

MỘT SỐ ĐIỂM CẦN ĐƯỢC NHÌN NHẬN LẠI TRONG CẤU TẠO TỪ TIẾNG VIỆT

PGS.TS VŨ ĐỨC NGHIÊU*

Trong các công trình nghiên cứu, miêu tả tiếng Việt, chúng ta thường gặp một hệ thống phân loại và định vị các từ về mặt cấu tạo, bao gồm từ đơn tiết và từ đa tiết; trong các từ đa tiết, bao gồm từ ngẫu hợp, từ láy và từ ghép; đến lượt mình, các từ ghép lại được phân định thành hai tiểu loại: từ ghép đẳng lập và từ ghép chính phụ; các từ láy cũng tiếp tục được phân chia chi tiết thành các tiểu loại khác nhau...

Hệ thống được đề xuất như vậy là khá đẹp và cân đối. Tuy nhiên, thực tế cho thấy: đằng sau nó có một số khoảng tối hiện chưa được các nhà nghiên cứu đồng thuận trong quan niệm và trong việc xử lý, định loại.

1. Trước hết là vấn đề của những đơn vị trước nay vẫn được gọi là từ ghép, gồm cả *từ ghép đẳng lập* lẫn *từ ghép chính phụ*.

Điều bất thuận đầu tiên bộc lộ chính ngay từ việc dùng thuật ngữ *đẳng lập* và *chính phụ*. Ai cũng biết rằng quan hệ giữa các yếu tố cấu tạo bên trong từ là *phối cú pháp* nhưng ở đây, tại những đơn vị như: *anh em, thuyền bè, làng xóm, đi lại, ăn nói, vật lộn, nhà máy, cá vàng, xe đạp, nhảy sào...* chúng ta hoàn toàn

không gặp khó khăn gì khi dựa vào những đặc điểm ngữ pháp và ngữ nghĩa để chỉ ra một cách rất cụ thể những mối quan hệ giữa các thành tố đó là: đẳng lập (*anh em, thuyền bè, làng xóm, đầu đuôi, đi lại, ăn nói, xoay trở, vật lộn...*), chính phụ (*nhà máy, cá vàng, xe đạp, nhảy sào...*), hay miêu tả chúng như những tổ hợp được tạo nên bằng những thành tố rất rõ ràng về mặt từ loại và đặc trưng ngữ pháp như: *danh - danh, danh - tính, danh - động, động - tân...*

Vấn đề đặt ra là: một khi những đơn vị như *thuyền bè, làng xóm, đi lại, ăn nói, vật lộn, đường sắt, nhà máy, cá vàng, xe đạp, nhảy sào...* mà đã là từ, thì các yếu tố như *thuyền, bè, làng, xóm, đi, lại, ăn, nói, vật, lộn, đường, sắt, nhà, máy, cá, vàng, xe, đạp, nhảy, sào...* đương nhiên phải là những *hình vị*. Thế nhưng, quả thật, hầu như chúng ta không thể chứng minh được những "yếu tố cấu tạo từ - hình vị" như: *thuyền, bè, làng, xóm, đi, lại, ăn, nói, vật, lộn, đường, sắt, nhà, máy, cá, vàng,*

* Phó hiệu trưởng Trường đại học KHXH & NV ĐHQG Hà Nội.

xe, đập, nhảy, sào... trong các “từ ghép” như thế khác với những từ đơn tiết: *thuyền, bè, làng, xóm, đi, lại, ăn, nói, vật, lộn, đường, sắt, nhà, máy, cá, vàng, xe, đập, nhảy, sào...* về tư cách hình thái học và về khái niệm được biểu thị.

Ở đây có ba điểm cần chú ý:

a. Sự khác nhau giữa các thành tố *thuyền, bè, nhảy, sào...* trong *thuyền bè, nhảy sào...* và *thuyền, bè, nhảy, sào...* với tư cách những từ đơn, tự do, nếu có chăng, chỉ là khác nhau về bối cảnh tồn tại và hoạt động của chính cùng một đơn vị: một đẳng, ‘chúng bị ràng buộc vào trong một bối cảnh (một tổ hợp) cố định hoặc tương đối cố định, làm thành một tên gọi biểu thị khái niệm, còn một đẳng thì chúng vẫn đứng một mình, hoàn toàn tự do, và sẵn sàng tham gia vào mọi kiểu tổ hợp ngữ pháp khi điều kiện cho phép.

b. Thực chất vấn đề là ở chỗ: Để đáp ứng nhu cầu đặt tên càng

ngày càng tăng và càng ngày càng phức tạp, những đơn vị từ vựng mới phải luôn luôn được tạo lập. Một số lượng từ ngữ khiêm tốn và chỉ gồm những hình thái (từ) đơn tiết sẽ không đủ để đáp ứng được nhu cầu đó. Phương thức cấu tạo các đơn vị từ vựng bằng cách ghép (tổ hợp) những từ sẵn có lại, đã được tiếng Việt lựa chọn; và chính những “từ ghép” như vừa nói trên: *cha mẹ, anh em, trường sở, lương bổng, đường sắt, mặt trời, cá vàng, cà chua, xe cộ, cỏ rả, trắng nõn, xanh lè, đối chác, đi lại, mua bán, đánh rửa, lau chùi...* đã hình thành bằng con đường như vậy.

Ngoài ra, cũng cần nói thêm là: tình hình như trên đây không phải chỉ xảy ra ở tiếng Việt. Quan sát những đơn vị từ vựng song phần có quan hệ đẳng lập trong hàng loạt ngôn ngữ khác gần gũi với tiếng Việt về mặt khu vực và loại hình, chúng ta có thể thấy, chẳng hạn như:

Tiếng Bru Vân kiều:

mpiq (mẹ) + mpaq (cha / bố) = mpiq mpaq (cha mẹ)
tăk (mặc) + cha (ăn) = tăk cha (ăn mặc)
panhiêih (đối) + khlak (khát) = panhiêih khlak (đối khát)
atùng (ngược) + atook (xuôi) = atùng atook (ngược xuôi)

Tiếng Chăm:

akok (đầu) + iku (đuôi) = akok iku (đầu đuôi)
lithay (cơm) + ia (nước) = lithay ia (cơm nước)
băng (ăn) + mưnhum (uống) = băng - mưnhum (ăn uống)
tangin (tay) + takai (chân) = tangin - takai (tay chân)...

Tiếng Dao:

luj2 (quần) + hăw4 (áo) = luj2 - hăw4 (quần áo)
căj4 (xây) + liêp6 (dựng) = căj4 - liêp6 (xây dựng)
pε76 (trắng) + kie75 (đen) = pε76 kie75 (trắng đen)
im2 (cay) + bjat6 (đắng) = im2 - bjat6 (cay đắng)...

Tiếng Katu:

<i>val</i> (trở lại)	+ <i>vôi?</i> (đi)	= <i>val vôi?</i> (đi lại)
<i>moh</i> (mũi)	+ <i>mạt</i> (mắt)	= <i>moh mạt</i> (mắt mũi / diện mạo)
<i>bhui</i> (vui)	+ <i>hal</i> (vui)	= <i>bhui hal</i> (vui vẻ)
<i>Jai</i> (giúp)	+ <i>por</i> (giúp)	= <i>Jai - por</i> (giúp đỡ)...

Tiếng Khmer:

/ <i>si</i> / (ăn)	+ / <i>chai</i> / (tiêu)	= / <i>si - chai</i> / (ăn tiêu)
/ <i>muk</i> / (mặt)	+ / <i>mwat</i> / (miệng)	= / <i>muk mwat</i> / (mặt mũi)
/ <i>că</i> / (tim)	+ / <i>thlaəm</i> / (gan)	= / <i>că - thlaəm</i> / (tắm lòng)
/ <i>sok</i> / (sương)	+ / <i>tuk</i> / (khô)	= / <i>sok tuk</i> / (sương khô)

Tiếng Lào:

/ <i>ai</i> / (anh)	+ / <i>nɔ:ŋ</i> / (em)	= / <i>ai - nɔ:ŋ</i> / (anh em)
/ <i>fɔ</i> / (cha)	+ / <i>mɛ</i> / (mẹ)	= / <i>fɔ - mɛ</i> / (cha mẹ)
/ <i>juŋ</i> / (rối)	+ / <i>ɲak</i> / (khó)	= / <i>juŋ - ɲak</i> / (khó khăn)
/ <i>cư</i> / (nhớ/ghi)	+ / <i>căm</i> / (nhớ)	= / <i>cư căm</i> / (nhớ/ghi nhớ)...

Tiếng Malay:

<i>tawa</i> (già)	+ <i>muda</i> (trẻ)	= <i>tawa muda</i> (già trẻ)
<i>cepat</i> (nhanh)	+ <i>ligat</i> (nhanh)	= <i>cepat ligat</i> (nhanh nhẹn)
<i>hutan</i> (rừng)	+ <i>rimba</i> (rừng)	= <i>hutan rimba</i> (rừng rú)
<i>tempik</i> (kêu)	+ <i>surak</i> (la)	= <i>tempik surak</i> (kêu la)...

Tiếng Miến điện:

/ <i>hyei</i> / (trước)	+ / <i>nau</i> / (sau)	= / <i>hyei nau</i> / (trước sau)
/ <i>kaun</i> / (tốt)	+ / <i>mun</i> / (tốt)	= / <i>kaun mun</i> / (tốt đẹp)
/ <i>badha</i> / (tiếng)	+ / <i>sakà</i> / (lời)	= / <i>badha sakà</i> / (tiếng nói)
/ <i>cí</i> / (nhìn)	+ / <i>hyu</i> / (ngó)	= / <i>cí hyu</i> / (nhìn ngó)...

Tiếng Pu péo:

<i>pê4</i> (cha)	+ <i>maj2</i> (mẹ)	= <i>pê4 - maj2</i> (cha mẹ)
<i>mjaw4</i> (mặt)	+ <i>tê1</i> (mắt)	= <i>mjaw4 - tê1</i> (mặt mũi)
<i>rưng3</i> (rừng)	+ <i>bu4</i> (núi/rú)	= <i>rưng3 - bu4</i> (rừng núi/rú)
<i>tăj5</i> (cây)	+ <i>mjak</i> (quả)	= <i>tăj5 mjak</i> (cây cối)...

Tiếng Tày Nùng:

<i>hua</i> (đầu)	+ <i>hang</i> (đuôi)	= <i>hua hang</i> (đầu đuôi)
<i>chan</i> (lười)	+ <i>mjət</i> (nhác)	= <i>chan mjət</i> (lười nhác)
<i>đông</i> (rừng)	+ <i>pá</i> (rừng)	= <i>đông pá</i> (rừng rú)
<i>doan</i> (sợ)	+ <i>mảo</i> (hãi)	= <i>doan - mảo</i> (sợ hãi)...

Tiếng Thái lan:

/ <i>phít</i> / (sai)	+ / <i>chɔp</i> /	= / <i>phít chɔp</i> / (đúng sai)
/ <i>cay</i> / (tim)	+ / <i>khɔ</i> / (cổ)	= / <i>cay khɔ</i> / (trái tim / tình cảm)
/ <i>cèp</i> / (ốm)	+ / <i>khây</i> / (sốt)	= / <i>cèp khây</i> / (ốm đau)
/ <i>ri:p</i> / (vội)	+ / <i>rɔ:n</i> / (nóng)	= / <i>ri:p rɔ:n</i> / (vội vàng)...

Kết quả là, trong tiếng Việt, những hệ lụy phức tạp và nhất là sự thiếu nhất quán trong việc xác định tư cách, cương vị của những đơn vị trước nay vẫn gọi là từ ghép, đã nảy sinh, gây nên những cuộc tranh luận triền miên.

c. Trong quá trình tạo lập và tồn tại của những đơn vị theo kiểu như vừa nêu trên, có một số thành tố trong đó đã dần dần bị "hao mòn" về nghĩa đến mức được coi là vô nghĩa. Chính trạng thái "vô nghĩa" của những thành tố như thế đã làm cho các đơn vị chứa chúng được nhìn nhận là có tính bền chặt, liền khối, và những đơn vị nào có thành tố "vô nghĩa" thì nghiêm nhiên được coi là "từ ghép", còn những đơn vị khác thì khi phân tích, định loại, lại phải cân nhắc, quyết đáp tùy từng trường hợp cụ thể, rất thiếu tính hệ thống và chặt chẽ. Thí dụ, tư liệu nghiên cứu ngôn ngữ học lịch sử, và nghiên cứu sự tiếp xúc giữa tiếng Việt với các ngôn ngữ lân cận cho thấy: phương thức tổ chức và tư cách của những đơn vị

như được dẫn dưới đây là không khác nhau:

a. *lúa má, cỏ rả, cá mú, tre pheo, chó má, xe cộ, áo xống, đối chác, súng ống, kiêng khem, xông pha, chửi bới, kiêng cử...*

và a'. *nhà cửa, áo quần, anh em, ăn ở, đầu đuôi, ăn nói, leo trèo, đánh rửa, lau chùi...*

b. *trắng nõn, đất thó, mưa phùn, lược bí, mặt nạ, đường sá, đòn càn...*

và b'. *đường sắt, xe đạp, khăn quàng, nhà lá, bàn gỗ, áo len, chậu nhôm...*

Thế nhưng hiện nay, các đơn vị như ở loại a. và b. thì nghiêm nhiên được coi là từ ghép, không hề gây một phản ứng gì, vì trong mỗi đơn vị ở đây đều có một thành tố "vô nghĩa". (Trong khi đó, các nghiên cứu hữu quan đã xác định được nghĩa của hàng loạt yếu tố "vô nghĩa" như: (*chung*) *chạ*, (*chửi*) *bới*, (*chợ*) *búa*, (*đòn*) *càn*, (*kém*) *cỏi*, (*xe*) *cộ*, (*bờ*) *cỏi*, (*đối*) *chác*, (*chùa*) *chiền*, (*chim*) *chóc*, (*dức*) *lác*, (*đất*)

đai, (theo) đòi, (giữ) gìn, (hỏi) han, (xanh) lè, (e) lệ, (láo) lếu, (chó) má, (mưa) móc, (cá) mú, ngật (nghèo), (bếp) núc, (inh) ỏi, (súng) ống, (xông) pha, (tre) pheo, (cỏ) rả, (sân) sườn, (tuổi) tác, (tối) tăm, (quét) tước, (vênh) váo, vì (nê), (xin) xỏ, (áo) xống...; còn những đơn vị như ở loại a'. và b'. thì các ý kiến tranh luận về chúng rất khó thuyết phục nhau, đến nay chưa có lời kết. Chính vì thế cho nên có tình trạng: những đơn vị như *đường sắt, nhà máy, mặt trời, trái đất, tủ lạnh, nhà khách...* thì được coi là từ ghép, mà những đơn vị như: *bàn gỗ, áo len, bát sắt, cốc nhựa, mặt đất, mặt biển...* lại không được coi là từ, trong khi thực chất cả hai loại đều là những ngữ, (những tổ hợp), vì bản chất ngữ pháp, ngữ nghĩa chúng hoàn toàn như nhau, chỉ chênh lệch nhau đôi chút về mức độ ràng buộc và hạn chế lẫn nhau giữa những thành tố tổ hợp với nhau mà thôi. Sự thiếu nhất quán này trong việc đánh giá tư cách từ / không phải là từ của những đơn vị được tổ chức theo kiểu như thế, rõ ràng cần phải được xem xét lại.

2. Về những từ láy vốn không phải là từ láy

Trên thực tế, có khá nhiều đơn vị được gọi là từ láy nhưng thật ra lại là những đơn vị đã hình thành không phải nhờ phương thức láy, mà do các từ tổ hợp với nhau, hoặc là theo quan hệ đẳng lập (quan hệ này là chủ yếu) hoặc là theo quan hệ chính phụ, rồi nhân những điều kiện nhất định, đã dần dần được định loại là từ láy. Cụ thể là: những đơn vị ấy vốn dĩ cũng hình thành bằng con đường tổ hợp hai từ với

nhau nhưng trải qua thời gian, cùng với những biến đổi lịch sử của ngôn ngữ và một số nhân tố khác nữa như xung đột đồng nghĩa, cận nghĩa chẳng hạn, nghĩa và quan hệ về nghĩa giữa các thành tố cấu tạo nên chúng đã bị che lấp hoặc bị xoá mờ đi, khiến cho người hiện đại không nhận ra được nữa; đồng thời, nếu giữa hai thành tố của mỗi đơn vị như thế mà ngẫu nhiên có bộ phận ngữ âm trùng lặp nhau theo kiểu của từ láy, thì chúng sẽ dễ dàng được đưa vào loại từ láy. Vì thế, những đơn vị như:

1- *đất đai, chùa chiền, tuổi tác, hỏi han, thân thích, thân thiết, quyến luyến, bệ vệ, tha thiết, giữ gìn...*

và: 2- *ngọc ngà, vải vóc, lụa là, rổ rá, cam quýt, co kéo, kìm kẹp, rơm rác, tướng tá, đến đài...* đã được đối xử khác nhau, chỉ đơn giản vì trong các đơn vị loại 1- đều có một thành tố bị "mờ nghĩa" để cho sự tương đồng (đến một mức đạt yêu cầu nào đó) về mặt ngữ âm giữa hai thành tố của chúng nổi lên và được coi trọng; còn sự tương đồng về ngữ âm giữa hai thành tố của các đơn vị loại 2- lại bị bỏ qua, bởi chúng còn "rõ nghĩa" cả.

Có thể chứng minh cho sự thiếu nhất quán như thế bằng một vài khảo sát cụ thể sau đây:

a. Trong *Hán Việt từ điển* [15], một từ điển khá lớn, thu thập khoảng 5000 đơn tự và khoảng 40.000 đơn vị đa tiết, nếu tạm gạt nghĩa của các đơn vị sang một bên, chỉ xét đến sự trùng lặp (tương đồng) về ngữ âm giữa hai thành tố của các đơn vị song tiết đó, thì thấy: 150 đơn vị có sự tương đồng hoàn toàn,

654 đơn vị tương đồng ở phần vần, 1571 đơn vị tương đồng ở âm đầu.

Trong số 2375 đơn vị này, có 94 đơn vị đã được *Từ điển từ láy tiếng Việt* [16] chính thức thừa nhận và thu thập (16 "từ láy hoàn toàn", 49 "từ láy âm đầu" và 29 "từ láy phần vần"). Thí dụ: *song song, thanh thanh, trùng trùng, nhất nhất, do dự, đảo đẽ, hỗn hào, lam lũ, lưỡng lự, nguy nga, tồi tàn, tung tích, thiết tha, bàng hoàng, quyến luyến, đạo mạo, ung dung, thổ lộ, tiêu điều ...*

Hơn chín chục "từ láy" trên đây rõ ràng vốn không có liên quan và chắc chắn, như đã chứng minh, vốn không phải là từ láy của tiếng Việt.

b. *Từ điển tiếng Việt* 1994 [18] thu thập hơn hai nghìn (2086) đơn vị từ vựng song tiết đẳng lập, có trên ba trăm năm mươi (355) đơn vị được tạo nên bằng những thành tố mang âm đầu trùng lặp nhau như: *bạn bè, bùa bả, ban bố, bụ bẫm, canh củi, câu cú, càn quét, cay cực, chấn chỉnh, đoán định, dạy dỗ, dạn dày, dàn dựng, đình đám, gánh gồng, lặn lội, lông lá, mệt mỏi, nghe ngóng, hàng hoá, học hành, hang hóc, phố phường, quê quán, ruộng rẫy, tóc tai, tán tưng, xem xét...*

Ở đây, nếu so sánh về phương diện tổ chức nghĩa, phương diện quan hệ giữa các thành tố và mô thức cấu tạo, thì rõ ràng những đơn vị này chẳng khác gì so với những đơn vị mà trước nay chúng ta vẫn đã gọi là từ láy âm đầu như: *con cái, ấm áp, chẳng chịt, đại đột, gớm ghiếc, rúc rười, nhọc nhằn, thêm thắt, thể thốt, xa xôi, bạn bè, tiệt tưng, ấm áp, vui vẻ, màu mè...*

Thật ra, những hiện tượng mập mờ và không rạch ròi về đường biên giới định loại như vừa trình bày, đều khởi nguồn từ sự tổ hợp các từ đơn (những đơn vị hình thái cơ bản) lại với nhau để tạo thành những đơn vị từ vựng đa tiết. Tuy nhiên, chính tình trạng bị hạn chế, ràng buộc của một số từ đơn nào đó với một từ đơn khác trong những tổ hợp đa tiết được tạo thành đã dẫn đến sự mơ hồ của cái gọi là tính ổn định, bền chặt về cấu trúc... của chúng; và khi trong các tổ hợp đa tiết (song tiết) bền vững hoặc khá bền vững đó, cả hai hoặc một trong hai thành tố của chúng bị mờ nghĩa đi, đồng thời giữa hai thành tố này có một tương đồng ngữ âm ngẫu nhiên nào đó (ở âm đầu hoặc phần vần) và quan hệ ngữ âm đó nổi dần lên thành quan hệ hàng đầu, thì lập tức các tổ hợp song tiết đó thường đã được đưa sang khu vực từ láy một cách "tự nhiên".

3. Trạng thái đa tiết của những đơn vị trước nay vẫn quen gọi là từ ngẫu hợp (ngẫu kết) như: *ba láp, ba trộn, bồ cào, bồ các, bồ câu, bồ chao, bồ (mồ) côi, bồ dài, bồ hòn, bồ kết, bồ ngót, bù (bồ) nhìn, bồ nông, bồ láo, mồ hôi...* cũng có rất nhiều điều còn ẩn khuất. Tuy vậy, có lẽ sự biến đổi của các tổ hợp phụ âm đầu trong quá trình phát triển lịch sử của tiếng Việt, nhất là những biến đổi theo con đường âm tiết hoá (một trong năm khả năng biến đổi của tổ hợp phụ âm đầu trong tiếng Việt - theo H. Maspero [10]), có thể là một chìa khoá hữu hiệu để giải quyết một số trường hợp thuộc loại này. Thí dụ:

a. Trường hợp từ *bồ câu*. Theo Jakhontov, [xem 1, 247] hình thái (từ) *bồ câu* của tiếng Việt hiện đại vốn bắt nguồn từ một hình thái cổ xưa có dạng tổ hợp phụ âm đầu trong ngôn ngữ Proto Đông Thuỷ là *pq.

Về vấn đề này, V.I. Gohman trong cuốn sách *Ngữ âm lịch sử các ngôn ngữ Thái* [8] cũng dẫn công trình của S.E. Jakhontov *Sự tương ứng giữa các phụ âm gốc lưỡi trong các ngôn ngữ Đông - Thái* (Tạp chí Đông phương học, số 7, Leningrad, 1980) và nhận xét:

"S.E. Jakhontov cho rằng ở đây có một tổ hợp phụ âm Proto Đông Thái kiểu *pq. Giải thuyết này là hoàn toàn có thể được, và có lẽ tổ hợp đó thoát kì thuỷ đã tồn tại trong ngôn ngữ Proto Đông Thuỷ" [8, 30]. "S.E. Jakhontov đề xuất tổ hợp phụ âm kiểu *pq và chúng tôi chia sẻ quan niệm ấy của ông" [8, 150]. Mặt khác, chính V.I. Gohman cũng cung cấp tư liệu [8, 30] cho biết: từ có nghĩa là *bồ câu* (từ *bồ câu* của tiếng Việt) được bảo lưu ở dạng *pau* trong 04 thổ ngữ, dạng *pjau* trong 01 thổ ngữ, dạng *peu* trong 01 thổ ngữ, dạng *kau* trong 01 thổ ngữ, dạng *qau* trong 01 thổ ngữ dạng *kuu* trong 01 thổ ngữ).

b. Từ điển Việt - Bồ Đào Nha - Latinh thế kỉ XVII còn ghi *tlòng; tlên...* mà ngày nay chúng ta có thể đối chiếu để thấy nguồn gốc của *tròng, lòng (lòng trắng, lòng đỏ), thông long; tên, then, trên, then len...* Như vậy, khả năng hình thành *thòng lọng* từ *tlòng*, của *thên len*

từ *tlên* dường như đã trở nên khá rõ ràng.

c. Trong tiếng Việt hiện đại, những tên gọi có nguồn gốc là từ phiên âm như Cà mau (< Khmău. Tiếng Khmer), cờ lê (< clé. Tiếng Pháp)... chính là những biểu hiện của cách thức biến đổi "truyền thống" vẫn còn hoạt động đến tận ngày nay (chen thêm một nguyên âm, thường là nguyên âm dòng giữa, vào tổ hợp phụ âm để cùng với thành tố phụ âm đứng phía trước làm thành một âm tiết).

Tuy nhiên, đây là những vấn đề vô cùng phức tạp, bởi vì không phải trường hợp nào cũng chỉ xảy ra theo một kiểu như vậy và tất cả đều đã kiểm chứng được một cách tường minh. Theo một nghiên cứu (chưa công bố) về các từ cổ trong tiếng Việt của GS Nguyễn Tài Cẩn mà chúng tôi được tham khảo thì từ *mồ hôi* của tiếng Việt ngày nay lại bắt nguồn từ dạng **pəhor* của ngôn ngữ Proto Việt Chứt (Proto Việt Mường), vì tư liệu của các nhà nghiên cứu hữu quan cho biết: tương ứng với **pəhor* đời xưa, hiện nay trong tiếng Mường có các dạng như *môl, pưon, bưon* ở những nhóm thổ ngữ Mường khác nhau, trong tiếng Arem là (*pəhujm*), trong tiếng Sách là (*puhu*), trong tiếng Rục là (*pahu*), trong tiếng Pacô - Taôi là (*pahoor*); còn trong tiếng Proto Bana, dạng phục nguyên của từ có nghĩa mồ hôi lại là **brahal*.

Như vậy, không giống như trường hợp *bồ câu* được hình thành từ một từ có tổ hợp phụ âm biến đổi do các nguyên nhân lịch sử (âm tiết hoá), từ *bồ (mồ) hôi* lại vốn là dạng

đa (song) tiết của Proto Việt Mường. Điều này chứng tỏ rằng với những gì đã có về mặt tư liệu, chúng ta đã có thể tự tin để nghĩ tới cách giải thích cho sự ra đời của nhiều hình thái trong hàng loạt hình thái (từ) như: *ba lớp, ba trợn, bồ cào, bồ các, bồ cắt, bồ câu, bồ chao, bồ (mồ) côi, bồ dài, bồ hòn, bồ kết, bồ ngót, bù (bồ) nhìn, bồ nông, mồ hôi, bố láo, cà chớn, cà khố, cà kheo, cà tàng...* Chúng, hoặc đã được hình thành bằng con đường âm tiết hoá các tổ hợp phụ âm, hoặc vốn dĩ đã là đa tiết từ thời xa xưa, chứ không phải là những đơn vị được hình thành nhờ hai thành tố "ngẫu hợp" với nhau.

4. Vấn đề tính đơn tiết và đa tiết

Một vấn đề vẫn thường đặt ra là: với tư cách là những đơn vị từ vựng, từ của tiếng Việt trước nay đã và đang được tổ chức, cấu tạo theo xu hướng đơn tiết hay đa tiết? Câu hỏi này không phải là khó lí giải, nhưng đã từng có những cách thức phân tích, biện minh không

giống nhau: có ý kiến khẳng định rằng: trong tiếng Việt, cứ mỗi tiếng là đúng một hình vị, đồng thời cũng chính là một từ; còn những ý kiến khác thì lại khẳng định: trong tiếng Việt tồn tại một cách hiển nhiên các từ đơn và từ ghép với những kiểu loại và cách thức tổ chức khác nhau.

Theo những nguồn tư liệu hiện biết và những kết quả nghiên cứu hữu quan [1], [4], [5], [6], [7], từ thời ngôn ngữ Proto Việt Mường (cũng gọi là Proto Việt Chút - [1]) và xa hơn, đến thời Proto Môn Khmer đã từng có những đơn vị đa tiết (từ cận song tiết *sesquisyllabic words*) bao gồm một âm tiết chính và một âm tiết phụ (tiền âm tiết: *presyllable*), trong đó, âm tiết chính luôn luôn được nhấn mạnh. Về sau, có một quá trình đơn tiết hoá đã diễn ra rất mạnh mẽ đối với những hình thái cận song tiết đó, khiến cho chúng trở thành đơn tiết và hiện diện trong tiếng Việt ngày nay. Thí dụ:

Proto MK.	V. hiện đại
* <i>gəm(a / i) ?</i>	<i>mưa</i>
* <i>lətə ?</i>	<i>đá</i>
* <i>Cətəm</i>	<i>(bên) đăm</i>
* <i>kəʔe:p</i>	<i>rết</i>
(...)	

Proto VM.	V. hiện đại
* <i>t(ə)ŋɔc</i>	<i>ngọt</i>
* <i>lətə ?</i>	<i>đá</i>
* <i>agas</i>	<i>gãi</i>
* <i>akă ?</i>	<i>cá</i>
* <i>kəlo</i>	<i>(ngôi) sao</i>
* <i>ahlɔ ?</i>	<i>lúa</i>
* <i>s / c (ə)lu</i>	<i>trâu</i>
* <i>k(ə)rəŋ</i>	<i>rừng</i>
(...)	

Proto MK.	V. hiện đại
* <i>səman</i>	<i>mượn</i>
* <i>Crɪŋ</i>	<i>sùng</i>
* <i>pəđəm</i>	<i>năm (5)</i>
* <i>j (-m-)ha: / iəm</i>	<i>rướm</i>

Proto VM.	V. hiện đại
* <i>usũk</i>	<i>tóc</i>
* <i>kəcet</i>	<i>giết</i>
* <i>acɔ ?</i>	<i>chó</i>
* <i>k(ə)mur</i>	<i>môi</i>
* <i>s(ə)rɔ ?</i>	<i>(khoai) sọ</i>
* <i>ahrɔ</i>	<i>rùa</i>
* <i>t(ə)ŋɔc</i>	<i>ngủi</i>
* <i>ate ?</i>	<i>đái</i>

Các tư liệu cho thấy cơ chế của xu hướng đơn tiết hoá này, thường chỉ đơn giản là cho rụng đi một âm tiết (thường là tiền âm tiết - presyllable). So sánh một số từ tiếng Việt với tiếng Rục, tiếng Thà Vưng, những ngôn ngữ có cùng quan hệ cội nguồn với tiếng Việt, chúng ta sẽ càng thấy rõ và khẳng định được điều này.

Rục	Thà Vưng	Việt
<i>Cuboj</i>	<i>apəɬ</i>	<i>môi</i>
<i>kakat</i>	<i>akayh</i>	<i>gãi</i>
<i>kũmul</i>	<i>khamool</i>	<i>mối</i>
<i>pəi</i>	<i>phaləə</i>	<i>trái (cây)</i>
<i>klɔi</i>	<i>kaləɬ</i>	<i>(chó) sói</i>
<i>kũluok</i>	<i>khalook</i>	<i>tróc</i>
(...)		

Ngày nay, cơ chế cho rụng bớt âm tiết đi, chỉ giữ lại một âm tiết khi có những từ đa tiết nào đó được đơn tiết hoá, vẫn tiếp tục hoạt động, tiếp tục được sử dụng như vậy. Thí dụ:

(Ông) Lý trưởng	> (Ông) Lý	(Cậu) Tú tài	> (Cậu) Tú
Đô la	> đô	Vải Simili	> Vải Si
Thái lan	> Thái	Ôstrâylia	> Úc...

Quả thực, xu thế đơn tiết hoá đã làm thay đổi diện mạo hình thái học của hàng loạt đơn vị hình thái cơ bản vốn là đa tiết và góp phần làm thay đổi cục diện phân bố tỉ lệ đơn tiết / đa tiết trong tiếng Việt: số lượng từ đơn tiết sẽ chiếm tỷ lệ cao trong toàn bộ từ vựng. Khảo sát ngẫu nhiên một vài nguồn ngữ liệu, chấp nhận những đơn vị như: *áo quần, con cháu, cùng nhau, lẫn quất, mền kép, thủy chung* ... là những từ đa tiết, chúng tôi thu được một số kết quả sau đây:

Bảng 1: Kết quả khảo sát một số tư liệu thành văn thế kỉ XV-XVII.

Nguồn ngữ liệu \ Tỷ lệ	Từ đơn tiết	Từ đa tiết
Quốc âm thi tập (thế kỷ XV)	1542 (68%)	693 (32%)
Hồng Đức quốc âm thi tập (thế kỉ XV)	1895 (59%)	1398 (41%)
Thơ Nôm Nguyễn Bỉnh Khiêm (thế kỉ XVI)	1573 (71%)	635 (29%)
Phép giảng tám ngày (thế kỉ XVII)	1145 (64%)	655 (36%)

Bảng 2: *Kết quả khảo sát một số tư liệu báo chí, từ cuối thế kỷ XIX đến năm 1920 [2]:*

<i>Thời gian</i>	<i>Số từ được khảo sát</i>	<i>Từ đơn tiết</i>	<i>Từ đa tiết</i>
Cuối t. kỉ XIX	4098	3592 (86,20%)	506 (13,80%)
1901 - 1910	6034	4042 (64,67%)	1992 (35,33%)
1911 - 1920	7662	4448 (57,82%)	3124 (42,18%)

Bảng 3: Một số nguồn tư liệu có tỷ lệ đa tiết cao hơn đơn tiết

<i>Nguồn ngữ liệu / Tỷ lệ</i>	<i>Từ đơn tiết</i>	<i>Từ đa tiết</i>
Tự vị An Nam - La tinh (thế kỉ XVIII).	4462 (16%)	23428 (84%)
7661 từ thống kê được trong Báo chí tiếng Việt năm 1921 - 1930. (xem [2])	3264 (45,7%)	3430 (54,3%)
Từ điển tiếng Việt 1994 (cuối thế kỉ XX; Chỉ tính riêng các đơn vị song tiết).		19366 (50%)

Các bảng thống kê trên đây cho thấy bức tranh phân bố đơn tiết / đa tiết rõ ràng là đa dạng, nhưng chúng ta không thể không nhận ra một xu hướng chung là, trong quá khứ lịch sử, tỷ lệ từ đơn tiết (chỉ ít là ở nguồn tư liệu được chúng tôi khảo sát) luôn cao hơn tỷ lệ đa tiết. Khi số liệu cho thấy tỷ lệ “từ đa tiết” cao hơn từ đơn tiết, thì điều này cũng không có gì đáng lạ, bởi vì nếu chúng ta nhìn nhận vấn đề “từ ghép” một cách chặt chẽ và có hệ thống hơn, có cái nhìn thực tế hơn đối với những đơn vị như: *áo quần, cha mẹ, đi lại, đầu đuôi, cà chua, áo dài, đường sắt, xe đạp...* coi chúng chỉ là những ngữ cố định do các từ đơn tổ hợp lại với nhau mà thành, thì chắc chắn sẽ không còn ai nghi ngờ gì về tính đơn tiết là phổ biến, điển hình, tỷ lệ từ đơn tiết là áp đảo trong từ vựng tiếng Việt.

Tuy nhiên, sự nổi trội, chủ đạo của tính đơn tiết không có nghĩa rằng tiếng Việt chỉ có từ đơn tiết. Trong tiến trình lịch sử đi tới tiếng Việt ngày nay, từ thời Proto Môn Khmer vẫn đã có những từ vốn là đa tiết. Đến thời Proto Katu cũng đã từng xuất hiện xu hướng đa tiết hoá những từ vốn là đơn tiết thời Proto Môn Khmer. Cụ thể là: có những từ vào thời Proto Môn Khmer là đơn tiết nhưng đến thời Proto Katu, Proto Việt Mường lại được song tiết hoá [1, 236]. Thí dụ:

X

Proto MK

Proto Katu

Proto VM

Việt hiện đại

* *ka?** *?aka** *akă?*

Cá

* *C (uə)?** *?aCɔ** *aCʔ*

Chó

Như vậy, có thể thấy, từ đa tiết cũng là những đơn vị tồn tại một cách hiển nhiên trong tiếng Việt và chúng đã được hình thành bằng những cách thức không giống nhau, chẳng hạn như: vốn là đa tiết từ nguyên thủy, âm tiết hoá một thành phần trong tổ hợp phụ âm (đó là chưa kể kết quả của xu thế ghép các đơn vị hình thái cơ bản - từ đơn - lại với nhau, tạo thành vô số đơn vị từ vựng đa tiết như đã trình bày bên trên, vẫn thường được coi là các từ). Thí dụ: Bảng số 3 trên đây đã trình bày: trong *Từ điển tiếng Việt 1994*, chỉ tính riêng các đơn vị từ vựng song tiết có quan hệ đẳng lập hoặc quan hệ chính phụ, chúng tôi đã kiểm được trên mười chín nghìn (19.366) đơn vị, chiếm tỷ lệ xấp xỉ 50% toàn bộ danh sách từ ngữ các loại của từ điển này).

Rõ ràng, tính đơn tiết của từ tiếng Việt là một thuộc tính phổ biến, điển hình và mang đặc trưng rất rõ nét về mặt loại hình. Tuy nhiên, phổ biến, điển hình vẫn không phải là thuộc tính tuyệt đối để đi tới chỗ khẳng định: từ tiếng Việt chỉ là đơn tiết. Mặt khác, tuy thấy sự hiện diện của các từ đa tiết là hiển nhiên, nhưng đối với những đơn vị mà trước nay vẫn quen được gọi là từ ghép đẳng lập, từ ghép chính phụ, thì tư cách từ của hai loại này cũng cần phải được xem xét lại.

5. Những phân tích trên đây tuy chưa hoàn toàn đủ chi tiết và cặn kẽ, nhưng cũng đủ để cho thấy rằng: trong lịch sử tiếng Việt nói chung, trong sự hình thành và phát triển của từ vựng, ngữ pháp tiếng Việt nói riêng, vẫn còn có nhiều điều cần phải được tiếp tục tìm hiểu, nhận thức một cách toàn diện hơn để có thể đi tới được cái nhìn chân thực hơn đối với các sự kiện ngôn ngữ hữu quan.

TÀI LIỆU THAM KHẢO

1. Nguyễn Tài Cẩn, Giáo trình lịch sử ngữ âm tiếng Việt, Nxb GD, H., 1995.
2. Trần Nhật Chính, Sự phát triển của từ vựng tiếng Việt hiện đại (30 năm đầu thế kỉ XX: 1900 - 1930). *Luận án tiến sĩ ngữ văn*, H., 2002.
3. Phạm Đức Dương, Nguồn gốc tiếng Việt: từ tiền Việt Mường đến Việt Mường chung, *Sách: Tiếp xúc ngôn ngữ ở Đông Nam Á*, Viện Đông Nam Á, 1983, tr. 76 - 133.
4. Diffloth G., The lexical evidence for Austric, so far, *Oceanic linguistics*, Vol. 33, No. 2, 1994, 309 - 321 pp.
5. Diffloth G., Vietnamese as a Mon - Khmer language, *Papers from the first annual meeting of the Southeast Asian linguistics Society*, 1991, 125 - 139 pp.
6. Ferlus M., L'infixe instrumental RN en Khamou et sa trace en Viet-

t /

ra

là

hi

ng

ột

ng

ng

lại

nh

ng

ĩa

ng

đa

oá

từ

oto

namien, *Cahier de linguistique Asie Orientale* No. 2, Sep., 1977, pp. 51 - 55.

7. Ferlus M., Lexique Thvung - Francais. *Cahiers de linguistique Asia Oriental*, No. 5, Mars 1979, pp. 83-94.

8. Gohman V.I., *Istoricheskaja Phonetika Taiskikh Jazykov, Izd. Nauka, Moscva, 1992.*

9. Huffman F.E., An examination of lexicon correspondences between Vietnamese and some other Austro-asiatic languages, *Lingua*, Vol. 43, No. 2-3, Amsterdam, 1977, 171 - 198 pp.

10. Maspero H., Etude sur la phonetique historique de la langue annamite, Les initiales, *BEFEO*, Vol. 12, No. 1, pp. 1 - 127 (Bản dịch của Khoa ngữ văn, Đại học tổng hợp Hà Nội).

11. Nguyễn Phú Phong, Trần Trí Dõi, M. Ferlus, Lexique Vietnamien-Ruc - Francais, *Universite de Paris VII.*

12. Nguyễn Văn Lợi, Trung tố cấu tạo danh từ Proto Việt Mường và dấu vết của chúng trong tiếng Việt hiện đại, *T/c Ngôn ngữ*, Số 2, 1992, tr. 29 - 36.

13. Võ Xuân Trang, Người Rục ở Việt Nam, *Nxb Văn hoá dân tộc*, H., 1998.

14. Luận văn, khoá luận tốt nghiệp đại học chuyên ngành ngôn ngữ học của các sinh viên: Phạm Xuân Chiến, Đặng Thị Tiên Phong, Trần Thu Hằng, Nguyễn Văn Hạnh, Nguyễn Thị Minh Ngọc, Trần Hương Thục (*Tư liệu Khoa ngôn ngữ học, Trường đại học KHXH-NV, Đại học quốc gia Hà Nội*)

15. Đào Duy Anh, Hán Việt từ điển, *Nxb Trường Thi, Sài gòn*, 1937.

16. Hoàng Văn Hành, Từ điển từ láy tiếng Việt, *Nxb GD, H.*, 1995.

17. Nguyễn Bình Khiêm, Thơ Nôm; Thơ văn Nguyễn Bình Khiêm; *Nxb Văn học, H.*, 1983.

18. Hoàng Phê (chủ biên), Từ điển tiếng Việt, *Nxb KHXH- Trung tâm từ điển học, H.*, 1994.

19. Pignau de Behain, Tự vị An nam - La tinh, *Nxb Trẻ*, 1999.

20. Rhodes. A.de, Phép giảng tám ngày... *Tư liệu Khoa ngữ văn, Đại học tổng hợp Hà Nội.*

21. Nguyễn Trãi, Quốc âm thi tập; Nguyễn Trãi toàn tập; *Nxb KHXH, H.*, 1976.

22. Hồng đức quốc âm thi tập, *Nxb Văn học, H.*, 1962, 1982.