

LỖI NGỮ NGHĨA VÀ KẾT HỢP TỪ TRONG SỬ DỤNG TỪ VỰNG TIẾNG TRUNG CỦA SINH VIÊN HỌC TIẾNG TRUNG NHƯ NGOẠI NGỮ HAI: MỘT NGHIÊN CỨU HỖN HỢP

BÙI THU PHƯƠNG
NGUYỄN THỊ THU

Khoa Ngoại ngữ, Đại học Bách khoa Hà Nội

Nhận bài ngày 05/7/2026. Sửa chữa xong 10/8/2026. Duyệt đăng 15/8/2026

Abstract

This study investigates semantic and collocational errors in Chinese vocabulary use among English-major students learning Chinese as a second foreign language. Using a quantitatively dominant mixed-methods design, data were collected from 194 students through a questionnaire, a vocabulary task, and interviews. The findings show that semantic errors are more prominent than collocational errors, with near-synonym/synonym confusion and verb-object combinations emerging as the most notable difficulties. Triangulated data suggest that the errors are associated with Vietnamese influence, incomplete knowledge of Chinese vocabulary, word-by-word learning outside context, and limited opportunities for practice. The study highlights the need to develop depth of lexical knowledge, particularly the ability to distinguish meanings, understand usage constraints, and master collocational patterns in teaching Chinese as a second foreign language.

Keywords: Error analysis, semantic errors, collocation, Chinese as a second foreign language, cross-linguistic influence, lexical knowledge.

1. Đặt vấn đề

Trong bối cảnh quan hệ kinh tế, giáo dục và giao lưu Việt Nam-Trung Quốc ngày càng mở rộng, tiếng Trung được giảng dạy tại nhiều cơ sở giáo dục đại học như một ngoại ngữ bổ sung. Đối với sinh viên (SV) chuyên ngành Ngôn ngữ Anh, tiếng Trung thường được tiếp nhận sau tiếng Việt và tiếng Anh, tạo nên một môi trường học tập đa ngôn ngữ, trong đó nhiều hệ thống ngôn ngữ có thể đồng thời tác động đến quá trình lựa chọn và sử dụng từ vựng.

Từ vựng là một thành tố cốt lõi của năng lực giao tiếp. Biết một từ không chỉ là nhận diện hình thức và nghĩa cơ bản mà còn bao gồm phạm vi nghĩa, sắc thái, điều kiện sử dụng, đặc điểm ngữ pháp và những từ thường kết hợp với nó (Nation, 2001; Schmitt, 2000; Webb & Nation, 2017). Vì vậy, người học có thể nhận biết đúng nghĩa từ điển nhưng vẫn tạo ra phát ngôn thiếu tự nhiên hoặc sai chuẩn khi lựa chọn sai từ gần nghĩa, mở rộng phạm vi nghĩa quá mức hoặc kết hợp từ theo thói quen của ngôn ngữ nền.

Phân tích lỗi, xem lỗi của người học như bằng chứng về một hệ thống ngôn ngữ đang phát triển, thay vì chỉ coi đó là biểu hiện thất bại (Corder, 1967, 1974; Ellis & Barkhuizen, 2005). Cùng với khái niệm ngôn ngữ trung gian (Selinker, 1972) và chuyển di liên ngôn ngữ (Odlin, 1989), cách tiếp cận này cho phép giải thích lỗi từ cả nguồn liên ngôn ngữ lẫn nội ngôn ngữ. Đối với người Việt học Tiếng Trung, mạng lưới từ Hán-Việt vừa có thể hỗ trợ việc nhận diện nghĩa, vừa có thể dẫn đến đồng nhất hóa quá mức về phạm vi nghĩa, sắc thái và thói quen kết hợp.

Các nghiên cứu về người học Việt Nam đã chỉ ra khó khăn đáng kể ở các cặp từ gần nghĩa, hiện tượng chuyển dịch nghĩa Hán-Việt và từ li hợp (Luu Vũ Hân, 2023; Mai Thị Huệ, 2018; Trần Quang Huy,

Email: phuong.bui@hust.edu.vn, thu.nguyenthi2@hust.edu.vn

2024). Ở bình diện kết hợp từ, nghiên cứu ngữ liệu cũng cho thấy những tổ hợp động từ-tân ngữ và tính từ-danh từ là khu vực dễ phát sinh lỗi và chịu ảnh hưởng của chuyển di ngôn ngữ (Li, 2016). Tuy nhiên, phần lớn các công trình khảo sát từng hiện tượng riêng lẻ hoặc tập trung vào SV chuyên ngành Tiếng Trung còn ít nghiên cứu xem xét đồng thời lỗi ngữ nghĩa và lỗi kết hợp từ trên cùng một mẫu SV Việt Nam học Tiếng Trung như ngoại ngữ hai, đồng thời đối chiếu dữ liệu lỗi thực tế, tự đánh giá và phỏng vấn.

2. Phương pháp nghiên cứu

2.1. Thiết kế và người tham gia

Nghiên cứu sử dụng thiết kế hỗn hợp theo hướng định lượng chiếm ưu thế. Dữ liệu định lượng được dùng để mô tả và so sánh mức độ lỗi, trong khi dữ liệu định tính được sử dụng để đối chiếu và giải thích các xu hướng nổi bật. Mẫu gồm 194 SV chuyên ngành Ngôn ngữ Anh đang học học phần Tiếng Trung Quốc IV như ngoại ngữ hai tại Khoa Ngoại ngữ, Đại học Bách khoa Hà Nội trong năm học 2025-2026, trong đó giới tính nam: $n=31$ (16%), giới tính nữ $n=163$ (84%). Toàn bộ 194 phiếu khảo sát hợp lệ và không có giá trị khuyết thiếu.

2.2. Công cụ thu thập dữ liệu

Nghiên cứu sử dụng 3 công cụ: Bảng hỏi tự đánh giá, bài tập khảo sát từ vựng và phỏng vấn bán cấu trúc. Các công cụ có chức năng bổ trợ cho nhau nhằm đối chiếu nhận thức của người học với biểu hiện lỗi thực tế và thông tin giải thích từ người tham gia phỏng vấn.

Bảng hỏi tự đánh giá: Bảng hỏi sử dụng thang Likert 5 mức và gồm 6 thang đo: Thực trạng học từ vựng (TV1-TV6), Tự đánh giá lỗi ngữ nghĩa (NN1-NN7), Tự đánh giá lỗi kết hợp từ (KH1-KH7), Thực trạng sử dụng từ (SD1-SD6), Nguyên nhân gây lỗi (NGN1-NGN8) và Đánh giá hoạt động dạy học (GD1-GD6). Độ tin cậy nội tại được kiểm tra bằng Cronbach's alpha.

Bảng 1: Độ tin cậy của các thang đo

Thang đo	Số biến	Cronbach's α
Thực trạng học từ vựng	6	0,840
Tự đánh giá lỗi ngữ nghĩa	7	0,831
Tự đánh giá lỗi kết hợp từ	7	0,848
Thực trạng sử dụng từ	6	0,848
Nguyên nhân gây lỗi	8	0,881
Đánh giá hoạt động dạy học	6	0,823

Nguồn: Tổng hợp từ kết quả xử lý dữ liệu khảo sát bằng SPSS, 2026 ($N = 194$).

Bài tập khảo sát từ vựng: Bài tập gồm các nhiệm vụ lựa chọn đáp án và dịch câu Việt-Trung, được thiết kế để thu thập trực tiếp biểu hiện lỗi ngữ nghĩa và lỗi kết hợp từ. Các câu trả lời sai được đối chiếu với cách dùng phù hợp trong tiếng Trung và mã hóa theo 8 tiểu loại lỗi đã xác lập. Dữ liệu từ bài tập là nguồn chính để tính số lỗi trung bình và tỷ lệ SV mắc lỗi.

Phỏng vấn bán cấu trúc: Phỏng vấn được thực hiện với 10 SV (SV1-SV10) và 4 giảng viên Tiếng Trung (GV1-GV4). Nội dung tập trung vào khó khăn trong học và sử dụng từ vựng, các lỗi thường gặp, chiến lược xử lý từ và những yếu tố có khả năng liên quan đến lỗi. Dữ liệu phỏng vấn được dùng để bổ trợ và giải thích, không thay thế cho kết quả định lượng.

2.3. Mã hóa và phân tích dữ liệu

Dữ liệu được mã hóa và xử lý bằng SPSS. Thống kê mô tả gồm tần số, tỷ lệ, trung bình (M) và độ lệch chuẩn (SD) được dùng để mô tả mẫu, mức độ khó khăn và phân bố lỗi. Với dữ liệu bài tập, mỗi câu trả lời sai được đối chiếu với đáp án/cách dùng phù hợp và gán vào một trong 8 tiểu loại lỗi. Hai chỉ số được sử dụng song song là số lỗi trung bình trên mỗi SV và tỷ lệ SV mắc ít nhất một lỗi ở từng tiểu loại. Do tổng lỗi ngữ nghĩa và tổng lỗi kết hợp từ không đáp ứng giả định phù hợp cho so sánh tham

số ghép cặp, nghiên cứu sử dụng kiểm định Wilcoxon Signed-Rank để so sánh hai tổng điểm lỗi ở mức ý nghĩa $\alpha = 0,05$. Dữ liệu phỏng vấn được đọc, mã hóa theo chủ đề và đối chiếu với các nhóm nguyên nhân đã xác lập trong khung lý thuyết. Việc kết hợp ba nguồn dữ liệu được sử dụng như một hình thức đối chiếu phương pháp nhằm phân biệt giữa lỗi quan sát được, nhận thức của người học và diễn giải của người tham gia.

3. Kết quả nghiên cứu

3.1. Cơ sở lý thuyết và tổng quan nghiên cứu

3.1.1. Phân tích lỗi, ngôn ngữ trung gian và chuyển di liên ngôn ngữ

Phân tích lỗi coi lỗi của người học là dữ liệu phản ánh quá trình phát triển ngôn ngữ. Corder (1967, 1974) và James (1998) hệ thống hóa quy trình từ thu thập ngữ liệu, nhận diện, mô tả/phân loại đến giải thích và đánh giá lỗi; Ellis và Barkhuizen (2005) tiếp tục nhấn mạnh giá trị của dữ liệu người học trong phân tích ngôn ngữ. Nghiên cứu này vận dụng cách tiếp cận đó để xác định lỗi lựa chọn và kết hợp từ, sau đó đối chiếu với bảng hỏi và phỏng vấn. Theo Selinker (1972), ngôn ngữ trung gian là một hệ thống đang được tái cấu trúc; lỗi vì vậy có thể bắt nguồn từ chuyển di, khái quát hóa quá mức, chiến lược học tập hoặc tri thức chưa hoàn thiện về ngôn ngữ đích. Odlin (1989) cho rằng chuyển di có thể hỗ trợ hoặc gây nhiễu tùy mức độ tương đồng thực sự giữa các ngôn ngữ. Ở người học nhiều ngoại ngữ, nguồn chuyển di không nhất thiết chỉ đến từ tiếng mẹ đẻ; khoảng cách ngôn ngữ và vị thế của ngoại ngữ đã biết cũng có thể tác động đến ngôn ngữ đang học (Cenoz, 2001). Vì vậy, ảnh hưởng của tiếng Việt và tiếng Anh trong nghiên cứu này chỉ được diễn giải khi có dữ liệu hỗ trợ.

3.1.2. Chiều sâu tri thức từ vựng và hai nhóm lỗi khảo sát

Tri thức từ vựng gồm cả chiều rộng và chiều sâu; chiều sâu bao hàm quan hệ nghĩa, hạn chế sử dụng, chức năng ngữ pháp và các từ thường đồng xuất hiện (Nation, 2001; Schmitt, 2000; Webb & Nation, 2017). Góc nhìn này phù hợp với lỗi ngữ nghĩa và kết hợp từ vì người học có thể biết nghĩa cơ bản nhưng chưa kiểm soát được ranh giới nghĩa và mạng lưới kết hợp. Nghiên cứu phân loại lỗi ngữ nghĩa thành 4 tiểu loại: nhầm từ gần nghĩa/đồng nghĩa, sai phạm vi nghĩa, sai sắc thái biểu cảm và sai phong cách ngữ thể. Lỗi kết hợp từ gồm kết hợp động từ-tân ngữ, tính từ-danh từ, kết hợp cố định/quen dùng và sử dụng từ li hợp. Hệ thống 8 tiểu loại cho phép xem xét đồng thời việc lựa chọn mục từ và khả năng đưa mục từ vào cấu trúc sử dụng.

Collocation có tính quy ước và không phải lúc nào cũng suy ra được từ nghĩa riêng lẻ của các thành tố (Firth, 1957; Sinclair, 1991). Li (2016) cho thấy lỗi kết hợp là khó khăn đáng kể ở người học Tiếng Trung. Với người học Việt Nam, nghiên cứu về từ gần nghĩa và từ li hợp cũng cho thấy biết nghĩa khái quát chưa bảo đảm sử dụng chính xác trong ngữ cảnh (Luu Vũ Hón, 2023; Mai Thị Huế, 2018).

3.2. Khoảng trống và câu hỏi nghiên cứu

Các nghiên cứu trước thường khảo sát riêng lỗi ngữ nghĩa hoặc lỗi kết hợp từ, tập trung vào từng cặp từ hay hiện tượng cụ thể; nhóm SV chuyên ngành Ngôn ngữ Anh học Tiếng Trung như ngoại ngữ hai cũng còn ít được khảo sát. Các nghiên cứu trong nước về từ gần nghĩa, từ li hợp và chuyển dịch nghĩa Hán-Việt cho thấy đây là những vùng khó khăn đáng chú ý (Luu Vũ Hón, 2023; Mai Thị Huế, 2018; Trần Quang Huy, 2024) nhưng chưa cung cấp bức tranh so sánh đồng thời 8 tiểu loại lỗi trên cùng một mẫu. Từ khoảng trống trên, nghiên cứu tập trung trả lời ba câu hỏi: 1) SV học Tiếng Trung như ngoại ngữ hai thường mắc những lỗi ngữ nghĩa và lỗi kết hợp từ nào? 2) Mức độ và đặc điểm phân bố của các lỗi ra sao, và tổng lỗi ngữ nghĩa có khác tổng lỗi kết hợp từ một cách có ý nghĩa thống kê hay không? 3) Những yếu tố nào có khả năng liên quan đến sự hình thành lỗi, và kết quả gợi ra hàm ý gì cho dạy học từ vựng tiếng Trung?

3.3. Thực trạng lỗi sử dụng từ vựng tiếng Trung

3.3.1. Bức tranh chung về 8 tiểu loại lỗi

Bảng 2: Xếp hạng các tiểu loại lỗi theo số lỗi trung bình

Tiểu loại lỗi	M lỗi/SV	Tỷ lệ SV mắc lỗi (%)
Nhầm từ gần nghĩa/đồng nghĩa	3,78	97,4
Kết hợp động từ–tân ngữ	2,30	92,8
Kết hợp cố định/quen dùng	1,06	70,6
Sai phạm vi nghĩa	1,02	74,2
Từ li hợp	0,88	68,0
Sai sắc thái biểu cảm	0,70	55,7
Kết hợp tính từ–danh từ	0,69	57,2
Sai phong cách ngữ thể	0,21	19,6

Nguồn: Tổng hợp từ kết quả xử lý dữ liệu khảo sát bằng SPSS, 2026 (N = 194).

Kết quả cho thấy lỗi tập trung mạnh ở hai khu vực: phân biệt từ gần nghĩa/đồng nghĩa và kết hợp động từ–tân ngữ. Gần như toàn bộ mẫu (97,4%) mắc ít nhất một lỗi nhầm từ gần nghĩa/đồng nghĩa, với trung bình 3,78 lỗi/SV. Ở nhóm kết hợp từ, 92,8% SV mắc ít nhất một lỗi động từ–tân ngữ, với trung bình 2,30 lỗi/SV. Sự tập trung này cho thấy khó khăn nổi bật không nằm ở một loại lỗi ngẫu nhiên mà ở các phương diện đòi hỏi tri thức về ranh giới nghĩa và giới hạn kết hợp.

3.3.2. Lỗi ngữ nghĩa

Bảng 3: Thống kê mô tả các tiểu loại lỗi ngữ nghĩa

Tiểu loại	Min	Max	M	SD
Nhầm từ gần nghĩa/đồng nghĩa	0	6	3,78	1,61
Sai phạm vi nghĩa	0	4	1,02	0,80
Sai sắc thái biểu cảm	0	2	0,70	0,70
Sai phong cách ngữ thể	0	3	0,21	0,44
Tổng lỗi ngữ nghĩa	0	11	5,70	2,27

Nguồn: Tổng hợp từ kết quả xử lý dữ liệu khảo sát bằng SPSS, 2026 (N = 194).

Nhầm từ gần nghĩa/đồng nghĩa chiếm ưu thế rõ rệt trong nhóm lỗi ngữ nghĩa. Kết quả chi tiết cho thấy 189/194 SV (97,4%) mắc ít nhất một lỗi, trong khi số lỗi dao động từ 0 đến 6. Mẫu lỗi cho thấy người học thường lựa chọn từ dựa trên nghĩa khái quát hoặc tương đương tiếng Việt mà chưa phân biệt đầy đủ đối tượng, ngữ cảnh, phạm vi nghĩa và điều kiện sử dụng. Lỗi sai phạm vi nghĩa đứng thứ hai về số lỗi trung bình; lỗi sắc thái và phong cách xuất hiện thấp hơn nhưng phản ánh những giới hạn ở tầng ngữ dụng của tri thức từ vựng.

3.3.3. Lỗi kết hợp từ

Bảng 4: Thống kê mô tả các tiểu loại lỗi kết hợp từ

Tiểu loại	Min	Max	M	SD
Động từ–tân ngữ	0	3	2,30	0,96
Tính từ–danh từ	0	3	0,69	0,68
Cố định/quen dùng	0	3	1,06	0,86
Từ li hợp	0	2	0,88	0,71
Tổng lỗi kết hợp từ	0	11	4,93	1,93

Nguồn: Tổng hợp từ kết quả xử lý dữ liệu khảo sát bằng SPSS, 2026 (N = 194).

Lỗi động từ–tân ngữ là dạng lỗi kết hợp nổi bật nhất: 180/194 SV (92,8%) mắc ít nhất một lỗi. Các trường hợp như *喝药 thay cho 吃药 hoặc *价格很贵 thay cho 价格高 cho thấy người học có thể biết

nghĩa của từng thành tố nhưng chuyển giao thói quen kết hợp từ tiếng Việt hoặc tự ghép các đơn vị dựa trên tính hợp lý về nghĩa. Lỗi từ li hợp phản ánh một cơ chế khác: người học có xu hướng xử lý những đơn vị có khả năng tách-hợp như động từ thông thường, cho thấy tri thức về cấu trúc và điều kiện kết hợp của ngôn ngữ đích chưa hoàn thiện.

3.3.4. So sánh tổng lỗi ngữ nghĩa và tổng lỗi kết hợp từ

Khi so sánh tổng số lỗi ngữ nghĩa và tổng số lỗi kết hợp từ trên cùng một mẫu nghiên cứu ($N = 194$), kết quả phân tích cho thấy có sự chênh lệch rõ rệt giữa hai loại lỗi này. Trong 194 SV thì 107 người có tổng lỗi ngữ nghĩa lớn hơn tổng lỗi kết hợp từ, 53 người có xu hướng ngược lại và 34 người có số lỗi bằng nhau. Kết quả kiểm định phi tham số Wilcoxon cho thấy khác biệt có ý nghĩa thống kê ($Z = -4,704$; $p < 0,001$). Như vậy, trong phạm vi bài khảo sát, lỗi ngữ nghĩa là khu vực khó khăn nổi trội hơn. Kết quả này không đồng nghĩa với việc lỗi kết hợp từ ít quan trọng, bởi tổng lỗi kết hợp vẫn đạt trung bình 4,93 lỗi/SV và một số tiểu loại có độ bao phủ rất cao.

3.4. Một số yếu tố liên quan đến lỗi và dữ liệu đối chiếu

3.4.1. Yếu tố được sinh viên tự đánh giá

Bảng 5: Thống kê mô tả thang đo nguyên nhân gây lỗi

Mã	Nội dung biến	M	SD
NGN1	Chưa hiểu đầy đủ cách sử dụng của một số từ tiếng Trung	3,01	1,09
NGN2	Ít có cơ hội sử dụng tiếng Trung trong thực tế	3,32	1,02
NGN3	Ghi nhớ từ chủ yếu dựa vào nghĩa tiếng Việt	3,02	1,07
NGN4	Khó phân biệt cách kết hợp của một số từ tiếng Trung	3,16	1,00
NGN5	Chịu ảnh hưởng của tiếng Việt khi sử dụng từ tiếng Trung	3,87	0,87
NGN6	Liên hệ từ Hán-Việt quen thuộc để ghi nhớ/lựa chọn từ	2,90	0,90
NGN7	Chưa dành đủ thời gian luyện tập sử dụng từ vựng	2,95	0,98
NGN8	Hoạt động và bài tập thực hành từ vựng trên lớp chưa đủ	3,01	0,96

Nguồn: Tổng hợp từ kết quả xử lý dữ liệu khảo sát bằng SPSS ($N = 194$).

Ảnh hưởng của tiếng Việt (NGN5) có điểm trung bình cao nhất ($M = 3,87$; $SD = 0,87$), tiếp theo là ít có cơ hội sử dụng tiếng Trung trong thực tế (NGN2: $M = 3,32$; $SD = 1,02$) và khó phân biệt cách kết hợp từ (NGN4: $M = 3,16$; $SD = 1,00$). Các biến còn lại dao động từ 2,90 đến 3,02. Đây là dữ liệu tự đánh giá, do đó phản ánh nhận thức của người học về các yếu tố có liên quan chứ không đủ để xác lập quan hệ nhân quả.

3.4.2. Đối chiếu với dữ liệu phỏng vấn

Dữ liệu phỏng vấn hỗ trợ 3 hướng giải thích chính: 1) Giảng viên ghi nhận xu hướng người học “tư duy từng từ” từ tiếng Việt sang tiếng Trung, sau đó ghép các đơn vị tương ứng theo thói quen của tiếng Việt. Hướng giải thích này phù hợp với điểm tự đánh giá cao của NGN5 và với các lỗi kết hợp như *喝茶 hoặc *价格很贵; 2) Các lỗi từ gần nghĩa và từ li hợp còn phản ánh nguyên nhân nội ngôn ngữ: người học chưa phân hóa đầy đủ nét nghĩa, điều kiện sử dụng và đặc điểm cấu trúc của từ tiếng Trung; 3) Việc học từ ở cấp độ đơn lẻ và cơ hội sử dụng tiếng Trung hạn chế làm giảm khả năng củng cố các mẫu kết hợp và ngữ cảnh sử dụng.

Một giảng viên nhấn mạnh rằng SV cần chuyển từ việc ghi nhớ một nghĩa tương đương sang học từ cùng câu/cụm từ điển hình và hoàn cảnh sử dụng; một giảng viên khác cho rằng học liệu cần cho người học tiếp xúc nhiều hơn với cặp từ gần nghĩa, cụm từ thường dùng và ngữ cảnh cụ thể. Các ý kiến này nhất quán với kết quả định lượng về NGN2, NGN4 và NGN8. Ngược lại, dữ liệu phỏng vấn chưa cung cấp đủ cơ sở để xem tiếng Anh là nguồn đáng kể của các lỗi từ vựng trong phạm vi khảo sát.

3.5. Thảo luận kết quả

3.5.1. Từ “biết nghĩa” đến “biết dùng”: chiều sâu tri thức từ vựng

Phát hiện trung tâm của nghiên cứu là khoảng cách giữa tri thức nghĩa cơ bản và năng lực sử dụng

từ. Khi 97,4% SV mắc lỗi từ gần nghĩa/đồng nghĩa và 92,8% mắc lỗi động từ-tân ngữ, khó khăn không thể được giải thích đơn thuần bằng quy mô vốn từ. Kết quả phù hợp với quan niệm cho rằng tri thức từ vựng bao gồm quan hệ nghĩa, hạn chế sử dụng, đặc điểm ngữ pháp và quan hệ kết hợp, chứ không chỉ là liên kết hình thức-nghĩa (Nation, 2001; Schmitt, 2000). Kết quả về từ gần nghĩa cũng tương thích với nghiên cứu của Lưu Vũ Hân (2023), trong đó người học Việt Nam có xu hướng đồng nhất các từ có cùng tương đương dịch nhưng khác về phạm vi và điều kiện sử dụng. Ở bình diện kết hợp, tỷ lệ lỗi động từ-tân ngữ cao nhất quán với hướng nghiên cứu collocation của Li (2016). Điều này cho thấy việc học từng mục từ tách rời có thể tạo ra vốn từ nhận biết tương đối lớn nhưng chưa đủ để sản sinh phát ngôn tự nhiên.

3.5.2. Vai trò nổi bật nhưng không tuyệt đối của tiếng Việt

Ảnh hưởng của tiếng Việt có điểm tự đánh giá cao nhất trong thang nguyên nhân ($M = 3,87$; $SD = 0,87$). Kết quả này phù hợp với lý thuyết chuyển di (Ogden, 1989) và với đặc điểm người học Việt Nam có sẵn mạng lưới tương ứng Hán-Việt. Các lỗi như chuyển trực tiếp cách kết hợp tiếng Việt sang tiếng Trung hoặc lựa chọn từ theo một tương đương dịch duy nhất là biểu hiện hợp lý của cơ chế này. Tuy nhiên, dữ liệu không ủng hộ cách giải thích đơn nguyên nhân. Biểu liên hệ Hán-Việt (NGN6) chỉ đạt $M = 2,90$; lỗi từ li hợp và một phần lỗi từ gần nghĩa còn phản ánh tri thức nội ngôn ngữ chưa hoàn thiện. Cách giải thích đa nguồn này phù hợp với phân tích lỗi hiện đại (Corder, 1967; Ellis & Barkhuizen, 2005; Richards, 1980). Dữ liệu phỏng vấn cũng chưa đủ để xem tiếng Anh là nguồn lỗi từ vựng đáng kể, vì vậy không nên mặc định mọi lỗi của người học ngoại ngữ hai đều bắt nguồn đồng thời từ L1 và L2.

3.5.3. Bối cảnh học ngoại ngữ hai và cơ hội sử dụng

Sinh viên tự đánh giá việc ít có cơ hội sử dụng tiếng Trung trong thực tế ở mức tương đối cao ($M = 3,32$; $SD = 1,02$). Trong bối cảnh tiếng Trung là ngoại ngữ hai, thời lượng tiếp xúc và cơ hội sử dụng thường phải cạnh tranh với các học phần chuyên ngành và ngoại ngữ khác. Điều này có thể làm giảm số lần người học gặp lại, lựa chọn, kết hợp và kiểm chứng cách dùng của từ trong những ngữ cảnh khác nhau. Tuy nhiên, mối liên hệ trên cần được diễn giải thận trọng. Dữ liệu về cơ hội sử dụng và nguyên nhân chủ yếu là tự báo cáo, trong khi thiết kế nghiên cứu là cắt ngang. Vì vậy, nghiên cứu không chứng minh rằng ít cơ hội sử dụng trực tiếp gây ra số lỗi cao. Kết quả cho thấy một mẫu hình nhất quán giữa tự đánh giá, biểu hiện lỗi và phỏng vấn, từ đó xác định những yếu tố cần ưu tiên kiểm tra bằng các thiết kế dọc hoặc can thiệp trong nghiên cứu tiếp theo.

3.6. Hàm ý sư phạm

Đó là: - *Dạy từ theo ngữ cảnh và mạng lưới nghĩa*: Với các cặp từ gần nghĩa, giảng viên nên tổ chức đối chiếu bằng các ngữ cảnh tối thiểu, làm rõ nét nghĩa khu biệt, đối tượng điển hình, sắc thái và trường hợp không thể thay thế. Bài tập nên yêu cầu người học lựa chọn và giải thích lý do thay vì chỉ nối từ với nghĩa tiếng Việt; - *Chuyển đơn vị dạy học từ “từ” sang “cụm”*: Mỗi mục từ mới nên đi kèm các kết hợp điển hình, cấu trúc cú pháp và ví dụ đầu ra. Hướng này phù hợp với lexical approach và nghiên cứu về collocation (Lewis, 1993; Sinclair, 1991), đồng thời trực tiếp đáp ứng nhóm lỗi động từ-tân ngữ và kết hợp cố định có tần suất cao; - *Sử dụng đối chiếu Việt-Trung có chọn lọc*: Tiếng Việt và yếu tố Hán-Việt có thể được khai thác như công cụ hỗ trợ nhận diện nhưng cần đánh dấu rõ các trường hợp chỉ tương đồng một phần hoặc khác về sắc thái, phạm vi nghĩa và khả năng kết hợp. Mục tiêu không phải loại bỏ tiếng Việt khỏi lớp học mà biến nó thành đối tượng so sánh có kiểm soát; - *Tăng cường đầu ra có phản hồi*: Các nhiệm vụ đặt câu, viết đoạn ngắn, dịch có định hướng, đóng vai và sửa lỗi theo cặp giúp người học chuyển tri thức nhận biết thành tri thức sử dụng. Phản hồi nên ưu tiên các mẫu lỗi có độ phủ cao, đặc biệt là từ gần nghĩa/đồng nghĩa và động từ-tân ngữ; - *Xây dựng ngân hàng lỗi người học*: Những lỗi thực tế trong bài làm có thể được ẩn danh, phân loại và tái sử dụng thành bài tập chẩn đoán. Theo thời gian, ngân hàng này có thể phát triển thành một kho ngữ liệu người học quy mô nhỏ phục vụ cả giảng dạy và nghiên cứu; - *Hướng dẫn người học tự kiểm chứng cách dùng bằng từ điển điện tử, corpus và công cụ AI một cách có kiểm soát*: Trọng tâm của hoạt động không phải là nhận một đáp án tự động mà

là so sánh ví dụ xác thực, kiểm tra cụm từ thường đi kèm và giải thích vì sao một lựa chọn phù hợp hơn lựa chọn khác. Hướng sử dụng này phù hợp với mục tiêu chuyển từ ghi nhớ nghĩa sang hình thành tri thức sử dụng.

4. Kết luận

Nghiên cứu cho thấy lỗi ngữ nghĩa và lỗi kết hợp từ là hai khó khăn rõ rệt trong sử dụng từ vựng tiếng Trung của SV học Tiếng Trung như ngoại ngữ hai. Tổng lỗi ngữ nghĩa cao hơn tổng lỗi kết hợp từ có ý nghĩa thống kê; nhầm từ gần nghĩa/đồng nghĩa và kết hợp động từ-tân ngữ là hai dạng nổi bật nhất. Dữ liệu tự đánh giá và phỏng vấn cho thấy ảnh hưởng của tiếng Việt, cơ hội sử dụng hạn chế, khó khăn trong phân biệt quan hệ kết hợp và cách học từ đơn lẻ là những yếu tố đáng chú ý; đồng thời, một bộ phận lỗi phản ánh tri thức nội ngôn ngữ tiếng Trung chưa hoàn thiện. Kết quả khẳng định rằng biết nghĩa chưa đồng nghĩa với biết dùng: tri thức từ vựng cần bao gồm ranh giới nghĩa, sắc thái, điều kiện sử dụng và mạng lưới kết hợp. Vì vậy, dạy học nên ưu tiên từ gần nghĩa, collocation, từ li hợp, các mẫu kết hợp và hoạt động đầu ra có phản hồi; đối chiếu Việt-Trung cần được sử dụng có chọn lọc để hỗ trợ nhận diện khác biệt thay vì củng cố tương đương 1-1. Nghiên cứu còn giới hạn ở một cơ sở đào tạo, một học phần và thiết kế cắt ngang; một phần dữ liệu về nguyên nhân là tự báo cáo nên không cho phép suy luận nhân quả. Nghiên cứu tiếp theo nên mở rộng cơ sở và trình độ, sử dụng thiết kế dọc hoặc thực nghiệm, đồng thời kiểm tra hiệu quả của dạy từ theo cụm, đối chiếu ngữ cảnh và tăng cường đầu ra có phản hồi.

Chú thích

(*) Nghiên cứu này được tài trợ bởi Đại học Bách khoa Hà Nội (HUST) trong đề tài mã số T2025-PC-005.

Tài liệu tham khảo

- Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition* (pp. 8-20). Multilingual Matters.
- Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, 5(1-4), 161-170.
- Corder, S. P. (1974). Error analysis. In J. P. B. Allen & S. P. Corder (Eds.), *The Edinburgh course in applied linguistics: Vol. 3. Techniques in applied linguistics* (pp. 122-154). Oxford University Press.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford University Press.
- Firth, J. R. (1957). *Papers in linguistics 1934-1951*. Oxford University Press.
- James, C. (1998). *Errors in language learning and use: Exploring error analysis*. Routledge. <https://doi.org/10.4324/9781315842912>.
- Lewis, M. (1993). *The lexical approach*. Language Teaching Publications.
- Li, S. (2016). A corpus-based analysis of collocational errors by American learners of Chinese and its implication for the teaching of vocabulary. *Chinese as a Second Language*, 51(1), 62-78.
- Lưu Vũ Hân. (2023). Phân tích lỗi nhầm lẫn các từ tiếng Trung Quốc “变”, “改变” và “变化” của sinh viên Việt Nam. *Tạp chí Khoa học Đại học Mở Thành phố Hồ Chí Minh - Khoa học Xã hội*, 18(2), 89-99.
- Mai Thị Huệ (2018). *Nghiên cứu quá trình thụ đắc từ li hợp trong tiếng Hán hiện đại của sinh viên Việt Nam*. Luận án Tiến sĩ Ngôn ngữ, Trường Đại học Ngoại ngữ, Đại học Quốc gia Hà Nội.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press.
- Richards, J. C. (1980). Second language acquisition: Error analysis. *Annual Review of Applied Linguistics*, 1, 91-107.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10, 209-231.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Trần Quang Huy.(2024). Sự chuyển dịch về mặt ngữ nghĩa của từ Hán Việt và những lỗi sai khi dịch Việt-Trung, Trung-Việt. *Tạp chí Khoa học và Công nghệ - Đại học Đà Nẵng*, 22(8), 65-68.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.