

Mark J. Alves | orcid: 0000-0001-7055-918
Montgomery College, Rockville, Maryland, USA
mark.alves@montgomerycollege.edu

Cộng Đồng Ngôn Ngữ Đông Sơn:

Bằng Chứng Về

Ngữ Chi Tiếng Việt (Vietic)

Ngô Bắc dịch

Đại Ý:

Bài viết này xem xét nhiều dòng dữ liệu nhằm cố gắng xác định tình trạng chủng tộc ngữ học vùng châu thổ sông Hồng ở miền Bắc Việt Nam thời Đông Sơn (khoảng 600 TCN [Trước Công Nguyên]–200 CN [Sau Công Nguyên]) trước khi có sự thành lập chính quyền Trung Quốc ở đó vào khoảng năm 200 TCN. Nhiều tình huống có thể xảy ra được xem xét dưới góc độ các dữ liệu văn bản ngôn ngữ học chủng tộc học, khảo cổ học, khảo cổ học và lịch sử. Một số kịch bản phải bị loại trừ vì chúng thiếu bằng chứng hỗ trợ, trong khi số ít còn lại được cân nhắc và xếp hạng với nhau. Tại thời điểm này, kịch bản có nhiều hỗ trợ nhất, bao gồm chủ yếu là dữ liệu ngôn ngữ học lịch sử và khảo cổ học, đó là một cộng đồng người nói tiếng Nam Á cư trú ở đồng bằng sông Hồng từ khoảng 4000 BP, nhưng vào thời điểm các nhóm người Trung Hoa xuất hiện, ngữ chi Việt (Vietic) (giai đoạn sau của nhóm Nam Á nguyên thủy ở đó) và các nhóm người Tày (Tai) sơ khai đã hiện diện ở vùng đất đó. Hơn nữa, bằng chứng ngôn ngữ học đối chiếu hỗ trợ mạnh mẽ nhất cho sự hiện diện của ngữ chi Việt (Vietic) thống trị ở khu vực đó vào thời điểm đó, phần của ngữ chi Việt (Vietic) đó cuối cùng tách ra để trở thành phân nhánh Việt-Mường và cuối cùng, trong đó, là Tiếng Việt.

Từ khóa

Miền Bắc Việt Nam – ngôn ngữ học lịch sử – khảo cổ học – lịch sử dân tộc học

Những Ý Tưởng Mở Đầu ¹

Các vấn đề về căn cước dân tộc trong thời kỳ tiền Tần ở miền nam Trung Quốc và Đông Nam Á Đại Lục giáp biên giới (Mainland Southeast Asia, sau đây gọi tắt là MSEA) từ lâu đã gặp khó khăn để trả lời. Các văn bản Trung Quốc cổ đại cung cấp rất ít chi tiết về các nhóm dân tộc tại miền nam Trung Quốc ngày nay, và những mô tả như vậy đã khái quát hóa quá mức với những thuật ngữ như *Yue* 越 *Việt* và *yi* 夷, *di*, trong khi đưa ra một số chi tiết để làm rõ bản sắc dân tộc của các nhóm dân khác nhau trong khu vực đó. ² Thách thức trong việc xác định căn cước văn hóa của các nhóm cổ xưa cũng một phần là do dân tộc không phải là một yếu tố hữu hình được sử dụng trong việc nghiên cứu tiền sử.

Ngược lại, một số khía cạnh của ngôn ngữ con người có khá nhiều các tài sản cụ thể hơn có thể truy tìm được trong quá khứ. Các yếu tố của ngôn ngữ chắc chắn thay đổi theo thời gian và có thể được trao đổi và biến hóa giữa các cộng đồng ngôn ngữ, ³ đôi khi theo những cách nhất quán nhưng cũng có những cách thức không nhất quán hoặc những phương cách che đậy các giai đoạn xưa hơn của các ngôn ngữ. Tuy nhiên, các ngôn ngữ có một số lượng lớn về tính liên tục lịch sử trong các từ ngữ, các lĩnh vực/hệ thống ngữ nghĩa (semantic), các âm thanh của lời nói và các khía cạnh cấu trúc ngôn ngữ kín đáo khác (ví dụ: cú pháp, hình thái học, các hệ thống đại từ (pronoun), thuật ngữ thân nhân trong họ hàng, hệ thống chữ số, v.v.), cũng như các dân số sử dụng những ngôn ngữ đó. Cả các ngôn ngữ lẫn các cộng đồng ngôn ngữ liên quan của chúng, chính vì thế, có thể truy nguyên về quá khứ thông qua việc tái cấu trúc ngôn ngữ. Và đối chiếu, cùng sự nghiên cứu ngôn ngữ học lịch sử, . Có thể là do kết quả của việc này, hơn là việc sử dụng sắc tộc như một công cụ để theo dõi các nền văn hóa khảo cổ, nhiều nghiên cứu khảo cổ học trong những thập kỷ gần đây đã sử dụng các ngữ hệ (language families) ⁴ như một điểm tham chiếu để kết nối dân cư hiện đại với dân cư quá khứ cổ xưa. ⁵ Ví dụ, trong các nghiên cứu khảo cổ học ở Thái Bình Dương, ngữ hệ Nam Đảo (Austronesian family) đã đóng một vai trò thiết yếu trong việc tìm hiểu sự chuyển động và sự tiếp xúc của người dân trong khu vực đó. ⁶

Khu vực được xem xét trong bài viết này là miền Bắc Việt Nam, tập trung ở Đồng bằng sông Hồng (sau đây gọi là ĐBSH) từ giữa thiên niên kỷ thứ nhất trước Công Nguyên (CN) đến vài thế kỷ đầu CN, thời kỳ văn hóa Đông Sơn ở vùng đó và thời kỳ mà người Trung Quốc đề cập đến các nhóm dân tộc. Địa lý ngôn ngữ hiện đại có thể được mô tả như sau. Tiếng Việt là một phần thuộc ngữ chi Việt (Vietic) ⁷ của ngữ hệ Nam Á, ⁸ là ngôn ngữ chiếm ưu thế của vùng ĐBSH, trong khi các ngôn ngữ Tây Trung Tâm (Tai-Kadai) ⁹ được nói ở địa lý ngoại vi phía bắc và phía Tây và ở Lào giáp biên giới. Những túi nhỏ của ngôn ngữ Mông-Miền (Hmong-Mien) và Hán-Tạng không phải tiếng Hoa (non-Chinese Sino-Tibetan languages) cũng nằm rải rác tương tự dọc theo rìa phía bắc của Việt Nam giáp ranh với Trung Quốc và Lào, một địa lý tiêu biểu cho sự di chuyển của người dân vào khu vực trong những thời kỳ sau này hơn so với các nhóm ngôn ngữ khác. Đối với các ngôn ngữ Nam Á khác, ngoài tiếng Việt, ở vùng này còn có các ngôn ngữ Việt khác cũng như các nhánh khác của Nam Á. Trong nhánh Việt, các biến thể

của tiếng Mường, là những ngôn ngữ có liên quan chặt chẽ nhất với tiếng Việt, chủ yếu tập trung vào ngoại vi phía tây và tây nam của ĐBSH, trong khi các ngôn ngữ Việt Nam Phương Nam (Southern Vietic languages), được nói bởi những nhóm còn nhỏ hơn, được nói xa hơn về phía nam ở các tỉnh Thanh Hóa, Nghệ An, và Quảng Bình và các vùng giáp ranh với Lào.

Cứu xét tình hình hiện đại này, các câu hỏi chính của nghiên cứu này là “Cộng đồng ngôn ngữ chiếm ưu thế ở vùng ĐBSH thời văn hóa Đông Sơn là gì?” và “Đâu là các nguồn dữ liệu khác nhau có thể phát hiện về tình hình ngôn ngữ trong khu vực vào khoảng thời gian đó?” Để cố gắng trả lời những câu hỏi này trong bài viết này, một sự kết hợp của dữ liệu khảo cổ học, lịch sử dân tộc và ngôn ngữ học được cung cấp và sử dụng để đánh giá các giả thuyết khác nhau. Câu trả lời ngắn gọn là rằng kịch bản có khả năng xảy ra nhất là trong các phương ngữ Việt (Vietic) sơ khai — nhóm ngôn ngữ tổ tiên của tiếng Việt và các ngôn ngữ liên hệ — cấu thành cộng đồng ngôn ngữ chiếm ưu thế, nhưng các ngôn ngữ nào thuộc nhánh tiếng Tày của Tày-Trung Tâm Kadai (Tai-Kadai) đã có mặt và tiếp xúc với ngữ chi Việt (Vietic). Tuy nhiên, các kịch bản khác không thể bị loại trừ hoàn toàn.

Bài viết có cấu trúc như sau: Nhiều giả thuyết trước đây được đưa ra trước đây, và kịch bản khả dĩ nhất đã được xác định. Sau đó, nhiều nguồn dữ liệu khác nhau được cung cấp: thứ nhất, các dữ liệu khảo cổ học (cả về văn hóa vật chất lẫn khảo cổ di truyền học (archaeogenetics)), và thứ nhì, các dữ liệu ngôn ngữ học (dữ liệu từ vựng đối chiếu, dữ liệu lịch sử âm vị học và dữ liệu ngôn ngữ dân tộc). Cuối cùng, các dữ liệu này được cứu xét dựa trên các kịch bản chính yếu.

Những giả thuyết trước đây về Thành Phần Ngôn Ngữ Dân Tộc của Văn Hóa Đông Sơn

Các giả thuyết trước đây về cấu trúc ngôn ngữ dân tộc của các dân tộc tại ĐBSH trong thời kỳ tiền nhà Tần (trước năm 221 TCN) thường được coi là có liên quan đến danh xưng của dân tộc và thể chế chính trị trong các văn bản cổ của Trung Quốc. Trong các văn bản Trung Hoa ban sơ từ thời Chiến Quốc (475 TCN–221 TCN) và các thế kỷ kế tiếp, từ đại loại (hay bao quát: hypernym) “Yue” 越/粵 (tiếng Trung cổ (Old Chinese): *[G]wat, tiếng Trung thời trung cổ (Middle Chinese) *hɟwot*) được sử dụng một cách bừa bãi để chỉ hầu như tất cả các nhóm thuộc phần mở rộng phía nam của nhà Hán (206 TCN–220 CN). Tại khu vực Quảng Đông, Quảng Tây và miền Bắc Việt Nam giáp ranh hiện nay, tài liệu tham khảo lịch sử được quy chiếu đến vương quốc Nam Việt (Nanyue 南越), có thủ đô là Phiên Ngung (Panyu 番禺) tại Quảng Đông ngày nay, trong khi người Lạc Việt (Luoyue” 雒越) (tiếng Việt: Lạc Việt) có Cổ Loa, một địa điểm khảo cổ gần Hà Nội ngày nay, là thủ đô của họ.¹⁰ Hai thuật ngữ lịch sử Âu Việt: (Ouyue: 甌越) và Lạc Việt đôi khi được coi là gắn liền lẫn lộn với các dân tộc thuộc nhánh tiếng Tày (Tai) và tiếng Việt (Vietic) sơ khai. Các sự trình bày lịch sử bằng tiếng Việt về Việt Nam, sự tham chiếu

đôi khi có nhắc đến người Lạc Việt, và đôi khi tới vương quốc Văn Lang 文郎.¹¹ Đã có nhiều nỗ lực được thực hiện để kết nối những thuật ngữ sau này với các từ bản địa hoặc từ dân tộc giữa các nhóm Nam Á (Austroasiatic),¹² nhưng không có cách nào để chứng minh hay bác bỏ những giả thuyết như thế, do đó làm suy yếu vai trò của chúng trong việc tái thiết lịch sử dân tộc. Các danh xưng dân tộc sẽ được thảo luận lại trong phần dữ liệu ngôn ngữ kết hợp với các bộ dữ liệu khác.

Trong những thập kỷ trước, một số giả thuyết đã tập trung vào thời kỳ ban sơ của khu vực này và những câu hỏi về vai trò của các nhóm ngữ chi tiếng Việt (Vietic) và tiếng Tày ở đó.

- Tác giả Stephen O'Harrow đề xuất rằng lãnh địa của các dân tộc vùng ĐBSH mở rộng về phía bắc vào khu vực Quảng Đông và Quảng Tây và rằng họ thực sự là tổ tiên của người Việt, điều mà ông gọi là “Nguyên Mẫu Việt: Proto-Việt”; rằng người dân là các người nói tiếng Nam Á (Austroasiatic); và rằng họ đã có một xã hội phát triển và tinh tế.¹³ Để hỗ trợ những tuyên bố này, ông đã duyệt xét các nghiên cứu khảo cổ học về các dụng cụ bằng đồng và cách thức canh tác khác nhau gần đây so với thời điểm xuất bản của ông và đã cung cấp các sự lượng định lịch sử dân tộc dựa trên thông tin từ các văn bản lịch sử.

- Tác giả Michel Ferlus đã gợi ý rằng quê hương của nhánh Vietic (dùng thuật ngữ “Việt-Mường”) nằm ở phần phía bắc của cao nguyên Khorat thuộc Miền Bắc Lào, và sau đó các nhóm này về sau lan tỏa ra phía bắc từ đó.¹⁴

- Trong khảo luận sâu rộng về lịch sử ngôn ngữ Việt Nam, tác giả Trần Trí Dõi đã khẳng định rằng lãnh thổ của nhánh tiếng Việt (Vietic) (mà ông gọi là tiền Việt-Mường) đã mở rộng lên phía bắc từ đèo Ngang, bao gồm các sông Hồng, sông Mã và sông Cả của vùng văn hóa khảo cổ Đông Sơn và vùng của sắc dân Lạc Việt lịch sử.¹⁵ Ngoài những vấn đề rộng lớn hơn trong cuốn sách về địa lý và lịch sử ngôn ngữ của nhánh tiếng Việt (Vietic), đối với giả thuyết này, các điểm tham chiếu của ông là các sự tham chiếu các tài liệu lịch sử về địa lý của vương quốc Văn Lang và về dân tộc Âu và dân tộc Lạc, cũng như về vị trí của văn hóa khảo cổ Đông Sơn.¹⁶

- Tác giả Keith Taylor đưa ra quan điểm rằng người dân vùng châu thổ ĐBSH đã nói tiếng Việt-Mường Nguyên Mẫu (Proto-Việt-Mường), nhưng ông cũng đưa ra viễn cảnh rằng đó là một ngôn ngữ liên quan đến tiếng Khmu hoặc rằng người Âu [Việt] chinh phục có khả năng là người nói tiếng Tày Trung-Tâm Kadai [Tai-Kadai].¹⁷ Ông cũng gợi ý rằng người nói tiếng Môn-Khmer (tức thuộc ngữ hệ Nam Á (Austroasiatic))¹⁸ di chuyển lên phía bắc ở các tỉnh Nghệ An và Hà Tĩnh, với một nguồn gốc ám chỉ về phía tây xuyên qua đèo Mụ Già.

- Trong luận văn về lịch sử ngôn ngữ Hán Việt, trong đó ông đưa ra giả thuyết về một ngôn ngữ “Hán-An Nam” trong khu vực vào thiên niên kỷ thứ nhất theo CN, tác giả John Phan trình bày một kịch bản trong đó tiếng Việt-Mường nguyên thủy (Proto-Việt--Mường) (mà ông sử dụng

hơn thuật ngữ “ngữ chi Việt: Vietic”) là một phần của văn hóa Đông Sơn, lan truyền khắp vùng bao gồm sông Hồng, sông Mã, sông Cả.¹⁹

– Tác giả James Chamberlain đã đưa ra giả thuyết rằng các dân tộc ở vùng ĐBSH và văn hóa Đông Sơn là những người nói tiếng Tày (Tai) trong thiên niên kỷ đầu tiên theo CE,²⁰ dựa trên một số đường hướng hỗ trợ. ²¹ Chamberlain lưu ý rằng Ngô Quyền 吳權 (897–944), nhân vật lịch sử lãnh đạo quân sĩ đánh bại chính quyền Trung Quốc và thành lập một vương quốc mới, được cho là đã đến từ phía nam ĐBSH, nơi mà tác giả Chamberlain cho là quê hương của ngữ chi Việt (Vietic). ²² Về các từ chỉ chủng tộc, ông lưu ý rằng từ *kew*, được nhận thấy ở nhiều ngôn ngữ Tày, ²³ để chỉ người Việt là một dân tộc. Thuật ngữ này được xem một cách tổng quát là có quan hệ với người *Jiao (Giao)* của Jiaozhi 交趾 (tiếng Việt là Giao Chi), và do đó điều này cho thấy một sự hiện diện lâu dài của tổ tiên người Việt hiện đại ở vùng ĐBSH. Tuy nhiên, ông cho rằng thuật ngữ này được sử dụng để đề cập đến các nhóm dân Tày (Tai) trong một bài thơ Tày cổ và trong một ngôn ngữ Hệ Trung Tâm (Kradai) và sau đó suy đoán rằng điều này cho thấy một cách sử dụng xưa hơn. ²⁴ Một điểm khác mà ông nêu ra là sự đa dạng về ngôn ngữ ở phía nam vùng ĐBSH. Khu vực của tiếng Mường và khoảng chục ngôn ngữ của ngữ chi Việt (Vietic) nằm ở phía nam và tây nam của vùng ĐBSH. Hơn nữa, tiếng Việt vùng Bắc Trung Bộ Việt Nam có tình trạng đa dạng về phương ngữ. ²⁵ Tác giả Chamberlain do đó gợi ý rằng vùng có sự đa dạng ngôn ngữ nhất chính là quê hương của ngữ chi Việt (Vietic). Điểm thứ tư liên quan đến những từ tiếng Hán vay mượn thời ban sơ. Ông đưa ra giả thuyết rằng, thay vì vay mượn những từ trực tiếp từ tiếng Hán ở vùng ĐBSH, những từ tiếng Hán được vay mượn một cách gián tiếp từ tiếng Tày vào thiên niên kỷ thứ nhất theo CN, ²⁶ mặc dù như được trình bày trong đoạn lượng giá các tình huống tiếp xúc ngôn ngữ, đây hầu như là một quan điểm không thể đứng vững được.

Tóm lại, các tác giả O’Harrow, Taylor, Ferlus, Trần và Phan đều cho rằng tổ tiên của tiếng Việt (tức là ngữ chi Việt (Vietic) của ngữ hệ Nam Á (Austroasiatic) là một phần chủ yếu của văn hóa Đông Sơn. Các tác giả Chamberlain, Ferlus và Trần cho rằng quê hương của ngữ chi Việt (Vietic) nằm ở phía nam, đâu đó trong khu vực chung của Lào và các vùng giáp ranh của Bắc và Trung Bộ Việt Nam. Cuối cùng, tác giả Chamberlain đưa ra sự tuyên xác rằng văn hóa Đông Sơn có nguồn gốc từ người dân tiếng Tày. Ở các phân đoạn tiếp theo, những quan điểm khác biệt này sẽ được lượng giá.

Các kịch bản ngôn ngữ dân tộc có thể xảy ra

Trước tiên, chúng ta xem xét các kịch bản có thể xảy ra về tình trạng ngôn ngữ ở vùng ĐBSH trong vài thế kỷ sau cùng của thiên niên kỷ thứ nhất trước Công Nguyên, như được liệt kê trong Bảng 1. Trong khi danh sách có thể chứa đựng nhiều tình huống suy đoán khác, những tình huống trong bảng cấu thành các khởi điểm khả dĩ.

BẢNG 1 Các kịch bản khả dĩ về sự cấu thành ngôn ngữ của văn hóa Đông Sơn

1. Các ngôn ngữ sớm hơn không thuộc tiếng Việt và tiếng Tày (ví dụ: các ngôn ngữ Australo-Melanesian [của dân ở quần đảo Melanesia, đông bắc Úc Đại Lợi, Ngô Bắc], các ngôn ngữ Nam Á (Austroasiatic) khác ngoài ngữ chi Việt (Vietic), Tạng-Miến, Hmong-Mien [Mèo và Dao}, v.v.)
2. Chủ yếu hoặc chuyên độc ngữ chi Việt (Vietic)
3. Chủ yếu hoặc chuyên độc tiếng Tày (Tai)
4. Chủ yếu là ngữ chi Việt (Vietic) với ít sự hiện diện của tiếng Tày..
5. Chủ yếu là tiếng Tày với ít sự hiện diện của ngữ chi Việt (Vietic).
6. Sự kết hợp ngang bằng giữa tiếng Tày (Tai) và ngữ chi Việt (Vietic).

Kịch bản 1 đặt ra một vấn đề quan trọng: các ngôn ngữ khác hẳn đã được sử dụng trong khu vực rất lâu trước kỷ nguyên được xem xét. Con người sớm nhất ở Đông Nam Á Đại Lục (MSEA) và Đông Nam Á Hải Đảo (insular) là những người thuộc các nhóm Australo-Melanesian, có lẽ là những người của văn hóa Hòa Bình (xem sự thảo luận trong đoạn về khảo cổ học). Những nhóm dân đó sống trong khu vực, nhưng ngôn ngữ của họ đã bị thay thế bằng các ngôn ngữ khác trên khắp Đông Nam Á Đại Lục (MSEA). Một số di tích có thể nhìn thấy của các nhóm sớm hơn bao gồm các nhóm “Negrito” ở Mã Lai (Malaysia), nhưng bất chấp ngoại hình thân thể khác biệt của họ, giờ đây họ nói các ngôn ngữ nhánh Aslian của ngữ hệ Nam Á (Austroasiatic) (ví dụ: ngôn ngữ Semang). Ở Philippines, những nhóm như thế nói tiếng thuộc ngữ hệ Nam Á (Austroasiatic).

27

Một khả tính khác là các nhóm ngôn ngữ Nam Á khác ngoài ngữ chi Việt (Vietic) đã có mặt. Như đã lưu ý, tác giả Taylor đã nêu lên khả tính của một ngôn ngữ liên quan đến tiếng Khmu (ngữ chi Khmuic của ngữ hệ Nam Á (Austroasiatic)), được nói ở phía tây của ngữ chi Việt (Vietic), phần lớn ở Lào.²⁸ Mặc dù không phải là không thể, nhưng không có bằng chứng xác thực, chẳng hạn như số lượng lớn các từ vay mượn từ tiếng Hán ban sơ trong tiếng Khmu. Các nhóm ngôn ngữ khác phần lớn nằm ở ngoại vi địa lý (ví dụ như tiếng Chăm về phía nam) và/hoặc là những ngôn ngữ tương đối mới đến khu vực này (ví dụ: ngôn ngữ Hmong-Mien và Tạng-Miến chủ yếu ở phía tây bắc).

Đặt những điều này sang một bên, tất cả đều thiếu bằng chứng hỗ trợ của một sự hiện diện tại ĐBSH vào giữa thiên niên kỷ thứ nhất Trước Công Nguyên (TCN), chúng ta chỉ còn lại tiếng Tày (Tai) và ngữ chi Việt (Vietic) là những ứng viên nguyên sơ. Câu hỏi sau đó sẽ trở thành liệu tiếng này, ngôn ngữ kia hay cả hai cộng đồng ngôn ngữ đều có mặt ở đó, và nếu cả hai đều ở trong khu vực, chúng đã đóng vai trò như thế nào trong xã hội lúc bấy giờ. Các kịch bản 2 và 3, chủ yếu hoặc độc quyền của ngôn ngữ này hay ngôn ngữ kia, đều không thể hỗ trợ được, như sẽ được phô bày ở các đoạn tiếp theo. Sau đó chúng ta phải tập trung vào các kịch bản 4, 5 và 6, trong mà cả hai cộng đồng ngôn ngữ có thể đã có mặt ở ĐBSH trong thời kỳ Đông Sơn.

Như đã lưu ý, các bộ dữ liệu khác nhau cuối cùng gợi ý rằng tình huống khả dĩ nhất là ngữ chi Việt (Vietic) là cộng đồng ngôn luận chính yếu, khiến kịch bản số 4 được hỗ trợ tốt nhất. Tuy nhiên, kịch bản 5 và đặc biệt là kịch bản 6 vẫn chưa thể bị loại trừ hoàn toàn, mặc dù bằng chứng về chúng yếu hơn nhiều.

Trong các đoạn tiếp theo, các dữ liệu khảo cổ học, ngôn ngữ học và lịch sử dân tộc được trình bày, được tiếp nối bởi việc duyệt xét các bộ dữ liệu này liên quan đến các kịch bản 4, 5 và 6.

Khảo cổ học: Văn Hóa Vật Chất và Khảo Cổ Di Truyền Học

Các nghiên cứu khảo cổ học cho thấy, trên khắp MSE A, những người săn bắn - hái lượm của văn hóa Hòa Bình đã cư trú hàng chục nghìn năm trước khi có sự bành trướng của các nông gia Thời Kỳ Đồ Đá Mới (hay Tân Thạch: Neolithic). Câu hỏi đặt ra là ai là những di dân canh nông thời kỳ đồ đá mới, và một giả thuyết khả dĩ là họ là các người nói tiếng Nam Á (Austroasiatic).

Nguồn gốc và sự phân bố của ngôn ngữ Nam Á đã được tranh luận qua cả một thế kỷ.²⁹ Các giả thuyết cơ bản hiện nay về quê hương nguyên thủy bao gồm (a) miền nam Trung Quốc³⁰ hoặc gần biên giới Assam-Miến Điện,³¹ (b) miền trung Đông Nam Á Đại Lục (MSEA),³² và (c) Bangladesh.³³ Câu trả lời cuối cùng cho câu hỏi về quê hương ngữ hệ Nam Á nằm ngoài phạm vi của bài viết này. Tuy nhiên, với tư cách là một giả thuyết đang được cứu xét, nó được suy đoán là vùng Đông Nam Á Đại Lục (MSEA) giáp ranh với miền nam Trung Quốc hoặc chính Đông Nam Á Đại Lục (MSEA). Sở dĩ có quan điểm này là do các nghiên cứu khảo cổ về văn hóa vật thể và khảo cổ di truyền học, như được thảo luận dưới đây.

Giả thuyết “Hai tầng lớp” tại Đông Nam Á Đại Lục (MSEA)³⁴ đề cập đến sự pha trộn của hai nhóm di truyền: những người săn bắt hái lượm Australo-Melanesian sớm hơn và các nông gia từ miền nam Trung Quốc thời đồ đá mới sau này.³⁵ Giả thuyết này được hỗ trợ bởi nhiều nghiên cứu cho thấy (a) các nhóm săn bắt hái lượm ở ĐNA Đại Lục (MSEA) dẫn tới việc mở rộng nông nghiệp và (b) sự xuất hiện của việc sản xuất ngũ cốc và kỹ thuật nông nghiệp ở miền đông Trung Quốc xung quanh khu vực sông Dương Tử,³⁶ cũng xảy ra trước khi có sự bành trướng nông nghiệp tại vùng ĐNA Đại Lục (MSEA). Có ý kiến cho rằng các nông gia Thời Đồ Đá Mới xuất phát từ miền nam Trung Quốc vào khoảng năm 2000 Trước Công Nguyên (TCN) (hoặc vài thế kỷ cuối của thiên niên kỷ thứ ba Trước CN) với “kiện hàng nông nghiệp Thời Tân Thạch” (ví dụ: gạo, hạt kê, lợn, chó, đồ gốm được chạm khắc, chôn cất nằm ngửa (subine burials), v.v.) là ~bg kẻ nói tiếng thuộc ngữ hệ Nam Á.³⁷ Xem xét sự hiện diện rộng rãi của ngôn ngữ Nam Á trên khắp ĐNA Đại Lục (MSEA) (không bao gồm lãnh thổ sau này bị chiếm cứ bởi việc mở rộng của các nhóm tiếng Tày (Tai), nó đã được khẳng định rằng một số nông gia nói tiếng Nam Á này đã du hành dọc theo các tuyến đường thủy của miền trung ĐNA Đại Lục (MSEA) và một số về phía nam dọc theo bờ biển Việt Nam ngày nay.³⁸ Gần đây hơn, các tác

giả Felix Rau và Paul Sidwell đã thêm vào đây giả thuyết cho rằng người nói tiếng Nam Á đã thực hiện hành trình xuyên đại dương, ³⁹ điều này càng có nhiều bằng chứng, và tác giả Sidwell đặc biệt chỉ ra rằng vùng Đồng Bằng Sông Hồng (ĐBSH) là một địa điểm khả dĩ cho sự phân tán. ⁴⁰ Bất kể các chi tiết về lịch sử ngôn ngữ Nam Á, những khẳng định trong bài viết này về các nhóm ở thời Đông Sơn vẫn có thể được duy trì.

Đặc biệt ở vùng ĐBSH, một trình tự văn hóa đã được thể hiện trong tài liệu khảo cổ. Các nghiên cứu về khảo cổ học Việt Nam đã cho thấy một sự phát triển liên tục sự của các phong cách và công nghệ văn hóa vật thể ngay từ đầu văn hóa Phùng Nguyên (2000–1500 TCN) đến văn hóa Đông Đậu (1500–1000 TCN) và văn hóa Gò Mun (khoảng 1100–800 TCN) sau này, và cuối cùng là các văn hóa thời kỳ Đông Sơn (600 TCN–200 CN). ⁴¹ Ngoài ra, các dữ liệu khảo cổ di truyền học gần đây và dữ liệu so sánh hình thái sọ người (morphocranial) và hình thái răng người (morphodontal) ở ĐBSH (RRD) hỗ trợ tương tự cho giả thuyết “Hai tầng lớp: Two-layers, cho thấy các yếu tố của cả các nhóm thuộc ngữ hệ Australo-Melanesian nguyên thủy lẫn các sự xuất hiện tiếp theo của các nhóm thuộc ngữ hệ khác. ⁴² Điều cốt yếu là vật liệu di truyền từ di tích ở di chỉ Mân Bắc từ thời Phùng Nguyên được so sánh với các ngôn ngữ hiện đại và chúng đã được chứng minh là có liên quan đến điều có thể được xem, một cách rộng rãi, đã tạo ra một nguồn nhiễm thể di truyền (gene) vùng Đông Á. ⁴³ Trong khi các kết quả của các cuộc nghiên cứu khảo cổ học nên được coi là giả thuyết tạm thời trong khi chờ thêm dữ liệu và tinh lọc cả về kỹ thuật và công nghệ, những nghiên cứu gần đây của đội ngũ các nhà khảo cổ học nổi tiếng trong lĩnh vực này phù hợp với nhiều loại các bộ dữ liệu (không chỉ về nhiễm thể di truyền (genes) mà còn về cả văn hóa vật thể và kiểu hình bên ngoài (phenotype) phải được xem là nghiêm túc.

Do đó, ở vùng ĐBSH, dữ liệu cho thấy các nhóm thuộc ngữ hệ Nam Á (Austroasiatic) thực hành nông nghiệp đã cư trú ở vùng ĐBSH từ khoảng 2000 năm Trước Công Nguyên và được hấp thụ sớm hơn những người săn bắt hái lượm thuộc ngữ hệ Australo-Melanesian. ⁴⁴ Các chi tiết về sự tiếp xúc văn hóa xã hội và ngôn ngữ xã hội giữa các nhóm Nam Á và các nhóm khác ở phía bắc tại miền nam Trung Quốc trong thiên niên kỷ thứ nhì Trước Công Nguyên thì rất mơ hồ. Tuy nhiên, Thời Đại Đồ Đồng (Bronze Age) bắt đầu ở vùng ĐBSH vào khoảng cuối thiên niên kỷ thứ nhất Trước Công Nguyên. ⁴⁵ Liệu những hoạt động như vậy có di chuyển về phía nam xuống bờ biển phía đông nam Trung Quốc hay từ cao nguyên Vân Nam cũng không chắc chắn, nhưng dù vậy, sự tiếp xúc liên văn hóa đã xảy ra và dẫn tới sự lan tỏa của ngành luyện kim. Thời điểm bắt đầu của Thời Đại Đồ Đồng dọc theo bờ biển phía nam Trung Quốc cho phép khả tính rằng công nghệ như vậy đã được chuyển xuống bờ biển. ⁴⁶ Kịch bản đó hàm ý sự tiếp xúc giữa bất kể cộng đồng ngôn ngữ Nam Á nào ở vùng ĐBSH và các nhóm đương thời ở vùng Lĩnh Nam 嶺南, như Quảng Đông, Quảng Tây và Hải Nam, được mô tả trong các văn bản lịch sử là quê hương của các nhóm Baiyue 百越 và vương quốc Nam Việt. Nhóm người nói tiếng Nam Á đó, sau hơn một nghìn năm tách biệt khỏi các nhóm Nam Á khác, đã có thời gian để trở

nên khác biệt với các cộng đồng ngôn ngữ Nam Á khác, là nguồn gốc tiềm năng của ngữ chi Việt (Vietic).

Thêm bằng chứng về sự liên hệ giữa vùng ĐBSH và Lĩnh Nam một lần nữa đến từ nghiên cứu khảo cổ học. Cuộc nghiên cứu năm 2018 của tác giả Lipson và cộng sự. nghiên cứu ghi nhận rằng các di tích của địa điểm Núi Nấp thời Đông Sơn cung cấp bằng chứng về một sự pha trộn các nguồn nhiễm thể di truyền (gene pools). Sự phát biểu ngắn gọn gợi ý rằng các nhiễm thể di truyền liên quan, gắn liền với “Dự án Núi Nấp thì gần gũi với người Việt và người Đại [chỉ sắc dân Đại nói tiếng Tày, Ngô Bắc] gần ”trung tâm...”.⁴⁷ Một vị trí trọng yếu có ý nghĩa khảo cổ học trong khu vực là địa điểm Cổ Loa, ngay ngoại ô Hà Nội. Khu vực có tường bao quanh này được giả thuyết là một đô thị nguyên mẫu (proto-urban), nơi cư ngụ cấp quốc gia, từ giữa thiên niên kỷ thứ nhất Trước Công Nguyên (TCN), với kiến trúc đồ sộ đòi hỏi một lực lượng lao động lớn và chính quyền tập trung trong một xã hội có thể đã trải qua sự phân tầng xã hội đáng kể.⁴⁸ Một nơi như vậy sẽ là địa điểm thích hợp cho việc trao đổi giữa các sắc tộc, và trong khi không có nghiên cứu khảo cổ học cụ thể nào ở Cổ Loa, các dữ liệu từ địa điểm Núi Nấp gợi ý về sự tiếp xúc. Sẽ có nhiều điều đáng lưu ý hơn về sự tiếp xúc ngôn ngữ giữa tiếng Việt và tiếng Tày ở phần về sự tiếp xúc ngôn ngữ tiếp xúc ngôn ngữ Tày-Việt.

Tóm lại, các nghiên cứu khảo cổ học về văn hóa vật thể, di truyền và hình thái sọ người và răng đều chỉ ra một sự hiện diện của ngữ hệ Nam Á ở vùng ĐBSH khoảng 4000 năm Trước Thời Hiện Đại (BP: Before the Present): không có nghiên cứu nào đặt để các nhóm Tày ở đó vào lúc đó. Tuy nhiên, đến giữa thiên niên kỷ thứ nhất Trước Công Nguyên, người nói tiếng Nam Á ở phần đó của vùng Đông Nam Á Đại Lục đã có sự tiếp xúc văn hóa xã hội với các nhóm người Tày.. Sự kiện rằng các nhóm nói tiếng Tày đã biến mất khỏi khu vực trung tâm này sau đó thì rõ ràng, nhưng lý do tại sao họ biến mất thì không rõ. Một khả năng đơn giản là họ không phải là các nhóm chính ở trung tâm của ĐBSH để khởi lập. Song, điều đó xảy ra khi nào và quá trình diễn ra như thế nào hoàn toàn thiếu cả bằng chứng cho phép suy đoán.

Các dữ liệu khảo cổ học và khảo cổ di truyền học được mô tả ở trên cung cấp thông tin vấn đề về các nhóm ngôn ngữ thời Đông Sơn. Trong các phần tiếp theo, chúng ta xem xét những gì mà các dữ liệu ngôn ngữ và dân tộc học có thể phát hiện.

Ngôn Ngữ Học:

Lịch Sử Ngôn Ngữ Học và Dữ Liệu Dân Tộc Học

Phần này trình bày các dữ liệu ngôn ngữ học có liên quan để cung cấp chung (a) bằng chứng nguồn gốc về sự tiếp xúc ngôn ngữ ban sơ ⁴⁹ thông qua sự hội tụ về mặt hình thái học giữa tiếng Hán (tức là cộng đồng ngôn ngữ nguồn gốc của tất cả các dạng tiếng Trung Hoa hiện đại), tiếng Tày (Tay) và ngữ chi Việt (Vietic); (b) phương tiện đánh giá các từ sớm vay mượn tiềm năng để

tăng độ tin cậy; (c) niên đại giả thuyết của những từ vay mượn ban đầu này dựa trên các đặc tính âm vị học; và (d) bằng chứng cho thấy vai trò hạn chế của tên gọi dân tộc và địa danh trong khu vực. Bằng chứng từ vựng và âm vị học này đặt tiếng Tày (Tai) và ngữ chi Việt (Vietic) ở vùng Lĩnh Nam và ĐBSH thời Đông Sơn, và nó chứng tỏ những tình huống và địa điểm khác nhau của sự tiếp xúc ngôn ngữ Hán-Tày và Hán-Việt ban đầu. Cuối cùng, các dữ liệu ủng hộ khẳng định rằng ngữ chi Việt (Vietic) được nói ở vùng ĐBSH vào thời điểm có sự bành trướng của nhà Hán ở vùng này vào khoảng năm 200 TCN. Bảng 6 cung cấp một tóm lược về các loại dữ liệu và những suy luận có thể được thực hiện về địa lý, niên đại và sự tiếp xúc văn hóa xã hội.

Các dữ liệu ngôn ngữ của các từ vay mượn tiếng Trung ban đầu hầu hết liên quan đến nửa đầu của thiên niên kỷ thứ nhất theo CN và thời kỳ Đông Sơn sau này. Từ dữ liệu này, chúng ta có thể đưa ra các giả định về vị trí của chúng tại thời điểm có tiếp xúc đầu tiên với các nhóm người Trung Quốc và do đó đưa ra giả thuyết về các vị trí gần đúng của các nhóm ngôn ngữ Tày và Việt trong thời kỳ đó.

Khu Vực Hội Tụ Loại Hình Ngôn Ngữ

Bằng chứng chính yếu của sự tiếp xúc ban đầu là lịch sử tiếp xúc ngôn ngữ trong khu vực. Sự bành trướng về phía nam của các dân tộc Trung Quốc vào thời nhà Hán (khoảng 200 TCN đến 200 CN) đã đưa những dân tộc nói tiếng Hán (Sinitic) tiếp xúc với cả cộng đồng nói tiếng Tày và tiếng Việt ở vùng Lĩnh Nam và vùng ĐBSH. Những mô tả lịch sử dân tộc hữu ích về sự tiếp xúc với các nhóm trong khu vực này bao gồm các công trình của tác giả Taylor (1980 và 2013) về miền bắc Việt Nam và của tác giả Churchman (2016) về vùng Lĩnh Nam.

Thông qua tiếp xúc văn hóa xã hội và ngôn ngữ xã hội lâu dài, tiếng Hán (Sinitic), tiếng Tày (Tai) và tiếng Việt (ngữ chi Việt: Vietic) đã phát triển các mô hình cấu trúc ngôn ngữ chung. Ngày nay, những ngôn ngữ này chia sẻ (a) tính đơn âm tiết (monosyllabicity) (tức là các từ gốc gồm một âm tiết), (b) các âm tiết có các cấu trúc đơn giản (ví dụ: các chuỗi nhiều phụ âm không thường có hoặc không tồn tại), (c) các hệ thống âm điệu phức tạp (ví dụ: có nhiều âm) và (d) thiếu hình thái biến tố (inflectional) hoặc phái sinh (derivational) (ví dụ: số nhiều được thể hiện thông qua các từ, không phải bằng hậu tố hay tiếp vĩ ngữ (suffixes)). Ngoài những đặc điểm về âm vị học và hình thái học này, các ngôn ngữ hiện đại ở khu vực này cũng có chung các hệ thống phân loại danh từ (tức là các từ xuất hiện giữa các từ số lượng và danh từ), các trợ từ (particles) cuối câu biểu thị các chức năng ngữ pháp và thực dụng (ví dụ: thì (tense)/khía cạnh, nghi vấn, phương thức, v.v.), và nhiều đặc điểm ngữ nghĩa-thực dụng khác nhau, làm nổi bật tính song ngữ (bilingualism) lâu dài ở khu vực này.

Vào thời điểm tiếp xúc ngôn ngữ ban sơ, tình trạng thể loại (typological) ngôn ngữ đã ngônkhá khác nhau. Hai nghìn năm trước, cả ba ngôn ngữ tổ tiên nguyên thủy đều thiếu âm điệu [hay thanh điệu]⁵⁰ và chúng đã được tái tạo ở những mức độ khác nhau với cấu trúc từ phức tạp

hơn và hình thái/phụ tố (affixes) chắp dính (agglutinative) (mặc dù Tiếng Tày Nguyên Mẫu (Proto-Tai) chưa được tái tạo bằng các phụ tố).⁵¹ Các nhà nghiên cứu đã trình bày cấu trúc từ phức tạp trước đó của tiếng Trung Hoa cổ,⁵² Tiếng Tày Nguyên Mẫu (Proto-Tai),⁵³ và Tiếng Việt (Vietic).⁵⁴ Sự tiếp xúc ngôn ngữ sớm giữa các nhóm ngôn ngữ này, sau đó dẫn đến sự hội tụ về loại hình, đã được các nhà nghiên cứu mô tả.⁵⁵ Nhìn chung, các đặc điểm chung về âm vị học, hình thái và cú pháp hiện đại gợi ý ảnh hưởng ngôn ngữ đa chiều lâu dài. Thay vì khẳng định rằng tiếng Hán (Sinitic) khiến các nhóm ngôn ngữ khác cũng giống mình, điều hợp lý hơn để giả định rằng, do việc tiếp xúc với nhau trong một quá trình kéo dài nhiều thế kỷ, tất cả nhóm đều phát triển những đặc điểm ngôn ngữ tương tự nhau. Một ví dụ hữu ích về điều này là mối quan hệ giữa tiếng Tày (Tai) và nhóm Việt Trung Hoa (Yue Chinese) (bao gồm tiếng Quảng Đông và một số phương ngữ khác được nói ở các tỉnh Quảng Đông và Quảng Tây): hai nhóm có chung một số đặc điểm ngôn ngữ khiến chúng khác biệt với các biến loại của tiếng nhiều Trung Quốc ở phương bắc.⁵⁶ Tuy nhiên, tiếng Hán (Sinitic) là một ngôn ngữ rrao tặng [hay cho vay] mạnh mẽ, để lại dấu ấn từ vựng rất rõ ràng đối với cả tiếng Tày (Tai) và tiếng Việt (Vietic), như đã thảo luận ở phần đánh giá các tình huống tiếp xúc ngôn ngữ.

Chi tiết về việc ấn định thời gian của những thay đổi này rất phức tạp và vượt xa phạm vi của bài nghiên cứu này, và trong bất kỳ trường hợp, chúng không thể chứng minh được các tình huống trong thời Đông Sơn, vì những thay đổi không thể xác định được thời gian một cách chính xác. Tuy nhiên, sự phát triển ban đầu của âm điệu trong tiếng Hán (Sinitic) nhiều phần có vẻ đã xảy ra vào nửa đầu thiên niên kỷ thứ nhất theo Công Nguyên (CN)⁵⁷, một vấn đề được thảo luận trong các phần tiếp theo về các từ vay mượn. Về sự hội tụ về thể loại (typological convergence), chúng ta có thể chỉ suy luận rộng rãi rằng có sự phổ biến song ngữ ở các mức độ khác nhau giữa tiếng Hán (Sinitic), tiếng Tày (Tai) và tiếng Việt (Vietic), và do đó tất cả chúng đều hiện diện ở các vùng Lĩnh Nam và ĐBSH trong thời kỳ được đề cập (với mối quan hệ chặt chẽ hơn giữa tiếng Tày (Tai) và tiếng Việt Trung Hoa (Yue Chinese)).⁵⁸

Lượng Giá Các Tình Huống Tiếp Xúc Ngôn Ngữ Thông Qua Các Từ Hán Vay Mượn Trong tiếng Tày Nguyên Mẫu và Tiếng Việt Nguyên Mẫu (Proto-Tai và Vietic)

Một lần nữa, sự hội tụ về thể loại cho thấy bằng chứng về ảnh hưởng hai chiều của ngôn ngữ. Cuối cùng, dân số nói tiếng Hán (Sinitic) đã trở thành một nhóm văn hóa xã hội thường trực, chiếm ưu thế. Hậu quả, trái ngược với các đặc điểm về thể loại, như cấu trúc của từ và âm tiết, và hệ thống âm điệu, ít nhất một phần là kết quả của sự hội tụ khu vực, việc vay mượn từ vựng chủ yếu là từ tiếng Hán (Sinitic) vào tiếng Tày (Tai) và tiếng Việt (Vietic), ngay từ thời kỳ tiếp xúc sớm nhất vào cuối giai đoạn của thời Đông Sơn. Những từ vay mượn như vậy rất hữu ích trong việc thiết lập sự hiện diện của tiếng Tày và tiếng Việt thời Đông Sơn, đồng thời chúng

cũng cung cấp bằng chứng rõ ràng rằng các trung tâm địa lý của các cộng đồng nói tiếng Tày và tiếng Việt thuộc các khu vực khác biệt.

Đối với cuộc nghiên cứu này, tôi đã sử dụng các bộ sưu tập từ mượn từ tiếng Trung Hoa thời ban sơ ở cả hai ngôn ngữ Tiếng Việt và tiếng Tày Nguyên Mẫu được lựa chọn dựa trên các mẫu âm vị.⁵⁹ Cơ sở dữ liệu các từ mượn tiếng Hán ban sơ trong tiếng Việt có trên 600 từ, ở mức độ chắc chắn trung bình (khoảng 200 mục) cho đến mức độ chắc chắn cao (khoảng 400 mục).⁶⁰ Trong khi các từ tái tạo được coi là nguồn dữ liệu hiệu quả nhất cho những câu hỏi về lịch sử dân tộc, những vấn đề tiềm ẩn trong việc sử dụng từ ngữ từ tiếng Việt hiện đại được giảm thiểu bởi (a) số lượng lớn các mục từ thể hiện các khuôn mẫu thông thường của các sự tương ứng về âm vị với các giai đoạn ban đầu của tiếng Hán, và (b) thực tế rằng những mục từ này được xác định là đã được vay mượn trong khoảng thời gian tương tự với những từ vay mượn tiếng Hán vào thời kỳ đầu của tiếng Tày (Tai). Về những từ mượn tiếng Hán trong tiếng Tày Nguyên Mẫu (Proto-Tai), tôi đã xác định được hơn 160 mục từ có độ chắc chắn từ trung bình đến cao là có nguồn gốc từ tiếng Trung Hoa. Có hàng trăm từ mượn trong các ngôn ngữ Tày khác nhau⁶¹ đã được vay mượn qua nhiều thế kỷ, như trường hợp của tiếng Việt, với hàng nghìn từ và hình vị Hán Việt trong các từ ghép hai âm tiết, nhưng những điều này không thể được xem xét trong nghiên cứu này do sự vay mượn chúng thuộc giai đoạn sau này. Chỉ những từ vay mượn gần thời Đông Sơn mới hữu ích trong bối cảnh hiện nay. Phụ lục B cung cấp một danh sách các tiêu chí và các phương pháp được sử dụng trong việc xác định và loại trừ các từ vay mượn khả dĩ trong các giai đoạn trước đó.

Những Từ Vay Mượn Từ Tiếng Hán Ban Sơ Được Chia Sẻ Chung

Những từ mượn từ tiếng Hán Sinitic là bằng chứng mạnh mẽ về sự tiếp xúc văn hóa xã hội ban đầu, nhưng tuyên bố rằng chúng được mượn từ 1.500 năm trước hoặc lâu hơn, các phương pháp lịch sử ngôn ngữ lịch sử- phải được áp dụng. Để xác định rằng các từ trong hai ngôn ngữ là liên quan, chúng phải được chứng minh là có các mô hình thường lệ của các sự tương ứng về âm vị học; mặt khác, những từ như vậy có thể chỉ là những điểm tương đồng một phần, ngẫu nhiên. Trong một số trường hợp, các mẫu âm vị học của các loại phụ âm, các nguyên âm và âm điệu có thể cung cấp thời biểu tương đối của các khoản vay mượn. Từ vay mượn tiếng Trung Hoa từ thiên niên kỷ thứ nhất CN ở cả tiếng Tày và tiếng Việt Nguyên Mẫu thực sự có chung các mô hình tương ứng âm vị với tiếng Hán cuối thời cổ và Tiếng Hán thời Trung cổ ban sơ (tức là tiếng Hán từ triều đại Tây Hán đến Đông Tấn, khoảng 200 năm trước Công nguyên đến khoảng 300 năm theo CN).

Vài chục từ mượn tiếng Hán này có mặt ở cả tiếng Tày lẫn tiếng Việt, bao gồm các thuật ngữ chỉ kim loại, các loại dụng cụ, ẩm thực và các sản phẩm mậu dịch, và các hoạt động liên quan như trong các ví dụ trong Bảng 2. Về thời điểm của các từ vay mượn, các từ trong tiếng

Tày Nguyên Mẫu (Proto-Tai) tất nhiên thuộc thời kỳ ban sơ của sự tiếp xúc ngôn ngữ. Tác giả Pittayawat Pittayaporn thừa nhận rằng những từ như vậy có niên đại trở lùi về thời ban sơ của thiên niên kỷ thứ nhất trong CN.⁶² Đối với các từ Hán cổ trong tiếng Việt, trong khi một số từ mượn có thể được coi là từ vay mượn của nửa đầu của thiên niên kỷ thứ nhất theo CN (dựa trên các loại âm điệu, như được mô tả bên dưới), không phải tất cả các từ có đặc điểm âm vị tạo thành bằng chứng mạnh mẽ đặt chúng trong thời kỳ sớm của này. Các mô tả kỹ thuật của các sự tương ứng âm vị học trong các dữ liệu nằm ngoài phạm vi của bài viết này. Nhiều nghiên cứu khác nhau được trích dẫn trong bài viết này giải quyết những vấn đề này, và mặc dù có nhiều điểm bất đồng, các ý tưởng tổng quát được trình bày ở đây không gây tranh cãi. Cuối cùng, cột cuối cùng của dữ liệu tiếng Mường từ biến thể Mường Bi⁶³ nêu bật cách thức ít nhất một số từ này đã được vay mượn bởi giai đoạn Việt-Mường Nguyên Mẫu (Proto Việt-Mường), do đó làm tăng độ dài thời gian.⁶⁴

Điều cần lưu ý là tính nhất quán của các loại âm điệu. Trong khi các bản tái tạo của tiếng Trung Hoa cổ đại (Old Chinese) có âm *-s cuối cùng, Tiếng Tày Nguyên Mẫu (Proto-Tai) thường có âm loại B, trong khi ở tiếng Việt, các từ mượn của tiếng Trung Hoa ban sơ (sau đây gọi tắt là ECL: Early Chinese Loans) có hoặc âm *ngã* [tiếng Việt trong nguyên bản, Ngô Bắc] (vay trước đó trước khi mất âm tiếng Hán cổ *-s/h) hoặc có âm độ *ngang* [tiếng Việt trong nguyên bản, Ngô Bắc] (các khoản vay sau này sau khi mất âm tiếng Hán cổ *-s/h nhưng trước khi có sự phát triển âm điệu (tonogenesis) trong tiếng trở thành Việt-Mường). Về âm tắc thanh môn (glottal stop) cuối cùng của tiếng Hán cổ, *-ʔ thường tương ứng với âm C trong Tiếng Tày Nguyên Mẫu (Proto-Tai), và các âm *sắc* và *nặng* trong Tiếng Việt. Tuy nhiên, các âm A và D không mang lại sự khác biệt về trình tự thời gian, do đó các phụ âm và các nguyên âm cũng phải được xem xét để xác định các sự vay mượn ban đầu này trong tiếng Việt. Trong một số trường hợp, độ chính xác về mặt trình từ niên đại không thể thu nhận được, nhưng tuy thế, chúng có chung đặc điểm của các phụ âm và các nguyên âm với ECL, và khác biệt về mặt âm vị với các mục từ Hán-Việt (Sino-Vietnamese, từ giờ về sau viết tắt là SV) phát sinh từ khoảng đầu thiên niên kỷ thứ hai theo CN.

Ngoài các mô hình lịch sử âm vị, thời điểm vay mượn có bản chất văn hóa và những thứ tương ứng với các tài liệu khảo cổ học có thể cung cấp thêm chi tiết.⁶⁵ Chỉ có thể đưa ra một niên đại tương đối dựa trên thông tin âm vị học, nhưng điều này dù sao cũng chỉ ra sự vay mượn theo hướng và vài thế kỷ sau khi kết thúc thời Đông Sơn.

Điều mà dữ liệu này không thể phát lộ là vị trí chính xác của các cộng đồng ngôn ngữ này, số lượng lớn. Tất cả các loại từ có thể được mượn ở bất cứ nơi nào có người nói tiếng Hán cư trú. Đối với các đặc điểm âm vị học, chúng cung cấp cái nhìn tổng quát quán triệt về niên đại, nhưng một lần nữa, không có đặc điểm nào, chẳng hạn như các sự khác biệt phương ngữ vùng miền, có thể chỉ ra địa điểm của sự vay mượn.. Tuy nhiên, dữ liệu có thể cho thấy các vị trí khác nhau, như được thảo luận trong phần lượng giá các tình trạng tiếp xúc ngôn ngữ.

Bảng 2: Mẫu Các Từ Vay Mượn Từ Tiếng Hán thời ban Sơ Trong Tiếng Tày Nguyên Mẫu và Tiếng Việt

[Các chữ viết tắt trong bảng 2: (**Gloss**: Anh ngữ), **Chin.**: Tiếng Hán; **OC**: Old Chinese: Tiếng Hán cổ; **MC**: Middle Chinese: Tiếng Hán thời Trung Cổ; **Proto Tai**: Tiếng Tày Nguyên Mẫu; **ECL**: Early Chinese Loans: Các từ Hán vay mượn ban đầu; **SV**: Sino-Vietnamese: từ Hán-Việt)]

TABLE 2 Sample early Chinese loanwords in Proto-Tai and Vietnamese^a

Gloss	Chin.	OC	MC	Proto-Tai	ECL	SV	Mường
harrow	<i>pá</i> 耙	*[b] ^ʰ ra	bae	*fiue ^A (Li)	bừa	bà	pừa
pincers/tongs	<i>qián</i> 鉗	*C.[g]<r>[e]m	gjem	*gi:m ^A (Pitt.)	kìm	kiềm	(kep)
needle	<i>zhēn</i> 針	*t.[k]əm	tsyim	*qem ^A (Pitt.)	kim	châm	kim
cage; coop	<i>lóng</i> 籠	*k.r ^ʰ oŋ	luwng	*kr ^ʰ oŋ ^A (Li)	lồng	lung	lồng
bag (noun)	<i>dài</i> 袋	*Cə.l ^ʰ ək-s	dojH	*dai ^B (Li)	đây	đại	(bao)
chopsticks	<i>zhù</i> 箸	*d<r>ak-s	drjoH	*thu ^B (Pitt.)	đũa	trợ	tũa
saw (noun)	<i>jù</i> 鋸	*k(r)a-s	kjoH	*ku ^B (Li)	cưa	cứ	khưa
to embroider	<i>xiù</i> 繡	*[s]iw(k)-s	sjuwH	*se:w ^B (Pitt.)	thêu	tú	thêu
scissors/shears	<i>jiǎo</i> 鉸	*mə-[k] ^ʰ r[a]w?	kaewX	*ke:w ^C (tentative)	kéo	giáo	NA
to dye	<i>rǎn</i> 染	*C.n[a]m?	nyemX	*jwu:m ^C (Pitt.)	nhuộm	nhiễm	nhuộm
to cut off, to behead	<i>zhǎn</i> 斬	*[ts]ram?	tsreamX	*tram ^C (Pitt.)	chém	trám	chém
cooked/ripe	<i>shú</i> 熟	*[d]uk	dzyuwk	*suk ^D (Pitt.)	thuộc	thục	thuộc

a Khái niệm các dạng tiếng Hán cổ đại (Old Chinese: OC trong bảng) và tiếng Hán thời Trung cổ (Middle Chinese:MC trong bảng) (Baxter và Sagart 2014) và Tiếng Tày Nguyên Mẫu (Proto-Tai) (Li 1977; Pittayaporn 2009) chứa đựng hỗn hợp các ký hiệu Phiên Âm Quốc Tế: IPA và các từ loại (categories) trích yếu nhằm mục đích tái tạo lịch sử. Về tiếng Việt thì chữ Quốc Ngữ theo mẫu tự La Mã được hiển thị. L Tầng lớp Hán Việt (SV) bao gồm các cách đọc tiêu chuẩn của các ký tự tiếng Trung, dễ dàng tìm thấy trong các Từ điển Hán Việt, trong khi ECL thì không được thể hiện và thay vào đó thường được coi là từ vựng “bản địa”.

Phân Biệt Các Tình Trạng Tiếp Xúc Ngôn Ngữ Ban Đầu Với Tiếng Trung Hoa

Mặc dù có những từ mượn tiếng Hán (Sinitic) được chia sẻ, nhưng vẫn có những khác biệt lớn về số lượng và các loại vay mượn từ vựng. Nhìn chung, các lĩnh vực văn hóa của từ ngữ trong từ vựng Hán-Tày và Hán-Việt Ban Sơ khá khác nhau, ngụ ý hoàn cảnh văn hóa xã hội khác nhau và nơi tiếp xúc với các nhóm người Trung Hoa ban đầu. Khoảng 160 từ mượn tiếng Hán trong Tiếng Tày Nguyên Mẫu (Proto-Tai) bao gồm một loạt các mục từ văn hóa khác nhau. các mục từ về mậu dịch, và một vài hạng từ về các bộ phận cơ thể, nhưng không có sự kết tụ thành chùm (hay cụm) lớn của các lĩnh vực văn hóa. Các ECL được xác định có số lượng lớn hơn nhiều, lên tới hàng trăm mục từ vựng và do đó bao trùm phạm vi rộng hơn của các lĩnh vực ngữ nghĩa (semantic), bao gồm các từ liên quan đến quan hệ họ hàng và hôn nhân; trình độ học vấn và hành chính; kim loại, các vũ khí kim loại và các dụng cụ kim loại khác, trong số các loại khác.

Từ vựng ngữ pháp và chữ số vay mượn cũng được tiết lộ. Như được thể hiện trong Bảng 3, cả từ vựng Tiếng Tày (Tai) lẫn ESV (từ Hán Việt Ban Sơ) đều chứa những số lượng và tự loại đáng kể các từ ngữ ngữ pháp trong tiếng Hán (ví dụ: các giới từ, các từ chỉ vị trí/chiều hướng, thuật ngữ về số lượng, v.v.), nhưng có nhiều hơn đáng kể số từ của ECL (tức có hơn 20 trong tiếng Tày, nhưng hơn 60 trong tiếng Việt). Điều được phát hiện hơn nữa là chỉ có một số mục từ được chia sẻ bởi cả hai ngôn ngữ, đáng chú ý là loại từ (classifier) tiếng Hán *gè* 個 (OC; Tiếng Hán cổ: *kʰa[r]-s, thời Hậu Hán *kai^C,⁶⁶ tiếng Việt: *cái*, tiếng Tày (Tai) *kai^C (Li)). Gần như tất cả những từ ngữ khác từ thời kỳ này thì khác biệt. Đối với các thuật ngữ về con số, một điều nổi bật là những con số trong Tiếng Tày Nguyên Mẫu từ số 2 đến 99 được thay thế hoàn toàn bởi chữ Hán, trong khi tiếng Việt giữ lại tất cả các con số trong tiếng Việt bản địa và/hay trong ngữ hệ Nam Á (Austroasiatic).

Chính vì thế, trong khi một số từ vay mượn từ chữ Hán được nhận thấy trong cả tiếng Tày Nguyên Mẫu (Proto-Tai) và Tiếng Việt, đa số đều không được cả hai chia sẻ. Sự phân bố khác biệt sâu sắc này sự các từ mượn từ chữ Hán làm nổi bật các tình huống tiếp xúc ngôn ngữ riêng biệt với tiếng Hán (Sinitic). Vì vậy, những từ này phải được vay mượn ở các vị trí địa lý khác nhau vào đầu thiên niên kỷ thứ nhất theo Công Nguyên (CN). Một lần nữa, những dữ liệu này không gợi ý rằng các ngôn ngữ này được sử dụng ở đâu, chỉ cho thấy rằng cộng đồng ngôn ngữ tiếng Tày và tiếng Việt ban sơ nằm ở các vùng địa lý khác nhau. Các dữ liệu khảo cổ học được mô tả trước đây phần lớn định vị ngữ chi Việt (Vietic) ở ĐBSH. Việc vay mượn số đếm từ tiếng Hán (Sinitic) trong tiếng Tày trái ngược với việc giữ nguyên các từ ngữ bản địa trong tiếng Việt không thôi đã là một chỉ dấu mạnh mẽ. Xét về bản chất của các từ vay mượn, đặc biệt trong lĩnh vực hành chính và trình độ học vấn, điều hợp lý khi ước định rằng địa điểm của Cổ Loa, khu định cư đô thị nguyên mẫu,⁶⁷ có nhiều khả năng là địa điểm chính của sự tiếp xúc ban sơ của ngôn ngữ Hán-Việt.

BẢNG 3 Sự khác biệt giữa các lĩnh vực ngữ nghĩa trong tiếng Tày và Các Từ Chữ Hán Vay Mượn Ban Sơ (ECL)

Thể loại	Tày	Hán-Việt Ban Sơ (ESV)
Miền văn hóa của các từ vay mượn (không bao gồm từ vay mượn chung)	Các mục từ văn hóa, mật dịch, cac; bộ phận cơ thể cơ thể khác nhau	Nhiều thuật ngữ liên quan đến quan hệ họ hàng/hôn nhân, trình độ học vấn, vũ khí hành chính, màu sắc (ví dụ: màu vàng/kim loại vàng & màu trắng/kim loại bạc)
Từ vựng ngữ pháp	hơn 20 từ trong tiếng Tày Nguyên Mẫu và sơ khai	hơn 60 từ
Các từ về số đếm	Số đếm chữ Hán từ 2 đến 99	Các số đếm bản địa được giữ lại

Điều này phù hợp với quan điểm của tác giả Churchman rằng người dân vùng ĐBSH dường như bị ảnh hưởng sâu sắc bởi văn hóa Trung Hoa hơn so với những người ở vùng Lĩnh Nam. Ở những chỗ khác nhau trong chuyên luận lịch sử dân tộc học của tác giả Churchman về vùng Lĩnh Nam từ những năm 200 đến 700 CN, thường xuyên đề cập đến ĐBSH lân cận, một mô hình tái diễn xảy ra cho thấy vùng Lĩnh Nam khá khác biệt so với ĐBSH.⁶⁸ Những vấn đề này bao gồm sự kiểm soát nhiều hơn của Trung Hoa ở vùng ĐBSH,⁶⁹ sự biệt lập về mặt địa lý và từ đó cả về văn hóa của vùng ĐBSH,⁷⁰ và người dân vùng ĐBSH dễ dàng chấp nhận hệ thống chính trị Trung Quốc hơn như thế nào.⁷¹

Một từ vay mượn là một ví dụ đặc biệt hữu ích. Các sự khai quật khảo cổ đã phát hiện ra hàng nghìn viên ngói kiểu Trung Hoa tại di chỉ Cổ Loa từ khoảng năm 200 trước Công Nguyên.⁷² Từ “ngói : gạch lợp mái nhà: roof tile” trong tiếng Việt có thể là một từ vay mượn chữ Hán ban sơ: ECL.⁷³ Nó đã mong đợi sự tương ứng về âm vị với tiếng Hán cổ về mặt ngữ nghĩa, phụ âm, nguyên âm, âm điệu của nó. Âm /-j/ cuối cùng (đánh vần bằng “i” trong chính tả tiếng Việt) dường như bắt nguồn từ tiếng Hán cổ *[-j],⁷⁴ một đặc điểm được hỗ trợ bởi các vần điệu trong Shijing 詩經 (Kinh Thi). Điều này cho thấy nó là một sự vay mượn rất cổ xưa và do đó hoàn toàn có thể - hoặc thậm chí có thể - từ thời gian và địa điểm được đề cập. Do đó, nó có khả năng là một từ vay mượn từ chữ Hán xuất hiện rất sớm trong tiếng Việt dựa trên sự kết hợp giữa các dữ liệu khảo cổ học và dữ liệu ngôn ngữ lịch sử.⁷⁵

BẢNG 4 Dữ liệu từ vựng của từ “ngói” (roof tile)

Các Ngôn Ngữ	Dữ Liệu Từ Vựng
Mặt chữ	瓦 wǎ
OC tiếng Hán cổ	*C.ŋ ^w ra[j]?
MC tiếng Hán trung cổ	ngwaeX
ECL từ mượn chữ Hán ban sơ	ngói
SVHán-Việt	ngoã
Tiếng Tày	Các phương ngữ Tày Trung Tâm (Leiping vaa ³ ; Lungming vaa ⁵ ; Lungchow, Pingsiang vaa ⁵ ; Ningming ŋvaa ⁶ (Hudak 200X:1Xc))

Trong khi nó cũng được vay mượn trong một số ngôn ngữ Tày nhưng không thể tái tạo được ở cấp độ Tiếng Tày Nguyên Mẫu (Proto-Tai) Nguyên vì nó chủ yếu chỉ được nhận thấy ở Trung Tâm tiếng Tày trong các vùng tại hoặc giáp ranh với miền nam Trung Quốc. Vì vậy chúng ta có thể suy đoán rằng đây là một từ được vay mượn tương đối gần đây hơn trong các ngôn ngữ Tày đó. Hơn nữa, trong các biến thể của tiếng Tày, không có từ nào trong số này có âm lướt [-j] (offglide) cuối từ.⁷⁶ Dựa trên dữ liệu, một kịch bản có thể xảy ra xuất hiện: từ chữ Hán (Sinitic) này được vay mượn bởi những người nói tiếng Việt tại Cổ Loa trong thời Đông Sơn.⁷⁷ Nó dường như không được vay mượn vào tiếng Tày (Tai) trong thời kỳ ban sơ này. Mặc dù đây chỉ là một từ nhưng nó bằng chứng tiềm tàng về sự tiếp xúc tiếng Hán - tiếng Việt (Sinitic-Vietic) rất sớm ở vùng ĐBSH trong thời kỳ thời Đông Sơn.

Tên Gọi Dân Tộc Và Địa Danh

Một loại dữ liệu ngôn ngữ khác cần xem xét là tên gọi dân tộc các địa danh. Như đã lưu ý trong phần giới thiệu, việc sử dụng các tên gọi dân tộc trong văn bản tiếng Hán để hiểu lịch sử dân tộc nêu lên những trở ngại lớn. Các dữ liệu trong văn bản tiếng Hán và ngữ học dân tộc đưa ra những hướng dẫn rộng rãi nhưng chỉ có thể được coi là bao quát và dữ liệu nền tảng có phần không đáng tin cậy. Như đã lưu ý, thuật ngữ “Yue: Việt” được sử dụng rộng rãi áp dụng cho nhiều nhóm ở miền nam Trung Quốc và ngày nay vẫn đề cập đến các biến thể của người Hoa tại và xung quanh tỉnh Quảng Đông. Ngoài ra, thuật ngữ được lấy làm tên của vương quốc “Đại Việt” (tiếng Trung: Dayue 大越) ở Việt Nam trong nhiều thế kỷ đầu tiên của thiên niên kỷ thứ nhất theo Công Nguyên (CN) và cuối cùng trong tên gọi “Việt Nam”, tiếng Hán: Yuenan 越南. Kết quả, thuật ngữ này có thể chỉ đề xuất sự tiếp xúc khu vực và sự trao đổi văn hóa rất tổng quát.

Vào cuối thời Đông Sơn, tên hành chính cũng được dùng để chỉ tới khu vực phía Bắc Việt Nam. Từ thời nhà Hán đến thời nhà Đường, các danh xưng hành chính có tên Jiao 交 (tiếng Hán cổ: *kraw; Tiếng Việt: Giao), nguồn gốc của nó vẫn còn gây tranh cãi. Trong khi thuật ngữ này là không còn được sử dụng trong tiếng Việt, nó được dùng trong các ngôn ngữ Tày để chỉ người dân Việt. Tác giả Nguyễn Văn Lợi lần đầu tiên trình bày bằng chứng đối chiếu bằng tiếng Hán, tiếng Tày, và tiếng Hmong-Mien cho dạng này, gợi ý sự tái tạo lần lượt là *Kraw, *Kliaw, và *Klaw.⁷⁸ Sau đó, các tác giả Ferlus và Chamberlain cũng đề xuất tương tự rằng tên của một số ngôn ngữ Tày (ví dụ: Gelao, Lào) gợi ý một dạng có thể tái tạo được *k.ra:w.⁷⁹ Các tác giả kể sau cũng lưu ý rằng, trong khi dạng liên quan *kew* trong các ngôn ngữ Tày là dùng để chỉ người Việt, trong tiếng Tày Pupeo, *kew* chỉ các nhóm người Tày ở Việt Nam.

Tác giả Chamberlain cũng đưa ra giả thuyết rằng việc sử dụng từ *jiao* trong một văn bản Tày cổ gợi ý rằng thuật ngữ này nguyên thủy được áp dụng cho các nhóm nói tiếng Tày và rằng nó chỉ được áp dụng cho các nhóm dân nói tiếng Việt muộn hơn sau khi các nhóm người Việt đã định cư ở vùng ĐBSH.⁸⁰ Quan điểm nêu sau này dường như bị bác bỏ bởi các dữ liệu khảo cổ học ở vùng ĐBSH trong cuộc nghiên cứu năm 2018 của tác giả Lipson và các cộng sự: ít nhất, cả hai nhóm đều ở trong vùng từ giữa thiên niên kỷ thứ nhất trước CN. Bao quát nhất, việc sử dụng tên gọi dân tộc (ethnonyms) để chỉ các nhóm dân Tày trong quá khứ và những nhóm dân đó ở Việt Nam cho thấy khả năng thuật ngữ này có thể đã có một tham chiếu địa lý tổng quát hơn, đặt cả những người nói tiếng Tày lẫn người nói tiếng Việt ở vùng đất *Giao* trong thời kỳ ban sơ đó.

Về tên gọi các thành phần địa lý, sông Hồng chắc chắn có ý nghĩa quan trọng. sự tốt đẹp. Tác giả họ Trần liệt kê mười tên lịch sử khác nhau của sông Hồng, nơi mà ông các tuyên bố bao gồm một hỗn hợp của ngôn ngữ Nam Á, Tày và Trung Hoa từ các thời đại khác nhau.⁸¹ Các danh xưng dân tộc và các địa danh không thể phục hồi được các chi tiết về thời cổ đại lịch sử, nhưng kết hợp với dữ liệu khảo cổ học và các từ vay mượn từ tiếng Trung thời ban sơ, chúng chỉ ra rằng cả người nói tiếng Việt và người nói tiếng Tày đều đã có mặt lâu dài ở ĐBSH. Thật không may, bằng chứng này không cung cấp chi tiết cụ thể.

Sự Tiếp Xúc Ngôn Ngữ Tày-Việt

Các trung tâm địa lý của các cộng đồng ngôn ngữ Tày và Việt được thể hiện khác nhau dựa trên dữ liệu từ vay mượn từ tiếng Hán (Sinitic) ban sơ. Tuy nhiên, các học giả đã lưu ý rằng tiếng Tày và tiếng Việt (Vietic) có chung những từ không được thấy rộng rãi trong các ngôn ngữ khác,⁸² mặc dù việc xem xét cẩn thận những từ này cho thấy có vấn đề với một số tuyên bố, ngụ ý về một vùng tiếp xúc được chia sẻ chung. Nhưng ngay cả khi xem xét những nghiên cứu trước đây, số lượng từ ngữ chia sẻ của Tai-Việt chỉ là một phần nhỏ so với những từ vay mượn từ tiếng Hán ban sơ trong tiếng Tày (Tai) hoặc tiếng Việt. Các từ này hầu hết được coi là từ mượn từ tiếng Tày (Tai) vào tiếng Việt (Vietic) nhưng trong một số trường hợp, chiều hướng vay mượn không rõ ràng. Về mức độ chắc chắn của lịch sử từ nguyên chung, một số lượng nhỏ của các mục từ làm

giảm khả năng xác định các mẫu âm vị và các lĩnh vực ngữ nghĩa của các từ, chẳng hạn như các từ được thấy trong các từ vay mượn từ tiếng Hán trong cả tiếng Tày và tiếng Việt. Chính vì thế, các từ ngữ Tày-Việt cấu thành bằng chứng tương đối yếu hơn trong việc khám phá sự tiếp xúc văn hóa xã hội giữa các nhóm nói tiếng Tày và tiếng Việt ban sơ.

Nói chung, khi những từ như vậy được chia sẻ bởi tiếng Tày và tiếng Việt hoặc Tiếng Việt Nguyên Mẫu (Proto-Vietic), nhưng sự tái tạo tiếng Nam Á Nguyên Mẫu (Proto-Austroasiatic) (các sự tái cấu trúc của tác giả Shorto)⁸³ lại khác biệt, chiều hướng vay mượn có nhiều khả năng được giả định là từ tiếng Tày (Tai) sang tiếng Việt hoặc ngữ chi Việt (Vietic). Tuy nhiên, nếu không có số lượng lớn các từ Tày-Việt được chia sẻ chung, sự sử dụng cả dữ liệu ngôn ngữ và ngoài ngôn-ngữ (extralinguistic) trở nên quan trọng. Ví dụ, thuật ngữ Nam Á Nguyên Mẫu (Proto-Austroasiatic) phổ biến *da[ʔ] “duck: vịt” khác với dạng trong Tiếng Việt Nguyên Mẫu (Proto-Vietic): *vi:t, thay vào đó, nhiều phần là từ dạng trong Tiếng Tày Nguyên Mẫu (Proto-Tai): *pjetD. Trong khi không có bằng chứng vững chắc của khảo cổ vững động vật học về việc thuần hóa vịt ở miền nam Trung Hoa, có thể đây là một vùng áp dụng phương pháp thuần hóa vịt được chia sẻ chung, ngoài từ ám chỉ con vịt đã được thuần hóa.⁸⁴ Về tuy nhiên, từ “gạo (rice)” là hình dạng từ phức tạp trong cả Tiếng Việt Nguyên Thủy và Tiếng Nam Á Nguyên Thủy cho rằng từ này được mượn vào tiếng Tày chứ không phải theo chiều ngược lại. Đối với từ “trống (drum), có một giả thuyết cho rằng đó là được vay mượn từ tiếng Tày vào tiếng Việt, vì dữ liệu khảo cổ học cho thấy nghề luyện kim đã di chuyển từ miền bắc xuống miền nam. Tuy nhiên, dưới quan điểm ngôn ngữ học, âm tắc thanh môn cuối từ (final glottal stop) trong Tiếng Việt Nguyên Mẫu kết hợp với sự thiếu vắng của nó trong tiếng Tày cho thấy từ này được vay mượn từ tiếng Việt (Vietic) vào tiếng Tày (Tai), sau đó âm tắc thanh môn cuối cùng bị biến mất. Nếu theo chiều ngược lại thì không thể giải thích được âm tắc thanh môn trong tiếng Việt Nguyên Mẫu (Proto-Vietic).

Bảng 5 chỉ liệt kê 19 mục từ. Có lẽ còn nhiều hơn nữa, nhưng đây là các mục từ duy nhất được xem xét trong cuộc nghiên cứu này sẽ được hỗ trợ đầy đủ bởi phương pháp so sánh. Bất kể số lượng nhỏ bé này, một phạm vi hữu ích của các loại ngữ nghĩa có thể được xác định: các địa điểm, các dụng cụ và sản phẩm nông nghiệp; các nhạc cụ; và các từ về quan hệ họ hàng. Không thể biết chính xác hoàn cảnh trong đó các từ được vay mượn, nhưng, ví dụ, việc vay mượn của các thuật ngữ chỉ họ hàng (cả hai dành cho nữ giới) tương ứng với các dữ liệu khảo cổ di truyền học hiển thị cả nhiễm thể di truyền (gene) của người Tày(Tai) và lẫn nhiễm thể di truyền của người Kinh (nhóm kể tên sau là một danh xưng trong tiếng Việt chỉ dân tộc học Việt Nam có. Tuy nhiên, trong khi những từ này được tái tạo trong ngữ chi Việt (Vietic), tiếng Việt có các từ nguyên (etyma) hoàn toàn khác biệt cho những mục từ này. Hoạt động buôn bán, tất nhiên, có thể được coi là phương cách trao đổi từ vựng khả thi cho các vật dụng và sản phẩm cụ thể, nhưng bao nhiêu các thuật ngữ nông nghiệp gợi ý về lao động hợp tác giờ này vẫn chưa được hay biết.

Các hình dạng tổng thể của từ ngữ tương ứng tốt và hầu hết phù hợp với các khuôn mẫu được nhận thấy trong âm vị lịch sử của tiếng Việt. Ví dụ: từ tiếng Việt có chữ cái đầu /m/ đến từ tiếng Tày Nguyên Thủy *m hoặc *b, một mẫu hình phổ biến, cũng như từ Tiếng Việt Nguyên Mẫu (Proto-Vietic) sang tiếng Việt. Hầu hết các âm điệu trong các dạng này đều tương ứng, mặc dù một số không tương ứng. Chỉ những từ Tiếng Tày có các âm điệu loại B và C mới có thể được xác định là từ nửa đầu thiên niên kỷ thứ nhất sau CN hoặc sớm hơn, như trường hợp của những từ vay mượn tiếng Hán (Sinitic) ban sơ.

Nhìn chung, những dữ liệu từ vựng này gợi ý sự tiếp xúc của Tiếng Tày-Tiếng Việt trong các lĩnh vực mậu dịch, nông nghiệp và hôn nhân khác chủng tộc (intermarriage). Tuy nhiên, chúng ta không thể biết chắc chắn liệu những từ ngữ này đã được trao đổi vào thời tiền Tần hoặc trong hoặc sau thời nhà Hán, và những khoản vay mượn này có thể đã trải dài qua nhiều thế kỷ, không nhất thiết trong một thời kỳ tiếp xúc duy nhất. Và một lần nữa, đây là một số ít từ vay mượn, do đó làm giảm mức độ chắc chắn rằng chúng thực sự là những từ mượn. Không phải tất cả các từ đều phù hợp hoàn hảo và với nhiều dữ liệu được thu thập hơn trong tương lai, một số có thể được chứng minh là có sự tương đồng ngẫu nhiên, nhưng nói chung, đây là những hình thức hợp lý để cứu xét. Có thể nói rằng chúng phản ánh sự tiếp xúc văn hóa xã hội, nhưng không sâu đậm như trường hợp của hàng trăm từ mượn Hán cổ trong tiếng Việt.

BẢNG 5: Các Từ Ngữ Tày-Việt Cùng Chia Sẻ

Từ Vựng	Việt	Tày	Ngữ Chi Việt (Vietic)	Ngữ hệ Nam Á
mountain village	<i>mường</i>	*muəŋ ^A (Pitt.)	NA	*kuŋ
canal; ditch	<i>mường</i>	* ^h muəŋ ^A (Pitt.)	*-miəŋ	* ^h baaŋ
field	<i>đồng</i>	*doŋ ^B (Li)	NA (but cf. [tò:ŋ] “field” Rục)	*sre?
water pipe of bamboo	<i>máng</i>	*baŋ ^{B/C} “tube, bamboo” (Pitt.)	* ^h ba:ŋ? “water pipe of bamboo”	NA
winnowing basket	<i>nong</i>	* ^h doŋ ^C (Li)	* ^h do:ŋ? “flat winnowing basket”	NA (Comparable word shapes in many AA languages)
knife	<i>rựa</i>	* ^h vra ^C (Li)	* ^h m-ra:?	NA (Comparable word shapes in many AA languages)
rope/string	<i>dây</i>	*sa:ŋ ^A “string” (Pitt.)	NA	(Other etyma)
rice (uncooked)	<i>gạo</i>	*xəu ^C (Li)	*r-ko:?	*rk[aw]?
kind of sticky rice	<i>lam</i>	*hlaŋ ^A (Li)	NA	NA
duck	<i>vịt</i>	*pjet ^D (Li)	*vit “canard, duck”	* ^h da[?]
orange	<i>cam</i>	PSW: khaam ^A “tamarind” (Jonsson)	NA (possible loans into Thavung, Ruc)	NA
water spinach	<i>muống</i>	* ^h buŋ ^C (Pitt.)	* ^h bo:ŋ? “water spinach”	NA
plum	<i>mận</i>	* ^h man ^C (Pitt.)	* ^h mən? (my reconstruction; only in Việt-Mường and Cuói)	NA (Comparable word shapes in many AA languages)
drum	<i>trống</i>	*klo:ŋ ^A (Pitt.)	*klo:ŋ? (my reconstruction)	*sgər, *sgəər
reed flute / khene	<i>kèn</i>	*gɛn ^{A2} (Li)	*gɛ:n “khene / reeded wind instrument”	NA (this word appears in many AA languages)
to pound / crush (rice)	<i>đâm</i>	*tam ^A (Li, Pitt.)	*təm “to pound (rice)”	(Other etyma)
to sweep	<i>quét</i>	*kwa:t ^D (Pitt.)	NA	(Other etyma)
daughter-in-law	NA	* ^h bəu ^C (Li)	* ^h bə:? “daughter-in-law”	(Other etyma)
aunt	NA	* ^h pa: ^C (Li) “aunt”	* ^h pa:? “father’s elder brother’s wife / father’s / mother’s elder sister”	(Other etyma)

BẢNG 6: Các Vấn Đề Ngữ Học Và Các Hàm Ý Về Tiếng Tày và Tiếng Việt

Các Vấn Đề Ngữ Học	Các Hàm Ý
a. Sự hội tụ về kiểu hình ngôn ngữ chung	- Biểu thị tình trạng song ngữ và đôi khi đa ngôn ngữ phổ biến, lâu dài giữa các cộng đồng nói tiếng Hán, Tày và Việt từ thời Nhà Hán
b. Các từ mượn tiếng Hán ban sơ chung	- Xác định vị trí Tiếng Tày và Tiếng Việt ở vùng Lĩnh Nam-ĐBSH từ thời Nhà Hán
c. Nhiều tình huống tiếp xúc ngôn ngữ ban sơ với tiếng Hán	- Chấm định các trung tâm Tiếng Tày và Tiếng Việt ở vav' vùng khác nhau - Nêu ý kiến một số người nói tiếng Việt ở (các) trung tâm Hành chính của ĐBSH, nơi nhiều phần có mật độ dân số đông hơn
d. Tên gọi dân tộc và địa danh	- Đặt định Tiếng Tày và Tiếng Việt trong vùng ĐBSH bao quát, nhưng cung cấp ít hay không có chi tiết về vị trí, thời biểu, hay quy mô dân số
e. Tiếp xúc ngôn ngữ Tày-Việt	- Đặt định Tiếng Tày và Tiếng Việt trong vùng ĐBSH - Nêu ý kiến sự tiếp xúc văn hóa xã hội sớm giữa Tiếng Tày và Tiếng Việt, nhưng với thời biểu không chắc chắn và vẫn còn hạn chế so với sự tiếp xúc của cả hai nhóm với Tiếng Hán

Tóm Tắt Các Hàm Ý Của Các Dữ Liệu Ngôn Ngữ

Các phần dữ liệu ngôn ngữ khác nhau cung cấp những hàm ý khác nhau, như được tóm lược trong Bảng 6. Sự tiếp xúc chung với tiếng Hán (Sinitic) từ thời kỳ trước các âm điệu trở thành các yếu tố âm vị học đầy đủ, như được thể hiện bằng các từ có Âm Diệu B và Âm Diệu C, chứng tỏ rằng cả Tiếng Tày (Tai) và Tiếng Việt (Vietic) đều thuộc vùng Lĩnh Nam-ĐBSH nói chung trong thời kỳ ban sơ đó, nhưng không cụ thể ở địa điểm nào. Ngoài ra, sự phủ trùm các danh xưng chủng tộc, các tên gọi sông Hồng và các từ ngữ Tày-Việt cùng chia sẻ đều gợi ý về sự hiện

diện của cả hai nhóm trong khu vực trước khi có sự xuất hiện của các nhóm người nói tiếng Hán (Sinitic), nhưng còn ít mức độ chắc chắn hơn về thời gian hoặc địa điểm. Mặt khác, các loại từ vay mượn tiếng Hán (Sinitic) khác nhau và có thể có nhiều loại hình tiếp xúc văn hóa khác nhau với các nhóm người Trung Hoa cho thấy rõ ràng rằng các trung tâm của Tiếng Tày và Tiếng Việt nằm ở các khu vực khác nhau, ít nhất là vào cuối thời Đông Sơn và thời Nhà Hán.. Hơn nữa, vị trí tiếp xúc ngôn ngữ này đối với Tiếng Việt ít có khả năng tập trung ở vùng đồi núi miền Bắc-Trung-Bộ Việt Nam. Đúng hơn điều hợp lý để giả định sự tiếp xúc với các viên chức và các dân định cư Trung Hoa khác chủ yếu ở trong và xung quanh khu cư trú Cổ Loa thời tiền đô thị, một trung tâm rõ ràng về mậu dịch và hành chính.

Lượng Giá Ba Kịch Bản Chính Và Địa Lý Ngụ Ý của Các Cộng Đồng Ngôn Ngữ

Bây giờ chúng ta quay lại ba kịch bản chính được mô tả trong phần dẫn nhập. về tình hình ngôn ngữ chủng tộc ở vùng ĐBSH thời kỳ tiền Tần. Không có thể loại trừ hoàn toàn bất kỳ kịch bản nào trong chúng, nhưng chúng có thể được xem xét về mức độ quan trọng của bằng chứng hỗ trợ mỗi kịch bản.

1. **Chủ yếu là Tiếng Việt ở Đông Sơn.** Kịch bản này là giả thuyết được hỗ trợ mạnh mẽ nhất. Dữ liệu khảo cổ học và các kiểu sọ và răng người gợi ý mạnh mẽ về dân số nói tiếng Nam Á (Austroasiatic) khoảng 2000 năm TCN, từ thời Phùng Nguyên (2000 đến 1500 TCN), như đã ghi nhận trong phần thông tin khảo cổ học. Những nhóm đó thay thế các cộng đồng ngôn ngữ trước đó của những người săn bắt hái lượm trong khu vực, mặc dù chi tiết về quá trình này vẫn chưa được biết đến.⁸⁵ Qua nhiều thế kỷ tiếp theo, ngữ chi Việt (Vietic) dần dần trở thành một nhánh riêng biệt của ngữ hệ Nam Á (Austroasiatic) và vào thời kỳ Đông Sơn, các nhóm ngữ chi Việt này có sự tiếp xúc văn hóa xã hội với các nhóm người Tày (Tai), phần lớn ở phía bắc vùng Lĩnh Nam. Sự tiếp xúc này nhiều phần đã xảy ra trong nhiều thế kỷ thông qua cả trong hoạt động mậu dịch lẫn sự trao đổi nhiễm thể [nhân của tế bào] di truyền (genes).

2. Sự hiện diện đồng đều của ngữ chi Việt (Vietic) và tiếng Tày ở Đông Sơn. Kịch bản này có phần được hỗ trợ bởi các dữ liệu. Vẫn có trường hợp rằng vùng ĐBSH nguyên thủy chủ yếu nói tiếng Nam Á (Austroasiatic), với ngữ chi Việt (Vietic) trở thành một ngữ chi khác biệt. Tuy nhiên, đến thời Đông Sơn, cộng đồng ngôn ngữ Tày và cộng đồng ngôn ngữ Việt có ảnh hưởng gần như ngang nhau trong khu vực. Kịch bản này phù hợp với giả thuyết về một cộng đồng đa sắc tộc khả dĩ ở một khu vực đô thị nguyên mẫu cấp quốc gia tại Cổ Loa. Tuy nhiên, các bộ phận của cộng đồng ngôn ngữ Tày dần dần bị hấp thụ và/hoặc chuyển đến vùng ngoại ĐBSH, như tình trạng hiện nay.

3. Nhóm người nói tiếng Tày chiếm ưu thế ở Đông Sơn . Khó duy trì được kịch bản này như một giả thuyết khả thi. Vào thời Phùng Nguyên, ĐBSH chủ yếu là người nói tiếng Nam Á (với một

nhóm ngôn ngữ khác giờ đây không tồn tại), nhưng trong kịch bản này, vào thời Đông Sơn, các nhóm nói tiếng Tày đã trở thành nhóm thống trị, còn nhóm ngữ chi Việt vẫn chỉ ở phía nam. Nhóm nói tiếng Tày chiếm ưu thế xuyên qua thiên niên kỷ thứ nhất của CN, sau đó các nhóm ngữ chi Việt (Vietic) chịu ảnh hưởng của chữ Hán (Sinitic) lại đã trở nên nhóm thống trị ở ĐBSH. Những bằng chứng về thời kỳ thống trị này của người Tày ở vùng ĐBSH bao gồm sự giải thích rộng rãi của các văn bản lịch sử và tên gọi chủng tộc của Trung Hoa, nhưng nhìn chung, không có bằng chứng văn bản cụ thể cho quan điểm này.

Một hậu quả của kịch bản đầu tiên là quê hương giả định của cộng đồng ngôn ngữ Việt (Vietic) sẽ mở rộng từ vùng ĐBSH đến Bắc Trung Bộ Việt Nam hoặc chủ yếu nằm trong vùng ĐBSH với các nhóm ngữ chi Việt (Vietic) khác sau này di chuyển về phía nam. Thực trạng đa dạng ngôn ngữ của ngữ chi Việt ở phía/ nam của ĐBSH và một quê hương ở phía bắc tương tự như trường hợp của chữ Hán (Sinitic) với một quê hương miền bắc nhưng sự đa dạng về ngôn ngữ thậm chí còn nhiều hơn ở vùng vươn tới phía nam đạt. Tuy nhiên, tại thời điểm này, không có bằng chứng mạnh mẽ nào cho thấy vị trí phân tán của ngữ chi Việt (Vietic), và chúng ta chỉ có thể mặc nhiên thừa nhận một khu vực tổng quát từ miền bắc đến bắc trung bộ Việt Nam và các vùng giáp ranh với Lào của ngữ chi Việt (Vietic) cho đến khi có cuộc nam tiến của người Việt sau khi vương quốc Chăm sụp đổ vào những năm 1400.

Một câu hỏi vẫn còn khó trả lời là sự hiện diện văn hóa của Trung Hoa ở vùng ĐBSH rộng lớn ra sao so với vùng xa hơn về phía nam trong khu vực của ngữ chi Việt phương nam (southern Vietic). Số lượng người nói tiếng Hán ở vùng ĐBSH nhiều hơn giải thích cho mức độ nhiều hơn của tác động từ vựng tiếng Hán trong tiếng Việt-Mường nhưng ít hơn trong ngữ chi Việt miền Nam. Các lăng mộ nhà Hán đã được đặt tại tỉnh Thanh Hóa nằm ở phía nam ĐBSH, do đó cung cấp ít nhất một vị trí điểm đầu ấn địa lý có thể xem xét.

Nhiều chi tiết dân tộc học khác không thể được xác định từ thông tin được cung cấp trong cuộc nghiên cứu này. Có lẽ có một loạt đồng đảo các thể chế chính trị xã hội trong số các cộng đồng ngôn ngữ Việt nguyên thủy: một cấu trúc cấp nhà nước tập trung ở Cổ Loa và có thể cả các lãnh địa (chiefdoms) (có lẽ như thế, nhưng không thể chứng minh được cũng như không thể bác bỏ được) và các bộ lạc (có thể cứu xét các nhóm hiện đại). Những nhóm ở trong hoặc gần trung tâm hành chính sẽ trở thành chịu ảnh hưởng nhiều nhất về mặt văn hóa xã hội và ngôn ngữ bởi sự hiện diện của văn hóa Trung Hoa và ngôn ngữ Hán, như có thể suy ra từ tiếng Việt hiện đại. Những kẻ trong các bộ lạc, hoặc những người bị di dời khỏi các lãnh địa, sẽ giữ được nhiều hơn. Các sự thực hành văn hóa xã hội nguyên thủy. Những điểm sau này mang tính suy đoán và được nêu lên chỉ là những gợi ý để xem xét khi có nhiều dữ liệu hơn được thu thập và phân tích. Sự thu thập và đánh giá bằng chứng lịch sử dân tộc học của một sự hiện diện của Trung Hoa ban sơ ở khu vực phía nam ĐBSH vào thời nhà Hán sẽ tiết lộ các chi tiết bổ sung và có khả năng ảnh hưởng đến các giả thuyết cứu xét được đưa ra trong nghiên cứu này./-

CHÚ THÍCH

1. Bài viết này nối tiếp hai bài trình bày của tôi vào tháng 5 năm 2019 (Alves 2019a và 2019b), bài thứ hai là bản dịch tiếng Việt của bài đầu tiên. Tôi đánh giá cao những phản hồi tôi nhận được từ những người ở cả hai cuộc nói chuyện. Tôi cũng đánh giá cao những bình luận về bài viết của Zev Handel thuộc Đại học Washington.

2. Tham chiếu Brindley (2015) để có cái nhìn tổng quan thực chất về các vấn đề lịch sử dân tộc của Bách Việt (Baiyue).

3. Việc phân biệt ngôn ngữ, phương ngữ và các nhóm ngôn ngữ hoặc phương ngữ – ngay cả những ngôn ngữ hiện đại – có thể khó khăn. Điều này tạo ra những thách thức trong việc đưa ra tuyên bố về lịch sử thời chưa được ghi chép. Vì vậy, trong bài viết này, thuật ngữ “cộng đồng ngôn ngữ” được sử dụng như một phương pháp phỏng đoán để đề cập, một cách không nhất thiết, đến những người nói một ngôn ngữ duy nhất, bất biến, mà đúng hơn, đến những người nói thuộc một nhóm các dạng giọng nói có liên quan với nhau với mức độ thông hiểu lẫn nhau khác nhau. Chúng có thể bao gồm các phương ngữ liên quan được sử dụng bởi các nhóm có chung một số tập tục văn hóa và các cảm thức khác nhau về căn cước văn hóa chung.

4. Thuật ngữ “ngữ hệ hay họ ngôn ngữ” dùng để chỉ một nhóm ngôn ngữ có chung ngôn ngữ gốc hoặc nhóm ngôn ngữ có quan hệ gần gũi trong quá khứ xa xưa. Thuật ngữ này đơn giản hóa các tình huống văn hóa xã hội nhưng cho phép đề cập đến một tình huống văn hóa xã hội phức tạp, ít được hiểu rõ.

5. Ví dụ: Bellwood (2005) về khảo cổ học và các ngữ hệ khác nhau trên thế giới, v.v. Về chủ đề các nhóm Bách Việt (Baiyue), Brindley (2015) dành một chương cho các câu hỏi lịch sử ngôn ngữ, đề cập chính xác đến lịch sử của các ngữ hệ hiện đại trong khu vực.

6. Ví dụ: Bellwood (2005); Matsumura và các cộng sự. (2018), trong số những công trình khác.

7 Thuật ngữ “ ngữ chi Việt (Vietic)” bao gồm cả nhánh phụ được gọi là “Việt-Mường” và hàng tá ngôn ngữ và nhóm phương ngữ liên quan khác. Mối quan hệ của các ngôn ngữ khác này với Việt-Mường đã được đưa ra giả thuyết (xem Sidwell và Rau 2015), nhưng hiện tại chưa có sự phân nhánh phát sinh loài nào được thống nhất. Vì mục đích của bài viết này, những ngôn ngữ này sẽ được gọi chung là “ngữ chi Việt Phương Nam (Southern Vietic)” (theo Ferlus 2004), vì chúng nằm ở phía nam của vùng Việt-Mường.

8. Ngữ hệ Nam Á (Austroasiatic) bao gồm hơn 160 ngôn ngữ trải rộng khắp Đông Nam Á Đại Lục (MSEA), vươn tới miền nam Trung Hoa và kéo dài đến miền đông Ấn Độ. Ngày nay, tiếng Việt và tiếng Khmer là ngôn ngữ quốc gia Nam Á duy nhất, mặc dù tiếng Môn trước đây là ngôn ngữ của các vương quốc Môn; tất cả các ngôn ngữ Nam Á khác đều là ngôn ngữ dân tộc thiểu số.

9. Tai-Kadai [Tiếng Tày-Trung Tâm Kadai], còn được gọi là ngữ hệ Kradai, bao gồm hơn 90 ngôn ngữ ở miền nam Trung Hoa và Đông Nam Á Đại Lục (MSEA). Tiếng Thái và tiếng Lào là các ngôn ngữ quốc gia duy nhất, trong khi tất cả các ngôn ngữ khác đều là ngôn ngữ dân tộc thiểu số.

10. Kim và các cộng sự. (2010).

11. Ví dụ: Đào (2005).

12. Tham chiếu. Ferlus (2006); Ferlus (2009a).

13. O’Harrow (1979), 142.

14. Ferlus (1989–1990), 58; Ferlus (2009b), 95.

15. Trần (2011), 219.

16. Sau khi xem bài thuyết trình của tôi tại Đại Học Quốc Gia Việt Nam vào tháng 5 năm 2019 (Alves 2019b), ông Trần (2019) sau đó đã xuất bản một bài viết (bằng tiếng Việt), trong đó ông trích dẫn bài thuyết trình của tôi và ủng hộ quan điểm của tôi.

17. Taylor (2013), 5–6, 19.

18. Trong khi thuật ngữ “Mon-Khmer” được sử dụng thường xuyên trong thế kỷ qua, nhưng nó đã được tranh luận một cách hữu hiệu là không có giá trị phát sinh loài (phylogenetic) (ví dụ Sidwell và Blench 2011), do đó khiến từ “Austroasiatic: Úc Á hay Nam Á” trở thành thuật ngữ duy nhất phù hợp và rõ ràng cho ngữ hệ này. Vì vậy, chỉ có từ Austroasiatic: Nam Á sẽ được sử dụng trong bài viết này.

19. Phan (2013), 6–7.

20. Chamberlain (2018), 20. Ông viết, “Tôi không thấy bằng chứng thuyết phục nào về sự hiện diện của ngữ hệ Nam Á ở miền nam Trung Quốc hoặc lãnh thổ ban sơ được gọi là Giao Chi...”

và hơn nữa: “Không có bằng chứng nào về ngữ chi Tiếng Việt (Vietic), Tiếng Việt-Mường Nguyên Mẫu (Proto-Việt-Mường) hoặc những người nói tiếng Nam Á khác sống trong và xung quanh Giao Chi ở hạ lưu sông Hồng trước thế kỷ 10 hoặc 11” (cùng nơi dẫn trên, 33). Dữ liệu trong nghiên cứu hiện tại dựa trên nhiều nguồn bằng chứng mà Chamberlain không có sự truy cập.

21. Chamberlain (1998); Chamberlain (2016), và ghi chú trong Kiernan (2017), 45–50.

22. Chamberlain (1998), 41.

23. Một số ví dụ về tiếng Tày là tiếng Lào, tiếng Nùng và tiếng Tày Trắng, tất cả đều có từ *kɛ:wɩ* để chỉ tiếng Việt, đôi khi được dùng với hàm ý tiêu cực.

24. Chamberlain (2016), 70.

25. Tham chiếu Hoàng (1989), 224–226; Alves (2007).

26. Chamberlain (2018), 158.

27. Reid (1994b).

28. Taylor (2013), 219.

29 Thuật ngữ được đặt ra sớm nhất cho ngữ hệ Austroasiatic: Nam Á là “Mon-Annam”, nhưng thuật ngữ “Austroasiatic: Úc-Á hay Nam Á” lần đầu tiên được Schmidt sử dụng trong bản in vào năm 1906. Để biết lịch sử nghiên cứu về ngữ hệ Nam Á, xem Sidwell (2009).

30. Tham chiếu Bellwood (2005), 221–226; Starosta (2005), 193–194; Higham (2017).

31. Tham chiếu Blust (1996), 124; Diffloth (2005). Norman và Mei (1976) đề xuất rằng các nhóm người Nam Á cư trú ở tận phía bắc Trung Quốc tới tỉnh Phúc Kiến. Giả thuyết của họ đã được trích dẫn nhiều lần, mặc dù dựa trên rất ít bằng chứng. Tác giả Sagart bác bỏ hầu hết các từ vựng bằng chứng họ cung cấp (Sagart 2008, 141–143). Hơn nữa, bờ biển Trung Quốc từ sông Dương Tử trở về phía nam có nhiều khả năng trước đây là nơi cư trú của người Austronesian: Nam Đảo và Các Nhóm Tiếng Tày - ngữ hệ Trung Tâm (Tai-Kadai). Nhìn chung, giả thuyết của Norman và Mei không còn được coi là có giá trị nữa. Cần có bằng chứng khác để xác định phạm vi của ngữ hệ Nam Á trong thời kỳ Đồ Đá Mới (Neolithic).

32. Sidwell và Blench (2011). Sidwell (2021) gần đây đã tuyên bố rằng ông nhận thấy một khu vực đáng quan tâm đối với vị trí phân tán ngữ hệ Nam Á ở Đòng Bằng Sông Hồng (viết tắt:

ĐBSH (RRD). Ông lưu ý một số trong cùng các địa điểm khảo cổ được thảo luận trong bài viết này là bằng chứng về một trong những điểm sớm nhất có sự hiện diện của người Nam Á trong khu vực.

33. Tham chiếu van Driem (2011).

34. McColl và các cộng sự. thừa nhận rằng cuộc nghiên cứu khảo cổ học quy mô lớn của họ không hoàn toàn ủng hộ cũng như không bác bỏ giả thuyết “Hai tầng lớp: Two-layer”, cũng như không tuyên bố mạnh mẽ rằng nông nghiệp đã được đưa đến Đông Nam Á Đại Lục (MSEA) và họ cho rằng đây là một vấn đề phức tạp hơn (McColl và các cộng sự 2018, 1) . Đây là một khẳng định có sắc thái hợp lý, nhưng việc chia sẻ một số hoạt động nông nghiệp với một số tác động văn hóa xã hội đáng kể dường như được ủng hộ mạnh mẽ, một lập trường có phần nhẹ nhàng hơn dường như có sự ủng hộ hợp lý trong tài liệu khảo cổ học.

35. Ví dụ: Koenigswald (1952); Matsumura và các cộng sự. (2019), trong số những công trình khác.

36. Ví dụ: Bellwood (2005).

37. Higham (2017); Higham (2021).

38. Ví dụ: Caoam (2014), 100.

39. Rau và Sidwell (2019); Sidwell (2021).

40. Cùng nơi đã dẫn trên.

41. Tác giả Kim tóm tắt các mốc thời gian và các nghiên cứu trước đây: Kim (2015), 106.

42. Ví dụ: Matsumura và các cộng sự. (2008); Corny và các cộng sự. (2017).

43. Điều này được hỗ trợ bởi Lipson và các cộng sự (2018). McColl và các cộng sự. (2018) cũng thực hiện nghiên cứu DNA trên các di tích khảo cổ ở Đông Nam Á. Tuy nhiên, mặc dù chúng cung cấp các thời đại và vị trí địa lý chung gần đúng nhưng các chi tiết lại khó diễn giải hơn. Ví dụ: họ mô tả Nhóm 2 bao gồm các cá nhân từ giai đoạn 4291 đến 2194 BP. (Trước Hiện Đại) (mà họ mô tả là cuối thời kỳ đồ đá mới và đầu thời kỳ đồ đồng) ở Việt Nam, Lào và bán đảo Mã Lai. Liên quan đến văn hóa Đông Sơn, trong phân tích di truyền của họ, Nhóm 3 thời Đông Sơn ở miền bắc Việt Nam (nhưng một lần nữa, không có vị trí khảo cổ chính xác) có điều mà họ gọi là nhiễm thể di truyền “Austro-Tai: Úc-Tày”, nhiễm thể di truyền Amis, Dai và Kradai. Điều này cung cấp bằng chứng khả hữu về sự hiện diện di truyền của người Tày, vấn đề này trở nên phức

tạp do ngày càng có nhiều bằng chứng cho rằng ngữ hệ Kra-dai là một phần của Austronesian: Nam Đảo. Mô tả của họ không đề cập đến bộ gen của người Austroasiatic, trái ngược với Hình 2, Sơ đồ C của họ, đánh dấu rõ ràng bộ (nhiễm thể di truyền) gen của người Austroasiatic. Điều đáng ngạc nhiên là Nhóm 2 và 4 trong vùng Đông Nam Á Đại Lục (MSEA) với các cá nhân đến từ Việt Nam (một lần nữa, không có các địa điểm cụ thể) đều có bộ gen Austroasiatic: Nam Á. Cần phải biết dữ liệu địa điểm cụ thể ở miền Bắc Việt Nam để giải thích điều đó. Tôi không đủ điều kiện để giải thích dữ liệu bộ gen trong cuộc nghiên cứu của tác giả McColl và các cộng sự, mặc dù sơ đồ của Nhóm 3 dường như chỉ ra bộ gen của cả hai nhóm Nam Á và Kra-dai, phù hợp với sự nghiên cứu của Lipson và các cộng sự.

44. Mất thời gian bao lâu để những người nói tiếng Australo-Melanesian chuyển sang hoặc truyền bá ngữ hệ Austroasiatic hoặc tác động của nó đến ngôn ngữ Austroasiatic là những vấn đề chưa được tìm hiểu. Không thể thừa nhận bất kỳ tác động nào nếu không có bất kỳ bằng chứng đối chiếu ngôn ngữ học nào. Tuy nhiên, một hiện tượng phổ biến khi hai nhóm ngôn ngữ có sự tiếp xúc ngôn ngữ mạnh mẽ và rộng rãi là sự đơn giản hóa cấu trúc ngôn ngữ (xem McWhorter 2007). Mối liên hệ giữa Nam Đảo và Nam Á, cụ thể là giả thuyết Austric (ngữ chi Nam Đảo [?], đã được Schmidt ghi nhận vào năm 1906, và các cuộc nghiên cứu về vấn đề này vẫn tiếp tục. (ví dụ Diffloth 1994, Reid 1994a, Reid 2005, Higham 1996, v.v.), nhưng vẫn chưa có sự đồng thuận rộng rãi giữa các nhà nghiên cứu. Một khả năng là các ngôn ngữ Nam Á (Austroasiatic), sau khi tiếp xúc với các ngôn ngữ sớm hơn, đã mất đi một số hình thái biến tố nguyên thủy và các hậu tố (suffixes) phổ biến trong các ngôn ngữ Nam Đảo (Austronesian), chỉ còn lại các tiền tố (prefixes), trung tố (infixes) và một số hình thái phái sinh. Một khả năng khác là các ngôn ngữ Nam Đảo (Austronesian) ở mỗi khu vực gặp phải những ngôn ngữ hoàn toàn khác nhau, như đã xảy ra sau hàng chục nghìn năm phát triển ngôn ngữ giữa các ngôn ngữ của các nhóm Australo-Melanesian, và mỗi nhóm sẽ để lại một chút từ vựng. còn sót lại, do đó làm tăng thêm sự khác biệt về ngôn ngữ giữa các nhánh của ngôn ngữ Nam Á (Austroasiatic). Tuy nhiên, một lần nữa, nếu không có bằng chứng đối chiếu ngôn ngữ học hiện đại còn tồn tại, câu hỏi này về sự tiếp xúc ngôn ngữ không thể trả lời được và việc xác định các từ khu vực được vay mượn hơn 3.000 năm trước là cực kỳ khó khăn, nếu không muốn nói là không thể làm được.

45. Ví dụ: Higham và. al. (2011); Higham (2014), 194; vân vân.

46. Tham chiếu Alves (2015), người đã tóm tắt dữ liệu khảo cổ học và cung cấp dữ liệu lịch sử ngôn ngữ hỗ trợ cho điều này.

47. Lipson và các cộng sự. (2018).

48. Tác giả Kim (2015) thảo luận nhiều lần về vấn đề giai tầng xã hội trong cuốn sách về công trình quan trọng của ông về sự xuất hiện của các chính thể ban đầu ở miền Bắc Việt Nam thời kỳ trước Công nguyên.

49. Thuật ngữ “sự tiếp xúc ngôn ngữ” đề cập đến nhiều tình huống khác nhau, từ song ngữ/đa ngôn ngữ trong các cộng đồng, với tác động ngôn ngữ đáng kể lên các ngôn ngữ, đến các tình huống trong đó hai hoặc nhiều ngôn ngữ được nói gần nhau với tình huống song ngữ ở mức độ thấp hơn nhưng vẫn có sự vay mượn từ vựng và đôi khi dẫn đến những đặc điểm ngôn ngữ chung.

50. Tác giả Haudricourt (1954) đã đưa ra giả thuyết chính yếu đầu tiên cho vấn đề này và nó đã được ủng hộ tích cực trong nhiều thập kỷ nhưng vẫn tiếp tục được thăm dò và cải tiến.

51. Đây cũng là trường hợp của tiếng Hmong-Mien Nguyên Mẫu: Proto-Hmong-Mien, vì nó nằm trong khu vực tiếp xúc chung với tiếng Hán (Sinitic) và tiếng Tày (Tai) (nhưng không phải ở miền bắc Việt Nam cho đến vài thế kỷ trước), và các ngôn ngữ con của nó đã chứng kiến sự phát triển như vậy. Độc giả có thể tham khảo Ratliff (2010) và các nghiên cứu khác để biết thêm thông tin.

52. Tôi đề cập đến sự tái thiết của các tác giả Baxter và Sagart (2014).

53. Những công trình tái thiết chính là của Li (1977) và Pittayaporn (2009). Tiếng Tày (Tai) là một nhánh của ngữ hệ Tai-Kadai (còn được gọi là Kradai), chứ không phải là toàn bộ ngữ hệ. Tuy nhiên, hiện tại chưa có sự tái thiết hoàn toàn của tiếng Tày Trung Tâm: Tai-Kadai, và khi xem xét mối liên hệ với ngữ chi tiếng Việt (Vietic), tiếng Tày (Tai) là bộ phận chính yếu của ngữ hệ Kradai gần cận nhất về mặt địa lý, lịch sử và có liên quan trực tiếp.

54. Các công trình liên quan bao gồm các cuộc nghiên cứu của Nguyễn Tài Cẩn (1995), Ferlus (2014), Shimizu (2000), Shimizu (2015), và Gong (2019).

55. Ví dụ: Ballard (1981); Delancey (2013); Ratliff (2010); Alves (2021).

56. Tham chiếu Matthews (2006).

57. Các sự mô tả bằng văn bản tiếng Hán về âm điệu (tones) có niên đại từ những năm 500 sau Công Nguyên CN. Tuy nhiên, chúng tôi không thể loại trừ khả năng phát triển âm sắc sớm hơn trong suốt thời kỳ Hán hóa (Sinitic). Song, sức nặng của bằng chứng từ những từ vay mượn từ tiếng Hán (Sinitic) ban sơ trong cả tiếng Tày (Tai) và tiếng Việt cho thấy rằng các âm điệu đã không xuất hiện trong những thế kỷ ban đầu, nếu không thì sẽ không có nhiều sự nhất quán về các khuôn mẫu âm điệu giữa các ngôn ngữ này. Đối chiếu với Ratliff (2010).

58. Một số tiếp xúc ngôn ngữ đã dẫn đến “sự tiếp thu không hoàn hảo”, điều mà McWhorter (2007) gọi là “Ngôn ngữ thứ nhì được quy ước không lai căng”, một hệ quả tự nhiên của việc

học ngôn ngữ mà không có sự giảng dạy hoặc hướng dẫn rõ ràng. Tình trạng này không dẫn đến sự pha chế hỗn tạp ngôn ngữ (creolization) hơn là các cấu trúc bị đơn giản hóa, chẳng hạn như mất đi chất liệu hình thái và đơn giản hóa cấu trúc âm tiết. Điều này được kết hợp với sự hội tụ về loại hình qua nhiều thế kỷ. Có bằng chứng cho thấy các âm điệu trong nhánh Việt-Mường của tiếng Việt xuất hiện ít nhất vài thế kỷ sau khi chúng phát triển trong tiếng Hán (xem Alves 2018). Hơn nữa, hệ thống âm điệu trong nhánh Việt-Mường đi theo hệ thống âm điệu tiếng Hán. Tuy nhiên, sẽ không chính xác khi cho rằng ngữ chi tiếng Việt (Vietic) nhận được hệ thống âm điệu từ tiếng Hán (Sinitic): thay vào đó, các ngôn ngữ thuộc ngữ chi tiếng Việt (Vietic) có các hệ thống âm vị (phonation) (ví dụ: giọng rõ ràng, giọng khàn khàn, giọng thờ dốc, giọng căng thẳng) dẫn đến các hệ thống âm điệu đồng đều cục bộ, qua đó hệ thống âm điệu tiếng Hán dường như đã thêm vào một số yếu tố phức tạp (Alves 2001). Sự phát triển nguyên thủy của các âm điệu trong tiếng Tày Trung Tâm Kadai (Tai-Kadai) và ảnh hưởng tiềm tàng của việc tiếp xúc với chữ Hán thậm chí còn ít rõ ràng hơn. Các trường hợp vay mượn từ vựng sang tiếng Hán từ các ngôn ngữ ở miền nam Trung Quốc và/hoặc ở vùng Đông Nam Á Đại Lục (MSEA) có số lượng nhỏ và với trạng thái không chắc chắn. Ví dụ, người ta đã đưa ra giả thuyết (ví dụ: Norman và Mei 1976, 295; Schuessler 2007, 404–405; Blench 2017, 221–223) rằng từ tiếng Hán có nghĩa là “nỏ (crossbow)” (tiếng Trung Hoa: 弩, OC (Tiếng Trung Hoa cổ): *C.nʰaʔ) nguyên thủy là một từ Nam Á (Austroasiatic), được phục chế thành *snaʔ (Shorto 2006). Nếu đúng như vậy, điều đó có nghĩa là công trình chế tác bằng gỗ nguyên thủy được chia sẻ ở phía bắc, nhưng sau đó, người Trung Hoa đã phát triển cò bắn (trigger) nỏ bằng kim loại (có niên đại khảo cổ sớm nhất là vào thế kỷ thứ bảy trước Công Nguyên). sau đó được chia sẻ trở lại miền nam. Sự ra đời của chiếc nỏ kim loại của nhân vật An Dương Vương là một phần quan trọng trong sự tiếp xúc đầu tiên mang tính huyền thoại lịch sử của tổ tiên người Trung Hoa và người Việt. Không có cách đơn giản nào để xác minh một cách chắc chắn tuyệt đối cả hai quan điểm vào thời điểm này, và các từ mượn khác được đề xuất vào tiếng Hán đều có những đề xuất thách thức không kém.

59. Trong thế kỷ XX, một trong những thách thức trong việc nghiên cứu lịch sử ngôn ngữ của tiếng Tày và tiếng Việt là việc xác định và phân biệt các từ cùng vay mượn chung từ tiếng Hán. Vào giữa thế kỷ 20, tác giả Haudricourt (1954) đã bác bỏ khẳng định của tác giả Maspero (1912) về mối liên hệ ngôn ngữ giữa tiếng Việt và tiếng Tày (tức là khẳng định rằng chúng thuộc cùng một ngữ hệ) bằng cách chứng minh âm điệu tiếng Việt phát triển như thế nào và bằng cách xác định nhiều từ mượn tiếng Hán được chia sẻ bởi cả tiếng Tày và tiếng Việt. Thực vậy, trong một số khẳng định về từ vay mượn tiếng Tày trong tiếng Việt (Maspero 1912; Nguyễn Tài Cẩn 1995, 322; Phạm Đức Dương 1992), nhiều khẳng định cũng là những từ mượn tiếng Hán ban sơ (Alves 2001; Alves 2006).

60. Dữ liệu đến từ sự kết hợp của nhiều công trình của các nhà nghiên cứu khác – với nhiều hạng mục đã phải loại trừ – và các hạng mục mà tôi đã xác định được. Những từ có trạng thái thấp hoặc bị loại trừ cũng được giữ lại để tham khảo, nhưng vì mục đích chức năng, chỉ những từ có

độ chắc chắn trung bình hoặc cao (dựa trên sự kết hợp giữa âm vị học, ngữ nghĩa, và các đặc điểm văn hóa xã hội) đã được sử dụng. Tôi tin rằng phần lớn trong số hàng trăm mục trong cơ sở dữ liệu của tôi là những từ mượn thực sự của tiếng Trung Hoa thời kỳ ban sơ (với vài trăm sự loại trừ), nhưng theo thống kê, ít nhất một tỷ lệ phần trăm có thể là những từ tương tự ngẫu nhiên. Tuy nhiên, có vài trăm mục từ đủ dùng cho mục đích chung và để đưa ra những tuyên bố trong bài viết này.

61. Ví dụ, tác giả Manaimovibol (1975) liệt kê vài trăm từ mượn tiếng Trung Hoa tiềm năng trong tiếng Tày. Tuy nhiên, các thời kỳ của những khoản vay từ vựng này biến đổi, có thể là vài thế kỷ sau thời kỳ Tiếng Tày Nguyên Mẫu (Proto-Tai).

62. Pittayaporn (2014).

63. Toàn bộ dữ liệu Mường Bi đều được lấy từ từ điển của Nguyễn Văn Khang và các cộng sự. (2005).

64. Trong khi hầu hết các từ Mường đều có những đặc điểm cho thấy giai đoạn vay mượn từ tiếng Hán sớm hơn, sẽ cần có nghiên cứu bổ sung để xác định xem từ nào, nếu có, trong số những từ này có thể là từ mượn gần đây hơn từ tiếng Việt.

65. Tham chiếu Alves (2016) về các từ tiếng Hán ban sơ và chứng thực bằng cơ văn bản và khảo cổ học.

66. Sự tái thiết tiếng Trung Hoa thời Hậu Hán này là của Schuessler (2007, 254).

67. Tác giả Kim (2015) thảo luận vấn đề này liên quan đến di chỉ Cổ Loa ở nhiều chỗ trong cuốn sách của ông.

68. Churchman (2016).

69. Cùng nơi đã dẫn trên, 29.

70. Cùng nơi đã dẫn trên, 3.

71. Cùng nơi đã dẫn trên, 30, 62, v.v.

72. Kim và các cộng sự. (2010).

73. Alves (2016), 287.

74. Trong bản tái tạo tiếng Hán cổ của Baxter và Sagart (2014), ký hiệu *-j cuối cùng được đặt trong ngoặc để biểu thị mức độ không chắc chắn liên quan đến hình thức âm vị học chính xác và những gì nó có thể đã trở thành ở những thời điểm khác nhau. Nó có thể là âm cuối cùng *-j tại thời điểm vay mượn, giống như trong nhánh Proto-Min của tiếng Hán (Sinitic) (Baxter và Sagart 2014, 268), nhưng nó cũng có thể là *-r cuối cùng (cùng nơi dẫn trên, 252). Nếu là trường hợp kể sau thì nó vẫn trở thành -j trong tiếng Việt, như âm *-r và *-l trong ngữ chi Việt Nguyên Mẫu (Proto-Vietic) trở thành -j trong tiếng Việt (Alves 2017).

75. Từ này không thể tái tạo được trong tiếng Việt Nguyên Mẫu (Proto-Vietic) vì các nhóm ngữ chi Việt miền nam không có cách kiến trúc này và do đó cũng không có từ này.

76. Cũng có trường hợp là âm [-j] cuối cùng không được thấy trong từ này trong bất kỳ dạng tiếng Trung Hoa hiện đại nào (chương trình cơ sở dữ liệu Starling và Cơ sở dữ liệu ngữ văn Xiaoxuetang mở rộng), dường như tất cả đều đã đánh mất đặc điểm đó. ECL này rõ ràng có một đặc điểm rất cổ xưa của từ này.

77. Một chi tiết khó xác định là quy mô dân số Trung Quốc ở các vùng khác nhau của Việt Nam dưới thời nhà Hán. Phải chăng các cộng đồng lớn nhất gần khu vực Cổ Loa là một địa điểm hợp lý như một nơi cư trú đô thị nguyên thủy, trong khi lại có ít cộng đồng hơn ở phía nam Đồng Bằng Sông Hồng (ĐBSH)? Tình trạng như vậy sẽ giải thích cho số lượng lớn các từ vay mượn tiếng Hán sơ khai trong tiếng Việt-Mường nhưng lại có số lượng nhỏ hơn nhiều trong các ngôn ngữ Việt Miền Nam (Southern Vietic).

78. Nguyễn Văn Lợi (1993), 42.

79. Ferlus (2006); Ferlus (2009a), 3–4; Chamberlain (2016).

80. Chamberlain (1998); Chamberlain (2016).

81. Trần (2008).

82. Ví dụ: Maspero (1912); Phạm (1962); Nguyễn Tài Căn (1995), 322; Nguyễn Ngọc San (2005).

83. Shorto (2006).

84. Alves (2015).

85. Một giả thuyết khác là các nhóm Nam Á là những người săn bắn hái lượm và tiếp nhận nền nông nghiệp từ một số nhóm từ miền nam Trung Hoa. Nếu vậy, câu hỏi đặt ra là những nhà nông

thời kỳ đồ đá mới đến từ miền nam Trung Hoa là ai. Ngoại trừ các ngôn ngữ hiện không tồn tại đã tuyệt chủng trong thời tiền sử – một mệnh đề không thể bác bỏ – một giả thuyết hợp lý là những nhóm đó là những người nói tiếng Tày Nguyên Mẫu thuộc ngữ hệ Kadai (Proto-Tai-Kadai). Mặc dù chắc chắn là có khả năng xảy ra, nhưng cho đến nay, các nghiên cứu về ngôn ngữ học, khảo cổ học hay khảo cổ học đều không đề xuất điều này. Trong nghiên cứu này, việc các nhóm Nam Á ở vùng Đồng Bằng Sông Hồng (ĐBSH) là những người săn bắn hái lượm hay những người làm nông nghiệp không ảnh hưởng đến điểm chính là các nhóm Nam Á đã ở vùng Đồng Bằng Sông Hồng (ĐBSH) 4000 Trước Thời Hiện Đại (BP). Giả thuyết thay thế sẽ cần được khám phá thêm để ủng hộ hoặc bác bỏ. Ở điểm này, nghiên cứu này chấp nhận quan điểm cho rằng người nói tiếng Nam Á (Austroasiatic) đã đưa nông nghiệp vào Đông Nam Á Đại Lục (MSEA)./-

Tài Liệu Tham Khảo

Aciri, Andrea, Roger Blench, và Alexandra Landmann (đồng biên tập), *Spirits and Ships: Cultural Transfers in Early Monsoon Asia* (Singapore: ISEAS – Yusof Ishak Institute, 2017).

Aikhenvald, Alexandra Y. và R. M. W. Dixon (đồng biên tập), *Grammars in Contact: A Cross-linguistic Typology* (Oxford, UK: Oxford University Press, 2006).

Alves, Mark J., “The Current Status of Vietnamese Genetic Linguistic Studies”, *The Fifth International Symposium on Languages and Linguistics, Hochiminh City* (Ho Chi Minh City: Vietnam National University, Ho Chi Minh City University of Social Sciences and Humanities, 2000), 6–17.

Alves, Mark J., “What’s So Chinese About Vietnamese?”, trong sách biên tập bởi Thurgood (2001), 221–242.

Alves, Mark, “Linguistic Research on the Origins of the Vietnamese Language: An Overview”, *Journal of Vietnamese Studies* 1:1–2 (2006), 104–130.

Alves, Mark J., “A Look at North-Central Vietnamese”, trong sách biên tập bởi Wayland và cộng sự (2007), 1–7.

Alves, Mark J., “Historical Notes on Words for Knives, Swords, and Other Metal Implements in Early Southern China and Mainland Southeast Asia”, *Mon-Khmer Studies* 44 (2015), 39–56.

Alves, Mark J., “Identifying Early Sino-Vietnamese Vocabulary via Linguistic, Historical, Archaeological, and Ethnological Data”, *Bulletin of Chinese Linguistics* 9 (2016), 264–295.

Alves, Mark J., “Does Vietnamese Have Evidence for OC *-r?”, *Cahiers de Linguistique Asie Orientale* 45 (2017), 151–173.

Alves, Mark J., “Data from Multiple Disciplines Connecting Vietic with the Dong Son Culture”, bản tin cho cuộc nói chuyện tại buổi hội thảo “Contact Zones & Colonialism in China’s South”, The Pennsylvania State University, USA, May 10–12, 2019a.
DOI: 10.13140/RG.2.2.32110.05446.

Alves, Mark J., “Dữ Liệu Liên Ngành: Chi Vietic Kết Nối Với Văn Hóa Đông Sơn” [Data from Multiple Disciplines Connecting Vietic with the Dong Son Culture], bản tin cho cuộc nói chuyện tại buổi hội thảo tại Đại Học Quốc Gia Hà Nội, Ban Khoa Học Xã Hội và Nhân Văn (the National University of Hanoi), Department of Social Sciences and Humanities,) Vietnam. May 5th, 2019b. DOI: 10.13140/RG.2.2.26690.22720.167

Alves, Mark J., “Linguistic Influence of Chinese in Southeast Asia”, trong sách đồng biên tập bởi Sidwell và Jenny (2021), 649–672.

Ballard, William L., “Aspects of the Linguistic History of South China”, *Asian Perspectives* 24:2 (1981), 163–185.

Baxter, William H. and Laurent Sagart, *Old Chinese: A New Reconstruction* (Oxford: Oxford University Press, 2014).

Bellwood, Peter, *First Farmers: The Origins of Agricultural Societies* (Hoboken, NJ: Wiley-Blackwell, 2005).

Blench, Roger, “Ethnographic and Archaeological Correlates for a Mainland Southeast Asia Linguistic Area”, trong sách biên tập bởi Acri và cộng sự (2017), 207–238.

Blust, Robert, “Beyond the Austronesian Homeland: The Austric Hypothesis and Its Implications for Archaeology”, trong sách biên tập bởi Goodenough (1996), 117–140.

Brindley, Erica Fox, *Ancient China and the Yue: Perceptions and Identities on the Southern Frontier, c. 400 BCE–50 CE* (Cambridge: Cambridge University Press, 2015).

Brunelle, Marc and James Kirby, “Tone and Phonation in Southeast Asian Languages”, *Language and Linguistics Compass* 10:4 (2016), 191–207.

Burusphat, Somsonge (biên tập), *Papers from the Eleventh Annual Conference of the Southeast Asian Linguistics Society 2001* (Phoenix: Arizona State University Programme for Southeast Asian Studies Monograph Series Press, 2004).

Chamberlain, James R., “The Origin of the Sek: Implications for Tai and Vietnamese History”, *Journal of the Siam Society* 86:1–2 (1998), 27–48.

Chamberlain, James R., “Kra-Dai and the Proto-History of South China and Vietnam”, *Journal of the Siam Society* 104 (2016), 27–77.

Chamberlain, James R., *A Kri-Mol (Vietic) Bestiary: Prolegomena to the Study of Ethnozoology in the Northern Annamites* (Kyoto Working Papers on Area Studies No. 133, Kyōto: Kyōto University, 2018).

Churchman, Catherine, *The People Between the Rivers: The Rise and Fall of a Bronze Drum Culture, 200–750 CE* (Lanham: Rowman & Littlefield Publishers, 2016).

Corny, Julien, và cộng sự, “Dental Phenotypic Shape Variation Supports a Multiple Dispersal Model for Anatomically Modern Humans in Southeast Asia”, *Journal of Human Evolution* 112 (2017), 41–56.

Đào Duy Anh, *Lịch Sử Cổ Đại Việt Nam* [History of Ancient Vietnam] (Hà Nội, Việt Nam: Nhà Xuất Bản Văn Hóa Thông Tin, 2005).

DeLancey, Scott, “The Origins of Sinitic”, trong sách biên tập bởi Jing-Schmidt (2013), 73–100.

Diffloth, Gerard, “The Lexical Evidence for Austric, So Far”, *Oceanic Linguistics* 33:2 (1994), 309–321.

Diffloth, Gerard, *The Contribution of Linguistic Palaeontology to the Homeland of Austro-Asiatic*, trong sách biên tập bởi Sagart và cộng sự (2005), 77–80.

Edmondson, Jerold A. và David B. Solnit (đồng biên tập), *Comparative Kadai: Linguistic Studies beyond Tai* (Summer Institute of Linguistics Publications in Linguistics, 168 Alves crossroads 19 (2020) 138–174 86. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington, 1988).

Enfield, N. J. (biên tập), *Enfield Dynamics of Human Diversity* (Canberra: Pacific Linguistics, School of Culture, History and Language, College of Asia and the Pacific, Australian National University, 2011).

Ferlus, Michel, “Sur L’Origine Géographique des Langues Việt-Mường”, *Mon-Khmer Studies* 18–19 (1989–1990), 52–59.

Ferlus, Michel, “The Origin of Tones in Viet-Muong”, trong sách biên tập bởi Burusphat (2004), 297–313.

Ferlus, Michel, “Sur L’Origine de Quelques Ethnonymes: Tai/Thai, Li/Hlai, Yi, Gelao, Lao (Chine du Sud, ASE)”, *XXèmes Journées de Linguistique Asie Orientale*, Jun 2006, Paris, France. <https://halshs.archives-ouvertes.fr/halshs-00932875>.

Ferlus, Michel, “Formation of Ethnonyms in Southeast Asia”, 42nd International Conference on Sino-Tibetan Languages and Linguistics, Nov 2009a, Chiang Mai, Thailand. <https://halshs.archives-ouvertes.fr/halshs-01182596>.

Ferlus, Michel, “A Layer of Dongsonian Vocabulary in Vietnamese”, *Journal of the Southeast Asian Linguistics Society* 1 (2009b), 95–110.

Ferlus, Michel, “Arem, a Vietic language”, *Mon-Khmer Studies* 43 (2014), 1–15.

Gong Xun, “Chinese Loans in Old Vietnamese with a Sesqui Syllabic Phonology”, *Journal of Language Relationship* 17.1–2 (2019), 55–72.

Goodenough, W. H. (ed.), Prehistoric Settlement of the Pacific, *Transactions of the American Philosophical Society* 86 (1996), 117–140.

Haudricourt, André G., “Comment Reconstruire le Chinois Archaïque”, *Word* 10:2–3 (1954), 351–364.

Higham, Charles, *Early Mainland Southeast Asia: From First Humans to Angkor* (Bangkok: River Books Press, 2014).

Higham, Charles F. W., “First Farmers in Mainland Southeast Asia”, *Journal of Indo-Pacific Archaeology* 41 (2017), 13–21.

Higham, Charles F. W., “The Neolithic Occupation of Southeast Asia”, trong sách biên tập bởi Sidwell và Jenny (2021), 21–33.

Higham, Charles, Thomas Higham, Roberto Ciarla, Katerina Douka, Ampham Kijngam, and Fiorella Rispoli, “The Origins of the Bronze Age of Southeast Asia”, *Journal of World Prehistory* 4:4 (2011), 227–274.

Hoàng Thị Châu, [*Tiếng Việt Trên Các Miền Đất Nước (Phương Ngữ Học)*] [Vietnamese in the Regions of the Country (Dialectology)] (Hanoi: Nhà Xuất Bản Khoa Học và Xã Hội, 1989).

Hudak, Thomas John, *William J. Gedney's Comparative Tai Source Book* (Honolulu: University of Hawai'i Press, 2008).

Jenny, Mathias và Paul Sidwell (đồng biên tập), *The Handbook of Austroasiatic Languages* (Leiden: Brill, 2015).

Jing-Schmidt, Zhuo (biên tập), *Increased Empiricism: Recent Advances in Chinese Linguistics* (Amsterdam: John Benjamins Publishing Company, 2013).

Kiernan, Ben, *Vietnam: A History from Earliest Times to the Present* (Oxford: Oxford University Press, 2017).

Kim, Nam, *The Origins of Ancient Vietnam* (Oxford Studies in the Archaeology of Ancient States, Oxford University Press, 2015).

Kim, Nam C., Lai Van Toi, và Trinh Hoang Hiep, “Co Loa: An Investigation of Vietnam's Ancient Capital”, *Antiquity* 84 (2010), 1011–1027.

Koenigswald, G. H. R., “Evidence of a Prehistoric Australomelanesoid Population in Malaya and Indonesia”, *Southwestern Journal of Anthropology* 8:1 (1952), 92–96.

Li Fang-Kuei, *A Handbook of Comparative Tai* (Honolulu: University Press of Hawaii, 1977).

Lipson, Mark et al., “Ancient Genomes Document Multiple Waves of Migration in Southeast Asian Prehistory”, *Science* 361:6397 (2018), 92–95.

Logan, J. R., “Ethnology of the Indo-Pacific Islands”, *Journal of the Indian Archipelago and Eastern Asia* VI (1852), VII (1853), VIII (1854), IX (1855).

Manomaivibool, Prapin, “A Study of Sino-Thai Lexical Correspondences”, Luận án Tiến Sĩ, University of Washington, 1975.

Maspero, Henri, “Études sur la phonétique historique de la langue annamite: Les initiales”, *Bulletin de l’École française d’Extrême-Orient* 12 (1912), 1–127.

Matsumura Hirofumi và cộng sự, “Cranio-metrics Reveal ‘Two Layers’ of Prehistoric Human Dispersal in Eastern Eurasia”, *Scientific Reports* 9:1451 (2019), 1–12.

Matsumura Hirofumi và cộng sự., “Morphometric Affinity of the Late Neolithic Human Remains from Man Bac, Ninh Binh Province, Vietnam: Key Skeletons with Which to Debate the ‘Two Layer’ Hypothesis”, *Anthropological Science* 116:2 (2008), 135–148.

Matsumura Hirofumi et al., “Cranio-morphometric and aDNA Corroboration of the Austronesian Dispersal Model in Ancient Island Southeast Asia: Support from Gua Harimau, Indonesia”, *PlosOne* 13.6: e0198689 (2018). <https://doi.org/10.1371/journal.pone.0198689>.

Matthews, Stephen, “Cantonese Grammar in Areal Perspective”, in Aikhenvald and Dixon (2006), 220–236.

McColl, Hugh, và cộng sự “The Prehistoric Peopling of Southeast Asia”. *Science* 361 (2018), 88–92.

McWhorter, John, *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars* (Oxford: Oxford University Press, 2007).

Nguyễn Ngọc San, *Tìm Hiểu Tiếng Việt Lịch Sử* [Exploring the History of the Vietnamese Language] (Hanoi: Nhà Xuất Bản Đại i Học Sư Phạm, 2003).

Nguyễn Tài Cẩn, *Giáo Trình Lịch Sử Ngữ Âm Tiếng Việt* [A Text on Vietnamese Historical Phonology] (Hanoi: Nhà Xuất Bản Giáo Dục, 1995).

Nguyễn Văn Lợi, “Tộc Danh Của Một Số Dân Tộc Ở Nam Trung Quốc Và Đông Nam Á. Vấn Đề Tên Gọi Giao Chi” [Ethonyms of some ethnicities in southern China and Southeast Asia: the question of the name Giao Chi], *Việt Nam – Những Vấn Đề Ngôn Ngữ Và Văn Hóa, Hội Ngôn Ngữ Học Việt Nam, Hà Nội, Đại học Ngoại ngữ Hà Nội* [Linguistic and Cultural Questions, Conference of Linguistics in Hanoi, Hanoi University of Foreign Languages] (1993), 35–44.

Norman, Jerry và Tsu-lin Mei, “The Austroasiatics in Ancient South China: Some Lexical Evidence”, *Monumenta Serica* 32 (1976), 274–301.

O’Harrow, Stephen, “From Co-Loa to the Trung-Sisters’ Revolt: Viet-Nam as the Chinese Found It”, *Asian Perspectives* 22:2 (1979), 140–164.

Phạm Đức Dương, “Cội Nguồn Mô Hình Văn Hóa – Xã Hội Lúa Nước Của Người Việt Qua Cứ Liệu Ngôn Ngữ” [The Origin of the Wet-Rice Culture of Vietnamese People through Linguistic Materials]. *Nghiên Cứu Lịch Sử* 5 (1982), 43–52.

Phan, John, “Lacquered Words: The Evolution of Vietnamese under Sinitic Influences from the 1st Century BCE through the 17th Century CE”, Luận án Tiến Sĩ, Cornell University, 2013.

Pittayaporn, Pittayawat, “The Phonology of Proto-Tai”, Luận án Tiến Sĩ, Cornell University, 2009.

Pittayaporn, Pittayawat, “Layers of Chinese Loanwords in Proto-Southwestern Tai as Evidence for the Dating of the Spread of Southwestern Tai”, *MANUSYA: Journal of Humanities, Special Issue* 20 (2014), 47–68.

Ratliff, Martha, *Hmong-Mien Language History* (Canberra: Pacific Linguistics, 2010).

Rau, Felix và Paul Sidwell, “The Munda Maritime Hypothesis”, *Journal of the Southeast Asian Linguistics Society* 12.2 (2019), 35–57.

Reid, Lawrence A., “Morphological Evidence for Austric”, *Oceanic Linguistics* 33.2 (1994a), 323–344.

Reid, Lawrence A., “Possible Non-Austronesian Lexical Elements in Philippine Negrito Languages”, *Oceanic Linguistics* 33.1 (1994b), 37–72.

Reid, Lawrence A., “The Current Status of Austric: A Review and Evaluation of the Lexical and Morphosyntactic Evidence”, trong sách biên tập bởi Sagart và cộng sự (2005), 132–160.

Sagart, Laurent, “The Expansion of Setaria Farmers in East Asia: a Linguistic and Archaeological Model”, trong sách biên tập bởi Sanchez-Mazas và cộng sự (2008), 133–157.

Sagart, Laurent, Roger Blench, và Alicia Sanchez-Mazas, *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics* (New York: Routledge Curzon, 2005).

Sanchez-Mazas, Alicia và cộng sự (đồng biên tập), *Past Human Migrations in East Asia* (New York: Taylor and Francis, 2008).

Schmidt, Wilhelm, “Die Mon – Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens”, *Archiv für Anthropologie, Braunschweig, new series* 5 (1906), 59–109

Shimizu Masaaki, “Khảo Sát Sơ Lược Về Cấu Trúc Âm Tiết Tiếng Việt Vào Thế Kỷ XIV–XV Qua Hai Cú Liệu Chữ Nôm” [A Brief Survey of Vietnamese Syllable Structure in the XIV–XVth Centuries in Two Nom Scripts], trong sách của Trung Tâm Khoa Học Xã Hội Và Nhân Văn Quốc Gia (2000), 252–265.

Shimizu Masaaki, “A Reconstruction of Ancient Vietnamese Initials Using Chữ Nôm Materials”, *Kokuritsu kokugo kenkyūjo ronshū* 国立国語研究所論集 (NINJAL Research Papers) 9 (2015), 135–158.

Shorto, Harry L., *A Mon-Khmer Comparative Dictionary* (Canberra: Pacific Linguistics, 2006).

Sidwell, Paul, *Classifying the Austroasiatic Languages: History and State of the Art* (Lincom Europa, 2009).

Sidwell, Paul, “Austroasiatic Dispersal: the AA ‘Water-World’ Extended”, phim ảnh trình bày tại SEALS 30 (2021). DOI: 10.5281/zenodo.5117960.

Sidwell, Paul and Roger Blench, “The Austroasiatic Urheimat: The Southeastern Riverine Hypothesis”, in Enfield (2011), 317–345.

Sidwell, Paul and Mathias Jenny, *The Languages and Linguistics of Mainland Southeast Asia: A Comprehensive Guide* (Boston: De Gruyter Mouton, 2021).

Sidwell, Paul and Felix Rau, “Austroasiatic Comparative-Historical Reconstruction: An Overview”, trong sách biên tập bởi Mathias và Sidwell (2015), 221–363.

The Starling Database Program, “Search for Data in: Chinese Dialects”.
<https://starling.rinet.ru/cgi-bin/query.cgi?basename=%5Cdata%5Cchina%5Cdoc&root=config>
(accessed 1 Jun 2021).

Starosta, Stanley, “Proto-East Asian and the Origin and Dispersal of the Languages of East and Southeast Asia and the Pacific”, trong sách biên tập bởi Sagart và cộng sự (2005), 182–197.

Taylor, Keith, *Birth of Vietnam* (Berkeley: University of California Press, 1980).

Taylor, Keith, *A History of the Vietnamese* (New York: Cambridge University Press, 2013).

Thurgood, Graham, “Notes on the Reconstruction of Proto-Kam – Sui”, trong sách đồng biên tập bởi Edmondson và Solnit (1988), 179–218.

Thurgood, Graham W. (biên tập), *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society* (Tempe: Arizona State University, Program for Southeast Asian Studies, 2001).

Trần Trí Dõi, “Tên Gọi Của Sông Hồng: Dấu Tí ch Biểu Hiện Nét Đa Dạng Văn Hoá Trong Lịch Sử Người Việt” [Names for the Red River: traces of cultural diversity in Vietnamese history], Paper of the 3rd International Vietnamese Conference, Hanoi, 4–7 April, 2008.

Trần Trí Dõi, *Giáo Trình Lịch Sử Tiếng Việt* [A textbook of the history of Vietnamese] (Nhà Xuất Bản Giáo Dục Việt Nam, 2011).

Trần Trí Dõi, “Ngôn Ngữ Của Cư Dân Văn Hóa Đông Sơn Ở Việt Nam” [The language of people of the Dong Son culture in Vietnam]. *Ngôn Ngữ* 8–9 (2019), 66–82.

Trung Tâm Khoa Học Xã Hội Và Nhân Văn Quốc Gia, Đại Học Quốc Gia Hà Nội [National Center of Social Sciences and Humanities, Hanoi National University], *Việt Nam Học – Kỷ Yếu Hội Thảo Quốc Tế Lần Thứ Nhất* [Vietnamese Studies: Proceedings of the First International Conference] (Hà Nội: Thế Giới, 2000).

van Driem, George, “The Ethnolinguistic Identity of the Domesticators of Asian Rice”, *Comptes Rendus Palevol* (Mainland and insular Asia: Current debates about first settlements) 11:2–3 (2011), 117–132.

Wayland, Ratreه và cộng sự., *SEALS XII Papers from the 12th Annual Meeting of the Southeast Asian Linguistics Society 2002* (Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University, 2007).

Xiaoxuetang wenzixue ziliaoku 小學堂文字學資料庫 [The Xiaoxue Philological Database].

<http://xiaoxue.ijs.sinica.edu.tw/> (tiếp cận ngày 1 Oct 2019).

Phụ Lục: Về Việc Lượng Định Các Từ Mượn Tiềm Năng

Bài viết này trình bày các dữ liệu liên quan đến các từ mượn từ tiếng Hán (Sinitic) vào tiếng Tày và ngữ chi Tiếng Việt (Vietic), và giữa tiếng Tày và các biến thể của tiếng Việt phát triển thành tiếng Việt-Mường và cuối cùng thành tiếng Việt hiện đại. Phụ lục này xem xét các thực tiễn và các giả định cơ bản về lịch sử âm vị học và phương pháp so sánh để người đọc có thể biết cách

thức tôi đã xác định các từ mượn tiềm năng và sau đó chọn lọc hoặc loại trừ chúng. Chi tiết cụ thể của các đặc điểm lịch sử âm vị học vượt quá những gì có thể được tóm tắt ở đây.

– Thu thập đầy đủ dữ liệu từ vựng. Lý tưởng nhất là một lượng lớn dữ liệu được thu thập, mặc dù số lượng và sự đa dạng của các loại nhất thiết phải thay đổi do phạm vi phức tạp của các yếu tố. Càng có nhiều dữ liệu thì càng có nhiều cơ hội để xác định các mô hình âm vị học làm gia tăng tính chắc chắn của việc nhận dạng, và một lần nữa, để loại trừ những từ được ghi nhận không theo mô hình. Các nguồn dữ liệu như vậy biến đổi sâu xa về chi tiết, nhưng ngày nay, có rất nhiều nguồn tài liệu số hóa hỗ trợ công việc này. Các nguồn tài liệu như thế được liệt kê dưới đây.

- Xác định các mô hình âm vị học. Xác định càng nhiều mô hình càng tốt các sự tương ứng âm thanh giữa các từ của hai ngôn ngữ cứu xét (ví dụ: các sự tương ứng của các phụ âm khởi đầu của hai từ trong hai ngôn ngữ, ví dụ: âm /b/ trong cả hai ngôn ngữ hoặc âm /b/ trong một ngôn ngữ và âm /p/ trong ngôn ngữ kia, v.v.). Số lượng các mô hình được xác định càng lớn thì dữ liệu từ vựng lựa chọn có thể được lượng định càng hiệu quả hơn. Liên quan đến vấn đề đầu tiên, số lượng từ để so sánh và tìm kiếm các mô hình càng lớn thì những xác quyết về các mô hình âm vị học có thể càng vững mạnh hơn.

– Tập chú vào ngôn ngữ nguyên mẫu. Thay vì lựa chọn ngẫu nhiên các từ từ các ngôn ngữ hiện đại để so sánh, người ta phải sử dụng các từ được tái thiết trong, ví dụ: tiếng Hán cổ (old Chinese) và tiếng Hán thời trung cổ (Middle Chinese), tiếng Tày nguyên mẫu (Proto-Tai) và tiếng Việt nguyên mẫu (Proto-Vietic). Một ngôn ngữ nguyên mẫu được tái thiết liên can đến sự so sánh nhiều ngôn ngữ và đã sẵn bao gồm một khối dữ liệu mạnh mẽ. Điều quan trọng là một hình thức tái thiết, được phóng chiếu quay ngược về quá khứ, có thể chỉ ra những đặc điểm âm vị học của các ngôn ngữ tại thời điểm có sự tiếp xúc ngôn ngữ. Cũng có thể hữu ích khi so sánh nhiều ngôn ngữ hiện đại (ví dụ: các danh sách so sánh các phương ngữ tiếng Hán, các ngôn ngữ tiếng Tày hoặc tiếng Mường) nếu chúng ta không có sự tái thiết, nhưng một lần nữa, điều này liên can đến nhiều ngôn ngữ và không chỉ là một ngôn ngữ được lựa chọn ngẫu nhiên. Một sai lầm mà những người không chuyên môn mắc phải là nhìn vào các từ chỉ trong hai ngôn ngữ (ví dụ: chỉ so sánh tiếng Quan Thoại hoặc tiếng Quảng Đông với tiếng Việt) và đưa ra các giả định về sự tương đồng, trong khi xem xét việc tái tạo và so sánh dữ liệu từ hàng chục ngôn ngữ là công việc cần thiết tối thiểu để đưa ra những tuyên bố xác đáng về lịch sử ngôn ngữ.

– So sánh với các ngôn ngữ khác và các nhóm ngôn ngữ khác. So sánh các từ trong chỉ có hai ngôn ngữ, hoặc thậm chí chỉ trong hai ngôn ngữ nguyên mẫu, có thể dẫn đến những giả định sai lầm. Ở Đông Nam Á Đại Lục (MSEA) và Đông Nam Á hải đảo (insular Southeast Asia), các ngôn ngữ đã được tiếp xúc trong nhiều thế kỷ và thường là hàng thiên niên kỷ, và ngôn từ đã có thời gian được truyền bá rộng rãi. Chính vì thế, các nhà nghiên cứu ngữ học phải sử dụng nhiều cơ sở dữ liệu từ vựng khác nhau và các nguồn từ vựng khác để xem xét tất cả các nhóm ngôn

ngữ trong một khu vực. Ở khu vực Đông Nam Á mở rộng về phía tây, phía bắc và phía đông, điều này có nghĩa là (a) không chỉ chữ Hán⁸⁶ mà còn cả chữ Hán-Tạng;⁸⁷ (b) Tiếng Tày (Tai)⁸⁸ và các nhóm liên quan như Kam-Sui,⁸⁹ Hlai,⁹⁰ và Kra;⁹¹ (c) Tiếng Việt (Vietic) và các nhánh khác của ngôn ngữ Nam Á (Austroasiatic);⁹² (d) Hmong-Mien;⁹³ và (e) Nam Đảo (Austronesian).⁹⁴ Các sự so sánh như thế không chỉ cho phép đánh giá các từ có thể có liên quan mà còn xác định vị trí địa lý của các từ vào lúc hiện diện hay vắng mặt của chúng trong các khu vực, điều có thể cung cấp những hiểu biết sâu sắc về sự tiếp xúc văn hóa xã hội, hoặc sự thiếu sót của nó.

– Nhận thức được những điểm tương đồng ngẫu nhiên. Áp dụng tất cả những điều trên không thể ngăn chặn khả năng tính rằng hai từ có thể giống nhau do trùng hợp ngẫu nhiên. Một lần nữa, càng nhiều dữ liệu từ vựng thu thập được, đặc biệt là trong một lĩnh vực văn hóa hoặc một hệ thống ngữ nghĩa nào đó như họ hàng thân thuộc, chúng ta càng có thể giảm khả năng một từ đơn độc tương đồng với một từ khác bởi có sự tình cờ. Ngoài ra, từ tượng thanh (onomatopoeia) là một phương lược hình thành từ tiềm năng dẫn đến những từ trông giống nhau trong các ngữ hệ.

– Xem xét âm điệu là thước đo thời gian. Như đã lưu ý trong nghiên cứu này, sự xuất hiện của âm điệu tiếng Hán trong thiên niên kỷ thứ nhất sau Công Nguyên được thiết lập vững chắc một cách hợp lý thông qua sự tham khảo văn bản và nó có thể được sử dụng để đánh giá trình tự thời gian tương đối và lượng định mức độ sự chắc chắn của các từ vay mượn. Những từ có âm điệu loại B và C là những từ có âm điệu tốt nhất có thể cung cấp trình tự thời gian tương đối này. Tuy nhiên, những từ này sau đó có các phụ âm và nguyên âm có thể được sử dụng song hành để xác định các từ vay mượn trước đó. Chủ đề này rất phức tạp và không thể đi sâu trong bài nghiên cứu này. Những tài liệu tham khảo các nghiên cứu như thế được ghi nhận trong suốt bài viết này.

– Xem xét trọng lượng của bằng chứng. Lịch sử ngôn ngữ học đưa ra ít điều chắc chắn, nhưng đôi khi, trọng lượng của bằng chứng cung cấp các trình tự thời gian tương đối và các hàm ý về của các tình huống văn hóa xã hội, sự tiếp xúc và các sự thay đổi. Nghiên cứu này áp dụng nguyên tắc này bằng cách kết hợp dữ liệu từ vựng, âm vị học, các phạm trù ngữ nghĩa, bằng chứng văn bản lịch sử, dân tộc học và khảo cổ học.

Chú Thích

86 Ví dụ, những công trình tái thiết tiếng Hán cổ (Old Chinese) bởi Baxter và Sagart (2015) và đôi khi bởi Schuessler (2007).

87 Cơ Sở Dữ Liệu The STEDT database.

88 Li (1977); Pittayaporn (2009).

89 Thurgood (1988).

90 Norquest (2010).

91 Ostapirat (1999).

92 Shorto (2006), *the Mon-Khmer Etymological Dictionary, and the Munda Etymological Dictionary*. Cũng xem chú thích 53 để có những công trình tái thiết chú tâm đến ngữ chi tiếng Việt (Vietic).

93 Ratliff (2010).

94 Tham chiếu Blust and Trussel's (2010) Proto-Austronesia

Nguồn: **Matrk Alves**, The Đông Sơn Speech Community: Evidence for Vietic, *Crossroads* 19 (2020) 138–174.

© Koninklijke Brill NV, Leiden, 2021 | doi:10.1163/26662523-bja100092