

VÀI NÉT VỀ TỪ ĐIỂN PHƯƠNG NGỮ VÀ VIỆC BIÊN SOẠN TỪ ĐIỂN PHƯƠNG NGỮ Ở VIỆT NAM

TRỊNH CẨM LAN *

Tóm tắt: Để tiến tới xây dựng cơ sở lý thuyết và phương pháp biên soạn từ điển phương ngữ, trên cơ sở nghiên cứu một số lý thuyết về từ điển phương ngữ cũng như khảo sát một số cuốn từ điển phương ngữ tiêu biểu của thế giới và Việt Nam, bài viết làm rõ tầm quan trọng của việc biên soạn từ điển phương ngữ nói chung và việc biên soạn từ điển phương ngữ ở Việt Nam nói riêng. Bên cạnh đó, bài viết cũng gợi ý một định hướng cho việc biên soạn một bộ từ điển phương ngữ bao phủ phạm vi toàn quốc trong tương lai với sự hỗ trợ của công nghệ số.

Từ khóa: Từ điển phương ngữ, cấu trúc vĩ mô, cấu trúc vi mô.

Abstract: In order to develop a theoretical basis and method of compiling dialect dictionaries, on the basis of studying a number of theoretical documents on dialect dictionaries as well as surveying a number of typical dialect dictionaries in the world and in Vietnam, the paper clarifies the importance of the compilation of dialect dictionaries in general and the compilation of dialect dictionaries in Vietnam in particular. Besides, the paper also suggests an orientation for compiling a dialect dictionary that covers the whole country in the future with the support of digital technology.

Keywords: Dialect dictionaries, macro-structure, micro-structure.

1. Đặt vấn đề

Sự quan tâm đến các phương ngữ Việt được cho là bắt đầu “có bóng dáng” từ cuối thế kỷ XIX qua *Đại Nam Quốc Âm tự vị* [3] của Huỳnh Tịnh Của. Trong cuốn tự vị này, tác giả đã có những quan tâm nhất định đến tình hình đa dạng trong cách phát âm một số từ tiếng Việt ở các vùng miền (ví dụ: *chánh - chính, nhất - nhứt, phúc - phước, thực - thật - thiệt,...*), mỗi từ đều được ghi ở một mục và giảng giải riêng. Như vậy, sự tồn tại của các biến thể phát âm địa phương này ít nhiều đã được thừa nhận và xem xét.

Ngược dòng thời gian hơn nữa, những khác biệt địa lý của tiếng Việt đã được ghi lại ít nhiều trong *Từ điển Việt - Bồ - La* (Alexandre de Rhodes, 1651), *Tự vị An Nam - Latin* (Pigneaux de Behaine, 1772) và *Nam Việt Dương hiệp tự vị* (Taberd, 1838) với sự hiện diện của nhiều từ ngữ địa phương thuộc hai phương ngữ lớn là phương ngữ Trung (PNT) như: *ả, mô, rứa, chi, chưởi, ngái, chừ,...* và phương ngữ Nam (PNN) như: *ghe, đui, cây sấu dàu, giùm, mắc, heo,...* Đặc biệt, cuốn *Nam Việt*

* PGS. TS - Trường Đại học Khoa học Xã hội và Nhân văn, Đại học Quốc gia Hà Nội; Email: tclan.vnu@gmail.com

Dương hiệp tự vị, phần III là một Phụ lục gần 50 trang ghi chép tên gọi các loài thảo mộc Đàng Trong (Taberd, 1838, tr.621-660) được trình bày theo cấu trúc vì mô của từ điển gồm: chữ Hán - Nôm, chữ Quốc ngữ và danh pháp tiếng Latin của các loài thực vật. Có thể nói, đây là một dạng sơ khai của từ điển tiếng địa phương. Tuy nhiên, đến đầu thế kỷ XX, phương ngữ Việt mới được đề cập một cách chính thức trong nghiên cứu của L. Cadieré năm 1902. Năm 1912, H. Maspero xuất bản cuốn *Ngữ âm học lịch sử tiếng Việt - các âm đầu* và đã tạo ra những ảnh hưởng quan trọng đối với phương ngữ học tiếng Việt. Tuy nhiên, phải đến nửa sau thế kỷ XX, các nghiên cứu phương ngữ Việt mới bắt đầu nở rộ.

Việc nghiên cứu phương ngữ Việt xưa nay tập trung chủ yếu vào mấy hướng: (1) Nghiên cứu phân vùng phương ngữ; (2) Mô tả đặc trưng của các vùng phương ngữ hay một phương ngữ cụ thể; (3) Nghiên cứu phương ngữ trong mối quan hệ với lịch sử tiếng Việt; (4) Nghiên cứu phương ngữ với việc chuẩn hóa tiếng Việt; (5) Nghiên cứu, biên soạn từ điển phương ngữ (TĐPN).

Bài viết này đề cập đến hướng thứ 5 với 3 vấn đề: 1. Vai trò của TĐPN; 2. Giới thiệu một vài cuốn TĐPN lớn trên thế giới; 3. Vài nét về việc nghiên cứu, biên soạn TĐPN ở Việt Nam và một vài đề xuất.

2. Vai trò của từ điển phương ngữ

2.1. Từ điển phương ngữ là nguồn tư liệu quý để nghiên cứu lịch sử của một ngôn ngữ

Với tư cách là sách công cụ, các từ điển phương ngữ đều có những giá trị nhất định đối với giới nghiên cứu ngôn ngữ học, đặc biệt là những người nghiên cứu ngôn ngữ học lịch sử. *Từ điển phương ngữ tiếng Anh* được Heinrich Ramisch cho rằng là một tác phẩm tham khảo kinh điển về phương ngữ

tiếng Anh và nó có thể được số hóa thành cơ sở dữ liệu quan trọng cho các nghiên cứu về phương ngữ học, ngôn ngữ học lịch sử và lịch sử văn hóa (Heinrich Ramisch, 2017). Nhiều nhà nghiên cứu, biên soạn TĐPN tin rằng các từ điển này có khả năng “khôi phục, sắp xếp, phân loại và bảo tồn di sản của ngôn ngữ” [11]. TĐPN, nhất là ở những đảo phương ngữ, thổ ngữ cổ thường được giới nghiên cứu xem như những bảo tàng ngôn ngữ quốc gia mà ở đó lưu giữ, phản ánh những diễn biến lịch sử của một ngôn ngữ. Khảo sát cuốn *Từ điển tiếng Nghệ* [6] của Trần Hữu Thung và Thái Kim Đình, *Từ điển tiếng địa phương Nghệ Tĩnh* [1] do Nguyễn Nhã Bản chủ biên, chúng tôi thấy một số lượng lớn các từ địa phương Nghệ Tĩnh (hiện vẫn đang hành chức tại địa phương) mà bằng việc tái lập những tương ứng âm vị học với các ngôn ngữ tiền thân hoặc cùng gốc, đã bộc lộ hàng loạt quá trình biến đổi ngữ âm trong lịch sử tiếng Việt như: quá trình xát hóa (*b > v; *th > d/gi, *th > s/x; *tr > s; *tr > l), xát hóa và hữu thanh hóa (*ph > v; *k > g; *ch > d/gi; *tr > d/gi), quá trình hữu thanh hóa (*kh > g; *ph > b), quá trình vô thanh hóa (*g > k), quá trình mũi hóa (*b > m),... [5]. Với những biến đổi được tìm ra từ kho từ vựng tại địa phương này, có thể nói, TĐPN Nghệ Tĩnh là một bảo tàng ngôn ngữ quý giá mà ở đó, các hóa thạch ngôn ngữ đã cho ta biết những gì đã từng diễn ra trong lịch sử tiếng Việt từ hàng trăm, thậm chí hàng ngàn năm về trước.

2.2. Từ điển phương ngữ là bằng chứng của sự đa dạng phương ngữ

Sự hiện diện của TĐPN cho ta bằng chứng về sự đa dạng ngôn ngữ. TĐPN cũng chính là công cụ bảo vệ sự đa dạng này cho một tương lai không còn phương ngữ. Vì vậy, việc nghiên cứu, biên soạn TĐPN sẽ giúp bảo vệ các biến thể ngôn ngữ có nguy

cơ bị mai một vì các lý do lịch sử hoặc dân tộc học. Tư liệu nghiên cứu ngôn ngữ của các thế hệ trong một số gia đình nhiều thế hệ ở Hà Nội, Hà Tây cũ, thành phố Hồ Chí Minh và thành phố Vinh cho thấy các biến thể địa phương giờ đây chỉ còn tìm thấy trong ngôn ngữ của người già, giảm dần ở thế hệ trung niên và gần như mất hẳn ở thế hệ thanh niên [5, tr.144, 171]. Đây là một trong rất nhiều bằng chứng về quá trình san bằng phương ngữ (dialect levelling) ở Việt Nam, một quá trình mà ở nhiều quốc gia Châu Âu về cơ bản là đã kết thúc và đang lan nhanh sang các nước đang phát triển (dữ liệu Google search từ khóa “dialect levelling” vào 9h37’ ngày 13.7.2022 cho trên 220.000 kết quả, trong đó có hàng trăm cuốn sách, bài nghiên cứu hiện tượng này ở nhiều quốc gia trên thế giới). San bằng phương ngữ càng cho thấy việc biên soạn TĐPN có ý nghĩa to lớn như thế nào.

2.3. Từ điển phương ngữ là nguồn ngữ liệu bổ sung cho từ điển ngôn ngữ quốc gia

Với tư cách là nơi lưu giữ bộ sưu tập từ vựng truyền miệng của các địa phương, TĐPN là nguồn ngữ liệu quan trọng bổ sung cho các từ điển ngôn ngữ quốc gia mà vốn xưa nay người ta chủ yếu biên soạn dựa trên kho ngữ liệu thành văn. Chẳng hạn, đánh giá về *Từ điển phương ngữ tiếng Anh* (*The English Dialect Dictionary - EDD*) [15] của Wright, nhiều nhà nghiên cứu cho rằng đây là một bằng chứng sống về tiếng Anh truyền miệng. Thực tế, khi biên soạn các loại từ điển ngôn ngữ, các soạn giả thường dựa vào nguồn tài liệu thành văn. Tuy nhiên, sự phát triển lịch sử của một ngôn ngữ, bao gồm cả các phương ngữ khu vực của nó, thường có cội nguồn từ lời ăn tiếng nói của người bản địa tồn tại dưới dạng truyền miệng, chứ không phải là ngôn ngữ văn học đã được chuẩn hóa. Chính vì vậy,

sự phong phú của tư liệu truyền miệng cũng như phương pháp sưu tập tư liệu qua điền dã của Wright khi biên soạn *Từ điển phương ngữ tiếng Anh* đã được đánh giá là không thể tốt hơn [15]. Tương tự, *Từ điển tiếng Anh khu vực Hoa Kỳ* (*Dictionary of American Regional English - DARE*) cũng hầu như không ghi lại ngôn ngữ chuẩn mà ghi lại những từ ngữ từ lời ăn tiếng nói dân gian ở các vùng miền, cách phát âm khác nhau giữa các vùng đất nước mà một người có thể học từ gia đình, bạn bè và những người đồng hương hơn là từ giáo viên và sách giáo khoa [10]. Trường hợp Việt Nam, các tác giả cuốn *Từ điển phương ngữ tiếng Việt* cũng cho biết một bộ phận đáng kể trong những cơ sở dữ liệu để biên soạn từ điển là “*Các tập ghi chép điền dã từ nhiều năm của Viện Ngôn ngữ học, đặc biệt là ở phương ngữ Bắc Bộ*” hay “*Những ghi chép cá nhân của các tác giả tham gia biên soạn từ điển*” [4, tr.7]. Khảo sát sơ bộ cuốn từ điển này, chúng tôi thấy một tỉ lệ không nhỏ các mục từ với các ví dụ minh họa từ khẩu ngữ.

2.4. Từ điển phương ngữ là nguồn tri thức bách khoa về các vùng miền

Do nhiều cuốn TĐPN, ở những mức độ khác nhau, có phản ánh các vấn đề văn hóa địa phương (chủ yếu của các cộng đồng nông thôn) đang biến mất cùng với nguồn từ vựng đặc trưng đi kèm nên đôi khi chúng có đáng giá của thành phần bách khoa thư trong cấu trúc vi mô của từ điển. Tính bách khoa thư này cũng làm cho TĐPN không giống với các loại từ điển ngôn ngữ khác.

Các nhà nghiên cứu thường nhắc đến hai loại TĐPN chính là từ điển tiếng địa phương của chỉ một địa phương hay một khu vực nhỏ nào đó (như *Từ điển tiếng Huế*, *Từ điển phương ngữ Quảng Nam*) và hai là từ điển vùng/khu vực lớn hơn (như *Từ điển tiếng địa phương Nghệ Tĩnh*, *Từ*

điển từ ngữ Nam Bộ). Cả hai loại này, đặc biệt là loại thứ nhất, có bản chất thực sự là những công cụ truyền tải kiến thức, cả ngôn ngữ và phi ngôn ngữ và do đó, chúng có xu hướng “làm mờ sự khác biệt giữa từ điển và bách khoa thư” [9], bởi chúng không chỉ giải thích từ địa phương mà còn chứa đựng những tri thức bách khoa về địa phương mà các từ chuyên chở. Những người biên soạn từ điển đôi khi cung cấp thành phần bách khoa trong cấu trúc vì mô khá phong phú với các ví dụ, mô tả lịch sử, mô tả kỹ thuật, hình minh họa và có thể nhiều tri thức khác nữa,... [15].

2.5. Từ điển phương ngữ là nơi lưu giữ thông tin địa lý của các biến thể ngôn ngữ

TĐPN là một loại sách công cụ đặc thù phản ánh không gian hành chức của các từ ngữ địa phương. Những thông tin địa lý này có thể được thể hiện dưới hai hình thức: chú thích tên địa phương và thể hiện bằng bản đồ.

Thứ nhất là chú thích tên địa phương (như một mục bắt buộc trong cấu trúc vi mô):

Ví dụ:

ABOUT, *prep.*, *conj.* and *adv.* In *gen.* use. See below,

[abut, aba t, abet, abeu't.]
L />•</>. Without ; lo get about a person, see hfiow. Also

co)tj. unless : usually by aphaeresis
Bout,q.v.w.Yks. Ah wor rairly off abaght it,
Treddlehoyle Bairnsla Ann. (i860) 39; 'E's
tekken t'dthrink w'ile 'e can't do about it (F. P.
T.). Lan. Aw cannot tell lies abeawt aw say
'at he's a pratty un, Waugh Ozvd Bodle 255.
Chs.^ To get about a person, is to get without
him, to get rid of him. Stf.' Abawt.[15, tr.8].

Thông tin địa lý được chú thích bao gồm: Yks. (Yorkshire), Lan. (Lancashire), Chs. (Cheshire), Stf. (Staffordshire) [15, tr.c3]. Đây là các vùng phương ngữ Anh mà

từ “about” được sử dụng với nghĩa được giải thích.

Một ví dụ khác:

ròi lằng (T. Hóa) d. Nhặng. Ròi lằng bay vo ve bên xác con chuột trương phềnh [4, tr.355].

Êênh ả d. Anh chị (V5) [1, tr.159].

Ở đây, địa phương sử dụng cách phát âm “Êênh ả” là Vùng 5 (được ghi chú là “bao gồm các huyện Hà Tĩnh” [1, tr.11].

Thứ hai là thể hiện bằng bản đồ. Đây là một trong những cách thể hiện khác với các loại từ điển khác. Các TĐPN tiếng Việt hiện hoàn toàn không dùng hình thức này mà chỉ dùng hình thức đầu tiên. Từ điển phương ngữ tiếng Anh [15] cũng không sử dụng bản đồ nhưng Từ điển tiếng Anh khu vực Hoa Kỳ [10] thì có tới hàng ngàn bản đồ hiển thị nơi mà các từ được tìm thấy trong quá trình điền dã.



Hình 1. Giao diện một trang bản đồ trong DARE được số hóa với các tùy chọn theo khu vực hành chức của từ địa phương

Tập VI của bộ từ điển này chứa hơn 1.700 bản đồ thể hiện sự phân bố các cặp/nhóm biến thể từ vựng đồng nghĩa ở các khu vực nước Mỹ [8].

Thông tin địa lý của các từ (nếu được phản ánh chính xác) thì không chỉ có ý nghĩa trong nghiên cứu ngôn ngữ nói chung, lịch sử ngôn ngữ nói riêng mà còn là một kênh tham khảo quý giá cho các nhà nghiên cứu sử học, dân tộc học về những quá trình vận động lịch sử, xã hội, những

cuộc tiếp xúc, di dân của quốc gia ở những địa phương, những vùng địa lý nhỏ hẹp mà đôi khi sử sách không ghi chép lại hết được.

3. Giới thiệu một vài bộ/cuốn từ điển phương ngữ lớn

TĐPN của bất kỳ quốc gia nào cũng đều có hai loại. Loại thứ nhất là từ điển bao phủ toàn đất nước. Đây là từ điển đa phương ngữ (multi-dialects), thường có dung lượng lớn (từ vài tập đến hàng chục tập) và việc chú thích thông tin địa lý của các mục từ đôi khi rất quan trọng và phức tạp. Loại thứ hai là từ điển về một khu vực nào đó (có thể là một tỉnh hoặc một vài tỉnh). Đối với loại này, việc chú giải thông tin địa lý thường theo các vùng nhỏ hơn hoặc có những mục từ vắng bóng thông tin này vì người dùng sẽ tự hiểu rằng đó là từ phổ biến tại địa phương/khu vực. Loại thứ 3 mới xuất hiện gần đây, đó là loại bao gồm các từ ngữ mới phản ánh các chiều cạnh xã hội. Những từ này được gọi là mới vì chúng được thể hệ trẻ sử dụng. Vì vậy, nó là một dạng của loại thứ hai nếu xét về phạm vi mà từ điển phản ánh. Đó không phải là một phạm vi địa lý mà là một phạm vi xã hội xác định [11, tr.82-83]. Dưới đây, bài viết sẽ chỉ giới thiệu một vài bộ từ điển cỡ lớn, tiêu biểu của loại một.

3.1. Từ điển phương ngữ tiếng Anh (*The English Dialect Dictionary - EDD*)

Đây là từ điển toàn diện nhất về phương ngữ tiếng Anh do nhà phương ngữ học Joseph Wright biên soạn và được xuất bản từ năm 1898 đến năm 1905. Phiên bản giấy của cuốn từ điển bao gồm khoảng 4.600 trang. Danh sách các vùng phương ngữ được điều tra gồm Anh và Xứ Wales (58 vùng), Scotland (43 vùng), Ireland (40 vùng) [15, tr.XX].

Từ điển này gồm 80.000 mục từ được sưu tập qua nghiên cứu điền dã trên toàn bộ các khu vực nói tiếng Anh. Các từ này được sử

dụng trong khoảng thời gian từ năm 1700 đến năm 1904, rất nhiều mục từ có thông tin từ nguyên trong tiếng Anh cổ và trung cổ, xuất hiện từ khoảng 1.000 năm trước [15].

Cấu trúc vĩ mô của từ điển bao gồm: (1) Các từ địa phương tiếng Anh (không có trong ngôn ngữ văn học) tại Anh, Ireland, Scotland và Wales; (2) Các từ có trong ngôn ngữ văn học nhưng có nghĩa đặc thù địa phương (không có những từ khác với ngôn ngữ văn học về cách phát âm, nhưng không khác về nghĩa); (3) Các từ địa phương tiếng Anh của Mỹ và các thuộc địa hiện được dùng ở Vương quốc Anh và Ireland, hoặc được tìm thấy trong các sách hay các bảng chú giải về phương ngữ được in từ rất sớm [15, tr.V]; (4) Các từ tiếng Anh thông tục, tiếng lóng, tiếng Anh “kỹ thuật” của các nhóm nghề nghiệp như nông dân, thợ mỏ,... [15]¹.

Cấu trúc vi mô (cấu trúc mục từ) của từ điển gồm 8 loại thông tin: (1) Đầu mục từ (2) Thông tin địa lý; (3) Thông tin từ loại; (4) Thông tin về cách phát âm; (5) Thông tin từ nguyên; (6) Thông tin giải nghĩa; (7) Ví dụ/nguồn trích dẫn; (8) Nhận xét/bình luận hoặc tham khảo chéo [15, tr.6].

Trên cơ sở phiên bản giấy của Wright, Markus đã xây dựng phiên bản *Từ điển phương ngữ tiếng Anh trực tuyến (EDD online)*. Đây là một phiên bản số hóa của EDD được đánh giá là một sự khởi đầu mới trong phương ngữ học tiếng Anh. Dự án đã được tiến hành từ năm 2006 và hiện đã có phiên bản cập nhật thứ ba vào 2021.

3.2. Từ điển tiếng Anh khu vực Hoa Kỳ (*Dictionary of American Regional English - DARE*)

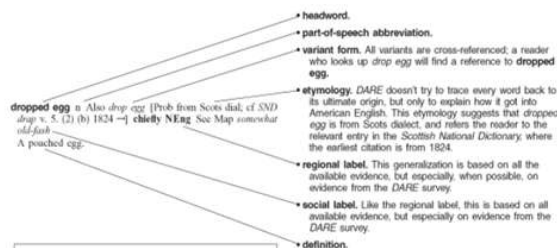
Đây là một bản ghi chép về tiếng Anh Mỹ được sử dụng tại Hoa Kỳ, từ những ngày đầu thành lập cho đến nay. Từ điển được biên

soạn trên cơ sở tư liệu thu thập từ 2.777 cuộc phỏng vấn, được thực hiện tại 1.002 cộng đồng trên khắp đất nước từ năm 1965 đến năm 1970. Ngoài ra, một số nguồn tài liệu thành văn như nhật ký, thư từ, tiểu sử, tiểu thuyết, tài liệu lịch sử, tài liệu chính phủ và báo chí cũng được thu thập [10, tr.XX]. Từ điển được đánh giá là “sự tổng hợp táo bạo của các tập bản đồ ngôn ngữ và từ điển lịch sử” [8]. Từ điển gồm 5 tập, 5.544 trang, chứa trên 60.000 từ đầu mục, gần 3.000 bản đồ cho thấy nơi các từ được tìm thấy trong quá trình nghiên cứu thực địa [12].

Phiên bản *DARE* kỹ thuật số cung cấp nhiều tính năng lựa chọn khác nhau. Người dùng có thể nhấp vào một tiểu bang trên bản đồ *DARE* hoặc chọn từ bất kỳ khu vực nào trong số 41 khu vực trên thanh công cụ kéo xuống để liên kết đến tất cả các mục từ có ở tiểu bang hoặc khu vực đó. Phiên bản *DARE* kỹ thuật số còn có hơn 5.000 đoạn âm thanh được đưa vào các mục. Người dùng có thể nhấp vào biểu tượng âm thanh để nghe các đoạn hội thoại được ghi lại từ năm 1965 đến năm 1970 với những người nói từ khắp địa phương trên nước Mỹ [12].

Cấu trúc vĩ mô của từ điển bao gồm: (1) Các từ ngữ mà hình thức (cách phát âm hoặc chính tả) hoặc ý nghĩa không được sử dụng phổ biến trên khắp đất nước mà bị hạn chế theo khu vực hoặc theo các nhóm xã hội; (2) Các từ ngữ theo cách sử dụng dân gian, truyền miệng được thu thập trong cộng đồng, không phải từ nhà trường, sách báo hay các phương tiện thông tin đại chúng; (3) Các từ ngữ lịch sử, hiện tại không còn được sử dụng (nhưng đã từng hành chức trong giao tiếp ở các địa phương) [12].

Các thông tin bao gồm: (1) Đầu mục từ; (2) Từ loại; (3) Dạng thức biến thể; (4) Từ nguyên; (5) Thông tin khu vực/địa lý; (6) Thông tin xã hội; (7) Giải nghĩa.



Hình 2. Cấu trúc mục từ trong *DARE* (nguồn: *DARE sample entry and map* (<https://dare.news.wisc.edu/dare-sample-entry-and-map>))

DARE được xem là một dạng biên niên sử bằng ngôn ngữ của người Mỹ [8]. *DARE* thực sự không phải là một cuốn từ điển, nó là một kho báu quốc gia/bảo vật quốc gia². Đồng đảo các tầng lớp trí thức như giáo viên, thủ thư, nhà nghiên cứu, bác sĩ, nhà ngôn ngữ học, nhà báo, nhà sử học, nhà viết kịch,... đã sử dụng, khai thác bộ sách này từ các góc độ, tính năng khác nhau phục vụ cho nghề nghiệp.

3.3. Từ điển phương ngữ Nhật Bản

(1) *Từ điển phương ngữ toàn Nhật Bản* (*Dialect Dictionary of All Japan*) của Misao Tojo

Khác với 2 bộ từ điển trên của Anh và Mỹ là được nghiên cứu, biên soạn từ những ghi chép trên thực địa, *Từ điển phương ngữ toàn Nhật Bản* là bộ từ điển xuất bản vào năm 1951, được hoàn thành từ năm triệu thẻ mục từ thu thập từ khoảng 500 TĐPN các khu vực, 350 tài liệu về phương ngữ các loại, bao gồm cả các báo cáo về lịch sử các địa phương [11, tr.76].

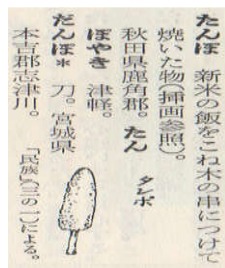
Cấu trúc vĩ mô của từ điển này, cũng như hai từ điển tiếng Anh nói trên, chỉ bao gồm từ vựng địa phương, những từ không có trong ngôn ngữ chuẩn. Các mục từ được sắp xếp theo thứ tự bảng chữ cái, theo sau là định nghĩa và tên các địa phương nơi từ được sử dụng. Tác giả từ điển rất chú trọng tập hợp những từ ngữ liên quan đến cuộc sống hàng ngày (văn hóa đời sống) của các

địa phương như trang phục, thực phẩm, dụng cụ và các đồ dùng trong sinh hoạt.

Cấu trúc vĩ mô, từ điển chỉ có 3 loại thông tin: (1) Đầu mục từ; (2) Thông tin địa lý; (3) Thông tin giải nghĩa. Tuy nhiên, cách giải nghĩa từ không chỉ bằng cách dùng từ tương đương trong vốn từ chung hay giải nghĩa bằng cách miêu tả (nếu từ không có đơn vị định danh tương đương trong vốn từ chung) mà một số từ còn được minh họa bằng hình ảnh. Yếu tố vẫn được gọi là có tính “bách khoa thư” này là cách mà nhiều TĐPN đã sử dụng.



Hình 3. Hình ảnh 'kera', một loại áo mưa (Tojo, 1951: 312, dẫn theo [11, tr.77])



Hình 4. Hình ảnh 'tampo', một loại bánh gạo được dán quanh que gỗ để nướng (Tojo, 1954: 604, dẫn theo [11, tr.77])

(2) Một bộ khác nữa là *Từ điển phương ngữ Nhật Bản* (1992-1994) của Hirayama. Bộ này gồm 6 tập, mô tả tổng thể từ ngữ của 72 địa phương ở Nhật Bản.

Cấu trúc vĩ mô của từ điển này về cơ bản giống như *Từ điển phương ngữ toàn Nhật Bản*.

Cấu trúc vi mô thì có vài nét khác biệt. Các mục từ được mô tả toàn diện về đặc điểm ngữ âm (phát âm), từ vựng (nghĩa) và ngữ pháp (khả năng kết hợp). Nét đặc biệt trong việc cung cấp mục từ của từ điển này là các từ trong ngôn ngữ chuẩn sẽ được sử dụng làm từ đầu mục. Các dạng biến thể phương ngữ (bao gồm cả biến thể ngữ âm, từ vựng, ngữ pháp,...) thu thập từ khảo sát thực địa được liệt kê dưới mỗi mục từ. Dựa trên từ

điển này, một loạt các từ điển tiếng địa phương cho mỗi tỉnh có tên *Series Language of Japan* đang được xuất bản [11, tr.78].

4. Vài nét về việc biên soạn từ điển phương ngữ ở Việt Nam và một vài đề xuất nhỏ

4.1. Một cái nhìn tổng thể

Nhìn một cách tổng thể, các TĐPN tiếng Việt cũng có 2 loại lớn.

Loại thứ nhất bao phủ các địa phương trên toàn quốc (*Từ điển phương ngữ tiếng Việt* [4]). Loại thứ 2 là từ điển từ ngữ ở một khu vực địa lý xác định. Đó có thể là một vùng, miền (*Sổ tay phương ngữ Nam Bộ*, Nguyễn Văn Ái, 1987; *Từ điển từ ngữ Nam Bộ* [7]), có thể chỉ là một vùng nhỏ hơn, hay một tỉnh (*Từ điển tiếng Nghệ Tĩnh* [6]; *Từ điển tiếng địa phương Nghệ Tĩnh* [1]; *Từ điển phương ngữ Quảng Nam*, Phạm Văn Hào, 2017;...). Dưới đây, chúng tôi sẽ trình bày kết quả khảo sát sơ bộ một vài cuốn để có cái nhìn tổng thể.

(1) *Từ điển phương ngữ tiếng Việt* [4] là từ điển đa phương ngữ, có dung lượng khoảng 14.000 mục từ được in trong 468 trang, khổ 16×24 cm.

Cấu trúc vĩ mô bao gồm các từ địa phương là biến thể về hình thức hoặc nội dung của từ trong ngôn ngữ chung hoặc các từ ngữ được dùng ở địa phương nhưng tác giả cho là không phải là các từ địa phương như *nhút*, *chẹo*, *lòn bon*, *nóp*,... [4, tr.8-9]. Ngoài ra, tác giả chủ trương không thu thập các cách nói khẩu ngữ như tiếng lóng, đặc ngữ, uyển ngữ dù đôi khi chúng mang đậm tính địa phương.

Cấu trúc vi mô của từ điển có 5 loại thông tin: (1) Đầu mục từ; (2) Thông tin địa lý; (3) Thông tin từ loại; (4) Giải nghĩa; (5) Ví dụ.

Về việc chú giải thông tin địa lý, từ điển chú thích ba vùng phương ngữ chính là Bắc

(B), Trung (T) và Nam (N). Ngoài ra, do tư liệu cho phép chủ yếu hơn về khu vực sử dụng, các tác giả “mạnh dạn chú tên tỉnh ở những từ có phạm vi sử dụng tương đối hẹp, có thể ước định được khu vực sử dụng một cách tương đối” [4, tr.9].

Ví dụ: **bầm** (HTây, PThọ) d. Mẹ (có thể dùng để xưng gọi).

Về cách định nghĩa từ, từ điển sử dụng song song 4 mô hình định nghĩa: 1. Đối với các từ có từ toàn dân tương đương: định nghĩa bằng từ toàn dân tương đương; 2. Đối với các từ chỉ có ở một phương ngữ nào đó: định nghĩa bằng điển giải; 3. Đối với các từ được dùng với các nét nghĩa riêng của từng vùng: định nghĩa bằng nêu các nét nghĩa ngay trong lời định nghĩa, hoặc chú trong ngoặc đơn, sau đó có thể đưa từ tương đương cơ bản; 4. Đối với các từ có từ đồng nghĩa, gần nghĩa: định nghĩa bằng chuyển chú.

Từ điển không có bản đồ hay hình ảnh minh họa.

(2) Một cuốn từ điển tiêu biểu cho loại thứ 2 là *Từ điển tiếng địa phương Nghệ Tĩnh* do Nguyễn Nhã Bản chủ biên. Phạm vi từ vựng trong từ điển được xác định rõ. Theo đó, từ địa phương được quan niệm là *vốn từ cư trú ở một địa phương cụ thể, có sự khác biệt với ngôn ngữ văn hóa hoặc địa phương khác về ngữ âm và ngữ nghĩa* [1, tr.6].

Cấu trúc vĩ mô bao gồm: a. Các từ địa phương là biến thể ngữ âm của từ trong ngôn ngữ văn hóa, tác giả gọi là *từ ngữ âm* như: *khải - gãi, ló - lúa, gấy - ghi - gái,...*; b. Các từ địa phương là biến thể ngữ nghĩa của từ trong ngôn ngữ văn hóa, tác giả gọi là *từ ngữ nghĩa* như (*gấy, ghi* có khi dùng với nghĩa là *gái*, có khi dùng với nghĩa là *vợ*...); c. Các tên gọi sản vật thuần túy chỉ có mặt ở địa phương (*nhút, toi, xà lì,...*).

Cấu trúc vi mô của từ điển gồm có: (1)

Đầu mục từ; (2) Thông tin từ loại; (3) Giải nghĩa; (4) Thông tin địa lý (nếu từ chỉ lưu hành ở một vài vùng nhỏ hẹp như xã, huyện).

Về việc chú giải thông tin địa lý, trong trường hợp mục từ không ghi chú gì thì đó là từ phổ biến hầu khắp cả vùng Nghệ Tĩnh. Còn những từ có phạm vi sử dụng hạn chế trong một vài địa phương nhỏ hẹp thì chú theo các địa danh khác nhau được quy định theo vùng, từ vùng 1 (V1) đến vùng 9 (V9). Các vùng ở đây chủ yếu là xã (1 hoặc nhiều xã) thuộc các huyện của Nghệ Tĩnh [1, tr.11].

Về cách định nghĩa từ, từ điển sử dụng 3 cách định nghĩa: 1. Các từ có sự tương ứng giữa tiếng địa phương với tiếng Việt văn hóa thì định nghĩa bằng đối chiếu. Vd: *tru - trâu, ló - lúa,...*; 2. Các từ có dạng thức ngữ âm gần nhau hay có từ đồng nghĩa thì chuyển chú; 3. Các từ là tên gọi những sản vật, hành động, tính chất chỉ có tại địa phương thì định nghĩa theo lối miêu tả, giải thích [1, tr.9-10].

Từ điển này cũng hoàn toàn không có bản đồ hay hình ảnh minh họa.

4.2. Một vài nhận xét

Về dung lượng, hầu hết các TĐPN của ta có dung lượng nhỏ. *Từ điển phương ngữ tiếng Việt* [4] của Phạm Văn Hảo có khoảng 14.000 mục từ, *Từ điển từ ngữ Nam Bộ* [7] của Huỳnh Công Tín khoảng 20.000 từ, *Từ điển tiếng địa phương Nghệ Tĩnh* [1] của Nguyễn Nhã Bản khoảng gần 6.000 mục từ, *Từ điển tiếng Nghệ* [6] của Trần Hữu Thung và Thái Kim Đình khoảng gần 3.000 mục từ.

Về việc lựa chọn mục từ, bên cạnh những từ địa phương như *bá cháy, chà bá, ghe bầu, giỡn mặt, mắt kiếng,...* lại có cả những từ như *anh em chú bác, anh em bạn dì, ẩm, gió nam, đèn khò, hàng xịn, lục đục, màu mỡ, ông nội, bà nội,...* [7] không phải từ địa phương Nam Bộ. Một vài từ điển vẫn có thể lọt vào một số từ mà trước được xem là

“địa phương” nhưng đã gia nhập vốn từ vựng chung từ lâu như *ăn dè, ăn chực, cào,...* [1] hay *bỏ bịch, chơi bời, lẩn quẩn, nổi điên,...* [7]. Như vậy, các từ được lựa chọn cũng chưa thật thỏa mãn những tiêu chí cơ bản.

Về cấu trúc vi mô, các thông tin về cơ bản gồm 4 hoặc 5 phần: *đầu mục từ, thông tin từ loại, giải nghĩa và ví dụ* [4]. Các phần này được sắp xếp theo trật tự nhất quán, tuy nhiên, nhiều mục từ không có ví dụ mặc dù nó được giải nghĩa bằng từ toàn dân chứ không phải bằng chuyên chú. Ngoài những thông tin như trên, mục từ cũng có thể có thêm *thông tin địa lý* [1]. Tất cả các mục từ đều có 3 hoặc 4 phần đầu giống nhau, còn 2 hoặc 1 phần sau ở tình trạng lưỡng khả.

Về mô hình định nghĩa, 3 hay 4 mô hình định nghĩa trong các từ điển về cơ bản là được tuân thủ. Có một vài cách giải nghĩa có thể gây tranh cãi (ví dụ: cách giải nghĩa từ “nhà báo” trong *Từ điển từ ngữ Nam Bộ* đã gây nhiều băn khoăn cho người sử dụng trong thời gian gần đây):

Nhà báo /na:² ba:w⁵/, [na:² ba:w⁵]
[01] (dt) (nb) người không có công
án việc làm, thất nghiệp, đang ăn
bám vào gia đình. "Ra trường rồi ở
nhà làm nhà báo chỗ có làm được
gì đâu."
Nhà cầm đồ /na:² kɿm:² do:²/.

Ngoài ra, có một cách giải nghĩa khá phổ biến trong các TĐPN với những từ ngữ định danh những sự vật, sản phẩm “đặc hữu” (ngoài cách miêu tả), đó là giải nghĩa bằng hình ảnh minh họa nhưng hầu như vắng bóng trong các TĐPN của ta.

Về cách hiển thị thông tin địa lý, các từ điển của ta chủ yếu hiển thị thông tin địa lý bằng cách chú vùng như Bắc (B), Trung (T), Nam (N), Thanh Hóa (THóa), Phú Thọ (PThọ) [4] hoặc vùng 1 (V1), vùng 8 (V8) [1] ở sau từ đầu mục hoặc cuối mục từ,...

Cách hiển thị thông tin địa lý bằng bản đồ hoàn toàn vắng bóng.

4.3. Một vài đề xuất nhỏ

Trong tương lai gần, việc biên soạn một cách có hệ thống, bài bản TĐPN tiếng Việt trên cơ sở kế thừa những thành tựu đã có là việc nên nghĩ tới. Để làm việc này một cách bài bản, cần cân nhắc thực hiện một số công việc cụ thể như:

(1) Nghiên cứu kỹ cơ sở lý thuyết và phương pháp biên soạn TĐPN.

(2) Trên cơ sở kết quả nghiên cứu đó, lập kế hoạch chi tiết các công việc:

- Tập hợp lực lượng, tập huấn người đi điền dã.

- Dự kiến cấu trúc vĩ mô và vi mô của từ điển.

- Thiết kế công cụ điều tra (bảng câu hỏi với sự thể hiện các thông tin trong cấu trúc vi mô).

- Tập huấn sử dụng các phương tiện hỗ trợ (máy ghi âm, ghi hình, các phần mềm...).

- Tập huấn xử lý các kết quả điều tra.

(3) Thiết kế cấu trúc vĩ mô và vi mô của từ điển.

Khi thực hiện việc này, cần lưu ý các nhận xét ở trên. Về dung lượng, nếu điều tra được tổng thể để có số lượng từ đầy đủ thì sẽ giúp tăng dung lượng. Về việc lựa chọn mục từ, cần xây dựng bộ tiêu chí lựa chọn từ chặt chẽ hơn. Về cấu trúc mục từ, bên cạnh việc kế thừa những thông tin đã có, nên có sự tham khảo hợp lý những thông tin mục từ có trong các TĐPN được giới chuyên môn thừa nhận, khảo cứu thông tin từ nguyên trong các từ điển cổ, tìm hiểu một số tri thức bách khoa thư liên quan đến các từ,... Về cách giải nghĩa từ, lưu ý đến cách giải nghĩa mà các từ điển của ta thiếu hụt (chẳng hạn: giải nghĩa bằng hình ảnh

trực quan). Về cách hiển thị thông tin địa lý, cần thấy hiệu quả và tính khoa học của bản đồ và tận dụng nó với sự hỗ trợ của công nghệ thông tin hiện đại.

Cuối cùng, để nâng cao hiệu quả sử dụng, việc số hóa và xuất bản dữ liệu cần được xem xét nghiêm túc.

5. Thay lời kết

Có thể tiến hành biên soạn các TĐPN theo con đường ngược lại: đi từ những từ điển khu vực đến những từ điển bao phủ cả quốc gia. Đây là cách mà người Nhật từng làm. Tất nhiên, việc nghiên cứu điền dã tổng thể để có dữ liệu bổ sung cũng vô cùng cần thiết. Trên thực tế, đặc điểm ngữ âm, từ vựng và đôi khi cả ngữ pháp ở một số địa phương đã được các học giả khác nhau mô tả và những nghiên cứu này sẽ là nền tảng đáng tin cậy để kế thừa cho một cuốn TĐPN Việt đầy đủ trong tương lai. Bởi vì TĐPN quốc gia sẽ là tổng hợp của từ vựng các phương ngữ, vì vậy, một con đường chậm nhưng chắc là chúng ta có thể bắt đầu từ việc mô tả hệ thống từ vựng ở các địa phương (kế thừa những địa phương đã có và biên soạn mới những địa phương còn thiếu) để tiến tới xây dựng từ điển từ vựng địa phương bao phủ phạm vi toàn quốc.

CHÚ THÍCH

¹ Loại thứ tư (d) không được mô tả trong Lời nói đầu của Wright nhưng được Manfred Markus nêu ra trong bản EDD số hóa.

² Nguyên văn trong *Language in Society*: DARE...is not a dictionary; it is a national treasure (Edward Callary, *Language in Society*).

TÀI LIỆU THAM KHẢO

- [1] Nguyễn Nhã Bản (Chủ biên), *Từ điển tiếng địa phương Nghệ Tĩnh*, Nxb. Văn hóa - Thông tin, Hà Nội, 1999.
- [2] Hoàng Thị Châu, *Phương ngữ học tiếng Việt*, Nxb. Đại học Quốc gia Hà Nội, Hà Nội, 2004.
- [3] Huỳnh Tịnh Paulus Của, *Đại Nam Quốc Âm tự vị*, Emprimerie REY, CURION Cie, Saigon, 1895-1896.
- [4] Phạm Văn Hào (Chủ biên), *Từ điển phương ngữ tiếng Việt*, Nxb. Khoa học xã hội, Hà Nội, 2009.
- [5] Trịnh Cẩm Lan, *Sự biến đổi ngôn từ của các cộng đồng chuyển cư đến Thủ đô - nghiên cứu trường hợp cộng đồng Nghệ Tĩnh tại Hà Nội*, Nxb. Khoa học xã hội, Hà Nội, 2007.
- [6] Trần Hữu Thung và Thái Kim Đình, *Từ điển tiếng Nghệ*, Nxb. Nghệ An, Nghệ An, 1997.
- [7] Huỳnh Công Tín, *Từ điển từ ngữ Nam Bộ*, Nxb. Chính trị Quốc gia, Hà Nội, 2009.
- [8] Adam, M., "Words of America: A Field Guide", *Humanity*, Vol.32, Issue 5, 2011.
- [9] Barbato, M. and Varvaro A., "Dialect dictionaries", *International Journal of Lexicography*, Vol.17, Issue 4, pp.429-439, 2004.
- [10] Cassidy, F. G. (Chief Editor), *Dictionary of American Regional English (DARE)*, Belknap Press of Harvard University Press, Cambridge, 1991.
- [11] Fukushima, C., "Dialect dictionary in Japan: Comparison with alcover's dictionary", *Dialectologia*, Special issue, VII, pp.73-81, 2017.
- [12] Greenbaum, S., "Review of Dictionary of American Regional English", *Journal of American Studies*, Vol.20, Issue 3, pp.462-464, 1986.
- [13] Markus, M., *English Dialect Dictionary Online: A New Departure in English Dialectology*, Cambridge University Press, Cambridge, 2021.
- [14] Trinh Cam Lan, "Ancient dialect of Area IV in two Nghe Tinh dialect dictionaries (Phonological characteristics of initial consonant system)", *31st Annual Meeting of The Southeast Asian Linguistics Society*, 2022.
- [15] Wright, J., *The English Dialect Dictionary (EDD)*, Oxford University Press, Oxford, 1898-1905.