



# ỨNG DỤNG NGÔN NGỮ HỌC NGỮ LIỆU VÀO DẠY HỌC NGOẠI NGỮ: DẠY HỌC TỪ VỰNG TIẾNG VIỆT CHO NGƯỜI NƯỚC NGOÀI QUA NGỮ LIỆU SONG NGỮ ANH – VIỆT

Lê Lâm Thi

Trường Đại học Ngoại Ngữ – Đại học Huế, 57 Nguyễn Khoa Chiêm, Huế, Việt Nam

**Đỗ Thị Xuân Dung**

Đại học Huế, 3 Lê Lợi, Huế, Việt Nam

\* Tác giả liên hệ: Lê Lâm Thi <lelamthi82@gmail.com>

(Ngày nhận bài: 14-6-2021; Ngày chấp nhận đăng: 05-7-2021)

**Tóm tắt.** Trong thời đại phát triển toàn cầu trên nền tảng công nghiệp 4.0 hiện nay, cùng với sự phát triển của các ngành khoa học có ứng dụng công nghệ thông tin khác, ngôn ngữ học ngữ liệu ngày càng khẳng định được ưu thế của mình trong việc nghiên cứu và dạy học ngoại ngữ. Nhờ có các kho ngữ liệu với các tính năng cung cấp nguồn ngữ liệu phong phú và có hệ thống, phương thức nghiên cứu ngôn ngữ đã có những thay đổi đáng kể. Từ việc nghiên cứu từng trường hợp đơn lẻ trong từ điển, các nhà ngôn ngữ học có thể tìm được trong khối liệu những trích dẫn về từ và cụm từ cần thiết trong hàng loạt văn bản trong một thời gian ngắn. Tuy nhiên, việc nghiên cứu những ứng dụng của kho ngữ liệu trong lĩnh vực dạy học ngoại ngữ vẫn còn hạn chế. Bài báo này sẽ thảo luận tiềm năng sử dụng, khai thác kho ngữ liệu song ngữ Anh – Việt trong việc dạy học tiếng Việt cho người nước ngoài, đặc biệt chú ý đến ứng dụng ngữ liệu song ngữ trong việc dạy học từ vựng tiếng Việt. Bài báo cũng sẽ mô tả và phân tích một số hoạt động thực tế về việc sử dụng ngữ liệu song ngữ trong việc dạy học từ vựng tiếng Việt đã được áp dụng tại trường Đại học Ngoại ngữ, Đại học Huế. Từ đó, nghiên cứu chỉ ra tính hữu dụng của kho ngữ liệu song ngữ Anh - Việt đối với việc dạy học từ vựng tiếng Việt.

**Từ khóa.** ngữ liệu, song ngữ, dạy - học, tiếng Việt, từ vựng

# APPLICATION OF CORPUS LINGUISTICS TO FOREIGN LANGUAGE TEACHING: TEACHING VIETNAMESE VOCABULARY TO FOREIGNERS THROUGH ENGLISH - VIETNAMESE CORPUS

Le Lam Thi

HU – University of Foreign Languages, 57 Nguyen Khoa Chiem St., Hue, Vietnam

Do Thi Xuan Dung

Hue University, 3 Le Loi St., Hue City, Vietnam

\* Correspondence to **Le Lam Thi** <lelamthi82@gmail.com>

(Received: June 14, 2021; Accepted: July 05, 2021)

**Abstract:** The development of Information Technology on the ground of the Industry 4.0 creates favorable opportunities for corpus linguistics to increasingly assert its advantages in language research and teaching. The method of research on language has, therefore, witnessed considerable changes. From studying single language samples in dictionaries, linguists now can find in corpus the necessary citations of words and phrases available in a series of texts just within a short time. However, literature shows that the research on the application of corpus in the field of language teaching is still limited. This article, with a hope to narrow the gap, will discuss the possibility of using and exploiting English-Vietnamese bilingual corpus in teaching Vietnamese to foreigners, especially teaching Vietnamese vocabulary. The article will discuss the process of organising practical activities on using bilingual corpus to teach Vietnamese vocabulary to foreigners at University of Foreign Languages, Hue University. It also attempts to point out the usefulness of bilingual corpus in teaching and learning Vietnamese vocabulary with the support of Computer-Assisted Language Learning.

**Keywords:** corpus, bilingual, teaching, Vietnamese, vocabulary

## 1. Kho ngữ liệu và những ứng dụng trong lĩnh vực giảng dạy ngoại ngữ

“Ngôn ngữ học ngữ liệu” hay “ngôn ngữ học khối liệu” (corpus linguistics) hẳn không còn xa lạ với những nhà nghiên cứu Việt ngữ học. Từ khi xuất hiện vào những năm 1960, ngôn ngữ học ngữ liệu ngày càng phát triển và phổ biến, có thể ứng dụng vào rất nhiều lĩnh vực, đặc biệt là lĩnh vực nghiên cứu ngôn ngữ. Là một phân ngành của ngôn ngữ học ứng dụng, ngôn ngữ học ngữ liệu nghiên cứu các hiện tượng ngôn ngữ có thật dưới dạng văn bản và ngôn bản thông qua các kho ngữ liệu với việc sử dụng công nghệ máy tính hiện đại. Ngày nay, ngôn ngữ học ngữ liệu phát triển rất nhanh và mạnh mẽ theo đà phát triển của khoa học máy tính và kĩ

thuật số. Nhiều kho ngữ liệu ra đời đã mở ra nhiều hướng nghiên cứu mới trong lĩnh vực ngôn ngữ cũng như giảng dạy ngoại ngữ.

Có nhiều định nghĩa khác nhau về “ngữ liệu”. Từ điển *Oxford Advanced Learner’s Dictionary with Vietnamese Translation* (2014) [1] định nghĩa “corpus” là “a collection of written and spoken texts” và được dịch là “kho ngữ liệu (nói/viết)”. Theo nhà nghiên cứu Sinclair (1991) [2], “ngữ liệu” là “tập hợp các mẫu văn bản được lựa chọn một cách có hệ thống theo những tiêu chí nhất định nhằm đại diện cho một thể loại ngôn ngữ cụ thể với mục đích nghiên cứu ngôn ngữ.” Theo Đinh Điền (2018), thuật ngữ “ngữ liệu” được tạm dịch từ thuật ngữ tiếng Anh “corpus” có gốc từ tiếng Latin với nghĩa là “thân thể” (body). Vì vậy, các từ “corps” (“thân thể” trong tiếng Pháp, “quân đoàn”, “đoàn” trong tiếng Anh) hay từ “corpse” (“tử thi” trong tiếng Anh) cũng là những từ dẫn xuất từ gốc “corpus” này [3, tr.1].

Như vậy, khái niệm “kho ngữ liệu” chúng tôi muốn đề cập ở đây được hiểu là tập hợp văn bản đơn ngữ, đa ngữ hay song ngữ (gồm các cặp văn bản đã được dịch thủ công, dịch tương ứng 1-1 về mặt ngữ nghĩa) và phù hợp với lĩnh vực, thể loại, niên đại mà ta cần nghiên cứu. Một kho ngữ liệu có thể chứa văn bản bằng một thứ tiếng đơn lẻ được gọi là kho ngữ liệu đơn ngữ (monolingual corpus), hoặc văn bản bằng nhiều thứ tiếng được gọi là kho ngữ liệu đa ngữ (multilingual corpus). Các kho ngữ liệu đa ngữ được định dạng đặc thù để có thể so sánh khi đặt cạnh nhau (side-by-side) được gọi là kho ngữ liệu song song có dóng hàng (aligned parallel corpora). Sự dóng hàng có thể ở các mức như: các từ được dóng hàng từng đôi một (từ của ngôn ngữ A là từ dịch của ngôn ngữ B); các ngữ được dóng hàng từng đôi một (ngữ của ngôn ngữ A là ngữ dịch của ngôn ngữ B); các câu được dóng hàng từng đôi một (câu của ngôn ngữ A là câu dịch của ngôn ngữ B)... [4].

Ngoài ra, dựa vào cách xây dựng ngữ liệu, người ta thường chia ngữ liệu thành các loại sau: Ngữ liệu thô (raw corpus): chỉ là tập hợp các dữ liệu mà không có xử lý gì thêm; Ngữ liệu được gán nhãn (tagged corpus): ngữ liệu trong corpus đã được xử lý như phân tích từ, phân tích cú pháp, gán nhãn từ loại,... Chẳng hạn, các từ trong cụm từ sẽ được chú giải thông tin về từ loại – gọi là gán nhãn từ loại (part-of-speech tagging, viết tắt: POS-tagging). Các cụm từ trong câu sẽ được phân tách và gán nhãn – gọi là phân đoạn cụm từ (chunking). Với các ngôn ngữ đơn lập như tiếng Việt, do ranh giới của từ không được xác định rõ ràng bằng hình thức, nên ngữ liệu thường phải trải qua quá trình xác định đơn vị từ và gán nhãn khu biệt – gọi là phân đoạn từ (word segmentation), v.v.

Từ khi ra đời cho đến nay, ngôn ngữ học khối liệu đã được ứng dụng trong rất nhiều lĩnh vực khác nhau, trong đó có lĩnh vực dạy học ngoại ngữ. Việc khai thác và ứng dụng ngữ liệu trong giảng dạy ngôn ngữ nói chung và giảng dạy một ngoại ngữ nào đó nói riêng là hết sức cần thiết. Các kho ngữ liệu có thể được sử dụng trong giảng dạy ngoại ngữ, cụ thể là trong các lĩnh vực biên soạn giáo trình dạy tiếng, thiết kế bài giảng ngữ âm, từ vựng, ngữ pháp, xây dựng

để thi đánh giá năng lực ngoại ngữ... Những công trình nghiên cứu khai thác các kho ngữ liệu trong lĩnh vực giảng dạy như *Classroom applications of corpus analysis* [5], *Advantages of using corpora to teach English* [4], *Sử dụng kho tư liệu giảng dạy tiếng Việt như một ngoại ngữ* [6], *Ứng dụng Ngữ liệu Song ngữ Anh-Việt trong Giảng dạy Ngôn ngữ* [7], *Ứng dụng Kho Ngữ liệu trong việc dạy tiếng Việt cho người nước ngoài* [8]... Các công trình này đã chỉ ra những khả năng ứng dụng kho ngữ liệu và đề xuất một số hoạt động giảng dạy ngôn ngữ và ngoại ngữ liên quan đến ngữ liệu. Tuy nhiên, vẫn chưa có công trình nào nghiên cứu việc ứng dụng ngữ liệu song ngữ Anh - Việt vào việc dạy học từ vựng tiếng Việt cho người nước ngoài. Chúng tôi nhận thấy việc dạy học từ vựng tiếng Việt cho người nước ngoài dựa trên ngữ liệu song ngữ là cần thiết bởi lẽ nó giúp người dạy tiết kiệm công sức, thời gian trong việc tìm kiếm ví dụ, thiết kế bài tập từ vựng. Mặt khác, so với ngữ liệu đơn ngữ, ngữ liệu song ngữ có nhiều tiện ích hơn. Ngữ liệu song ngữ có thể giúp người học xác định rõ nghĩa của từ trong các ngữ cảnh cụ thể, xác định quy tắc kết hợp của từ thông qua so sánh các chuỗi câu đồng hiện. Hơn nữa, trong một số trường hợp, ngữ liệu song ngữ còn giúp người học tránh được những chuyển di tiêu cực do ảnh hưởng của tiếng mẹ đẻ trong quá trình học ngoại ngữ.

## **2. Ứng dụng tiện ích của kho ngữ liệu vào việc dạy học từ vựng tiếng Việt cho người nước ngoài qua ngữ liệu song ngữ Anh – Việt (Trường hợp của trường Đại học Ngoại ngữ, Đại học Huế)**

Một trong những ứng dụng của kho ngữ liệu vào việc giảng dạy ngoại ngữ đó chính là ứng dụng vào việc dạy học từ vựng cho người nước ngoài học tiếng Việt. Trong quá trình giảng dạy tiếng Việt cho người nước ngoài ở Khoa Việt Nam học, trường Đại học Ngoại ngữ, Đại học Huế, chúng tôi đã tổ chức một số hoạt động dạy – học từ vựng qua ngữ liệu song ngữ Anh – Việt. Kho ngữ liệu mà chúng tôi lựa chọn để thực hiện là kho ngữ liệu song ngữ Anh – Việt, gồm 60.000 câu song ngữ. Đây là kho ngữ liệu được xây dựng trên cơ sở ngữ liệu song ngữ Anh – Việt EVC của trung tâm Ngôn ngữ học tính toán, trường Đại học Khoa học Tự nhiên, Đại học Quốc gia TP. Hồ Chí Minh đã được chỉnh sửa, bổ sung cho phù hợp với mục đích giảng dạy tiếng Việt cho người nước ngoài. Theo đó, các bản dịch tương ứng của mỗi ngôn ngữ (Anh – Việt) được đặt song song với nhau hay còn được gọi là dóng hàng với nhau (alignment). Mức độ dóng hàng này ở cấp độ câu (sentence alignment), cấp độ ngữ (phrase alignment) và cấp độ từ (word alignment); nghĩa là từng câu, từng ngữ, từng từ trong ngôn ngữ nguồn được dóng (liên kết) với câu dịch tương ứng trong ngôn ngữ đích. Có thể xem ví dụ ở hai hình dưới đây:

Search Word - □ X

Language  
 English  
 Vietnamese

Key search

Match cases  Morphological cases  Phrase

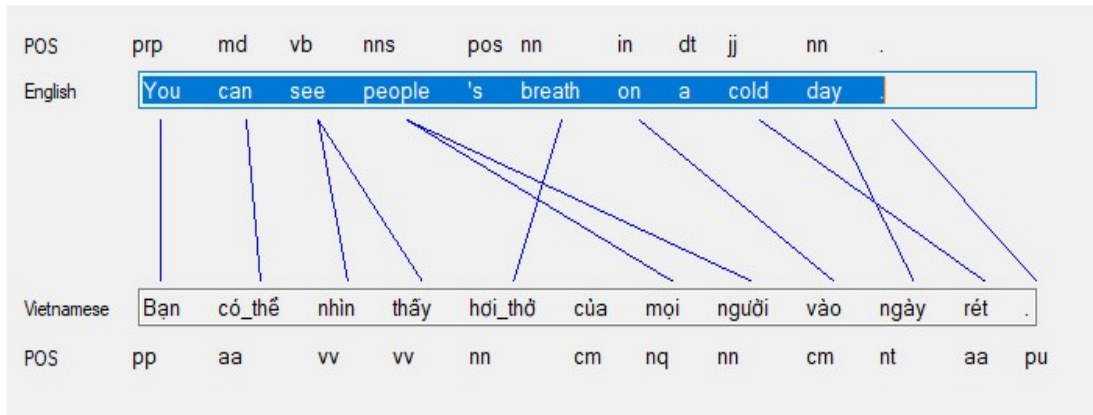
**English** Found total 624 results.

Left	Key	Right
A guest has come to	see	Mr. Brown from the United States.
A man whose like we shall not	see	again.
A Mrs Green is waiting to	see	you.
A purblind oligarchy that flatly refused to	see	that history was condemning it to the dustbin.
A screech of brakes made us took to	see	what had happened in the street.
A strange man, fred by name came to	see	me.
A throng of fans waiting to	see	the star.
A traveler in Scotland, for instance, might not	see	any man wearing kilts.
After all, you can't	see	this country very well with your eyes closed!
a message that travel over the nerves are those that let us	see	hear, taste, smell, and feel

**Vietnamese**

Left	Key	Right
Một người khách đến	thăm	ông Brown đến từ Mỹ.
Một người giống như thế chúng ta sẽ không	gặp	lại.
Một bà Green nào đó đang chờ	gặp	anh.
Một tập đoàn đầu sỏ chính trị u tối cứ nhất định không chịu	thấy	rằng lịch sử đang muốn loại họ vào thùng rác.
Tiếng thắng xe kêu rít khiến chúng tôi chạy ra	xem	ngoài phố đã xảy ra chuyện gì.
Một người lạ có tên là Fred đã đến	gặp	tôi.
Một đám đông những kẻ ái mộ đứng đợi để	nhìn thấy	ngôi sao đó.
Chẳng hạn một du khách có thể không	nhìn thấy	bất cứ một người đàn ông nào mặc váy ngắn tại Scotlar
Rốt cuộc, ông không thể	tham quan	được đất nước này kỉ với đôi mắt buồn ngủ như thế đâu
n hình của chúng ta có những thông điệp cho nhận chúng	nhìn	nghe, nếm, ngửi và sờ thấy

Hình 1: Dóng hàng cấp độ câu



Hình 2: Dóng hàng cấp độ từ

Những hoạt động dạy - học từ vựng qua ngữ liệu song ngữ đã đem đến một số hiệu quả nhất định. Chẳng hạn như kho ngữ liệu song ngữ Anh - Việt rất hữu ích trong việc dạy tiếng Việt cho người nước ngoài thông qua các so sánh trực quan như khảo sát cách dùng từ qua chuỗi đồng hiện, khảo sát chuỗi ngôn từ, chuyển ngữ Anh - Việt cho một số cụm từ cố định... Với kho ngữ liệu lớn, được lấy mẫu hợp lý, hầu hết các hiện tượng ngôn ngữ sẽ được phản ánh chính xác, sinh động và cập nhật hơn so với từ điển [3]. Trong quá trình giảng dạy tiếng Việt cho người nước ngoài tại trường Đại học Ngoại ngữ, Đại học Huế, chúng tôi đã tổ chức 3 hoạt động dạy học từ vựng thông qua việc sử dụng kho ngữ liệu song ngữ Anh - Việt. Những hoạt động này đều được áp dụng cho người học học tiếng Việt ở trình độ trung cấp.

#### *Hoạt động 1: Xác định ý nghĩa của từ từ ngữ liệu song ngữ Anh - Việt*

*Gới thiệu:* Đối với người dạy ngoại ngữ, việc giúp người học nắm được nghĩa của từ, đặc biệt đối với từ đa nghĩa là rất quan trọng. Thông thường chúng ta phải căn cứ vào ngữ cảnh của từ (context) để xác định nghĩa. Khi xem xét nghĩa hoặc cách dùng của một từ nào đó, ta cần xem xét ngữ cảnh tương ứng của nó [1], [2]. Việc người dạy tìm hoặc biên soạn những câu ví dụ về nghĩa của từ trong những ngữ cảnh khác nhau có thể mất rất nhiều thời gian và công sức. Với kho ngữ liệu song ngữ, giảng viên có thể tìm kiếm ngữ liệu dễ dàng hơn rất nhiều.

*Mục đích của hoạt động:* Giúp người học xác định nghĩa và cách dùng của từ thông qua chuỗi đồng hiện các cặp câu song ngữ Anh - Việt.

*Mô tả hoạt động:* Người dạy chuẩn bị tài liệu để phát cho người học 20 câu song ngữ có chứa từ muốn dạy được trích xuất từ kho ngữ liệu song ngữ Anh - Việt và yêu cầu người học phân loại các câu đó thành những nhóm nghĩa hoặc những trường hợp sử dụng khác nhau của từ đó.

Chẳng hạn như trường hợp của từ “nhà”, người dạy có thể cung cấp cho người học 20 cặp câu trích từ ngữ liệu song ngữ có chứa từ “nhà” trong câu tiếng Việt và yêu cầu người học

phân loại các câu trên theo những nhóm nghĩa của từ “nhà”. Sau khi phân loại xong các câu song ngữ, người dạy và người học sẽ cùng nhau xác định nghĩa của từ “nhà” trong tiếng Việt. Kết quả phân loại thể hiện như sau.

Nhóm một: Từ “nhà” trong tiếng Việt có nghĩa là “nơi ở”, tương đương với “house”/ “home”/ “door” trong tiếng Anh. Các cặp câu thuộc nhóm này sẽ là:

1. *You have a very nice **house**.* / *Bạn có một ngôi **nhà** rất đẹp.*
2. *The **house** was filled the aroma of coffee.* / *Ngôi **nhà** đầy hương cà phê.*
3. *You would be well advised to stay **indoors**.* / *Anh nên ở **trong nhà** thì tốt hơn.*
4. *Yes, and the **next door** neighbor is a good friend of mine.* / *Vâng, người láng giềng **kế bên nhà tôi** là một người bạn tốt của tôi.*
5. *The children were already asleep when we **got home**.* / *Khi chúng tôi **về nhà** thì bọn trẻ đã ngủ rồi.*
6. *The prisoner alleges that he was **at home** on the night of the crime.* / *Người tù viện lẽ rằng anh ta **ở nhà** vào cái đêm xảy ra án mạng.*

Nhóm hai: Nếu như từ “nhà” trong tiếng Việt có thể kết hợp với những từ khác để tạo thành từ ghép như “tòa nhà” (building)/ “trần nhà” (ceiling), khu nhà (block), “mái nhà” (roof), “sàn nhà” (floor) thì tiếng Anh lại không có từ tương đương với “nhà” như trong các từ ghép tiếng Việt mà ý nghĩa của “nhà” nằm ngay trong các từ “building”, “ceiling”, “block”, “roof”, “floor”. Ví dụ:

7. *The exterior of the **building** is very unattractive.* / *Mặt ngoài của **toà nhà** trông không có gì lôi cuốn.*
8. *The **ceiling** collapsed and rubble poured into the room.* / ***Trần nhà** sụp xuống và gạch vụn tuôn xuống căn phòng.*
9. *The **block** is still just a skeleton of girders.* / ***Khu nhà** vẫn chỉ là một bộ khung sắt.*
10. *The beams receive the full weight of the walls and **roof**.* / *Các cây đà đỡ chịu toàn bộ sức nặng của những bức tường và **mái nhà**.*
11. *The books were scattered pell - mell over the **floor**.* / *Sách rải bừa bãi khắp **sàn nhà**.*

Nhóm ba: Từ “nhà” được sử dụng để diễn đạt ý nghĩa sở hữu. Thay vì nói “con trai” của anh, ta có thể nói thành “con trai nhà anh”, “vườn của bà” có thể nói thành “vườn nhà bà”... Tuy nhiên, người dạy cần lưu ý với người học không phải tất cả những trường hợp diễn đạt ý nghĩa sở hữu đều có thể dùng từ “nhà” mà chỉ có những con người, những sự vật có sự liên quan đến “nhà” mới dùng được. Ví dụ:

12. *Your son is the despair of all his teachers.* / *Cậu con trai **nhà anh** là nỗi thất vọng của tất cả thầy giáo dạy nó.*
13. *Your children are disturbing my wife, can you speak to them, please?* / *Bọn trẻ **nhà anh** quấy rầy vợ tôi quá, làm ơn máng chúng nó giùm!*

14. *I don't want to alarm you, but there's a strange man in **your** garden.* / Tôi không muốn làm bà sợ hãi đâu, nhưng có một người lạ trong vườn **nhà bà**.

Nhóm bốn: Từ “nhà” trong tiếng Việt có nghĩa là “người” gắn liền với nghề nghiệp hay chức vụ nào đó. Nếu đem nó so sánh với tiếng Anh thì ta thấy không có từ tiếng Anh nào tương đương với từ “nhà” trong tiếng Việt. Ví dụ:

15. *A clever **politician** who knows how to manipulate public opinion.* / Một **nhà chính trị** khôn ngoan biết cách chi phối công luận.

16. *A **diplomat** subverted by foreign power.* / Một **nhà ngoại giao** bị thế lực ngoại bang làm biến chất.

17. *Even an experienced **climber** can get into trouble.* / Ngay cả một **nhà leo núi** kinh nghiệm cũng có thể gặp khó khăn.

18. *It appears that the two **leaders** are holding secret talks.* / Hình như là hai **nhà lãnh đạo** đang có cuộc hội đàm bí mật.

19. *The **historian** discovers the past by the judicious use of such a heuristic device as the ideal type.* / **Nhà sử học** đó khám phá quá khứ bằng việc sử dụng đúng đắn một phương pháp suy đoán tới mức được coi là loại lý tưởng.

Nhóm năm: Từ “nhà” có nghĩa là những người trong gia đình, tương ứng với “family” trong tiếng Anh. Ví dụ:

20. *I really don't feel much like cooking, but the **family** must eat.* / Em thật sự không thích nấu nướng lắm đâu, nhưng **cả nhà** phải ăn.

Hoạt động này cũng có thể được áp dụng như một bài tập ở nhà, theo đó người học sẽ được cung cấp ngữ liệu song ngữ và tự tìm hiểu nghĩa của một số từ đa nghĩa khác trong tiếng Việt.

**Hoạt động 2: Phân biệt sự khác nhau về ngữ nghĩa và cách dùng của các danh từ chỉ loại thể**

*Giới thiệu:* Danh từ chỉ loại thể hay còn gọi là loại từ là những danh từ dùng để chỉ đơn vị rời khi được kết hợp với các danh từ có ý nghĩa tổng loại, ví dụ: *cái, con, quả, củ, tấm, bức, sợi, quyển, con, trận, viên, hòn, bộ, vị, ngài, cây, người, đàn, làn...* Những từ này được sử dụng cùng với các danh từ có ý nghĩa tổng loại để thể hiện tính chất cụ thể của danh từ đó [4]. Tiếng Việt có một số lượng lớn từ chỉ loại thể. Việc xác định cách sử dụng những từ này tương đối phức tạp và có thể gây khó khăn, nhầm lẫn cho người nước ngoài học tiếng Việt. Người học có thể gặp khó khăn trong việc lựa chọn và sử dụng các từ này khi kết hợp với danh từ. Hoạt động với kho ngữ liệu song ngữ sẽ phần nào giúp người học phân biệt được các trường hợp sử dụng các từ này.

*Mục đích của hoạt động:* Giúp người học phân biệt được các trường hợp sử dụng các danh từ chỉ loại thể.



*Mô tả hoạt động:* Sau khi giới thiệu về các từ loại thể trong tiếng Việt, người dạy chuẩn bị 40 cặp câu song ngữ, trong đó câu tiếng Việt có chứa các từ chi loại thể rồi yêu cầu người học sắp xếp, phân loại đối tượng, xem những từ nào có thể kết hợp được với những từ chi loại thể và rút ra quy tắc kết hợp giữa chúng. Như vậy, từ kết quả phân loại, người dạy và người học cùng khái quát hóa cách dùng của các danh từ loại thể trong tiếng Việt. Chẳng hạn như để phân biệt về ngữ nghĩa của từ “con” và từ “cái”, hai từ chi loại phổ biến nhất trong tiếng Việt, người dạy có thể cung cấp cho người học 20 câu có chứa có từ loại thể “con” trong tiếng Việt và 20 câu có chứa từ “cái” kèm câu tiếng Anh tương đương. Sau khi phân loại ngữ liệu, người dạy và người học cùng thảo luận một số nội dung chính như sau:

Trường hợp từ “con” trong tiếng Việt

Thứ nhất, từ “con” có thể kết hợp với các từ chỉ động vật.

1. *A bee stung me on the cheek.* / **Con ong** đốt vào má tôi.
2. *A fox darted out of the midst of the thicket.* / **Con cáo** phóng ra từ giữa bụi rậm.
3. *A cat is luxuriating in the warm sunshine.* / **Con mèo** đang tận hưởng ánh nắng ấm áp.
4. *A lion has escaped from its cage.* / Một **con sư tử** đã sống chuồng.
5. *An escaped tiger is roaming free in the town.* / Một **con cọp** sống chuồng đang đi lang thang thoải mái trong thành phố.

Thứ hai, từ “con” có thể kết hợp với các từ như “đường”, “phố”, “hẻm”, “đê” ...

6. *A high wall made the street a cul - de - sac.* / Bức tường cao làm **con phố** đó thành một phố cụt.
7. *A narrow lane off the main road.* / Một **con đường** nhỏ tách ra từ con đường chính.
8. *A dyke was built to retain the floods.* / Một **con đê** được xây lên để ngăn lụt.
9. *All the approaches to the palace were guarded by troops.* / Tất cả **con đường** dẫn tới cung điện đều được quân đội canh gác.

Thứ ba, từ “con” có thể kết hợp với các từ chỉ đồ vật như “dao”, “búp bê”, “át”, “thuyền” ...

10. *A knife blunted by years of use.* / Một **con dao** đã cùn qua nhiều năm sử dụng.
11. *Fashion a doll from a piece of wood.* / Làm một **con búp bê** từ một mẩu gỗ.
12. *He took my ace with a low trump.* / Anh ta ăn **con át** của tôi bằng một con chủ bài nhỏ.
13. *A strong wind arose and blew our boat on to the rocks.* / Một con gió mạnh nổi lên thổi băng **con thuyền** của chúng tôi vào các tảng đá.
14. *Are there any stamps in the drawer?* Trong ngăn kéo đó có **con tem** nào không?
15. *She ran to the station only to find that the train had left.* / Cô ấy đã chạy đến nhà ga chỉ để thấy rằng **con tàu** đã rời đi.

Thứ tư, từ “con” có thể kết hợp với những sự vật trong thế giới tự nhiên như “sóng”, “sông”, “suối”, “kênh”....

16. *The breaking waves left the beach covered with foam.* / Những **con sóng** vỗ vào bờ biển phủ đầy bọt sóng.
17. *After the storm all the rivers were in spate.* / Sau cơn bão mọi **con sông** đều ngập lũ.
18. *Can you see a stream in the sun beside the rock?* / Ông có thể nhìn thấy **con suối** dưới ánh mặt trời bên cạnh tảng đá đó không?
19. *A canal connects it to the Rhine River.* / Nó nối với sông Rhine bằng một **con kênh** đào.

Thứ năm, từ “con” có thể kết hợp với các từ chỉ những đối tượng không có trong thực tế nhưng được “thiên hóa” như “ma”, có thể mở rộng ra trong thực tế sử dụng người Việt còn dùng “con quý”, “con yêu tinh”...

20. *Was what you saw something substantial or only a ghost?* / Cái anh thấy là cái có thật hay chỉ là **con ma**?

Đối với từ “cái” thì những trường hợp kết hợp có thể như sau:

Thứ nhất, “cái” dùng trước những đồ vật có hình dáng cụ thể.

1. *A baby stifled by a pillow.* / Một đứa bé bị **cái gối** làm ngộp thở.
2. *A key ring is for carrying keys on.* / **Cái vòng đeo khoá** dùng để móc các chìa khoá vào.
3. *A ladder leaning against the wall.* / Một **cái thang** tựa vào bức tường.

Thứ hai, từ “cái” dùng trước một động từ để biến động từ đó thành danh từ. Trong trường hợp này “cái” sẽ có nghĩa trừu tượng hơn. Ví dụ:

4. *A light tap on the shoulder.* / Một **cái vỗ** nhẹ trên vai.
5. *A mother weeping over the death of her child.* / Một người mẹ khóc than về **cái chết** đứa con của cô ta.
6. *A mutinous look is on one's face.* / **Cái nhìn** nổi loạn trên gương mặt ai đó.
7. *A scornful toss of the head.* / **Cái hất đầu** tỏ vẻ khinh miệt.
8. *A twinkle in his eye which belied the gravity of his demeanour.* / Một **cái nháy mắt** làm mắt hết vẻ trang nghiêm trong cách cư xử của ông ta.
9. *All that betrayed his surprise was a sharply indrawn breath.* / Tất cả những gì làm bại lộ sự ngạc nhiên của anh ta là một **cái hít** vào nhanh.
10. *By the way, please bring a date if you want.* / Nhân tiện hãy cho một **cái hẹn** nếu bạn muốn.
11. *The idea of making a speech in public petrified him.* / **Cái ý nghĩ** đọc một bài diễn văn trước công chúng đã làm anh ta sững sờ.

Thứ ba, từ “cái” có thể dùng trước những từ chỉ bộ phận cơ thể như “miệng”, “chân”, “mặt”, “mũi” ...

12. *A sensitive **mouth** is necessary to appreciate good wine. / Phải có một **cái miệng** sành ăn mới thường thức được món rượu ngon.*

Thứ tư, “cái” có thể đứng trước những danh từ có nghĩa trừu tượng như “có”, “giá”, “cách” hoặc các đại từ không xác định “gì”, “đó”, “kia” ...

13. *You want to tell me **something**? / Anh muốn nói với tôi **cái gì** phải không?.*  
 14. *A minor offence which provided a **peg** to hang their attack on. / Một sự xúc phạm nhỏ làm **cái cớ** cho sự tấn công của họ.*  
 15. *Do you like **the way** I 've rearranged the room? / Bạn có thích **cái cách** tôi sắp xếp lại căn phòng không?*  
 16. *Being recognized wherever you go is **the price** you pay for being famous. / Đi đâu cũng bị nhận ra ngay là **cái giá** anh phải trả cho sự nổi tiếng.*

Thứ năm, “cái” có thể đứng trước những tính từ để danh từ hóa tính từ đó chẳng hạn như “tốt”, “xấu”, “đẹp”, “ác” ...

17. *A twilight zone between **good** and **evil**. / Một lĩnh vực khó phân biệt giữa **cái tốt** và **cái xấu**.*  
 18. ***The good** and **harm** cancel each other out. / **Cái tốt** và **cái xấu** triệt tiêu lẫn nhau.*  
 19. *The forces of **evil** still at work today. / Ảnh hưởng của **cái ác** vẫn còn hoành hành.*  
 20. *You must learn the difference between **right** and **wrong**. / Anh phải biết phân biệt giữa **cái phải** với **cái trái**.*

Người dạy có thể giải thích cho người học từ “con” trong tiếng Việt vốn dĩ là loại từ chỉ động vật tuy nhiên trong thực tế sử dụng, người Việt cũng có thể nói *con đường, con phố, con dao, con sông, con búp bê, con ma...* Những cách dùng này được hình thành trong quá trình nhận thức của người Việt: con vật khác đồ vật ở đặc điểm cơ bản là chuyển động được. Người Việt gọi là “con” cho vật nào mang đặc điểm động (chúng mang dáng dấp hoặc phản ánh sự vận động, chuyển động). Chẳng hạn, người ta gọi những vật do con người làm ra rồi điều khiển chúng chuyển động là “con” như *con dao, con át (bàn), con thuyền, con búp bê ...* Người Việt cũng gọi *con đường, con hẻm, con phố...* vì trong đó có dòng người qua lại không ngừng cũng gọi ra sự chuyển động. Tương tự, *sông, suối, mương, máng, kênh, rạch* có dòng nước chảy gọi ra một đối tượng chuyển động nên chúng được gọi là “con”. Trong lúc đó, những đối tượng kết hợp với “cái” hầu như không hoặc ít có sự chuyển động. Ví dụ, chúng ta gọi *cái bàn, cái tủ, cái chậu, cái sân...* Như vậy, người dạy có thể phân biệt cho người học từ “con” và từ “cái” bằng ý niệm “chuyển động” và “không chuyển động” của sự vật.

**Hoạt động 3: Phân biệt từ đồng nghĩa qua việc so sánh cách sử dụng từ của người học với cách sử dụng từ của người bản ngữ**

*Giới thiệu:* Trong quá trình dạy ngoại ngữ nói chung và học tiếng Việt như một ngoại ngữ nói riêng, người dạy sẽ gặp những trường hợp người học nhầm lẫn trong việc sử dụng những từ đồng nghĩa. Hoạt động với kho ngữ liệu Anh – Việt sẽ giúp người học dễ dàng tiếp cận cách sử dụng từ ngữ của người bản ngữ thông qua việc so sánh với tiếng mẹ đẻ của họ.

*Mục đích của hoạt động:* Giúp người học sử dụng đúng những từ đồng nghĩa trong tiếng Việt.

*Mô tả hoạt động:* Người dạy cung cấp 20 mẫu ngữ liệu tiếng Anh trích từ kho ngữ liệu song ngữ Anh – Việt đã có và yêu cầu người học dịch các câu đó sang tiếng Việt. Sau khi người học dịch xong, người dạy mới cung cấp phần ngữ liệu tiếng Việt tương ứng để người học so sánh sự giống nhau và khác nhau giữa bản dịch của người học và ngữ liệu tiếng Việt được cung cấp, chú trọng vào việc chuyển dịch từ cần dạy. Từ sự so sánh này, người học sẽ nghiệm ra được cách dùng của các từ tương ứng trong tiếng Việt.

Chẳng hạn như người dạy cung cấp cho người học 15 câu tiếng Anh có chứa từ “wear” được lựa chọn từ kho ngữ liệu song ngữ Anh – Việt và yêu cầu người học thử chuyển dịch những câu đó sang tiếng Việt rồi so sánh những câu người học vừa dịch với câu tiếng Việt trong ngữ liệu. Ngữ liệu song ngữ có thể lựa chọn trong trường hợp này là:

1. *Did you have to wear a school uniform in your high school?/ Bạn có phải mặc đồng phục trong trường trung học không?*
2. *Do you wear glasses or contact lenses? / Bạn đeo kính hay kính sát tròng?*
3. *Do you wear perfume or cologne? / Bạn có xức nước hoa hay tinh dầu không?*
4. *In many Muslim countries, the women wear veils. / Ở nhiều nước Hồi giáo, phụ nữ đeo mạng che mặt.*
5. *It is advisable to wear a safety belt when you 're driving. / Nên thắt dây an toàn trong khi bạn lái xe.*
6. *My trousers only stay up if I wear a belt. / Cái quần của tôi chỉ không tụt xuống khi đeo thắt lưng.*
7. *She continued, perversely, to wear shoes that damaged her feet. / Cô ấy khăng khăng tiếp tục đi đôi giày đã làm đau chân cô.*
8. *She had her ears pierced so that she could wear earring. / Cô ta đã xỏ lỗ tai để đeo bông tai.*
9. *We were advised to wear mittens to avoid frost bit in winter./ Chúng tôi được khuyên đeo găng tay ấm để tránh cái giá lạnh của mùa đông.*
10. *What does a Scotsman wear (underneath) his kilt? / Người Scotland mặc gì bên dưới (bên trong) cái váy của anh ta?*
11. *Which shoes shall I wear, the red shoes or the brown ones?/ Em đi đôi giày nào , đôi đỏ hay đôi nâu?*

12. *Why don't you wear your pink skirt? / Tại sao bạn không mặc cái váy màu hồng?*
13. *You can go out on condition that you wear an overcoat. / Cậu có thể đi ra ngoài nếu như cậu chịu khoác áo ngoài.*
14. *It's a contradiction to love animals and yet wear furs. / Thật là một điều mâu thuẫn khi yêu thương súc vật mà lại mặc áo lông thú.*
15. *I don't have the courage to wear something like that. / Tôi không có can đảm mặc thứ gì tương tự như thế.*

Sau khi đối chiếu bản dịch của người học và bản dịch trong kho ngữ liệu, người học có thể nhận ra một số lỗi thường gặp như tiếng Việt sẽ không nói *Cô ta xỏ lỗ tai để mặc bông tai* mà phải nói *Cô ta xỏ lỗ tai để đeo bông tai*; hoặc không nói *Nên mặc dây an toàn trong khi bạn lái xe* mà phải nói *Nên đeo dây an toàn trong khi bạn lái xe*. Qua sự hướng dẫn của người dạy, người học nhận ra từ “wear” trong tiếng Anh có thể được chuyển dịch thành nhiều từ khi kết hợp với danh từ đứng đằng sau nó như *mặc (đồng phục, váy, áo...)*, *đeo (kính, găng tay, mạng che mặt...)*, *xức (nước hoa, dầu...)*, *thắt (dây an toàn, cà vạt...)*, *đi (giày, tất...)*...

### 3. **Tiềm năng ứng dụng ngữ liệu song ngữ Anh – Việt trong việc dạy học tiếng Việt cho người nước ngoài**

Trong phần trước, chúng tôi đã đề cập đến ba hoạt động sử dụng ngữ liệu song ngữ từ kho ngữ liệu song ngữ Anh – Việt như một công cụ đặc lực để hỗ trợ quá trình dạy – học tiếng Việt cho người nước ngoài. Với kho ngữ liệu song ngữ Anh – Việt, người dạy có thể tổ chức nhiều hoạt động dạy học tiếng Việt nói chung và dạy học từ vựng tiếng Việt nói riêng. Tùy theo mục đích bài học, trình độ người học, người dạy có thể lựa chọn những hoạt động phù hợp.

Việc sử dụng ngữ liệu song ngữ trong quá trình dạy học tiếng Việt cho người nước ngoài có rất nhiều lợi ích. Qua ngữ liệu song ngữ, người học biết cách người bản ngữ sử dụng các từ cụ thể trong ngôn ngữ của họ như thế nào. Điều này đã cung cấp cho người dạy và người học những thông tin quý giá về ý nghĩa và cách sử dụng từ ngữ tiếng Việt (hoạt động 1). Kho ngữ liệu song ngữ cũng có thể cho phép người học xác định được cách dùng của các từ chỉ loại (cái, con,...) (hoạt động 2), giúp người học lựa chọn những từ phù hợp trong một bối cảnh hoặc tình huống nhất định. Ngữ liệu song ngữ cũng có thể giúp người học hiểu cách dùng của một số từ đồng nghĩa trong ngữ cảnh qua việc so sánh để tìm ra điểm giống và điểm khác trong cách sử dụng từ tiếng Việt của người học (không phải bản ngữ) với cách sử dụng từ của người Việt (người bản ngữ) (hoạt động 3). Tóm lại, hoạt động với ngữ liệu song ngữ đã đề cập ở trên có thể giúp người học sử dụng từ tiếng Việt một cách thành thạo, phù hợp ngữ cảnh và chuẩn mực hơn.

Không những vậy, bản thân người học cũng có thể sử dụng kho ngữ liệu song ngữ để khám phá các vấn đề khác thuộc ngôn ngữ học, chẳng hạn như cấu trúc cụm từ, sự có mặt hay

vắng mặt của các từ chỉ loại..., giúp giao tiếp tiếng Việt tốt hơn. Do đó, tiềm năng khai thác của kho ngữ liệu đối với việc dạy và học từ vựng tiếng Việt là rất lớn. Hơn nữa, việc nghiên cứu các kho ngữ liệu song ngữ sẽ cho phép người học hiểu cơ chế sử dụng từ vựng, để đưa ra giải pháp cho nhiều vấn đề từ vựng khác nhau và chính xác hơn trong tiếng Việt bằng cách khám phá ý nghĩa, cách dùng từ và các đơn vị tương đương (cụm từ) qua các mẫu câu được lặp đi lặp lại trong việc sử dụng ngôn ngữ hàng ngày. Như vậy, có thể nói rằng, kho ngữ liệu là những công cụ và tài nguyên rất hữu ích cho việc dạy và học từ vựng nói chung và từ vựng tiếng Việt nói riêng.

Với việc tiếp cận kho ngữ liệu song ngữ như vừa phân tích ở trên, người học sẽ dễ dàng nắm bắt được cách dùng từ của người bản ngữ để sử dụng chúng một cách chuẩn xác trong những ngữ cảnh giao tiếp cụ thể. Đồng thời, người học cũng có thể nhận ra rằng không phải tất cả mọi thứ trong ngôn ngữ đều được sắp xếp ngăn nắp như đôi khi chúng được mô tả trong lớp học. Người học nắm được cách thức hoạt động thực sự của từ vựng tiếng Việt bằng cách phân tích ngôn ngữ từ kho ngữ liệu có sự đối chiếu với ngôn ngữ mà mình đang sử dụng (tiếng Anh). Trong trường hợp này, sau khi người học phải đối mặt với sự phức tạp của nhiều khía cạnh thực tế của từ vựng đang sử dụng, sẽ dần nghiệm ra những “quy tắc”, những “công thức” sử dụng từ tiếng Việt. Đây có thể được xem là sự thúc đẩy học tập theo hướng khám phá. Học ngôn ngữ dựa trên ngữ liệu cũng góp phần phát triển một số kỹ năng cơ bản và thúc đẩy người học chủ động hơn trong việc tự học. Hơn nữa, phân tích ngữ liệu có thể được cho là lý tưởng trong việc chuẩn bị cho người học thâm nhập vào cuộc sống thực tế khi giao tiếp với người bản ngữ. Người học rất tích cực trong các hoạt động này và được khuyến khích khám phá ngôn ngữ bằng mọi cách rồi tự mình đưa ra kết luận. Trong một số trường hợp, kho ngữ liệu cho phép người học phân tích ý nghĩa văn bản và ngữ cảnh tình huống sử dụng từ. Điều này cung cấp cho người học một bức tranh thực tế hơn về cách sử dụng ngôn ngữ nói chung và từ vựng nói riêng. Với việc sử dụng phương pháp tiếp cận dựa trên ngữ liệu để nghiên cứu từ vựng, người học nhận thức được tầm quan trọng của ngữ cảnh trong giao tiếp. Ngoài ra, người học có thể cảm thấy rằng họ đang tiếp xúc với việc sử dụng ngôn ngữ trong ngữ cảnh thực, họ thực sự có thể nghe thấy những người thật đang nói/viết trong một số ngữ liệu.

Tóm lại, chúng ta có thể thấy rằng các đóng góp của kho ngữ liệu nói chung và kho ngữ liệu song ngữ Anh – Việt nói riêng vào việc giảng dạy và nghiên cứu từ vựng là rất đáng chú ý. Kho ngữ liệu đã mang ngôn ngữ của đời sống thực vào lớp học và đề cao tầm quan trọng của việc học tập một cách tự chủ. Qua việc giới thiệu ba hoạt động dạy – học từ vựng tiếng Việt nêu trên, chúng tôi hi vọng kết quả nghiên cứu của bài báo sẽ có những đóng góp hữu ích trong việc đổi mới phương pháp nhằm nâng cao chất lượng dạy – học từ vựng tiếng Việt cho người nước ngoài. Tuy vậy, tiềm năng sử dụng, khai thác kho ngữ liệu song ngữ Anh – Việt trong việc dạy học ngoại ngữ là rất lớn nên vấn đề ứng dụng kho ngữ liệu vào việc dạy học tiếng Việt nói

chung, việc dạy học tiếng Việt cho người nước ngoài nói riêng vẫn còn nhiều khía cạnh cần nghiên cứu, phân tích để hoạt động giảng dạy và học tập hiệu quả hơn.

## TÀI LIỆU THAM KHẢO

1. Oxford Advanced Learner's Dictionary With Vietnamese Translation (2014), Nhà xuất bản Trẻ.
2. Sinclair, J. (1991), *Corpus Concordance Collocation*, HK: Oxford University Press.
3. Đinh Điền (2018), *Sách chuyên khảo Ngôn ngữ học ngữ liệu*, Nhà xuất bản Đại học Quốc gia Thành phố Hồ Chí Minh.
4. Alcántar Díza Carlota de Jesús, José Luis Quintero Carrillo, “Advantages of using corpora to teach English”, truy cập từ: <https://cenedic.uco.es/fieel/pdf/1.pdf>, ngày truy cập: 01/06/2021.
5. Thomas Cobb and Alex Boulton (2015), “Classroom applications of corpus analysis”, truy cập từ: [https://www.researchgate.net/publication/280856533\\_Classroom\\_applications\\_of\\_corpus\\_analysis](https://www.researchgate.net/publication/280856533_Classroom_applications_of_corpus_analysis), ngày truy cập: 16/05/2021.
6. Đào Hồng Thu (2009), “Sử dụng khối tư liệu giảng dạy tiếng Việt như một ngoại ngữ”, truy cập từ: <https://corling.wordpress.com/2009/12/30/>, ngày truy cập: 15/05/2021.
7. Đinh Điền, Lý Ngọc Minh (2015), “Ứng dụng Ngữ liệu Song ngữ Anh-Việt trong Giảng dạy Ngôn ngữ”. *Kỷ yếu hội thảo Liên ngành NNH Ứng dụng & Giảng dạy Ngôn ngữ*, 11/2015, Huế.
8. Đinh Điền, Hồ Xuân Vinh (2016), “Ứng dụng Kho Ngữ liệu trong việc dạy tiếng Việt cho người nước ngoài”. *Kỷ yếu hội thảo quốc tế Giảng dạy, Nghiên cứu Việt Nam học và tiếng Việt*, 1/2016, Bình Châu, Bà Rịa – Vũng Tàu, tr. 172-180. Huế.