

GÓP Ý VỀ BIÊN SOẠN TIÊU ĐỀ ĐỀ MỤC « Việt Nam - Các ngôn ngữ »

PGS. TS. VƯƠNG TOÀN

Phòng Nghiệp vụ

Viện Thông tin Khoa học Xã hội

1. Mở đầu

Công tác phân loại – biên mục tài liệu ở một thư viện, dù theo cách nào, thì cũng có đề mục *Ngôn ngữ*. Có dịp đọc bản dịch ra tiếng Việt (của Phạm Thị Lệ Hương) *Tiêu đề "Việt Nam" trong bản tiêu đề đề mục của Thư viện Quốc hội Hoa Kỳ* (của Thomas Mann), xin được góp đôi ý kiến về Bảng số 3, dành cho *Vietnam - Languages*, mà trước hết, chúng tôi muốn thay *những* bằng *các* và xin dịch là *Việt Nam - Các ngôn ngữ*. Nhân đó, cũng muốn được góp một phần cụ thể vào công việc Định chủ đề & Biên soạn Khung Tiêu đề đề mục mà ngành thư viện đang đặt ra.

Chúng ta biết rằng tên ngôn ngữ và tên dân tộc thuộc phạm vi nghiên cứu của môn tên gọi (hay danh học - onomastics). Tên ngôn ngữ thường gắn với tên một dân tộc (hay còn gọi là tộc danh - ethnonyme)¹. Một ngôn ngữ có thể được sử dụng ở nhiều nước khác nhau (tiếng Anh, tiếng Pháp, tiếng Tây Ban Nha,...) và một dân tộc có thể sử dụng các ngôn ngữ khác ngoài tiếng mẹ đẻ của mình làm ngôn ngữ chính thức, ngôn ngữ làm việc, ngôn ngữ giao tiếp chung ... nhất là trong điều kiện của một cộng đồng đa dân tộc (như tiếng Việt được dùng làm "tiếng phổ thông" cho cộng đồng các dân tộc Việt Nam).

Tên dân tộc có thể có nguồn gốc khác nhau: tên tự gọi hay được gọi. Khi xác định thành phần dân tộc, các nhà nghiên cứu có cách phân định của mình, họ thường dựa trên những cứ liệu khoa học và đưa ra một tên gọi, có thể không hoàn toàn trùng với tên tự gọi hoặc thường gọi.

Xét theo nguồn gốc, có tên tự gọi, tên do (các) dân tộc khác gọi, tên do (các) nhà nghiên cứu đặt (và / hoặc dịch ra tiếng nước ngoài) và quan trọng là tên chính thức được ghi trong các văn bản có tính pháp quy của một nhà nước. Chính do vậy mà những tên gọi này có thể thống nhất là một, song cũng có thể khác nhau.

Theo *Danh mục* tên gọi của Ủy ban Dân tộc và Miền núi² ta thấy những tên gọi này có thể thống nhất là một (Cơ ho, Mơ nông), song thường rất khác nhau (ví dụ: người *La Chí* có tên tự gọi là *Cù Tê*, và các tên gọi khác : *Thổ Đen*, *Mán Xá*; *Mông* = *Hmông*,

¹ Xem : « *Về tên gọi các ngôn ngữ* ». Trong “Những vấn đề văn học và ngôn ngữ học (30 năm thông tin các khoa học ngữ văn)”. Hà Nội, Nxb Khoa học xã hội, 2003, tr.851-871.

² *Sổ tay công tác dân tộc và miền núi*. Ủy ban Dân tộc và Miền núi & Ban Tư tưởng – Văn hoá Trung ương, Hà Nội, 2000, tr. 78-82.

Na miếu, Mẹo, Mèo, Miếu ha, Mán Trắng). Sự khác biệt ở cách gọi tên, và có khi đã có sự thay đổi, nghĩa là tên gọi mang tính lịch sử (ví dụ: *Mán* nay được gọi là *Dao*, *Thổ* trước đây nay gọi là *Tày*, và *Thổ* hiện nay chỉ một dân tộc vốn có các tên tự gọi là : *Người Nhà làng. Mường, Con Kha, Xá Lá Vàng*, và các tên gọi khác là *Kẹo, Mọn, Cuối, Họ, Đan Lai, Li Hà, Tày Poọng*, thuộc nhóm nhóm Việt – Mường).

Đó là chưa kể cách phát âm tên gọi theo tiếng dân tộc được ghi lại trên các ấn phẩm đã xuất bản trong và ngoài nước cũng có thể khác nhau. Nghĩa là không chỉ có các cách gọi mà cả cách đọc (phỏng theo ngôn ngữ dân tộc) khác nhau dẫn đến cách viết khác nhau: *Chàm* và *Chăm*, hoặc *Tiam* ; *Mông*, *Hơ mông* và *Hmông*, *H'mông*, hoặc *Mèo*, *Miao*,... ; hoặc dùng con chữ khác nhau: *Gia rai* và *Jrai* , hoặc *Jorai*; *Cơ tu*, *Cờ tu*, *Ca tu* và *Ka tu*, *Khơ me* và *Khmer*; *Ba na* và *Bahnar*; *Xinh mun* và *Ksingmul*, *Xơ tiêng* và *Stiêng*, hoặc *Stien*,,...).... Đó là chưa kể cách viết liền hay rời, có hay không dùng gạch nối, vốn còn chưa được qui chuẩn thống nhất trong phiên chuyển ngôn ngữ ở Việt Nam hiện nay.

Sự khác biệt có khi không dừng ở cấp độ ngữ âm - từ vựng như trên mà còn ở cả sắc thái tu từ (trung tính hay miệt thị). Ví dụ như khi hiện nay ta không còn dùng cách gọi người *Mường* là *Mọi* (tựa như gọi người Việt Nam là *annamite*) vì nay chúng đã mang hàm ý không tôn trọng, thậm chí là miệt thị .

Đó là chưa kể việc xác định thành phần dân tộc và ngôn ngữ là vẫn đang tiếp diễn. Do vậy những tên gọi ở các thời kỳ lịch sử khác nhau, được ghi lại trên các loại xuất bản phẩm mà các thư viện đã và đang lưu trữ có thể khác nhau nhưng lại là một, hoặc ngược lại có thể vẫn tên gọi ấy nay đã được dùng để chỉ dân tộc và ngôn ngữ khác.

Như vậy, sự khác biệt có thể mang tính lịch sử, nghĩa là tên gọi (kể cả tên chính thức) được thay đổi khi sử dụng (và ghi nhận) ở những thời điểm khác nhau. Hơn nữa, vấn đề tộc danh và cùng với nó là tên gọi các ngôn ngữ, cũng như cùng với cách phân chia ngữ hệ là tên gọi các nhóm ngôn ngữ, các ngữ hệ còn không ít ý kiến khác nhau. Do đó, con số ngôn ngữ được ghi nhận ở mỗi nước và trên thế giới không phải luôn có sự chính xác tuyệt đối. Do đó, việc định chủ đề cho tài liệu có liên quan đến ngôn ngữ và dân tộc cần phù hợp với những tên gọi sử dụng đương thời. Người làm công việc định chủ đề cho tài liệu liên quan đến Việt Nam cũng cần theo dõi những gì đã, đang và sẽ diễn ra trong việc xác định thành phần dân tộc và ngôn ngữ ở Việt Nam.

2. Tên ngôn ngữ theo cách phân định thành phần dân tộc

Theo dòng thời gian, các nhà nghiên cứu đã ba lần đưa ra danh mục thành phần dân tộc ở Việt Nam. Đó là vào các năm : 1959, 1974 và 1979³.

³ Xem : Không Diễn . *Dân số và dân số tộc người ở Việt Nam*. Hà Nội, Nxb Khoa học xã hội 1995, 315 tr.

Bản danh mục về thành phần các dân tộc ở Việt Nam được công bố tương đối đầy đủ và khá chi tiết lần đầu tiên trong cuốn *Các Dân Tộc Thiểu Số ở Việt Nam* của nhóm tác giả: Lã Văn Lô, Nguyễn Hữu Thấu, Mạc Đường... (Hà Nội, Nxb Văn hoá, 1959). Theo đó, nước ta có 64 dân tộc, thuộc ba ngữ hệ (linguistic family) khác nhau, là *Hán-Tạng*, *Môn-Khơ me*, và *Malayô-Pôlinêdiêng*. Kết quả nghiên cứu này là một trong những cơ sở, căn cứ khoa học về mặt tộc người để tiến hành cuộc Tổng điều tra dân số toàn miền Bắc (ngày 1/3/1960). Các tài liệu được công bố kể từ đó, theo danh mục này.

Thế rồi với mục đích nhằm khắc phục những nhầm lẫn và đôi chỗ sắp: xếp chưa được thỏa đáng, giới nghiên cứu đã có nhiều lần họp bàn các vấn đề cụ thể ở địa phương và trung ương, để có những điều chỉnh trong việc xây dựng danh mục các dân tộc, phục vụ cho cuộc tổng điều tra dân số lần thứ hai. Ủy ban Khoa học Xã hội Việt Nam (tên gọi Viện Khoa học Xã hội Việt Nam lúc đó), trực tiếp là Viện Dân tộc học, đã phối hợp với các ngành hữu quan mở hội nghị khoa học xây dựng danh mục các dân tộc ở miền Bắc nước ta vào tháng 6 và tháng 11/1973 thảo luận các vấn đề cần thiết. Hội nghị đã làm sáng rõ nhiều vấn đề cơ bản về công tác xác định dân tộc ở miền Bắc lúc bấy giờ, trên cơ sở đó xây dựng một danh mục các dân tộc. Bản danh mục thành phần các dân tộc ở Việt Nam lần này xếp theo dân số các dân tộc, đăng trong tạp chí *Dân tộc học*, số 1/1974, tr.57-63; Danh mục *xếp theo ngôn ngữ các dân tộc* đăng trong cuốn *Các dân tộc ít người ở Việt Nam các tỉnh phía Bắc*, Hà Nội, 1978.

So với bản danh mục lần đầu (1959), thì chỉ còn 59 dân tộc (trước là 64). Các dân tộc vẫn được phân thành 3 hệ ngôn ngữ, nhưng ở lần này, *ngữ hệ Môn-Khơ me* được thay bằng *ngữ hệ Nam Á*. Trong ngữ hệ đó ngoài nhóm Môn-Khơ me còn bao gồm cả các nhóm Việt - Mường, Tày - Thái, Mèo - Dao và nhóm ngôn ngữ khác (nhóm Ka Đai).

- Những tên dân tộc được bớt đi là :

Xá, Co Chơ, Pa Dí, Thu Lao, Xạ Phang, Đan Lai-Ly Hà, Mày, Khùa, Rục, Vân Kiều, Lạt, Tơ Lốp, Cơ Don, Tơ La, La Gia, Rơ Ngao, Bih, La Oang, Cao Lan và Sán Chỉ⁴.

- Những tên dân tộc mới xuất hiện là:

Thổ, Chứt, Kháng, Khơ Mú, Ô Đu, Hà Lăng, Co, Thủy, Pà Thén, Tống, La Ha, Cống, Chăm Hroi, Cao Lan-Sán Chỉ, Cơ Ho.

- Những dân tộc vẫn có nhưng thay đổi tên gọi :

U Ní → Hà Nhì, Cò Sung → La Hủ, Khả Pẹ → Si La, Mạng Ư → Mảng, Chi La → Si La, Lừ → Lự, Nhắng → Giáy, Chùng Trà → Bố Y, Quý Châu → Pú Nà, Kha Tu → Ca Tu, Chăm → Hrê, Mán → Dao, Puộc → Xinh Mun.

⁴ Ngoài nhóm Xá đã phân chia thành một số dân tộc khác nhau, còn phần lớn là qui hợp một vài nhóm nhỏ thành một nhóm, một dân tộc, chẳng hạn Cao Lan và Sán Chỉ thành dân tộc Cao Lan - Sán Chỉ hoặc Sán Chay.

Cách xác định này được sử dụng để phân định các dân tộc trong hai cuộc điều tra dân số: ngày 1/ 4/1974 ở miền Bắc và ngày 5/2/1976 ở miền Nam.

So với danh mục dân tộc được sử dụng trong cuộc điều tra dân số miền Bắc năm 1960, thì ngoài những điểm trên, ở danh mục điều tra dân số (miền Bắc) năm 1974 còn có một số trường hợp vài nhóm nhỏ được gộp thành một dân tộc lớn hơn như :

Hoa + Xạ Phang = Hoa;

Mày + Sách + Rục + Arem + Mã Liềng = Chứt;

Khùa + Tri + Ma Cong = Bru;

Tày Poọng + Đan Lai + Ly Hà + Kẹo + Mọn + Cuối + Họ = Thổ.

Như thế, trên các tài liệu công bố kể từ thời điểm này, một số tên ngôn ngữ và dân tộc đã có sự thay đổi. Đặc biệt, xin lưu ý rằng cũng từ đó, tộc danh *Thổ* nay thuộc nhóm Việt - Mường, chứ không còn dùng để chỉ người *Tày*, thuộc nhóm *Tày – Thái*, như trước kia.

Hội nghị 1973 cũng đã nêu ra những vấn đề cần tiếp tục đi sâu nghiên cứu, làm sáng tỏ trên cơ sở khoa học có tính thuyết phục cao hơn. Cũng do vậy, để tiếp tục phối hợp tiến hành nghiên cứu xác định thành phần các dân tộc ở Việt Nam, ngày 22/12/1978, Ủy ban Khoa học xã hội Việt Nam (nay là Viện Khoa học xã hội Việt Nam) cùng Ủy ban Dân tộc Trung ương (nay là Ủy ban Dân tộc) đã có tờ trình lên Thủ tướng Chính phủ về những kết quả nghiên cứu đã đạt được. Và ngày 2/3/1979, được sự ủy nhiệm của Chính phủ, Tổng cục Thống kê đã ra quyết định số 121 - TCTK/PPCĐ chính thức ban hành Danh mục các dân tộc Việt Nam, làm cơ sở cho cuộc Tổng điều tra dân số năm 1979. Thuật ngữ *dân tộc* ở đây được xác định là những cộng đồng được xác định dựa trên ba tiêu chuẩn cơ bản: đặc điểm ngôn ngữ, đặc điểm sinh hoạt - văn hóa và ý thức tự giác dân tộc.

Theo quyết định này, nước ta có 54 dân tộc. So với bản danh mục (lần thứ hai) năm 1974, có những thay đổi như sau:

1. Kết hợp hai, ba "dân tộc" thành một dân tộc:

- Giẻ + Triêng = Giẻ - Triêng
- Tu Dí + Bồ Y = Bồ Y
- Chăm + Chăm Hroi = Chăm
- Xơ Đăng + Hà Lãng = Xơ Đăng
- Xrê + Cơ Ho = Cơ Ho
- Pú Nà + Giáy = Giáy
- Thủy + Pà Thên = Pà Thên
- Tồng + Dao = Dao

2. Xuất hiện những dân tộc có tên gọi mới :

- Ngái (tách từ Hoa thành 2 dân tộc : Hoa và Ngái).
- Chơ Ro
- Brâu
- Rơ Măm.

3. Xác định lại một số tên dân tộc cho chính xác :

- Pa Cô → Tà Ôi
- Ca Tu → Cơ Tu
- Cao Lan + Sán Chỉ → Sán Chay.
- Mèo → Hmông.

Đáng chú ý là Danh mục này tiếp tục được sử dụng cho các cuộc hai cuộc Tổng điều tra dân số 1989, 1999 và trong các tài liệu thống kê ở các cấp từ trung ương đến địa phương suốt từ đó đến nay. Các nhà nghiên cứu ở Viện Khoa học Xã hội Việt Nam vừa hoàn thành một Dự án nhằm góp phần trả lời khách quan hơn cho những câu hỏi còn và đang được đặt ra với một số nhóm dân tộc và ngôn ngữ, một khi cuộc sống thay đổi cũng cần có những điều chỉnh thích hợp. Song cuộc Tổng điều tra dân số sắp tới tiến hành vào 2009 có lẽ vẫn theo Danh mục này. .

3. Tên ngôn ngữ theo cách phân loại ngôn ngữ

Việc phân loại các ngôn ngữ ở Việt Nam trải qua một thời dài nghiên cứu và tranh luận, trao đổi. Thực tế cho thấy có những ngôn ngữ gồm nhiều phương ngữ tộc người nhưng những phương ngữ tộc người ấy lại là một ngôn ngữ riêng rẽ, độc lập của một tộc người/dân tộc nhất định. Và ngược lại cũng rất có thể có những nhóm thổ ngữ chỉ là một bộ phận của một phương ngữ nào đó. Trong loại hình học ngôn ngữ, bằng vào cứ liệu ngôn ngữ học so sánh - lịch sử, cách phân chia các nhóm (groups) ngành hay nhánh (branches), ngữ hệ hay họ ngôn ngữ (families)) cũng có thể theo những quan điểm khác nhau. Vì vậy cách gọi tên đối tượng nghiên cứu cũng có thể khác nhau. Chẳng hạn theo Atsushi Kasuga thì thuật ngữ "Vietnamuong" được M. E. Barker dùng năm 1963 để chỉ nhóm ngôn ngữ bao gồm tiếng Việt và tiếng Mường; thuật ngữ "Viet-Muong" được M. Ferlus đề cập đến năm 1975 và được L. C. Thompson dùng năm 1976; thuật ngữ "Viettic" được G. Diffloth sử dụng năm 1990.

Cho đến nay, ý kiến đã khá thống nhất rằng ở Việt Nam có mặt các dân tộc thuộc cả 5 ngữ hệ (được GS. TS Trần Trí Dõi gọi là *họ ngôn ngữ*), và nếu ta xếp theo số lượng ngôn ngữ giảm dần, đó là : Nam Á, Thái – Ka đai (hay Kam – Thai), Hán – Tang, Nam đảo (hay Mã lai - Đa đảo) và Mông – Dao (hay Mèo - Dao).

Cụ thể ta có thể lập bảng như sau:

I. Ngữ hệ Nam Á (nhánh Môn – Khor me), có 25 ngôn ngữ, trong 5 nhóm:				
	<i>I. a. Nhóm Khor me, có 2 ngôn ngữ :</i>			
1.	1.	1.	Khor me	
2.	2.	2.	Rơ măm	
	<i>I. b. Nhóm Ba na, có 11 ngôn ngữ, chia thành 2 tiểu nhóm:</i>			
		<i>I. b. 1. Tiểu nhóm Ba na Bắc có 6 ngôn ngữ</i>		
3.	3.	1.	Ba na	
4.	4.	2.	Xơ đăng	
5.	5.	3.	Hơ rê	
6.	6.	4.	Gié – Triêng	
7.	7.	5.	Co	
8.	8.	6.	Bơ râu	
		<i>I. b. 2. Tiểu nhóm Ba na Nam có 5 ngôn ngữ</i>		
9.	9.	1.	Kơ ho	
10.	10.	2.	Mơ nông	Bài
11.	11.	3.	Xơ tiêng	Bài
12.	12.	4.	Mạ	Bài
13.	13.	5.	Chơ ro	Bài
	<i>I. c. Nhóm Ka tu có 3 ngôn ngữ:</i>			
14.	14.	1.	Bru – Vân Kiều	Bài 5
15.	15.	2.	Cơ tu	Bài 5, 8
16.	16.	3.	Tà ôi	Bài

	<i>I. d. Nhóm Việt - Mường có 4 ngôn ngữ:</i>			
17.	17.	1.	Việt	
18.	18.	2.	Mường	Bài
19.	19.	3.	Thổ	Bài
20.	20.	4.	Chứt	Bài 6
	<i>I. e. Nhóm Khor mú có 5 ngôn ngữ:</i>			
21.	21.	1.	Khor mú	Bài
22.	22.	2.	Xinh mun	Bài
23.	23.	3.	Kháng	Bài
24.	24.	4.	Mảng	Bài
25.	25.	5.	Ố đu	Bài 4, 8
II. Ngũ hệ Thái – Ka đai, có 12 ngôn ngữ, trong 4 nhóm:				
	<i>II. a. Nhóm Ka đai có 4 ngôn ngữ:</i>			
26.	1.	1.	La chí	
27.	2.	2.	La ha	
28.	3.	3.	Cơ lao	
29.	4.	4.	Pu péo	
	<i>II. b. Nhóm Day Sec chỉ có 1 ngôn ngữ</i>			
30.	5.		Giáy	
	<i>II. c. Nhóm Cao Lan chỉ có 1 ngôn ngữ</i>			
31.	6.		Sán Chay (Cao lan – Sán chỉ)	
	<i>II. d. Nhóm Thái - Tày gồm hai tiểu nhóm:</i>			

		<i>II. d. 1. Tiểu nhóm Thái có 3 ngôn ngữ:</i>		
32.	7.	1.	Thái	
33.	8.	2.	Lào	
34.	9.	3.	Lự	
		<i>II. d. 2. Tiểu nhóm Tày cũng có 3 ngôn ngữ:</i>		
35.	10.	1.	Tày	
36.	11.	2.	Nùng	
37.	12.	3.	Bố Y	
III. Ngũ hệ Hán – Tạng, có 9 ngôn ngữ, trong 2 nhóm:				
	<i>III. a. Nhóm Hán có 3 ngôn ngữ:</i>			
38.	1.	1.	Hoa	
39.	2.	2.	Sán diu	
40.	3.	3.	Ngái	
	<i>III. b. Nhóm Tạng (cùng thuộc tiểu nhóm Lô lô) có 6 ngôn ngữ:</i>			
41.	4.	1.	Hà nhì	
42.	5.	2.	Phù lá	
43.	6.	3.	La hủ	
44.	7.	4.	Lô lô	
45.	8.	5.	Cống	
46.	9.	6.	Si la	
IV. Ngũ hệ Nam Đảo (cùng thuộc nhóm Chăm), có 5 ngôn ngữ :				
47.		1.	Gia rai	

48.		2.	Ê đê	
49.		3.	Chăm	
50.		4.	Ra glai	
51.		5.	Chu ru	
V. Ngũ hệ Mông - Dao, có 3 ngôn ngữ, trong 2 nhóm :				
	<i>V. a. Nhóm Mông có 2 ngôn ngữ:</i>			
52.	1.	1.	Mông	
53.	2.	2.	Pà Thẻn	
	<i>V. b. Nhóm Dao có tiếng Dao</i>			
54.	3.		Dao	

Tuy nhiên, cần lưu ý rằng các nhà nghiên cứu ngôn ngữ có thâm niên như trong lĩnh vực này như GS. TS Trần Trí Dõi (ở Khoa Ngôn ngữ học, Đại học Quốc gia Hà Nội) quan niệm đây chỉ là "con số mang tính quản lý hành chính... Những kết luận đã có đó chỉ phù hợp với điều kiện nghiên cứu lúc bấy giờ, phản ánh khả năng tiếp cận vấn đề được đặt ra vào thời kỳ ấy. Hiện nay chúng ta có cơ sở để nói rằng trong thực tế bức tranh ngôn ngữ dân tộc còn phức tạp hơn nhiều."⁵ GS. TSKH Lý Toàn Thắng và GS. TS Nguyễn Văn Lợi (ở Viện Ngôn ngữ học) cũng nhận xét rằng “Hiện nay, về mặt ngôn ngữ, trong nhiều trường hợp, chưa xác định đủ rõ thứ tiếng này hay khác là ngôn ngữ độc lập hay chỉ là các biến thể địa phương (phương ngữ) của một ngôn ngữ; do vậy, chưa ai đưa ra được con số chính xác, đủ sức thuyết phục về số lượng các ngôn ngữ ở Việt Nam”⁶.

PGS. TS Hoàng Văn Ma đã từng lưu ý đến “sự khác biệt giữa sự phân loại các cộng đồng dân tộc về phương diện ngôn ngữ học và dân tộc học. Nếu phân loại **dưới** góc độ dân tộc học, tiêu chuẩn chính là lịch sử - văn hoá, trong

⁵ Trần Trí Dõi. *Nghiên cứu ngôn ngữ các dân tộc thiểu số ở Việt Nam*. Hà Nội, Nxb Đại học Quốc gia Hà Nội, 1999, tr. 14.

⁶ Lý Toàn Thắng & Nguyễn Văn Lợi. « Về sự phát triển của ngôn ngữ các dân tộc thiểu số ở Việt Nam trong thế kỷ XX. » *Ngôn ngữ*, số 2, 2001, tr. 1.

đó ngôn ngữ là một yếu tố rất quan trọng; nếu phân loại dưới góc độ ngôn ngữ học, tiêu chuẩn ngôn ngữ là duy nhất”⁷. Các ví dụ cụ thể được chuyên gia về ngôn ngữ các dân tộc này (ở Viện Ngôn ngữ học) đưa ra là:

+ Ngay "Danh mục ngôn ngữ trong nhóm Việt – Mường, nhóm được nhiều nhà nghiên cứu trong cũng như ngoài nước chú ý nhất, vì gắn với vấn đề lịch sử tiếng Việt, cho đến nay vẫn chưa được khẳng định về số lượng" (tr. 10). Về dân tộc, có 4: Kinh, Mường, Thổ, Chứt. Nhưng về ngôn ngữ, có nhà nghiên cứu khẳng định có đến 7: Thổ được tách thành Cuối và Pong, Chứt được chia thành 3: Arem, Sách (gồm cả Mày, Rục⁸) và Mã Liềng.

+ Nhóm Tày – Thái (Kam-Thai), xét dưới góc độ dân tộc học, có 8 dân tộc: Tày, Thái, Nùng, Sán Chay (Cao Lan – Sán Chỉ), Giáy, Lào, Lự, Bô Y. Thế nhưng xét dưới góc độ ngôn ngữ học thì phải kể thêm: tiếng Thủy, tiếng Tống (sắp trở thành tử ngữ, vì họ đã dùng tiếng Dao trong giao tiếp hàng ngày). Tiếng Pa Dí và tiếng Thu Lao có thể là những ngôn ngữ độc lập vì nhiều khía cạnh, chúng gần tiếng Thái hơn tiếng Tày. Đặc biệt là với dân tộc Sán Chay⁹, gồm hai nhóm địa phương thì ngôn ngữ của người Cao Lan thuộc nhóm Tày – Thái, trong khi ngôn ngữ của người Sán Chỉ lại nằm trong nhóm Hán. Như vậy, số lượng ngôn ngữ trong nhóm Tày – Thái lên đến 13, trong khi số lượng dân tộc học là 8.

+ Trong khi nhóm dân tộc Hán ở Việt Nam có Hoa và Ngái. Họ đều là những người từ Trung Hoa sang nhưng những người sống ở thành phố và đồng bằng được gọi là Hoa, còn những người sống ở nông thôn miền núi – có thể do cung cách sinh hoạt không như cả thành phố - nên được gọi là Ngái, Xạ Phang... Như vậy, về mặt ngôn ngữ, rất có thể xem Ngái và Hoa là một.

3. Thay cho kết luận: Đòi điều gợi ý cho việc định chủ đề

Từ thực tế trên đây, chúng tôi xin được đưa ra một số ý kiến, có liên quan đến việc định chủ đề:

- Trước hết là xin có mấy nhận xét về Bảng số 3 *Vietnam - Languages* của Thư viện Quốc hội Hoa Kỳ như sau:

⁷ Hoàng Văn Ma . *Ngôn ngữ dân tộc thiểu số Việt Nam: Một số vấn đề về quan hệ cội nguồn và loại hình học*. Hà Nội, Nxb Khoa học xã hội, 2002, tr. 9 trở đi.

⁸ NGUYỄN VĂN LỢI . *Tiếng Rục*. Hà Nội, Nxb Khoa học xã hội, 1993, 172 tr.

⁹ Ta vẫn gặp tên riêng *Cao Lan* chẳng hạn trong những cuốn sách in ở Nxb Khoa học xã hội có tên là *Văn hoá dân gian của dân tộc Cao Lan* của Đặng Đình Thuận (2005) và *Văn hoá Cao Lan* của Lâm Quý (2004)..

Có ngôn ngữ mà Thomas Mann đưa vào Bảng này, chúng tôi không gặp cách gọi tên này khi tiếp cận các tài liệu về Việt Nam. Đó là *Biat language*, cho nên người dịch đã không tìm thấy từ tiếng Việt tương đương.

Thuật ngữ *Banaric* nên được dịch là *nhóm Ba na* (gồm ít ra là 11 ngôn ngữ).

Cũng xin lưu ý rằng, cứ theo bảng trên, thì *Bru* và *Brâu* là hai dân tộc : Bơ râu là một trong 6 ngôn ngữ thuộc *Tiểu nhóm Ba na Bắc*, còn *Bru – Vân Kiều*¹⁰ – cùng với *Cơ tu* và *Tà ôi* là 3 ngôn ngữ thuộc *Nhóm Ka tu*¹¹

- Việc định chủ đề cho từng ngôn ngữ là xem ra không thực sự cần thiết, một khi còn khá nhiều ngôn ngữ ở nước ta chưa được nghiên cứu.

Khảo sát tài nguyên thông tin trong nước về các ngôn ngữ ở Việt Nam cho thấy kết quả nghiên cứu về các ngôn ngữ này được công bố tại các hội thảo, hội nghị khoa học, có thể được gộp thành sách. Một số công trình được in thành sách. Phần lớn tài nguyên thông tin bao gồm bài viết công bố trên các tạp chí, tập san khoa học của các trung tâm nghiên cứu và giảng dạy ngôn ngữ, đặc biệt là các tạp chí "Ngôn ngữ" do Viện Khoa học Xã hội Việt Nam quản lý và "Ngôn ngữ và đời sống" của Hội Ngôn ngữ học Việt Nam.

Do tiếng Việt có vị trí đặc biệt, nên số lượng các công trình nghiên cứu về ngôn ngữ này luôn chiếm ưu thế tuyệt đối (khoảng 94%) khi ta khảo sát khối lượng công trình trong kho tài nguyên thông tin về các ngôn ngữ ở Việt Nam.

Thật vậy, thống kê trong 30 năm tạp chí "Ngôn ngữ" từ khi ra đời (1969-1999), thì chỉ có 81 bài về lĩnh vực ngôn ngữ các dân tộc thiểu số (chiếm 6,2% số bài công bố trên tạp chí này)¹². Theo khảo sát cho đến 2007 của Nguyễn Thị Kim Thoa¹³ trong tổng số 2.511 bài viết được công bố suốt 38 năm trên tạp chí này, đã có 147 bài, chiếm 5,85%, là nghiên cứu về các ngôn ngữ thiểu số ở Việt Nam và Đông Nam Á, còn lại 94,15% (2.364 bài) đề cập đến các bình diện khác nhau của tiếng Việt.

¹⁰ So sánh : HOÀNG VĂN MA, TẠ VĂN THÔNG . *Tiếng Bru - Vân Kiều*, Hà Nội, Nxb Khoa học xã hội, 1998.

¹¹ Trong khi cuốn sách của Lưu Hùng có tên là *Góp phần tìm hiểu văn hoá Cơ tu* (Hà Nội, Nxb Khoa học xã hội, 2006) hay bài viết của Trần Nguyễn Khánh Phong in trên tạp chí *Dân tộc & Thời đại*, số 102, tháng 5-2007, tr. 5-6, có tên là *Vách nhà bằng vỏ cây của đồng bào Tà-ôi và Cơ-tu* cuốn sách của Nguyễn Văn Lợi & Nguyễn Hữu Hoàn có tên là *Tiếng Katu* (Hà Nội, Nxb Khoa học xã hội, 1998).

¹² *Tổng mục lục Tạp chí Ngôn ngữ* (1969-1999), Hà Nội, Nxb Khoa học xã hội, 2000, tr.6.

¹³ Nguyễn Thị Kim Thoa. *Tiểu luận Chuyên đề NCS 2: Tình hình nghiên cứu ngôn ngữ các dân tộc thiểu số ở Việt Nam hiện nay*. Hà Nội, 2008.

Về số lượng các ngôn ngữ được nghiên cứu thì trong *Nghiên cứu ngôn ngữ các dân tộc thiểu số ở Việt Nam từ những năm 90¹⁴*, chúng tôi nhận thấy còn 23 ngôn ngữ chưa được nghiên cứu thì cũng theo khảo sát của Nguyễn Thị Kim Thoa, cho đến 2007 vẫn còn 12 ngôn ngữ chưa được nghiên cứu:

- 5 ngôn ngữ thuộc họ Nam Á là: Rơ Măm, Chơ Ro, Co, Brâu và Ó Đu.
- 3 ngôn ngữ thuộc họ Hán - Tạng là: Cống, Si La và La Hủ.
- 4 ngôn ngữ thuộc họ Thái - Kadai là: Cơ Lao, Bô Y, Lự và Giáy.

Như vậy, việc định chủ đề cho cả mấy chục ngôn ngữ ở Việt Nam xem ra không hẳn đã là cần thiết. Do vậy, chúng tôi cho rằng chỉ nên định chủ đề riêng cho *Tiếng Việt*. Còn về các ngôn ngữ khác có thể theo 5 *ngữ hệ* (họ) hoặc *nhóm* hay *tiểu nhóm* ngôn ngữ (như trong bảng tóm tắt ở trên), không đi sâu vào chi tiết, tránh được những rắc rối còn đang đặt ra với chính các nhà khoa học chuyên ngành.

Như thế, dù tác giả tài liệu trước đây và hiện nay có sử dụng cách gọi tên nào (theo ngôn ngữ, tiếng địa phương, nhóm...; tên tự gọi,...) thì công việc định chủ đề cũng vẫn dễ dàng, một khi ta bổ sung thêm Bảng đối chiếu các tên gọi cũ và mới.

- Như vậy, điều đáng chú ý là qua ba lần công bố thành phần dân tộc, ngay cả tên gọi chính thức thì cũng đã có thể thay đổi. Bên cạnh những tên gọi này, trong các *Danh mục* còn có (những) tên tự gọi và tên gọi khác. Người làm công tác xử lý tài liệu ở các thư viện không những cần tránh sử dụng những tên gọi không còn phù hợp, nhất là khi phải xử lý hồi cố tư liệu, mà còn cần biết những tên gọi tương ứng cho một (nhóm) dân tộc.

Nhắc lại những thay đổi trong quá trình xác định thành phần dân tộc ở Việt Nam trên đây, chúng tôi muốn lưu ý việc xác định tiêu đề đề mục các ngôn ngữ cần xác định cho đúng đối tượng, và cũng nên có phần so sánh với những tên gọi đã dùng trước, để tránh "*ông nói gà, bà nói vịt*".

*

* *

Vấn đề nêu ra trên đây, trước hết xin được xem như những thông tin gợi ý, nên thận trọng khi đọc và cũng như xử lý, phân định tài liệu về các ngôn ngữ và các dân tộc ở Việt Nam.

Thêm nữa, chúng tôi cũng muốn từ hiểu biết chuyên sâu của mình, có một số ý kiến cụ thể được góp phần vào công việc Định chủ đề & Biên soạn Khung Tiêu đề đề mục mà ngành thư viện đang quan tâm.

Tháng 4 năm 2008

¹⁴ Thông tin KHXH chuyên đề, 2002, 240 tr.