



## QUAN HỆ CAO LAN, SÁN CHÍ XÉT VỀ MẶT NGÔN NGỮ

NGUYỄN VĂN LỢI

### **Dẫn nhập**

Trong công tác xác định thành phần dân tộc ở Việt Nam vào những năm 70 của thế kỷ XX, khái niệm dân tộc (hay tộc ng- òi) th- òng đ- ợc hiểu theo nghĩa hẹp: Là một cộng đồng ng- òi hình thành trong quá trình lịch sử, có những đặc điểm văn hoá chung t- ơng đối ổn định và một số đặc điểm tâm lý t- ơng ứng, cũng có ý thức về sự thống nhất của mình và về sự khác biệt giữa mình với các cộng đồng khác cùng loại. Theo các nhà dân tộc học, những đặc tr- ng về ngôn ngữ, đặc điểm sinh hoạt văn hoá, ý thức tự giác dân tộc là những tiêu chuẩn chủ yếu để xác định thành phần dân tộc.

Quả vậy, nếu nh- sự thống nhất lãnh thổ và hoạt động kinh tế là nhân tố quan trọng để hình thành tộc ng- òi, thì ngôn ngữ cùng với những đặc điểm văn hoá là những tiêu chí chính để thống nhất và phân biệt tộc ng- òi này với tộc ng- òi khác. Ngôn ngữ cũng là yếu tố bên trong, nh- sợi dây liên kết các thành viên của tộc ng- òi.

Trong danh mục thành phần dân tộc ở n- ớc ta, Cao Lan và Sán Chí đ- ợc coi là một dân tộc trong cộng đồng 54 dân tộc Việt Nam, với tên gọi Sán Chay (hoặc Cao Lan - Sán Chí). Tuy nhiên hiện nay, câu hỏi Cao Lan và Sán Chí là một tộc ng- òi hay là hai tộc ng- òi độc lập đang đ- ặt ra

nhiều vấn đề cần giải quyết. Thực tế là ngày nay, cộng đồng Cao Lan và Sán Chí sử dụng 2 ngôn ngữ hoàn toàn khác nhau, không thể hiểu nhau. Trong tr- òng hợp nh- vậy, thì tiêu chí sự thống nhất ngôn ngữ cần đ- ợc hiểu và xem xét nó trong mối quan hệ với các tiêu chí khác (ý thức tự giác tộc ng- òi, thống nhất về văn hoá...) nh- thế nào, khi xác định lại thành phần dân tộc Sán Chay (Cao Lan Sán Chí) ? Để giải quyết vấn đề này, chúng ta phải làm sáng tỏ một loạt các câu hỏi sau đây:

Ngôn ngữ của cộng đồng Cao Lan và cộng đồng Sán Chí đang sử dụng có đặc điểm gì? Nguồn gốc, quá trình phát triển, và mối quan hệ giữa chúng trong lịch sử nh- thế nào? Về mặt đồng đại, các ngôn ngữ này hành chức ra sao ? Quan niệm, thái độ, nguyện vọng những ng- òi sử dụng các ngôn ngữ đó nh- thế nào?

Để trả lời những câu hỏi trên, trong năm 2003, Viện Ngôn ngữ học đã tiến hành điều tra tiếng Cao Lan và Sán Chí ở các tỉnh Quảng Ninh, Bắc Giang, Tuyên Quang, Thái Nguyên là những địa ph- òng có đông đồng bào Cao Lan, Sán Chí c- trú. Mục đích điều tra nhằm:

1. Xác định những đặc điểm cấu trúc ngữ âm, ngữ pháp, vốn từ cơ bản của hai ngôn ngữ để tìm hiểu nguồn gốc và quá trình phát triển lịch sử của chúng.

2. Tìm hiểu cảnh huống ngôn ngữ (tình hình sử dụng ngôn ngữ, chữ viết, trạng thái song ngữ, đa ngữ, ý thức tự giác tộc ngữ- ời và ý thức tự giác ngôn ngữ, thái độ nguyện vọng của nhân dân thuộc hai cộng đồng Cao Lan và Sán Chí về những vấn đề liên quan.

T- liệu về cấu trúc ngôn ngữ đ- ọc thu thập bằng ph- ơng pháp ngôn ngữ học điền dã. Tại mỗi địa điểm điều tra, chúng tôi thu thập bảng 1.000 từ cơ bản, một số hiện t- ợng ngữ pháp. Riêng đối với tiếng Cao Lan ở xã Đèo Gia, huyện Lục Ngạn tỉnh Bắc Giang, t- liệu đ- ọc thu thập gồm 2.000 từ, 2.000 câu theo 10 bảng hỏi dự báo về các hiện t- ợng ngữ pháp cơ bản, một số văn bản truyện cổ tích, bài hát dân gian. Ph- ơng pháp điều tra xã hội học bằng cách điều tra trực tiếp theo bảng hỏi đ- ọc áp dụng để thu thập t- liệu về cảnh huống ngôn ngữ. Tại mỗi địa điểm chúng tôi tiến hành điều tra ở 100-150 ng- ời (tổng số 850 ng- ời ở 8 địa điểm), theo bảng hỏi gồm 20 câu hỏi. Việc điều tra về cấu trúc và cảnh huống ngôn ngữ Cao Lan, Sán Chí đ- ọc tiến hành ở các điểm sau đây:

	<b>Cao Lan</b>	<b>Sán Chí</b>
Quảng Ninh	Xã Thanh Sơn, huyện Ba Chẽ	Xã Đại Dực, huyện Tiên Yên
Bắc Giang	Xã Đèo Gia, huyện Lục Ngạn	Xã Kiên Lao, huyện Lục Ngạn
Tuyên Quang	Xã Đông Thọ, huyện Sơn D- ơng	Bình Yên, huyện Sơn D- ơng
Thái Nguyên	Xã Phấn Mễ, huyện Phú L- ơng	Xã Yên Lạc, huyện Phú L- ơng

T- liệu điều tra cấu trúc ngôn ngữ đ- ọc số hoá, phân tích bằng máy tính và l- u

giữ trong cơ sở dữ liệu về ngôn ngữ Cao Lan, Sán Chí.

**Kết quả điều tra**

**I. Quan hệ Cao Lan, Sán Chí xét về mặt lịch đại**

T- liệu điều tra về cấu trúc ngôn ngữ Cao Lan và Sán Chí cho phép chúng tôi đ- a ra một số nhận định về nguồn gốc, sự phát triển lịch sử và quan hệ của hai ngôn ngữ.

**A. Ngôn ngữ của cộng đồng Cao Lan**

Về quan hệ cội nguồn của tiếng Cao Lan đã đ- ọc A. G. Haudricourt, David Streker, Jerold Edmondson thảo luận. Các tác giả thống nhất xem Cao Lan thuộc dòng Tai, họ ngôn ngữ Tai Ka Đai, nh- ng ch- a thống nhất về vị trí của ngôn ngữ này trong dòng Tai: thuộc nhóm Tai Bắc, hay Tai Trung. Để xác định mức độ của quan hệ cội nguồn và phân nhóm các ngôn ngữ Tai, các tác giả th- ờng dựa vào từ vựng và các hiện t- ợng cách tân (innovation), đặc biệt là hiện t- ợng cách tân trong biến đổi âm đầu và thanh điệu. Về từ vựng và các biến đổi thanh điệu, tiếng Cao Lan, Nùng An (ở Cao Bằng) và Ts- n Va (Quảng Ninh) có nhiều điểm dị th- ờng, gây nên những bất đồng ý kiến trong việc xác định quan hệ của chúng với các ngôn ngữ Tai.

**Từ vựng Cao Lan**

1.Tr- ớc hết, cần phải khẳng định rằng, từ vựng tiếng Cao Lan thuộc về từ vựng chung các ngôn ngữ Tai.

2.Tỷ lệ từ chung giữa Cao Lan và Giáy (Tai Bắc) là 80%, trong khi với Tày (Tai Trung tâm) là 75%. Nếu xét các từ có tính chất chỉ báo để nhận diện các nhóm Tai

Bắc (Giáy, Pu Giay, Bó Y, Choang Bắc, Seak), Tai Trung (Tày, Nùng, Choang Nam) và Tai Tây Nam (Thái Đen, Thái Trắng, Lào, Lự, Thái Xiêm), dựa theo công trình của Lý Ph- ơng Quế và các tác giả khác, thì một số từ trong tiếng Cao Lan chung với các ngôn ngữ Tai Bắc, nh- ng một số từ lại thuộc về nhóm Tai Trung tâm.

3. Trong từ vựng Cao Lan có một số từ chung với tiếng Sán Chí, gốc Hán, có hình thức ngữ âm Sán Chí. Điều này chứng tỏ ng- òi Cao Lan cũng nh- Sán Chí đã từng sử dụng cùng một ph- ơng ngữ tiếng Hán.

4. Trong lớp từ cơ bản tiếng Cao Lan, chúng ta tìm thấy từ *bùn* /pam/ có nguồn gốc từ ngôn ngữ Dao. Từ này cũng có trong tiếng Sán Chí.

Trong từ vựng Cao Lan cũng xuất hiện một số từ vay m- ợn tiếng Việt. Dạng thức ngữ âm các từ này chứng tỏ sự vay m- ợn mới xảy ra.

### Về các quá trình biến đổi ngữ âm trong ngôn ngữ Cao Lan

1. Tiếng Cao Lan cũng nh- Nùng An và Ts- n Va có các quá trình biến đổi ngữ âm vừa đặc tr- ng cho Tai Bắc, vừa đặc tr- ng cho Tai Trung. Một số quá trình biến đổi phụ âm đầu nói chung và quá trình biến đổi liên quan đến tiêu chí bật hơi chứng tỏ Cao Lan gần với Tai Bắc. Đồng thời, trong Cao Lan cũng có quá trình \*f>ph t- ơng tự nh- Tai Trung. Quá trình biến đổi các tổ hợp phụ âm có yếu tố thứ hai là \*-l/\*-r- trong tiếng Cao Lan vừa giống với tiếng Giáy (Tai Bắc), vừa gần với tiếng Tày (Tai Trung). Về đặc điểm này, d- ờng nh- Cao Lan nằm ở giữa của hai

cực: cực biến đổi tối đa - tiếng Giáy, tất cả các tổ hợp phụ âm đều biến thành /Δ/, và cực kia là tiếng Tày Cao Bằng: bảo l- u tối đa các đối lập, các tổ hợp phụ âm t- ơng ứng trong Proto Tai trở thành các phụ âm khác nhau trong Tày hiện đại, nh- nhận xét của Jerold Edmondson .

Một điểm cần l- u ý là, xét về quá trình biến đổi phụ âm đầu, Cao Lan và Nùng An có nhiều điểm t- ơng đồng, và khác với các ph- ơng ngữ Tày, Nùng ở Việt Nam.

2. Hệ thống thanh điệu tiếng Cao Lan ngày nay - kết quả của quá trình biến đổi lịch sử, có nhiều đặc điểm khác th- ờng so với các ngôn ngữ Tai. Ngôn ngữ này có hệ thống thanh điệu nh- sau:

	A	B	C	DS	DL
<b>1</b>	44	32/	35	44	32
<b>2</b>	32	32-	4540	32	32-

3. Trong ngôn ngữ Cao Lan, sự phát triển hệ thanh chủ yếu liên quan đến tiêu chí có/không hữu thanh của phụ âm đầu trong Proto Tai. Nói cách khác, quá trình biến đổi thanh điệu Cao Lan chủ yếu là kết quả của quá trình vô thanh hoá phụ âm đầu hữu thanh, theo nguyên tắc hữu thanh-âm vực thấp (voiced-low principle). Đặc điểm này khác với tiếng Giáy (Tai Bắc) và phần nào giống với các ngôn ngữ Tai Trung.

4. Quá trình vô thanh hoá không chỉ làm hệ thống thanh điệu phát triển theo kiểu nhân đôi, mà còn ảnh h- ưởng đến phẩm chất của các nguyên âm. Đây là hiện t- ợng khác th- ờng, không thấy trong các ngôn ngữ Tai mà th- ờng thấy trong một số ngôn ngữ Môn Khơme và Nam Đảo.

5. Một hiện tượng đặc biệt khác trong ngôn ngữ Cao Lan là hiện tượng biến thanh trong ngữ l- u (tone sandhi). Hiện tượng biến thanh trong ngữ l- u phổ biến ở các ngôn ngữ Hán, Hmông, Miên (Dao), nh- ng hiếm thấy trong các ngôn ngữ Tai nói riêng và Tai Ka Đại nói chung. Sự có mặt của hiện tượng này trong tiếng Cao Lan đ- ọc giải thích bằng sự ảnh hưởng từ tiếng Hán mà cộng đồng Cao Lan đã từng sử dụng nh- ngôn ngữ thứ hai.

Từ những cứ liệu trên cho phép đi đến kết luận:

1. Ngôn ngữ mà ng- ời Cao Lan sử dụng là một ngôn ngữ có những đặc điểm bất thường, kết quả của quá trình biến đổi lịch sử do tiếp xúc ngôn ngữ. Xét về quan hệ cội nguồn, tiếng Cao Lan thuộc về ngôn ngữ Choang Bắc, nhóm Tai Bắc, phân bố ở Quý Châu và Bắc Quảng Tây. Nh- vậy, cộng đồng ng- ời Cao Lan, tr- ớc khi di cư xuống phía nam, đã cư trú ở khu vực ranh giới Quý Châu, Quảng Tây- vùng hiện nay phân bố các ngôn ngữ Choang Bắc, PuGiay thuộc nhóm Tai Bắc. Điều này phù hợp với các tài liệu dân tộc học: ng- ời Cao Lan vẫn nhớ về quê hương cũ của mình ở khu vực Quý Châu, Quảng Tây; bộ phận Cao Lan ở Ba Chẽ Quảng Ninh vẫn tự gọi (dân tộc) mình là Chùng Nhàn - "ng- ời Choang".

2. Lớp từ vựng vốn đặc trưng cho ngôn ngữ Tai Trung, cũng nh- một số quá trình biến đổi âm đầu và thanh điệu mang đặc điểm của Tai Trung có trong Cao Lan là kết quả của quá trình tiếp xúc với các cư dân nói các phương ngữ, ngôn ngữ Tai Trung (Choang Nam) khi cộng đồng Cao Lan di chuyển xuống phía nam. Những dấu vết của Tày Nùng có trong Cao Lan

không nhiều, điều đó chứng tỏ những đặc điểm của ngôn ngữ nhóm Tai Trung trong Cao Lan không phải do ảnh hưởng trực tiếp từ các phương ngữ Tày, Nùng ở Việt Nam, mà đã có từ tr- ớc khi cộng đồng Cao Lan di cư vào Việt Nam.

3. Cộng đồng Cao Lan đã từng là cộng đồng song ngữ; ngôn ngữ thứ hai họ sử dụng (nói và viết) là một thổ ngữ thuộc nhóm phương ngữ Việt (Yue) tiếng Hán. Ngôn ngữ này hiện nay cộng đồng Sán Chí vẫn đang sử dụng. Kết quả của quá trình tiếp xúc ngôn ngữ là sự tồn tại trong tiếng Cao Lan hiện nay, một lớp từ vựng Hán, và một vài đặc điểm ngữ âm do sự tác động của tiếng Hán (hiện tượng biến thanh trong ngữ l- u).

### *B. Ngôn ngữ của cộng đồng Sán Chí*

Ngôn ngữ của cộng đồng Sán Chí sử dụng thuộc ngôn ngữ Hán, hệ Hán - Tạng.

Tiếng Hán bao gồm nhiều phương ngữ, và nhóm phương ngữ, trong đó có nhóm phương ngữ Việt (Yue) đ- ọc cư dân ở Quảng Đông và Quảng Tây sử dụng. Nhóm phương ngữ Việt lại bao gồm nhiều phương ngữ và thổ ngữ, tiêu biểu là phương ngữ Quảng Châu, tỉnh Quảng Đông. Tiếng Hán (phương ngữ Việt) ở Quảng Tây lại phân thành 4 vùng thổ ngữ, có những đặc điểm riêng về ngữ âm, từ vựng. Ngoài ra, ở Quảng Tây tồn tại một thứ tiếng đ- ọc gọi là Bình Thoại (Ping Hua).

### *Một số đặc điểm ngữ âm tiếng Sán Chí và nguồn gốc của ngôn ngữ này*

Ngôn ngữ của ng- ời Sán Chí mang những đặc điểm điển hình của phương

ngữ Việt, trong đó đặc điểm của nổi trội nhất về ngữ âm là :

1. Sự có mặt của các âm cuối tắc vô thanh: -p, -t, -k (không có trong tiếng Hán Quan Thoại và các ph- ơng ngữ Hán khác).

2. Hệ thống thanh điệu phong phú gồm 10 thanh điệu, với sự bảo l- u đối lập các thanh trong tiếng Hán Trung đại (phân thành các loại thanh: Bình, Th- ơng, Khứ Nhập và loạt thanh: Âm (Cao), D- ơng (Thấp)).

Hệ thống thanh điệu Sán Chí Quảng Ninh gồm 10 thanh, có đặc điểm ngữ âm và quá trình biến đổi đ- ợc phản ánh trong bảng sau đây:

Loại Loạt	A (Bình Thanh)	B (Th- ơng Thanh)	C (Khứ Thanh)	D (Nhập Thanh)	
				DS	DL
Âm (Cao)	41	33	35	33	44
D- ơng (Thấp)	55	23	21	23	21

Có thể rút ra một số nhận xét sau về tiếng Sán Chí:

1. Quá trình phân chia thành hai loại của các thanh nhập trong tiếng Sán Chí, t- ơng tự nh- tiếng Quảng Đông, phụ thuộc vào tr- ờng độ nguyên âm.

2. Trong 4 vùng thổ ngữ ph- ơng ngữ Việt, tiếng Sán Chí gần với thổ ngữ vùng thứ t- -vùng bờ biển phía nam. Điều này phù hợp với tài liệu dân tộc học: cộng đồng Cao Lan và Sán Chí vẫn nhớ rằng tổ tiên họ đã từng c- trú ở vùng Khâm Châu, Liêm Châu, Lôi Châu thuộc vùng bờ biển Quảng Đông-Quảng Tây, Trung Quốc.

3. Một số hiện t- ợng ngữ âm khác của tiếng Sán Chí lại chứng tỏ rằng, ngôn ngữ này gần với Bình Thoại hơn là tiếng Quảng Đông.

Theo một số tác giả, một trong những đặc điểm để phân biệt Bình Thoại và Quảng Đông là quá trình biến đổi phụ âm đầu hữu thanh cổ thành vô thanh không bật hơi trong Bình Thoại và vô thanh bật hơi trong tiếng Quảng Đông. Theo tiêu chí này, tiếng Sán Chí gần với Bình Thoại hơn Quảng Đông: phụ âm \*b trở thành p, \*d trở thành t , \*g trở thành k, nh- Bình Thoại.

Nh- vậy, xét về mặt ngữ âm, tiếng Sán Chí thuộc về ph- ơng ngữ Việt, trong đó có một số đặc điểm nghiêng về tiếng Quảng Đông, (vùng thổ ngữ bờ biển phía Nam), một số đặc điểm nghiêng về tiếng Bình Thoại Nam Quảng Tây.

### Một số đặc điểm từ vựng, ngữ pháp tiếng Sán Chí

Hệ thống từ vựng Sán Chí thuộc về tiếng Hán.

Tuy nhiên, trong tiếng Sán Chí có một số từ (thuộc cả lớp từ cơ bản và từ văn hoá) có nguồn gốc Tai, dạng thức ngữ âm hoàn toàn đồng nhất với Cao Lan.

Cấu trúc ngữ nghĩa của một số tr- ờng từ vựng ngữ nghĩa theo mô hình Tai.

Một điều đáng l- u ý là, trong từ vựng Sán Chí cũng có một số từ gốc Dao, chẳng hạn từ có nghĩa *bùn, đất*.

Về ngữ pháp, khác với các ph- ơng ngữ Hán và t- ơng tự ngôn ngữ Tai, trong tiếng Cao Lan từ chỉ định đứng sau danh từ đ- ợc chỉ định. Nhóm danh từ (danh ngữ) tiếng Sán Chí có mô hình t- ơng tự ngôn

ngữ Tai: số từ - loại từ- danh từ chính-từ chỉ định; còn tiếng Hán có trật tự khác hẳn: Từ chỉ định-số từ-loại từ-danh từ chính.

Có thể xem các hiện tượng bất thường về từ vựng- ngữ nghĩa và ngữ pháp trên là dấu vết cơ tầng Tai trong ngôn ngữ Sán Chí.

Những đặc điểm mang tính bất thường trong ngôn ngữ Sán Chí cho phép chúng ta đi đến kết luận rằng:

1. Cũng như Cao Lan, cộng đồng Sán Chí trước đây vốn nói một phương ngữ thuộc ngôn ngữ Choang Bắc, nhóm Tai Bắc.

2. Trong quá trình di cư xuống phía nam, trong sự tiếp xúc với các dân nói ngôn ngữ Hán (vùng thổ ngữ bờ biển phía Nam, phương ngữ Việt), họ đã tiếp nhận ngôn ngữ này làm ngôn ngữ thứ hai và trở thành cộng đồng song ngữ. Dần dần, cộng đồng Sán Chí từ bỏ ngôn ngữ thứ nhất của mình, tiếng Hán trở thành ngôn ngữ giao tiếp hàng ngày, và trở thành ngôn ngữ mẹ đẻ của các thế hệ sau. Ngôn ngữ thứ nhất (tiếng Choang Bắc) dần dần đi ra khỏi đời sống của người Sán Chí, những dấu vết (cơ tầng) trong hệ thống ngữ âm, từ vựng, ngữ pháp trong tiếng Sán Chí.

## **II. Quan hệ ngôn ngữ Cao Lan, Sán Chí xét về mặt đồng đại**

### **A. Tình trạng sử dụng ngôn ngữ trong cộng đồng Cao Lan**

Kết quả điều tra xã hội ngôn ngữ học cho phép chúng tôi đi đến một số kết luận:

1. Cộng đồng Cao Lan hiện nay là cộng đồng song ngữ, đa ngữ. Phổ biến hơn cả là

song ngữ Cao Lan - Việt. Kết quả điều tra chỉ ra rằng, người Cao Lan thuộc các giới, lứa tuổi khác nhau đều sử dụng thành thạo tiếng Việt.

2. Hình thức song ngữ khác, khá phổ biến là hình thức song ngữ Cao Lan -Tày (Thái Nguyên, Tuyên Quang, Bắc Giang), Cao Lan - Hoa (Quảng Ninh).

3. Trong cộng đồng Cao Lan ở hầu hết các tỉnh đều có những người còn sử dụng được tiếng Hán-Sán Chay (ngôn ngữ Hán mà người Sán Chí đang sử dụng) - dạng ngôn ngữ viết và nói. Số người biết chữ Hán-Sán Chay (chữ vuông, đọc theo âm Sán Chí) không nhiều, đều có độ tuổi trên 45, thường trong những gia đình có truyền thống làm thầy cúng. Số người có biết (chỉ biết đọc hoặc chỉ biết viết, đọc, viết ở trình độ trung bình) chiếm tỉ lệ 3% ở Thái Nguyên, 5% ở Tuyên Quang, 4% ở Quảng Ninh. Số người vừa biết đọc vừa biết viết tương đối thành thạo có tỉ lệ rất thấp - chỉ 1% ở các địa phương. Số người biết tiếng Hán-Sán Chay khác nhau ở các vùng. Những vùng cộng đồng Cao Lan và Sán Chí phân bố gần nhau, số người biết tiếng Hán-Sán Chay khá cao, những vùng ít có điều kiện tiếp xúc, thì số người biết tiếng Hán-Sán Chay chiếm tỉ lệ thấp. Theo độ tuổi, những người trên 45 tuổi còn nhớ và biết thành thạo tiếng Hán - Sán Chay chiếm tỉ lệ cao hơn. Số người có thể sử dụng (nghe nói, hoặc có thể nghe được, không biết nói) ngôn ngữ nói Hán - Sán Chay có tỉ lệ 100% ở Quảng Ninh; 95 % ở độ tuổi trên 45 và 43% ở độ tuổi dưới 45 ở Thái Nguyên; 10 % ở Tuyên Quang.

### **B. Tình trạng sử dụng ngôn ngữ trong cộng đồng Sán Chí**

1. Cộng đồng Sán Chí cũng là cộng đồng đa ngữ. Trạng thái song ngữ phổ

biến hơn cả là song ngữ Sán Chí -Việt; tuy nhiên có sự khác nhau theo địa ph- ơng và lứa tuổi, giới.

2. Trạng thái song ngữ Sán Chí-Tày, Sán Chí-Nùng, phổ biến ở các tỉnh Tuyên Quang, Thái Nguyên, Bắc giang. Trạng thái song ngữ Sán Chí- Hoa khá phổ biến ở Quảng Ninh, và phụ thuộc vào độ tuổi, giới.

3. Số l- ợng ng- ời biết (đọc và viết, đọc không biết viết) chữ Hán –Sán Chay có tỉ lệ: 6% ở Quảng Ninh, 8% ở Tuyên Quang, 11% ở Thái Nguyên. Những ng- ời này phần lớn ở độ tuổi trên 45 và hầu hết là nam.

4. Số ng- ời Sán Chí biết tiếng Cao Lan nói chung, chiếm tỉ lệ thấp và phụ thuộc vào điều kiện tiếp xúc với ng- ời Cao Lan. ở Quang Ninh, nơi có ít ng- ời Cao Lan sinh sống, 100% ng- ời đ- ọc đ- ọc hỏi không biết tiếng Cao Lan, ở Thái Nguyên, có 4,28 % có thể nghe, nói và 17,4% có thể nghe, không biết nói ; ở Tuyên Quang có 28% nghe, nói đ- ọc và 27% nghe đ- ọc, không nói đ- ọc bằng tiếng Cao Lan. Hầu hết những ng- ời này ở độ tuổi trên 45 .

*C. Ý thức tự giác tộc ng- ời và ý thức tự giác ngôn ngữ trong cộng đồng Cao Lan và Sán Chí*

Để tìm hiểu ý thức tự giác tộc ng- ời, ý thức tự giác ngôn ngữ, chúng tôi điều tra những vấn đề sau:

- Tên gọi chính thức (tên gọi dân tộc đ- ọc khai báo trong các cuộc điều tra dân số, trong giấy khai sinh, lý lịch, chứng minh th- ).
- Tên tự gọi dân tộc.

- Tên gọi dân tộc mà đồng bào cho là phù hợp với nguyện vọng của mình.

- Ý kiến về tên gọi chung của hai cộng đồng Cao Lan và Sán Chí (Sán Chay, hoặc Cao Lan –Sán Chí).

- Quan niệm về việc nhận biết những ng- ời đồng tộc với mình (cùng tiếng nói, cùng dòng họ, cùng phong tục tập quán, có quan hệ hôn nhân, cùng di c- đến...).

- Quan niệm về tiếng nói của mình, giống, khác với ngôn ngữ nào mà họ biết.

- Quan niệm về ngôn ngữ Cao Lan và Sán Chí (giống /khác nhau, có thể giao tiếp chung).

- Ý kiến về việc đồng nhất hay tách biệt tộc ng- ời Cao Lan và Sán Chí.

#### \* Trong cộng đồng Cao Lan

**Về tên gọi tộc ng- ời và tên gọi ngôn ngữ**

*Tên gọi chính thức* (tên tự nhận, khai báo trong điều tra dân số hoặc giấy tờ chính thức) trong cộng đồng Cao Lan là tên gọi **Cao Lan** và **Sán Chay**, trong đó phổ biến là **Cao Lan**.

	Thái nguyên	Tuyên Quang	Quảng Ninh
Cao Lan	93%	100%	100%
Sán Chay	6,8%		

*Tên tự gọi* phổ biến trong cộng đồng Cao Lan là: **Hờn Bán** (/hôn21/ng- ời,/ ban35/ bản (ng- ời trong bản), (có khi chỉ là Bán "bản"), **Cao Lan, Sán Chay và Trại Cao**. Ngoài ra, đồng bào còn biết đến các tên gọi khác nh- **Chùng Nhàn** (Choang Nhân) ng- ời Chùng (Choang). Ở Bắc Giang, Thái Nguyên, Quảng Ninh,

tên tự gọi phổ biến nhất là **Hồn Bán** và **Cao Lan**; ở Tuyên Quang là **Sán Chay**.

	Bắc Giang	Thái Nguyên	Tuyên Quang	Quảng Ninh
Cao Lan	5,33	41,9%		
Hồn Bán	73,33%	41,9%		100%
Sán Chay		1,4%	100%	
Trại Cao	1,59%	1,4%		

*Tên gọi đ-ợc coi là phù hợp với ý nguyện của ng-ời Cao Lan*

Trong số các tên gọi Cao Lan, Hồn Bán, Sán Chay, Trại Cao, tên gọi **Cao Lan** và **Hồn Bán** đ-ợc nhiều ng-ời cho là phù hợp với nguyện vọng của cộng đồng mình. Riêng ở Tuyên Quang 53% ng-ời Cao Lan đ-ợc hỏi chấp nhận tên gọi **Sán Chay**.

	Bắc Giang	Thái Nguyên	Tuyên Quang	Quảng Ninh
Cao Lan	38%	60%	35%	
Hồn Bán	62%	1%		100%
Sán Chay		1%	53%	

**\* Trong cộng đồng Sán Chí**

*Tên gọi chính thức:*

**Sán Chay** :165/300 ( 55%)

**Sán Chí**: 134/300 (44,67%)

**Cao Lan- Sán Chí**: 1/300 (03%)

*Tên tự gọi*

**Sán Chay**: 297/300 (99%)

**Sán Chí**: 3/300 (1%)

*Tên gọi đ-ợc coi là phù hợp với ý nguyện của ng-ời Sán Chí:*

**Sán Chay**: 83% ở Thái Nguyên, 100% ở Tuyên Quang, 100% ở Quảng Ninh.

**Ý thức tự giác tộc ng-ời và tự giác ngôn ngữ**

Trong cả cộng đồng Cao lan và Sán Chí đều thống nhất ở cách đánh giá các "tiêu chí", những đặc điểm t-ơng đồng để họ nhận biết những ng-ời đồng tộc với mình. Trong các tiêu chí thống nhất về ngôn ngữ, về phong tục tập quán, cùng dòng họ, có quan hệ hôn nhân, thì sự thống nhất về ngôn ngữ có vị trí quan trọng. D-ối đây là tỉ lệ ý kiến đánh giá các tiêu chí trong cộng đồng Cao Lan ở Đèo Gia và Sán Chí ở Kiên Lao, huyện Lục Ngạn tỉnh Bắc Giang.

	Cao Lan	Sán Chí
Thống nhất về ngôn ngữ	92,6%	1,33%
Giống về phong tục tập quán	46,6%	9,33%
Cùng dòng họ	19,3%	39,33 %
Có quan hệ hôn nhân	0,13%	0,06%

Hiện nay, ng-ời Cao Lan vẫn ý thức rằng họ có nhiều nét t-ơng đồng về phong tục, tập quán, sinh hoạt văn hoá, lễ nghi, trang phục, nhà cửa với cộng đồng Sán Chí. Số ng-ời Cao Lan thừa nhận sự t-ơng đồng này có tỉ lệ khá cao: 89,7 % ng-ời Cao Lan ở Thái Nguyên, 95 % ở Quảng Ninh.

Trong cộng đồng Sán Chí, ý kiến thống nhất (giống nhau) về phong tục tập quán Cao Lan và Sán Chí có tỉ lệ nh- sau: 88% ng-ời Sán Chí ở Tuyên Quang, 58, 6% ở Thái Nguyên.

Trong khi đó hầu hết ng- ời Cao Lan đ- ợc hỏi lại cho rằng ngôn ngữ của họ hoàn toàn khác với ngôn ngữ của ng- ời Sán Chí: 97,4% ở Thái Nguyên, 100% ở Quảng Ninh, 89% ở Tuyên Quang. Số ý kiến của ng- ời Cao Lan ở 4 tỉnh đã điều tra cho rằng tiếng Cao Lan hoàn toàn khác tiếng Sán Chí chiếm tỉ lệ 93,30 %.

Đồng thời, ý kiến trong cộng đồng Sán Chí cho rằng ngôn ngữ của họ hoàn toàn khác ngôn ngữ của cộng đồng Cao Lan cũng có tỉ lệ cao: 85,5% ng- ời Sán Chí ở Thái Nguyên, 46,6% ở Bắc Giang.

Trong quan niệm nh- vậy, phần lớn ng- ời Cao Lan và Sán Chí coi Cao Lan và Sán Chí là hai tộc ng- ời riêng biệt. Sau đây là tỉ lệ ý kiến trong cộng đồng Cao Lan và Sán Chí cho rằng Cao Lan và Sán Chí là những tộc ng- ời độc lập:

	Cao Lan	Sán Chí
Bắc Giang	100%	100%
Thái Nguyên	75,6%	91%
Tuyên Quang	47%	76,1%
Quảng Ninh	96%	

**Ghi chú:** 100% ng- ời Sán Chí ở Tiên Yên, Quảng Ninh không biết ngoài cộng đồng Sán Chí của họ còn có cộng đồng Cao Lan, trong dân tộc Sán Chay (Cao Lan -Sán Chí).

## Kết luận

Từ kết quả điều tra, phân tích t- liệu về cấu trúc và cảnh huống ngôn ngữ của cộng đồng Cao Lan và Sán Chí cho phép chúng tôi đi đến một số nhận định về quan hệ Cao Lan và Sán Chí nh- sau:

1. Trong lịch sử, cộng đồng Cao Lan và Sán Chí có chung một nguồn gốc, sử dụng chung một ngôn ngữ thuộc Choang Bắc, nhóm Tai Bắc, dòng Tai, họ ngôn ngữ Tai Ka Đai. Khi di c- xuống phía nam

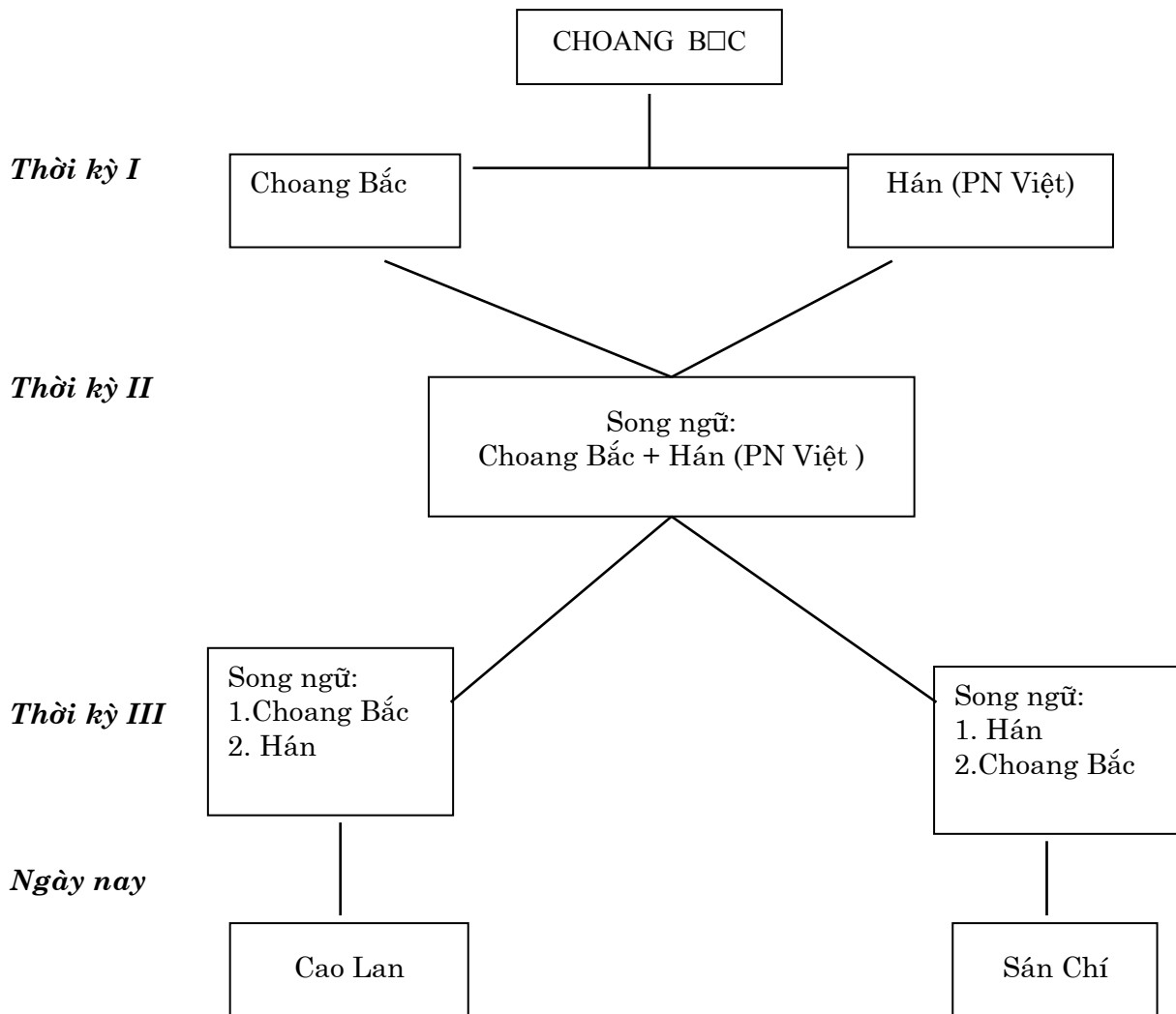
(Quảng Tây, Quảng Đông), do tiếp xúc với các cộng đồng tộc ng- ời khác (Choang Nam, Hán), họ trở thành cộng đồng song ngữ (đa ngữ) và tiếp nhận tiếng Hán (ph- ơng ngữ Việt (Yue) làm ngôn ngữ thứ hai. Sau đó, nhất là khi đã di c- vào Việt Nam (khoảng 400-500 năm tr- ớc), họ sống t- ơng đối phân tán, bộ phận ng- ời Cao Lan tiếp tục duy trì ngôn ngữ mẹ đẻ (Choang Bắc) của mình làm ngôn ngữ giao tiếp chính, đồng thời, một số ng- ời (tr- ớc hết là các thầy thuốc, thầy cúng) vẫn sử dụng tiếng (chữ Hán) trong một số lĩnh vực nh- lễ nghi, tôn giáo, sinh hoạt văn hoá (hát sinh ca)... Trong khi đó, bộ phận Sán Chí lại dùng tiếng Hán là ngôn ngữ chính và dần quên tiếng mẹ đẻ (ngôn ngữ Choang Bắc) của mình.

Những t- liệu về ngôn ngữ hoàn toàn phù hợp với t- liệu về dân tộc học về sự thống nhất tộc ng- ời Cao Lan và Sán Chí trong lịch sử:

- Thống nhất về sinh hoạt và văn hoá: phong tục tập quán, lễ nghi, cách thức sản xuất, canh tác.
- Thống nhất về quan hệ dòng họ và hôn nhân.
- Thống nhất về một số quan hệ khác, đặc biệt trong sinh hoạt văn hoá tinh thần nh- hát sinh ca, cúng bái, ý thức về nơi c- trú tr- ớc đây, về việc di c- ; hình thức tổ chức xã hội.

Nh- vậy, những tài liệu ngôn ngữ học và tài liệu dân tộc học chứng tỏ rằng xét về mặt lịch đại, Cao Lan và Sán Chí có nguồn gốc chung và tr- ớc khi di c- đến Việt Nam, họ đã từng có một quá trình tộc ng- ời và quá trình ngôn ngữ - tộc ng- ời chung. Quá trình ngôn ngữ - tộc ng- ời Cao Lan và Sán Chí có thể khái quát theo sơ đồ sau:

**Quá trình ngôn ngữ - tộc ngữ - ời Cao Lan và Sán Chí**



**Thời kỳ I:** Cộng đồng ngôn ngữ thống nhất, thuộc Choang Bắc, nhóm Tai Bắc, dòng Tai, ngữ hệ Tai Ka Đai, phân bố ở biên giới Quý Châu, Quảng Tây.

**Thời kỳ II:** Di c- xuống phía nam, tiếp xúc với tiếng Hán (ph- ơng ngữ Việt). Vẫn là cộng đồng ngôn ngữ thống nhất, nh- ng là cộng đồng song ngữ: ngôn ngữ thứ nhất Choang Bắc, ngôn ngữ thứ hai - Hán.

**Thời kỳ thứ III:** Tr- ớc và sau khi di c- vào Việt Nam. Vẫn là một cộng đồng song ngữ, nh- ng bắt đầu có sự phân chia thành

hai bộ phận, mỗi bộ phận có cách ứng xử khác nhau đối với tiếng Choang Bắc và tiếng Hán. Một bộ phận vẫn giữ tiếng Choang Bắc là ngôn ngữ thứ nhất, tiếng Hán là ngôn ngữ thứ hai, nh- ng ngày càng ít sử dụng và quên dần. Bộ phận thứ hai, lấy tiếng Hán là ngôn ngữ thứ nhất, còn tiếng Choang Bắc bị đẩy xuống vị thế ngôn ngữ thứ hai, dần ít đ- ợc sử dụng và bị quên ở thế hệ sau.

**Tình trạng hiện nay:** Hai cộng đồng ngôn ngữ độc lập: **Cao Lan** và **Sán Chí**

2. Tộc ng- ời (và ngôn ngữ) là sản phẩm lịch sử, đ- ợc xác định trong điều kiện lịch sử cụ thể. Có tr- ờng hợp trong lịch sử, hai cộng đồng tộc ng- ời (cộng đồng ngôn ngữ) tồn tại nh- hai tộc ng- ời (ngôn ngữ) độc lập, nh- ng sau đó, đ- ợc hợp lại thành một tộc ng- ời (ngôn ngữ) thống nhất; đồng thời, có không ít tr- ờng hợp, một tộc ng- ời (một ngôn ngữ) thống nhất, nh- ng trong quá trình phát triển lịch sử, lại tách ra thành các tộc ng- ời (ngôn ngữ) độc lập. Do đó, những cứ liệu về lịch đại chỉ có giá trị giải thích về lịch sử tộc ng- ời, lịch sử ngôn ngữ, không thể xem là tiêu chí để xác định tộc ng- ời (ngôn ngữ). Để xác định tộc ng- ời (ngôn ngữ) cần phải căn cứ vào tình trạng đồng đại hiện nay của các tộc ng- ời (ngôn ngữ) đó.

3. Để xác định tộc ng- ời, sự thống nhất về ngôn ngữ có vai trò quan trọng. Đây không chỉ là tiêu chí đ- ợc các nhà khoa học nhấn mạnh, mà thể hiện ngay trong cảm thức để nhận biết những ng- ời đồng tộc của ng- ời Cao Lan và Sán Chí. Đa số ng- ời Cao Lan và Sán Chí đều xem sự giống nhau về ngôn ngữ (có thể giao tiếp với nhau) là tiêu chí để nhận biết ng- ời đồng tộc.

4. Trong việc xác định sự thống nhất ngôn ngữ, tiêu chí về ý thức tự giác ngôn ngữ có vai trò quan trọng hàng đầu. Mỗi cá thể cũng nh- mỗi cộng đồng có cách quan niệm, cách đánh giá về việc họ nói ngôn ngữ gì, ngôn ngữ này gần với ngôn ngữ nào, khác với ngôn ngữ nào, mức độ giống/khác nhau và quan niệm của họ về các ngôn ngữ khác có liên quan. Ý thức tự giác ngôn ngữ cũng bao gồm trong tên gọi

ngôn ngữ, nhất là tên gọi họ mong muốn đ- ợc dùng để chỉ ngôn ngữ (lingonyme) của mình.

Tuyệt đại bộ phận ng- ời Cao Lan và Sán Chí ý thức rằng tiếng Cao Lan và Sán Chí là hai ngôn ngữ khác biệt và có tên gọi riêng: tiếng Cao Lan và tiếng Sán Chay.

5. Khi coi ngôn ngữ là đặc tr- ng tộc ng- ời, là tiêu chí xác định thành phần tộc ng- ời, cũng cần chú ý đến sự phát triển của mỗi ngôn ngữ và sự phát triển của quan hệ giữa các ngôn ngữ.

Sự thống nhất về ngôn ngữ (có chung ngôn ngữ thứ nhất-ngôn ngữ Choang Bắc và ngôn ngữ thứ hai - tiếng Hán) của cộng đồng Cao Lan và Sán Chí vốn có trong lịch sử, nay đã hoặc đang mất dần, tùy thuộc từng cá thể, và ở từng vùng c- trú. Những cá thể và nhóm ng- ời song ngữ Cao Lan (Choang Bắc) - Sán Chí (ph- ơng ngữ Việt, tiếng Hán) đang giảm dần và hiện nay có tỉ lệ rất thấp. Nhìn chung, tiếng Cao Lan và Sán Chí không còn là ph- ơng tiện giao tiếp chung của cả cộng đồng Cao Lan và Sán Chí.

6. Ý thức tự giác tộc ng- ời là tiêu chí quan trọng (nếu không muốn nói là tiêu chí quan trọng nhất, nh- quan niệm của nhiều nhà nhân học hiện đại), trong việc xác định tộc ng- ời. Ý thức tự giác dân tộc là quan niệm của một cộng đồng ng- ời về họ nh- một tộc ng- ời và khác với cộng đồng khác. Mỗi thành viên tự ý thức mình thuộc về một tộc ng- ời nhất định, và khác với những cộng đồng tộc ng- ời khác. Ý thức dân tộc có tính quần chúng và tính bền vững cao. Các nhà nhân học hiện đại

th- ờng gắn ý thức tự giác tộc ng- ời với cái gọi là quyền tự xác định của mỗi cá nhân, mỗi cộng đồng.

Khi bàn đến các vấn đề dân tộc, Lênin cũng đã nói đến ý thức tự giác dân tộc.

Khi nhấn mạnh tiêu chí ý thức tự giác tộc ng- ời, chúng ta muốn đề cao tính dân chủ, tính quần chúng trong công tác dân tộc nói chung, và trong việc xác định thành phần dân tộc nói riêng, nh- chủ tr- ơng của Đảng và Nhà n- ớc ta. Ngay trong những cuộc thảo luận về tiêu chí xác định thành phần dân tộc vào thập kỷ bảy m- ời của thế kỷ XX, các nhà dân tộc học n- ớc ta cũng đã rất đề cao tiêu chí ý thức tự giác tộc ng- ời: "nh- tiêu chí hàng đầu trong xác định thành phần dân tộc nói chung, và nh- yếu tố quyết định, khi thành phần dân tộc và ngôn ngữ không phù hợp nhau, nói riêng" (Bế Viết Đăng, 1975, tr 81).

6.1. Kết quả điều tra chỉ ra rằng, tuyệt đại bộ phận ng- ời Cao Lan và Sán Chí ý thức rằng hai cộng đồng Cao Lan và Sán Chí là hai tộc ng- ời độc lập.

6.2. Ý thức tự giác tộc ng- ời đ- ợc biểu hiện rõ qua tên gọi tộc ng- ời (tộc danh) trong đó có tên tự gọi và tên gọi mà họ mong muốn đ- ợc nhận. Bản thân sự tồn tại một tên gọi chung, thống nhất đã chứng tỏ rằng các thành viên tộc ng- ời nhận thức rõ về sự thống nhất tộc ng- ời của mình và sự khác biệt với tộc ng- ời khác. Tên gọi dân tộc có tính t- ợng tr- ơng cao.

Trong cộng đồng Cao Lan tên tự gọi phổ biến hơn cả là **Hòn Bán**, sau đó là

tên gọi **Cao Lan**; tên gọi họ mong muốn đ- ợc nhận là **Cao Lan**.

Trong cộng đồng Sán Chí tên tự gọi phổ biến và tên gọi họ mong muốn đ- ợc nhận là **Sán Chay**.

### Tài liệu tham khảo

Bauer Robert, Matthews Stephen. 2003. Cantonese. *The Sino-Tibetan Languages*, London 2003.

Bonifacy, Augst, 1905. *Monographie des Man Cao Lan*, Revue Indochinoise. No13, pp 899-928.

Bonifacy, Augst, 1907. *Étude sur les Cao Lan. T'oung Pao ou Archives concernant l' histoire, les langues, la géographie et l' ethnographie de l' asie orientale*. Serie ii. Vol.viii. Leiden: E.J Brill, 429-38.

Khổng Diễn (Chủ biên). 2002. *Dân tộc Sán Chay ở Việt Nam*. Hà Nội.

Diller, Anathony. *The Tai language family and the comparative method*. The Internationnal Conference on Tai Studies, July 29-31, 1998.

Bế Viết Đăng, 1975. *Về danh mục các dân tộc thiểu số ở miền bắc n- ớc ta*, trong sách "Về vấn đề xác định thành phần các dân tộc thiểu số ở miền Bắc Việt Nam". 57-83. Hà Nội.

Edmondson, Jerold A. 1992. *Fusion and diffusion in E,Guangxi, Province, China*. Paccific Linguistic Series, C-110. Australian National University, Canberra, 131-40.

Edmondson, Jerold A., Zheng Gúoqiao, Yang Quán, Gui Mingchao. 1997. *Yángan*

*Hua of Luocheng County, Guangxi Province.* Kam-Tai Studies 7-18.

Edmondson, Jerold A. 1998. *The language corridor: New evidence from Vietnam.* The International Conference on Tai Studies, July 29-31, 1998.

Edmondson, Jerold A., Gui Mingchao. 2001. *Tuguai Hua, a Type of PingHua of Northern Guangxi.* South East Asia Linguistics Studies XI

Li Fangkuei 1977. *A handbook of comparative Tai.* Honolulu: The University Press Hawaii.

Gokhman V.I. 1992. *Ngữ âm lịch sử các ngôn ngữ Tai.* Moskova. (Tiếng Nga)

Gregerson Kenneth J. and Edmondson Jerold A. 1998. *Some puzzles in Cao Lan.* The International Conference on Tai Studies, July 29-31, 1998.

Haudricourt, André. 1960. *Note sur les dialectes de la région de Moncay.* BEFEO. 50.167-77.

Haudricourt, André 1973. *Mấy điều nhận xét về lí luận và thực tiễn nhân một chuyến đi thăm các dân tộc thiểu số ở Tây Bắc và Việt Bắc.* Tc. Ngôn ngữ, Số 3

Mechkovskaja N. B. 1996. *Ngôn ngữ học xã hội.* Moskova (Tiếng Nga).

Nicolson Beth, 1998. *The Nung An Language of Vietnam: Stepchild or Aberrant son.* International Conference on Vietnamese Studies,

Norman Jerry . 2003. *The Chinese dialects Phonology.* The Sino-Tibetan Languages. London 2003. 72-83.

Pranee Kullavanijaya and Theraphan L-thongkum. 1998. *Linguistic criteria for deterring Tai ethnic groups: Case studies on central and south-western Tais.* The International Conference on Tai Studies, July 29-31, 1998.

Strecker, David. 1985. *The classification of the Caolan languages.* Southeast Asian Linguistic Studies presented to Andre G. Haudricourt. Bangkok: Mahidol University.

Đoàn Thiện Thuật. 1998. *Hệ thống thanh điệu tiếng Sán Chí.* Hội nghị quốc tế Việt Nam học.

Thurgood Graham, Lapolla Randy J. (edit.) 2003. *The Sino-Tibetan Languages.* London 2003.

Nguyễn Nam Tiến. 1975. *Lại bàn về hai nhóm Cao Lan và Sán Chí,* trong sách "Về vấn đề xác định thành phần các dân tộc thiểu số ở miền Bắc Việt Nam", Hà Nội, 274-286.

Nguyễn Khánh Toàn. 1975. *Một vài quan điểm cơ bản cần quán triệt trong quá trình xây dựng danh mục các dân tộc thiểu số ở miền Bắc nước ta,* trong sách "Về vấn đề xác định thành phần các dân tộc thiểu số ở miền bắc Việt Nam", Hà Nội, tr. 19-26.

Viện Dân tộc học. 1973. *Các dân tộc ít người ở Việt Nam (Các tỉnh phía Bắc),* Hà Nội.

Zang Yansheng, Wei Xinggyun 1997. *Regional Variants and Vernaculars in Zhuang.* Comparative Kadai: the Tai Branch. Dallas.