1 *Tạp chí Khoa học - Trường Đại học Mở Hà Nội 85 (11/2021) 10-19*

*Nghiên cứu trao đổi* ● *Research-Exchange of opinion*

# BÀN VỀ ỨNG DỤNG KHO NGỮ LIỆU TRỰC TUYẾN VÀO GIẢNG DẠY TIẾNG TRUNG QUỐC

### DISCUSSION ON THE APPLICATION OF ONLINE CORPUS TO TEACHING CHINESE

#### Dương Tiến Lê\*

Ngày tòa soạn nhận được bài báo: 03/05/2021 Ngày nhận kết quả phản biện đánh giá: 02/11/2021 Ngày bài báo được duyệt đăng: 29/11/2021

***Tóm tắt:*** *Sử dụng kho ngữ liệu hỗ trợ việc giảng dạy tiếng Trung Quốc đã trở thành một phương pháp phổ biến hiện nay. Do hạn chế về tính chất giáo trình tiếng Trung Quốc cơ sở, trong quá trình dạy học luôn tồn tại một số vấn đề nhất định. Sử dụng dữ liệu trong kho ngữ liệu có thể cung cấp cho việc dạy học một lượng lớn tài nguyên, tăng cường khả năng ứng dụng ngôn ngữ văn tự của người học. Ứng dụng sử dụng kho ngữ liệu trực tuyến vừa có thể nâng cao tính trực quan của việc dạy học lý thuyết, vừa có thể tăng cường khả năng ứng dụng ngôn ngữ của người học.*

***Từ khóa:*** *Kho ngữ liệu trực tuyến, giảng dạy tiếng Trung Quốc, ứng dụng ngôn ngữ học, giảng dạy lý thuyết, ngôn ngữ văn tự tiếng Hán.*

***Abstract:*** *Using corpus to support teaching Chinese has become a popular method today. Due to the limitations of basic Chinese textbooks, there are always certain problems in the teaching process. Using data in the corpus can provide teaching with a large amount of resources, increase students’s ability to apply written language. The application using an online corpus can both improve the visualization of theoretical teaching and enhance the language application of learners.*

***Keywords:*** *Online corpus, teaching Chinese, applied linguistics, theory teaching, Chinese language and character.*

###### Dẫn nhập

Thu thập nguồn ngôn ngữ trong hoàn cảnh ngôn ngữ thực tế và đưa chúng trở thành tài liệu để giải thích và chứng minh cho hiện tượng ngôn ngữ hiện có hoặc phát hiện quy luật ngôn ngữ mới là một trong những biện pháp cơ bản của

nghiên cứu ngôn ngữ. Đối với các nghiên cứu ngôn ngữ khi chưa có sự ứng dụng của công nghệ, người nghiên cứu ngôn ngữ chỉ có thể chỉnh lý, thu thập, lưu trữ tài liệu một cách thủ công qua những nguồn ngữ liệu hữu hạn và nhiều hạn chế. Trong những năm qua, khi công nghệ thông tin

* Khoa Tiếng Trung Quốc - Trường Đại học Mở Hà Nội

và Internet được ứng dụng vào việc nghiên cứu ngôn ngữ, các kho ngữ liệu trực tuyến cũng ngày càng được hoàn thiện và nâng cấp, quy mô và chức năng của chúng ngày càng mạnh mẽ, việc tìm kiếm ngày càng thuận tiện, việc ứng dụng kho ngữ liệu để tiến hành nghiên cứu ngôn ngữ cũng được phát triển mạnh mẽ. Giống như nhà ngôn ngữ học M.A.K Halliday đã nói, việc xuất hiện của các phương pháp khai thác ngữ liệu cùng với khả năng tích trữ to lớn của ngữ liệu hiện đại đã làm sâu sắc nhận thức của chúng ta về ngôn ngữ và sức mạnh to lớn của hiểu biết.” [1]

Những thành quả nghiên cứu gần đây đã liên tục thay đổi quan niệm và mô hình nghiên cứu ngôn ngữ trước đây, đồng thời cũng giúp việc nghiên cứu ngôn ngữ bằng ngữ liệu ngày càng được phổ biến rộng rãi. Ngày nay, kho ngữ liệu đã phát triển thành một nhánh khá riêng trong hệ thống ngôn ngữ học – ngôn ngữ học kho ngữ liệu. Nó mang đến ảnh hưởng ngày càng sâu rộng cho nhiều lĩnh vực trong nghiên cứu ngôn ngữ, được ứng dụng mạnh mẽ trong các nghiên cứu về các phương diện từ vựng, ngữ pháp, ngữ nghĩa, ngữ dụng, ngữ thể, thậm chí cả nghiên cứu ngôn ngữ xã hội học, nghiên cứu khẩu ngữ, biên soạn từ điển và nghiên cứu tổng hợp ngôn ngữ tự nhiên, trí tuệ thông minh, phiên dịch bằng máy và nhận biết ngôn ngữ.

Dạy tiếng Trung Quốc thuộc môn khoa học ứng dụng ngôn ngữ học, liên quan đến nghiên cứu giao thoa giữa nhiều môn học, việc lấy thành quả ứng dụng và phương pháp nghiên cứu từ môn khoa học mới – ngôn ngữ học kho ngữ liệu này rất cần thiết để phong phú lý luận nghiên cứu, thúc đẩy sự phát triển của bản thân môn

khoa học này. Phương pháp nghiên cứu ngôn ngữ lấy ngữ liệu trực tuyến làm cơ sở là một phương pháp hỗ trợ hiện đại không thể thiếu trong việc nghiên cứu của bất kỳ môn học nào. Vì thế, việc giảng dạy tiếng Trung Quốc cũng cần thu thập tài nguyên ngôn ngữ chân thực, khách quan, phong phú từ kho ngữ liệu để tiến hành nghiên cứu về giảng dạy và học thuật. Người viết cho rằng, kho ngữ liệu tiếng Trung Quốc trực tuyến cần được ứng dụng rộng rãi hơn nữa trong việc giảng dạy và nghiên cứu tiếng Trung Quốc tại Việt Nam.

###### Cơ sở lý luận

* 1. Giá trị của kho ngữ liệu đối với việc giảng dạy tiếng Hán

Tiếng Trung Quốc cơ sở là môn cơ sở, cũng là môn công cụ, là nền tảng để học các môn chuyên ngành chuyên sâu hơn. Nhưng đa số người học cho rằng tiết tiếng Trung Quốc cơ sở khô khan, khó học khó nhớ. Tuy nhiên, nền tảng tiếng Trung Quốc không chắc chắn sẽ ảnh hưởng nghiêm trọng đến việc học tập các môn học chuyên sâu hơn sau này, như lỗi ngữ pháp câu hay lỗi tu từ trong môn Viết, lỗi phát âm sai thanh điệu trong môn Nói, thậm chí là gặp khó khăn trong xác định cấu trúc câu trong môn Dịch song ngữ Việt – Trung, Trung Việt. Vì thế, sử dụng kho ngữ liệu sẽ hỗ trợ rất tốt trong quá trình dạy học, có thể nâng cao hiệu quả dạy tiếng Trung Quốc cơ sở.

* + 1. *Mang đến số lượng lớn ngữ liệu mới mẻ trong việc giảng dạy tiếng Trung Quốc cơ sở*

Cùng với sự phát triển của thời đại, lượng thông tin và kiến thức mới liên tục xuất hiện, nhưng nội dung giáo trình tiếng Trung Quốc cơ sở hiện có đã cách khá xa

với thực tế xã hội hiện tại, lý luận môn học và thực tế ngôn ngữ cũng đã có những sự chênh lệch nhất định. Ví dụ như một số từ hiện đã không được dùng trong thực tế do thay đổi bởi thói quen sử dụng của người bản địa như 小 姐 (xiǎojiě: nghĩa gốc là cô gái trẻ, giờ đã thành tiếng lóng chỉ từ tiêu cực) (Giáo trình tiếng Trung Quốc cơ sở 1 - quyển Thượng), hay thời gian trích các dẫn bài viết trong giáo trình đã quá lâu. Hay lượng từ 支(zhī: lượng từ của bài hát) (Giáo trình tiếng Trung Quốc cơ sở 3 - quyển Thượng), hiện giờ cũng rất ít sử dụng, thực tế về cơ bản chỉ sử dụng từ 首 (shǒu). Ngoài ra, do điều kiện giới hạn của khu vực, thời đại của tác giả khiến cho ví dụ trong giáo trình cũng mang tính địa phương, ví dụ như trong cuốn “Giáo trình tiếng Trung Quốc cơ sở” của Dương Ký Châu [2] thường xuyên lấy ví dụ ở các địa điểm của Bắc Kinh (về Trường Thành 长城, Di Hòa Viên 颐和园, Đại học Ngôn Ngữ Bắc Kinh 北京语言大学, Tây Trực Môn 西直门, Phục Hưng Môn 复兴门, Học viện Hý Kịch Trung Ương 中 央 戏剧学院, còn của Hoàng Bá Vinh, Lý Vĩ

1. thường lấy ví dụ ở Quảng Châu, như:

去北京路 (Đường Bắc Kinh)逛街, 我的中学同学六级肯六继科去年已经从中山大学 (Đại học Trung Sơn) 毕业了(trang 44), 他俩漫步在波光粼粼的东湖(Đông Hồ) 边 (trang 46). Những ví dụ này ít nhiều khiến người học ở nơi khác không dễ nắm bắt được nghĩa của câu. Vì thế, người viết cho rằng trong điều kiện đảm bảo lý thuyết cơ bản không đổi, tìm ra những ví dụ gần với đời sống, thú vị, thích hợp với người học hiện đại là rất cần thiết.

Như phần dạy học ngữ pháp, có thể so sánh các cách biểu đạt giữa tiếng phổ thông và tiếng địa phương, vùng miền

khác nhau, kết cấu ngữ pháp cũng khác nhau, giúp người học phân biệt được rõ ràng hơn khi học tập. So với các ví dụ trong giáo trình, khi người dạy tận dụng được những nguồn ngữ liệu khác nhau cũng mang tới sự phong phú và hấp dẫn, giúp người học cảm nhận được sự mới mẻ, hứng thú học tập cũng mạnh mẽ hơn. Ví dụ khi sử dụng kho Ngữ liệu trực tuyến BBC (BLCU Corpus Center) của Đại học Ngôn ngữ Bắc Kinh, chúng ta có thể lựa chọn hoàn cảnh xuất hiện từ mới đó trong “Đa lĩnh vực”, “Văn học”, “Báo chí”, “Đối thoại”, “Trích dẫn trong tác phẩm”, “Tiếng Hán cổ đại” hoặc “Tùy chọn”, từ đó lấy được ví dụ trong nhiều ngữ cảnh, nhiều bối cảnh không gian thời gian, thậm chí so sánh làm nổi bật được sự thay đổi về ngữ nghĩa của từ theo chiều dài lịch sử.

* + 1. *Mang đến phương pháp công nghệ mới cho quá trình dạy học tiếng Trung Quốc cơ sở*

Lý luận phương pháp dạy học của tiếng Trung Quốc truyền thống cơ bản áp dụng hình thức người dạy giảng bài, người học nghe bài; tiêu chuẩn đánh giá với người học là nắm được kiến thức lý luận coi như đã hoàn thành tốt môn học này, từ đó khiến người học cảm thấy môn học này nặng về lý thuyết, thiếu đi sự thú vị, hiệu quả giảng dạy không cao. Tuy nhiên, cùng với sự phát triển của thời đại, mô hình nhu cầu của xã hội về nguồn nhân lực cũng có những thay đổi, không chỉ yêu cầu người học phải có kiến thức chuyên môn chắc chắn, còn phải có năng lực thực tiễn mạnh mẽ, ngoài ra còn phải có ý thức sáng tạo nhất định. Nhưng mô hình dạy học truyền thống hiển nhiên không thể thích ứng với sự thay đổi này, còn kho ngữ liệu đã mang đến điều kiện thực hiện cải cách cho việc giảng dạy tiếng Hán.

Kho ngữ liệu tiếng Trung Quốc mang đến khả năng mới về tư liệu cho việc giảng dạy tiếng Hán. So với mô hình giảng dạy kiến thức truyền thống lấy người dạy làm trung tâm, mô hình dạy học lấy người học làm trung tâm để tìm kiếm tri thức đã càng làm nổi bật tính chủ động của người học. Khi người học đối diện với nguồn tư liệu ngôn ngữ phong phú đa dạng, có thể thỏa thích tìm kiếm tư liệu mình muốn tìm hiểu theo mục đích cá nhân, thông qua việc phân tích tài liệu để có được quy luật ngôn ngữ mà không còn cần người dạy trực tiếp truyền thụ điểm kiến thức cụ thể nữa.[3]

###### Yêu cầu về giảng dạy và học tập tiếng Trung Quốc từ kho ngữ liệu

Có thể nói rằng, học tập tiếng Trung Quốc qua kho ngữ liệu đã mang đến không gian khai thác rộng mở cho việc giảng dạy, cả người dạy và người học đều có thể tìm thấy tư liệu cần thiết theo yêu cầu của nội dung giảng dạy đưa ra, tiết kiệm được một lượng lớn thời gian tìm kiếm. Trong điều kiện dạy học với kho ngữ liệu, người học càng dễ nắm bắt 7 loại hình ý nghĩa của từ vựng tiếng Hán, gồm: ý nghĩa ngữ pháp, ý nghĩa khái niệm, ý nghĩa nội hàm, ý nghĩa ngữ thể, ý nghĩa tình thái và ý nghĩa kết hợp. Trong đó, ý nghĩa nội hàm là cơ sở chính để tiến hành phân biệt các từ gần nghĩa, ý nghĩa ngữ thể, ý nghĩa tình thái và ý nghĩa kết hợp lại có mối quan hệ rất lớn với đặc trưng ngữ dụng của từ. Ngoài ra, còn có nhà nghiên cứu nêu ra, người dạy và người học đều cần đặc biệt chút ý tới các khái niệm liên quan mật thiết với việc sử dụng từ ngữ, gồm: kết hợp, liên kết theo từ loại và ngữ nghĩa.[4] Khi học đến một điểm ngôn ngữ quan trọng của một phần nào đó, người dạy có thể thực

hiện tìm kiếm trước các thông tin quan trọng tạo thành bài giảng để trình bày trước người học, dẫn dắt người học cùng tổng kết các đặc điểm của từ quan trọng trong bài. Như vậy, kiến thức và năng lực về từ vựng được hình thành trên cơ sở một lượng lớn ngữ liệu chân thực, tránh được những thiếu sót hoặc sai lầm từ việc người dạy dựa vào trực giác cá nhân, giải thích của từ điển và số lượng ví dụ ít ỏi gây ra. Sau khi sử dụng thuần thục công cụ tìm kiếm, người dạy có thể để người học tự tiến hành quy nạp các kiến thức từ vựng có liên quan, chuyển hóa biến đổi dần sang tự chủ học tập hoàn toàn.

###### Ứng dụng kho ngữ liệu vào giảng dạy tiếng Trung Quốc cơ sở

Kho ngữ liệu mang đến không gian rộng mở trong việc cập nhật nội dung giảng dạy, thay đổi mô hình giảng dạy, đổi mới phương pháp giảng dạy. Vì thế chúng ta cần phát huy mạnh mẽ vai trò của kho ngữ liệu trong quá trình giảng dạy, dùng kho ngữ liệu để khai thác các tài nguyên dạy học có tính mục tiêu, để bù đắp những thiếu sót trong giáo trình; đưa ra nhiệm vụ giảng dạy, triển khai các hoạt động giảng dạy về ngữ pháp, từ vựng, tu từ hay thậm chí sửa lỗi sai của câu; khơi gợi lên cho người học sự hứng khởi và nhiệt tình với việc chủ động học tập, nâng cao năng lực thực tiễn về ngôn ngữ của người học.

* 1. Dùng kho ngữ liệu để khai thác nguồn tài liệu hỗ trợ giảng dạy có tính mục tiêu

Cùng với sự phát triển của thời đại, đời sống và ngôn ngữ của con người cũng có những sự thay đổi to lớn, hiện tượng ngôn ngữ mới liên tục xuất hiện. Kho ngữ liệu thu thập lượng lớn các tài nguyên

ngôn ngữ trong đời sống, có thể quan sát được sự thay đổi trong quá trình diễn biến của một số ngôn ngữ từ kho ngữ liệu trực tuyến. Đồng thời, ưu thế của kho ngữ liệu còn thể hiện ở việc thống kê định lượng tần suất sử dụng của một hình thức ngôn ngữ nào đó, ví dụ như từ “ 来 着 ” trong “Tiếng Trung Quốc hiện đại” của Hoàng Liêu giải thích rằng “thường chỉ việc vừa phát sinh chưa lâu, chỉ dụng với việc trong quá khứ, nhấn mạnh hành vi động tác.” Nhưng Triệu Trí Thanh (2010) [5] qua thống kê kho ngữ liệu cho rằng câu chỉ “quá khứ” chỉ chiếm 48.4%, ngoài ra có 45.2% chỉ “người nói từng tiếp xúc với một thông tin nào đó, sau đó trong khi nói nhất thời không nhớ ra, muốn có được xác nhận hoặc nhắc nhở của đối phương”. Có thể thấy lớp nghĩa sau này có tỉ lệ lớn trong ứng dụng sử dụng ngôn ngữ. Vì thế, bằng các chuyên đề tìm kiếm rà soát kho ngữ liệu, lựa chọn các tài liệu ngôn ngữ theo chuyên đề nhất định có thể khiến người học hiểu vì sao lại thế, bù đắp cho những thiếu sót của giáo trình, cũng giúp cho người học dễ tiếp nhận và yêu thích kiến thức hơn.

* 1. Dùng kho ngữ liệu triển khai các yếu tố giảng dạy chuyên ngành

Lý luận cơ sở ngôn ngữ của tiếng Trung Quốc hiện đại đề cập đến nhiều phương diện như ngữ âm, từ vựng, chữ Hán, ngữ pháp và tu từ. Chúng ta cần tận dụng tối đa ưu thế của kho ngữ liệu để thể hiện hiện tượng ngữ âm, kết cấu và ví dụ từ vựng, kết cấu ngữ pháp, quy luật tu từ thậm chí cả câu sai để người học có cái nhìn trực quan, từ đó phát hiện được quy luật ẩn giấu phía sau ngữ liệu, nâng cao hiệu quả giảng dạy.

Ví dụ như trong giảng dạy ngữ pháp, thực tế hay xuất hiện dạng “ 有 + VP”, nhưng trong giáo trình không hề nhắc tới. Bằng việc tìm kiếm kho ngữ liệu, có thể phát hiện ra nguyên nhân và nguồn gốc tạo ra hiện tượng này: Có thể do ảnh hưởng của phương ngữ các vùng phía nam Trung Quốc, sau đó các nhân vật “nổi tiếng” cũng sử dụng gây lan tỏa đến khẩu ngữ trong giới trẻ, rồi tăng dần sức ảnh hưởng.

* 1. Dùng kho ngữ liệu để kích thích sự nhiệt tình trong học tập của người học, nâng cao khả năng ứng dụng ngôn ngữ

Cùng với sự phát triển nhanh chóng của khoa học công nghệ những năm gần đây, việc sử dụng thậm chí là dựa vào công nghệ của người học ngày càng mạnh mẽ. Vì thế, dẫn dắt, định hướng người học khai thác những nguồn ngữ liệu mới liên quan mật thiết với đời sống hàng ngày, thông qua các biện pháp kỹ thuật số, chỉnh lý, gia công các ngữ liệu đó cũng là một hướng đi mới. Như vậy vừa có thể bồi đắp năng lực tự chủ học tập của người học, vừa có thể dùng kho ngữ liệu đó để phục vụ cho việc điều tra nắm bắt ngôn ngữ đời sống, khuyến khích người học phát hiện vấn đề, giải quyết vấn đề, quan trọng hơn nữa là có thể đưa một nguồn học liệu riêng đó vào giảng đường, làm sống động không khí lớp học, tăng sự yêu thích học tập của người học, đạt được mục đích “dạy và học cùng phát triển” cho người dạy và người học.

Ngôn ngữ là công cụ giao tiếp quan trọng. Vì thế, tạo dựng thói quen và ý thức tự giác quan sát, năng lực tổng hợp phân tích hiện tượng ngôn ngữ cho người học là điều vô cùng quan trọng, giúp ích cho

việc nâng cao trình độ tiếng Trung Quốc của người học, là một trong những nhiệm vụ quan trọng của giảng dạy tiếng Hán. Kho ngữ liệu cũng phát huy mạnh mẽ vai trò của mình trong việc giúp người học nâng cao khả năng khẩu ngữ và viết của mình. Đại học Ngôn ngữ Bắc Kinh đã và đang khai thác “Kho ngữ liệu bài viết HSK động”[6], thu thập bài làm văn của người học học tiếng Trung Quốc trên toàn thế giới, mang đến lượng tài nguyên vô cùng dồi dào cho việc giảng dạy tiếng Hán. Bằng kho dữ liệu đó, người dạy có thể phát hiện ra các vấn đề khác nhau, có thể lấy làm ví dụ khi giảng dạy, để người học tránh được những lỗi mắc phải, từ đó nâng cao khả năng viết tiếng Trung Quốc của mình.

###### Kết luận

Kho ngữ liệu mang đến cho chúng ta một không gian mở mới về lượng tài nguyên có thể tiếp cận được, tất nhiên không phải tất cả môn học và bài khóa đều thích hợp sử dụng kho ngữ liệu để hỗ trợ dạy học, để thực sự thấu hiểu và cảm thụ về một đoạn bài viết cần hiểu các lớp kết cấu từ câu, đoạn văn, phân tích quan hệ lô-gíc. Về phương diện này kho ngữ liệu không thể nào thay thế cách dạy học truyền thống, chỉ là trong một số trường hợp cụ thể, cần triển khai những mục tập trung vào những điểm ngữ pháp, từ vựng, tu từ phức tạp, thường dùng mà dễ nhầm,

muốn từ góc độ khảo sát thực tế để biết được tần suất sử dụng, thói quen sử dụng, các lớp nghĩa, cách sử dụng ra sao, kho ngữ liệu có thể trở thành tài liệu dạy học hỗ trợ bổ sung và tham khảo. Tóm lại, kho ngữ liệu không chỉ thích hợp cho bản thể ngôn ngữ và nghiên cứu dạy học, chỉ cần sử dụng chính xác, kết cấu rõ ràng, thiết kế đúng đắn, kho ngữ liệu chính là trợ thủ đắc lực cho việc giảng dạy và học tập tiếng Hán.

Tài liệu tham khảo:

[1]. M.A.K. Halliday (1991). *Corpus studies and probabilistic grammar, in Aijimer and Altenberg (eds)*. English Corpus Linguistics: Studies in Honour of Jan Svartvik. London: Longman. P41~42.

[2]. 杨继洲（2007），《汉语教程》，北京语言大学出版社

[3]. 黄伯荣、李炜（2013），《现代汉语》，北京大学出版社

[4]. 高晓梅，金乃茹2012，《 基于语料库的现代汉语课程改革探析》， 佳木斯大学社会科学学报

[5]. 赵志清 2007， 再谈“来着”——基于语料库的考察，临沂师范学院学报

[6]. “HSK动态作文语料库”课题组 2009，

《 “HSK动态作文语料库”正式开通》，世界汉语教学

***Địa chỉ tác giả: Khoa Tiếng Trung Quốc – Trường Đại học Mở Hà Nội***

[***Email: dtle@hou.edu.vn***](mailto:dtle@hou.edu.vn)