

NGÔN NGỮ HỌC VÀ VIỆT NGỮ HỌC

ĐỘ KHÓ CỦA TỪ VỰNG TRONG CÁC GIÁO TRÌNH DẠY TIẾNG VIỆT CHO NGƯỜI NƯỚC NGOÀI

NGUYỄN THỊ THANH VÂN*

TÓM TẮT: Từ vựng giữ vị trí quan trọng trong công tác biên soạn giáo trình, đặc biệt là giáo trình dạy tiếng Việt cho người nước ngoài. Yêu cầu đặt ra cho người biên soạn phải có sự sàng lọc, đo lường kĩ càng, nhằm mục đích cung cấp được lượng từ vừa phải, đáp ứng được yêu cầu và trình độ của người học. Trong bài viết này, chúng tôi tập trung khảo sát, thống kê, phân loại vốn từ vựng được cung cấp trong các giáo trình dạy tiếng Việt cho người nước ngoài từ trình độ sơ cấp đến cao cấp theo các tiêu chí liên quan đến cấu tạo từ và sự thông dụng của từ nhằm xác định được độ khó vốn từ vựng giữa các cấp bậc học, đồng thời đề xuất thang đo độ khó của từ làm cơ sở cho việc hệ thống hóa lượng kiến thức, tạo thuận lợi trong công tác biên soạn, kiểm tra và đánh giá.

TỪ KHOẢ: từ vựng; cấu tạo từ; độ thông dụng; độ khó từ vựng; thang đo độ khó.

NHẬN BÀI: 9/12/2021.

BIÊN TẬP-CHỈNH SỬA-DUYỆT ĐĂNG: 6/4/2022

1. Đặt vấn đề

Những năm gần đây, việc dạy tiếng Việt cho người nước ngoài ngày càng phổ biến, nhiều trung tâm dạy tiếng Việt từ đó ra đời, trải dài từ Bắc đến Nam. Các đối tượng học viên theo học tiếng Việt rất đa dạng, xuất phát từ các châu lục khác nhau với mục đích học tập khác nhau. Theo xu hướng đó, các công trình nghiên cứu về việc dạy và học tiếng Việt như một ngoại ngữ lần lượt xuất hiện và nhận được nhiều sự quan tâm của các nhà chuyên môn cũng như độc giả. Trong khi việc biên soạn giáo trình được xem là vấn đề cốt lõi, là tiền đề ảnh hưởng đến kết quả đào tạo và khâu sàng lọc vốn từ đưa vào các giáo trình là hết sức quan trọng song vẫn chưa có kết quả khoa học nào đánh giá về độ khó vốn từ trong các giáo trình dạy tiếng Việt cho người nước ngoài cũng như xây dựng thang đo độ khó của từ, vì vậy nội dung chúng tôi khai thác trong bài viết này là cần thiết và mới mẻ.

2. Kết quả khảo sát vốn từ trong các giáo trình dạy tiếng Việt cho người nước ngoài**2.1. Phương pháp và ngữ liệu khảo sát**

Phương pháp miêu tả, phân tích, so sánh được chúng tôi áp dụng nhằm phân tích và xác định độ khó của vốn từ trong các giáo trình dạy tiếng Việt dành cho người nước ngoài, từ đó phản ánh các hiện trạng và đề xuất hướng giải quyết.

Để phục vụ cho việc khảo sát vốn từ, nguồn ngữ liệu được chúng tôi sử dụng là phần Từ vựng của bài đọc và Bảng từ chung của 3 bộ giáo trình dạy tiếng Việt cho người nước ngoài trình độ sơ cấp đến cao cấp được sử dụng phổ biến hiện nay: [1] *Tiếng Việt 123 (Tiếng Việt cho người nước ngoài - Vietnamese for beginners)* (Tập thể giáo viên Vietnamese, 2020), Nxb Thế giới; [2] *Tiếng Việt vui* (Nguyễn Minh Thuyết (Chủ biên), 2015), quyển 3 và quyển 4 (sách thử nghiệm), Nxb Giáo dục; [3] *Thực hành tiếng Việt - Trình độ C* (Đoàn Thiện Thuật (Chủ biên), 2014), Nxb Thế giới.

Tổng số từ khảo sát trong ba bộ giáo trình là 2.396 đơn vị, bao gồm 791 từ đơn đơn tiết, 13 từ đơn đa tiết, 1.530 từ ghép, 62 từ láy.

2.2. Khảo sát độ khó vốn từ trong các giáo trình dạy tiếng Việt cho người nước ngoài xét trên phương diện cấu tạo từ

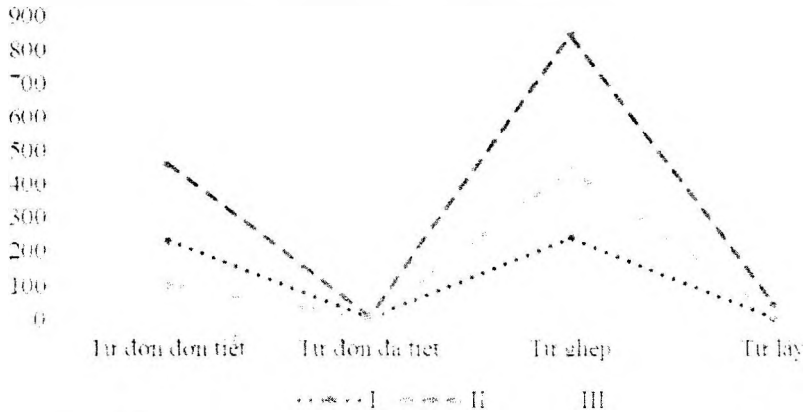
Các loại cấu tạo từ được khảo sát trong các giáo trình gồm: Từ đơn đơn tiết (ví dụ: áo, ăn, đẹp...), từ đơn đa tiết (ví dụ: cà phê, cà vạt, xích lô...), từ ghép (ví dụ: tàu hỏa, thị lợn, khen thưởng...), từ láy (ví dụ: may mắn, thắc mắc, lung tung...). Sở dĩ chúng tôi tiến hành thực hiện khảo sát theo hướng cấu tạo từ vì nhận thấy rằng số lượng từ và âm tiết ảnh hưởng rất nhiều đến việc học từ vựng. Đây cũng là một trong các tiêu chí ảnh hưởng đến độ khó vốn từ. Đối với người học khi tiếp cận với một ngôn ngữ mới, họ sẽ dễ thuộc và ghi nhớ các từ đơn âm so với các từ đa âm hoặc các từ có kết cấu phức tạp như từ ghép, từ láy, thậm chí là cả cụm từ.

Bảng 2.2.1. Thống kê số lượng từ trên phương diện cấu tạo trong các giáo trình dạy tiếng Việt cho người nước ngoài

* Trường Đại học Văn Lang; Email: van.ntt@vlu.edu.vn

Kí hiệu	Trình độ	Giáo trình	Số từ được cung cấp	Cấu tạo từ			
				Đơn đơn tiết	Đơn đa tiết	Ghép	Láy
I	Sơ cấp	[1]	481	232	5	240	4
II	Trung cấp	[2]	1.360	460	8	846	46
III	Cao cấp	[3]	555	99	0	444	12

Để thể hiện rõ tình hình cung cấp từ vựng trên phương diện cấu tạo trong ba bộ giáo trình, từ các số liệu có được, chúng tôi thể hiện kết quả dưới hình 2.2.1.



Hình 2.2.1. Biểu đồ phân bố vốn từ trên phương diện cấu tạo trong các giáo trình dạy tiếng Việt cho người nước ngoài

Từ số liệu trên có thể nhận thấy rằng tình trạng phân bố từ ghép và từ đơn đơn tiết chiếm số lượng lớn. Trong khi đó số lượng từ láy và từ đơn đa tiết chiếm tỉ lệ rất thấp. Sự phân bố như thế là hợp lí, trên thực tế trong giao tiếp vốn từ đơn và từ ghép được vận dụng phổ biến trong hầu hết các diễn đạt, cả trong viết và nói. Vốn là những từ đơn và từ ghép, người học có thể phái sinh thành các cụm từ và câu. Khi cung cấp số lượng từ chủ yếu là từ đơn đơn tiết và từ ghép, người học dễ dàng phát triển nhận thức về ngôn ngữ khi có thể tự tách, ghép các từ với nhau để sử dụng chúng trong giao tiếp. Nhờ đó, chúng ta có thể xác định được sự vận động, đặc trưng của các sự vật, hiện tượng. Đối với từ đơn đa tiết, từ láy sơ dĩ xuất hiện rất ít trong các giáo trình dạy tiếng Việt cho người nước ngoài là do liên quan đến thực tiễn giao tiếp, ví dụ các từ đơn đa tiết như *cà vạt, căn tin, ban công, v.v* đều là các từ vay mượn từ tiếng nước ngoài và được Việt hóa để đưa vào sử dụng trong giao tiếp hoặc các từ chỉ vật, loài vật xuất hiện ở một lãnh thổ nhất định, nếu không có kiến thức ngữ cảnh sẽ khó nhận diện được các loại từ này. Về từ láy, chúng có quan hệ về mặt âm, vần, dùng để mô phỏng, gọi tả hình ảnh, âm thanh trong thực tế. Ví dụ: các từ như *dong dong, trực trặc, ngát ngưỡng, sừng sốt, rành rành,...* đòi hỏi chúng ta phải tiếp xúc với hoàn cảnh thực tiễn mới có thể hiểu được ý nghĩa của các từ láy. Đối với người nước ngoài, từ láy thật sự là khó khăn để có thể lĩnh hội và vận dụng. Nếu khi người học thực hiện ghi nhớ một từ với hai, ba âm tiết thay vì một từ đơn một âm tiết, yêu cầu người học phải nhớ nhiều hơn, có nghĩa là độ khó của từ đó đã tăng lên gấp hai đến ba lần so với từ đơn đơn tiết. Nói như thế để thấy rằng, các đơn vị từ đơn đơn tiết và từ ghép xuất hiện nhiều trong các giáo trình dạy tiếng Việt cho người nước ngoài là dễ hiểu và hoàn toàn hợp lí.

Song về mặt định lượng, kết quả cho thấy số lượng từ trong giáo trình trung cấp cao hơn rất nhiều so với giáo trình sơ cấp và cao cấp, vì thế ở hầu hết các kiểu cấu tạo từ ở giáo trình trung cấp đều vượt trội hơn so với 2 bộ giáo trình còn lại. Từ đơn đơn tiết trong bộ trung cấp là 460 từ, gấp 2 lần so với bộ sơ cấp với 232 từ và gấp 4.6 lần so với bộ cao cấp với 99 từ, từ ghép trong bộ trung cấp có 846 từ, gấp 3.5 lần bộ sơ cấp với 240 từ và gấp gần 2 lần bộ cao cấp với 444 từ. Các cấu tạo từ còn lại có số lượng nhỏ, chúng tôi không thực hiện so sánh ở đây. Kết quả nghiên cứu trên chỉ ra rằng nếu đặt bộ giáo trình

trung cấp với khối lượng kiến thức được biên soạn trong hai quyển giáo trình cùng với bộ giáo trình sơ cấp và cao cấp nhưng nội dung chỉ tóm gọn trong một quyển thật sự thiếu sự cân bằng, theo nguyên tắc thường thấy ở các cấp bậc học, trình độ càng cao thì vốn từ cung cấp trong các giáo trình càng nhiều và khó hơn, thế nhưng trong trường hợp này, chúng không tuân theo quy tắc tăng dần độ khó, trình độ trung cấp nhưng vốn từ vượt nhiều lần so với cao cấp là bất hợp lý, số lượng từ giữa trình độ sơ cấp và cao cấp cũng chỉ chênh nhau 74 từ, hoàn toàn không có sự phân cấp rõ ràng và thống nhất ở các giáo trình dạy tiếng Việt cho người nước ngoài hiện nay.

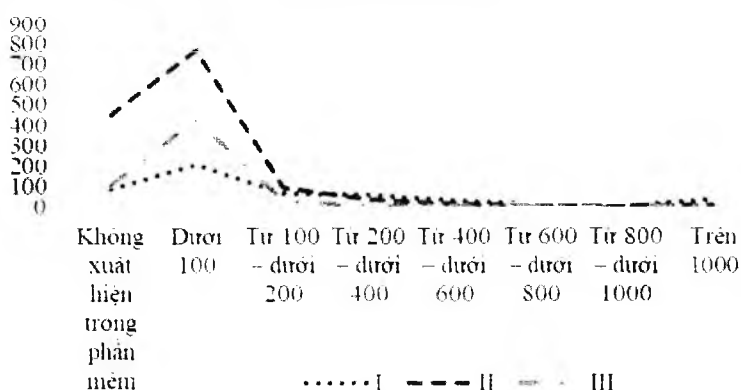
2.3. Khảo sát độ khó vốn từ trong các giáo trình dạy tiếng Việt cho người nước ngoài xét trên phương diện độ thông dụng của từ

Từ ngữ thông dụng đóng vai trò rất quan trọng trong việc hỗ trợ người học ghi nhớ và tiếp thu kiến thức, từ vựng càng thông dụng, càng phổ biến người học càng dễ tiếp cận và biến chúng thành vốn kiến thức ngôn ngữ, ngược lại từ vựng có tần số xuất hiện thấp, ít gặp trong thực tế hoặc vốn từ có nguồn gốc từ các ngôn ngữ khác, được vay mượn và Việt hóa trở thành vốn từ vựng tiếng Việt sẽ phần nào gây cản trở trong việc lĩnh hội đối với người học tiếng Việt như một ngoại ngữ. Để thực hiện khảo sát về độ thông dụng của từ, chúng tôi sử dụng phần mềm CLC PARALLEL CORPUS TOOL V2.0 của tác giả Đinh Điền và khảo sát nguồn gốc của từ qua Từ điển Hán Việt (<https://hvdic.thivien.net/>).

Bảng 2.3.1. Bảng thống kê tần suất của từ vựng trong các giáo trình dạy tiếng Việt cho người nước ngoài

Kí hiệu	Trình độ	Giáo trình	Số từ được cung cấp	Tần số xuất hiện của từ							
				Không xuất hiện trong phần mềm	Dưới 100	Từ 100 đến dưới 200	Từ 200 đến dưới 400	Từ 400 đến dưới 600	Từ 600 đến dưới 800	Từ 800 đến dưới 1000	Trên 1000
I	Sơ cấp	[1]	481	83	200	65	57	29	10	5	32
II	Trung cấp	[2]	1.360	448	767	94	31	8	4	3	5
III	Cao cấp	[3]	555	105	424	25	0	1	0	0	0

Nhằm thể hiện tổng quan về kết quả sau khi khảo sát tần suất từ trong các giáo trình, các số liệu trong bảng 2.3.1 được thể hiện qua hình 2.3.2 dưới đây:



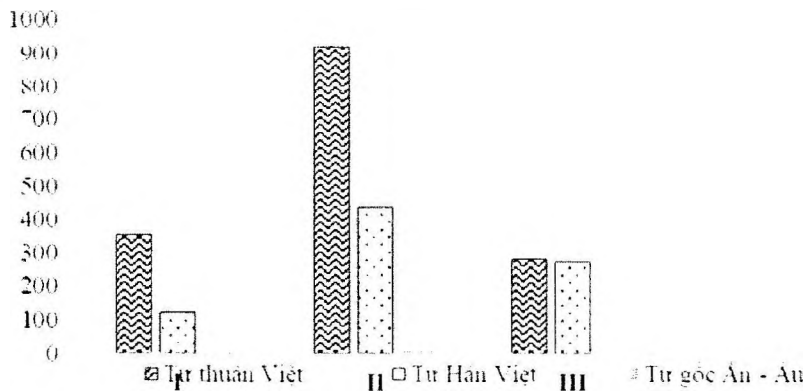
Hình 2.3.1. Biểu đồ tổng quan về tần số thông dụng của từ vựng trong giáo trình dạy tiếng Việt cho người nước ngoài

Sau khi tiến hành khảo sát độ thông dụng dựa trên tần suất của từ vựng, chúng tôi tiếp tục khảo sát độ thông dụng dựa trên nguồn gốc của vốn từ vựng trong 3 bộ giáo trình, kết quả chúng tôi thu về như sau:

Bảng 2.3.2. Bảng thống kê nguồn gốc vốn từ vựng trong các giáo trình dạy tiếng Việt cho người nước ngoài

Kí hiệu	Trình độ	Giáo trình	Số từ được cung cấp	Nguồn gốc của từ		
				Từ thuần Việt	Từ Hán Việt	Từ gốc Ấn-Âu
I	Sơ cấp	[1]	481	353	124	4
II	Trung cấp	[2]	1.360	918	434	8
III	Cao cấp	[3]	555	280	272	3

Dưới đây là biểu đồ tổng quan về nguồn gốc vốn từ vựng trong các giáo trình dạy tiếng Việt cho người nước ngoài dựa trên các dữ liệu trong bảng 2.3.2.



Hình 2.3.3. Biểu đồ nguồn gốc vốn từ vựng trong các giáo trình dạy tiếng Việt cho người nước ngoài

Tần suất dưới 100 lần chiếm số lượng tối đa trong ba bộ giáo trình, vốn từ không xuất hiện trong phần mềm chiếm tỉ lệ khá cao. Tần suất từ 100 trở lên chiếm số lượng rất thấp, sau đó giảm dần đến tần suất trên 1000 lần. Có một điểm đáng chú ý, ở bộ giáo trình sơ cấp *Tiếng Việt 123*, số lượng từ xuất hiện trên 1000 lần có xu hướng cao hơn so với bộ trung cấp và cao cấp, điều này cho thấy trong giáo trình sơ cấp, số lượng từ thông dụng khá cao, tương đối phù hợp với trình độ sơ cấp.

Về nguồn gốc của từ, ở ba bộ giáo trình, số lượng từ gốc Ấn-Âu rất thấp, chiếm tỉ lệ 0,6% (15/2.396). Trong giáo trình sơ cấp, số lượng từ Hán Việt chiếm hơn 1/3 số lượng từ thuần Việt, ở giáo trình bậc trung cấp số lượng từ Hán Việt chiếm gần 1/2 số lượng từ thuần Việt, mặt khác, ở giáo trình cao cấp, số lượng từ Hán Việt và thuần Việt có số lượng tương đương nhau. Kết quả phân bố như thế là hợp lí và phù hợp với trình độ cao cấp.

Chúng ta có thể dễ dàng nhận thấy độ dễ và khó của từ nằm ở việc sử dụng chúng trong giao tiếp, đặc biệt là đối với bản thân người mới tiếp xúc với ngôn ngữ thứ hai. Vì vậy, việc cung cấp vốn từ phải thực hiện theo nguyên tắc từ cơ bản đến nâng cao, từ những từ thông dụng đến những từ chuyên sâu vào các lĩnh vực chuyên môn. Tuy nhiên, ở mỗi cột mốc về tần suất của từ và nguồn gốc của từ, vốn từ trong giáo trình trung cấp đều nhiều hơn hẳn bộ sơ cấp và cao cấp, điều này đồng nghĩa người học không được tiếp xúc lượng từ có tần suất ổn định từ dễ đến khó.

3. Thang đo độ khó từ vựng

Để xây dựng được thang đo phù hợp, cần phải xuất phát từ đối tượng người học và yêu cầu chuẩn đánh giá năng lực từ vựng tiếng Việt 6 bậc dựa trên Khung năng lực tiếng Việt và Khung tham chiếu chung cho ngôn ngữ Châu Âu (CEFR):

Bảng 3.1. Chuẩn đánh giá năng lực từ vựng tiếng Việt

Bậc 6	Có khả năng làm chủ vốn từ ngữ rất rộng, bao gồm cả thành ngữ và tục ngữ; có khả năng sử dụng và nhận thức ý nghĩa hàm ẩn; Sử dụng và kiểm soát vốn từ vựng nhất quán, chính xác, phù hợp.
Bậc 5	Có khả năng làm chủ vốn từ vựng rộng để có thể khắc phục được những từ ngữ còn

	chưa biết hoặc quên bằng những từ ngữ đã được thụ đắc trước đó; làm chủ được các thành ngữ và tục ngữ; Khả năng kiểm soát từ vựng tốt các chủ đề thuộc nhiều lĩnh vực chuyên môn nhưng có mắc một số lỗi không đáng kể về từ vựng.
Bậc 4	Có vốn từ vựng để trình bày các chủ đề liên quan đến lĩnh vực chuyên môn; có khả năng sử dụng các từ ngữ đồng nghĩa để tránh lặp từ vựng; Có khả năng kiểm soát từ vựng ở mức độ chính xác cao, nhưng vẫn còn nhầm lẫn từ vựng, giải thích dài dòng nhưng không gây trở ngại trong giao tiếp.
Bậc 3	Có vốn từ vựng để trình bày về các chủ đề liên quan đến cuộc sống hằng ngày (như gia đình, sở thích, điều quan tâm, việc làm, du lịch, và các sự kiện hiện tại); Có khả năng kiểm soát tốt vốn từ vựng về các chủ đề gần gũi với cuộc sống hằng ngày, nhưng vẫn xảy ra lỗi khi thể hiện suy nghĩ phức tạp hoặc khi trình bày các chủ đề không quen thuộc.
Bậc 2	Có vốn từ ngữ đủ để thực hiện nhu cầu giao tiếp cơ bản, thường xuyên, hằng ngày liên quan đến các tình huống và chủ đề quen thuộc; Có thể kiểm soát một số vốn từ hẹp đáp ứng được nhu cầu giao tiếp cụ thể, hằng ngày.
Bậc 1	Có vốn từ ngữ cơ bản liên quan đến tình huống giao tiếp cụ thể; Khả năng kiểm soát từ vựng hạn chế.

(Nguồn: Hồ Trần Ngọc Oanh (2019), *Xây dựng thang đo đánh giá năng lực từ vựng tiếng Việt cho học sinh người dân tộc thiểu số*. Educational Sciences, 2019, Volume 64, Issue 7)

3.1. Các yếu tố ảnh hưởng đến độ khó từ vựng

Việc dạy từ vựng cho người học là tất yếu trong quá trình giảng dạy ngôn ngữ, tuy nhiên làm thế nào để đánh giá được độ khó của từ vựng tiếng Việt để có thể đem đến cho người học lượng từ vừa phải với trình độ và phản ánh được năng lực tiếp thu và sử dụng từ vựng của họ. Dựa trên nghiên cứu của các nhà khoa học đi trước, nhóm chúng tôi xác định các yếu tố ảnh hưởng đến độ khó của từ bao gồm: Phương thức cấu tạo từ tiếng Việt; Nghĩa của từ và Độ thông dụng của từ.

Bảng 3.1.1. Các yếu tố ảnh hưởng đến độ khó từ vựng

Yếu tố	Mô tả	Độ khó tăng dần
1. Phương thức cấu tạo từ tiếng Việt	Bao gồm đa dạng các phương thức: đơn, ghép, láy. Người học có khả năng nhận diện được các phương thức và phái sinh từ	Từ đơn - Từ ghép - Từ láy
2. Nghĩa của từ	Gồm: Từ mang ý nghĩa từ vựng (thực từ) và từ mang ý nghĩa ngữ pháp (hư từ), giúp người học phát triển khả năng kết nối giữa hình thức từ và nghĩa của từ; kết nối từ với khái niệm và sở chỉ; sắp xếp, tổ chức, kết hợp các từ trong cụm và câu	Thực từ - Hư từ
3. Độ thông dụng của từ	Bao gồm tần số xuất hiện của từ và nguồn gốc của từ. Người học được tiếp cận với vốn từ thông dụng, phổ biến đến vốn từ chuyên sâu trong các lĩnh vực, đồng thời có khả năng nhận diện sử dụng từ vựng phù hợp bối cảnh giao tiếp cụ thể	Tần suất Từ toàn dân - Từ thông dụng Tần suất Từ thuần Việt - Từ vay mượn Tần suất từ Từ địa phương - Từ chuyên biệt Tần suất Từ cổ - Từ mới

3.2. Các chỉ số về độ khó của từ

Sau khi xác định các yếu tố ảnh hưởng đến độ khó của từ và tiêu chuẩn đánh giá với mỗi cấp bậc, bước tiếp theo cần xây dựng các chỉ số về độ khó của từ, tương ứng với mỗi chỉ số là các bộ tiêu chí biểu hiện đáp ứng các chỉ số đó với mục đích phân biệt và đánh giá được khả năng đáp ứng của từ

đối với người học.

Các chỉ số này cho thấy mức độ khó của từ được diễn ra như thế nào, từ đó người biên soạn giáo trình có thể quan sát và đưa ra đánh giá về khối lượng kiến thức cung cấp cho người học.

Bảng 3.2.1. Các chỉ số về độ khó của từ

Yếu tố	Chỉ số về độ khó
1. Phương thức cấu tạo từ tiếng Việt	1.1. Từ có cấu tạo đơn giản, chỉ có một âm tiết. 1.2. Từ có cấu tạo phức tạp, có hai âm tiết trở lên
2. Nghĩa của từ	2.1. Từ mang ý nghĩa từ vựng (thực từ) 2.2. Từ mang ý nghĩa ngữ pháp (hư từ)
3. Độ thông dụng của từ	3.1. Tần suất Từ toàn dân - Từ thông dụng 3.2. Tần suất Từ thuần Việt - Từ vay mượn 3.3. Tần suất từ Từ địa phương - Từ chuyên biệt 3.4. Tần suất Từ cổ - Từ mới

3.3. Thang đo độ khó từ vựng

Sau khi xác định được các yếu tố ảnh hưởng đến độ khó của từ và các chỉ số về độ khó, chúng tôi tiến hành xây dựng thang đo độ khó từ vựng dựa trên các kết quả đã phân tích. Về yếu tố độ thông dụng của từ, chúng tôi chọn mốc tần suất 100 để làm chuẩn phân loại mức độ khó, sở dĩ có sự chọn lựa như vậy là vì thông qua khảo sát các ngữ liệu trong các giáo trình dạy tiếng Việt cho người nước ngoài trên phần mềm CLC PARALLEL CORPUS TOOL V2.0 về tần suất từ, kết quả cho thấy có trường hợp từ không xuất hiện trong phần mềm và có từ thậm chí xuất hiện trên 3.000 lần. Đa số các từ có tần suất trên 100 lần được chúng tôi đánh giá là thông dụng, xuất hiện nhiều trong giao tiếp ở các chủ đề quen thuộc như âm thực, chào hỏi, gia đình... (ví dụ: *món, tên, con trai, bố...*), các từ xuất hiện dưới 100 được liệt vào nhóm ít thông dụng hơn, đa số các từ này bị giới hạn ở một số chủ đề như phong tục, lễ Tết, thể thao... (ví dụ: *truyền thống, hôn lễ, câu thù, khăn dài...*).

Bảng 3.3.1. Mức độ biểu hiện của các chỉ số thể hiện độ khó từ vựng tiếng Việt dành cho người nước ngoài

Yếu tố	Chỉ số	Xếp loại	Mức độ		
			1	2	3
1. Phương thức cấu tạo từ tiếng Việt	1.1. Từ có cấu tạo đơn giản, chỉ có một âm tiết.	1	Từ có một âm tiết, rất dễ ghi nhớ khi vừa tiếp xúc.	Từ có một âm tiết, có thể ghi nhớ khi vừa tiếp xúc.	Từ có một âm tiết, khó ghi nhớ khi vừa tiếp xúc.
	1.2. Từ có cấu tạo phức tạp, có hai âm tiết trở lên	2	Từ được xây dựng theo phương thức ghép (chính phụ, đẳng lập), phương thức láy (láy hoàn toàn, láy bộ phận), có hai âm tiết trở lên, rất dễ ghi nhớ khi vừa tiếp xúc.	Từ được xây dựng theo phương thức ghép (chính phụ, đẳng lập), phương thức láy (láy hoàn toàn, láy bộ phận), có hai âm tiết trở lên, có thể ghi nhớ khi vừa tiếp xúc.	Từ được xây dựng theo phương thức ghép (chính phụ, đẳng lập), phương thức láy (láy hoàn toàn, láy bộ phận), có hai âm tiết trở lên, khó ghi nhớ khi vừa tiếp xúc.
2. Nghĩa của từ	2.1. Từ mang ý nghĩa từ vựng	1	Từ mang nghĩa cụ thể chỉ sự vật, hiện tượng gần gũi trong cuộc sống.	Từ mang nghĩa chỉ sự vật, hiện tượng trong cuộc sống, giới hạn ở một vài lĩnh vực.	Từ mang nghĩa trừu tượng, chuyên sâu ở một lĩnh vực nhất định.

Yếu tố	Chỉ số	Xếp loại	Mức độ		
			1	2	3
	2.2. Từ mang ý nghĩa ngữ pháp	2	Từ mang chức năng ngữ pháp, thường xuyên xuất hiện trong diễn đạt.	Từ mang chức năng ngữ pháp, ít xuất hiện trong diễn đạt.	Từ mang chức năng ngữ pháp, hiếm gặp trong diễn đạt.
3. Độ thông dụng của từ	3.1. Tần suất Từ toàn dân - Từ thông dụng	1	Từ có tần số xuất hiện trên 100.	Từ có tần số xuất hiện dưới 100.	Từ không xuất hiện trong từ điển tần số tiếng Việt.
	3.2. Tần suất Từ thuần Việt - Từ vay mượn	2	Từ có tần số xuất hiện trên 100.	Từ có tần số xuất hiện dưới 100.	Từ không xuất hiện trong từ điển tần số tiếng Việt.
	3.3. Tần suất từ Từ địa phương - Từ chuyên biệt	3	Từ có tần số xuất hiện trên 100.	Từ có tần số xuất hiện dưới 100.	Từ không xuất hiện trong từ điển tần số tiếng Việt.
	3.4. Tần suất Từ cổ - Từ mới	4	Từ có tần số xuất hiện trên 100.	Từ có tần số xuất hiện dưới 100.	Từ không xuất hiện trong từ điển tần số tiếng Việt.

Quy ước cách tính điểm: Độ khó của từ = (Xếp loại chỉ số x Mức độ)/tổng chỉ số được tính.

Bảng 3.3.2. Bảng ma trận xác định số điểm của chỉ số

Xếp loại	Mức độ		
	1	2	3
1	1	2	3
2	2	4	6
3	3	6	9
4	4	8	12

Bảng 3.3.3. Bảng phân loại độ khó của từ

Mức độ khó	Điểm ưu tiên
Thấp	Từ 1 đến 3
Trung bình	Từ 4 đến 6
Cao	Từ 7 đến 8

4. Kết luận

Khi thực hiện khảo sát vốn từ trong giáo trình từ sơ cấp đến cao cấp, chúng ta có thể thấy rằng giữa các giáo trình này có sự khác nhau trong việc cung cấp vốn từ kể cả về định lượng cũng như định tính. Về mặt định lượng, giáo trình trung cấp lại có số lượng từ nhiều hơn trình độ cao cấp, lí do là nội dung kiến thức trình độ trung cấp được biên soạn thành nhiều bài học và phân bố trong hai quyển giáo trình, nội dung bài học ở trình độ sơ cấp và cao cấp có số lượng ít hơn, chỉ được phân bố trong một quyển, tình trạng như vậy chưa đảm bảo được khối lượng kiến thức đưa đến người học. Tuy nhiên về mặt định tính, độ thông dụng của vốn từ trong giáo trình sơ cấp vẫn cao hơn hai trình độ còn lại, nguồn gốc vốn từ ở trình độ cao cấp cũng chiếm đa số các từ Hán Việt so với trình độ sơ cấp và trung cấp. Xét về mặt thông dụng của từ, sự phân bố như vậy tương đối phù hợp với kế hoạch giảng dạy.

Thang đo độ khó từ vựng tiếng Việt là thang đo khả năng đáp ứng của từ tiếng Việt đối với người học tiếng Việt như một ngoại ngữ căn cứ trên các chuẩn đánh giá 6 bậc được căn cứ theo khung tham chiếu chung châu Âu (CEFR) và khung năng lực tiếng Việt dành cho người nước ngoài (ngày 1 tháng 9 năm 2015). Với các yếu tố, chỉ số và các mức độ biểu hiện cụ thể, những đề xuất xây dựng thang đo phần nào phản ánh khối lượng kiến thức mà các tác giả đem đến cho người học, là căn cứ hữu ích

trong việc đo lường nội dung biên soạn các giáo trình dạy tiếng Việt cho người nước ngoài.

Trên thực tế, giáo trình dạy tiếng Việt cho người nước ngoài cần phải được đặt vào một hệ thống để đánh giá và xem xét mức độ phù hợp từ trình độ sơ cấp đến cao cấp. Khi thực hiện biên soạn một cách riêng lẻ không tránh khỏi nội dung kiến thức có nhiều lỗ hổng, đặt biệt công tác biên soạn phải quan tâm đến số lượng giáo trình cho từng cấp bậc, số bài học và vốn từ trong từng trình độ.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Bộ Giáo dục & Đào tạo (2015), Thông tư 17/2015/TT-BGDĐT *Ban hành khung năng lực tiếng Việt dùng cho người nước ngoài.*
2. Đinh Điền (2018), *Ngôn ngữ học ngữ liệu.* Nxb Đại học Quốc gia, Thành phố Hồ Chí Minh.
3. Đinh Điền, & Hồ 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 – Những vấn đề lý thuyết và thực tiễn, ngày 15-17/01/2016, Tp. HCM. Nxb Đại học Quốc gia Tp. Hồ Chí Minh.
4. Đỗ Hữu Châu (2004), *Giáo trình từ vựng học tiếng Việt.* Nxb Đại học Sư phạm.
5. Hồ Trần Ngọc Oanh (2019), *Xây dựng thang đo đánh giá năng lực từ vựng tiếng Việt cho học sinh người dân tộc thiểu số.* Educational Sciences, 2019, Volume 64, Issue.
6. Mai Xuân Huy, & Hứa Ngọc Tân (2017), "Về việc giảng dạy và biên soạn giáo trình tiếng Việt cho người nước ngoài". Tạp chí *Ngôn ngữ*, số 3.
7. Nguyễn Thị Như Diệp (2020), *Các yếu tố ngôn ngữ ảnh hưởng đến độ khó của văn bản tiếng Việt (Đối chiếu với tiếng Anh).* Luận án tiến sĩ Ngôn ngữ học.
8. Trần Trọng Nghĩa (2018), *Bàn về độ khó trong bài tập tiếng Việt cho người nước ngoài.* Kỷ yếu: Giảng dạy, nghiên cứu Việt Nam học và tiếng Việt, Nxb Đại học Quốc gia TP HCM.

Tiếng Anh

10. Cervetti, G.N., Hiebert, E.H., Pearson, P.D., & McClung, N.A. (2015), *Factors that influence the difficulty of science words.* Journal of Literacy Research, 47, tr.153-185.
11. Syeda, A. I., & Syeda, A. K. (2017), *Analyzing the Effectiveness of Vocabulary Knowledge Scale on Learning and Enhancing Vocabulary through Extensive Reading.* Canadian Center of Science and Education. Vol. 10. No. 9.

Difficulty level of vocabulary in Vietnamese teaching textbooks for foreigners

Abstract: Vocabulary plays a crucial role in the compilation of textbooks, especially in teaching Vietnamese to foreigners. The requirements for textbook compilers are that vocabulary must be carefully screened and measured in order to provide a moderate amount of words to meet the levels of learners. In this article, surveys, statistics, and vocabulary classification were conducted in Vietnamese language teaching textbooks for foreigners from beginner to advanced level according to criteria of word structure and word popularity to determine the difficulty of vocabulary between learning levels. Simultaneously, a scale of difficulty level of vocabulary is also proposed as a basis for systematizing the amount of knowledge, which facilitates the editing work, testing and evaluation.

Key words: vocabulary; word structure; popularity; difficulty level of vocabulary; a scale of difficulty level of vocabulary.