

Ứng dụng phần mềm AntConc vào phân tích lỗi sai trong giảng dạy tiếng Trung Quốc- Nghiên cứu trường hợp liên từ “和(và)”

Nguyễn Thị Thúy Hòa*

* ThS. Đại học Kinh tế quốc dân

Received: 25/10/2024; Accepted: 31/10/2024; Published: 7/11/2024

Abstract: Currently, the application of corpora in language research is increasingly popular. This article builds an intermediate-level corpus from the writings of students studying Chinese at the National Economics University, who are learners at the intermediate level. Then use Antconc software (version 4.2.0) for analysis, helping to analyze errors in the article more effectively.

Keywords: Linguistic corpus, Antconc software, Chinese language teaching

1. Mở đầu

Kho ngữ liệu là một phương pháp nghiên cứu ngôn ngữ dựa trên việc thu thập và phân tích các tập hợp dữ liệu ngôn ngữ lớn, gọi là kho ngữ liệu. Kho ngữ liệu có thể bao gồm bản viết hoặc nói và được sử dụng để nghiên cứu các hiện tượng ngôn ngữ học như tần suất từ vựng, các kết hợp từ, cấu trúc ngữ pháp, và các lỗi thường gặp. Phần mềm như Antconc có thể được sử dụng để phân tích kho ngữ liệu, giúp nhận diện các mẫu và xu hướng trong việc sử dụng ngôn ngữ, từ đó dễ dàng phát hiện những lỗi sai nổi trội, cơ bản và mang tính đặc thù.

2. Nội dung nghiên cứu

2.1. Giới thiệu về phần mềm AntConc phiên bản 4.2.0

Phần mềm AntConc 4.2.0 cung cấp các chức năng như tìm kiếm từ (Frequency), danh sách từ (Word), cụm từ (Cluster), kết hợp từ (Collocate), phân tích ngữ cảnh (Plot) và tạo danh sách từ khóa (Keyword). AntConc là phần mềm được thiết kế và phát triển bởi Laurence, với ưu điểm lớn nhất là thao tác đơn giản, thu hút sự quan tâm của nhiều nhà nghiên cứu. Các chức năng chính của AntConc 4.2.0 bao gồm:

- Tìm kiếm từ (Concordance): Có thể trích xuất tất cả các cụm từ hoặc câu chứa từ cần tìm kiếm từ một kho ngữ liệu lớn. Kết quả có thể được xuất và lưu dưới dạng tệp văn bản (text), thuận tiện cho việc phân tích từ vựng.

- Thống kê tần suất từ (Frequency) và tạo danh sách từ (Word): Hai chức năng này bổ sung cho nhau, cho phép trích xuất tất cả từ vựng trong văn bản và tính toán số lần xuất hiện của từng từ. Kết quả có thể được sắp xếp từ cao đến thấp hoặc ngược lại, đồng thời tạo biểu đồ để tham khảo.

- Cụm từ (Cluster): Chức năng này giúp tìm kiếm các từ liên tiếp hoặc đoạn văn bản chứa một từ hoặc cụm từ nhất định. Thông qua tính toán, chức năng này có thể tự động tìm ra các cụm từ có ý nghĩa bao gồm hai từ hoặc nhiều từ lặp lại cùng một hình thức.

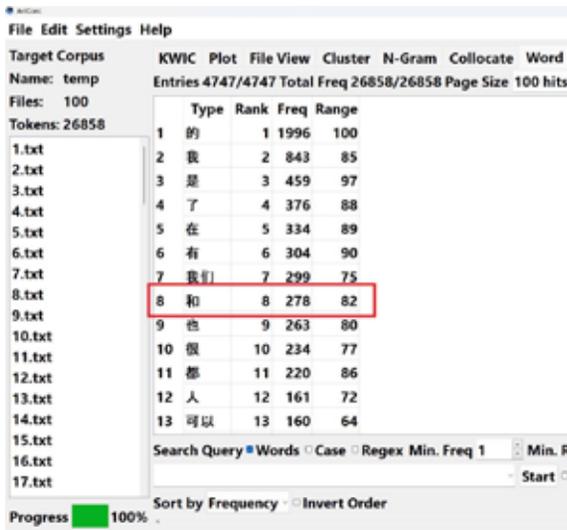
- Kết hợp từ (Collocate): Chức năng này cho phép tìm ra các từ đứng trước và sau một từ cụ thể, đồng thời đánh dấu vị trí của chúng trong câu. Bản chất của chức năng này là phân tích sự kết hợp từ trong văn bản.

- Chức năng từ khóa (Keyword): Đây là một chức năng rất quan trọng, có thể tìm kiếm các từ khóa trong ngữ liệu nghiên cứu (Study Corpus) so với ngữ liệu tham chiếu (Reference Corpus). Nó cũng tính toán tính chất đặc thù (Keyness) của từng từ khóa và sắp xếp từ cao đến thấp. Danh sách từ khóa (Keyword List) được tạo ra có vai trò quan trọng trong việc phân tích văn bản.

2.2. Ứng dụng phần mềm AntConc vào việc phân tích lỗi sai trong bài viết tiếng Trung Quốc

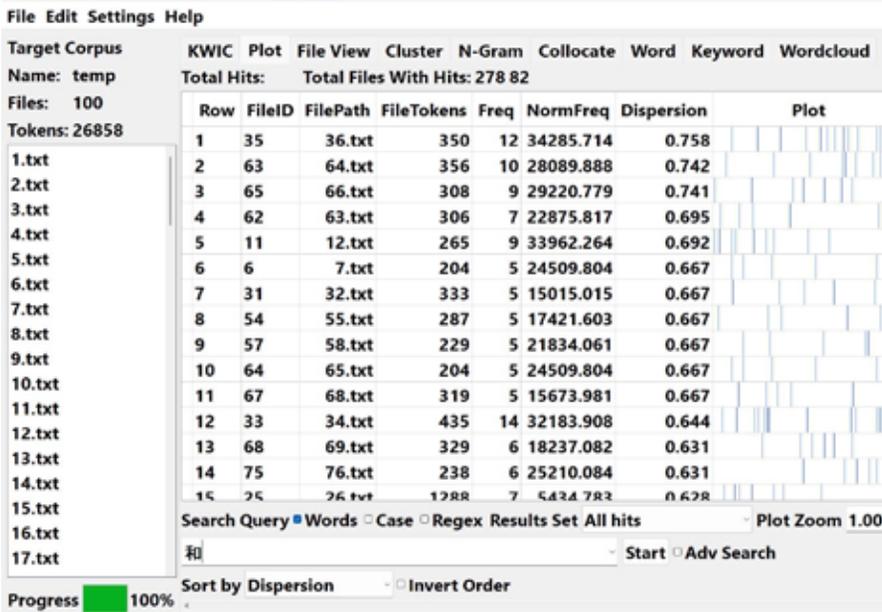
2.2.1. Phân tích lỗi sai bằng phần mềm AntConc

Chúng tôi sau khi thu thập bài viết của SV Đại học Kinh tế Quốc dân Việt Nam ở giai đoạn viết tiếng Trung Quốc trình độ trung cấp, định dạng file Word, tổng cộng 100 bài, mỗi tài liệu được đặt tên (Tên file) từ 1 đến 100 theo số thứ tự. Sử dụng phần mềm SegmentAnt để phân tách từ, sau đó chuyển mỗi văn bản sang mã ký tự UTF-8 và định dạng txt. Tiếp theo, các tập tin được nhập vào phần mềm AntConc 4.2.0, tạo thành một kho ngữ liệu có quy mô 26.858 từ. Ngữ liệu này là ngữ liệu bài viết tiếng Trung Quốc- trình độ trung cấp của SV Đại học Kinh tế Quốc dân, hay còn gọi là ngữ liệu trung gian.



Ảnh 2.1. Bảng Word trên phần mềm AntConc

Từ ảnh 2.1 có thể thấy quy mô ngữ liệu là 100 bài viết (Files), với tổng số từ là 26.858 (tokens). Nhấp vào Word có thể thống kê được 100 từ có tần suất xuất hiện cao nhất, xếp hạng từ cao đến thấp. 10 từ có tần suất cao nhất bao gồm: “的, 我, 是, 了, 在, 和, 我们, 有, 也, 都”. Trong KWIC có thể quan sát sơ bộ cách sử dụng của từng từ, chúng tôi phát hiện rằng tỷ lệ sai sót khi sử dụng liên từ “和(và)” là cao nhất.



Ảnh 2.2. Bảng Plot trên phần mềm AntConc

Tiến hành kiểm tra thêm số lần xuất hiện của từ “和(và)” trong văn bản (ảnh 2.2), chúng tôi sử dụng chức năng chỉ mục (Index Slot), nhấp vào Plot để có kết quả: trong 82 văn bản (trong tổng số 100 văn bản của ngữ liệu), từ “和(và)” xuất hiện 278 lần. Có thể

thấy tần suất sử dụng từ “和(và)” của trong bài viết của SV là khá cao, điều này có thể coi là một trọng điểm lỗi sai của SV.

Để phân tích lỗi sử dụng từ “和(và)” trong bài viết của SV, chúng tôi đã sao chép 278 kết quả từ KWIC và nhập vào một tệp Excel. Sau đó, bắt đầu tiến hành phân tích từng kết quả.

2.2.2. Kết quả phân tích lỗi sai

Sau khi phân tích, chúng tôi phát hiện ra một số lỗi phổ biến khi sử dụng liên từ này bao gồm:

(a) Lỗi sử dụng liên từ “和(và)” nối các tính từ, cụm tính từ đóng vai trò thành phần vị ngữ

Lu Qing He (2006) trong cuốn “Ngữ pháp giảng dạy tiếng Trung Quốc cho người nước ngoài” chỉ ra rằng, liên từ “和(và)” có thể kết nối các tính từ và cụm tính từ, đóng vai trò là thành phần phi vị ngữ (chủ ngữ, tân ngữ hoặc định ngữ), thường có các thành phần có liên quan chung. Qua việc khảo sát bảng từ KWIC, phát hiện SV khi viết thường mắc phải lỗi sử dụng liên từ “和(và)” nối các tính từ, cụm tính từ đóng vai trò thành phần vị ngữ. Ví dụ:

(1) *北江人非常善良和好客. (Mã văn bản: 94.txt)

(2) *这里的环境很安静和安全. (Mã văn bản: 59.txt)

Ở ví dụ (1) SV đã dùng sai liên từ “和(và)” để nối tính từ đảm nhiệm vị ngữ. Trong ví dụ này hai từ “善良” (luong thiện) và “好客” (hiếu khách) đều là tính từ đảm nhiệm vai trò vị ngữ, cho nên không thể dùng “和(và)” để nối. Nên sửa thành: “北江人又善良又好客”. Ví dụ (2) hai từ “安静” (yên tĩnh) và “安全” (an toàn) cũng là tính từ đảm nhiệm vị ngữ cho nên cũng không thể dùng “和(và)” liên kết, nên sửa thành: “这里

的环境又安静又安全。”

(b) Lỗi sử dụng liên từ “和(và)” nối động từ, động từ hoặc cụm động từ đóng vai trò thành phần vị ngữ

Liên từ “和(và)” có thể nối động từ và cụm động từ, đóng vai trò thành phần phi vị ngữ. Tuy vậy qua khảo sát phát hiện SV do ảnh hưởng bởi tiếng mẹ đẻ

nên đã dùng để nói động từ, cụm động từ đóng vai trò vị ngữ. Ví dụ:

(3)*这个办法以控制和快速解决交通拥堵问题。(Mã văn bản: 54.txt)

Ví dụ (3) hai cụm động từ “控制” (kiểm soát) và “快速解决” (nhanh chóng giải quyết) này không phải là quan hệ đẳng lập, mà là quan hệ tăng tiến, cho nên câu này nên đổi là: “这个办法可以使其得到控制而快速解决交通拥堵问题。”

(4)*这门课很难, 学习的时候, 我们要记住和写出正确的句子结构。(Mã văn bản: 83.txt)

Ví dụ (4) hai cụm từ “记住” (ghi nhớ) và “写出正确的句子结构” (Viết ra cấu trúc chính xác của câu) không phải là quan hệ đẳng lập, cũng không có thành phần chung, từ “记住” phía sau thiếu tân ngữ. Cả câu nên chuyên thành: “这门课很难, 学习的时候, 我们要记住生词, 也要写出正确的句子结构。”

(c) Lỗi sử dụng sai nghĩa

Liên từ “和(và)” dùng để nối các từ hoặc cụm từ, biểu thị mối quan hệ đẳng lập. Qua khảo sát, phát hiện SV khi sử dụng “和/và” để nối từ hoặc cụm từ, do hiểu sai nghĩa, dẫn đến những lỗi sử dụng trong quá trình viết.

(5)*每个地方都有属于自己的特色和文化。(Mã văn bản: 3.txt)

Ví dụ (5) hai cụm từ “自己的特色”(đặc sắc của mình) và “文化”(văn hóa) hai thành phần này không bình đẳng về mặt ngữ nghĩa, cho nên không thể dùng “和(và)” liên kết. Nên sửa thành: “每个城市都拥有自己的特色文化。”

(6)*要克服这种局面, 首先需要有很多组织和每一个人来共同保护。(Mã văn bản: 41.txt)

Ví dụ (6) hai cụm từ “很多组织” (rất nhiều tổ chức) và “一个人” (một người) không bình đẳng về mặt nghĩa. Nên sửa thành: “要克服这种局面, 首先需要有很多组织以及每一个人来共同保护。”

(7)*道路使用者素质差和不遵守交通规则。(Mã văn bản: 43.txt)

Ví dụ (7) hai cụm từ “道路使用者素质差” (ý thức người tham gia giao thông kém) và “不遵守交通规则” (không tuân thủ luật giao thông) bị hiểu sai về nghĩa, không phải là quan hệ đẳng lập, mà là quan hệ nhân quả. Câu này nên sửa thành: “因为道路使用者素质差, 所以不遵守交通规则。”

2.2.3. Biện pháp khắc phục lỗi sai

Về cách sử dụng liên từ “和/và”, khi chữa bài viết, GV có thể tìm ra một số câu sai khi sử dụng “和(và)” và cùng SV phân tích. Có thể sử dụng phần mềm Antconc, nhập văn bản viết của SV, từ đó dễ dàng tìm ra nhiều ví dụ câu. Phần mềm này cho phép

xác định các lỗi sử dụng liên từ nhanh chóng và hiệu quả. Qua đó, hướng dẫn học sinh đâu là cách diễn đạt đúng, đâu là sai.

Khuyến khích SV cùng nhau xây dựng một ngữ liệu viết bài, nếu muốn tìm cách sử dụng của một từ, có thể nhập từ đó vào phần mềm Antconc, nhấp vào bảng từ KWIC để tham khảo nhiều ví dụ câu. Sau đó, trong bảng từ Word, phân tích một số từ có tần suất cao và tìm hiểu cách sử dụng của chúng, bởi vì các từ có tần suất cao thường cần sử dụng nhiều trong bài viết.

Giảng viên cần tạo cơ hội cho SV thực hành viết thường xuyên, đồng thời cung cấp phản hồi chi tiết về các lỗi sai, đặc biệt là cách sử dụng liên từ “和”. Giảng viên có thể thiết kế các bài tập tập trung vào việc sử dụng liên từ trong nhiều loại câu khác nhau, từ đơn giản đến phức tạp, để SV dần quen và hiểu sâu cách dùng chính xác.

3. Kết luận

Kho ngữ liệu trung gian có thể được ứng dụng trong giảng dạy tiếng Trung Quốc, đặc biệt là môn Việt. Phần mềm AntConc sẽ giúp SV có cái nhìn tổng quát hơn về những lỗi sai trọng điểm. Bài báo này cung cấp cho các GV giảng dạy tiếng Trung Quốc một giải pháp nâng cao hiệu quả môn học hơn.

Ngoài ra, bài viết này còn chỉ ra các lỗi sai phổ biến trong cách sử dụng liên từ “和” của SV học tiếng Trung Quốc giai đoạn trung cấp tại Đại học Kinh tế quốc dân, và cung cấp những biện pháp khắc phục cụ thể. Với sự hỗ trợ của công nghệ, giảng viên có thể nâng cao hiệu quả giảng dạy, giúp SV cải thiện đáng kể kỹ năng viết, đồng thời hiểu sâu hơn về ngữ pháp tiếng Trung Quốc trong ngữ cảnh thực tế. Đây là một giải pháp mang tính ứng dụng cao, góp phần thúc đẩy việc học tập và giảng dạy tiếng Trung Quốc trở nên chất lượng và hiệu quả hơn.

Tài liệu tham khảo

[1]. Liu Song Hao (2007), *Thụ đắc ngôn ngữ thứ hai*. [M]. NXB Thế giới, Bắc Kinh.

[2]. Lu Qing He (2006), *Giảng dạy ngôn ngữ*. [M]. NXB Đại học Bắc Kinh, Bắc Kinh.

[3]. Nguyễn Thị Thu Hằng (2010), *Phân tích lỗi sai của SV Việt nam khi sử dụng giới từ*. [J] Tạp chí Ngôn ngữ học Quốc tế Cát Lâm, Trung Quốc 2021 (1): 19-27.

[4]. Xiao Zhonghua (2010), *Ứng dụng kho ngữ liệu trong giảng dạy ngôn ngữ*. [J] Tạp chí Đại học Chiết Giang, Trung Quốc 2010 (4): 190-200.

[5]. Xiao Zhonghua (2008), *Kho ngữ liệu và giáo dục ngôn ngữ*. [J] Giáo dục Ngoại ngữ Trung Quốc, 2008 (2): 48-58.