

## CÁC NGÔN NGỮ DÂN TỘC Ở VIỆT NAM TỪ GÓC NHÌN KHU VỰC HỌC

Mai Ngọc Chú<sup>(\*)</sup>

1. Khu vực học (area studies) là một khái niệm mới, một ngành học mới xuất hiện gần đây. Ngành học này, như nghĩa đen vốn có của nó, lấy đối tượng nghiên cứu là khu vực.

Xét về mặt không gian, khái niệm khu vực có thể hiểu theo phạm vi rộng hẹp khác nhau. Có người quan niệm khu vực là khái niệm chỉ một không gian rộng lớn, gồm nhiều quốc gia (chẳng hạn khu vực Đông Nam Á, khu vực Tây Á, khu vực Trung Cận Đông, v.v.) do vậy đã phân biệt khu vực học với đất nước học (chẳng hạn Việt Nam học, Trung Quốc học, v.v.). Chúng tôi không đi theo cách hiểu này.

Theo một quan niệm khác thì khu vực có thể được hiểu theo cả nghĩa rộng (như quan niệm trên) lẫn nghĩa hẹp. Rộng thì có thể bao gồm cả châu lục hoặc liên châu lục như châu Á, châu Âu, châu Mĩ, phương Đông, phương Tây, v.v. Hẹp có thể khuôn lại trong một làng, một bản. Chính vì vậy, có người đã nhận diện khu vực theo 4 mức: trường hợp (case), tiểu vùng (subregion), vùng (region) và khu vực (area)[17]. Với cách quan niệm như thế thì đất nước học cũng thuộc phạm trù khu vực học.

Tuy nhiên dù có thể hiểu rộng hẹp như thế nào thì đối tượng xem xét của khu vực học vẫn là một không gian địa lí - xã hội nhất định, tức là không gian trong đó có sự tồn tại của xã hội, của con người.

Trong một khu vực, người ta có thể tiến hành nhiều nghiên cứu khác nhau. Nói cách khác, đó có thể là đối tượng của nhiều ngành khoa học như ngôn ngữ học, sử học, văn hoá học, xã hội học, kinh tế học, v.v.

Với tư cách là một ngành khoa học thực thụ, riêng biệt, *khu vực học xem xét một khu vực nào đó một cách toàn diện và tổng hợp*. Khu vực học là *một khoa học đa ngành và liên ngành* bởi muốn nghiên cứu toàn diện văn hoá của một nhóm người nào đó trong một không gian văn hoá - xã hội nào đó thì không thể không nghiên cứu cả tiếng nói, phong tục tập quán, văn hoá dân gian, lịch sử, rồi cả những vấn đề kinh tế, xã hội, chính trị, v.v. của họ. Như vậy để có thể tìm hiểu sâu về một khu vực thì cần có sự phối hợp của nhiều ngành khoa học, nhiều nhà khoa học thuộc các lĩnh vực khác nhau.

2. Trong lịch sử nghiên cứu các khu vực và đất nước trên thế giới, người ta thường bắt đầu từ những nghiên cứu ngôn ngữ (tiếng bản địa), chính trị (khu vực chính trị) và văn hoá - văn minh (các nền văn minh). Trong bài viết này, từ một cách tiếp cận, chúng tôi bắt đầu từ các ngôn ngữ dân tộc ở Việt Nam.

Từ góc nhìn khu vực - văn hoá học, các ngôn ngữ ở Việt Nam có thể được phân về 6 khu vực văn hoá:

Khu vực Tây Bắc

<sup>(\*)</sup> PGS - TS, Khoa Đông phương học, Đại học Khoa học Xã hội & Nhân Văn, Đại học Quốc gia Hà Nội

Khu vực Việt Bắc

Khu vực châu thổ Bắc Bộ

Khu vực Trung Bộ

Khu vực Tây nguyên - Trường Sơn

Khu vực Nam Bộ

Trong các khu vực trên, khu vực châu thổ Bắc Bộ chủ yếu là địa bàn cư trú của người Việt (Kinh), là phạm vi phổ biến của tiếng Việt, không có ngôn ngữ dân tộc thiểu số, vì vậy sẽ không được đề cập đến ở đây.

### 2.1. Khu vực Tây Bắc

Tây Bắc, xét theo đặc trưng văn hoá - tộc người, là khu vực gồm trọn 5 tỉnh Lai Châu, Điện Biên, Sơn La, Lào Cai, Yên Bái và một phần của tỉnh Hoà Bình. Đây là xứ sở của hoa ban, của xoè hoa. Đặc trưng tiêu biểu của vùng này là *ngôn ngữ - văn hoá Thái, Mèo* mà dấu ấn đậm nét còn lưu giữ ở những thiên tình sử như "Xông chu xôn sao" (Tiễn dặn người yêu) của dân tộc Thái và "Lux txax mô nhangx" (Tiếng hát lâm dâu) của dân tộc Mèo. Tiếng Thái và tiếng Mèo trở thành ngôn ngữ khu vực. Điều này lí giải vì sao trước đây, có lúc, khu Tây Bắc còn được gọi là khu tự trị Thái - Mèo. Ngoài hai ngôn ngữ này, Tây Bắc còn có hàng loạt ngôn ngữ khác như Dao, Khơ mú, Giáy, Hànhì, Xinhmul, Lào, Phulá, Lahù, Kháng, Lự, Mảng, Cống, Sila, Mường, Laha, Tây - Nùng, Bôy, Láchí, Sánchay (Caolan - Sánchì). Tổng số ngôn ngữ ở khu vực này lên tới con số 22. *So với các khu vực khác trên toàn quốc, Tây Bắc là khu vực đậm đặc ngôn ngữ nhất.* Ở Lai Châu - Điện Biên có tới 15 ngôn ngữ, Lào Cai có 12 ngôn ngữ. Đây là những tỉnh có nhiều ngôn ngữ nhất Việt Nam. Dưới đây là sơ phân bố ngôn ngữ ở các tỉnh:

**Lai Châu - Điện Biên:** Thái, Mèo, Dao, Khơ mú, Giáy, Hànhì, Xinhmul, Lào, Phulá, Lahù, Kháng, Lự, Mảng, Cống, Sila. (15)

**Lào Cai:** Thái, Tây - Nùng, Mèo, Dao, Sánchay (Caolan - Sánchì), Khơ mú, Giáy, Hànhì, Láchí, Phulá, Bôy, Laha. (12)

**Sơn La:** Thái, Mèo, Mường, Dao, Khơ mú, Xinhmul, Lào, Kháng, Laha. (9)

**Yên Bái:** Thái, Mèo, Tây, Mường, Dao, Khơ mú, Phulá. (7)

**Hoà Bình:** Thái, Mèo, Mường, Khơ mú. (4)

Xét về mặt nguồn gốc, các ngôn ngữ thuộc khu vực Tây Bắc chủ yếu thuộc về các ngữ hệ Nam Á, Thái - Kadai và Hán Tạng.

Các ngôn ngữ thuộc về họ Hán Tạng đều nằm ở nhóm Tạng - Miến, đó là Hànhì, Phulá, Lahù, Cống, Sila.

Các ngôn ngữ thuộc ngữ hệ Thái là Thái, Tây - Nùng, Sánchay, Giáy, Lự, Lào, Bôy; và nhóm Kadai bao gồm Laha<sup>(1)</sup>, Láchí.

Các ngôn ngữ thuộc họ Nam Á bao gồm tiếng Mường nhóm Việt - Mường; tiếng Mèo, tiếng Dao thuộc nhóm Mèo - Dao<sup>(2)</sup>; các ngôn ngữ Khơ mú, Kháng, Xinhmul, Mảng thuộc nhóm Môn - Khơme.

Các ngôn ngữ thuộc nhóm Tạng - Miến, ngữ hệ Hán - Tạng thường có một số từ mới được vay mượn từ tiếng Thái vốn là một ngôn ngữ mang tính khu vực. Ngoài ra có thể có một số từ thuộc gốc Miến (như trong tiếng Cống) hoặc thuộc gốc Tạng (như trong tiếng Lahù). Một số từ được mượn từ tiếng quan hoá.

<sup>(1)</sup> Trong cuốn "Những nhóm dân tộc thuộc ngữ hệ Nam Á ở Tây Bắc Việt Nam", NXB KHXH, 1972, các tác giả Đặng Nghiêm Vạn, Nguyễn Trúc Bình, Nguyễn Văn Huy, Thanh Thiên xếp tiếng Laha vào ngữ hệ Nam Á. Tác giả Giáo sư Paul K. Benedict (Mỹ) từ 1941 đã xếp tiếng này cùng với các ngôn ngữ Láchí, Colao, Púpéo vào nhóm Kadai. Một số tác giả khác xếp 4 ngôn ngữ vừa nêu vào nhóm "hỗn hợp".

<sup>(2)</sup> Cũng có tác giả tách Mèo - Dao thành một ngữ hệ riêng.

Trong số các ngôn ngữ thuộc nhóm Môn - Khơme, tiếng Khơ mú có sự khác biệt đáng kể so với ba ngôn ngữ còn lại (Kháng, Xinhmul, Mảng) là những ngôn ngữ có lịch sử lâu đời nhất ở Tây Bắc như A.G. Haudricourt, H. Maspéro và một số nhà khoa học khác đã khẳng định. Ba dân tộc này định cư ở Tây Bắc sớm hơn Khơ mú và sớm hơn cả cộng đồng Thái. Theo các tác giả của "Những nhóm dân tộc thuộc ngữ hệ Nam Á ở Tây Bắc Việt Nam", trong tiếng Khơ mú, ngoài những từ cơ bản cùng gốc với các ngôn ngữ Kháng, Xinhmul, Mảng, còn có cả những từ cơ bản gần với nhóm ngôn ngữ Paluang, Va và Lamet ở Lào. (Điều này càng góp phần khẳng định ý kiến của các nhà khoa học về tổ tiên xa xưa của người Khơ mú ở bên Lào).

Sự khác biệt giữa tiếng Khơ mú với các ngôn ngữ trên còn có thể tìm thấy ở hệ thống phụ âm đầu và ở thanh điệu. Trong tiếng Khơ mú, tổ hợp phụ âm đầu còn rất nhiều, trong khi ở các ngôn ngữ kia các tổ hợp như thế có xu hướng giảm đi rõ rệt. Ví dụ: prok (con sóc), ptê (đất), Kămbra (vợ), hla (lá), hdroi (gió), pdrcăn (gan, mật), v.v. Về thanh điệu, nếu như ở các ngôn ngữ Kháng, Xinhmul, Mảng sự thể hiện của các thanh tương đối rõ ràng thì ở tiếng Khơ mú, chúng đang ở trong giai đoạn "manh nha" gần giống như tình hình ở các ngôn ngữ Paluang - Va và Môn [1, tr.41].

Các ngôn ngữ thuộc nhóm Môn - Khơme đều chịu ảnh hưởng lớn từ tiếng Thái, đặc biệt là ngôn ngữ Kháng. Trong vốn từ Kháng, số từ mượn từ tiếng Thái càng ngày càng tăng. Thậm chí ở một số vùng, toàn bộ số đếm đã được người Kháng mượn từ tiếng Thái, mặc dù trong tiếng Kháng không phải không có. Tiếng Kháng cũng có khá nhiều từ giống tiếng Việt và tiếng Mường. Điều này khẳng định mối

quan hệ chặt chẽ giữa Kháng và nhánh Việt - Mường vốn cùng ngữ hệ Nam Á.

Cũng như tiếng Kháng, trong tiếng Xinhmul có khá nhiều từ vay mượn từ tiếng Thái và có các từ chung gốc Môn - Khơme. Ảnh hưởng của tiếng Thái mạnh đến mức tuyệt đại đa số người Xinhmul đều nói được tiếng Thái. So với tiếng Kháng và Xinhmul thì tiếng Mảng có phần biệt lập hơn do dân tộc này vừa quá nhỏ bé về số lượng vừa định cư ở những nơi rất hẻo lánh.

Điểm chung của các ngôn ngữ nhóm Môn - Khơme Tây Bắc là đang diễn ra quá trình rụng dần các tổ hợp phụ âm đầu, làm đơn giản hoá âm tiết; phương thức cấu tạo từ bằng phụ tố, nhất là tiền tố và trung tố, hầu như không còn nữa.

Mèo - Dao là hai ngôn ngữ anh em ruột thịt. Quan hệ này được xác định dựa trên sự thống nhất một số lượng lớn từ vựng cơ bản và những đặc trưng cơ bản về ngữ pháp của hai ngôn ngữ, đồng thời điều này cũng phù hợp với sự thống nhất về các đặc điểm dân tộc Mèo và Dao, phù hợp với những tài liệu thư tịch cổ Trung Quốc về nước Tam Miêu trong lịch sử cổ đại Trung Quốc [8, tr.141]. Xét về mặt ngữ âm, đặc điểm nổi bật của tiếng Mèo là số lượng phụ âm đầu và thanh điệu rất lớn, trái lại, số lượng nguyên âm làm âm chính và âm cuối thì lại rất ít. Ở các phương ngữ khác nhau, số lượng phụ âm đầu, nói chung, đều vượt quá con số 50, thậm chí đến 70; số lượng thanh điệu 7 - 8; trong khi nguyên âm chỉ ở mức 5 - 6. Về từ vựng, có một điểm đáng chú ý là trong vốn từ của tiếng Mèo, có những từ có quan hệ với các ngôn ngữ Tạng - Miến, có những từ gần gũi với các ngôn ngữ Thái - Kadai, có rất nhiều từ có quan hệ với các ngôn ngữ Môn - Khơ me, và gần đây là những từ mượn từ tiếng Việt.

Bức tranh từ vựng như thế là lí do chính dẫn đến việc ngôn ngữ này (cùng với tiếng Dao) khi thì được xếp vào họ Hán - Tạng, khi thì được đưa vào họ Nam Á, khi thì lại được coi là một ngôn ngữ riêng biệt trong họ Nam phương, v.v.

Như trên đã nói, các phương ngữ Mèo rất đa dạng, phong phú. Tuy nhiên tính thống nhất của chúng cũng rất cao. Chính vì thế hoàn toàn có cơ sở để xác định một vùng phương ngữ cơ sở, là đại diện để xây dựng hệ thống chữ viết Mèo. Đến nay, tiếng cơ sở được xác định là tiếng Mèo Hoa ở Sapa và hệ thống chữ viết Mèo cũng đã được xây dựng và đưa vào sử dụng.

Thái là dân tộc thuộc loại lớn nhất Việt Nam và cư trú chủ yếu ở Tây Bắc. Từ mười thế kỉ trở lại đây, với vai trò chủ thể trong lịch sử phát triển của vùng, văn hoá Thái, với những yếu tố tiếp biến từ văn hoá Đông Nam Á, đã nổi lên như một sắc thái đại diện cho văn hoá Tây Bắc [14, tr.159]. Trong vai trò này, không thể không kể đến tiếng Thái - một ngôn ngữ đã được khá nhiều dân tộc sử dụng trong giao tiếp hoặc vay mượn từ. Một ví dụ có thể nêu lên là 100 % số người thuộc nhóm La ha và Kháng biết thông thạo tiếng Thái [2, tr.37].

Tiếng Thái, giống như tiếng Việt, là ngôn ngữ đơn lập điển hình. Đó là ngôn ngữ âm tiết tính, với 5 thanh điệu như tiếng Thái ở Thái Lan. Sự khác nhau giữa tiếng Thái Tây Bắc Việt Nam với tiếng Thái Thái Lan chủ yếu là ở vốn từ vay mượn: vốn từ vay mượn của tiếng Thái Tây Bắc chủ yếu là từ tiếng Việt, còn vốn từ vay mượn của tiếng Thái ở Thái Lan thì chủ yếu từ Pali - Sanskrit, ngôn ngữ gắn liền với văn hoá Phật giáo.

Từ lâu dân tộc Thái đã có chữ viết riêng của mình. Tuy nhiên đây là một hệ

thống chữ viết cổ, rất khó học đối với thế hệ trẻ. Để gánh vác nhiệm vụ là ngôn ngữ khu vực và theo nguyện vọng của đông đảo đồng bào Thái, hệ thống chữ viết Thái đã được cải tiến. Hệ thống này dựa trên cơ sở của tiếng Mường La. Việc cải tiến chữ Thái rõ ràng rất có ý nghĩa đối với sự nghiệp phát triển văn hoá - giáo dục không chỉ cho dân tộc Thái mà còn cho toàn khu vực.

## 2.2. Khu vực Việt Bắc

Việt Bắc, xét theo cách hiểu truyền thống, bao gồm 6 tỉnh: Cao Bằng, Bắc Cạn, Thái Nguyên, Lạng Sơn, Tuyên Quang và Hà Giang. Tuy nhiên xét theo khu vực văn hoá - tộc người thì nó còn gồm cả phần đồi núi của các tỉnh Vĩnh Phúc, Phú Thọ, Bắc Giang và Quảng Ninh.

Cư dân chủ yếu của Việt Bắc là các dân tộc Tày, Nùng. Ngoài ra còn có các dân tộc khác như Mèo, Dao, Lô Lô, Sánchay, Sánđiù, Ngái, Púpéo, Bóy, Cólao, Pàthên, Phùlá, Láchí, Giáy, Hoa, Mường.

Ở Việt Bắc, mật độ các ngôn ngữ cũng khá dày. Tổng số ngôn ngữ của khu vực này là 16, đứng sau khu Tây Bắc. Hà Giang là tỉnh có nhiều ngôn ngữ nhất: 12. Tình hình phân bố ngôn ngữ ở các tỉnh là như sau.

**Hà Giang:** Tày - Nùng, Hoa, Mèo, Sánchay (Caolan - Sánchì), Giáy, Láchí, Phùlá, Pàthên, Lô Lô, Cólao, Bóy, Púpéo. (12)

**Cao Bằng:** Tày - Nùng, Mèo, Dao, Sánchay, Lô Lô, Ngái. (6)

**Bắc Cạn:** Tày - Nùng, Mèo, Dao, Sánchay, Sánđiù. (5)

**Lạng Sơn:** Tày - Nùng, Dao, Sánchay, Lô Lô, Ngái. (5)

**Tuyên Quang:** Tày - Nùng, Hoa, Dao, Sánchay, Pàthên. (5)

**Thái Nguyên:** Tày - Nùng, Mèo, Dao, Sánchay, Sánđiù. (5)

**Bắc Giang:** Tày- Nùng, Hoa, Sánchay, Sánđiù, Ngái. (5)

**Quảng Ninh:** Tày, Hoa, Sánđiù, Ngái. (4)

**Phú Thọ:** Mường, Sánđiù. (2)

**Vĩnh Phúc:** Sánđiù. (1)

Xét về mặt ngữ hệ, các ngôn ngữ khu vực Việt Bắc thuộc về 3 họ: Thái - Kđai, Hán - Tạng và Nam Á

*Ngữ hệ Thái - Kđai:*

Nhóm Tày - Thái: Tày-Nùng, Sánchay, Giáy, Bóy

Nhóm Kđai: Láchí, Cólao, Pupéo.

*Ngữ hệ Hán - Tạng:*

Nhóm Hán: Hoa, Sánđiù, Ngái

Nhóm Tạng - Miến: Phùlá, Lôlô.

*Ngữ hệ Nam Á:*

Nhóm Việt - Mường: Mường

Nhóm Mèo - Dao: Mèo, Dao, Pàthên.

Nếu như ở Tây Bắc, tiếng Thái là ngôn ngữ khu vực thì ở Việt Bắc, ngôn ngữ Tày - Nùng giữ vị trí ấy. Ở khu vực này, nhiều người thuộc các dân tộc khác nhau đã sử dụng tiếng Tày-Nùng trong giao tiếp xã hội.

Người Tày và người Nùng xưa kia là những tộc người có chung một nguồn gốc lịch sử: cùng thuộc khối Bách Việt. Ngày nay tuy là hai dân tộc nhưng ngôn ngữ của họ, về cơ bản, vẫn giống nhau. Đây chính là lí do để các nhà ngôn ngữ coi Tày - Nùng là một ngôn ngữ thống nhất. Sự thống nhất được biểu hiện ở tất cả các mặt: ngữ pháp, ngữ âm và từ vựng. Tất nhiên, do địa bàn cư trú của cư dân khác nhau, tiếng Tày và tiếng Nùng vẫn có những đặc điểm riêng, có những khác biệt nhất định về ngữ âm và từ vựng. Riêng đối với tiếng địa phương Nùng thì sự khác biệt còn có phần cao hơn giữa các ngành Nùng: Nùng Cháo,

Nùng An, Nùng Inh, Nùng Dín, Nùng Lôi, Nùng Giang, Nùng Phần Sinh, v.v. Tuy vậy tiếng Tày ở địa phương này có thể gần với tiếng Nùng ở cùng địa phương đó hơn so với tiếng Tày ở một địa phương khác; tiếng Nùng của hai ngành khác nhau cùng ở một địa phương có thể gần nhau hơn so với tiếng Nùng của cùng ngành ở khác địa phương [3, tr.72]. Điều này có nghĩa là tiêu chí cùng địa phương quan trọng hơn, và Tày - Nùng hoàn toàn có thể coi là một ngôn ngữ.

Trong lịch sử, ở vào thời kì cận đại, ngôn ngữ Tày - Nùng đã có hình thức chữ viết là chữ Nôm. Ngày nay, do yêu cầu phát triển và do vai trò là ngôn ngữ khu vực, chữ viết Latinh Tày - Nùng đã được xây dựng và đưa vào sử dụng. Chữ viết Latinh Tày - Nùng, khi mới xây dựng, vì chiếu cố đến tình hình đoàn kết giữa các vùng Tày - Nùng là chính, đã không dựa vào một tiếng địa phương nhất định làm cơ sở mà dựa vào những âm phổ biến nhất trong toàn khu vực. Do vậy, trong quá trình sử dụng đã nảy sinh những bất cập (chẳng hạn, việc biên soạn sách vở), ảnh hưởng đến quá trình hình thành ngôn ngữ Tày - Nùng thống nhất [6, tr.3]. Từ tình hình trên đây, hiện nay, vấn đề cải tiến chữ viết Tày - Nùng đã được đặt ra và tiếng địa phương được chọn làm cơ sở chính là tiếng của một vùng lấy Đông Khê (Cao Bằng) - Thất Khê (Lạng Sơn) làm trục chính.

### 2.3. Khu vực Trung Bộ

Khu vực Trung Bộ là vùng đất thuộc các tỉnh Quảng Bình, Quảng Trị, Thừa Thiên - Huế, Quảng Nam, Đà Nẵng, Quảng Ngãi, Bình Định, Phú Yên, Khánh Hoà, Ninh Thuận, Bình Thuận. Ngoài ra, xét về mặt văn hoá - tộc người có thể kể cả vùng núi của các tỉnh Nghệ An, Hà Tĩnh.

Trung Bộ là một dải đất vừa dài vừa hẹp. Đây cũng là địa bàn cư trú của nhiều dân tộc: Pacô - Tàôi, Cốt, Bru - Vânkiêu, Chứt, Co, Giê - Triêng, Xơđăng, Hơrê, Bana, Raglai, Chăm, Êđê, Churu, Cơho, Ođú, Khơmú, Thổ, Mèo, Thái.

Dưới đây là tình hình phân bố ngôn ngữ ở các tỉnh:

**Nghệ An:** Ođú, Khơmú, Thổ, Mèo, Thái. (5)

**Quảng Bình:** Chứt, Bru - Vân kiêu. (2)

**Thừa Thiên - Huế:** Pacô - Tàôi, Cốt, Bru - Vân kiêu. (3)

**Quảng Trị:** Pacô-Tàôi, Bru-Vân kiêu. (2)

**Quảng Nam:** Co, Giê - Triêng, Cốt, Xơđăng. (4)

**Quảng Ngãi:** Co, Hơrê, Xơđăng. (3)

**Bình Định:** Hơrê, Bana. (2)

**Phú Yên:** Raglai, Chăm, Bana, Êđê. (4)

**Khánh Hoà:** Raglai, Bana, Êđê. (3)

**Ninh Thuận:** Churu, Raglai, Cơho, Chăm. (4)

**Bình Thuận:** Raglai, Chăm. (2)

Xét về mặt ngữ hệ, các ngôn ngữ thuộc khu vực Trung Bộ thuộc về 3 họ ngôn ngữ dưới đây.

*Ngữ hệ Nam Á:*

Nhóm Việt - Mường: Thổ, Chứt

Nhóm Mèo - Dao: Mèo

Nhóm Môn - Khơme: Pacô-Tàôi, Cốt, Bru-Vânkiêu, Co, Giê - Triêng, Xơđăng, Hơrê, Bana, Cơho, Ođú, Khơmú

*Ngữ hệ Nam đảo:* Raglai, Chăm, Êđê, Churu

*Ngữ hệ Thái:* Thái.

Khác với khu Tây Bắc và khu Việt Bắc, ở khu vực Trung Bộ không có một ngôn ngữ chung cho toàn khu như tiếng Thái,

tiếng Mèo ở Tây Bắc, Tiếng Tày - Nùng ở Việt Bắc. Tuy nhiên các ngôn ngữ phổ biến của khu vực này chủ yếu thuộc về nhóm Môn - Khơme (ngữ hệ Nam Á), rải suốt từ Quảng Bình đến Bình Định; và nhóm Mã Lai - Đa đảo (ngữ hệ Nam Đảo) phân bố từ Phú Yên đến Bình Thuận.

Về mặt lịch sử, khu vực này gắn liền với nhà nước Chăm pa, do vậy có nhiều dấu ấn đặc trưng của văn hoá - ngôn ngữ Chăm. Tiếng Chăm có một vị trí đặc biệt trong nghiên cứu ngôn ngữ học, dân tộc học, sử học, tiếp xúc văn hoá Việt Nam. Ở Việt Nam, tiếng Chăm có quan hệ gần gũi với các ngôn ngữ thuộc nhóm Nam đảo lục địa như Êđê, Raglai, Churu, Giarai, và cũng có quan hệ với các ngôn ngữ như Malay, Tagalog, v.v. thuộc nhóm Nam đảo hải đảo. Cộng đồng người Chăm Trung Bộ sống đan xen với cộng đồng người Việt và tiếng Chăm chịu ảnh hưởng khá mạnh của tiếng Việt. Có thể nói rằng, so với các ngôn ngữ cũng thuộc nhóm Nam đảo lục địa thì tiếng Chăm chịu ảnh hưởng của tiếng Việt mạnh hơn. Ngoài việc mượn từ của tiếng Việt với một khối lượng lớn, ảnh hưởng của tiếng Việt đến tiếng Chăm còn ở cả cấu trúc âm tiết. Khác với các ngôn ngữ Nam đảo khác, *tiếng Chăm đang trong quá trình hình thành thanh điệu*. Và chính vì quá trình hình thành thanh điệu đang diễn ra nên có nhiều quan điểm về vấn đề này: người cho rằng ở đó mới chỉ có sự đối lập giữa âm vực cao và âm vực thấp (E. Lee, Đoàn Văn Phúc), người cho rằng đã có thanh điệu: hoặc là 2 thanh (Bùi Khánh Thế) hoặc là 4 thanh (G. Moussay, Hoàng Thị Châu), v.v.

Trong tiếng Chăm, trọng âm có giá trị khu biệt âm tiết mở (âm tiết phụ) / âm tiết tỏ (âm tiết chính). Âm tiết mở đang có xu hướng mất dần vì hiện tượng đơn tiết hoá

diễn ra khá mạnh. Hiện tượng này làm cho nhiều từ ngữ âm gồm 1 âm tiết phụ và 1 âm tiết chính chuyển thành từ đơn tiết, ví dụ: kinong > nong (giận), talej > tlej (đây), v.v. Ở tiếng Chăm phương ngữ miền Trung (Ninh Thuận, Bình Thuận), do xu hướng đơn tiết hoá mạnh, từ đơn tiết phổ biến hơn. Điều này khác với tiếng Chăm phương ngữ Nam Bộ vốn có nhiều từ đa tiết. Từ ba âm tiết trở lên ngày càng hiếm trong tiếng Chăm hiện đại [11].

Dân tộc Chăm sớm có chữ viết. Đến nay có 3 hệ thống chữ viết đang song song tồn tại: 1) Akhãr Thrah, 2) Akhãr Jawi Chăm, 3) Chữ Chăm Latinh. Hiện nay hệ thống chữ Chăm Latinh được phổ biến rộng rãi. Nhiều sách giáo khoa, từ điển được biên soạn theo hệ thống này.

#### 2.4. Khu vực Tây Nguyên - Trường Sơn

Khu vực Tây Nguyên - Trường Sơn nằm ở phía Tây Trung Bộ, bao gồm các tỉnh Gia Lai, Kon Tum, Đắc Lắc, Đắc Nông, Lâm Đồng. Đây là địa bàn cư trú của hàng chục dân tộc. Trừ một số dân tộc mới di cư đến, các dân tộc tồn tại lâu đời ở mảnh đất này đều thuộc về hai nhóm ngôn ngữ chủ yếu: Nhóm Môn - Khơme (ngữ hệ Nam Á) và nhóm Mã Lai - đa đảo (ngữ hệ Nam đảo).

Các ngôn ngữ ở khu vực này là: Cơho, Giarai, Raglai, Mơnông, Mạ, Churu, Êđê, Rơmăm, Bana, Xơđăng, Giê - Triêng, Brâu. Tình hình phân bố ở các tỉnh là:

**Lâm Đồng:** Cōho, Raglai, Mơnông, Mạ, Churu. (5)

**Đắc Lắc - Đắc Nông:** Giarai, Êđê, Mơnông, Mạ. (4)

**Gia Lai:** Giarai, Bana, Rơmăm, Xơđăng. (4)

**Kon Tum:** Giarai, Bana, Xơđăng, Giê - Triêng, Brâu. (5)

**Nhóm Môn - Khơme:** Bana, Xơđăng, Cōho, Mơnông, Giê - Triêng, Mạ, Rơmăm, Brâu. (8)

**Nhóm Mã Lai - Đa đảo:** Giarai, Êđê, Raglai, Churu. (4)

Ngôn ngữ - văn hoá của các dân tộc Tây Nguyên gắn liền với những Trường ca (mà tiếng Êđê gọi là *Khan*) bất hủ như *Trường ca Đăm San* (Khan Đăm San), *Trường ca Xinh Nhã* (Khan Xinh Chơ Nga), v.v.

Vai trò của các ngôn ngữ Môn - Khơme ở Tây Nguyên rất lớn. Ở Lâm Đồng, các dân tộc đều biết tiếng Cōho và tiếng Mạ; ở Đắc Lắc, tiếng Mơnông được sử dụng rộng rãi; ở Gia Lai và Công Tum, các ngôn ngữ Xơđăng, Bana trở thành tiếng phổ thông. Các ngôn ngữ Bana, Mơnông, Cōho, Mạ đã được người Pháp trước đây và người Mỹ sau này xây dựng chữ viết trên cơ sở chữ cái Latinh. Những chữ viết này vẫn được sử dụng cho đến ngày nay.

Ở các ngôn ngữ Nam đảo Tây Nguyên người ta có thể tìm thấy một số lượng khá lớn những từ cơ bản, gốc Nam đảo chung với các ngôn ngữ Mã lai - đa đảo ngày nay như tiếng Malay, tiếng Tagalog. Điều này rất có ý nghĩa đối với các nhà ngôn ngữ học.

Cũng như tiếng Chăm, các ngôn ngữ Mã Lai - Đa đảo Tây Nguyên đang phát triển theo hướng đơn tiết hoá. Điều này khác với xu hướng phát triển phụ tố, làm gia tăng các từ đa tiết trong các ngôn ngữ Mã Lai - Đa đảo ở khu vực hải đảo như tiếng Malay, tiếng Tagalog. Trong tiếng Malay, số từ đa tiết chiếm hơn 90%, trong khi đó tỉ lệ cao nhất ở các ngôn ngữ Mã Lai - Đa đảo Việt Nam cũng chỉ dưới 50%.

Âm tiết trong các ngôn ngữ Nam đảo Tây Nguyên hiện nay có thể xem là một đơn vị nhất thể, có ý nghĩa từ vựng và ý

ngôn ngữ pháp, tức là một đơn vị hình thái học. Xét về mặt đồng đại, đó là sự phản ánh quá trình biến đổi hình thái học từ một ngôn ngữ đa tiết có phụ tố đến một ngôn ngữ đơn tiết không phụ tố. Đối với các ngôn ngữ này, sự biến đổi phần đầu của âm tiết thường gây ra những cuộc tranh luận chưa có hồi kết. Đây là phần liên quan trực tiếp đến quá trình đơn tiết hoá, và đồng thời cũng là hệ quả trực tiếp của quá trình này. Liên quan đến vấn đề này, người này cho rằng có tổ hợp 2 thậm chí 3 phụ âm đi liền nhau (kiểu  $C_1C_2$  hoặc  $C_1C_2C_3$ ), người khác lại cảm thấy giữa 2 hoặc 3 phụ âm ấy vẫn còn một nguyên âm chuyển sắc, nghĩa là vẫn tồn tại một âm tiết mà chữ chưa phải đã mất hẳn. Đây có lẽ là điểm sẽ còn gây bất đồng giữa các ý kiến. Vì điều này cũng hoàn toàn dễ hiểu vì quá trình đơn tiết hoá chưa chấm dứt mà đang diễn ra hàng ngày. Rõ ràng chỉ có thể giải thích rằng các ngôn ngữ Nam đảo Tây Nguyên đang phát triển rất mạnh. Và phát triển theo con đường mà tiếng Việt đã trải qua.

### 2.5. Khu vực Nam Bộ

Khu vực Nam Bộ gồm các tỉnh Đồng Nai, Bình Dương, Bình Phước, Tây Ninh, Bà Rịa - Vũng Tàu, Long An, Tiền Giang, Bến Tre, Vĩnh Long, Trà Vinh, Đồng Tháp, Cần Thơ, Hậu Giang, Sóc Trăng, An Giang, Kiên Giang và Minh Hải. Ở khu vực này, ngoài tiếng Việt, ngôn ngữ có sự phổ biến hơn cả là tiếng Khơme. Ngoài ra còn có một số ngôn ngữ khác: Hoa, Chăm, Xtiêng Mạ, Chơro, Mơnông. Các dân tộc Xtiêng Mạ, Mơnông, Chơro cư trú ở các vùng đồi từ phần cuối của dãy Trường Sơn đổ về vùng ven đồng bằng Đông Nam Bộ.

Tiếng Khơme là ngôn ngữ đứng hàng thứ 6 về số lượng người nói, sau các ngôn ngữ Việt, Tày - Nùng, Thái, Mường, Hoa.

Tiếng Khơme phổ biến ở 10 tỉnh: Tây Ninh, Sóc Trăng, Trà Vinh, Kiên Giang, An Giang, Vĩnh Long, Cần Thơ, Hậu Giang, Bạc Liêu và Cà Mau.

Tiếng Khơme có chữ viết từ khoảng thế kỷ thứ VI. Cũng như chữ Thái và chữ Lào, chữ Khơme bắt nguồn từ hệ thống Pali - Sanskrit của Ấn Độ. Tuy nhiên trước đây, ở Nam Bộ, tiếng Khơme chủ yếu chỉ được sử dụng trong giao tiếp hàng ngày. Nói cách khác, chức năng giáo dục của chữ viết Khơme chưa được đề cao. Người Mỹ không mở trường học tiếng Khơme cho con em dân tộc Khơme. Ngày nay địa vị của tiếng Khơme được đề cao cả ở phạm vi giao tiếp xã hội lẫn chức năng giáo dục trong nhà trường. Đặc biệt, tiếng Khơme có vai trò lớn trong chức năng tôn giáo: hàng trăm bộ Kinh Phật bằng tiếng Pali đã được dịch ra tiếng Khơme; ngoài ra còn có hàng loạt những tập kinh sách lá buông được sao chép, tồn trữ ở hàng trăm ngôi chùa Khơme Nam Bộ [13, tr.14]. Chương trình giáo dục trong nhà chùa Khơme vẫn được duy trì và phát triển cho dù đã có hàng trăm trường học của nhà nước giảng dạy bằng tiếng Khơme cho con em đồng bào Khơme. Vai trò của tiếng Khơme còn được đặc biệt chú ý trên các phương tiện thông tin đại chúng như đài phát thanh, truyền hình, báo chí, biển quảng cáo, v.v. Tuy nhiên trong lĩnh vực văn học nghệ thuật, những sáng tác bằng tiếng Khơme còn hạn chế.

Tiếng Khơme là ngôn ngữ đơn tiết không thanh điệu. Đặc điểm đặc trưng nhất của tiếng Khơme xét về mặt ngữ âm là hệ thống nguyên âm vô cùng đa dạng và phức tạp. Thêm nữa, so với hệ thống nguyên âm tiếng Khơme ở Campuchia thì hệ thống nguyên âm tiếng Khơme Nam Bộ còn phức tạp hơn nhiều. Vấn đề này, do vậy, vẫn đang được đặt ra cho các nhà ngôn ngữ học.



### 3. Một vài nhận xét tổng quát

Từ việc xem xét tình hình phân bố và sử dụng ngôn ngữ các dân tộc thiểu số ở Việt Nam, chúng ta có thể nêu lên một số nhận xét như sau.

1. Ở tất cả các khu vực miền núi Việt Nam không có khu vực nào chỉ sử dụng một ngôn ngữ duy nhất. Nói cách khác, trong một khu vực cư trú thường có nhiều dân tộc sống xen kẽ với nhau theo thế cài răng lược. Tình hình này tất yếu dẫn đến *sự tiếp xúc ngôn ngữ khu vực mà kết quả cụ thể của nó là hiện tượng vay mượn từ của nhau và nói tiếng nói của nhau*. Thậm chí có những dân tộc, trong một số vùng, thông thạo hai thứ tiếng như nhau đến mức không thể phân biệt được đâu là tiếng mẹ đẻ đâu là tiếng phổ thông của khu vực [2, tr.40].

2. Không chỉ ngữ hệ mà các nhóm hay các ngành ngôn ngữ trong mỗi ngữ hệ cũng không có một địa bàn cư trú duy nhất. Mỗi nhóm ngôn ngữ thường phân bố ở nhiều khu vực khác nhau. Nhóm Môn - Khơme chẳng hạn, phân bố ở cả Tây Bắc, Tây Nguyên, Trung Bộ và Nam Bộ.

3. Nói chung, ở mỗi khu vực đã được xem xét ở trên, thường có một hoặc một vài ngôn ngữ trở thành ngôn ngữ khu vực,

mang tính phổ thông cho toàn vùng, là công cụ giao dịch giữa các dân tộc trong vùng, ví dụ, tiếng Thái, tiếng Mèo ở Tây Bắc, tiếng Tày - Nùng ở Việt Bắc, tiếng Khơme ở Nam Bộ. Nếu không có một ngôn ngữ mang tính toàn khu vực thì cũng có thể có những ngôn ngữ phổ biến ở khu vực hẹp hơn như trong một tỉnh, một vùng liên huyện, v.v. Ngôn ngữ phổ thông của khu vực thường là ngôn ngữ của cộng đồng đông người nhất hoặc có trình độ văn minh cao nhất.

4. Sự cộng cư và tiếp xúc ngôn ngữ, sự phát triển lán át của ngôn ngữ khu vực và một số nguyên nhân khác nữa *có thể làm tiêu biến một số ngôn ngữ của những cộng đồng quá nhỏ bé*. Ở nước ta hiện nay, tình trạng báo động đang xảy ra đối với các ngôn ngữ như Arem, Măliêng, Chứt, Brâu, Rơmăm, Sila, Ođú, Bôy, Cólao, v.v.

5. Nhiều ngôn ngữ dân tộc ở Việt Nam đã có chữ viết. Các hệ thống chữ viết này thực sự có vai trò quan trọng trong sự nghiệp phát triển văn hoá - giáo dục của các dân tộc. Tuy nhiên phần lớn các hệ thống chữ viết hiện hành cần phải được cải tiến cho khoa học hơn nữa. Đó là nhiệm vụ đang được đặt ra cho các nhà ngôn ngữ học.

### TÀI LIỆU THAM KHẢO

1. Đặng Nghiêm Vạn, Nguyễn Trúc Bình, Nguyễn Văn Huy, Thanh Thiên, *Những nhóm dân tộc thuộc ngữ hệ Nam Á ở Tây Bắc Việt Nam*, NXB KHXH, 1972.
2. Đặng Nghiêm Vạn, Bước đầu tìm hiểu tình hình sử dụng nhiều tiếng nói ở miền núi miền Bắc nước ta, *Ngôn ngữ*, số 2, 1971.
3. Đoàn Thiện Thuật, *Hệ thống ngữ âm tiếng Tày - Nùng*, trong cuốn: "Tìm hiểu ngôn ngữ các dân tộc thiểu số ở Việt Nam", Tập 1, Viện Ngôn ngữ học ấn hành, 1972.
4. Đoàn Văn Phúc, *Ngữ âm tiếng Êđê*, NXB KHXH, 1996.
5. Bùi Khánh Thế, *Về cơ cấu tiếng Chăm*, Luận án phó Tiến sĩ, 1981.
6. Hoàng Ma, Tình hình nghiên cứu ngôn ngữ các dân tộc thiểu số ở nước ta trong 30 năm qua, *Ngôn ngữ*, số 4, 1975.

7. Nguyễn Văn Lợi, Loại hình học đồng đại và lịch đại: hiện tượng tiền mũi trong các ngôn ngữ Đông Nam Á, *Ngôn ngữ*, số 1 - 2, 1987.
8. Nguyễn Văn Lợi, *Một số nét về tiếng Mèo ở Việt Nam*, trong cuốn: "Tìm hiểu ngôn ngữ các dân tộc thiểu số ở Việt Nam", Tập 1, Viện Ngôn ngữ học ấn hành, 1972.
9. Nguyễn Văn Lợi, Sự phân loại và tình hình phân bố ngôn ngữ các dân tộc ở miền Nam nước ta, *Ngôn ngữ*, số 1, 1977.
10. P. K. Benedict, Thai, Kadai and Indonesian - A new alignment in Southeastern Asia, *Amer. Anthr. New series*, Vol. 11, No 4.
11. Phú Văn Hán, *Cơ cấu ngữ âm và chữ viết tiếng Chăm Việt Nam và tiếng Melayu Malaysia*, Luận án Tiến sĩ Ngữ Văn, 2003.
12. Tạ Văn Thông - Lê Đông, *Tiếng Hà Nhi*, NXB Văn hoá Dân tộc, 2001.
13. Thái Văn Chải, *Các chức năng của tiếng Khmer Nam Bộ*, trong cuốn: "Tiếng Việt và các ngôn ngữ Đông Nam Á", NXB KHXH, 1988.
14. Trần Quốc Vương (chủ biên), *Cơ sở văn hoá Việt Nam*, NXB ĐHQG Hà Nội, 1997.
15. Vương Hoàng Tuyên, *Các dân tộc nguồn gốc Nam Á ở miền Bắc Việt Nam*, NXB Giáo dục, 1963.
16. Vương Hoàng Tuyên, *Sự phân bố dân tộc và cư dân ở miền Bắc Việt Nam*, NXB Giáo dục, 1966.
17. Vũ Minh Giang, *Khu vực học với nghiên cứu Phương Đông*, trong cuốn: *Đông phương học Việt Nam*, Kỷ yếu hội thảo Đông phương học Việt Nam lần thứ nhất, NXB Đại học Quốc gia Hà Nội, 2001, tr.55.

VNU JOURNAL OF SCIENCE, SOC. SCI. HUMAN, T XXI, N<sup>o</sup>1, 2005

## ETHNIC MINORITY LANGUAGES IN VIETNAM IN THE ASPECT OF AREA STUDIES

Assoc.Prof. Dr. Mai Ngoc Chu

*College of Social Sciences and Humanities, VNU*

Ethnic minority languages in Vietnam are always researched in historical - comparative and typological - comparative aspects. In this article, they are considered from the view of area studies.

Ethnic minority languages in Vietnam belong to 5 areas: Tay Bac, Viet Bac, Tay Nguyen, Trung Bo and Nam Bo. There are often one or two most popular languages in each area, for example, Thai and Hmong in Tay Bac, Tay - Nung in Viet Bac, Khmer in Nam Bo.

The author has tried to mention the distribution of ethnic minority languages in each area, their characteristics and the problems which have to be solved, for example writing systems, tones, etc.