NGOẠI NGỮ VỚI BẢN NGỮ

NGHIÊN CỨU TẦN SUẤT VÀ CÁCH SỬ DỤNG CÁC PHƯƠNG TIỆN LIÊN KẾT TRONG BÀI VIẾT IELTS TASK 2 CỦA SINH VIÊN NĂM THỨ NHẤT KHÔNG CHUYÊN TẠI TRƯỜNG ĐẠI HỌC HÀ NỘI

TRẦN QUỐC VIỆT* - TRẦN THỊ THANH TÂM - ĐẶNG NAM THẮNG*****

TÓM TẮT: Nghiên cứu khảo sát tần suất và cách thức sử dụng các phương tiện liên kết trong bài viết IELTS Task 2 của sinh viên không chuyên năm thứ nhất Trường Đại học Hà Nội, dựa trên khung lí thuyết năm phạm trù liên kết của Halliday và Hasan (1976) và Halliday và Matthiessen (2014). Ngữ liệu nghiên cứu gồm 103 bài viết (tương ứng 32.102 từ), thu thập theo phương pháp chọn mẫu nghiên cứu thuận tiện và được mã hóa bằng phương pháp phân tích ngữ liệu kết hợp phân tích định tính và định lượng. Kết quả cho thấy, toàn bộ ngữ liệu chứa 5.160 phương tiện liên kết ngữ pháp, trong đó quy chiếu chiếm tỉ lệ cao nhất (51,05%), tiếp theo là liên từ (chiếm 46,72%), và cuối cùng là phép thế (chiếm 2,23%). Xét theo tiểu phạm trù, quy chiếu ngôi ba và liên từ bổ sung là hai phương thức liên kết được sử dụng phổ biến nhất. Kết quả nghiên cứu không chỉ góp phần làm rõ đặc điểm sử dụng phương tiện liên kết mà đồng thời còn gợi mở một số định hướng nhằm nâng cao chất lượng giảng dạy kỹ năng viết học thuật tại Trường Đại học Hà Nội.

TỪ KHÓA: Liên kết; phương tiện liên kết; tần suất sử dụng; sinh viên năm thứ nhất không chuyên; bài viết IELTS Task 2.

NHẬN BÀI: 26/07/2026. BIÊN TẬP-CHỈNH SỬA-DUYỆT ĐĂNG: 04/11/2026

1. Đặt vấn đề

Trong những năm qua, bài kiểm tra đánh giá năng lực sử dụng tiếng Anh của người học tiếng Anh như một ngoại ngữ (IELTS) đã trở thành một trong những chứng chỉ tiếng Anh quốc tế được công nhận phổ biến tại Việt Nam, không chỉ phục vụ mục đích du học hay tuyển dụng mà còn được nhiều cơ sở giáo dục đại học sử dụng như một điều kiện xét miễn học phần tiếng Anh trong chương trình đào tạo hoặc điều kiện xét tốt nghiệp. Trong bốn kĩ năng được đánh giá, viết IELTS Task 2 thường được xem là phần thi có yêu cầu cao đối với người học, bởi thí sinh phải đồng thời đáp ứng các yêu cầu về tổ chức và triển khai lập luận, tính mạch lạc của văn bản, cũng như khả năng sử dụng các phương tiện ngôn ngữ chính xác và đa dạng trong điều kiện thời gian cho phép. Trong bốn tiêu chí chấm điểm của bài viết IELTS Task 2, mạch lạc và liên kết chiếm 25% tổng điểm, tương đương với ba tiêu chí còn lại, và được xem là tiêu chí phản ánh trực tiếp khả năng tổ chức, kết nối ý tưởng bằng các phương tiện ngôn ngữ - trong đó phương tiện liên kết giữ vai trò quan trọng.

Mặc dù tầm quan trọng của phương tiện liên kết đối với bài viết học thuật bằng tiếng Anh đã được khẳng định qua nhiều công trình nghiên cứu kể từ mô hình của Halliday và Hasan (1976); tuy nhiên, phần lớn các nghiên cứu tại Việt Nam về chủ đề này vẫn tập trung chủ yếu vào nhóm sinh viên chuyên ngành tiếng Anh hoặc đối tượng là người học tiếng Anh tại các trung tâm ngoại ngữ. Trong khi đó, sinh viên không chuyên - nhóm đối tượng chiếm số lượng áp đảo trong các trường đại học tại

**TS; Trường Đại học Hà Nội; Email: viettq@hanu.edu.vn (Tác giả liên hệ)*

***Email: tamtran121205@gmail.com*

****Trường Đại học Hà Nội; Email: thangdn@hanu.edu.vn*

Việt Nam và ngày càng có nhu cầu cao đối với chứng chỉ IELTS phục vụ các mục đích học tập và nghiên cứu, lại chưa nhận được nhiều sự quan tâm, đặc biệt là ở giai đoạn năm thứ nhất, khi năng lực viết học thuật bằng tiếng Anh còn đang trong quá trình định hình.

Xuất phát từ khoảng trống nói trên, nghiên cứu của chúng tôi tiến hành khảo sát và mô tả một cách có hệ thống tần suất và cách thức sử dụng các phương tiện liên kết trong bài viết IELTS Task 2 của sinh viên không chuyên năm thứ nhất Trường Đại học Hà Nội, dựa trên khung lí thuyết năm phạm trù liên kết của Halliday và Hasan (1976) và Halliday và Matthiessen (2014). Cụ thể, nghiên cứu tập trung trả lời ba câu hỏi sau: (i) Các phương tiện liên kết ngữ pháp được sử dụng với tần suất và tỉ lệ như thế nào trong toàn bộ ngữ liệu được khảo sát? (ii) Các tiểu phạm trù liên kết được phân bố ra sao trong từng phạm trù và trong toàn bộ ngữ liệu được khảo sát? và (iii) Mối quan hệ giữa mức độ sử dụng liên kết ngữ pháp và mức độ đa dạng từ vựng trong bài viết của sinh viên là như thế nào?

Kết quả nghiên cứu được kì vọng không chỉ góp phần bổ sung cho khoảng trống nghiên cứu về đối tượng sinh viên không chuyên, mà còn cung cấp cơ sở thực chứng cho việc thiết kế các hoạt động giảng dạy kĩ năng viết học thuật phù hợp hơn với đặc điểm ngôn ngữ thực tế không chỉ của nhóm sinh viên không chuyên này.

2. Một số vấn đề lí luận

2.1. Khái niệm liên kết

Liên kết là một khái niệm trung tâm của ngôn ngữ học văn bản, đề cập đến những quan hệ giúp các câu trong văn bản tạo thành một chỉnh thể có tổ chức. Halliday và Hasan (1976, tr.2-4) định nghĩa liên kết "là quan hệ ngữ nghĩa trong đó việc diễn giải một yếu tố ngôn ngữ phụ thuộc vào một yếu tố khác trong văn bản". Theo đó, các quan hệ liên kết tạo nên sự gắn kết giữa các thành tố và góp phần phân biệt một văn bản với một chuỗi câu rời rạc.

Trong nghiên cứu tiếng Việt, Diệp Quang Ban (2006, tr.119) xem liên kết là "tập hợp các phương tiện ngữ pháp và từ vựng có chức năng vượt qua ranh giới câu, tạo nên sự thống nhất của văn bản; xét về bản chất, đó là quan hệ nghĩa giữa các yếu tố ngôn ngữ mà việc diễn giải một yếu tố phụ thuộc vào yếu tố kia". Tương tự, Nguyễn Thiện Giáp (2009, tr.173) nhấn mạnh đến phương diện hình thức khi định nghĩa liên kết là "những dấu hiệu cho thấy các kiểu quan hệ giữa các câu trong diễn ngôn". Như vậy, mặc dù có sự khác biệt về cách diễn đạt, các quan niệm trên đều thống nhất ở việc xem liên kết là những quan hệ ngôn ngữ có chức năng kết nối các thành tố của văn bản.

Liên kết cần được phân biệt với mạch lạc, một khái niệm gần gũi nhưng không đồng nhất. Theo De Beaugrande và Dressler (1981, tr.3-10), liên kết liên quan đến các quan hệ giữa những thành tố được biểu hiện trên bề mặt văn bản, trong khi mạch lạc liên quan đến các quan hệ khái niệm được kiến tạo và diễn giải ở cấp độ ý nghĩa. Vì vậy, liên kết chủ yếu có thể được nhận diện thông qua các dấu hiệu ngôn ngữ và được lượng hóa, trong khi mạch lạc phụ thuộc nhiều hơn vào cách các ý tưởng được tổ chức và diễn giải trong toàn bộ văn bản. Một văn bản có thể sử dụng nhiều phương tiện liên kết nhưng vẫn thiếu mạch lạc nếu các ý tưởng không được triển khai theo một trình tự hợp lí. Sự phân biệt này có ý nghĩa trực tiếp đối với nghiên cứu của chúng tôi, bởi tần suất sử dụng phương tiện liên kết có thể được khảo sát như một chỉ báo định lượng, nhưng không đồng nhất với mức độ mạch lạc của toàn bộ bài viết. Điều này cũng lí giải vì sao trong các khung đánh giá viết học thuật, trong đó có bài viết IELTS Task 2, mạch lạc và liên kết thường được xem xét đồng thời nhưng vẫn là hai khía cạnh khác nhau.

2.2. Khái niệm phương tiện liên kết

Nếu liên kết là quan hệ ngữ nghĩa tổng quát nói trên, thì phương tiện liên kết là đơn vị ngôn ngữ cụ thể hiện thực hóa quan hệ đó trong văn bản. Halliday và Hasan (1976) gọi mỗi trường hợp liên kết đơn lẻ là một "tie" - tức một cặp quan hệ giữa yếu tố tiền giả định và yếu tố được giả định. Một phương tiện liên kết; do đó, không tồn tại một cách độc lập mà luôn cần được xác định trong quan hệ với một yếu tố khác trong văn bản, chẳng hạn đại từ "it" chỉ trở thành một phương tiện liên kết khi nó quy chiếu về một danh từ cụ thể đã xuất hiện trước đó, và ý nghĩa của "it" chỉ được xác định đầy đủ khi người đọc truy ngược lại yếu tố được quy chiếu đó (tr.4).

Bên cạnh đó, các tác giả cũng phân biệt hai hướng quy chiếu của phương tiện liên kết: (1) tham chiếu ngoại chỉ là khi yếu tố ngôn ngữ hướng ra tình huống giao tiếp bên ngoài văn bản để được diễn giải, và (2) tham chiếu nội chỉ là khi yếu tố ngôn ngữ hướng đến một yếu tố khác nằm trong chính văn bản. Chỉ tham chiếu nội chỉ mới được tính là phương tiện liên kết theo nghĩa hẹp, vì chỉ trường hợp này mới tạo ra một mối nối giữa hai điểm trong văn bản, góp phần vào kết cấu chung của toàn bộ đoạn văn hay bài viết. Tham chiếu nội chỉ, đến lượt nó, còn được chia thành hồi chỉ và khứ chỉ (tr.33). Như vậy, cách hiểu này về phương tiện liên kết được coi như một đơn vị ngôn ngữ mang tính quan hệ, và chỉ có ý nghĩa đầy đủ khi được đặt trong liên hệ với một yếu tố khác trong văn bản và đây chính là cơ sở lí thuyết cho việc xây dựng các hệ thống phân loại và mã hóa phương tiện liên kết được áp dụng trong nghiên cứu này.

2.3. Phân loại phương tiện liên kết

Dựa trên các quan niệm khác nhau về liên kết, các nhà nghiên cứu đã đề xuất nhiều cách phân loại phương tiện liên kết, chẳng hạn, Quirk và cộng sự (1985) phân chia liên từ thành ba nhóm: liên từ tương quan, liên từ phụ thuộc và liên từ kết hợp. Trong nghiên cứu này, khung phân loại của Halliday và Hasan (1976, tr.4-5) được chúng tôi lựa chọn do có tính hệ thống và phù hợp với việc phân tích định lượng các phương tiện liên kết trong văn bản. Theo đó, liên kết được chia thành hai nhóm lớn là liên kết ngữ pháp và liên kết từ vựng.

Liên kết ngữ pháp bao gồm bốn phạm trù, gồm: (1) quy chiếu là quan hệ trong đó việc diễn giải một yếu tố ngôn ngữ phụ thuộc vào một yếu tố khác trong văn bản hoặc ngữ cảnh, gồm quy chiếu cá nhân, quy chiếu chỉ định và quy chiếu so sánh; (2) phép thế là việc sử dụng một yếu tố thay thế cho yếu tố đã xuất hiện nhằm hạn chế sự lặp lại về hình thức, gồm thế danh từ, thế động từ và thế mệnh đề; (3) tỉnh lược là sự lược bỏ một thành tố có thể được khôi phục từ ngữ cảnh, bao gồm tỉnh lược danh từ, động từ và mệnh đề, và (4) liên từ là sự biểu thị các quan hệ logic, ngữ nghĩa giữa các đơn vị văn bản, thường được phân thành các quan hệ bổ sung, tương phản, nguyên nhân và thời gian.

Liên kết từ vựng được tạo lập thông qua sự lựa chọn các từ ngữ có quan hệ ngữ nghĩa với nhau và không phụ thuộc vào một cấu trúc ngữ pháp cụ thể. Halliday và Hasan (1976, tr.274) phân chia liên kết từ vựng thành hai dạng chính là phép quy hồi và phép kết hợp từ.

Mặc dù được đề xuất từ năm 1976, mô hình năm phạm trù của Halliday và Hasan vẫn được sử dụng rộng rãi trong nghiên cứu về liên kết, đặc biệt trong các nghiên cứu định lượng về liên từ trong bài viết của người học ngoại ngữ. Nhờ tính hệ thống, khả năng thao tác hóa thành các đơn vị có thể dễ nhận diện và đếm được, cũng như khả năng đối chiếu kết quả giữa các nghiên cứu, mô hình Halliday và Hasan (1976) được sử dụng làm khung lí thuyết phân tích chính trong nghiên cứu của chúng tôi.

2.4. Liên kết trong bài viết tiếng Anh của người học tại Việt Nam

Tại Việt Nam, nghiên cứu về phương tiện liên kết trong bài viết tiếng Anh của người học EFL còn tương đối hạn chế, nhưng cũng đã được triển khai trên nhiều nhóm đối tượng và loại hình văn bản khác nhau. Tran (2005) là một trong những nghiên cứu đầu tiên về liên kết khi tập trung vào việc phân tích lỗi sử dụng phương tiện liên kết trong bài viết của sinh viên năm nhất chuyên ngành tiếng Anh. Tiếp đó, Hung và Thu (2014) cũng chỉ ra rằng liên kết từ vựng chiếm tỷ lệ cao nhất, tiếp theo là quy chiếu và liên từ, đồng thời phát hiện một số lỗi sử dụng có tính hệ thống. Ở phạm vi liên ngôn ngữ, Bui và cộng sự (2021) khảo sát việc sử dụng phương tiện liên kết của giáo viên EFL Việt Nam trong bối cảnh so sánh với người học EFL ở châu Á. Bui (2022), tiến hành nghiên cứu 168 báo cáo học thuật của sinh viên năm cuối chuyên ngành tiếng Anh, cũng cho thấy liên kết từ vựng chiếm tỉ lệ cao nhất (39,68%), tiếp theo là quy chiếu (37,48%) và liên từ (15,52%), nghiên cứu cũng chỉ ra sự thiếu nhất quán giữa nhận thức và việc vận dụng phương tiện liên kết của người học. Ngược lại, Nguyen và cộng sự (2023) khi phân tích 400 đoạn văn của 80 người học trình độ tiền trung cấp, nhận thấy liên kết ngữ pháp chiếm ưu thế rõ rệt (trong đó quy chiếu và liên từ lần lượt chiếm 51% và 38,5%), còn liên kết từ vựng chỉ chiếm 9,3%.

Sự khác biệt giữa các kết quả trên cho thấy việc phân bố và sử dụng phương tiện liên kết có thể chịu ảnh hưởng trực tiếp của trình độ người học, chuyên ngành học, thể loại văn bản và nhiệm vụ yêu cầu của bài viết, cũng như phương pháp phân tích. Vì vậy, các nghiên cứu này cũng chưa thể xác lập một khuynh lí thuyết thống nhất về việc sử dụng phương tiện liên kết của người học tại Việt Nam. Bên cạnh đó, phần lớn các công trình này cũng chủ yếu tập trung vào đối tượng nghiên cứu là nhóm sinh viên chuyên ngành tiếng Anh hoặc người học tại các trung tâm ngoại ngữ, trong khi sinh viên không chuyên tiếng Anh vẫn chưa nhận được nhiều sự quan tâm. Đặc biệt, mối quan hệ giữa việc sử dụng phương tiện liên kết và chất lượng viết luận theo định dạng bài viết IELTS Task 2 còn ít được khảo cứu. Đây chính là khoảng trống tạo cơ sở cho nghiên cứu của chúng tôi tập trung vào việc xác lập tần suất và cách thức sử dụng phương tiện liên kết trong bài viết IELTS Task 2 của sinh viên không chuyên tiếng Anh tại Việt Nam.

2.5. Liên kết và viết IELTS Task 2

Theo hướng dẫn chấm thi chính thức của IELTS năm 2023, bài viết luận của sinh viên với tối thiểu 250 từ được đánh giá theo bốn tiêu chí có trọng số ngang nhau, gồm: (1) đáp ứng yêu cầu của nhiệm vụ; (2) mạch lạc và liên kết; (3) nguồn từ vựng; và (4) sự đa dạng và chính xác của ngữ pháp. Trong đó, tiêu chí mạch lạc và liên kết đánh giá mức độ tổ chức và phát triển logic của bài viết, bao gồm việc sắp xếp ý tưởng, phân đoạn văn bản và sử dụng phù hợp các phương tiện liên kết. Ở mức yêu cầu cao, thí sinh được kì vọng sử dụng các phương tiện liên kết một cách linh hoạt, đa dạng và tự nhiên, thay vì phụ thuộc vào việc lặp lại một số từ nối cơ bản.

Do chiếm 25% tổng điểm, tiêu chí mạch lạc và liên kết đã trở thành một trong những khía cạnh được quan tâm trong nghiên cứu về viết IELTS Task 2 tại Việt Nam. Nguyen và Nguyen (2022) khảo sát 205 người học IELTS tại hai trung tâm ngoại ngữ ở Thành phố Hồ Chí Minh nhằm xác định những vấn đề và nguyên nhân liên quan đến kết quả bài viết Task 2. Tương tự, Nguyen, Nguyen và Phan (2022), khảo sát nhận thức của người học IELTS tại Việt Nam về viết IELTS Task 2 và đặc điểm vốn từ trong bài viết thực tế. Nhìn chung, các nghiên cứu này cho thấy việc sử dụng phương tiện liên kết phù hợp và linh hoạt đóng vai trò quan trọng đối với chất lượng bài viết và sự phân hóa giữa các mức điểm. Đây cũng là cơ sở để lựa chọn phương tiện liên kết làm đối tượng khảo sát trong nghiên cứu này.

3. Phương pháp nghiên cứu

Nghiên cứu sử dụng kết hợp phương pháp nghiên cứu mô tả, phương pháp phân tích ngữ liệu, phương pháp chọn mẫu thuận tiện, phương pháp mã hóa dữ liệu, phân tích định tính và phân tích định lượng nhằm mô tả tần suất và cách sử dụng các phương tiện liên kết trong bài viết IELTS Task 2 của sinh viên không chuyên năm thứ nhất tại Trường Đại học Hà Nội.

3.1. Nguồn ngữ liệu nghiên cứu

Ngữ liệu khảo sát ban đầu gồm 105 bài viết IELTS Task 2 của sinh viên năm thứ nhất không chuyên, được thu thập bằng phương pháp chọn mẫu thuận tiện từ các lớp học phần kĩ năng viết do nhóm nghiên cứu trực tiếp giảng dạy. Sau khi đối chiếu và loại hai bài không có văn bản gốc, 103 bài được đưa vào phân tích, với tổng cộng 32.102 từ; độ dài trung bình là 311,67 từ/bài (SD = 84,72), dao động từ 181 đến 778 từ. Phương pháp chọn mẫu thuận tiện phù hợp với mục tiêu mô tả đặc điểm ngôn ngữ của một nhóm sinh viên cụ thể này, song cũng có hạn chế đó là chưa thể khái quát hóa cho toàn bộ sinh viên không chuyên năm thứ nhất tại Trường Đại học Hà Nội nói riêng hay sinh viên không chuyên tại Việt Nam nói chung.

3.2. Khung lí thuyết và hệ thống mã hóa nguồn ngữ liệu

Các phương tiện liên kết trong nguồn ngữ liệu khảo sát được chúng tôi xác định và mã hóa theo khung năm phạm trù của Halliday và Hasan (1976) và Halliday và Matthiessen (2014) gồm quy chiếu, phép thế, tỉnh lược, liên từ và liên kết từ vựng. Các phạm trù này được tiếp tục phân chia thành những tiểu loại tương ứng, cụ thể: (1) quy chiếu gồm quy chiếu cá nhân, quy chiếu chỉ định và quy chiếu so sánh; (2) phép thế gồm thế danh từ, thế động từ và thế mệnh đề; và (3) liên từ gồm liên từ bổ sung, liên từ tương phản, liên từ nguyên nhân, và liên từ thời gian.

Tần suất sử dụng được xác định bằng phương pháp gán nhãn tự động dựa trên danh sách từ, với nguyên tắc khớp theo ranh giới từ và ưu tiên các cụm từ dài hơn nhằm hạn chế hiện tượng đếm trùng giữa các tiểu loại. Một số điều chỉnh về phạm vi mã hóa được áp dụng, chẳng hạn như loại *the* khỏi quy chiếu chỉ định do tần suất xuất hiện quá cao và loại bỏ các đại từ (*I, we, you)* ra khỏi quy chiếu cá nhân vì chúng chủ yếu mang chức năng ngoại chỉ trong văn nghị luận.

Do đặc điểm của ngữ liệu nghiên cứu và công cụ khảo sát, phạm trù tỉnh lược không được đưa vào phân tích định lượng trong kết quả nghiên cứu của chúng tôi vì không thể được xác định trực tiếp thông qua phương thức khớp từ hoặc cụm từ mà cần dựa vào việc khôi phục thành phần bị tỉnh lược từ ngữ cảnh. Đối với liên kết từ vựng, công cụ phân tích tự động chỉ tập trung vào phép lặp lại ở dạng lặp lại nguyên văn và tỉ lệ loại/ lượt từ (*Type-Token Ratio - TTR*) được sử dụng như một chỉ số gián tiếp về mức độ đa dạng của từ vựng. Các quan hệ ngữ nghĩa như đồng nghĩa, trái nghĩa, từ khái quát và kết hợp từ chưa được mã hóa tự động do đòi hỏi cần thêm sự phán đoán về ngữ nghĩa. Vì vậy, các kết quả về liên kết từ vựng được sử dụng trong nghiên cứu này chỉ là những số liệu ban đầu và cần được bổ sung bằng mã hóa thủ công trong các nghiên cứu tiếp theo.

3.3. Cách thức phân tích ngữ liệu

Phân tích định tính được chúng tôi sử dụng để diễn giải chức năng của các phương tiện liên kết trong ngữ cảnh cụ thể và minh họa các kiểu sử dụng liên từ tiêu biểu. Phân tích định lượng trong nghiên cứu được thực hiện bằng thống kê mô tả, bao gồm tần số, tỉ lệ phần trăm, giá trị trung bình, độ lệch chuẩn và tần suất chuẩn hóa trên 100 từ nhằm tạo cơ sở so sánh giữa các bài viết có độ dài khác nhau. Hệ số tương quan Pearson được sử dụng để khảo sát mối quan hệ giữa liên kết ngữ pháp và chỉ số đa dạng từ vựng. Việc kết hợp phân tích định tính và định lượng cho phép nghiên cứu vừa mô tả đặc điểm phân bố của các phương tiện liên kết, vừa diễn giải chức năng và ý nghĩa của các khuôn mẫu sử dụng được nhận diện trong ngữ liệu khảo sát.

4. Kết quả nghiên cứu và thảo luận

4.1. Tần suất sử dụng phương tiện liên kết

Kết quả mã hóa tự động trong tổng số 103 bài viết (tương ứng 32.102 từ) có 5.160 phương tiện liên kết ngữ pháp, tương ứng với mật độ trung bình 16,07 phương tiện/100 từ tính trên toàn ngữ liệu được khảo sát (cụ thể 16,10 phương tiện/ 100 từ nếu tính theo trung bình cộng của từng bài, độ lệch chuẩn = 2,41; dao động từ 11,31 đến 21,69 phương tiện/ 100 từ giữa các bài viết). Đây là những điểm rất đáng chú ý, cho thấy trung bình cứ khoảng sáu từ trong bài viết của sinh viên lại xuất hiện một phương tiện liên kết ngữ pháp. Kết quả nghiên cứu này được chúng tôi trình bày cụ thể dưới bảng sau:

Bảng 1*. Tần suất sử dụng phương tiện liên kết ngữ pháp theo từng phạm trù*

Phạm trù	Tổng số lần xuất hiện	TB/ bài (SD)	Tỉ lệ % trong LK ngữ pháp	TB/ 100 từ
Quy chiếu	2.634	25,57 (9,51)	51,05%	8,21
Liên từ	2.411	23,41 (7,61)	46,72%	7,51
Phép thế	115	1,12 (0,93)	2,23%	0,36
TỔNG	**5.160**	**50,10 (15,66)**	**100%**	**16,07**

Kết quả Bảng 1 cho thấy sự phân bố không đồng đều giữa ba phạm trù liên kết ngữ pháp được mã hóa tự động, cụ thể: (1) quy chiếu là phạm trù được sử dụng phổ biến nhất, chiếm tỉ lệ 51,05% (trung bình 25,57 lượt xuất hiện mỗi bài); (2) liên từ đứng ở vị trí thứ hai với tỉ lệ 46,72% (trung bình 23,41 lượt/ bài); (3) trong khi phép thế chỉ chiếm tỉ lệ rất nhỏ 2,23% (trung bình 1,12 lượt/ bài). Cũng từ bảng tổng hợp trên, chúng ta có thể thấy tổng số phương tiện liên kết ngữ pháp của quy chiếu và liên từ chiếm tỉ lệ rất lớn (tương đương 97,77%), gấp hơn 40 lần với số phương tiện liên kết của phép thế

(2,23%). Điều này cho thấy nhóm sinh viên này có xu hướng sử dụng chủ yếu hai phương thức liên kết quy chiếu và liên từ nhằm đảm bảo sự tương đối an toàn và dễ sử dụng trong quá trình viết. Bởi lẽ, quy chiếu và liên từ là hai phương thức không đòi hỏi về cao mặt cấu trúc ngữ pháp.

Kết quả tổng hợp từ Bảng 1 cũng cho thấy việc sử dụng phương tiện liên kết quy chiếu và liên từ chiếm ưu thế rõ rệt so với phép thế này khá tương đồng với những phát hiện của Nguyen và cộng sự (2023), song lại có phần khác biệt với kết quả nghiên cứu của Liu và Braine (2005) nơi liên kết từ vựng được ghi nhận chiếm ưu thế hơn liên kết ngữ pháp.

4.2. Tần suất và phân bố các tiểu phạm trù phương tiện liên kết ngữ pháp

Kết quả nghiên cứu tần suất và sự phân bố các tiểu phạm trù liên kết ngữ pháp được chúng tôi tổng hợp dưới bảng sau:

***Bảng 2**. Tần suất và sự phân bố các tiểu phạm trù liên kết ngữ pháp*

Phạm trù	Tiểu phạm trù	Tổng số lần	TB/bài	Tỉ lệ (%) trong từng phạm trù	Tổng
Quy chiếu	Ngôi ba	1.121	10,88	42,56%	21,72%
	Chỉ định	976	9,48	37,05%	18,91%
	So sánh	537	5,21	20,39%	10,41%
Phép thế	Danh từ	114	1,11	99,13%	2,21%
	Động từ	0	0,00	0,00%	0,00%
	Mệnh đề	1	0,01	0,87%	0,02%
Liên từ	Bổ sung	1.497	14,53	62,09%	29,01%
	Tương phản	377	3,66	15,64%	7,31%
	Nguyên nhân	234	2,27	9,71%	4,53%
	Thời gian	303	2,94	12,57%	5,87%

Đi sâu vào từng tiểu phạm trù, Bảng 2 cho thấy bức tranh phân bố chi tiết hơn, cụ thể:

Trong phạm trù quy chiếu, tiểu loại phạm trù quy chiếu ngôi ba có số lượng sử dụng nhiều nhất (chiếm 42,56%), tiếp theo là tiểu loại quy chiếu chỉ định (chiếm 37,05%) và tiểu loại quy chiếu so sánh đứng cuối cùng (chiếm 20,39%). Trong khi đó ở phạm trù liên từ, tiểu phạm trù liên từ bổ sung có số lượng sử dụng chiếm ưu thế rõ rệt (tương đương 62,09%). Ba tiểu loại phạm trù liên từ còn lại (tương phản, nguyên nhân và thời gian) có tỉ lệ sử dụng khá tương đồng với nhau (lần lượt là 15,64%, 9,71%, và 12,57%).

Đáng chú ý nhất là sự phân bố cực kì chênh lệch trong phạm trù phép thế. Từ bảng tổng hợp chúng ta có thể thấy, tiểu phạm trù thế danh từ có số lượng sử dụng nhiều nhất (chiếm 99,13%); trong khi, tiểu loại thế mệnh đề chỉ xuất hiện đúng duy một lần trong toàn bộ 103 bài viết, và tiểu loại thế động từ hoàn toàn không xuất hiện. Kết quả này cho thấy trong phạm vi ngữ liệu khảo sát, phạm trù phép thế vốn đòi hỏi người viết vừa nắm vững cấu trúc hình thức vừa phải theo dõi sát sao thực thể được thay thế qua nhiều mệnh đề gần như chỉ được sinh viên khai thác ở dạng đơn giản nhất, trong khi hai tiểu loại phức tạp hơn về mặt cú pháp gần như hoàn toàn vắng bóng trong ngữ liệu khảo sát. Đây là một phát hiện có ý nghĩa sư phạm đáng lưu ý, gợi ý rằng phép thế động từ và thế mệnh đề có thể là những mục tiêu giảng dạy cần được chú trọng hơn đối với nhóm sinh viên này.

4.3. So sánh liên kết ngữ pháp với liên kết từ vựng

Kết quả so sánh liên kết ngữ pháp với liên kết từ vựng được chúng tôi trình bày cụ thể dưới bảng sau:

Bảng 3: *Liên kết ngữ pháp so sánh với liên kết từ vựng*

Chỉ số	Giá trị
Tổng số từ trong ngữ liệu	32.102
Tổng số phương tiện liên kết ngữ pháp	5.160
Liên kết ngữ pháp/ 100 từ (toàn ngữ liệu)	16,07
TB liên kết ngữ pháp/ 100 từ (theo từng bài, SD)	16,10 (2,41)
Tổng lượt thực từ lặp lại	8.108 (TB 78,72/bài)
TB Tỷ lệ loại/ lượt từ - TTR (SD)	0,571 (0,060)
Hệ số tương quan r (LK ngữ pháp/100 từ và TTR)	-0,186

Như đã trình bày ở phần trên, liên kết từ vựng được mã hóa tự động trong nghiên cứu này giới hạn ở phép lặp lại nguyên văn thực từ và tỉ lệ loại/ lượt từ (TTR) như một chỉ số gián tiếp nhằm phản ánh mức độ đa dạng từ vựng (không bao gồm các cơ chế liên kết từ vựng đòi hỏi phán đoán ngữ nghĩa như từ đồng nghĩa, từ trái nghĩa, từ khái quát hay kết hợp từ). Kết quả từ bảng 3 cho thấy tổng cộng có 8.108 lượt thực từ được lặp lại trong toàn ngữ liệu (trung bình 78,72 lượt/ bài); trong khi đó, TTR trung bình đạt 0,571 (độ lệch chuẩn = 0,060, dao động từ 0,405 đến 0,715 giữa các bài viết) nghĩa là trung bình khoảng 57% số lượt từ trong mỗi bài viết là các kiểu từ khác nhau, phần còn lại là các lượt lặp lại của những từ loại đã xuất hiện.

Khi đối chiếu mức độ liên kết ngữ pháp (tính trên 100 từ) với TTR ở cấp độ từng bài viết, kết quả nghiên cứu chỉ ra hệ số tương quan Pearson r = -0,186. Đây là một mối tương quan nghịch ở mức độ yếu. Xét theo khía cạnh này thì hệ số này phần nào cho thấy sinh viên trong đối tượng khảo sát có xu hướng nghiêng về việc sử dụng liên từ và quy chiếu trong liên kết ngữ pháp hơn là khai thác một vốn từ vựng đa dạng (liên kết từ vựng thông qua TTR) để tạo lập tính liên kết cho bài viết; hay nói cách khác, liên kết ngữ pháp được sử dụng nhiều hơn so với sự đa dạng từ vựng trong cách sinh viên không chuyên năm thứ nhất dùng để tạo lập kết cấu của văn bản.

Tuy nhiên, kết quả này cũng cần được nhìn nhận một cách rất thận trọng vì: (1) đây là một hệ số tương quan, không hàm ý quan hệ nhân quả: số liệu từ kết quả khảo sát không khẳng định việc sử dụng nhiều liên kết ngữ pháp là nguyên nhân trực tiếp khiến cho việc sử dụng từ vựng kém đa dạng hơn, hay ngược lại; và (2) quan trọng hơn, TTR vốn được xem xét là một chỉ số nhạy cảm đối với độ dài văn bản nghĩa là bài viết càng dài, thì tỷ lệ xác suất một từ đã dùng bị lặp lại càng cao, khiến TTR có xu hướng giảm một cách cơ học chứ không nhất thiết phản ánh sự nghèo nàn về việc sử dụng vốn từ. Trong ngữ liệu nghiên cứu của chúng tôi, độ dài bài viết cũng có mối tương quan khá rõ với TTR (r = -0,573), cho thấy một phần biến thiên của TTR giữa các bài viết có thể xuất phát từ chênh lệch độ dài hơn là những khác biệt thực chất về năng lực sử dụng vốn từ vựng của sinh viên.

5. Kết luận

Nghiên cứu tiến hành khảo sát tần suất và cách thức sử dụng các phương tiện liên kết trong bài viết IELTS Task 2 của 103 bài của sinh viên không chuyên năm thứ nhất Trường Đại học Hà Nội, dựa trên khung lí thuyết năm phạm trù liên kết của Halliday và Hasan (1976) và Halliday và Matthiessen (2014). Kết quả cho thấy liên kết ngữ pháp - đặc biệt là quy chiếu và liên từ chiếm ưu thế rõ rệt trong ngữ liệu (16,07 phương tiện/100 từ), trong khi phép thế gần như không được đề cập, nhất là ở hai tiểu loại thế động từ và thế mệnh đề; đồng thời, mức độ liên kết ngữ pháp có sự tương quan nghịch ở mức độ yếu với chỉ số đa dạng từ vựng (TTR).

Về mặt lí luận, kết quả nghiên cứu góp phần bổ sung dữ liệu thực chứng cho một nhóm đối tượng sinh viên không chuyên năm thứ nhất, vốn chưa nhận được nhiều sự quan tâm và khảo sát trong các

nghiên cứu về việc sử dụng phương tiện liên kết trong quá trình học kỹ năng viết IELTS Task 2 tại Việt Nam, nơi mà các nghiên cứu trước đây chủ yếu tập trung vào sinh viên chuyên ngành tiếng Anh hoặc tại các trung tâm ngoại ngữ. Về mặt thực tiễn, những phát hiện về sự mất cân đối giữa các tiểu phạm trù liên kết ngữ pháp, cũng như giữa liên kết ngữ pháp và sự đa dạng trong việc sử dụng vốn từ vựng đã cung cấp nhiều luận cứ khoa học không chỉ đối với cơ sở giáo dục mà còn cho giảng viên trong việc điều chỉnh nội dung giảng dạy kĩ năng viết học thuật tại Trường Đại học Hà Nội theo hướng bám sát hơn với đặc điểm ngôn ngữ thực tế của sinh viên, thay vì chỉ dựa trên kinh nghiệm giảng dạy.

Kết quả nghiên cứu cũng chỉ ra một số hạn chế nhất định. Thứ nhất, do giới hạn của công cụ mã hóa tự động, nghiên cứu chưa thể xác định và phân tích đầy đủ phạm trù tỉnh lược và các dạng liên kết từ vựng; do cần đòi hỏi sự phán đoán về ngữ nghĩa, bao gồm quan hệ đồng nghĩa, trái nghĩa, từ khái quát và kết hợp từ. Việc bổ sung mã hóa thủ công đối với các phạm trù này sẽ được chúng tôi triển khai trong các nghiên cứu tiếp theo nhằm nhận diện đầy đủ hơn các phương tiện liên kết trong ngữ liệu. Bên cạnh đó, việc sử dụng phương pháp chọn mẫu thuận tiện cũng dẫn đến nhiều hạn chế trong việc khái quát hóa kết quả; do vậy, trong nghiên cứu tiếp theo chúng tôi sẽ cân nhắc mở rộng cỡ mẫu và áp dụng phương pháp chọn mẫu ngẫu nhiên hoặc có tính đại diện cao hơn nhằm tăng tính đại diện của mẫu nghiên cứu, nâng cao độ tin cậy của kết quả và cải thiện khả năng khái quát hóa các phát hiện cho ngữ liệu nghiên cứu. Cuối cùng, nghiên cứu sử dụng tỷ lệ loại/lượt từ (*Type-Token Ratio - TTR*) để mô tả mức độ đa dạng từ vựng, trong khi chỉ số này lại chịu ảnh hưởng đáng kể bởi độ dài văn bản. Vì vậy, chúng tôi sẽ cân nhắc sử dụng những chỉ số đa dạng từ vựng ít phụ thuộc hơn vào độ dài văn bản nhằm nâng cao độ tin cậy và khả năng so sánh của kết quả trong những nghiên cứu tiếp theo.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Diệp Quang Ban (2006), *Văn bản và liên kết trong tiếng Việt*, Nxb Giáo dục, Hà Nội.
2. Nguyễn Thị Giáp (2009), *Dụng học Việt ngữ*, Nxb ĐHQG Hà Nội.

Tiếng Anh

3. Bui, H. P. (2022), "Vietnamese EFL students' use and misconceptions of cohesive devices in writing", *SAGE Open, 12*(3). https://doi.org/10.1177/21582440221126993.
4. Bui, H. P., Nguyen, N. Q., Nguyen, T. L., & Nguyen, T. V. (2021), "A cross-linguistic approach to analysing cohesive devices in expository writing by Asian EFL teachers", *3L: Language, Linguistics, Literature, 27*(2), 16-30.
5. De Beaugrande, R., & Dressler, W. U. (1981), *Introduction to text linguistics*. Longman.
6. Halliday, M. A. K., & Hasan, R. (1976), *Cohesion in English*. Longman.
7. Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014), *Halliday's introduction to functional grammar* (4th ed.). Routledge.
8. Hung, D. M., & Thu, V. T. A. (2014), "Vietnamese learners' attention and use of cohesive devices in English essay writing at Dong Thap University", *Asian Journal of Educational Research, 2*(2), 1-14.
9. IELTS (British Council, IDP: IELTS Australia, & Cambridge Assessment English). (2023), *IELTS writing key assessment criteria*. https://ielts.org/cdn/Guides/ielts-writing-key-assessment-criteria.pdf.
10. Liu, M., & Braine, G. (2005), "Cohesive features in argumentative writing produced by Chinese undergraduates", *System, 33*(4), pp.623-636. https://doi.org/10.1016/j.system.2005.02.002.
11. Nguyen, B. T. T., Nguyen, T. T. B., & Phan, Q. N. (2022), *Vietnamese IELTS learners' perceptions of IELTS Writing Task 2 and their performance-based lexical resource* (IELTS Research Reports Online Series, No. 5/22). British Council, Cambridge Assessment English and IDP: IELTS Australia.

12. 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. https://doi.org/10.54855/ijte.222111.
13. Nguyen, Q. T., Nguyen, T. B. T., & Nguyen, V. Q. N. (2023), "Use of cohesive devices in paragraph writing by EFL students at English language centers in Vietnam", *VNU Journal of Foreign Studies, 39*(3), pp.152-169. https://doi.org/10.63023/2525-2445/jfs.ulis.5099.
14. Quirk, R., Greenbaum, S., Leech, G., Svartvik, J., & Crystal, D. (1985), *A comprehensive grammar of the English language* (Vol. 397), London: Longman.
15. Tran, T. B. (2005), *An error analysis on the use of cohesive devices in writing by freshmen majoring in English at Thang Long University*, Vietnam National University, Hanoi.

A study on the frequency and use of cohesive devices in IELTS Task 2 essays by first-year non-English-major students at Hanoi University

Abstract: The study investigates the frequency and use of cohesive devices in IELTS Task 2 essays produced by first-year non-English-major students at Hanoi University, drawing on the theoretical framework of the five categories of cohesion proposed by Halliday and Hasan (1976) and Halliday and Matthiessen (2014). The research corpus comprises 103 essays (totaling 32.102 words) collected through convenience sampling and analyzed using a corpus-based approach that combines quantitative and qualitative analyses. The findings indicate that the corpus contains a total of 5.160 grammatical cohesive devices, with reference accounting for the highest proportion (51,05%), followed by conjunction (46,72%) and substitution (2,23%). At the subcategory level, third-person reference and additive conjunctions are the two most frequently used cohesive devices. The findings not only contribute to a clearer understanding of students' use of cohesive devices but also suggest several pedagogical directions for improving the quality of academic writing instruction at Hanoi University.

Key words: Cohesion; cohesive devices; frequency of use; first -year non-English-major students; IELTS Task 2 essay.