

# CẬP NHẬT THÔNG TIN CHO XỬ LÝ TÀI LIỆU CÓ LIÊN QUAN ĐẾN CÁC NGÔN NGỮ Ở VIỆT NAM

VƯƠNG TOÀN<sup>(\*)</sup>

Trong hoạt động nghiệp vụ thư viện, công tác phân loại tài liệu dù theo cách nào thì cũng có mục *Ngôn ngữ*. 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 thời điểm không phải luôn có sự chính xác tuyệt đối. Khi tiến hành phân loại và biên mục tài liệu liên quan đến các ngôn ngữ ở Việt Nam, người xử lý cần được cập nhật thông tin về các kết quả nghiên cứu chuyên ngành ngôn ngữ học.

Có dịp biên tập liên tục 166 tập thư mục “Thông báo sách mới nhập” vào các thư viện thuộc Viện KHXH Việt Nam (kể từ 1998 cho đến nay là số 12-2011, rút từ cơ sở dữ liệu (CSDL) tích hợp gần 100 ngàn biểu ghi) và theo dõi thông tin ngôn ngữ học từ nhiều năm, chúng tôi nhận thấy người xử lý tài liệu rất cần theo sát một số vấn đề đang đặt ra cho nghiên cứu các ngôn ngữ ở Việt Nam, đành rằng có những vấn đề chỉ được giải đáp dần dần và vẫn còn những vấn đề “treo lơ lửng”. Một trong số các vấn đề cần cập nhật thông tin để xử lý tài liệu theo quan điểm mới nhất là số lượng và cùng với nó là danh sách các ngôn ngữ tồn tại ở Việt Nam. Nói cách khác, phải

chỉnh sửa *đề mục* “Việt Nam - các ngôn ngữ” sao cho phù hợp với những thành tựu mới nhất của ngành ngôn ngữ học.

Trong số những vấn đề mà ngôn ngữ học đang tìm lời giải đáp có việc xác định số lượng và danh sách các ngôn ngữ (và các phương ngữ) hiện đang tồn tại ở Việt Nam và đặc biệt là trong tương quan với thành phần dân tộc đã qua nhiều lần được nghiên cứu và xác lập, với những thay đổi tách/gộp. Xây dựng danh mục các dân tộc ở nước ta là kết quả đánh dấu từng giai đoạn phát triển của ngành dân tộc học và các ngành hữu quan, *trong đó có ngôn ngữ học*. Theo dòng thời gian, chúng ta thấy các nhà nghiên cứu đã đưa ra *danh mục dân tộc* vào các năm 1959, 1974, 1979 và hiện đang tiếp tục nghiên cứu,...

Trong khi đó, việc phân loại ngôn ngữ có thể xét theo các phương diện khác nhau: tộc người, văn hoá, địa lý hay lịch sử. PGS. TS. Hoàng Văn Ma lưu ý rằng: “Trong giới khoa học xã hội và nhân văn, một số người thường đồng nhất, mà không thấy có 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

---

<sup>(\*)</sup> PGS. TS., Viện Thông tin KHXH.

hoá, trong 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. Vì hoàn cảnh lịch sử, do sống xen kẽ, nên có trường hợp ngôn ngữ không trùng hợp với dân tộc. Đó là chưa kể các nhóm đã quên tiếng mẹ đẻ, mà vẫn nhận tộc người gốc, như một số bộ phận dân tộc thiểu số ở Tây Bắc và phía Tây Nghệ An" (4, tr.11).

Là một chuyên gia về các ngôn ngữ ở Việt Nam, GS. TS. Trần Trí Dõi đã chỉ rõ rằng "Một ngôn ngữ, theo cách biểu hiện thông thường *bao giờ cũng bao gồm các phương ngữ*. Đến lượt mình, *các phương ngữ lại có các thổ ngữ khác nhau*" nhưng vì đây là một vùng "có nhiều ranh giới giữa ngôn ngữ và phương ngữ còn chưa được phân chia một cách rạch ròi" (1, tr.46) cho nên trong nhiều trường hợp tên gọi ngôn ngữ/tiếng nào đó chỉ là *mang tính quy ước*. Như thế, "con số 53 dân tộc thiểu số và cùng với dân tộc là cách gọi ước định, 53 ngôn ngữ là con số mang tính quản lý hành chính. Đây là kết quả của một quá trình nghiên cứu công phu của các nhà dân tộc học nước ta từ khi nước nhà được độc lập,... 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" (1, tr.14).

Trong công trình của mình, ông "chấp nhận" rằng ở Việt Nam có mặt các dân tộc thuộc cả 5 ngữ hệ (mà ông 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 – Tạng, Nam đảo (hay Mã lai - Đa đảo) và Mông – Dao (hay Mèo - Dao). Cụ thể như sau:

**I. Ngữ hệ Nam Á** (nhánh Môn – Khơ me), có 25 ngôn ngữ, trong 5 nhóm:

- *Nhóm Khơ me*, có 2 ngôn ngữ: Khơ me, Rơ măm.

- *Nhóm Ba na*, có 11 ngôn ngữ, chia thành 2 tiểu nhóm:

+ *Tiểu nhóm Ba na Bắc* có 6 ngôn ngữ: Ba na, Xơ đăng, Hơ rê, Gié – Triêng, Co, Bơ râu.

+ *Tiểu nhóm Ba na Nam* có 5 ngôn ngữ: Kơ ho, Mơ nông, Xơ tiêng, Mạ, Chơ ro.

- *Nhóm Ka tu* có 3 ngôn ngữ: Bru – Vân Kiều, Cơ tu, Tà ôi.

- *Nhóm Việt Mường* có 4 ngôn ngữ: Việt, Mường, Thổ, Chứt.

- *Nhóm Khơ mú* có 5 ngôn ngữ: Khơ mú, Xinh mun, Kháng, Mảng, Ở đu.

**II. Ngữ hệ Thái – Ka đai**, có 12 ngôn ngữ, trong 4 nhóm:

- *Nhóm Ka đai* có 4 ngôn ngữ: La chí, La ha, Cơ lao, Pu pèo.

- *Nhóm Day Sec* chỉ có 1 ngôn ngữ là Giáy.

- *Nhóm Cao Lan* cũng chỉ có 1 ngôn ngữ là Sán Chay (Cao lan – Sán chỉ)

- *Nhóm Thái - Tày* gồm hai tiểu nhóm:

+ *Tiểu nhóm Thái* có 3 ngôn ngữ: Thái, Lào, Lự

+ *Tiểu nhóm Tày* cũng có 3 ngôn ngữ: Tày, Nùng, Bố y.

**III. Ngữ hệ Hán – Tạng**, có 9 ngôn ngữ, trong 2 nhóm:

- *Nhóm Hán* có 3 ngôn ngữ: Hoa, Sán diu, Ngái.

- *Nhóm Tạng* (cùng thuộc tiểu nhóm Lô lô) có 6 ngôn ngữ: Hà nhì, Phù lá, La hủ, Lô lô, Cống, Si la.

**IV. Ngữ hệ Nam Đảo** (cùng thuộc nhóm Chàm), có 5 ngôn ngữ: Gia rai, Ê

đê, Chăm, Ra glai, Chu ru.

**V. Ngữ hệ Mông - Dao**, có 3 ngôn ngữ, trong 2 nhóm :

- *Nhóm Mông* có 2 ngôn ngữ: Mông, Pà Thén.

- *Nhóm Dao* có tiếng Dao.

Những kiến thức cơ bản về 53 ngôn ngữ dân tộc thiểu số được tác giả trình bày trong công trình này theo danh sách xếp theo thứ tự dân số (1, tr.49-118).

Tuy nhiên, vấn đề đặt ra cần có lời giải là xác định số lượng và danh sách các ngôn ngữ hiện đang tồn tại ở Việt Nam. Và giới ngôn ngữ học đã đi theo hướng nghiên cứu này. GS. TSKH. Lý Toàn Thắng và GS. TS. Nguyễn Văn Lợi đã 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” (xem: 6).

Năm 2008, vào dịp tổng kết 40 năm nghiên cứu của Viện Ngôn ngữ học, PGS. TS. Tạ Văn Thông cho rằng “Số lượng các ngôn ngữ ở Việt Nam chưa được chính thức xác nhận, nhưng chắc chắn nhiều hơn số các dân tộc, có khoảng trên dưới 90 ngôn ngữ” (7, tr.43).

Kết quả nghiên cứu gần đây hơn ở Viện Ngôn ngữ học cho thấy: “Việt Nam có 54 dân tộc, nhưng có đến gần 100 ngôn ngữ/phương ngữ dân tộc thiểu số hành chức” (12, tr.23), mỗi dân tộc và nhóm dân tộc lại có thể có nhiều tên gọi (cũ và mới). Điều này được PGS. TS. Vũ Thị Thanh Hương xác nhận lại ở bài viết *Những vấn đề của ngôn ngữ học ứng dụng năm 2010* rằng: “Theo các tài liệu chính thức, Việt Nam có 54 dân tộc cùng chung sống và sử dụng khoảng

gần 100 ngôn ngữ khác nhau (xem: 9, tr.19).

Trước tình hình như vậy, câu hỏi cần trả lời chính xác là có bao nhiêu ngôn ngữ ở Việt Nam và cụ thể tên các ngôn ngữ được gọi và viết như thế nào - cùng với câu hỏi đó là sự phân định các nhóm địa phương trong một ngôn ngữ - để hiệu chỉnh lại *Tiêu đề đề mục* “Việt Nam - các ngôn ngữ”.

Tuy có thể trả lời là ngành ngôn ngữ học còn đang tiếp tục nghiên cứu, nhưng nếu theo dõi thông tin chuyên ngành, chúng ta có thể biết những kết quả nghiên cứu gần đây, được xem như lời giải đáp dần dần: từ số lượng và danh sách các ngôn ngữ ở Việt Nam đến tên gọi trong các cách phân loại ngôn ngữ: họ/hệ, ngành, nhánh, theo “tiếng” và “ngôn ngữ”...

Chỉ xin dẫn ra ví dụ về một số ngôn ngữ thuộc hệ Nam Á (là một trong 5 hệ ngôn ngữ) có mặt ở Việt Nam.

Hai nhà nghiên cứu Quế Lai và Lê Văn Dương đã từng phân tích để đi đến kết luận rằng “Tiếng *Cọi* và tiếng *Mã liêng* có mối quan hệ thân thuộc; nhưng mối quan hệ này không đủ để gộp chúng thành một ngôn ngữ mà phải coi chúng là hai ngôn ngữ thuộc nhánh *Chứt*” (xem: 3). Và như vậy, “về mặt ngôn ngữ học, đến nay chúng ta có thể kể được những ngôn ngữ thuộc nhánh Chứt ở Việt Nam gồm có: *Mày, Rục, Sách, Arem, Mã liêng, Cọi*” và “Có thể cho ba ngôn ngữ *Mã liêng, Rục, Cọi* là bộ ba ngôn ngữ cổ của nhánh *Chứt*” (3, tr.60).

Hai mươi năm sau, GS. TS. Trần Trí Dõi cũng “lấy trường hợp dân tộc *Chứt* và tiếng *Chứt* là ví dụ (xem thêm: 4). Trong danh mục thành phần các dân tộc công bố năm 1979, tiếng *Chứt* của dân tộc *Chứt* bao gồm những nhóm địa phương khác nhau có tên là *Mày, Rục,*

*Sách, Mã liêng và Arem*. Cả năm nhóm này được coi là thành viên địa phương của tiếng *Chứt*. Có thể thấy quan niệm ấy được chấp nhận trong gần hai thập kỷ qua. Thế nhưng qua những kết quả nghiên cứu ngôn ngữ gần đây, người ta đã có thể chứng minh rằng trong thực tế năm nhóm địa phương mà trước đây được coi là những bộ phận của tiếng *Chứt* ấy có thể là những ngôn ngữ khác biệt nhau, tức là chúng có thể là *những ngôn ngữ riêng lẻ chứ không phải là các phương ngữ* của ngôn ngữ. Cụ thể là, người ta cho rằng cái gọi là tiếng *Chứt* chỉ có thể bao gồm các nhóm *Mày, Rục, Sách*. Hai nhóm còn lại, tiếng *Arem* và tiếng *Mã liêng* là những ngôn ngữ riêng lẻ có bà con họ hàng với tiếng *Chứt* ấy. Đây là cách đặt vấn đề được trình bày trong các nghiên cứu của *M. Ferlus*, của chúng tôi và phần nào trong cả nghiên cứu của GS. TS. Nguyễn Văn Lợi trong thời gian gần đây” (1, tr.15). Với cách nhìn ấy, khi “Giới thiệu tóm tắt một vài ngôn ngữ dân tộc thiểu số ở Việt Nam”, sau tiếng *Chăm* là tiếng *Rục* - được hiểu “là một ngôn ngữ đặc trưng cho tiếng *Chứt*” (1, tr.181) - rồi đến tiếng *Nùng*.

Mười năm sau, cũng về trường hợp này, theo Báo cáo tổng quan kết quả đề tài nghiên cứu cấp Bộ năm 2010: *Bức tranh toàn cảnh về các ngôn ngữ ở Việt Nam* do PGS. TS. Đoàn Văn Phúc làm chủ nhiệm (5, tr.14) có ghi chú: “Theo kết quả của Chương trình *Điều tra các ngôn ngữ ở Việt Nam với vấn đề xác định thành phần dân tộc*, Viện Ngôn ngữ học đã nghiệm thu năm 2006, thì các ngôn ngữ Nam Á ở Việt Nam có thể có tới hơn 40 các thứ tiếng khác nhau. *Tiểu chi Việt – Chứt* có thể có 12 ngôn ngữ (*Việt, Mường, Thổ, Poọng, Cuối, Đan lai - Li hà, Rục, Mày, Sách, Mã liêng, Kha phoọng* (Krih), *Arem*).

Rồi trong công trình nghiên cứu gần đây nhất (xem: 8), ngoài việc cho biết những điều đã có thể khẳng định về loại hình (tr.148) sau đó là những câu hỏi chưa thể kết luận (tr.254), các tác giả đã cung cấp *Danh sách các ngôn ngữ Nam Á ở Việt Nam* (tr.334), gồm 43 ngôn ngữ (với các phương ngữ) thuộc 6 nhóm: *Việt* (Vietic), *Mảng* (Mangic), *Khơ mú* (Khmuic), *Cơ tu* (Katuic), *Ba na* (Bahnaric, gồm 3 tiểu nhóm) và *Khơ me* (Khmeric). Cụ thể như sau:

**I. Nhóm Việt (Vietic)** (13 ngôn ngữ): Kinh, Mường, Nguồn, Poọng, Đan lai, Li hà, Cuối, Rục, Mày, Sách, Mã liêng, Krih (Kha Phoọng), Arem.

**II. Nhóm Mảng (Mangic)** (1 ngôn ngữ): Mảng.

**III. Nhóm Khơ mú (Khmuic)** (4 ngôn ngữ): Khơ mú, Xinh mun (Puộc, Pu Hoóc), Kháng (Kháng Sơn La, Kháng Quảng Lâm), Ó đu (Tày hạt, I đuh).

**IV. Nhóm Cơ tu (Katuic)** (4 ngôn ngữ): Bru, Pa cô, Tà ôi, Cơ tu (Ca tu, Ka tu).

**V. Nhóm Ba na (Bahnaric)** (20 ngôn ngữ).

1. *Tiểu nhóm Ba na Bắc* (14 ngôn ngữ): Ba na (Bahnar), Co (Cua, Kor), Ca dong, Ha lăng, Giẻ (Jeh), Triêng, Bhnoong (Pơ noong), Xơ dăng (Xê dăng, Hteang, Htea, Steang), Tơ đra (Xơ Rá, Mơ đrá), Rơ ngao, Hrê (Đá Vách), Mơ nâm (Pơ noong, Bơ nâm), Ve (Veh, La ve), Pơ măm (Rơ măm, La măm).

2. *Tiểu nhóm Ba na Tây* (1 ngôn ngữ): Brâu.

3. *Tiểu nhóm Ba na Nam* (5 ngôn ngữ): Cơ ho (Cơ ho, Kơ ho), Mnông (Mơ nông), Mạ (Châu mạ), Xtiêng (Stiêng), Chrau (Châu ro, Chro, Ro, Ta mun).

**VI. Nhóm Khơ me** (1 ngôn ngữ): Khơ me Nam bộ.

Như thế, đề mục “Việt Nam - các ngôn ngữ” cần dựa theo kết quả nghiên cứu mới nhất này của Viện Ngôn ngữ học, đặc biệt cần lưu ý rằng *Poọng, Đan lai, Li hà* và *Cuối* được tách riêng thành 4 ngôn ngữ trong *nhóm Việt* (Vietic) và không có cái gọi là ngôn ngữ *Thổ*. Và cũng thuộc nhóm này, bên cạnh *Rục, Mày, Sách, Mã liêng, A rem* có thêm *Krih* (Phoọng) và không có cái gọi là ngôn ngữ *Chứt*. *Pa cô* tách khỏi *Ta ôi* trở thành 2 ngôn ngữ cùng thuộc nhóm *Cơ tu*. Các nhóm *Giẻ, Triêng, Ve* và *Bnoong* (thuộc dân tộc *Giẻ - Triêng*) là 4 ngôn ngữ riêng thuộc *Tiểu nhóm Ba na Bắc*.

Không chỉ các nhà nghiên cứu hay giảng dạy về ngôn ngữ mà những người làm việc có động chạm đến ngôn ngữ - trong đó có những người xử lý sơ bộ tài liệu thư viện - đều mong có lời giải đáp chính thức của giới chuyên môn về một số vấn đề, ví như việc xác định số lượng và danh sách các ngôn ngữ hiện đang tồn tại ở Việt Nam, để hiệu chỉnh *Tiêu đề đề mục* “Việt Nam - các ngôn ngữ” mà chúng tôi đã có dịp nói tới.

Kết quả nghiên cứu thể hiện quá trình đi đến chân lý. Và chúng ta cũng phải tạm bằng lòng với kết quả nghiên cứu của từng giai đoạn lịch sử. Hơn nữa, vấn đề này không thể tách rời hoàn toàn khỏi vấn đề dân tộc (như ở Hội thảo ngôn ngữ học nọ, có nhà nghiên cứu nọ đã lầm tưởng đây là hai chuyện hoàn toàn khác nhau) nên các ý kiến và giải pháp luôn cần hết sức tế nhị.

Nhu cầu hiện nay của người dùng tin là cần có thông tin cập nhật về kết quả nghiên cứu chuyên sâu, khẳng định thành tựu khoa học của từng thời kỳ. Mong rằng các họ ngôn ngữ khác ở Việt Nam cũng sẽ được nghiên cứu theo hướng này.

Nhắc lại những thay đổi trong quá

trình nghiên cứu, 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 còn muốn được xem như những thông tin gợi ý, nên thận trọng khi đọc cũng như xử lý tài liệu về các ngôn ngữ và các dân tộc ở Việt Nam, mà chúng tôi đã có dịp bàn tới tại Hội thảo “Thống nhất công việc Định chủ đề & Biên soạn Khung Tiêu đề đề mục”, do Liên Chi hội Thư viện Đại học phía Nam phối hợp và tổ chức tại Trung tâm Thông tin - Học liệu, Đại học Đà Nẵng, từ 14-16/5/2008. Nay với những thành tựu mới của nghiên cứu ngôn ngữ học, cần có những thay đổi.

#### TÀI LIỆU THAM KHẢO

1. Trần Trí Dõi. Nghiên cứu ngôn ngữ các dân tộc thiểu số ở Việt Nam. H.: Đại học Quốc gia Hà Nội, 301 tr., 2000 (tái bản).
2. Nguyễn Văn Khang. Sự tác động của xã hội đối với ngôn ngữ và những vấn đề đặt ra đối với chính sách ngôn ngữ ở Việt Nam hiện nay. Tạp chí *Ngôn ngữ*, số 8, 2010.
3. Quế Lai, Lê Văn Dương. Về tiếng Cọi và tiếng Mã liêng – hai ngôn ngữ Chứt ở miền Tây Nghệ Tĩnh. Tạp chí *Dân tộc học*, số 4, 1979.
4. Hoàng Văn Ma. Ngôn ngữ dân tộc thiểu số Việt Nam: Một số vấn đề về quan hệ cội nguồn và loại hình học. H.: Khoa học xã hội, 2002, 341 tr.
5. Đoàn Văn Phúc (chủ nhiệm đề tài). Bức tranh toàn cảnh về các ngôn ngữ ở Việt Nam. Kết quả đề tài nghiên cứu cấp Bộ, 403 tr., 2010.

(Xem tiếp trang 57)