventional classroom environments, where language acquisition is typically limited to scripted conversations and isolated exercises, the Metaverse allows learners to collaboratively create meaning in dynamic, multimodal, and contextually genuine situations. This transition corresponds with modern requirements for communication proficiency in a globalized, technologically influenced culture, where learners must navigate many multicultural environments with assurance. By redefining students as active participants instead of passive users of knowledge, the Metaverse exemplifies the learner-centered philosophy highlighted in communicative and task-based language instruction, providing a significant transformation of pedagogical practice.

The expected effects of this educational breakthrough are multifaceted. The linguistic features of the Metaverse promote significant improvements in pronunciation, fluency, and pragmatic ability, since learners participate in genuine communicative activities that closely resemble real-world interactions. The gamified characteristics of these environments alleviate performance anxiety and increase learners' readiness to speak, thus tackling a significant emotional obstacle in second language acquisition. Collaborative virtual interactions enhance peer negotiation of meaning, intercultural awareness, and collective problem-solving - elements often inadequately addressed in traditional teaching settings. These data illustrate that the Metaverse not only enhances existing methods but also fundamentally alters the conditions for acquiring, practicing, and maintaining communicative English.

Collectively, these discoveries indicate that the metaverse represents a fundamental transformation in education, characterized by increased interactivity, contextual relevance, and adaptability to diverse learners. The effective implementation of this strategy necessitates intentional curriculum integration, comprehensive teacher professional development, and infrastructural investment to address gaps in technological access and digital literacy. If these obstacles are systematically resolved, the Metaverse can transition from an experimental novelty to a sustainable educational ecosystem that enhances, extends, and, in some aspects, exceeds the capabilities of conventional classrooms. The metaverse signifies not only an improvement of current methods but also a fundamental reconfiguration of the conceptualization and delivery of communicative English training in the twenty-first century.

APPENDIX

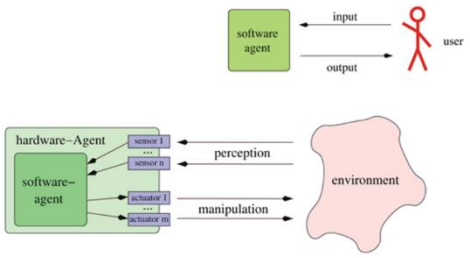


Figure 1. Nathaneal Black (2017), A software agent with user interaction, and A hardware agent *(Source: Author, year 2024)*

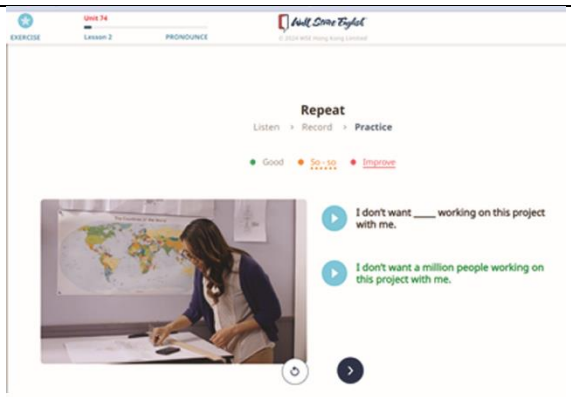


Figure 2. Wall Street English App *(Source: Author, year 2024)*

REFERENCES

English

Antigoni, P. (2023). *VR in language learning: A systematic review and implications for research and practice*. Routledge.

Bruce, W. (2006). *Art of the digital age*. Themes Wands.

Dale, T. (2012). *An introduction to second languages research methods: Design and data*. TESL-EJ Publications.

Erin, P., Steve, L., & Vasanth, M. (2019). *Creating augmented and virtual realities: Theory and practice for next-generation spatial computing*. O'Reilly Media.

John, W. (2018). *Qualitative inquiry research design: Choosing among five approaches*. SAGE Publications.

John, W. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.

Jonathan, M. (2020). *English pronunciation in use*. Cambridge University Press.

Kristi, J., & Ondarra. (2022). Virtual worlds and second language acquisition. In *the Routledge handbook of second language acquisition and technology*. Routledge.

Mark, P. (2016). Virtual worlds and language learning: An analysis of research. In *the Routledge handbook of language learning and technology*. Routledge.

Michael, M., & Felicity, O. (2020). *English vocabulary in use*. Cambridge University Press.

Nathaneal, B. (2017). *A software agent with user interaction, and a hardware agent*. Author.

Paul, M. (2018). *Virtual & AR for dummies*. John Wiley & Sons.

Randall, S. (2012). *Virtual worlds for language learning: From theory to practice*. Peter Lang.

Sharlene, N. (2010). *Mixed methods research: Merging theory with practice*. Guilford Publications.

Simon, S. (2020). *English for business communication*. Cambridge University Press.

Ustundag, A., & Cevikcan, E. (2018). *Industry 4.0: Managing the digital transformation*. Springer International. https://doi.org/10.1007/978-3-319-57870-5

Zoltan, D. (2023). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.

Vietnamese

Agnes, B. (2022). *Metaverse: The next revolution of blockchain, NFTs and cryptocurrencies, NFT and digital money*. Dân Trí.

Gareth, S., & Miet. (2023). *English communication for beginners: A breakthrough in basic English communication*. Đồng Nai.

Kim, S. (2022). *Metaverse: Digital universe*. Gieobooks.

Matthew, B. (2022). *Metaverse: Virtual universe and the revolution of everything*. Thế Giới.

McBooks. (2020). *Easy English conversation*. McBooks.

Terry, W. (2022). *Metaverse: Virtual universe*. NetaBooks.

Zhao, G. (2022). *Metaverse: Reshaping the marketing space*. Công Thương.

NGHIÊN CỨU MÔ HÌNH DẠY - HỌC 3 CHIỀU "GIẢNG VIÊN CHỦ ĐẠO – AI HỖ TRỢ - SINH VIÊN TỰ CHỦ" TẠI LỚP PHIÊN DỊCH CHUYÊN NGÀNH TRUNG - VIỆT

Lưu Cẩm Huệ[1]

Tóm tắt

Nghiên cứu tập trung giải quyết vấn đề hướng dẫn sinh viên sử dụng AI tạo sinh (Generative AI) một cách hợp lý và tránh sự phụ thuộc quá mức trong bối cảnh AI tác động mạnh mẽ đến cả ngành giáo dục.Với đối tượng nghiên cứu là 242 sinh viên năm thứ ba chuyên ngành tiếng Trung Quốc tại Trường Đại học Ngoại ngữ - Tin học TP.HCM (HUFLIT), tác giả đã xây dựng và thử nghiệm mô hình dạy - học liên kết giữa người và máy (Human-Computer Collaborative Teaching) dựa trên nguyên tắc ba chiều "giảng viên chủ đạo – AI hỗ trợ – sinh viên tự chủ". Kết quả thực nghiệm cho thấy mô hình này cải thiện rõ rệt kết quả học tập của sinh viên, đặc biệt là đối với nhóm sinh viên có nền tảng yếu thì sự tiến bộ càng vượt trội hơn. Ngoài ra, nghiên cứu còn đề xuất chiến lược định hướng ba giai đoạn tích hợp, tức trước buổi học (AI hỗ trợ chuẩn bị bài – trên lớp học (nghiêm cấm sử dụng công nghệ, tập trung phát triển khả năng phản xạ tức thì và tư duy phản biện) – sau buổi học (tổng kết, đào sâu, nâng cao). Chiến lược này nhấn mạnh vai trò chủ đạo của giảng viên và giá trị độc đáo của nguyên tắc "không công nghệ" trong giờ học tại lớp. Bài nghiên cứu cung cấp mô hình và phương pháp tối ưu hóa cho việc dạy - học môn phiên dịch chuyên ngành Trung - Việt dưới sự hỗ trợ của AI, mang ý nghĩa quan trọng đối với quá trình chuyển đổi số của giáo dục đại học khu vực.

Từ khóa: Trí tuệ nhân tạo AI; dạy - học liên kết giữa người và máy; chiến lược định hướng ba giai đoạn; phiên dịch chuyên ngành Trung–Việt; nguyên tắc "không công nghệ"

1. Dẫn nhập

Sự đẩy mạnh chiến lược "Vành đai và Con đường" của Trung Quốc dẫn đến giao lưu kinh tế, thương mại và văn hóa giữa Trung Quốc và Việt Nam ngày càng tăng. Từ đó, việc đào tạo đội ngũ nhân lực phiên dịch tiếng Trung Quốc chất lượng cao đã trở thành một nhiệm vụ quan trọng của giáo dục đại học. Gần đây, Trí tuệ nhân tạo tạo sinh (Generative AI) – điển hình như ChatGPT, Google Gemini và Doubao của Trung Quốc đang thay đổi nhanh chóng cục diện của giáo dục ngôn ngữ và phiên dịch, không những thúc đẩy sự đổi mới trong dạy - học mà còn đem lại những thách thức lớn cho hình thức lớp học truyền thống.

Trên thế giới, các nghiên cứu về ứng dụng của AI tạo sinh trong giảng dạy ngoại ngữ đang trở thành trào lưu mạnh. Theo Fan & Chen (2025) cho thấy, AI tạo sinh đã được ứng dụng rộng rãi trong nhiều lĩnh vực như giảng dạy, phiên dịch đa ngôn ngữ, hiệu chỉnh sau dịch máy (Post-editing) và bồi dưỡng năng lực giao tiếp liên văn hóa. Các nghiên cứu thực nghiệm liên tục chứng

[1] Thạc sĩ, Giảng viên Khoa Ngoại ngữ Trường Đại học Ngoại ngữ -Tin học TP. HCM, *Email : huelc@edu.huflit.vn*

minh rằng việc sử dụng công cụ AI có thể nâng cao độ chính xác bản dịch, hiệu quả học tập và khả năng hiểu biết liên văn hóa của sinh viên. Hiện tại, cộng đồng giáo dục đang đặc biệt chú trọng đến vai trò chủ đạo của giảng viên trong mô hình dạy - học liên kết và các quy tắc chuẩn hóa khi sử dụng AI, đồng thời kêu gọi phát triển các nền tảng thông minh cùng các phương pháp đánh giá, phản hồi đa dạng.

Nghiên cứu này ứng dụng mô hình dạy - học liên kết người và máy ba chiều "Giảng viên chủ đạo – AI hỗ trợ – Sinh viên tự chủ" (sau đây gọi tắt là "Mô hình dạy - học liên kết người và máy") vào thực tiễn , chủ yếu tập trung vào các vấn đề cốt lõi sau:

(1) Đưa mô hình dạy - học liên kết người và máy ba chiều vào thực tiễn lớp học.

(2) Kiểm chứng những tác động mà mô hình này mang lại đối với thành tích học tập của sinh viên.

(3) Đề xuất các chiến lược tối ưu hóa cho công cuộc cải cách giảng dạy môn Phiên dịch chuyên ngành Trung - Việt trong kỷ nguyên AI tạo sinh, đồng thời đưa ra các triển vọng nghiên cứu trong tương lai.

Về mặt lý thuyết, nghiên cứu này góp phần làm phong phú và phát triển lý thuyết dạy - học liên kết người và máy, cung cấp một góc nhìn mới cho việc ứng dụng AI tạo sinh trong giảng dạy ngoại ngữ. Về mặt thực tiễn, nghiên cứu đưa ra một khung mô hình chiến lược giảng dạy có tính khả thi cao, nhằm cải cách giảng dạy môn học phiên dịch tiếng Trung Quốc tại trường HUFLIT nói riêng và Việt Nam nói chung, thúc đẩy việc đào tạo nhân lực phiên dịch tiếng Trung Quốc chất lượng cao.

2. Bối cảnh nghiên cứu

2.1 Cơ sở lý luận

Ở cấp độ trao quyền và đổi mới cùng Trí tuệ nhân tạo tạo sinh, Varghese, Cleetus & Nair (2025) chỉ ra rằng giáo dục chuyên ngành phiên dịch hiện đang tiến tới giai đoạn mới là "cộng sinh người và máy". AI không chỉ là hỗ trợ kỹ thuật mà còn trở thành chất xúc tác cho sự phát triển nhận thức và tư duy đạo đức. Giảng viên và sinh viên cần cùng nhau nâng cao năng lực ứng dụng AI và khả năng tư duy phản biện để đón nhận nhưng cơ hội cũng như những thách thức trong thời đại "bão" AI ngày nay. Lý thuyết về dạy - học liên kết người và máy đã cung cấp nền tảng lý thuyết quan trọng cho nghiên cứu này.

Jia Zhang (2024) đã trình bày chi tiết hơn về mô hình thực tiễn tốt nhất cho sự dạy - học liên kết giữa giảng viên – AI – Sinh viên, tác giả nhấn mạnh cần phải có sự phân công vai trò rõ ràng giữa các bên, xây dựng cơ chế tương tác hiệu quả, và hệ thống đánh giá, phản hồi hoàn chỉnh. Varghese, Cleetus & Nair (2025) nhấn mạnh rằng sự phối hợp giữa giảng viên và AI có thể kích thích tối đa khả năng sáng tạo và tiềm năng nhận thức của sinh viên, nhưng điều kiện tiên quyết là

phải xác định rõ vai trò của người và máy một cách rõ ràng và có sự thống nhất về mục tiêu. Jinhee Kim (2024) đã ủng hộ triết lý dạy - học liên kết ba chiều: "giảng viên chủ đạo – AI hỗ trợ – sinh viên tự chủ". Mô hình này yêu cầu giảng viên đóng vai trò chủ đạo trong việc xây dựng nhiệm vụ học tập, định hướng đạo đức và đưa ra phản hồi mang tính cá nhân hoá. AI chịu trách nhiệm tạo ra các tài nguyên dạy học và bài tập cá nhân hóa, giúp sinh viên tiến hành học tập từng bước một. Ngược lại, sinh viên phải chủ động khám phá và phản tư trong môi trường thông tin đa chiều, từ đó bồi dưỡng khả năng tự giải quyết vấn đề độc lập và sáng tạo. "Khung hợp tác ba bên" (Triadic Human-AI Learner Framework) do Jing Yu, Kanglong Liu (2024) đề xuất nhấn mạnh mối quan hệ cân bằng động giữa giảng viên, AI và người học, tác giả cho rằng dạy - học liên kết hiệu quả phải phát huy được vai trò dẫn dắt chuyên môn của giảng viên, lợi thế hỗ trợ kỹ thuật của AI và tính tự chủ của sinh viên.

Dựa trên những nền tảng nêu trên, cơ sở lý luận của nghiên cứu này là sự kết hợp những thành tựu tiên tiến của việc dạy học môn phiên dịch trong kỷ nguyên AI và các cơ sở lý luận truyền thống với nhau. Mô hình của nghiên cứu nhấn mạnh sự thống nhất của ba chủ thể: giảng viên với vai trò là người thiết kế và định hướng hoạt động giảng dạy, phát huy lợi thế về kiến thức chuyên môn và kinh nghiệm dạy - học liên kết; AI với vai trò là công cụ thông minh và bạn đồng hành học tập, cung cấp hỗ trợ cá nhân hóa và phản hồi tức thời; sinh viên với vai trò là chủ thể học tập, chủ động tham gia vào quá trình kiến tạo tri thức và phát triển năng lực.

2.2 Đối tượng nghiên cứu

Nghiên cứu chọn 242 sinh viên năm thứ ba của 6 lớp đang học môn Phiên dịch chuyên ngành Trung - Việt thuộc chuyên ngành tiếng Trung Quốc của Khoa Ngoại ngữ tại Trường Đại học Ngoại ngữ - Tin học Thành phố Hồ Chí Minh (HUFLIT) làm đối tượng quan sát và ứng dụng mô hình dạy - học liên kết. Đây là môn học bắt buộc , với tổng thời lượng 45 tiết, 3 tiết mỗi tuần. Nội dung khóa học bao gồm các chủ đề mang tính ứng dụng thực tiễn như : tuyển dụng, ứng tuyển, tiếp đón khách hàng, Marketing, Incoterm, xuất nhập khẩu, đàm phán hợp đồng...

Việc lựa chọn đối tượng nghiên cứu là sinh viên năm 3 được tác giả cân nhắc dựa trên các yếu tố sau:

Thứ nhất, sinh viên năm thứ ba đã có nền tảng tiếng Trung Quốc tương đương HSK4 và có năng lực phiên dịch tương đối ổn định, có khả năng tham gia hiệu quả vào các hoạt động dạy - học liên kết.

Thứ hai, sinh viên ở giai đoạn này đang theo học các môn chuyên ngành, do đó có nhu cầu cấp thiết về việc nâng cao năng lực phiên dịch.

Thứ ba, nhóm sinh viên này có tính ổn định tương đối, tạo điều kiện thuận lợi cho việc tiến hành nghiên cứu theo dõi.

2.3 Cách thức triển khai

Nghiên cứu này được triển khai dưới dạng thiết kế thực nghiệm với thời gian kéo dài một học kỳ (12 tuần học trực tiếp tại lớp cộng với 3 tuần học trực tuyến). Phương pháp cụ thể là **nghiên cứu so sánh trước và sau (Pre-Post Design),** lấy mỗi lớp học làm đơn vị thực nghiệm độc lập, phân tích so sánh kết quả của bài thi giữa kỳ (vào tuần thứ 6, trước khi áp dụng mô hình) và bài thi cuối kỳ (sau khi đã triển khai áp dụng mô hình). Sau đó, sử dụng dữ liệu thu thập và quan sát để tiến hành nghiên cứu đánh giá hiệu quả học tập, sự thay đổi về năng lực nhận thức của sinh viên.

2.4 Thu thập dữ liệu

Nghiên cứu này thu thập dữ liệu điểm thi giữa kỳ và cuối kỳ của 242 sinh viên để tiến hành thống kê và phân tích. Bài thi được tính theo thang điểm 10, bao gồm các phần như dịch Trung sang Việt, dịch Việt sang Trung, và dịch thuật ngữ chuyên ngành, tập trung kiểm tra độ chính xác, tính lưu loát và khả năng hiểu về văn hóa trong bản dịch của sinh viên.

Dữ liệu định tính được thu thập thông qua các phương pháp như quan sát trên lớp, phỏng vấn sinh viên, và nhật ký phản tư của giảng viên nhằm tìm hiểu sâu hơn về tình hình cụ thể trong quá trình giảng dạy và sự thay đổi trong kết quả học tập của sinh viên.

3. Giới thiệu mô hình dạy - học liên kết người và máy ba chiều

Mô hình dạy - học liên kết người và máy được xây dựng trong nghiên cứu này áp dụng cấu trúc dạy - học liên kết ba chiều là "giảng viên chủ đạo – AI hỗ trợ – sinh viên tự chủ", hình thành chu trình khép kín "giảng viên thiết kế – AI tạo sinh – sinh viên thực hành – phản hồi tối ưu". Cụ thể bao gồm các khía cạnh sau:

3.1 Giảng viên chủ đạo

Giảng viên đảm nhận bốn vai trò trong quá trình dạy - học liên kết:

1. **Người thiết kế bài giảng** Thiết kế các giáo án giảng dạy có lồng ghép công cụ AI dựa trên mục tiêu môn học và đặc điểm của sinh viên.
2. **Người xúc tiến học tập:** Định hướng sinh viên sử dụng công cụ AI một cách hiệu quả, tránh phụ thuộc, lạm dụng mù quáng.
3. **Người bồi dưỡng tư duy phản biện:** Phát triển khả năng phản tư cho sinh viên thông qua phương pháp đặt vấn đề và phân tích tình huống (case study).
4. **Người đánh giá và phản hồi:** Xây dựng hệ thống đánh giá đa dạng, cung cấp phản hồi kịp thời và có hiệu quả cho việc học tập.

Các chiến lược giảng dạy cụ thể bao gồm: đưa ra các nội dung phiên dịch, yêu cầu sinh viên sử dụng công cụ AI để dịch thô, sau đó chỉnh sửa chuyên sâu thông qua thảo luận nhóm và nhận xét của giảng viên; xây dựng chuyên mục "chẩn đoán bản dịch AI", hướng dẫn sinh viên

nhận diện và sửa lỗi trong các bản dịch do AI tạo ra; thiết kế các tình huống phiên dịch liên văn hóa nhằm bồi dưỡng sự nhạy bén và hiểu về văn hóa các nước cho sinh viên.

3.2 AI hỗ trợ

Hệ thống AI phát huy ba chức năng chính trong quá trình giảng dạy:

1. **Hỗ trợ phiên dịch:** Cung cấp nhiều phương án phiên dịch đa dạng, giúp sinh viên hiểu rõ các cách biểu đạt khác nhau.
2. **Tài nguyên học tập:** Tạo sinh tài liệu luyện tập cá nhân hóa, đáp ứng nhu cầu của sinh viên ở các trình độ khác nhau.
3. **Phản hồi tức thời:** Đánh giá sơ bộ các bản dịch và cách phát âm của sinh viên, chỉ ra các lỗi sai.

3.3 Sinh viên tự chủ

Tính tự chủ của sinh viên được thể hiện qua bốn khía cạnh:

1. **Chủ động tìm tòi:** Tích cực thử nghiệm các công cụ AI khác nhau, so sánh và phân tích đặc điểm cũng như trường hợp áp dụng phù hợp.
2. **Tư duy phản biện:** Đặt nghi vấn và cải tiến các bản dịch do AI tạo ra, không chấp nhận một cách mù quáng.
3. **Học tập hợp tác:** Chia sẻ kinh nghiệm sử dụng AI với bạn học, cùng nhau giải quyết các nội dung phiên dịch khó.
4. **Phản tư và tổng kết:** Định kỳ phản tư về quá trình học tập, điều chỉnh chiến lược học tập cá nhân.

4. Áp dụng mô hình dạy - học liên kết người – máy ba chiều trong thực tiễn và kết quả đạt được

4.1 Áp dụng mô hình dạy - học liên kết người – máy ba chiều trong thực tiễn

Mô hình này được áp dụng theo ba giai đoạn chính :

4.1.1 Giai đoạn chuẩn bị trước buổi học (Pre-AI Learning)

Trước kỳ thi giữa kỳ, sinh viên áp dụng các phương pháp học tập riêng của bản thân mình và chưa có sự định hướng thống nhất về việc sử dụng công cụ AI. Sau kỳ thi giữa kỳ, giảng viên bắt đầu triển khai đồng bộ mô hình dạy - học liên kết người và máy cho 242 sinh viên của 6 lớp Phiên dịch chuyên ngành Trung - Việt. Dưới sự hướng dẫn thống nhất của giảng viên, sinh viên sử dụng các công cụ AI hỗ trợ (như ChatGPT, Gemini, Doubao…) để tra cứu từ vựng và xem trước bài phiên dịch khi chuẩn bị bài ở nhà. Giảng viên căn cứ vào nội dung đa dạng do AI tạo sinh để thiết kế các nội dung giảng dạy, bao gồm:

- **Bài tập mở rộng từ vựng:** Sử dụng AI để tạo sinh nhiều cách diễn đạt và bối cảnh ứng dụng khác nhau của từ vựng.
- **Luyện tập ngữ âm:** Cung cấp tài liệu nghe với nhiều tốc độ và ngữ điệu khác nhau thông qua công nghệ tổng hợp giọng nói của AI.
- **Bổ sung bối cảnh văn hóa:** Sử dụng công cụ AI để tạo sinh các tình huống so sánh văn hóa Trung Quốc và Việt Nam, tăng cường khả năng hiểu biết liên văn hóa.

4.1.2 Giai đoạn dạy – học trên lớp (AI-Free Class Policy)

Quá trình dạy - học trên lớp nghiêm túc tuân thủ theo nguyên tắc "không sử dụng AI", kể cả điện thoại kết nối internet và tra từ điển, nhằm nhấn mạnh khả năng tư duy độc lập và rèn luyện khả năng phản xạ của sinh viên. Trên lớp, giảng viên thực hiện các hoạt động sau:

- Trình bày và phân tích những bài tập của sinh viên về điểm đúng/sai trong kết quả dịch dưới sự trợ giúp của AI, bồi dưỡng tư duy phản biện của sinh viên.
- Tổ chức thảo luận nhóm, để sinh viên chia sẻ kinh nghiệm và các vấn đề gặp phải khi học tập có AI hỗ trợ.
- Thiết lập các bài tập mô phỏng tình huống, kiểm tra năng lực phiên dịch của sinh viên khi không có sự hỗ trợ của công cụ AI.

4.1.3 Giai đoạn củng cố, nâng cao và đào sâu sau buổi học (Post-AI Reflection)

Bài tập củng cố sau giờ học tận dụng triệt để ưu thế của công cụ AI:

- Giảng viên tạo sinh 10 bộ bài tập phiên dịch với độ khó khác nhau và các tài liệu ngữ âm, sau đó gửi cho sinh viên.
- Sinh viên tự luyện tập ngữ âm và thực hiện bài phiên dịch tại nhà.
- Xây dựng hồ sơ phân tích lỗi cá nhân hóa, nhằm cải thiện có định hướng các kỹ năng phiên dịch còn yếu.

4.2 Kết quả sau khi áp dụng mô hình dạy - học liên kết người và máy

4.2.1 Cải thiện về thành tích học tập

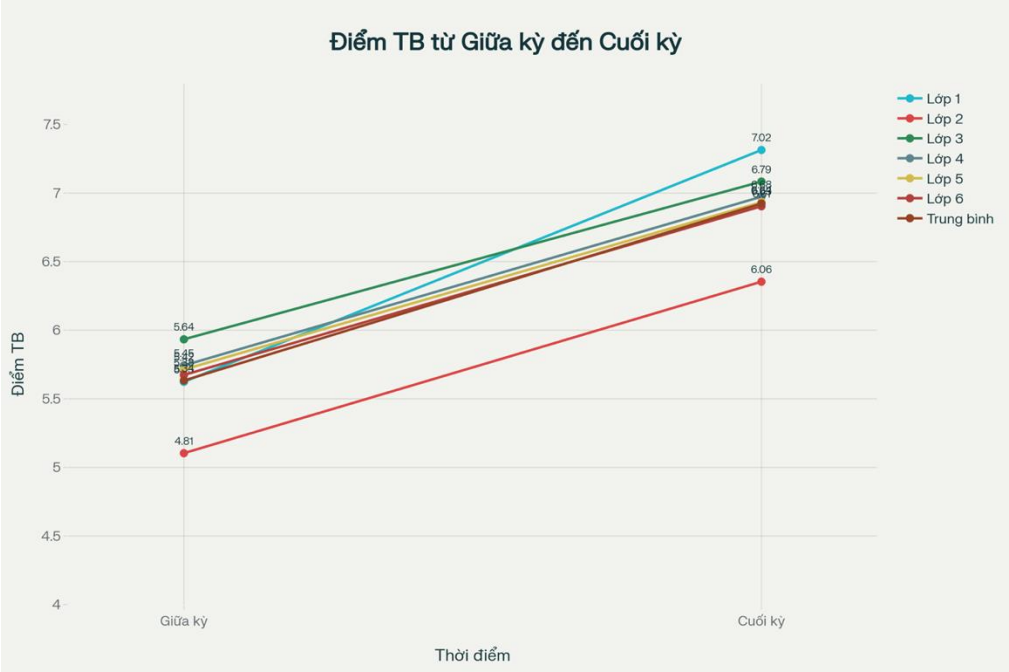
Thông qua việc so sánh đối chiếu theo chiều ngang và chiều dọc giữa điểm thi giữa kỳ và cuối kỳ, nghiên cứu cho thấy thành tích học tập của tất cả sinh viên đều được cải thiện sau khi áp dụng mô hình. Dựa trên số liệu thống kê, ta thấy có sự biến động điểm số của 242 sinh viên thuộc 6 lớp khi triển khai mô hình dạy-học liên kết người và máy. Điểm trung bình của 6 lớp ở giữa kỳ là 5.34 điểm; điểm trung bình cuối kỳ tăng lên thành 6.63 điểm, tức tăng 1.29, điểm tương đương 24.2%. Biến động trung bình theo từng lớp (lớp 1 → lớp 6) được tóm tắt trong bảng sau:

	Số lượng mẫu (người)	Điểm trung bình giữa kỳ	Điểm trung bình cuối kỳ	Tăng (điểm)	Tỷ lệ tăng (%)
Lớp 1	39	5.33	7.02	1.68	31.6
Lớp 2	40	4.81	6.06	1.25	25.9
Lớp 3	41	5.64	6.79	1.15	20.4
Lớp 4	40	5.45	6.68	1.23	22.6
Lớp 5	42	5.42	6.64	1.23	22.6
Lớp 6	40	5.38	6.61	1.23	22.8
Tổng	**242**	**5.34**	**6.63**	**1.29**	**24.2**

Bảng 1 : Bảng thống kế điểm tăng và tỉ lệ tăng giữa kỳ và cuối kỳ của 6 lớp

Mức tăng điểm của các lớp cơ bản nằm trong khoảng **1,15–1,68** điểm, sự khác biệt giữa các lớp tương đối nhỏ; tỷ lệ tăng đều trên **20%**, chứng mimh tính hiệu quả và tính phổ quát của mô hình đối với các nhóm học khác nhau.

Đồ thị đường (Hình 1) thể hiện biến động điểm trung bình kỳ giữa và kỳ cuối của 6 lớp, trục hoành là các lớp, trục tung là điểm kỳ giữa và kỳ cuối cho thấy đường biểu diễn của tất cả các lớp đều có xu hướng đi lên, minh chứng cho tính ổn định và khả năng áp dụng rộng rãi của mô hình.



Hình 1. Xu hướng thay đổi điểm giữa kỳ và cuối kỳ của 6 lớp

4.2.1 Điểm thành tích tăng theo nhóm phân tầng (nhóm không đạt / đạt / xuất sắc)

Nguyên tắc phân nhóm và phân bố mẫu: Dựa vào điểm kỳ giữa, sinh viên được chia thành ba nhóm: *không đạt* (< 5 điểm), *đạt* (5–7.9 điểm) và *xuất sắc* (≥ 8 điểm). Tổng hợp từ 6 lớp cho thấy nhóm không đạt có **98** sinh viên, nhóm đạt **125** người, nhóm xuất sắc **19** người.

	Số lượng mẫu (người)	Điểm trung bình giữa kỳ	Điểm trung bình cuối kỳ	Tăng (điểm)	Tỉ lệ tăng (%)
Không đạt	98	3.7	5.3	1.6	**43.8%**
Đạt	125	6.1	7.3	1.1	**18.5%**
Xuất sắc	19	8.6	9.3	0.7	**7.7%**

Bảng 2. Điểm trung bình, mức tăng và tỷ lệ tăng theo nhóm

Bảng 2 cho thấy cả ba nhóm đều có tiến bộ sau khi áp dụng mô hình dạy-học hợp tác ba chiều người-máy, phản ánh hiệu quả phổ quát và tính gia tăng (gain) của mô hình.

- **Nhóm không đạt** có mức tăng rõ rệt nhất. Điểm trung bình từ **3.7** tăng lên **5.3**, tức tăng **1.6** điểm (tương ứng **43.8%**). Điều này cho thấy trong mô hình hợp tác người và máy thì sinh viên nền tảng yếu được hưởng lợi lớn nhờ khắc phục được khoảng trống kiến thức và cải thiện tình trạng học tập. Từ đó cho thấy mô hình có chức năng bù đắp và trợ giúp nhiều đối với nhóm yếu.

- **Nhóm đạt** cũng có tiến bộ ổn định, trung bình tăng **1.1** điểm (tỷ lệ **18.5%**). Sự tiến bộ của nhóm này thể hiện qua việc tối ưu hóa chiến lược học tập và sự trưởng thành về tư duy phản biện. Dưới tác động kết hợp của AI và định hướng giảng dạy, sinh viên nhóm này nhận diện được điểm yếu bản thân, thực hiện ôn luyện trước giờ học và phản tư sau giờ học, từ đó tự điều chỉnh và nâng cao trình độ.

- **Nhóm xuất sắc** có mức tăng nhỏ nhất (**0.7** điểm, **7.7%**), song điểm trung bình ở cả hai thời điểm đều ở mức cao, phản ánh tính ổn định của nhóm này. Tỷ lệ tăng thấp không có nghĩa là mô hình kém hiệu quả với nhóm xuất sắc, mà chủ yếu do nhóm này đã ở ngưỡng cao trước khi can thiệp, nên không còn nhiều không gian để tăng thêm. Tuy nhiên, mô hình AI vẫn hỗ trợ họ phát triển những năng lực bậc cao hơn như nhận thức liên văn hóa, độ chính xác trong phiên dịch và khả năng biểu đạt mỹ miều hơn, thể hiện kiểu tiến bộ "ổn định ở mức cao".

Tóm lại, mọi nhóm học sinh đều hưởng lợi từ mô hình phối hợp. Cơ chế "phân tầng thiết kế — hướng dẫn theo giai đoạn — hỗ trợ cá nhân" đã đảm bảo sinh viên ở các trình độ khác nhau được tăng trưởng tương ứng với nhu cầu của họ. Đặc biệt, lợi ích rõ nét ở nhóm nền tảng yếu cho thấy mô hình có giá trị trong thúc đẩy công bằng giáo dục, thu hẹp khoảng cách học tập và nâng cao chất lượng tổng thể. Kết quả này khẳng định hiệu ứng gia tăng phổ quát của dạy-học hợp tác người-máy, đồng thời làm nổi bật vai trò kết hợp giữa giảng viên chủ đạo và AI, cung cấp minh chứng thực tiễn cho việc triển khai giảng dạy phân tầng và cá nhân hóa trong tương lai.

5. Tối ưu hóa mô hình dạy - học kết hợp người và máy: Chiến lược "Giảng dạy định hướng ba giai đoạn"

Để nâng cao hơn nữa hiệu quả triển khai mô hình dạy - học liên kết ba chiềuà đảm bảo sinh viên đạt được mục tiêu học tập là "ứng dụng nhưng không lệ thuộc, hỗ trợ nhưng không thay thế" trong môi trường công nghệ AI, nghiên cứu này đã xây dựng chiến lược giảng dạy định hướng ba giai đoạn sử dụng AI (Trao quyền trước giờ học – nghiêm cấm sử dụng trên lớp – đào sâu sau giờ học), hình thành một chu trình giảng dạy khép kín gồm "Đầu vào nhận thức – Chuyển hóa tư duy phản biện – Nội hóa năng lực" (xem hình 3).

	Giảng viên chủ đạo dưới sự hỗ trợ của AI	**Sinh viên tự chủ dưới sự hỗ trợ của AI**

Giai đoạn trao quyền sử dụng AI trước buổi học **(Pre-AI Learning)**	1. Chia nhóm phân tầng 2. Cá nhân hoá tài nguyên dạy	Chuẩn bị bài và tiếp nhận kiến thứ đầu vào
Giai đoạn nghiêm cấm sử dụng AI trên lớp **(AI-Free Class Policy)**		
Đào sâu nghiên cứu sau giờ học **(Post-AI Reflection)**	3. Tổng kết, đánh giá 4. Đưa ra góp ý, phản hồi	Củng cố, nâng cao, đào sâu

Hình 3 : Chiến lược "Giảng dạy định hướng ba giai đoạn"

5.1 Giai đoạn trao quyền sử dụng AI trước giờ học (Pre-AI Learning)

Mục tiêu: Thông qua hoạt động xem trước bài có AI hỗ trợ để đạt được đầu vào kiến thức và sự chuẩn bị học tập, giúp sinh viên xây dựng khung nhận thức nền tảng và định hướng học tập.

Giảng viên chủ đạo và AI hỗ trợ: Giảng viên tiến hành chẩn đoán phân nhóm theo trình độ sinh viên và thiết kế bài tập mang tính cá nhân hoá. Nhờ các công cụ AI như ChatGPT, Gemini, Doubao… hỗ trợ. Giảng viên cung cấp tài nguyên giảng dạy cá nhân hóa cho từng cấp độ sinh viên, bao gồm: bài tập mở rộng từ vựng, file ngữ âm, các trường hợp về khác biệt văn hóa. Đồng thời, Giảng viên cần xác định rõ giới hạn sử dụng AI, hướng dẫn sinh viên chuẩn hóa hành vi xem trước bài, ngăn chặn tình trạng phụ thuộc, lạm dụng AI.

Sinh viên tự chủ và AI hỗ trợ: Dưới sự định hướng của giảng viên, sinh viên sử dụng AI để hoàn thành các nhiệm vụ xem trước bài như tra cứu từ vựng, so sánh đối chiếu câu cú, tìm hiểu bối cảnh văn hóa, từ đó bước đầu kiến tạo hệ thống kiến thức. Sinh viên cần sàng lọc và suy ngẫm về nội dung do AI tạo sinh, hình thành ghi chú cá nhân, đặt nền móng cho việc học tập tư duy phản biện trên lớp.

5.2 Giai đoạn nghiêm cấm sử dụng AI trên lớp (AI-Free Class Policy)

Mục tiêu: Thông qua lớp học "phi kỹ thuật số" để thúc đẩy khả năng tư duy độc lập, suy luận logic, khả năng phản xạ và vận dụng ngôn ngữ của sinh viên, tăng cường tư duy phản biện xem người học làm trung tâm.

Giảng viên chủ đạo : Giảng viên quản lý sát sao, nghiêm ngặt việc thực hiện "chính sách không sử dụng AI" (No AI in Class), cấm sinh viên sử dụng điện thoại di động hoặc kết nối mạng để tra cứu nền tảng AI và từ điển, nhằm đảm bảo lớp học tập trung vào việc kiến tạo tri thức và tương tác khám phá. Giảng dạy trên lớp chủ yếu dưới các hình thức: phân tích các bản dịch mẫu của AI, chẩn đoán lỗi dịch, tranh luận nhóm, và mô phỏng tình huống, thúc đẩy sinh viên rèn luyện tư duy phản biện vàkhả năng phản xạ tức thời trong môi trường không có AI.

Sinh viên tự chủ và rèn luyện tư duy phản biện: Sinh viên cần tuân thủ quy định của lớp học, hoàn thành nhiệm vụ trên lớp một cách độc lập, vận dụng kiến thức tích lũy và khả năng phán đoán logic của bản thân để trả lời và dịch. Thông qua quá trình suy nghĩ phản xạ tức thời khi "thoát ly AI", sinh viên có thể nội hóa thành quả học tập AI trước đó thành khả năng hiểu và áp dụng vào thực tế, từng bước hình thành tư duy phản biện và tư duy sáng tạo.

5.3 Giai đoạn Đào sâu sau giờ học (Post-AI Reflection)

Mục tiêu: Thông qua phản hồi của AI và học tập phản tư để đạt được tái kiến tạo năng lực, củng cố và mở rộng những kiến thức đã tiếp thu trên lớp.

Giảng viên chủ đạo và AI hỗ trợ: Giảng viên cung cấp các nhiệm vụ theo nhóm được cá nhân hóa cho từng cấp độ sinh viên, đồng thời sử dụng nền tảng AI để đề xuất các bài tập mở rộng

và tài liệu ngữ âm, giúp sinh viên củng cố kiến thức một cách có trọng tâm. Giảng viên cần nhận xét và hướng dẫn nhật ký phản tư và kết quả nhiệm vụ AI của sinh viên, thúc đẩy khả năng chuyển giao học tập và tự điều chỉnh của họ.

Sinh viên tự chủ và Phản tư với AI: Sinh viên sử dụng AI để nhận phản hồi tức thời và tự hoàn thành việc phân tích lỗi dịch và ghi chép tổng kết, iết "Nhật ký phản tư khi sử dụng AI", tổng kết các vấn đề học tập và phương án cải thiện. Thông qua chu trình phản tư này, sinh viên dần chuyển từ "phụ thuộc vào AI" sang "tự chủ kiểm soát", từ đó nội hóa và nâng cao năng lực học tập một cách bền vững.

6. Những hạn chế và định hướng nghiên cứu trong tương lai

6.1 Hạn chế

(1) **Phạm vi mẫu còn hạn chế.** Nghiên cứu này chỉ lựa chọn đối tượng là sinh viên chuyên ngành tiếng Trung Quốc của Trường Đại học Ngoại ngữ – Tin học Thành phố Hồ Chí Minh, do đó tính khu vực và tính đặc thù chuyên ngành có thể ảnh hưởng đến khả năng khái quát hóa và suy rộng kết quả nghiên cứu.

(2) **Chu kỳ thực nghiệm ngắn.** Thời gian tiến hành nghiên cứu chỉ kéo dài trong một học kỳ, chưa đủ để phản ánh biến động dài hạn và tính bền vững của hiệu quả mô hình dạy học hợp tác người–máy.

(3) **Tính biến động của công nghệ cao.** Các nền tảng AI thay đổi nhanh chóng, việc cập nhật, nâng cấp tính năng và giao diện có thể ảnh hưởng đến mức độ thích ứng trong giảng dạy cũng như tính ổn định của kết quả nghiên cứu.

(4) **Dữ liệu còn đơn nhất**. Mặc dù nghiên cứu đã kết hợp khảo sát bằng bảng hỏi và quan sát lớp học với phân tích thống kê, song thiếu dữ liệu định lượng sâu hơn, điều này có thể hạn chế khả năng phân tích chi tiết quá trình học tập.

6.2 Định hướng nghiên cứu trong tương lai

Nghiên cứu trong tương lai có thể được mở rộng theo các hướng sau:

1. Mở rộng quy mô mẫu và nghiên cứu so sánh liên khoa, liên trường: Thu hút người học đến từ nhiều khu vực, nhiều chuyên ngành và đa ngôn ngữ, thực hiện nghiên cứu thực tiễn liên trường với phân tích định lượng chuyên sâu, nhằm nâng cao tính phổ quát và khả năng ứng dụng của kết quả.
2. Theo dõi dọc và giám sát động: Tiến hành nghiên cứu kéo dài qua nhiều học kỳ hoặc nhiều năm học, nhằm khám phá ảnh hưởng dài hạn của mô hình đối với chuyển giao tri thức, năng lực đổi mới sáng tạo và khả năng thích ứng nghề nghiệp.

3. Nghiên cứu bản địa hóa đạo đức AI: Tập trung phân tích sự khác biệt giá trị trong ứng dụng AI dạy học giữa các bối cảnh văn hóa Trung – Việt, từ đó đề xuất bộ quy tắc đạo đức sử dụng AI phù hợp với thực tiễn dạy học.

7. Kết luận

Nghiên cứu này lấy mô hình "Giảng viên chủ đạo – AI hỗ trợ – Sinh viên tự chủ" làm trung tâm, nhằm tìm ra cách ứng dụng AI hợp lý trong giảng dạy phiên dịch trong bối cảnh Ai tạo sinh phát triển nhanh chóng. Kết quả thi giữa kỳ và cuối kỳ cho thấy sinh viên có sự tiến bộ rõ rệt về thành tích học tập. Đặc biệt, nhóm sinh viên có nền tảng yếu đạt mức tăng mạnh nhất, trung bình tăng 1,6 điểm (tương đương 43,8%). Kết quả này khẳng định hiệu quả phổ quát của mô hình trong việc nâng cao chất lượng học tập, tạo cơ hội học tập công bằng và khơi gợi động lực học tập.

Nghiên cứu cũng cho thấy, việc sử dụng AI không nhằm thay thế tư duy của người học, mà giúp họ học cách sử dụng công cụ một cách tỉnh táo và chủ động. Dưới chiến lược ba giai đoạn "AI hỗ trợ trước giờ học – tư duy phản biện trong lớp học – phản tư thông minh sau giờ học", sinh viên từng bước chuyển từ vai trò người học thụ động sang người kiến tạo tri thức. Trong quá trình này, giảng viên đảm nhiệm nhiều vai trò: người thiết kế, người hướng dẫn, người dẫn dắt tư duy phản biện và người phản hồi đánh giá. Chiến lược "không dùng AI trong lớp học" giúp sinh viên rèn khả năng phản xạ nhanh, tức thời, tự dịch và diễn đạt,vduy trì sự cân bằng giữa công nghệ và yếu tố con người trong học tập.

Tóm lại, nghiên cứu đã gợi mở một hướng tiếp cận mới cho đổi mới dạy - học môn phiên dịch trong thời đại AI. Trong tương lai, mô hình có thể được mở rộng sang nhiều ngành học khác, triển khai theo dõi dài hạn và kết hợp với các nền tảng thông minh để kiểm chứng hiệu quả lâu dài, hướng tới xây dựng một môi trường giáo dục lấy con người làm trung tâm và đồng hành cùng trí tuệ nhân tạo.

Tài liệu tham khảo

1. Varghese, J., Cleetus, B., & Nair, S. (2025). *The power duo: Unleashing cognitive potential through human-AI synergy in STEM and non-STEM education. Frontiers in Education, 10*, 1534582.
2. Xu, S., Su, Y., & Liu, K. (2024). *Integrating AI for enhanced feedback in translation revision: A mixed-methods investigation of student engagement. arXiv preprint*, arXiv:2410.08581.
3. Duan, H. (2025). *The application of AI translation tools in improving translation quality: A comparative study in higher education. SSRN Electronic Journal.*
4. Fan, K., & Chen, J. (2025). *人工智能在翻译教学中的应用、问题及对策—以西南石油大学为例 [Application, problems and countermeasures of AI in translation teaching: A case study of Southwest Petroleum University]. 现代教育论坛 [Modern Education Forum]*, 5(3), 114–123.
5. Jing Yu, Kanglong Liu *重塑翻译研究：AI 技术影响下的范式转换与未来方向探索 [Reshaping translation studies: Paradigm shifts and future directions under AI influence]. 外语学刊 [Journal of Foreign Languages]*, 47(4), 15–27.
6. Jinhee Kim (2024). *Types of teacher-AI collaboration in K-12 classroom instruction: Chinese teachers' perspective. Education and Information Technologies*, 29(10), 12145–12163.
7. Jia Zhang (2024). *Investigating novice translation students' AI literacy in translation education. The Interpreter and Translator Trainer*, 18(3), 1–20.

Thông tin tác giả

- Thạc sĩ Lưu Cẩm Huệ
- Giảng viên Khoa Ngoại ngữ Trường Đại học Ngoại ngữ -Tin học TP. HCM
- Địa chỉ nơi công tác : số 828 Sư Vạn Hạnh, Phường 3, Quận 10, TP. HCM
- Email : huelc@edu.huflit.vn

SỬ DỤNG AI VÀO CÁC HOẠT ĐỘNG HỌC TẬP DẠNG TRÒ CHƠI, NHẰM MỤC ĐÍCH NÂNG CAO ĐỘNG LỰC HỌC TẬP CHO SINH VIÊN TRONG MÔN BIÊN DỊCH 1 TẠI KHOA NGOẠI NGỮ, HUFLIT

Nguyễn Đức Châu[1]

Tóm tắt : Nghiên cứu này được thực hiện nhằm khám phá tiềm năng của việc ứng dụng trí tuệ nhân tạo (AI) để tạo ra các hoạt động học tập dạng trò chơi, với mục đích nâng cao động lực học tập cho sinh viên trong học tập ngoại ngữ, đặc biệt là chuyên ngành Biên dịch tại Khoa Ngoại ngữ, HUFLIT. Chúng tôi sử dụng bối cảnh học tập tập trung vào việc sử dụng các công cụ và nền tảng AI để thiết kế các tiết mục, phần thưởng, và yếu tố tương tác mang tính trò chơi, từ đó tăng cường sự hứng thú và chủ động của sinh viên đối với môn học Biên dịch. Phương pháp nghiên cứu kết hợp khảo sát sinh viên về mức độ động lực và trải nghiệm học tập, cùng với phân tích hiệu quả học tập thông qua các hoạt động được trò chơi hóa do AI tạo ra. Kết quả nghiên cứu dự kiến sẽ làm sáng tỏ vai trò của AI trong việc cá nhân hóa trải nghiệm học tập và thúc đẩy động lực nội tại của sinh viên trong bối cảnh học ngoại ngữ chuyên sâu.

Từ khóa: Trí tuệ nhân tạo, Trò chơi, Động lực học tập, Học ngoại ngữ, Biên dịch.

I. GIỚI THIỆU

1.1 Bối cảnh nghiên cứu (Background to the Study)

Trong thời đại số hóa hiện nay, trí tuệ nhân tạo (AI) đang đóng vai trò ngày càng quan trọng trong lĩnh vực giáo dục, đặc biệt là trong việc đổi mới phương pháp giảng dạy và học tập. Các nghiên cứu hiện nay cho thấy việc tích hợp AI vào quá trình học tập không chỉ mang lại sự tiện lợi, mà còn mở ra cơ hội cá nhân hóa việc học, tăng tính tương tác và nâng cao hiệu quả tiếp thu kiến thức. Một trong những hướng đi nổi bật hiện nay là việc kết hợp AI với phương pháp trò chơi hóa (gamification) – tức là biến các hoạt động học tập trở nên sinh động, thú vị như một trò chơi nhằm khơi gợi và duy trì động lực học tập.

Trong bối cảnh giảng dạy ngoại ngữ tại Việt Nam nói chung và tại Trường Đại học Ngoại ngữ - Tin học TP.HCM (HUFLIT) nói riêng, sinh viên chuyên ngành Biên dịch thường gặp nhiều khó khăn trong việc ghi nhớ từ vựng, cấu trúc câu cũng như rèn luyện kỹ năng chuyển ngữ hiệu quả. Các nghiên cứu hiện nay cũng cho thấy việc học từ vựng theo lối truyền thống dễ gây nhàm chán, thiếu tính gợi mở và có thể làm giảm động lực học tập lâu dài.

Do đó, việc ứng dụng AI để xây dựng các trò chơi học thuật – chẳng hạn như trò chơi từ vựng có kết hợp ngữ pháp và cấu trúc câu – hứa hẹn sẽ là một giải pháp tiềm năng, góp phần cải

[1] Tiến sĩ, Trường Đại học Ngoại ngữ - Tin học TP.HsCM, *Email: chaund@huflit.edu.vn*

thiện chất lượng học tập cũng như tạo hứng thú cho sinh viên chuyên ngành Biên dịch tại Khoa Ngoại ngữ, HUFLIT.

1.2 Vấn đề nghiên cứu (Statement of the Problem)

Mặc dù có nhiều công cụ và ứng dụng học tập hỗ trợ sinh viên học ngoại ngữ, nhưng việc sử dụng chúng một cách hiệu quả và có hệ thống, đặc biệt là trong chuyên ngành Biên dịch, vẫn chưa được triển khai đồng bộ. Một vấn đề đặt ra là:

Làm thế nào để ứng dụng AI trong việc tạo ra các hoạt động học tập mang tính trò chơi hóa, từ đó nâng cao động lực và hiệu quả học tập cho sinh viên chuyên ngành Biên dịch tại HUFLIT?

1.2.1 Mô tả trò chơi

Trong quá trình học ngoại ngữ, đặc biệt là chuyên ngành Biên dịch, sinh viên cần rèn luyện khả năng sử dụng từ vựng trong ngữ cảnh cụ thể và cấu trúc ngữ pháp chính xác. Một giải pháp khả thi là sử dụng AI để thiết kế các trò chơi học tập, trong đó AI có thể:

- Phân tích nội dung bài học để tự động tạo ra trò chơi từ vựng tương ứng với các chủ đề.
- Kết hợp từ vựng vào các cấu trúc câu có độ dài 22–25 từ, giúp sinh viên rèn luyện khả năng viết và dịch sát ngữ cảnh.
- Cung cấp phản hồi tức thì để sinh viên có thể tự điều chỉnh và cải thiện.
- Phần thưởng đa dạng từ quà tặng, điểm hoạt động, đến điểm thái độ.

1.3 Câu hỏi nghiên cứu (Research Questions)

1. Việc học thông qua trò chơi với sự hỗ trợ của AI có ảnh hưởng như thế nào đến động lực học tập của sinh viên?
2. Sinh viên đánh giá hiệu quả của các trò chơi học tập do AI tạo ra như thế nào trong việc nâng cao vốn từ vựng và cấu trúc ngữ pháp trong môn Biên dịch 1?

1.4 Ý nghĩa của nghiên cứu (Significance of the Study)

Nghiên cứu này có ý nghĩa thiết thực trong bối cảnh đổi mới phương pháp giảng dạy tại các trường đại học. Việc chứng minh được tính hiệu quả của việc trò chơi hóa hoạt động học tập với sự hỗ trợ của AI sẽ giúp các giảng viên, nhà quản lý giáo dục:

- Có thêm cơ sở để triển khai các công cụ học tập hiện đại, giúp sinh viên học tập chủ động và hiệu quả hơn.

- Góp phần đổi mới phương pháp dạy và học ngoại ngữ, phù hợp với xu hướng chuyển đổi số trong giáo dục.
- Nâng cao chất lượng đào tạo ngành Biên dịch, tạo ra nguồn nhân lực có kỹ năng vững vàng và thích ứng nhanh với công nghệ.

1.5 Phạm vi và giới hạn của nghiên cứu (Scope and Delimitation of the Study)

Nghiên cứu này tập trung vào sinh viên chuyên ngành Biên-Phiên dịch tại Khoa Ngoại ngữ, HUFLIT. Phạm vi nghiên cứu bao gồm:

- Việc sử dụng một số công cụ AI cụ thể như ChatGPT, Gemini để xây dựng hoạt động trò chơi trong học tập.
- Đánh giá tác động của hoạt động này đối với động lực học tập và khả năng sử dụng từ vựng, cấu trúc câu, và kỹ năng chuyển ngữ của sinh viên.

Giới hạn của nghiên cứu:

- Không khảo sát sinh viên các môn Biên dịch khác hay các chuyên ngành ngoại ngữ khác ngoài Biên dịch 1.
- Tác động dài hạn của phương pháp này chưa được nghiên cứu sâu, do giới hạn về thời gian thực hiện đề tài.

II. TỔNG QUAN NGHIÊN CỨU

2.1. Giới thiệu

Chương này nhằm mục đích cung cấp một cái nhìn tổng quan về các nghiên cứu hiện có liên quan đến việc ứng dụng yếu tố trò chơi (gamification) trong giáo dục, đặc biệt là trong bối cảnh học ngoại ngữ và Biên dịch. Chương sẽ khám phá các khái niệm cốt lõi, các lý thuyết nền tảng, và các nghiên cứu thực nghiệm đã được thực hiện về hiệu quả của gamification trong việc nâng cao động lực học tập và kết quả học tập của sinh viên. Cuối cùng, chương này sẽ làm nổi bật sự cần thiết của nghiên cứu này trong bối cảnh tạo động lực học khi giảng dạy môn Biên dịch 1 tại Khoa Ngoại ngữ, HUFLIT.

2.2. Khái niệm về Yếu tố Trò chơi

- **Định nghĩa:** Trình bày các định nghĩa khác nhau về gamification từ các học giả uy tín (ví dụ: Deterding et al., 2011; Kapp, 2012), nhấn mạnh vào việc sử dụng các yếu tố thiết kế trò chơi trong các bối cảnh không phải trò chơi.
- **Các yếu tố chính:** Liệt kê và mô tả các yếu tố trò chơi thường được sử dụng trong giáo dục, chẳng hạn như điểm số (points), huy hiệu (badges), bảng xếp hạng (leaderboards),

thử thách (challenges), phần thưởng (rewards), cốt truyện (narratives), và sự hợp tác (collaboration).

- **Phân biệt Trò chơi hóa với Học từ trò chơi:** Làm rõ sự khác biệt giữa gamification (tích hợp yếu tố trò chơi vào hoạt động hiện có) và game-based learning (sử dụng trò chơi được thiết kế sẵn cho mục đích học tập).

2.3. Lý thuyết Nền tảng

- **Lý thuyết Tự quyết (Self-Determination Theory - SDT):** Giải thích cách các yếu tố trò chơi có thể đáp ứng các nhu cầu tâm lý cơ bản của sinh viên về năng lực (competence), tự chủ (autonomy), và sự liên quan (relatedness), từ đó thúc đẩy động lực nội tại và sự tham gia vào quá trình học tập (Deci & Ryan, 2000).
- **Lý thuyết Dòng chảy (Flow Theory):** Thảo luận về cách các yếu tố trò chơi có thể tạo ra trạng thái "dòng chảy" (flow), một trạng thái tập trung cao độ và thích thú khi người học cảm thấy thử thách vừa đủ so với kỹ năng của họ (Csikszentmihalyi, 1990).
- **Lý thuyết Học tập Trải nghiệm (Experiential Learning Theory):** Phân tích cách các hoạt động mang tính trò chơi có thể cung cấp trải nghiệm học tập tích cực, cho phép sinh viên học hỏi thông qua hành động và phản ánh (Kolb, 1984).

2.4. Nghiên cứu về Gamification trong Giáo dục và Học Ngoại ngữ

- **Ứng dụng Gamification trong Giáo dục Đại học:** Trình bày các nghiên cứu đã khám phá việc sử dụng gamification trong các môn học khác nhau ở bậc đại học, tập trung vào tác động của nó đến động lực, sự tham gia, và kết quả học tập (ví dụ: Buckley & Doyle, 2016; Hamari et al., 2014).
- **Gamification trong Học Ngoại ngữ:** Đánh giá các nghiên cứu cụ thể về việc áp dụng gamification trong việc dạy và học ngoại ngữ, bao gồm việc cải thiện từ vựng, ngữ pháp, kỹ năng nghe, nói, đọc và viết (ví dụ: Chinnappan & Chandler, 2010; Lee & Hammer, 2011).
- **Trò chơi trong Dạy và Học Biên dịch:** Rà soát các nghiên cứu (nếu có) về việc sử dụng yếu tố trò chơi trong việc dạy và học biên dịch. Nếu có ít hoặc không có nghiên cứu trực tiếp, chúng ta sẽ thảo luận về tiềm năng và sự phù hợp của Trò chơi trong việc phát triển các kỹ năng biên dịch.

2.5. Động lực Học tập trong Môn Biên dịch 1

- **Đặc điểm của Động lực Học tập Biên dịch 1:** Thảo luận về các yếu tố cụ thể có thể ảnh hưởng đến động lực học tập của sinh viên trong môn Biên dịch 1, chẳng hạn như sự phức tạp của nhiệm vụ, mức độ tự tin, nhận thức về giá trị của môn học, và sự tương tác trong lớp học.

- **Tầm quan trọng của Động lực:** Nhấn mạnh tầm quan trọng của động lực học tập trong việc đạt được thành công trong môn biên dịch, bao gồm việc nâng cao sự tham gia, sự kiên trì, và khả năng tiếp thu kiến thức và kỹ năng.

2.6. Khoảng trống Nghiên cứu và Mục tiêu Nghiên cứu

- **Xác định khoảng trống:** Chỉ ra những lĩnh vực mà nghiên cứu hiện tại còn hạn chế, đặc biệt là trong bối cảnh ứng dụng Trò chơi cụ thể cho môn Biên dịch 1.
- **Mục tiêu nghiên cứu:** Nêu rõ mục tiêu chính của nghiên cứu này là khám phá tác động của việc tích hợp các yếu tố trò chơi vào môn Biên dịch 1 tại Khoa Ngoại ngữ, HUFLIT đối với động lực học tập của sinh viên.

Tóm tắt các điểm chính của chương này, nhấn mạnh sự liên quan của các nghiên cứu trước đây đến đề tài hiện tại và tái khẳng định tầm quan trọng của việc nghiên cứu tác động của Trò chơi đối với động lực học tập của sinh viên môn Biên dịch 1. Chương tiếp theo sẽ trình bày phương pháp nghiên cứu được sử dụng để đạt được các mục tiêu đã đề ra.

III. PHƯƠNG PHÁP NGHIÊN CỨU

Trong chương này, chúng tôi trình bày chi tiết phương pháp nghiên cứu được sử dụng để thu thập và phân tích dữ liệu liên quan đến việc ứng dụng AI tạo trò chơi hóa nhằm nâng cao động lực học tập cho sinh viên môn Biên dịch 1 tại Khoa Ngoại ngữ, HUFLIT. Nghiên cứu này sử dụng phương pháp nghiên cứu định lượng để đo lường một cách khách quan mức độ động lực học tập của sinh viên trước và sau khi tham gia các hoạt động được trò chơi hóa do AI tạo ra. Dữ liệu định lượng được thu thập thông qua bảng câu hỏi khảo sát gửi cho SV qua nền tảng trực tuyến Zalo.

3.1 Thiết kế Nghiên Cứu

Nghiên cứu này sử dụng thiết kế tiền thực nghiệm theo dạng tiền kiểm và hậu kiểm cùng một nhóm (one-group pretest-posttest design). Thiết kế này bao gồm việc đo lường mức độ động lực học tập của nhóm sinh viên trước khi triển khai các hoạt động trò chơi hóa (giai đoạn tiền can thiệp), sau đó giới thiệu các hoạt động này (giai đoạn can thiệp), và cuối cùng đo lường lại mức độ động lực học tập của cùng nhóm sinh viên sau khi kết thúc giai đoạn can thiệp (giai đoạn hậu can thiệp). Sự khác biệt giữa kết quả đo lường trước và sau can thiệp sẽ được phân tích để đánh giá tác động của việc sử dụng AI tạo trò chơi hóa.

3.2 Đối Tượng Nghiên Cứu và Mẫu

Đối tượng nghiên cứu là sinh viên đang theo học môn Biên dịch 1, Học kỳ 2, năm học 2024-2025 tại Khoa Ngoại ngữ, HUFLIT. Mẫu nghiên cứu bao gồm 160 sinh viên đang theo học 4 lớp Biên dịch 1 năm học 2024-2025, dựa trên sự đồng ý tham gia của 126 sinh viên trong các lớp nói trên, thông qua phiếu trả lời khảo sát của họ. Số lượng khảo sát này được xác định dựa trên số lượng sinh viên hiện có và đảm bảo đủ lớn để tiến hành phân tích thống kê.

3.3 Công Cụ Thu Thập Dữ Liệu

Công cụ chính để thu thập dữ liệu trong nghiên cứu này là bảng câu hỏi khảo sát. Bảng câu hỏi sẽ được thiết kế dựa trên các thang đo đã được kiểm chứng về động lực học tập, có thể bao gồm các khía cạnh như:

- Động lực nội tại: Mức độ hứng thú, tò mò và niềm vui khi tham gia vào hoạt động học tập.
- Động lực ngoại tại: Mức độ bị thúc đẩy bởi phần thưởng, điểm số hoặc sự công nhận từ bên ngoài.
- Giá trị nhiệm vụ: Mức độ sinh viên nhận thức được tầm quan trọng và tính hữu ích của nhiệm vụ học tập.
- Kỳ vọng thành công: Mức độ sinh viên tin tưởng vào khả năng hoàn thành tốt nhiệm vụ học tập.
- Nỗ lực: Mức độ sinh viên sẵn sàng đầu tư thời gian và công sức vào hoạt động học tập.

Bảng câu hỏi sẽ sử dụng thang đo Likert với các mức độ đánh giá từ hoàn toàn không đồng ý đến hoàn toàn đồng ý (ví dụ: 1 - Hoàn toàn không đồng ý, 5 - Hoàn toàn đồng ý). Bảng câu hỏi sẽ bao gồm các câu hỏi đo lường động lực học tập, bao gồm các câu hỏi được thiết kế để đánh giá các khía cạnh của động lực học tập như đã nêu ở trên.

Bảng câu hỏi sẽ được trình bày rõ ràng, dễ hiểu và được thí điểm trên một nhóm nhỏ sinh viên không thuộc mẫu nghiên cứu chính để đảm bảo tính tin cậy và giá trị của công cụ.

3.4 Quy Trình Thu Thập Dữ Liệu

Quy trình thu thập dữ liệu sẽ diễn ra theo các bước sau:

1. Giai đoạn tiền can thiệp: Bảng câu hỏi khảo sát về mức độ động lực học tập sẽ được phát trực tiếp hoặc thông qua nền tảng trực tuyến Zalo cho 160 sinh viên tham gia nghiên cứu vào đầu học kỳ hoặc trước khi triển khai các hoạt động trò chơi hóa. Sinh viên sẽ được hướng dẫn đầy đủ về mục đích của nghiên cứu và đảm bảo về tính bảo mật của thông tin cá nhân và câu trả lời.
2. Giai đoạn can thiệp: Các hoạt động học tập môn Biên dịch 1 được trò chơi hóa, do AI tạo ra, sẽ được tích hợp vào quá trình giảng dạy trong một khoảng thời gian nhất định 10 tuần, tính từ tuần 2 đến hết tuần 11, trong hệ thống học trực tiếp 12 tuần của HUFLIT. Các hoạt động này có thể bao gồm các thử thách dịch thuật, các thử thách về từ vựng, các thử thách về ngữ pháp, tất cả đều có tính điểm hoạt động hoặc các yếu tố tương tác khác được thiết kế để tăng tính hấp dẫn và cạnh tranh lành mạnh.

3. Giai đoạn hậu can thiệp: Sau khi kết thúc giai đoạn can thiệp, bảng câu hỏi khảo sát có điều chỉnh chút đỉnh để phù hợp với bối cảnh sau can thiệp sẽ được phát lại cho cùng 160 sinh viên đã tham gia khảo sát ở giai đoạn trước.

3.5 Phân Tích Dữ Liệu

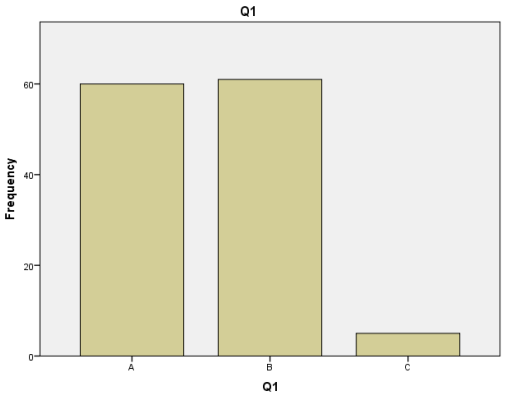
Dữ liệu thu thập được từ bảng câu hỏi khảo sát sẽ được xử lý và phân tích bằng phần mềm thống kê SPSS. Kết quả phân tích sẽ cho thấy liệu việc sử dụng AI tạo trò chơi hóa có tác động đáng kể đến mức độ động lực học tập của sinh viên môn Biên dịch 1 hay không. Chương tiếp theo sẽ trình bày kết quả chi tiết của quá trình thu thập và phân tích dữ liệu này.

IV. PHÂN TÍCH VÀ DIỄN GIẢI DỮ LIỆU

Chương này trình bày kết quả phân tích dữ liệu thu thập từ khảo sát về việc ứng dụng AI vào các hoạt động học tập được trò chơi hóa (Gamification) nhằm nâng cao động lực học tập cho sinh viên môn Biên dịch 1 tại Khoa Ngoại ngữ, HUFLIT. Dữ liệu bao gồm 12 câu hỏi (Q1–Q12) với các lựa chọn trả lời khác nhau, phản ánh thái độ và trải nghiệm của sinh viên đối với phương pháp học tập mới này. Thông qua phân tích định lượng và diễn giải chi tiết, chương này sẽ làm rõ xu hướng chung, sự khác biệt trong ý kiến của sinh viên, cũng như những yếu tố ảnh hưởng đến hiệu quả của Gamification trong học tập. Kết quả phân tích sẽ là cơ sở quan trọng để đưa ra các đề xuất cụ thể trong chương 5 tiếp theo.

1. Q1: I am interested in game-based learning activities in my Translation-1 class.

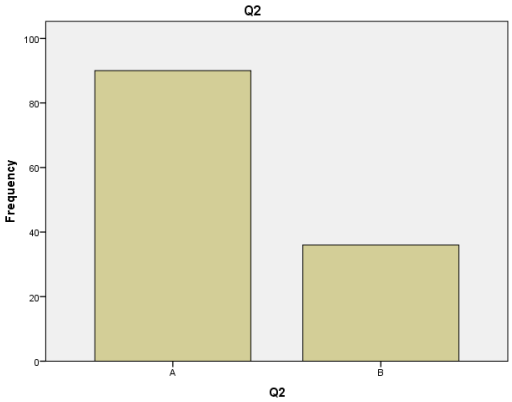
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	60	47.6	47.6	47.6
	B	61	48.4	48.4	96.0
	C	5	4.0	4.0	100.0
	Total	126	100.0	100.0	



Dữ liệu này cho thấy kết quả khảo sát với tổng số 126 phản hồi, được phân thành ba nhóm: A, nhóm chọn B, và nhóm chọn C. Nhóm chọn B có số lượng 61 (48.4%), theo sau là nhóm A với 60 (47.6%). Nhóm chọn C có số lượng ít nhất với 5 (4.0%). Không có ai chọn D và E. Tất cả dữ liệu đều hợp lệ và được bao gồm trong phân tích.

1. Q2: I feel more motivated to participate in class when games are included.

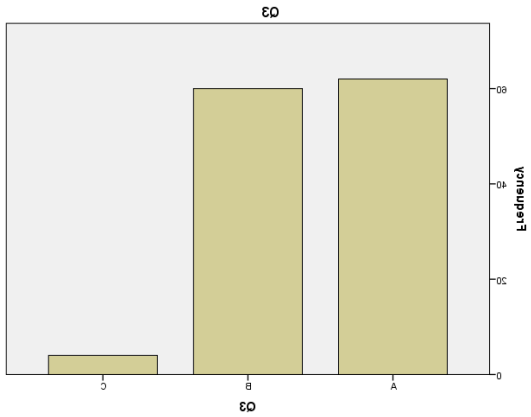
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	90	71.4	71.4	71.4
	B	36	28.6	28.6	100.0
	Total	126	100.0	100.0	



Dữ liệu này cho thấy rằng trong tổng số 126 khảo sát, nhóm chọn A chiếm phần lớn với 90 lần (71.4%), trong khi nhóm B có 36 lần (28.6%). Không có dữ liệu nào bị thiếu hoặc không hợp lệ trong tập dữ liệu này.

Q3: Games help me feel more confident in my translation skills.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	62	49.2	49.2	49.2
	B	60	47.6	47.6	96.8
	C	4	3.2	3.2	100.0
	Total	126	100.0	100.0	

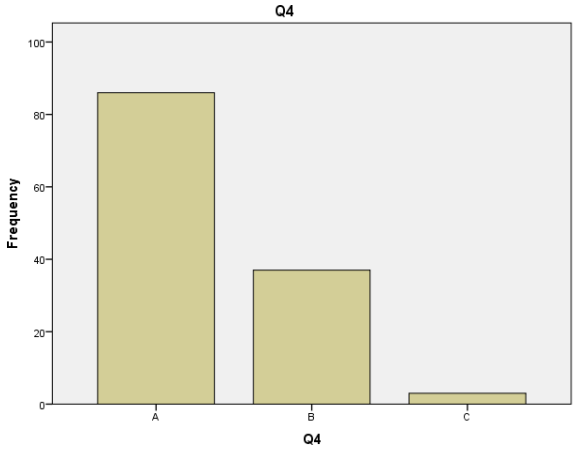


- Lựa chọn A nhận được 62 phản hồi, chiếm 49.2% tổng số phản hồi hợp lệ.
- Lựa chọn B có 60 người chọn, tương đương với 47.6% tổng số phản hồi hợp lệ.
- Lựa chọn C là lựa chọn ít phổ biến nhất với chỉ 4 người chọn, chiếm 3.2% tổng số phản hồi hợp lệ.

  Không có phản hồi nào cho các lựa chọn D hoặc E.

3. Q4: I am more likely to raise my hand during game activities. 4.

		Frequency	Percent	Valid Percent	Cumulative Percent	
Valid	A	86	68.3	68.3	68.3	
	B	37	29.4	29.4	97.6	
	C	3	2.4	2.4	100.0	
	Total	126	100.0	100.0		



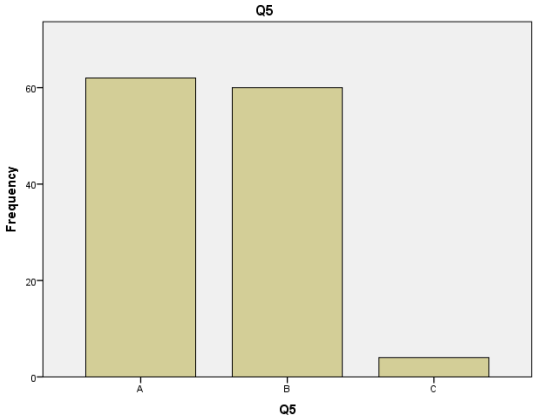
- Lựa chọn A lần này nhận được 86 phản hồi, chiếm một tỷ lệ đáng kể là 68.3% tổng số phản hồi hợp lệ.
- Lựa chọn B có 37 người chọn, tương đương với 29.4% tổng số phản hồi hợp lệ.
- Lựa chọn C vẫn là lựa chọn ít được ưa chuộng nhất với chỉ 3 người chọn, chiếm 2.4% tổng số phản hồi hợp lệ.

  Không có phản hồi nào cho các lựa chọn D hoặc E.

Q5: Games help me remember new vocabulary better than traditional exercises.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	62	49.2	49.2	49.2
	B	60	47.6	47.6	96.8
	C	4	3.2	3.2	100.0

	Total	126	100.0	100.0	

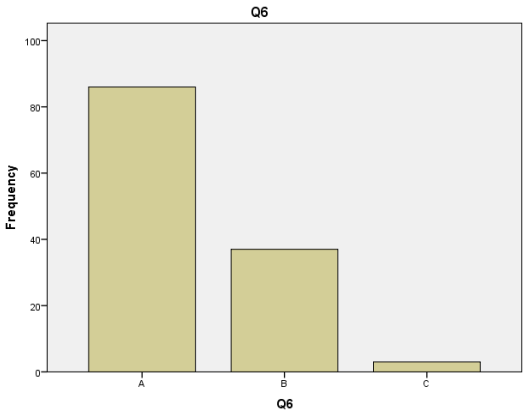


- Lựa chọn A nhận được 62 phiếu, chiếm 49.2% tổng số phản hồi.
- Lựa chọn B có 60 người lựa chọn, tương đương 47.6% tổng số phản hồi.
- Lựa chọn C có số lượng người chọn ít nhất, với chỉ 4 phiếu, chiếm 3.2% tổng số phản hồi.

Không có ai chọn các phương án D hoặc E. Tất cả dữ liệu đều hợp lệ và đã được đưa vào phân tích.

5. Q6: Games improve my understanding of grammar rules in translation.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	86	68.3	68.3	68.3
	B	37	29.4	29.4	97.6
	C	3	2.4	2.4	100.0
	Total	126	100.0	100.0	

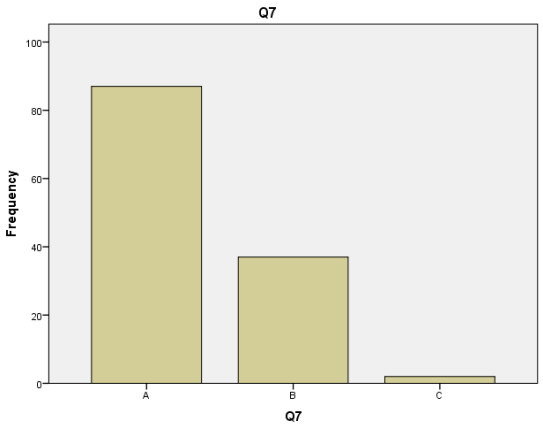


- Lựa chọn A chiếm ưu thế rõ rệt với 86 người chọn, tương đương 68.3% tổng số phản hồi.
- Lựa chọn B có 37 người lựa chọn, chiếm 29.4% tổng số phản hồi.
- Lựa chọn C là lựa chọn ít phổ biến nhất, chỉ có 3 người chọn, chiếm 2.4% tổng số phản hồi.

Tương tự như các bộ dữ liệu khác, không có phản hồi nào cho các lựa chọn D hoặc E, và tất cả các phản hồi đều được xác nhận là hợp lệ.

Q7: Games encourage me to practice translation more frequently.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	87	69.0	69.0	69.0
	B	37	29.4	29.4	98.4
	C	2	1.6	1.6	100.0
	Total	126	100.0	100.0	



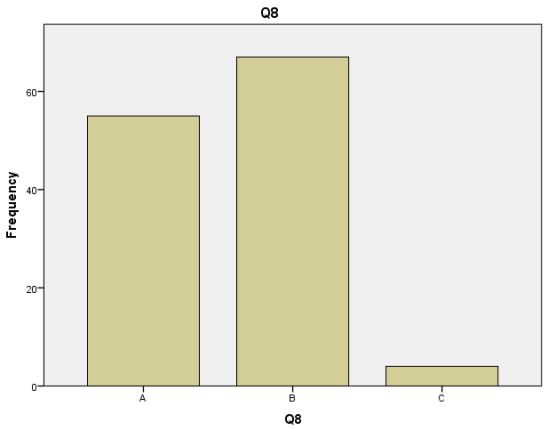
Phân bố phản hồi cho từng lựa chọn như sau:

- Lựa chọn A nhận được 87 phiếu, chiếm tỷ lệ cao nhất là 69.0% tổng số phản hồi.
- Lựa chọn B có 37 người lựa chọn, tương đương 29.4% tổng số phản hồi.
- Lựa chọn C có số lượng người chọn ít nhất, chỉ với 2 phiếu, chiếm 1.6% tổng số phản hồi.

Tương tự như các bộ dữ liệu trước, không có phản hồi nào dành cho các lựa chọn D hoặc E, và tất cả 126 phản hồi đều được xác nhận là hợp lệ và được đưa vào phân tích.

7. Q8: I learn more when the instructor gives feedback during/after games.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	55	43.7	43.7	43.7
	B	67	53.2	53.2	96.8
	C	4	3.2	3.2	100.0
	Total	126	100.0	100.0	



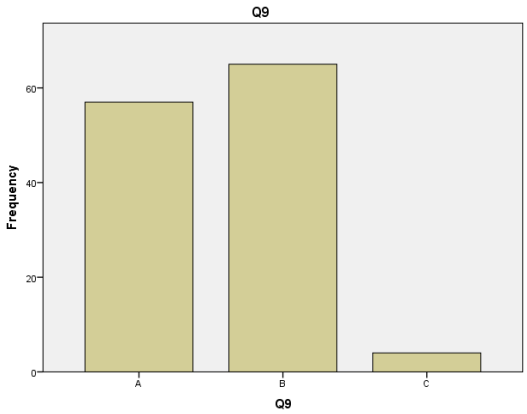
- Lựa chọn A nhận được 55 phiếu, chiếm 43.7% tổng số phản hồi.
- Lựa chọn B là lựa chọn phổ biến nhất trong bộ dữ liệu này với 67 người chọn, tương đương 53.2% tổng số phản hồi.
- Lựa chọn C có số lượng người chọn ít nhất, với chỉ 4 phiếu, chiếm 3.2% tổng số phản hồi.

Tương tự như các bộ dữ liệu trước, không có phản hồi nào cho các lựa chọn D hoặc E, và tất cả 126 phản hồi đều được xác nhận là hợp lệ và đã được đưa vào phân tích.

Q9: I enjoy the following game format the most: *(more than 1 choice)*

- A. Translation quizzes
- B. Word races
- C. Role-play scenarios
- D. Flashcards

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	57	45.2	45.2	45.2
	B	65	51.6	51.6	96.8
	C	4	3.2	3.2	100.0
	Total	126	100.0	100.0	

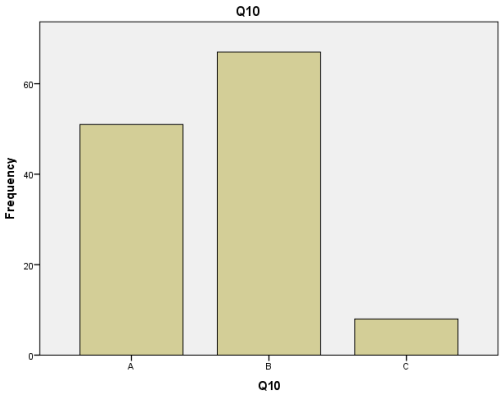


- Lựa chọn A nhận được 57 phiếu, chiếm 45.2% tổng số phản hồi.
- Lựa chọn B tiếp tục là lựa chọn phổ biến nhất với 65 người chọn, tương đương 51.6% tổng số phản hồi.
- Lựa chọn C có số lượng người chọn ít nhất, với chỉ 4 phiếu, chiếm 3.2% tổng số phản hồi.

Tương tự như tất cả các bộ dữ liệu trước, không có phản hồi nào cho các lựa chọn D hoặc E, và tất cả 126 phản hồi đều được xác nhận là hợp lệ và đã được đưa vào phân tích.

9. Q10: Game activities help me prepare for exams.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	51	40.5	40.5	40.5
	B	67	53.2	53.2	93.7
	C	8	6.3	6.3	100.0
	Total	126	100.0	100.0	

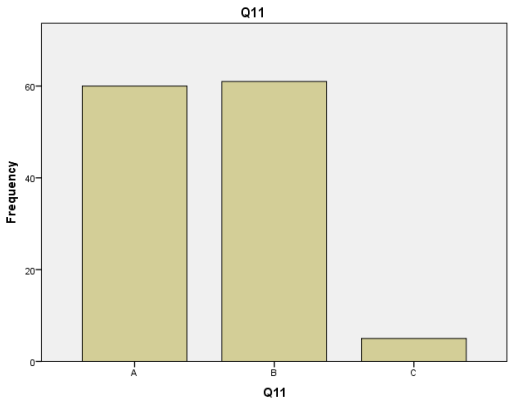


- Lựa chọn A nhận được 51 phiếu, chiếm 40.5% tổng số phản hồi.
- Lựa chọn B tiếp tục là lựa chọn phổ biến nhất với 67 người chọn, tương đương 53.2% tổng số phản hồi.
- Lựa chọn C có số lượng người chọn cao hơn so với các bộ dữ liệu trước, với 8 phiếu, chiếm 6.3% tổng số phản hồi.

Tương tự như tất cả các bộ dữ liệu trước, không có phản hồi nào cho các lựa chọn D hoặc E, và tất cả 126 phản hồi đều được xác nhận là hợp lệ và đã được đưa vào phân tích.

Q11: Game-based learning has improved my translation skills in Translation 1.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	60	47.6	47.6	47.6
	B	61	48.4	48.4	96.0
	C	5	4.0	4.0	100.0
	Total	126	100.0	100.0	

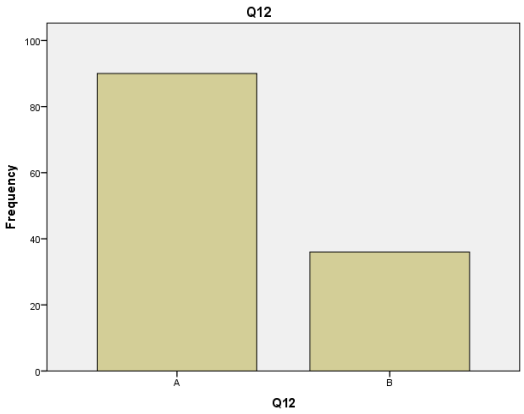


- Lựa chọn A nhận được 60 phiếu, chiếm 47.6% tổng số phản hồi.
- Lựa chọn B là lựa chọn phổ biến nhất trong bộ dữ liệu này với 61 người chọn, tương đương 48.4% tổng số phản hồi.
- Lựa chọn C có số lượng người chọn ít nhất, với chỉ 5 phiếu, chiếm 4.0% tổng số phản hồi.

Tương tự như tất cả các bộ dữ liệu trước, không có phản hồi nào cho các lựa chọn D hoặc E, và tất cả 126 phản hồi đều được xác nhận là hợp lệ và đã được đưa vào phân tích.

Q12: More games should be used in translation courses.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	90	71.4	71.4	71.4
	B	36	28.6	28.6	100.0
	Total	126	100.0	100.0	



- Lựa chọn A nhận được 90 phiếu, chiếm tỷ lệ rất cao là 71.4% tổng số phản hồi.
- Lựa chọn B có 36 người lựa chọn, tương đương 28.6% tổng số phản hồi.
- Lựa chọn C có số lượng người chọn là 0.
- Lựa chọn D có số lượng người chọn là 0.
- Lựa chọn E có số lượng người chọn là 0.

Tương tự như tất cả các bộ dữ liệu trước, không có phản hồi nào cho các lựa chọn D hoặc E, và tất cả 126 phản hồi đều được xác nhận là hợp lệ và đã được đưa vào phân tích.

4.4. Tổng hợp kết quả

- Xu hướng chung: Đa số sinh viên (trên 68% ở nhiều câu hỏi) ủng hộ việc sử dụng AI và Gamification trong học tập, đặc biệt là chuyên ngành Biên dịch.
- Khác biệt ý kiến: Một số câu hỏi cho thấy sự phân hóa ý kiến giữa lựa chọn A và B, phản ánh nhu cầu cần có thêm thông tin hoặc trải nghiệm thực tế để sinh viên đánh giá chính xác hơn.
- Lựa chọn C: Luôn chiếm tỷ lệ thấp, chứng tỏ hầu hết sinh viên đều có quan điểm rõ ràng về chủ đề nghiên cứu.

4.5. Kết luận chương 4

Dữ liệu khảo sát cho thấy:

1. AI và Gamification được đánh giá cao về khả năng nâng cao động lực học tập, với tỷ lệ ủng hộ lên đến 71.4% ở một số câu hỏi.
2. Cần tăng cường nhận thức và trải nghiệm để giảm bớt sự phân hóa ý kiến ở một số khía cạnh cụ thể.

3. Ứng dụng thực tế cần được triển khai linh hoạt, phù hợp với nhu cầu và phản hồi của sinh viên.

Kết quả này đã trả lời phù hợp và mang tính tích cực cho 3 câu hỏi nghiên cứu ở mục 1.3. Kết quả này cũng là cơ sở quan trọng để đề xuất các giải pháp cụ thể trong Chương 5, nhằm tối ưu hóa việc sử dụng AI và Gamification trong giảng dạy ngoại ngữ trong các môn Biên dịch tại khoa Ngoại ngữ, HUFLIT.

V. KẾT LUẬN VÀ ĐỀ XUẤT

5.1 Tóm tắt kết quả nghiên cứu

Nghiên cứu này đã khám phá tiềm năng của việc ứng dụng trí tuệ nhân tạo (AI) vào các hoạt động học tập được trò chơi hóa (Gamification) nhằm nâng cao động lực học tập cho sinh viên môn Biên dịch 1 tại Khoa Ngoại ngữ, HUFLIT. Kết quả khảo sát với 126 sinh viên cho thấy:

- Đa số sinh viên (71.4%) đánh giá cao vai trò của AI và Gamification trong việc tăng cường động lực học tập.
- Đa số sinh viên (71.4%) tỏ ra thú vị với các hoạt động trò chơi hóa do AI tạo ra, như thử thách dịch thuật và trò chơi từ vựng, và được sinh viên tiếp nhận tích cực, đặc biệt trong việc cải thiện kỹ năng ngữ pháp và từ vựng.
- Một bộ phận sinh viên (28.6%) còn băn khoăn về hiệu quả thực tế của phương pháp này, phản ánh sự cần thiết phải có thêm trải nghiệm thực tế.

5.2 Kết luận

Tại Khoa Ngoại ngữ, Trường Đại học Ngoại ngữ - Tin học TP.HCM (HUFLIT), nghiên cứu đã chứng minh rằng việc tích hợp AI vào các hoạt động học tập được trò chơi hóa mang lại nhiều lợi ích thiết thực. Cụ thể:

- Tăng động lực học tập: Các yếu tố trò chơi như điểm số, huy hiệu và bảng xếp hạng đã kích thích sự hứng thú và chủ động của sinh viên.
- Cải thiện hiệu quả học tập: Sinh viên ghi nhớ từ vựng và cấu trúc ngữ pháp tốt hơn thông qua các thử thách tương tác do AI thiết kế.
- Cá nhân hóa trải nghiệm: AI giúp điều chỉnh nội dung học tập phù hợp với năng lực của từng sinh viên, đáp ứng nhu cầu học tập đa dạng.

5.3 Đề xuất

5.3.1 Đối với nhà thiết kế trò chơi mang tính kiểm tra

- Phát triển các công cụ AI có khả năng tạo ra nhiều dạng trò chơi đa dạng, phù hợp với từng chủ đề học tập.

- Tích hợp hệ thống phản hồi tức thì để sinh viên có thể điều chỉnh và cải thiện kỹ năng ngay lập tức.
- Đảm bảo tính thân thiện và dễ sử dụng của các nền tảng Gamification để tối đa hóa trải nghiệm người dùng.

5.3.2 Đối với sinh viên

- Chủ động khám phá và sử dụng các công cụ học tập trò chơi hóa để nâng cao kỹ năng ngoại ngữ.
- Phản hồi ý kiến với giảng viên về hiệu quả của các hoạt động Gamification, giúp cải thiện chất lượng học tập.
- Kết hợp phương pháp truyền thống với các công cụ AI để đạt hiệu quả học tập toàn diện.

5.3.3 Đối với nhà nghiên cứu

- Tiếp tục mở rộng nghiên cứu về tác động dài hạn của AI và Gamification trong các môn học khác nhau.
- Khám phá tiềm năng của các công nghệ AI mới như ChatGPT, Gemini trong việc thiết kế trò chơi học tập.
- Phối hợp với giảng viên và sinh viên để thu thập dữ liệu đa chiều, từ đó đề xuất giải pháp tối ưu.

5.3.4 Đối với giảng viên

- Tích hợp các hoạt động trò chơi hóa vào giáo án giảng dạy, đặc biệt trong các môn đòi hỏi sự sáng tạo như Biên dịch.
- Tổ chức các buổi tập huấn để nâng cao kỹ năng sử dụng công cụ AI và Gamification trong giảng dạy.
- Khuyến khích sinh viên tham gia vào quá trình thiết kế trò chơi học tập, tạo sự gắn kết và hứng thú.

5.4 Hạn chế của nghiên cứu

Nghiên cứu này có một số hạn chế cần lưu ý:

- Phạm vi hẹp: Nghiên cứu chỉ tập trung vào sinh viên môn Biên dịch 1 tại khoa Ngoại ngữ, HUFLIT, chưa khảo sát các môn Biên dịch khác, các chuyên ngành khác trong khoa Ngoại ngữ, hay các chuyên ngành ngoại ngữ khác.
- Thời gian ngắn: Tác động dài hạn của phương pháp chưa được đánh giá sát sao do giới hạn về thời gian thực hiện đề tài.

- Công cụ AI: Nghiên cứu mới chỉ sử dụng một số công cụ AI phổ biến như ChatGPT, Gemini, Deep Seek R1, chưa khám phá hết tiềm năng của các nền tảng khác cũng như sự phát triển rất nhanh chóng của các công cụ AI.

Kết luận chung

Nghiên cứu này cùng với một số nghiên cứu trước đây đã khẳng định tiềm năng to lớn của việc ứng dụng AI vào các hoạt động học tập được trò chơi hóa, đặc biệt trong lĩnh vực dạy và học ngoại ngữ. Kết quả này đã trả lời thỏa đáng cho 3 câu hỏi nghiên cứu ở mục 1.3. Kết quả này góp phần củng cố việc mở ra hướng đi mới cho việc đổi mới phương pháp giảng dạy, góp phần nâng cao chất lượng đào tạo tại HUFLIT nói riêng và các cơ sở giáo dục nói chung. Để phát huy tối đa hiệu quả, cần có sự phối hợp chặt chẽ giữa giảng viên, sinh viên, nhà nghiên cứu và nhà thiết kế công cụ học tập.

TÀI LIỆU THAM KHẢO

Buckley, P., & Doyle, E. (2016). *Gamification and student motivation. *Interactive Learning Environments**, *24*(6), 1162-1175.

Chinnappan, M., & Chandler, P. (2010). *Applying cognitive load theory to multimedia learning in mathematics.* *Computers in Human Behavior*, *26*(6), 1305-1309.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.

Deci, E. L., & Ryan, R. M. (2000). *The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior.* *Psychological Inquiry*, *11*(4), 227-268.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. E. (2011). *From game design elements to gamefulness: Defining "gamification*". In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15).

Hamari, J., Koivisto, J., & Sarsa, S. (2014). *Does gamification work?–A literature review of empirical studies on gamification.* In *2014 47th Hawaii international conference on system sciences* (pp. 3025-3034). IEEE.

Kapp, K. M. (2012). **The gamification of learning and instruction: Game-based methods and strategies for training and education**. Pfeiffer.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.

Lee, J. J., & Hammer, J. (2011). *Gamification in education: What, how, why bother?* *Academic Exchange Quarterly*, *15*(2), 146.

Thông tin tác giả:

TS Nguyễn Đức Châu là giảng viên tiếng Anh tại Khoa Ngoại ngữ, HUFLIT. Ông từng giữ chức vụ Chủ tịch Hội đồng Khoa học Khoa Ngoại ngữ (HUFLIT) năm 2011-2021. Ông từng giữ chức Phó Trưởng Đề án Ngoại ngữ Quốc gia 2020 tại HUFLIT và là thành viên chính thức của Hội đồng Khoa học HUFLIT (2006-2015). Từ năm 2015 đến nay, ông giữ chức vụ Trưởng bộ môn Biên Phiên dịch.

NHẬN DIỆN QUAN HỆ BAO THUỘC BẰNG CÁC DẤU HIỆU HÌNH THỨC

Phạm Văn Lam [1]

TÓM TẮT: Quan hệ bao thuộc (hyponymy) là một trong những quan hệ xương sống, quan trọng nhất trong tổ chức từ vựng của bất kì ngôn ngữ nào. Đó là quan hệ giữa một từ có tính khái quát, được gọi là từ bao, và một từ có tính cụ thể, được gọi là từ thuộc. Quan hệ bao thuộc luôn được xem là một trong những vấn đề cơ bản, cốt lõi trong ngữ nghĩa học nói chung và ngữ nghĩa học từ vựng, đặc biệt là ngữ nghĩa học quan hệ nói riêng. Chúng ta có thể nhận diện quan hệ bao thuộc trong tiếng Việt bằng một loạt dấu hiệu có tính hình thức, được gom thành các nhóm: các dấu hiệu logic, các dấu hiệu cú pháp, các dấu hiệu từ pháp và các dấu hiệu ngữ nghĩa. Các dấu hiệu nhận diện có tính hình thức ở đây đều là những đặc điểm hay yếu tố ngôn ngữ tồn tại trong ngôn ngữ tự nhiên hoặc được dùng để giải thích ngôn ngữ tự nhiên (trong lời nói hay trong các từ điển ngữ văn). Trong số các dấu hiệu này, kết cấu từ vựng – ngữ pháp A là một loại/ loài của B được xem là dấu hiệu hình thức cơ bản, trung tâm và quan trọng nhất. Tất cả các dấu hiệu nhận diện khác đều có thể được xem là những biểu hiện cụ thể của kết cấu từ vựng – ngữ pháp này. Các dấu hiệu hình thức được trình bày ở đây đều có thể được áp dụng một cách thiết thực và hữu dụng vào việc phát hiện tự động, đo lường độ tương tự nghĩa,… trong xử lí ngôn ngữ tự nhiên, trong định nghĩa nghĩa từ ở tử điển, giải thích nghĩa từ cho học sinh, dạy học các tổ chức và triển khai văn bản/ diễn ngôn,…

Từ khoá: quan hệ bao thuộc, từ bao, từ thuộc, nhận diện quan hệ bao thuộc, dấu hiệu hình thức

1. ĐẶT VẤN ĐỀ

Quan hệ bao thuộc là một trong những quan hệ xương sống, quan trọng nhất trong tổ chức từ vựng của bất kì ngôn ngữ nào. Đó là quan hệ giữa một từ bao với các từ thuộc; từ thuộc được hình dung như là một loại của từ bao. Mối quan hệ bao thuộc là mối quan hệ được trải dài từ các từ chung/ danh từ chung cho đến các từ riêng/ danh từ riêng. Số lượng các đơn vị từ vựng có quan hệ bao thuộc ở trong mỗi một ngôn ngữ thường chiếm khoảng 50%; trong tiếng Việt, số lượng đơn vị từ vựng có quan hệ này là 50,44% (Phạm Văn Lam, Nguyễn Phương Thái, 2015). Chính vì vậy, hiện nay quan hệ bao thuộc luôn được xem là một trong những vấn đề cơ bản, cốt lõi trong ngữ nghĩa học nói chung và ngữ nghĩa học từ vựng, đặc biệt là ngữ nghĩa học quan hệ nói riêng. Tuy nhiên, quan hệ này cũng như nhiều quan hệ ngữ nghĩa có tính hệ hình ở bậc từ khác, chưa được Việt ngữ học đề cập và nghiên cứu nhiều. Do đó, bài viết này đặt ra nhiệm vụ chủ đạo là xác định các dấu hiệu có tính hình thức được sử dụng để nhận diện quan hệ bao thuộc. Các dấu hiệu này có tầm quan trọng đặc biệt trong công tác xử lí ngôn ngữ tự nhiên (phát hiện tự động quan hệ nghĩa, đo lường độ tương tự nghĩa,…), trong định nghĩa nghĩa từ của từ điển, trong dạy học tổ chức và triển khai văn bản/ diễn ngôn,… Ngoài phần đặt vấn đề và kết luận, bài viết được kết cấu gồm hai phần: phần 2. trình bày sơ lược lịch sử nghiên

[1] Tiến sĩ, Viện Ngôn ngữ học, *email: phamvanlam1999@gmail.com*

cứu và quan niệm về quan hệ bao thuộc; phần 3 trình bày về các dấu hiện hình thức được sử dụng để nhận diện quan hệ bao thuộc.

2. LỊCH SỬ NGHIÊN CỨU VÀ QUAN NIỆM VỀ QUAN HỆ BAO THUỘC

2.1. Lịch sử nghiên cứu

Quan hệ bao thuộc (hyponymy) là một quan hệ hiển nhiên trong ngôn ngữ. Lâu nay, nó được thường xem xét qua khái niệm bao hàm (inclusion), một khái niệm đã được nghiên cứu nhiều cả trong và ngoài ngôn ngữ học ngay từ thời cổ đại. Tuy nhiên, khái niệm quan hệ bao thuộc chỉ mới xuất hiện gần đây trong ngôn ngữ học. J. Lyons thường được cho là người đầu tiên phổ biến khái niệm này vào lí thuyết ngữ nghĩa học (Lyons, *Structural Semantics*,1963, tr. 69 – 71).

Quan hệ bao thuộc, giống như quan hệ bất tương hợp, "là loại quan hệ ngữ nghĩa có tính hệ hình cơ bản nhất để cấu trúc nên hệ thống từ vựng của ngôn ngữ" (Lyons, 1968, tr. 466). "Khi chúng ta muốn bao hàm nghĩa của một từ nào đó trong một từ khác, chúng ta nói đến quan hệ bao thuộc. Khi chúng ta muốn loại trừ nghĩa của một từ nào đó khỏi một từ khác, chúng ta nói đến quan hệ bất tương hợp" (Crystal, 1994, tr. 267). Tuy nhiên, mặc dù quan trọng như vậy nhưng hiểu biết của chúng ta về quan hệ này, so với các quan hệ ngữ nghĩa truyền thống như đồng nghĩa hay trái nghĩa, quả là còn rất khiêm tốn; và quan hệ này là loại quan hệ quan trọng hơn nhiều lần so với quan hệ đồng nghĩa và trái nghĩa song nó lại "ít quen thuộc" với hầu hết mọi người (Crystal, 1994, tr. 266). "Quan hệ này là loại quan hệ đặc biệt quan trọng đối với các nhà ngôn ngữ học" (Crystal, 1994, tr. 266), song, thậm chí ngay cả tên gọi của quan hệ này và các thuật ngữ được sử dụng để mô tả nó, ngay từ đầu, ở trong ngôn ngữ học, đã có sự nhập nhằng, chưa rõ nghĩa.

Trong lịch sử của ngữ nghĩa học từ vựng tiếng Việt, Đỗ Hữu Châu (1986, 2006) là người đầu tiên đề cập một cách sơ lược đến quan hệ này. Trong *Cơ sở ngữ nghĩa học từ vựng* (1986), ông gọi quan hệ bao thuộc là quan hệ cấp loại, và trong *Giáo trình từ vựng học tiếng Việt* (2006) ông lại gọi là bằng một khái niệm có nhiều nhầm lẫn là quan hệ bao gồm – nằm trong. Tiếp sau Đỗ Hữu Châu, quan hệ ngữ nghĩa này cũng đã được một số tác giả khác nhắc đến dưới tên gọi quan hệ bao nghĩa, chẳng hạn như trong Cao Xuân Hạo và Hoàng Dũng (2005), Hoàng Dũng (2007), Vũ Đức Nghiệu (2010) và Nguyễn Thiện Giáp (2014). Tuy là những người đi sau Đỗ Hữu Châu, nhưng khi đề cập đến quan hệ bao thuộc, tất cả những tác giả này đều không đề cập thêm được điểm nào vượt hay bổ sung cho Đỗ Hữu Châu, nếu không muốn nói là còn đề cập sơ sài hơn và kém sâu sắc hơn so với Đỗ Hữu Châu.

2.2. Quan niệm về quan hệ bao thuộc

Quan hệ bao thuộc là quan hệ giữa một từ có tính khái quát, được gọi là từ bao, và một từ có tính cụ thể, được gọi là từ thuộc. Từ thuộc là từ có tính cụ thể hơn từ bao. Thí dụ, *ghế – ghế đẩu, đũa – đũa con, ghế – đồ dùng, đũa – đồ dùng, đồ dùng – đồ vật, gà – gia cầm, trâu – gia súc, gia súc – động vật,..* Chính vì thế mà quan hệ bao thuộc được xem là loại quan hệ giữa

những đơn vị từ vựng mà có sự "khác biệt mang tính cấp độ về sự cụ thể" (Cruse, 1975, tr. 153). Và sự cụ thể ở đây được hiểu là:

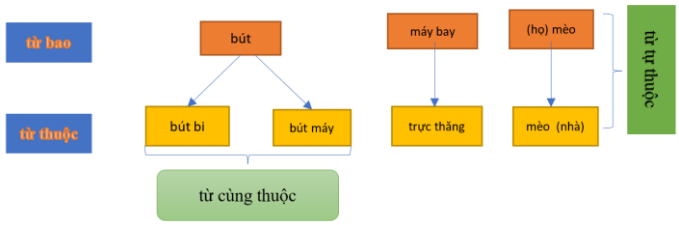
– Phạm vi ứng dụng hay quy chiếu (về mặt ngoại diên và chỉ vật) của từ thuộc nhỏ hơn từ bao; nhỏ hơn tức là cụ thể hơn. Chẳng hạn, tập hợp những thực thể được gọi là *gia cầm* (từ bao) lớn hơn hay nhiều hơn tập hợp những thực thể được gọi là *gà* hay *vịt* (các từ thuộc), vì *gia cầm* gồm cả *gà, vịt, ngan, ngỗng, chim bồ câu*…. cộng lại. Về mặt này, từ bao *gia cầm* bao hàm các từ thuộc *gà, vịt, ngan, ngỗng, chim bồ câu,….;* các từ thuộc *gà, vịt, ngan, ngỗng, chim bồ câu…* bị bao hàm trong *gia cầm.*

– Lượng thông tin có thể có (các nét nghĩa hay nét nội hàm) trong từ thuộc nhiều hơn từ bao; nhiều hơn tức là chi tiết hay cụ thể hơn. Khía cạnh thứ hai của sự cụ thể này còn có thể được diễn giải là nghĩa hay nội hàm của từ bao nằm gọn trong nghĩa hay nội hàm của từ thuộc; hoặc nghĩa hay nội hàm của từ thuộc bao hàm nghĩa hay nội hàm của từ bao. Ví dụ, từ bao *bút* được hiểu là "đồ dùng, để viết", còn từ thuộc *bút bi* được hiểu là *bút + X:* "đồ dùng, để viết" + "có ngoài bằng viên bi nhỏ bằng kim loại gắn ở đầu một ống mực đặc". Xét trên phương diện này, từ thuộc *bút bi* bao hàm *bút*, từ bao *bút* bị bao hàm trong *bút bi*; từ bao *gia cầm* bị bao hàm trong các từ thuộc *gà, vịt, ngan, ngỗng, chim bồ câu,….;* các từ thuộc *gà, vịt, ngan, ngỗng, chim bồ câu…* bao hàm từ bao *gia cầm* (Phạm Văn Lam, 2025).

Hình 1. *Gia cầm* bao hàm *gà*, xét về mặt ngoại diên hay chỉ vật

Hình 2. *Gà* bao hàm *gia cầm*, xét về mặt nội hàm hay nghĩa

Như vậy, quan hệ bao thuộc chính là quan hệ giữa một từ bao (hypernym) với các từ thuộc (hyponym); từ thuộc được hình dung như là một loại của từ bao. Mối quan hệ bao thuộc là mối quan hệ được trải dài từ các từ chung/ danh từ chung cho đến các từ riêng/ danh từ riêng (như *làng – Mộ Trạch*). Một từ bao có thể có nhiều từ thuộc. Một từ thuộc có thể thuộc về nhiêu từ bao. Các từ thuộc cùng cấp xét trong mối quan hệ với từ bao cấp trên được gọi là từ cùng thuộc (cohyponym) (như *gia súc – trầu/ bò/ lợn/ chó,…*). Từ mà vừa có nghĩa thuộc cấp trên và vừa có nghĩa thuộc về cấp dưới được gọi là từ tự thuộc (autohyponym) (như *mèo (nhà) – (động vật họ) mèo*).



Hình 3. Từ tự thuộc và từ cùng thuộc trong quan hệ bao thuộc

3. NHẬN DIỆN QUAN HỆ BAO THUỘC BẰNG CÁC CẤU HIỆU HÌNH THỨC

Gọi X và Y là hai yếu tố có quan hệ bao thuộc, trong đó X là từ thuộc, và Y là từ bao. Ta đọc như sau:

X có từ bao là Y

Y có từ thuộc là X

Hoặc:

X thuộc Y

Y bao X

Có thể có nhiều (loại) dấu hiệu có tính hình thức khác nhau được dùng vào việc nhận diện quan hệ bao thuộc. Song do đây là một quan hệ phức tạp, ít được nghiên cứu, dưới đây chúng tôi xin đề xuất/ giới thiệu một số dấu hiệu nhận diện hình thức quan trọng của quan hệ bao thuộc. Những đề xuất này chỉ là những gợi ý hay kết quả thu thập ban đầu; và bước đầu cũng đã được chúng tôi vận dụng vào trong một số công việc tính toán, nhận diện tự động (2019). Chắc chắn là khi chúng ta đi sâu vào từng khu vực quan hệ bao thuộc cụ thể, chúng ta vẫn có thể tiếp tục phát hiện ra được nhiều (loại) dấu hiệu hình thức quan trọng, có hiệu lực nhận diện mạnh. Công việc này vẫn sẽ được chúng tôi tiến hành thu thập, tinh chỉnh trong thời gian tới.

Các dấu hiệu nhận diện ở đây đều là những đặc điểm hay yếu tố ngôn ngữ tồn tại trong ngôn ngữ tự nhiên hoặc được dùng để giải thích ngôn ngữ tự nhiên (trong lời nói hay trong các từ điển ngữ văn). Cũng nên lưu ý rằng quan hệ bao thuộc là bộ khung của tổ chức trí nhớ ngữ nghĩa, do đó, một mặt, nó hằng tồn trong trí não người bản ngữ, và mặt khác, nó cũng thường tồn trong sử dụng ngôn ngữ hằng ngày, trong văn bản, trong diễn ngôn, trong các hoạt động tư duy logic, trong các hoạt động tâm lí,...

3.1. Các dấu hiệu logic

Quan hệ bao thuộc là loại quan hệ phi đối xứng, chuyển tiếp (còn gọi là chuyển cấp hay bắc cầu). Do đó, quan hệ này cũng hàm ý rằng từ bao luôn có khả năng thay thế cho từ thuộc trong những ngữ cảnh quy chiếu xác định. Quan hệ đồng nghĩa là quan hệ đối xứng, không chuyển tiếp, nhưng cũng là quan hệ mà ở đó hai yếu tố nằm trong thế quan hệ đồng nghĩa hoàn toàn có thể thay thế cho nhau được (về mặt logic thuần tuý). Tính khả thế này chính là một đặc tính mà chúng ta phải đem ra cân nhắc khi xác định đâu là quan hệ đồng nghĩa và đâu là quan

hệ bao thuộc. Vì thế, ở đây hoàn toàn có thể sử dụng lại công thức chẩn đoán tính đồng nghĩa (Cruse, 1986) vào việc xác định tính bao thuộc.

– X là danh từ$_1$, do đó X là danh từ$_2$.

– X là vị từ$_1$, do đó X là vị từ$_2$.

Vì quan hệ bao thuộc là một quan hệ phi đối xứng, do đó, công thức kiểm tra trên cũng là công thức không có tính đối xứng. Nghĩa là:

X là danh từ$_1$, do đó X là danh từ$_2$ *khả chấp,*

nhưng X là danh từ$_2$, do đó X là danh từ$_1$ *cũng không khả chấp.*

Mở rộng quan hệ bao thuộc sang quan hệ cách, chúng ta cũng có thể phát biểu:

X là vị từ$_1$, do đó X là vị từ$_2$ *khả chấp,*

nhưng X là vị từ$_2$, do đó X là vị từ$_1$ *cũng không khả chấp.*

Ví dụ:

(lúa, lúa tẻ)	*lúa tẻ là lúa* **lúa là lúa tẻ*
	lúa tẻ có từ bao là *lúa* *lúa* có (một) từ thuộc là *lúa tẻ*

Như thế, có thể nói rằng công thức chẩn đoán của quan hệ đồng nghĩa được áp dụng vào quan hệ bao thuộc là công thức có tính một chiều, phi phản chiếu, không đối xứng. Tính chất một chiều của quan hệ bao thuộc cũng giả định rằng các đơn vị từ vựng trong một cặp quan hệ bao thuộc không đồng nhất về điều kiện chân trị.

Trong quan hệ bao thuộc, một từ bao có thể có nhiều từ cùng thuộc. Bởi vậy, đối với quan hệ bao thuộc, có thể mở rộng phát biểu về tính khả thế như sau:

Nếu một cặp từ T_1 và T_2 vận hành được trong khung (chẩn đoán) thay thế thì sẽ tiềm tại ít nhất một từ T_3 nào đó mà nó vận hành tốt trong khung thay thế có quan hệ với từ T_2 ; lúc này T_1 và T_3 được gọi là những từ cùng thuộc của T_2.

Trong chuỗi từ ngữ có quan hệ bao thuộc, từ bao nào cũng có từ cùng thuộc. Như vậy, có thể nói rằng:

Sự hiện diện của từ cùng thuộc là một sự tất yếu để xác lập một quan hệ bao thuộc chân thực.

Ngay cả khi chúng ta tưởng như không có từ cùng thuộc (*gà – gà rừng; lợn – lợn lòi, (nước) đá – đá viên*),… thì thực chất chúng vẫn có các từ cùng thuộc, chẳng qua tính chất từ của những đơn vị cùng thuộc đó có thể yếu hơn từ thuộc đang xét mà thôi. Cảm giác không có từ cùng thuộc này thường đến từ việc khi chúng ta chọn một góc nhìn và chúng ta thường chỉ thấy có một ứng viên đã được từ vựng hoá dành cho từ thuộc, ví dụ *ô tô điện* (phân biệt với

những đơn vị mà có tính từ vựng hoá yếu như *ô tô chạy bằng dầu, ô tô chạy bằng xăng*). Về nguyên tắc, trong một cặp quan hệ tôn ti từ bao – từ thuộc, không bào giờ tồn tại duy nhất một từ thuộc, vì từ thuộc này chỉ tồn tại trong sự đối lập, phân lập, khu biệt với từ thuộc kia.

Từ góc nhìn của logic, có thể xác định quan hệ bao thuộc hoặc là về mặt ngoại diên, hoặc là về mặt nội hàm.Từ góc độ ngoại diên, có thể nhận diện quan hệ bao thuộc của X, Y qua phát biểu:

X là một từ thuộc của Y khi và chỉ khi tồn tại một định đề ngữ nghĩa liên quan đến X' và Y' có dạng: ∀[X' (x) → Y"(x)], nhưng lại không có dạng: ∀[Y'(x) → X'(x)].

Phát biểu này là phát biểu được chúng tôi điều chỉnh từ gợi ý của Cann (1993).

Ở đây, X' và Y' là những hằng số logic tương ứng với các biểu thức từ vựng X và Y; trong cách diễn đạt này thì X là từ thuộc của Y và ngoại diên của X' phải nằm trong ngoại diên của Y'.

Ở đây có một vấn đề nhỏ đối với cách diễn đạt logic này. Đó là, từ góc nhìn của ngữ nghĩa học từ vựng, người ta chỉ biết được ngoại diên, tức chỉ vật, chứ người ta không biết được nghĩa, tức nội hàm của X và Y. Như vậy, đối lập ngoại diên với nội hàm, cần bổ sung cách tiếp cận nội hàm để khắc phục nhược điểm này. Về mặt nội hàm, quan hệ bao thuộc có thể được khái quát trong một quy tắc nhận diện:

Nếu X là một từ thuộc của Y thì nội hàm của Y phải là một phần nội hàm không thể thiếu của X.

Ví dụ, nội hàm của *tứ giác* (đa giác có bốn cạnh) là một bộ phận nội hàm của *hình thang* (tứ giác [đa giác có bốn cạnh] có hai cạnh đối song song). *Tứ giác* là từ bao, còn *hình thang* là từ thuộc.

Nhận định này là một nhận định được chứng thực mạnh mẽ trong sử dụng ngôn ngữ tự nhiên cũng như trong lời giải thích của các từ điển. Nội hàm của Y được hiện diện trong nội hàm của X chủ yếu theo một trong hai cách: hoặc từ bao Y được sử dụng lặp lại, hoặc một ngữ biểu thức giải thích có giá trị quy chiếu tương đương với Y. Vẫn tiếp cận vấn đề theo hướng nội hàm, nếu sử dụng thuật ngữ suy ra/ kéo theo, ta có thể có một công thức khác để nhận diện quan hệ bao thuộc:

X là từ thuộc của Y khi và chỉ khi F(X) suy ra/ kéo theo, chứ không được suy ra/ kéo theo từ (F)Y.

Đây là một công thức nhận diện được sử dụng phổ biến trong ngữ nghĩa học từ vựng. Người đầu tiên giới thiệu công thức này từ logic vào trong ngôn ngữ học là Lyons (1963). Công thức này thường được nhắc đến trong logic học. Ngữ nghĩa học hình thức là lĩnh vực sử dụng, khai thác triệt để công thức nhận diện này (Cann, 2010, 2011).

Cách tiếp cận nội hàm này gần với ngữ nghĩa học từ vựng, vì nó được dẫn xuất trực tiếp từ nghĩa, gần với cảm thức bản ngữ và diễn đạt ngôn ngữ tự nhiên. Ví dụ:

Đây là gà suy ra *Đây là gia cầm.*

**Đây là gia cầm* suy ra *Đây là gà*

Đây là lợn suy ra *Đây là gia súc.*

**Đây là gia súc* suy ra *Đây là lợn.*

Tuy nhiên, như đã nói, quan hệ bao thuộc là quan hệ của ngôn ngữ tự nhiên, chứ không phải quan hệ logic thuần tuý. Do đó, không phải lúc nào công thức này cũng có hiệu quả. Không phải lúc nào *Đây là bút bi* đơn phương/ một chiều cũng suy ra *Đây là bút.* Cho nên, sẽ có trường hợp mà ở đó ***F(Y) suy ra F(X),*** như ví dụ sau:

Đây là bút bi suy ra *Đây là bút.*

Đây không phải là bút suy ra *Đây không phải là bút bi.*

Tất cả các loại bút đều có ngòi suy ra *Tất cả các loại bút bi đều có ngòi.*

Nếu đây là bút, bút thì có ngòi suy ra *Nếu đây là bút bi, bút bi thì có ngòi.*

Có một cấu trúc ở cạnh bờ đê suy ra *Có một sự vật ở cạnh bờ đê.*

Không có sự vật nào ở cạnh bờ đê suy ra *Không có cấu trúc nào ở cạnh bờ đê.*

Không có cấu trúc nào ở cạnh bờ đê suy ra *Không có công trình xây dựng nào ở cạnh bờ đê.*

Không có công trình xây dựng nào ở cạnh bờ đê suy ra *Không có ngôi nhà nào ở cạnh bờ đê.*

Không có ngôi nhà nào ở cạnh bờ đê suy ra *Không có ngôi nhà sàn nào ở cạnh bờ đê.*

Các cách nói này, trên thực tế, là một sự hình thức hoá tôn ti kế thừa của quan hệ bao thuộc. Tất cả những cái này đều có sẵn trong trí nhớ ngữ nghĩa của người bản ngữ. Đó là quan hệ logic ngầm ẩn của tri thức và cảm thức ngôn ngữ. Trong *bút bi, công trình xây dựng, nhà, nhà sàn,* một cách lần lượt, đều có sẵn thông tin của *bút, cấu trúc, công trình xây dựng* và *nhà.*

Nói theo chủ nghĩa hoài nghi thì dường như công thức nhận diện trên không có giá trị thao tác trong các mệnh đề mà vị từ của nó bị phủ định. Tuy nhiên, về mặt logic thuần tuý thì ***F(Y) suy ra F(X),*** không có gì là khó hiểu, và không mâu thuẫn với ***X là từ thuộc của Y khi và chỉ khi F(X) suy ra, chứ không được suy ra từ (F)Y*** bởi lẽ quan hệ giữa X và Y là quan hệ bao hàm, có tính kế thừa. Cho nên, những gì là thuộc tính của Y đều có thể được lặp lại ở X.

3.2. Các dấu hiệu cú pháp

Khái niệm cú pháp được hiểu ở đây một cách lỏng lẻo. Những gì không thuộc về từ pháp thì ở đây đều được gán nhãn cú pháp. Các dấu hiệu cú pháp là những dấu hiệu xuất hiện trong phạm vi câu. Tất nhiên, ngoài phạm vi câu, vẫn có thể có các dấu hiệu nhận diện quan hệ bao thuộc. Từ bao thuộc là một cách tổ chức văn bản, cho nên, nó hoạt động trên toàn văn bản; do đó, dấu hiệu nhận diện ngoài câu là những dấu hiệu hết sức phức tạp. Ở đây, tạm thời chưa xét

đến khả năng này. Các dấu hiệu cú pháp ở đây có thể là một khung câu, có thể là một kết cấu từ vựng – ngữ pháp, có thể là một từ trong một mô hình sử dụng xác định nào đó,… Nhận diện các từ ngữ có quan hệ bao thuộc xét về mặt cú pháp, chúng ta có các dấu hiệu quan trọng sau.

Dấu hiệu kết hợp thông thường:

Có thể sử dụng các kết hợp ngữ pháp thông thường để nhận diện quan hệ bao thuộc giữa X và Y. Các kết hợp này phải là những kết hợp ngữ pháp bình thường, đáp ứng được cảm thức bản ngữ.

Một quan hệ ngữ nghĩa có tính hệ hình là quan hệ phải thoả mãn các điều kiện cơ bản như: các từ trong một quan hệ luôn cùng phạm trù ngữ pháp (thường là phạm trù từ loại), cùng chức năng ngữ pháp, có thể cùng kết hợp, có thể thay thế cho nhau được,… Quan hệ bao thuộc, dù là một quan hệ hệ hình nhưng nó khác với các quan hệ khác, như quan hệ trái nghĩa – một quan hệ có tính điển hình về tính chất hệ hình – quan hệ bao thuộc lại là quan hệ tôn ti, do đó, chức năng và khả năng kết hợp ngữ pháp của các yếu tố ngôn ngữ trong một cặp quan hệ bao thuộc có thể không như nhau.

Nói rằng *đàn chim đang bay lượn trên bầu trời* thì được, nhưng *đàn cu gáy bay lượn trên bầu* trời thì đã có thể khiến người ta cau mày (so với *đôi/ con cu gáy đang bay lượn trên bầu trời)*, đến *đàn cánh cụt đang bay lượn trên bầu* trời thì đã là bất thường. Nói rằng *chú đà điểu lao đến bổ thằng bé* thì được, nhưng nếu *chú chim sâu lao về phía thằng bé* thì cũng đã phải xem lại.

Cảm thức bản ngữ về tính bình thường của một kết hợp ngữ pháp nào đó cho các đơn vị từ vựng có quan hệ bao thuộc có thể khác nhau. Nguyên lí chung trong ngữ nghĩa học từ vựng là các từ càng có nghĩa đặc biệt thì dường như kết hợp của nó càng đặc biệt. Cùng là những đơn vị cùng thuộc nhưng rõ ràng người ta chỉ nói *nghển* với *cổ, kiễng* với *chân, với* với *tay, vươn* với *vai, ưỡn* với *ngực*,… Trong quan hệ bao thuộc, X và Y khác nhau về mức độ cụ thể của nghĩa, do đó, cũng khác nhau về mức độ cụ thể của ngữ cảnh. Từ đây, có thể đề xuất một quy tắc về kết hợp của X và Y trong quan hệ bao thuộc như sau:

X là từ thuộc của Y khi và chỉ khi tập hợp các ngữ cảnh thông thường/ bình thường của X là một tập hợp con của các ngữ cảnh thông thường/ bình thường của Y.

Bình thường hay không bình thường về mặt ngữ pháp trong cảm thức bản ngữ chính là tiêu chí để xem xét khả năng kết hợp của X và Y trong quan hệ bao thuộc. Xuất phát từ thuộc tính tôn ti, có thể xem tính bình thường của sự kết hợp của X và Y trong phát biểu trên là có sự phân cấp. Do vậy, có thể có một cách phát biểu khác với phát biểu trên về tính bình thường như sau:

X là từ thuộc của Y khi và chỉ khi có những ngữ cảnh mà ở đó sự có mặt của Y hoặc X được xem làm bình thường hơn sự có mặt của X hoặc Y một cách lần lượt.

Cho nên, cùng là *chim,* theo cách hiểu thông thường, nhưng các con chim thuộc lớp *chim lặn* thì có thể *lặn, bơi;* các con chim thuộc lớp *chim bay* thì *bay, liệng* là bình thường; *chim*

chạy thì *đi, chạy, nhảy, chui vào hang* là bình thường. Dù rằng, cảm thức bản ngữ cứ nói đến *chim* là gán ngay cho *chim* khả năng *bay, hót,…*

Ăn đào thì phải nhè hạt thì được, nhưng *ăn đào lộn hột mà nhè hạt* hay *ăn thanh long mà nhè hạt* thì thậm vô lí. Nói *ăn táo thì phải nhè hạt* thì được vì *táo* nào (các từ cùng thuộc) cũng có *hạt*. Nhưng nói *ăn chuối thì không phải nhè hạt* có thể chấp nhận được, vì thông thường *chuối* không có *hạt*. Tuy nhiên, *chuối hột* hay *chuối cô đơn (chuối cô độc)* cũng là một loại *chuối,* giống như *chuối tây, chuối tiêu, chuối ngự, chuối trứng, chuối lùn, chuối tím,….* Song do tính chất đặc biệt của nét nghĩa "có hột" mà có thể nói *ăn chuối hột/ chuối cô đơn thì phải nhè hạt.*

Ví dụ minh hoạ:

- **Chim cánh cụt thì biết bay* và *Chim thì biết bay.*
- **Gà mái thì biết gáy* và *Gà thì biết gáy.*
- **Xe đạp thì chạy bằng xăng* và *Xe thì chạy bằng xăng.*
- *Trâu thì biết cày ruộng* và **Gia súc thì biết cày ruộng.*
- *Chó thì sủa* và **Động vật thì sủa.*
- *Cá thì có mang* và **Động vật thì có mang.*

Việc xác định các tập hợp ngữ cảnh như thế này đặc biệt có giá trị đối với các công việc liên quan đến tính toán, nhận diện tự động, miêu tả cú pháp, đo lường độ tương tự về nghĩa,…

Các khung cú pháp:

Bản thân quan hệ bao thuộc là một sự phân loại tự nhiên, một sự phân loại tự nhiên được thể hiện ngay trong ngôn ngữ tự nhiên. Chính vì thế mà ngôn ngữ nào cũng có những yếu tố siêu ngôn ngữ có thể thể hiện được bản thân sự phân loại này. Các từ như *loại, loài, giống, chủng, dạng, biến thể, biến dạng,…. kiểu, cách, kiểu cách, phong cách, cách thức, phương pháp, phương thức,…* là những từ như vậy. *Loại, loài, giống* là những từ điển hình dùng để phân loại các thực thể được gọi tên bằng danh từ. *Cách, cách thức* là những từ điển hình dùng để phân loại các quá trình được gọi tên bằng động từ. Những từ như *dạng, kiểu* là những từ có thể được sử dụng để phân loại cho cả thực thể lẫn quá trình.

Nếu X và Y có quan hệ bao thuộc thì X và Y có thể xuất hiện trong công thức nhận diện:

– ***X (X_1, X_2, X_3, …X_n) là một loài/ loại/ kiểu của Y***

Y thì có/ gồm có/ bao gồm X (X_1, X_2, X_3, …X_n)

****Y là một loài/ loại/kiểu của X (X_1, X_2, X_3, …X_n)***

Ví dụ:

	bom bi là một loại bom
(bom bi, bom)	*bom gồm có bom bi (bom chùm, bom từ trường…)*

	**bom là một loại bom bi*
	bom bi có từ bao là *bom* *bom* có (một) từ thuộc là *bom bi*

Đơn giản hơn, có thể diễn dịch ví dụ này như sau:

Bom bi là từ thuộc của bom vì bom bi là một loại bom

Bom là một từ bao của bom bi vì bom gồm có bom bi (bom chùm, bom từ trường…)

Loài, loại, kiểu,… trong các công thức nhận diện trên được xem như là những từ chứng.

Trong thực tế, đi vào từng khu vực từ vựng, từng chủ đề giao tiếp, chúng ta còn có thể có hàng loạt các từ chứng đặc trưng khác nữa. Chẳng hạn, trong tiếng Việt, đối lập một cách tự nhiên giữa tính chất hoang dã và tính chất đã được thuần hoá, tất cả các loại cây cối con vật đã được con người thuần hoá, quan trọng đối với đời sống của con người đều được người Việt gọi một cách tự nhiên là *cây trồng, vật nuôi/ con vật*. Đây vừa là những từ chứng để có thể đi vào các khung cú pháp nhận diện, vừa là những từ bao tự nhiên ghi nhận giá trị toàn hình cấu trúc của từ ngữ (trong phân loại dân gian theo chức năng), bên cạnh toàn hình cấu trúc (trong phân loại sinh học).

Bằng chứng hùng hồn nhất cho việc này chính là trong các từ điển ngữ văn tiếng Việt, nhất là trong *Từ điển tiếng Việt* của Viện Ngôn ngữ học (do Hoàng Phê cb.), người ta đã đối lập từ bao *cây trồng* với *cây dại* hay *cây mọc hoang*, *vật nuôi/ con vật* với *động vật hoang dã*. Trong tiếng Việt, các từ như *đồ, đồ vật, đồ dùng, vật dụng,…* cũng là người từ có chức năng tương tự như *cây trồng, vật nuôi.* Cùng là cây có *hạt* và có thể lấy *hạt* sử dụng làm lương thực nhưng *kê, láng, mì, bo bo* thì được quy loại là *cây trồng,* nhưng *lồng vực* thì được quy loại là *cỏ* hay *cây dại/ cây mọc hoang.* Hạt *mần ri* không khác gì hạt *cải,* nhưng *cải* là *cây trồng*, *mần ri* là *cây dại*. Chính vì thế mà *gà, lợn, chó, mèo,…* thì có thể nói và giải thích là *vật nuôi*, nhưng *hổ, báo, cáo, cú,…* thì không thể nói và giải thích là *vật nuôi* được. *Mèo nhà* và *hổ, báo, mèo rừng* đều là những từ cùng thuộc, thuộc họ *mèo* của từ bao *mèo*. Quan hệ giữa *mèo nhà*, *hổ, báo, mèo rừng* (từ cùng thuộc) và *mèo* (từ bao) trong trường hợp này là quan hệ phân loại học, quan hệ mang tính cấu trúc. Nhưng quan hệ giữa *mèo nhà*, *hổ, báo, mèo rừng* (từ cùng thuộc) và *vật nuôi* hay *thú cưng* (từ bao) là quan hệ không có tính phân loại, mang tính chức năng.

Công thức trên cũng được áp dụng cho các từ bao và từ thuộc trong quan hệ cách của động từ. Ví dụ:

(cười khẩy, cười)	*cười khẩy là một kiểu cười* *cười gồm có cười khẩy* **cười là cười khẩy*
	cười khẩy có từ bao là *cười* *cười khẩy* có (một) từ thuộc là *cười*

Các ngữ giải thích:

Nếu X và Y có quan hệ bao thuộc thì trong ngôn ngữ tự nhiên chúng có thể xuất hiện trong công thức sử dụng như:

– ***X là Y + P***

Nếu A là X thì A là Y

****Nếu A là Y thì A là X***

P: là một ngữ giải thích

Ví dụ: *(hình bình hành, hình tứ giác)*	*hình bình hành là hình tứ giác có hai cạnh đối song song và bằng nhau* *hình bình hành là hình tứ giác* **hình tứ giác là hình bình hành*
	hình bình hành có từ bao là *hình tứ giác* *hình tứ giác* có (một) từ thuộc là *hình bình hành*

Nếu xem cách là một tiểu loại của quan hệ bao thuộc thì chúng ta có thể có công thức:

X là Y + T

****Y là X + T***

T: là một tính từ

Công thức này dành cho những trường hợp kiểu như *đi* và *đi cà nhắc, nói* và *nói liến thoắng, cười* và *cười hềnh hệch, khóc* và *khóc tỉ ti, chết* và *tự tử, chết* và *chết đuối, tự tử* và *treo cổ, táng* và *đánh, tẩn* và *đánh, đả* và *đánh, nện* và *đánh,…* Tuy nhiên, cần lưu ý là những trường hợp như thế này mặc dù chúng ta vẫn sử dụng thuật ngữ từ bao và từ thuộc để chỉ chúng, nhưng về bản chất, chúng là quan hệ cách, được khu biệt với quan hệ bao thuộc. Riêng cách trường hợp như *táng* và *đánh, tẩn* và *đánh, đả* và *đánh, nện* và *đánh*…. thì *táng, tẩn, đả, nện* cũng là những từ tự thuộc (tức vừa là từ bao, vừa là từ cách, trong động từ).

Chính vì các từ ngữ các quan hệ bao thuộc với nhau như vậy cho nên trong sử dụng ngôn ngữ hằng ngày chúng ta mới bắt gặp các cách nói như:

Điện thoại thông minh là điện thoại….

Gà là loại gia cầm….

Trâu, bò là loại gia súc….

Bút là đồ dùng học tập….

Gà lợn là động vật nuôi trong nhà…

Chính vì thế mà Lyons (1977) đã lập luận rằng, quan hệ bao thuộc có thể được định nghĩa bằng một ngữ đoạn trong đó có một từ cấp trên và một định ngữ. Vì vậy, *bơi là di chuyển trong nước, (gà) trống là con gà đực, nghé là con trâu nhỏ, bê là con bò nhỏ.*

> **Quan hệ bao thuộc là quan hệ hệ hình của nghĩa dựa trên sự bao chứa trong từ thuộc của một định ngữ trên trục ngữ đoạn với nghĩa của một từ cấp trên**. (Lyons, 1977, tr. 294)

Lyons còn cho rằng người ta có thể hỏi một câu như *Đây là con gà gì?* Và câu trả lời có thể là *Đây là con gà mái/ gà trống/ gà ri/ gà ta/ gà tây/ gà gô/ gà tre/ gà đen/ gà chọi,...* Cách nói này gợi ý rằng trong nhiều trường hợp một từ thuộc có thể bao chứa cả nghĩa của một định ngữ (đối với danh từ) và kết hợp nó với một nghĩa của một đơn vị từ vựng cấp trên, tức từ bao. "(N)ghĩa của một từ bao có thể xem như là sản phẩm của nghĩa của từ bao và một định ngữ nào đó " (Lyons, 1977, tr. 293). Điều này có nghĩa là từ thuộc tương đương với một từ bao + một định ngữ bổ nghĩa cho từ bao đó:

Nghé = trâu + con/ non

Bê = bò + con/ non

Bàn học = bàn + (dùng để) học

Áo ba lỗ = áo + (có) ba lỗ

Định nghĩa này của Lyons đã đẩy quan hệ hệ hình sang phạm vi của quan hệ ngữ đoạn (Cann, 2011). Và đây có thể xem là một manh nha của cách tiếp cận ngữ cảnh luận đối với quan hệ bao thuộc. Sau này, Cruse (1986) đã có nhiều kế thừa từ gợi ý này của Lyons. Và có điều thú vị là, ngôn ngữ học tính toán gần đây khi phát hiện tự động các đơn vị từ vựng có quan hệ bao thuộc đã tận dụng triệt để các manh mối ngữ cảnh này (đặc biệt là các các kết cấu đồng hiện).

Các dấu hiệu cú pháp khác:

Trong ngôn ngữ tự nhiên có nhiều cách nói mà ở đó có chứa các dấu hiệu báo hiệu các từ ngữ có quan hệ bao thuộc. Dưới đây là một vài ví dụ như vậy:

_ và những _ khác *(chó và những loài động vật khác, lan phi điệp và những loài lan khác).* Các từ ngữ xuất hiện trong cách nói này luôn là các từ ngữ có quan hệ bao thuộc.

chẳng hạn như _ *(Ở đó có nhiều loài hoa quả nhiệt đới, chẳng hạn như chuối, cam, bưởi).* Các từ xuất hiện trường hợp này thường là những từ cùng thuộc.

_ cũng giống như _ *(trâu thì cũng như bò).* Các từ xuất hiện trường hợp này thường là những từ cùng thuộc.

_ và/ hay/ hoặc/ lẫn _ *(hoa hồng và hoa huệ, cả giáo viên và học sinh).* Trong văn bản, dấu phẩy cũng là một dấu hiệu dùng để nhận diện các từ cùng thuộc; chức năng của chúng không khác gì các liên từ như *và, hoặc, hay…* Các từ xuất hiện trường hợp này thường là những từ cùng thuộc.

_ là một trường hợp điển hình/ đặc biệt *(Bình Dương là một trường hợp điển hình trong công tác xoá đói giảm nghèo/ Bình Dương là một tỉnh điển hình trong công tác xoá đói giảm nghèo)*. *Bình Dương* là từ thuộc trường hợp, *tỉnh* là từ bao lớp. Cách nói này thường áp dụng cho các từ thuộc trường hợp.

_ là một ví dụ/ thí dụ của/ về/ cho _ *(gà là một ví dụ về gia cầm).* Các từ ngữ xuất hiện trong cách nói này luôn có quan hệ bao thuộc.

trong số/ giữa _ thì _ *(trong số các học sinh của trường thì A…).* Các từ ngữ xuất hiện trong cách nói này luôn có quan hệ bao thuộc; từ thuộc thường là từ thuộc trường hợp.

_ so sánh với _ *(khi so sánh công nhân với trí thức).* Các từ xuất hiện trường hợp này thường là những từ cùng thuộc.

ngoại trừ _, đặc biệt là _. Các từ xuất hiện các cách nói này thường là những từ cùng thuộc.

bao gồm_, gồm có_ *(những lựa chọn nghề nghiệp tương đối an toàn gồm giáo viên, luật sư, nhà khoa học, lập trình viên,…; quy định khu chức năng bao gồm: khu kinh tế, khu công nghiệp, khu chế xuất, khu công nghệ cao, khu du lịch, khu nghiên cứu, đào tạo, khu thể dục thể thao,…).* Những từ cũng xuất hiện trong vị trí này luôn là những từ cùng thuộc, cùng thuộc trong ngữ cảnh đó, bất luận là trong các liên tưởng thực tế các các đơn vị từ vựng đó có thể tạo thành một vào cặp bao thuộc khác cấp logic.

Nếu _ thì _, đừng _ mà hãy _ (*nếu đã thi công an thì thôi bộ đội, đừng học kiến trúc mà hãy học âm nhạc).* Những từ xuất hiện trong các ô này là những từ cùng thuộc.

_ như _ *(phát triển các dịch vụ như cầu đường, giao thông công cộng, công viên, giáo dục, trông trẻ, y tế, dưỡng lão chăm sóc sức khoẻ; các món ăn phổ biến như phở, bún, bánh cuốn, cháo, xôi,…)*. Những từ xuất hiện trong các vị trí này luôn tạo thành một cặp từ bao và các từ cùng thuộc. *V.v.*

Lưu ý rằng, nếu đối lập giữa ngôn ngữ và lời nói, xã hội và cá nhân, từ vựng tinh thần và từ vựng trong ngữ cảnh,… thì quan hệ bao thuộc có ở các đơn vị từ vựng xuất hiện trong các cách nói năng như đã dẫn ra trên đây thường là những quan hệ mang tính ngôn ngữ, xã hội, từ vựng tinh thần hơn là lời nói, cá nhân, từ vựng trong ngữ cảnh, một cách lần lượt. Tuy vậy, chúng ta không nên quên rằng ngôn ngữ tồn tại qua lời nói, qua các cách sử dụng thực tế, các quan hệ ổn định trong từ vựng tinh thần của một ngôn ngữ được bắt đầu từ những cách sử dụng ngôn ngữ cá nhân đơn lẻ. Trong thực tế, những biểu thức như trên rất có giá trị trong việc giúp chúng ta hình dung ra được bộ mặt thực tế của quan hệ bao thuộc trong sử dụng ngôn ngữ hằng ngày, vai trò của chúng trong việc tổ chức, biểu diễn tri thức, văn hoá, xã hội, tư duy, tâm lí và cấu trúc quyền lực xã hội,…

Trong nói năng hằng ngày, chúng ta có thể không để ý đến những dấu hiệu này, nhưng trong xử lí ngôn ngữ tự nhiên, các cách sử dụng ngôn ngữ như thế này là những manh mối rất quan trọng để huấn luyện máy phát hiệu và sàng lọc tự động quan hệ bao thuộc…

Đối lập toàn hình cấu trúc với toàn hình chức năng, quan hệ bao thuộc có các từ bao thuộc cấu trúc và các từ bao thuộc chức năng. Tất cả các dấu hiệu nhận diện cú pháp ở trên là các dấu hiệu được áp dụng một cách triệt để cho việc nhận diện các từ ngữ bao thuộc cấu trúc. Có một vài dấu hiệu cú pháp ở trên có thể được áp dụng để nhận diện các từ ngữ bao thuộc chức năng. Và mặt khác, các dấu hiệu trên cũng là các dấu hiệu có nhiêu giá trị trong việc nhận diện quan hệ bao thuộc lớp. Việc nhận diện quan hệ bao thuộc trường hợp trong sử dụng ngôn ngữ tự nhiên không khó. Chẳng hạn, người ta hoàn toàn có thể sử dụng các dấu hiệu chính tả, chẳng hạn như viết hoa tên riêng, để nhận diện các từ thuộc trường hợp. Tuy thế, cũng không nên chủ quan để cho rằng chỉ cần căn cứ vào những dấu hiệu chinh như vậy là có thể khái quát nên được các mối quan hệ bao thuộc trường hợp ổn định tĩnh tại, có tính chất đại diện trong một ngôn ngữ. Trong xử lí ngôn ngữ tự nhiên, bài toán nhận diện các thực thể trong đó có cả bài toán nhận diện các tên riêng (NER) là một bài toán lớn, phức tạp.

Để nhận diện quan hệ bao thuộc chức năng, chúng ta có thể sử dụng cách nói đặc trưng như:

– ***X (X_1, X_2, X_3, …X_n) có thể được xem/ coi như/ xếp vào một loài/ loại/ kiểu của Y***

Y thì có/ gồm có/ bao gồm X (X_1, X_2, X_3, …X_n)

****Y là một loài/ loại/kiểu của X (X_1, X_2, X_3, …X_n)***

Ví dụ:

(chó)	*chó là một loại thú cưng* *thú cưng thì có chó* **thú cưng là một loại chó*
	chó có từ bao là *thú cưng* *thú cưng* có (một) từ thuộc là *chó*
(chó)	*chó có thể được xem/ coi như/ xếp vào một loài/ loại/ kiểu thú cưng* *thú cưng thì có chó* **thú cưng là một loại chó*
	chó có từ bao là *thú cưng* *thú cưng* có (một) từ thuộc là *chó*

Quan hệ bao thuộc chức năng là quan hệ không có được sự ổn định như quan hệ bao thuộc cấu trúc. Vì thế, trong công thức nhận diện trên, từ *có thể* là từ thể hiện không ổn định đó. Trong ngôn ngữ tự nhiên, chúng ta thường bắt gặp công thức nhận diện này. Trong ngôn ngữ tự nhiên, chúng ta cũng có nhiều các từ chứng giống như từ *có thể*. *Xét theo khía cạnh này, từ góc*

*độ này, theo tiêu chí này, bây giờ, ở đây, lúc này, trong chừng mực này, trong phạm vi này, đối với, dành cho, ...*là những biểu thức ngôn ngữ báo hiệu các góc nhìn mà ở đó có thể quyết định sự phân loại, phân loại của những thứ đang được đưa ra xem xét.

Chẳng hạn:

Xe hạng sang là đồ chơi của nhà giàu. So với *Xe hạng sang là một loại xe.*

Ô tô là một loại tiêu sản. So với *Ô tô là một loại phương tiện giao thông.*

Thỏ là thú cưng đối với trẻ con ngày này. So với *Thỏ là một loài động vật/ loài gậm nhấm.*

Đỉa là thú cưng dành cho những kẻ dị hợm. So với *Đỉa là một loài giun đốt/ động vật.*

Khoai lang đối với ngày xưa chỉ là thức ăn của lợn. Khoai lang ngày nay là thức ăn của nhà giàu. So với *Khoai lang là một loại khoai.*

Vàng bạc có thể được xem như là một loại gia vị trong đồ ăn thức uống của giới thượng lưu. So với *Vàng bạc là đồ trang sức/ là một loại kim loại.*

Đất, sỏi, đá, khúc tre, nõn tre, lá dứa, lá đa, lá mít, dâm bụt, ống lúa, ống niễng, dây mơ, chuồn chuồn,... đối với trẻ em ngày xưa thì đều là *đồ chơi* chính hiệu, nhưng đối với trẻ con ngày nay thì đều không phải là *đồ chơi,* là *thứ bỏ đi, bẩn thỉu….*

Hột thị, đòng đòng, hoa dâm bụt,... đối với trẻ em ngày xưa thì đều là *đồ ăn thức uống,* nhưng đối với trẻ em ngày nay thì đều là những *thứ bỏ đi, không ăn được.*

Có thể tạm phân loại các biểu thức ngôn ngữ biểu thị quan hệ bao thuộc chức năng như sau:

- Các biểu thức ngôn ngữ dành cho chủ thể như *dành cho, đối với, là của,...*
- Các biểu thức ngôn ngữ nói về phương diện, cách thức, tiêu chí, góc nhìn,… như: *xét về mặt này, trên phương diện này, mặt này, mặt khác, đứng từ góc nhìn, dựa trên tiêu chí,...*
- Các biểu thức ngôn ngữ nói về không gian, thời gian: *ở đây, ở đó, bây giờ, lúc đó, hiện nay, ngày xưa,...*
- Các biểu thức ngôn ngữ nói về công dụng, chức năng, nhiệm vụ,… như *có thể được sử dụng như, có thể được dùng thay cho,....* Chẳng hạn, Cái *nơm*, cái *đó*,… vốn là *ngư cụ, nông cụ* (bao thuộc cấu trúc), nhưng ngày nay thường được sử dụng như là *vật trang trí* hoặc *bộ phận một chiếc đèn* (sử dụng như *cái chụp*) (bao thuộc chức năng). Cái *nong, nia, sàng, giần, thúng, mủng, mẹt, giá vo gạo, giành, sọt, quang gánh,...* hiện giờ vẫn có thể đặc điểm xem là *nông cụ* với *nhà nông,* nhưng với người ở phố, đã thoát li *nông nghiệp* thì những *vật dụng* đó đã trở thành *vật trang trí, vật có chức năng bảo tồn giá trị văn hoá của dân tộc*. Chính cái *có thể* này là cái rất linh hoạt. Các đồ vật nhân tạo thường được sử dụng trong các kết cấu diễn đạt như vậy. Đây cũng chính là lí thuyết giải thích tại sao các đồ vật tạo tác có xu hướng gắn với

toàn hình chức năng (bao thuộc chức năng), còn các vật tự nhiên (cây, con, hiện tượng tự nhiên) lại có xu hướng được phân loại gắn với các toàn hình cấu trúc (bao thuộc cấu trúc).

3.3. Các dấu hiệu từ pháp

Trong số khoảng 40.000 đơn vị từ vựng đã được các từ điển tiếng Việt thu thập, có khoảng khoảng 22.000 tổ hợp từ song tiết mà trong đó cả hai yếu tố cấu thành nên chúng đều có nghĩa (trong tổng số khoảng 30.000 tổ hợp song tiết). Trong các tổ hợp từ song tiết mà cả hai yếu tố đều có nghĩa có thể có nhiều loại quan hệ ngữ nghĩa khác nhau: bao thuộc, tổng phân, cách, nhân quả, thuộc tính, vai, đồng nghĩa, trái nghĩa… Những quan hệ này là những quan hệ cơ bản, quan trọng trong tổ chức từ vựng của ngôn ngữ nói chung và trong tổ chức của các tổ hợp từ song tiết nói riêng.

Xem xét sự tồn tại của các quan hệ ngữ nghĩa trong các tổ hợp song tiết mà cả hai yếu tố cùng có nghĩa, chúng ta thấy có hai khả năng:

– Giữa hai thành tố cấu tạo nên tổ hợp từ song tiết có thể có quan hệ ngữ nghĩa nào?

– Giữa từng thành tố cấu tạo với toàn bộ tổ hợp song tiết có thể có quan hệ ngữ nghĩa nào?

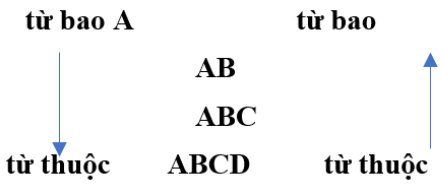
Trong tiếng Việt, quan hệ bao thuộc là quan hệ có mặt trong các tổ hợp song tiết mà cả hai yếu tố cùng có nghĩa và đều có khả năng hoạt động độc lập. Do đó, chúng ta hoàn toàn có thể nói cấu tạo từ tiếng Việt luôn tiềm tại quan hệ bao thuộc. Hay nói cách khác, quan hệ bao thuộc là một quan hệ có thể được sử dụng để cấu tạo từ tiếng Việt. Như vậy, rõ ràng là trong từ pháp tiếng Việt đã có sẵn những dấu hiệu nhận diện quan hệ bao thuộc.

Với quan hệ bao thuộc trong hệ thống danh từ, đối với các tổ hợp có cấu tạo ngữ pháp xuôi, ta có các khả năng ngữ nghĩa quan trọng là:

A ⊃ B và **A/B ⊃ AB.**

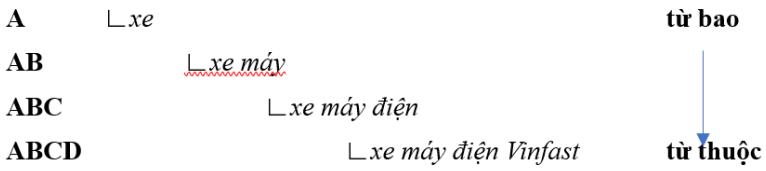
Ví dụ, xét từ các tổ hợp *áo dài, chó sói, cá rô, hoa hồng, thợ tiện, xe thồ,…,* ta thấy: Với khả năng thứ nhất, *áo dài, chó sói, thợ tiện, xe thồ,…,* một cách lần lượt, là một loại *áo, chó, cá, hoa, thợ, xe,…* Như vậy, *áo dài, chó sói, thợ tiện, xe thồ,…,* là những từ thuộc của các từ bao *áo, chó, cá, hoa, thợ, xe,…* Với khả năng thứ hai, *sói, rô, hồng,…,* một cách lần lượt, là một loại *chó, cá, hoa,…* Như vậy, *sói, rô, hồng* là những từ thuộc của các từ bao *chó, cá, hoa.* Như vậy, nếu mở rộng vấn đề, chúng ta thấy các trong các cụm từ định danh (danh từ) có cấu tạo chính phụ nói chung, bất luận là thành tố phụ là mang giá trị hạn định hay miêu tả, đều có thể có chứa quan hệ bao thuộc mà trong đó yếu tố chính đứng đầu có tư cách là một từ bao và toàn bộ cụm từ định danh đó có tư cách là từ thuộc. Có thể khái quát như sau:

A ⊃ AB ⊃ ABC ⊃ ABCD.



Hình 4. Cơ chế ghép chính phụ và từ bao thuộc

Ví dụ:



Hình 5.1 Cơ chế ghép chính phụ biệt loại và từ bao thuộc

3.4. Các dấu hiệu ngữ nghĩa

Có một sự tương ứng giữa ngôn ngữ học và logic: từ và khái niệm, chỉ vật và ngoại diên, nội hàm và nghĩa. Ở trên đã nói, đứng từ góc độ nội hàm, nội hàm của từ bao luôn là một bộ phận của từ thuộc. Tương tự như vậy, đứng từ góc độ ngữ nghĩa, nghĩa của từ bao luôn là một bộ phận của từ thuộc. Do đó, về mặt ngữ nghĩa, quan hệ bao thuộc có thể được khái quát theo một công thức:

> ***Nếu X là một từ thuộc của Y thì nội dung ngữ nghĩa của Y phải là một phần nội dung ngữ nghĩa không thể thiếu của X.***

Điều này có nghĩa là nội hàm hay nghĩa của từ bao luôn nằm trong nghĩa hay nội hàm của từ thuộc; nội hàm hay nghĩa của từ thuộc luôn rộng hơn nghĩa hay nội hàm của từ bao. Ở đây, không nên nhầm lẫn giữa sự thực này với những cách nói năng hằng ngày có sử dụng từ "nghĩa rộng, nghĩa hẹp" (như cách nói *xét theo nghĩa rộng).* Rộng hẹp trong những cách nói như vậy là rộng hẹp về phạm vi quy chiếu, phạm vi ứng dụng của nghĩa trong hiện thực chứ không phải là sự rộng hẹp của bản thân nghĩa.

Ví dụ, ở cặp bao thuộc *sinh vật – động vật* và *thực vật,* nghĩa "sinh vật [thực thể có sự sống]/ thực thể có cảm giác và có thể tự di chuyển được" của *động vật,* "sinh vật [thực thể có sự sống]/ thực thể không có khả năng cảm giác và không thể di chuyển được" của *thực vật* luôn bao hàm nghĩa "thực thể có sự sống" của *sinh vật.*

Trong khi đó, ở cặp bao thuộc *động vật – gia súc* và *gia cầm* thì nghĩa "thực thể có cảm giác và có thể tự di chuyển được" lại nằm gọn hay bị bao hàm trong nghĩa "động vật [thực thể có cảm giác và có thể tự di chuyển được] bốn chân được nuôi trong nhà" của *gia súc,* trong nghĩa "động vật [thực thể có cảm giác và có thể tự di chuyển được] hai chân được nuôi trong nhà" của *gia cầm.*

Nói cách khác, đứng từ bình diện thao tác trong việc giải thích nghĩa hay định nghĩa nội hàm, người ta có thể sử dụng từ bao để giải thích hay định nghĩa từ thuộc. Từ bao là từ có thể có mặt trong lời giải thích hay định nghĩa nghĩa của từ thuộc. Chính vì vậy mà trong từ điển học thực hành, loại định nghĩa dùng từ bao là loại định nghĩa thường được sử dụng phổ biến nhất. Loại định nghĩa này không chỉ được sử dụng để định nghĩa các từ bao thuộc chân chính trong danh từ mà còn được sử dụng phổ biến để định nghĩa các từ cách trong động từ (*di chuyển – chạy, chạy – chạy marathon, bơi – bơi ếch,…*) và các từ giả bao thuộc trong tính từ (*màu – xanh, xanh – xanh lè, đỏ – đỏ ửng,…*). Giở bất kì từ điển ngữ văn một thứ tiếng nào ra, ta cũng đều gặp lối định nghĩa sử dụng từ bao. Điều này đặc biệt đúng trong các từ điển giải thích thuật ngữ khoa học.

Nếu trong giải thích ngôn ngữ tự nhiên người ta thường chọn một từ bao dễ hiểu nhất theo khía cạnh đang xét (Ví dụ: *gậy: công cụ; gậy: đồ chơi; gậy: dụng cụ thể thao; gậy: hung khí; gậy: vật chống đỡ;…*) thì trong định nghĩa khái niệm của khoa học người ta luôn dùng một từ bao cấp trên liền kề.

Hình 6. Bài tập dùng từ bao để giải thích từ thuộc (Phạm Văn Lam, 2015)

Việc sử dụng từ bao để để giải thích từ thuộc không chỉ được sử dụng trong từ điển, mà còn thường xuyên được sử dụng trong hoạt động ngôn ngữ hằng ngày.

Khi gặp một từ nào đó mới nghe hay mới tiếp xúc lần đầu, chưa hiểu được nó thì người ta sẽ đặt câu hỏi *là cái gì, là gì, có nghĩa là gì, là như thế nào,…* và câu trả lời một cách tự nhiên thường nhận được là *là loại/ là một loại/ là một kiểu/ là cách/ là con vật/ là đồ dùng, là thứ, là cái, là vật, là đồ chơi, là dụng cụ, là thức ăn, là đồ uống,…* Từ đứng liền sau trong các từ *loại/ loài/ cách/ kiểu/ con vật/ đồ dùng….* này chính là từ bao của từ thuộc cần được giải thích. Ví dụ:

- *Tàu ngầm là loại tàu…*
- *Thuốc lá điện tử là loại thuốc…*
- *Truyện trinh thám là loại truyện…*
- *Hiến pháp là một loại (văn bản) luật…*
- *Cây mắc ca là loại cây…*
- *Kì lân là con vật…*

Như vậy, trong việc nhận diện quan hệ bao thuộc, các dấu hiệu ngữ nghĩa được thể hiện ở trong các lời giải thích hay định nghĩa nghĩa của các từ điển, trong cách nói năng có tính giải thích hằng ngày. Từ được giải thích hay định nghĩa là từ thuộc trong mối quan hệ với từ bao vốn xuất hiện và đứng ở vị trí đầu tiên trong lời giải thích về từ thuộc.

4. KẾT LUẬN

Quan hệ bao thuộc là một trong những quan hệ xương sống, quan trọng nhất trong tổ chức từ vựng của bất kì ngôn ngữ nào. Quan hệ bao thuộc là quan hệ giữa một từ có tính khái quát, được gọi là từ bao, và một từ có tính cụ thể, được gọi là từ thuộc. Từ thuộc là từ có tính cụ thể hơn từ bao. Đó là quan hệ giữa một từ bao với các từ thuộc; từ thuộc được hình dung như là một loại của từ bao. Quan hệ bao thuộc luôn được xem là một trong những vấn đề cơ bản, cốt lõi trong ngữ nghĩa học nói chung và ngữ nghĩa học từ vựng, đặc biệt là ngữ nghĩa học quan hệ nói riêng.

Chúng ta có thể nhận diện quan hệ bao thuộc trong tiếng Việt bằng một số dấu hiệu có tính hình thức. Các dấu hiệu nhận diện có tính hình thức ở đây đều là những đặc điểm hay yếu tố ngôn ngữ tồn tại trong ngôn ngữ tự nhiên hoặc được dùng để giải thích ngôn ngữ tự nhiên (trong lời nói hay trong các từ điển ngữ văn). Các dấu hiện hình thức được sử dụng để nhận diện quan hệ bao thuộc được gom thành các nhóm: các dấu hiệu logic, các dấu hiệu cú pháp, các dấu hiệu từ pháp và các dấu hiệu ngữ nghĩa. Trong số các dấu hiệu này, kết cấu từ vựng – ngữ pháp *A là một loại/ loài của B* được xem là dấu hiệu hình thức cơ bản, trung tâm và quan trọng nhất. Tất cả các dấu hiệu nhận diện khác có thể được xem là những biểu hiện cụ thể của kết cấu từ vựng – ngữ pháp này.

Các dấu hiệu hình thức được trình bày ở đây đều có thể được áp dụng một cách thiết thực và hữu dụng vào việc phát hiện tự động, đo lường độ tương tự nghĩa,… trong xử lí ngôn ngữ tự nhiên, trong định nghĩa nghĩa từ ở từ điển, giải thích nghĩa từ cho học sinh, dạy học cách tổ chức và triển khai văn bản/ diễn ngôn cho học sinh,…

TÀI LIỆU THAM KHẢO

1. Cao Xuân Hạo, Hoàng Dũng (2005), *Từ điển thuật ngữ ngôn ngữ học đối chiếu Anh – Việt, Việt Anh,* Nxb Khoa học Xã hội, Hà Nội.
2. Đỗ Hữu Châu (1986), *Cơ sở ngữ nghĩa học từ vựng,* NXB Giáo dục, Hà Nội.
3. Đỗ Hữu Châu (2006), *Giáo trình từ vựng học tiếng Việt,* NXB Đại học Sư phạm, Hà Nội.
4. Cruse, D. A. (1986), *Lexical Semantics*, Cambridge: Cambridge University Press.
5. Crystal D. (1985), *A Dictionary of Linguistics and Phonetics,* Oxford: Blackwell.
6. Hoàng Dũng, Bùi Mạnh Hùng (2007), *Giáo trình dẫn luận ngôn ngữ học,* Nxb ĐHSP, Hà Nội.
7. Nguyễn Thiện Giáp (2014), *Nghĩa học Việt ngữ,* NXB Giáo dục, Hà Nội.
8. Phạm Văn Lam, Nguyễn Phương Thái (2015), "Phác thảo *Mạng từ tiếng Việt", Đỗ Hữu Châu: hành trình và tiếp nối,* NXB Đại học Quốc gia Hà Nội, Hà Nội, tr. 418 – 438.
9. Phạm Văn Lam (2020), *Từ trái nghĩa tiếng Việt,* NXB Giáo dục Hà Nội Hà Nội.
10. Phạm Văn Lam (coauthor) (2019), "Hypernymy Detection for Vietnamese Using Dynamic Weighting Neural Network", *Proceedings of The 19th International Conference on Computational Linguistics and Intelligent Text Processing*, in lại trong *Lecture Notes in Computer Science*, Springer.
11. Phạm Văn Lam (2020), *Nghiên cứu quan hệ bao thuộc trong tiếng Việt,* Viện Ngôn ngữ học.
12. Phạm Văn Lam (2025), *Tiếng Việt: Cơ sở ngữ nghĩa học quan hệ* (đang in).
13. Lyons, J. (1977), *Semantics,* Cambridge: Cambridge University Press.
14. Vũ Đức Nghiệu (cb.), Nguyễn Văn Hiệp (2009), *Dẫn luận ngôn ngữ học,* NXB Đại học Quốc gia Hà Nội, Hà Nội.

KỶ YẾU HỘI THẢO KHOA HỌC QUỐC GIA LẦN 3

NGÔN NGỮ HỌC TÍNH TOÁN

CÔNG NGHỆ VÌ CON NGƯỜI

NHÀ XUẤT BẢN TRI THỨC
Địa chỉ: Lô D20, Ngõ 19 Duy Tân, Phường Cầu Giấy, TP.Hà Nội
E-mail: lienhe@nxbtrithuc.com.vn
Điện thoại: (024) 6687 8415

ISBN 978-632-628-163-7
978-632-628-163-7