

# TRAO ĐỔI



## Ý KIẾN

### VẤN ĐỀ XÁC ĐỊNH LẠI THÀNH PHẦN MỘT SỐ DÂN TỘC Ở VIỆT NAM

## VẤN ĐỀ NGÔN NGỮ TRONG XÁC ĐỊNH THÀNH PHẦN DÂN TỘC Ở VIỆT NAM

NGUYỄN VĂN LỢI

### I. TIÊU CHÍ NGÔN NGỮ TRONG XÁC ĐỊNH THÀNH PHẦN DÂN TỘC

Trong mấy chục năm qua, nghiên cứu xác minh thành phần ngôn ngữ - tộc người ở Việt Nam luôn là một hướng nghiên cứu quan trọng, tập trung nỗ lực của những người nghiên cứu ngôn ngữ học, dân tộc học, lịch sử và đã thu được những kết quả đáng khích lệ. Tuy nhiên hiện nay, trong lĩnh vực này, đang đặt ra nhiều vấn đề lý luận và thực tiễn cần giải quyết. Đến nay chúng ta vẫn chưa có được câu trả lời thống nhất cho câu hỏi tưởng như đơn giản: ở Việt Nam có bao nhiêu dân tộc, bao nhiêu ngôn ngữ?

Vào những năm cuối của thập niên 50 của thế kỷ XX, một bảng danh mục thành phần dân tộc nước ta được công bố khá chi tiết trong công trình *Các dân tộc thiểu số ở Việt nam* của tập thể tác giả Lã Văn Lô, Nguyễn Hữu Thấu, Mai Văn Trí, Ngọc Anh, Mạc Như Đường. Theo các tác giả trên, ở Việt Nam có 64 dân tộc, trong đó có 63 dân tộc thiểu số, nói các ngôn ngữ thuộc 3 ngữ hệ: Hán Tạng, Môn - Khơ me, Malayô-Pôlinêdiêng. Điều đáng chú ý là, trong công trình này, các ngôn ngữ Tày - Thái, Mèo - Dao, Lạc - Việt (Việt - Mường) được coi thuộc ngữ hệ Hán Tạng. Trong từng nhóm cũng có một số nhầm lẫn, ví dụ tiếng Mảng và các ngôn ngữ của cư dân được gọi là Xá (Khơ mú, Kháng, La Ha)

được xếp vào Tạng - Miến; Phù Lá xếp vào Tày; Rục, Mày được coi thuộc nhóm Môn - Khơ me. Mặc dù còn một số hạn chế, công trình của các tác giả trên là công trình đầu tiên hệ thống được một danh sách các dân tộc ở nước ta và nó là cơ sở khoa học về mặt tộc người để tiến hành cuộc Tổng điều tra dân số Miền Bắc tháng 3 năm 1960.

Những năm 60-70, vấn đề ngôn ngữ tộc người cũng được một số tác giả đề cập đến trong các công trình miêu tả những tộc người, một số tộc người ở các khu vực như: *Các dân tộc nguồn gốc Nam Á ở Miền Bắc Việt Nam* của Vương Hoàng Tuyên, *Các nhóm dân tộc Tày Nùng Thái ở Việt Nam* của Lã Văn Lô, Đặng Nghiêm Vạn, *Các dân tộc miền núi Bắc Trung Bộ* của Mạc Đường, *Những nhóm dân tộc thuộc ngữ hệ Nam Á ở Tây Bắc* của Đặng Nghiêm Vạn, Nguyễn Trúc Bình..., các tộc người Tây Nguyên trong *Cao nguyên Miền Thượng* của Cửu Long Giang và Toan Ánh...).

Tại cuộc Hội thảo về xác định thành phần dân tộc, tháng 6 và tháng 11 năm 1973, do Viện Dân tộc học thuộc Ủy ban KHXH phối hợp với các vụ thuộc Ủy ban Dân tộc Trung ương, khoa Sử trường Đại học Tổng hợp Hà Nội tổ chức, các tiêu chí xác định thành phần dân tộc ở nước ta,

trong đó có tiêu chí ngôn ngữ, lần đầu tiên được nghiên cứu, thảo luận một cách khoa học.

Trong các Hội thảo, một số quan điểm sau đây được nhiều nhà nghiên cứu dân tộc học ủng hộ và được xem như quan điểm chỉ đạo trong việc xác định thành phần dân tộc ở Việt Nam:

Đơn vị cơ bản để tiến hành xác định thành phần dân tộc là cộng đồng tộc người hay là dân tộc. Đó là một tập đoàn người tương đối ổn định, được hình thành trong quá trình lịch sử dựa trên những mối liên hệ chung về địa vực cư trú, sinh hoạt kinh tế, tiếng nói, những đặc điểm sinh hoạt văn hoá, và dựa trên ý thức về thành phần và tên gọi chung.

*Cộng đồng tộc người ngôn ngữ* là khái niệm rộng hơn, và nhóm địa phương là khái niệm hẹp hơn.

Vận dụng các tiêu chuẩn phân loại tộc người do các nhà dân tộc học Xô Viết đưa ra vào điều kiện Việt Nam, các nhà dân tộc học Việt Nam cho rằng trong công tác xác minh dân tộc ở Việt Nam, tiêu chí về lãnh thổ hay địa vực cư trú không thể coi là tiêu chí quan trọng; còn lại tiêu chí về *ngôn ngữ, đặc điểm văn hoá và ý thức tự giác dân tộc* là những tiêu chí chủ yếu để xác minh thành phần dân tộc ở nước ta.

Mặc dù có sự tương đối thống nhất trong quan niệm về dân tộc, về các tiêu chí cơ bản để xác định thành phần dân tộc, nhưng trong thực tế áp dụng cho từng trường hợp cụ thể vẫn không thống nhất, trong nhiều trường hợp bộc lộ rõ tính chủ quan, chưa đủ tư liệu, không sát với tình hình thực tế khách quan, chưa chú ý đúng mức tới tính phức tạp, tình trạng phân bố đan xen các tộc người (ngôn ngữ), nhóm địa phương (phương ngữ), và tính đa dạng về trình độ phát triển kinh tế xã hội, văn hoá, ngôn ngữ của các dân tộc ở nước ta.

Như đã nhận xét ở trên, trong một số công trình xuất bản trước năm 1950, miêu thuật về một dân tộc cụ thể, một nhóm dân tộc ở một khu vực nào đấy, một số tác giả không phân biệt phân loại dân tộc và phân loại ngôn ngữ; nhiều bảng phân loại thực chất là phân loại ngôn ngữ. Trong các công trình nghiên cứu dân tộc học sau này, nhất là trong các cuộc thảo luận về xác định thành phần dân tộc ở nước ta, hầu hết các nhà nghiên cứu đều nhấn mạnh cần phân biệt *phân loại dân tộc* và *phân loại ngôn ngữ* (hay còn gọi là phân loại ngôn ngữ tộc người- Ethnolinguistics Classification). Trong khi thảo luận về những tiêu chí xác định thành phần dân tộc cũng như khi phải giải quyết những trường hợp cụ thể, các nhà nghiên cứu dân tộc đều nhấn mạnh tầm quan trọng của vấn đề ngôn ngữ trong công tác xác minh dân tộc. Một số tác giả đề nghị xếp tiêu chí về ngôn ngữ lên vị trí đầu tiên trong 3 tiêu chí để xác định dân tộc: **1- Ngôn ngữ, 2. Văn hoá và 3. Ý thức tự giác dân tộc.** Tuy nhiên, các tác giả không thống nhất trong việc hiểu vấn đề ngôn ngữ trong xác minh dân tộc. Trong khi thảo luận không tác giả nào giải thích rõ vận dụng tiêu chí ngôn ngữ trong xác định thành phần dân tộc là như thế nào? Khi giải quyết những trường hợp cụ thể, các tác giả tỏ ra không nhất quán và lúng túng. Mặc dù ai cũng thừa nhận tầm quan trọng của tiêu chí ngôn ngữ, nhưng trong nhiều trường hợp, vấn đề ngôn ngữ bị bỏ qua khi nhập hoặc tách các nhóm cư dân thành một (hoặc hai) dân tộc. Người Tóng và Dao, Thuỷ và Pà Thẻn nói các ngôn ngữ thuộc các ngữ hệ hoàn toàn khác nhau (Tóng, Đông, Kam và Thuỷ (Sủi) thuộc nhánh Kam (Đông) - Thuỷ ngữ hệ Tai Ka Đại; Dao và Pà Thẻn thuộc ngữ hệ Hmông Miên). Hầu hết các tác giả tỏ ra lúng túng khi vận dụng tiêu chí ngôn ngữ: lấy gì làm cơ sở để

nói thứ tiếng này và tiếng kia là cùng một ngôn ngữ hay là 2 ngôn ngữ độc lập. Các nhà nghiên cứu không phân biệt so sánh cấu trúc (so sánh loại hình học - cấu trúc ngữ âm, cấu trúc câu, cấu trúc danh ngữ) và so sánh xác định quan hệ cội nguồn (so sánh từ vựng cơ bản) và thường không biết lấy tiêu chí nào làm cơ sở để đồng nhất hay phân biệt về ngôn ngữ. Các tác giả đặc biệt lúng túng khi chứng minh sự giống khác nhau giữa các ngôn ngữ rất gần về cội nguồn như Tày và Thái, Na Mèo và Hmông. Trong các trường hợp này, các tác giả chỉ đưa ra một vài từ cơ bản, hay sơ đồ cấu trúc câu, trật tự chủ ngữ, vị ngữ, bổ ngữ, trật tự yếu tố hạn định và được hạn định trong cấu trúc nhóm danh từ. Nhưng đối với các ngôn ngữ gần nhau về cội nguồn như Tày và Thái, Nguồn và Việt Mường, Rục và A rem, Na Mèo và Hmông... thì chúng có tỉ lệ khá lớn từ vựng cơ bản chung, và các đặc điểm về cấu trúc là hoàn toàn đồng nhất. Tiếng Pa Dí, Tày Nặm ở Lào Cai có vốn từ cơ bản, cấu trúc câu gần với tất cả các ngôn ngữ nhánh Tai (Tai Trung tâm như Tày Nùng, Choang Nam, Tai Bắc như Giáy, Đioi, Choang Bắc và cả với Tai Tây Nam như Thái Đen, Thái Trắng hay Lào, Thái Lan; không thể căn cứ vào một số từ vựng và cấu trúc câu có trật tự S-P-O (Chủ ngữ-Vị ngữ-Bổ ngữ) để nói rằng các ngôn ngữ này gần với Tày, và xếp nhóm người này vào dân tộc Tày.

Ngoài ra, cũng cần nói thêm rằng các bảng phân loại ngôn ngữ tộc người các dân tộc ở nước ta, kể cả các bảng phân loại được đưa ra trong các tài liệu chính thức, cũng có nhiều điều cần thảo luận. Chưa nói đến sự phân loại ngôn ngữ, phương ngữ, sắp xếp vị trí các ngôn ngữ, mà cả cách phân chia thành họ (ngữ hệ), dòng, nhánh, nhóm... cũng không hợp lý, không dựa trên các nguyên tắc, phương pháp khoa học, không phản ánh được những

thành tựu, kết quả nghiên cứu của giới khoa học trong và ngoài nước.

Do sự không thống nhất trong việc vận dụng các tiêu chí (trong đó có tiêu chí về ngôn ngữ) để xác định thành phần dân tộc nên danh mục các dân tộc Việt Nam rất khác nhau giữa các tác giả, kể cả danh mục chính thức đã công bố ở các thời kỳ khác nhau, mà trước hết là khác nhau về số lượng dân tộc: 64 (1959), 59 (1975), 54 (1979).

Việc thay đổi danh mục dân tộc qua các thời kỳ còn có một nguyên nhân khác: thành phần tộc người ở nước ta cũng như bất cứ quốc gia nào khác không phải bất biến mà có thể biến đổi qua các thời kỳ lịch sử. Cùng với những thay đổi rất lớn trong đời sống kinh tế, xã hội văn hoá của các dân tộc ở nước ta, trong mấy thập niên gần đây, các quá trình tộc người cũng có những biến động, có những thay đổi lớn. Đặc biệt, thay cho xu hướng cố kết tộc người, hoặc xích lại gần nhau giữa các tộc người, nhóm địa phương xảy ra khoảng vài chục năm trước, thì hiện nay đang có xu hướng các nhóm địa phương muốn tách thành các tộc người riêng.

Hiện nay, trong các tài liệu, văn bản Nhà nước, trên các phương tiện thông tin đại chúng khi đề cập đến vấn đề thành phần các dân tộc ở nước ta, có những ý kiến không thống nhất về thành phần, tên gọi các dân tộc. Đồng thời, nhiều địa phương đã có những kiến nghị đến các cơ quan hữu trách ở Trung ương về việc xem xét lại việc xác định thành phần dân tộc. Trong giai đoạn phát triển mới của đất nước, để thực hiện tốt chính sách dân tộc của Đảng và Nhà nước, việc xác định lại thành phần một số dân tộc ở nước ta là yêu cầu khá cấp thiết.

Chúng tôi nhất trí rằng, việc xác định thành phần dân tộc, hay nói cách khác là phân loại dân tộc về mặt dân tộc học

không đồng nhất với việc phân loại ngôn ngữ (dù là phân loại ngôn ngữ tộc người); Danh mục các ngôn ngữ, phương ngữ được phân loại theo quan hệ cội nguồn ở Việt Nam chắc chắn không trùng với danh mục các dân tộc, các nhóm địa phương (và cũng được phân loại về ngôn ngữ tộc người - theo quan hệ cội nguồn ngôn ngữ). Phân loại cội nguồn (thân tộc) ngôn ngữ là phân loại theo quan hệ lịch sử xa / gần của các ngôn ngữ: ngữ hệ, ngữ chi, nhánh, nhóm, ngôn ngữ, phương ngữ, thổ ngữ. Để phân loại cội nguồn ngôn ngữ, cần áp dụng các thủ pháp của phương pháp nghiên cứu ngôn ngữ học lịch sử. Tiêu chí phân loại trong trường hợp này, thuần túy là các tiêu chí về ngôn ngữ học. Trong khi đó, phân loại dân tộc trong công tác xác định thành phần dân tộc phải dựa vào một loạt các tiêu chí, trong đó ngôn ngữ chỉ là một trong các tiêu chí.

Cần phải nhắc lại rằng, vấn đề ngôn ngữ trong xác định thành phần dân tộc chính là vấn đề xác định ranh giới ngôn ngữ / phương ngữ: các thứ tiếng của cư dân đang bàn đến là: 1. Các ngôn ngữ độc lập, 2. Các phương ngữ của một ngôn ngữ chung, và 3. Phương ngữ của một ngôn ngữ khác.

Trong khi áp dụng các tiêu chí trên vào việc phân định phương ngữ ngôn ngữ cũng như vận dụng tiêu chí ngôn ngữ vào việc xác định thành phần dân tộc ở nước ta cần chú ý đến những đặc điểm của Việt Nam về ngôn ngữ, lịch sử, tộc người. Phần lớn các ngôn ngữ ở Việt Nam là các ngôn ngữ chưa có chữ viết, chưa hình thành ngôn ngữ văn học, không có một hình thức siêu phương ngữ chung. Như vậy, đối với việc xác định ngôn ngữ và phương ngữ, chúng ta không thể vận dụng tiêu chí thứ hai, mà chỉ có thể vận dụng tiêu chí thứ nhất (về mức độ hiểu tiếng nói lẫn nhau của các cư dân) và tiêu chí thứ 3

(về ý thức tự giác đồng tộc của các cư dân đó).

Ngoài ra, khi chúng ta đang bàn đến vấn đề ngôn ngữ trong xác định thành phần dân tộc, có nghĩa là, chúng ta chưa có câu trả lời đủ rõ về ý thức tự giác đồng tộc, thì đương nhiên, chúng ta lại chỉ có thể dựa vào tiêu chí thứ 1: về khả năng hiểu biết tiếng nói lẫn nhau của các cư dân.

Vấn đề xác định dân tộc / nhóm địa phương thực chất chỉ xảy ra ở các nhóm người nói các ngôn ngữ có quan hệ gần gũi về cội nguồn (về mặt cội nguồn thuộc cùng một nhóm, ví dụ, Cuối và Thổ cùng thuộc nhóm Thổ-Poọng trong các ngôn ngữ Việt Mường; Xá Phó và Phù Lá thuộc nhóm Lô Lô trong các nhánh Miến Điện - Lô Lô, dòng Tạng - Miến; Rơ Ngao và Xơ Đăng cùng thuộc tiểu nhóm Ba Na Bắc trong nhóm Bahnar, nhánh Môn - Khơ me...), mà ít khi xảy ra đối với các cư dân nói các ngôn ngữ xa về cội nguồn (ví dụ, giữa các cư dân nói ngôn ngữ Tạng - Miến với cư dân nói ngôn ngữ Môn - Khơ me). Như vậy, xác định mức độ của sự giống/ khác nhau chính là xác định mức độ của quan hệ cội nguồn.

Trong các tiêu chí để xác định dân tộc được một số nhà khoa học nói đến có tiêu chí về lịch sử của các cư dân (số phận lịch sử chung, hay nguồn gốc chung). Lịch sử ngôn ngữ phản ánh lịch sử tộc người. Mức độ quan hệ cội nguồn các ngôn ngữ phản ánh quan hệ lịch sử và tiến trình biến đổi ngôn ngữ, đồng thời cũng phản ánh quan hệ lịch sử các cư dân nói các ngôn ngữ đó. Do đó, bản chất của việc xác định khả năng hiểu biết tiếng nói lẫn nhau (mức độ gần gũi ngôn ngữ) chính là việc xác định mức độ quan hệ cội nguồn của chúng.

Hơn nữa, về phương pháp, việc xác định khả năng hiểu biết tiếng nói lẫn nhau của các nhóm người có liên quan

chính là xác định mức độ giống nhau và khác nhau giữa các thứ tiếng đó. Sự giống / khác nhau thể hiện trên tất cả các cấp độ: ngữ âm, từ vựng, ngữ pháp. Các ngôn ngữ ở Việt Nam đều thuộc cùng một loại hình: ngôn ngữ đơn lập, đơn tiết. Sự khác nhau về ngữ pháp không thật lớn. Xác định mức độ giống và khác nhau giữa các ngôn ngữ, về phương pháp chỉ là xác định sự giống và khác nhau về từ vựng (từ vựng cơ bản và từ vựng văn hoá) và ngữ âm (các quá trình biến đổi ngữ âm đã xảy ra, tạo nên hệ thống ngữ âm nói chung và diện mạo ngữ âm của từng từ - những tương ứng ngữ âm, nói riêng). Xác định sự giống / khác nhau về từ vựng cũng như sự giống / khác nhau về ngữ âm do các quá trình biến đổi đã xảy ra (những cách tân trong ngôn ngữ) chính là các thủ pháp để xác định mức độ gần gũi và phân loại các ngôn ngữ thân thuộc.

Như vậy, vấn đề ngôn ngữ trong xác định thành phần dân tộc chính là vấn đề phân định ngôn ngữ / phương ngữ. Do tình hình thực tế ở nước ta, việc phân định ngôn ngữ / phương ngữ, xét cả về bản chất của vấn đề (bản thể), cũng như phương pháp giải quyết vấn đề (nhận thức), thực chất là việc xác định mức độ quan hệ thân thuộc ngôn ngữ. Chúng ta có thể vận dụng các phương pháp xác định mức độ thân thuộc các ngôn ngữ để giải quyết vấn đề ngôn ngữ khi xác định thành phần dân tộc. Hai thủ pháp quan trọng để xác định và phân loại ngôn ngữ thân thuộc là thống kê từ vựng và xác định những cách tân trong ngôn ngữ.

## II. ĐẶC ĐIỂM HÌNH THÀNH CÁC NGÔN NGỮ Ở VIỆT NAM VÀ VẤN ĐỀ NGÔN NGỮ TRONG XÁC ĐỊNH THÀNH PHẦN DÂN TỘC

Do những điều kiện lịch sử, địa lý đặc biệt, từ xa xưa, Việt Nam là nơi tụ hội của nhiều các bộ tộc, bộ lạc thuộc các thành

phần nhân chủng, văn hoá, ngôn ngữ khác nhau. Các quá trình hình thành tộc người, hình thành ngôn ngữ diễn ra khá phức tạp. Kết quả là, hiện nay ở nước ta có khá nhiều tộc người, nói các thứ tiếng thuộc các ngôn ngữ, nhóm ngôn ngữ, ngữ chi, ngữ hệ khác nhau.

Ở Việt Nam đối tượng xác minh thành phần dân tộc khá phức tạp, bởi vì quá trình hình thành tộc người diễn ra trong thời gian dài, trên một địa bàn tụ hội nhiều luồng di cư. Một trong những đặc điểm hình thành tộc người (ngôn ngữ) ở nước ta là sự tồn tại rất nhiều nhóm địa phương (phương ngữ). Mỗi nhóm địa phương (phương ngữ) hình thành theo một cách riêng, phụ thuộc vào điều kiện lịch sử, địa lý, xã hội. Có thể khái quát thành 2 kiểu hình thành nhóm địa phương (phương ngữ).

**Kiểu thứ nhất** là các nhóm địa phương tách ra từ các cư dân đang trên đường qui tụ thành bộ tộc. Trước Cách mạng Tháng Tám, các nhóm cư dân này chưa trải qua hình thái bộ tộc, chưa trải qua hình thái nhà nước phôi thai. Những nhóm người này thường là cư dân bản địa, đã cư trú lâu đời ở địa bàn đang sinh sống. Trong đời sống kinh tế, xã hội, văn hoá, họ chịu ảnh hưởng sâu đậm của các thể chế thị tộc bộ lạc. Những nhóm người này thường khép kín xã hội, văn hoá, ngôn ngữ của mình trong phạm vi, khu vực hẹp, thậm chí từng buôn, làng. Họ có những mối liên hệ yếu ớt với những người đồng tộc, tuy nói cùng một thứ tiếng, nhưng cư trú cách xa. Về mặt ngôn ngữ, quá trình biến đổi chủ yếu theo hướng phân ly: vốn từ một ngôn ngữ chung (ngôn ngữ mẹ) tách thành các ngôn ngữ độc lập (ngôn ngữ con). Quan hệ thân thuộc giữa các ngôn ngữ theo kiểu hình chuỗi: các ngôn ngữ phân bố gần nhau thì có mức độ gần

gũi cội nguồn hơn. Do đó, mỗi nhóm người này thường tự xem mình là tộc người độc lập, với một tộc danh riêng. Về ngôn ngữ, ranh giới giữa ngôn ngữ và phương ngữ, nhất là giữa những tiếng phân bố liên kề, không thật rõ ràng. Việc xác minh thành phần tộc người các nhóm này khá phức tạp, đòi hỏi vận dụng đồng thời nhiều tiêu chí. Quá trình hình thành nhóm địa phương theo kiểu trên là đặc trưng cho các tộc người ở Bắc Trường Sơn, Tây Nguyên như các nhóm: Pacoh, Pahi, Taoih, Ka Tang, Kan Tu, Cơ Tu ở Quảng Trị, Thừa Thiên - Huế; Ba Na, Rơ Ngao, Xơ Đăng, Giẻ Triêng, Pơ Nông, Ka Dong, Ha Lăng ở Quảng Nam, Tây Nguyên; các nhóm địa phương của Ê đê, Gia rai ở Plây Cu, Đăk Lăk; các nhóm địa phương của Mơ nông, Mạ, Cơ ho, Xtiêng ở Lâm Đồng, Bình Phước... Hiện nay, một loạt vấn đề liên quan đến việc xác định thành phần tộc người đang đặt ra đối với các nhóm trên. Ở khu vực Bình Trị Thiên cần giải quyết vấn đề ranh giới dân tộc / nhóm địa phương (ngôn ngữ / phương ngữ) giữa các nhóm Trì, Măng Coong, Vân Kiều, Bru, giữa nhóm Pa kôh và Tà Ôih, giữa Pa hi và Pa koh; giữa các nhóm Ka Tu, Kan Tu, Phương ở Thừa Thiên- Huế, Quảng Nam; giữa các nhóm Ka Dong, Ha Lăng, Tơ Đạ, Xơ Đăng, giữa Giẻ, Triêng, Ve ... ở Quảng Nam, Quảng Ngãi và Bắc Tây Nguyên; vấn đề các nhóm địa phương trong dân tộc Ê đê, Gia rai ở Đăk Lăk, Plâycu cũng như nhóm Mơ Nông, Cơ Ho, Mạ, ở Lâm Đồng, Bình Phước cũng cần được thảo luận thêm.

**Kiểu thứ hai** trong sự hình thành các nhóm địa phương là nhóm địa phương được tách ra từ các tộc người đã trải qua hoặc đang bước vào quá trình qui tụ thành bộ tộc. Họ từ các vùng khác nhau ở Nam Trung Quốc, di cư đến Việt Nam và hoà vào môi trường tộc người mới ở nước ta.

Nơi xuất phát của các cuộc di cư thường là các tỉnh Nam và Tây Nam Trung Quốc như Quảng Đông (các nhóm Dao), Tứ Xuyên, Quý Châu, Quảng Tây, Vân Nam (các nhóm Nùng, Giáy, Bố Y, Hmông, Di). Thời gian di cư cũng khác nhau, có những nhóm vào Việt Nam khá sớm như La Chí, Cơ Lao, Pu Páo, Dao, một số nhóm Nùng; có những nhóm vào muộn hơn, như Hmông, Tống, Thủy, Pà Thên... Họ có thể đi thành từng nhóm, gồm những người đồng họ, đồng tộc, có khi gồm những người khác (ví dụ nhóm Pà Thên, Tống, Thủy ở Chiêm Hoá, Tuyên Quang). Trong quá trình di cư, một số yếu tố văn hoá, thậm chí cả ngôn ngữ của họ đã thay đổi (nhóm Cao Lan). Khi đã định cư ở Việt Nam, họ mau chóng hoà vào các tộc người khác, văn hoá, ngôn ngữ của họ có những thay đổi lớn. Những biến đổi này là kết quả các quá trình giao thoa, vay mượn, thống hợp, thay thế trong tất cả các bình diện ngôn ngữ từ ngữ âm, từ vựng, đến kết cấu ngữ pháp. Tính chất không thuần nhất, pha trộn, lai tạp về văn hoá và ngôn ngữ làm cho việc xác minh thành phần ngôn ngữ tộc người khá phức tạp.

Quá trình hình thành các nhóm địa phương theo kiểu thứ hai xảy ra phổ biến ở nhiều tộc người thuộc khu vực Tây Bắc, Việt Bắc, đặc biệt khu vực biên giới Việt - Trung.

Cần chú ý một số đặc điểm về ngôn ngữ của các nhóm này khi xác minh thành phần dân tộc:

1. Số lượng người nói các ngôn ngữ thường rất ít, các nhóm địa phương lại phân bố đan xen, do đó các ngôn ngữ của họ nằm ở bên bờ của nguy cơ tiêu vong (ví dụ các nhóm Cơ Lao, Tống, Thủy...). Nếu gộp họ vào các dân tộc (ngôn ngữ) lớn hơn, càng đẩy nhanh quá trình tiêu vong các ngôn ngữ của họ.

2. Tình hình phân chia các phương ngữ, thổ ngữ khá phức tạp. Ở một số nhóm, tiếng nói vừa khác biệt theo ngành (vốn từ các vùng khác nhau ở Trung Quốc di cư đến), vừa khác nhau theo vùng lãnh thổ cư trú ở Việt Nam. Hiện nay, các nhóm địa phương (các ngành) của cùng một tộc người, mặc dù cư trú cận kề, vẫn không thể giao tiếp với nhau bằng ngôn ngữ của mình. (Trường hợp các nhóm Dao (như Dao Đỏ, Dao Tiền, Dao Lô Gang...) thuộc ngành Kim Mùn và các nhóm Dao (như Dao Thanh Y, Dao Quần Trắng...) thuộc Kim Miên; Hmông Lênh và Hmông Sua; Cơ Lao Đỏ và Cơ Lao Trắng).

3. Sự thống nhất tộc người chủ yếu ở ý thức tự giác dân tộc, các đặc trưng văn hoá.

4. Căn cứ chủ yếu để xác định sự giống / khác nhau về ngôn ngữ của các nhóm địa phương là so sánh từ vựng cơ bản và tìm hiểu các quá trình biến đổi ngữ âm (các cách tân trong ngôn ngữ).

5. Dựa vào cứ liệu ngôn ngữ có thể tìm được khu vực cư trú gốc của các cư dân. Đây là nguồn cứ liệu quan trọng để nghiên cứu lịch sử tộc người và xác minh dân tộc.

Dưới đây là một số trường hợp minh hoạ:

**Tiếng Tống:** Người Tống được coi là một bộ phận của dân tộc Dao. Họ tự nhận là người Kam, ngôn ngữ của họ là một biến thể của tiếng Kam (Đông) thuộc dòng Đông Tai, họ Tai-Ka Đai. So sánh từ vựng cơ bản và tìm hiểu quá trình biến đổi thanh điệu có thể xác định được rằng người Kam ở Việt Nam vốn di cư từ vùng LiPing Shuikou thuộc tỉnh Quý Châu di cư đến khoảng 150 năm trước.

**Tiếng Thuỷ:** Người Thuỷ được coi là một bộ phận của dân tộc Pà Thẻn. Tiếng Thuỷ thuộc nhóm Đông Thuỷ, dòng Đông

Tai, ngữ hệ Tai Ka Đai. Cứ liệu ngôn ngữ chỉ ra rằng, người Thuỷ vốn từ huyện lý Sandong, châu Sandu, tỉnh Quý Châu di cư đến (có lẽ cùng với người Pà Thẻn).

**Tiếng Giáy:** Người Giáy tự nhận là Pu Jăj. Ngôn ngữ của họ là biến thể của ngôn ngữ Pu Ji, thuộc nhóm Tai Bắc, nhánh Tai, dòng Đông Tai, họ Tai - Kaday. Dựa vào cứ liệu ngôn ngữ có thể xác định họ từ khu vực Tây Nam Quý Châu di cư đến.

**Tiếng Xá Phó:** Hiện người Xá Phó được coi là một bộ phận của dân tộc Phù Lá. Tuy nhiên, ngôn ngữ của họ khác xa tiếng Phù Lá và khá gần với tiếng Di, phân bố ở Tứ Xuyên, Vân Nam - Trung Quốc. Trong ngôn ngữ này còn lưu giữ những đặc điểm rất cổ của các ngôn ngữ Tạng Miến (như các tổ hợp phụ âm kl, khl, gl...).

**Tiếng Ngạn:** Người Ngạn hiện được coi là một bộ phận của người Tày. Từ vựng cơ bản, các quá trình biến đổi thanh điệu trong tiếng Ngạn chỉ ra rằng ngôn ngữ này thuộc nhóm Tai Bắc (Tày Nùng, Choang Nam thuộc nhóm Tai Trung tâm).

**Tiếng Na Mèo:** Người Na Mèo nói một ngôn ngữ thuộc nhóm Mèo-ngôn ngữ Hmu, ở Quý Châu, Trung Quốc; di cư đến Việt Nam khoảng 3 thế kỷ trước. Sau khi đến Việt Nam, họ hoà vào với cư dân Tày ở Lạng Sơn, Tuyên Quang. Họ thay đổi trong cách ăn mặc, cư trú, một số phong tục, tập quán. Tuy nhiên, họ vẫn ý thức họ là một tộc người độc lập, với một tên tự gọi (tộc danh) *Na mẻo sla* (người Mèo đen). Ngôn ngữ của họ tuy chịu ảnh hưởng khá rõ của tiếng Tày, nhưng vẫn được các thành viên sử dụng trong giao tiếp gia đình và thôn bản giữa những người đồng tộc. Ngôn ngữ, ý thức tự giác dân tộc (thông qua tên tự gọi) là những tiêu chí để

nhận diện người Na Mèo và phân biệt họ với các tộc người khác.

Trong quá trình nghiên cứu, chúng tôi thấy rằng, vẫn còn 45 nhóm địa phương thuộc 21 dân tộc cần làm sáng rõ trong xác minh thành phần dân tộc liên quan đến ngôn ngữ. Đó là những dân tộc với các nhóm ngôn ngữ như sau:

**Các dân tộc ở miền Bắc**

1. Dân tộc Tày

- 1.1. Ngôn ngữ của nhóm Thu Lao ở Lao Cai
- 1.2. Ngôn ngữ của nhóm Tày Nặm ở Lao Cai
- 1.3. Ngôn ngữ của nhóm Pa Dí ở Lao Cai
- 1.4. Ngôn ngữ của nhóm Ngạn ở Quảng Ninh
- 1.5. Ngôn ngữ của nhóm Thổ ở Đà Bắc Hoà Bình.

2. Dân tộc Nùng

- 2.1. Ngôn ngữ nhóm Nùng Vền ở Cao Bằng
- 2.2. Ngôn ngữ nhóm Nùng Giang ở Cao Bằng

3. Dân tộc Giáy và Bố Y

- 3.1. Ngôn ngữ nhóm Giáy
- 3.2. Ngôn ngữ nhóm Bố Y
- 3.3. Ngôn ngữ nhóm Pu Nà (Quý Châu)
- 3.5. Ngôn ngữ nhóm Tu Dí

4. Dân tộc Hmông

- 4.1. Ngôn ngữ nhóm Na Miểu

5. Dân tộc Dao

- 5.1. Ngôn ngữ nhóm Tống

6. Dân tộc Pa Thên

- 6.1. Ngôn ngữ nhóm Thuỷ

7. Dân tộc Phù Lá

- 7.1. Ngôn ngữ nhóm Xá Phó

- 7. 2. Ngôn ngữ các nhóm địa phương của dân tộc Hoa

**Các tỉnh miền Trung và Bắc Trường Sơn**

- 9. Ngôn ngữ nhóm Nguồn (thuộc dân tộc Kinh)

10. Dân tộc Thổ

- 10.1. Ngôn ngữ nhóm Đan Lai
  - 10.2. Ngôn ngữ nhóm Li Hà
  - 10.3. Ngôn ngữ nhóm Tày Poọng
11. Dân tộc Chứt
- 11.1. Ngôn ngữ nhóm Mày
  - 11.2. Ngôn ngữ nhóm Sách
  - 11.3. Ngôn ngữ nhóm Mã Liềng
  - 11.4. Ngôn ngữ nhóm Kri
  - 11.5. Ngôn ngữ nhóm A Rem

13. Dân tộc Bru-Vân Kiều

- 13.1. Ngôn ngữ nhóm Vân Kiều
- 13.2. Ngôn ngữ nhóm Trì
- 13.3. Ngôn ngữ nhóm Ma Coong

14. Dân tộc Tà ôi

- 14.1. Ngôn ngữ nhóm Pa Hi
- 14.2. Ngôn ngữ nhóm Pa Koh
- 14.3. Ngôn ngữ nhóm Ta Ôih

15. Dân tộc Cơ tu

- 15.1. Ngôn ngữ nhóm Kan Tu (Phương) ở A Lưới

**Các dân tộc ở Trường Sơn-Tây Nguyên**

16. Dân tộc Giẻ -Triêng

- 16.1. Ngôn ngữ nhóm Ve (La Veh, Vê)
- 16.2. Ngôn ngữ nhóm Pơ Noong
- 16.3. Ngôn ngữ nhóm Triêng

17. Dân tộc Xơ Đăng

- 17.1. Ngôn ngữ nhóm Tơ Dơ (Xơ rá)
- 17.2. Ngôn ngữ nhóm Ka Dong
- 17.3. Ngôn ngữ nhóm Ha Lăng
- 17.4. Ngôn ngữ nhóm Rơ Ngao

18. Dân tộc Cơ ho

- 18.1. Ngôn ngữ nhóm Xrê
- 18.2. Ngôn ngữ nhóm Lát
- 18.3. Ngôn ngữ nhóm Chil

19. Dân tộc Raglai

- 19.1. Ngôn ngữ nhóm Rai

20. Dân tộc Xtiêng

- 20.1. Ngôn ngữ nhóm Ta Mun

21. Dân tộc Chăm

- 21.1. Ngôn ngữ nhóm Hroi