

# Nguồn gốc những điểm dị biệt trong bảng chữ cái tiếng Việt\*

*André-Georges Haudricourt*

Nguyên tác: Haudricourt, A. G. (1949), “L’origine des particularités de l’alphabet vietnamien”, *Dân Việt Nam* (3), pp. 61-68.

## Lời giới thiệu cho bản dịch tiếng Việt<sup>1</sup>

Nghiên cứu về chữ Quốc ngữ sẽ thật thiếu sót nếu như không nghiên cứu lịch sử của nó. Trong giới nghiên cứu lịch sử chữ Quốc ngữ ở thế kỉ 20 (thế kỉ thiết lập vị thế “danh chính ngôn thuận” của chữ Quốc ngữ), ta không thể không nhắc tới A.G. Haudricourt (1911 – 1996) với bài báo nổi tiếng “*Nguồn gốc những điểm dị biệt trong bảng chữ cái tiếng Việt*”, đăng trên *Dân Việt Nam*, tập san bằng tiếng Pháp của Viện Viễn Đông Bác cổ Pháp (EFEO), năm 1949 tại Hà Nội. Vào thời điểm ra đời bài báo, A.G. Haudricourt hẳn còn chưa vang danh thế giới với giả thuyết về nguồn gốc tiếng Việt cũng như thanh điệu tiếng Việt (cf. Haudricourt 1953, 1954). Tuy nhiên, phải nói rằng, ngay từ thời điểm bấy giờ, như Alexis Michaud (tác giả của bản dịch tiếng Anh) nhận xét, A.G. Haudricourt đã chứng tỏ được “*niềm đam mê [...] trong việc phục dựng nguồn gốc lịch sử của những hiện tượng phức tạp, cũng như kỹ năng thu thập các bằng chứng từ nguồn tư liệu cực kì phong phú*”, mà cụ thể ở đây là với bộ chữ La-tinh của tiếng Việt.

Bộ chữ La-tinh hóa này, như A.G. Haudricourt đã chỉ ra, chịu ảnh hưởng sâu sắc từ các sản phẩm văn tự La-tinh hóa của các ngôn ngữ Roman vì chính tác giả của nó – những linh mục Công giáo dòng Tên. Cụ thể, giống như hầu hết các bộ chữ viết La-tinh của các ngôn ngữ Roman, chữ Quốc ngữ của tiếng Việt cũng chứa đựng những điểm bất đối xứng (từ của Haudricourt là “*dị biệt*” – “*peculiarity*”) giữa âm đọc và con chữ, tức trong nhiều trường hợp một chữ cái lại tương ứng với nhiều âm đọc, hoặc ngược lại, một âm đọc lại tương ứng với nhiều chữ cái. Haudricourt đã dành phần lớn bài báo để truy tìm nguồn gốc những điểm bất đối xứng này trong chữ Quốc ngữ. Phần còn lại của bài báo, tác giả đã cố gắng giải thích nguyên lai của những kí hiệu, dấu phụ đáng chú ý trong chữ Quốc ngữ, ví dụ như tại sao ta lại viết là Đ cho phụ âm /d/, nguồn gốc dấu nón của Ô - /o/ v.v. Những tìm tòi và khám phá của Haudricourt thực sự đáng để bất cứ nhà nghiên cứu tiếng Việt nói chung và chữ Quốc ngữ nói riêng nào tham khảo. Và trên thực tế, bài báo này vẫn được coi là “*một trong những nghiên cứu sâu sắc và có giá trị nhất về nguồn gốc văn*

---

\*Bản dịch năm 2020 này hiệu chỉnh, bổ sung bản dịch năm 2015 của cùng tác giả. Tham khảo bản dịch tiếng Anh của Alexis Michaud năm 2010 được lưu trữ tại kho tài liệu số hóa HAL của CCSD tại địa chỉ.

<sup>1</sup>Bản dịch tiếng Việt năm 2015 được in trong tập kỷ yếu Hội thảo *Chữ Quốc ngữ: Sự hình thành, phát triển và những đóng góp vào văn hóa Việt Nam* do Đại học Phú Yên và Đại học KHXH&NV TP. Hồ Chí Minh đồng tổ chức tại Phú Yên.

tự tiếng Việt hiện đại” (xem lời giới thiệu của Alexis Michaud). Đó cũng chính là động lực lớn nhất để dịch giả quyết tâm cho ra đời bản dịch tiếng Việt để đông đảo hơn các độc giả ở Việt Nam có điều kiện tiếp cận tới công trình giá trị này.

Dịch giả xin chân thành cảm ơn TS. Alexis Michaud (nghiên cứu viên CNRS và MICA, Đại học Bách Khoa Hà Nội) vì đã giới thiệu bài báo cũng như động viên dịch giả hoàn thành bản dịch. Bên cạnh đó, dịch giả cũng chân thành gửi lời cảm ơn đến NCS. Phạm Thị Kiều Ly (người đang viết luận án về quá trình La-tinh hóa tiếng Việt tại trường Đại học Paris 3 - Pháp) vì những góp ý giá trị của chị về bản dịch qua đối sánh với bản gốc tiếng Pháp.

Hà Nội, 30/4/2015

Dịch giả: Cao Thành Việt

### Lời giới thiệu của Alexis Michaud<sup>2</sup>

Đóng góp của André-Georges Haudricourt vào Đông Nam Á học trên tầm quốc tế là không thể phủ nhận (tham khảo cuốn Haudricourt Festschrift (Suriya, Thomas và Suwilai 1985)). Tuy nhiên, khá nhiều công trình của Haudricourt vẫn chưa được biên dịch sang Anh ngữ để đông đảo hơn các độc giả có thể tiếp cận được. Chính vì lẽ đó, một nhóm các học giả đến từ nhiều nước đang tiến hành biên dịch và sẽ sớm cho ra mắt một tuyển tập các công trình nghiên cứu quan trọng nhất của Haudricourt bằng tiếng Anh. Mục đích của nó không gì hơn là nhằm chia sẻ với cộng đồng học thuật dùng Anh ngữ các trước tác của Haudricourt mà phần nhiều trong số chúng là đề cập tới các vấn đề liên quan đến các ngôn ngữ Đông Nam Á, ngôn ngữ học và nhân học xã hội.

“*Nguồn gốc những điểm dị biệt trong bảng chữ cái tiếng Việt*” thực chất không phải là một trong những bài báo nổi tiếng nhất của Haudricourt, bởi vậy, nó dự kiến sẽ không được đưa vào tuyển tập tiếng Anh các trước tác của Haudricourt sắp được xuất bản tới đây. Tuy nhiên, cho tới tận ngày này, bài báo này vẫn được coi là một trong những nghiên cứu sâu sắc và có giá trị nhất về nguồn gốc văn tự tiếng Việt hiện đại. Nó truy tìm nguồn gốc những kí tự dị biệt có trong chính tả của ngôn ngữ này bằng cách lần theo dấu vết của chúng trong các hệ thống chính tả của các ngôn ngữ Rô-man gần gũi với chính Roman đã chế tác ra chúng. Đây là một công trình thể hiện niềm đam mê của Haudricourt trong việc phục dựng nguồn gốc lịch sử của những hiện tượng phức tạp, cũng như kỹ năng thu thập các bằng chứng từ nguồn tư liệu cực kì phong phú. Bài báo này cũng khác với hầu hết các công trình khác của Haudricourt ở việc ngay từ ban đầu, nó đã xác định cho mình đối tượng mục tiêu đó là độc giả phổ thông. Vì vậy, tính đại chúng của nó thể hiện rất rõ trong văn phong, cách dùng từ tránh những thuật ngữ quá chuyên môn. Công trình đã ra mắt độc giả trong số thứ ba, và cũng là số cuối cùng của tập san *Dân Việt Nam*, một tập san do Viện Viễn Đông Bác Cổ Pháp xuất bản trong những năm 1948, 1949 tại Việt Nam. Tài liệu nguyên gốc hiện tại rất khó tiếp cận, hơn nữa, đa số độc giả có khả năng quan tâm đến vấn đề này, ngày nay, rất có thể không có vốn tiếng Pháp tốt như đối tượng độc giả của Haudricourt vào năm 1949, thời điểm xuất bản bài báo. Vì vậy, hy vọng bản

<sup>2</sup>Tham khảo Haudricourt, A.G. (2010), “The origin of the peculiarities of the Vietnamese alphabet” (Alexis Michaud trans.), *Mon-Khmer Studies* (39), pp. 89–104.

dịch này có thể giúp những ai quan tâm đến tiếng Việt và vấn đề chữ viết nói chung tiếp cận được dễ dàng hơn tới tài liệu này.

Nhà xuất bản ở Việt Nam bấy giờ rõ ràng đã gặp rất nhiều khó khăn trong việc biên tập bài báo vốn sử dụng nhiều kí tự đặc biệt này. Vì thế, bản dịch lần này cũng sẽ đính chính luôn những lỗi soạn thảo có trong tài liệu gốc. Cụ thể, C được thay thế bằng G ở trang 64; *p'*, *t'*, *k'* lần lượt thay thế cho *pc*, *tc*, *kc* trong nguyên bản .v.v. Phiên âm trong bản dịch cũng được sửa chữa lại theo đúng hệ thống IPA hiện thời. Các chú thích của dịch giả được để trong ngoặc vuông hoặc để ở phần cước chú.

Dịch giả xin chân thành cảm ơn Michel Ferlus vì đã giới thiệu bài báo này, cảm ơn Boyd Michailovsky, Martine Mazaudon đã bỏ công nhuận sắc cho bản dịch và cảm ơn Jean-Michel Roynard đã giúp đỡ dịch giả trong phần minh họa.

### Tóm lược

Những nhà truyền giáo có công sáng tạo bảng chữ cái La-tinh cho tiếng Việt là những người Bồ Đào Nha, Ý và Pháp. Sản phẩm văn tự họ tạo ra kế thừa những đặc điểm dị biệt bắt nguồn từ chính các văn tự thuộc nhóm ngôn ngữ Roman này.

Tuy vậy, những phụ âm bật hơi (aspirated)<sup>3</sup> trong chữ Quốc ngữ bao gồm H, PH, TH, KH [IPA: /h/, /p<sup>h</sup>/, /t<sup>h</sup>/, /k<sup>h</sup>/] lại không tồn tại trong những ngôn ngữ Roman này. Trong những ngôn ngữ này, các tổ hợp con chữ PH, TH, KH chỉ xuất hiện trong những từ có nguồn gốc Hy Lạp, tương ứng với các phụ âm đầu bật hơi trong tiếng Hy Lạp là phi, theta, khi (φ, θ, χ) và theo đó mà những tổ hợp này được dùng để ghi các phụ âm tắc bật hơi trong tiếng Việt.

Hai kí tự ghi âm tắc mặt lưỡi C, G chỉ được sử dụng trước các nguyên âm /a/, /o/ và /u/. Lí do là trong những ngôn ngữ Roman thì đó là vị trí duy nhất mà những phụ âm trên còn bảo lưu được nét *chặn* (obstruent) từ ngôn ngữ mẹ – tiếng La-tinh; GHE, GHI thì được dùng với giá trị ngữ âm mà chúng có trong tiếng Ý; tổ hợp KE và KI thì cần được lí giải từ chữ cái K trong tiếng Hy Lạp (*kappa*, κ) và các thứ tiếng German.

Âm tắc môi-ngạc mềm (labiovelar) QU và GU thì được mượn từ văn tự tiếng Ý và tiếng La-tinh. Trong số những âm tiền ngạc (prepalatal), âm tắc vô thanh CH được lấy từ văn tự Bồ Đào Nha và Tây Ban Nha, những văn tự mượn tổ hợp đó từ *tiếng Pháp thượng cổ* (Old French) để ghi một âm môi, không có trong tiếng La-tinh.

Âm tắc hữu thanh D được dùng như một ký hiệu phiếm chỉ ghi lại một âm không có trong các ngôn ngữ châu Âu, nơi D luôn cùng với T tạo thành một cặp đối lập về tính thanh. Trong tiếng Việt, một chữ cái hoàn toàn mới là Đ được chế tác [dùng để ghi âm tắc lợi-tiền thanh hầu hóa (preglottalised alveolar stop): /d̥/], với nét ngang ở giữa giúp liên tưởng đến âm T gần gũi với nó.

Âm xát<sup>4</sup> hữu thanh GI mượn từ tiếng Bồ Đào Nha và tiếng Pháp (bấy giờ J vẫn chưa được sử dụng ở châu Âu).

Âm xát vô thanh X thì được mượn từ tiếng Bồ Đào Nha và các phương ngữ tiếng Tây Ban Nha miền Bắc: do trong những thứ tiếng này, âm S là một âm quặt lưỡi [IPA: /ɣ/] giống như trong tiếng Việt, trong khi ở các ngôn ngữ châu Âu khác, S chỉ là một âm xuýt-tiền (anterior sibilant) giống như trong tiếng Pháp [IPA: âm lợi /s/].

NH [kí hiệu âm mũi ngạc cứng /ɲ/] thì được mượn từ tiếng Bồ Đào Nha trong khi TR là một kí tự hoàn toàn không tìm thấy trong các thứ tiếng châu Âu [IPA: /t̪/].

Ô, Ê thì được mượn từ tiếng Bồ Đào Nha và đó là lí do tại sao cặp Ê, E không có được giá trị ngữ âm như vốn có trong tiếng Pháp. [Trong tiếng Pháp, Ê kí hiệu cho /ɛ/, E thay thế /e/, nhưng trong tiếng Việt tình hình lại ngược lại, Ê thay thế cho /e/ và E thay thế cho /ɛ/.]

Y được dùng trong chữ Quốc ngữ theo cách tương tự như trong văn tự tiếng Tây Ban Nha, nơi nó thay thế cho I ở vị trí giữa các nguyên âm hoặc ở cuối từ. Chữ cái Y có nguồn gốc từ bảng chữ cái Hy Lạp (*upsilon*, υ). O và U

<sup>3</sup>ND: Các thuật ngữ tiếng Việt trong bài được dùng theo Cao Xuân Hạo & Hoàng Dũng (2005).

<sup>4</sup>Thuật ngữ “*spirant*” được sử dụng trong bài báo này với nội hàm tương đương với “*fricative*” (“xát”). Xem chú thích tiếp theo để hiểu kĩ hơn về sự khác biệt có thể có giữa “*spirants*” và “*fricatives*”.

thì được sáng tạo để ghi hai nguyên âm của tiếng Việt không có trong các thứ tiếng châu Âu [IPA /ɤ/, /u/].

## Dẫn nhập

Hệ thống chữ viết “*alphabet*” là sản phẩm dùng để ghi lại cách phát âm của từ. Về nguyên tắc, mỗi chữ cái chỉ tương đương với một âm [một âm vị], và một chữ cái đã cho luôn luôn chỉ thay thế cho một và chỉ một âm; ngược lại, một âm đã cho chỉ có thể được ghi bởi một và chỉ một chữ cái. Tuy nhiên, sự lí tưởng này không hề được tìm thấy trong bất cứ một bộ chữ “*alphabet*” của một ngôn ngữ nào có lịch sử và truyền thống văn chương lâu bền. Bởi lẽ theo thời gian, hệ thống chữ viết thường vẫn giữ nguyên trong khi cách phát âm thì biến đổi. Các văn bản viết vẫn được truyền từ thế hệ này qua thế hệ khác mà không có sự sửa đổi (có chăng chỉ sửa đổi chút ít, và thế là “tụt hậu” so với diễn tiến của ngôn ngữ nói). Vì vậy, mối tương quan giữa một bên là chữ viết và một bên là âm đọc cũng thay đổi dần theo quá trình diễn tiến của ngôn ngữ nói. Cái tôi gọi là những điểm dị biệt của một bảng chữ cái chính là những sự bất tương đồng giữa chữ cái với cách phát âm: một chữ cái lại tương ứng với nhiều âm, hoặc ngược lại, một âm lại tương ứng với nhiều chữ cái. Những điểm dị biệt đó bắt nguồn từ diễn tiến lịch sử cách phát âm của chính ngôn ngữ đó.

Khi một ngôn ngữ được ghi lại bằng chữ viết lần đầu tiên, hệ thống văn tự dùng để ghi lại ngôn ngữ đó thông thường sẽ dựa trên một hệ thống văn tự sẵn có của một ngôn ngữ khác và theo đó mà kế thừa luôn những đặc điểm dị biệt từ văn tự điển mẫu. Vì vậy, do được xây dựng dựa trên hệ thống chính tả của một số ngôn ngữ Roman mà chữ Quốc ngữ của tiếng Việt cũng thừa hưởng luôn những điểm dị biệt có thể được kiến giải căn cứ trên diễn tiến ngữ âm của tiếng La-tinh ở châu Âu. Hãy nhớ lại rằng, tiếng La-tinh, thứ tiếng vào thế kỉ thứ 3 trước CN vẫn chỉ là thổ ngữ của thị dân Roma, đã lan tỏa ra khắp châu Âu cùng với sự bành trướng của đế chế Roman (cho đến thế kỉ thứ 5 CN) và chỉ để lại sau lưng một số ít khu vực không nói tiếng La-tinh, ví dụ như vùng phía Tây Pyrénées nơi ngôn ngữ bản địa là tiếng Basque vẫn còn được bảo lưu. Sau đó, những biến thể tiếng La-tinh cũng từ đó mà diễn tiến theo những cách khác nhau: tiếng La-tinh vùng Florence trở thành tiếng Ý, tiếng La-tinh vùng Paris trở thành tiếng Pháp và tiếng La-tinh vùng Burgos trở thành tiếng Tây Ban Nha. Mặc dù vậy, trên toàn cõi Roman, tiếng La-tinh cổ điển vẫn được coi là ngôn ngữ của tôn giáo và khoa học bất chấp nay không ai còn phát âm theo cách nguyên thủy của nó nữa. (Ngày nay, tiếng La-tinh cổ điển vẫn là ngôn ngữ của Giáo hội Công giáo Roman và các nhà nghiên cứu lịch sử tự nhiên vẫn dùng tiếng La-tinh để đặt tên các cây trồng và vật nuôi). Các ngôn ngữ như tiếng Pháp, tiếng Ý, tiếng Tây Ban Nha và tiếng Bồ Đào Nha vì vậy mà được gọi với những cái tên như “*các ngôn ngữ La-tinh*”, “*các ngôn ngữ Tân La-tinh*” hay “*các ngôn ngữ Roman*”. Bảng chữ cái tiếng La-tinh còn được dùng cho những ngôn ngữ khác của châu Âu (tiếng Celtic, tiếng German và tiếng Slavơ) nhờ học giả thành thạo việc phiên chuyển từ ngữ tiếng Hy Lạp sang tiếng La-tinh, tiếng Hy Lạp cũng là ngôn ngữ văn hóa của thời Cổ đại (Antiquity).

Trong bài báo này, tôi sẽ chỉ ra những nét dị biệt mà bảng chữ cái tiếng Việt [hiện nay gọi là Quốc ngữ] đã thừa hưởng từ các ngôn ngữ Roman. Bài báo cũng sẽ trả lời cho câu hỏi tại sao bảng chữ cái tiếng Việt lại khác với bảng chữ cái tiếng Pháp mặc dù tiếng Pháp cũng là một ngôn ngữ Roman và sự khác nhau giữa cả bảng chữ cái tiếng Việt, tiếng Pháp với bảng chữ cái ngữ âm của các nhà ngôn ngữ học cụ thể ra sao.

Thứ tự sắp xếp các đơn vị trong một bảng chữ cái có một lịch sử lâu đời. Chúng ta biết rằng bảng chữ cái “*alphabet*” là phát minh của người Phoenicia (Tây Bắc Palestine ngày nay) và từ đây lan tỏa đi khắp các xứ sở nói tiếng Hy Lạp. Ở Đông Hy Lạp, người ta đã

cải tiến nó để cho ra đời bảng chữ cái tiếng Hy Lạp như ngày nay. Ở Tây Hy Lạp nó thoát thai sang một dạng thức khác và thâm nhập vào Ý, nơi người ta đã cải tiến nó để cho ra đời bảng chữ cái La-tinh. Thứ tự các chữ cái ngày nay trong bảng chữ cái Roman – ví dụ, thứ tự mà chúng ta dùng trong từ điển – là giống hệt với thứ tự trong bảng chữ cái tiếng Hy Lạp và tiếng Do Thái (ngôn ngữ của các cư dân sống ở khu vực Palestine cổ đại). Vì thế có thể thấy rằng thứ tự này đơn thuần mang tính quy ước truyền thống. Ta khó có thể giải thích hay chứng minh điều gì ở đây. Thứ tự trong bảng chữ cái tiếng Hindu thì hoàn toàn khác, chúng được sắp xếp một cách khoa học, phân thành hai nhóm, phụ âm và nguyên âm, và sắp xếp nhóm phụ âm theo vị trí cấu âm, bắt đầu từ vị trí sâu trong miệng rồi tiến dần ra đến môi. [Các phụ âm sẽ được trình bày đầu tiên (**phần I, mục 1**) sau đến nguyên âm (**mục 2**); **phần II** sẽ dành để nói về thanh điệu.]

## Phần I: Các phụ âm và nguyên âm

### 1. Phụ âm

Tiếng La-tinh có hai dãy *phụ âm tắc* (phụ âm khi phát ra bị cản trở hoàn toàn tại một vị trí nào đó trong miệng) nếu căn cứ theo tính thanh: dãy hữu thanh (tức thanh quản rung): /b/, /d/, /g/, viết là B, D, G, và dãy vô thanh (tức thanh quản không rung): /p/, /t/, /k/, viết là P, T, C, Q. Hệ thống hai loạt phụ âm này giờ vẫn được bảo lưu trong tiếng Pháp. Mặt khác, trong tiếng Hy Lạp cổ đại có tới ba dãy phụ âm tắc: hữu thanh (/b/, /d/, /g/, viết là B, Δ, Γ), vô thanh không bật hơi (/p/, /t/, /k/, viết là Π, Τ, Κ), và vô thanh bật hơi (/p<sup>h</sup>/, /t<sup>h</sup>/, /k<sup>h</sup>/, viết là Φ, Θ, Χ). Giới học giả La-tinh ghi lại dãy phụ âm thứ ba này (vô thanh bật hơi) bằng cách thêm con chữ H vào sau các phụ âm dãy thứ hai (vô thanh không bật hơi), kết quả thu được các tổ hợp PH, TH, KH. Tuy nhiên không lâu sau khi thời kì Cổ đại kết thúc, cách phát âm tiếng Hy Lạp bắt đầu biến đổi: tất cả các phụ âm tắc đều bị xát hóa, tức thay vì bị cản trở toàn bộ, luồng khí vẫn thoát ra được từ một điểm nào đó thuộc bộ vị phát âm. Do vậy, khi chữ cái Roman được dùng để ghi âm cho các ngôn ngữ German vào giai đoạn bắt đầu thời kì Trung cổ, các tổ hợp PH, TH, KH vẫn được sử dụng trong khi các từ ngữ mà chúng biểu thị bấy giờ đều đã biến thành âm xát. Hệ quả là PH, TH, KH được các học giả sử dụng với giá trị ngữ âm mới, khác với giá trị nguyên gốc của chúng trong tiếng La-tinh cổ điển: chúng được dùng để ghi các âm xát. Đây là lí do vì sao cho đến ngày nay, TH vẫn còn giá trị xát trong tiếng Anh và CH vẫn còn giá trị xát trong tiếng Đức. Trong tiếng La-tinh thì PH (/p<sup>h</sup>/) phải được phát âm giống F (như /f/) và TH phải được phát âm giống T (/t/). Chính vì thế mà tiếng Pháp phát âm: /t/ cho TH và /f/ cho PH. Trong tiếng Việt, tình hình lại tương đồng với tiếng Hy Lạp: TH vẫn là một âm tắc bật hơi giống như trong tiếng Hy Lạp cổ đại; còn PH lại là một âm xát giống tiếng Hy Lạp hiện đại. Trong kí pháp ngữ âm học, âm xát luôn cần phải được tách biệt rõ với âm tắc bật hơi. Âm xát thường được miêu tả bằng các con chữ Hy Lạp φ, θ, χ đối với tập hợp vô thanh (môi: /φ/, răng: /θ/, ngạc mềm: /χ/) và β, δ, γ với tập hợp hữu thanh (môi: /β/, răng: /δ/, ngạc mềm: /γ/). Với bộ âm tắc bật hơi ph, th, kh [IPA: /p<sup>h</sup>/, /t<sup>h</sup>/, /k<sup>h</sup>/], một vài nhà nghiên cứu ghi chúng với một dấu móc lưng góc phải bên trên là p', t', k': đó chính là *dấu thờ hắt* (spiritus asper) trong chính tả tiếng Hy Lạp, sáng tạo bởi người Hy Lạp thời Alexandria để miêu tả tính bật hơi có trong một số biến thể phương ngữ tiếng Hy Lạp (ngoại trừ tiếng Hy Lạp vùng Athen), những nơi mà con chữ H đã được dùng để ghi một nguyên âm [viết hoa là H, viết thường là η]. Để ghi các âm trên, các nhà Ấn Độ học dùng chữ *ph*, *th*, *bh*, *dh* còn các nhà Hán học thì dùng *p'*, *t'*, *b'*, *d'*.

Tiếng La-tinh có bốn kiểu âm tắc: âm tắc môi (labials) B, P (/b/, /p/), khi phát âm hai môi chạm nhau; âm tắc đầu lưỡi (apicals), tức là âm răng T, D (/t/, /d/), khi phát âm đầu

lưỡi (tiếng La-tinh: apex) chạm vào răng hàm trên; âm tắc mặt lưỡi (dorsals) (tức âm ngạc cứng) C, G (/c/, /ɟ/), khi phát âm mặt lưỡi tiếp xúc với ngạc cứng. Dãy thứ tư là âm tắc ngạc mềm (velar stop), viết bằng chữ cái Q, khi phát âm mặt lưỡi tiếp xúc với phần ngạc mềm; trong tiếng La-tinh, âm này luôn bị môi hóa [IPA: /kʷ/], tức là khi phát âm phải tròn môi (giống như khi phát âm nguyên âm U [IPA: /u/] trong tiếng Việt). Đó chính là lí do tại sao sau Q lại có thêm chữ cái U trong hệ thống chính tả La-tinh. QU [IPA: /kʷ/] là một âm tắc vô thanh tương liên với âm tắc hữu thanh GU [IPA: /gʷ/]. Đối lập này giờ vẫn được bảo lưu trong tiếng Rumani, nơi mà chữ La-tinh AQUA “nước” biến thành *apa* trong khi chữ LINGUA “lưỡi” lại biến thành *limba* [nói cách khác, cặp đối lập /kʷ/ và /gʷ/ trong tiếng La-tinh vẫn còn được bảo lưu nét đối lập về tính thanh trong tiếng Rumani hiện tại, cụ thể là giữa /p/ (vô thanh) và /b/ (hữu thanh)]. Trong tiếng La-tinh, QUI là một âm tiết đơn [kʷi/], trong khi CUI lại là một tổ hợp hai âm tiết [IPA: /ku.i/].

Ngoài ra, cách phát âm cổ của cặp đối lập QU và GU trong tiếng La-tinh giờ vẫn còn được bảo lưu trong tiếng Ý, và chính từ tiếng Ý mà âm đọc này được áp dụng cho tiếng Việt [nơi mà QU là cách viết của /kʷ/]. Trong tiếng Pháp và tiếng Tây Ban Nha thì các âm (tắc) môi – ngạc mềm đều biến thành âm (tắc) mặt lưỡi (ngoại trừ trường hợp khi đứng trước A). Vì thế, từ La-tinh QUAMQUAM được phát âm là CANCAN [IPA: /kãkã/, phụ âm đầu là /k/, chứ không phải /kʷ/] trong trường học ở Paris thế kỉ thứ 16.

## C, G

Lịch sử của C rất phức tạp. Đầu tiên, khi đem so sánh bảng chữ cái tiếng La-tinh và bảng chữ cái tiếng Hy Lạp với nhau, người ta thấy rằng chữ cái thứ ba của bảng chữ cái tiếng Hy Lạp là một âm mặt lưỡi hữu thanh (voiced dorsal), Γ (viết thường: γ), không phải âm mặt lưỡi vô thanh như trong tiếng La-tinh (C). Sở dĩ có đặc điểm này là bởi bảng chữ cái tiếng La-tinh bị ảnh hưởng bởi “người hàng xóm” của mình - người Etruscan, Roman phân biệt hai dãy âm tắc bật hơi và không bật hơi [vd: /tʰ/ và /t/] nhưng lại không phân biệt hai dãy âm tắc hữu thanh và vô thanh với nhau [ví dụ: /d/ và /t/]. (Người Trung Quốc ở Bắc Kinh và Quảng Đông cũng vậy). Vì vậy, người Etruscan đã dùng chữ cái thứ ba của bảng chữ cái tiếng Hy Lạp để biểu thị một âm mặt lưỡi vô thanh. Người La-tinh thì khác, họ vẫn phân biệt rạch ròi giữa âm tắc vô thanh và hữu thanh nên nếu sử dụng cùng một chữ cái cho cả hai âm tắc này sẽ là một việc không hợp lí. Bởi vậy, họ đã thêm một dấu ngang vào giữa C để phân biệt G và C. Vậy là C vẫn tiếp tục nằm ở vị trí thứ ba trong bảng chữ cái trong khi G “rớt xuống” vị trí ban đầu thuộc về Z [viết thường: ζ] trong bảng chữ cái tiếng Hy Lạp, kí tự mà vào giai đoạn đó người La-tinh vẫn chưa sử dụng. (Hai chữ cái tiếng Hy Lạp là Y và Z nằm ở những vị trí cuối cùng trong bảng chữ cái tiếng La-tinh vì mãi về sau chúng mới được nhập hệ).

Trong tiếng La-tinh, C từng có cùng một cách phát âm trong cả âm tiết CE lẫn CO. Tình hình này vẫn còn được bảo lưu trong những ngôn ngữ Celtic thuộc quần đảo Anh (Ai-len và xứ Wales), những ngôn ngữ phi Roman đầu tiên đi vay mượn bảng chữ cái La-tinh. Tuy nhiên, trong tiếng La-tinh, từ thế kỉ thứ 4 trở đi, khi phụ âm mặt lưỡi đi trước các nguyên âm hàng trước E, I (/e/, /i/), chúng đều trở thành những phụ âm tiền ngạc (prepalatal), tức là khi phát âm thì đầu lưỡi đều đồng thời tiếp xúc với phần đầu ngạc. Những phụ âm trên có mặt trong hệ thống âm vị các ngôn ngữ Đông Dương từ rất sớm và có vị trí vững chắc trong những ngôn ngữ này. Trái lại, ở châu Âu, những âm môi trên rất khó thụ cảm và khó phát âm vì khi phát âm chúng, lưỡi phải có hình dạng giống như một “chiếc máng”: nếu như “chiếc máng” nghiêng về phía cuống lưỡi, ta được âm tắc-xuýt-phi tiền (non-anterior sibilant stop) [kí hiệu IPA ngày nay là /tʃ/, một âm tắc xát-lợi], giống như CI, CE trong tiếng Ý; nếu máng nghiêng về phía đầu lưỡi, ta được âm tắc-xuýt-tiền (anterior sibilant stop) [IPA: /ts/, âm tắc xát-lợi]: đây là cách phát âm CI, CE

trong tiếng Pháp thượng cổ và *tiếng Tây Ban Nha thượng cổ* (Old Spanish). Âm tắc xát-lợi này hiện giờ vẫn còn tìm thấy trong tiếng Bắc Kinh và tiếng Quảng Đông, nhưng hoàn toàn vắng bóng trong các ngôn ngữ Đông Dương.

Trong khi đó, các phụ âm trong các âm tiết tiếng La-tinh như TIA, TIO đều đã trở thành những âm tắc-xuýt-tiền [tức là âm tắc xát /ts/] trong tiếng Ý cũng như các ngôn ngữ Roman khác. Chữ cái Z của tiếng Hy Lạp khi này được vay mượn để biểu thị âm tắc-xuýt-tiền-hữu thanh [tức âm /dz/]. Trong tiếng Ý, Z biểu thị cho cả hai âm hữu thanh lẫn vô thanh này [tức cả /dz/ and /ts/].

## K, KH, GH

Khi bảng chữ cái La-tinh được cải tiến để ghi các ngôn ngữ Germanic (vào thế kỉ thứ 7), rồi tới các ngôn ngữ Slavic vốn dĩ có những âm tắc này, Z và C được dùng tới. Người ta đã phải mượn chữ cái K của tiếng Hy Lạp (kí tự này trong bảng chữ cái tiếng La-tinh vẫn giữ nguyên vị trí như trong bảng chữ cái tiếng Hy Lạp) để biểu thị âm tắc-mặt lưỡi (dorsal stop).

Chính điều này giải thích tại sao trong chữ Quốc ngữ, người ta viết CA, CÔ, CU [biểu thị cho /ka/, /ko/, /ku/] trong khi lại viết KÊ, KI [biểu thị cho /ke/, /ki/]: trong tiếng Việt, QU không thể sử dụng theo cách mà nó được sử dụng trong tiếng Pháp và tiếng Tây Ban Nha [tức để biểu thị cho /k/] vì QU đã được dùng để biểu thị một âm khác [IPA /kʷ/]; âm /k/ tiếng Việt cũng không thể được kí hiệu là CH (như cách làm trong tiếng Ý) vì CH đã được dùng để biểu thị âm tắc-vô thanh-ngạc cũng như tình hình ta thấy sau đây. Cũng chính vì lí do trên mà ta loại trừ luôn khả năng sử dụng CH cho âm tắc bật hơi [IPA: /kʰ/]. Từ đó dẫn tới việc sử dụng KH để biểu thị âm tắc bật hơi này [IPA: /kʰ/], khi nó đứng trước các nguyên âm. Tính chất mặt lưỡi hữu thanh cũng tạo ra hệ quả tương tự với: GA, GÔ, GU [cho các âm tiếng Việt /ya/, /yo/, /yu/], nhưng lại là GHÊ và GHI [cho các âm /ge/ và /gi/ mà ngày nay đọc là /ye/ và /yi/] chứ không phải là GÊ và GI; người ta cũng không thể dùng được tổ hợp GUÊ và GUI, vốn được dùng trong trường hợp phụ âm đầu là một âm ngạc mềm-môi, trong khi kí pháp theo cách của tiếng Ý có thể được sử dụng, tức GHÊ, GHI.

Trong khoa học thì kí pháp k vẫn thường được dùng cho âm tắc-mặt lưỡi-vô thanh (unvoiced dorsal stop) còn g được dùng cho âm tắc-mặt lưỡi-vang (sonorant dorsal stop).

## CH

Trong các ngôn ngữ Roman phía Tây có một âm tắc tiền ngạc (prepalatal stop) phái sinh và sau đó dần chuyển thành một âm tắc xuýt – phi tiền (non-anterior sibilant stop) [IPA: âm xát sau lợi (postalveolar affricate), /tʃ/]. Trong tiếng Pháp, nó bắt nguồn từ âm tiết CA của tiếng La-tinh: ví dụ, *calidum* “nóng” thành *chaud* (tiếng Pháp), *caballum* “ngựa” thành *cheval* (tiếng Pháp). Âm tắc xuýt – phi tiền này [IPA: âm xát sau lợi (postalveolar affricate), /tʃ/] được viết là CH, trong đó chức năng của H là chỉ báo rằng âm C này được phát âm khác với C trong CO và CE. Trong thời kì Charle Đại đế (Charlemagne) - thế kỉ thứ 8, âm này vẫn là một âm tắc tiền ngạc (prepalatal stop); tuy nhiên chỉ một thời gian ngắn sau đó (thế kỉ thứ 9), trong tiếng Pháp thượng cổ, nó đã biến thành một âm tắc xuýt-phi tiền [IPA: /tʃ/] và bảo lưu từ bấy đến nay trong tiếng Anh. Trong tiếng Bồ Đào Nha thì nó lại bắt nguồn từ tổ hợp phụ âm đầu PL, CL; còn trong tiếng Tây Ban Nha thì nó bắt nguồn từ tổ hợp *nằm giữa từ* (word-medial) CT, ví dụ tiếng La-tinh là *noctem* “đêm” thì ban đầu chuyển thành *noite*, như hiện dạng trong tiếng Bồ Đào Nha, sau đó chuyển thành *noche*. Trong tiếng Basque, âm tắc tiền ngạc này vẫn còn được bảo lưu cho đến ngày nay, ví dụ trong từ *éche* “ngôi nhà”.



Với tình hình xảy ra ở các ngôn ngữ gần gũi đó, thật dễ hiểu khi Roman sáng tạo ra chữ Quốc ngữ lại sử dụng CH để kí hiệu cho âm tắc tiền ngạc của tiếng Việt.

Kí pháp khoa học của âm tắc ngạc (palatal stop) chưa thực sự thống nhất [IPA chuẩn: /c/] do đây là một âm khá hiếm trong các ngôn ngữ châu Âu; đôi khi nó được kí hiệu là k' hoặc t' trong đó dấu ' báo hiệu tính chất mềm hóa, một hiện tượng được gọi tên trong tiếng Pháp là *mouillure*, nghĩa đen là 'làm ướt', vì khi phát âm âm này vị trí nằm ngang của mặt dưới lưỡi chạm vào, tức 'làm ướt' ngạc. Kí hiệu này được mượn từ một ngôn ngữ Slavơ - tiếng Séc - và có trong tên thương hiệu giấy *Bat'a* nổi tiếng khắp châu Âu. Các nhà ngữ âm học tiếng Anh và tiếng Hán thì phản đối cách thêm dấu phụ ở bên cạnh như vậy. Thay vào đó, họ thêm một dấu lượn vào bên dưới con chữ. Các học giả Ấn Độ thì ghi nó là C. Thực chất, âm tắc tiền ngạc này trong tiếng Sanskrit đã từng được bảo lưu ở Đông Dương, khi kí tự chữ Sanskrit được vay mượn để ghi âm tiếng Khmer và tiếng Lào. Còn ở Ấn Độ thì âm này dần chuyển thành một âm tắc xuýt-phi tiền [IPA: âm tắc xát sau lợi (postalveolar affricate), /tʃ/]. Với những âm tắc xuýt (sibilant stop), cách kí âm trong tiếng Séc là rất thông dụng: C dùng cho âm tắc xuýt hàng trước [tức /ts/], Č cho âm tắc xuýt phi tiền [tức /tʃ/] (vì vậy mà "Czech" được ghi là *česky*). Kí hiệu Č là một cách viết tắt của CZ, một tổ hợp vẫn còn được sử dụng trong tiếng Ba Lan. Tiếng Séc hiện tại thì không có âm tắc xuýt-tiền-hữu thanh (voiced anterior sibilant stop) nào, và Z chỉ được sử dụng để ghi một âm xát hữu thanh [tức âm xát /z/] giống như trong tiếng Pháp.

Tiếng Pháp thượng cổ từng có những âm tắc giống như vậy [IPA: tắc xát (affricate)] mà từ thế kỉ thứ 13 chúng suy yếu dần và biến thành những âm xát (spirant). Âm tắc-xuýt-tiền trở thành âm xát-xuýt-tiền. Kể từ đó, CE và CI phát âm giống như SE và SI. Đồng thời, âm hữu thanh tương ứng cũng trở thành âm hữu thanh đối lập với /s/, tức /z/. Âm tắc xuýt-phi tiền [tức âm tắc xát /tʃ/] biến thành một âm xát-phi tiền: trong tiếng Pháp hiện tại, CH không còn phát âm giống CH trong tiếng Anh [tức /tʃ/] nữa, mà phát âm giống như SH trong tiếng Anh [tức /ʃ/]. Một tổ hợp đầu từ trong tiếng La-tinh có I theo sau là một nguyên âm theo đó cũng diễn biến thành một âm tắc-ngạc-hữu thanh (voiced palatal stop) (giống như G trong GE, GI /ge/, /gi/), sau đó lại chuyển thành một âm tắc xuýt-phi tiền-hữu thanh (voiced non-anterior sibilant stop) [IPA /dʒ/] (chữ cái đầu từ J, nguyên gốc trong tiếng La-tinh là I, nay vẫn được bảo lưu giá trị ngữ âm trong tiếng Anh), và cuối cùng, đến thế kỉ thứ 13 thì chuyển thành một âm xát-phi tiền-hữu thanh (voiced non-anterior spirant) như trong tiếng Pháp hiện nay [IPA: /ʒ/].

#### GI, D, Đ

Kí hiệu của những âm xát này (/ʒ/) là J trong tiếng Pháp mới chỉ xuất hiện từ thế kỉ thứ 17. Trước đó, J đơn thuần là một biến thể của I khi ở vị trí đầu từ. Vì thế người ta từng viết *jnde* (tương đương trong tiếng Pháp hiện tại là *Inde* "Ấn Độ"), *jure* (tương đương với hiện tại là *ivre*, "say"), *jean* (tương đương với danh từ riêng *Jean* trong tiếng Pháp; *John* trong tiếng Anh bây giờ) .v.v. Trong tiếng Ý, tổ hợp GI từng được sử dụng trong tất cả các trường hợp âm tương ứng với nó là một phụ âm, ví dụ *Giovannis* "John". Vì lẽ đó, vào thế kỉ thứ 17, khi được áp dụng vào tiếng Việt, J không tồn tại với tư cách như một chữ cái riêng đối lập I; cho nên ta có thể dễ dàng hiểu được tại sao GI được dùng để làm kí hiệu cho một âm xát hữu thanh tiền ngạc (/ʒ/). Khi kí âm cho âm tắc tiền ngạc hữu thanh (voiced prepalatal stop)<sup>5</sup>, mà người châu Âu nghe tựa như một âm ngạc hóa /d/, một chữ cái D đơn giản đã được dùng tới. Trong tiếng Việt do có một âm tắc lợi [và tiền thanh hầu

<sup>5</sup>Kể từ sau khi bài báo của Haudricourt được xuất bản (1949), các nhà nghiên cứu đã chứng minh rằng âm vị kí hiệu là D trong chữ Quốc ngữ **không phải** là một âm tắc tiền ngạc-hữu thanh (voiced prepalatal stop) (tương đương gần nhất trong IPA là /ɖ/) khi chữ Quốc ngữ mới ra đời, mà là một âm xát răng (dental spirant), được Ferlus (1982) kí hiệu là ɖ (cf. Ferlus 2001), trên cơ sở các tư liệu phục nguyên và từ mượn trong

hóa] hữu thanh [IPA: âm hút vào-hữu thanh (voiced implosive) /ɗ/] nên để kí âm cho nó, những người sáng chế ra bộ chữ cái tiếng Việt đã nghĩ ra một chữ cái mới, Đ, bằng cách thêm vào D một dấu ngang để loại suy nó với T, vốn nằm trong thể đối lập tính thanh với Đ.

## S, TR

Âm tắc-xuýt-tiền (anterior sibilant stop) trong tiếng Tây Ban Nha thượng cổ, được viết là C và Z, đã bị nhuộc hóa, nhưng không hề bị đồng quy với S giống như trong tiếng Pháp. Nó biến thành một âm xát răng (dental spirant) giống như TH trong tiếng Anh [IPA: /θ/]. Theo đó, khi phát âm âm này, lưỡi sẽ nhích ra trước hơn so với âm S tiếng Pháp. Ở một vài khu vực (đặc biệt là vùng Soule Basque), Z tương ứng với S tiếng Pháp: đây là một âm xát răng [IPA: /s/], trong khi S được phát âm (lưỡi) quặt ra đằng sau hơn và trở thành một âm xát quặt lưỡi (cacuminal/retroflex spirant) [IPA: /ʂ/]. Trong tiếng Việt cũng có âm xát quặt lưỡi như vậy, và được kí hiệu bằng chỉ một chữ S vì có sự tương đồng với âm S vừa nhắc đến ở trên, trong khi âm tắc quặt lưỡi tương ứng [IPA: /tʃ/] được phẩm định bằng tổ hợp TR. Những phụ âm này hiếm thấy trong các ngôn ngữ châu Âu và người châu Âu thường nhầm chúng với những phụ âm xuýt-phi tiền (non-anterior sibilant consonants). Ví dụ, tiếng Trung có một phụ âm kí hiệu là TCH trong tài liệu La-tinh hóa của EFEO (Viện Viễn Đông Bác cổ Pháp), và CH trong hệ thống Wade-Giles. Những tổ hợp này dường như biểu thị rằng đây là một phụ âm tắc-xuýt-phi tiền [IPA: /tʃ/], phát âm bằng cách uốn lưỡi ra phía sau đồng thời chu môi ra phía trước, trong khi thực tế nó chỉ là một âm quặt lưỡi [IPA: /tʂ/], phát âm bằng cách uốn đầu lưỡi về phía ngạc cứng. Những phụ âm giống như thế này [IPA: âm quặt lưỡi (retroflex)] rất hay gặp ở các ngôn ngữ Ấn Độ; các học giả Ấn Độ kí hiệu chúng bằng cách thêm một dấu chấm vào bên dưới: ण ण ण. Còn các nhà ngữ âm học tiếng Trung và tiếng Anh lại thêm một dấu móc vào dưới các kí tự phụ âm: ण ण ण [cách này giờ đã thành quy chuẩn của IPA]. Những âm quặt lưỡi là hiếm gặp trong các ngôn ngữ vùng Đông Dương: ở đồng bằng Bắc bộ, chúng đều đã chuyển thành những phụ âm tiền-ngạc (prepalatals).

## X

Trong tiếng La-tinh, chữ cái X được dùng để chỉ hai âm: nó là chữ viết tắt của nhóm kí tự CS. Trong các ngôn ngữ Roman, nhóm phụ âm này diễn tiến theo cách giống với tổ hợp CT: chúng biến thành IS (vd: tiếng La-tinh *coxa*, tiếng Pháp *cuisse* [/kɥis/] “đùi”). Tuy nhiên vào thời Trung cổ, nảy ra một thói quen đó là coi US và IS là một và nhìn qua đều tựa tựa như chữ cái X của tiếng La-tinh, thêm vào đó, từ có nhóm kí tự IS trong tiếng Pháp thường tương ứng với những từ có chữ cái X trong tiếng La-tinh; vì vậy, X dần dần được sử dụng thay thế cho S sau U hoặc I. Đó là lí do tại sao tiếng Pháp phát âm *soixante* “sáu mươi” y như *soissante* [IPA: /swasɑ̃t/], *Auxerre* (tên địa danh) y như *Ausserre* [IPA: /osɛʁ/], và *Bruxelles* (Brussels) y như *Brusselles* [IPA: /bʁɥsɛl/]; nó cũng giải thích tại sao dạng số nhiều của *cheval* “ngựa” là *chevaux*, dạng số nhiều của *genou* “đầu gối” là *genoux*. Đây đơn giản là một vấn đề về thói quen chính tả. Tuy nhiên, kể từ thế kỷ thứ 16, khi tiếng

tiếng Việt. Tương đương gần gũi nhất trong IPA của âm này hiện tại là /ð/. Tuy nhiên, Ferlus phân biệt rạch ròi giữa âm xát “*spirant*” và âm xát “*fricative*” theo quan điểm của Martinet 1956:24-25. Theo Martinet, “*spirant*” và “*fricative*” là hai nhóm âm khác nhau: “*cần thiết phải minh định sự khác biệt giữa phương thức cấu âm lỏng (relaxed) – có xu hướng gần với phương thức nguyên âm mà sản phẩm của nó ta gọi là “spirant” – với các phụ âm được cấu âm theo phương thức phụ âm đúng nghĩa, đặc trưng bởi sự ma sát khi luồng hơi thoát qua vị trí cản trở: đó chính là những âm “fricative” đúng nghĩa.*” (cf. Martinet 1981, 1985 và Thomas, Bouquiaux and Cloarec-Heiss 1976:29-31). Tuy nhiên, Haudricourt không có ý thức phân biệt hai thuật ngữ trên trong bài báo này: ông nhất loạt dùng “**spirant**” trong toàn bài.

Pháp vay mượn một số từ ngữ của tiếng Hy Lạp, cách phát âm của chữ cái X lại được phục nguyên như cũ, tức là như KS, GZ [IPA: /ks/, /gz/].

Chúng ta đều biết trong tiếng Tây Ban Nha, nhóm kí tự IT đã chuyển thành những âm tắc-tiền ngạc và tình hình tương tự cũng xảy ra với nhóm IS, viết là X. Theo đó, trong tiếng Tây Ban Nha thượng cổ, X kí hiệu cho một âm xát tiền-ngạc; tuy nhiên đối với người Pháp đây lại là một âm xuýt phi tiền (non-anterior sibilant): *Don Quixote* và *Ximena* được ghi trong tiếng Pháp là *Don Quichotte* và *Chimène* [cả hai đều phát âm là /ʃ/]. Tuy nhiên, tên của nhà truyền giáo nổi tiếng *Xavier* (một cái tên trong tiếng Tây Ban Nha xứ Basque, tương đương với trong tiếng Pháp xứ Basque là *echeverry*, “nhà mới”) lại được xem như một từ bác học và được phát âm là *Gzavié* [IPA: /gzavje/].

Những dữ kiện này cho phép chúng ta giải thích tại sao âm xát tiền ngạc (prepalatal) trong tiếng Việt lại được ghi là X.

Kí pháp của âm xát này trong bảng IPA là ɕ (chữ cái c với dấu móc câu nhỏ bên dưới), nhưng khi nó được phát âm gần giống với một âm xuýt (tiền âm, giống như trong tiếng Việt, hoặc phi-tiền), dấu phụ biểu thị nét ngạc hóa được thêm vào kí hiệu phụ âm xuýt tương ứng (như trong tiếng Ba Lan: s', c'). Những âm xát xuýt-tiền luôn được ghi là S, Z [IPA: /s/, /z/]. ( ... )<sup>6</sup> Theo các nhà ngữ âm học Anh ngữ và Hán ngữ, âm tắc xuýt thường được ghi bằng tổ hợp hai kí tự, theo thứ tự tắc + xát: /ts/, /dz/ cho những âm xuýt-tiền âm và /tʃ/, /dʒ/ cho những âm xuýt phi-tiền. ( ... ) Âm tắc uốn lưỡi trong tiếng Việt và tiếng Hán đều được ghi là ʈʂ, ʈʂ hoặc ʈʂ, ʈʂ. Theo quy ước của các nhà ngữ âm học, những âm tắc được ghi bằng hai chữ cái là những âm bị tắc-xát hóa.

## N, NH, NG

Âm mũi (âm tắc hữu thanh cộng với việc lưỡi con hạ thấp, để luồng khí từ phổi thoát ra một phần theo mũi) N, M có cùng một giá trị trong tất cả các ngôn ngữ. Tuy nhiên trong tiếng Pháp, khi N, M ở vị trí cuối âm tiết, chúng không được phát âm theo cách bình thường nữa: thay vào đó các nguyên âm ở trước chúng lại bị mũi hóa (lưỡi con hạ thấp trong khi phát âm nguyên âm). Trong tiếng Bồ Đào Nha, hiện tượng tương tự cũng xảy ra, nhưng âm N, được viết trên đầu các nguyên âm, biến thành ã. Các nhà ngôn ngữ học sau này tận dụng dấu phụ này để chỉ các nguyên âm mũi, ví dụ /ã/. Trong tiếng Tây Ban Nha, chữ *N kép* (chuỗi hai chữ N) tương tự cũng được viết thêm dấu ã, tức ñ; sau này chữ N kép chuyển thành một âm mũi-tiền ngạc (prepalatal nasal); các nhà Ấn Độ học đã vay mượn kí tự ñ này cho âm này. Trong tiếng Pháp và tiếng Ý, âm mũi-tiền ngạc bắt nguồn từ tổ hợp phụ âm GN của tiếng La-tinh; quy tắc chính tả này được bảo lưu trong tiếng Pháp và tiếng Ý, trong khi đó, trong phương ngữ Provençal và Gascon của tiếng Bồ Đào Nha, âm này được kí hiệu là NH vì loại suy từ âm tắc tương ứng, CH. Quy tắc chính tả NH này của tiếng Bồ đã được áp dụng cho tiếng Việt. Các nhà ngữ âm học Anh ngữ và Hán ngữ thì kéo dài và uốn cong nét đầu của N để biểu thị âm này: /ɲ/.

Cuối cùng, trong các thứ tiếng Roman thì âm mũi mặt lưỡi (dorsal nasal) rất hiếm gặp, trừ khi N đi trước C hay G; âm này chỉ xuất hiện ở vị trí cuối từ trong tiếng Anh hay Đức, và được viết là NG. Vì vậy hoàn toàn tự nhiên khi trong tiếng Việt người ta cũng ghi âm này là NG. Các nhà ngữ âm học Hán ngữ và Anh ngữ thì gộp 2 kí tự này làm một bằng cách kéo dài nét thứ hai của n sao cho giống với nét thứ hai của g, và kết quả ta được: ɳ. Các nhà Ấn Độ ngữ học thì ghi âm này với một dấu chấm trên đầu của n, tức ṇ.

## V

<sup>6</sup>Trong bản dịch có hai câu bị xóa được kí hiệu bằng ( ... ).

Chúng ta hãy điểm lại nguồn gốc của V. Trong tiếng La-tinh, V và U không hề tách biệt nhau: V là kí tự dùng để ghi nguyên âm môi-ngạc mềm (labiovelar), còn U là dạng viết tháu của V. Khi theo sau U là một nguyên âm, nó trở thành một phụ âm môi-ngạc mềm [IPA: /w/], sau đó nó biến tiếp thành một phụ âm môi-răng (đối lập về tính thanh với F [IPA: /v/], đối lập với /f/]); biến âm này xảy ra vào khoảng thế kỉ thứ 6, trong khi đó, tình hình này được giữ nguyên cho đến thế kỉ 17 nên, mãi cho đến thời điểm đó, hai âm tách biệt này vẫn được ghi bằng một chữ cái. Cụ thể, khi chữ cái V dùng ở vị trí đầu từ, nó được hiểu là một nguyên âm (giống như trong *vn* “một” [trong tiếng Pháp hiện đại được ghi là *un*]) hoặc là một phụ âm (như trong *vent* “gió” [giờ vẫn được viết là *vent* trong tiếng Pháp hiện đại]). Khi ở vị trí giữa từ, U được dùng thay thế [cho V], như trong *chanure* “sợi gai dầu” [trong tiếng Pháp hiện đại là *chanvre*] hay *mur* [tiếng Pháp hiện đại: *mur*]. Tuy nhiên, trong các thứ tiếng khác, có một kí tự mới được sáng tạo trước thế kỉ 17 từ rất sớm: vì rằng phụ âm môi-ngạc mềm này đã chuyển thành một âm môi-răng trong tiếng La-tinh phương ngữ Roman vào thế kỉ thứ 6, các học giả Roman, Roman đã chế tác hệ thống chính tả cho các ngôn ngữ German, bèn sáng tạo ra một kí tự mới cho phụ âm môi-ngạc mềm xuất hiện trong các ngôn ngữ German: chữ cái “double U” (“U kép” trong tiếng Anh ngày nay vẫn gọi như vậy), có hình dạng ngày nay là W (gọi là “double V” trong tiếng Pháp [ND: cách gọi “V kép” này được người Việt vay mượn]). Các phương ngữ của tiếng Anh và tiếng Pháp miền Bắc (Picard, Wallon) ngày nay vẫn bảo lưu âm môi-ngạc mềm này, như cách phát âm từ *Wallon* [/walɔ̃/, chú không phải /valɔ̃/]. Trong khi đó, trong tiếng Đức, Hà Lan và Ba Lan, W lại được dùng để ghi một âm môi-răng (vì vậy mà phụ âm đầu *Wagram*, *Württemberg* được phát âm như một âm môi-răng [/v/]). Các nhà ngôn ngữ học thì tách bạch V – một âm môi-răng với W – một âm môi-ngạc mềm.

## H

Về phụ âm cuối cùng xét đến trong bài này, chúng ta biết rằng trong tiếng Pháp có một phụ âm có chữ viết nhưng không được phát âm: đó là âm “*H bật hơi*” (“aspirated H”) nhằm ngăn việc đọc nuốt mạo từ [theo đó phải ghi là *la hache*, *le hibou* “cái rìu”, “con cú” chứ không ghi là *l'hache*, *l'hibou*]. Âm này là âm câm, ngoại trừ trong một số thổ ngữ (của vùng Lorraine hay Normandie). Ngược lại, tiếng Việt lại có một phụ âm được phát âm nhưng lại không có chữ viết: âm tắc thanh hầu (đột ngột mở thanh quản) giữ vị trí âm đầu trong các từ bắt đầu bằng nguyên âm. Trong kí pháp khoa học, âm này được ghi bởi một kí tự gọi là “*spiritus lenis*” trong hệ thống chữ viết Hy Lạp, dấu này có hình móc câu với phần eo uốn ra trước: ‘ [ví dụ: *Ἱπποκράτης* “Hippocrates”].

Thay bởi vậy, các nhà ngữ âm học tiếng Anh và tiếng Hán thì dùng dấu hỏi để biểu thị phụ âm này: ? [kí pháp IPA hiện thời: ʔ]

## 2. Nguyên âm

A, E, I, O, U

Tiếng La-tinh từng có 5 nguyên âm:

- (i) A : một nguyên âm mở (tức khoảng cách giữa lưỡi và vòm miệng lớn nhất)
- (ii) và (iii): hai nguyên âm đóng cực đại. I, một nguyên âm tiền ngạc (prepalatal), hay miêu tả chính xác hơn là một nguyên âm hàng trước căn cứ vào vị trí lưỡi; I không phải là âm tròn môi, tức khi phát âm môi không mở rộng. U, một nguyên âm ngạc mềm, chính xác hơn là một nguyên âm hàng sau (xét theo vị trí lưỡi) và là một âm tròn môi;
- (iv) và (v): hai nguyên âm trung bình (ND: xét theo độ nâng của lưỡi): E nằm giữa I và A; O nằm giữa U và A.

U từng thay đổi giá trị ngữ âm trong tiếng Pháp thượng cổ vào quãng thế kỉ thứ 10: trở thành một nguyên âm hàng trước tròn môi [IPA: /y/]. Kí pháp khoa học của nguyên âm tiếng Pháp này là ü (còn ö thì được dùng để kí hiệu cho nguyên âm nằm giữa ü và A); kí pháp này bắt nguồn từ tiếng Đức, nơi xuất hiện dấu *Umlaut* [ND: hai chấm ở phía trên nguyên âm] nhằm thay cho E: vì vậy ü thực chất thay cho ue và ö thì thay cho oe. Trong khi đó, các nhà ngữ âm học tiếng Anh thì chuộng kí pháp Y của người Scandinavia hơn.

Y

Trong tiếng Hy Lạp cổ đại, một nguyên âm đóng hàng sau tròn môi [IPA: /u/], ghi là Y, đã từng chuyển thành một phụ âm tròn môi-tiền âm (anterior rounded consonant). Các học giả La-tinh đã mượn chữ cái Y của Hy Lạp để kí hiệu cho nguyên âm này. Âm này giờ không còn trong tiếng La-tinh. Trong tiếng Hy Lạp, nguyên âm này sau đó bị đồng quy với I khiến trong tiếng La-tinh Y được phát âm như I. Do đó mà người Pháp ngày nay vẫn kêu chữ cái Y bằng tên gọi “*chữ I Hy Lạp*” [ND: hay còn được gọi là “chữ I cò-rét” ở Việt Nam]. Về vai trò của nó trong bảng chữ cái tiếng Việt: trong tiếng Pháp, Y thường được dùng để kí hiệu âm xát tiền ngạc hữu thanh (voiced prepalatal spirant) và quan hệ giữa Y với I giống hệt như quan hệ giữa W với U; đây là kí pháp mà các nhà Ấn Độ học [và Châu Phi học] sử dụng. Tuy nhiên, các nhà ngữ âm học tiếng Anh lại ưa dùng j để kí hiệu cho âm này hơn. [Kí pháp /j/ nay đã trở thành kí pháp chuẩn trong IPA.]

Tiếng La-tinh phân biệt nguyên âm ngắn và nguyên âm dài. Kí pháp được quy ước như sau: a cho nguyên âm dài, ă cho nguyên âm ngắn tương ứng; chính kí pháp ă đã được sử dụng cho tiếng Việt. Các nhà Ấn Độ học thì chỉ ghi nguyên âm dài mà bỏ qua các nguyên âm ngắn. Các nhà ngữ âm học tiếng Anh và tiếng Hán thì biểu thị trường độ dài của các nguyên âm bằng dấu hai chấm viết kế bên hoặc viết kép: *a:* hoặc *aa*. [IPA: /a:/]

Ê, Ô, Â

Trong các ngôn ngữ Roman, nguyên âm dài xuất hiện khi hai nguyên âm đi cùng nhau; chúng được ghi tắt bằng cách chỉ ghi kí hiệu một nguyên âm đồng thời thêm dấu mũ ^ phía trên nó, ví dụ trong tiếng Pháp *aage* được viết là *âge*, *meur* thì viết là *mûr*. (Trong tiếng Pháp, ^ thường thế cho âm s đứng đằng sau, ví dụ *fête* “ngày lễ” có dạng thức trước đây là *feste*, *pâte* “hồ dán” là *paste*.) Tình hình tương tự cũng xảy ra trong tiếng Bồ Đào Nha, khi *oo* viết thành *ô*, *ec* viết thành *ê*. Trong tiếng Tây Ban Nha, các nguyên âm mới có cách phát âm đóng nhiều hơn: ô nằm ở vị trí trung gian giữa o và u, còn ê nằm giữa e và i. Đây chính là nguồn gốc các kí pháp được sử dụng cho các nguyên âm tiếng Việt. Trong tiếng Tây Ban Nha, tiếng Ý hay tiếng vùng Provence (Pháp) đều không có kí hiệu thích hợp để phân biệt hai nguyên âm e và hai nguyên âm o này.

Trong tiếng Pháp, ô [IPA: /o/] có khẩu độ đóng nhiều hơn so với o [IPA: /ɔ/], và ê [IPA: /e/] cũng có khẩu độ đóng nhiều hơn so với e [IPA: /ɛ/]; căn cứ vào đó, một vài nhà ngôn ngữ học đã sử dụng dấu sắc: ó, é để biểu thị mức độ đóng nhiều hơn của các nguyên âm, và dấu huyền, tức ò, è để biểu thị mức độ mở nhiều hơn của các nguyên âm này. Một cách khác đó là dùng dấu chấm vào bên dưới để chỉ các nguyên âm đóng nhiều hơn và dấu móc vào bên dưới để chỉ các nguyên âm mở nhiều hơn. Cuối cùng, các nhà ngữ âm học tiếng Anh sử dụng e, o cho các nguyên âm đóng nhiều hơn và ɔ, ɛ (tức chữ C ngược và chữ *epsilon* của tiếng Hy Lạp) cho các nguyên âm mở nhiều hơn. Kí hiệu æ được dùng cho nguyên âm nằm giữa a và ɛ, giống như nguyên âm trong từ *cat* của tiếng Anh. Một chữ a nghiêng (i.e., *a*) thì được dùng cho nguyên âm giống như nguyên âm đầu trong từ *pâte* của tiếng Pháp.

Trong tiếng Việt có một số nguyên âm hàng sau không tròn môi [IPA: /u/, /ɤ/]. Chúng được viết là u và ơ, có thể là do lấy cảm hứng từ kí pháp ù, ồ đã nhắc đến ở trên, cho dù thực tế là u khá đối lập với ù xét trên vị trí môi và lưỡi: u [IPA: /u/] là một nguyên âm hàng sau và không tròn môi, trong khi ù [IPA: /y/] lại là một nguyên âm hàng trước tròn môi. Những nguyên âm này (ND: tức /u/, /ɤ/ không có trong các ngôn ngữ Roman phía Tây, tuy nhiên chúng lại có trong các ngôn ngữ Roman phía Đông: ví dụ trong tiếng Rumani, từ mang nghĩa “chó” trên thực tế có thể được ghi theo chính tả tiếng Việt là *cưn*. Những nguyên âm này đôi khi được ghi với hai dấu “*umlaut*”: i, e hoặc là ghi với một dấu tròn. Các nhà ngữ âm học tiếng Anh thì ghi chúng là /u/, /ə/ [IPA /ə/ ngày này được coi đích thị là một nguyên âm hàng giữa, còn /ɤ/ thì biểu thị một nguyên âm hàng sau, không tròn môi, độ nâng lưỡi trung bình], và /ʌ/, tức những kí pháp đảo ngược, theo thứ tự, của m, e và v. Nguyên âm cuối /ʌ/ chính là âm â trong tiếng Việt.

## Phần II: Thanh điệu

Cuối cùng, chúng ta hãy tìm hiểu các kí pháp thanh điệu được mượn từ tiếng Hy Lạp cổ đại. Tiếng Hy Lạp cổ đại có hai thanh, ghi là ´ và ~, còn kí hiệu thứ ba được dùng để mô tả thanh điệu của những từ vô thanh: ˘. Những kí hiệu này sau được bổ sung thêm những kí hiệu đánh dấu câu. Trên thực tế, trong các ngôn ngữ Roman, thanh điệu không thể được sử dụng để phân biệt ý nghĩa các từ, chúng chỉ có chức năng phân biệt mục đích phát ngôn, ví dụ, trong tiếng Pháp, câu nghi vấn *C'est vrai?* “Việc đó là thật à?” nhờ dấu câu mà phân biệt với câu trần thuật *C'est vrai* “Việc đó là thật.” Theo đó, người ta đặt một dấu chấm nhỏ (giống với dấu chấm câu) bên dưới nguyên âm và một dấu hỏi trên đầu nguyên âm để biểu thị kí pháp thanh điệu [minh họa cho hệ thống 6 thanh trên nguyên âm A hiện nay là: a à á ạ ã ã].

Các nhà ngữ âm tiếng Anh và tiếng Hán thì sử dụng một dấu đứng đặt bên phải từ kèm một dấu mô tả đường nét và âm vực của thanh điệu để kí hiệu chúng [ví dụ ˧ kí hiệu thanh cao, ˨ kí hiệu thanh trung bình, .v.v.].

## Kết luận

Bảng chữ cái tiếng Việt là một sản phẩm có nguồn gốc dễ nhận diện. Những giáo sĩ châu Âu đầu tiên đặt chân đến vùng Viễn Đông bằng đường biển đều đến từ bán đảo Iberia: họ là người Bồ Đào Nha, Tây Ban Nha và xứ Basque. Trong số đó, những học giả đã sử dụng bảng chữ cái tiếng La-tinh để ghi âm tiếng Việt chính là các giáo sĩ Thiên chúa giáo – Roman tinh thông tiếng La-tinh, tiếng Ý và tiếng Hy Lạp. Tất cả những điểm dị biệt của bảng chữ cái tiếng Việt vì vậy mà có thể được soi sáng nhờ những dữ kiện này.

Nhìn vào những bộ kí pháp khoa học dựa trên bảng chữ cái Roman đã trình bày ở những phần trước, ta có thể chia chúng ra làm hai nhóm sau: (i) những bộ được sử dụng cho các ngôn ngữ Roman phía Đông đã có chữ viết “alphabet” riêng trước đó, trong trường hợp này việc ghi âm tiếng La-tinh cơ bản chỉ là việc chuyển tự, tức thay thế một chữ cái của bảng chữ cái bản địa bằng một chữ cái La-tinh sao cho tạo điều kiện thuận lợi cho việc in ấn ở châu Âu; (ii) những bộ chữ được chế tác bởi các nhà ngữ âm học, Roman có mục đích miêu tả tất cả các sắc thái của cách phát âm. Loại bảng chữ cái thứ hai này chủ yếu được dùng ở các nước không có bộ chữ cái “alphabet” (ví dụ: Trung Quốc) hoặc các nước mà chính tả cách xa ngữ âm (ví dụ như nước Anh: rất khó có thể đoán được cách phát âm nguyên âm trong từ tiếng Anh, cho dù tiếng Anh cũng dùng bảng chữ cái La-tinh). Tuy nhiên, phải nói rằng bảng chữ cái ngữ âm được thiết kế vốn nhằm phục vụ việc viết tay, và không thích hợp cho việc in ấn.

**Bảng 1:** Từ bảng chữ cái tiếng Do Thái và tiếng Hy Lạp đến bảng chữ cái tiếng La-tinh.

Hebrew	Greek		Latin	
א ALEF	A α	ALPHA	A	
ב BET	B β	BETA	B	
ג GIMEL	Γ γ	GAMMA	C	
ד DALET	Δ δ	DELTA	D	
ה HE	E ε	E-pilon	E	
ו VAV			F	
ז ZAYIN	Ζ ζ	ZETA		G
ח HET	Η η	ETA	H	
ט TET	Θ θ	THETA		
י YOD	Ι ι	IOTA	I	J
כ KAF	Κ κ	KAPPA	K	
ל LAMED	Λ λ	LAMBDA	L	
מ MEM	Μ μ	MU	M	
נ NUN	Ν ν	NU	N	
ס SAMEKH	Ξ ξ	XI		
ע AYIN	Ο ο	O-micron	O	
פ PE	Π π	PI	P	
צ TSADI				
ק KUF Q			Q	
ר RESH	Ρ ρ	RHO	R	
ש SHIN	Σ σ ς	SIGMA	S	
ת TAV	T τ	TAU	T	
	Υ υ	U-pilon	U	V
	Φ φ	PHI	X	
	Χ χ	CHI	Y	
	Ψ ψ	PSI	Z	
	Ω ω	O-mega		

**Bảng 2:** Đối chiếu các bảng chữ cái: tiếng Việt, tiếng Pháp, và kí pháp khoa học (bao gồm kí pháp của giới Ấn Độ học và giới Hán học).

Vietnamese	French	Scientific notations :	
		indianists and ethnologists	phoneticians and sinologists
th		th	t'
ph			φ
	f	f	f
	ch	š	s
	tch	č or tš	ts
ch		c or t'	t or t̥
d		j or d'	d or d̥
x		ç or ś or s' or š'	ç or ś
gi	z' or ž'		ž
s		ş	ş
tr		tş	ts
nh	gn	ñ	ɲ
ng		ṇ	ŋ
vocalic initial		'	ʔ
	eu	ö œ	ø
	u	ü	y
ư			
u	ou	u	u
ơ		ə	ə
	ill	y	j
ê	é	ẹ	e
e	è	e	ε
ô	o	ọ	o
	o		ɔ



## Tài liệu tham khảo

1. Alexandre de Rhode (1991), *Từ điển Annam – Lusitan – Latinh* (bản dịch của Hoàng Xuân Việt, Thanh Lăng, Đỗ Quang Chính), NXB Khoa học xã hội, TP. Hồ Chí Minh.
2. Cao Xuân Hạo & Hoàng Dũng (2005), *Từ điển thuật ngữ ngôn ngữ học đối chiếu Anh Việt - Việt Anh*, NXB Khoa học Xã hội.
3. Ferlus, M. (1982), “Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien”, *Cahiers de linguistique - Asie Orientale* (11:1), pp. 83-106.
4. Ferlus, M. (2001), “Les hypercorrections dans le thổ de Làng Lỗ (Nghệ An, Vietnam) ou les pièges du comparatisme”, Paper presented at the *Quinzièmes Journées de Linguistique de l’Asie Orientale*, Ecole des Hautes Etudes en Sciences Sociales, Paris.
5. Haudricourt, A.G. (1949), “L’origine des particularités de l’alphabet vietnamien”, *Dân Việt Nam* (3), pp. 61-68.
6. Haudricourt, A.G. (1953), “La place du vietnamien dan les langues austroasiatique” (Vị trí của tiếng Việt trong các ngôn ngữ Nam Á), *BSLP* (49), pp. 122-128.
7. Haudricourt, A.G. (1954), “De L’origine Des Tons En Vietnamien” (Nguồn gốc thanh điệu tiếng Việt), *Journal Asiatique* (242), pp. 69–82.
8. Haudricourt, A.G. (2010), “The origin of the peculiarities of the Vietnamese alphabet” (Alexis Michaud trans.), *Mon-Khmer Studies* (39), pp. 89–104.
9. Martinet, A.(1956), *La Description phonologique avec application au parler franco-provençal d’Hauteville (Savoie)*, Genève: Droz.
10. Martinet, A. (1981), “Fricatives and spirants”, *Suniti Kumar Chatterji commemoration volume ed. by B. P. Mallik, Burdwan*, West Bengal, India: Burdwan University Press, pp. 145-151.
11. Martinet, A. (1985), “Two proposals”, *The Study of Sounds (Onsei no Kenkyuu), Commemorative volume for the 50th anniversary of the Phonetic Society of Japan XXI*, pp. 67-72.
12. Suriya, R. et al. (1985), *Southeast Asian Linguistic Studies presented to André-G. Haudricourt*, Mahidol University, Bangkok.
13. Thomas, J. et al. (1976), *Initiation à la phonétique : phonétique articulatoire et phonétique distinctive*, Presses Universitaires de France, Paris.