

tiếng việt và ngôn ngữ học hiện đại
sơ khảo về cú pháp



DŨNG VŨ

tiếng việt và ngôn ngữ học hiện đại
sơ khảo về cú pháp



V I E T

STUTTGART

2003

Copyright © 2000-2004 by Dung Vu.

Tác giả giữ bản quyền. Không một phần nào của cuốn sách này được phép tái dụng hoặc sao chép dưới bất kỳ hình thức hoặc phương tiện nào mà không có sự đồng ý của tác giả.

Xin chỉ sử dụng tài liệu này cho mục đích vô vụ lợi (giáo dục, nghiên cứu, ...). Xin liên lạc với tác giả: Tien-Dung.Vu@siemens.com

In lần thứ nhất, VIET Stuttgart – Germany, 2003

Phiên bản V02.00, VIET Stuttgart – Germany, 2004

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronically or mechanically, including photocopying, recording or any information storage or retrieval system, without written permission of the author.

First printing, VIET Stuttgart – Germany, 2003

Version V 02.00, VIET Stuttgart – Germany, 2004

MỤC LỤC

LỜI NÓI ĐẦU	I
CHƯƠNG 1: NGÔN NGỮ VÀ KHOA HỌC	1
<i>NGÔN NGỮ LÀ GÌ ?</i>	<i>1</i>
<i>ĐỐI TƯỢNG CỦA NGÔN NGỮ HỌC</i>	<i>4</i>
1. Đối tượng chất liệu và đối tượng hình thức của ngôn ngữ học	4
2. Sự hình thành đối tượng khoa học	5
3. Khái niệm ngôn ngữ theo Ferdinand de Saussure	9
4. Đối tượng của ngôn ngữ học theo Chomsky	10
<i>PHƯƠNG HƯỚNG NGHIÊN CỨU NGÔN NGỮ</i>	<i>12</i>
<i>NGÔN NGỮ KHÔNG BẮT ĐỊNH – SỰ THÀNH THẠO NGÔN NGỮ</i>	<i>13</i>
<i>CÚ PHÁP VÀ KHOA HỌC NHẬN THỨC</i>	<i>14</i>
CHƯƠNG 2: NGÔN NGỮ VÀ DIỄN ĐẠT – ĐẶC ĐIỂM CỦA TIẾNG VIỆT	17
<i>ĐỐI TƯỢNG</i>	<i>17</i>
<i>CÁCH PHẢN ÁNH ĐỐI TƯỢNG QUYẾT ĐỊNH CÁCH HÀNH NGÔN</i>	<i>17</i>
<i>VÀI ĐẶC ĐIỂM CHÍNH TRONG TIẾNG VIỆT</i>	<i>21</i>
1. Tính tĩnh tiến	21
2. Tính giàu thông tin cụ thể	27
3. Tính tĩnh lược	29
4. Tính cảm đề	32
5. Tính đề diễn	36
CHƯƠNG 3: SỰ THAY ĐỔI CỦA NGÔN NGỮ	41
<i>TIẾNG VIỆT XƯA</i>	<i>43</i>
1. Tính tĩnh tiến	50
2. Tính giàu thông tin cụ thể	52
3. Tính tĩnh lược	53
4. Tính cảm đề	54
5. Tính đề diễn	55
<i>TIẾNG VIỆT NAY</i>	<i>57</i>
<i>VĂN NÓI VÀ VĂN VIẾT</i>	<i>65</i>
CHƯƠNG 4: VẬT LIỆU VÀ CẤU TRÚC NGÔN NGỮ	67
<i>VẬT LIỆU NGÔN NGỮ</i>	<i>67</i>
<i>CẢM GIÁC VỀ MỘT CÂU NÓI ĐƯỢC</i>	<i>77</i>

MỤC LỤC

<i>VỊ TRÍ VÀ QUAN HỆ LỆ THUỘC CỦA TỪ</i>	81
<i>PHƯƠNG PHÁP PHÂN TÍCH CÂU</i>	89
1. Phép thế	90
2. Phép thử câu hỏi	91
3. Phép bố trí	91
4. Phép tỉnh lược	93
5. Phép hoán vị	94
6. Phép biến hình	95
<i>TÓM LƯỢC</i>	97
<i>ĐIỂM MỞ</i>	98
CHƯƠNG 5: CẤU TRÚC NGỮ ĐOẠN	99
<i>PHẠM TRÙ NGỮ PHÁP</i>	99
<i>PHẠM TRÙ TỪ VỰNG</i>	101
<i>PHẠM TRÙ CÚ PHÁP</i>	109
1. Ngữ đoạn danh từ NP	116
2. Ngữ đoạn tính từ AP	124
3. Ngữ đoạn giới từ PP	127
4. Ngữ đoạn động từ VP	128
5. Câu	132
<i>VÀI VẤN ĐỀ TRONG NGÔN NGỮ</i>	137
1. Tính hồi quy	137
2. Tính đa nghĩa	138
<i>TÓM LƯỢC</i>	140
<i>NHẬN XÉT VÀ ĐIỂM MỞ</i>	141
CHƯƠNG 6: CÂY CÚ PHÁP	143
<i>CÂY CÚ PHÁP</i>	143
1. Hình học của cây	143
2. Tính chi phối	144
3. Thế tiền vị	148
4. Lệnh-C	150
<i>CÁCH VẼ CÂY CÚ PHÁP</i>	151
1. Vẽ cây từ dưới lên trên	151
2. Vẽ cây từ trên xuống dưới	156
<i>NHẬN XÉT VÀ ĐIỂM MỞ</i>	159
CHƯƠNG 7: TÍNH RÀNG BUỘC	161
<i>SỰ RÀNG BUỘC</i>	163
<i>NHẬN XÉT VÀ ĐIỂM MỞ</i>	170
CHƯƠNG 8: LÝ THUYẾT X-GẠCH	171
<i>NGUYÊN TẮC X-GẠCH</i>	173

<i>VÀI VÍ DỤ X-GẠCH</i>	178
1. N-gạch	178
2. V-gạch	182
3. A-gạch	185
4. P-gạch	186
<i>TỔNG KẾT CÁC QUY TẮC X-GẠCH CỦA TIẾNG ANH</i>	187
1. Bình diện chính	188
2. Bình diện trung gian	188
3. Bình diện tận cùng	189
4. Tổng kết các quy tắc	189
<i>TÓM LƯỢC</i>	190
<i>ĐIỂM MỞ</i>	191
CHƯƠNG 9: LÝ THUYẾT X-GẠCH – BỔ NGỮ, PHỤ NGỮ	193
<i>BỔ NGỮ, PHỤ NGỮ</i>	193
<i>BIỆT ĐỊNH NGỮ</i>	215
<i>NGỮ ĐOẠN CHỈ ĐỊNH TỪ</i>	216
<i>TÓM LƯỢC</i>	220
<i>NHẬN XÉT VÀ ĐIỂM MỞ</i>	220
CHƯƠNG 10: LÝ THUYẾT X-GẠCH – CÁC LOẠI NGỮ ĐOẠN KHÁC	221
<i>THÔNG SỐ VÀ THỨ TỰ CỦA TỪ</i>	221
<i>NGỮ ĐOẠN THÌ VÀ NGỮ ĐOẠN TÁC TỬ BỔ NGỮ HÓA</i>	222
1. Mệnh đề	223
2. Ngữ đoạn tác tử bổ ngữ hóa CP	228
3. Ngữ đoạn thì	240
<i>TÓM LƯỢC</i>	242
<i>NHẬN XÉT VÀ ĐIỂM MỞ</i>	243
CHƯƠNG 11: TIỂU PHẠM TRÙ, TỪ VỰNG – X-GẠCH HẠN CHẾ	245
<i>TỪ VỰNG</i>	247
1. Vị từ và tham số	247
2. Tiểu phạm trù	250
3. Quan hệ chủ đề và vai trò “theta”	258
<i>TIỂU PHẠM TRÙ HÓA</i>	262
<i>X-GẠCH HẠN CHẾ</i>	264
<i>MÔ HÌNH NGỮ PHÁP</i>	265
<i>NGUYÊN TẮC CHIẾU MỞ RỘNG</i>	267
<i>TÓM LƯỢC</i>	268
<i>NHẬN XÉT VÀ ĐIỂM MỞ</i>	270
CHƯƠNG 12: PHÉP BIẾN HÌNH	271
<i>Ý NGHĨA CỦA PHÉP BIẾN HÌNH</i>	271

MỤC LỤC

<i>CHUYỂN VỊ ĐẦU TỚI ĐẦU</i>	279
1. Chuyển vị động từ V ($V \rightarrow T$)	279
2. Chuyển vị thì T ($T \rightarrow C$)	285
<i>QUY TẮC LỒNG</i>	289
<i>CHUYỂN VỊ NP</i>	290
<i>CHUYỂN VỊ - WH</i>	301
<i>TÓM LƯỢC</i>	309
<i>NHẬN XÉT VÀ ĐIỂM MỞ</i>	310
PHỤ LỤC	311
<i>BẢNG ĐỐI CHIẾU THUẬT NGỮ VIỆT - ANH</i>	313
<i>BẢNG ĐỐI CHIẾU THUẬT NGỮ ANH - VIỆT</i>	327
<i>TÀI LIỆU THAM KHẢO</i>	341
CHỈ LỤC	347

LỜI NÓI ĐẦU

Cú pháp học vốn là một bộ môn khoa học ngôn ngữ xuất xứ từ Âu châu, được xây dựng trên nền tảng ngôn ngữ Âu châu và được dùng để giải thích cấu trúc ngôn ngữ Âu châu.

Được du nhập vào Á Đông, trong đó có Việt Nam, bộ môn cú pháp học đã gặp nhiều trở ngại trong việc giải thích những ngôn ngữ lạ. Điều này rất tự nhiên, bởi lẽ không phải ngôn ngữ nào cũng giống ngôn ngữ Âu châu. Ngay cả trong nội bộ ngôn ngữ Âu châu cũng có điểm khác biệt.

Ứng dụng một lý thuyết không có giá trị toàn thể với hy vọng giải quyết được vấn đề thực tế của mình một cách hoàn hảo hiển nhiên là không thực tế. Lại càng không thực tế khi chưa hiểu rõ nó. Ngược lại, nhìn nó bằng đôi mắt thành kiến cũng không hẳn là đúng.

Thiết nghĩ, việc cần làm trước tiên khi muốn sử dụng một lý thuyết là nên nắm bắt thông tin về nó càng nhiều càng tốt, phân tích cho thật kỹ lưỡng. Chỉ bằng cách đó, mới mong rút tỉa được những điều bổ ích, thích hợp cho hoàn cảnh của mình, có thể giải quyết được vấn đề riêng của mình.

Trong tinh thần thông tin xây dựng và đồng thời để các bạn không chuyên có dịp tìm hiểu về tiếng Việt, chúng tôi xin giới thiệu các bạn đọc tập tài liệu này. Đây là tập thứ nhất thuộc chuỗi chuyên luận *Tiếng Việt và Ngôn Ngữ Học hiện đại*. Ngay bước đầu, chúng ta sẽ tìm cách:

- Nắm bắt những kiến thức căn bản về cú pháp học.
- Tìm hiểu về lý thuyết ngữ pháp tạo sinh (Generative Grammar).
- Tìm hiểu sự khác biệt của cấu trúc tiếng Việt so với cấu trúc ngoại ngữ.
- Phân tích những điểm mạnh, điểm yếu của lý thuyết ngữ pháp phương Tây. Đặc biệt, ghi nhớ những điểm không ứng dụng được cho tiếng Việt để từ đó nên tìm hướng nghiên cứu thích hợp hơn.
- Khai thác những điểm có thể ứng dụng tin học vào ngôn ngữ cho những mục đích cần thiết và hữu ích.

Để đối chiếu với tiếng Việt, chúng tôi chọn hai ngôn ngữ tiêu biểu có cấu trúc tương đối khác nhau: tiếng Anh và tiếng Đức. Tiếng Anh là ngôn ngữ thông dụng và tuyến tính gần giống tiếng Việt, trái lại là tiếng Đức.

Tài liệu dùng ký hiệu ngôn ngữ học theo truyền thống quốc tế. Có phần phụ trương, đối chiếu thuật ngữ Việt-Anh, Anh-Việt. Các thuật ngữ tạm dịch chỉ là ý kiến riêng trong khi chờ đợi sự góp ý, thống nhất. Ngoài ra, còn có nhiều điểm mở cho bạn đọc tự đặt vấn đề, tìm hiểu và mở rộng.

Ở những phần chưa vào sâu, chúng tôi sẽ cố gắng viết giản dị, kèm theo nhiều ví dụ để mọi người đọc được. Đối với các bạn đọc cho vui, nhất là giới văn thi sĩ, các bạn có thể đọc từ chương 1 đến chương 5. Các chương ấy sẽ cho chúng ta thấy sự khác biệt của ngôn ngữ là do đâu, cấu trúc của nó ra sao. Đặc biệt, chương 2 và 3 sẽ cho chúng ta một cái nhìn bao quát về một vài đặc điểm chính của tiếng Việt, nguyên tắc hành ngôn, cách diễn tả đối tượng của người Việt khác dân tộc khác thế nào, sự biến đổi của tiếng Việt xưa và nay, hiện tượng lai tiếng nước ngoài, ... hy vọng nhờ đó mà sẽ đỡ lẫn lộn cấu trúc tiếng Việt với cấu trúc ngôn ngữ khác. Tất nhiên tiếng Việt cũng có vấn đề và sẽ được nêu ra. Phần nào quá lý thuyết, bạn đọc không hiểu, cứ lướt qua. Các chương còn lại dành cho các bạn thích đào sâu, khảo cứu.

Sau hết, chúng tôi xin chân thành cảm ơn các viện, đại học, diễn đàn văn hóa/văn học nghệ thuật, các anh chị văn nghệ sĩ và bạn hữu đã nhiệt tình giúp đỡ, khuyến khích chúng tôi làm công việc này; đặc biệt, xin chân thành cảm tạ Prof. Dr. Wagner, Prof. Dr. Carnie, Prof. Dr. Lê Văn Đặng và chị Nguyễn Phương Lan (Viện Việt Học, Hoa Kỳ), Dr. Tanaka, Dr. Lohnstein, Bộ Văn Hóa Pháp, Dr. Trần Duy Trác (trang Văn Hóa Việt Nam, Hoa Kỳ), chị Phạm Chi Lan (trang Văn Học Nghệ Thuật VHNT, Hoa Kỳ), anh Phạm Việt, bạn Nguyễn Văn Quân, bạn Ian Bùi. Chúng tôi cũng mong đón nhận và cảm ơn tất cả mọi ý kiến đóng góp, phê bình xây dựng của bạn đọc.

Dũng Vũ

Stuttgart, 02.2003

Chương 1

NGÔN NGỮ VÀ KHOA HỌC

NGÔN NGỮ LÀ GÌ ?

Thượng Đế đã ban cho con người nhiều thứ thông minh, thế nhưng có lẽ một trong những thứ thông minh thuần túy người nhất vẫn là thẩm năng ngôn ngữ.

Là loài động vật cao cấp nhất hành tinh, con người tồn tại và phát triển nhờ ngôn ngữ. Không có diễn đạt nào không nhờ ngôn ngữ. Người bình thường dùng *ngôn ngữ tự nhiên (natural language)*. Hai người cầm nói chuyện với nhau bằng những ngón tay: *ngôn ngữ cơ thể (body language)*. Một bài toán đầy ký hiệu ngoằn ngoèo: *ngôn ngữ hình thức (formal language)*. Một chương trình được mã hóa bằng Basic, Java, C/C++, ...: *ngôn ngữ lập trình (programming language)*. Đến cả âm nhạc, hội họa, ... cũng đều có ngôn ngữ riêng. Có thể nói tổng quát, con người dùng ngôn ngữ như một phương tiện nhằm phản ánh vũ trụ, diễn đạt tình cảm, tư duy và để thông tin.

Trong thực tế, nếu nhìn thật trừu tượng, thì tình cảm, tư duy gì đi nữa, tất cả đều là *vấn đề*. Muốn trình bày vấn đề, con người dùng ngôn ngữ. Song ngôn ngữ không chỉ được dùng để *trình bày* vấn đề mà còn để *giải quyết* luôn cả vấn đề. Bởi lẽ một giải pháp cũng được trình bày thông qua ngôn ngữ.

Đã là vấn đề, hẳn phải có *độ phức tạp*. Vấn đề càng phức tạp, sự trình bày vấn đề càng phức tạp. Nhiều khi ngôn ngữ được sử dụng không thể diễn tả hết hoặc giải quyết hết mức độ phức tạp của vấn đề, bắt buộc người ta phải dùng loại ngôn ngữ khác phức tạp hơn. Vì thế mà ngôn ngữ còn được ví như một cây thước đo lường độ phức tạp của vấn đề.

Dùng ngôn ngữ để trình bày vấn đề, vậy muốn hiểu vấn đề, thì phải hiểu ngôn ngữ ? Đúng ! Thế thì làm thế nào để hiểu được ngôn ngữ ? Suy ra ngôn ngữ phải nhận diện được, ít nhất là phải có *quy tắc*.

Bước chân vào khoa học ngôn ngữ, công việc đầu tiên phải làm là đi định nghĩa “*Ngôn ngữ là gì ?*”. Trong lịch sử ngôn ngữ học đã có nhiều câu trả lời khác nhau. Edward Sapir định nghĩa:

“Ngôn ngữ là một *phương pháp truyền thông* thuần túy người và *phi bản năng*, được dùng nhằm biểu đạt *ý tưởng*, *cảm xúc* và *ước vọng* nhờ một *hệ thống ký hiệu tự tạo* (*system of voluntarily produced symbols*)” (Sapir 1921: 8).

Điểm nổi bật nhất của định nghĩa trên là *phương pháp truyền thông* (*method of communication*) và *hệ thống ký hiệu* (*system of symbols*). Song John Lyons lại cho rằng, suy nghĩ của Sapir vẫn chưa ổn lắm:

“Định nghĩa này còn vướng mắc nhiều khuyết điểm. Cho dù có cắt nghĩa các từ như “*ý tưởng*”, “*cảm xúc*” và “*ước vọng*” bao quát đến mấy, rõ ràng vẫn có những cái được truyền thông qua ngôn ngữ mà không nằm trong sự giải thích ấy; đặc biệt nữa là “*ý tưởng*”, tự bản chất nó đã không chính xác. Mặt khác, còn nhiều hệ thống ký hiệu tự tạo chỉ có thể được coi là ngôn ngữ theo nghĩa nới rộng hoặc ẩn dụ của từ “*ngôn ngữ*”. Ví dụ, cái thường được gọi là *ngôn ngữ cơ thể* (*body language*) – tức dùng cử chỉ, điệu bộ, ánh mắt – dường như thỏa điểm Sapir định nghĩa. Tính *thuần túy người* và *phi bản năng* có lẽ vẫn là điểm đáng ngờ. Và cũng đáng ngờ y thế là câu hỏi *liệu mọi ngôn ngữ nói chung đều thuần túy người và phi bản năng ?* Đây là điểm cần lưu ý nơi định nghĩa của Sapir” (Lyons 1981: 3f.).

Leonard Bloomfield, người đồng thời với Sapir, lại hiểu ngôn ngữ theo kiểu khác:

“Toàn thể mọi phát ngôn của một *cộng đồng tiếng nói* (*speech-community*) là ngôn ngữ của cộng đồng tiếng nói ấy” (Bloomfield 1926: 153).

Đối với Bloomfield, ngôn ngữ có nghĩa như một tập hợp bao gồm mọi phát ngôn chấp nhận được của một cộng đồng tiếng nói. Nhìn sơ, suy nghĩ ấy không khác suy nghĩ của Chomsky cho lắm:

“Kể từ nay, tôi sẽ coi mỗi ngôn ngữ là một tập hợp (hữu hạn hoặc vô hạn) chứa mọi câu, mỗi câu có độ dài giới hạn” (Chomsky 1957: 13).

Định nghĩa của Chomsky rộng và tổng quát. Theo quan điểm của Chomsky, mỗi ngôn ngữ trong trường hợp cụ thể còn tùy thuộc vào những phần tử đã dựng nên nó. Ví dụ, hằng đẳng thức $(a+b)^2 = a^2 + 2ab + b^2$ là một *câu* (*sentence*) của ngôn ngữ toán, trong khi đó $^2(a+b) = ^2a + +2ab ^2b$ không phải. ***Ngôn ngữ tự nhiên*** (***natural language***), tức ngôn ngữ con người, chỉ là một trường hợp đặc biệt:

“Mọi ngôn ngữ tự nhiên dưới dạng nói hay viết đều là những ngôn ngữ theo nghĩa này, vì mỗi ngôn ngữ tự nhiên đều có một số âm vị (phoneme) giới hạn (hay số chữ cái trong bộ mẫu tự) và mỗi câu đều có thể được trình bày dưới dạng một chuỗi giới hạn chứa những âm vị (hoặc chữ), mặc dù số câu không giới hạn. Tương tự vậy, một tập hợp “câu” trong vài hệ thống lập thức của toán cũng có thể được xem là ngôn ngữ” (Chomsky 1957: 13).

Sự khác biệt của Chomsky đối với những nhà ngôn ngữ học khác được John Lyons nhận xét:

“Điều đó không nói gì đến chức năng truyền thông của ngôn ngữ tự nhiên hoặc phi tự nhiên (non-natural); nó cũng không nói gì đến yếu tố *tự nhiên có tính cách biểu tượng* của những phần tử hoặc những chuỗi phần tử. Mục đích của nó là chú trọng vào những tính chất thuần cấu trúc của ngôn ngữ nhằm nói lên rằng, người ta có thể khảo sát những tính chất này từ những quan điểm chính xác của toán học. Điều đóng góp chính yếu của Chomsky cho ngôn ngữ học là ông đã đặc biệt nhấn mạnh được cái ông gọi là mối quan hệ *lệ thuộc cấu trúc (structure dependence)* trong tiến trình tạo thành câu của ngôn ngữ tự nhiên, và đã đưa ra được một lý thuyết ngữ pháp tổng quát tựa vào cơ sở định nghĩa cá biệt về tính chất ấy” (Lyons 1981: 7f.).

Nhà ngôn ngữ học Halliday lại xem ngôn ngữ là một cái gì “năng động” hơn một chút:

“Ngôn ngữ không hiện hữu, nó chỉ xảy ra. Nó không phải là bộ phận hữu cơ như nhiều nhà ngôn ngữ học trong thế kỷ 19 đã trông thấy vậy; nó cũng chẳng phải là một tòa nhà như người ta đã nhìn ngắm nó vào đầu thời “*cấu trúc luận*” hiện đại của ngành ngôn ngữ học. Ngôn ngữ là một hoạt động căn bản bao gồm bốn thứ: *nói, nghe, viết và đọc*” (Halliday et al. 1964: 9).

Tựu trung, mỗi nhà khoa học đều có cách định nghĩa ngôn ngữ từ những góc độ nhìn riêng:

- Hệ thống ký hiệu dùng để thông tin (Sapir 1921)
- Toàn thể phát ngôn chấp nhận được trong cộng đồng tiếng nói (Bloomfield 1926)
- Tập hợp câu (Chomsky 1957)
- Một hoạt động (Halliday et al. 1964)

“Ngôn ngữ là gì?”, tra cứu quanh câu hỏi này, chúng ta sẽ thấy còn nhiều quan điểm khác. (Xem từ điển ngôn ngữ học của Theodor Lewandowski 1990).

ĐỐI TƯỢNG CỦA NGÔN NGỮ HỌC

Đối với ngôn ngữ tự nhiên, nhiều nhà khoa học đã cố tránh định nghĩa ngôn ngữ một cách hiển ngôn. Thay vì thế, họ nhìn ngôn ngữ như một *đối tượng* nghiên cứu của ngành ngôn ngữ học. Nhà cấu trúc luận H. A. Gleason quan niệm:

“Ngôn ngữ có nhiều mối quan hệ với nhiều khía cạnh trong đời sống con người và có thể được nghiên cứu từ nhiều quan điểm khác biệt. Tất cả mọi quan điểm đều có giá trị và hữu ích, và tự chúng cũng rất lý thú. Ngôn ngữ học là một khoa học luôn cố gắng đi tìm hiểu ngôn ngữ dưới quan điểm cấu trúc nội tại của nó” (Gleason 1961: 2).

Hoặc R.H. Robins:

“Ngôn ngữ và tất cả các hình thức biểu lộ của nó, tức tất cả các ngôn ngữ trên thế giới và cách dùng chúng trong những trạng huống khác nhau của loài người, là những cái tạo ra môi trường cho nhà ngôn ngữ học. Nhà ngôn ngữ học là kẻ đi tìm hiểu một cách khoa học chỗ đứng của ngôn ngữ trong đời sống con người, những cách nó được tổ chức nhằm thỏa đúng những điều cần thiết nó cần làm và thỏa đúng chức năng nó giữ” (Robins 1964: 2 f.).

1. Đối tượng chất liệu và đối tượng hình thức của ngôn ngữ học

Vào khoảng đầu thế kỷ 20, Ferdinand de Saussure, người sáng lập chính của *cấu trúc luận* Âu châu, đã phân biệt hai yếu tố *chất liệu* và *đối tượng* của ngành ngôn ngữ học trong giáo trình *Cours de linguistique générale* (1916) của ông như sau:

“La *matière* de la linguistique est constituée d'abord par toutes les manifestations du langage humain” (Saussure 1916: 20).

“Bien loin que *l'objet* précède le point de vue, on dirait que c'est le point de vue qui crée l'objet” (Saussure 1916: 23).

Ferdinand de Saussure cho *cách nhìn* của người nghiên cứu ngôn ngữ là cái đặc biệt của khoa học có liên quan đến ngôn ngữ. Thực ra ý nghĩ này chẳng có gì

đặc biệt, bởi lẽ từ thời trung cổ, người ta đã biết phân biệt giữa ***đối tượng chất liệu (obiectum materiale)*** cụ thể và ***đối tượng hình thức (obiectum formale)*** trừu tượng.

Hiểu theo Leont'ev (1971: 15ff.), một nhà ngôn ngữ học tâm lý người Nga, thì ***đối tượng chất liệu (material object)*** của một ngành khoa học, cái mà bao gồm mọi hiện tượng thực tế khách quan cần được khảo sát một cách cụ thể, thực ra đã tồn tại trước khi có ngành khoa học. Nó hoàn toàn độc lập đối với ngành khoa học, người nghiên cứu, nhận thức và cách nhìn của người nghiên cứu. Dựa trên lý lẽ này, có thể nói, nhiều ngành khoa học khác nhau có thể có chung một đối tượng chất liệu, mặc dù góc độ nhìn nó hoặc nhận thức về nó có thể khác nhau. Ngôn ngữ là một hiện tượng đối tượng chất liệu điển hình đối với nhiều ngành như triết học, tâm lý học, xã hội học, ... Chính vì sự khác biệt về cách nhìn mà mỗi ngành khoa học mới có mục đích riêng, từ đó mới có một đối tượng hình thức riêng và được xem như đối tượng nghiên cứu của ngành khoa học ấy.

Xét về mặt cụ thể là như thế. Ngược lại, nếu xét về mặt ***hình thức (formal)***, thì hình thức chẳng qua là lập thức một cách trừu tượng những tính chất thật của một đối tượng dưới cái nhìn bao quát. Cũng theo Leont'ev, ***đối tượng hình thức (formal object)*** là tất cả mọi cái trừu tượng được dựng nên nhằm quan sát đối tượng chất liệu bằng nhiều góc độ nhìn và hướng nhận thức khác nhau.

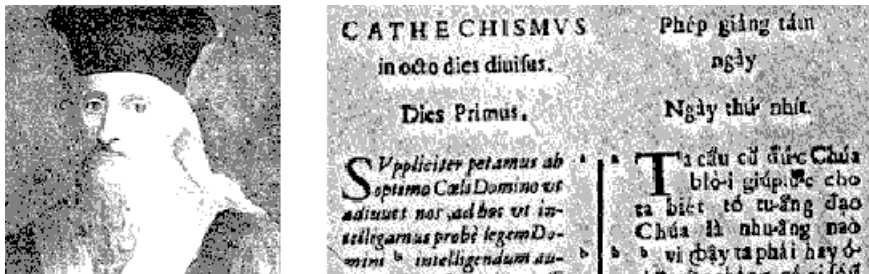
2. Sự hình thành đối tượng khoa học

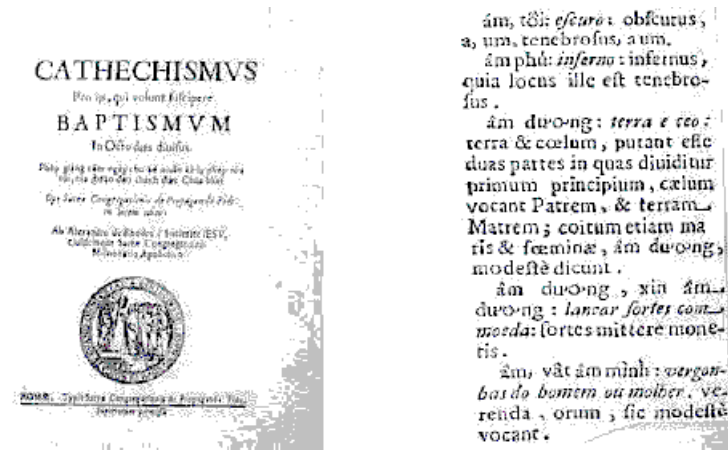
Sự hình thành đối tượng hình thức của một môn khoa học vào một thời điểm nhất định còn lệ thuộc vào những điều kiện khách quan như ***ý định chủ quan và khách quan, tình hình phát triển khoa học, cơ sở khoa học và nhận thức lý thuyết của người nghiên cứu.***

Từ thuở ban đầu, nhu cầu phát triển nghiên cứu ngôn ngữ đã nảy sinh từ những ***ý định chủ quan và khách quan*** của con người. Một mặt, nghiên cứu ngôn ngữ nhằm thỏa mãn tính hiếu kỳ về khoa học, mặt khác, không thiếu những câu hỏi đặc thù, quan trọng đối với xã hội được đặt ra và cần được trả lời. Một vài ví dụ lịch sử:

Ngay thời đại cổ Trung Hoa, khoa học ngôn ngữ đã ra đời từ sự khảo cứu những bản cổ văn nhằm thiết lập một hệ thống có thể diễn tả được hình tự tiếng Tàu. Ở Ấn Độ, cũng trong thời đại cổ, ý định chính của khoa học ngôn ngữ là làm sao

mô tả ngữ nghĩa, ngữ pháp và ngữ âm của tiếng *Phạn* (*Sanskrit*) hầu bảo tồn nét tinh nguyên của những “*văn bản thánh*” đã được truyền khẩu vào khoảng 1100 năm trước Công nguyên. Cũng chính nhờ công lao của các nhà ngữ âm học Ấn Độ xưa kia mà ngày nay người ta mới có thể hiểu cách phát âm của tiếng Phạn nhiều hơn bất kỳ cách phát âm của một loại tiếng cổ nào khác. Về phía Ả Rập, khoa học ngôn ngữ đã phát triển từ nhu cầu tôn giáo và hoàn cảnh thế giới thời bấy giờ. Sự bành trướng của Hồi giáo đến các xứ không theo đạo Hồi đã giúp tiếng Ả Rập thu thập thêm nhiều kiến thức ngôn ngữ mới, rất hữu ích cho mục đích chính trị, thương mại, quản lý và pháp luật. Cũng một trường hợp tương tự ở thời đại cổ nữa là Hy Lạp. Vào khoảng 300 năm trước Công nguyên, sau cuộc xâm lăng Tiểu Á và Ai Cập của Đại đế Alexandre, nền văn minh và văn hóa Hy Lạp đã lan rộng khắp nơi. Lúc đó, tiếng Hy Lạp đã được dùng phổ biến như một thứ tiếng mẹ đẻ thứ hai ngay tại thuộc địa. Thế nhưng dần dà những ngôn ngữ trong vùng lại trà trộn phát triển rồi biến thành một ngôn ngữ chung gọi là *Koiné*. Sự phát triển này dĩ nhiên đã gặp sức kháng cự của giới bảo thủ đang ráng giữ gìn và truyền bá tiếng Hy Lạp cổ (s.s. Robins 1973:8ff). Ở Việt Nam tựa vậy, xưa kia chữ Nôm đã ra đời thay cho chữ Hán. Ý nghĩa của nó dường như chỉ có tính biểu tượng muốn nói lên nền độc lập nước Nam nhiều hơn là tiện dụng. Là một loại hình tự vay mượn từ chữ Hán, chữ Nôm có khả năng phối triển thêm những âm vị mà chữ Hán không có. Chữ Nôm khá phức tạp và hầu như chỉ phổ biến trong giới học giả, trí thức, văn thi sĩ thời bấy giờ. Không thiếu những tác phẩm văn chương được viết bằng chữ Nôm còn để lại cho tới ngày nay: thơ Nguyễn Trãi, thơ Hồ Xuân Hương, Kiều của Nguyễn Du, Lục Vân Tiên của Nguyễn Đình Chiểu, ... Sau thời kỳ chữ Nôm, hình dáng chữ Việt lại thay đổi thêm lần nữa qua cách dùng mẫu tự La tinh do các nhà truyền giáo Tây phương đem tới. Một trong số các giáo sĩ như Christoforo Borri, Gasparo d’Amiral, Antonio Barbosa, ... , có thể nói, *Alexandre de Rhodes* là người được biết đến nhiều nhất qua quá trình La tinh hóa chữ Việt. (Đỗ Quang Chính 1972).





Giáo sĩ Alexandre de Rhodes và hai tác phẩm: *Cathéchismus* (Phép giảng tám ngày) (Roma 1651), *La Dictionarium Annamaticum Lusitium et Latinum* (Roma 1651). (Tài liệu: Bộ Văn Hóa Pháp. Paris, 1999)

Khách quan mà nói, dù là công lao của bất kỳ người đi trước nào đi nữa, việc phiên âm tiếng Việt dựa vào mẫu tự La tinh xưa kia chỉ thuần mục đích truyền giáo, cụ thể là để giúp cho các giáo sĩ hiểu tiếng Việt dễ dàng hơn khi giao lưu với dân bản xứ, trên thực tế chỉ là một cộng đồng rất nhỏ gồm giáo dân và giáo sĩ ở làng trong cũng như làng ngoài. Các nhà truyền giáo đã không hề nghĩ rằng, La tinh hóa tiếng Việt là nhằm lật đổ chữ Nôm và thay vào đó mô hình alphabet của phương Tây. Chính việc phát triển chữ Quốc Ngữ sau này chủ yếu do người Việt chủ động vì sự tiện dụng của nó.

Suốt hành trình phát triển lâu dài, ngôn ngữ còn lệ thuộc vào *tình hình phát triển chung của khoa học*. Ngành ngôn ngữ học đã thụ hưởng nhiều thành tựu của các ngành như vật lý, tâm lý học, tin học, ngôn ngữ học điện toán, ... hoặc sự chế tạo thành công máy thu âm, máy phân tích âm, máy vi tính, ... Trong các ngành có liên quan đến ngôn ngữ, đáng kể nhất, phải nói là ngành *tin học (informatics)* và *ngôn ngữ học điện toán (computational linguistics)*. Hàng loạt những ngôn ngữ lập trình đã nối tiếp nhau chào đời, từ những *ngôn ngữ máy (machine language)* thô sơ dưới dạng *Assembler* cho đến những *ngôn ngữ lập trình cao (high programming language)* Basic, Cobol, Fortran, PL/M, Pascal, Modula II, Prolog, ADA, C, C++, Java, C#, ... Như một phương tiện, ngôn ngữ lập trình đã đẩy mạnh mọi ngành khoa học cùng phát triển. Ngày nay không có ngành nào không hưởng lợi nhờ sự tiến bộ vượt bậc của ngành tin học, ngôn ngữ học và ngôn ngữ học điện toán. Từ trường học cho tới kỹ nghệ, tất cả đều cần ngôn ngữ lập trình cho mục đích giáo dục, tự động hóa, nâng cao chất lượng, năng suất sản xuất. Từ

văn phòng, nhà băng, nhà thương, nhà thờ, ... cho tới nhà tù, đâu đâu cũng cần cái máy computer cho việc quản trị, xử lý dữ liệu, trao đổi thông tin. Đặc biệt nữa là ngành ngôn ngữ học điện toán đã và đang phát triển tối đa việc xử lý ngôn ngữ tự nhiên. Bằng những phương pháp *trí thông minh nhân tạo (artificial intelligence)*, ngành này dần dần đã đạt tới khả năng làm cho máy có thể nhận diện, hiểu và nói được tiếng người. Việc ứng dụng ngôn ngữ tự nhiên không chỉ thấy qua các hệ thống phức tạp như *hệ nhận diện ngôn ngữ (language recognizing system)*, *hệ chuyên gia (expert system)*, mà còn thấy ở những chiếc xe lăn biết nói, dịch vụ điện thoại hiểu nhiều thứ tiếng, ... Ngay đến cả đồ chơi trẻ con, ví dụ, một con búp bê cũng biết nói chuyện với đứa bé đang chơi với nó. Phải nói, mọi phát triển đều dựa vào ngôn ngữ. Thời đại thông tin đã bùng nổ dữ dội nhờ ngôn ngữ. Một ứng dụng cụ thể hẳn ai cũng thấy là **Internet**. Sức mạnh của nó như thế nào hẳn ai cũng thấy. Nhờ Internet mà ngày nay quả đất đã được thu nhỏ lại. Nhờ nó mà con người có thể giao lưu với nhau dễ dàng hơn, học hỏi dễ dàng hơn, đỡ tốn kém hơn, ... Nhờ ngôn ngữ, khoa học phát triển, thông tin phát triển, kỹ nghệ phát triển, kinh tế phát triển và đòi hỏi tri thức phải phát triển. Chưa bao giờ nhân loại làm chủ nhiều tri thức như hôm nay. Tri thức là chất xám của con người, làm sao có thể lập thức, ứng dụng được chất xám ấy nếu không nhờ ngôn ngữ? Dùng phương tiện gì để cất giữ, chuyển tải chất xám nhanh chóng nhất nếu không nhờ khoa học thông tin. Nền kinh tế tri thức đã hình thành mà tri thức là nhân tố chính. Tin học sẽ còn thúc đẩy kinh tế tri thức phát triển mạnh mẽ hơn nữa vào thế kỷ 21. Nhưng hãy nên cẩn thận, tin học vừa có thể là một năng tiên vừa có thể là con ma của nhân loại. Xưa nay nhân loại đã không ngừng phát triển và mọi phát triển – bao giờ cũng thế – vẫn được biện minh là nhằm phụng sự con người.

Có vài câu hỏi. Nếu xem ngôn ngữ là đối tượng hình thức của nhiều ngành khoa học khác nhau, thì ngôn ngữ sẽ được xử lý thế nào? Cái gì biện minh cho đặc điểm của từng ngành? Các câu hỏi này còn tùy vào *cơ sở khoa học và nhận thức lý thuyết của người nghiên cứu*.

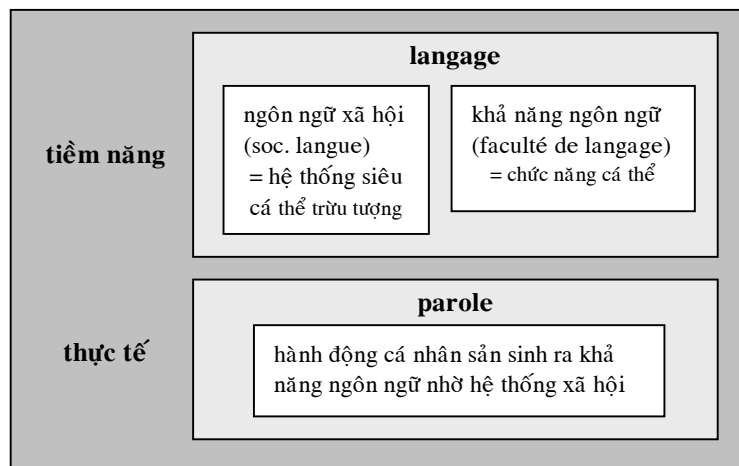
Để giải thích ý nghĩa của ngôn từ, ngành ngôn ngữ học có bộ môn **ngữ nghĩa học (semantics)**. Dĩ nhiên làm rõ ý nghĩa nội dung là một việc không thể thiếu đối với những ngành có liên quan đến ngôn ngữ văn tự, ví dụ như triết học, tâm lý học, xã hội học, sinh học, ... Tuy nhiên, cách xử lý ngữ nghĩa ở đó không giống như ở ngành ngôn ngữ học. Đối tượng ở đó không phải là chính bản thân ngôn ngữ mà là một cơ sở lý thuyết khoa học khác. Về phía ngôn ngữ học, bản thân ngôn ngữ chính là cái để người ta thắc mắc và tìm hiểu, nói chung là nhận thức. Xét cho cùng, mục đích của những ngành khác ngoài ngôn ngữ học không

phải là đi giải thích ngôn ngữ hoặc cung cấp một cách có hệ thống những gì có liên quan đến ngôn ngữ. Công việc đó là của ngành ngôn ngữ học.

Ngoài các ngành có liên quan đến *ngôn ngữ* vừa nêu, trên thực tế còn có cả các ngành có liên quan đến *ngôn ngữ học* như ngôn ngữ học triết học, ngôn ngữ học xã hội, ngôn ngữ học tâm lý, ngôn ngữ học sinh học, ... Trong khi ngôn ngữ học cố gắng giới hạn ngành học của mình bằng cách, một là chú trọng vào phạm vi nghiên cứu cấu trúc nội tại của ngôn ngữ, hai là phân biệt ranh giới với những ngành khác như thể muốn tự cô lập và thu nhỏ đối tượng (ngôn ngữ) của mình, thì các ngành phụ khác, ít nhất là trong vài thập niên qua, lại tìm cách “xào nấu ngôn ngữ” bằng đủ loại công thức, nào là ngôn ngữ và xã hội, nào là ngôn ngữ và tư duy, nào là ngôn ngữ và sinh học, nào là ngôn ngữ và chính trị, ... Kết quả, có rất nhiều điểm chông chéo ngoài những đặc điểm tiêu biểu đại diện cho bản thân một ngành. Có thể nói, các ngành vừa nêu “cũng” làm việc với “ngôn ngữ”. Ngôn ngữ học vốn dĩ chỉ làm việc với ngôn ngữ, giờ đây cũng phải phụ tay giúp những anh đầu bếp ấy.

3. Khái niệm ngôn ngữ theo Ferdinand de Saussure

Ảnh hưởng của Saussure chủ yếu bắt nguồn từ tập *Cours de linguistique générale* đã được học trò của ông phổ biến sau khi ông mất vào năm 1916. Chỉ có điều, cho đến nay vẫn chưa ai rõ là văn bản này có thực sự phản ánh cái nhìn của ông về ngôn ngữ như vậy hay không ngoại trừ Leont’ev (1971: 19ff) đã cố gắng trình bày lại như sau:



Trong mô hình Saussure, *ngôn ngữ (la langue)* là một hệ thống siêu cá thể trừu tượng bao gồm mọi *ký tự (character)*; còn *khả năng ngôn ngữ (faculté de langage)* là chức năng cá thể (Saussure dùng chữ “*langage*” thay vì “*langue*”). Ngôn ngữ là cái sở hữu chung của cộng đồng tiếng nói và là một bản thể độc lập đối với từng cá nhân cũng như đối với cộng đồng. Mặt khác, nó lại là sản phẩm của khả năng ngôn ngữ cá thể, bởi mỗi thành viên của cộng đồng tiếng nói đều có khả năng sao chép ngôn ngữ. Bao hàm hai ý nghĩa “*faculté de langage*” và “*langue*”, khái niệm “*langage*” được hiểu như tiềm năng thành lập “*lời nói (parole)*” mà trong thực tế mỗi con người đều có. Một mặt, hành động nói đòi hỏi khả năng ngôn ngữ, mặt khác, đòi hỏi sự hiểu biết về hệ thống ngôn ngữ. Theo Saussure, nhiệm vụ chính của ngôn ngữ học đúng nghĩa là nên khảo sát ngôn ngữ như một hình thức “*langue*” trong cấu trúc nội tại của nó.

Hệ thống ngôn ngữ của Saussure đã ảnh hưởng sâu rộng đến tư duy của nhiều nhà ngôn ngữ học thời sau. Đặc biệt, bài bản của Saussure đã được nhà ngôn ngữ học người Đan Mạch Louis Hjelmslev (1899-1965) nghiên cứu rất kỹ lưỡng.

4. Đối tượng của ngôn ngữ học theo Chomsky

Từ cuối những năm 50, ngôn ngữ học Âu châu đã phát triển rực rỡ nhờ công lao của nhiều người. Ngoài một tên tuổi tiêu biểu như Saussure, hẳn phải kể đến Chomsky, cha đẻ của lý thuyết ***ngữ pháp biến hình tạo sinh (Generative Transformation Grammar)***.

Đối với Chomsky, đối tượng của một lý thuyết ngôn ngữ được hiểu như một *người nói/nghe lý tưởng* sống ở một cộng đồng tiếng nói hoàn toàn đồng nhất, rất thành thạo ngôn ngữ của mình và khi sử dụng kiến thức ngôn ngữ để hành ngôn, dù có gặp những điều kiện trở ngại như kém trí nhớ, lẩn thẩn, nhầm lẫn, thiếu tập trung, song ngữ pháp vẫn không bị ảnh hưởng lắm (s.s. Chomsky 1965: 13). Cộng đồng tiếng nói đồng nhất được hiểu là một cộng đồng không dùng *tiếng địa phương (dialect)*, *tiếng xã hội* (tiếng bồi, tiếng lóng).

Trong cuốn *Aspects of the Theory of Syntax* (1965), Chomsky đã đưa ra hai khái niệm *competence* và *performance* để giải thích khả năng thành thạo và hiệu năng sử dụng ngôn ngữ của con người. Ông phân biệt, ***thẩm năng ngôn ngữ (language competence)*** là sự thành thạo ngôn ngữ của một người nói/nghe lý tưởng sống trong một cộng đồng tiếng nói đồng nhất. Những kiến thức này tạo nên một hệ thống quy tắc và nguyên tắc tinh thần cho phép người nói/nghe nhờ

làm chủ một tập hợp hữu hạn của “âm”, “từ” để từ đó có khả năng thành lập và hiểu đúng vô số những phát biểu cũng như đánh giá được tính đúng sai của ngữ pháp, tính đa nghĩa, đồng nghĩa của câu. ***Dụng năng ngôn ngữ (language performance)*** được hiểu là hiệu năng sử dụng đúng ngôn ngữ của người nói/nghe lúc hành ngôn. Khả năng này thường bị giới hạn do kém trí nhớ, thiếu tập trung, mệt mỏi, say rượu, ...

Theo Chomsky là vậy. Thế nhưng trong thực tế vẫn có trường hợp sử dụng ngôn ngữ không sai ngữ pháp do say rượu, lú lẫn, thất tình hay chán đời gì cả mà vẫn đòi hỏi người nghe phải thành thạo ngôn ngữ lắm mới hiểu nổi. Một ví dụ:

(1) Phượng hiểu Mai hơn Cúc.

Câu này có hai cách hiểu:

(2) Phượng hiểu Mai hơn là Cúc *hiểu Mai*.

(3) Phượng hiểu Mai hơn *hiểu* Cúc.

Hai ý nghĩa khác nhau hoàn toàn chỉ vì thiếu chữ. Thiếu *ngữ cảnh*, câu nói trở nên khó tưởng tượng. Nếu cũng cùng một cấu trúc y hệt vậy mà nói:

(4) Hùng thích Mai hơn Cúc.

thì có lẽ xác suất hiểu: “*Hùng thích Mai hơn thích Cúc*” cao hơn, bởi chuyện con trai yêu thích con gái chắc (vẫn còn) bình thường hơn chuyện con gái yêu thích con gái. Nhờ ngữ nghĩa phái tính nam nữ, câu nói dễ tưởng tượng hơn. Dù vậy, điểm quyết định vẫn là sự đòi hỏi người nghe phải vận dụng trí thông minh để nhận đoán ý nghĩa.

Vì những hiện tượng kể trên, khái niệm *competence* của Chomsky thường bị phê phán là thiếu để ý đến những câu tối nghĩa thường gặp ở ngoài đời. Từ đó đã xuất hiện một khái niệm mới: ***communicative competence (thẩm năng giao lưu)*** (Habermas 1971; Hymes 1972). Chính Chomsky cũng tự thú nhận, lý thuyết hiểu biết và tiếp thụ ngôn ngữ phải được bổ túc từ những lý thuyết ứng dụng ngôn ngữ. Với suy nghĩ ấy, ông đưa ra thêm khái niệm *pragmatic* để phân biệt giữa ***grammatical competence (thẩm năng ngữ pháp)*** và ***pragmatical competence (thẩm năng ngữ dụng)***. *Thẩm năng ngữ dụng* được hiểu là khả năng ứng dụng sự thành thạo ngữ pháp để lập câu chính xác, hợp lý cho mỗi trường hợp và mục đích tương ứng.

PHƯƠNG HƯỚNG NGHIÊN CỨU NGÔN NGỮ

Những gì nói được, những gì không, bao giờ, đối với ai, ... chỉ là vài vấn đề của ngôn ngữ con người. Còn rất nhiều vấn đề ngữ nghĩa khác có liên quan đến tính đúng sai của câu cú. Nghiên cứu ngôn ngữ, phải nói, cú pháp là một đề tài sâu rộng và không dễ nuốt. Tuy vậy, trước khi đi sâu vào lý thuyết, chúng ta hãy thử tìm hiểu sơ qua hướng nghiên cứu của các nhà ngôn ngữ học xưa nay thế nào.

Trong những công việc nghiên cứu tương đối mới sau này, Chomsky đã giới thiệu hai khái niệm đại diện cho hai quan điểm đối lập của giới ngôn ngữ học tạm gọi là *ngôn ngữ ngoại hóa (externalized language)* và *ngôn ngữ nội hóa (internalized language)* (Stechow, 1988: 28).

Giới *ngôn ngữ ngoại hóa* hầu hết được biết đến qua truyền thống *cấu trúc luận* theo kiểu *phân loại (taxonomy)* của Mỹ. Mục đích của nhóm này là cố gắng gom tụ thật nhiều những thử nghiệm ngôn ngữ ngoài đời (corpora) rồi tìm cách giải thích. Ngôn ngữ ngoại hóa được hiểu như một cuốn album gồm vô số những câu cú nói được *hoàn toàn độc lập với tâm trí con người*. Dựa vào đây, công việc kế tiếp của người nghiên cứu là thiết lập một ngữ pháp tương ứng. Nói cách khác, nhiệm vụ của nhà ngôn ngữ học ngoại hóa là dùng ngữ pháp để đưa mọi dữ kiện nằm rải rác bên ngoài vào vòng trật tự. Loại ngữ pháp này không có gì khác hơn là một tập hợp toàn những *mẫu cấu trúc (structure pattern)* được dựng nên nhằm giải thích những dữ kiện ngôn ngữ.

Giới *ngôn ngữ nội hóa*, ngược lại, đặt câu hỏi, người nói hiểu gì về ngôn ngữ và những hiểu biết ấy từ đâu ra. Giới này cho ngôn ngữ là một *tính chất tinh thần nằm bên trong con người* chứ chẳng phải là cái gì “ngoại hóa” (bên ngoài) cả. Ngữ pháp xác định được nhờ *bài bản và thông số*.

Chomsky còn khẳng định thêm, cách nhìn về ngôn ngữ của ngành ngôn ngữ học trong suốt lịch sử cận đại đã có đà chuyển hướng từ ngôn ngữ ngoại hóa sang ngôn ngữ nội hóa, ở đó, ngôn ngữ nội hóa được hiểu như một hệ thống mô tả từng cá thể tinh thần. *Ngữ pháp* mô tả sự hiểu biết về ngôn ngữ chứ chẳng phải *câu*. Tiền đề này đã tự chứng minh được sự thành công của nó qua nhận định, ngữ pháp có khả năng nắm rõ và giải thích được một khả năng tinh thần của con người, đó là sự hiểu biết về ngôn ngữ.

Nói tóm lại, hoạt động của giới ngôn ngữ ngoại hóa là cố gắng gom nhặt thật nhiều cách nói có thật ngoài đời, trong khi đó, giới ngôn ngữ nội hóa lại đưa ra

câu nói được, câu không nói được rồi nhờ người hành ngôn kiểm chứng. Giới ngôn ngữ ngoại hóa trách giới ngôn ngữ nội hóa coi thường dữ kiện thực. Ngược lại, giới ngôn ngữ nội hóa cười phía kia là chỉ biết nghiên cứu toàn những thứ tầm thường dễ như đếm đậu. Chomsky thuộc giới “*nội hóa*”.

NGÔN NGỮ KHÔNG BẮT ĐỊNH – SỰ THÀNH THẠO NGÔN NGỮ

Ngôn ngữ là một tập hợp câu cú nói được. Thử tưởng tượng, có một số từ, ta có thể hoán vị thành nhiều chuỗi từ. Hoán vị 5 từ, ta có 120 chuỗi từ (theo công thức $5!$ (5 giai thừa)). Gấp đôi 5 là 10, kết quả sẽ nhảy vọt từ 120 lên đến 3.628.800 ($10!$). Tuy nhiên, trong số lượng khổng lồ hơn ba triệu chuỗi từ ấy, không phải chuỗi từ nào cũng đều nói được. Xem một phần kết quả thử nghiệm với 5 từ của Lê Văn Lý (1948):

- (5) Sao nó bảo không đến ?
 Sao bảo nó không đến ?
 Sao không đến bảo nó ?
 Sao không bảo nó đến ?
 Sao ? Đến bảo nó không ?
 Sao ? Bảo nó đến không ?
 Nó đến, sao không bảo ?
 Nó đến, không bảo sao ?
 Nó đến bảo không sao.
 Nó bảo sao không đến ?
 Nó đến, bảo sao không ?
 Nó bảo đến không sao.
 Nó bảo không đến sao ?
 Nó không bảo, sao đến ?
 Nó không bảo đến sao ?
 Nó không đến bảo sao ?
 Bảo nó sao không đến ?
 Bảo nó: Đến không sao.
 Bảo sao nó không đến ?
 Bảo nó đến, sao không ?
 Bảo nó không đến sao ?
 Bảo không, sao nó đến ?
 Bảo ! Sao, nó đến không ?
 Không bảo, sao nó đến ?

Không đến bảo nó sao ?
 Không sao, bảo nó đến.
 Không bảo nó đến sao ?
 Không đến, bảo nó sao ?
 Không đến, nó bảo sao?
 Đến bảo nó không sao.
 Đến không ? Bảo nó sao ?
 Đến không ? Nó bảo sao ?
 Đến, sao không bảo nó ?
 Đến bảo nó sao không ?
 Đến, sao nó không bảo.
 Đến, nó bảo không sao.
 Đến, nó không bảo sao ?
 Đến, sao bảo nó không ?

Tạm cho 38 chuỗi từ trên đều nói được, số còn lại không, song thế nào là nói được ? Người Việt còn có cảm giác ấy, nhưng người ngoại quốc không rành tiếng Việt thì sao ?

Theo Chomsky, người có thẩm năng ngôn ngữ được hiểu như người nói/nghe lý tưởng thành thạo ngôn ngữ, có khả năng xác nhận một lời nói có thực sự là một câu của tiếng nói đó hay không. Tuy vậy, sự thành thạo ấy không hiển ngôn mà chỉ *thuần trực giác*. Bởi lẽ nếu đặt câu hỏi tại sao câu nói được, chắc chắn sẽ có người chẳng biết phải giải thích thế nào. Hoặc, cho dù một người có đủ kiến thức nói đúng, viết đúng, nhưng vẫn có thể chẳng hiểu tại sao cái cấu trúc ngữ pháp lại có hình thù và đặc tính ấy. Dựa vào đâu để có thể giải thích sự khác biệt giữa hai câu “*Nó đến, sao bảo không ?*” và “*Nó đến, sao không bảo ?*” một cách hiển ngôn ? Dựa vào đâu để giải thích, người Việt không ai nói “*Nó đến không sao bảo*” ? Cho nên cái gọi là khả năng thành thạo ngôn ngữ của người nói/nghe mà Chomsky gọi là lý tưởng hầu như chỉ là một thẩm năng thuộc về vô thức chứ không phải là tri thức để có thể giải thích thật khoa học cái cấu trúc ngôn ngữ mà mình thông thạo. Chính vì vậy, khoa học cú pháp đã hình thành.

CÚ PHÁP VÀ KHOA HỌC NHẬN THỨC

Cú pháp là gì ? Thuật ngữ *cú pháp* (*syntax*) được Nguyễn Kim Thản giải thích trong cuốn *Nghiên cứu ngữ pháp tiếng Việt* như sau:

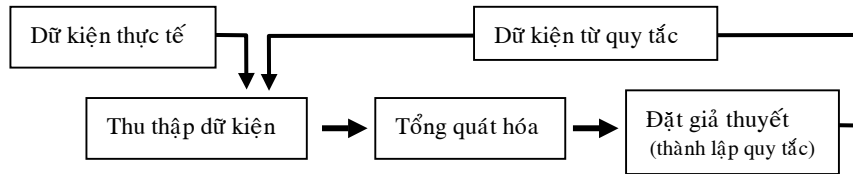
“Cú pháp nghiên cứu cách ghép từ thành ra những từ tổ và câu”. (Nguyễn Kim Thản 1997: 12)

Nguyễn Kim Thản đã giải thích cú pháp theo góc nhìn về cách tổ chức ngôn ngữ. Rộng rãi hơn, cú pháp trong ngôn ngữ học hiện đại còn được hiểu như một ***khoa học nhận thức (cognitive science)*** (Carnie 2000). Khoa học nhận thức bao gồm mọi ngành khoa học có cùng mục đích chuyên mô tả và cắt nghĩa khả năng suy nghĩ của con người. Có những cái, con người làm đúng nhưng chẳng biết phải cắt nghĩa thế nào cho đúng. Có những cái trừu tượng chứa *logic mờ (fuzzy logic)*, con người hiểu nhưng chẳng biết phải cắt nghĩa thế nào cho người khác hiểu. Hoặc giả như có khả năng cắt nghĩa, thì cũng chẳng biết phải dựa vào một hệ thống lý thuyết khoa học nào để mà cắt nghĩa. Ngôn ngữ tự nhiên là một vấn đề tiêu biểu. Làm thế nào để cắt nghĩa cái suy nghĩ của con người được phản ánh qua ngôn ngữ tự nhiên? Ngôn ngữ tự nhiên vốn là một sản phẩm có cấu trúc tinh vi, phức tạp do chính con người làm chủ hoặc để ra, nhưng ngộ cái là con người đã không hề nghĩ ra bất kỳ lý thuyết, bài bản hay quy tắc gì trước đó. Vì thế mà khoa học nhận thức đã đảm nhận công việc đi tìm hiểu, mô tả và giải thích cái suy nghĩ của con người. Nói cụ thể, mục đích của khoa học ***cú pháp (syntax)*** là đi khám phá các quy tắc đã tạo thành ngôn ngữ và sau đó sẽ tìm cách mô tả, giải thích chúng cho thật khoa học và có hệ thống. Những quy tắc ấy chính là ***ngữ pháp (grammar)***. Mục đích sau cùng mà khoa học cú pháp phải đạt tới là xác định tính đúng sai của cách hành ngôn; tất cả mọi quy tắc hình thức được tìm thấy phải phản ánh chính xác tinh thần ngôn ngữ trong thực tế và phải trùng hợp với tập hợp ngôn ngữ tự nhiên.

Có lẽ cách nhìn về cú pháp trong giới ngôn ngữ học quốc tế đồng nghĩa với *ngữ pháp học* của Nguyễn Kim Thản:

“Ngữ pháp là thuật ngữ dịch từ “*grammaire*” (Pháp), “*grammar*” (Anh), v.v... mà ra. Thuật ngữ “*grammaire*”, “*grammar*” ngày nay có hai ý nghĩa: một là hệ thống ngữ pháp của một ngôn ngữ cụ thể (như: ngữ pháp tiếng Việt ...), hai là khoa học nghiên cứu về ngữ pháp nói chung. Các nhà ngôn ngữ Trung Quốc phân biệt rõ ràng hai ý nghĩa đó, và khi nói đến cái nghĩa thứ hai, thì gọi là “*ngữ pháp học*”. Chúng tôi cũng chủ trương phân biệt hai nghĩa trên bằng hai thuật ngữ khác nhau”. (Nguyễn Kim Thản 1997 : 11)

Làm việc với cú pháp theo nghĩa khoa học nhận thức, người nghiên cứu ngôn ngữ sẽ trải qua ba bước. Bước một: *nắm bắt dữ kiện thực tế bên ngoài*, bước hai: *tổng quát hóa*, bước cuối cùng: *đặt giả thuyết*.



Những giả thuyết được xem như những quy tắc sẽ được thử nghiệm tới lui nhằm tìm thêm dữ kiện mới. Có thêm dữ kiện mới từ quy tắc cũ, chu trình mới lại bắt đầu, người quan sát lại tiếp tục tìm kiếm thêm những dữ kiện thực tế từ thế giới bên ngoài. Đã nắm bắt xong mọi dữ kiện mới và cũ, khâu tổng quát hóa lại hồi quy, người thử nghiệm lại đặt giả thuyết “chắc là đúng”, rồi lại tiếp tục thử nghiệm. Chu trình ấy cứ được lặp tới lặp lui cho tới khi nào những giả thuyết đặt ra được xem như đúng thì thôi. Nếu giả thuyết đúng, người ta nghĩ, những quy tắc lập thành phải đúng (tức quy tắc ngữ pháp). Đây là nguyên tắc làm việc của khoa học kinh nghiệm thường thấy.

Đã có nhiều người dấn thân vào công việc kể trên. Một trong những người tiên phong phải kể đến là Chomsky, nhà ngôn ngữ học nổi tiếng của thế kỷ 20 người gốc Do Thái. Vào giữa những năm 50, Chomsky đã cho ra đời một lý thuyết ngữ pháp được gọi là *ngữ pháp tạo sinh* (*Generative Grammar*) (Chomsky 1965, 1979). *Ngữ pháp tạo sinh* là một mô hình ngữ pháp diễn tả một tập hợp quy tắc tạo thành ngôn ngữ. Khái niệm *tạo sinh* đã bắt nguồn từ ý nghĩa ấy. Lý thuyết này đã một thời gây tiếng vang trong làng ngôn ngữ học và vẫn còn tạo ảnh hưởng cho tới ngày nay. Hàng loạt những lý thuyết ngữ pháp theo sau của ông và của những người khác đều dựa vào nguyên tắc ấy: *Ngữ pháp biến hình tạo sinh* (*Generative Transformation Grammar*), *Lý thuyết tiêu chuẩn* (*Standard Theory*), *Lý thuyết tiêu chuẩn mở rộng* (*Extended Standard Theory*), *Lý thuyết chi phối và ràng buộc* (*Government and Binding Theory (GB)*), *Nguyên tắc và thông số* (*Principles and Parameters (P&P)*), *Tối thiểu luận* (*Minimalism (MP)*), *Ngữ pháp cấu trúc ngữ đoạn tổng quát hóa* (*Generalized Phrase Structure Grammar (GPSG)*) (Gazdar et al. 1985).

Song, không phải mọi người đều đồng ý với cách nhìn ngôn ngữ theo kiểu cấu trúc của Chomsky. Vào khoảng cuối thập niên 70, Bresnan và Kaplan đã bày tỏ tư tưởng đối lập của mình qua lý thuyết *Ngữ pháp chức năng từ vựng* (*Lexical Functional Grammar*) (Bresnan 1982). Dung hòa hơn, Pollard và Sag đã cho ra đời lý thuyết lừng danh *Head-Driven Phrase Structure Grammar (HPSG)* (Pollard, Sag 1994).

Chương 2

NGÔN NGỮ VÀ DIỄN ĐẠT –
ĐẶC ĐIỂM CỦA TIẾNG VIỆT**ĐỐI TƯỢNG**

Một trong những mục đích của con người khi dùng ngôn ngữ là nhằm diễn đạt đối tượng. Đối tượng có thể hiểu là mọi thứ trong vũ trụ có thể tưởng tượng được đối với mỗi con người còn khả năng cảm nhận. Có thứ sờ mó được, có thứ không. Có thứ không sờ mó được nhưng thấy được. Có thứ không thấy được nhưng nghe được. Có thứ không nghe được nhưng cảm được, hình dung được, ... Cái bàn, cái ghế, ... là những thứ sờ mó được. Mặt trời, mặt trăng là những thứ xa tầm tay nhưng không vượt khỏi tầm mắt. Cả âm thanh cũng không thể thoát khỏi lỗ tai một người không bị điếc. Trừu tượng nữa là những gì không thấy, không nghe, không sờ mó, ngửi, nếm được. Ví dụ: tình yêu. Làm sao thấy được, nghe được, sờ mó, ngửi, nếm được tình yêu ? Chỉ nhờ cảm giác, con người mới mong cảm nhận được đối tượng này. Hoặc là tư duy. Làm sao nhìn thấy được tư duy ?

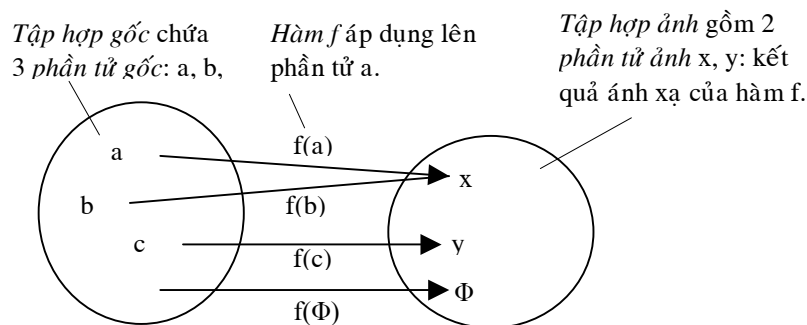
Đối tượng không những chỉ là cái bàn, cái ghế, mặt trời, mặt trăng, ... tình yêu, mà là mọi thứ không gian, thời gian, là điều kiện, sự tình, hiện tượng, là tư tưởng, tinh thần, hỉ, nộ, ái, ố, ... là vân vân và vân vân, nói chung là mọi thứ tồn tại trong cái thế giới khả giác của con người, bất kỳ cụ thể hay trừu tượng mà người ta có thể xem đó là một yếu tố đóng vai khởi điểm cho việc thuyết minh, hoặc một *chủ đề* và bằng cách nào đó, muốn *diễn đạt* nó.

CÁCH PHẢN ÁNH ĐỐI TƯỢNG QUYẾT ĐỊNH CÁCH HÀNH NGÔN

Ngôn ngữ khác nhau một phần cũng là do cách phản ánh khác nhau về đối tượng, từ đó, cách hành ngôn, câu cú mới khác nhau.

Cùng một đối tượng, có người diễn tả kiểu này, có người diễn tả kiểu khác. Cách nói của người Việt: “*một cô gái đẹp*” chắc chắn phải khác cách nói của người Anh: “*a pretty girl*”, hay người Đức: “*ein schönes Mädchen*”. Đối tượng ở đây là *cô gái*.

Khi muốn mô tả một đối tượng, dĩ nhiên người mô tả phải biết đối tượng ấy, nghĩa là trước hết phải nhận diện được nó, rồi sau đó mới xem nó như thế nào, nghĩa là nhận thức về nó. Nhận diện không những chỉ được hiểu là sự nhận thấy bằng mắt mà còn được hiểu là sự hình dung được đối tượng xuyên qua trí tưởng, cảm giác, trí nhớ, ... Trình tự nhận diện và nhận thức của người Việt thường mang tính **tuyến tính (linear)**. Khi người Việt nói: “*một cô gái đẹp*”, thì trước nhất phải nhìn thấy “*một cô gái*” rồi mới xem có “*đẹp*” không. Nói cách khác là trước nhất phải nhận diện được cái thực thể, rồi mới đánh giá được nó theo nhận thức của mình, dù là chủ quan. Trong trường hợp, cái thực thể ấy không có thực trước mắt mà chỉ nằm trong trí tưởng tượng, tâm hồn, thì sự hình dung về nó cũng là một hình thức nhận diện: *nhận diện trừu tượng*. Hành động miêu tả trình tự “*nhận diện* $\rightarrow$ *nhận thức*” chính là hành động dùng ngôn ngữ để sao chép lại nó. Đối với người Việt nói riêng, bởi quá trình nhận diện và nhận thức đã theo trình tự tuyến tính, cho nên từ ngữ trong lời nói cũng được xếp đặt một cách tuyến tính y như thế. Đặt câu hỏi “*cô gái thế nào*”? Trả lời: À, “*cô gái đẹp*”. Từ “*cô gái*” mới có “*cô gái đẹp*”. Quan điểm này giống như một *phép áp* (*phép ánh xạ*) trong toán học. Người ta chỉ có thể áp dụng một phép áp cho một phần tử, khi phần tử ấy đã tồn tại trong một tập hợp gốc nhất định.



Hàm f sẽ không có tác dụng đối với một phần tử không tồn tại, bởi kết quả chỉ là ảnh trống ($f(\Phi)$). Không ai muốn dùng nó một cách vô nghĩa thế. Hàm f phải chứng tỏ được chức năng của mình là biến phần tử gốc thành phần tử mới. Các phần tử ảnh x, y mang tính chất f .

Diễn tiến của lối suy nghĩ này mang tính tuyến tính. Tiến trình này xuất phát từ một khởi điểm (gốc) và dẫn đến một kết quả (ảnh).

Ví dụ một hàm số thực tế trong toán: Biến tất cả mọi số nguyên thành số bình phương:

$$\forall n \in \mathbb{N}; \exists f \in F:$$

$$f: n \rightarrow f(n) = n^2$$

Giả sử, trước hết đã tồn tại một tập hợp số nguyên, tức những phần tử gốc n . Những phần tử n sẽ được hàm f biến hóa thành phần tử ảnh dưới dạng n^2 .

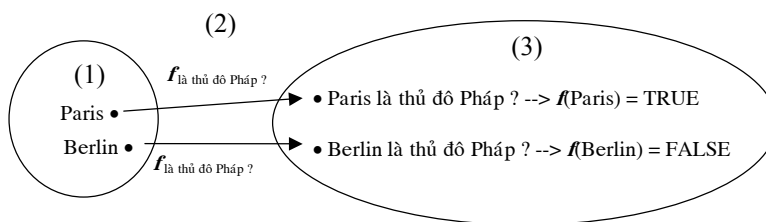
Hoặc xem một mô hình toán logic khác: *hàm đặc tính (characteristic function)*.

Nếu nói Paris là thủ đô của Đức, thì không ai công nhận. Điều đó *sai (false)*. Nhưng nếu nói Paris là thủ đô của Pháp, thì ai cũng công nhận, bởi điều đó *đúng (true)*. Giả sử:

- (i) Trước nhất đã tồn tại hai phần tử: thành phố Paris và Berlin.
- (ii) Đặt câu hỏi: “*Thành phố nào là thủ đô nước Pháp ?*”
- (iii) Hãy trả lời: Một đúng, hai sai. Thành phố nào ?

Lập mô hình hàm đặc tính, ta có 3 phần (1), (2), (3):

- (1) : tập hợp gốc gồm hai phần tử: *Paris* và *Berlin*.
- (2) : hàm đặc tính f , hỏi, có phải ... là *thủ đô nước Pháp ?*
- (3) : tập hợp ảnh mang hai trị *đúng/sai* { *TRUE, FALSE* }



Hàm đặc tính là một hình thức *hàm chọn lựa (selecting function)*. Nó có chức năng xác định tính đúng sai của một phần tử khi được gán cho một tính chất được

chọn lựa. Nếu nói “*Paris là thủ đô Pháp*”, thì đúng; còn nói “*Berlin là thủ đô Pháp*” thì là sai. Thay vì đặt câu hỏi “*có phải ... là thủ đô nước Pháp*”, người ta có thể chọn lựa các đặc tính khác như “*đẹp*”, “*xấu*”, “*lớn*”, “*nhỏ*”, “*mập*”, “*ốm*”, ... Nói chung, cho dù có chọn lựa bất cứ tính chất gì đi nữa, nếu muốn ứng dụng được hàm đặc tính, thì điều kiện đòi hỏi trước hết vẫn là phải tồn tại những phần tử mà người ta muốn đánh giá. Hàm số không thể ứng dụng được cho chân không, cho một phần tử chưa tồn tại, chưa được biết đến.

Chúng ta thấy, ngôn ngữ toán là một sản phẩm của con người và được dùng để diễn tả tư tưởng của con người trong lĩnh vực toán học. Từ cách vẽ mô hình cho tới cách lập thức trong toán học, tất cả đều đi theo trình tự “*nhận diện rồi nhận thức*” của người diễn tả: một là chọn thực thể, hai là đánh giá nó hoặc nói tiếp.

Phức tạp hơn, song ngôn ngữ tự nhiên không khác. Ví dụ tiếng Việt. Dùng tiếng Việt, trên nguyên tắc, trước hết người nói phải có *chủ thể* (*subject*) hoặc *chủ đề* (*theme*) (tức phần tử gốc), nói chung là *khởi điểm* (*start point*) rồi sau đó mới nói tiếp, diễn giải tiếp. Nếu khởi điểm là một đối tượng cần được diễn tả thì tiến trình mô tả nó cũng giống như một hàm số. Khi nghe một ai nói có “*một người đàn bà đẹp*”, thì người nghe hẳn thừa hiểu, có một “*người đàn bà*” (là điểm chính, tức chủ thể) và người đó “*đẹp*” (là điểm phụ, tức lời bình theo sau).

Thực ra dân tộc nào cũng có cách hành ngôn theo nguyên tắc có khởi điểm rồi diễn giải. Chỉ có điều, theo thời gian, nhiều dân tộc không còn giữ đúng cách phản ánh cái trật tự thực tế của mô hình đó, ví dụ “*chủ đề → diễn giải*”, bằng ngôn ngữ nữa. Không theo trình tự này khi diễn tả đối tượng, ngôn ngữ phải được hình thức hóa bằng cách khác và từ đó cấu trúc lời nói cũng thay đổi theo.

Hình thức hóa ngôn ngữ cũng có nguyên do và lợi điểm của nó. Bằng cách này, người nói có thể hành ngôn cô đọng hơn, diễn tả thời gian, không gian chính xác hơn, hoặc làm rõ được những tính chất sở hữu, phái tính, ... nhờ những hình vị (*morpheme*) tương ứng ...

Đa số ngôn ngữ Âu châu là vậy. Thế còn tiếng Việt thì sao ? Cách hành ngôn của tiếng Việt khác các ngôn ngữ khác ở điểm nào ?

Sau đây, chúng ta hãy thử xét sơ qua vài đặc điểm chính trong tiếng Việt, đồng thời so sánh với vài ngôn ngữ khác.

VÀI ĐẶC ĐIỂM CHÍNH TRONG TIẾNG VIỆT

1. Tính tịnh tiến

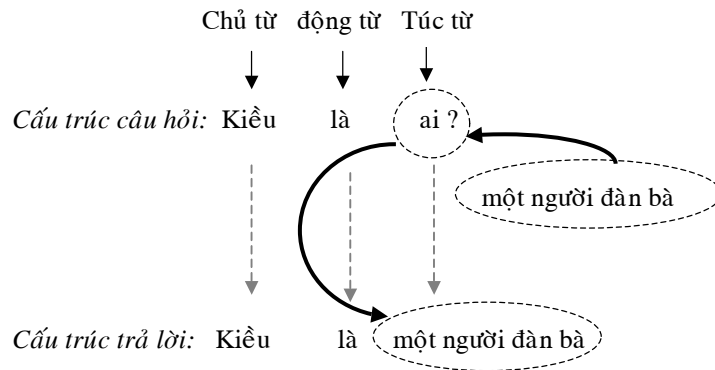
Tiếng Việt là một loại ngôn ngữ tuyến tính. Không những tuyến tính, tức là chứa tính “*tiến*”, tiếng Việt còn chứa cả tính “*tĩnh*”. Tính chất “*tĩnh*” xuất hiện trong cấu trúc của các thể câu thường ít bị thay đổi. Tiêu biểu là tính chất đồng dạng của cấu trúc câu hỏi và câu trả lời. Giả sử hỏi:

(1) Kiều là ai ?

Trong câu hỏi, từ nghi vấn “*ai*” được hiểu như một *biến số* (*variable*), nghĩa là một ẩn số chưa có trị số rõ ràng. Câu hỏi sẽ biến thành câu trả lời, nếu gán cho biến số “*ai*” một giá trị, ví dụ thay thế “*ai*” bằng cụm từ “*một người đàn bà*”:

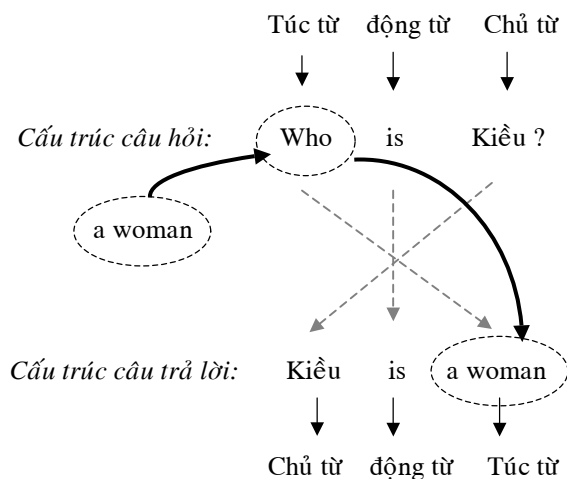
(2) Kiều là một người đàn bà.

Kết quả: Cấu trúc câu hỏi và câu trả lời không thay đổi. Vị trí của các thành phần trong câu nói (thường được gọi là chủ từ, động từ, túc từ) vẫn giữ nguyên (giả sử xem “*là*” ở đây như một động từ):



Tiếng Anh cũng được xem là một ngôn ngữ tuyến tính, thế nhưng lại không “*tĩnh*” bằng tiếng Việt. Không tịnh được hiểu là cấu trúc câu hỏi và câu trả lời của tiếng Anh không giống nhau. Người Anh không hỏi “*Kiều is who ?*” và trả lời “*Kiều is a woman*”. Lập câu hỏi như vậy là sai ngữ pháp. Hỏi đúng phải là

“Who is Kiều?”. Câu hỏi ngược hẳn câu trả lời vì vị trí của chủ từ và túc từ bị đảo ngược:



Để giải thích cái ý nghĩa nguyên thủy từ sự nhận diện và nhận thức vốn chưa bị đảo ngược, nay do cấu trúc câu hỏi bị đảo ngược, các nhà ngôn ngữ học Âu châu đã phải nhờ đến một quy tắc được gọi là quy tắc **chuyển vị wh (wh-movement)**. Đây là một loại *quy tắc biến hình (transformation rule)* thường gặp trong hướng *ngữ pháp biến hình tạo sinh (Generative Transformation Grammar)* của Chomsky (sẽ nói sau).

Thử xem một câu hỏi tiếng Việt khác cũng mang cấu trúc “Chủ từ-Động từ-Túc từ” giống “Kiều là ai?”, bên cạnh là các câu tiếng Anh/Đức đồng nghĩa:

- (3) Kiều thương ai ?
- (4) Whom does Kiều love ? (tiếng Anh)
- (5) Wen liebt Kiều ? (tiếng Đức)

Trong câu hỏi tiếng Việt, động từ “thương” đứng trực tiếp trước “ai”. Nếu gán cho biến số này một giá trị có lý, ví dụ như “Kim Trọng”, ta sẽ được câu trả lời mà cấu trúc câu vẫn giữ nguyên. Trong khi đó ở câu hỏi tiếng Anh, từ nghi vấn “whom” lại nằm xa động từ “love”, hoặc ở tiếng Đức, “wen” (ai) lại nằm trực tiếp trước động từ “liebt” (yêu). Muốn được một câu xác định mà cấu trúc không bị đổi (như tiếng Việt), người Anh, người Đức không thể thay thế các từ “whom”, “wen” bằng “Kim Trọng” một cách dễ dàng:

- (6) * ¹⁾ *Kim Trọng* does *Kiều* love. (tiếng Anh)
 (7) *Kim Trọng* liebt *Kiều*. (tiếng Đức)
 (8) *Kiều* thương *Kim Trọng*.

Ta thấy, câu trả lời tiếng Anh sai ngữ pháp, còn câu trả lời tiếng Đức lại có nghĩa khác hoàn toàn: “*Kim Trọng thương Kiêu*” chứ không phải “*Kiều thương Kim Trọng*”. Muốn cả hai trường hợp đều đúng ngữ pháp cũng như ngữ nghĩa, bắt buộc người Anh và người Đức phải đảo ngược cấu trúc câu. Đây là một đặc điểm của **ngôn ngữ nghịch đảo (inversion language)**.

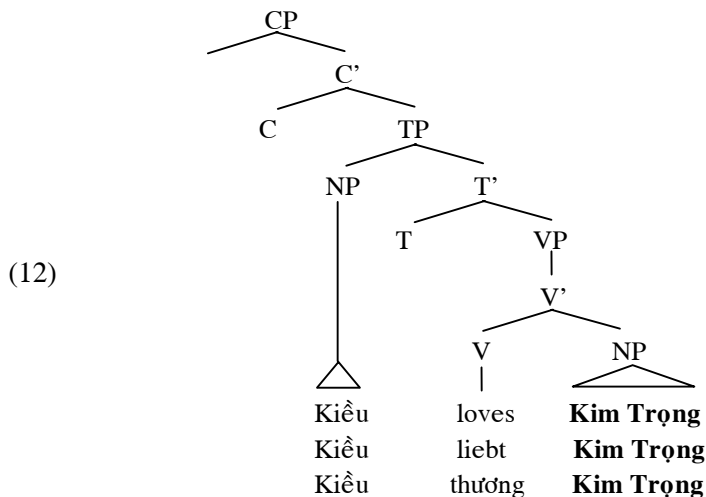
Ý nghĩa của ba câu xác định sau:

- (9) *Kiều* thương ***Kim Trọng***.
 (10) *Kiều* loves ***Kim Trọng***. (tiếng Anh)
 (11) *Kiều* liebt ***Kim Trọng***. (tiếng Đức)

có dạng cấu trúc được gọi là **cấu trúc D (D-structure)**. (Xin lưu ý: Các bạn không chuyên có thể lướt qua phần này. Phần lý thuyết được đóng trong ngoặc [[.]]).

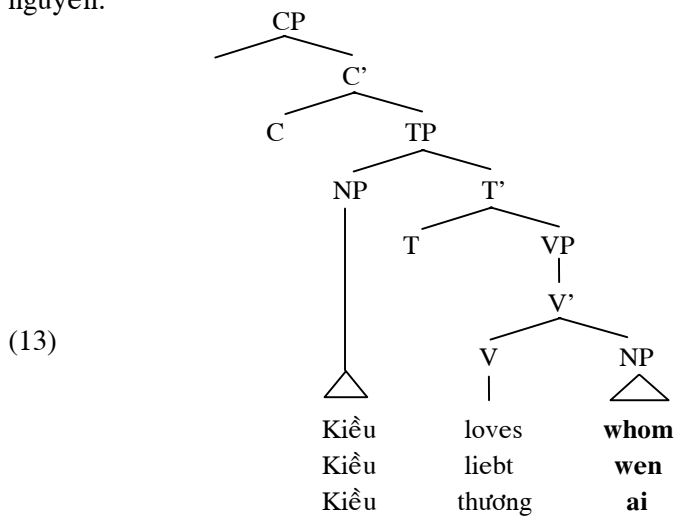
[[

Mô tả bằng *cây cú pháp (syntax tree)* (sẽ được nói rõ sau), ta có:



¹⁾ Trường hợp không nói được sẽ được đánh dấu *.

Cấu trúc D nói lên ngữ nghĩa “*Kiều thương Kim Trọng*” theo đúng thứ tự của từng dữ kiện tuần tự được thu nhận. Song đây không hẳn là cấu trúc của câu nói thật ở ngoài đời. Câu nói thật mang một cấu trúc khác: **cấu trúc S (S-structure)**. Nguyên thủy, cái được nhận diện và nhận thức chính là cái ngữ nghĩa thật sự tuần tự được ghi lại dưới dạng cấu trúc D, thế nhưng khi trình bày, người nói lại lập câu theo kiểu khác. Ba câu xác định (Việt, Anh, Đức) ở ví dụ này tình cờ có cấu trúc D giống nhau, kể cả cấu trúc S. Song khi lập câu hỏi, người Anh, người Đức lại lập câu theo kiểu khác mặc dù cấu trúc D vẫn giữ nguyên:



Nhìn lại cấu trúc D trên, chỉ có tiếng Việt nói được; tiếng Anh và tiếng Đức thì không:

- (14) Kiều thương **ai**
 (15) * Kiều loves **whom** (tiếng Anh)
 (16) * Kiều liebt **wen** (tiếng Đức)

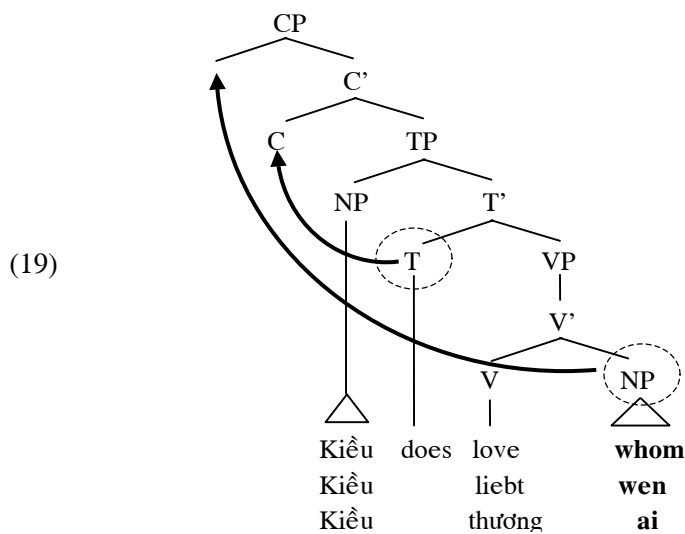
Muốn nói được, các câu tiếng Anh, tiếng Đức có cấu trúc D bên trên phải được chuyển sang dạng cấu trúc S như sau:

- (17) **Whom** does Kiều love ? (tiếng Anh)
 (18) **Wen** liebt Kiều ? (tiếng Đức)

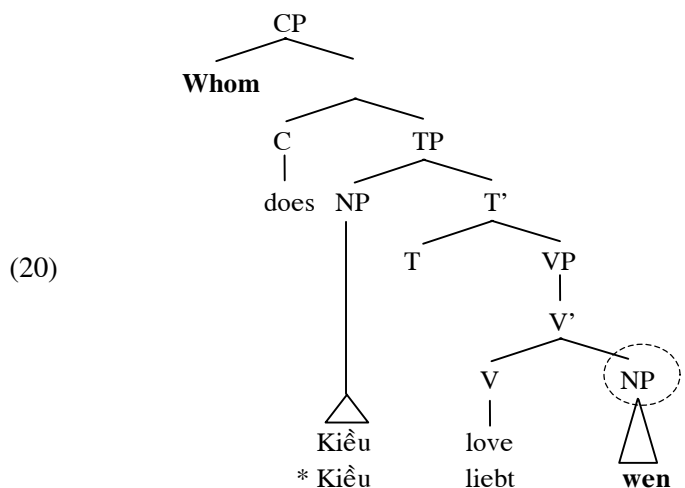
Trong khi đó, cấu trúc S và D của câu hỏi và câu trả lời trong tiếng Việt giống hệt nhau. Muốn giải thích sự khác biệt giữa hai cấu trúc D (ngữ nghĩa từ nhận diện/nhận thức) và cấu trúc S (cách hành ngôn thực tế), tiếng Anh và tiếng Đức

phải nhờ đến quy tắc “chuyển vị *wh*” . Quy tắc chuyển vị có nhiệm vụ chuyển động từ và bổ ngữ (tức từ) của câu hỏi trong cấu trúc D vào vị trí của câu hỏi trong cấu trúc S.

Quá trình chuyển vị tương đối phức tạp. Đối với tiếng Anh, trước nhất, phải lồng thêm *động từ phụ* (thường gọi là *trợ động từ*) “*does*” vào để diễn tả thì hiện tại. Quá trình chuyển vị trải qua ít nhất là hai bước. Bước một, đưa thông tin diễn tả thì (*tense*) lên C. Bước hai, chuyển “*whom*”, “*wen*” lên đầu câu (một hình thức *chuyển vị lên cấp* (*raising movement*)):



Sau khi chuyển vị, cấu trúc D biến thành cấu trúc S có thực ngoài đời:



Câu tiếng Việt “*Kiều thương ai*” vốn đã nói được, cho nên không cần giải thích bằng lý thuyết chuyển vị. Chỉ có câu tiếng Anh nằm dưới dạng cấu trúc D “*Kiều loves whom*” do không nói được, nên mới cần chuyển sang dạng S. Trong cấu trúc này, tiếng Đức không thể chuyển vị tương đương như tiếng Anh.

]]

Đọc đến đây, có lẽ có nhiều bạn đọc không hiểu hết. Điều này cũng không quan trọng. Chúng ta cứ tạm hiểu sơ rằng, cái cấu trúc phản chiếu ý nghĩa thực sự của sự nhận diện và nhận thức của con người về một đối tượng muốn diễn tả là cấu trúc D. Còn khi nói, người hành ngôn sẽ lập thức ý nghĩa này bằng cấu trúc S. (Xem thêm *Deep Structur* trong *Language and Responsibility* (Chomsky 1979: 169) hoặc chương 12).

Mọi cắt nghĩa sơ quát bên trên chỉ nhằm cho thấy sự khác biệt giữa ngôn ngữ tĩnh tiến và không tính tiến. Giới ngôn ngữ học Tây phương đã đưa ra những lý thuyết giải thích (đúng theo nghĩa khoa học nhận thức của bộ môn cú pháp). Tuy nhiên, không phải một lý thuyết bao giờ cũng đủ khả năng giải thích được cấu trúc của mọi ngôn ngữ, điển hình là cách giải thích theo thuyết “chuyển vị” bên trên.

Nói một cách ngắn gọn, hai tính chất “tĩnh” và “tiến” nói lên tính “*tĩnh tiến*”²⁾ của tiếng Việt.

Tính tĩnh tiến có mặt khắp nơi trong ca dao, tục ngữ, lời nói bình dân, văn chương bác học, ...:

(21) “Đêm qua gió bắc mưa lằm,
Đền lằm với bóng, bóng lằm với ai ?
Đêm qua gió bắc mưa dầm,
Đền lằm với bóng, bóng lằm với anh”.

(22) “Lá gì không nhánh không cành ?
Lá gì chỉ có tay mình trao tay ?
Lá thư không nhánh không cành,
Lá thư chỉ có tay mình trao tay”.
(Ca dao - Thái Doãn Hiếu, Hoàng Liên 1994)

²⁾ Tính tĩnh tiến ở đây không đồng nghĩa với *phép tĩnh tiến* (*translation*) trong toán học. Phép tĩnh tiến trong toán học là phép biến hình, trong đó, một hình được di chuyển hay trượt đến một vị trí khác mà không quay, trục mới của nó song song với trục cũ. (Tạ Hùng 1996)

Cấu trúc tương đối không thay đổi của câu diễn tả tính “tĩnh”. Mặt khác, tính “tĩnh” cho phép câu tuần tự được mở rộng một cách *hợp lý* về phía phải với từng thông tin thu thập được từ thế giới bên ngoài.

- (23) “Lạy trời” ... (rồi sao ?)
 “Lạy trời mưa xuống” ... (Để làm gì ?)

Xin lưu ý, dù thiếu động từ “*roi*”, cách hành ngôn vẫn diễn tả được tính và hướng chuyển động của “*mưa*”. Vả lại không thể thêm “*roi*” vào, vì cấu trúc thơ 4 chữ sẽ bị hủy.

- (24) “Lấy nước”, ... (để làm gì ?)

Không ai hỏi “*nước gì*”, bởi đã biết là “*nước mưa*”.

- (25) “Lấy nước tôi uống,
 Lấy ruộng tôi cày,
 Lấy đầy bát cơm,
 Lấy rơm đun bếp”.

Đó là kiểu hành ngôn tuyến tính của người Việt. Ở đó không cứ phải có động từ, liên từ, ... giống các ngôn ngữ Âu châu thì mới được coi là đúng ngữ pháp.

2. Tính giàu thông tin cụ thể

Ngoài tính tĩnh tiến, tiếng Việt còn một đặc tính khác là giàu thông tin cụ thể. Một trường hợp điển hình là tiếng xưng hô hoặc *đại danh từ nhân vật* (*personal pronoun*)³⁾ nói chung. Cách xưng hô trong tiếng Việt có lẽ là khó xài nhất thế giới. Chỉ riêng cho ngôi thứ nhất, số ít, tiếng Việt đã có hơn 40 cách dùng khác nhau:

- (26) Tôi, tui, tớ, tao, tau, ta, mỗ, người ta, mình, cụ, cố, ông, ôn, bà, mẹ, nội, ngoại, bác, ba, tía, bố, cha, mạ, mẹ, u, chú, cô, cậu, dì, o, dượng, thím, mợ, thầy, bố, vú, anh, chị, em, con, cháu, ...

³⁾ Thay vì dịch là “*nhân xưng đại danh từ*”, chúng tôi chọn nguyên tắc dịch theo lối ghép từ của tiếng Việt với bổ/phụ ngữ nằm bên phải (right attributed): “*đại danh từ nhân vật*”. Nguyên tắc này cũng sẽ được áp dụng cho các từ khác. “*Nhân vật*” được dịch sát nghĩa từ “*personal*”. Không có lý do gì phải dịch “*personal*” thành “*nhân xưng*” trong khi nó hoàn toàn không chứa ý nghĩa “*xưng*”.

Tính giàu thông tin còn nằm dưới dạng điệp ngữ, từ ghép: “*ruộng đồng*”, “*sông nước*”, “*mênh mông*”, “*trái cây*”, “*gạo lúa*”, “*dồi dào*”, “*nơi nơi*”, ... (Xem thêm *Từ tiếng Việt* - nhiều tác giả 1998, *Nghiên cứu ngữ pháp tiếng Việt* – Nguyễn Kim Thản 1997, *Ngữ pháp tiếng Việt* – Nguyễn Tài Cẩn 1999)

Tính giàu thông tin còn xuất hiện qua cách dùng khá nhiều động từ trong câu, tiêu biểu là động từ ghép hoặc động từ xen kẽ với danh từ.

- (27) Đi + chơi → đi chơi
Đi + làm → đi làm
- (28) “Nước sông Nhị Hà mới bắt đầu *lên to*, cuộn cuộn *chảy*, *tưởng muốn lôi phăng* cái cù lao giữa sông đi”. (*Anh phải sống* - Khái Hưng).
- (29) Hấn *xoay* sang, *nhíp* mắt, *nhìn* tôi, *cười* không *nói*, *làm* ra vẻ *hiếu* hết những gì tôi đang *nghĩ*.
- (30) “Khéo khéo *đi* đâu lữ *ngẩn* ngơ,
Lại đây cho chị *dạy* làm thơ.
Ong non *ngứa* nọc *châm* hoa rửa.
Đê cồn *buồn* sừng *húc* đậu thưa.”
(*Măng học trò đốt* - Hồ Xuân Hương)

(Xem thêm *Động từ trong tiếng Việt* – Nguyễn Kim Thản 1999)

Tính giàu thông tin cụ thể còn xuất hiện trong cách thành lập danh từ:

- (31) Nhà → cái nhà, ngôi nhà
(32) Cửa → cái cửa, cánh cửa
(33) Vườn → cái vườn, góc vườn, khu vườn, mảnh vườn
(34) Thuyền → cái thuyền, con thuyền, chiếc thuyền

Cách thức mô tả sự vật bằng cách “*tiểu tập hợp hóa*” cái tập hợp toàn thể là một điểm tiêu biểu trong tiếng Việt mà ít có ngôn ngữ Ấn-Âu nào có được. Ví dụ “*nhà*” là một tập hợp toàn thể. Chia nhỏ ra, ta có các dạng cá thể có tính cách đơn vị như “*cái nhà*”, “*ngôi nhà*”, “*dãy nhà*”. Đây là những phần tử đếm được thuộc về tập hợp “*nhà*”.

3. Tính tỉnh lược

Một mặt, cách hành ngôn của người Việt khá phong phú qua cách dùng thật nhiều thông tin miêu tả, lắm lúc vòng vo:

- (35) “Đêm vắng,
Buồng thơm.
Em soi gương trộm,
Ngọn đèn mượn gió, nghiêng nghiêng”
(*Đêm vắng* - Phùng Cung)
- (36) “Trên trời có đám mây xanh,
Chính giữa mây trắng, chung quanh mây vàng.
Ước gì anh lấy được nàng ...”
(Ca dao)

Mặt khác, nhiều khi câu trúc câu nói rất cô đọng, từ ngữ bị lược bỏ khá nhiều, đặc biệt là các từ “*thì*”, “*là*”, “*mà*” có công dụng phân đoạn về mặt cú pháp:

- (37) “Trên thuận dưới hòa”
(Trên có thuận *thì* dưới mới hòa)
- (38) “Đất lành chim đậu”
(Đất có lành *thì* chim mới đậu)
- (39) “Chim lồng cá chậu”
(Chim *thì* ở trong lồng, cá *thì* ở trong chậu)
- (40) “Cha nào con nấy”
(Cha thế nào *thì* con sẽ như thế nấy)
- (41) “Ăn gì bổ đó”
(Ăn cái gì *thì* bổ cái đó)

Tính tỉnh lược có thể làm hư cả cấu trúc câu:

- (42) “Thằng Bờm có cái quạt mo,
Phú ông xin đổi *ba* bò *chín* trâu”.
(Ba con bò, chín con trâu)

hoặc hiện tượng câu thiếu chủ ngữ (chủ từ):

- (43) Ê ! Đi đâu đó ? Cơm nước gì chưa ?
 (44) Sao, khỏe không ?

Hiện tượng câu thiếu chủ ngữ là một hình thức nói tắt, bởi người được hỏi đã biết câu hỏi dành cho ai. Hiển nhiên, nói tắt, nói cộc lốc làm giảm tính lịch lãm, song đó là lối dụng ngôn phổ biến trong xã hội Việt Nam.

Hiện tượng thiếu chủ ngữ xảy ra lắm khi cũng là do bỏ ngữ qua cuộc đối thoại lần đầu, khi người đối diện không biết phải xưng hô thế nào cho hợp tình, hợp lý. Nhiều ý kiến cho rằng, do ảnh hưởng nền văn hóa Nho giáo, người Việt mới có lối xưng hô theo tôn ti, vai bậc, trong khi các dân tộc Âu châu chỉ có cách xưng hô trừu tượng, ví dụ *I, you* của tiếng Anh, *Ich, Du, Sie* của tiếng Đức, *Je, Tu, Vous* của tiếng Pháp. Luận cứ này có lẽ không được vững lắm. Một dân tộc như người Hoa cũng bị Nho giáo ảnh hưởng khá nặng, thậm chí còn nặng hơn dân tộc Việt, song người Hoa vẫn có các tiếng xưng hô trừu tượng *Ngô, Lị, ...* Đây là một điểm thú vị. Sáng tạo được cho người Việt những tiếng xưng hô trừu tượng, có lẽ không phải là một điểm dở.

Đặc biệt, tính tỉnh lược thường xuất hiện trong thơ, ca dao, tục ngữ, vè, câu đối, ... nói chung là ở mọi thể thơ có luật. Lý do dễ hiểu là vì số chữ dùng bị giới hạn một cách nghiêm ngặt. Chẳng hạn thể lục bát chỉ cho phép câu đầu 6 chữ, câu sau 8 chữ. Hình thức bị giới hạn khiến người làm thơ khó có thể đặt tất cả câu cú đúng theo ngữ pháp. Song nhiều khi chính người làm thơ cũng không muốn vậy vì sợ câu cú sẽ trở nên cứng nhắc, mất hồn thơ. Mặt khác, tính tỉnh lược trong tiếng Việt cũng có nguyên do lịch sử của nó. Cả ngàn năm bị Tàu đô hộ, biết bao nhiêu cổ văn bác học viết đúng câu, đúng cú đã bị đốt sạch. Muốn bảo vệ và lưu truyền tư tưởng, ông bà tổ tiên của chúng ta chỉ còn nước là truyền miệng:

- (45) “Trăm năm bia đá cũng mòn.
 Ngàn năm bia miệng vẫn còn trơ trơ.”
 (Ca dao)

Đã gọi là truyền miệng thì phải làm sao cho ngắn gọn, dễ nhớ. Vì thế mà văn chương truyền khẩu Việt Nam thường ở thể thơ có luật, có vần, dễ nghe, dễ cảm, dễ thuộc, nhớ lâu. Hình thức bị cô đọng tối đa nhưng nội dung lại dồi dào, sâu sắc nhờ ẩn dụ, thâm thúy nhờ chơi chữ, ... Nhiều khi cách hành ngôn bất buộc

người nghe phải hiểu ngầm hoặc có lúc phải cần diễn giải thì mới hiểu. Ví dụ một câu tục ngữ, người Bắc hay dùng:

(46) “Nát dẻo sống bùi”

Thực, khi mới nghe câu tục ngữ này lần đầu, người Nam hoặc cả người Việt nói chung khó có thể hiểu nổi ý muốn nói gì, nếu không nhờ giải thích. Ý câu này muốn nhắc nhở đại khái rằng, hễ có ai mời mình tới nhà dùng cơm, mà nhờ cơm có nát, thì cũng nên khen là dẻo, hoặc nhờ cơm có sống, thì cũng nên khen là bùi, là béo quá cho chủ nhà lên tinh thần. Theo thời gian, cứ truyền miệng mãi cốt cho câu cú được cô đọng, bất cần luật, chỉ cần nội dung, nhiều người Việt đâm ra mắc một cái tật: bỏ được gì cứ bỏ. Rõ ràng câu “*Nát dẻo sống bùi*” sai ngữ pháp nhưng may mắn thay, người nghe vẫn hiểu. Hiểu được là nhờ trí thông minh lành mạnh của con người.

Hiện tượng tỉnh lược thực ra không chỉ riêng ở Việt Nam mới có mà còn có ở các dân tộc khác. Một trường hợp thường gặp nhất là cách chạy những hàng tít lớn trên nhật báo. Ví dụ một hàng tít to tướng trên một nhật báo Đức sau khi công chúa Diana tử nạn: “*Diana tot*” (Diana chết) thay vì: “*Diana ist tot*”. Đúng ngữ pháp tiếng Đức, câu phải có động từ “*ist*”. Hoặc cách hỏi vắn tắt trong lúc đàm thoại có tính cách bạn bè, thân mật, chẳng hạn như hỏi thăm sức khỏe người đối diện: “*Wie geht ?*” thay vì: “*Wie geht es Dir ?*” (Bạn khỏe không ?), hoặc: “*Weißt ?*” thay vì: “*Weißt Du ?*” (Bạn biết không ?). Đúng ngữ pháp, câu phải có đại danh từ nhân vật “*Du*”. Cách hành ngôn vắn tắt này trong tiếng Đức không khác cách hỏi tắt trong tiếng Việt: “*Đi đâu đó. Cơm nước gì chưa ?*”, “*Sao, khỏe không ?*”. Chúng ta nên nhớ rằng, người Anh, người Đức, người Pháp, người Ý, người Thổ, người Tàu, người Nhật, người Phi Châu, ... gì đi nữa dù nói tiếng mẹ đẻ, nhưng chưa chắc là ai cũng nói đúng ngữ pháp hoàn toàn. Người Việt cũng vậy, khi hành ngôn, có câu chuẩn, có câu không. Ngôn ngữ bao la nhưng không bất định. Mức độ hoàn hảo của ngôn ngữ lệ thuộc vào thẩm năng và dụng năng ngôn ngữ của người sử dụng. Cho nên, định rõ tiêu chuẩn ngữ pháp là một việc cần thiết, đặc biệt là đối với người nghiên cứu ngữ pháp. Dựa trên tiêu chuẩn này, nên phân biệt cách hành ngôn bị tỉnh lược và cách hành ngôn đúng *ngữ pháp chuẩn hình (well-formed grammar)*.

4. Tính cảm đề

Hiện tượng bỏ chủ ngữ (chủ từ) trong tiếng Việt lắm khi dẫn đến tình trạng câu bị thiếu mà người Âu châu cho là sai ngữ pháp. Kể cả trong nhiều trường hợp khác, mặc dù có chủ ngữ nhưng câu vẫn có thể bị cho là sai vì lý do: cách dùng vị từ (*predicate*) không thích hợp với chủ ngữ (vị từ tương đương với động từ, tính từ thường dùng trong ngôn ngữ học cổ điển).

(47) Bức tranh *treo* ở trên tường.

Theo người Đức, câu này vô lý quá. Làm sao bức tranh có thể tự nó treo lên tường được, phải có ai treo nó lên chứ? Với suy nghĩ ấy, họ sẽ cho là tiếng Việt không logic, ngữ pháp lộn xộn. Người Đức không dùng động từ “*treo*” cho chủ ngữ “*bức tranh*” ở trạng thái *chủ động* (*active*) mà phải dùng ở trạng thái *bị động* (*passive*), đại để, câu phải có dạng như sau:

(48) Bức tranh *được treo* ở trên tường.

Thực chất của vấn đề này là cách nhận diện và nhận thức về trạng thái của đối tượng cần được diễn tả chứ không hẳn là ngữ pháp. Theo suy nghĩ của người Âu châu, *hoặc chủ thể để ra hành động, hoặc chủ thể bị một hành động tác động lên nó*. Chủ thể phải xuất hiện trong câu qua hình thức chủ ngữ. Hành động được diễn tả bằng động từ. Động từ phải ở *thái chủ động* (*active voice*), nếu chủ thể gây ra hành động. Ngược lại, động từ phải ở *thái bị động* (*passive voice*), nếu chủ thể bị một hành động tác động. Rất logic. Thế nhưng cách diễn tả “*chủ động*” và “*bị động*”, nói chung là “*động*”, “*tĩnh*” của người Việt khác. Xét về “*động*”, người Việt có thể nói y như một người Đức:

(49) Ich hänge das Bild an die Wand auf.

(50) Tôi treo bức tranh lên tường.

Ở đây, chủ ngữ “*Tôi*”, “*Ich*” tạo ra hành động. Còn nếu xét về “*tĩnh*”, người Việt sẽ nói:

(51) Tôi thấy bức tranh treo ở trên tường.

dịch nội dung này sang tiếng Đức, ta có:

(52) Ich sehe, das Bild ist an der Wand aufgehängt.

Câu tiếng Đức này hoàn toàn đúng ngữ pháp song trên thực tế, hầu như không ai nói vậy mà là nói:

(53) Ich sehe das Bild an der Wand.

(54) Tôi thấy bức tranh ở trên tường.

Nói câu trên, người Đức không dùng động từ “*treo*” (*aufhängen*). Đã trông thấy *bức tranh ở trên tường* (das Bild an der Wand) thì dĩ nhiên là phải có ai treo nó lên đó, không cần nói đến. Như vậy, họ đã tự công nhận, có một chủ thể nào đó đã làm công việc ấy mà không có mặt trong câu. Người Việt suy nghĩ không khác lắm, khi nói:

(55) Bức tranh *treo* ở trên tường.

Nghĩa là ai đó đã treo bức tranh lên mà không có mặt trong câu. Suy ra, chủ ngữ trong câu chỉ mang tính hình thức cho đúng cấu trúc ngữ pháp của câu, chứ chưa chắc đã, hoặc không nhất thiết phải làm chủ động từ. Vậy thì làm sao có thể xác định được mối quan hệ giữa động từ và đối tượng đã gây ra hành động đó ? Chắc chắn không phải chủ ngữ “*bức tranh*” mà là một *ngoại chủ thể* (*external subject*). Nếu vậy thì làm sao giải thích được động từ “*treo*” có quan hệ gì với chủ ngữ “*bức tranh*” ?

Mối quan hệ giữa “*bức tranh*” và “*treo*” phản ánh 1-1 cách nhận diện của người Việt về hiện trạng của thế giới thực và cách diễn tả. Nói nôm na là thấy sao nói vậy.

(56) *Hỏi:* Bức tranh thế nào ?

Trả lời: Bức tranh (thì) treo. (Rõ ràng khi quan sát thì thấy bức tranh đang ở vị thế treo chứ không nằm)

Hỏi: Bức tranh (thì) treo ở đâu ?

Trả lời: Bức tranh (thì) treo ở trên tường.

Từ lối quan sát tuyến tính, từ ngữ được xếp đặt thứ tự theo nguyên tắc tịnh tiến, cuối cùng ta có câu “*Bức tranh treo ở trên tường*”. Cả câu diễn tả *tình trạng tĩnh* của đối tượng “*bức tranh*”. Đối với người Việt, khía cạnh đúng sai ngữ pháp ở đây tùy thuộc vào sự hợp lý của ý nghĩa tịnh tiến chứ không nhất thiết phải dựa vào một *chủ ngữ ngữ pháp* (*grammatical subject*), tức chủ từ (như thường gọi),

bắt buộc phải nằm trong câu và làm chủ động từ cho đúng ngữ pháp theo cách nhìn của người Âu châu.

Kết luận, động từ trong tiếng Việt có một đặc tính là lệ thuộc vào *chủ đề* hoặc *đề tài* đang nói tới thay vì bị cưỡng bách phải lệ thuộc vào chủ ngữ trong câu như thường thấy ở những ngôn ngữ Âu châu. Nói cách khác, động từ có thể diễn tả nội dung có liên quan đến đề tài, thay vì bắt buộc phải diễn tả hành động của chủ ngữ trong câu. Điểm này nói lên tính (tạm gọi là) ***cảm đề (subject/theme sensitive)*** trong tiếng Việt. Nhiều khi không chỉ động từ mới cảm đề mà cả những loại từ khác. Thực tế cho thấy, có những từ theo sau chủ ngữ chẳng ăn nhập gì đến chủ ngữ trong câu, ngược lại, nội dung của câu gắn bó với *ngoại chủ thể* hoặc *chủ đề* đang được nói tới. Hãy xem vài ví dụ:

- Trường hợp động từ gắn bó với ngoại chủ thể trong chủ đề “*quần áo*” như:

(57) Quần áo đang bán xon kìa. Đi mua đi !

Người nghe có thể hiểu ngầm, ngoại chủ thể là “*người ta*”, “*siêu thị*”, “*chợ trời*”... :

(58) Quần áo (người ta) đang bán xon kìa. Đi mua đi !

(59) Quần áo (siêu thị) đang bán xon kìa. Đi mua đi !

(60) Quần áo (chợ trời) đang bán xon kìa. Đi mua đi !

- Hoặc trong chủ đề “*nhạc nào*”, “*món gì*”, “*cái gì*”:

(61) Nhạc nào hát cũng được.

(62) Nhạc nào (cổ) hát cũng được.

(63) Nhạc nào (ổng) hát cũng được.

(64) Nhạc nào (chúng tôi) hát cũng được.

(65) Nhạc nào (họ) hát cũng được.

(66) Nhạc nào (ca đoàn) hát cũng được.

(67) Nhạc nào (đoàn văn công) hát cũng được.

(68) Món gì ăn cũng được.

(69) Món gì (nó) ăn cũng được.

(70) Món gì (bả) ăn cũng được.

(71) Món gì (tụi tôi) ăn cũng được.

(72) Món gì (kể thiếu ăn) ăn cũng được.

(73) Cái gì (ông tỉnh trưởng) ăn cũng được.

(74) Cái gì (trâu bò) ăn cũng được.

Không dùng chủ ngữ (chủ từ) ngữ pháp, người Việt có cơ hội diễn tả nội dung bằng nhiều ngoại chủ thể khác nhau. Đó là điểm lợi, ví dụ hình thức nói bóng gió, thế nhưng cũng là điểm hại, bởi sự sử dụng ngôn ngữ một cách phóng khoáng thường dễ sinh ra những câu thiếu sắc bén, đa nghĩa. Ví dụ câu:

(75) Người nào cũng chửi.

Đây là một câu hoàn chỉnh. Chủ ngữ “*người nào*” thực sự làm chủ động từ “*chửi*”. Nhưng trong một hoàn cảnh nào đó, nếu hiểu theo kiểu, câu có ngoại chủ thể, thì “*người nào*” sẽ không còn làm chủ động từ “*chửi*” nữa:

(76) Người nào (nó) cũng chửi.

(77) Người nào (Chí Phèo) cũng chửi.

Một ví dụ khác có liên quan đến tính cảm đề, chẳng hạn hỏi tên ai. Người Anh hỏi:

(78) What is your name ?

Trong văn viết tiếng Việt, đôi khi cũng nghe hỏi vậy:

(79) Tên anh là gì ?

Lối nói này có lẽ lai ngữ pháp Âu châu: “*Tên của anh là gì ?*” (chứ không phải “*anh*”). Trong trường hợp này (theo cách nhìn của người Việt), phần chủ thể, tức cái căn cước thực tế, là con người đáng lý phải được nhấn mạnh, nhưng cách hỏi lại nhắm vào cái tên để bảo đảm cách dùng cho đúng chủ ngữ trong câu hỏi, tức một chủ thể chỉ có chức năng *ngữ pháp* như đã nói. Tương tự vậy, người Đức hỏi:

(80) Wie ist Ihr Name ? (Dich sát: Tên anh thế nào ?)

(81) Wie war Ihr Name ? (Dich sát: Tên anh đã thế nào ?)

Bên cạnh đấy, họ cũng thường hay hỏi:

(82) Wie heißen Sie ? (Dich sát: Anh tên thế nào ?)

Có lẽ cách hỏi này gần giống cách hỏi của người Việt:

(83) Anh tên gì ?

Nếu vậy, “*tên*” ở đây phải giữ chức năng giống động từ “*heißen*” trong tiếng Đức. Tuy nhiên, cách hỏi thông thường nhất của người Việt vẫn là:

(84) Anh tên là gì ?

Phân tích câu theo kiểu tiếng Anh, câu hỏi ấy có thể bị coi là quái đản, văn phạm sai quá: “*You what is the name ?*”. Thế nhưng người Việt vẫn dùng mà chẳng thấy gì là sai cả. Chúng ta sẽ làm rõ vấn đề này qua nguyên tắc đề diễn sau đây.

5. Tính đề diễn

Ở nhiều ngôn ngữ Âu châu, hiện tượng chủ ngữ không tương ứng với vị từ (động từ, tính từ) trong câu ít xảy ra hơn so với tiếng Việt. Đối với người Việt, khi bắt đầu muốn diễn tả điều gì, họ sẽ nghĩ ngay đến một đối tượng. Đối tượng có thể là một người gây ra hành động, tức một **tác nhân (actor)**. Loại đối tượng này được coi là **chủ thể logic (logical subject)**. Đối tượng cũng có thể là một **chủ đề (theme)**. Loại đối tượng này được gọi là **chủ thể tâm lý (psychological subject)**. Hoặc đối tượng chính là chủ thể trong câu. Loại này được gọi là **chủ thể ngữ pháp (grammatical subject)** hay **chủ ngữ ngữ pháp** (Halliday 1994:32). Có chủ thể, người ta sẽ tìm cách *diễn giải*. *Diễn giải* là sự thuyết minh, sự định nghĩa, miêu tả, giải thích, bổ sung ý nghĩa cho cái muốn nói hoặc giới thiệu phần thông báo của khởi điểm ấy. Hãy tạm gọi đây là nguyên tắc **Đề Diễn (Theme-Rheme)**. Một ví dụ, tương tự như câu “*Anh tên là gì ?*”, người Việt còn hay hỏi:

(85) Anh nhà ở đâu ?

(86) Anh quê ở đâu ?

(87) Anh người nước nào ?

(88) Anh người gì ?

Chủ thể tâm lý ở đây là nhân vật “*anh*”, tức chủ thể muốn nói. Đó là đối tượng trước nhất phải có. Kế tiếp, người nói sẽ lập thêm một câu, mệnh đề, thành tố

(tạm hiểu là một cụm từ nói được) theo sau để nói thêm về đối tượng đó như một hình thức bổ nghĩa. “*Nhà*”, “*quê*” ở đây được coi như chủ thể ngữ pháp.

Để phân biệt giữa hai phần *chủ đề* và *diễn giải*, người Việt thường dùng “*thì*”, “*là*”, “*mà*”. Ví dụ đối với câu hỏi “*Anh người nước nào ?*”, với:

(89) “*Anh*” là phần chủ đề.

(90) “*Người nước nào*” là phần diễn giải.

Nếu trả lời: “*Người Việt Nam*”, ta có câu “*Tôi người Việt Nam*” mà chúng ta vẫn thường nghe. Hoặc nếu dùng chữ “*là*” nối hai phần này lại, ta sẽ có một câu nguyên vẹn:

(91) Anh là người nước nào ?

Trả lời:

(92) Tôi là người Việt Nam.

“*Là*” giữ chức năng phân phần chủ đề và phần diễn giải. Tương tự là “*thì*”:

(93) Tham thì thâm.

Nói chung, đây là cấu trúc đề diễn rất tiêu biểu trong tiếng Việt. Chỉ vì người Việt thường hay bỏ “*thì*”, “*là*”, “*mà*”, cho nên chúng ta mới khó nhận ra cấu trúc chuẩn hình của câu. Đặc biệt là giới văn thơ có thể do không hiểu “*thì*”, “*là*”, “*mà*” dùng để làm gì, họ thường hay bỏ. Mặt khác, họ cũng có lý riêng: câu có “*thì*” nghe nặng quá, thôi bỏ.

Trên thực tế, cấu trúc “... *thì* ... *là*” không xa lạ gì với chúng ta cho lắm:

(94) Theo chỗ tôi biết *thì* cô ấy là người Việt chứ không phải người Hoa.

(95) Phải quen *thì* mới biết cô ấy là người Việt.

(96) Thế *thì* cô ấy không phải là người mà tôi muốn nói.

Từ “*thì*” rất thường bị lược bỏ:

(97) Đất lành chim đậu.

Giải thích cho có câu cú hẫng hời, thì câu trên có nghĩa:

(98) Đất có lành *thì* chim mới đậu.

Vì câu tục ngữ quá ngắn, không có mặt “*thì*”, nên chúng ta khó nhận ra cái ranh giới giữa chủ đề “*đất lành*” và phần diễn giải “*chim đậu*”. Ở một trường hợp khác, nếu không lược bỏ “*thì*”, chúng ta sẽ dễ nhận ra cấu trúc của nó:

(99) Gắn mực *thì* đen, gắn đèn *thì* sáng.

Câu tục ngữ trên chứa hai mệnh đề. Trong mệnh đề đầu, “*gắn mực*” là phần đề, “*đen*” là phần diễn. Trong mệnh đề sau, “*gắn đèn*” là phần đề, “*sáng*” là phần diễn. Người sáng chế ra câu tục ngữ trên đã không thể lược bỏ “*thì*”, bởi nếu bỏ đi, ta sẽ được một ý nghĩa nằm ngoài ý muốn:

(100) Gắn mực đen, gắn đèn sáng.

Tựu trung, người Việt chỉ có thể lược bỏ các diễn từ “*thì*”, “*là*”, “*mà*”, nếu ngữ nghĩa của lời nói không bị tổn thương.

Ngoài các trường hợp vừa nêu, chủ thể tâm lý cũng thường đứng trước một câu/mệnh đề:

(101) Con bé Thu nó không chịu ăn cơm.

“*Con bé Thu*” là chủ thể tâm lý, trong khi đó “*nó*” là chủ ngữ ngữ pháp của câu “*nó không chịu ăn cơm*”. Cả đoạn diễn giải về đối tượng “*con bé Thu*” được coi như một hình thức bổ nghĩa/phụ nghĩa⁴⁾. Trong trường hợp này, người Việt không dùng “*thì*”, “*là*”, “*mà*” vì ranh giới giữa phần đề và diễn đã khá rõ ràng. “*Nó*” chỉ là sự lặp lại nhằm chỉ định “*con bé Thu*”.

Có nhiều trường hợp không thích sử dụng “*thì*”, “*là*”, “*mà*”, người viết sẽ thay vào đó bằng một *dấu phẩy*:

(102) “Nga, con ông phủ họ Lê”.

(*Lá ngọc cành vàng* – Nguyễn Công Hoan)

⁴⁾ bổ nghĩa và phụ nghĩa khác nhau, sẽ giải thích ở các chương sau.

Dấu phẩy có công dụng phân giới giữa hai phần: đề và diễn. “*Nga*” là phần đề. “*Con ông phủ họ Lê*” là phần diễn. Dấu phẩy thay thế cho “*là*” trong câu “*Nga là con ông phủ họ Lê*”. Thử phân tích các bài viết của các nhà văn hay dùng dấu phẩy thì sẽ thấy. Thay vì dùng dấu phẩy trong văn viết, khi nói, người ta ngưng lại một chút.

Nguyên tắc đề diễn là một đặc điểm dễ nhận thấy trong tiếng Việt, song - như đã nói - không phải chỉ có người Việt mới hành ngôn theo nguyên tắc này. Hầu hết các dân tộc khác cũng hành ngôn như thế chỉ có điều là có thể bằng hình thức khác. Ví dụ người Anh thường dùng *câu chẻ* (*cleft sentence*)⁵⁾:

(103) It was I who knocked at your door.

“*It was I*” là phần đề. “*Who knocked at your door*” là phần diễn. Cũng bởi vì - như đã nói - câu cú của tiếng Anh không chấp nhận sự thiếu vắng của chủ ngữ ngữ pháp, cho nên người ta phải đưa “*it*” vào làm chủ ngữ cho động từ “*was*”, đồng thời dùng “*who*” đại diện cho “*I*” (tức chủ thể logic). Chủ ngữ “*who*” làm chủ phần động từ “*knocked*”. Phần diễn “*Who knocked at your door*” có chức năng bổ nghĩa cho phần đề. Phần này bắt buộc phải có, chứ không sẽ không ai hiểu được “*It was I*” có liên quan đến việc gì.

Cấu trúc câu trên không phải là một cấu trúc thông dụng của tiếng Đức. Thay vì thế, muốn bổ nghĩa đối tượng với một hình thức na ná vậy, người Đức thường dùng *mệnh đề phụ* (*subordinate clause*)⁶⁾:

(104) Das Haus, in dem Sie wohnen, ist geisterhaft.

(Cái nhà) (trong đó) (bà) (ở) (thì) (có ma quỷ)

(105) Cái nhà bà đang ở có ma.

Mệnh đề phụ “*in dem Sie wohnen*” là một hình thức chỉ định “*das Haus*” (cái nhà). Hoặc bằng một hình thức khác:

(106) “Der Abend wiegte schon die Erde,
Und an den Bergen hing die Nacht”.
(Goethe)

⁵⁾ xem chi tiết ở các chương sau.

⁶⁾ sẽ giải thích sau.

Tạm dịch:

Chiều tối đọng đưa mặt đất,
Và bên núi, đêm treo.

Đối tượng trong câu thơ thứ hai mà Goethe muốn diễn tả là những ngọn núi bị màn đêm che phủ. Để phản ngược câu đầu mang *thái chủ động (active voice)*, Goethe muốn câu sau mang *thái bị động (passive voice)*. Nếu lập câu đại khái như “*núi bị màn đêm che phủ*” thì dở quá. Vả lại Goethe muốn chọn một động từ mang tính tĩnh “*hing*” (*treo*) để đối lại động từ mang tính động của câu đầu “*wiegte*” (*đu đưa, đong đưa, ru ngủ*). Thay vì dùng trạng thái bị động để đối lại trạng thái chủ động, Goethe đặt đối tượng “*núi*” vào một *ngữ đoạn giới từ (prepositional phrase)*⁷⁾ “*an den Bergen*” (*bên núi*) rồi diễn tả tiếp bằng đoạn “*hing die Nacht*” (*đêm treo*).

Nói tóm lại, Đề-Diễn là một cách thức diễn đạt rất phổ biến của người Việt. Nhiều người ngoại quốc học tiếng Việt do không quen cách ngữ dụng này của người Việt, cho nên cách hành ngôn của họ đối với người Việt nhiều khi nghe không được tự nhiên cho lắm. Chỉ cần giải thích cho họ hiểu rõ nguyên tắc này là họ đã nắm bắt được phần lớn cách nói giống như người Việt.

Nguyên tắc đề diễn là một đề tài khá thú vị đã được nhiều người nghiên cứu. Tiêu biểu là các công trình về *Ngữ pháp chức năng (Functional Grammar)* của Halliday (1994).

Tựu trung, mỗi dân tộc đều có cách nhìn và mô tả đối tượng khác nhau, song mục đích cuối cùng vẫn là làm sao mô tả cho được đối tượng bằng một hình thức hành ngôn nào đó. Dù khác biệt thế nào đi nữa, ngôn ngữ vẫn không bất định, kể cả sự đổi thay của nó theo thời gian mà chúng ta sẽ thấy ở chương sau.

⁷⁾ sẽ giải thích sau.

Chương 3

SỰ THAY ĐỔI CỦA NGÔN NGỮ

Từ ngữ của một ngôn ngữ có thể thay đổi nhanh chóng theo đà phát triển của khoa học, xã hội, tư duy. Con người cần từ để biểu thị đối tượng. Có đối tượng trước kia chưa từng có, đến lúc nó xuất hiện, người ta sẽ tìm cách đặt cho nó một cái tên. Ngày xưa chưa có các phương tiện Internet, Computer, Hardware, Software, ... đến khi phát minh ra được những thứ ấy, người ta mới đặt tên, từ đó mới có thêm những từ mới cho những khái niệm ấy.

Có thể nói, từ mới xuất hiện ngày càng nhiều. Dĩ nhiên cũng có nhiều từ đã lỗi thời, không còn thông dụng.

Ngữ pháp của một ngôn ngữ cũng thay đổi theo thời gian nhưng không quá nhanh chóng như từ ngữ. Điều này không có gì khó hiểu, bởi cấu trúc ngữ pháp phản ánh cái nhận diện và nhận thức tự nhiên của con người về đối tượng muốn diễn tả hoặc đã được quy ước. Trật tự của hệ thống này không thể thay đổi liên miên, tùy cảm hứng. Lấy tiếng Đức làm ví dụ, giờ bất kỳ một tác phẩm nào đã được viết cách đây mấy trăm năm về trước ra đọc, người đọc vẫn nhận thấy rằng, cấu trúc ngữ pháp xưa của tiếng Đức không khác ngữ pháp ngày hôm nay cho lắm. Ví dụ một đoạn thư của Schiller gửi Goethe kể về kịch phẩm “*Maria Stuart*” của mình:

“Ich werde Ihnen, wenn Sie hier sind, einige tragische Stoffe, von freier Erfindung, vorlegen, um nicht in der ersten Instanz, in dem Gegenstande, einen Mißgriff zu tun. Neigung und Bedürfnis ziehen mich zu einem frei phantasierten, nicht historischen, und zu einem bloß leidenschaftlichen und menschlichen Stoff, denn Soldaten, Helden und Herrscher habe ich vor jetzt herzlich satt ” (Trích thư Schiller gửi Goethe, 19.03.1799, Nxb. Reclam 1986)

Ngữ pháp tiếng Đức mà Schiller đã dùng cách đây hai thế kỷ vẫn y hệt như ngữ pháp bây giờ.

Hoặc một đoạn thơ (Brief-Gedicht) của Goethe gửi cho nhóm bạn Riese ở Frankfurt năm 1766, lúc Goethe 16 tuổi:

“Da sah ich erst, daß mein erhabener Flug,
Wie er mir schien, nichts war als das Bemühn
Des Wurms im Staube, der den Adler sieht
Zur Sonn sich schwingen und wie der hinauf ...”
(Reclam 1967)

Ngữ pháp của Goethe dùng cách đây gần 250 năm về trước không khác ngữ pháp ngày hôm nay.

Không kể những cổ ngữ như tiếng Phạn, tiếng La tinh, là những ngôn ngữ cổ, song nhiều ngôn ngữ ngày nay, sống và sống mạnh, chẳng hạn như tiếng Anh, ngữ pháp cũng không thay đổi nhiều bằng sự bùng nổ từ ngữ. Các ngôn ngữ Âu châu khác (Pháp, Tây Ban Nha, Ý, ...) tương tự vậy.

Thế còn ngữ pháp tiếng Việt ?

Giống mọi ngôn ngữ khác trên thế giới, ngữ pháp tiếng Việt cũng không thay đổi nhanh chóng như sự phát triển của từ. Thế nhưng so với tốc độ thay đổi ngữ pháp của các ngôn ngữ khác, chưa chắc tiếng Việt đã chậm hơn và điều đó cũng không có nghĩa là tích cực.

Sau đây, chúng ta hãy thử xem sự thay đổi ngữ pháp tiếng Việt bằng cách khảo sát một vài văn bản xưa và nay, văn viết cũng như văn nói. Cụ thể, hãy chọn một vài văn bản tiêu biểu của *Lý Thường Kiệt* (thế kỷ 11), *Nguyễn Trãi* (thế kỷ 14-15), *Hồ Xuân Hương* (thế kỷ 18), *Nguyễn Du* (thế kỷ 18), *Lê Quý Đôn* (thế kỷ 18), *Nguyễn Công Trứ* (thế kỷ 19), *Trần Tế Xương* (thế kỷ 19), *Trần Trọng Kim* (thế kỷ 20). Xin lưu ý, chúng ta sẽ không dùng lý thuyết để phân tích ngữ pháp, bởi điều này đòi hỏi rất nhiều kiến thức về lý thuyết ngữ pháp mà chương này chưa thể đi vào chi tiết. Thay vì thế, chúng ta dùng phương pháp so sánh cách hành ngôn. Mục đích của việc so sánh là để chúng ta có cảm giác, đó là ngữ pháp của tiếng Việt, dù đã trải qua nhiều thế kỷ song vẫn không có nhiều thay đổi. Cách hành ngôn vẫn mang những đặc điểm của tiếng Việt đã được đề cập ở chương trước. Những đặc điểm này xuất hiện rõ nhất trong các văn bản chữ Nôm được dịch sang chữ Quốc Ngữ.

Phân tích xong tiếng Việt ngày xưa, chúng ta sẽ xét đến tiếng Việt ngày nay. Do điều kiện lịch sử, cấu trúc ít nhiều gì cũng đã chịu ảnh hưởng của ngôn ngữ Âu châu. Thay vì xác định chính xác, tiếng Việt đã bị ảnh hưởng từ bên ngoài đến độ nào, bao nhiêu phần trăm, chúng ta cũng sẽ dựa vào các đặc điểm của tiếng Việt để so sánh sự khác biệt giữa cách hành ngôn của người Việt và cách hành ngôn của các dân tộc khác. Hãy chọn tiếng Anh và tiếng Đức làm ví dụ.

TIẾNG VIỆT XƯA

Lý Thường Kiệt (1019-1105): Năm 1077, Lý Thường Kiệt đã cho ra đời bản tuyên ngôn độc lập đầu tiên trong lịch sử dân tộc Việt: *Nam Quốc Sơn Hà*. Tương truyền rằng, bài thơ đã được đọc lên ngay cuộc chiến giữa quân Đại Việt và mấy chục vạn quân Tống do Quách Quỳ chỉ huy trên trận tuyến sông Như Nguyệt (khúc sông Cầu thuộc huyện Yên Phong, tỉnh Hà Bắc ngày nay).

南 國 山 河

南	國	山	河	南	帝	居
截	然	定	分	在	天	書
如	何	逆	虜	來	侵	犯
汝	等	行	看	取	敗	虛

Nam Quốc Sơn Hà – Lý Thường Kiệt, 1077 (bản Hán). (Tài liệu: Trang Văn Hóa Việt Nam, www.vhvn.com)

Bản Hán:

“Nam Quốc sơn hà Nam Đế cư
 Tiệt nhiên định phận tại thiên thư
 Như hà nghịch lỗ lai xâm phạm ?
 Nhữ đẳng hành khan thủ bại hư !”

Hoa Đào Thi

1

Một đoá đào hoa khá tốt tươi
Cánh xuân mơn mớn thấy xuân cười
Đông phong ắt có tình hay nữa
Kín tiền mùi hương dễ động người

蔑朶桃花窖卒鮮
隔春嗽嗽体春嗤
東風乙固情哈女
建羨味香易動得

2

Động người hoa khéo tỏ tình thần
Ít bởi vì hoa, ít bởi xuân
Dễ sử chim xanh đừng chuốc lối
Bù trì đã có khí hồng quân

動得人窖烁精神
少摆爲花少摆春
吧使貼撐停祝磊
逋持色固氣洪鈞

3

Khí hồng quân hãy sá tài qua
Chở phụ xuân này chở phụ hoa
Hoa có ý thì xuân có ý
Đâu đâu cũng một khí dương hoà

氣洪鈞矣舍才戈
渚負春尼渚負花
花固意時春固意
兜兜共蔑氣陽和

4

Khí dương hoà há có tư ai
Năng một hoa này nhần mọi loài
Tính kể chẵn còn ba tháng nữa
Kịp xuân mưa để má đào phai

氣陽和可固私埃
能蔑花尼忍每類
算計參群匹胸女
及春馬底騰桃派

5

Má đào phai hết bởi xuân qua
Nếu lại đâm thì liền luống hoa
Yến sở Dao trì đà có hẹn
Chở cho Phương Sóc đến lân la

顰桃派歇摆春戈
裊吏宄郛連隴花
宴所瑤池它固限
渚朱方朔旦鄰羅

6

Phương Sóc lân la đã hờ cơ
Ba phen trộm được há tình cờ
Có ai ướm hỏi tiên Vương Mẫu
Tin khá tin thì ngờ khá ngờ

方朔鄰羅色許機
匹番濫特呵情期
固埃厭晦仙王每
信可信郛疑可疑

Hoa Đào Thi – Nguyễn Trãi. (Tài liệu: Viện Việt Học, www.viethoc.com)

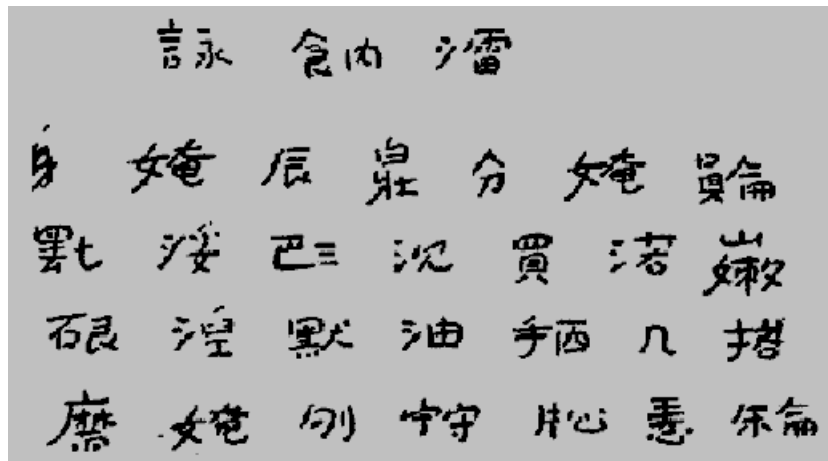
So với bài Nam Quốc Sơn Hà của Lý Thường Kiệt, ta cảm thấy ngữ pháp của Nguyễn Trãi trong bài Hoa Đào Thi Đọc có vẻ gần gũi với chúng ta hơn, dù đó là loại ngữ pháp cách đây hơn 5 thế kỷ (thế kỷ 14-15). Càng gần gũi nữa là ngữ

pháp tiếng Việt của Hồ Xuân Hương (thế kỷ 18), cũng không khác xa ngữ pháp tiếng Việt bây giờ cho lắm:

Hồ Xuân Hương (1706-1783): Đánh Đu

“Bốn cột khen ai khéo khéo trồng
 Người thì lên đánh, kẻ ngồi trông
 Trai đu gối hạc khom khom cật
 Gái uốn lưng ong ngửa ngửa lòng.
 Bốn mảnh quần hồng bay phấp phới,
 Hai hàng chân ngọc duỗi song song.
 Chơi xuân đã biết xuân chẳng tá ?
 Cọc nhỏ đi rồi, lỗ bỏ không !”

Hoặc *Vịnh Bánh Trôi* được dịch từ nguyên bản chữ Nôm sang chữ Quốc Ngữ:



Bảng chữ Nôm *Vịnh Bánh Trôi* - Hồ Xuân Hương. (Tài liệu: *L'oeuvre de la poétesse vietnamienne Hồ Xuân Hương*, Maurice Durand. EFEO, 1968, p. 155)

“Thân em thì trắng phận em tròn
 Bảy nổi ba chìm với nước non
 Dấn nát mặc dầu tay kẻ nặn
 Mà em vẫn giữ tấm lòng son”
 (*Vịnh bánh trôi* – Hồ Xuân Hương)

Nguyễn Du (1765-1820): Ngữ pháp tiếng Việt của Nguyễn Du cũng không làm cho người Việt ngày nay bỡ ngỡ:

“Trăm năm trong cõi người ta,
Chữ tài chữ mệnh khéo là ghét nhau.
Trải qua một cuộc bể dâu,
Những điều trông thấy mà đau đớn lòng.
Lạ gì bỉ sắc tư phong,
Trời xanh quen thói má hồng đánh ghen.
Cảo thơm lần giở trước đèn,
Phong tình cổ lục còn truyền sử xanh ...”

Đó là cách dùng ngữ pháp trong thơ, nay hãy thử xem trong văn xuôi thế nào.

Một vài ví dụ văn viết và văn nói trong tác phẩm *Đại Việt Thông Sử* của **Lê Quý Đôn (1759)** do dịch giả Lê Công Liệu (1973) dịch lại. Có những đoạn viết về triều Lê (quyển 1):

Văn viết:

Thái Tổ Thượng [tờ 7a]: Vuy Thái Tổ, thụy hiệu << Thống Thiên khải vận thánh – đức thần – công duệ – văn anh – vũ khoan – minh dũng – trí hoàng – nghĩa trí – nhân đại – hiếu Cao Hoàng Đế >>. Vua họ Lê, tên húy là Lợi, người Lam Giang, huyện Lương giang, phủ Thanh Hoa. Cụ Tăng Tổ của vua tên húy là Hối, sau truy tôn là “Cao Thượng – tổ Minh Hoàng Đế”. Tính cụ chất phác ngay thẳng, giữ mình như người ngu, nhưng hiểu biết rất sâu xa, có thể biết trước những sự chưa thành hành. [tờ 7b]: Nguyên trước ở thôn Như Áng, một hôm, cụ đi chơi, thấy đàn chim liệng vòng quanh trên một khoảng đất nơi dưới núi Lam Sơn, trông hình như một đám người tụ hội. Cụ tự nghĩ: “Chỗ này tất là nơi đất lành”, bèn dời nhà đến ở đấy, rồi khai phá ruộng vườn, tự chăm lo cày cấy, được 3 năm, trở thành một sản nghiệp, tự đấy, đời đời đều là hùng trưởng một phương. Sau này vua dựng đô mở nước, thực cũng có căn cơ tự đấy vậy.”

“Ngày Hoàng đế sanh thì trong nhà có hào quang đỏ chiếu sáng rực, và mùi thơm ngào ngạt khắp làng. [tờ 8b]: Khi lớn tuổi, ngài thông minh dũng lược, độ lượng hơn người, vẻ người tươi đẹp hùng vĩ, mắt sáng mồm rộng; sống mũi cao, xương my mắt gồ lên; bả vai bên tả có 7 nốt ruồi, bước đi như rồng như hổ; tiếng nói vang vang như tiếng chuông. Các bậc thức giả biết ngay là một người phi thường”

Văn nói:

Một đoạn viết về Lương Nhữ Hốt khai báo với nhà Minh về hoạt động của Lê Lợi:

“[tờ 10b]: ...“Người chủ Lam Sơn chiêu nạp những kẻ vong mạng, và đãi ngộ sĩ tốt rất hậu, chí của người ấy không phải là nhỏ. Nếu không sớm tính đi, để cho con rồng có lúc gặp mây gặp mưa, thì khi ấy nó sẽ không còn là một con vật trong ao nữa đâu. Vậy xin trừ ngay đi, đừng để lưu tai vạ về sau này”.

Một đoạn viết về Lê Lợi hỏi các tướng sĩ:

[tờ 11b]: ... “Có ai dám bắt chước việc Kỷ Tín thời xưa không ?”

Một đoạn viết về Lê Lợi sai hai tướng Trịnh Khả và Lê Lôi mang thông điệp sang Vương Quốc Lào, dọa sẽ đánh nếu không nạp lương thực, vũ khí và voi cho quân Lê Lợi:

[tờ 12a]: ... “Quốc gia tôi phụng thờ thông điệp của triều Đại Minh ban cho nhà vua. Vậy nhà vua hãy đem số lương thực đủ quân sĩ dùng trong 5 tháng, và khí giới cùng voi trận tới yết kiến, rồi nhận tờ điệp văn về thi hành, để khỏi phải bắt giải. Nếu không tuân mệnh, lập tức sai nước Xa Lý và Lão Qua hợp quân 6 nước tiến đánh”.

Một đoạn viết về Lê Lợi bàn với các tướng sĩ:

[tờ 15b]: ... “Hiện tình, bên địch nhiều quân, bên ta thì ít, bên địch hành quân vất vả, mà bên ta thì quân sĩ được thanh thoi. Binh pháp có câu nói rằng: “Sự thắng địch cốt ở tướng giỏi, chứ không quan hệ ở số quân nhiều hay ít”. Nay chúng ta tuy nhiều quân, nhưng ta dùng quân được thanh thoi đánh quân vất vả, nhọc mệt, thì tất phá được ”

Giả như dịch giả dịch đúng, thì ngữ pháp Lê Quý Đôn dùng cho văn viết và văn nói xưa kia cũng không xa lạ gì mấy đối với người Việt ngày hôm nay.

Xem thêm một vài ví dụ nữa:

Bài tựa bằng Hán văn do Bùi Kỳ và Trần Trọng Kim dịch và in ở đầu quyển Kim Vân Kiều, Vĩnh Hưng long thư quán, Hà Nội, 1925:

“Trong trời đất đã có người tài tình tuyệt thế, tất thế nào cũng có việc khảm kha bất bình. Tài mà không được gặp gỡ, tình mà không được hả hê, đó là cái căn

nguyên của đoạn trường vậy. Thế mà lại có kẻ thương tiếc tài tình, xem thấy việc, trông thấy người thì còn nhin thế nào được mà không thở than rên rĩ !

Nghĩa là bậc thánh mới quên được tình, bậc ngu không hiểu tới tình, tính chung chú vào đâu, chính là chung chú vào bọn chúng ta vậy. Cho nên phàm người đã ít tình tất là không có tài, chỉ nửa lừa nửa sáng, sống chết ở trong vòng áo mũ, trong cuộc no say, dù có gặp cái cảnh thanh nhã như hoa thơm buổi sáng, trắng tỏ ban đêm, cũng chỉ trơ trơ như cây cỏ, như chim cá vậy.

[...]

Tháng hai năm Canh Thìn niên hiệu Minh Mệnh (*), viết ở Thán hoa hiên đất Hạc giang. Tiên phong Mộng Liên đường chủ nhân”

(*): Tài liệu của Sài Phi thư trang có ghi rõ: Tháng hai niên hiệu Minh-mệnh năm đầu (1820).

Nguyễn Công Trứ (1778-1858): Vịnh Cảnh Nghèo

“Chẳng phải rằng ngậy chẳng phải đàn
Bởi vì nhà khó hóa bần thân.
Mấy đời thầy kiện mà thua mọ,
Nghĩ phận thằng nghèo phải biết thân.
...”

Trần Tế Xương (1870-1907): Đi Thi Tự Vịnh

“Tấp tễnh người đi tớ cũng đi,
Cũng lều, cũng chõng, cũng vô thi.
Tiễn chân, cô mất ba đồng lẻ;
Sờ bụng, thầy không một chữ gì.
...”

Đã lướt qua vài ví dụ, hẳn chúng ta ai cũng có cảm giác rằng, ngữ pháp tiếng Việt ngày xưa không khác ngày hôm nay cho lắm. Với thẩm năng ngôn ngữ, người Việt có thể nhận ra ngay đó là cách hành ngôn của người Việt chứ không phải của người Anh, người Đức, người Pháp. Có lạ chăng là tác giả dùng nhiều chữ cổ, hoặc trong câu cú, nhất là ở thể văn xuôi, “*thì*”, “*là*”, “*mà*” không bị tính lược nhiều. Mặt khác, cách dùng ngữ pháp của nhiều tác giả không khác nhau nhiều. Lý do này cũng dễ hiểu bởi cách hành văn chứa nhiều đặc điểm tiêu biểu

của tiếng Việt. Chúng ta sẽ kiểm chứng lại dưới đây thêm lần nữa. Điều này quan trọng, bởi cần nhận diện rõ cái căn cước của tiếng Việt trước khi dùng lý thuyết để so sánh với các ngôn ngữ khác. Chỉ khi đã nắm bắt được những đặc điểm của tiếng Việt, chúng ta mới hy vọng hiểu được tại sao những lý thuyết của ngôn ngữ học hiện đại có thể giải thích thành công các ngôn ngữ khác nhưng lại gặp khó khăn với tiếng Việt trong nhiều trường hợp.

1. Tính tịnh tiến

Có thể nói, cách phản ánh đối tượng của người xưa khá gắn bó với nguyên tắc tịnh tiến. Nguyên tắc tịnh tiến hiện ra trong cách hành ngôn như một hình thức sao chép lại cách nhận diện và nhận thức của con người về một đối tượng đang được để ý tới. Trình tự này quyết định trật tự của từ ngữ trong chuỗi từ được nói ra. Nếu xem sự mô tả một đối tượng như một tiêu điểm, ta có thể có nhiều kết quả xuyên qua nhiều cách đặt câu hỏi hữu lý khác nhau. Song chắc chắn sẽ có một kết quả thỏa một câu hỏi, nghĩa là thỏa nguyên tắc tịnh tiến. Ví dụ:

- (1) Sông núi nước Nam vua Nam ở.
(Lý Thường Kiệt)

Khi nghe câu này, người nghe có thể đặt câu hỏi bằng nhiều cách khác nhau:

- (2) *Hỏi: Sông núi thế nào ?*
(3) *Trả lời: Sông núi hùng vĩ và thơ mộng.*

hoặc:

- (4) *Hỏi: Sông núi ở đâu ?*
(5) *Trả lời: Sông núi ở nước Nam.*

Thế nhưng tác giả đã tự chọn một câu hỏi hữu lý khác và trả lời:

- (6) *Hỏi: Sông núi nào ?*
(7) *Trả lời: Sông núi nước Nam.*

Tương tự vậy:

- (8) *Hỏi: Sông núi nước Nam thế nào ?*

(9) *Trả lời: Sông núi nước Nam vua Nam ở.*

Tính tịnh tiến cũng xuất hiện trong những câu thơ còn lại của Lý Thường Kiệt:

(10) Rành rành định phận tại sách trời.

(11) Cớ sao lũ giặc sang xâm phạm ?

(12) Chúng bay sẽ bị đánh tơi bời !

Tính tịnh tiến cũng xuất hiện trong thơ Nguyễn Trãi. Tác giả Lê Văn Đăng đã dịch 1-1 từ Nôm sang Quốc Ngữ, bảo đảm sao đúng ngữ pháp của văn bản gốc. Đọc thử hai câu đầu của Hoa Đào Thi:

(13) Một đoá đào hoa khá tốt tươi

Cánh xuân mơn mẩn thấy xuân cười

Thử kiểm chứng lại những cách đặt câu hỏi hữu lý sau đây:

Hỏi: Một đoá gì ?

Trả lời: Một đoá đào hoa.

Chúng ta thấy, giống Lý Thường Kiệt, Nguyễn Trãi cũng bị “nhiễm” Hán phần nào. Cấu trúc của cụm từ “*đào hoa*” là cấu trúc Hán; thuần Việt phải là “*hoa đào*”: “*Một đoá hoa đào*”.

Hỏi: Một đoá đào hoa ra sao ?

Trả lời: Một đoá đào hoa khá tốt tươi.

V.v.

Sau nhiều thế kỷ, tính tịnh tiến không biến mất mà vẫn tồn tại trong ngôn ngữ. Hồ Xuân Hương cũng có cùng cách lập câu tịnh tiến giống như Lý Thường Kiệt, Nguyễn Trãi, thậm chí còn thuần Việt hơn nữa, vì Hồ Xuân Hương không dùng cấu trúc Hán:

(14) Trai đu gối hạc khom khom cật.

Gái uốn lưng ong ngửa ngửa lòng.

Bốn mảnh quần hồng bay phấp phới,

Hai hàng chân ngọc duỗi song song.

Thử kiểm chứng lại những cách đặt câu hỏi hữu lý sau đây:

Hỏi: Trai thế nào ?

Trả lời: Trai đu gối hạc khom khom cật.

Hỏi: Gái làm sao ?

Trả lời: Gái uốn lưng ong giữa giữa lòng.

Hỏi: Bốn cái gì ?

Trả lời: Bốn mảnh quần hồng.

Hỏi: Bốn mảnh quần hồng thế nào ?

Trả lời: Bốn mảnh quần hồng bay phấp phới.

Hỏi: Hai hàng gì ?

Trả lời: Hai hàng chân ngọc

Hỏi: Hai hàng chân ngọc thế nào ?

Trả lời: Hai hàng chân ngọc đuổi song song.

Như một trò chơi, chúng ta có thể xét những câu khác cũng bằng cách đặt câu hỏi và trả lời y thế. Hễ thấy cấu trúc câu hỏi và câu trả lời giống nhau thì đấy chính là cách lập câu tịnh tiến. Thế nhưng nên lưu ý - như đã nói - là chỉ có thể kiểm chứng tính tịnh tiến bằng cách lấy ra một câu trả lời có sẵn rồi xem câu trả lời ấy đi từ câu hỏi nào.

2. Tính giàu thông tin cụ thể

Tính giàu thông tin cụ thể trong *hoạt động ngôn ngữ* của người Việt đã xuất hiện từ xưa chứ không phải gần đây mới thấy. Hiện tượng mô tả tỉ mỉ về một đối tượng theo ngôn ngữ học hiện đại Âu châu có thể diễn ra với 3 hình thức (sẽ được nói rõ hơn ở các chương sau):

- dùng *bổ ngữ (complement)* để bổ sung thêm nghĩa ý nghĩa cần thiết cho đối tượng cần được diễn giải.
- dùng *phụ ngữ (adjunct)* để phụ thêm nghĩa cho đối tượng nhưng không bắt buộc.
- dùng *biệt định ngữ (specifier)* để chỉ định tính chất cá biệt của đối tượng ấy.

Xem thử một câu khá dài của Lê Quý Đôn mô tả về Lê Lợi:

“ Khi lớn tuổi, ngài thông minh dững lược, độ lượng hơn người, vẻ người tươi đẹp hùng vĩ, mắt sáng mồm rộng; sống mũi cao, xương my mắt gồ lên; bả vai bên tả có 7 nốt ruồi, bước đi như rồng như hổ; tiếng nói vang vang như tiếng chuông. Các bậc thức giả biết ngay là một người phi thường”

Cách mô tả chi tiết về một đối tượng theo kiểu Việt Nam này rất khó dịch sang ngôn ngữ Âu châu chỉ bằng một câu duy nhất. Ví dụ như trong tiếng Đức, muốn mô tả chi tiết như vậy, người ta phải dùng nhiều *liên từ* (*conjunction*) “und” (“và”) để nối từng thông tin một thay vì dùng dấu phẩy như người Việt. Đường nhiên cũng có trường hợp họ dùng dấu phẩy, nhưng cách hành văn với quá nhiều yếu tố phụ/bổ nghĩa này, họ không quen; vả lại dùng quá nhiều “und”, “und”, “und”, ... sẽ làm câu nói trở nên rườm rà, đơn điệu, không hay lắm. Thay vì vậy, họ chia ra từng câu nhỏ hoặc dùng mệnh đề phụ để nói thêm về đối tượng muốn diễn tả. Hoặc họ dùng một cụm tính từ, hoặc mệnh đề như một hình thức tính từ đặt trước một danh từ, thế nhưng độ phức tạp cũng có giới hạn. Nhìn chung, cách mô tả này không phổ biến trong thực tế dùng tiếng Đức. Trong các tiếng Âu châu khác có lẽ cũng thế.

Câu sau cũng không khác, chứa nhiều thông tin về đối tượng:

“Thế mà lại có kẻ thương tiếc tài tình, xem thấy việc, trông thấy người thì còn nhin thế nào được mà không thở than rên rĩ”

Chỉ có một chủ thể “kẻ” thôi mà lại có nhiều *vị ngữ* (*predicate phrase*) (tạm thời hiểu nôm na là phần động từ + bổ ngữ (tức từ), sẽ được nói rõ sau). Từng vị ngữ được tách ra bằng dấu phẩy. “Kẻ” ở đây là một nhân vật xưa nay thường được gọi là chủ từ:

- (15) thương tiếc tài tình
- (16) xem thấy việc
- (17) trông thấy người thì còn nhin thế nào được mà không thở than rên rĩ

3. Tính tỉnh lược

Xem cách sử ngữ dụng pháp của người xưa, chúng ta còn thấy cả hiện tượng lược từ (giống ngày nay). Tính tỉnh lược thường thấy trong các thể thơ có luật.

- (18) Sờ bụng, thầy không (có) một chữ gì.
(*Đi Thi Tự Vịnh* - Trần Tế Xương)

“*Có*” bị lược bỏ. Đúng ra là “*thầy không có*”. Động từ “*có*” đòi hỏi một bổ ngữ (tức túc từ trực tiếp như thường gọi).

- (19) Thân em thì trắng, phận em (*thì*) tròn.
(*Vịnh cái quạt* – Hồ Xuân Hương)

Lối nói tỉnh lược này là một hình thức ngữ dụng (pragmatic), ở đó, một từ nằm trong mệnh đề theo sau thường bị lược bỏ vì đã nằm trong mệnh đề phía trước. Đúng ra là “*phận em **thì** tròn*”. Nhưng “*thì*” đã xuất hiện ở mệnh đề “*thân em **thì** trắng*” phía trước, cho nên bỏ được. Cũng có trường hợp, Hồ Xuân Hương không lược từ:

- (20) Hai bên *thì* núi, giữa *thì* sông.
(*Kẽm Trống* – Hồ Xuân Hương)

Trong thơ, do số chữ bị thể thơ giới hạn một cách nghiêm ngặt khiến tác giả không thể tuân thủ đúng ngữ pháp, song trong văn xuôi, hiện tượng tỉnh lược vẫn xảy ra dù không quá phổ biến như ngày nay. Vài ví dụ trong Đại Việt Thông Sử của Lê Quý Đôn:

- (21) ... Hiện tình, bên địch (*thì/có*) nhiều quân, bên ta thì ít.
(22) Nếu không tuân mệnh, (*chúng tôi sẽ*) lập tức sai nước Xa Lý và Lão Qua hợp quân 6 nước tiến đánh.

4. Tính cảm đề

Lược từ làm câu mất thông tin, khiến người nghe phải hiểu ngầm:

- (23) Nếu không tuân mệnh, lập tức sai nước Xa Lý và Lão Qua hợp quân 6 nước tiến đánh.

Câu trên của Lê Quý Đôn muốn diễn tả rằng, nếu vương quốc Lào không tuân theo mệnh lệnh của Lê Lợi, lập tức Lê Lợi sẽ sai nước Xa Lý và Lão Qua hợp quân 6 nước tiến đánh nước Lào ngay. Ở chỗ trống ấy, người nghe phải hiểu ngầm là “*chúng tôi*”, một ngoại chủ thể không có mặt trong câu, một thành tố zéro (sẽ nói sau).

5. Tính đề diễn

Nguyên tắc đề diễn của tiếng Việt đã xuất hiện vào thế kỷ thứ 11 qua *Nam Quốc Sơn Hà*:

- (24) Sông núi nước Nam vua Nam ở.
(Nam Quốc Sơn Hà Nam Đế cư)

Câu trên có thể hiểu là:

- (25) Sông núi nước Nam *thì* vua Nam ở.

Chủ đề ở đây là “*Sông núi nước Nam*” (vì vậy mà Nam Quốc Sơn Hà còn là tựa đề của bài thơ). “*Sông núi nước Nam*” còn được diễn giải thêm bằng những câu thơ kế tiếp:

- (26) Rành rành định phận tại sách trời.
(Tiệt nhiên định phận tại thiên thư)

Nghĩa là: “*Sông núi nước Nam*” rõ ràng đã được sách trời quyết định ở nơi chốn đó.

- (27) Cớ sao lũ giặc sang xâm phạm ?
(Như hà nghịch lỗ lai xâm phạm ?)

“*Xâm phạm*” được hiểu là xâm phạm “*sông núi nước Nam*”. Và cuối cùng là lời cảnh cáo, nếu xâm phạm nước Nam:

- (28) Chúng bay sẽ bị đánh tơi bời
(Nhữ đẳng hành khan thủ bại hư)

Có thể trong thơ Lý Thường Kiệt, chúng ta chưa thấy rõ ranh giới giữa phần đề và phần diễn, thế nhưng trong thơ Hồ Xuân Hương, ranh giới của cấu trúc đề diễn hiện ra khá rõ qua cách phân đoạn bằng “*thì*”, “*là*”:

- (29) Người *thì* lên đánh, kẻ ngồi trông.
(30) Thân em *thì* trắng phận em tròn.
(31) Tất cả những *là* thu với vén.
(32) Chồng con cái nợ *là* như thế.

Các đối tượng, tức chủ đề, cần được diễn giải là “*người*”, “*thân em*”, “*tất cả những*”, “*chồng con cái nợ*”.

Đặc biệt trong Kiều của Nguyễn Du, cấu trúc đề diễn có mặt khắp nơi. Hai phần “*đề*” và “*diễn*” được Nguyễn Du phân biệt rất chuẩn bằng “*thì*”, “*là*”, “*mà*”:

- (33) “Bên *thì* mấy ả mày ngài,
 Bên *thì* ngồi bốn năm người làng chơi.
 Giữa *thì* hương án hắt hoi,
 Trên treo một tượng trắng đôi lông mày.
 Lầu xanh quen lối xưa nay,
 Nghề này *thì* lấy ông này tiên sư.”
- (34) “Chẳng qua đồng cốt quàng xiên
 Người đâu *mà* lại thấy trên côi trần ?”
- (35) “Đã đem mình bán cửa tao,
 Lại còn khùng khỉnh làm cao thế này !
 Nào *là* gia pháp nọ bay !
 Hãy cho ba chục biết tay một lần”

Tất nhiên cấu trúc đề diễn cũng xuất hiện trong văn viết. Lê Quý Đôn đã diễn tả ngày vua ra đời như sau:

- (36) Ngày Hoàng đế sanh *thì* trong nhà có hào quang đỏ chiếu sáng rực, và mùi thơm ngào ngạt khắp làng.

Đoạn “*ngày Hoàng Đế sanh*” là phần đề. Đoạn theo sau là phần diễn.

Hoặc:

- (37) Vua họ Lê, tên húy là Lợi, người Lam Giang, huyện Lương giang, phủ Thanh Hoa.

Trong câu này, “*vua*” là phần đề, những đoạn còn lại là phần diễn giải về “*vua*”. Bởi thiếu chữ “*thì*”, thành ra hơi khó phân biệt. Nếu lược bỏ “*họ Lê*” và “*húy*” ở hai đoạn đầu, chúng ta sẽ được một câu quen thuộc mà ngày nay người Việt vẫn hay dùng:

- (38) Vua tên là Lợi.

“*Vua*” chính là chủ thể tâm lý. “*Tên*” là chủ ngữ ngữ pháp. Tương tự vậy:

(39) *Cụ Tăng Tổ của vua tên húy là Hối.*

“*Cụ Tăng Tổ của vua*” là phần đề, đoạn còn lại là phần diễn (“*tên húy là Hối*”).
Lược bỏ “*húy*”, ta có:

(40) *Cụ Tăng Tổ của vua tên là Hối.*

“*Cụ Tăng Tổ của vua*” là chủ thể tâm lý.

TIẾNG VIỆT NAY

Đã khảo sát sơ quát về ngữ pháp tiếng Việt xưa, vậy, tiếng Việt ngày hôm nay thế nào ? Có một điều có thể tiên đoán được là ngữ pháp thay đổi theo lịch sử. Tất nhiên. Xã hội Việt Nam trong thế kỷ 20 đã thay đổi rất nhiều. Về ngôn ngữ nói riêng, tiếng Việt đã không thoát khỏi ảnh hưởng của những ngôn ngữ khác, đáng kể nhất là tiếng Pháp và tiếng Anh.

Nhiều nhà ngôn ngữ học Việt Nam đã báo động rằng, tiếng Việt ngày nay đã bị lai cấu trúc tiếng Anh, tiếng Pháp khá nhiều. Phải thừa nhận là đúng, thế nhưng khó có thể xác định được mức độ trầm trọng của sự kiện này một cách chính xác. Cấu trúc tiếng Việt bị lai chủ yếu là qua hoạt động phiên dịch ngoại ngữ sang tiếng Việt, cũng như tác giả viết sách đã bị ngữ pháp ngoại ngữ ảnh hưởng nặng nề. Ví dụ tác giả là người sinh sống lâu năm ở nước ngoài, sử dụng thành thạo tiếng bản xứ nhưng lại ít có cơ hội sử dụng tiếng Việt trong đời sống thường ngày.

Không thể trách một người là đã đi sao chép cả cấu trúc ngữ pháp tiếng Anh, tiếng Pháp vào tiếng Việt khi người ấy không hiểu *cấu trúc ngữ pháp là gì*, không phân biệt được *cấu trúc ngữ pháp của hai ngôn ngữ*. Sự khác biệt ngữ pháp nói lên đặc điểm của từng ngôn ngữ. Làm sao con người có thể nhận biết được ngôn ngữ của mình có *đặc điểm* nào, *tính hoa* nào mà không phải ngôn ngữ nào cũng có ? Tự trung, nguyên do chính khiến cấu trúc tiếng Việt bị lai vẫn là do thiếu kiến thức tổng quát về ngôn ngữ học, một vấn đề giáo dục. Bởi không được trang bị những kiến thức ngôn ngữ cần thiết, con người thường phạm những sai lầm một cách vô tình trong sinh hoạt ngôn ngữ.

Chúng ta không thể lập bảng thống kê để cho thấy bao nhiêu phần trăm người Việt đã bị ngữ pháp ngoại ngữ ảnh hưởng và bao nhiêu phần trăm cấu trúc ngữ pháp tiếng Việt đã bị lai. Chỉ có một cách duy nhất để mỗi người tự kiểm chứng là *so sánh* cách hành ngôn tiếng Việt với cách hành ngôn tiếng nước ngoài. Đối với những điểm khác biệt, tất nhiên không có gì để nói, nhưng đối với những điểm tương đồng, người so sánh phải có khả năng trả lời câu hỏi, cái cấu trúc tương đồng ấy xưa nay người Việt vẫn dùng hay là mới được du nhập sau này. Để trả lời, người so sánh phải vận dụng thẩm năng và dụng năng ngôn ngữ của mình, đặc biệt là nên để ý đến những đặc điểm của tiếng Việt.

Sau đây, hãy thử so sánh tiếng Việt với tiếng Anh và tiếng Đức. Tiếng Anh là một ngôn ngữ tiêu biểu chứa tính chất tuyến tính và rất thông dụng, ngược lại, tiếng Đức không phải là một ngôn ngữ tuyến tính tiêu biểu. Để so sánh cấu trúc ngôn ngữ của hai trường hợp ấy, chúng ta sẽ dùng phương pháp dịch 1-1, nghĩa là dịch sát từng từ/cụm từ một, như có thể.

Như đã biết, có rất nhiều dân tộc trên thế giới sử dụng nguyên tắc đề diễn để diễn đạt điều muốn nói. Tuy nhiên, lối diễn đạt của họ chưa chắc đã hoàn toàn giống lối dùng “*thì*”, “*là*”, “*mà*” của người Việt. Chẳng hạn người Đức nói:

(41) Die Frau, *die* Sie zum Kaffee trinken eingeladen haben, ist meine Frau !

(42) Cái bà *mà* ông mời đi uống cà phê là bà xã tôi !

Dịch theo nội dung là vậy, song nếu dịch sát từng chữ trong câu tiếng Đức sang tiếng Việt, chúng ta sẽ được một câu rất buồn cười như sau:

(43) * Cái bà, cái ông để cà phê uống mời đã, là tôi bà xã.

Câu này ngàn đời người Việt không thể nào nói được, bắt chước được. Hoặc một câu tiếng Anh:

(44) I must go and look at the house before I take it.

Nếu dịch câu trên như sau, phải nói là người dịch đã sao chép y cấu trúc tiếng Anh vào tiếng Việt:

(45) Tôi phải đi và xem cái nhà trước khi tôi lấy nó.

(I) (must) (go) (and) (look at) (the house) (before) (I) (take) (it)

Câu tiếng Anh trên cũng mang cấu trúc tuyến tính, nhưng nếu dịch 1-1 sang tiếng Việt như thế thì kỳ quá. Nên nhớ rằng, nhiều cấu trúc câu xác định của tiếng Anh rất giống cấu trúc câu xác định của tiếng Việt, song như thế nào đó vẫn khác. Thay vì dịch như vậy, vẫn có cách dịch dễ nghe hơn:

(46) Tôi phải đi xem căn nhà ấy trước khi thuê (hoặc mua).

Hiện tượng chuyển dịch câu cú sát nghĩa thường thấy ở thể hệ trẻ Việt Nam sinh ra và lớn lên ở hải ngoại. Khi sử dụng tiếng Việt, các em thường có tâm lý lập câu tiếng nước ngoài trước rồi sau đó mới chuyển 1-1 sang tiếng Việt. Gặp ví dụ trên, nên cắt nghĩa cho các em hiểu, thay vì dùng động từ một cách rời rạc, người Việt có cách ghép 2 động từ lại với nhau mà không cần liên từ “và”. Thay vì nói “*đi và xem*” (*go and see*), người Việt nói “*đi xem*”. Cả về mặt *ngữ dụng* (*pragmatic*), cũng cần nên giải thích cho các em hiểu, động từ “*lấy*” của tiếng Việt không phải là một động từ *rất thường dùng* cho danh từ “*nhà*”. Thay vì vậy, có những cách dùng tiêu biểu như “*mua nhà*”, “*thuê nhà*”, ...

Có những trường hợp dịch hầu như 1-1 mà người dịch tưởng rằng người Việt nào nghe cũng lọt lỗ tai, điển hình là những *câu bị động* (*passive sentence*):

(47) My father was fined **by** the police.

Đây là cách lập câu bị động bình thường trong tiếng Anh. Nhiều người hay dịch sát chữ câu trên sang tiếng Việt như sau:

(48) Bố tôi bị phạt **bởi** công an.

Câu này bị lai cấu trúc tiếng Anh. Cứ thử kiểm lại bằng tính tịnh tiến, ta sẽ thấy không ai đặt một câu hỏi khó chấp nhận như dưới đây:

(49) Bố tôi bị phạt **bởi** ai ?

Trên thực tế, người Việt không nói “*Bố tôi bị phạt bởi công an*” mà là nói:

(50) Bố tôi bị công an phạt.

bởi cách trả lời này mang tính tịnh tiến chấp nhận được. Xem tiếp trường hợp:

(51) Bố anh bị gì ?

Từ đây có nhiều câu trả lời:

(52) Bố tôi bị *bệnh* (hoặc *thất nghiệp, phạt, lường gạt*, v.v.)

Giả sử bị ai đó làm cái gì, người Việt sẽ hỏi:

(53) Bố anh bị *ai làm cái gì* ?

Từ đây mới có câu trả lời tịnh tiến, ví dụ:

(54) Bố tôi bị công an phạt.

(55) Bố tôi bị hăng sa thải.

(56) Bố tôi bị người ta lường gạt.

Thành thử không ai nói:

(57) Bố tôi bị sa thải **bởi** hăng.

(58) Bố tôi bị lường gạt **bởi** người ta.

Tương tự vậy, người Việt không nói:

(59) Bả bị cắn **bởi** chó.

(60) Đứa con út được cưng chiều **bởi** bố mẹ.

(61) Cuộc thi hoa hậu được tổ chức **bởi** hăng nước hoa Chanel.

(62) Tội ác sẽ bị lên án **bởi** xã hội.

mà là nói:

(63) Bả bị chó cắn.

(64) Đứa con út được bố mẹ cưng chiều.

(65) Cuộc thi hoa hậu do hăng nước hoa Chanel tổ chức.

(66) Tội ác sẽ bị xã hội lên án.

Giả như “chó”, “bố mẹ”, “hăng nước hoa Chanel”, “xã hội” không phải là các yếu tố gây ra hành động (tác nhân¹⁾) đi chăng nữa, vẫn có trường hợp muốn diễn tả trạng thái bị động mà người Việt không thể bắt chước y hệt cách lập câu của người Anh:

¹⁾ Chủ ngữ logic (sẽ giải thích chi tiết sau)

- (67) A mountain was seen in the distance.
 (68) Một ngọn núi đã được trông thấy ở đằng xa.

Câu dịch sát ngữ pháp như vậy không được tự nhiên vì chúng ta đã sao chép nguyên cấu trúc tiếng Anh vào câu tiếng Việt. Cách hành ngôn đó là của người Anh. Còn cách hành ngôn của người Việt thông thường là:

- (69) Người ta thấy có một ngọn núi ở đằng xa.
 (70) Người ta thấy một ngọn núi ở đằng xa.
 (71) Có một ngọn núi ở đằng xa.

Những câu trên mang tính tĩnh tiến *chấp nhận được*. Ví dụ:

- (72) *Hỏi:* Người ta thấy **gì**?
 (73) *Trả lời:* Người ta thấy **một ngọn núi**.
 (74) *Hỏi tiếp:* ... ngọn núi **ở đâu**?
 (75) *Cuối cùng là câu trả lời:* Người ta thấy một ngọn núi **ở đằng xa**.

Ta thấy, mọi câu nghi vấn đều được trả lời hữu lý mà người Việt chấp nhận. Còn câu “*Một ngọn núi đã được trông thấy ...*” khó được chấp nhận theo kiểu hỏi và trả lời:

- (76) *Hỏi:* Một ngọn núi **thế nào**?
 (77) *Trả lời:* Một ngọn núi **đã được trông thấy**.

Dùng động từ “*trông thấy*” ở thái bị động cho chủ ngữ “*một ngọn núi*” không thích hợp. Người Việt không chấp nhận tính tĩnh tiến này. Đành rằng câu không sai ngữ pháp, nhưng không ổn về cách dùng từ. Đây là một vấn đề thuộc phạm trù ngữ nghĩa (sẽ nói chi tiết ở các chương sau). Lý thuyết thì có, song có lẽ cũng chẳng cần, bởi chắc chắn mỗi người dịch rành tiếng Việt cũng thừa biết động từ nào dùng được cho danh từ nào.

Cả những cấu trúc đề diễn chứa mệnh đề chính và mệnh đề phụ trong tiếng Anh tưởng chừng rất dễ dịch sang tiếng Việt, thế nhưng không hẳn thế:

- (78) It was I who knocked at your door.

Nguyên câu trên là *mệnh đề chính (main clause)*. Phần “*who knocked at your door*” là *mệnh đề phụ (subordinate clause)*. Đây là hình thức *câu chẻ (cleft*

sentence) thường thấy trong tiếng Anh. Nếu sao chép cả cấu trúc này vào tiếng Việt, chúng ta sẽ được một câu nghe khá lạ tai:

(79) Đó là tôi người đã gõ cửa phòng anh.

Người Việt không nói thế. Cao lăm họ nói:

(80) Chính tôi đã gõ cửa phòng anh.

Hình thức câu chẻ trên hầu như không thấy trong tiếng Việt. Một trường hợp khác là hình thức dùng “*to*”:

(81) You forget *to* answer his letter.

(82) * Anh quên *để* trả lời bức thư của hắn.

Người Việt không nói vậy mà là nói:

(83) Anh quên trả lời thư hắn.

(84) Anh quên trả lời bức thư của hắn.

Hoặc:

(85) I’ll come *to* see you tomorrow.

(86) ? Tôi sẽ tới *để* thăm anh ngày mai.

Thay vì vậy, vẫn có cách nói thuần túy Việt Nam:

(87) Mai tôi sẽ ghé thăm anh.

(88) Ngày mai tôi sẽ ghé thăm anh.

Hoặc:

(89) She loves *to* read novels.

(90) * Cô ta yêu thích *để* đọc tiểu thuyết.

Thay vì vậy, người Việt nói:

(91) Cô ta thích đọc tiểu thuyết.

Nói chung, thay vì dùng “to” giữa hai động từ như người Anh, người Việt thường ghép 2 động từ lại một: “*quên trả lời*”, “*ghé thăm*”, “*thích đọc*”, ... Tất nhiên trong tiếng Việt cũng có cách dùng “để” mà thường bị hiểu nhầm là hoàn toàn đồng nghĩa với “to” của tiếng Anh hoặc “zu” của tiếng Đức. Trong tiếng Việt, “để” nằm giữa hai động từ cốt để diễn tả *một hành động nhằm mục đích* gì, hoặc để diễn tả ý nghĩa *nhằm mục đích* gì. Hiểu nôm na là “*để chi vậy ?*”:

- (92) Ăn để mà sống chứ không phải sống để mà ăn.
- (93) Sở dĩ tôi phải giải thích dông dài là để anh đừng hỏi nữa.

Tiếng Anh có nhiều cấu trúc câu khá giống tiếng Việt. Không những chỉ câu xác định mà còn cả *câu hỏi có/không* (yes/no question). Một trường hợp điển hình là câu hỏi “*phải không*”. Có nhiều ý kiến cho rằng, đây không phải là cấu trúc thuần tiếng Việt.

- (94) They are Vietnamese, aren't they ?
- (95) Họ là người Việt phải không ?

Người Đức cũng có cách hỏi tương tự:

- (96) Sie sind Vietnamesen, nicht wahr ?
(Họ) (là) (người Việt), (không) (phải) ?

Có lẽ người Việt, người Đức không có nhiều cách hỏi như người Anh. Người Anh còn có những cách hỏi khác:

- (97) He can swim, can't he ?
- (98) He will play football, won't he ?
- (99) He speaks English, doesn't he ?

mà người Việt, người Đức không thể dịch đúng cấu trúc. Ví dụ một câu:

- (100) He can swim, can't he ?
- (101) *Anh ấy có thể bơi, có thể không ?
- (102) *Er kann schwimmen, kann er nicht ?

Khó có thể quả quyết rằng, “*phải không*” là cấu trúc lai tiếng Anh hoặc tiếng Đức. Người ta chỉ biết xưa nay người Việt vẫn dùng lối hỏi này, đại để nó tương đương với “*à*”, “*chứ gì*”, ... nhưng nghiêm chỉnh hơn:

(103) Họ là người Việt à ?

(104) Họ là người Việt chứ gì ?

Có thể đây chỉ là sự trùng hợp ngẫu nhiên. “*Phải không*” có nghĩa như “*đúng không*”, “*đúng chứ gì*”, một ngữ đoạn phụ có chức năng đặt nghi vấn về câu xác định “*Họ là người Việt*” phía trước.

Tiếng Việt, tiếng Anh, tiếng Pháp là những ngôn ngữ tuyến tính. Trái hẳn lại, tiếng Đức là một ngôn ngữ phi tuyến tính. Tiêu biểu là động từ trong câu tiếng Đức thường nằm ở cuối câu, nếu câu có hai động từ trở lên, hoặc bị đảo ngược lên trước chủ ngữ (chủ từ), nếu phần bổ ngữ (túc từ) nằm ở đầu câu.

Động từ nằm cuối câu:

(105) Er möchte mit ihr sprechen.

(Hắn) (muốn) (với) (cô ta) (nói chuyện)

Động từ bị đảo ngược:

(106) Mit ihr möchte er sprechen.

(Với) (cô ta) (muốn) (hắn) (nói chuyện)

Nghe một ngôn ngữ tuyến tính, nhiều khi người nghe có thể đoán trước người nói muốn nói gì dù câu chưa chấm dứt. Còn khi nghe tiếng Đức, người nghe phải kiên nhẫn đợi người nói nói hết câu vì động từ nằm mãi tí đằng sau. Ví dụ “*Er möchte mit ihr ...*” (*hắn muốn ... với cô ta*) nhưng hắn muốn gì ? Muốn *nói chuyện với cô ta* (*mit ihr sprechen*), hay muốn *đi xem cinê với cô ta* (*mit ihr ins Kino gehen*) ? Còn biết bao động từ khác nữa: *sống, chết, thức, ngủ*, v.v.. Nên cẩn thận bởi không phải bao giờ người nghe cũng có thể hiểu ngầm hoặc đoán trước được ý người Đức khi họ chưa nói hết. Vả lại câu cú tiếng Đức thường dài, người nghe lại càng nên “*lịch sự*” kiên nhẫn lắng nghe. Ngay cả trường hợp động từ phụ bị đảo ngược, người nghe cũng phải đợi câu chấm dứt vì động từ cũng có thể nằm ở cuối câu.

Một ví dụ không tuyến tính nữa trong tiếng Đức là cách đọc số. Ví dụ 64, người Việt đọc theo thứ tự tuyến tính:

(107) Sáu mươi bốn

Người Đức đọc là:

(108) Vier und sechzig
(bốn) (và) (sáu mươi)

Có lẽ người Đức nhìn thấy con số 4 trước khi thấy con số 60 ? Trong khi đó, người Anh đọc theo thứ tự như người Việt.

(109) Sixty four
(sáu mươi) (bốn)

Nhận xét: Chỉ cần dịch 1-1 tiếng Đức sang tiếng Việt cũng đủ thấy câu cú hai ngôn ngữ khác nhau một trời một vực. Kết quả dịch thường đưa đến một câu rất lạ mà người Việt không bao giờ nói thế. Điều này cũng đúng luôn cho những người Việt đã sinh sống lâu năm ở Đức, nhuần nhuyễn tiếng Đức như người Đức nhưng trong thực tế, khi nói chuyện với nhau bằng tiếng Việt hoặc viết bằng tiếng Việt, cách hành ngôn của họ vẫn không thể lai ngữ pháp tiếng Đức một cách dễ dàng. Điều này dễ hiểu, bởi nguyên tắc hành ngôn của người Việt là cố vẽ lại trình tự nhận diện và nhận thức càng trung thực càng tốt.

Bằng cách so sánh bên trên, chúng ta mới hiểu tại sao người Việt dễ tiếp nhận những cấu trúc của các ngôn ngữ tuyến tính như tiếng Anh, tiếng Pháp. Còn phi tuyến tính như tiếng Đức, rất khó. Tuy vậy, tất cả đều tương đối.

Nói chung, đối với ngoại ngữ nào cũng vậy, muốn xác định sự khác biệt giữa cấu trúc tiếng Việt và cấu trúc của một ngoại ngữ nào đó, *người nói/nghe lý tưởng* (s.s. chương 1) chỉ cần dịch ra từng chữ một theo cách dịch 1-1, rồi sau đó dùng thẩm năng/dụng năng ngôn ngữ của mình, xem thử chuỗi từ vừa dịch, người Việt (*người nói/nghe lý tưởng*) có ai *nói* vậy không. Xin lưu ý là *nói* chứ không phải *viết*. Một câu văn nói mang tính đề diễn thường đúng ngữ pháp hơn văn viết. Tuy vậy, không có nghĩa là văn nói bao giờ cũng chuẩn.

VĂN NÓI VÀ VĂN VIẾT

Viết và nói là hai hình thức sinh hoạt ngôn ngữ. Cả hai đều đòi hỏi ngữ pháp, song chưa chắc cách hành ngôn đã giống nhau. Thử xem vài khác biệt (đặt điều kiện, người sử dụng ngôn ngữ là người nói/nghe lý tưởng).

Khi nói, nhất là khi đối thoại, người nói không có nhiều thì giờ suy nghĩ. Với thẩm năng và dụng năng ngôn ngữ, người nói phải lập câu cho đúng và nói ngay. Thêm vào đó, người nói còn phải chuẩn bị trình bày những ý tưởng tiếp theo. Tiến trình nói diễn ra với một vận tốc cao đến nỗi những quy tắc ngữ pháp vẫn ở dạng vô thức, trung thực nhất mà người nói không có thì giờ biến hóa. Trong khi viết, người viết có nhiều thời gian hơn. Có thời gian, người viết thường có tâm lý chọn lựa từ ngữ, trau chuốt câu cú. Hành động ngôn ngữ cố ý này chủ yếu là muốn cho câu viết thêm văn chương, mạch lạc. Thế nhưng nếu không cẩn thận, hành động ấy rất dễ dẫn đến hậu quả là ngữ pháp bị biến hóa, bị phức tạp hóa và không còn đúng như những quy tắc ngữ pháp nguyên thủy đã lóe ra từ vô thức ngay ban đầu. Một câu văn viết có thể dài, được soạn kỹ, được văn chương hóa, ngược lại, một câu văn nói thường ngắn hơn, trung thực hơn, bình dân hơn hoặc dễ hiểu hơn. Tuy nhiên, trong khi nói cũng có hiện tượng vụng về do thiếu thẩm năng và dụng năng ngôn ngữ, hoặc do sống động, ví dụ hiện tượng lặp lại một đoạn câu hoặc từ nhiều lần:

(110) Chào bác, chào bác ạ. Mời bác, mời bác vào nhà chơi, vào nhà chơi.

(111) Ê ! Ê ! Làm cái gì đó ? Làm cái gì đó ?

Kể cả văn nói tiếng Việt trong văn bản nhiều khi cũng không còn trung thực và tự nhiên như câu nói thực ở ngoài đời. Dĩ nhiên những hiện tượng này cũng thường xảy ra ở những ngôn ngữ khác.

Một hiện tượng khác cũng thường xảy ra là trong khi nói, ngữ pháp ít bị lược bỏ, trừ những trường hợp nói cộc lốc. Câu nói của tiếng Việt được phân giới khá rõ bằng “thì”, “là”, “mà”. Những từ này đóng vai ngữ pháp quan trọng như “to be”, “have” của tiếng Anh hoặc “sein”, “haben” của tiếng Đức. Hầu như không có người Việt nào lược bỏ các từ “thì”, “là”, “mà” trong khi nói như khi viết. So với người Anh hoặc người Đức, có lẽ hành động ngôn ngữ của người Việt khác họ ở chỗ này. Trong văn viết, thay vì dùng “thì”, “là”, “mà”, người Việt thường dùng dấu phẩy để phân đoạn.

Có một cách tránh sự xa rời đặc tính ngữ pháp tiếng Việt trong văn viết là nên chép ra điều mình muốn diễn tả bằng văn nói, rồi sau đó dùng các phương pháp như *phép tỉnh lược*, *phép thế*, *phép đặt câu hỏi*, *phép bố trí*, *phép hoán vị* mà chọn lựa, thêm bớt từ ngữ, trau chuốt câu cú cho hoàn chỉnh như ý muốn. Các phương pháp này sẽ được trình bày chi tiết ở chương sau.

Chương 4

VẬT LIỆU VÀ CẤU TRÚC NGÔN NGỮ

VẬT LIỆU NGÔN NGỮ

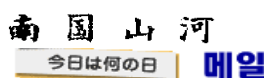
Cho dù mỗi nhà ngôn ngữ học có cách nhìn và định nghĩa khác nhau về ngôn ngữ, song nếu rút ra những ý chung, thì *ngôn ngữ tự nhiên (natural language)*, tức *ngôn ngữ con người (human language)*, bao giờ cũng là một tập hợp gồm tất cả các *câu, chuỗi từ, cụm từ, từ, ký tự, âm* có giới hạn, có ý nghĩa hoặc sử dụng được để tạo ra ý nghĩa.

Không bàn về ngôn ngữ đặc biệt như ngôn ngữ toán, ngôn ngữ lập trình, ngôn ngữ của người câm, ... thì mỗi ngôn ngữ tự nhiên bao giờ cũng bao hàm hai 2 yếu tố: *âm thanh* và *chữ viết*.

Trong ngôn ngữ tự nhiên, những đơn vị nhỏ nhất được dùng để biểu thị âm thanh và tạo thành chữ viết - có thể nói – là *ký tự (character)*. Ký tự là tập hợp của mọi *mẫu tự (alphabet)*, *hình tự (ideo-graph)* và *dấu (sign)*.

(1) Mẫu tự: La Tinh (a, b, c, ...), Hy Lạp (α , β , χ), ...

(2) Hình tự tiếng Hoa, Nhật, Đại Hàn, ...:



(3) Dấu: ‘, ` , ˆ , ~ , . , ^ , ? , ! , + , - , * , / , \$, ...

Ký tự được dùng để thành lập *cụm ký tự (group of characters)*. Đối với trường hợp ngôn ngữ dùng hình tự như tiếng Hoa, Nhật, Đại Hàn, ... tự bản thân mỗi ký tự có thể đã là một đơn vị ngôn ngữ có ý nghĩa. Hình tự chẳng qua là một hình tượng được vẽ bằng nhiều nét ngang, dọc, chấm, phẩy, ...

(4) 南 (Nam), 国 (Quốc)

Một ký tự được xem như thuộc về một ngôn ngữ, nếu nó có khả năng tạo thành chữ viết của ngôn ngữ ấy. Có nhiều ngôn ngữ dùng ký tự giống nhau.

- (5) @ \$ ^ % & ~ # ß *: một cụm ký tự vô nghĩa.
 Chào: một cụm ký tự tiếng Việt.
 Grüßen: một cụm ký tự tiếng Đức.
 House: một cụm ký tự tiếng Anh.
 Español: một cụm ký tự Tây Ban Nha.
 Latinum: một cụm ký tự La Tinh.

Một cụm ký tự có thể có ý nghĩa, có thể không. Có trường hợp, một cụm ký tự không thể đứng một mình mà phải kết hợp với một cụm ký tự khác để thành lập một đơn vị ngôn ngữ có ý nghĩa. Tiêu biểu là *tiền tố* và *hậu tố* của những ngôn ngữ đa âm (Anh, Đức, Pháp, Ý, ...):

Tiền tố (prefix)¹⁾ là cụm ký tự có ý nghĩa nhưng không thể đứng độc lập mà phải được gắn phía trước một cụm ký tự khác để thành một đơn vị ngôn ngữ có ý nghĩa.

Hậu tố (postfix)²⁾ có thể có ý nghĩa nhưng không thể đứng độc lập mà phải được gắn phía sau một cụm ký tự khác để thành một đơn vị ngôn ngữ có ý nghĩa.

Một hay nhiều cụm ký tự có thể sinh ra một ***từ (word)***. Trên bình diện cú pháp, giới ngôn ngữ học Tây phương xem *từ* như một đơn vị ngôn ngữ nhỏ nhất, có âm, ý nghĩa và loại *từ vựng (lexical)* nhất định (danh từ, động từ, tính từ, liên từ, giới từ, ...). Người ta gọi đơn vị ngôn ngữ nhỏ nhất này là ***đơn vị cú pháp (syntax unit)*** nhỏ nhất.

- (6) a. “Ăn” là một cụm ký tự có nghĩa. Vậy, “ăn” là một từ. Nói rõ hơn, “ăn” là một đơn vị cú pháp nhỏ nhất, có ý nghĩa, là một động từ, đứng độc lập dưới dạng một ***từ đơn (single word)***.
 b. “Năn” là một cụm ký tự không có nghĩa, không thể coi là một từ.
 c. “Ăn” ghép với “năn” thành “ăn năn”, một từ có nghĩa nhưng lại không dính dáng gì tới “ăn”. “Ăn năn” là một đơn vị cú pháp nhỏ nhất có ý nghĩa, và là một động từ.

¹⁾ còn được gọi là *tiếp đầu ngữ*.

²⁾ còn được gọi là *tiếp vĩ ngữ (suffix)*.

- d. “Internet” (*liên mạng*) là một từ được ghép từ “inter” (*liên*) + “net” (*mạng*). “Net” là một từ có thể đứng một mình như một danh từ. “Inter” là một *tiền tố* không thể đứng một mình
- e. “Schönheit” (*cái đẹp*) là một từ được ghép từ “schön” (*đẹp*) + “heit” (*cái*) trong tiếng Đức, là một danh từ. “Schön” là một tính từ có thể đứng một mình. “Heit” là một *hậu tố* để thành lập danh từ.
- f. Từ còn có thể chứa nhiều hơn 2 cụm ký tự: “thảo cầm viên”, “câu lạc bộ”, “hàng không mẫu hạm”, “trung cầu dân ý”, ...

Đối với tiếng Việt, một cụm ký tự như vậy được coi như một chữ/tiếng. Viết thì gọi là chữ, nói thì gọi là tiếng. Tổng quát, từ là kết quả của nhiều chữ hợp lại. Nếu từng chữ là từ, thì kết quả ấy là một **từ ghép (compound word)**.

Từ ghép rất phổ biến trong ngôn ngữ độc âm, song không phải là không có trong những tiếng đa âm, chẳng hạn như tiếng Đức, một ngôn ngữ nổi tiếng về nghệ thuật ghép từ (Wortbildung).

Đối với tiếng Việt, dù ngày nay chúng ta không dùng hình tự như chữ Hán, chữ Nôm để thành lập một đơn vị ngôn ngữ có ý nghĩa như xưa, thế nhưng tiếng Việt vẫn khác các ngôn ngữ sử dụng *hình vị (morpheme)*. Hình vị được hiểu là cụm ký tự tự tạo có hình thức nhất định, có ý nghĩa và cách phát âm riêng biệt. Trong tiếng Việt, cấu tạo của mỗi đơn vị ngôn ngữ nhỏ nhất bao gồm một số âm nhất định. Đơn vị này cố định. Nói cách khác, một từ tiếng Việt bao giờ cũng ở thể nguyên mẫu. Trong khi đó, như một cách *hình thức hóa từ ngữ*, một từ nguyên mẫu trong ngôn ngữ Âu châu có thể được chia ra thành nhiều hình vị khác nhau nhằm mô tả ngôi thứ, số lượng, thời gian, tính sở hữu, phái tính, ... Cụ thể, *động từ nguyên mẫu (infinitive verb)* được chia ra thành nhiều hình vị nhằm diễn tả nhân vật số nhiều/số ít, ngôi thứ, hoặc thời gian (hiện tại, quá khứ).

- (7) a. “Go” (*đi*) là một động từ nguyên mẫu của tiếng Anh.
 b. “Go” (*đi*) diễn tả ngôi 1/2 của số ít, 1/2/3 của số nhiều, hiện tại.
 c. “Goes” (*đi*) diễn tả ngôi ba, số ít, hiện tại.
- (8) a. “Gehen” (*đi*) là một động từ nguyên mẫu của tiếng Đức.
 d. “Gehe” (*đi*) diễn tả ngôi một, số ít, hiện tại.
 e. “Gehst” (*đi*) diễn tả ngôi hai, số ít, hiện tại.
 f. “Ging” (*đi*) diễn tả ngôi một, số ít, quá khứ.
 g. “Gingst” (*đi*) diễn tả ngôi hai, số ít, quá khứ.
 h. V.v.

Tiếng Việt không có hình thức chia động từ. Động từ “đi” bao giờ cũng ở thể nguyên mẫu “đi”.

Khoan xét đến tính “nói được hay không” trong ngôn ngữ, thì nhiều từ/cụm ký tự có khả năng sinh ra một **cụm từ**; nhiều cụm từ có khả năng sinh ra một **chuỗi từ**. Xin lưu ý, chúng ta sẽ dùng hai khái niệm này khá thường xuyên. *Cụm từ* được hiểu như vài ba từ/cụm ký tự ghép lại. *Chuỗi từ* lớn hơn, được hiểu như đoạn câu hoặc cả câu (không xét về mặt cú pháp). Chỉ cần hiểu tương đối vậy thôi.

- (9) a. “Thằng Bờm có cái quạt mo” là một chuỗi từ.
 b. “thằng Bờm”, “cái quạt mo”, “Bờm có cái”, ... là những cụm từ.
 c. “thằng”, “Bờm”, “có”, ... là những từ.
 d. “Thằng Bờm có cái @&~#*” là một chuỗi từ nhưng “@&~#*” là một cụm ký tự không có trong tiếng Việt (những trường hợp không thực tế này không bàn tới).

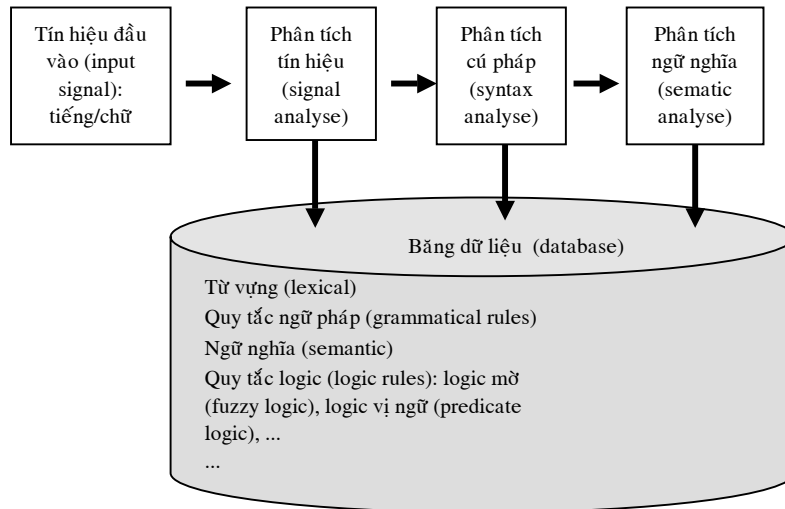
Nếu ứng dụng tin học, đặc biệt là những phương pháp *trí thông minh nhân tạo* (*artificial intelligence*), để thiết lập hệ nhận diện ngôn ngữ (*language recognizing system*), chắc hẳn tiếng Việt sẽ có nhiều lợi điểm so với những ngôn ngữ thiên về hình vị. Dễ thấy nhất là ngay công đoạn *phân tích tín hiệu* (*signal analyse*) và *phân tích cú pháp* (*syntax analyse*). Do không cần phải chia động từ, máy tính có thể tiết kiệm được rất nhiều chỗ trong *bảng từ vựng* (*lexical bank*). Tốc độ xử lý dữ liệu cũng tối ưu hơn vì đỡ mất thì giờ tìm kiếm hoặc phải xuyên qua nhiều bước, nhiều thao tác. Giải thuật lập trình nhờ vậy mà đỡ phức tạp hơn.

Định nghĩa từ là một vấn đề thường gây tranh cãi. Một trong những nguyên nhân chính là do cách cấu tạo từ của từng ngôn ngữ khác nhau. Khó có thể định nghĩa hiển ngôn và tổng quát “từ là gì” cho mọi ngôn ngữ trên trần thế, song rõ ràng, cách cấu tạo từ tiếng Việt không thiên về hình vị như ngôn ngữ Âu châu. Từ là sản phẩm của một cụm từ hoặc nhiều cụm ký tự. Chúng ta tạm gọi mỗi cụm ký tự giới hạn này là một “*chữ*” hay một “*tiếng*”. Tất nhiên còn nhiều cách gọi khác. Ví dụ, theo Phan Khôi (1997), mỗi cụm ký tự là một “*tự*”.

Việc xem “từ là sản phẩm của chữ hay tiếng” xuất phát từ sự quan sát cách làm việc thực tế của một hệ nhận diện ngôn ngữ có khả năng hiểu tiếng Việt. Điều này có thể giải thích được. Thử nhìn sơ nguyên tắc của hệ thống này.

Một hệ nhận diện ngôn ngữ thô sơ nhất gồm có 4 thành phần:

- ***Phân tích tín hiệu (signal analyse)***
- ***Phân tích cú pháp (syntax analyse)***
- ***Phân tích ngữ nghĩa (semantic analyse)***
- ***Bảng dữ liệu (database)***



Trong một hệ nhận diện ngôn ngữ, ***phân tích tín hiệu*** là công việc đầu tiên mà máy phải làm khi nhận được một chuỗi từ được gửi vào bất kể dưới dạng nói hay dạng viết. Đây là những *tín hiệu đầu vào (input signal)*. Ví dụ gửi chuỗi từ “Tôi đi” vào máy. Tín hiệu đầu tiên máy sẽ nhận được là “tôi”, kế tiếp là “đi”. Thoạt tiên, máy sẽ kiểm tra xem tiếng/chữ “tôi” có nằm trong bảng từ vựng (thuộc bảng dữ liệu) hay không. Nếu không, nó sẽ không hiểu và sẽ hỏi lại “***tôi có nghĩa là gì ?***”. Quá trình này tự nhiên giống như một đứa bé học một chữ mới, vì chưa hiểu, nó sẽ hỏi lại bố mẹ. Hỏi là để học. Đối với một ứng dụng trí thông minh nhân tạo thì đây là tiến trình thu thập kiến thức.

Trong quá trình học, có trường hợp máy không hiểu vì người nói phát âm sai, hoặc không rõ ràng, hoặc máy nghe chưa quen, hoặc máy bị ... hỏng. Loại bỏ trường hợp máy hỏng, chúng ta thấy, “nghe không rõ” là một chuyện rất bình thường trong khi đàm thoại. Chẳng hạn đối với người Việt, giọng người Trung nói, người Bắc nghe không quen, giọng người Nam nói, người Trung nghe không hiểu. Ban đầu là vậy, phải nghe quen dần, tức là học, là thu thập kiến thức, rồi mới hiểu. Máy cũng thế, cũng phải học như người.

Nói một cách lý thuyết, người hoặc máy không hiểu một dữ liệu là do dữ liệu ấy chưa bao giờ tồn tại trong bộ não của người hoặc trong băng dữ liệu của máy. Muốn biết một dữ liệu có tồn tại hay không, thì phải coi lại, nghĩa là *tìm* (*search*). Tìm nghĩa là đi so sánh cái *dữ liệu muốn tìm* (*sample*) với tất cả những *dữ liệu mẫu* (*pattern*) đã tồn tại trong bộ não, trong băng dữ liệu xem có trùng hoặc xứng hay không, có *trùng hợp* (*matching*) hay không.

Ở trong máy, mọi dữ liệu mẫu đều được lưu trữ dưới dạng *mã nhị phân* (*binary code*), cho nên một dữ liệu mới được đưa vào máy cũng phải được mã hóa nhị phân, thì máy mới hiểu, mới làm việc được. Đơn vị tiêu chuẩn của mã được tính bằng Byte. Dữ liệu càng nhiều, số Byte càng nhiều.

Tìm mẫu-trùng hợp (*pattern-matching*) là đối chiếu *chuỗi Byte dữ liệu được đưa vào* với *chuỗi Byte dữ liệu mẫu*. Hai Bytes giống nhau nghĩa là *trùng nhau* (*match*). Còn không, hai dữ liệu không hoàn toàn giống nhau.

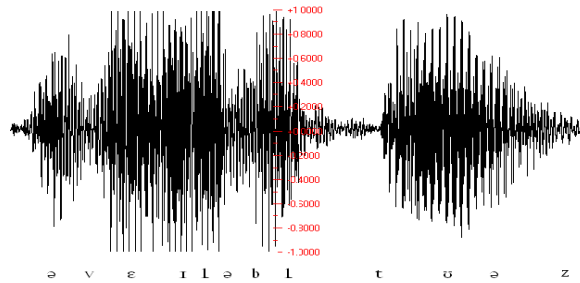
Nhận diện *chữ* đối với máy không khó và nhanh. Chỉ cần một dòng lệnh cũng đủ biết chữ ấy có nằm trong băng dữ liệu hay không, máy có biết chữ này không, ví dụ dùng lệnh SELECT của *SQL* (*Structured Query Language*)³⁾:

```
SELECT * FROM Lexical WHERE Word = "tôi"
```

(Nghĩa là giả sử đã tồn tại một *bảng* (*list*) tên là *Lexical* chứa mọi dữ liệu giống một cuốn tự điển khổng lồ. Hỏi: có chữ "tôi" trong bảng ấy hay không? Nếu có, nó sẽ liệt kê mọi tính chất của chữ "tôi", nếu không, nó sẽ báo là không có phần tử ấy; mọi tính chất của chữ "tôi" đều trống (NULL)).

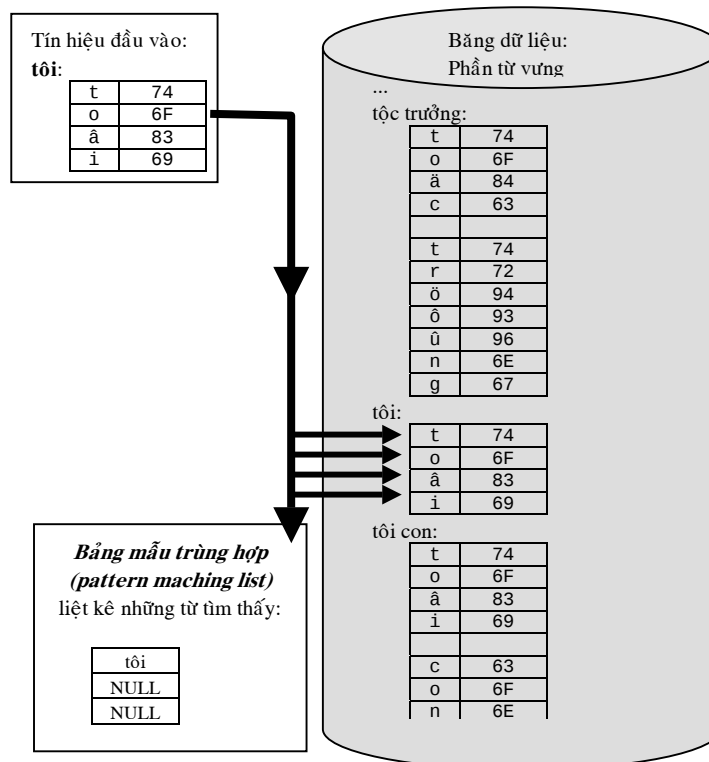
Nhận diện *tiếng* (*sound*) mới khó bởi vì đây là âm thanh. Mỗi âm thanh đều nằm dưới dạng sóng sine/cosine có biên độ, tần số, ... khác nhau. Hàm lượng giác sine/cosine đối với máy tính là một loại hàm phức tạp, đòi hỏi rất nhiều thời gian tính. Tất cả mọi yếu tố biên độ, tần số, ... phải được mã hóa. Khi so sánh, bắt buộc phải so sánh tất cả mọi yếu tố, rất tốn thời gian. Hơn nữa, từng tiếng nói được gửi vào máy không phải bao giờ cũng cố định như một hằng số mà có lúc nhẹ, lúc mạnh, lúc nhanh, lúc chậm, lúc trong, lúc đục, lúc trầm, lúc bổng, Những yếu tố này làm thay đổi biên độ sóng, tần số sóng, ... Vì thế, kết quả so sánh sóng chỉ có thể gần đúng. Bởi vì chỉ gần đúng, cho nên độ chính xác nhận diện âm thanh đương nhiên không thể cao bằng nhận diện chữ.

³⁾ Trên thực tế, PROLOG là ngôn ngữ lập trình thường dùng cho hệ nhận diện ngôn ngữ.



Sóng của các cụm ký tự “available”, “to”, “us”. Giọng của Noam Chomsky. (Tài liệu: Wagner 1999). Mỗi ký tự có một sóng khác nhau.

Khác với âm thanh, một chữ viết cũng là một cụm ký tự nhưng được mã hóa trực tiếp bằng đơn vị Byte thay vì hàm số. Mỗi ký tự (character) là một Byte (8 Bits), chẳng hạn theo mã ASCII hoặc 2 Bytes theo mã UNICODE. So sánh một chữ viết với một chữ viết nghĩa là so sánh từng Byte của chữ này với từng Byte của chữ kia cho đến hết chuỗi Byte (một trong hai). Một ví dụ, hãy xem chữ “tôi” có nằm trong băng dữ liệu từ vựng hay không ?



Chữ “*tôi*” ở đầu vào được so sánh với những từ gần giống trong bảng dữ liệu. Từ trong bảng dữ liệu đã được xếp đặt theo thứ tự alphabet. Ta thấy, thao tác so sánh chạy theo mũi tên từ trên xuống dưới. Xuyên qua nhiều chữ không giống, máy tiếp tục so sánh “*tôi*” với “*tộc trưởng*”. Cũng không giống. Đi tiếp sang từ kế, cuối cùng gặp “*tôi*”. Giống. Tìm thấy. Ngừng.

Tương tự cách tìm chữ viết, tiếng cũng có thể được truy tìm bằng nguyên tắc so sánh mẫu như trên, chỉ có điều là phức tạp hơn nhiều.

Bất kỳ một chữ hay một tiếng sau khi được tìm thấy, dù chính xác hay gần đúng, cũng sẽ được ghi vào *bảng mẫu trùng hợp* (*pattern matching list*). Bảng này cần cho công đoạn phân tích cú pháp kế tiếp.

Chúng ta có thể nhận thấy rằng, tìm kiếm một từ tiếng Việt – nhìn chung – chắc chắn phải nhanh hơn tìm kiếm một từ tiếng Anh hoặc tiếng Đức. Thử tưởng tượng, truy tìm động từ “*gehen*” của tiếng Đức sẽ lâu thế nào, vì động từ được chia cho đủ mọi trường hợp (*gehen* → *gehe, gehst, geht, gingen, gingst, ging, gegangen,...*). Ngay ví dụ này, bảng dữ liệu tiếng Đức cần nhiều chỗ để chứa tất cả mọi hình vị của từ nguyên mẫu “*gehen*”, trong khi đó tiếng Việt chỉ chiếm một chỗ độc nhất là: “*đi*”.

Trên thực tế, khâu phân tích tín hiệu không chỉ đón nhận từng tín hiệu một mà đợi một chốc để nắm bắt luôn những tín hiệu theo sau. Cách thức *nhìn trước* (*look ahead*) nhằm mục đích thu thêm nhiều dữ liệu có liên quan đến dữ liệu máy đang để ý (*focussing*). Giả sử máy vừa nghe ai nói tiếng “*Nó*” và đã hiểu, kế tiếp là nhận tiếng “*ăn*”, máy sẽ truy tìm chữ này trong bảng từ vựng và song song đó, nó sẽ chờ xem còn tiếng nào theo sau nữa hay không. Nếu không, coi như câu chấm dứt. Nếu có, nó sẽ phối hợp “*ăn*” với tín hiệu sau, xem hai tín hiệu này có lệ thuộc nhau không. Giả sử theo sau “*ăn*” là “*năn*”, máy sẽ kiểm tra xem “*ăn năn*” có phải là một *từ* hay không. Nếu đúng, nó sẽ quên chữ “*ăn*” và chấp nhận tín hiệu đầu vào là từ “*ăn năn*”. Còn không, ví dụ, nghe thêm được “*phở*”, nó sẽ tìm trong bảng dữ liệu xem có từ “*ăn phở*” hay không. Nếu không, nó sẽ xem “*ăn*” là một động từ và sẽ tiếp tục đi tìm từ “*phở*”. Giả như tìm thấy “*phở*” là một danh từ, nó sẽ ghi chú vào bảng. Hai từ này sẽ được khâu phân tích cú pháp công nhận là một động từ ghép qua quy tắc ngữ pháp VP → V N (sẽ nói sau). Dưới đây là cách ghi nhớ dữ liệu của hai bảng mẫu trùng hợp:

(10) Nó ăn năn.

Nó ăn phở.

từ *chuẩn hình* (*well-formed*) đủ nói được sẽ được coi như một **thành tố** (**constituent**) (tạm hiểu vậy, chúng ta sẽ định nghĩa một cách lý thuyết hơn).

- (11) a. “*thằng Bờm*”, “*cái quạt mo*” là hai cụm từ đủ nói được và được xem như hai thành tố.
 b. “*Bờm*”, “*có*” cũng là những thành tố nhưng chỉ có một từ duy nhất.
 c. “*Bờm có cái*” chưa đủ nói được, vì thiếu “*cái gì*”, cho nên chưa phải là một thành tố.
 d. “*Bờm thằng cái mo quạt*”, “*cái thằng quạt*” là những cụm từ lộn xộn, không thể coi là thành tố được.

Như vậy **thành tố** còn có thể được hiểu như một *đơn vị cú pháp có quy tắc ngữ pháp* (*grammatical rule*), nói được. Những gì nói được, trước nay chúng ta thường gọi là đúng *văn phạm*. Thay vì dùng khái niệm “*văn phạm*”, kể từ đây, ta sẽ gọi là *ngữ pháp* (*grammar*). Xin mượn lời giải thích của Nguyễn Kim Thản:

“Trước đây ta thường gọi ngữ pháp là “mẹo” hoặc “văn phạm”. Gọi là “mẹo” không nói lên bản chất của ngữ pháp. Gọi là “văn phạm” cũng không ổn, bởi vì qui tắc ngữ pháp không phải chỉ áp dụng vào việc viết văn, và yếu tố thứ nhất của văn không phải là cái gì khác ngoài ngôn ngữ” (Nguyễn Kim Thản 1997 : 10)

Đã thông qua các khái niệm cần thiết, nay đặt câu hỏi, thế còn **câu** (**sentence**) là gì ?

Lewandowski định nghĩa:

“Câu là một đơn vị có ngữ pháp, cách phát âm và nội dung theo quy tắc của từng ngôn ngữ được tổ chức một cách tuyến tính và có tôn ti như một phương tiện dùng để diễn đạt, mô tả và phát ngôn, (và) để truyền thông những hình dung hoặc suy nghĩ về tình huống”.

“([Der Satz ist eine] grammatisch, intonatorisch und inhaltlich nach den Regularitäten der jeweiligen Sprache linear und hierarchisch organisierte Einheit als Mittel zu Ausdruck, Darstellung und Appell, zur Kommunikation von Vorstellungen oder Gedanken über Sachverhalte)”. (Lewandowski 1990, x. *Satz*).

Không thiếu sách vở nói về câu: *Was ist ein Satz* (Ries 1931), *Geschichte und Kritik der wichtigsten Satzdefinitionen* (Seidel 1935), v.v.. Liệt kê cho đủ, còn khoảng 200 định nghĩa khác nhau (Wagner 1999).

Định nghĩa khái niệm *câu* cho thật tổng quát, hiển ngôn và hoàn hảo không phải là chuyện dễ dàng. Cũng một chuỗi từ có cùng cấu trúc, có ngôn ngữ cho là câu, có ngôn ngữ lại không chấp nhận. Thay vì vậy, các nhà ngôn ngữ học đã tìm cách dựng nên một lý thuyết nhằm xác định cấu trúc của một phát biểu xem nó có thực sự thuộc về ngôn ngữ ấy hay không. Lý thuyết này được gọi là **ngữ pháp (grammar)**. Để xác định, theo các nhà khoa học, cần phải có một tập hợp chứa đựng mọi quy tắc mẫu đã sản sinh ra ngôn ngữ ấy. Tập hợp này chính là *quy tắc ngữ pháp*. Giả như tất cả các từ nói ra đều thuộc về ngôn ngữ ấy, thì dựa vào ngữ pháp, người ta có thể xác nhận chuỗi từ nào đối với một ngôn ngữ nào được công nhận là một câu. Nếu được công nhận, câu ấy đúng ngữ pháp, còn không, nghĩa là sai hoặc có nghĩa khác.

(12) “*Mary is a pretty girl*”, dịch 1-1 theo đúng vị trí của câu tiếng Anh này sang tiếng Việt, ta có: “*Mary là một đẹp cô gái*”. Câu này sai ngữ pháp tiếng Việt. Nói đúng phải là “*Mary là một cô gái đẹp*”.

(13) Giả sử người Việt nói “*Hắn muốn hôn cô ta*”, người Đức nói: “*Er möchte sie küssen*”. Dịch sát từng chữ một theo đúng vị trí trong câu tiếng Đức sang tiếng Việt:

Er möchte sie küssen.
(Hắn) (muốn) (cô ta) (hôn)

Ý nghĩa khác hoàn toàn.

Cho nên, đối với ngữ pháp, vị trí của từ là chìa khóa quyết định xem một chuỗi từ có đúng là một câu của ngôn ngữ đang được xét đến hay không. Chúng ta sẽ tìm hiểu kỹ hơn về điểm này.

CẢM GIÁC VỀ MỘT CÂU NÓI ĐƯỢC

Theo Chomsky, một người có thẩm năng ngôn ngữ là người nói/nghe lý tưởng sống trong một cộng đồng tiếng nói đồng nhất và thành thạo ngôn ngữ của mình. Giả sử không mệt mỏi, say rượu, người ấy sẽ có thừa khả năng nói, nghe và đánh giá một câu nói có thực sự thuộc về thứ ngôn ngữ mà mình thành thạo hay không. Song trên thực tế, cho dù người ấy có khả năng xác nhận một câu nói sai hay đúng ngữ pháp, nhưng khi hỏi tại sao thế, có thể anh ta cũng mù tịt, chẳng biết phải giải thích thế nào.

Thẩm năng ngôn ngữ ở hầu hết mọi con người đều nằm dưới dạng vô thức. Con người thu thập kiến thức ngôn ngữ từ xã hội như một hình thức sao chép rồi nhớ, nói nôm na là người ta nói sao, mình nói vậy. Theo thời gian, mọi kiến thức ngôn ngữ nhập vào đầu và được cất giữ dưới dạng những dữ liệu, quy tắc ngữ pháp khuôn mẫu. Đó là những mẫu dữ liệu duy nhất mà con người có thể dùng để so sánh với dữ liệu nhận vào, xem đó có phải là ngôn ngữ của mình hay không. Thao tác so sánh này giống giống như cái ví dụ đi truy tìm từ của hệ nhận diện ngôn ngữ đã được trình bày ở bên trên.

Thực ra, quá trình nhận diện thông thường ấy chỉ thuần cảm giác chứ không phải nhờ tri thức khoa học để có thể giải thích một cách hiển ngôn, tại sao câu này nói được, câu kia không nói được. Xã hội chỉ dạy cho chúng ta cách nói đúng theo những mẫu ngữ pháp, chứ không dạy cho chúng ta cách giải thích cấu trúc ngôn ngữ. Công việc ấy là của nhà trường.

Thế thì chúng ta đã học được gì ở nhà trường ?

Có lẽ nhiều người, nhất là những người sống ở miền Nam Việt Nam trước 75, vẫn còn nhớ, hồi còn đi học, chúng ta không được học nhiều về ngữ pháp tiếng Việt. Trong giờ Việt văn, chúng ta học làm luận nhiều hơn. Chẳng hạn thầy cô cho chúng ta một cái đề đại khái như:

- (14) “Bầu ơi thương lấy bí cùng.
Tuy rằng khác giống nhưng chung một giàn”
(Ca dao)

rồi bảo, các em hãy bình luận. Thế là chúng ta lo bình. Có người thì bình, ý hai câu ca dao muốn nói, dù con người khác giống, khác hoàn cảnh sống, nhưng hãy yêu thương nhau, đùm bọc lẫn nhau. Có người thì bình, bầu và bí là hai loại quả nấu canh ăn rất là ngon. Chẳng thầy cô nào bảo chúng ta hãy phân tích “văn phạm” hai câu ca dao ấy bao giờ. Nói thế không có nghĩa là chúng ta không được học về ngữ pháp. Có lẽ ai cũng đều biết sơ qua những khái niệm như *chủ từ*, *động từ*, *túc từ*, *mạo từ*, *danh từ*, *tính từ*, ..., và có thể phân tích được câu dưới đây:

- (15) Vị bác sĩ có một tấm lòng nhân hậu.

- “Vị” (tạm gọi) là một *mạo từ* (*article*).

- “*bác sĩ*” là một *danh từ (noun)*.
- “*có*” là một *động từ (verb)*.
- “*tắm*” (tạm gọi) là một *mạo từ*.
- “*lòng*” là một *danh từ*.
- “*nhân hậu*” là một *tính từ (adjective)*.
- “*Vị bác sĩ*” là *chủ từ*.
- “*một tắm lòng nhân hậu*” là *túc từ*.

Hơn thế nữa, chúng ta còn biết:

1. Có những từ dính chùm thành cụm từ nói được, như “*vị bác sĩ*”, “*một tắm lòng*”, ...
2. Trong câu, mỗi từ đều có vị trí nhất định, ví dụ chủ từ phải đứng trước động từ, động từ thường đứng trước túc từ⁴⁾.
3. Hễ một chuỗi từ có chủ từ và động từ, thì đó là một câu. Có trường hợp câu cần túc từ, có trường hợp không. Có trường hợp chỉ có một từ duy nhất: “Im !”, “Đi !”, ...

Đại khái đó là những kiến thức ngôn ngữ của chúng ta. Đành rằng chưa giải thích được tại sao mình có khả năng lập câu đúng, suy nghĩ đúng, song ít nhất đó đã là một căn bản cần thiết giúp chúng ta tìm hiểu thêm về khoa học cú pháp, và tự nó cũng khá lý thú.

Trước nhất, thử tìm hiểu tại sao một câu được coi là nói được. Ví dụ câu: “*Vị bác sĩ có một tắm lòng nhân hậu*”.

Vỡ đại một cụm từ, ví dụ cụm từ “*vị bác sĩ*”, ra xem. Cụm từ này gồm có hai từ: “*vị*” và “*bác sĩ*”. Theo cảm giác, chúng ta thấy cụm từ này mang tính *danh từ*, vì “*vị bác sĩ*” là một nhân vật. Nói theo ngôn ngữ học truyền thống, cụm từ “*vị bác sĩ*” làm chủ từ trong câu. Điều này có lý, bởi chúng ta thường thấy chủ từ có dạng danh từ. Vậy thì – suy nghĩ một cách logic - nếu thay thế cụm từ “*vị bác sĩ*” bằng một danh từ hay một cụm từ nào đó cũng mang tính danh từ giống vậy mà nói được, chúng ta sẽ có một câu nói được:

- (16) a. Ông bác sĩ có một tắm lòng nhân hậu.
b. Bà bác sĩ có một tắm lòng nhân hậu.

⁴⁾ Cứ tạm gọi là chủ từ, động từ, túc từ như đã quen gọi trong ngôn ngữ học truyền thống trước nay cho dễ hình dung. Sau này chúng ta sẽ gọi chủ từ là chủ ngữ, động từ là vị từ (predicate), túc từ là bổ ngữ (complement).

- c. *Cô y tá* có một tấm lòng nhân hậu.
- d. *Chúa* có một tấm lòng nhân hậu.
- e. *Phật* có một tấm lòng nhân hậu.

Quả thực, mọi câu đều đúng. Song, nếu đưa một *tính từ* vào vị trí ấy, câu sẽ bị hỏng ngay:

- (17) a. * *Từ bi* có một tấm lòng nhân hậu.
 b. * *Thông minh* có một tấm lòng nhân hậu.
 c. * *Khờ khạo* có một tấm lòng nhân hậu.

Vậy chứ chỉ có trường hợp danh từ mới dùng được hay sao ? Thử tiếp tục thay thế bằng vài cụm từ khác cũng mang tính danh từ:

- (18) a. * *Cái bàn* có một tấm lòng nhân hậu.
 b. * *Cái ghế* có một tấm lòng nhân hậu.
 c. * *Con gà* có một tấm lòng nhân hậu.
 d. * *Cái nôi* có một tấm lòng nhân hậu.
 e. * *Con số không* có một tấm lòng nhân hậu.

Kết quả, không câu nào nói được. Tại sao vậy ? Cùng là danh từ mà không làm chủ từ được ?

Chắc chắn sẽ có ý kiến cho rằng, “*cái bàn*”, “*cái ghế*”, “*con gà*”, ... không cùng *ngữ cảnh (context)* như những cụm từ “*ông bác sĩ*”, “*bà bác sĩ*”, “*cô y tá*”, ..., tại vì chúng không chỉ người. Có lý. Thế nhưng nếu một ông/bà nhà văn nào đó nói:

- (19) a. *Quê hương tôi* có một tấm lòng nhân hậu.
 b. *Thơ ai* có một tấm lòng nhân hậu.

có lẽ nghe vẫn êm tai, mặc dù những cụm từ được đem vào thay thế ấy không chỉ người.

Lối nói này mang tính văn chương bay bướm, bởi lẽ ai cũng biết “*tấm lòng*” là một danh từ dành riêng cho người hoặc các đấng thần linh. “*Tấm lòng*” không thích hợp cho bàn ghế, gà vịt, nôi niêu xoong chảo, ... Tuy nhiên, nếu người ta vẫn chấp nhận nói được, thì âu cũng là do tính thần của những cụm từ đem vào thế đã được nhân cách hóa. Hiện tượng sử dụng ngôn ngữ không đại chúng như một hình thức *siêu ngôn ngữ (metalinguage)* này, chúng ta khoan bàn đến.

Kết luận: Chuyện “*nói được hay không*” không những chỉ lệ thuộc vào vị trí và tính loại của từ/cụm từ mà còn lệ thuộc vào *ngữ cảnh*.

VỊ TRÍ VÀ QUAN HỆ LỆ THUỘC CỦA TỪ

Nhận xét đầu tiên khi phân tích một câu là *vị trí của từ trong câu phải đúng thì câu mới đúng*. Thử lấy vài ví dụ và đổi vị trí của từ trong câu:

Tiếng Việt:

- (20) a. Ông ta là một giáo sư âm nhạc: đúng ngữ pháp và là câu xác định.
 b. * Là ông ta một giáo sư âm nhạc.
 c. * Ông ta là một âm nhạc giáo sư.
 d. * Âm nhạc là một ông ta giáo sư.

Tiếng Anh:

- (21) a. * He is a professor music.
 b. Is he a music professor: đúng ngữ pháp và là câu hỏi
 c. He is a music professor: đúng ngữ pháp và là câu xác định.
 d. * Music is a he professor.

Tiếng Đức:

- (22) a. * Er ist ein Professormusik.
 b. Ist er ein Musikprofessor: đúng ngữ pháp và là câu hỏi
 c. Er ist ein Musikprofessor: đúng ngữ pháp và là câu xác định.
 d. * Musik ist ein er Professor.

Các ví dụ trên cho thấy, vị trí của từ trong câu đúng ngữ pháp không bất kỳ mà có một trật tự nhất định. Trật tự này quyết định một cấu trúc hợp lý. Cùng một cấu trúc câu tương ứng, có ngôn ngữ cho là đúng ngữ pháp, có ngôn ngữ lại cho là sai hoặc mang ý nghĩa khác hoặc có cường độ khác:

- Trong mọi trường hợp (a.): Cụm từ “*giáo sư âm nhạc*” của tiếng Việt đúng ngữ pháp, nhưng nếu dịch 1-1 sang tiếng Anh “*professor music*” hoặc tiếng

Đức thành “*Professormusik*”, cả hai đều sai. Đúng ngữ pháp phải là “*music professor*”, “*Musikprofessor*”.

- Trong mọi trường hợp (b.): Tiếng Việt sai ngữ pháp. Tiếng Anh, Đức đúng và là câu hỏi.
- Trong mọi trường hợp (d.): Tất cả đều sai ngữ pháp.
- Nếu bỏ 3 danh từ “*âm nhạc*”, “*music*” và “*Musik*” ra khỏi 3 câu (a.), tất cả các ví dụ đều đúng ngữ pháp:

- (23) a. Ông ta là một giáo sư.
 b. He is a professor.
 c. Er ist ein Professor.

Vị trí từ rất quan trọng đối với ngữ pháp và ý nghĩa của một chuỗi từ. Xét đến vị trí từ trong một chuỗi từ là xét đến *tính tuyến tính (linearity)* của cấu trúc ngôn ngữ. Thế nhưng vẫn chưa đủ, bởi từ còn có mối quan hệ đối với nhau. Có những từ liên kết với nhau chặt chẽ hơn những từ khác. Có từ quan trọng, đóng vai ý chính, tức là làm trung tâm ý nghĩa. Có từ không quan trọng, chỉ giữ nhiệm vụ thêm nghĩa cho từ khác. Trong mối quan hệ này, các từ phải đứng đúng vị trí để tạo nên *thành tố* (hiểu nôm na là một cụm từ/chuỗi từ nói được). Khi đổi vị trí của một từ, những từ khác trong thành tố phải được đổi theo. Chỉ bằng cách ấy, toàn thể cấu trúc còn lại và ý nghĩa của chuỗi từ mới không bị hủy. Điểm này nói lên tính *quan hệ lệ thuộc* của từ.

Quan hệ lệ thuộc (dependency) của từ là mối quan hệ *tôn ti (hierarchy)* giữa hai từ trong chuỗi từ, ở đó, sự hiện diện, hình thức cũng như hoạt động cú pháp của một từ sẽ bị từ kia kiểm soát.

Thử nghiệm bằng cách lược từ, chúng ta sẽ thấy, có những từ phải được bỏ *cùng lúc* thì câu mới còn giữ đúng ngữ pháp.

- (24) Trị quốc là một nghệ thuật rất cao quý.

Bỏ tính từ “*cao quý*”:

- (25) * Trị quốc là một nghệ thuật rất ~~cao quý~~. (Không nói được)

Nhưng bỏ trạng từ “*rất*”:

- (26) Trị quốc là một nghệ thuật ~~rất~~ cao quý. (Nói được)

Bỏ trạng từ “*rất*” và tính từ “*cao quý*” cùng một lúc:

(27) Trị quốc là một nghệ thuật ~~rất cao quý~~. (Nói được)

Bỏ danh từ “*nghệ thuật*”:

(28) * Trị quốc là một ~~nghệ thuật~~ rất cao quý. (Không nói được)

Bỏ mạo từ “*một*” và trạng từ “*rất*” cùng lúc:

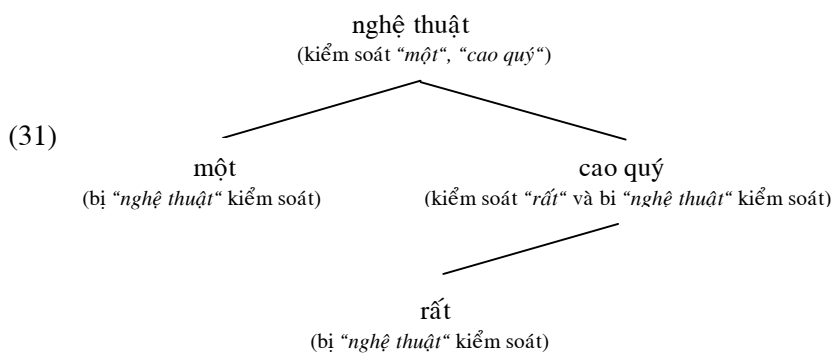
(29) Trị quốc là ~~một~~ nghệ thuật ~~rất~~ cao quý. (Nói được nhưng hơi lạ)

Bỏ mạo từ “*một*”, trạng từ “*rất*” và tính từ “*cao quý*” cùng lúc:

(30) Trị quốc là ~~một~~ nghệ thuật ~~rất cao quý~~. (Nói được).

Qua trò chơi gạch bỏ từ, chúng ta đã nhìn thấy sơ mối quan hệ lệ thuộc của từ trong tiếng Việt:

- Trạng từ phải đi đôi với tính từ (“*rất*” + “*cao quý*”).
- Mạo từ đi đôi với danh từ (“*một*” + “*nghệ thuật*”).
- Trạng từ “*rất*” bị tính từ “*cao quý*” kiểm soát.
- Tính từ “*cao quý*”, mạo từ “*một*” bị danh từ “*nghệ thuật*” kiểm soát.



Hình vẽ miêu tả các phần tử bị kiểm soát đều có vị trí thấp hơn phần tử kiểm soát chúng (sẽ trình bày chi tiết về cây cú pháp sau). Xin lưu ý, “**kiểm soát**” là cách nói của giới ngôn ngữ học. Ngoài ra còn thấy hay dùng từ “**chi phối**”.

Quan sát tiếng Việt trong thực tế kỹ hơn, chúng ta còn thấy vài tính chất kiểm soát sau đây:

Tính từ kiểm soát trạng từ: Xét về mặt *quan hệ lệ thuộc*, trạng từ đi đôi với tính từ là để thêm nghĩa cho tính từ chứ không quan trọng bằng tính từ. Còn xét về mặt *vị trí*, thì trạng từ tùy loại có thể đứng sau hay đứng trước tính từ (tùy ngôn ngữ, trường hợp). Trong cụm từ “*rất cao quý*”, trạng từ “*rất*” đứng trước. Trong cụm từ “*cao quý lắm*” thì trạng từ “*lắm*” lại đứng sau.

Tính từ kiểm soát tính từ: Tính từ còn thêm nghĩa cho cả tính từ:

- (32) Cô ấy hát *hay* ghê.
- (33) Cô ấy hát *hay* dễ sợ.

“*Hay*” kiểm soát “*ghê*” và “*dễ sợ*”. Bỏ “*hay*”, câu sẽ sai ngữ pháp vì trong trường hợp này, “*hay*” có nhiệm vụ bổ nghĩa cho động từ, không thể thiếu:

- (34) * Cô ấy hát ghê.
- (35) * Cô ấy hát dễ sợ.

Nếu thêm “*thật*” đằng sau, thì câu lại nói được; “*ghê*” và “*dễ sợ*” kiểm soát “*thật*”:

- (36) Cô ấy hát ghê *thật*.
- (37) Cô ấy hát dễ sợ *thật*.

Bỏ “*ghê*” và “*dễ sợ*”, câu vẫn nói được, vì trong trường hợp này, “*thật*” thêm nghĩa cho động từ, tức động từ kiểm soát tính từ (xem thêm bên dưới):

- (38) Cô ấy hát *thật*.

Động từ kiểm soát trạng từ: Trạng từ còn thêm nghĩa cho cả động từ, tức bị động từ kiểm soát. Mỗi quan hệ lệ thuộc cũng giống y như thế:

- (39) Ông ta *rất* thương trẻ mồ côi.

- (40) Ông ta thương trẻ mồ côi *lắm*.
- (41) Tay cán bộ ấy *rất* ham ăn hối lộ.
- (42) Tay cán bộ ấy ham ăn hối lộ *lắm*.

Danh từ kiểm soát tính từ: Xét về mặt quan hệ lệ thuộc, tính từ không quan trọng bằng danh từ. Tính từ chỉ được dùng để thêm nghĩa cho danh từ. Nếu xét về vị trí, chắc chắn nó phải đứng sau danh từ. Chỉ những từ/cụm từ Hán-Việt mới cho phép tính từ đứng phía trước, ví dụ “*mỹ nhân*”. “*Mỹ*” (đẹp) là tính từ. “*Nhân*” (người) là danh từ. Người Việt nói “*người đẹp*” với tính từ “*đẹp*” đứng sau danh từ “*người*”.

Động từ kiểm soát tính từ: Tính từ còn phụ nghĩa cho cả động từ. Trong tiếng Đức cũng vậy:

- (43) Cô ấy hát *hay* ghê.
- (44) Sie singt so *schön*.

Tính từ “*hay*” phụ nghĩa cho động từ “*hát*”. Tính từ “*schön*” phụ nghĩa cho động từ “*singt*”.

Động từ kiểm soát danh từ: Danh từ cũng có thể phụ nghĩa cho động từ:

- (45) “Con cò mà đi ăn *đêm*.
Đâu phải cảnh mềm, lộn cổ xuống ao”.
(Ca dao)
- (46) Ra tiệm phở, ăn *sáng*.

Danh từ “*đêm*” và “*sáng*” đều phụ nghĩa cho động từ “*ăn*”. Xin lưu ý, “*sáng*” ở đây được hiểu là buổi sáng.

Danh từ kiểm soát danh từ: Cả danh từ cũng có thể phụ nghĩa cho một danh từ:

- (47) Thành phố *Huế*
- (48) Sông *Hương*
- (49) Cánh đồng *lúa*
- (50) Cây cầu *khỉ*

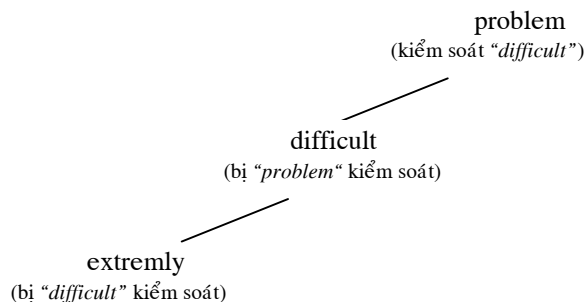
“Huế”, “Hương” là danh từ riêng. “Huế” phụ nghĩa cho “thành phố”, “Hương” phụ nghĩa cho “sông”. “Lúa”, “khỉ” là danh từ chung. “Lúa” phụ nghĩa cho “cánh đồng”, “khỉ” phụ nghĩa cho “cây cầu”.

Một từ hoặc một cụm từ có khả năng thêm nghĩa cho một từ khác bằng cách làm biến đổi ý nghĩa nguyên thủy của từ ấy, sẽ được gọi là **phụ ngữ (adjunct)**. Trong khi đó, cũng bằng hình thức thêm nghĩa, một **bổ ngữ (complement)** có tính cách bắt buộc hơn, gắn bó hơn phụ ngữ. Bỏ bổ ngữ, chuỗi từ sẽ sai ngữ pháp hoặc có ý nghĩa khác. Chúng ta sẽ bàn chi tiết về sự khác biệt này ở các chương sau.

Sơ lược về mối quan hệ lệ thuộc của từ trong Việt là thế. Tiếng Anh tương tự:

(51) An extremely difficult problem

- Trạng từ “*extremely*” bị tính từ “*difficult*” kiểm soát.
- Tính từ “*difficult*” bị danh từ “*problem*” kiểm soát.



Tiếng Đức phức tạp hơn. Không những tính từ bị danh từ kiểm soát mà cả *hình vị* của nó cũng bị kiểm soát. Tính từ nguyên mẫu thay đổi tùy theo vị trí và thuộc tính của danh từ, như *phái tính* (giống đực, giống cái, giống trung), *lượng* (số nhiều, số ít), *thể cách* (*danh cách (nominative)*, *sinh cách (genitive)*, *tặng cách (dative)*, *đối cách (accusative)*, ...), ...

Danh từ kiểm soát tính từ:

(52) Ein *schönes* Bild (das Bild) (Giống trung, số ít, danh cách).
(Một bức tranh đẹp)

Bởi “*Bild*” (*bức tranh*) là số ít, giống trung, danh cách (tức danh từ ở vị trí chủ ngữ), nên tính từ “*schön*” (*đẹp*) được thêm “*es*” thành ra “*schönes*”.

- (53) Ein *schöner* Garten (der Garten) (Giống đực, số ít, danh cách).
(Một cái vườn đẹp)

Bởi “*Garten*” (*vườn*) là số ít, giống đực, danh cách, nên tính từ “*schön*” (*đẹp*) được thêm “*er*” thành ra “*schöner*”.

- (54) Eine *schöne* Blume (die Blume) (Giống cái, số ít, danh cách).
(Một bông hoa đẹp)

Bởi “*Blume*” (*bông hoa*) là số ít, giống cái, danh cách, nên tính từ “*schön*” (*đẹp*) được thêm “*e*” thành ra “*schöne*”.

- (55) Die *schönen* Blumen (die Blumen) (Giống cái, số nhiều, danh cách/đối cách).
(Những bông hoa đẹp)

Bởi “*Blumen*” (*những bông hoa*) là số nhiều, giống cái, nên tính từ “*schön*” (*đẹp*) được thêm “*en*” thành ra “*schönen*”.

Khá nhức đầu nhưng chưa hết.

Động từ kiểm soát bổ ngữ:

- (56) Ich danke Ihnen (Tặng cách).
Tôi cảm ơn ông/bà.

Bởi động từ “*danke*” (*cảm ơn*, nguyên mẫu: *danken*) chỉ đi với một bổ ngữ đóng vai *tặng cách*, thường gọi là *túc từ gián tiếp* (*indirect object*), nên “*ông/bà*” có hình vị “*Ihnen*”. Nếu động từ đi với một bổ ngữ đóng vai *đối cách*, thường gọi là *túc từ trực tiếp* (*direct object*), thì “*ông/bà*” lại có hình vị “*Sie*”:

- (57) Ich verstehe Sie (Đối cách).
Tôi hiểu ông/bà.

Xin hiểu túc từ trực tiếp là loại túc từ bị động từ ảnh hưởng trực tiếp, ngược lại là túc từ gián tiếp.

(58) Chị tôi mua *một cái áo*.

“*Một cái áo*” là *túc từ trực tiếp* của động từ “*mua*”. Trong khi đó:

(59) Chị tôi mua cho *tôi* một cái áo.

“*Một cái áo*” là *túc từ trực tiếp* của hành động “*mua*”. Còn “*tôi*” là *túc từ gián tiếp*. Bỏ thành tố “*cho tôi*”, câu vẫn nói được như thường. Nhưng bỏ thành tố “*một cái áo*”, câu sẽ bị thiếu, vì động từ “*mua*” thiếu món hàng gì đó, tức *túc từ trực tiếp*.

Một trường hợp đối cách khác của tiếng Đức:

(60) Herr Schmidt hat gerade *einen neuen Mercedes gekauft* (Đối cách).
Ông Schmidt vừa tâu *một* chiếc Mercedes mới toanh.

Động từ “*gekauft*” (*đã mua*, nguyên mẫu: *kaufen*) chỉ đi với bổ ngữ đóng vai *đối cách*, thành thử khi sử dụng mạo từ “*một*”, người Đức nói “*einen*” (*ein* + ***en***). Nếu động từ đi với một bổ ngữ đóng vai *tặng cách*, ví dụ động từ “*fahren*” (*lái*), họ lại nói “*einem*” (*ein* + ***em***):

(61) Herr Schmidt fährt mit *einem* Mercedes (Tặng cách).
(Ông) Schmidt (lái) (với) (*một*) (Mercedes) (Dịch 1-1)

Tựu trung, hình vị của mạo từ, tính từ và danh từ ở vị trí bổ ngữ trong tiếng Đức đều bị động từ kiểm soát.

Đại danh từ nhân vật kiểm soát động từ:

Hình vị của động từ nguyên mẫu tiếng Đức còn biến đổi theo từng đại danh từ nhân vật:

(62) Ich *singe* (đại danh từ nhân vật ngôi 1, số ít).
Tôi hát.

(63) Du *singst* (đại danh từ nhân vật ngôi 2, số ít).
Mi hát.

(64) Er *singt* (đại danh từ nhân vật ngôi 3, số ít).
Nó hát.

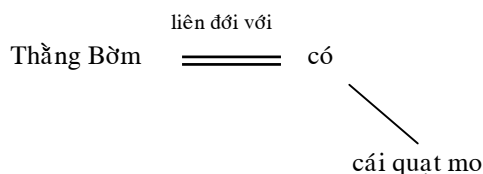
- (65) Sie *singen* (đại danh từ nhân vật ngôi 3, số nhiều).
Họ hát.

Tóm lại, quan hệ lệ thuộc của từ có thể quyết định một hình vị tương ứng, tiêu biểu là ở các ngôn ngữ Âu châu, như chúng ta vừa thấy, còn trong tiếng Việt thì không.

Sự lệ thuộc của từ không những chỉ một chiều mà có thể cả hai. Trường hợp này nói lên **sự liên đới (interdependency)** giữa hai phần, ở đó, cả hai đều có vị trí ngang nhau. Dễ thấy nhất là sự liên đới giữa phần **chủ ngữ (subject)** và phần **vị ngữ (predicate phrase)** trong câu. Xin lưu ý, **vị ngữ** là phần **động từ cộng túc từ**, tức phần còn lại của câu (nếu có) (các phần này sẽ được nói rõ hơn ở các chương sau).

- (66) Thăng Bờm có cái quạt mo.

“Thăng Bờm” là **chủ ngữ**. “Cái quạt mo” là **vị ngữ**.



Đã nắm rõ ý nghĩa của **vị trí** và **quan hệ lệ thuộc** của từ, nay ta có thể bắt tay vào việc phân tích câu để xem cụm từ, chuỗi từ nào là thành tố.

PHƯƠNG PHÁP PHÂN TÍCH CÂU

Phải nói, những cố gắng chính của các nhà ngôn ngữ học thiên về cấu trúc luận kiểu Mỹ sau thời đại Bloomfield là bao giờ cũng tìm cách nghĩ ra những phương pháp tiêu chuẩn nhằm phân tích câu theo lối cấu trúc. Họ cho rằng, ngôn ngữ là cái gì đó có cấu trúc phức tạp cần được khám phá. Cụm từ “*discovery procedures*” đã đi ra từ đó và đã trở thành chủ đích của công việc nghiên cứu ngôn ngữ. Đại để trong công việc này là hãy khám phá được những cái giống nhau, cái khác nhau trong lời lẽ thực tế ngoài đời, thì thể nào cũng tìm ra được quy tắc ngữ pháp.

Sáng kiến chính của phương pháp này là **phân đoạn (segment)**. Một lời nói sẽ được phân ra làm nhiều đoạn tuyến tính và được xếp loại theo những chuẩn tắc nhất định. Mục đích của phân đoạn là khảo sát xem cụm từ/chuỗi từ nào trong câu là *thành tố*. Cụ thể là 6 phương pháp thử nghiệm sau (s.s. Glinz 1965):

1. Phép thế (substitution)
2. Phép thử câu hỏi (question test)
3. Phép bố trí (coordination)
4. Phép tính lược (reduction)
5. Phép hoán vị (permutation)
6. Phép biến hình (transformation)

1. Phép thế

Một trong những phương pháp phân tích câu thông dụng nhất là **phép thế (substitution)**. Điều kiện dùng phép thế đòi hỏi mọi phần tử được đem vào thế phải hợp lý và phải ở cùng môi trường ngữ cảnh. Những phần tử này có thể là những phần tử **đối lập (opposition)**. “Ông cụ” đối lập với “đứa bé”, “hoa hồng” đối lập với “hoa lan”, ... Từng cặp phần tử có mối quan hệ đối lập được xem như **đối vị (paradigmatic)** và được xếp chung vào một tập hợp được gọi là **hệ hình (paradigm)**. Xin lưu ý, tính chất đối lập ở đây không gắt gao đến độ hai ý nghĩa đối lập bắt buộc phải hoàn toàn trái ngược, như “đàn ông” đối lập với “đàn bà”, “trời” đối lập với “đất”, ...

Mục đích của phép thế là nhằm xác định *thành tố cùng loại* qua cách thử thay thế *một cụm từ trong chuỗi từ bằng một cụm từ khác*; phần còn lại giữ nguyên. Cụm từ mới không được phép làm sai ngữ pháp. Cụm từ mới với ngữ cảnh tương tự, hoặc cùng loại ngữ pháp, hệ hình có thể tạo nên *thành tố*. Ví dụ thay thế cụm từ “đứa bé” trong câu “đứa bé đọc sách” bằng những cụm từ khác như “ông cụ”, “ông giáo sư âm nhạc” ...:

(67) Đứa bé đọc sách.

{	Ông cụ Ông giáo sư âm nhạc Hùng Hùng và Hưng Nhiều người	}	đọc sách.
---	--	---	-----------

Tất cả các cụm từ trong ngoặc {} đều làm thành những câu đúng ngữ pháp. Đặt câu hỏi, vậy, phải chăng những cụm từ ấy đều là những thành tố cùng loại? Nếu giả thiết này đúng, cũng bằng phép thế, hãy thử xét câu:

(68) Cô sinh viên vừa ăn vừa đọc sách.

Câu này đúng ngữ pháp nhưng chuỗi từ “*Cô sinh viên vừa ăn vừa*” không thể coi là thành tố, vì không nói được, nếu đứng độc lập.

Trên thực tế, cần cẩn thận. Phép thế không bảo đảm kết quả tìm thấy bao giờ cũng là một thành tố đúng đối với tiếng Việt. Tiếng Việt có những cấu trúc mà ngôn ngữ khác không có, điển hình là tiếng Đức không có cách nói hoàn toàn giống như cấu trúc “*vừa làm cái này vừa làm cái nọ*” như trên.

2. Phép thử câu hỏi

Điểm yếu của phép thế là đòi hỏi nhiều cảm giác thử nghiệm và không dễ xác định thành tố cùng loại. Để khắc phục khó khăn này, người ta dùng thêm phương pháp “*thử câu hỏi*” (*question test*). Cụ thể, cũng với ví dụ “*đứa bé đọc sách*”, nếu đặt câu hỏi:

(69) Ai đọc sách?

và trả lời bằng những cụm từ có sẵn:

- (70) a. Ông cụ.
b. Ông giáo sư âm nhạc.
c. Hùng.
d. Hùng và Hưng.
e. Nhiều người.
f. * Cô sinh viên vừa ăn vừa.

Các cụm từ thử nghiệm trên đều là các thành tố cùng loại giống như “*đứa bé*”, trừ trường hợp “*cô sinh viên vừa ăn vừa*”. Có lý phải là cụm từ “*cô sinh viên*”.

3. Phép bố trí

Với **phép bố trí (coordination)**, một câu có thể được mở rộng bằng cách dùng *liên từ (conjunction)*, ví dụ “và”, để nối thêm một cụm từ. Nhiều nhà ngôn ngữ học theo hướng cấu trúc luận tin rằng, cụm từ được nối phía sau sẽ là một thành tố cùng loại.

(71) a. *Hút thuốc có hại cho sức khỏe.*

mở rộng câu a. bằng cách thêm cụm từ “*uống rượu*” và liên từ “và” ở giữa:

b. *Hút thuốc và uống rượu có hại cho sức khỏe.*

Câu b. nói được. Vậy, hai cụm từ “*hút thuốc*” và “*uống rượu*” là hai thành tố cùng loại. Tương tự:

- (72) a. *Lan đặt ngọn nến lên bàn.*
 b. *Lan đặt bình hoa lên bàn.*
 c. *Lan đặt ngọn nến và bình hoa lên bàn.*

Tiếng Anh:

- (73) a. *Peter runs up the hill.*
 b. *Peter runs up the stairs.*
 c. *Peter runs up the hill and up the stairs.*

Nhưng:

- (74) a. *Peter rings up his mother.*
 b. *Peter rings up his sister.*
 c. * *Peter rings up his mother and up his sister.*

Cách bố trí của “*up his mother*” và “*up his sister*” cho thấy, không một chuỗi từ nào nói được. Mặt khác, mặc dù “*up the hill*”, “*up the stairs*” là hai thành tố chuẩn, nhưng cũng cùng một cấu trúc giống hệt vậy, hai chuỗi từ “*up his mother*” và “*up his sister*” lại không phải là thành tố. Cách bố trí của vế [*Peter*] [*runs*] [*up the hill*] khác với [*Peter*] [*rings up*] [*his mother*]. Xin lưu ý, dấu ngoặc [] cũng thường được dùng cho mục đích phân đoạn câu.

Vài ví dụ tiếng Anh khác:

- (75) a. Peter wrote to Mary and to Bill.
 b. Peter wrote a postcard and a letter.
 c. * Peter wrote to Mary and a letter.
 d. * Peter wrote a postcard and to Bill.

4. Phép tỉnh lược

Với cách tìm thành tố theo **phép tỉnh lược (reduction)**, chúng ta xem cụm từ trùng lặp nào trong chuỗi từ có thể lược bỏ được mà ý nghĩa chính của nó vẫn giữ nguyên nhờ ngữ cảnh phía trước.

- (76) a. Tiếng Anh: Bill won't *help me*, but John will ~~*help me*~~.
 b. Tiếng Đức: Die Mutter *schenkt* dem Sohn eine Hose und ~~*schenkt*~~ der Tochter einen Rock.
 (Bà mẹ tặng đứa con trai một cái quần còn đứa con gái (thì một)) chiếc váy)
 c. Tiếng Việt:
 “Lúc chết, hai chân anh *đi* đằng trước,
 mắt ~~*đi*~~ đằng sau còn vương vấn cuộc đời.
 Em có còn gì để lại trên môi,
 hay tình yêu chẳng qua là ảo ảnh ?”
 (Nguyễn Sa)

Phép tỉnh lược tương đối hữu ích cho việc định *cấu trúc cú pháp tối thiểu (minimal syntax structure)* của một câu:

- (77) a. Nhà cách mạng ôn hòa Gandhi đã từ trần hôm qua.
 b. Nhà cách mạng ~~ôn hòa~~ Gandhi đã từ trần hôm qua.
 c. Nhà cách mạng ~~ôn hòa~~ Gandhi đã từ trần ~~hôm qua~~.
 d. ~~Nhà cách mạng~~ ~~ôn hòa~~ Gandhi đã từ trần ~~hôm qua~~.
 e. ~~Nhà cách mạng~~ ~~ôn hòa~~ Gandhi ~~đã~~ từ trần ~~hôm qua~~.

“Gandhi từ trần” là câu nói được và ngắn gọn nhất, không thể bỏ thêm.

Đối với tiếng Việt, phương pháp lược từ này giúp chúng ta dễ nhận diện, từ nào đại diện cho từ nào, nghĩa là từ nào đóng vai cú pháp mà không thể bỏ được:

- (78) “Anh ~~đồ~~ tỉnh, anh ~~đồ~~ say,

Sao *anh* ghẹo nguyệt giữa ban ngày ?
 Này này chị bảo cho mà biết,
 Chốn ấy hang hằm chớ mó tay”
 (Trách Chiêu Hồ (1) - Hồ Xuân Hương)

Ở dòng 2 của ví dụ trên, “*đồ*” bỏ được nhưng “*anh*” thì không:

(79) * Sao *đồ* ghẹo nguyệt giữa ban ngày ?

“*Anh*” phải giữ vai trò đại diện cho “*anh đồ*” như một đơn vị cú pháp.

Đối với ngôn ngữ Âu châu như tiếng Anh, tiếng Đức, phương pháp này thường được dùng để lược bỏ động từ; câu lược bỏ nghe vẫn ổn. Song đối với tiếng Việt, không phải bao giờ cũng đơn giản thế. Bỏ động từ dễ làm sai ngữ pháp:

- (80) Mary *ißt* eine Orange, Hans ~~*ist*~~ einen Apfel.
 (Mary ăn một trái cam, còn Hans một trái táo.)
 (81) * Trèo lên cây búi *hái* hoa, bước xuống vườn cà ~~*hái*~~ nụ tầm xuân.
 (Câu không nói được, nếu bỏ “*hái*” ở đoạn sau).
 (82) * Tôi *đi* chơi, còn nó ~~*đi*~~ làm.
 (Câu không nói được vì thiếu “làm cái gì”).

Hoặc làm thay đổi ý nghĩa:

- (83) Tôi *đi xem* triển lãm, còn nó ~~*đi xem*~~ đá banh.
 (Ý nghĩa “*đi xem đá banh*” khác “*đá banh*”).

5. Phép hoán vị

Phương pháp **hoán vị** (*permutation*) cho phép chuyển đổi thứ tự những cụm từ trong câu. Cụm từ đổi được là thành tố.

Tiếng Anh:

- (84) a. I don't like *your teacher*.
 b. *Your teacher*, I don't like.

Tiếng Đức (cùng nghĩa):

- (85) a. Ich mag *Deinen Lehler* nicht so sehr.
b. *Deinen Lehler* mag ich nicht so sehr.

Tiếng Việt (cùng nghĩa):

- (86) a. Tớ không thích ông thầy cậu lắm.
b. Ông thầy cậu, tớ không thích lắm

Các cụm từ “*your teacher*”, “*Deinen Lehler*”, “*ông thầy cậu*” đều là thành tố. Xin lưu ý, câu b. cũng là một câu đề diễn trong tiếng Việt: “*Ông thầy cậu* (thì) tớ không thích lắm”.

Đối với tiếng Anh, Đức, hoán vị câu có lẽ là chuyện bình thường. Nhưng dùng phép này trong tiếng Việt, câu dễ trở nên lẩn cẩn:

- (87) a. Anh ấy chưa có việc làm, vẫn còn thất nghiệp.
b. ? ⁵⁾Việc làm, anh ấy chưa có, vẫn còn thất nghiệp. (Khó nghe).
c. ? Chưa có việc làm, anh ấy vẫn còn thất nghiệp. (Khó nghe).

Người Âu châu thường dùng phép hoán vị để nhấn mạnh. Trong trường hợp này, thành tố sẽ được chuyển lên đầu câu.

Điểm đáng lưu ý trong khi hoán vị là không được tách rời từ ngữ trong thành tố một cách tùy tiện:

- (88) a. * *Làm*, anh ấy chưa có *việc*, vẫn còn thất nghiệp.
b. * *Có* việc làm, anh ấy *chưa*, vẫn còn thất nghiệp.
c. * Chưa có *việc*, anh ấy *làm*, vẫn còn thất nghiệp.

6. Phép biến hình

Ngoài những phương pháp thử nghiệm kể trên còn có ***phép biến hình (transformation)***. Phép biến hình thường được dùng để chẩn đoán những cấu trúc từng phần. Ví dụ những hiện tượng sau đây trong tiếng Anh:

⁵⁾ Trường hợp không rõ sẽ được đánh dấu “?”

Câu chẻ (cleft sentence): Trong cấu trúc câu chẻ, chủ ngữ không đứng trực tiếp trước động từ mà được mệnh đề chính “*It is/was*” giới thiệu như một hình thức chủ thể. Chủ thể sẽ được diễn giải bằng một mệnh đề phụ:

(89) It was Mary who rang up me.

Câu chẻ giả (pseudo-cleft sentence): Tương tự như câu chẻ, bổ ngữ (túc từ) trong cấu trúc *câu chẻ giả* sẽ được đưa lên đầu câu và được diễn giải ở phần cuối câu.

(90) What stood on the table was the vase.

Quan hệ chủ động-bị động (active-passive relation):

(91) The student found the solution to the problem.

(92) The solution to the problem was found by the student.

Ví dụ, từ một câu nguyên thủy:

(93) The vase stood on the table.

ta có thể tìm thấy những thành tố (trong “...”) sau khi biến hình:

- (94) a. “*It*” stood on the table. (*It* = The vase, phép đại danh từ hóa)
 b. What stood on the table was “*the vase*”. (Câu chẻ giả)
 c. Where did the vase stand ? “*On the table*”. (Phép thử câu hỏi)
 d. What the vase stood on was “*the table*”. (Câu chẻ giả)
 d. The vase “*stood on the table*” and “*looked splendid*”. (Phép bố trí)

Tiếng Đức (cùng nghĩa):

(95) Die Vase stand auf dem Tisch.

- (96) a. “*Sie*” stand auf dem Tisch. (*Sie* = die Vase, phép đại danh từ hóa)
 b. ? Was auf dem Tisch stand war “*die Vase*”. (Câu chẻ giả)⁶⁾

⁶⁾ Đúng ngữ pháp nhưng trên thực tế hầu như không ai nói.

- c. Wo stand die Vase ? “*Auf dem Tisch*”. (Phép thử câu hỏi)
- d. ? Wo(rauf) die Vase stand war “*der Tisch*”. (Câu chẻ giả)⁷⁾
- e. Die Vase “*stand auf dem Tisch*” und “*sah prächtig aus*”. (Phép bố trí)

Tiếng Việt nguyên chất hình như không có hiện tượng “*câu chẻ*”, ngoại trừ cách dịch thuật đã bị lai ngôn ngữ Âu châu.

(97) Tuổi trẻ ngày nay rất thích sống tự do.

- (98) a. “*Chúng nó*” ngày nay rất thích sống tự do. (*Chúng nó* = tuổi trẻ, cách đại danh từ hóa)
- b. Cái mà tuổi trẻ ngày nay rất thích là “*sống tự do*”. (Gần giống câu chẻ giả)
- c. Tuổi trẻ ngày nay rất thích gì ? “*Sống tự do*”. (Cách đặt câu hỏi)
- d. Tuổi trẻ ngày nay rất thích “*sống tự do*” và “*tự lập*”. (Cách bố trí)

Những cụm từ “*the vase*”, “*on the table*”, “*the table*”, “*stood on the table*”, “*die Vase*”, “*auf dem Tisch*”, “*der Tisch*”, “*stand auf dem Tisch*”, “*sống tự do*” đều là những thành tố có thể xác định được nhờ phép biến hình. Trong tiếng Việt và tiếng Đức, cách biến hình không phổ biến như trong tiếng Anh.

Phép biến hình còn được dùng khá nhiều trong việc giải thích cấu trúc sâu và cấu trúc bề mặt của ngôn ngữ trong lý thuyết ngữ pháp tạo sinh (sẽ vào chi tiết ở các chương sau).

TÓM LƯỢC

Cấu trúc mỗi ngôn ngữ khác nhau là do cách xếp đặt *vị trí từ*. Mặt khác, trong một thành tố, có từ quan trọng đóng vai chính, có từ không quan trọng đóng vai phụ để bổ nghĩa/phụ nghĩa cho vai chính. Đây là *mối quan hệ lệ thuộc* giữa các từ đối với nhau.

Từ là một đơn vị ngôn ngữ nhỏ nhất có âm, nghĩa, loại từ vững hoặc chức năng ngữ pháp

⁷⁾ Đúng ngữ pháp nhưng trên thực tế hầu như không ai nói.

Chủ ngữ được hiểu là phần *chủ từ* như trong ngữ pháp truyền thống. *Vị ngữ* là *động từ cộng túc từ*, tức phần còn lại của câu.

Thành tố được hiểu là một từ/cụm từ/chuỗi từ nói được (đúng ngữ pháp).

Các phương pháp phân tích câu thường dùng:

- Phép thế
- Phép thử câu hỏi
- Phép bố trí
- Phép tính lược
- Phép hoán vị
- Phép biến hình

ĐIỂM MỞ

- 1) Thử dùng phương pháp tính lược để tìm những thành tố nhỏ nhất trong câu tiếng Việt. Lấy các đoạn văn trong sách báo.
- 2) Thử dùng phương pháp tính lược để lược từ trong câu tiếng Việt và câu tiếng Anh, tiếng Pháp tương ứng, xem khác biệt thế nào. Phần tử nào của ngôn ngữ nào có thể lược bỏ được ?
- 3) Thử dùng phép hoán vị cho trạng từ, tính từ trong trường hợp thêm nghĩa động từ (đặt trước/sau động từ, đặt ở cuối câu, đặt sau bổ ngữ (túc từ)).
- 4) Thử xem đối với tiếng Việt từ có phải là phần tử cú pháp nhỏ nhất không, hay chữ/tiếng ?
- 5) Thử lấy một câu văn xưa và dùng mọi phương pháp trên thêm bớt biến thành dạng câu tiếng Việt hiện đại. Xem chuỗi từ nào là thành tố (xưa, hiện đại). Cấu trúc thành tố xưa và nay khác biệt ra sao ?
- 6) Thử chuyển một câu văn nói sang câu văn viết và ngược lại. Xem đặc tính của văn nói và văn viết khác nhau chỗ nào. Điểm khác biệt nào tiêu biểu nhất ?
- 7) Thử tìm hiểu thêm về quan hệ chủ động, bị động (bị/được) trong tiếng Việt. Tiếng Việt khác những tiếng Âu châu ở điểm nào ? Tại sao ?

Chương 5

CẤU TRÚC NGỮ ĐOẠN

PHẠM TRÙ NGỮ PHÁP

Khi xét về mặt cấu trúc của những thành phần trong ngôn ngữ, điều trước hết cần được làm rõ là cái gì được coi như phần tử nhỏ nhất đối với một lĩnh vực nhất định (ngữ âm học, hình thái học, cú pháp học, ...). Đối với bộ môn cú pháp, *từ* được xem là phần tử nhỏ nhất theo quan điểm của giới ngôn ngữ học Tây phương.

Từ được hiểu như một đơn vị cú pháp nhỏ nhất, có ý nghĩa, hoặc có loại từ vững, hoặc chức năng ngữ pháp. Từ tham gia cấu tạo, tổ chức làm thành những thành phần lớn hơn (một cụm từ, một chuỗi từ) sao cho nói được, nghĩa là đúng ngữ pháp. Xin nhắc lại bằng vài ví dụ:

- (1) “Ăn” là một từ có ý nghĩa và là một động từ.
- (2) “Năn” là tiếng/chữ không có nghĩa. Nhưng nếu ghép “ăn” với “năn”, ta sẽ được một động từ có ý nghĩa: “ăn năn”. Từ này được ghép bằng 2 tiếng/chữ. Như vậy, từ có thể là sản phẩm của hơn hai tiếng/chữ hoặc của từ.
- (3) “Net” là một từ tiếng Anh có nghĩa là “mạng”(lưới) và là một danh từ.
- (4) “Inter” là một tiền tố dù có nghĩa (“liên”) nhưng không thể đứng một mình. Nếu ghép “Inter” với “net”, ta được “Internet”, một từ có nghĩa là “liên mạng”. Đây là một từ hoàn chỉnh có đủ nghĩa không cần thêm bớt gì nữa cả.

Có một số từ trong tay, chúng ta có thể tạo ra rất nhiều cụm từ, chuỗi từ. Thế nhưng không phải cụm từ, chuỗi từ nào cũng đều nói được. Một cụm từ, chuỗi từ nói được được xem là một *thành tố*. Sở dĩ nói được là nhờ thành tố đúng ngữ pháp. Nói cách khác là trong nó đã tồn tại một quy tắc ngữ pháp. Chúng ta sẽ bàn dưới đây.

Trong thực tế, câu nói được nhiều vô số kể, cho nên những chuỗi từ nói được cũng nhiều vô số kể. Đó là chưa kể trong một chuỗi từ nói được còn có những cụm từ nhỏ hơn nói được, và cứ thế càng có nhiều thành tố nữa.

Hỏi: Có cách nào nắm bắt được tất cả các thành tố không ? Trả lời: Không, ngoài một việc khả thi là xem một thành tố thuộc loại nào.

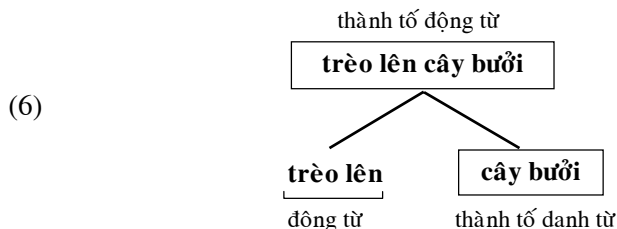
Mỗi từ đều có *loại từ vựng (lexical type)* nhất định như danh từ, động từ, Thành tố cũng thế, cũng có loại mang tính danh từ, có loại mang tính động từ, ... Xem một ví dụ cho dễ hiểu:

- (5) Trèo lên cây bưởi, hái hoa,
Bước xuống vườn cà, hái nụ tầm xuân.
(Ca dao)

Khi nghe chuỗi từ “trèo lên cây bưởi”, chúng ta sẽ có cảm giác rằng, chuỗi từ này diễn tả hành động. Đây là một chuỗi từ nói được, một thành tố mang tính động từ. Tựa vậy, những thành tố khác như “hái hoa”, “bước xuống vườn cà”, “hái nụ tầm xuân” đều diễn tả hành động, nghĩa là cũng mang tính động từ.

Thành tố cũng có loại từ vựng giống như từ. Thành tố chỉ khác từ ở chỗ là nó có thể chứa nhiều từ hơn. Tất nhiên, một thành tố cũng có thể là một từ. Còn nếu bao gồm nhiều từ, thì các từ phải được xếp đặt theo thứ tự sao cho nói được. Người ta nói, nó phải có *cấu trúc cú pháp (syntax structure)*.

Nhìn lại ví dụ bên trên, ta thấy “trèo lên cây bưởi” là một thành tố động từ. Nhưng nó không chỉ chứa động từ “trèo lên” mà còn các từ khác, như “cây”, “bưởi”. Bản thân “cây bưởi” cũng là một thành tố và có tính danh từ. Kết quả, thành tố động từ “trèo lên cây bưởi” là một tổng hợp của hai thành phần: động từ “trèo lên” và thành tố danh từ “cây bưởi”. (Xin lưu ý, do sự khác biệt về tổ chức, từ và thành tố là hai yếu tố khác nhau). Thành tố động từ “trèo lên cây bưởi” có cấu trúc như sau:



Cấu trúc cú pháp của thành tố “trèo lên cây bưởi” có một trật tự nhất định. Nếu đảo ngược trật tự ấy, nghĩa là thay đổi vị trí các thành phần trong thành tố, cấu trúc cú pháp sẽ bị hủy ngay và sẽ khiến chuỗi từ không còn nói được. Chẳng hạn “cây bưởi trèo lên” là một chuỗi từ không nói được. Đối với các thành tố “hái hoa”, “bước xuống vườn cà”, “hái nụ tầm xuân” cũng thế. Ngoài ra, còn một điểm cần để ý nữa là trong thành tố động từ, phần tử *động từ* luôn xuất hiện trước *thành tố danh từ* (xem hình vẽ bên trên).

Những gì vừa được xét sơ chính là công việc phải làm khi xét về **phạm trù ngữ pháp (grammatical category)**. Có thể hiểu một cách ngắn gọn rằng, mục đích của phạm trù ngữ pháp là nhằm *phân loại từ và thành tố*. Cụ thể:

1. Phân loại từ về mặt từ vựng, tức là xác định từ theo **phạm trù từ vựng (lexical category)**,
2. Sau đó, dựa trên loại từ, phân loại thành tố về mặt cú pháp, tức là xác định thành tố theo **phạm trù cú pháp (syntax category)**.

Xét về phạm trù từ vựng là để nhận diện loại từ (danh từ, động từ, ...). Xét về phạm trù cú pháp là để xem những từ đã được nhận diện có mối quan hệ gì với nhau khi tham gia vào việc tạo nên thành tố, và thành tố ấy thuộc loại nào (danh từ, động từ, ...), có cấu trúc ra sao. Chẳng hạn, nếu xét về phạm trù từ vựng thì “trèo lên” là một từ thuộc loại động từ, còn xét về phạm trù cú pháp, thì động từ “trèo lên” là một phần tử cú pháp đã tham gia vào việc tạo nên thành tố “trèo lên cây bưởi”. Phần tử động từ này làm cho cả thành tố cũng mang tính động từ.

Sau đây chúng ta sẽ đi vào chi tiết từng phạm trù một. Trước nhất, hãy xét về phân loại từ, tức phạm trù từ vựng.

PHẠM TRÙ TỪ VỰNG

Từ (word), theo ngôn ngữ học Tây phương, được phân thành nhiều loại: *mạo từ (article)*, *danh từ (noun)*, *đại danh từ nhân vật (personal pronoun)*, *đại danh từ sở hữu (possessive pronoun)*, *đại danh từ chỉ thị (demonstrative pronoun)*, *đại danh từ quan hệ (relative pronoun)*, *đại danh từ nghi vấn (interrogative pronoun)*, *đại danh từ tái quy (reflexive pronoun)*, *đại danh từ hỗ tương (reciprocal pronoun)*, ..., *động từ (verb)*, *tính từ (adjective)*, *tính từ sở hữu (possessive adjective)*, *trạng từ (adverb)*, *liên từ (conjunction)*, *giới từ (preposition)*, *tán thán từ (interjection)*, ...

Ít nhiều gì, ai trong chúng ta cũng đã biết sơ qua một số loại từ kể trên. Thiết nghĩ, chỉ cần hiểu vài loại từ (danh từ, động từ, tính từ, trạng từ, liên từ, giới từ, mạo từ) cũng đã đủ giúp chúng ta hiểu những lý thuyết theo sau, thành thử xin miễn vào chi tiết từng loại từ một. Thay vì vậy, có thể xem thêm định nghĩa trong từ điển ngôn ngữ học của Lewandowski (1990).

Bởi có quá nhiều loại từ, cho nên các nhà ngôn ngữ học phải tìm cách phân loại từ theo từng *lớp tương đương*.

Lớp tương đương (equivalent class) được hiểu là một tiểu tập hợp bao gồm mọi phần tử có cùng tính chất tương đương đã được chọn lọc theo những chuẩn tắc nhất định.

Một ví dụ thực tế:

- *Lớp hoa Lan* là một lớp chỉ chứa các loại hoa được công nhận là thuộc họ Lan: phong lan, mộc lan, ... Hoa hồng, hoa cúc không thuộc họ Lan, nên không được kể là thuộc lớp hoa Lan.
- *Lớp xe cơ giới* chỉ chứa những phương tiện được định nghĩa là xe và có động cơ: xe gắn máy, xe hơi, xe vận tải, ... Xe đạp, xe xích lô đạp không được xếp vào lớp này.

Vài ví dụ *lớp từ tương đương*:

- *Lớp danh từ (noun class)* là một lớp chỉ gồm những từ có tính chất và chức năng danh từ, như danh từ riêng, danh từ chung, đại danh từ nhân vật, đại danh từ quan hệ, đại danh từ sở hữu, đại danh từ nghi vấn, đại danh từ hỗ tương, ...
- *Lớp chỉ định từ (determiner class)* gồm có mạo từ, đại danh từ tái quy, đại danh từ chỉ thị, ...
- *Lớp tính từ (adjective class)* gồm có tính từ và trạng từ.
- *Lớp động từ (verb class)* chỉ có động từ.

Kể ra chi tiết, thì còn nhiều. Trên thực tế, giới ngôn ngữ học Tây phương đã cố thu gọn như sau:

Phạm trù từ vựng (lexical category) được hiểu là tập hợp bao gồm 6 *lớp từ (word class)* sau đây (Robins 1966):

1. Danh từ (noun)
2. Động từ (verb)
3. Tính từ (adjective)
4. Giới từ (preposition)
5. Liên từ (conjunction)
6. Chỉ định từ (determiner)

Lớp từ (word class) được hiểu như một lớp tương đương chứa mọi từ coi như cùng loại và chức năng. (Xem ví dụ vừa nêu)

Những lớp từ kể trên có thể được tóm tắt như sau:

Lớp	Ký hiệu	Ví dụ
Danh từ (danh từ riêng, danh từ chung, đại danh từ nhân vật, đại danh từ sở hữu, đại danh từ nghi vấn, đại danh từ hồ tương, đại danh từ quan hệ, ...)	N	house, dog, I, you, mine, yours, each other, one another, what, who, whom, which, ...
Động từ	V	live, eat, ...
Tính từ (tính từ, trạng từ)	A	my, your, warm, hot, slowly, extremely, ...
Giới từ	P	to, for, on, ...
Liên từ	C	and, or, ...
Chỉ định từ (mạo từ, đại danh từ tái quy, ...)	D	the, a, this, that, myself, yourself, ...
Những loại từ của ngôn ngữ khác.		

Cách phân loại trên thực ra chỉ có giá trị cho những ngôn ngữ Ấn-Âu, đặc biệt là tiếng Anh, chứ không đầy đủ cho mọi ngôn ngữ, chẳng hạn như tiếng Việt. Một ví dụ tiêu biểu:

Cũng là danh từ như tiếng Anh, người Việt nói một “con gà”, một “bức tranh” một “chiếc thuyền”, ... Những từ đặc biệt “con”, “bức”, “chiếc”, ... tiếng Anh

không có. Trong tiếng Việt, loại từ này có thể đi với một từ (danh từ, động từ, tính từ) để thành lập một *thành tố mang tính danh từ* và có thể coi như một thực thể hầu như đếm được. Nếu đứng một mình, loại từ này có thể đại diện cho đối tượng trong thành tố phía trước, nghĩa là đã được nói trước. Loại từ này có khả năng:

1. diễn tả *cá thể người*: ông giám đốc, anh kỹ sư, ...
2. diễn tả *cá thể đồ vật*: cái nhà, cánh cửa, ...
3. diễn tả *cá thể thực vật*: cây cam, trái quýt, ...
4. diễn tả *cá thể thú vật*: con chó, con mèo, ...
5. diễn tả *phái tính nam nữ*: cô dâu, chú rể, ...
6. diễn tả *sự già trẻ*: bà lão, cháu bé, ...
7. diễn tả *tính cao quý, sang hèn, kính trọng, khinh khi*: đáng anh hùng, nhà tư bản, ngài đại sứ, kẻ ăn mày, ...
8. diễn tả *thần linh*: đức Chúa Trời, đức Phật, ...
9. diễn tả *sự kiện*: cuộc biểu tình, sự bất mãn, ...
10. diễn tả *tâm tư*: niềm hy vọng, nỗi nhớ thương, ...
11. diễn tả *số lượng*: bảy trẻ, đàn gà, đám tang, ...
12. diễn tả *khối lượng*: giọt mưa, trái núi, ...
13. diễn tả *tính từng phần*: thửa vườn, mảnh ruộng, ...
14. diễn tả *thứ loại*: loại nghệ thuật đất tiền, thứ văn chương rẻ tiền, ..., vân vân và vân vân ...

Tiếng Anh có **mạo từ bất định (non-specified article)** “a”, tiếng Đức có “ein”, “eine”, tiếng Pháp có “un”, “une”. Tương tự, tiếng Việt có “một”.

Tiếng Anh có **mạo từ xác định (specified article)** “the”, tiếng Đức có “der”, “die”, “das”, tiếng Pháp có “le”, “la”. Tương tự vậy, tiếng Việt có “cái” và còn nhiều nữa.

- (7) “Bà già đi chợ cầu Đông,
Hỏi thăm thầy bói có chồng lợi chăng.
Thầy bói xem quẻ nói rằng,
Lợi thì có lợi nhưng răng không còn”
(Ca dao)

Để diễn tả cụm từ “bà già”, người Đức thường nói “die Alte”. “Die” là mạo từ xác định, giống cái. “Alte” (già) là tính từ. Vậy, “die Alte” xem như tương đương với cách nói “bà già” trong tiếng Việt.

Thế còn “*thầy bói*” thì sao ? Người Anh và kể cả người Đức khó có thể diễn tả được chữ “*thầy*” (*thầy lang, thầy thuốc, thầy võ, thầy đồ, thầy chùa, thầy mo, thầy dùi, thầy số, ...*) bằng một mạo từ trong ngôn ngữ họ. “*Thầy*” theo người Anh, người Đức là một danh từ chỉ một người làm nghề dạy học, như thầy dạy tiếng Anh chẳng hạn (english teacher, Englischer Lehler).

Giả sử (xin nhấn mạnh là giả sử) xem những tiếng dùng chỉ định danh từ như một hình thức “*mạo từ*”, thì tiếng Việt có đến ba loại mạo từ và cách dùng khác nhau. Ta có một tập hợp DET gồm 3 tập hợp con DET₁, DET₂, DET₃:

$$DET = DET_1 \cup DET_2 \cup DET_3$$

Tạm gọi:

- DET là tập hợp tất cả mạo từ. (DET viết tắt của *determiner*)
- DET₁ là mạo từ bậc một (article level 1)
- DET₂ là mạo từ bậc hai (article level 2) và
- DET₃ là mạo từ bậc ba (article level 3).
- DET₁ = { Một }
- DET₂ = { Cái }
- DET₃ = { Áng, Anh, Ánh, Bà, Bác, Bài, Bãi, Bản, Bao, Bận, Bầy, Bề, Bó, Bọn, Bộ, Bờ, Buồng, Bùn, Bức, Cách, Cánh, Cảnh, Cẩn, Cặp, Câu, Cậu, Cây, Chàng, Chị, Chiếc, Chiều, Chồng, Chú, Chùm, Chuối, Chương, Con, Cô, Cỗ, Cú, Cục, Cùm, Cuộc, Cuốn, Dây, Dịp, Dòng, Đám, Đàn, Đạo, Đấng, Đóa, Đoàn, Đoạn, Đôi, Đồ¹⁾, Đồng, Đợt, Đứa, Đức, Em, Gã, Gian, Giọt, Góc, Gối, Hạt, Hòn, Hột, Hộp, Kê, Khẩu, Khóm, Khu, Khúc, Kiện, Kỳ, Làn, Lão, Loại, Loài, Loạt, Lô, Lờ, Lũ, Luồng, Mảnh, Mảng, Mẫu, Mé, Miếng, Món, Mỡ, Mụ, Mùa, Mụn, Nải, Nàng, Nấm, Nén, Nền, Ngọn, Ngôi, Ngụm, Người, Nhà, Nhóm, Niềm, Nổi, Nụ, Ông, Phần, Pho, Quả, Quyển, Sự, Sợi, Tay, Tảng, Tắm, Tấn, Tập, Tên, Thang, Thanh, Thằng, Thầy²⁾, Thể, Thứ, Thửa, Tiếng, Tính, Toán, Tốp, Tờ, Trái, Tụi, Viên, Vĩ, Việc, Vỡ, Vóc, Vững, Xấp, Xó, ... }

¹⁾ không đếm được.

²⁾ đếm được còn tùy ngữ cảnh.

Xin lưu ý, tập hợp trên chưa phải là toàn bộ kho tàng mạo từ tiếng Việt, nhất là mạo từ bậc ba.

Tập hợp DET_1 chứa *mạo từ bất định*, chỉ cái muốn nói thì không xác định, một cái gì đó. Ví dụ “*một làng*”, nhưng “*làng*” nào không rõ. DET_1 tương đương với “*a*” của tiếng Anh (*a village*), “*ein*”, “*eine*” của tiếng Đức (*ein Dorf*).

(8) Có *một làng* nằm bên sông Hương.

Tập hợp DET_2 chứa *mạo từ xác định*, chỉ cái muốn nói thì xác định. Ví dụ “*cái làng*”. DET_2 chỉ tương đương một phần với “*the*” của tiếng Anh (*the village*), “*der*”, “*die*”, “*das*” của tiếng Đức (*das Dorf*).

(9) À ! *Cái làng* anh nói là làng An Vân đấy.

Tập hợp DET_3 cũng bao gồm những *mạo từ xác định*, tương tự DET_2 , chỉ cái muốn nói thì xác định. Ví dụ “*ngôi làng*”. DET_3 là *mạo từ tận cùng* (*terminal article*) vì theo sau đó chỉ còn có *danh từ*, *tính từ*, *động từ* để biến thành một thành tố mang tính *danh từ* có thể ví như một danh từ hầu như đếm được. Hãy xem cách thành lập danh từ của DET_3 qua ba ví dụ dưới đây. Xin lưu ý: Dấu “?” là trường hợp không có hoặc không biết.

A. Ví dụ DET_3 kết hợp với *danh từ*:

Anh lính, Bà hoàng hậu, Bài thơ, Bản nhạc, Bầy con nít, Bọn học trò, Bộ luật, Bức tranh, Cánh cửa, Cảnh đời, Cặp vợ chồng, Câu chuyện, Cây cau, Chàng chiến sĩ, Chị sinh viên, Chiếc lá, Chú học trò, Con bò, Cô y tá, Cục đá, Cuộc đời, Cuốn sách, Đám bèo, Đàn trẻ nhỏ, Đạo luật, Đấng trưởng phu, Đống rác, Đôi lời, Đứa bạn, Đức Phật, Em gái, Gã thủy thủ, Giọt nước, Hạt cát, Hòn đảo, Hột đậu, Kẻ ?, Khẩu súng, Khúc sông, Lão phường trưởng, Loại xe, Lời nhạc, Lũ quỉ, Mảnh vườn, Mảng đời, Mẩu truyện, Miếng cơm, Món cá chiên, Mụ tú bà, Nàng công chúa, Nền văn hóa, Ngôi chùa, Ngọn núi, Người lính, Nhà văn, Nhóm bạn, Niềm ?, Nỗi lòng, Ông vua, Pho tượng, Quả địa cầu, Quyển truyện, Sự đời, Tay anh chị, Tấm hình, Tấn tuồng, Tên tài xế, Thanh gươm, Thằng em, Thứ trái cây, Thửa vườn, Tiếng chuông, Tờ giấy, Trái chuối, Viên Thống đốc, Vị tổng thống, ...

B. Ví dụ DET_3 kết hợp với *động từ*:

Anh bán nước, Bà bán chiếu, Bài ca, Bản phụ họa, Bầy ăn cướp, Bọn cá độ, Bộ ?, Bức vẽ, Cánh quạt, Cảnh sinh hoạt, Cặp ?, Câu nói, Cây ?, Chàng hát dạo, Chị giúp việc, Chiếc ?, Chú phụ rể, Con ở đợ, Cô phục vụ, Cục nợ, Cuộc sống, Cuốn ?, Đám biếu tình, Đàn ăn cướp, Đạo làm người, Đấng cứu thế, Đồng nợ, Đôi ?, Đứa bội tình, Đức ?, Em đứng đường, Gã mua vui, Giọt ?, Hạt ?, Hòn ?, Hột ?, Kể chán đời, Khẩu ?, Khúc hát, Lão ăn mày, Loại tuyên truyền, Lời hứa, Lũ ăn bám, Mảng ?, Mảnh ?, Mẩu ?, Miếng ăn, Món nhậu, Mụ mai mối, Nàng tiếp tân, Nền ?, Ngôi ?, Ngọn ?, Người gây chiến, Nhà lãnh đạo, Nhóm truyền giáo, Niềm thương, Nỗi nhớ, Ông đập xích lô, Pho ?, Quả ?, Quyển ?, Sự bố thí, Tay tham nhũng, Tắm ?, Tấn ?, Tên đâm chém, Thanh ?, Thằng buôn lậu, Thứ ăn chơi, Thừa ?, Tiếng thở than, Tờ báo cáo, Trái cấm, Vị chỉ huy, Viên thông dịch, ...

C. Ví dụ DET₃ kết hợp với *tính từ*:

Anh cả, Bà diên, Bài ?, Bản ?, Bầy trẻ, Bọn nhỏ, Bộ ?, Bức ?, Cánh ?, Cảnh éo le, Cặp ?, Câu ?, Cây ?, Chàng hào hiệp, Chị lớn, Chiếc ?, Chú nhỏ, Con nhỏ, Cô bé, Cục ?, Cuộc vui, Cuốn ?, Đám du côn, Đàn trẻ, Đạo ?, Đấng linh thiêng, Đôi trẻ, Đứa dốt, Đức ?, Em bé, Gã khùng, Giọt sầu, Hạt mềm, Hòn ?, Hột ?, Kể khốn cùng, Khẩu ?, Khúc thần tiên, Lão mù, Loại tốt, Lời ai oán, Lũ lai căng, Mảng tình, Mảnh tình, Mẩu tình ?, Miếng ngon, Món rẻ, Mụ ác độc, Nàng quê, Nền ?, Ngôi ?, Ngọn ?, Người nông cạn, Nhà thông thái, Nhóm trí thức, Niềm vui, Nỗi buồn, Ông già, Pho ?, Quả ?, Quyển ?, Sự lạnh lùng, Tay cứng đầu, Tắm ?, Tấn ?, Tên lưu manh, Thanh ?, Thằng ngu, Thứ thật, Thừa ?, Tiếng xấu, Tờ ?, Trái đắng, Viên ?, Vị thần linh, ...

So với những ngôn ngữ khác như Anh, Pháp, Đức, ... mạo từ tiếng Việt dồi dào là nhờ nhóm bậc 3.

Mạo từ bậc 3 không thể đứng một mình. Loại từ (word type) này không phải là mạo từ bình thường. Thực chất của nó là một danh từ. Nói đúng hơn là một loại danh từ có khả năng thành lập một đơn vị danh từ. Hãy tạm gọi là ***danh từ hình thức (formal noun)***³⁾. Danh từ hình thức có chức năng:

- kết hợp với một từ để thành lập một thành tố mang tính danh từ. Kết quả diễn tả một đối tượng đã được hình thức hóa mà đa số đều đếm được.

³⁾ Có tác giả còn gọi là *phó danh từ* (Nguyễn Kim Thản), *tiền danh từ* (Phan Khôi), *loại từ* (Lê Văn Lý). Trước khi đời sự thống nhất về thuật ngữ ngôn ngữ, chúng tôi tạm gọi là vậy, kể cả các thuật ngữ khác.

- đại diện cả thành tố được thành lập ấy về mặt cú pháp ở một môi trường ngữ cảnh đã được làm rõ phía trước.

(10) Cô ấy vừa ngâm hai *bài thơ* hay quá. Một *bài* của Quang Dũng, một *bài* của Hữu Loan.

Chúng ta biết “*thơ*” là một danh từ cụ thể, thế nhưng không đếm được. “*Thơ*” là một tập hợp rất chung chung. Muốn đếm được, người ta phải giới hạn nó bằng một hình thức, một đơn vị nào đó như một “*bài thơ*”, “*câu thơ*”, “*văng thơ*”, ... Những hình thức này chính là những đối tượng cụ thể cũng giống như những tập hợp con đã được phân loại từ một tập hợp rất lớn và tổng quát. Vì “*thơ*” không đếm được, cho nên không thể nói:

(11) * Cô ấy vừa ngâm *hai thơ* hay quá.

Nhờ đơn vị hóa “*thơ*” bằng hình thức “*bài*”, chúng ta mới có một loại thực thể đếm được, đó là “*bài thơ*”. Đây là khả năng thứ nhất của danh từ hình thức.

Khả năng thứ hai của danh từ hình thức (“*bài*”) là đại diện cho đối tượng (“*bài thơ*”) đã được làm rõ trong môi trường ngữ cảnh phía trước. Nhờ ngữ cảnh “*thơ*” đã xuất hiện ở câu đầu, cho nên ở câu sau, chúng ta mới có thể lược bỏ chữ “*thơ*” bằng cách nói tắt là “*một bài*”.

(12) Một *bài* ~~thơ~~ của Quang Dũng, một *bài* ~~thơ~~ của Hữu Loan.

Giữ chức năng như một danh từ, “*bài*” đại diện cho thực thể “*bài thơ*” để câu sau đừng sai ngữ pháp. Nếu không nhờ ngữ cảnh trong câu, chúng ta không thể nói khờ khạo “*một bài*”. Không ai hiểu “*bài*” đây là bài gì cả.

Bởi vì danh từ hình thức có khả năng đại diện luôn cho thành tố ở môi trường ngữ cảnh đã được làm rõ phía trước, cho nên nó có khả năng làm trung tâm ý nghĩa cho cái thành tố bị tỉnh lược phía sau.

Tương tự danh từ hình thức, mạo từ bậc hai “*cái*” cũng có thể tạo thành một thành tố mang tính danh từ đếm được: “*một cái bàn*”, “*hai cái ghế*”, “*ba cái nhà*”, ... Chính bản thân “*cái*” cũng mang tính chất danh từ. Nếu đứng trước nó là mạo từ bậc nhất “*một*” hoặc số đếm, thì chức năng của nó lại hao hao như một danh từ hình thức.

Sơ lược là vậy. Đây là một đề tài không đơn giản. Tốt nhất là bàn chi tiết ở một chuyên luận khác. Để đi tiếp, hãy xếp danh từ hình thức vào lớp danh từ.

Sau khi đã xét sơ về loại từ vựng của từ, nay hãy xét xem từ có khả năng tạo nên những thành tố có loại cấu trúc cú pháp nào. Đây công việc thuộc *phạm trù cú pháp*.

PHẠM TRÙ CÚ PHÁP

Đã hiểu nôm na, thành tố là một từ, cụm từ, chuỗi từ nói được, nay ta định nghĩa rõ hơn:

Thành tố là một tập hợp từ có cấu trúc cú pháp.

Nếu một tập hợp chỉ chứa một từ, thì từ ấy chính là thành tố ấy. Trường hợp đơn giản này chúng ta không bàn đến. Hãy xét đến trường hợp một thành tố chứa nhiều từ.

Trong một thành tố chứa nhiều từ, chắc chắn sẽ có một từ giữ vai trò *trung tâm ý nghĩa* của cả thành tố. Từ này sẽ quyết định luôn loại từ vựng của thành tố. Ví dụ một cụm từ nói được:

(13) Truyện hay

“*Truyện hay*” là một thành tố. Khi nghe nói đến “*truyện hay*”, lập tức chúng ta sẽ để tâm ngay đến “*truyện*”, bởi đây là thông tin quan trọng nhất trong mọi thông tin được nói ra. Người ta nói, từ này là từ chủ chốt của thành tố, nó làm trung tâm thành tố, nó nói lên đối tượng. Đối tượng ở đây là “*truyện*”. Trong khi đó, từ “*hay*” không quan trọng. Nó chỉ phụ nghĩa thêm để biết truyện thế nào thôi, ví dụ “*truyện hay*”, “*truyện dở*”, “*truyện chương*”, “*truyện tiểu lâm*”, “*truyện tình cảm xã hội*”, “*truyện trinh thám*”, ... Là truyện gì đi nữa, tất cả đều là “*truyện*”. Từ “*truyện*” mang ý nghĩa trung tâm.

“*Truyện*” là một *danh từ*. Nói bao quát hơn, nó thuộc lớp danh từ (noun class). “*Hay*” là một *tính từ*, thuộc lớp tính từ (adjective class). Bởi danh từ “*truyện*” là từ chủ chốt của thành tố, cho nên cả thành tố cũng mang tính chất danh từ. Người ta nói, từ chủ chốt này là *đầu thành tố*.

Có lẽ ví dụ trên đã gợi cho chúng ta nhớ đến mối *quan hệ lệ thuộc* của từ trong thành tố ở chương trước. Đúng vậy. Xét đến ***phạm trù cú pháp (syntax category)*** là xét đến mối *quan hệ lệ thuộc* của từ trong thành tố. Nguyên tắc của quan hệ lệ thuộc khá đơn giản:

“Những từ lệ thuộc vào một *đầu từ vựng* sẽ tạo nên một thành tố”.

Đầu từ vựng (lexical head) theo ngôn ngữ học Tây phương được hiểu như một từ đóng vai trung tâm, đại diện cho một trong 4 phạm trù từ vựng. Nói một cách tổng quát là 4 *lớp từ*:

- 1) danh từ (N)
- 2) động từ (V)
- 3) tính từ (A)
- 4) giới từ (P)

Đầu từ vựng là yếu tố quyết định tính đặc trưng của thành tố. Trong một thành tố, nếu đầu từ vựng thuộc lớp danh từ, thì đó là thành tố danh từ; nếu đầu từ vựng thuộc lớp động từ, thì đó là thành tố động từ; nếu đầu từ vựng thuộc lớp tính từ, thì đó là thành tố tính từ; nếu đầu từ vựng thuộc lớp giới từ, thì đó là thành tố giới từ. Người ta gọi một thành tố được phân loại cụ thể như danh từ, động từ, ... là một ngữ đoạn. Định nghĩa rõ hơn:

Ngữ đoạn (phrase) là một thành tố đã được định loại. Loại thành tố là tên của *đầu từ vựng*.

Sau đây, để dễ giải thích cấu trúc ngữ pháp, chúng ta sẽ dùng mô hình ***cây (tree)***. Trong cây, mỗi ***nút (node)*** được đặt tên bằng một ký hiệu biểu thị từng lớp từ riêng biệt, ví dụ như *tính từ* là A (adjective), *danh từ* là N (noun)⁴⁾, ... Ngoài ra, bởi quan hệ lệ thuộc còn nói lên cái *tôn ti* của cấu trúc, cho nên từng nút một sẽ được diễn tả theo vị trí cao thấp nhất định.

Trong cụm từ “*truyện hay*”, “*truyện*” là danh từ, “*hay*” là tính từ. “*Hay*” lệ thuộc vào “*truyện*”, tức mối quan hệ “*tính từ lệ thuộc vào danh từ*”. Nói tổng quát hơn, *lớp tính từ* lệ thuộc vào *lớp danh từ*. Vị trí tính từ thấp hơn vị trí danh từ. N là đầu từ vựng:

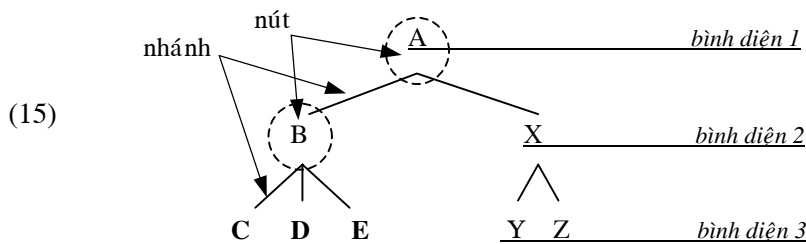
⁴⁾ Theo truyền thống ngôn ngữ học, những ký hiệu viết tắt A, N, V, ... được dùng như những ký hiệu quốc tế. Chúng tôi cũng dùng như vậy.



“*Truyện hay*” là một thành tố thuộc loại *danh từ* (noun, viết tắt: N). Thành tố này được đặt tên là *ngữ đoạn danh từ* (*noun phrase*). Cấu trúc trên mô tả *ngữ đoạn* này.

Thay vì vẽ thô sơ theo kiểu trên, chúng ta vẽ cây theo kiểu khác thường được gọi là ***cây cú pháp*** (*syntax tree*).

Cây cú pháp có *nút* và *nhánh*. Cây có thể có nhiều *phẳng diện* (*level*). Một nút ở phẳng diện trên có thể bao gồm nhiều nút ở phẳng diện dưới. Hai nút (trên, dưới) nối với nhau bằng một nhánh:

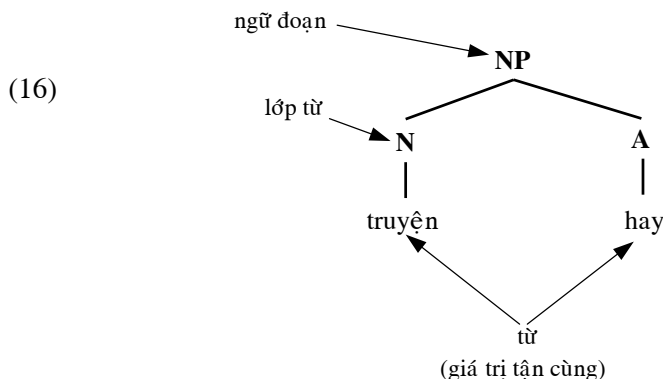


Nút A nằm ở phẳng diện cao nhất. Nút A sinh ra nút B và X. Nút A nối với nút B và X qua hai nhánh. Nút B sinh ra nút C, D, E. Nút X sinh ra Y, Z. Tương tự, các nhánh nối các nút trên/dưới với nhau.

Một nút sinh ra nút khác sẽ được gọi là *ngữ đoạn*. A, B, X là các ngữ đoạn. Ngữ đoạn là một phần tử *không tận cùng* (*non-terminal*)⁵⁾. Ngược lại, một nút không sinh ra nút khác, sẽ được coi như phần tử *tận cùng* (*terminal*)⁶⁾. Giá trị tận cùng là một từ (chương sau sẽ nói chi tiết về cây cú pháp):

⁵⁾ Có tác giả gọi là *không kết thúc*.

⁶⁾ Có tác giả gọi là *kết thúc*.



Nút ngữ đoạn danh từ NP (noun phrase) chứa nút danh từ N và nút tính từ A. Nút N và nút A tận cùng bằng hai từ cụ thể “truyện” và “hay” (giá trị tận cùng).

Nút và nhánh của cây cú pháp mô tả một loại quy tắc được gọi là *quy tắc cấu trúc ngữ đoạn*:

Quy tắc cấu trúc ngữ đoạn (phrase structure rule) là những quy tắc có chức năng xếp đặt những thành tố trong chuỗi từ và những phần tử liên hệ trong thành tố vào từng phạm trù cú pháp nhất định. Quy tắc cấu trúc ngữ đoạn có dạng chung: $A \rightarrow B$

A là ngữ đoạn; B là một hay nhiều ngữ đoạn hoặc một lớp từ.

Xin lưu ý: Trong tin học, quy tắc cấu trúc ngữ đoạn đồng nghĩa với **quy tắc sản sinh (production rule)**, ở đó, mỗi thành tố A, B được xem như một *biến số (variable)* hoặc phần tử không tận cùng, tức còn sản sinh được nữa. Ngược lại, B được coi như một *nguyên tử (atom)* hoặc phần tử tận cùng, nếu B tận cùng bằng một từ cụ thể.

Ví dụ quy tắc cấu trúc ngữ đoạn danh từ NP trong tiếng Việt đã nói trên:

(17) $NP \rightarrow N A$

NP là phần tử không tận cùng vì NP sản sinh ra N và A. Hai phần tử N (noun) và A (adjective) là hai phần tử tận cùng. N tận cùng bằng “truyện”, A tận cùng bằng “hay”.

Dựa vào quy tắc cấu trúc ngữ đoạn, các nhà ngôn ngữ học Tây phương đã dựng nên một mô hình ngữ pháp được gọi là *ngữ pháp cấu trúc ngữ đoạn*.

Ngữ pháp cấu trúc ngữ đoạn (phrase structure grammar, viết tắt *PSG*) của một ngôn ngữ là một tập hợp bao gồm tất cả các quy tắc cấu trúc ngữ đoạn của ngôn ngữ đó.

Trên cơ sở phạm trù từ vựng, ngữ đoạn được chia làm 4 loại:

1. **Ngữ đoạn danh từ** hoặc **danh ngữ (noun phrase, viết tắt *NP*)** là ngữ đoạn có đầu từ vựng thuộc lớp danh từ.
2. **Ngữ đoạn tính từ**⁷⁾ hoặc **tính ngữ (adjective phrase, viết tắt *AP*)** là ngữ đoạn có đầu từ vựng thuộc lớp tính từ.
3. **Ngữ đoạn giới từ** hoặc **giới ngữ (preposition phrase, viết tắt *PP*)** là ngữ đoạn có đầu từ vựng thuộc lớp giới từ.
4. **Ngữ đoạn động từ** hoặc **động ngữ**⁸⁾ (**verb phrase, viết tắt *VP***) là ngữ đoạn có đầu từ vựng thuộc lớp động từ.

Ngữ pháp cấu trúc ngữ đoạn, theo như Chomsky (1977: 127) kể lại trong cuốn *Language and Responsibility*, là một mô hình toán đã được các nhà toán học dựng nên nhằm khảo sát cấu trúc của ngôn ngữ con người theo kiểu toán. Schützenberger là một nhà toán học tiêu biểu trong lĩnh vực này. Mô hình này được coi như một loại **ngữ pháp tạo sinh (Generative Grammar)**. Cả Chomsky cũng đã dự phần vào việc nghiên cứu chung với Schützenberger, song cạnh đó, ông cũng có một lý thuyết ngữ pháp tạo sinh riêng.

Cũng theo lời Chomsky, ngữ pháp tạo sinh của ông thực ra không phải mãi tới năm 1965 mới được phổ biến. Ý tưởng này đã được nhiều người biết đến qua luận án thạc sĩ *Morphophonemics of Modern Hebrew* của ông làm tại đại học Pennsylvania vào năm 1951.

Ngữ pháp tạo sinh (generative grammar) của một ngôn ngữ là một tập hợp quy tắc hiển ngôn có khả năng xác định chuỗi từ nào là câu của ngôn ngữ đó và đồng thời xếp sẵn cho mỗi câu một cách mô tả cấu trúc.

Mô tả cấu trúc (structure description) câu là sự xác định cấu trúc của thành tố bằng cách làm rõ mối quan hệ của các phần tử trong đó theo phạm trù ngữ pháp.

⁷⁾ Có tác giả còn gọi là *ngữ đoạn tính từ và trạng từ*

⁸⁾ Ở các chương sau, chúng tôi gọi là *vị ngữ*.

Tạo sinh được hiểu một cách giản đơn là đẻ ra, sản sinh ra. Chomsky cho rằng ngôn ngữ là một tập hợp (hữu hạn hoặc vô hạn) chứa mọi câu (Chomsky 1957: 13). Tất cả mọi câu đều có cấu trúc nhất định. Mỗi cấu trúc đều mô tả được nhờ quy tắc. Mỗi quy tắc chứa đựng nhiều phần tử ngôn ngữ khác nhau, trong đó một phần tử có thể sản sinh ra những phần tử khác. Với suy nghĩ ấy, ông đã tìm cách lập một mô hình. Khởi đầu bằng một chân lý và một tập hợp những quy tắc chuẩn hình, những kết quả sẽ tuần tự được tạo sinh. Mỗi quy tắc đều có dạng:

$$(18) \quad \varphi \rightarrow \psi$$

φ (thành phần nằm *bên trái* mũi tên) sinh ra ψ (thành phần nằm *bên phải* mũi tên). Chomsky đã giải thích ý tưởng của ông bằng một ví dụ như sau:

Giả sử có một ngôn ngữ chỉ gồm có 3 từ “*Jules*”, “*Chloe*”, “*loves*”. Người nói ngôn ngữ này sẽ biết ngay những tổ hợp nào của 3 từ này thì nói được còn những tổ hợp khác thì không:

- (19) * Jules Chloe loves.
- (20) Jules loves Chloe.
- (21) * loves Jules Chloe.
- (22) Chloe loves Jules.
- (23) * Chloe Jules loves.

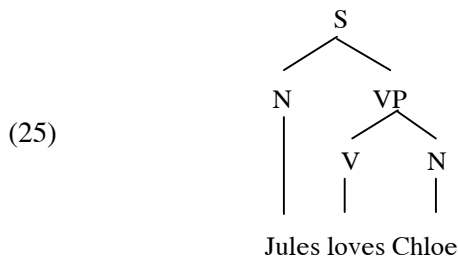
Câu nào nói được sẽ được công nhận là câu của ngôn ngữ ấy. Nhiệm vụ của ngữ pháp tạo sinh là phải cắt nghĩa được những kiến thức hiển ngôn và hàm ngôn của người nói, hoặc sự thông minh của người đọc. Để cắt nghĩa cái kiến thức hiển ngôn, người ta có thể đề nghị một ngữ pháp như sau (Chomsky 1977: 104):

$$(24) \quad \left\{ G \right\} \left| \begin{array}{l} S \rightarrow N VP \\ VP \rightarrow V N \\ N \rightarrow \left\{ \begin{array}{l} Jules \\ Chloe \end{array} \right\} \\ V \rightarrow loves \end{array} \right. \right\} = \left\{ \begin{array}{l} G = \text{grammar} \\ S = \text{sentence, được viết lại là:} \\ \quad \text{danh từ} + \text{ngữ đoạn động từ} \\ \quad (\text{noun } N) \quad (\text{verb phrase } VP) \\ VP = \text{verb phrase, được viết lại là:} \\ \quad \text{động từ} + \text{danh từ} \\ \quad (\text{verb } V) \quad (\text{noun } N) \\ N = \text{noun, được viết lại: } Jules \text{ hoặc } Chloe \\ V = \text{verb, được viết lại: } loves \end{array} \right.$$

Từ sự sản sinh của các quy tắc trên (mà Chomsky gọi là suy diễn (*deduction*)⁹⁾ (nếu S thì N + NP, v.v.)), ta được những kết quả được phái sinh sau:

- Thế S bằng N VP
- Thế VP bằng V N, ta có: N V N
- Thế N thứ nhất bằng “Jules”, ta có: Jules V N
- Thế V bằng “loves”, ta có: Jules loves N
- Thế N thứ hai bằng “Chloe”, ta có: Jules loves Chloe

Kết quả, chuỗi từ đã tận cùng; không còn *phái sinh* (*desire*) được nữa. Người ta có thể vẽ lại *sự phái sinh* (*derivation*) của câu trên dưới dạng cây cú pháp. Cây cú pháp minh họa cấu trúc của những phần tử tận cùng trong câu:



Đó là nguyên tắc của ngữ pháp tạo sinh. Cùng một nguyên tắc như ngữ pháp cấu trúc ngữ đoạn, đây cũng là một mô hình toán. Từng quy tắc và thông tin có thể lập trình được một cách dễ dàng, chẳng hạn như viết một *bộ phân tích* (*parser*) để nhận dạng cấu trúc cú pháp của ngôn ngữ. Trên thực tế, bàn bản ngữ pháp tạo sinh được ứng dụng khá phổ biến trong việc thiết lập *ngôn ngữ lập trình* (Pascal, C/C++, Java, ...), *trình biên dịch* (*compiler*).

Xét về mặt cấu trúc ngôn ngữ, chắc hẳn mỗi ngôn ngữ đều có những quy tắc cấu trúc ngữ đoạn đặc thù. Không ai dám quả quyết, mọi ngôn ngữ trên thế giới chỉ có 4 loại cấu trúc ngữ đoạn căn bản thuộc phạm trù cú pháp như *danh từ*, *động từ*, *giới từ*, *tính từ*, hoặc nói ngược lại là chỉ cần 4 loại này thôi cũng đủ giải thích được tất cả cấu trúc của mọi ngôn ngữ. Cả phạm trù từ vựng cũng thế. Chắc chắn sẽ có một loại từ mà một ngôn ngữ nào đó có nhưng tiếng Anh không có. Bởi vì ngữ pháp tạo sinh nguyên thủy của Chomsky được dựng từ tiếng Do Thái và tiếng Anh, cho nên không thể nói, mọi tính chất được rút ra từ đó đều thỏa

⁹⁾ Chomsky gọi là *quy tắc suy diễn* (*deduction rule*), đồng nghĩa với *quy tắc sản sinh* (*production rule*) thường được nói tới trong tin học.

mọi đặc điểm của mọi ngôn ngữ. Lý thuyết chỉ là nguyên tắc. Muốn ứng dụng nó, mỗi ngôn ngữ phải tự xác định mọi cấu trúc ngữ đoạn của mình. Song cũng chưa chắc là hạt nhân của lý thuyết sẽ thỏa mọi trường hợp. Đây cũng là lý do khiến Chomsky đã nghĩ đến sự *biến hình (transformation)* của cấu trúc ngôn ngữ. Sáng kiến này là một tiền đề cho việc mở rộng lý thuyết ngữ pháp tạo sinh thành *Ngữ pháp biến hình tạo sinh (Generative Transformation Grammar)* của ông sau này.

Sau đây, dựa vào lý thuyết ngữ pháp tạo sinh, chúng ta hãy thử xét sơ qua những cấu trúc ngữ đoạn căn bản của 3 ngôn ngữ Anh, Việt, Đức, và qua đó, thử so sánh sự khác biệt của chúng. Gặp trường hợp các ngôn ngữ có cùng cấu trúc thì đó chỉ là sự tình cờ chứ không hẳn bao giờ cũng thế. Mặt khác, giả như tiếng Việt có đòi hỏi một cấu trúc tương ứng mà tiếng Anh hoặc tiếng Đức không có, thì đó chỉ là một đề nghị mở rộng cho tiếng Việt chứ không hẳn là quy tắc đã được mọi người đồng ý. Trong khuôn khổ bài viết, chúng ta chỉ nói tổng quát về nguyên tắc chứ không vào chi tiết. Khảo sát thêm về cấu trúc ngữ đoạn tiếng Việt là điểm mở dành cho các bạn có quan tâm.

Bằng ví dụ cụ thể, chúng ta sẽ thử xuyên qua 4 loại ngữ đoạn căn bản thuộc phạm trù cú pháp:

- Ngữ đoạn danh từ NP
- Ngữ đoạn tính từ (/trạng từ) AP
- Ngữ đoạn giới từ PP
- Ngữ đoạn động từ VP

1. Ngữ đoạn danh từ NP

Một ngữ đoạn danh từ NP đơn giản nhất chỉ có một danh từ:

(26) NP → N

đọc là ngữ đoạn danh từ NP (noun phrase) sản sinh danh từ N (noun). Cây cú pháp tương ứng:

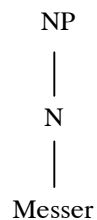


Ví dụ cụ thể với 3 danh từ (N) “*knife*”, “*Messer*”, “*dao*”:

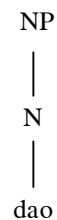
a. NP tiếng Anh



b. NP tiếng Đức



c. NP tiếng Việt



Các cây cú pháp trên mô tả hai quy tắc:

(27) $\text{NP} \rightarrow \text{N}$

(28) $\text{N} \rightarrow \text{knife} \mid \text{Messer} \mid \text{dao}$

Ngữ đoạn danh từ NP sản sinh danh từ N. Danh từ N tận cùng bằng “*knife*” hoặc “*Messer*” hoặc “*dao*”. Ký tự $\mid$ đọc là “*hoặc*” (phép OR của logic).

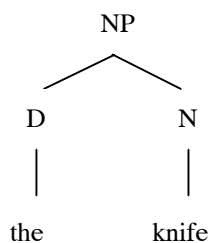
Mở rộng quy tắc $\text{NP} \rightarrow \text{N}$ bằng một *chỉ định từ* (*determiner*, viết tắt: D) đứng trước danh từ N, ta có:

(29) $\text{NP} \rightarrow \text{D N}$

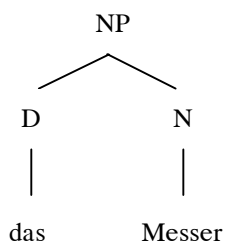
Ví dụ với mạo từ xác định:

(30) $\text{D} \rightarrow \text{the} \mid \text{das}$

a. NP tiếng Anh



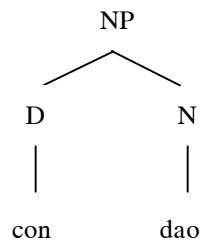
b. NP tiếng Đức



Trong tiếng Việt, nếu hiểu một cách đơn giản như trên, ta có:

(31) $D \rightarrow \text{con}$

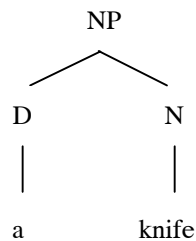
Cây cú pháp:



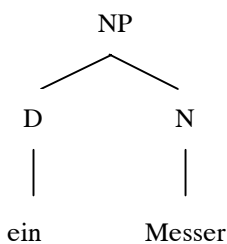
Nhưng gặp trường hợp mạo từ bất định:

(32) $D \rightarrow a \mid \text{ein} \mid \text{một}$

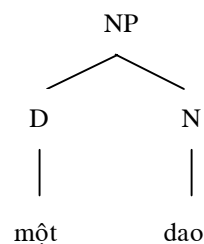
a. NP tiếng Anh



b. NP tiếng Đức



c. NP tiếng Việt



tiếng Anh và tiếng Đức đúng, còn tiếng Việt sẽ gặp vấn đề ngay. Nghe một người Anh nói “*I give you a knife*” hoặc người Đức nói “*Ich gebe Dir ein Messer*”, không sao. Nghe một người Việt nói: “*Tao cho mày một dao*” thì nguy quá.

Thực ra ý nghĩa mà người Anh, người Đức muốn nói ở đây là “*một con dao*”, hoàn toàn vô hại, nhưng người Việt lại hiểu tắt là “*một (nhát) dao*”. Tại sao? Rất dễ hiểu, bởi danh từ “*knife*” trong tiếng Anh, “*Messer*” trong tiếng Đức đã nói lên cá thể “*con dao*” và tất nhiên đếm được, trong khi đó ở tiếng Việt thì chưa. Danh từ “*dao*” có vẻ cụ thể tưởng chừng đếm được, vậy mà không. “*Dao*” là một tập hợp toàn thể. Muốn đếm được, cần phải giới hạn nó bằng cách hình thức hóa, mô tả thêm “*cái gì dao*”. Ví dụ ngay đây, phải thêm “*con*” phía trước “*dao*” thì mới đủ diễn tả cái cá thể của “*dao*”, tức chính là “*con dao*”, và mới

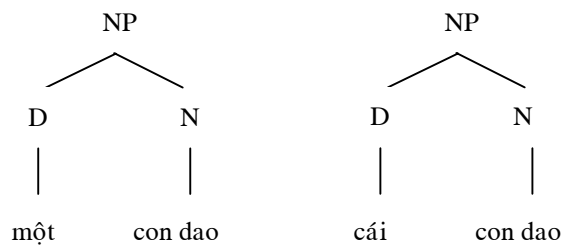
đếm được. Từ đó mới nói được “*một con dao*” và mới đồng nghĩa với tinh thần “*a knife*”, “*ein Messer*” của tiếng Anh, tiếng Đức.

Tương tự như trường hợp “*dao*”, còn có rất nhiều danh từ khác tưởng chừng đếm được, nhưng không: “*bầu*”, “*cua*”, “*cá*”, “*cọp*”, “*mai*”, “*lan*”, “*cúc*”, “*trúc*”, “*ngheo*”, “*sò*”, “*ốc*”, “*hến*”, v.v.. Không ai nói “*bán cho tôi một cua*”, “*tôi muốn mua một cá*” mà phải nói “*bán cho tôi một con cua*”, “*tôi muốn mua một con cá*” theo công thức:

(33) ***danh từ hình thức + danh từ*** = ngữ đoạn danh từ

Nếu đơn giản hóa vấn đề bằng cách xem cụm từ “*con dao*” như một danh từ, thì quy tắc NP → D N của tiếng Anh, Đức cũng dùng được cho tiếng Việt:

(34) D → một | cái



Còn nếu xem “*con dao*” như một ngữ đoạn danh từ, thì bắt buộc phải mở rộng quy tắc NP → D N thành:

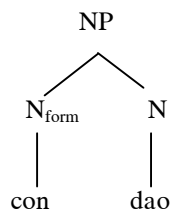
(35) NP → D NP (tiếng Việt)

NP của vế phải diễn tả quy tắc thành lập “*con dao*”:

(36) NP → N_{form} N (tiếng Việt)

(N_{form} ký hiệu cho *danh từ hình thức (formal noun)*),

(37) N_{form} → con (tiếng Việt)



Thế nhưng chưa hết, bởi còn có trường hợp nói được như “*một cái con dao*”, “*một cái con đường*”, “*một cái bức tranh*”, ... (mặc dù hơi rườm rà). Ở đây, “*một*” là *mạo từ bất định (mạo từ bậc 1)*, “*cái*” là *mạo từ xác định (mạo từ bậc 2)* và “*con*”, “*bức*” là *danh từ hình thức*. Vị trí của chúng thỏa hệ thức:

(38) mạo từ bất định > mạo từ xác định > danh từ hình thức

cụ thể là “*một*” đứng trước “*cái*”; “*cái*” đứng trước “*con*”, “*bức*”.

Trong tiếng Anh, không ai nói “*a the*”, hoặc tiếng Đức “*ein der/die/das*”. Thỉnh thoảng người Đức cũng có nói nhưng lại ngược (đối với người Việt):

(39) *die eine Frau* (Dịch 1-1: cái một đàn bà)

(40) *der eine Mann* (Dịch 1-1: cái một đàn ông)

Trong tiếng Việt, không những chỉ có mạo từ “*một*” đứng trước “*cái*” mà nói chung còn cả số đếm “*hai*”, “*ba*”, “*bốn*”, ... Ngược lại không được. Không ai nói “*cái một con dao*” như người Đức. Mạo từ “*cái*” trong hệ thức trên có thể bỏ được nếu danh từ hình thức đã tồn tại: “*một con dao*”, “*một con đường*”, “*một bức tranh*”. “*Con*” và “*bức*” thừa sức cho ta những đơn vị danh từ đếm được mà không cần đến mạo từ “*cái*”.

Vậy, quy tắc $NP \rightarrow D\ NP$ cần được mở rộng cho tiếng Việt bằng cách thêm *mạo từ bất định* “*một*” (ký hiệu: D_{unspec}) và *mạo từ xác định* “*cái*” (ký hiệu: D_{spec}):

(41) $NP \rightarrow D_{unspec} D_{spec} NP$ (tiếng Việt)

Trong tiếng Việt, danh từ ghép cũng là một trường hợp tương tự như trường hợp ghép từ với danh từ hình thức:

- Ghép đôi: “*đường phố*”, “*vợ chồng*”, “*con cái*”, ...
- Ghép ba: “*thảo cầm viên*”, “*hàng không mẫu hạm*” ...
- Ghép tư: “*thủy quân lục chiến*”, ...

Việc thành lập một danh từ dùng được như một *đơn vị từ* là một vấn đề thuộc về bộ môn từ ngữ, chúng ta sẽ không bàn ở đây. Đây là một điểm mở cho bạn đọc có quan tâm tự suy nghĩ và đào sâu thêm. Hãy trở về quy tắc tổng quát:

(42) $NP \rightarrow D N$

Không bắt buộc chỉ định từ D bao giờ cũng phải đi với danh từ. Ta có thể cho D vào ngoặc (), coi như một *lựa chọn* (option):

(43) $NP \rightarrow (D) N$

(44) $NP \rightarrow (D_{\text{unspec}} (D_{\text{spec}})) NP$ (tiếng Việt)

Để đơn giản hóa vấn đề, hãy xem:

- $D_{\text{unspec}} (D_{\text{spec}})$ như D
- NP như N

trong quy tắc $NP \rightarrow D N$ (nếu cần, ta sẽ vào chi tiết sau).

Danh từ thường được bổ nghĩa bằng tính từ A. Quy tắc cần được mở rộng. Tương tự chỉ định từ D, việc dùng tính từ A chỉ là một lựa chọn:

(45) a. Tiếng Việt: $NP \rightarrow (D) N (A)$

b. Tiếng Anh: $NP \rightarrow (D) (A) N$

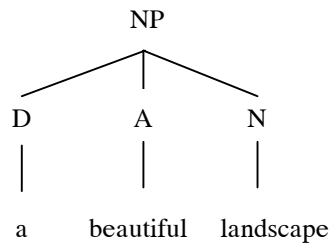
c. Tiếng Đức: $NP \rightarrow (D) (A) N$

a. Một phong cảnh đẹp.

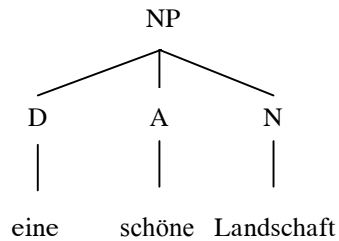
b. A beautiful landscape.

d. Eine schöne Landschaft.

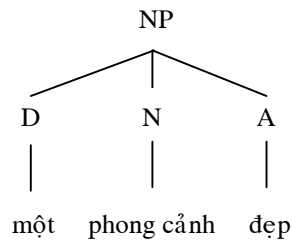
a. NP tiếng Anh:



b. NP tiếng Đức:



c. NP tiếng Việt:



Ngữ đoạn danh từ NP có thể được mở rộng bằng cách cộng thêm một ngữ đoạn giới từ PP và cũng mang tính lựa chọn:

(46) a. Tiếng Việt: NP → (D) NP (A) (PP)

b. Tiếng Anh: NP → (D) (A) N (PP)

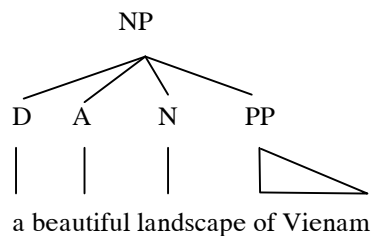
c. Tiếng Đức: NP → (D) (A) N (PP)

a. Một phong cảnh đẹp của Việt Nam.

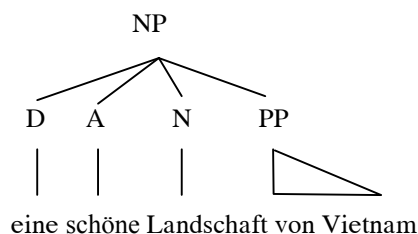
b. A beautiful landscape of Vietnam.

c. Eine schöne Landschaft von Vietnam.

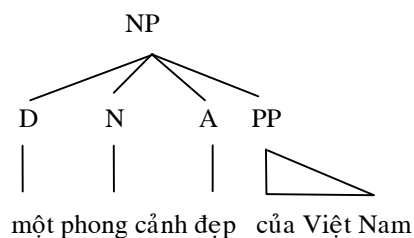
a. NP tiếng Anh:



b. NP tiếng Đức:



c. NP tiếng Việt:



Xin lưu ý: Vẽ hình tam giác có nghĩa là không vào chi tiết.

Thay vì mô tả ngữ đoạn danh từ NP bằng cây cú pháp, người ta có thể đơn giản hóa bằng cách dùng dấu ngoặc [] và *nhãn hiệu* (*label, mark*) như N, V, A, D, PP, ...:

- (47) a. [D Một] [N phong cảnh] [A đẹp] [PP của Việt Nam]
 b. [D A] [N beautiful] [N landscape] [PP of Vietnam]
 c. [D Eine] [N schöne] [N Landschaft] [PP von Vietnam]

Cũng có thể thêm nghĩa một *danh từ* bằng nhiều *tính từ* và *ngữ đoạn giới từ* cùng một lúc:

- (48) Một phong cảnh đẹp, (và) trữ tình trên vùng cao nguyên (của) Việt Nam.

Muốn diễn tả tính lặp lại của quy tắc NP, có thể thêm dấu + sau A và PP:

- (49) a. Tiếng Việt: NP → (D) N (A+) (PP+)
 b. Tiếng Anh: NP → (D) (A+) N (PP+)

c. Tiếng Đức: NP $\rightarrow$ (D) (A+) N (PP+)

Xin lưu ý: Ở đây chỉ nói đến quy tắc tiếng Việt dưới dạng đơn giản. Trong trường hợp dùng liên từ, chúng ta phải mở rộng ngữ đoạn tính từ AP bên dưới.

2. Ngữ đoạn tính từ AP

Mục đích của tính từ là thêm nghĩa cho danh từ.

- (50) a. Một phong cảnh đẹp, trữ tình.
b. Một phong cảnh rất đẹp.

Trong trường hợp a., hai tính từ “*đẹp*” và “*trữ tình*” thêm nghĩa cho danh từ “*phong cảnh*”. Trong trường hợp b., chúng ta dùng một trạng từ và tính từ. Có sự khác biệt:

- *Trường hợp a.*: Cả hai tính từ có vị thế ngang nhau, không thêm nghĩa cho nhau mà cho danh từ “*phong cảnh*”. Một trong hai tính từ đều có thể bỏ được mà vẫn không ảnh hưởng lẫn nhau. Ngoài ra, chuỗi từ vẫn không sai ngữ pháp:

a. Một phong cảnh đẹp, ~~trữ tình~~.
b. Một phong cảnh ~~đẹp~~, trữ tình.
- *Trường hợp b.*: Trạng từ “*rất*” thêm nghĩa cho tính từ “*đẹp*” chứ không cho danh từ. Bỏ “*rất*” thì được nhưng bỏ “*đẹp*” sẽ sai ngữ pháp:

a. Một phong cảnh ~~rất~~ đẹp.
b. * Một phong cảnh rất ~~đẹp~~.

Tương tự ngữ đoạn danh từ NP, một ngữ đoạn tính từ AP đơn giản nhất chỉ sản sinh một tính từ duy nhất:

(51) AP $\rightarrow$ A

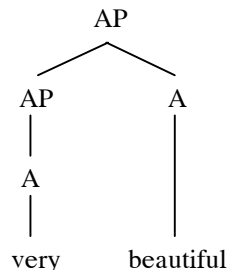
Nếu muốn thêm nghĩa cho tính từ bằng một trạng từ, quy tắc AP $\rightarrow$ A phải được mở rộng thành:

(52) $AP \rightarrow (AP) A$

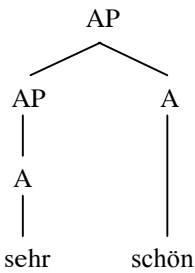
(53) $AP \rightarrow Adv$

(Adv viết tắt của *Adverb* (trạng từ))

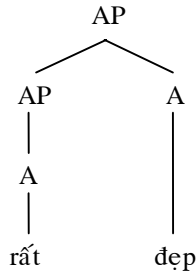
a. AP tiếng Anh



b. AP tiếng Đức



c. AP tiếng Việt



Đã mở rộng quy tắc $AP \rightarrow A$, nay những quy tắc liên hệ khác cũng cần được sửa đổi theo. Cụ thể, ta thay thế A bằng AP:

(54) a. Tiếng Việt: $NP \rightarrow (D) N (AP+) (PP+)$

b. Tiếng Anh: $NP \rightarrow (D) (AP+) N (PP+)$

c. Tiếng Đức: $NP \rightarrow (D) (AP+) N (PP+)$

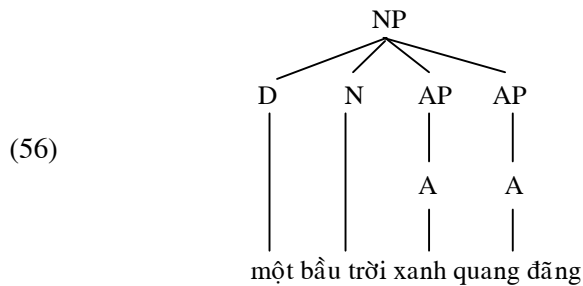
Trạng từ của tiếng Anh được phân biệt với tính từ bằng hậu tố “ly” phía sau. Tiếng Đức không có hình thức này. Nhìn chung, trạng từ trong tiếng Đức giống như hình thức “trợ tính từ”, có thể coi như thuộc *lớp tính từ*.

Trạng từ không thể đứng một mình mà đi đôi với tính từ hoặc động từ. Một ví dụ về mối quan hệ lệ thuộc của tính từ và trạng từ:

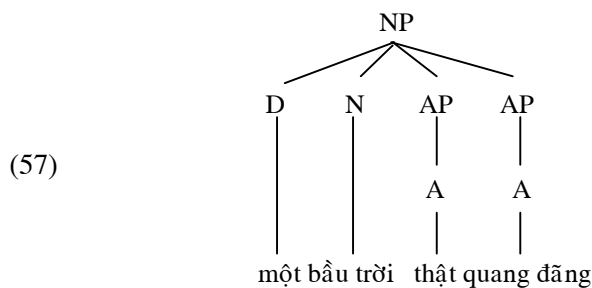
(55) a. Một bầu trời xanh quang đãng.

b. Một bầu trời thật quang đãng.

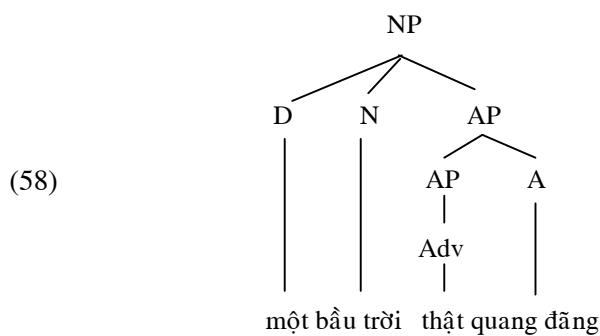
Trong trường hợp a., nên hiểu là “bầu trời vừa xanh vừa quang đãng”. Hai tính từ “xanh” và “quang đãng” độc lập với nhau, nghĩa là không có mối quan hệ lệ thuộc đối với nhau:



Trong trường hợp b., nếu hiểu theo tinh thần của a., thì có nghĩa là “*bầu trời vừa thật, vừa quang đãng*”:



Thế nhưng không ai muốn hiểu vậy. “*Thật*” đây đồng nghĩa với “*rất*”, là một trạng từ có công dụng thêm nghĩa cho tính từ “*đẹp*”. Cho nên phải diễn giải ý nghĩa này theo cấu trúc như sau:



Muốn dùng liên từ, bạn đọc có thể tự mở rộng quy tắc AP. Ví dụ với liên từ “*và*”:

(59) Một bầu trời xanh và quang đãng.

3. Ngữ đoạn giới từ PP

Giả sử nói:

(60) a. Có một phong cảnh.

Ứng dụng nguyên tắc *tịnh tiến*¹⁰⁾ của tiếng Việt, ta có thể đặt câu hỏi:

b. Có một phong cảnh *thế nào* ?

Nếu trả lời bằng một tính từ có lý, chẳng hạn “*đẹp*”, thì tức là đã phụ nghĩa cho ngữ đoạn danh từ “*một phong cảnh*” bằng ngữ đoạn tính từ “*đẹp*”.

c. Có một phong cảnh *đẹp*.

Câu đã hoàn tất, thế nhưng nếu thích hỏi tiếp như “*ở đâu* ?” và trả lời chẳng hạn như “*ở Việt Nam*”, thì cụm từ “*ở Việt Nam*” chính là một ngữ đoạn giới từ, trong đó giới từ “*ở*” đóng vai quyết định danh từ chỉ nơi chốn “*Việt Nam*”.

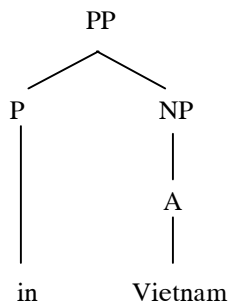
d. Có một phong cảnh đẹp *ở đâu* ?

e. Có một phong cảnh đẹp *ở Việt Nam*

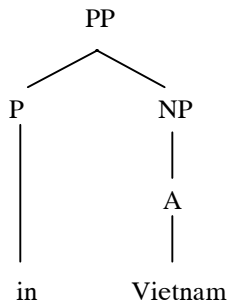
Bằng cách *tịnh tiến* tự nhiên này, chúng ta có quy tắc:

(61) PP → P NP

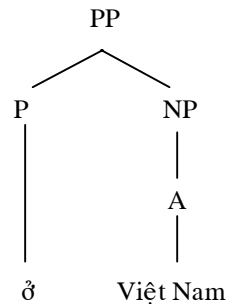
a. PP tiếng Anh



b. PP tiếng Đức



c. PP tiếng Việt



¹⁰⁾ Nguyên tắc *tịnh tiến* không chỉ ra quy tắc đặt câu hỏi mà người sử dụng phải tự quyết định. Câu hỏi tất nhiên mang tính chủ quan.

4. Ngữ đoạn động từ VP

Tương tự ngữ đoạn danh từ NP, một ngữ đoạn động từ VP đơn giản nhất chỉ có một động từ. Đây là trường hợp **động từ bất xuyên tính (intransitive verb)**¹¹⁾. Loại động từ này không hướng tới một túc từ mà chỉ có tác động trong nội bộ chủ từ. Nói theo nghĩa của ngôn ngữ học hiện đại là nó không hướng tới một mục tiêu ngoài chủ ngữ mà chỉ có tác động trong nội bộ chủ ngữ. Ví dụ “khóc”, “chơi”, “ngủ”, “đi”, “đứng”, “ngồi” (“đứa bé khóc”, “đứa bé chơi”, “đứa bé ngủ”), ... Ngược lại là **động từ xuyên tính (transitive verb)**¹²⁾. Loại động từ này có tác động bên ngoài chủ ngữ, tức xuyên qua chủ ngữ và hướng tới một mục tiêu bên ngoài. Ví dụ “nhớ”, “thương”, “có”, “cho”, “mua”, “bán” (“đứa bé nhớ mẹ nó”, “đứa bé thương mẹ nó”, “đứa bé có nhiều đồ chơi”), ... “Mẹ nó”, “đồ chơi” là mục tiêu, có chức năng bổ nghĩa cho “nhớ”, “thương”, “có”.

Quy tắc ngữ đoạn động từ VP chỉ có một động từ V:

(62) VP → V

a. VP tiếng Anh

VP
|
V
|
go

b. VP tiếng Đức

VP
|
V
|
gehen

c. VP tiếng Việt

VP
|
V
|
đi

Ta có thể dùng trạng từ phụ nghĩa thêm cho động từ. Trạng từ được xếp vào lớp tính từ (như đã nói):

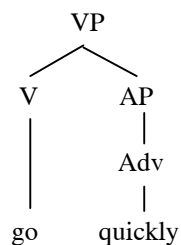
(63) VP → V AP

- a. They go quickly.
- b. Sie gehen schnell.
- c. Họ đi nhanh.

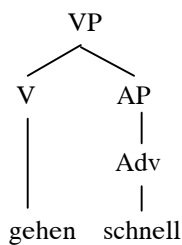
¹¹⁾ Tạm gọi vậy. Có tác giả gọi là *nội động từ*, *tự động từ*.

¹²⁾ Tạm gọi vậy. Có tác giả gọi là *ngoại động từ*, *tha động từ*.

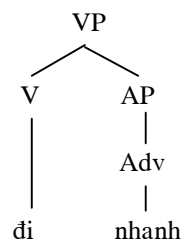
a. VP tiếng Anh



b. VP tiếng Đức



c. VP tiếng Việt



Động từ không nhất thiết phải cần đến trạng từ (xem như lựa chọn):

$$(64) \text{ VP} \rightarrow \text{V (AP)}$$

Trong tiếng Anh và tiếng Việt, trạng từ có thể nằm bên trái động từ:

$$(65) \text{ VP} \rightarrow (\text{AP}) \text{V}$$

- a. Anh ta *thường* nói ...
- b. He *often* speaks ...

Trạng từ cũng có thể xuất hiện cùng lúc bên trái và bên phải động từ:

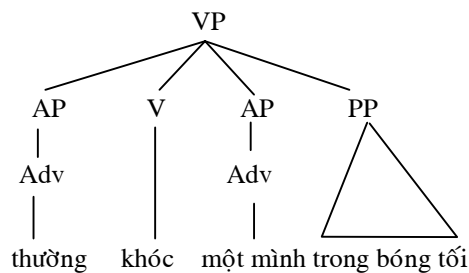
$$(66) \text{ VP} \rightarrow (\text{AP}) \text{V (AP)}$$

- a. Anh ta *thường* nói *chậm rãi* ...
- b. He *often* speaks *slowly* ...

và cũng có thể để cho một ngữ đoạn giới từ PP phụ nghĩa:

$$(67) \text{ VP} \rightarrow (\text{AP}) \text{V (AP)(PP)}$$

Lan *thường* khóc *một mình* *trong bóng tối*.



Có khi trạng từ được lặp lại nhiều lần:

(68) $VP \rightarrow (AP+) V (AP+) (PP)$

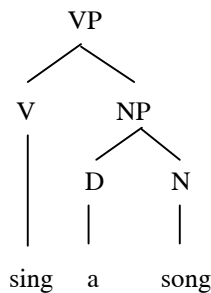
Lan _{[AP+[thường] [hay]]} khóc _{[AP+[lẻ loi] [một mình]]} trong bóng tối.

Đó là những quy tắc được ứng dụng cho *động từ bất xuyên tính*. Đối với *động từ xuyên tính*, quy tắc cần được mở rộng cho bổ ngữ. Bổ ngữ nằm dưới dạng một ngữ đoạn danh từ NP:

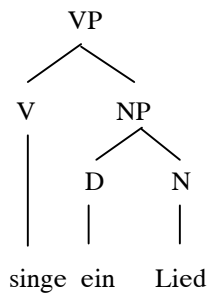
(69) $VP \rightarrow V (NP)$

- a. Tôi hát *một bài ca*.
- b. I sing *a song*.
- c. Ich singe *ein Lied*.

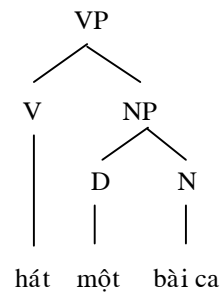
a. VP tiếng Anh



b. VP tiếng Đức



c. VP tiếng Việt



Tiếng Việt không thiếu những cấu trúc như:

(70) $VP \rightarrow V (NP) (AP)$

Bà Hai _{[VP[V ghét] [NP bà Ba] [AP khủng khiếp]]}

(71) $VP \rightarrow (AP+) V (NP)$

Hấn _{[VP[AP thực sự] [V căm thù] [NP ông chủ hấn]]}

Hùng _{[VP[AP rất] [V thương] [NP vợ]]}

Hoặc còn nhiều trường hợp khó hơn nữa mà tiếng Anh, tiếng Đức không có:

- Hấn thương vợ lắm *đấy*.
“*Đấy*” đứng sau trạng từ.
- Hấn thương vợ lắm *mà*.
“*Mà*” đứng sau trạng từ.
- Hấn *mà* thương vợ hấn.
“*Mà*” đứng trước động từ.
- Hấn *có* thương gì vợ hấn.
“*Có*” đứng trước động từ. “*Gì*” đứng sau động từ.
- Hấn thương *đếch* gì vợ hấn.
“*Đếch* gì” đứng sau động từ.

Cách thành lập động từ trong tiếng Việt thường chứa tính mục đích. Cụ thể, người Việt có thể xếp hai ba động từ đi chung:

- Đi + tu → đi tu
- Đi + chơi → đi chơi
- Đi + làm việc → đi làm việc
- Đi + ăn trộm → đi ăn trộm
- Đi + làm ăn + buôn bán → đi làm ăn buôn bán

hoặc động từ thường xen kẽ với danh từ, cho ra một chuỗi từ chứa nhiều động từ:

- Hấn *xoay* sang, *nhíp* mắt, *nhìn* tôi, *cười* không *nói*, *làm* ra *vẻ* *hiểu* hết những điều tôi đang *nghĩ*.
- Hấn *ôm* vợ, *vuốt* ve, *xin* lỗi, *thề* bỏ rượu chè cờ bạc, *ráng* đi *làm*, *kiếm* tiền, *mua* gạo về *nuôi* vợ, *nuôi* con.

Động từ cũng là một đề tài khá thú vị và phức tạp không thua kém gì danh từ hình thức. Đây là một điểm mở dành cho các bạn thích tìm hiểu.

Tóm lại, quy tắc ngữ đoạn VP của tiếng Anh có thể được tổng quát hóa cho hầu hết mọi trường hợp đơn giản:

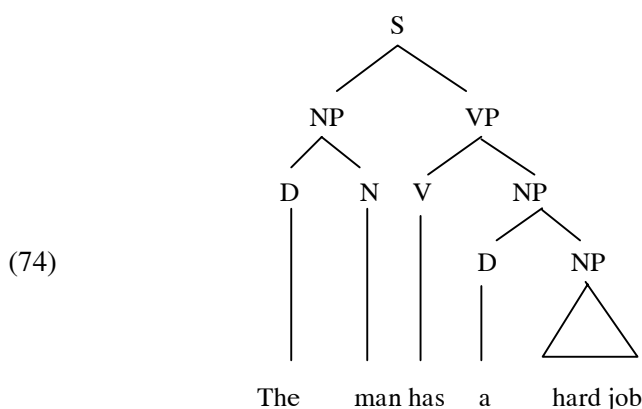
(72) VP → (AP+) V (NP) (AP+) (PP)

5. Câu

Câu, theo ngôn ngữ học cổ truyền, thường được hiểu là một cấu trúc bao gồm *chủ từ*, *động từ* hoặc có thêm *túc từ*. Kể từ đây, ta nên hiểu, câu là một chuỗi từ gồm có hai phần: *chủ ngữ* (*subject*), tức chủ từ, và *vị ngữ* (*predicate phrase*), tức những gì theo sau và bổ nghĩa cho chủ ngữ. Chủ ngữ thường là một ngữ đoạn danh từ NP. Vị ngữ là một ngữ đoạn động từ VP. Quy tắc cấu trúc ngữ đoạn của một câu có hình dạng tiêu biểu:

- (73) $S \rightarrow NP VP$
(S viết tắt của *sentence*, ký hiệu cho câu)

Ví dụ câu “*The man has a hard job*” có thể được phân tích như sau:

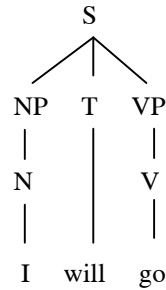


Câu của ngôn ngữ Âu châu thường có ***động từ phụ* (*auxiliary verb*)**¹³⁾. Một ứng dụng là dùng để diễn tả ***thì* (*tense*)**, nếu cần:

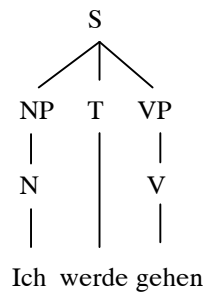
- (75) $S \rightarrow NP (T) VP$
(T viết tắt của *tense*)
- a. I will go.
 - b. Ich werde gehen.
 - c. Tôi sẽ đi.

¹³⁾ Thường gọi là *trợ động từ*. Chúng tôi sửa lại cách dịch cho thống nhất với cách bổ/phụ nghĩa bên phải (*right attributed*) theo kiểu tiếng Việt.

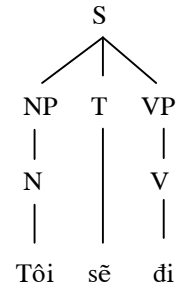
a. S tiếng Anh



b. S tiếng Đức



c. S tiếng Việt



Tiếng Việt cũng có những từ có công dụng diễn tả thì: “*đã*”, “*đang*”, “*sẽ*”, “*sắp*”... Thế nhưng đấy không phải là động từ mà có lẽ chỉ thuần là loại trạng từ chỉ thời gian.

Diễn tả thì bằng động từ phụ là một ứng dụng. Ngoài ra còn những ứng dụng khác như dùng “*should*”, “*would*”, “*could*”, “*must*”, ... trong tiếng Anh. Tương tự, người Việt cũng có thể dùng “*muốn*”, “*phải*”, “*nên*”, ... như động từ.

Dùng động từ phụ thực ra là một hình thức dùng một cụm động từ (động từ phụ + động từ chính) mà ngôn ngữ nào cũng có. Tiếng Việt cũng có những mẫu câu đơn giản như những tiếng Âu châu:

- a. Tôi *không* đi.
- b. Tôi *không muốn* đi.
- c. Tôi *phải đi* ngay.

Hoặc những câu phức tạp dùng nhiều động từ:

- d. Tôi *không biết phải đi* đâu.
- e. Tôi *chưa biết có nên đi* không.
- f. Tôi còn *đang phân vân không biết có nên đi* không.

Trên thực tế còn rất nhiều trường hợp phức tạp khác. Tiêu biểu là những câu phức tạp chứa *câu nhỏ* hơn. Ví dụ câu “*Tôi biết tôi không thể nào đi được*”. Câu này chứa câu “*tôi không thể nào đi được*”:

[_S tôi không thể nào đi được]

Thực chất của câu “tôi không thể nào đi được” là một **mệnh đề phụ** (*subordinate clause*) (thường gọi là câu phụ), trong tiếng Anh được đánh dấu bằng “that”, trong tiếng Đức bằng “daß”, trong tiếng Việt bằng “rằng” (tình cờ giống vậy ?):

- a. I know that I can't go.
- b. Ich weiß, daß ich nicht gehen kann.
- c. Tôi biết rằng tôi không thể nào đi được.

“That”, “daß”, “rằng” được xem như những **tác tử bổ ngữ hóa** (*complementizer*) (ký hiệu bằng C):

- a. [s I know [s' [C that] [s I can't go]]]
 - b. [s Ich weiß [s' [C daß] [s ich nicht gehen kann]]]
 - c. [s Tôi biết [s' [C rằng] [s tôi không thể nào đi được]]]
- (S' ký hiệu cho mệnh đề phụ)

Vì lý do này, giới ngôn ngữ học Âu châu đã đưa ra một quy tắc đặc biệt:

$$(76) \quad S' \rightarrow (C) S$$

Mệnh đề phụ S' giữ vai trò bổ ngữ. Có trường hợp mệnh đề phụ được chuyển sang vị trí chủ ngữ, một hình thức **chuyển vị** (*movement*).

- a. [s [s' That I can't go] is my problem]
- b. [s [s' Daß ich nicht gehen kann] ist mein Problem]
- c. [s [s' Rằng tôi không thể nào đi được] là vấn đề của tôi]

Câu tiếng Việt trên hiếm thấy sử dụng ngoài đời hiện tại. Không rõ cấu trúc của nó có bị lai ngôn ngữ Âu châu hay không, cần xét kỹ lại. Thế nhưng cấu trúc “Rằng ...” rất thường gặp trong Kiểu của Nguyễn Du.

Đối với trường hợp chuyển vị, quy tắc S dành cho tiếng Anh phải được mở rộng:

$$(77) \quad S \rightarrow NP (T) VP$$

$$(78) \quad S \rightarrow S' (T) VP$$

Viết gọn thành:

(79) $S \rightarrow \{NP \mid S'\} (T) VP$

($\{NP \mid S'\}$ đọc là hoặc NP hoặc S' , tức chỉ được chọn một trong hai trường hợp).

(80) $VP \rightarrow (AP+) V (NP) (PP+) (AP+)$

(81) $VP \rightarrow (AP+) V (S') (PP+) (AP+)$

Viết gọn thành:

(82) $VP \rightarrow (AP+) V (\{NP \mid S'\}) (PP+) (AP+)$

Xét cho cùng, những gì vừa được trình bày mới chỉ là vài quy tắc cấu trúc ngữ đoạn được đưa ra nhằm giải thích ngữ pháp. Trên thực tế, vấn đề không đơn giản thế, nhất là đối với tiếng Việt. Có thể nói, nếu giải thích tiếng Việt bằng ngữ pháp cấu trúc ngữ đoạn, chúng ta phải đương đầu với nhiều quy tắc phức tạp mà tiếng Anh không có. Rất nhức đầu. Cả tiếng Đức cũng không dễ như tiếng Anh, nhưng đầu sao vẫn còn đỡ nhức đầu hơn. Dụng ý đưa tiếng Anh, tiếng Đức vào làm ví dụ so sánh là để cho thấy điều đó.

Khách quan mà nói, người Việt sử dụng ngôn ngữ khá uyển chuyển. Đây là một trong những nguyên do làm cho cấu trúc ngữ pháp tiếng Việt quá đa dạng. Sự đa dạng không những chỉ nằm trên bình diện cú pháp mà còn trên bình diện ngữ nghĩa. Trong tiếng Việt còn có những trường hợp hành ngôn mà tưởng chừng như một loại *ngôn ngữ vũ trụ* (*universal language*), khó giải thích vô cùng. Nó xuất hiện khắp nơi qua ca dao, tục ngữ, thành ngữ, thơ, văn từ bình dân cho tới bác học. Tính ngữ dụng phóng khoáng cũng có nguyên do của nó, đặc biệt là trong thể giới văn chương. Có nhiều trường hợp, nó đã không tuân thủ quy tắc ngữ pháp như tuân thủ thi pháp vì lý do “luật bất thành văn”, hoặc làm ngữ pháp phức tạp thêm, hoặc còn đẻ ra các kiểu ngữ pháp mới, lạ kỳ. Phi logic hay không, không dám nói, nhưng rõ ràng những kiểu dùng từ ngữ, ngữ pháp ấy đôi khi mờ mờ ảo ảo, lung linh như chính tâm hồn nó vậy.

(83) “Bà già đi chợ cầu Đông,
Hỏi thăm thầy bói có chồng *lợi* chăng.
Thầy bói xem quẻ nói rằng,
Lợi thì có *lợi* nhưng rằng không còn”.
(Ca dao)

“*Lợi*” ở dòng 2 được hiểu là động từ, nhảy sang dòng cuối, “*lợi*” biến thành danh từ.

- (84) “Thằng Bờm có cái quạt mo.
Phú ông xin đổi ba bò, chín trâu”.
(Ca dao)

Câu đầu nghe đúng, qua câu thứ nhì, ngữ pháp sai ngay. Đúng ngữ pháp, không ai nói “ba bò”, “chín trâu” mà phải nói là “ba con bò”, “chín con trâu”. Rõ ràng đây là một bằng chứng cho thấy, văn chương đã có lý do miễn phục tùng ngữ pháp vì phải tôn trọng thi pháp (“lục bát”).

Tóm lại, *phạm trù ngữ pháp* bao gồm hai phạm trù: *phạm trù từ vựng* và *phạm trù cú pháp*. Phạm trù từ vựng ấn định tên gọi của từng lớp từ riêng biệt. Mỗi lớp từ chứa những từ tương cận. Phạm trù cú pháp ấn định tên gọi của một thành tố. Tên gọi của thành tố là tên của đầu từ vựng đại diện cho nó. Ngữ đoạn là một thành tố có tên.

	Phạm trù ngữ pháp (grammatical category)	Ký hiệu
Phạm trù từ vựng (lexical category)	Danh từ (danh từ riêng, danh từ chung, đại danh từ nhân vật, đại danh từ sở hữu, đại danh từ nghi vấn, đại danh từ hô tương, đại danh từ quan hệ, ...)	N
	Động từ	V
	Tính từ (tính từ, trạng từ)	A
	Giới từ	C
	Liên từ	P
	Chỉ định từ (mạo từ, đại danh từ tái quy, ...)	D
	Phần mở rộng cho những ngôn ngữ khác	v.v.
Phạm trù cú pháp (syntax category)	Câu	S
	Ngữ đoạn danh từ	NP
	Ngữ đoạn tính từ	AP
	Ngữ đoạn giới từ	PP
	Ngữ đoạn động từ	VP
	Phần mở rộng cho những ngôn ngữ khác	v.v.

VÀI VẤN ĐỀ TRONG NGÔN NGỮ

1. Tính hồi quy

Mệnh đề phụ thường dẫn đến một vấn đề, đó là tính ***hồi quy (recursive)*** (còn gọi là *đệ quy*). Quy tắc hồi quy được hiểu là quy tắc lại ứng dụng chính nó và cứ thế tiếp tục chạy vòng vo trong cái vòng vô tận (*endless loop*). Tính hồi quy xuất hiện trong tiếng Anh qua quy tắc $S \rightarrow S' VP$, rồi $S' \rightarrow (C) S$, rồi $S \rightarrow S' VP$, rồi $S' \rightarrow (C) S \dots$:

- (85) The dog that chased the cat that chased the mouse ...
(Con chó mà đuổi con mèo mà đuổi con chuột ...)

Tuy vậy, cần phân biệt tính hồi quy với tính lặp lại liên tục cấu trúc *động từ + danh từ* trong tiếng Việt:

- (86) $S \rightarrow NP VP +$

- (87) $VP \rightarrow V N$

- (88) Con mèo mà [_{VP}trèo cây cau]
[_{VP}Hỏi thăm chú chuột] [_{VP}đi đâu] vắng nhà.
[_{NP}Chú chuột] [_{VP}đi chợ] đường xa,
[_{VP}Mua mắm] [_{VP}mua muối] [_{VP}giỗ cha chú mèo]
(Ca dao)

- (89) Ông ta nhờ tôi dạy cô con gái học đàn.

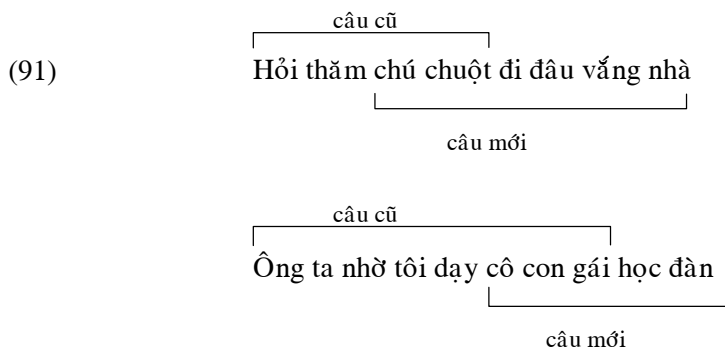
Hoặc muốn nói dài hơn, người ta cũng có thể lập câu đại để như sau:

- (90) Nó nghe Bả kể Người ta đồn Cổ nói Chả nghi Chỉ nói Ông báo công an
Mụ là ... vân vân và vân vân.

Đó là hiện tượng bổ ngữ làm chủ ngữ cho một vị từ (động từ) kế tiếp. Vị từ này hướng tới một bổ ngữ, rồi bổ ngữ ấy lại trở thành chủ ngữ của một vị từ mới tiếp theo, ...

Ở ví dụ đầu, “*hỏi thăm*” là vị từ hướng tới bổ ngữ “*chú chuột*”. Tiếp theo, “*chú chuột*” lại trở thành chủ ngữ của vị từ “*đi*”.

Ở ví dụ cuối, trong đoạn “*Nó nghe Bà*”, “*nó*” là chủ ngữ, “*nghe*” là vị từ, “*bà*” là bổ ngữ. Sau đó, “*bà*” lại làm chủ ngữ cho “*kể*”. “*Kể*” hướng tới bổ ngữ là “*người ta*”. Sau đó “*người ta*” lại biến thành chủ ngữ, ... Nói theo ngôn ngữ học truyền thống là túc từ của câu trước lại có khả năng đi với động từ để tạo thành câu mới.



Điểm đáng chú ý nơi cách hành ngôn này là câu tiếp tục được mở rộng theo nguyên tắc tịnh tiến:

Hỏi: Ông ta nhờ ai ?
 Trả lời: ... nhờ tôi
 Hỏi: ... làm gì ?
 Trả lời: ... dạy
 Hỏi: ... dạy ai ?
 Trả lời: ... dạy cô con gái
 Hỏi: ... làm gì ?
 Trả lời: ... học
 Hỏi: ... học gì ?
 Trả lời: ... học đàn.

2. Tính đa nghĩa

Một câu nói trong ngôn ngữ tự nhiên không phải bao giờ cũng chỉ có một nghĩa duy nhất mà có thể chứa tính **đa nghĩa (ambiguity)**. Đây là hiện tượng một câu có thể được hiểu theo hai nghĩa với hai cấu trúc ngữ pháp khác nhau. Xem thử một câu:

(92) a. Anh họa sĩ vẽ cô gái khỏe thân quá đẹp.

Câu này có thể hiểu theo hai nghĩa, hoặc:

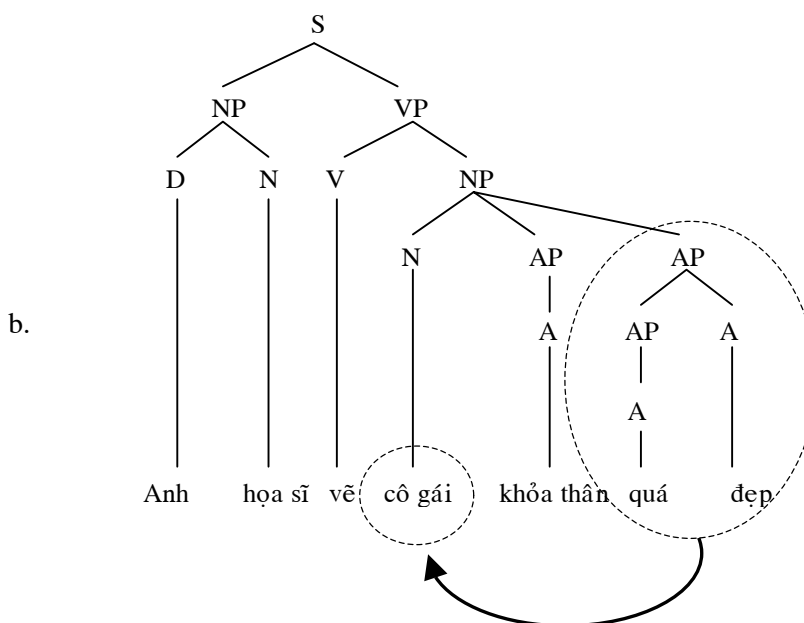
b. Cô gái khỏe thân quá đẹp.

hoặc chưa chắc gì cô ta đã đẹp mà là:

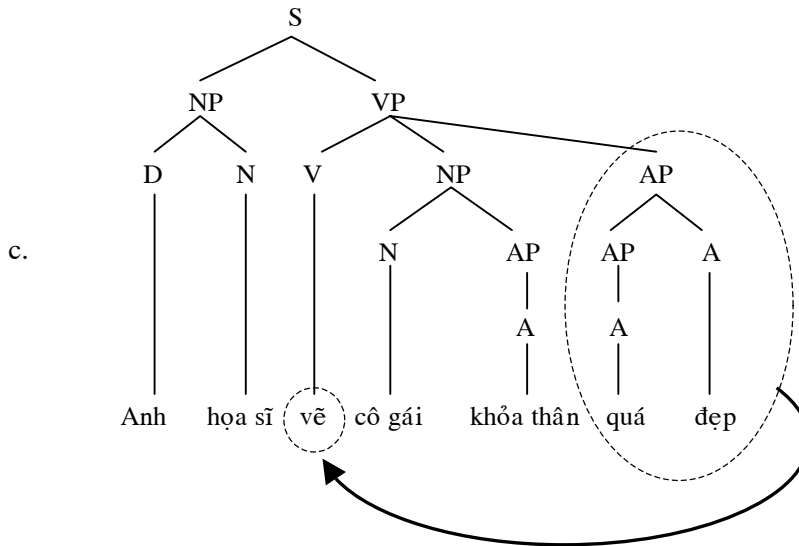
c. Chàng họa sĩ vẽ quá đẹp.

Vẽ ra cây cú pháp, ta có hai cấu trúc khác nhau:

Ở trường hợp b., ngữ đoạn tính từ AP “*quá đẹp*” phụ nghĩa cho “*cô gái*”:



Ở trường hợp c., “*quá đẹp*” phụ nghĩa cho “*vẽ*”:



Cấu trúc ngữ pháp trên là một trong những cấu trúc tiêu biểu nhất của tiếng Việt chứa tính đa nghĩa. Ở cấu trúc này, người ta không biết phụ ngữ được sử dụng cho phần tử nào. Một ví dụ khác:

(93) Ê ! Anh họa sĩ vẽ cô gái khỏa thân kìa.

Trong câu này, chắc gì cô gái đang khỏa thân mà là anh họa sĩ thì sao ? Nếu thực vậy, thì nên thêm dấu phẩy vào để tách rời hai hành động “vẽ” và “khỏa thân” của anh ta:

(94) Ê ! Anh họa sĩ vẽ cô gái, khỏa thân kìa.

Có điều viết còn phẩy được, nhưng nói thì hơi ... kệt.

TÓM LƯỢC

- 1) Vị trí từ trong câu quyết định tính đúng sai ngữ pháp. Vị trí từ đúng, câu sẽ đúng. Vị trí từ sai, câu sẽ sai.
- 2) Từ có quan hệ lệ thuộc lẫn nhau. Từ cho ra cụm từ. Cụm từ nói được là một thành tố.

- 3) Một thành tố có tên danh từ, động từ, tính từ, giới từ được gọi là một ngữ đoạn.
- 4) Có 4 loại ngữ đoạn cơ bản (theo ngôn ngữ học Âu châu,): ngữ đoạn danh từ NP (noun phrase), ngữ đoạn động từ VP (verb phrase), ngữ đoạn tính từ AP (adjective phrase), ngữ đoạn giới từ PP (prepositional phrase).
- 5) Quy tắc cấu trúc ngữ đoạn mô tả ngữ đoạn. Quy tắc có thể được mô tả bằng cây cú pháp. (Xem chi tiết ở chương sau)
- 6) Nhiều ngữ đoạn tạo thành câu. Cấu trúc cơ bản: $S \rightarrow NP VP$
- 7) Mục đích của cú pháp là xếp đặt các từ theo quy tắc nhất định để tạo nên thành tố và câu.
- 8) Ngữ pháp là một tập hợp gồm tất cả các quy tắc diễn tả vị trí và mối quan hệ lệ thuộc của các từ do cú pháp tìm ra.
- 9) Mô hình ngữ pháp được trình bày bên trên là mô hình *ngữ pháp cấu trúc ngữ đoạn* (Phrase Structure Grammar). Loại ngữ pháp này là căn bản của *ngữ pháp tạo sinh* (Generative Grammar). Sáng kiến của ngữ pháp tạo sinh là dùng quy tắc ngữ pháp để tạo ra ngôn ngữ. Từ “*tạo sinh*” đã bắt nguồn từ ý nghĩa ấy.
- 10) Một câu có thể chứa tính hồi quy hoặc tính đa nghĩa.

NHẬN XÉT VÀ ĐIỂM MỞ

Sau khi đã khảo sát sơ về ngữ pháp tạo sinh căn bản, có lẽ ai cũng nhận thấy vài điểm yếu, điểm mạnh của mô hình này đối với tiếng Việt. Có những điểm mở mà chúng ta có thể khảo sát thêm. Thử để ý đến những điểm sau đây:

- 1) Ngữ pháp tạo sinh căn bản có thể giải thích được những cấu trúc nhỏ, tương đối ít phức tạp. Đối với các cấu trúc phức tạp, căn bản này có thích hợp không. Thử xét các trường hợp dùng những tiếng tạm gọi là “tiếng đệm” (đó, mà, chứ, nhé, ...): “*Đi đâu đó*”, “*Ăn đi mà*”, “*Ngủ đi chứ*”, ...
- 2) Ngữ pháp tạo sinh căn bản có thích hợp cho tính tịnh tiến của tiếng Việt không ?
- 3) Có thể dùng sáng kiến của quy tắc cấu trúc ngữ đoạn để lập trình. Thử dựng một bộ phân tích (parser) nho nhỏ để phân tích các cấu trúc đơn giản tiếng Việt. Xem thử mức độ ứng dụng tin học của ngữ pháp này tối ưu ra sao.
- 4) Qua phân tích trên, chúng ta đã nhận thấy vài khác biệt giữa cấu trúc tiếng Việt và cấu trúc các ngôn ngữ khác, như Anh, Đức. Thử liệt kê thật nhiều cấu trúc tiếng Việt và so sánh với tiếng Anh, tiếng Đức, tiếng Pháp

để xem tiếng Việt bị các tiếng ấy ảnh hưởng đến độ nào. Tại sao các ngoại ngữ ấy khó bị lược bỏ tự do như tiếng Việt ? Trong thơ tự do Việt Nam, người làm thơ thường lược bỏ từ hoặc ngắt đoạn câu. Thử phân tích ngữ đoạn, thành tố, xem trường hợp nào hữu lý, trường hợp nào sai, thành tố sau có quan hệ gì với thành tố trước và có hợp lý không ? Nếu không, phải lược bỏ, ngắt đoạn trong câu thơ như thế nào mà cấu trúc ngữ pháp không bị tổn thương ?

- 5) Tìm tính đa nghĩa trong tiếng Việt trong trường hợp một phụ ngữ (tính từ/trạng từ) nằm ở cuối câu. Tính đa nghĩa trong tiếng Việt còn xuất hiện dưới dạng nào, tại sao ?
- 6) Chúng ta đã giải thích được phần nào về cấu trúc đề diễn, thế nhưng hoàn toàn không đề cập gì đến phạm trù ngữ nghĩa. Thử tìm những ví dụ sử dụng sai tính đề diễn của tiếng Việt trong văn viết (sách vở, phương tiện thông tin đại chúng, ...)

Chương 6

CÂY CÚ PHÁP

Một trong những hình thức được dùng để mô tả *cấu trúc cú pháp* (*syntax structure*) là *cây cú pháp* (*syntax tree*). Cây cú pháp minh họa mọi quy tắc được gọi chung là *quy tắc sản sinh* (*production rule*) hoặc *P-marker* (theo cách gọi của Chomsky 1965). Cây cú pháp được dùng rất nhiều trong ngôn ngữ học. Cả trong tin học, mô hình này cũng được dùng để mô tả *ngôn ngữ hình thức* (*formal language*), ví dụ, rất thường gặp trong bộ môn *thiết lập trình biên dịch* (*compiler design*). Sau đây chúng ta sẽ tìm hiểu chi tiết hơn về cây cú pháp theo nghĩa của ngôn ngữ học. Việc làm này cần thiết cho những lý thuyết về sau.

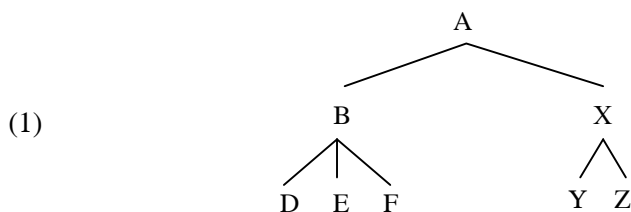
CÂY CÚ PHÁP

1. Hình học của cây

Quy tắc cấu trúc ngữ đoạn có thể được mô tả bằng *cây cú pháp*. Vị trí và tôn ti của mỗi từ được diễn tả qua hình ảnh những nút và cành nhánh. Hình ảnh này minh họa *hình học của cây* (*geometry of tree*):

- Mỗi cây khác loại nhờ tên gọi.
- Nhờ phân loại cây, mối quan hệ của cây trong toàn thể cấu trúc sẽ được làm sáng tỏ.

Thử quan sát một cây cú pháp trừu tượng:



Cây trên mô tả các quy tắc:

- (2) a. $A \rightarrow B X$
 b. $B \rightarrow D E F$
 c. $X \rightarrow Y Z$

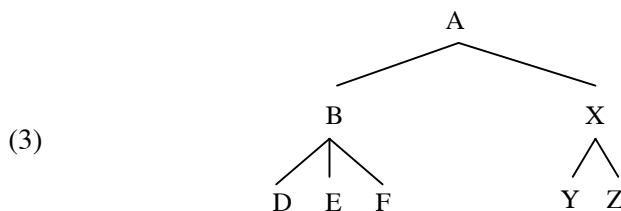
Một vài định nghĩa:

- i. **Nhánh (branch)** là đường thẳng nối hai điểm trong cây.
- ii. **Nút (node)** là những điểm nối trong cây: A, B, ...
- iii. **Nhãn hiệu (label)** là tên của mỗi nút: A, B, ...
- iv. **Nút đỉnh (root node)**¹⁾ là nút cao nhất. Trên nó không có nhánh nào nối với nó: A.
- v. **Nút tận cùng (terminal node)** là nút thấp nhất, dưới nó không còn nhánh nào đâm ra: D, E, F, Y, Z.
- vi. **Nút không tận cùng (non-terminal)** là nút có nhánh nối tới nút khác: A, B, X.

2. Tính chi phối

Đó là mặt hình học. Ngoài ra, còn phải xét về mối quan hệ chiều dọc giữa những nút với nhau. Mối quan hệ chiều dọc được hiểu như tôn ti của cấu trúc, trong đó có nút nằm trên cao kiểm soát nút nằm phía dưới; có nút nằm ngang nhau. Mối quan hệ này nói lên **tính chi phối (dominance)** hoặc **tính ưu thế** của từng nút trong cây.

Tính chi phối (dominance): Một nút A *chi phối* một nút B nếu và chỉ nếu vị thế A trong cây cao hơn vị thế B. Nhánh chỉ nối từ A trở xuống. Tất nhiên A chi phối tất cả các nút nằm dưới B và nối với B.



¹⁾ Root (tiếng Anh) là rễ. Lý do không dịch root node là nút rễ vì nút rễ không rõ nghĩa hơn nút đỉnh.

Trong cây, A chi phối tất cả các nút nằm phía dưới. B chi phối D, E, F. Nút X chi phối Y, Z, nhưng không chi phối D, E, F vì không có nhánh nào từ X nối xuống D, E, F.

Tính chi phối được Higginbotham (1985) định nghĩa một cách hình thức như sau:

Chân lý tính chi phối (Axioms of Dominance):

Với hệ thức $x \leq y$, có nghĩa x chi phối y, ta có:

- a) $x \leq x$
- b) nếu $x \leq y \leq z$, thì $x \leq z$
- c) nếu $x \leq y \leq x$, thì $x = y$
- d) nếu $x \leq z$ và $y \leq z$, thì hoặc $x \leq y$ hoặc $y \leq x$ hoặc $x = y$

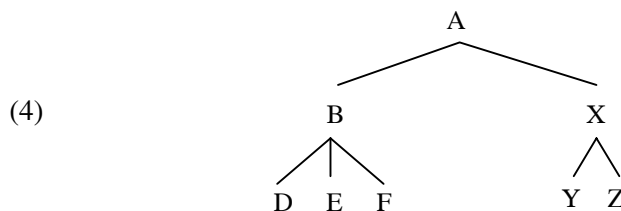
Định nghĩa *nút* dựa trên tính chi phối:

- i. ***Nút mẹ (mother node):*** A là *nút mẹ* của B nếu A chi phối B.
- ii. ***Nút con [gái]²⁾ (daughter node):* B là *nút con* của A nếu B bị A chi phối.**
- iii. ***Nút đỉnh:*** A là *nút đỉnh* nếu A chi phối tất cả mọi nút, nhưng không bị nút nào chi phối.
- iv. ***Nút tận cùng:*** Y là *nút tận cùng* nếu Y không chi phối bất kỳ nút nào.
- v. ***Nút không tận cùng:*** X là *nút không tận cùng* nếu X chi phối những nút khác.

Ngoài ra:

Chi phối trực tiếp (immediately dominate): A chi phối *trực tiếp* B nếu A trực tiếp nối B nằm dưới.

Trong cây:



²⁾ Đọc là “con”. Dịch sát nghĩa là “con gái”.

- A chi phối trực tiếp B và X.
- B chi phối trực tiếp D, E, F.
- X chi phối trực tiếp Y, Z.

Dựa vào tính chi phối trực tiếp, các *nút* còn được định nghĩa nghiêm ngặt hơn:

- Nút mẹ*: A là nút mẹ của B nếu A chi phối trực tiếp B.
- Nút con [gái] (daughter)*: B là nút con của A nếu B bị A chi phối trực tiếp.
- Nút chị [em] (sisters)*: Hai nút là chị em nếu có cùng nút mẹ.

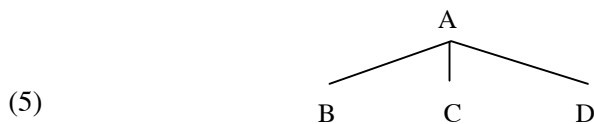
Ví dụ trong cây:

- A là nút mẹ của B, X. Ngược lại, B, X là nút con của A.
- B là nút mẹ của D, E, F. Ngược lại, D, E, F là nút con của B.
- X là nút mẹ của Y, Z. Ngược lại, Y, Z là nút con của X.
- B, X là 2 nút chị em.
- D, E, F là 3 nút chị em.
- Y, Z là 2 nút chị em.

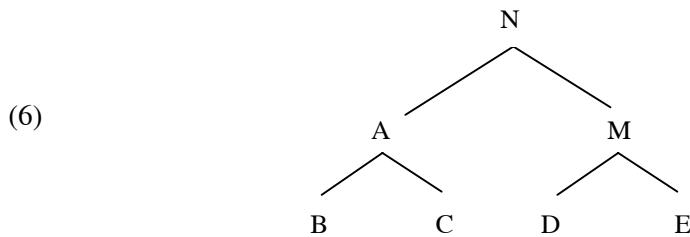
Thêm một góc nhìn khác về tính chi phối:

Sự chi phối tận mức (exhaustive domination): Một nút A *chi phối tận mức* một tập hợp nút $\{B, C, \dots\}$ nếu A chi phối trực tiếp tất cả mọi phần tử trong tập hợp và không một nút nào bị A chi phối trực tiếp mà không nằm trong tập hợp này.

Trường hợp 1: Giả sử ta có một tập hợp $SET = \{B, C, D\}$. Sự kiện “A chi phối trực tiếp tất cả các phần tử” chính là mối quan hệ cấu trúc:



Trường hợp 2: Giả sử mặt khác:



A chỉ chi phối trực tiếp B, C chứ không chi phối trực tiếp D. Vậy, A không được kể là có khả năng chi phối tận mức tập hợp {B, C, D}.

Dựa vào những định nghĩa hình thức trên, một thành tố sẽ được định nghĩa chặt chẽ như sau:

Thành tố (constituent): Thành tố là một tập hợp bao gồm mọi nút bị một nút chi phối tận mức.

Trong cây (6), thành tố N bao gồm A, M. Thành tố A bao gồm B, C. Thành tố M bao gồm D, E.

Cần phân biệt giữa hai khái niệm “thành tố” và “thành tố của”. Hai khái niệm này dễ gây lẫn lộn, thế nhưng tiếc rằng lý thuyết dùng chúng, cho nên không thể bỏ qua:

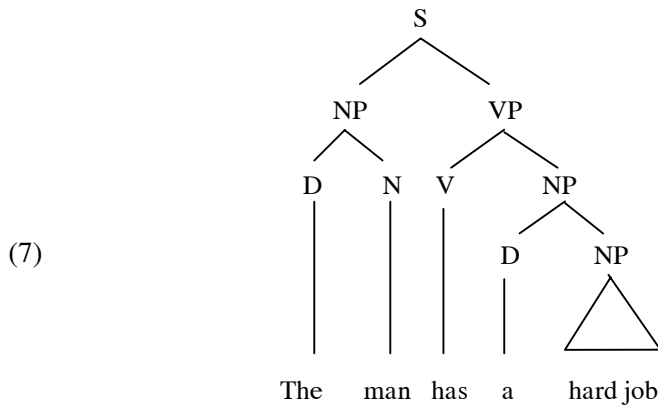
Thành tố của (constituent of): A là thành tố của B nếu và chỉ nếu B chi phối A.

Trong cây (6), B và C là thành tố của A. Trong khi đó, D, E là thành tố của M. Cả hai nút A và M là thành tố của N. Những nút B, C, D, E đều là thành tố của N.

Tất nhiên còn có mối quan hệ chặt chẽ hơn:

Thành tố trực tiếp của (immediate constituent of): B là thành tố trực tiếp của A nếu và chỉ nếu A chi phối trực tiếp B.

Một ví dụ cụ thể:



- S chỉ phối trực tiếp NP và VP, cho nên ở vị thế đó, NP và VP là *thành tố trực tiếp của S*
- NP là một *thành tố* vì D, N là *thành tố của NP*

3. Thế tiên vị

Tính chi phối diễn tả tôn ti của cấu trúc cây; đại để là có trên có dưới. Thế nhưng không phải chỉ xét đến quan hệ cấu trúc theo *chiều dọc* là đủ mà còn phải xét theo *chiều ngang*, xem vị thế từng nút trong cấu trúc nằm trước sau thế nào. Đó là *thế tiên vị (precedence)*.

Một ví dụ: Quy tắc $S \rightarrow NP VP$ không những chỉ nói rằng:

- (a) câu S bao gồm hai ngữ đoạn NP và VP

mà còn nói:

- (b) ngữ đoạn danh từ NP *đứng trước* ngữ đoạn động từ VP

Dĩ nhiên chúng ta đang xét đến những ngôn ngữ viết từ trái qua phải như trường hợp tiếng Việt. Mỗi quan hệ “*cái gì đứng trước*” được hiểu là *thế tiên vị*, nói đơn giản là “*đứng trước*”. Định nghĩa một cách hiển ngôn:

Thế tiên vị (precedence): Nút A ở *thế tiên vị* đối với nút B nếu và chỉ nếu

1. A ở bên trái B, không có trường hợp A vừa chi phối B, B chi phối A
và
2. Mỗi nút chi phối A hoặc B

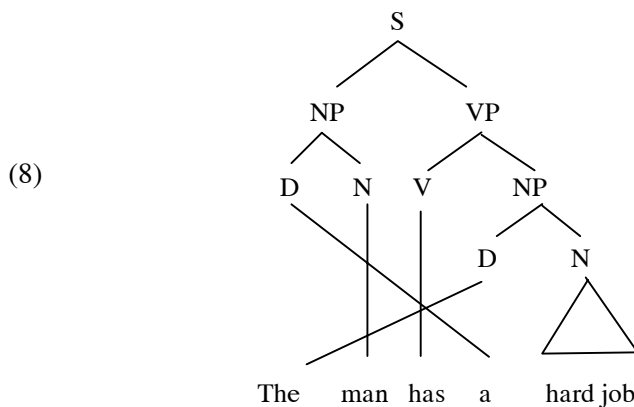
Thế tiền vị được Higginbotham (1985) định nghĩa một cách hình thức như sau:

Chân lý thế tiền vị (Axioms of Precedence):

Với phép $\neg$ có nghĩa *đứng trước* (hay ở thế thế tiền vị đối với), ta có:

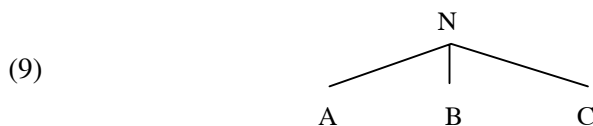
1. nếu $x \neg y$, thì NOT ($y \neg x$)
2. nếu $x \neg y \neg z$, thì $x \neg z$
3. nếu $x \neg y$ hoặc $y \neg x$, thì NOT ($x \leq y$) và NOT ($y \leq x$)
4. $x \neg y$, đối với tất cả phần tử tận cùng u, v , nếu $x \leq u$ và $y \leq v$, suy ra $u \neg v$

Sở dĩ cần định nghĩa thế tiền vị thật chặt chẽ là để tránh tình trạng hiểu không rõ, tương đối. Chẳng hạn không thể hiểu và vẽ một cấu trúc cây treó ngược như này:



Tương tự tính chi phối trực tiếp, ta có:

Thế tiền vị trực tiếp (immediate precedence): Nút A ở thế tiền vị trực tiếp đối với nút B, nếu không có nút nào đứng giữa A và B.



Nút A ở thế tiền vị trực tiếp đối với nút B, tức đứng trực tiếp trước nút B.

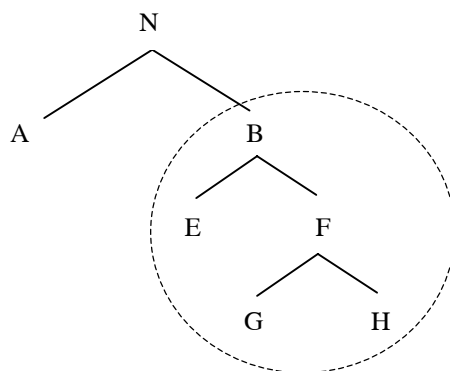
4. Lệnh-C

Trong tất cả các quan hệ cấu trúc, có lẽ loại quan hệ được gọi là ***lệnh-C (c-command)***³⁾ là quan hệ trừu tượng nhất. Cái tên nghe hơi lạ kỳ nhưng lại là một quan hệ hữu ích, đặc biệt đối với ***lý thuyết ràng buộc (binding theory)*** mà chúng ta sẽ gặp sau này. Vậy, ***lệnh-C*** là gì ?

Hiểu nôm na thì lệnh-C là mối quan hệ giữa một nút A đối với một nút chị em B và những nút con của B. Định nghĩa kỹ hơn:

Lệnh-C (c-command): Nút A có mối quan hệ ***lệnh-C*** với nút chị em B, nếu một nút chi phối A, cũng sẽ chi phối B và không có trường hợp A hoặc B chi phối nút khác.

(10)



- Cây diễn tả mối quan hệ lệnh-C giữa nút A và nút chị em B. Mối quan hệ này còn có giá trị đối với tất cả các nút nằm dưới B.

³⁾ Cái tên ***Lệnh C (c-command)*** đã làm không ít người thắc mắc về nguồn gốc của nó. Người ta chỉ biết, chữ “*command*” được vay mượn từ môn ngữ pháp truyền thống. Có người đồn rằng, “*c*” trong chữ “*c-command*” là ký hiệu viết tắt từ chữ “*constituency*”. Có người thì bảo, bởi trong ngữ pháp tạo sinh có hai khái niệm hàm nghĩa mệnh lệnh: *Command* và *Kommand*, để phân biệt “*C*” với “*K*”, người ta gọi là “*c-command*” với “*c*” thay vì với “*k*”. Thế nhưng đó chỉ là truyền thuyết trong giới ngôn ngữ học.

Mối quan hệ lệnh-C còn được phân biệt giữa *đối xứng (symetric)* và *bất đối xứng (asymetric)*:

Lệnh-C đối xứng (symetric c-command) được hiểu là hai nút chị em có cùng mối quan hệ với nhau.

Lệnh-C bất đối xứng (asymetric c-command) được hiểu là chỉ có một nút có mối quan hệ với nút chị em kia, ngược lại thì không.

Tính nội bộ của quan hệ cấu trúc đã được làm rõ dựa vào hai định nghĩa “*tính chi phối trực tiếp*” và “*thế tiên vị trực tiếp*”. Tương tự vậy còn có mối quan hệ trực tiếp của lệnh-C. Quan hệ này được gọi là **sự chi phối (government)**. Quan hệ chi phối rất phổ biến trong những công trình nghiên cứu tiếp theo về ngữ pháp tạo sinh vào khoảng những năm 1981-1991. Trên thực tế còn có cả lý thuyết riêng được gọi là **lý thuyết chi phối và ràng buộc (Government and Binding Theory)** (s.s. Chomsky 1993), (Haegeman 1991). Phần này sẽ được bàn chi tiết ở các chương sau.

CÁCH VẼ CÂY CÚ PHÁP

Quy tắc cấu trúc ngữ đoạn và cây cú pháp đã được trình bày chi tiết. Dùng cây cú pháp, các nhà ngôn ngữ học có thể mô tả quy tắc cấu trúc ngữ đoạn của một chuỗi từ. Chúng ta cũng có thể làm giống vậy bằng hai phương pháp dưới đây.

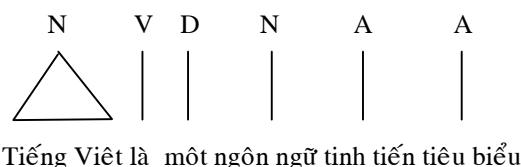
1. Vẽ cây từ dưới lên trên

Vẽ cây từ dưới lên trên (*bottom up*) có lẽ là cách dễ nhất để mô tả cấu trúc một chuỗi từ. Theo nguyên tắc tổng quát, trước nhất định loại từng từ một, xem nó thuộc lớp từ nào. Kế tiếp là đi tìm những cụm từ nói được. Sau đó xét mối quan hệ của từ rồi xác định một loại ngữ đoạn cho cụm từ ấy. Lặp lại thao tác này cho đến khi đạt tới nút đỉnh (tức đã hoàn tất một câu). Cụ thể làm từng bước như sau:

Bước 1: Viết ra chuỗi từ muốn phân tích, ví dụ câu:

“*Tiếng Việt là một ngôn ngữ tịnh tiến tiêu biểu*”.

Trước khi định loại từ, cũng xin lưu ý, để đơn giản hóa vấn đề, chúng ta sẽ không phân tích chi tiết một danh từ ghép mà hãy xem nó như một đơn vị danh từ. Ví dụ “tiếng Việt” là một đơn vị danh từ. Trong trường hợp này, chúng ta vẽ hình tam giác thay vì một đường thẳng. Giả sử xem “là” là động từ:



So sánh ký hiệu viết tắt trong bảng tóm tắt:

	Phạm trù ngữ pháp (grammatical category)	Ký hiệu
Phạm trù từ vựng (lexical category)	Danh từ (danh từ riêng, danh từ chung, đại danh từ nhân vật, đại danh từ sở hữu, đại danh từ nghi vấn, đại danh từ hô tương, đại danh từ tương đối, ...)	N
	Động từ	V
	Tính từ (tính từ, trạng từ)	A
	Giới từ	C
	Liên từ	P
	Chỉ định từ (mạo từ, đại danh từ chỉ định, đại danh từ tái quy, , ...)	D
Phạm trù cú pháp (syntax category)	Câu	S
	Ngữ đoạn danh từ	NP
	Ngữ đoạn tính từ	AP
	Ngữ đoạn giới từ	PP
	Ngữ đoạn động từ	VP

Bước 2: Tìm tất cả mọi cụm từ nói được và xét kỹ mối quan hệ của những từ ấy đối với nhau. Kế tiếp, xác định loại ngữ đoạn, tốt nhất, nên theo thứ tự:

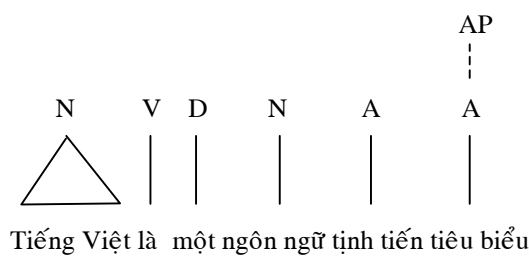
1. Ngữ đoạn tính từ AP.
2. Ngữ đoạn danh từ NP và ngữ đoạn giới từ PP.

3. Ngữ đoạn động từ VP.
4. Câu.

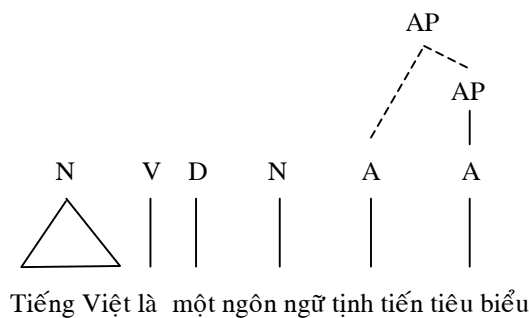
Lặp lại bước này cho đến khi đạt được nút đỉnh S.

Bắt đầu: “*Tiếng Việt*” đã được coi là danh từ, không cần xét. Hãy xét cụm từ nói được: “*một ngôn ngữ tịnh tiến tiêu biểu*”.

Trước nhất, xét về mối quan hệ của tính từ. Giữa hai tính từ “*tĩnh tiến*” và “*tiêu biểu*”, thì tính từ “*tĩnh tiến*” quan trọng hơn, bởi ý muốn nhấn mạnh “*tiếng Việt là một ngôn ngữ tĩnh tiến*”. Tính từ “*tiêu biểu*” chỉ là một phụ ngữ có chức năng phụ nghĩa cho tính từ “*tĩnh tiến*”. Do đó vị thế của “*tiêu biểu*” phải thấp hơn vị thế của “*tĩnh tiến*”. Để diễn tả “*tiêu biểu*” có vị thế thấp hơn một bậc, chúng ta chêm thêm một nhãn hiệu AP vào mà phải là ngữ đoạn tính từ AP để đại diện cho một tính từ:

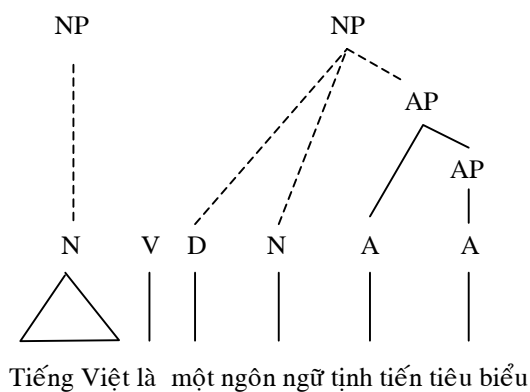


Gộp hai nút A và AP, ta được một ngữ đoạn tính từ AP:

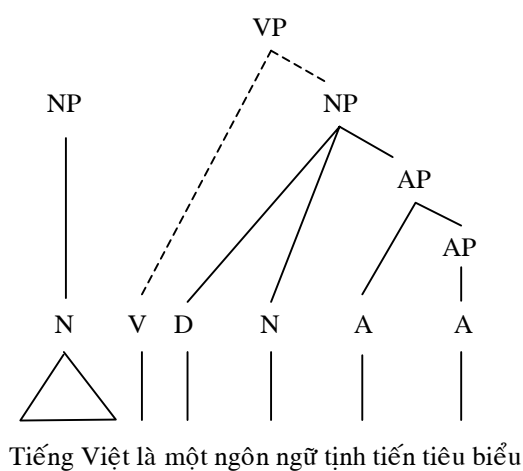


Đã xong loại ngữ đoạn tính từ AP, ta xét qua loại ngữ đoạn danh từ NP. Xem cum từ “*ngôn ngữ tịnh tiến tiêu biểu*” như một ngữ đoạn danh từ NP là hợp lý.

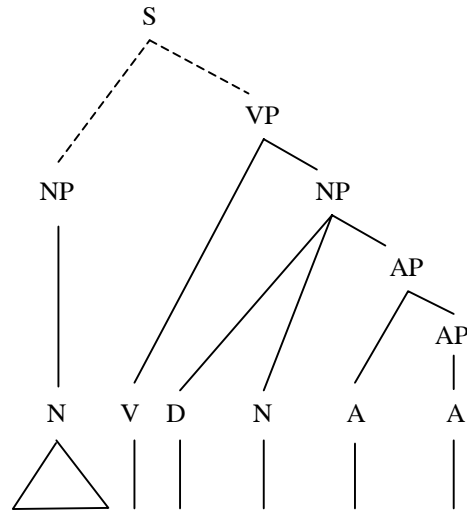
Tốt nữa, gộp luôn mạo từ “*một*” vào cho gọn. Đừng quên, nút danh từ N “*Tiếng Việt*” cũng là một NP:



Vẽ xong phần NP, chỉ còn sót lại nút động từ V (“*là*”). Động từ V kết hợp với NP cho ra một ngữ đoạn động từ VP ($VP \rightarrow V\ NP$):



Xong 3 loại ngữ đoạn AP, NP và VP, ta chỉ còn hai nút hở: NP và VP. Nối hai nút này lại là xong:



Tiếng Việt là một ngôn ngữ tinh tiến tiêu biểu

Cây cú pháp hoàn tất. Tất cả các nút từ trên xuống dưới diễn tả những quy tắc cấu trúc ngữ đoạn sau:

- 1) $S \rightarrow NP VP$
- 2) $VP \rightarrow V NP$
- 3) $NP \rightarrow N$
- 4) $NP \rightarrow D N AP$
- 5) $AP \rightarrow A AP$
- 6) $AP \rightarrow A$

Kiểm chứng lại, ta sẽ thấy, các quy tắc được tìm thấy ở đây đều trùng lặp với những quy tắc mà ta đã làm quen.

Phương pháp vẽ cây cấu trúc từ dưới lên trên còn được xem như một phương pháp tìm quy tắc ngữ pháp. Trên thực tế, phương pháp này được ứng dụng rất phổ biến trong công đoạn *phân tích cú pháp* (*syntax analyse*) của *trình biên dịch* (*compiler*) hoặc của các *hệ thống nhận diện ngôn ngữ* (*language recognizing system* (LRS)). (LRS là một hệ thống được thiết lập trên cơ sở ứng dụng phương pháp *trí thông minh nhân tạo* (*artificial intelligence*) nhằm nhận diện, hiểu và nói được tiếng người. LRS là một lĩnh vực khoa học thuộc ngành *ngôn ngữ học điện toán* (*computational linguistics*). Trí thông minh nhân tạo là một lĩnh vực khoa học thuộc ngành *tin học* (*informatics*)).

2. Vẽ cây từ trên xuống dưới

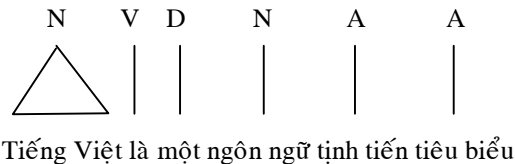
Ngược với cách vẽ cây từ dưới lên trên (bottom up) là vẽ cây *từ trên xuống dưới* (top down). Các nhà ngôn ngữ học chuyên về cú pháp thường ưa chuộng phương pháp này vì nó nhanh hơn. Nhưng bù lại, nó đòi hỏi người sử dụng phải nắm mọi quy tắc ngữ pháp; nếu không cẩn thận, rất dễ sai. Cụ thể ở phương pháp này, chúng ta sẽ bắt đầu từ nút đỉnh S, từ đó khai triển ra những ngữ đoạn mới. Có những ngữ đoạn mới, phải khéo léo chọn lựa những quy tắc thích hợp kế tiếp sao cho đúng đường dẫn đến sự trùng lặp với mẫu chuỗi từ hoặc mẫu câu cho sẵn. Nút nào mới sinh ra, phải thanh toán ngay nút đó. Lặp lại chu trình này cho tới khi nào nối được tất cả các nút tận cùng.

Tương tự cách vẽ từ dưới lên trên, hãy thử nghiệm bằng hai bước cụ thể sau đây:

Bước 1: Viết ra chuỗi từ muốn phân tích, ví dụ câu:

“Tiếng Việt là một ngôn ngữ tịnh tiến tiêu biểu”.

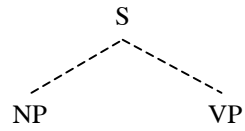
rồi định loại:



Bước 2: Bắt đầu từ quy tắc cao nhất $S \rightarrow NP VP$ rồi lần lần đi xuống với điều kiện là trong đầu đã thuộc lòng những quy tắc:

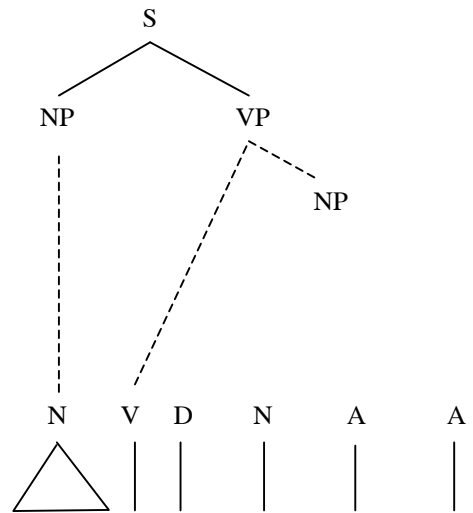
- 1) $S \rightarrow NP VP$
- 2) $VP \rightarrow V NP$
- 3) $NP \rightarrow N$
- 4) $NP \rightarrow D N AP$
- 5) $AP \rightarrow A AP$
- 6) $AP \rightarrow A$

Theo quy tắc $S \rightarrow NP VP$, ta có:



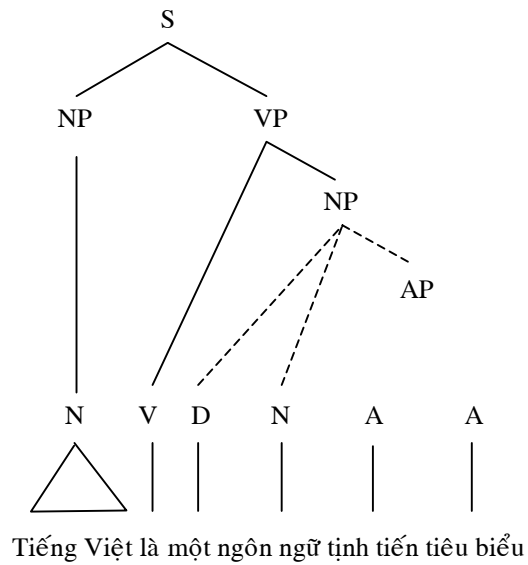
Tiếng Việt là một ngôn ngữ tịnh tiến tiêu biểu

Kế tiếp, dùng 2 quy tắc $NP \rightarrow N$ và $VP \rightarrow V NP$ cho 2 nút hở NP và VP vừa vẽ ra:

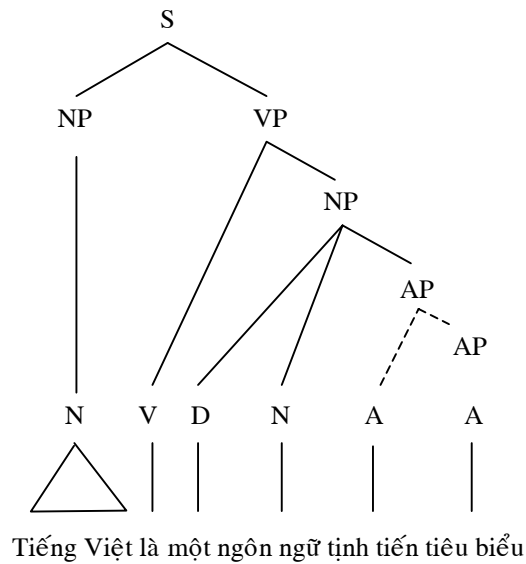


Tiếng Việt là một ngôn ngữ tịnh tiến tiêu biểu

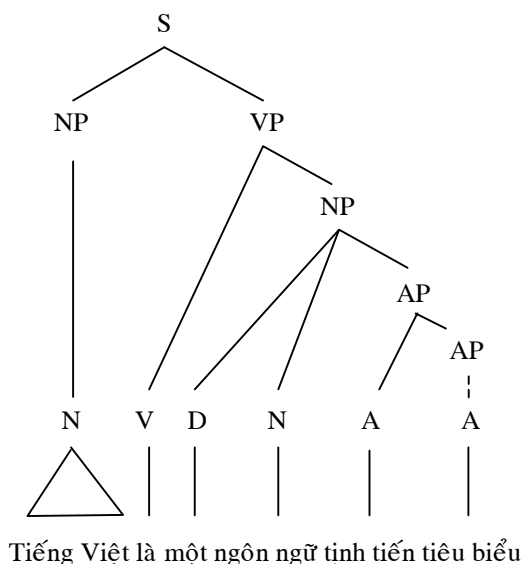
Gấp nút D, N và NP, vậy dùng $NP \rightarrow D N AP$ là hợp lý nhất:



Còn 3 nút hở là A, A, AP. Xem hai nút tính từ A, nút nào phụ nghĩa nút nào. Nhận thấy tính từ “*tiêu biểu*” là *phụ ngữ*, vậy dùng quy tắc $AP \rightarrow A AP$ là đúng nhất (thay vì $AP \rightarrow A$):



Cuối cùng chỉ còn lại 2 nút hở A và AP. Nối hai nút lại bằng quy tắc $AP \rightarrow A$ là xong:



NHẬN XÉT VÀ ĐIỂM MỞ

Cây cú pháp rất hữu ích trong việc mô tả quy tắc cấu trúc ngữ đoạn, nói chung là các đoạn câu. Thay vì phân tích bằng tay, chúng ta có thể tự động hóa bằng cách lập trình. Ví dụ:

- 1) Lập bộ phân tích cú pháp (syntax analyse).
- 2) Định nghĩa cấu trúc ngữ đoạn: Khi đưa một chuỗi từ vào, chương trình sẽ cho phép định nghĩa từng từ một theo phạm trù ngữ pháp và sau đó tự động sản sinh ra những quy tắc cấu trúc.

Từ hai ứng dụng trên, thử lập một hệ nhận diện tiếng Việt nho nhỏ mà chưa cần đến thành phần phân tích ngữ nghĩa. Hệ thống này sẽ giúp chúng ta nhận diện rất nhanh chóng những dạng ngữ đoạn tồn tại trong tiếng Việt. Ví dụ khi đưa một bất kỳ một văn bản nào vào máy, nó sẽ tự động tìm những ngữ đoạn nó biết và sẽ mô tả rõ ràng trên màn hình. Gặp những trường hợp lạ, máy không hiểu, chúng ta phải định nghĩa hoặc sửa lại quy tắc ngữ pháp như một hình thức dạy cho nó hiểu những thông tin mới. Cuối cùng chúng ta sẽ được một kho tàng quy tắc cấu trúc ngữ đoạn. Sau đó có thể tối ưu hóa chúng. Vài đề nghị:

- Nên thực hiện bằng từ vựng bằng băng dữ liệu mạnh (ví dụ ORACLE), vì số từ rất lớn và sẽ tăng theo thời gian.
- Để chứa quy tắc, không nhất thiết phải dùng băng dữ liệu.
- Cần cẩn thận khi định nghĩa quy tắc ngữ pháp, bởi vì nếu máy hiểu sai, nó có thể điều chỉnh những quy tắc đã đúng thành ra sai hết. Đó là một trong những tai hại của việc sử dụng phương pháp trí thông minh nhân tạo trong tin học.

Chương 7

TÍNH RÀNG BUỘC

Con người có thể mô tả đối tượng bằng ngôn ngữ. Đối tượng có thể là người, vật, suy nghĩ, tâm tư, sự tình, biến cố, ... , không phân biệt cụ thể hay trừu tượng. Ví dụ, tình yêu là một đối tượng thường được giới nghệ sĩ dùng làm chủ đề sáng tác. Một bài thơ tình, một bản nhạc tình, câu chuyện tình, v.v.. Mục đích của toàn bài thơ, bản nhạc, cuốn truyện là nhằm diễn tả đối tượng tình yêu. Thế nhưng có thể trong đó còn có nhiều chi tiết *trời, trăng, mây, gió, nhân vật này, nhận vật nọ*, ... Mọi thứ này cũng là đối tượng và có thể được diễn tả chi tiết hơn trong câu.

Thử xét một đoạn thơ *Màu tím hoa sim* của Hữu Loan được Phạm Duy phổ nhạc như sau:

- (1) “Nàng có ba người anh đi bộ đội lâu rồi.
 Nàng có đôi người em,
 Có em chưa biết nói.
 Tóc nàng hãy còn xanh.
 Tóc nàng hãy còn xanh.
 ...
 Tôi là người chiến binh, xa gia đình đi chiến đấu.
 Tôi yêu nàng như yêu,
 người em gái tôi yêu,
 người em gái tôi yêu ...”

Văn bản trên chứa nhiều đối tượng; câu nào cũng có: “*nàng*”, “*người em gái*”, “*tôi*”, “*tóc nàng*”... Nhìn kỹ, ta sẽ thấy, có chuỗi từ/câu không gây hiểu nhầm, nhưng có chuỗi từ/câu không được sắc bén nếu không để ý tới ngữ cảnh phía trước. Ví dụ:

- (2) Người em gái tôi yêu.

Cách diễn đạt này dễ bị hiểu nhầm nếu hiểu “*người em gái tôi*” là một cụm từ viết tắt của “*người em gái của tôi*” đại để như “*nhà tôi*” của Văn Cao:

- (3) “Nhà tôi bên chiếc cầu soi nước”.
(*Bến Xuân* – Văn Cao)

Nếu nói khơi khơi “*người em gái tôi yêu*” mà không để ý đến ngữ cảnh, người nghe rất dễ đâm thắc mắc: cô ấy yêu ai ? Chuỗi từ “*người em gái tôi yêu*” bao gồm hai cá thể “*người em gái*” và “*tôi*” hay chỉ một cá thể duy nhất: “*người em gái tôi*” ?

Chỉ nhờ môi trường ngữ cảnh, chúng ta mới hiểu rằng cụm từ “*người em gái tôi yêu*” có nghĩa đại khái là “*tôi xem nàng như một người em gái*” và “*tôi yêu*” cô. Nói cho rõ là:

- (4) Người em gái mà tôi yêu.

Ngắn nữa là:

- (5) Người em gái, tôi yêu (với dấu phẩy).

Theo cảm xúc, câu này ngắn gọn và hay nhất, thành thử tác giả mới bỏ chữ “*mà*”¹⁾ đi.

Vài ví dụ khác trong tiếng Anh:

- (6) Mary and Rosi visited Viet nam.

Nếu hỏi ai về Mary và được biết, Mary nói rằng ...:

- (7) Mary said that *she* visited Vietnam.

Nếu hỏi Mary: “*Rosi làm gì ?*” và được Mary trả lời ...:

- (8) Mary said that *she* visited Vietnam.

¹⁾ “*Thì*”, “*là*”, “*mà*” trong tiếng Việt là những tiếng có công dụng phân giới cấu trúc cú pháp. Đối với cú pháp, đây một điểm mạnh của tiếng Việt, thế nhưng người Việt thường không thích dùng, đặc biệt là giới văn chương. Hiện tượng thiếu “*thì*”, “*là*”, “*mà*” rất dễ gây lẫn lộn trong câu.

hắn ta phải thấy ý nghĩa hai câu trả lời khác nhau, dù hai câu giống hệt nhau. “*She*” trong câu (7) ám chỉ chính Mary, “*she*” trong câu (8) ám chỉ Rosi. Giống như người Anh, người Đức cũng có cách tường thuật tương tự:

- (9) Mary und Rosi haben Vietnam besucht.
(Mary và Rosi đã đi thăm Việt Nam)
- (10) Mary sagte, daß *sie* Vietnam besucht hat.
(Mary kể, cô đã đi thăm Việt Nam)
- (11) Mary sagte, daß *Rosi* Vietnam besucht hat.
(Mary kể, Rosi đã đi thăm Việt Nam)
- (12) Mary sagte, daß *sie* Vietnam besucht hat.
Nghĩa là: Mary kể, Rosi đã đi thăm Việt Nam. Người Việt không dùng “*cô*” trong câu này để ám chỉ Rosi bằng cách nói: “*Mary kể, cô đã đi thăm Việt Nam*”.

Các đối tượng được đề cập tới trong những ví dụ trên đều có mối quan hệ ràng buộc lẫn nhau. Sự ràng buộc không những chỉ tồn tại trên bình diện cú pháp mà còn trên bình diện ngữ nghĩa. Trong chương này, ta thử xem **lý thuyết ràng buộc (*binding theory*)** sẽ giải thích thế nào về ý nghĩa này.

SỰ RÀNG BUỘC

Cách dễ nhất để nhận ra tính ràng buộc của một đối tượng đối với một đối tượng là trả lời câu hỏi, đối tượng nào *dẫn chỉ* (*refer*) tới đối tượng nào. Trong câu “*Mary and Rosi visited Vietnam*”, ngữ đoạn danh từ NP “*Mary and Rosi*” dẫn chỉ tới “*Vietnam*”. Loại ngữ đoạn danh từ NP này giữ vai trò như một *biểu thức dẫn chỉ*.

Biểu thức dẫn chỉ (*referring expression*) là một ngữ đoạn danh từ NP nhận ý nghĩa bằng cách dẫn chỉ vào một sự vật.

Tính dẫn chỉ không những chỉ nói lên sự ràng buộc của một chủ thể đối với sự vật bên ngoài mà còn đối với chính chủ thể ấy.

- (13) *Lão* đã hại *mình* bằng trò chơi cờ phiếu.

Ở ví dụ trên, “*lão*” và “*mình*” ám chỉ một đối tượng duy nhất. “*Lão*” dẫn chỉ tới “*mình*” và ngược lại “*mình*” dẫn chỉ ngược về “*lão*”. Nói nôm na là “*mình*” tức

là cái “lão” ấy. “*Mình*” giữ chức năng như một *trùng ngữ* và cần ngữ cảnh phía trước.

Trùng ngữ (anaphor) là một ngữ đoạn danh từ NP nhận ý nghĩa từ một ngữ đoạn danh từ NP khác trong câu.

Tiếng Anh có những trùng ngữ như *myself, yourself, himself, herself, themselves, each other*. Ví dụ:

- (14) The murder killed *himself* with a gun.

Tính ràng buộc của ngữ đoạn danh từ NP còn xuất hiện ở dạng *đại danh từ*.

Đại danh từ (pronoun) là một ngữ đoạn danh từ NP có thể nhận ý nghĩa từ một từ khác trong câu.

- (15) Mary said that *she* visited Vietnam.

“*She*” là đại danh từ nhân vật và nhận ý nghĩa từ Mary.

Quan hệ ràng buộc chỉ cho phép ba phần tử *biểu thức dẫn chỉ, trùng ngữ* và *đại danh từ* nằm ở vị trí nhất định trong câu. Đối vị trí từ, câu có thể bị tổn thương:

- (16) a. The murder killed *himself* with a gun.
b. * *Himself* killed murder with a gun.

Câu a. nói được. Câu b. không nói được. Ngữ đoạn danh từ NP “*the murder*” trong câu a. là *tiền ngữ* đứng trước trùng ngữ “*himself*”.

Tiền ngữ (antecedent) là một ngữ đoạn danh từ NP nhận ý nghĩa từ một đại danh từ hoặc một trùng ngữ.

Riêng tiếng Việt, cần cẩn thận khi dùng “*mình*”:

- (17) a. Lão đã hại *mình* bằng trò chơi cổ phiếu.
b. *Mình* đã hại lão bằng trò chơi cổ phiếu.

Ý nghĩa hai câu trên khác nhau. “*Mình*” trong tiếng Việt có hai nghĩa, một là chính cá nhân đó, hai là “*tôi*”, “*ta*”, “*chúng ta*”. Ngữ đoạn danh từ NP “*lão*” trong câu a. là *tiền ngữ* vì đứng trước trùng ngữ “*mình*”.

Bởi có sự lặp lại đối tượng bằng những ngữ đoạn khác nhau, cho nên cần định rõ chức năng và vị thế của từng ngữ đoạn để phân biệt cho rõ từng đối tượng một. Muốn vậy, người ta dùng **chỉ số (index)**. Bằng cách này, ngữ đoạn diễn tả một đối tượng giống nhau sẽ được đánh chỉ số giống nhau. Loại chỉ số này còn được gọi là **chỉ số phụ (coindex)**. (được ghi chú bằng *i, j, k, l, m, ...* bên phải, phía dưới ngoặc vuông ‘*]*’).

(18) [Mary]_i und [Rosi]_j haben [Vietnam]_k besucht.

Câu trên bao gồm ba ngữ đoạn danh từ NP khác nhau: “*Mary*”, “*Rosi*”, “*Vietnam*”. Mỗi ngữ đoạn mang một chỉ số khác nhau: *i, j, k*. Tựa vậy:

(19) [Lão]_i đã hại [mình]_i bằng [trò chơi cổ phiếu]_j.

bao gồm ba ngữ đoạn danh từ NP: “*lão*”, “*mình*”, “*trò chơi cổ phiếu*”. “*Mình*” và “*lão*” chỉ cùng đối tượng, được ghi chú bằng *i*. “*Trò chơi cổ phiếu*” (không ám chỉ “*mình*” hoặc “*lão*”) mang chỉ số *j*. Cấu trúc câu này rất giống cấu trúc câu tiếng Anh dưới đây:

(20) [The murder]_i killed [himself]_i with [a gun]_j.

Đã mừng tượng được tính ràng buộc, chúng ta thử đi sâu vào vấn đề. Giả sử lấy câu tiếng Anh vừa nêu làm cơ sở phân tích:

(21) [The murder]_i killed [himself]_i with a gun.

Mở rộng câu trên bằng hai ví dụ:

- (22) a. [The murder_i’s lawyer]_j killed [himself]_j with a gun.
b. * [The murder_i’s lawyer]_j killed [himself]_i with a gun.

Câu a. nói được. “*Himself*” không có cùng chỉ số như “*murder*” mà cùng chỉ số như ngữ đoạn “[The murder_i’s lawyer]_j”. Đối tượng chính không phải là *kẻ sát nhân (murder)* mà là *ông luật sư (lawyer)*. Mặt khác, nếu đánh chỉ số như ở trường hợp câu b., cấu trúc câu sẽ hỏng ngay. “*Himself*” chỉ có thể dẫn chỉ tới

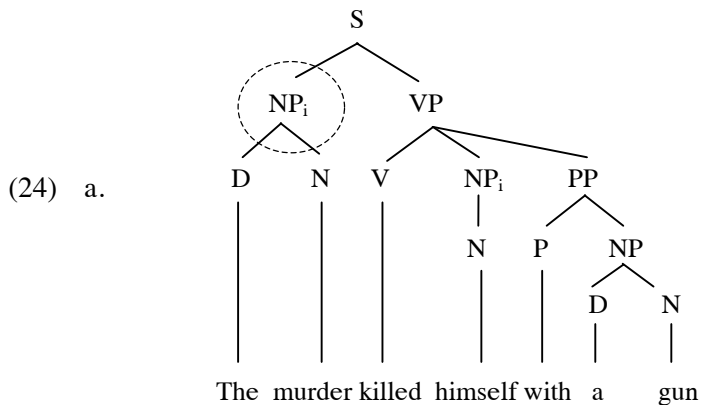
toàn thể chủ ngữ “*The murder’s lawyer*” chứ không thể dẫn chỉ tới một ngữ đoạn danh từ NP (ở đây là “*the murder*”) trong một chủ ngữ NP. Tương tự vậy:

- (23) a. [The lawyer of the murder]_i killed [himself]_j with a gun.
 b. * [The lawyer of the murder]_i killed [himself]_i with a gun.

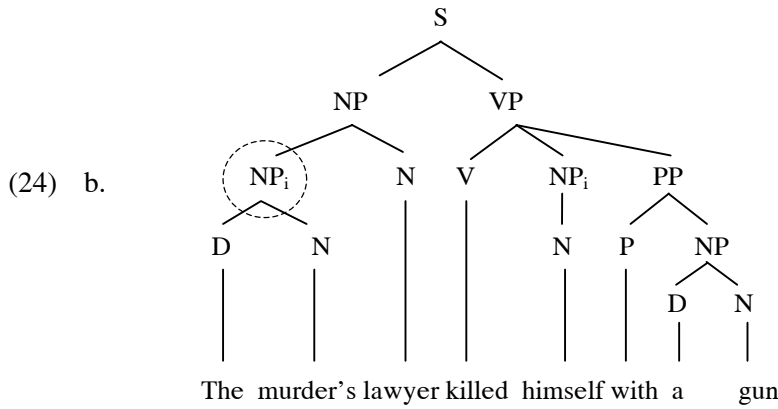
Các ví dụ nói được đều thỏa các quy tắc cấu trúc ngữ đoạn:

1. $S \rightarrow NP VP$
 2. $NP \rightarrow (D) (NP) (AP+) N$
 3. $VP \rightarrow V (NP) (PP+)$
 4. $PP \rightarrow P NP$
- $D \rightarrow \text{the} \mid a$
 - $N \rightarrow \text{murder} \mid \text{lawyer} \mid \text{himself} \mid \text{gun}$
 - $V \rightarrow \text{killed}$
 - $P \rightarrow \text{with}$

Cây cú pháp của câu “[*The murder*]_i killed [*himself*]_i with a gun”:



Ngữ đoạn danh từ NP_i bên trái (được khoanh tròn) chỉ phối trực tiếp “*murder*” và có quan hệ lệnh-C với ngữ đoạn động từ VP (tức nút chị em). Và như đã biết, nó cũng có cùng mối quan hệ với những nút con của VP bên dưới, ngay chỗ NP_i chỉ phối trực tiếp “*himself*”. Trong khi đó, cây cú pháp của “[*The murder*]_i’s lawyer]_j killed [*himself*]_i with a gun” lại có dạng khác:



NP_i bên trái chi phối trực tiếp “murder”. NP_i bên phải chi phối trực tiếp “himself”. Tuy vậy, NP_i bên trái không có quan hệ lệnh-C với NP_i bên phải. Kết quả: hai NP có cùng chỉ số phụ *i* nhưng cấu trúc lại không ổn.

Hãy đặt điều kiện chặt chẽ hơn:

Ràng buộc (bind): Một nút A bị ràng buộc vào nút B, nếu và chỉ nếu:

1. A là tiền ngữ đối với B **và**
2. A có quan hệ lệnh-C với B **và**
3. A và B có cùng chỉ số.

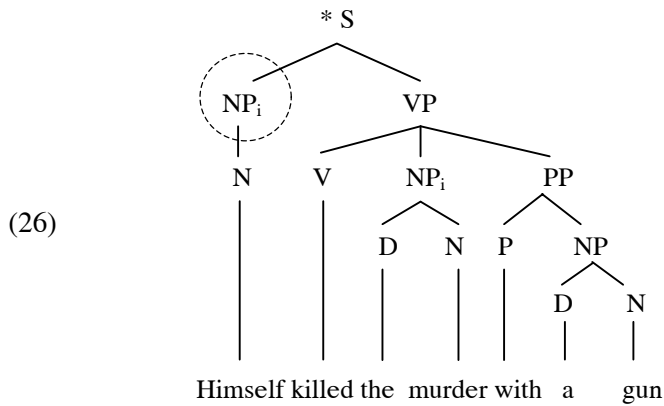
Với các điều kiện kể trên, ta có thể giải thích hai câu sau đây là sai ngữ pháp:

- (25) a. * Himself killed murder with a gun.
 b. * [The murder_i 's lawyer]_j killed [himself]_i with a gun.

bởi không có trùng ngữ nào bị ràng buộc. Tạm thời tổng quát hóa bằng nguyên tắc:

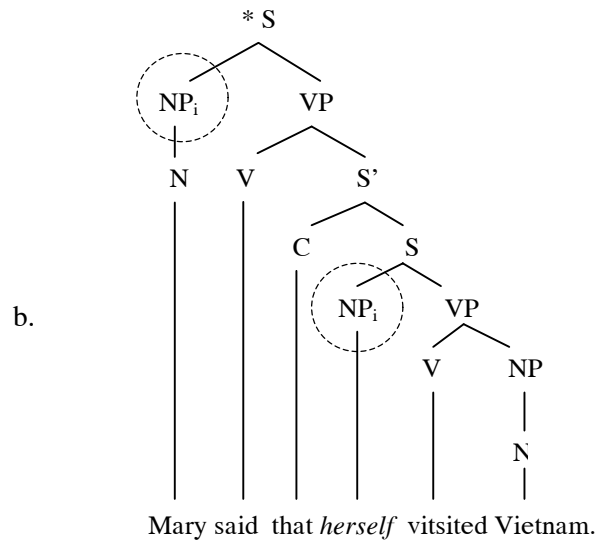
Nguyên tắc ràng buộc A (binding principle A): Một trùng ngữ bao giờ cũng bị ràng buộc.

Xin nhắc lại, “bị ràng buộc” có nghĩa là vừa có cùng chỉ số với một NP, vừa là tiền ngữ, vừa có mối quan hệ lệnh-C với chính mình. Trường hợp dưới đây không được. Ngữ đoạn “himself” không phải là tiền ngữ của “murder”:



Tính ràng buộc của tiếng Anh hay gặp rắc rối ngay mệnh đề phụ:

- (27) a. Mary said that *she* visited Vietnam.
 b. * [Mary]_i said that [*herself*]_i visited Vietnam.



Trong chuỗi từ “*Mary said that herself visited Vietnam*”, [*Mary*] vừa có cùng chỉ số như [*herself*], vừa có mối quan hệ lệnh-C với “*herself*”, và vừa là tiền ngữ. Ví dụ này hoàn toàn thỏa nguyên tắc ràng buộc A, thế nhưng lại sai ngữ pháp. Sự khác biệt giữa hai trường hợp a. và b. chỉ nằm ở chỗ “*she*” và “*herself*”. Như vậy, có thể phỏng đoán, “*herself*” không thể đóng vai trò chủ ngữ. Nói cách khác là trùng ngữ “*herself*” cần một tiền ngữ trong câu mà chính nó phải nằm

trong đó, dù đó là mệnh đề phụ. Tính chất này đã chỉ ra tính ***hạn chế cục bộ*** (*locality constrain*) trong ***phạm vi ràng buộc*** (*binding domain*), ở đó, một trùng ngữ bắt buộc phải có tiền ngữ đứng gần nó. Vì lý do này, tính ràng buộc đòi hỏi thêm điều kiện:

Nguyên tắc ràng buộc A (mở rộng): Một trùng ngữ bao giờ cũng bị ràng buộc trong phạm vi ràng buộc của nó.

Tuy vậy, những điều kiện của tính ràng buộc kể trên có lẽ chỉ thỏa cho tiếng Anh và tiếng Đức.

- (28) a. Mary said that *she* visited Vietnam (tiếng Anh).
 b. Mary sagte, daß *sie* Vietnam besucht hat (tiếng Đức).
 c. * [Mary]_i said that [*herself*]_i visited Vietnam (tiếng Anh).
 d. * [Mary]_i sagte, daß [*selbst*]_i Vietnam besucht hat (tiếng Đức)..

Tiếng Việt ít bị vấn đề hơn. Thay vì dùng trùng ngữ như *myself*, *yourself*, *himself*, *herself*, ..., tiếng Việt có rất nhiều đại danh từ nhân vật có đủ khả năng diễn tả chính xác từng nhân vật một. Nhờ thế mà một câu nói không bị hiểu nhầm đối tượng hoặc bị sai ngữ pháp.

Cụ thể, khi nói:

- (29) Mary và Rosi đã từng đến Việt Nam.

Nếu hỏi một người, Mary nói gì và nghe kể:

- (30) Mary nói là *cô* đã từng đến Việt Nam.

thì người Việt sẽ hiểu ngay, “*cô*” đây là Mary chứ không ai khác. Còn nếu hỏi Mary về Rosi, thì câu trả lời phải là:

- (31) Mary nói là Rosi đã từng đến Việt Nam.
 (32) *Cô* nói là bà/chị ấy đã từng đến Việt Nam.

Đối với những trường hợp dưới đây, người Việt cũng có thể diễn tả rõ ràng hơn. Thay vì nói:

- (33) Lão đã hại *mình* bằng trò chơi cổ phiếu.

người ta thường nói:

(34) Lão đã hại *đời mình* bằng trò chơi cổ phiếu.

Hoặc muốn diễn tả tinh thần của câu tiếng Anh:

(35) The murder killed himself with a gun.

người Việt có thể nói theo cách dịch 1-1:

(36) Tên sát nhân đã tự sát bằng súng.

hoặc dịch xuôi theo tính chất thời gian của tính tịnh tiến:

(37) Tên sát nhân đã cầm súng, tự sát.

NHẬN XÉT VÀ ĐIỂM MỞ

Tính ràng buộc là một vấn đề phức tạp trong các tiếng Âu châu. Chomsky (1981), Aoun (1985) đã phân tích góc cạnh này rất chi tiết qua lý thuyết *chi phối* và ràng buộc (*Government and Binding Theory*).

Tiếng Việt có một điểm lợi là nhờ sự phong phú của đại danh từ nhận vật, nên dễ tránh được vấn đề ấy. Thế nhưng hãy thử xem tính ràng buộc còn xuất hiện ở đâu? Những kiến thức này có thể sử dụng được không?

Chương 8

LÝ THUYẾT X-GẠCH

Ngôn ngữ không bất định mà có cấu trúc ngữ pháp. Tính tuyến tính của cấu trúc ngữ pháp xác định được nhờ vị trí từ. Ngoài tính tuyến tính, cấu trúc ngữ pháp còn có tôn ti. Tôn ti của cấu trúc ngữ pháp làm rõ mối quan hệ của từ đối với từ trong ngữ đoạn. Có từ quan trọng đóng vai ý chính của thành tố. Có từ chỉ giữ nhiệm vụ thêm nghĩa cho từ khác chứ không quan trọng. Ví dụ nói một câu:

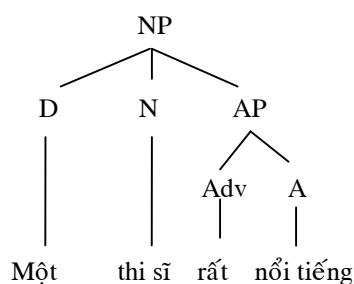
(1) Nguyễn Du là một thi sĩ rất nổi tiếng

thì cụm từ:

(2) Một thi sĩ rất nổi tiếng

là một ngữ đoạn danh từ NP chiếu theo quy tắc:

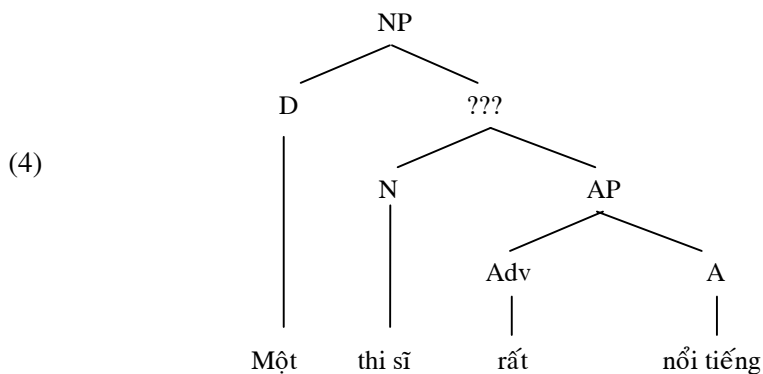
(3) $NP \rightarrow (D) N (N+) (AP+)^{1)}$



Đó là cách phân tích theo bài bản cú pháp cấu trúc ngữ đoạn (phrase structure syntax) truyền thống.

¹⁾ Xin nhắc lại: Phần tử mang dấu + có nghĩa là có thể lặp lại nhiều lần. Phần tử được đóng trong ngoặc () chỉ là một lựa chọn, có thể có, có thể không.

Ta thấy, hình thức của cụm từ “*thi sĩ rất nổi tiếng*” trong cây cú pháp trên chưa phải là một thành tố chuẩn. Muốn thấy dạng thành tố, ta phải có một cây chi tiết đại để như sau:



Phạm trù “???” chứa chuỗi “N AP”, thế nhưng nó là gì ? Dù có thể đoán “???” là một ngữ đoạn danh từ NP, nhưng chắc chắn nó không phải là một NP thường gặp.

Thử xem một câu khác:

(5) *Một thi sĩ rất nổi tiếng* được nhiều người biết tới là Nguyễn Du

Câu này nói được. “*Một thi sĩ rất nổi tiếng*” là một ngữ đoạn danh từ NP chuẩn của câu. Song nếu bỏ “*một*” ngay đầu câu:

(6) ~~Một~~ *thi sĩ rất nổi tiếng* được nhiều người biết tới là Nguyễn Du

ta sẽ có cảm giác là câu bị thiếu ngay. Nghĩ kỹ lại, ta sẽ thấy rằng cụm từ “*thi sĩ rất nổi tiếng*” trong trường hợp này không phải là một ngữ đoạn danh từ NP bình thường. Như thế nào đó, nó có vẻ lớn hơn danh từ nhưng lại nhỏ hơn ngữ đoạn danh từ NP thường gặp, đại để là nằm chính giữa. Quả thực vậy, đó là một *bình diện trung gian (intermediate level)* nằm giữa 2 phạm trù: *phạm trù từ vựng* và *phạm trù cú pháp* (tức phạm trù ngữ đoạn).

Đưa thêm *bình diện trung gian* vào nhằm mô tả chi tiết cấu trúc ngữ pháp chính là sáng kiến của **lý thuyết X-gạch (X-Bar Theory)**²⁾ (Chomsky 1970), Jackendoff (1977). Sau đây, chúng ta sẽ tập trung tìm hiểu nguyên tắc này.

(Cần lưu ý: các quy tắc ngữ pháp trong lý thuyết này nguyên thủy dựa trên tiếng Anh và được dùng chủ yếu cho tiếng Anh, cho nên chưa chắc đã đúng cho mọi ngôn ngữ khác. Tuy vậy, cứ thử xem cách giải thích về tiếng Anh và cạnh đó xem tiếng Việt thế nào)

NGUYÊN TẮC X-GẠCH

Nhờ bài bản của thuyết X-gạch, việc phân tích cú pháp không còn bị giới hạn trong 4 loại ngữ đoạn NP, VP, AP, PP, ngược lại, người phân tích có thể tự định nghĩa một ngữ đoạn mới cho ngôn ngữ đang được xét đến. Theo lý thuyết, người ta có thể định nghĩa vô số phạm trù. Mỗi *phạm trù gạch (bar-category)* tương ứng với một phạm trù từ vựng X tự tạo được gọi là **bóng chiếu X-gạch (X-bar projection)** của X.

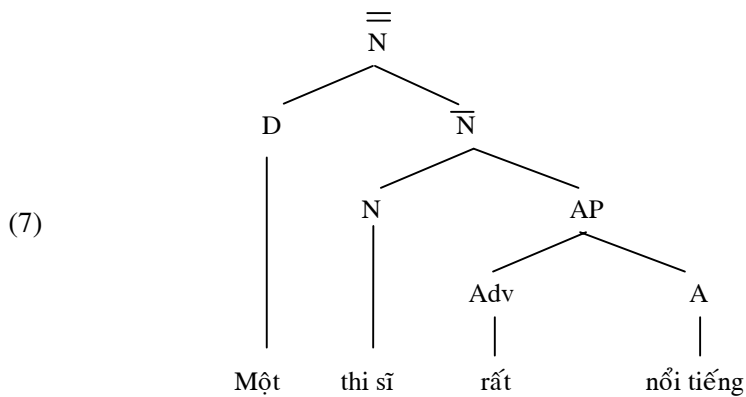
1. X (X không gạch hoặc X^0 là phạm trù từ vựng, ví dụ N, V, A, P, ...)
2. X' (X một gạch³⁾ hoặc X^1 là phạm trù X-gạch đơn (X-single-bar))
3. X'' (X hai gạch hoặc X^2)
4. X''' (X ba gạch hoặc X^3)
5. X'''' (X bốn gạch hoặc X^4)
6. ...

Qua sáng kiến tổng quát hóa sự phân đoạn đó, cú pháp cấu trúc ngữ đoạn truyền thống chỉ còn là một tiểu tập hợp của cấu trúc X-gạch.

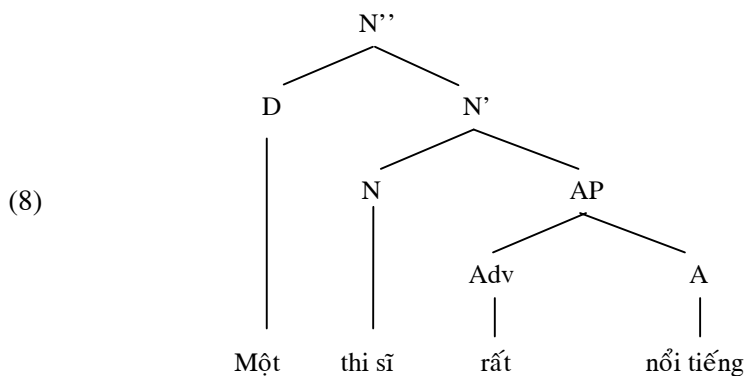
Thử xem một ví dụ cho dễ hiểu bằng cách vẽ lại cây cú pháp (4) theo sơ đồ X-gạch như sau. Phạm trù N có nghĩa là N-gạch (như X-gạch):

²⁾ Tạm dịch là lý thuyết X-gạch. Gạch có nghĩa là cái gạch trên đầu mẫu tự X ($\overline{X}$). Lý thuyết X-gạch được Chomsky giới thiệu lần đầu tiên và được những người theo sau (như Jackendoff) mở rộng.

³⁾ Dấu ' thay cho dấu gạch trên.



Xin lưu ý, để tiện việc đánh máy, kể từ đây, chúng tôi dùng dấu phẩy (') thế cho dấu gạch (-):

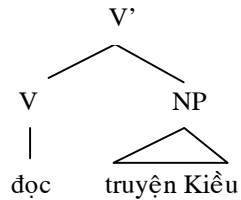


Như đã biết, mỗi ngữ đoạn đều có một *đầu* (*head*). N là đầu của NP, V là đầu của VP, v.v.. Trong cú pháp X-gạch cũng thế. Mỗi ngữ đoạn X-gạch phải có một X nhất định làm *đầu*.

Nguyên tắc đầu (*head*) X-gạch: Mỗi ngữ đoạn X-gạch chỉ có một *đầu*.

1. X-gạch đơn (X' tức X^1) phải có X (X^0) làm đầu
2. X-hai gạch (X'' tức X^2) phải có X-gạch đơn (X^1) làm đầu
3. X-ba gạch (X''' tức X^3) phải có X^2 làm đầu
4. X-bốn gạch (X'''' tức X^4) phải có X^3 làm đầu
5. ...

(9) Ông ấy đọc truyện Kiều.



V' có đầu là V (tức “đọc” là đầu của ngữ đoạn V' “đọc truyện Kiều”).

Lập thức theo quy tắc sản sinh, ta có:

1. $X^1 \rightarrow \dots X \dots$
2. $X^2 \rightarrow \dots X^1 \dots$
3. $X^3 \rightarrow \dots X^2 \dots$
4. ...

Tổng quát:

5. $X^n \rightarrow \dots X^{n-1} \dots$

có nghĩa (theo Chomsky), một ngữ đoạn thuộc phạm trù X với n gạch phải chứa một đầu là một thành tố thuộc phạm trù X với n-1 gạch (Radford 1985: 103).

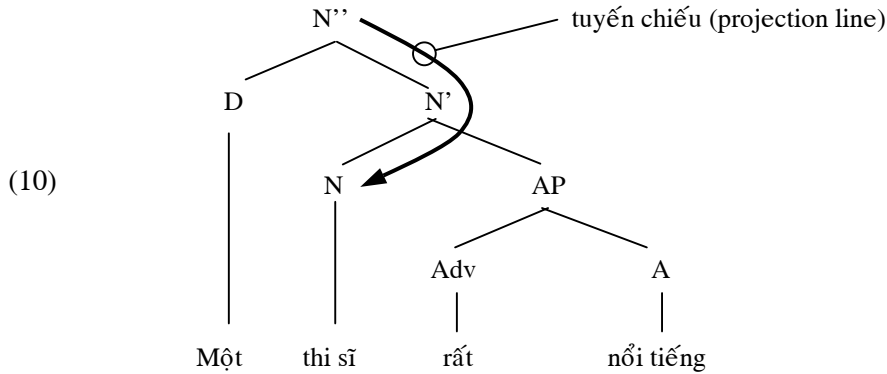
Xin nhắc lại, đầu có thể được định nghĩa tự do tùy phạm trù từ vựng của từng ngôn ngữ chứ không nhất thiết phải là N, V, A, P, ...

Trên thực tế, phạm trù từ vựng X^0 có thể không nằm trực tiếp dưới bình diện X^1 mà là dưới nhiều bình diện. Từ đỉnh ngữ đoạn muốn đi xuống X^0 , ta phải đi xuyên qua nhiều bình diện. Tuyến đường này được gọi là *tuyến chiếu* (projection line).

Nguyên tắc tuyến chiếu (projection line): Đường từ đỉnh ngữ đoạn dẫn xuống một đầu được gọi là tuyến chiếu.

Cụ thể, trong ngữ đoạn danh từ NP “một thi sĩ rất nổi tiếng” bên trên, “thi sĩ” là một từ thuộc phạm trù từ vựng danh từ N (làm trung tâm cho ngữ đoạn).

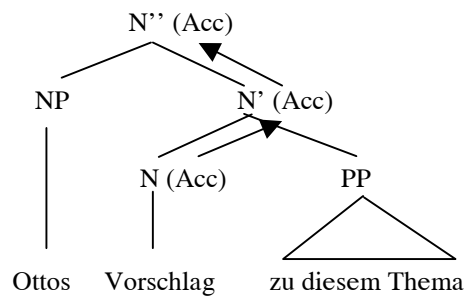
Ta có thể nhận ra tuyến chiếu (từ đỉnh ngữ đoạn N'' (NP) dẫn xuống danh từ “*thi sĩ*” như sau:



Nguyên tắc di truyền của đầu trong ngữ đoạn: Đặc điểm hình vị của ngữ đoạn dẫn đến đầu. Nói cách khác, đặc điểm này sẽ được thực hiện ở đầu.

Một ví dụ về đặc điểm hình vị *đối cách* (*accusative*, viết tắt: Acc) của một ngữ đoạn dẫn đến đầu (Stechow 1988):

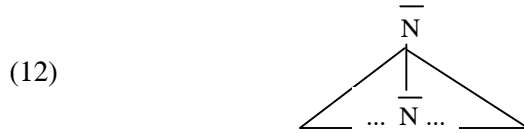
- (11) Ottos Vorschlag zu diesem Thema
(Đề nghị của Otto về đề tài này)



Nguyên tắc ngữ đoạn: Ngoại trừ đầu, phần còn lại là ngữ đoạn.

Nguyên tắc bình diện: Đầu của một phạm trù X^i là một phạm trù X^j ;
 $0 \leq j \leq i$.

Cùng quan điểm với Hornstein và Lightfoot, Radford cho rằng còn có những trường hợp một cấu trúc chứa *cấu trúc phụ* (*substructure*):



Không hẳn:

6. $X^n \rightarrow \dots X^{n-1} \dots$
ví dụ $\overline{\overline{N}} \rightarrow \dots \overline{N} \dots$

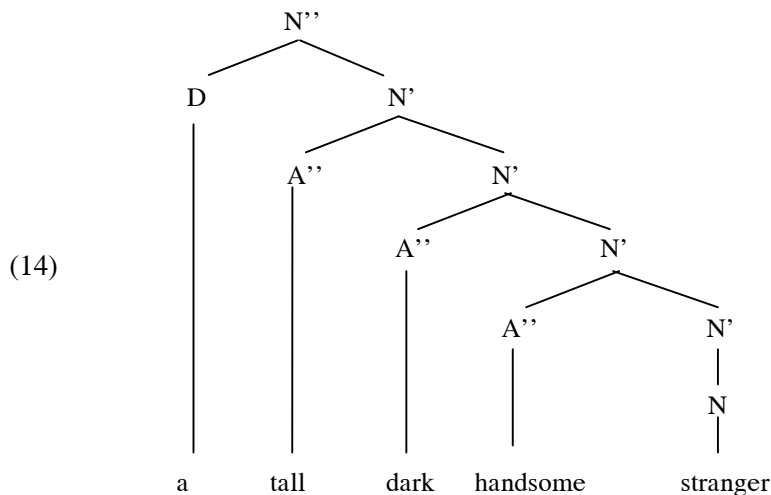
mà:

7. $X^n \rightarrow \dots X^m \dots$ ($m = n$ hoặc $m = n-1$)
ví dụ $\overline{N} \rightarrow \dots \overline{N} \dots$

Đây là trường hợp cấu trúc có tính hồi quy. Chẳng hạn một danh từ có nhiều tính từ giữ chức năng phụ nghĩa (Radford 1985: 103):

(13) a tall, dark, handsome stranger

Trong chuỗi từ, cả 3 tính từ đều ở vị thế ngang nhau, tức cùng nằm ở bình diện A''. Tính từ không chỉ phối lẫn nhau. Cây cú pháp có dạng:



Nguyên tắc X-gạch là như thế, sau đây hãy xem vài ví dụ.

VÀI VÍ DỤ X-GẠCH

Phép thế - như đã biết - là một trong những phương pháp phân tích cấu trúc ngữ pháp thường dùng trong ngôn ngữ học. Phương pháp này thực ra không có gì lạ. Ngoài đời, chúng ta vẫn thường dùng nó một cách vô tình để nói tắt:

(15) Nó quen ông xếp lớn của Microsoft, tôi cũng vậy.

“Vậy” là cách nói tắt (thay thế) cho “*quen ông xếp lớn của Microsoft*”, nghĩa là “*cũng quen ...*”. Nếu chưa nghe câu “*Nó quen ông xếp lớn của Microsoft*” mà nói trống không “*tôi cũng vậy*”, thì không ai hiểu được. “Vậy” ở đây đòi hỏi một ngữ cảnh đi trước. Chuỗi từ “*quen ông xếp lớn của Microsoft*” chính là ngữ cảnh đó. Người ta nói, chuỗi từ này đã *di truyền ý nghĩa* đó tới “vậy”.

Cấu trúc câu có tồn tại, trong đó, từng ngữ đoạn có *phạm vi* (*level*) cao thấp khác nhau. Di truyền ý nghĩa được hiểu như một *phép chiếu* (*projection*). Ví dụ ngữ đoạn “*quen ông xếp lớn của hãng Microsoft*” nằm ở một phạm vi nhất định nào đó và đã *chiếu* (*project*) sang “vậy”.

Một trong những điểm đáng chú ý trong lý thuyết X-gạch là một ngữ đoạn của một *phạm vi gạch* (*bar-level*) chính là cái cấu trúc hoàn hảo của phần nói tắt phía sau.

Sau đây chúng ta hãy thử khảo sát sơ qua những cách “chiếu” trong tiếng Anh, cụ thể là *bóng chiếu phạm vi gạch* (*bar-level projection*) của 4 trường hợp:

1. **N-gạch (N-bar)**
2. **V-gạch (V-bar)**
3. **A-gạch (A-bar)**
4. **P-gạch (P-bar)**

1. N-gạch

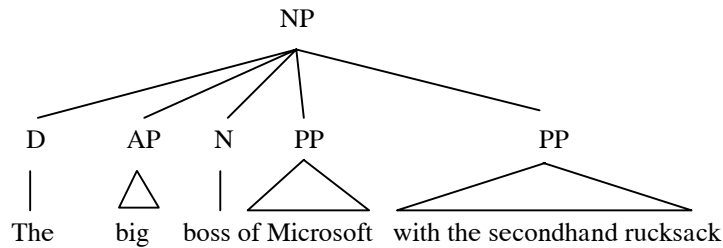
Xét đến *N-gạch* nghĩa là xét đến hình thức X-gạch mô tả một *ngữ đoạn danh từ NP*.

(16) The big boss of Microsoft with the secondhand rucksack

Ví dụ này thỏa quy tắc:

(17) $NP \rightarrow (D) N (AP+) (PP+)$

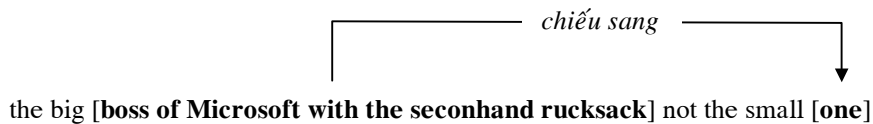
Vẽ cây cú pháp theo cấu trúc ngữ đoạn truyền thống, ta có:



Giả sử trong bối cảnh câu phát biểu nguyên thủy trên, có ai đó nói:

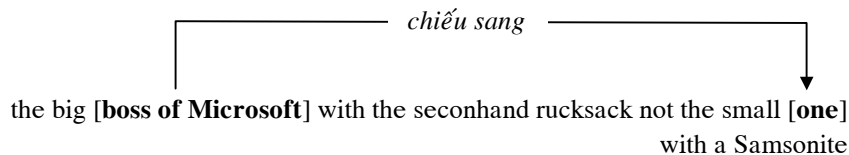
(18) I see the big [boss of Microsoft with the secondhand rucksack] *not the small [one]*.

ta sẽ thấy, “*not the small*” không thể đi với ngữ đoạn “*the big boss of Microsoft with the secondhand rucksack*” mà là đi với ngữ đoạn “*boss of Microsoft with the secondhand rucksack*”. Bằng **phép thế-“one” (one-replacement)**, tiếng Anh mới có câu trên.



Còn có trường hợp một phép chiếu đi từ “*boss of Microsoft*” vào “*one*”:

(19) I see the big [boss of Microsoft] with the secondhand rucksack *not the small [one] with a Samsonite*.



Chuỗi từ “*the big boss of Microsoft with the secondhand rucksack*” thỏa các quy tắc N-gạch :

(20) $NP \rightarrow (D) N'$

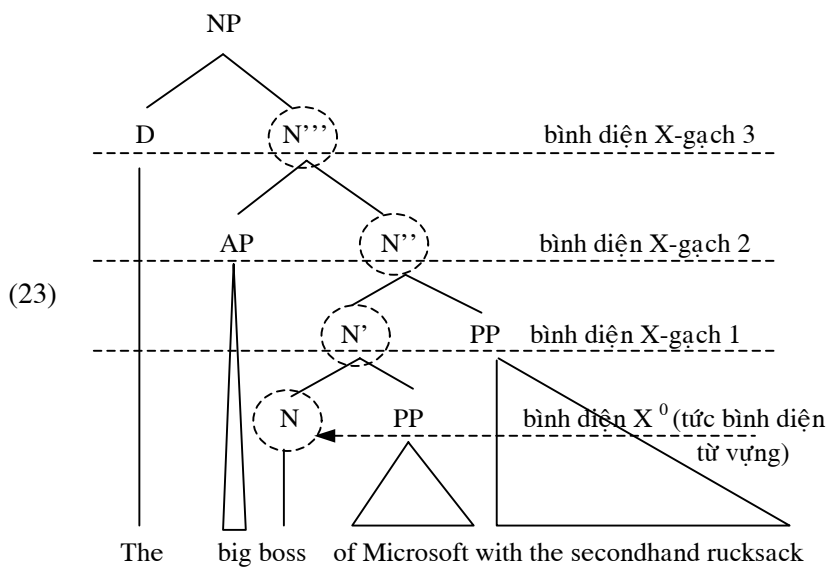
(21) $N' \rightarrow (AP) N' \mid N' (PP)$

(22) $N' \rightarrow N (PP)$

(| : phép *hoặc* (OR) trong logic)

Nhận xét: Các quy tắc X-gạch ở trên được soạn tổng quát cho ngữ đoạn danh từ NP của tiếng Anh, ở đó, bình diện N' đóng vai một ngữ đoạn trung gian dưới dạng NP.

Cây cú pháp X-gạch của “*The big boss of Microsoft with the secondhand rucksack*” có dạng:



Cây mô tả 3 bình diện X-gạch (1, 2, 3). Mỗi ngữ đoạn trung gian nằm ở một bình diện nhất định. Hai ngữ đoạn N' và N'' ở bình diện 1 và 2 là hai ngữ đoạn mà “one” có thể thế được.

Bình diện của N''' được dành sẵn cho mạo từ “*the*”. Có trường hợp, thay vì dùng “*the*”, người ta dùng “*that*” (*that one*):

(24) I see [_{NP} this [_{N'} secondhand rucksack]] not [_{NP} that [_{N'} one]].

Rộng hơn nữa là trường hợp sử dụng *liên từ*, ví dụ “*and*”:

(25) The big boss of Microsoft with the secondhand rucksack *and* the old hut

Đó là vài trường hợp thể bằng *phép thế* “*one*” trong tiếng Anh. Nay ta hãy thử xem tiếng Việt có tương tự thể không. Giả sử nói:

(26) Tôi quen ông xếp lớn của Microsoft.

(27) Tôi quen ông xếp lớn của Microsoft chứ không quen ông xếp nhỏ.

Hai câu đều nói được. Ngữ đoạn danh từ “*ông xếp*” tương ứng với “*one*” của tiếng Anh. Tuy vậy, trong tiếng Việt, người nói không thể nói tắt hoặc lược từ dễ dàng, chẳng hạn như lược bỏ từ “*xếp*” trong câu phía dưới:

(28) ? Tôi quen ông xếp lớn của Microsoft chứ không quen ông ~~xếp~~ nhỏ.

Câu này nghe khá lấn cấn vì mất tính chất “*xếp*”. Có phải vì thiếu chữ “*xếp*”, một vai trò trung tâm không thể thiếu ? Vậy, thử bỏ “*ông*”:

(29) Tôi quen ông xếp lớn của Microsoft chứ không quen ~~ông~~ xếp nhỏ.

Như thế nào đó, câu nghe vẫn không được tự nhiên cho lắm. Trên nguyên tắc, danh từ hình thức (formal noun) “*ông*” đại diện cho “*xếp*”, nhưng chưa chắc bao giờ nó cũng khẳng định được vai trò trung tâm.

Một ví dụ khác:

(30) Tôi quen bà vợ cũ của ông ta.

(31) Tôi quen bà vợ cũ của ông ta chứ không quen bà vợ mới.

Lược bỏ danh từ “*vợ*” (giống trường hợp “*xếp*” ở trên), ta có:

(32) Tôi quen bà vợ cũ của ông ta chứ không quen bà ~~vợ~~ mới.

Câu nói được, nhưng sẽ khó nghe nếu bỏ “*bà*”:

(33) ? Tôi quen bà vợ cũ của ông ta chứ không quen ~~bà~~ vợ mới.

Trong trường hợp này, ta lại thấy, danh từ hình thức “*bà*” giữ vai trò trung tâm một cách hữu hiệu. Xem thêm vài trường hợp gần giống:

- (34) Tôi quen thằng con nhỏ của ông ta.
- (35) Tôi quen thằng con nhỏ của ông ta chứ không quen thằng con lớn.
- (36) Tôi quen thằng con nhỏ của ông ta chứ không quen thằng ~~con~~ lớn.

Tương tự trường hợp lược bỏ “*bà*” ở ví dụ trên, nếu lược bỏ “*thằng*”, câu sẽ bị vấn đề:

- (37) * Tôi quen thằng con nhỏ của ông ta chứ không quen ~~thằng~~ con lớn.

Kết quả: Cả hai danh từ hình thức “*bà*”, “*thằng*” không bỏ được vì phải giữ vai trò đại diện cho trung tâm ngữ đoạn. Ngoài ra còn có cả trường hợp *danh từ hình thức* lẫn *danh từ chính* đều giữ chức năng ấy:

- (38) Tôi quen ông cha sở họ Vườn Xoài.
- (39) Tôi quen ông cha sở họ Vườn Xoài chứ không quen ông cha sở bên xứ Bùi Phát.
- (40) Tôi quen ông cha sở họ Vườn Xoài chứ không quen ông ~~cha-sở~~ bên xứ Bùi Phát.
- (41) Tôi quen ông cha sở họ Vườn Xoài chứ không quen ~~ông~~ cha sở bên xứ Bùi Phát.

Một lần nữa, qua phép tỉnh lược, ta lại thấy tính phức tạp và uyển chuyển của cấu trúc ngữ đoạn danh từ trong tiếng Việt. Phép tỉnh lược là một đề tài khá thú vị, thế nhưng xin miễn đào sâu vào chi tiết ở đây mà xin nhường cho một chuyên luận khác.

2. V-gạch

Xét đến *V-gạch* nghĩa là xét đến hình thức X-gạch mô tả một *ngữ đoạn động từ VP*.

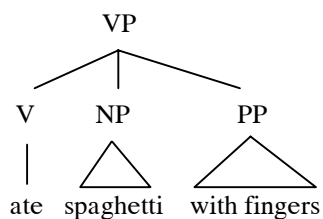
Một ví dụ trong tiếng Anh:

- (42) Peter [ate spaghetti with fingers].

Cấu trúc ngữ đoạn động từ VP “*ate spaghetti with fingers*” có quy tắc truyền thống:

$$(43) \text{ VP} \rightarrow \text{V (NP) (PP)}$$

Cây cú pháp:



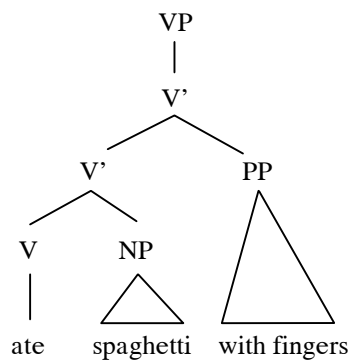
Mở rộng cấu trúc trên bằng quy tắc V-gạch:

$$(44) \text{ VP} \rightarrow \text{V'}$$

$$(45) \text{ V'} \rightarrow \text{V' (PP)}$$

$$(46) \text{ V'} \rightarrow \text{V (NP)}$$

ta có cây cú pháp:



Nhờ cấu trúc X-gạch, người Anh có thể giải thích được trường hợp ***phép thế*** “*do so*” (*do-so-replacement*):

(47) Peter [ate spaghetti] with fingers but Mary [did so] with forks.

(48) Peter [ate spaghetti with fingers] and Mary [did so] too.

Ngay ví dụ này, người Việt cũng có cách diễn tả tương tự song không dùng một loại động từ giống “*do so*” để nói tắt mà dùng “*cũng vậy*”, “*cũng thế*”, “*cũng rứa*”:

(49) Người ta [ăn cơm bằng đũa], nó [cũng vậy].

[*Cũng vậy*] thế cho ngữ đoạn động từ [*ăn cơm bằng đũa*] nhưng không thể thế cho [*ăn cơm*] (giống [*ate spaghetti*] trong tiếng Anh):

(50) * Người ta [ăn cơm] bằng đũa, nó [cũng vậy] bằng đũa.

Muốn diễn tả “*cũng ăn cơm*”, người Việt thường nói:

(51) Người ta ăn cơm bằng đũa, còn nó (ăn) bằng muống.

(52) Người ta ăn cơm bằng đũa, còn nó (thì) bằng muống.

(53) Người ta ăn cơm bằng đũa, nó (thì) bằng muống.

hoặc ngắn nữa:

(54) Người ta ăn cơm bằng đũa, nó thì muống.

[*thì*] đại diện cho cấu trúc [*(thì) ăn cơm bằng*].

Hiện tượng nói tắt quá đáng thường làm câu nói bị cảm ngữ cảnh (context sensitive). Thay vì nói, ăn cơm bằng đũa, hoặc thịt bò “*beaf steak*” bằng dao, chúng ta vẫn thường nghe nói tắt:

(55) Ăn đũa nhá ?

(56) Ăn dao nhá ?

Đây là một trường hợp cảm ngữ cảnh tiêu biểu: Muốn *ăn thịt bò* hay *ăn dao* ?

Hà tiện chữ quá khiến ngôn ngữ dễ trở nên cộc lốc, thiếu trong sáng. Thế nhưng cũng có trường hợp dù không bỏ chữ, ngữ nghĩa vẫn có thể bị tổn thương vì dùng chữ không đúng. Chẳng hạn dùng giới từ “*với*” ngay ví dụ này. Nếu nói:

(57) Ăn thịt bò với dao

suy ra, cũng có thể nói:

(58) Ăn cơm với dưa.

Cấu trúc hai kiểu nói trên giống hệt như:

(59) Ăn cơm với nước mắm.

Nghĩ logic, “*nước mắm*” nuốt vào bụng được, suy ra “*dưa*”, “*dao*” cũng nuốt vào bụng được !?

Muốn tránh hiểu nhầm, chỉ còn cách, thay vì dùng “*với*” thì nên dùng “*bằng*”.

3. A-gạch

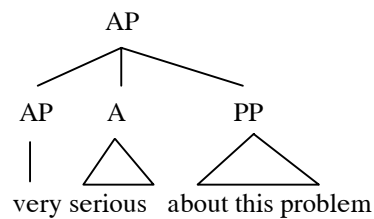
Xét đến *A-gạch* nghĩa là xét đến hình thức X-gạch mô tả một *ngữ đoạn tính từ AP*.

(60) Peter is [very serious about this problem].

Với cấu trúc ngữ đoạn truyền thống:

(61) $AP \rightarrow (AP) A (PP)$

ta có cây cú pháp:



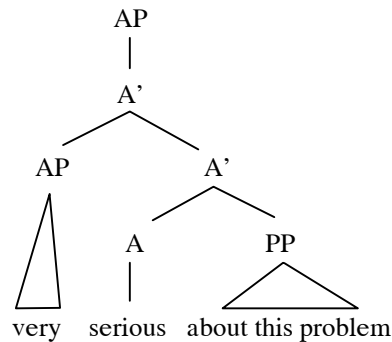
Mở rộng cấu trúc bằng sơ đồ A-gạch:

(62) $AP \rightarrow A'$

(63) $A' \rightarrow (AP) A'$

(64) $A' \rightarrow A (PP)$

ta có:



Nhờ cấu trúc A-gạch, tiếng Anh có thể thay thế ngữ đoạn tính từ AP [*very serious about this problem*] bằng **phép thế-“so” (so-replacement)**:

(65) Peter is [very serious about this problem] but Mary is not [so].

Người Việt cũng có thể diễn tả tương tự bằng cách dùng “*cũng vậy*”, “*cũng thế*”, “*cũng rứa*”:

(66) Mùa Đông ở Thụy Điển [lạnh lắm], mùa Đông ở Canada [*cũng vậy*].

(67) Cảnh [buồn], người [*cũng rứa*].

(68) Món ăn Việt [rất ngon và đa dạng], món ăn Tàu [*cũng thế*].

Thế nhưng gặp trường hợp phủ định, ta lại phải thêm v.d. “*thế*”, “*phải thế*” vào:

(69) Ông bố rất nghiêm khắc với con cái nhưng bà mẹ thì không.

(70) Ông bố rất nghiêm khắc với con cái nhưng bà mẹ ~~thì~~ không **thế**.

(71) Ông bố rất nghiêm khắc với con cái nhưng bà mẹ ~~thì~~ không **phải thế**.

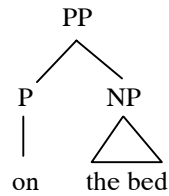
4. P-gạch

Xét đến *P-gạch* nghĩa là xét đến hình thức X-gạch mô tả một *ngữ đoạn giới từ* PP.

(72) Peter slept [on the bed].

Cấu trúc truyền thống:

(73) $PP \rightarrow (PP) P (NP)$

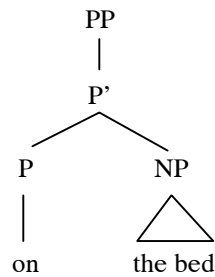


Cấu trúc P-gạch:

(74) $PP \rightarrow P'$

(75) $P' \rightarrow P' (PP)$

(76) $P' \rightarrow P (NP)$



TỔNG KẾT CÁC QUY TẮC X-GẠCH CỦA TIẾNG ANH

Đã xem qua những ví dụ X-gạch của 4 phạm trù ngữ pháp *danh từ N*, *động từ V*, *tính từ A* và *giới từ P*, hẳn chúng ta đã thấy chiều sâu của cấu trúc đã được làm rõ như thế nào. Xin nhắc lại, cú pháp X-gạch không bị giới hạn trong 4 phạm trù ấy mà có thể là những phạm trù trừu tượng do người sử dụng tự định nghĩa nhằm giải thích một cấu trúc ngôn ngữ không nhất thiết phải tồn tại trong tiếng Anh.

Đối với tiếng Anh, các quy tắc X-gạch được dùng để giải thích 4 phạm trù ngữ pháp *danh từ N*, *động từ V*, *tính từ A* và *giới từ P* bên trên có thể được tóm gọn như sau:

(77) $NP \rightarrow (D) N'$

(78) $N' \rightarrow (AP) N' \mid N' (PP)$

(79) $N' \rightarrow N (PP)$

(80) $VP \rightarrow V'$

$$(81) \quad V' \rightarrow V' (PP)$$

$$(82) \quad V' \rightarrow V (NP)$$

$$(83) \quad AP \rightarrow A'$$

$$(84) \quad A' \rightarrow (AP) A'$$

$$(85) \quad A' \rightarrow A (PP)$$

$$(86) \quad PP \rightarrow P'$$

$$(87) \quad P' \rightarrow P' (PP)$$

$$(88) \quad P' \rightarrow P (NP)$$

Tất cả các ngữ đoạn NP, VP, AP, PP trên đều có nút **đầu (head)**. Khác với N, V, A, P, các nút N', V', A', P' thuộc **bình diện trung gian (intermediate level)**. Bình diện trung gian không tận cùng bằng một từ mà dẫn xuống ngữ đoạn sâu hơn. Toàn thể cấu trúc gồm có 3 bình diện: *bình diện chính* (NP, VP, AP, PP), *bình diện trung gian* (N', V', A', P') và *bình diện tận cùng* (N, V, A, P). Tổng quát hóa các quy tắc bằng biến số X, Y, W, Z, ta có:

1. Bình diện chính

$$(89) \quad NP \rightarrow (D) N'$$

$$(90) \quad VP \rightarrow V'$$

$$(91) \quad AP \rightarrow A'$$

$$(92) \quad PP \rightarrow P'$$

NP, VP, AP, PP là bốn bình diện chính. Dùng biến số X thế cho những phạm trù ngữ pháp N, V, A, P, với tập hợp $X = \{N, V, A, P\}$, ta sẽ được một quy tắc gọn gàng:

$$(93) \quad XP \rightarrow (YP) X'$$

(XP: X-Phrase, YP: Y-Phrase). Xin lưu ý, YP là biến số của D trong quy tắc NP.

2. Bình diện trung gian

Các nút N', V', A', P' mang tính hồi quy:

- (94) $N' \rightarrow (D) N' \mid N' (PP)$
 (95) $V' \rightarrow V' (PP)$
 (96) $A' \rightarrow (AP) A'$
 (97) $P' \rightarrow P' (PP)$

Thế N, V, A, P bằng biến số X, Z, ta có:

$$(98) \quad X' \rightarrow (ZP) X' / X' (ZP)$$

Z đại diện cho D, PP, AP, PP.

3. Bình diện tận cùng

Cuối cùng, thế các quy tắc sau:

- (99) $N' \rightarrow N (PP)$
 (100) $V' \rightarrow V (NP)$
 (101) $A' \rightarrow A (PP)$
 (102) $P' \rightarrow P (NP)$

bằng biến số X và W, ta được quy tắc:

$$(103) \quad X' \rightarrow X (WP)$$

W đại diện cho PP, NP.

4. Tổng kết các quy tắc

Cuối cùng chỉ còn 3 quy tắc được rút gọn nhờ biến số X, Y, Z, W:

- (104) $XP \rightarrow (YP) X'$
 (105) $X' \rightarrow (ZP) X' / X' (ZP)$
 (106) $X' \rightarrow X (WP)$

Trên lý thuyết, chỉ cần ba quy tắc tổng quát, giới ngôn ngữ học Tây phương đã có thể giải thích được hầu hết các câu tiếng Anh. Tuy vậy, không phải là không còn vấn đề. Mặt khác, chưa chắc các quy tắc trên đều dùng được cho mọi ngôn

ngữ dù cũng có các phạm trù *danh từ N*, *động từ V*, *tính từ A* và *giới từ P*. Chúng ta sẽ phân tích chi tiết vấn đề này ở các chương sau.

TÓM LƯỢC

Bài bản quy tắc cấu trúc ngữ đoạn truyền thống không thể mô tả chiều sâu cấu trúc của nhiều ngôn ngữ. Từ đó, nó đã được mở rộng bằng lý thuyết X-gạch.

Sáng kiến chính của lý thuyết X-gạch là dùng nút trung gian X-gạch. Đầu của từng ngữ đoạn X không nhất thiết phải là danh từ N, động từ V, tính từ A, giới từ P, ... như trước nay mà có thể là một phạm trù trừu tượng được người sử dụng tự định nghĩa.

Các đặc điểm chính của lý thuyết X-gạch:

1. **Nguyên tắc đầu ngữ đoạn:** Mỗi ngữ đoạn chỉ có một đầu, có thể được định nghĩa tự do tùy phạm trù từ vựng của từng ngôn ngữ.
2. **Nguyên tắc di truyền đầu trong ngữ đoạn:** Đặc điểm hình vị của ngữ đoạn dẫn đến đầu.
3. **Nguyên tắc ngữ đoạn:** Phần phụ của đầu là một ngữ đoạn.
4. **Nguyên tắc bình diện:** Đầu một phạm trù X^i là một phạm trù X^j ; $0 \leq j \leq i$.

Tuyến chiếu (projection line) là đường một phạm trù phức tạp dẫn xuống đầu từ vựng. Nói ngược lại, đầu từ vựng sẽ chiếu đặc điểm của nó qua tuyến này, thế nhưng không chiếu lên toàn thể cây cú pháp mà chỉ từng phần.

Độ phức tạp của bình diện (level complexity): Đối với mọi phạm trù từ vựng $X \in \{N, V, A, P, \dots\}$:

- X^0 : **bình diện từ vựng (lexical level).**
- $X^{\max}$: **bình diện ngữ đoạn (phrase level = XP)**

Ví dụ: bổ ngữ, tức phần tử được coi như phạm trù phụ, nằm ở bình diện 1 (X^1); phụ ngữ ở bình diện 2 (X^2), ... Xin xem thêm chi tiết về bổ ngữ, phụ ngữ ở chương sau.

ĐIỂM MỞ

- 1) **Ngữ đoạn danh từ NP:** Ứng dụng hoặc nếu cần, phải khai triển lý thuyết X-gạch thế nào để giải thích cách thành lập ngữ đoạn danh từ chứa danh từ hình thức (“cái”, “chiếc”, “con”, “cây”, ...). Có trường hợp nào có cùng cấu trúc ngữ pháp mà không nói được? Còn sự khác biệt nào chưa rõ?
- 2) **Ngữ đoạn động từ VP:** Giải thích cách bổ nghĩa trong ngữ đoạn động từ có trạng từ, tính từ, giới từ, danh từ (mệnh đề phụ).
- 3) Thử xét sâu hơn về phép tỉnh lược trong tiếng Việt. Phép chiếu có thể ứng dụng cho những trường hợp bị lược bỏ nào? Thử lập quy tắc X-gạch, xem cách nói tắt nào hợp lý, cách nói tắt nào không.
- 4) Có thể tóm gọn những cấu trúc căn bản của tiếng Việt thành những quy tắc X-gạch như tiếng Anh được không?
- 5) Thử phân tích vài ví dụ cụ thể sau đây để xem ta có thể tự định nghĩa được những phạm trù khác N, V, A, P đã nêu trên hay không. Xét kỹ các từ như “hả”, “đâu”, “à”, “nhá”, “kia”, “đi”, “nhỉ”... xem chúng có chức năng gì. Nếu tỉnh lược những từ này, câu còn nói được không hoặc ý nghĩa bị thay đổi thế nào?
 - Ăn cơm đi chứ hả.
 - Anh chưa ăn đâu.
 - Đói thấy mụ nội mà chưa ăn à.
 - Lạng quạng là ăn đòn đấy nhá.
 - Đi chơi nhá.
 - Đi chơi vui nhá.
- 6) Thử phân tích chiều sâu ngữ nghĩa của những ví dụ dưới đây. Những danh từ nào thực sự làm chủ ngữ logic trong câu? Ứng dụng sơ đồ X-gạch như thế nào để giải thích các trường hợp: chủ ngữ logic không có mặt trong câu hoặc câu bị tỉnh lược?
 - Bức tranh treo ở trên tường.
 - Bức tranh nằm ở trên bàn.
 - Bức tranh rơi xuống đất.
 - Ê ! Quần áo đang bán xon kìa. Đi mua đi.
 - Bà Ba mới mua cái nhà mà đã bán mất tiêu.
 - Bà Ba mới mua cái nhà mà đã cháy mất tiêu.
 - Hùng phải đi cắt tóc.
 - Ông Năm đang cắt tóc cho khách.

- “Cứ đến ngày Chủ nhật là phải chuẩn bị tiền để thối. Khổ lắm, không có tiền lẻ là khách bỏ đi mua chỗ khác ngay”.
(Trích báo Tuổi Trẻ, 21.09.2001)
- “Lô hàng giả cổ của DN tư nhân Văn Lê (Đồng Nai) xuất qua Mỹ đầu tuần rồi không cập cảng được, phải neo lại ngoài khơi, mãi đến đầu tuần này mới giải tỏa được hàng”.
(Trích báo Tuổi Trẻ, 21.09.2001)
- “Năm nào cũng vậy, sau khi đã đưa được con vào trường là phụ huynh lại bắt đầu nơm nớp chờ đợi “không biết năm nay đóng bao nhiêu ?” mà lo ngại nhất có lẽ vẫn là “tiền xây dựng trường”. Một trường tiểu học ở quận 1, TP.HCM mới xây vài năm nay với một cơ ngơi khá đồ sộ ...”.
(Trích báo Tuổi Trẻ, 21.09.2001)
- “Ngoài mỏ Ruby đã đưa vào khai thác và mỏ mới phát hiện, Petronas cũng đã khoan thăm định và phát hiện khí ở mỏ Emerald nằm trong lô này”
(Trích báo Tuổi Trẻ, 21.09.2001)

Chương 9

LÝ THUYẾT X-GẠCH – BỔ NGỮ, PHỤ NGỮ

BỔ NGỮ, PHỤ NGỮ

Trong một ngữ đoạn, có từ đóng vai chính, có từ đóng vai phụ. Từ đóng vai chính được coi là từ làm trung tâm. Từ này bao hàm ý chính của ngữ đoạn và được xem như chủ điểm muốn nói. Từ đóng vai phụ chỉ giữ nhiệm vụ thêm nghĩa. Song, thêm nghĩa không chỉ thuần là thêm nghĩa. Có hình thức thêm nghĩa nhằm bổ túc cho đầy đủ nghĩa, có hình thức nhằm phụ thêm nghĩa chứ không quan trọng. Một ví dụ:

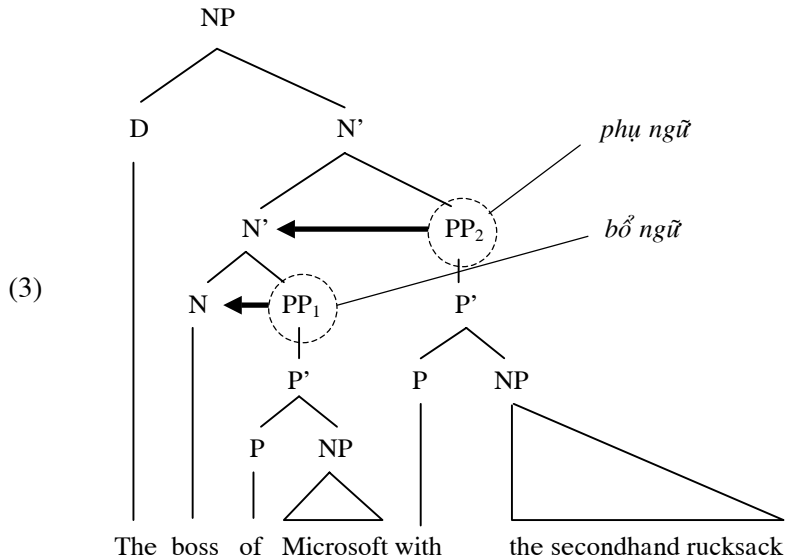
- (1) Người con gái Việt Nam da vàng

Trong ví dụ này, từ “*Việt Nam*” được coi như giữ chức năng **bổ ngữ** (*complement*), bởi vì nó bổ nghĩa cho đối tượng “*người con gái*” để đối tượng này được đầy đủ ý nghĩa. Trong khi đó, “*da vàng*” chỉ là **phụ ngữ** (*adjunct*). Công dụng của nó là phụ nghĩa thêm chứ không cần thiết. Thành thử khi nghe cụm từ “*Người con gái Việt Nam da vàng*”, chúng ta sẽ cảm thấy “*Việt Nam*” có vẻ gắn bó với cụm từ “*người con gái*” hơn so với cụm từ “*da vàng*”. Đây là ý nghĩa của bổ ngữ và phụ ngữ trong ngôn ngữ học. Sau đây chúng ta sẽ đi vào chi tiết.

Chép lại ví dụ tiếng Anh ở chương rồi:

- (2) The boss of Microsoft with the secondhand rucksack

Nguyên tắc X-gạch cho phép giải thích các ngữ đoạn trừu tượng. Ứng dụng sơ đồ này để giải thích hai loại ngữ đoạn bổ ngữ, phụ ngữ, ta có cây cú pháp:



Cấu trúc trên miêu tả hai ngữ đoạn giới từ PP nằm ở hai bình diện khác nhau. PP₂ ở vị thế cao hơn PP₁. Trong khi PP₂ là *chị em* của N', tức *bình diện gạch đơn* (*single bar-level*), thì PP₁ là *chị em* của N. Nút N' kiểm soát nút N. Nút PP₂ thỏa quy tắc:

$$(4) \quad X' \rightarrow X' (ZP)$$

mặt khác, PP₁ thỏa:

$$(5) \quad X' \rightarrow X (WP)$$

Đó là hai loại ngữ đoạn giới từ PP. Mỗi loại giữ nhiệm vụ hoặc bổ nghĩa hoặc phụ nghĩa. Định nghĩa:

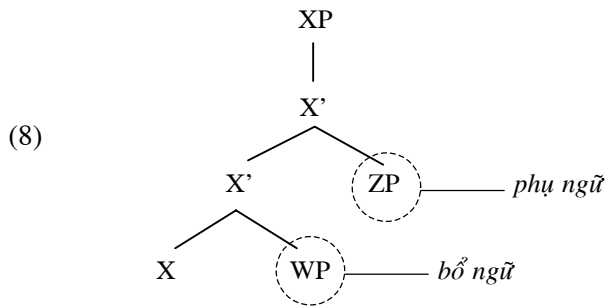
Phụ ngữ (*adjunct*) là ngữ đoạn XP có quan hệ chị em với những bình diện gạch đơn (N', V', A', P') và đồng thời là con của một bình diện gạch đơn (N', V', A', P').

Bổ ngữ (*complement*) là ngữ đoạn XP có quan hệ chị em với đầu (N, V, A, P) và đồng thời là con của của một bình diện gạch đơn (N', V', A', P').

Quy tắc ngữ đoạn diễn tả phụ ngữ và bổ ngữ trong tiếng Anh có thể được lập thức như sau:

- (6) Quy tắc phụ ngữ: $X' \rightarrow X' (ZP)$
 (7) Quy tắc bổ ngữ: $X' \rightarrow X (WP)$

Cây cú pháp tổng quát:



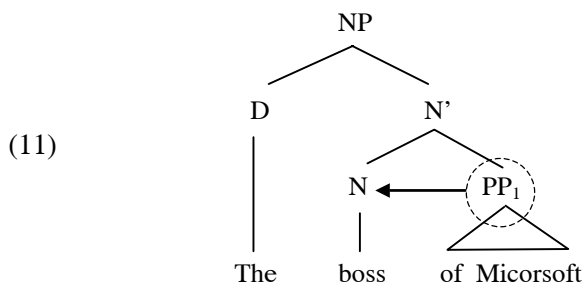
Sau đây, chúng ta sẽ tuần tự xét thêm chi tiết về tính chất bổ ngữ và phụ ngữ trong các loại ngữ đoạn NP, VP, AP, PP.

Trước nhất, hãy xét đến bổ ngữ và phụ ngữ trong ngữ đoạn danh từ NP. Ta dùng lại hai ví dụ tiếng Anh:

- (9) The boss of Microsoft
 (10) The boss with the rucksack

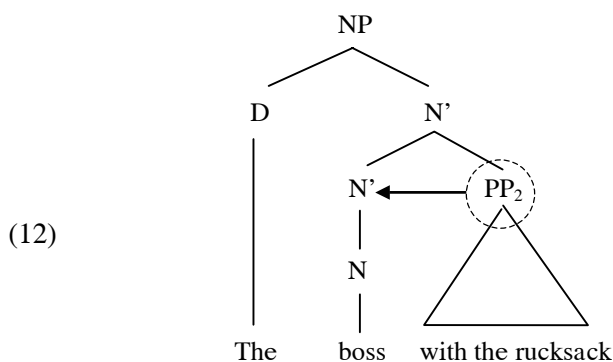
Mặc dù tính cách thêm nghĩa cho “the boss” trong hai ví dụ có vẻ giống nhau, thế nhưng giới ngôn ngữ học Tây phương lại cho là không giống. Sự khác biệt nằm ở độ sâu.

Trường hợp bổ nghĩa:



Ngữ đoạn giới từ PP₁ làm bổ ngữ cho “the boss”:

Trường hợp phụ nghĩa:



Ngữ đoạn giới từ PP₂ làm phụ ngữ cho “the boss”. PP₂ không phải là *chị em* của *đầu ngữ đoạn N* như ở trường hợp trên mà là *chị em* của *bình diện gạch đơn N'*. Mới nhìn sơ thì thấy PP₁ và PP₂ có vẻ giống nhau, thế nhưng nhìn sâu hơn thì khác: PP₁ nằm sát “the boss” để bổ túc cho đầy nghĩa. Nghe nói “The boss”, nhưng “the boss” nào ? À, “the boss of Microsoft”. Trong khi đó, PP₂ chỉ giữ nhiệm vụ phụ nghĩa, chứ không xác định “the boss” rõ ràng như PP₁. Trong hãng Microsoft có nhiều “boss”, có “boss” xách Samsonite đi làm, có “boss” đeo ba lô, có “boss” đi tay không.

Sự phán đoán một từ hoặc một ngữ đoạn đóng vai trò gì, bổ ngữ hay phụ ngữ, thực ra chỉ có tính cách tương đối. Hầu hết các từ hoặc ngữ đoạn tiếng Anh theo sau “of” đều là bổ ngữ (dĩ nhiên cũng có trường hợp ngoại lệ). Phụ ngữ thường theo sau những giới từ như *from*, *as*, *to*, *with*, *under*, ...

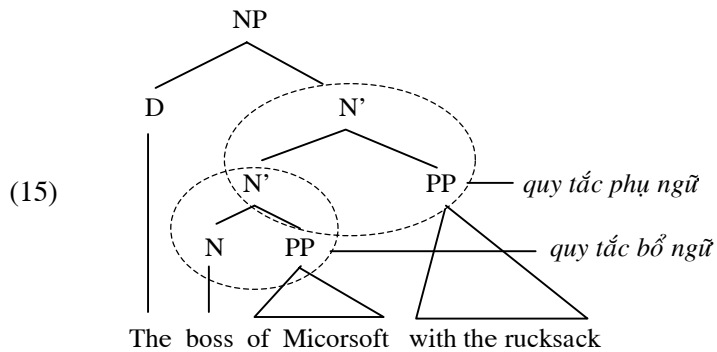
Theo kinh nghiệm, bổ ngữ thường đứng trực tiếp bên cạnh đối tượng muốn nói, tức là đầu, ví dụ ngữ đoạn danh từ NP dưới đây:

- (13) [NP The boss [of Microsoft] [with the rucksack]]
 đầu bổ ngữ phụ ngữ

Câu này nói được, nhưng nếu đổi vị trí bổ ngữ và phụ ngữ, câu sẽ có vấn đề:

- (14) * [NP The boss [with the rucksack] [of Microsoft]]
 đầu phụ ngữ bổ ngữ

Cây cú pháp của (13):

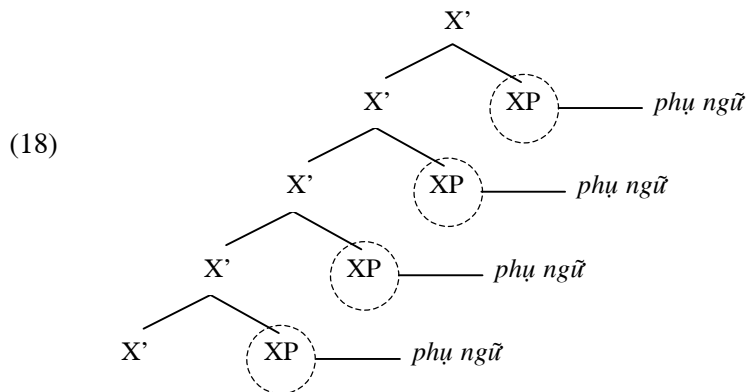


Quy tắc phụ ngữ còn chứa thêm một đặc điểm là tính *hồi quy*. Ngược lại, quy tắc bổ ngữ dường như không có tính này:

(16) Quy tắc phụ ngữ: $X' \rightarrow X' (XP)$

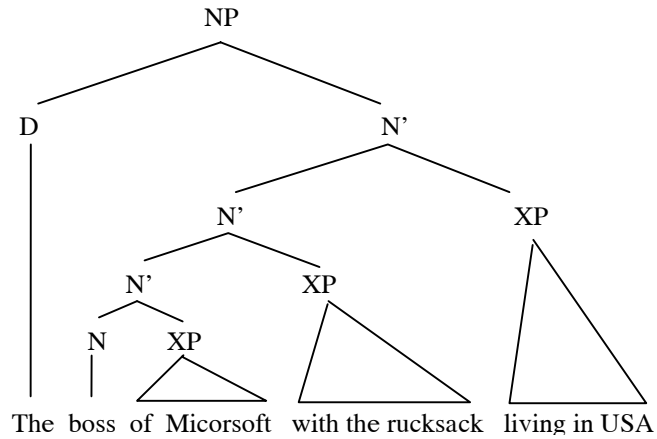
(17) Quy tắc bổ ngữ: $X' \rightarrow X (XP)$

Quy tắc bổ ngữ đi từ X' vào X và dừng ở đó. Trong khi đó, quy tắc phụ ngữ đi từ X' vào X' , X' lại phải sinh ra X' và cứ thế tiếp tục lặp lại chính nó, làm cấu trúc nở rộng:



Một trường hợp dùng nhiều phụ ngữ:

(19) [_{NP} The boss [of Microsoft] [with the rucksack] [living in USA]]
 đầu bổ ngữ phụ ngữ phụ ngữ



Xin lưu ý: XP được hiểu một cách tổng quát, không riêng gì PP trong ví dụ trên.

Thứ tự của bổ ngữ và phụ ngữ khá quan trọng. Ví dụ trên nói được, nhưng nếu đổi vị trí bổ ngữ và phụ ngữ như dưới đây, cả chuỗi từ sẽ không còn nói được:

- (20) * [_{NP} The boss [with the rucksack] [living in USA] [of Microsoft]]
 đầu phụ ngữ phụ ngữ bổ ngữ

Nhiều nhà ngôn ngữ học Tây phương cho rằng, vị trí giữa *bổ ngữ* và *phụ ngữ* không đổi được, nhưng hoán vị phụ ngữ thì được (Carnie 2000).

- (21) The book [of poems] [with a red cover] [from Blackwell] [by R. Burns]
 đầu bổ ngữ phụ ngữ phụ ngữ phụ ngữ

- (22) The book [of poems] from Blackwell with a red cover by R. Burns
 (23) The book [of poems] from Blackwell by R. Burns with a red cover
 (24) The book [of poems] by R. Burns from Blackwell with a red cover
 (25) The book [of poems] by R. Burns with a red cover from Blackwell
 (26) The book [of poems] with a red cover by R. Burns from Blackwell
 (27) * The book with a red cover [of poems] from Blackwell by R. Burns
 (28) * The book with a red cover from Blackwell [of poems] by R. Burns
 (29) * The book with a red cover from Blackwell by R. Burns [of poems]

Nhận định này có lẽ chỉ tương đối, bởi ngay trường hợp “[*The boss [of Microsoft] [with the rucksack] [made in China]*”, nếu đổi vị trí phụ ngữ, chúng ta sẽ gặp vấn đề:

- (30) * The boss [of Microsoft] [made in China] [with the rucksack]
 đầu bổ ngữ phụ ngữ phụ ngữ

Phép hoán vị còn dẫn đến trường hợp đa nghĩa. Chẳng hạn dùng ngữ đoạn giới từ PP “*from China*”:

- (31) The boss [of Microsoft] [with the rucksack] [from China]
 (32) The boss [of Microsoft] [from China] [with the rucksack]

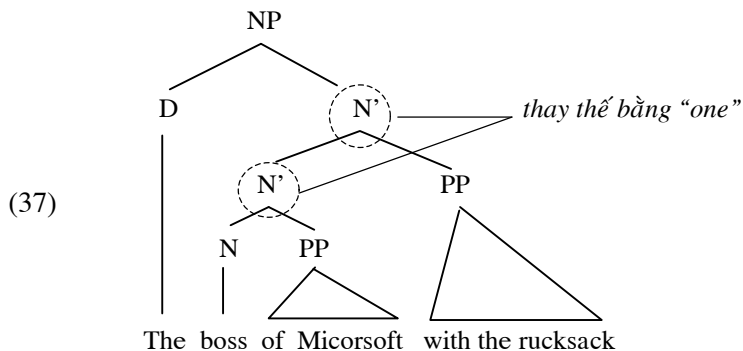
Có hai cách hiểu: hoặc cái ba lô của ông xếp là ba lô “*made in China*”, hoặc ông xếp xuất xứ từ Tàu. Hai đối tượng “*ba lô*” và “*ông xếp*” khác nhau.

Phép thế – “*one*” có thể thay thế ngữ đoạn trung gian N’:

- (33) The [_{N'} boss of Microsoft with the rucksack]
 (34) The *one*

Hoặc thay thế N’ thấp hơn:

- (35) The [_{N'} boss of Microsoft] with the rucksack
 (36) The *one* with the rucksack



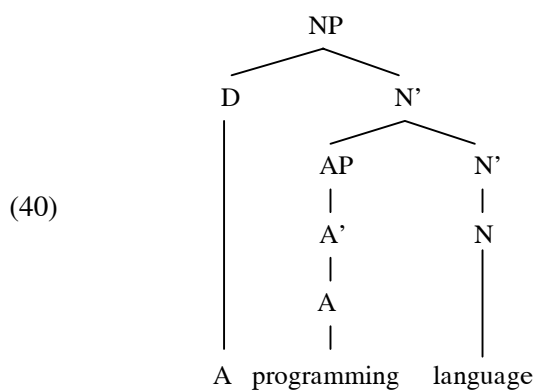
Phép thế – “*one*” không dùng được cho N:

- (38) * The one of Microsoft with the rucksack

Lưu ý: Khi giải thích ngữ pháp bằng sơ đồ X-gạch, một lỗi thường gặp là sự lẫn lộn giữa bổ ngữ và phụ ngữ.

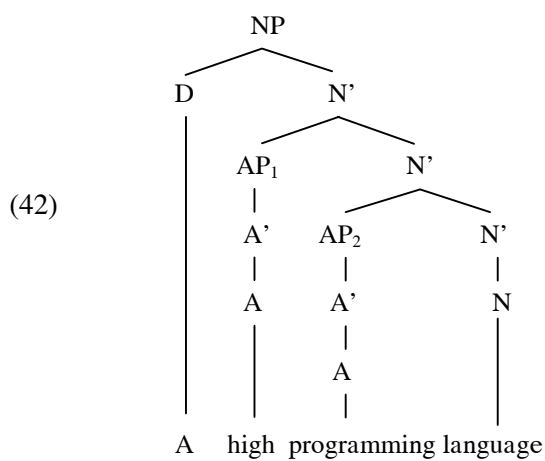
(39) A programming language

Nếu xem “*programming*” là bổ ngữ, thì cấu trúc X-gạch phải là:



Phụ thêm một tính chất ví dụ như “*high*” vào, thì “*high*” sẽ đóng vai phụ ngữ:

(41) A high programming language
 phụ ngữ bổ ngữ



Sau đây là vài nhận xét của ngôn ngữ học Tây phương về bổ ngữ và phụ ngữ trong tiếng Anh:

1. Một ngữ đoạn danh từ NP có thể có nhiều phụ ngữ nhưng chỉ một bổ ngữ:

- (43) An efficient high programming language

2. Bổ ngữ đứng gần đầu ngữ đoạn hơn phụ ngữ:

- (44) A high programming language
(45) * A programming high language

3. Tùy quan hệ lệ thuộc, phụ ngữ hoán vị được, nhưng vị trí bổ ngữ vẫn ở gần đầu ngữ đoạn:

- (46) A modern, efficient programming language
(47) An efficient, modern programming language

4. Phụ ngữ có thể đứng trước “one” trong phép thế- “one”, còn bổ ngữ thì không:

- (48) The efficient programm language but not the inefficient one
(49) * The programm language but not the natural one

5. Phụ ngữ có thể liên kết với phụ ngữ, bổ ngữ có thể liên kết với bổ ngữ, nhưng bổ ngữ không thể liên kết với phụ ngữ:

- (50) A strong and efficient high programming language
(51) A efficient projecting and programming language
(52) * A high and programming language

Nói thêm: Nếu kể về ngôn ngữ điện toán, thì *ngôn ngữ lập trình (programming language)* là một loại *ngôn ngữ hình thức (formal language)* dùng để viết chương trình. Mặt khác còn có loại ngôn ngữ đặc biệt được gọi là *ngôn ngữ lập đồ (projecting language)*. Loại này thường được dùng để lập đồ án trong kỹ thuật tự động hóa (automation technique). Hoặc có loại vừa dùng cho lập trình và vừa dùng cho lập đồ, ví dụ STEP 7. Trong trường hợp này, “*programming*” và “*projecting*” đều có chức năng bổ nghĩa: *projecting and programming language*. Trường hợp này rõ ràng, nhưng ở trường hợp dưới đây lại có vấn đề. Nếu nói:

- (53) A high programming language

thì “*high*” lại có chức năng phụ nghĩa. Thế nhưng nếu nói:

(54) An efficient high programming language

có lẽ “*high*” phải được coi là yếu tố bổ nghĩa. “*High*” không thể liên kết với “*efficient*” bằng cách nói:

(55) * An efficient and high programming language

Nhiều khi một danh từ có thể được bổ nghĩa cùng lúc. Khẳng định một từ làm bổ ngữ hoặc phụ ngữ có lẽ chỉ tương đối. Chẳng hạn như người Anh nói:

(56) the German teacher

Có hai cách hiểu, hoặc “ông thầy” (*teacher*) là người Đức (*German*), hoặc ông ta dạy tiếng Đức. Song nếu nói:

(57) the English German teacher

có lẽ xác suất hiểu “ông thầy Đức dạy tiếng Anh” cao hơn. Hay nếu nói:

(58) the English and German teacher

thì càng rõ nữa: ông thầy ấy vừa dạy tiếng Anh vừa dạy tiếng Đức, chứ khó có thể hiểu ông ta vừa lai Anh, vừa lai Đức. Ở đây, “*English*” bổ nghĩa cho “*teacher*”, “*German*” cũng vậy. “*English and German*” là một tổ hợp bổ ngữ. Trường hợp dùng “*and*” này có lý, nhưng ở trường hợp:

(59) A high programming language

“*programming*” lại bổ nghĩa cho “*language*”. “*High*” cũng vậy, vì khi nói đến “*high language*” nghĩa là muốn ám chỉ “*high programming language*”, song “*high*” không thể liên kết với “*programming*” bằng liên từ “*and*” để thành:

(60) * A high and programming language

Khá rắc rối. Mặt khác, nếu xem “*high*” làm bổ ngữ cho “*programming language*”, thì lại sai quy tắc X-gạch. Chỗ này không ổn lắm.

Tương tự tiếng Anh, những quy tắc X-gạch của ngữ đoạn danh từ NP cũng dùng được cho tiếng Việt. Chỉ khác là ngữ đoạn tính từ AP trong tiếng Việt nằm bên phải danh từ:

(61) $NP \rightarrow (D) N'$

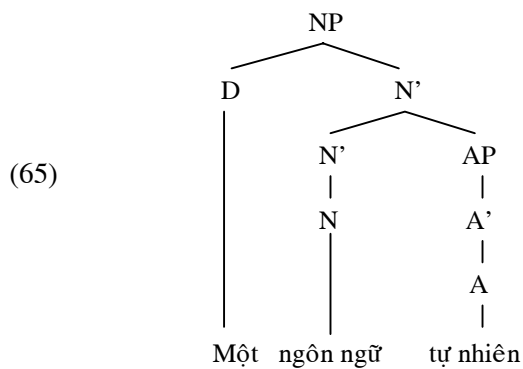
(62) $N' \rightarrow N' (AP) \mid N' (PP)$

(63) $N' \rightarrow N (PP)$

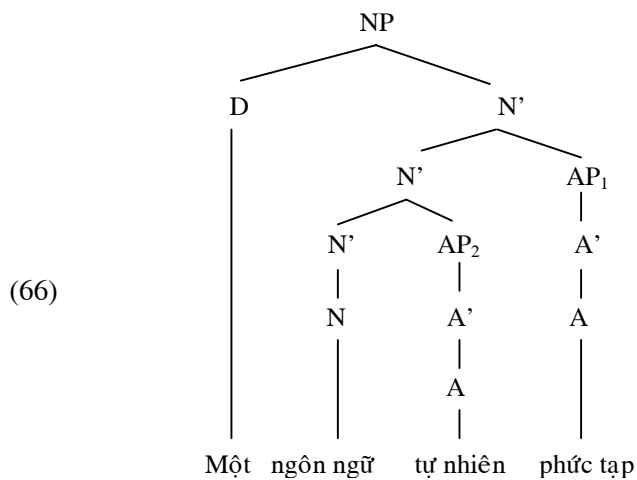
Xem một trường hợp na ná như ví dụ tiếng Anh bên trên:

(64) Tiếng Việt là một ngôn ngữ tự nhiên phức tạp.

Nếu xem “*tự nhiên*” làm bổ ngữ cho “*ngôn ngữ*”, ta có cấu trúc X-gạch như sau:

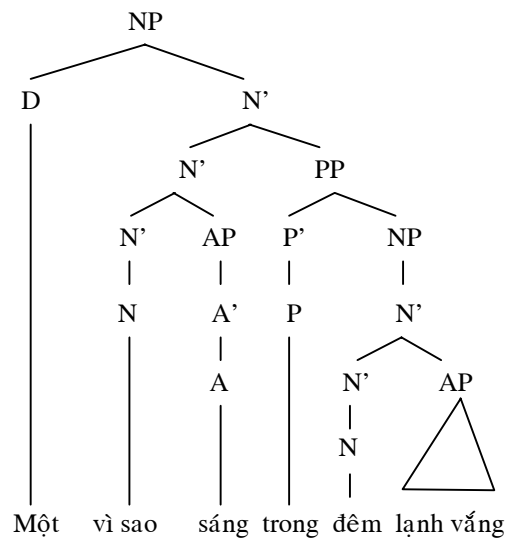


Cộng thêm tính từ “*phức tạp*” làm phụ ngữ:



Trong trường hợp này, quy tắc X-gạch thỏa tiếng Việt giống tiếng Anh. Xem một ví dụ khác chi tiết hơn:

- (67) “Một vì sao sáng trong đêm lạnh vắng”
(Thanh Trang)

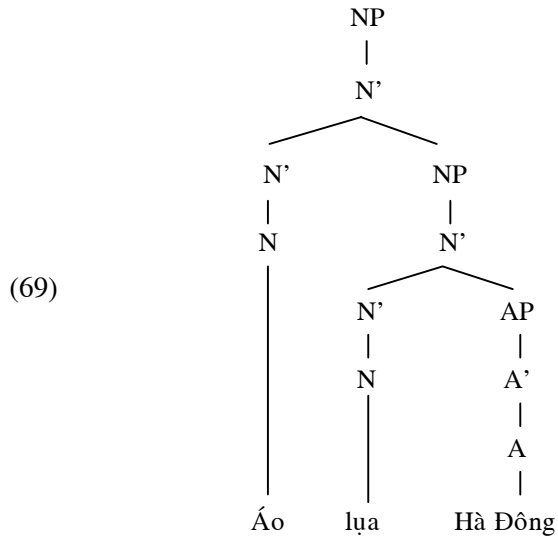


Cả trường hợp phức tạp chứa ngữ đoạn giới từ PP cũng thỏa.

Không những tiếng Việt chỉ thỏa những trường hợp trên mà còn có trường hợp một đối tượng bổ nghĩa cho một đối tượng:

- (68) “Nắng Sài Gòn, anh đi mà chợt mát,
Bởi vì em mặc áo lụa Hà Đông”
(Nguyễn Sa)

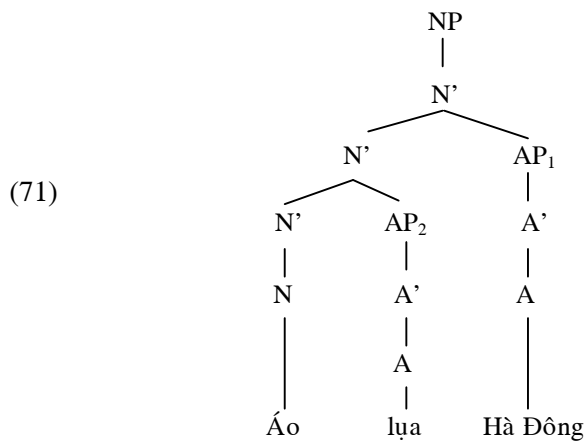
Ngữ đoạn “áo lụa Hà Đông” chứa hai đối tượng “áo” và “lụa”. “Lụa” bổ nghĩa cho “áo” để “áo” được đầy đủ ý nghĩa như một loại áo (*áo the, áo sơ mi, áo dài, áo bà ba, áo đầm, áo vét, áo cà sa, ...*). Mặt khác, nên hiểu rằng “Hà Đông” ở đây không phụ nghĩa cho đối tượng “áo” mà là cho đối tượng “lụa”. Tương tự áo, lụa cũng có nhiều loại khác nhau: “lụa Hà Đông”, “lụa Pháp”, “lụa Ấn”, ... Cả cụm từ “lụa Hà Đông” bổ nghĩa cho “áo”, chứ không riêng gì “lụa”. Xem cây cú pháp bên dưới, chúng ta sẽ thấy rõ chiều sâu ngữ nghĩa trong cấu trúc tiêu biểu này của tiếng Việt.



Cây cú pháp trên cần thêm một quy tắc mới:

(70) $N' \rightarrow N' (NP)$

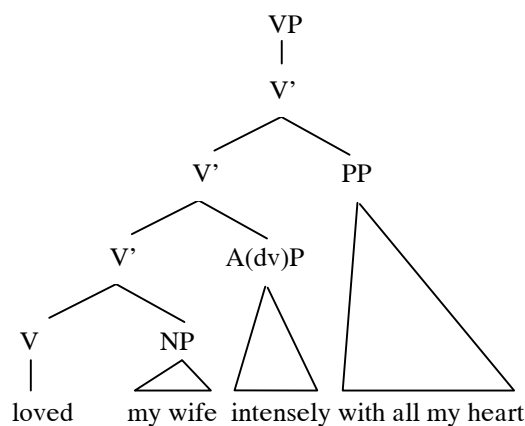
Hiểu như trên nghĩa là hiểu một loại “áo” được làm bằng vật liệu “lụa Hà Đông”. Ý nghĩa này khác với trường hợp: “Hà Đông” bổ nghĩa cho đối tượng “áo lụa”. Bởi áo lụa ấy chưa chắc là áo làm bằng “lụa Hà Đông” mà là mua ở Hà Đông thì sao. Trong trường hợp này, ta có một cấu trúc khác:



Vấn đề xác định bổ ngữ và phụ ngữ trong ngữ đoạn danh từ NP đôi khi không rõ ràng. Có rõ ràng hơn, có lẽ là ở ngữ đoạn động từ VP.

Theo các nhà ngôn ngữ học Tây phương, trong tiếng Anh, phần tử nào đứng ngay bên đông từ, phần tử ấy sẽ được xem là bổ ngữ.

- (72) He loved [my wife] [intensely] [with all my heart].
động từ ngữ đoạn danh từ NP trạng từ ngữ đoạn giới từ PP
bổ ngữ phụ ngữ phụ ngữ



Trong cấu trúc này, “*my wife*” đứng ngay sau động từ và làm bổ ngữ. Câu trên đúng vì không thể nói:

- (73) * I loved intensely my wife with all my heart.

Nói một cách tổng quát, bổ ngữ của tất cả phạm trù từ vựng (N, V, A, ...) là đối tượng ngữ nghĩa của đầu trong ngữ đoạn.

Vài ví dụ bổ ngữ của từ vựng (động từ, tính từ, danh từ) trong tiếng Anh:

- (74) He *fears* his wife. (verb)
 (75) He *is* afraid of his wife. (adjective)
 (76) He *has* a fear of his wife. (noun)

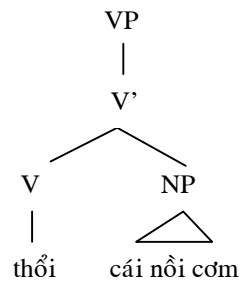
“(Of) his wife” trong tất cả các câu trên đều là bổ ngữ.

Tính phụ/bổ nghĩa trong Tiếng Anh có lẽ đơn giản hơn tiếng Việt. Mặc dù chương này không chủ ý bàn chi tiết, thế nhưng chúng ta hãy thử nhìn sơ qua vài hiện tượng phụ/bổ nghĩa cho động từ trong tiếng Việt:

1. Phụ/bổ ngữ là danh từ N hoặc ngữ đoạn danh từ NP:

(77) Lão Hạc thối [_{NP} cái nồi cơm], châm đóm. Tôi đã thông điều và bỏ điều rồi. (*Lão Hạc* – Nam Cao)

(78) Quy tắc X-gạch: $V' \rightarrow V (NP)$

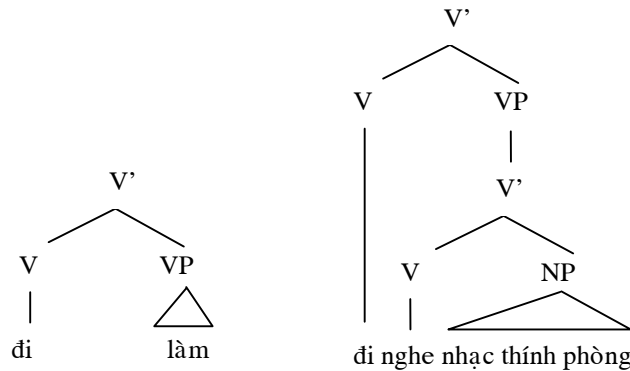


2. Phụ/bổ ngữ là động từ V hoặc ngữ đoạn động từ VP:

(79) Đi [_V làm], đi [_V chơi], đi [_V học], đi [_{VP} xem triển lãm], đi [_{VP} nghe nhạc thính phòng], đi [_{VP} coi đá banh], ...

(80) “Em đi [_{VP} thấy hội non sông],
Cười mê ánh sáng muôn lòng xuân xanh”.
(*Bên kia sông Đuống* – Hoàng Cầm)

(81) Quy tắc X-gạch: $V' \rightarrow V (VP)$



3. Phụ/bổ ngữ là tính từ, trạng từ hoặc ngữ đoạn tính từ AP:

(82) Câu chuyện này nghe $[_{AP} \text{hấp dẫn quá}]$.

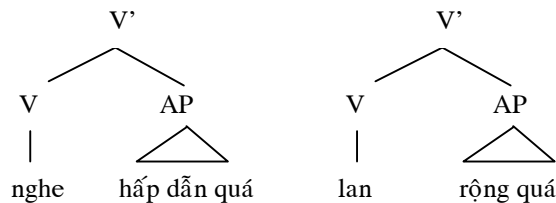
Xin nhắc lại. Tiếng Việt có một điểm đặc biệt cần lưu ý là nhiều khi động từ trong câu không hẳn thuộc về *chủ ngữ của câu* (tức chủ ngữ ngữ pháp (*grammatical subject*)). Động từ có thể liên quan tới một *chủ thể* (*subject*) nào đó bên ngoài câu. Ví dụ ở đây, “*câu chuyện*” (là chủ ngữ trong câu) tự nó không “*nghe*” được. Câu này ám chỉ là “chúng ta”, tức chủ thể đứng bên ngoài, mới “*nghe*” được. Hiện tượng này ít thấy ở những ngôn ngữ Âu châu. Động từ trong những ngôn ngữ Âu châu thường xuất phát từ chủ ngữ trong câu.

- Trường hợp chủ ngữ mang tính động:

(83) Câu chuyện này lan $[_{AP} \text{rộng quá}]$.

Để giải thích 2 câu trên, ta cần một quy tắc bổ ngữ thuộc dạng ngữ đoạn tính từ AP đứng sau động từ:

(84) Quy tắc X-gạch: $V' \rightarrow V \text{ AP}$



Trường hợp này cần bổ ngữ, vì câu sẽ không nói được nếu thiếu AP bổ nghĩa cho động từ:

(85) * Câu chuyện này *nghe*.

(86) * Câu chuyện này *lan*.

Tổng quát hơn, có trường hợp động từ không cần AP đứng sau, có thể đóng ngoặc ():

(87) Quy tắc X-gạch: $V' \rightarrow V (AP)$

- (88) Nó ngáy (*khò khò*).
 (89) Nó ngủ (*say sưa*).

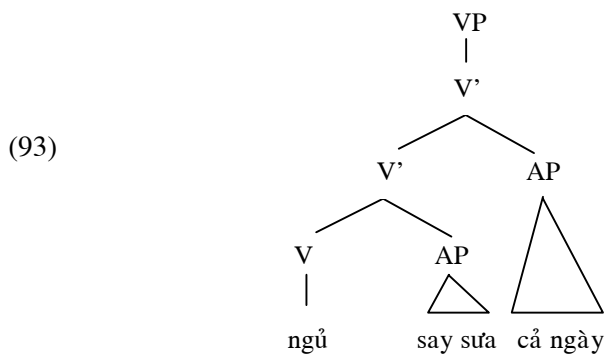
• Trường hợp AP phụ/bổ nghĩa cho động từ:

- (90) Nó uống *say sưa cả đêm*.
 (91) Nó ngủ *say sưa cả ngày*.

“Cả đêm”, “cả ngày” giữ nhiệm vụ phụ nghĩa. “Say sưa” giữ nhiệm vụ bổ nghĩa bởi “say sưa” thường đi với “ngủ”, “uống” hơn là “cả đêm”, “cả ngày”. Để diễn tả tính phụ nghĩa, chúng ta cần quy tắc:

- (92) Quy tắc X-gạch: $V' \rightarrow V' (AP)$

Cây cú pháp tương ứng của câu “nó ngủ say sưa cả ngày”:



4. Phụ/bổ ngữ là giới từ P hoặc ngữ đoạn giới từ PP:

- (94) Quy tắc X-gạch: $V' \rightarrow V' (PP)$ (giống tiếng Anh)
 (95) Hai vợ chồng ấy sống [_{PP} với nhau] [_{AP} đã lâu rồi].

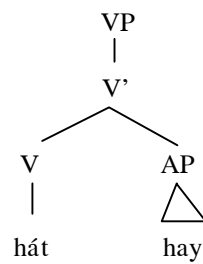
5. Phụ/bổ nghĩa bên phải động từ:

- (96) Quy tắc X-gạch: $V' \rightarrow V (NP)$

(97) Hương đã trở thành *một ca sĩ nổi tiếng*.

(98) *Quy tắc X-gạch:* $V' \rightarrow V (AP)$

(99) Cô ta hát *hay*.

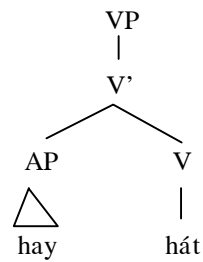


6. Trạng từ/tính từ phụ/bổ nghĩa *bên trái* động từ:

(100) Cô ta *hay* hát.

(101) Cô ta *chỉ* hát nhạc tiền chiến.

(102) *Quy tắc X-gạch:* $V' \rightarrow (AP) V'$



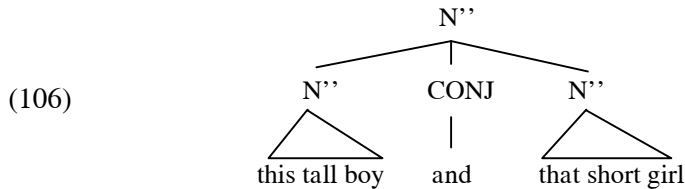
7. Trạng từ/tính từ phụ/bổ nghĩa *hai bên* động từ:

(103) Cô ta *hay* hát *lắm*.

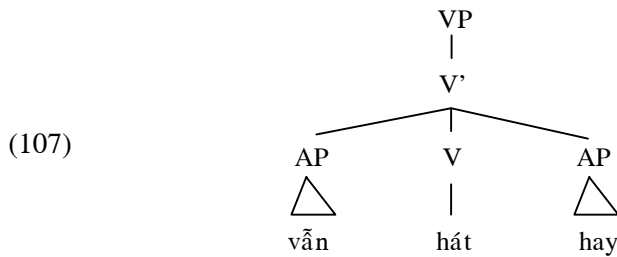
(104) Cô ta *vẫn* hát *hay*.

(105) Cô ta *vẫn* luyện giọng *thường xuyên*.

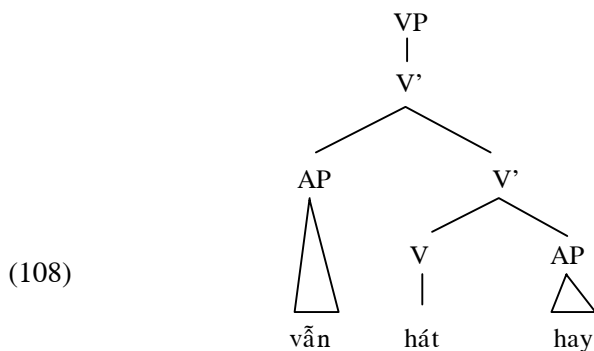
Cấu trúc X-gạch thường có dạng *nhị phân (binary)*, tức một nút không tận cùng (non-terminal) chỉ có hai nhánh. Trên thực tế còn có trường hợp phải cần đến ba nhánh, tiêu biểu là ngữ đoạn chứa *liên từ (conjunction)* (Radford 1985: 105):



Đây là một điểm chưa rõ vì có đến hai đầu N'' ("*this tall boy*" và "*that short girl*") trong khi cú pháp X-gạch chỉ cho phép có một đầu. Mặt khác, ở trường hợp trạng từ/tính từ phụ/bổ nghĩa "*vẫn hát hay*", nếu diễn tả cây cú pháp bằng quy tắc X-gạch: $V' \rightarrow (AP) V' (AP)$ như sau:



X-gạch sẽ mất ngay tính nhị phân, bởi vì nút V' sinh ra ba nhánh AP, V và AP. Không thấy thuyết X-gạch nhắc đến trường hợp này. Muốn hai nút AP không được nằm cùng bình diện, chỉ còn cách là đặt một AP vào bình diện phụ ngữ. Có lẽ AP bên trái thỏa tính chất này. Thử vẽ cây cú pháp cho giả thiết ấy:



Xét mặt ngữ nghĩa, hành động “*hát*” trong ví dụ nói lên hai ý nghĩa của hai trạng thái:

- Nếu nói “*hát hay*”, thì đó là một hành động “*hát*” có chất lượng, khác với “*hát dở*”, “*hát nghêu ngao*”, ... Từ “*hát*” được từ “*hay*” bổ túc ý nghĩa, nói lên chất lượng tốt. Cho nên “*hay*” có thể giữ chức năng như một bổ ngữ, nếu người ta muốn thế.
- Mặt khác, trạng thái “*vẫn hát*” diễn tả một hành động còn đang tiếp diễn. Ý nghĩa này khác với “*vẫn hát hay*”. “*Vẫn hát hay*” diễn tả hành động còn hát nhưng phải là hát hay. Ý nghĩa này hiển nhiên khác ý nghĩa “*vẫn hát*” nhưng chưa chắc còn hay nữa. Thế nên, “*vẫn*” trong cụm từ “*vẫn hát hay*” hình như giữ chức năng phụ nghĩa cho hành động “*hát hay*” nhiều hơn là bổ nghĩa. Tự bản thân “*hát hay*” – hiểu tổng quát - đã diễn tả tính liên tục của chất lượng tốt.

Nếu chấp nhận cách giải thích trên, ta có:

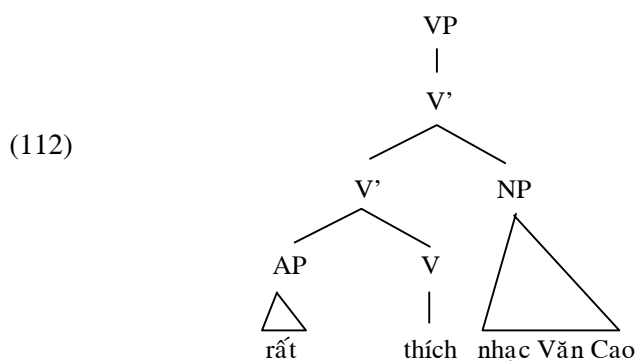
$$(109) \text{ Quy tắc X-gạch: } \begin{array}{l} V' \rightarrow (AP) V' \\ V' \rightarrow V (AP) \end{array}$$

Còn một hiện tượng khá phổ biến là trạng từ bổ nghĩa bên trái động từ. Tiêu biểu là trạng từ “*rất*”:

(110) Cô ta *rất* thích nhạc Văn Cao.

(111) Cô ta *rất* mê nhạc Văn Cao.

Trạng từ “*rất*” đứng trước động từ với cường vị của một bổ ngữ nhằm diễn tả hành động có cường độ. Trạng từ “*rất*” kết hợp với động từ thành một tổ hợp chặt chẽ. Cách nhấn mạnh ấy khiến yếu tố bổ nghĩa theo sau động từ yếu đi, làm như chỉ còn tác dụng phụ nghĩa. Theo thuyết X-gạch, ngữ đoạn “*nhạc Văn Cao*” lại được xem như bổ ngữ. Tiếng Việt lại vướng phải trường hợp bổ ngữ nằm bên trái lẫn bên phải động từ. Muốn thỏa thuyết X-gạch, chỉ còn cách, một là xem ngữ đoạn danh từ NP “*nhạc Văn Cao*” như một phụ ngữ của động từ và đương nhiên không thể chấp nhận được, hai là coi cả cụm từ “*rất + động từ*” như một VP và được NP “*nhạc Văn Cao*” bổ nghĩa.



- (113) *Quy tắc X-gạch:*
- $$V' \rightarrow V' (NP)$$
- $$V' \rightarrow (AP) V$$

8. Phụ/bổ nghĩa đầu câu:

- (114) *Thường thường*, cô ấy chỉ thích nhạc Văn Cao.
[Thường thường] là trạng ngữ (coi như một AP)

- (115) *Cả thơ tiền chiến*, cô ta cũng thích.

- (116) *Nhạc này* tôi không thích lắm.

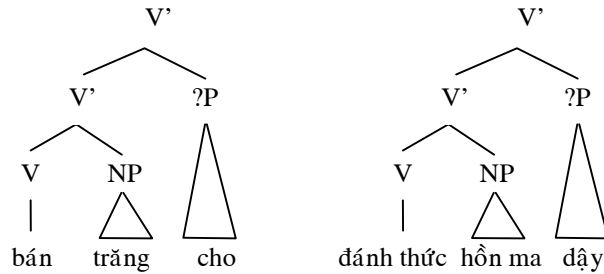
Xem phép chuyển vị (chương 12).

9. Phụ/bổ nghĩa cuối NP:

- (117) “Ai mua trăng, tôi *bán* trăng *cho*.
 Trăng nằm im trên cành liễu, đợi chờ”
 (Hàn Mặc Tử)

- (118) “Trời đã thu rồi em ở đâu.
 Nằm trong đáy mộ, chắc em sâu.
 Thu ơi, *đánh thức* hồn ma *dậy*.
 Ta muốn vào thăm đáy huyết sâu”
 (Đinh Hùng)

- (119) *Quy tắc X-gạch:* $V' \rightarrow V' (?P)$ (?P = cần định nghĩa)
 $V' \rightarrow V (NP)$



10. Ngữ đoạn giới từ PP phụ/bổ nghĩa sau VP:

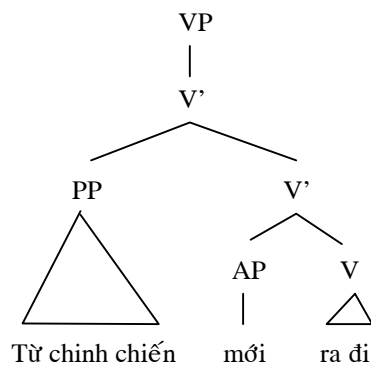
(120) Cô ta *đã học nhạc hồi còn nhỏ*.

(121) *Quy tắc X-gạch:* $V' \rightarrow V' (PP)$
 $V' \rightarrow V (NP)$

11. Ngữ đoạn giới từ PP phụ/bổ ngữ trước VP:

(122) “Em ở thành Sơn chạy giặc về.
 Tôi *từ chinh chiến* mới ra đi.
 Cách biệt bao ngày quê Bất Bạt.
 Chiều xanh không thấy bóng Ba Vì”.
 (Quang Dũng)

(123) *Quy tắc X-gạch:* $V' \rightarrow (PP) V'$
 $V' \rightarrow (AP) V$



Như trên là một vài đặc điểm của tính phụ/bổ nghĩa trong tiếng Việt. Sau khi thử nghiệm dùng các quy tắc được lập riêng cho tiếng Anh, chúng ta thấy tiếng Việt có lẽ không đơn giản, cụ thể là ngữ đoạn danh từ và động từ. Cần lưu ý là những quy tắc đưa ra cần được kiểm nghiệm kỹ càng hơn. Có nhiều trường hợp cấu trúc giống nhau song ngữ nghĩa lại khác.

BIỆT ĐỊNH NGỮ

Chúng ta đã xem qua hai quy tắc cấu trúc ngữ đoạn X-gạch dùng cho tiếng Anh:

(124) *Quy tắc phụ ngữ:* $X' \rightarrow X' (ZP) \mid (ZP) X'$

(125) *Quy tắc bổ ngữ:* $X' \rightarrow X (WP)$

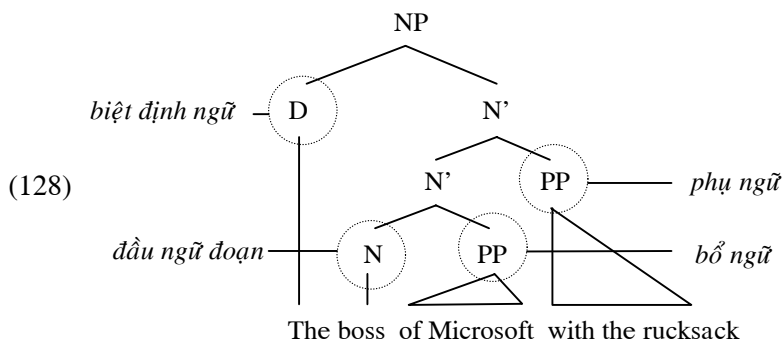
Cùng mục đích thêm nghĩa cho cái muốn nói, các nhà ngôn ngữ học Tây phương còn đề nghị thêm một quy tắc, tạm gọi là *biệt định ngữ* (*specifier*):

(126) *Quy tắc biệt định ngữ:* $X' \rightarrow (YP) X'$

Với ví dụ cũ:

(127) [The] [boss] [of Microsoft] [with the rucksack].
biệt định ngữ đầu bổ ngữ phụ ngữ

ta có cây cú pháp:



Định nghĩa:

Biệt định ngữ (*specifier*) là một ngữ đoạn X-gạch có quan hệ *chị em* với một bình diện X' và là *con* của một ngữ đoạn XP.

Ở ví dụ trên, D có quan hệ chị em với N' nằm bên phải cùng bình diện. D đồng thời là con của ngữ đoạn danh từ NP.

Sự khác biệt của biệt định ngữ trong tiếng Anh đối với bổ ngữ và phụ ngữ là không có tính hồi quy. Không thể nói:

(129) * The this big boss

Bởi muốn phân tích chi tiết cấu trúc ngữ pháp, người ta đã giải thích trường hợp có chỉ định từ “*the*” của tiếng Anh bằng cách coi như *biệt định ngữ* (ngoài “*the*” còn có *a, that, this, those, these, ...*). Ở đây, tiếng Anh sẽ gặp vấn đề. Trong ví dụ trên, cả ngữ đoạn danh từ NP “*boss of Microsoft with the rucksack*” làm bổ ngữ cho “*the*”, tức “*the*” là đầu trong ngữ đoạn. Điều này mâu thuẫn với lý thuyết X-gạch, bởi tất cả mọi phần tử không phải là *đầu* đều là *ngữ đoạn*. Đầu trong ngữ đoạn danh từ NP chỉ có thể là một danh từ. Thế thì thực chất của “*the*” trong ngữ đoạn danh từ NP là gì ? Chúng ta sẽ tìm hiểu dưới đây.

NGỮ ĐOẠN CHỈ ĐỊNH TỪ

Theo khám phá của giới ngôn ngữ học Tây phương, có một điều đáng ngạc nhiên là trong tiếng Anh, chỉ định từ “*the*” không nằm trong ngữ đoạn danh từ NP. Điểm này có thể nhìn thấy trong sinh cách (còn gọi là sở hữu cách):

- **sinh cách-“*of*” (*of genitive*)** hoặc **sinh cách tự do (*free genitive*)**
- **sinh cách-“*'s*” (*'s genitive*).**

Sinh cách tự do sử dụng giới từ “*of*” để diễn tả nét *quan hệ sở hữu* giữa 2 ngữ đoạn danh từ NP:

(130) The rucksack of the boss

(131) The rucksack of the boss standing over there

Mặt khác, sinh cách-“*'s*” đóng vai trò như một *dấu hiệu sở hữu* nằm ngay sau ngữ đoạn danh từ NP:

(132) [_{NP} The boss standing over there]’s rucksack

Vì thế mà không thể nói:

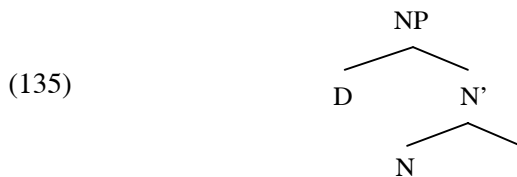
(133) * The boss’s standing over there rucksack

hoặc:

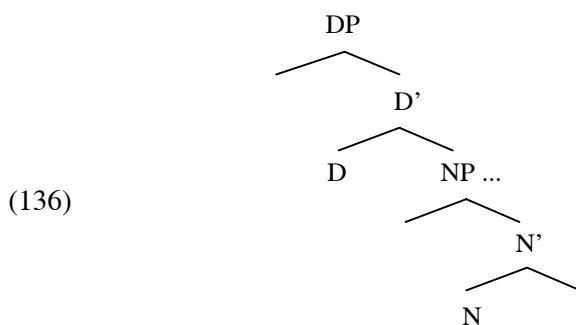
(134) * The boss’s the rucksack

Sinh cách-“’s” không cho phép 2 danh từ đều có *chỉ định từ* (*determiner*). Để giải thích điểm này, Abney đã đưa ra giả thuyết, chỉ định từ không nằm trong ngữ đoạn danh từ NP mà chỉ là *đầu* của một phép chiếu ngữ đoạn của chính nó (Abney 1987):

Nếu nhìn theo cách cũ:



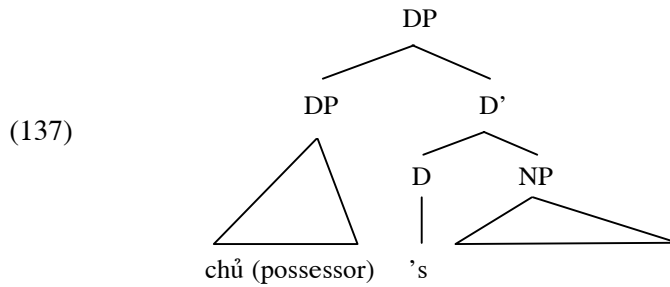
thì D là nút tận cùng và chỉ có thể tận cùng bằng một chỉ định từ. D không sản sinh ra quy tắc chứa biến số. Theo giả thuyết của Abney, D chính là ngữ đoạn DP:



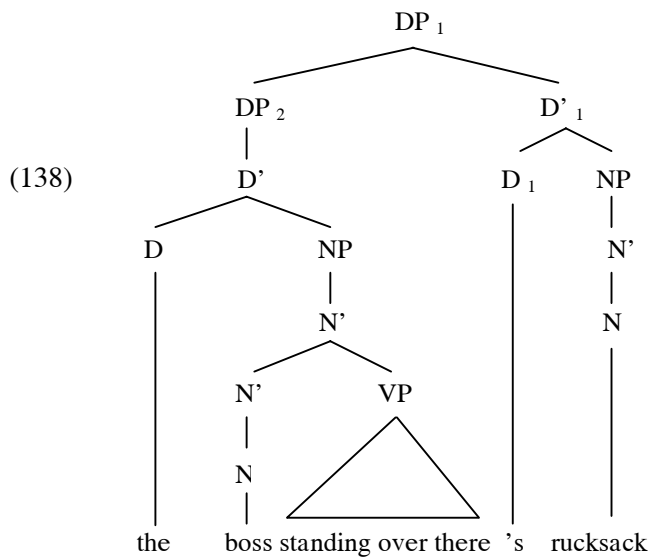
Giả thuyết này xem ra có phần hữu lý, bởi người ta có thể giải thích trường hợp sinh cách-“’s” dựa vào nhận định: chỉ định từ và sinh cách-“’s” cùng nằm trong

thế *phân bố bổ sung* (*complementary distribution*). Đối với ngôn ngữ học, hai phần tử nằm trong thế phân bố bổ sung khi chúng cùng là thực thể của một vật giống nhau. Ví dụ, đàn ông, đàn bà là hai thực thể cùng nằm trong một *lớp* (*class*): lớp con người. Đàn ông, đàn bà khác nhau thế nào, ai cũng biết, thế nhưng sự khác biệt chỉ là hình thức, chức năng chi tiết thuộc về cá thể không đáng nói. Điều muốn nói là những điểm chung, con người nào cũng có như tóc, tai, mặt, mũi, đầu mình, tứ chi, ... Chỉ định từ và “’s” cũng thế. Chúng chỉ khác nhau về danh tính thuộc một loại giống nhau.

Giả sử “’s” là chỉ định từ và giả sử giả thuyết ngữ đoạn chỉ định từ DP đúng, thì vị trí tương đối của “’s” đối với phần *chủ* (*possessor*) có thể giải thích được. Phần tử “’s” được đưa vào vị trí của đầu D. Phần chủ còn lại nằm trong biệt định ngữ:



Một ví dụ cụ thể:



Phần chủ DP_2 thuộc biệt định ngữ của DP_1 có đầu là “’s”. Nói cách khác, “*the boss standing over there*” thế cho biệt định ngữ của DP_1 . Cách giải thích thỏa điều kiện lý thuyết X-gạch: tất cả phần nào không phải là *đầu*, đều là *ngữ đoạn*. Ngữ đoạn có đầu là một chỉ định từ được gọi là **ngữ đoạn chỉ định từ** (**determiner phrase**), viết tắt **DP**, hoặc chỉ định ngữ. Quy tắc X-gạch tương ứng:

(139) $DP \rightarrow DP D'$

(140) $D' \rightarrow D NP$

Vấn đề không đơn giản. Kể cả tiếng Việt.

Tiếng Việt cũng có sở hữu cách như mọi ngôn ngữ khác. Có trường hợp, tiếng Việt dùng giới từ “*của*” (tương đương “*of*” của tiếng Anh), có trường hợp không:

(141) “*Nhà tôi* bên chiếc cầu soi nước,
Em đến tôi một lần”
(*Bến xuân* – Văn Cao)

“*Nhà tôi*” được hiểu là “*nhà của tôi*”.

(142) Tôi đã ra đời và lớn lên suốt cả thời thơ ấu trên một mỏm đất đầu hiu nằm bên bờ biển Đông miền Nam nước Việt. Thực ra đây không phải là *nguyên quán của tôi*, vì Ba tôi người Bắc, còn Mẹ tôi người Huế.

Mặc dù tiếng Việt không gặp rắc rối vì *sinh cách*–“’s” như tiếng Anh hoặc tiếng Đức nhưng không phải là không có trường hợp phức tạp đòi hỏi người nghe phải hiểu ngầm cái nào thuộc về cái nào. Ví dụ ở đoạn văn trên, câu đầu không có “*của*”. Cấu trúc câu có thể được diễn giải bằng 2 cách khác nhau:

(143) [NP mỏm đất đầu hiu nằm bên bờ biển Đông] (*của*) [NP miền Nam nước Việt]

(144) mỏm đất đầu hiu nằm bên [NP [NP bờ biển Đông] (*của*) [NP miền Nam nước Việt]]

Điểm cần được làm nổi bật là 2 đối tượng chính: “*mỏm đất*” và “*bờ biển*”. Một cấu trúc diễn tả “*mỏm đất ...*” của “*miền Nam nước Việt*”, một cấu trúc diễn tả “*bờ biển ...*” của “*miền Nam nước Việt*”. Ranh giới của ngữ đoạn thiếu sắc bén dễ dẫn đến hiện tượng đa nghĩa.

TÓM LƯỢC

Nguyên tắc X-gạch cho phép giải thích các ngữ đoạn trừu tượng (bổ ngữ, phụ ngữ, biệt định ngữ, chỉ định ngữ)

Bổ ngữ có chức năng bổ nghĩa cho cái muốn nói đầy đủ ý nghĩa. Phụ ngữ có chức năng thêm nghĩa nhưng không cần thiết. Bổ ngữ quan trọng hơn phụ ngữ.

NHẬN XÉT VÀ ĐIỂM MỞ

- 1) Biệt định ngữ, chỉ định ngữ trong tiếng Việt ra sao ?
- 2) Ngược với tiếng Anh, tiếng Đức, một ngữ đoạn danh từ NP tiếng Việt được thêm nghĩa bên phải; tất cả làm thành một ngữ đoạn danh từ NP lớn hơn. Vậy, làm sao biết được đó là một ngữ đoạn danh từ hay một câu/mệnh đề. Ví dụ:
 - a. Ba người anh đi bộ đội lâu rồi.
 - b. Nàng có [_{NP} ba người anh đi bộ đội lâu rồi].

Chương 10

LÝ THUYẾT X-GẠCH –
CÁC LOẠI NGỮ ĐOẠN KHÁC

THÔNG SỐ VÀ THỨ TỰ CỦA TỪ

Bằng nhiều ví dụ cụ thể, chúng ta đã xoay quanh 3 quy tắc X-gạch trong chương rồi:

- (1) *Quy tắc biệt định ngữ:* $X' \rightarrow (YP) X'$
- (2) *Quy tắc phụ ngữ:* $X' \rightarrow X' (ZP) \mid (ZP) X'$
- (3) *Quy tắc bổ ngữ:* $X' \rightarrow X (WP)$

Những quy tắc trên chủ yếu dựa vào tiếng Anh. Trên thực tế, có nhiều ngôn ngữ không theo các quy tắc đó, điển hình là tiếng Việt. Hay vài ví dụ khác:

Tiếng Nhật:

- (4) [Phong-ga] [sensei] [desu].
Chủ ngữ bổ ngữ vị từ
(Phong là thầy giáo)
- (5) Hung-ga ookii desu.
(Hùng (thì) lớn)
- (6) Phong-ga Hung-o mita.
(Phong (đã) thấy Hùng)
- (7) Hung-ga yonda.
(Hùng đọc)
- (8) Hung-ga hon-o yonda.
(Hùng đọc cuốn sách)

Tiếng Thổ Nhĩ Kỳ:

- (9) [Can] [kitab-i] [oku-du].
Chủ ngữ bổ ngữ vị từ
 (Can (đã) đọc cuốn sách)

Thứ tự vị trí bổ ngữ và vị từ của câu tiếng Nhật và tiếng Thổ Nhĩ Kỳ không giống như của tiếng Anh. Đó là những ngôn ngữ có cấu trúc Subject-Object-Verb (SOV) thay vì Subject-Verb-Object (SVO). Trên thực tế, vị trí của bổ ngữ, phụ ngữ hoặc biệt định ngữ của từng ngôn ngữ có thể khác nhau mà các quy tắc trên không giải thích được. Do đó, quy tắc X-gạch đã được đề nghị tổng quát hóa như sau (Carnie 2000):

- (10) *Quy tắc biệt định ngữ:* $X' \rightarrow X' (YP) \mid (YP) X'$
 (11) *Quy tắc phụ ngữ:* $X' \rightarrow X' (ZP) \mid (ZP) X'$
 (12) *Quy tắc bổ ngữ:* $X' \rightarrow X (WP) \mid (WP) X$

Quy tắc X-gạch tổng quát bao gồm hai trường hợp, vị trí của ngữ đoạn (YP, ZP, WP) hoặc nằm *bên phải* hoặc nằm *bên trái* biến số X' .

Xét cho cùng, sáng kiến chính của lý thuyết X-gạch là dùng bình diện trung gian để giải thích chiều sâu của cấu trúc ngữ pháp. Muốn ứng dụng nó, mỗi ngôn ngữ phải tự giải thích những trường hợp của mình sao cho thích hợp. Vị trí của từ chỉ là **thông số** (*parameter*) (Travis 1984).

NGỮ ĐOẠN THÌ VÀ NGỮ ĐOẠN TÁC TỬ BỔ NGỮ HÓA

Trên lý thuyết, lý thuyết X-gạch có thể giải thích được hầu hết các loại ngữ đoạn giữ chức năng thêm nghĩa. Tuy vậy vẫn còn lại hai vấn đề gây trở ngại, kể cả đối với bản thân tiếng Anh: *mệnh đề phụ* và *câu có thì*:

- (13) *Mệnh đề phụ:* $S' \rightarrow (C) S$
 (14) *Câu có thì:* $S \rightarrow NP (T) VP$

(Xin nhắc lại: T viết tắt của “*tense*” (thì). C viết tắt của “*complementizer*”, tạm dịch là *tác tử bổ ngữ hóa*).

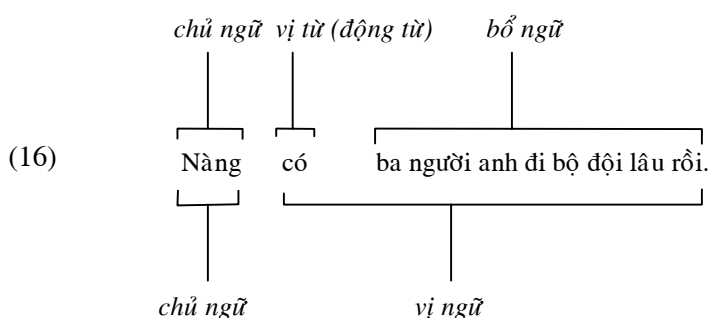
Sau đây, chúng ta sẽ thông qua một vài khái niệm trước khi vào chi tiết.

1. Mệnh đề

Theo quan niệm của giới ngôn ngữ học Tây phương thì *mệnh đề* (*clause*) được hiểu như một chuỗi từ gồm **chủ ngữ** (*subject*) cộng **vị ngữ** (*predicate phrase*). Chủ ngữ thường là một ngữ đoạn danh từ NP. Vị ngữ là ngữ đoạn động từ VP, thường được hiểu nôm na là động từ cộng với túc từ trong ngôn ngữ học cổ điển.

- (15) “*Nàng có ba người anh đi bộ đội lâu rồi*”
(*Màu tím hoa sim* -Thơ Hữu Loan, Phạm Duy phổ nhạc)

Trong câu trên, “*nàng*” là chủ ngữ, “*có ba người anh đi bộ đội lâu rồi*” là vị ngữ. Vị ngữ gồm động từ “*có*” cộng với phần bổ ngữ còn lại (“*ba người anh đi bộ đội lâu rồi*”). Xin lưu ý, thay vì gọi là động từ, chúng ta sẽ gọi là *vị từ* (*predicate*).



Một câu đầy đủ chủ ngữ, vị ngữ (không cần thêm gì nữa) được coi là **mệnh đề chính** (*main clause*). Một mệnh đề chính có thể chứa **mệnh đề phụ** (*subordinate clause*) (thường gọi là *câu phụ*). Mệnh đề phụ còn được (tạm) gọi là **mệnh đề lồng** (*embedded clause*), tức được lồng vào mệnh đề chính. Vài ví dụ tiếng Anh, tiếng Đức (đồng nghĩa):

- (17) She said [that she had three brothers].
(18) Sie sagte [dass sie drei Brüder hatte].

Cả hai câu trên đều là mệnh đề chính. Phần đóng trong ngoặc [...] là mệnh đề phụ.

Nếu hiểu câu “*Nàng có ba người anh đi bộ đội*” là có một người con gái nào đó mà có ba người anh đều đi bộ đội cả, ta có thể tưởng tượng ra một câu theo kiểu tiếng Anh, tiếng Đức như sau:

(19) She had three brothers, [*who* were soldiers]. (tiếng Anh)

(20) Sie hatte drei Brüder, [*die* Soldaten waren]. (tiếng Đức)

Sao chép cấu trúc cú pháp trên sang tiếng Việt, ta có:

(21) Nàng có ba người anh, [(*những*) người đi bộ đội lâu rồi].

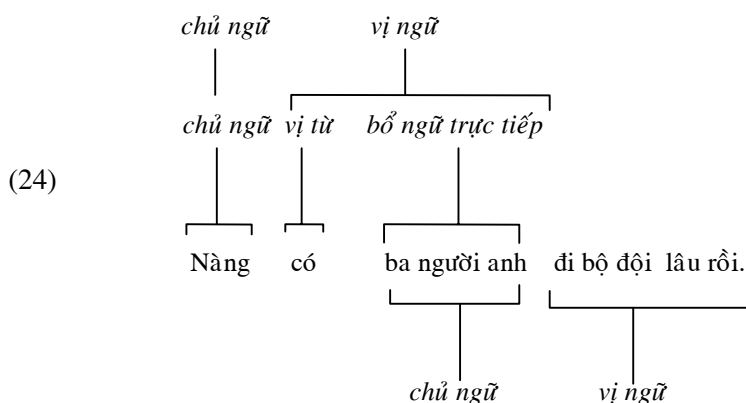
Thế nhưng người Việt không nói vậy mà thường là:

(22) Nàng có ba người anh [(*mà* đã) đi bộ đội lâu rồi].

“Mà” ở đây có chức năng giới thiệu phần nói thêm (phần diễn) cho “ba người anh”. Còn trong tiếng Anh, Đức, người ta dùng “*who*”, “*die*” để ám chỉ “ba người anh” đó và được dùng làm chủ ngữ trong mệnh đề phụ. Đây là một điểm khác biệt về chức năng ngữ pháp. Điểm khác biệt thứ hai là nếu lược bỏ “mà”, ta sẽ được một cấu trúc khác:

(23) Nàng có ba người anh đi bộ đội lâu rồi.

Cách lược bỏ “mà” như trên dẫn đến một cấu trúc đặc biệt là *bổ ngữ trực tiếp* tức “*direct object*”¹⁾ lại làm *chủ ngữ* cho mệnh đề (phụ) theo sau:



Câu/mệnh đề trong tiếng Anh chủ yếu bắt đầu bằng một chủ ngữ ngữ pháp và chấm dứt bằng vị ngữ. Nếu muốn diễn tả thêm về một đối tượng, người ta có thể

¹⁾ Xưa nay thường gọi là túc từ trực tiếp. Có tác giả còn gọi là *tân ngữ*. Tân nghĩa là “khách” (*object*) (Phan Khôi 1997).

dùng mệnh đề phụ. Mệnh đề phụ cũng có cấu trúc chủ vị với tính chất như trên. Tiếng Việt cũng có cấu trúc chủ vị song lại có những điểm khác biệt so với ngôn ngữ Âu châu.

Thử xét thêm vài ví dụ nữa để xem cấu trúc chủ vị của lý thuyết Tây phương có thực sự giải thích được hết cấu trúc tiếng Việt hay không.

Trong tiếng Việt có trường hợp, đối tượng trong câu chưa chắc đã giữ đúng vai trò của một bổ ngữ hoặc chủ ngữ đối với vị từ (động từ):

(25) Bà Năm mới mua cái nhà mà đã cháy mất tiêu.

Động từ “*mua*” trong câu dùng cho chủ ngữ “*bà Năm*”. “*Đã cháy*” dùng cho “*cái nhà*”. Ngay đây, cấu trúc sâu và cấu trúc bề mặt giống nhau. Cấu trúc sâu quyết định ngữ nghĩa thực của phát biểu. Cấu trúc bề mặt là cấu trúc của câu nói thực tế.

Cùng một cấu trúc ngữ pháp y hệt vậy, nếu thay thế động từ “*cháy*” bằng “*bán*”, thì động từ “*bán*” lại không diễn tả hành động của “*cái nhà*” mà là hành động của “*bà Năm*”.

(26) Bà Năm mới mua cái nhà mà đã *bán* mất tiêu.

Sự vắng mặt của chủ thể thật (ngoại chủ thể) cũng là một lối nói thông thường của người Việt:

(27) Mới mua cái nhà mà đã *bán* mất tiêu.

Nghĩa là ai đó (ví dụ như bà Năm) mới mua cái nhà nhưng đã bán mất rồi. Ai đó là chủ thể thật nhưng không nằm trong câu, chỉ có người hành ngôn mới biết.

(28) Mới mua cái nhà mà đã *cháy* mất tiêu.

Trong câu, vị từ “*cháy*” dùng cho “*cái nhà*”, nhưng vị từ “*mua*” lại thiếu chủ ngữ. Ngay đây, thuyết vị-ngữ cũng không giải thích được.

Tương tự là những cách nói:

(29) Quần áo đang bán xon kìa. Đi mua đi !

Người nghe có thể hiểu ngầm, ngoại chủ thể là “*người ta*”, “*siêu thị*”, “*chợ trời*”... :

- (30) Quần áo (người ta) đang bán xon kìa. Đi mua đi !
- (31) Quần áo (siêu thị) đang bán xon kìa. Đi mua đi !
- (32) Quần áo (chợ trời) đang bán xon kìa. Đi mua đi !

Hoặc:

- (33) Nhạc nào hát cũng được.
- (34) Nhạc nào (cổ) hát cũng được.
- (35) Nhạc nào (ổng) hát cũng được.
- (36) Nhạc nào (chúng tôi) hát cũng được.
- (37) Nhạc nào (họ) hát cũng được.
- (38) Nhạc nào (ca đoàn) hát cũng được.

- (39) Món gì ăn cũng được.
- (40) Món gì (nó) ăn cũng được.
- (41) Món gì (bả) ăn cũng được.
- (42) Món gì (tụi tôi) ăn cũng được.
- (43) Món gì (kẻ thiếu ăn) ăn cũng được.
- (44) Cái gì (ông tỉnh trưởng) ăn cũng được.
- (45) Cái gì (trâu bò) ăn cũng được.

Không dùng chủ ngữ (chủ từ) ngữ pháp, người Việt có cơ hội diễn tả nội dung bằng nhiều ngoại chủ thể khác nhau. Đó là điểm lợi, ví dụ hình thức nói bóng gió, thế nhưng cũng là điểm hại, bởi sự sử dụng ngôn ngữ một cách phóng khoáng thường dễ sinh ra những câu thiếu sắc bén, đa nghĩa. Ví dụ câu:

- (46) Người nào cũng chửi.

Đây là một câu hoàn chỉnh. Chủ ngữ “*người nào*” thực sự làm chủ động từ “*chửi*”. Nhưng trong một hoàn cảnh nào đó, nếu hiểu theo kiểu, câu có ngoại chủ thể, thì “*người nào*” sẽ không còn làm chủ động từ “*chửi*” nữa:

- (47) Người nào (nó) cũng chửi.
- (48) Người nào (Chí Phèo) cũng chửi.

Hoặc còn có trường hợp chẳng biết ai là ai; chủ thể được hiểu chung chung là người ta, chúng ta, một người quan sát chung chung nào đó đang tường thuật:

- (49) “Trải qua một cuộc bể dâu.
Những điều trông thấy mà đau đớn lòng”
(*Kiều* – Nguyễn Du)

Ai trải qua một cuộc bể dâu ? Ai trông thấy mà đau đớn lòng ? Câu không nói. Chắc chắn “*trông thấy*” không phải là vị từ của “*những điều*”.

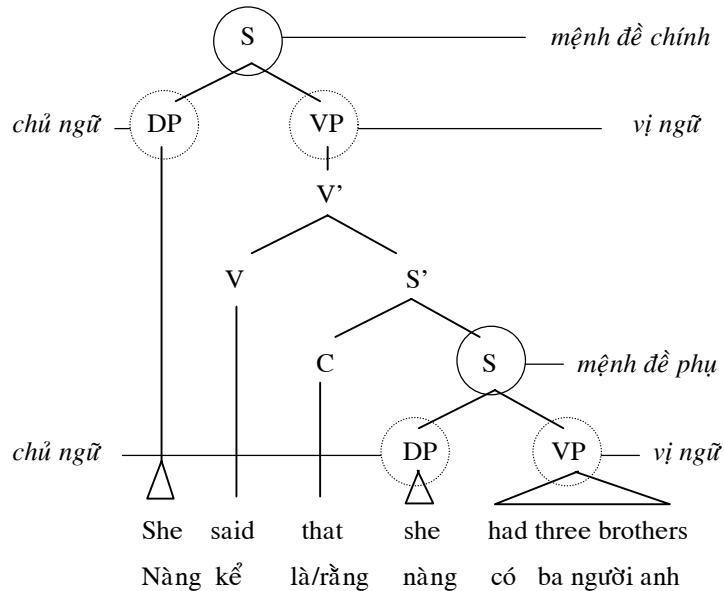
Nói tóm lại, câu/mệnh đề của tiếng Việt có những điểm khác những tiếng Âu châu. Nói thế không có nghĩa là khác hoàn toàn. Tất nhiên cũng có trường hợp, tiếng Việt vẫn nói được y hệt như tiếng khác. Ví dụ, người Anh nói:

- (50) She said [that she had three brothers].

Người Việt cũng có thể nói:

- (51) Nàng kể [rằng nàng có ba người anh].
(52) Nàng kể [là nàng có ba người anh].

Cây cú pháp tương ứng:



2. Ngữ đoạn tác tử bổ ngữ hóa CP

Bây giờ chúng ta hãy trở lại hai vấn đề vẫn còn gây trở ngại cho lý thuyết X-gạch để xem các nhà ngôn ngữ học Tây phương giải quyết ra sao.

(53) *Mệnh đề phụ:* $S' \rightarrow (C) S$

(54) *Câu có thì:* $S \rightarrow NP (T) VP$

Quy tắc mệnh đề phụ đã chỉ ra rằng, bản thân S' là một câu nằm dưới dạng mệnh đề phụ, có nhiệm vụ bổ nghĩa cho mệnh đề chính. Trong mệnh đề phụ S' , phần tử C , hãy tạm gọi, là *tác tử bổ ngữ hóa* (*complementizer*) và loại ngữ đoạn này là **ngữ đoạn tác tử bổ ngữ hóa** (***complementizer phrase***), ký hiệu là **CP**.

Để ứng dụng sơ đồ X-gạch, chúng ta thay thế S' bằng CP:

(55) $CP \rightarrow C'$

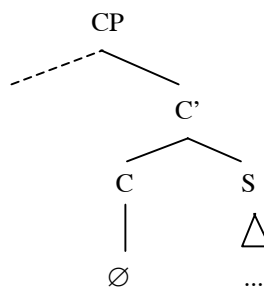
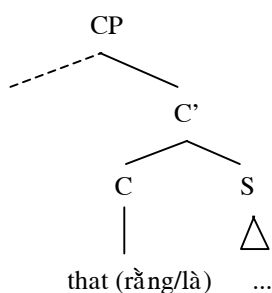
(56) $C' \rightarrow (C) S$

Ngữ đoạn trung gian C' được xem như bình diện X-gạch thỏa quy tắc $C' \rightarrow (C) S$. Phần tử C mang giá trị “*that*” (*rằng/là*). (Giá trị này chỉ là một chọn lựa, không bắt buộc phải có. Nếu không cần “*that*” (*rằng/là*), C coi như **trống** (*null*) hoặc có trị $\emptyset$ (zero)).

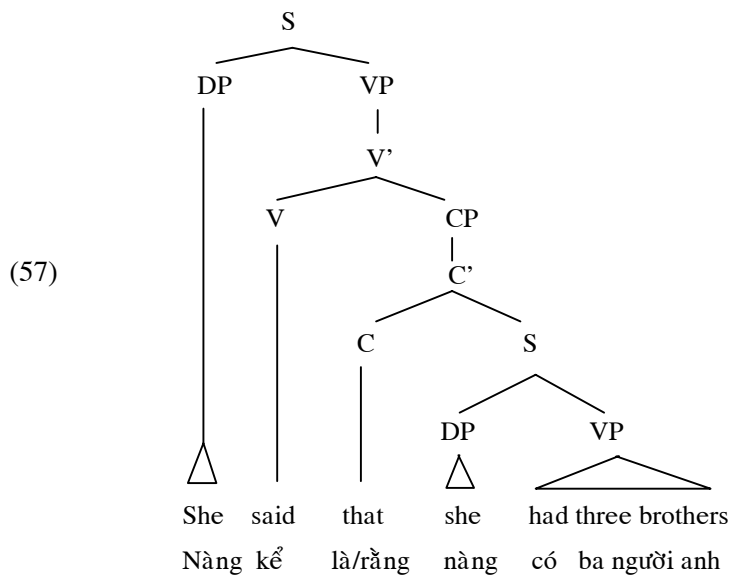
Cây cú pháp tương ứng:

a. C có “*that*” (*rằng/là*)

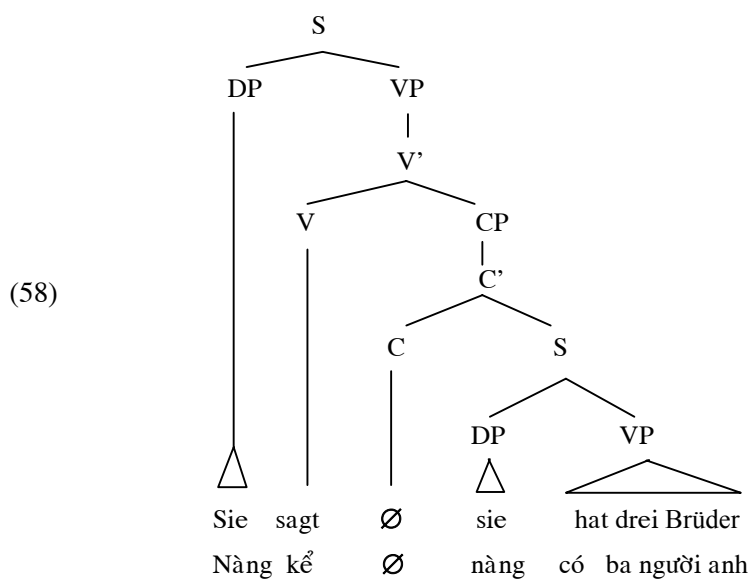
b. C có trị trống (không cần “*that*” (*rằng/là*)):



Một ví dụ tiếng Việt và tiếng Anh:



Một ví dụ tiếng Việt và tiếng Đức không có “*dass*” (*rằng/là*), tức C có trị $\emptyset$:



Ngoài việc dùng *tác tử bổ ngữ hóa C* để giải thích trường hợp “*that*”, các nhà ngôn ngữ học Tây phương tin rằng, C còn dùng được cho việc giải thích sự chuyển vị động từ trong câu xác định lên đầu câu hoặc đặt động từ ở đầu câu để

được **câu hỏi có/không** (*yes/no question*). Hiện tượng này thường thấy ở những ngôn ngữ **ngịch đảo chủ ngữ/động từ phụ** (*subject/aux inversion*)²⁾, điển hình là tiếng Anh, tiếng Đức. Trong tiếng Anh, muốn thành lập câu hỏi có/không, người ta phải đặt **động từ phụ** (*auxiliary verb*) ở đầu câu. Tiếng Đức tương tự, song không chỉ động từ phụ mà cả động từ thứ nhất trong câu. Vài ví dụ đối chiếu có nghĩa gần giống:

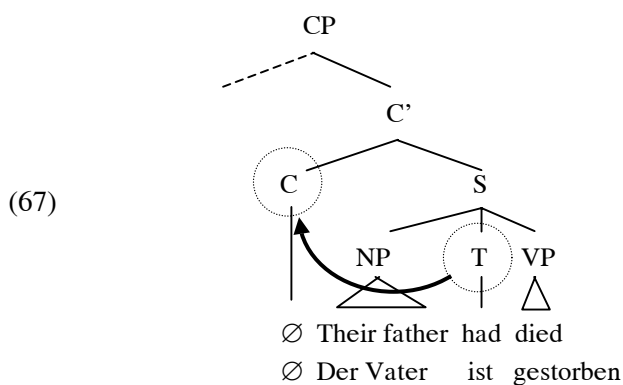
Câu xác định:

- | | | |
|------|-----------------------------------|-------------|
| (59) | The sons looked for the treasure. | (tiếng Anh) |
| (60) | Their father had died. | (tiếng Anh) |
| (61) | Die Kinder wollten die Erbe. | (tiếng Đức) |
| (62) | Der Vater ist gestorben. | (tiếng Đức) |

Câu hỏi có/không:

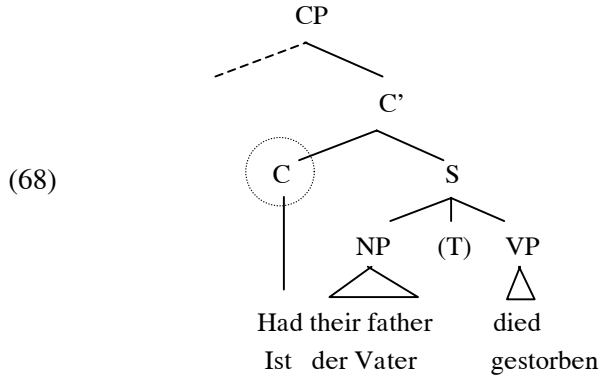
- | | | |
|------|--------------------------------------|-------------|
| (63) | Did the sons look for the treasure ? | (tiếng Anh) |
| (64) | Had their father died ? | (tiếng Anh) |
| (65) | Wollten die Kinder die Erbe ? | (tiếng Đức) |
| (66) | Ist der Vater gestorben ? | (tiếng Đức) |

Cây cú pháp của câu xác định:



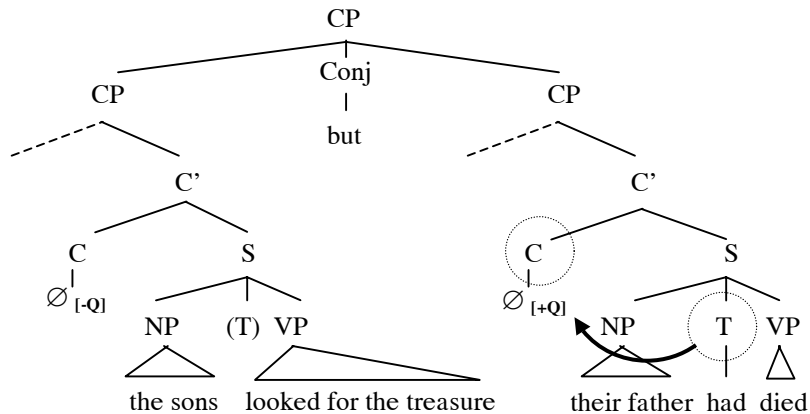
Cây cú pháp của câu hỏi có/không (ja/nein):

²⁾ Thuật ngữ **subject/aux inversion**, theo chúng tôi, có vẻ hơi hạn hẹp. Dùng **subject/verb inversion** có lẽ tổng quát hơn.



Bằng cách chuyển vị động từ phụ “*had*”, “*ist*” từ vị trí T lên vị trí C, câu xác định sẽ biến thành câu hỏi. Cấu trúc câu hỏi được giải thích tương đối dễ dàng. Trong câu trúc câu hỏi, vị trí T trống. Điều này chấp nhận được bởi T là một lựa chọn, không bắt buộc phải có. Cách giải thích trên cũng thỏa cho cả những trường hợp phức tạp, ví dụ có liên từ:

(69) The sons looked for the treasure but had their father died ?



Cùng một cấu trúc $C' \rightarrow C S$, tác tử bổ ngữ hóa C vừa mang tính chất câu xác định $[-Q]$ ³⁾, vừa mang tính chất câu hỏi $[+Q]$. Ở phần cây bên trái, C có trị trống $\emptyset_{[-Q]}$, trong khi đó, ở phần cây bên phải, C có trị trống $\emptyset_{[+Q]}$. Nếu chuyển vị “*had*” vào vị trí C, mệnh đề “*their father had died*” sẽ trở thành câu hỏi: “*had their father died ?*”.

³⁾ Q viết tắt của Question. +Q có nghĩa có tính chất câu hỏi, ngược lại là -Q.

Cấu trúc câu hỏi có/không của nhiều tiếng Âu châu là vậy, song tiếng Việt khác. Câu hỏi có/không trong tiếng Việt có nhiều loại. Xem thử vài loại thường gặp:

1. Câu xác định + “*không*”

Loại này bao gồm một câu xác định kèm theo trạng từ “*không*” ở phía sau. Phối hợp với từ chỉ thời gian thường không thích hợp. Vị từ (động từ, tính từ) thường được dùng để trả lời vắn tắt, hoặc để phủ nhận (đi kèm với “*không*”).

Câu hỏi:

- (70) Anh thích học tiếng Việt *không* ?
- (71) Câu hỏi tiếng Việt đa dạng *không* ?
- (72) Anh hiểu kỹ vấn đề này *không* ?

Câu xác định:

- (73) Tôi thích học tiếng Việt. (Trả lời vắn tắt: (Dạ/vâng/ừ, ...) thích)
- (74) Câu hỏi tiếng Việt đa dạng. ((Dạ/vâng/ừ, ...) đa dạng)
- (75) Tôi hiểu kỹ vấn đề này. ((Dạ/vâng/ừ, ...) hiểu)

Câu phủ định:

- (76) (Không). Tôi *không* thích học tiếng Việt. (Phủ định vắn tắt: không)
- (77) (Không). Câu hỏi tiếng Việt *không* đa dạng. (Phủ định vắn tắt: không)
- (78) (Không). Tôi *không* hiểu kỹ vấn đề này. (Phủ định: không/không hiểu)

(Không) : Nói thêm cũng được. Có thể phủ định vắn tắt bằng “*không*” + vị từ.

Thông thường, câu xác định không đi cùng với từ chỉ thời gian “*đang*”, “*đã*”, “*sẽ*”. Không thể nói:

- (79) * Anh *đang* thích học tiếng Việt *không* ?
- (80) * Câu hỏi tiếng Việt *đã* đa dạng *không* ?
- (81) * Anh *sẽ* hiểu kỹ vấn đề này *không* ?

Từ chỉ thời gian dùng trong phần bổ ngữ trực tiếp của câu hỏi:

Câu hỏi:

- (82) Anh nghĩ *tình hướng sẽ thay đổi* không ?
 (83) Anh tin *tình hướng sẽ thay đổi* không ?

Câu xác định:

- (84) Tôi nghĩ *tình hướng sẽ thay đổi*.
 (85) Tôi tin *tình hướng sẽ thay đổi*.

Ở câu phủ định, chúng ta sẽ gặp hai trường hợp đều nói được. Một mặt, “*không*” đứng trước động từ của chủ ngữ:

- (86) (Không). Tôi *không* nghĩ *tình hướng sẽ thay đổi*.
 (87) (Không). Tôi *không* tin *tình hướng sẽ thay đổi*.

Mặt khác, “*không*” đứng trước động từ chính của bổ ngữ; nói theo lý thuyết X-gạch, “*không*” giữ chức năng bổ nghĩa cho động từ.

- (88) (Không). Tôi nghĩ *tình hướng sẽ không thay đổi*.
 (89) (Không). Tôi tin *tình hướng sẽ không thay đổi*.

Mệnh đề phụ “*Tình hướng sẽ không thay đổi*” có khả năng đứng một mình như một câu, suy ra chúng ta có thể đặt câu hỏi:

- (90) *Tình hướng sẽ thay đổi không ?*

Hiện tượng dùng được từ chỉ thời gian “*sẽ*” này đâm ra mâu thuẫn với các trường hợp không nói được:

- (91) * Anh *đang* thích học tiếng Việt *không* ?
 (92) * Câu hỏi tiếng Việt *đã* đa dạng *không* ?
 (93) * Anh *sẽ* hiểu kỹ vấn đề này *không* ?

Phải chăng đây là trường hợp ngoại lệ ?

Xét tiếp, chúng ta còn thấy, từ chỉ thời gian còn dùng được cả trong trường hợp có ngữ đoạn giới từ PP nằm ở đầu câu:

Câu hỏi:

- (94) *Theo anh, (thì) tình huống sẽ thay đổi không ?*
- (95) *Theo anh, (thì) tình huống sẽ sáng sủa hơn không ?*
- (96) *Theo anh, (thì) họ sẽ hòa giải không ?*

Xin lưu ý, dấu phẩy ở đây thay thế cho “*thì*”.

Câu xác định:

- (97) *Theo tôi, (thì) tình huống sẽ thay đổi.* (Trả lời vấn tắt: (Sẽ) thay đổi)
- (98) *Theo tôi, (thì) tình huống sẽ sáng sủa hơn.* ((Sẽ) sáng sủa (hơn))
- (99) *Theo tôi, (thì) họ sẽ hòa giải* ((Sẽ) hòa giải).

Câu phủ định:

- (100) *Theo tôi, (thì) tình huống sẽ không thay đổi.*
- (101) *Theo tôi, (thì) tình huống sẽ không sáng sủa hơn.*
- (102) *Theo tôi, (thì) họ sẽ không hòa giải.*

Bỏ ngữ đoạn giới từ, chúng ta lại gặp trường hợp “*Họ sẽ hòa giải*” qua câu hỏi:

- (103) *Họ sẽ hòa giải không ?*

Một trường hợp tương tự cũng nói được:

- (104) *Lịch sử sẽ lập lại không ?*

Một giả thuyết: Phải chăng những từ có tính cách diễn biến đều dùng được với từ chỉ thời gian ?

2. Câu xác định với “*có ... không*”

Loại câu hỏi này cùng tính chất như loại trên, chỉ khác là có thêm từ “*có*” sau chủ ngữ.

Câu hỏi:

- (105) *Anh có thích học tiếng Việt không ?*

(106) Câu hỏi tiếng Việt *có* đa dạng *không* ?

(107) Anh *có* hiểu kỹ vấn đề này *không* ?

Câu xác định:

(108) Tôi thích học tiếng Việt. (Trả lời vắn tắt: (Dạ/vâng/ừ, ...) thích)

(109) Câu hỏi tiếng Việt đa dạng. ((Dạ/vâng/ừ, ...) đa dạng)

(110) Tôi hiểu kỹ vấn đề này. ((Dạ/vâng/ừ, ...) hiểu)

Câu phủ định:

(111) (Không). Tôi *không* thích học tiếng Việt. (Phủ định vắn tắt: không)

(112) (Không). Câu hỏi tiếng Việt *không* đa dạng. (Phủ định vắn tắt: không)

(113) (Không). Tôi *không* hiểu kỹ vấn đề này. (Phủ định vắn tắt: không)

(Không) : Nói thêm cũng được. Có thể phủ định vắn tắt bằng “*không*” + vị từ.

Thông thường, câu xác định không đi cùng với từ chỉ thời gian “*đang*”, “*đã*”, “*sẽ*”. Không thể nói:

(114) * Anh *đang có* thích học tiếng Việt *không* ?

(115) * Câu hỏi tiếng Việt *đã có* đa dạng *không* ?

(116) * Anh *sẽ có* hiểu kỹ vấn đề này *không* ?

Từ chỉ thời gian dùng được trong phần bổ ngữ trực tiếp:

Câu hỏi:

(117) Anh có nghĩ *tình hướng sẽ thay đổi* không ?

(118) Anh có tin *tình hướng sẽ thay đổi* không ?

Câu xác định:

(119) Tôi nghĩ (tình hướng) *sẽ thay đổi*.

(120) Tôi tin (tình hướng) *sẽ thay đổi*.

Những trường hợp còn lại đều giống như loại kể trên. Nay, hãy thử đặt “*có*” vào trường hợp có ngữ đoạn giới từ PP ở đầu câu:

Câu hỏi:

- (121) *Theo anh, (thì) tình huống có thay đổi không ?*
 (122) *Theo anh, (thì) tình huống có sáng sủa hơn không ?*
 (123) *Theo anh, (thì) họ có hòa giải không ?*

Xin lưu ý, dấu phẩy ở đây thay thế cho “thì”.

Câu xác định:

- (124) *Theo tôi, (thì) tình huống có thay đổi.* (Trả lời vấn tắt: (Có) thay đổi)
 (125) *Theo tôi, (thì) tình huống có sáng sủa hơn.* ((Có) sáng sủa (hơn))
 (126) *Theo tôi, (thì) họ có hòa giải* ((Có) hòa giải).

Câu phủ định:

- (127) *Theo tôi, (thì) tình huống không thay đổi.*
 (128) *Theo tôi, (thì) tình huống không có thay đổi.*
 (129) *Theo tôi, (thì) tình huống không có gì thay đổi.*
 (130) *Theo tôi, (thì) tình huống không sáng sủa hơn.*
 (131) *? Theo tôi, (thì) tình huống không có sáng sủa hơn.* (Câu này nghe khá lấn cấn).
 (132) *Theo tôi, (thì) tình huống không có gì sáng sủa hơn.* (Nghe được nhưng câu dưới đây vẫn tự nhiên hơn)
 (133) *Theo tôi, (thì) tình huống không có gì sáng sủa (cả).*
 (134) *Theo tôi, (thì) họ không hòa giải.*
 (135) *Theo tôi, (thì) họ không có hòa giải.*

Thử nghiệm dùng chữ cho thấy, mặc dù đúng ngữ pháp nhưng không phải câu nói bao giờ cũng dễ nghe. Ngôn ngữ nào cũng vậy.

3. Câu xác định với “*phải không*”

Loại câu hỏi này gồm “*câu xác định*” + “*phải không*” ở cuối câu. Tất cả các câu có diễn từ “là” đều có câu hỏi dưới dạng này. Cách sử dụng từ chỉ thời gian ở đây không bị giới hạn như ở các trường hợp trên.

Câu hỏi:

- (136) Anh thích học tiếng Việt *phải không* ?
 (137) Câu hỏi tiếng Việt rất đa dạng *phải không* ?
 (138) Chị là người Việt *phải không* ?

Câu xác định:

- (139) Tôi thích học tiếng Việt. (Trả lời vắn tắt: (Dạ/vâng/ừ/đúng, dạ đúng...))
 (140) Câu hỏi tiếng Việt rất đa dạng. (Dạ/vâng/ừ/đúng, dạ đúng, dạ phải, ...)
 (141) Tôi là người Việt. (Dạ/vâng/ừ/đúng, dạ đúng, dạ phải, ...)
 (142) (Dạ/vâng/ừ/đúng, dạ đúng...). Tôi thích học tiếng Việt.
 (143) (Dạ/vâng/ừ/đúng, dạ đúng, dạ phải, ...). Câu hỏi tiếng Việt rất đa dạng.
 (144) (Dạ/vâng/ừ/đúng, dạ đúng, dạ phải, ...). Tôi là người Việt.

Câu phủ định:

- (145) (Không). Tôi *không* thích học tiếng Việt. (Phủ định vắn tắt: không)
 (146) (Không). Câu hỏi tiếng Việt *không* đa dạng. (Phủ định vắn tắt: không (phải))
 (147) (Không phải). Tôi *không phải* là người Việt. (Phủ định vắn tắt: không (phải))

(Không), (Không phải) : Nói thêm cũng được. “*Không*” đứng trước động từ bình thường hoặc tính từ. “*Không phải*” đứng trước “là”.

Loại câu hỏi này có thể đi cùng với từ chỉ thời gian “*đang*”, “*đã*”, “*sẽ*”, “*đã từng*”, ...:

- (148) Anh *đã* học tiếng Việt *phải không* ?
 (149) Anh *đã từng* học tiếng Việt *phải không* ?
 (150) Anh *sẽ* học tiếng Việt *phải không* ?
 (151) Anh *đang* học tiếng Việt *phải không* ?
 (152) Anh *đang có ý định* học tiếng Việt *phải không* ?

“*Có*” giữ chức năng như một động từ và dùng được với “*phải không*”, khác với trường hợp dưới đây không nói được:

- (153) * Anh *có* thích học tiếng Việt *phải không* ?

Thế nhưng, trường hợp sau lại nói được:

(154) Anh *có* học nhạc *phải không* ?

Động từ “*thích*”, “*muốn*” phải có một tính chất gì đấy, cho nên mới không dùng “*có*” được. “*Sẽ*” tương tự vậy:

(155) * Anh *sẽ* muốn học tiếng Việt *phải không* ?

Câu hỏi có phần bổ ngữ trực tiếp:

(156) Anh nghĩ tình huống sẽ thay đổi *phải không* ?

(157) Anh tin tình huống sẽ thay đổi *phải không* ?

Câu trả lời bằng cách xác nhận:

(158) Tôi nghĩ tình huống sẽ thay đổi.

(159) Tôi tin tình huống sẽ thay đổi.

Câu hỏi có ngữ đoạn giới từ PP ở đầu câu:

(160) *Theo anh, (thì)* tình huống sẽ thay đổi *phải không* ?

(161) *Theo anh, (thì)* tình huống sẽ sáng sủa hơn *phải không* ?

(162) *Theo anh, (thì)* họ sẽ hòa giải *phải không* ?

Ngoài dạng “*phải không*”, người Việt còn dùng “*đúng không*”, “*chứ gì*”, ...

4. Câu xác định với “***Có phải ... không***”

Loại câu hỏi này gồm “*Có phải*” + *câu xác định* + “*không*”. Dạng này thường đi với “*là*”. Cách sử dụng từ chỉ thời gian ở đây không bị giới hạn nhiều.

Câu hỏi:

(163) *Có phải* anh thích học tiếng Việt *không* ?

(164) *? *Có phải* câu hỏi tiếng Việt rất đa dạng *không* ? (Không dùng với “*rất*” được)

(165) *Có phải* chị là người Việt *không* ?

Câu xác định:

- (166) Tôi thích học tiếng Việt. (Trả lời vắn tắt: (Dạ/vâng/ừ/đúng, phải, dạ đúng, dạ phải, ...))
 (167) Tôi là người Việt. (Dạ/vâng/ừ/đúng, dạ đúng, dạ phải, ...)

Câu phủ định:

- (168) (Không). Tôi *không* thích học tiếng Việt. (Phủ định vắn tắt: không)
 (169) (Không phải). Tôi *không phải* là người Việt. (Phủ định vắn tắt: không (phải))

(Không), (Không phải) : Nói thêm cũng được. “*Không*” đứng trước động từ bình thường hoặc tính từ. “*Không phải*” đứng trước “là”.

Loại câu hỏi này có thể đi cùng với từ chỉ thời gian “*đang*”, “*đã*”, “*sẽ*”, “*đã từng*”, ...:

- (170) *Có phải* anh *đã* học tiếng Việt *không* ?
 (171) *Có phải* anh *đã từng* học tiếng Việt *không* ?
 (172) *Có phải* anh *sẽ* học tiếng Việt *không* ?
 (173) *Có phải* anh *đang có ý định* học tiếng Việt *phải không* ?

5. Các loại câu hỏi có/không khác:

Ngoài ra, trong tiếng Việt còn có nhiều cách lập câu hỏi có/không khác:

- (174) Cô Khánh Ly hát bài đó chứ gì ?
 (175) “Năng có hồng bằng đôi môi em ?” (không dùng “*không*” cuối câu)
 (Trịnh Công Sơn)
 (176) Phải chăng nền văn học Việt Nam không tiến nổi ?
 (177) Đi chơi nhé ?
 (178) Có ăn hay không ?
 (179) Đi ngủ chưa ?
 (180) Đó là việc của giới thi văn ư/hở/hả/à/sao/chẳng/... ?

v.v.

Sau khi đã lướt sơ qua các trường hợp câu hỏi có/không, chúng ta nhận thấy rằng, tiếng Việt không có trường hợp chuyển vị động từ lên đầu câu như tiếng Anh, tiếng Đức. Như vậy quy tắc X-gạch $C' \rightarrow C S$ không thích hợp cho tiếng Việt trong việc giải thích loại câu hỏi này.

3. Ngữ đoạn thì

Cấu trúc mệnh đề phụ $S' \rightarrow (C) S$ có thể chuyển sang sơ đồ X-gạch, ở đó, CP thay thế S' :

(181) $CP \rightarrow C'$

(182) $C' \rightarrow (C) S$

Vấn đề còn lại là cấu trúc câu có *thì* (*tense*):

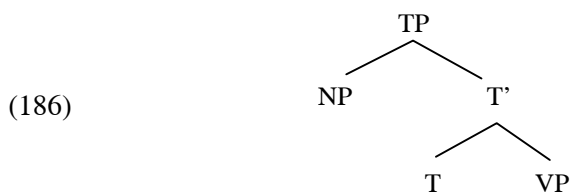
(183) $S \rightarrow NP (T) VP$

Muốn thỏa lý thuyết X-gạch, điều cần để ý là xem phần tử nào có thể đóng vai đầu ngữ đoạn. Các nhà ngôn ngữ học Tây phương đã đề nghị một mô hình: thay thế S bằng TP:

(184) $TP \rightarrow NP T'$

(185) $T' \rightarrow T VP$

TP là ngữ đoạn thì (*tense phrase*).



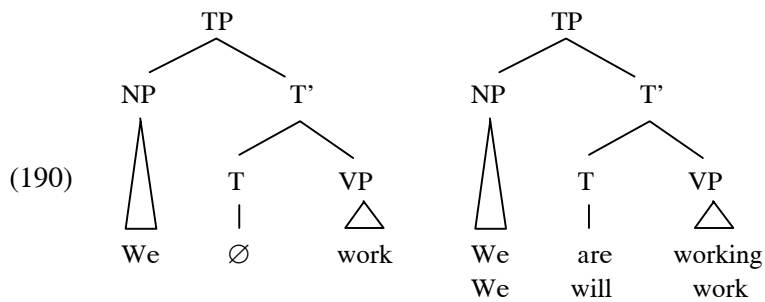
Cây cho thấy, NP nằm trong TP, và VP làm bổ ngữ. Với cấu trúc này, người ta có thể giải thích cách sử dụng thì trong ngôn ngữ Âu châu, nhất là tiếng Anh. Động từ chính nằm ở vị trí VP, động từ phụ nằm ở vị trí T. Trong trường hợp không dùng động từ phụ, T sẽ có trị zero. Cụ thể để diễn tả thì hiện tại, động từ chính có thể ở thể nguyên mẫu hoặc ở thể *biến thể* (*inflection*). Ví dụ “*going*” là

một biến thể của động từ “go” nhằm diễn tả hành động “đang đi”, hoặc “*slept*”, “*worked*” là các biến thể của “sleep”, “work” được dùng cho quá khứ, v.v..

(187) We work.

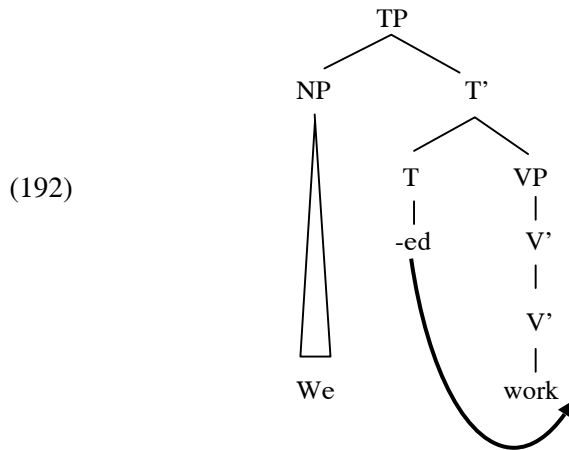
(188) We are working.

(189) We will work.



Thế nhưng có trường hợp, cần chuyển vị dấu hiệu chỉ thì “ed”:

(191) We worked.



Còn có trường hợp phức tạp hơn nữa như:

(193) We have worked.

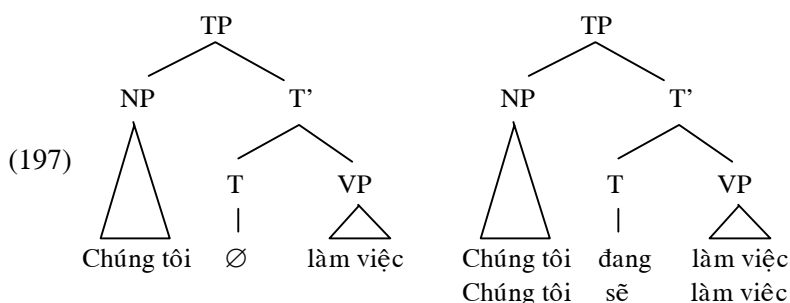
Trong trường hợp này, quy tắc X-gạch bắt buộc phải được mở rộng (ở đây xin miễn bàn chi tiết, có thể xem thêm Pollock 1989).

So với các ngôn ngữ Âu châu, cách dùng thì trong tiếng Việt không theo hình thức chia động từ. Ngay vị trí T, chúng ta chỉ có thể đưa những trạng từ chỉ thời gian như “đang”, “đã”, “sẽ”, “sắp”, ... vào:

(194) Chúng tôi làm việc.

(195) Chúng tôi đang làm việc.

(196) Chúng tôi sẽ làm việc.



Tuy nhiên, sự giải thích ấy cần được xét lại, bởi một là, theo nguyên tắc, trạng từ phải nằm trong VP với tư cách phụ nghĩa, hai là cách mô tả thời gian trong tiếng Việt còn phụ thuộc vào các từ chỉ thời gian khác, như “hôm qua”, “bây giờ”, “ngày mai”, “ngày xưa, ngày xưa” ... Xin miễn bàn sâu về chủ đề này ở đây mà coi đó nhưng một điểm mở.

TÓM LƯỢC

Có nhiều ngôn ngữ không có cấu trúc Subject-Verb-Object (SVO). Quy tắc X-gạch đã được tổng quát hóa như sau:

(198) Quy tắc biệt định ngữ: $X' \rightarrow X' (YP) \mid (YP) X'$

(199) Quy tắc phụ ngữ: $X' \rightarrow X' (ZP) \mid (ZP) X'$

(200) Quy tắc bổ ngữ: $X' \rightarrow X (WP) \mid (WP) X$

Ngoài các quy tắc trên, người ta có thể dùng sơ đồ X-gạch để giải thích thì và mệnh đề/câu.

Theo ngôn ngữ Âu châu, *mệnh đề* được hiểu là một chuỗi từ gồm *chủ ngữ* cộng *vị ngữ*.

Vấn đề *mệnh đề phụ* và câu *có thì* được giải thích xuyên qua *ngữ đoạn tác tử bổ ngữ hóa* và *ngữ đoạn thì*.

Ngoài việc dùng *tác tử bổ ngữ hóa C* để giải thích trường hợp “*that*”, các nhà ngôn ngữ học Tây phương còn dùng ngữ đoạn này để giải thích sự chuyển vị động từ trong câu xác định thành *câu hỏi có/không* (*yes/no question*) ở các ngôn ngữ nghịch đảo. Tiếng Việt không phải là ngôn ngữ nghịch đảo (Câu hỏi có/không trong tiếng Việt không bắt buộc chuyển vị động từ).

NHẬN XÉT VÀ ĐIỂM MỞ

Lý thuyết X-gạch có bài bản và được lập thức rõ ràng. Sơ đồ X-gạch cho phép thành lập các ngữ đoạn trừu tượng. Các quy tắc sản sinh cụ thể không những chỉ giúp chúng ta nhận diện ngữ pháp của ngôn ngữ mình mà còn có thể dùng trong việc lập trình, nói chung là ứng dụng được tin học vào ngôn ngữ.

Khởi thủy, lý thuyết X-gạch được dùng để giải thích tiếng Anh. Những quy tắc sản sinh từ đó không thỏa đáp mọi ngôn ngữ. Muốn sử dụng lý thuyết này, mỗi ngôn ngữ phải tự tìm cho mình những lời giải thích hợp. Một ví dụ điển hình: Quy tắc $C' \rightarrow C S$ dùng giải thích cấu trúc câu hỏi có/không của ngôn ngữ Âu châu chắc chắn không thích hợp cho tiếng Việt. Động từ trong câu tiếng Việt không bị chuyển vị để trở thành câu hỏi.

Nói chung, xuyên qua các chương bàn về lý thuyết X-gạch, hẳn ai trong chúng ta cũng nhận thấy rằng, tiếng Việt có nhiều cái đặc biệt mà tiếng Anh không có. Đó là những điểm mở mà ta có thể đào sâu, nghiên cứu thêm. Trong khuôn khổ chương này, thử xét vài điểm sau đây:

- 1) ***Quan hệ chủ ngữ-động từ (Subject-Verb)***: Lập quy tắc X-gạch giải thích sự vắng mặt của chủ thể trong câu. Làm thế nào để xác định được động từ nào thuộc chủ ngữ nào (chủ ngữ thật, chủ thể (chủ ngữ) ngữ pháp) ?
- 2) ***Thì, là, mà***: Thử giải thích các trường hợp chứa “*thì*”, “*là*”, “*mà*” bằng lý thuyết X-gạch.
- 3) ***Mệnh đề phụ***: Phân tích thêm mệnh đề phụ bằng lý thuyết X-gạch.

- 4) **Câu hỏi có/không:** Thử giải thích câu hỏi có/không của tiếng Việt bằng quy tắc X-gạch.
- 5) Thử giải thích cấu trúc câu phủ định bằng quy tắc X-gạch. Thử xem trường hợp dùng tiếng đệm (ví dụ “*đâu*”) để nhấn mạnh câu.
 - Em *đâu* có đi ***đâu***.
 - Nó *đâu* đã ăn gì ***đâu***.
- 6) Thử dùng lý thuyết X-gạch giải thích sâu hơn về thì trong tiếng Việt.
- 7) Thử xét trường hợp mệnh lệnh cách pha lẫn câu hỏi khi sử dụng tiếng đệm:
 - Ăn cơm đi ***chứ hả***
 - Ngủ đi ***chứ***
- 8) Thử xét trường hợp câu hỏi khi sử dụng tiếng đệm:
 - Ăn cơm ***hả***?
 - Ăn cơm ***nhé***?
 - Đi chơi ***nhé***?
- 9) Thử xét trường hợp tính chất câu hỏi có tiếng đệm sẽ biến mất khi sử dụng thêm trạng từ:
 - Ăn cơm ***ngon*** *nhé*.
 - Đi chơi ***vui*** *nhé*.

Chương 11

TIỂU PHẠM TRÙ, TỪ VỰNG –
X-GẠCH HẠNG CHẾ

Như đã biết, ngôn ngữ Âu châu thiên về cấu trúc *chủ-vị* (*chủ ngữ-vị ngữ, subject-predicate*). Cấu trúc này xuất hiện trong câu/mệnh đề. Tiêu biểu là quy tắc cú pháp:

(1) $S \rightarrow NP VP$

Câu/mệnh đề S là một kết quả của một chủ ngữ dưới dạng NP và một vị ngữ dưới dạng VP. Vị ngữ lệ thuộc chủ ngữ. Nói một cách đơn giản, vị từ (động từ) thuộc về chủ ngữ trong câu. Chủ ngữ câu được hiểu là ***chủ ngữ ngữ pháp (grammatical subject)***.

(2) She sings. (tiếng Anh)

(3) Sie singt. (tiếng Đức)

(4) Cô ấy hát.

Động từ "hát" (*sings, singt*) xuất phát từ chủ ngữ "cô ấy" (*she, sie*): "cô ấy hát", chứ không ai khác.

Tiếng Việt cũng có cấu trúc chủ-vị, song không phải chỉ có cấu trúc ấy. Có trường hợp, động từ không thuộc về chủ ngữ trong câu mà lại liên quan đến một ngoại chủ thể không có mặt trong câu mà ta đã thấy ở chương rồi.

Cấu trúc câu/mệnh đề của từng ngôn ngữ có thể khác nhau. Song nếu xét về vị ngữ, thì có lẽ mọi ngôn ngữ đều có những điểm chung. Vị từ có khi không cần bổ ngữ hoặc có khi cần.

(5) She sings a song. (tiếng Anh)

(6) Sie singt ein Lied. (tiếng Đức)

(7) Cô ấy hát một bài ca. (tiếng Việt)

“*Một bài ca*” (*a song, ein Lied*) là bổ ngữ trực tiếp có nhiệm vụ bổ nghĩa cho vị từ. Hoặc một hình thức khác:

- | | | |
|------|--------------------|--------------|
| (8) | She sings with me. | (tiếng Anh) |
| (9) | Sie singt mit mir. | (tiếng Đức) |
| (10) | Cô ta hát với tôi. | (tiếng Việt) |

“*Với tôi*” (*with me, mit mir*) là một ngữ đoạn giới từ PP.

Dựa vào quy tắc cú pháp truyền thống:

- (11) $VP \rightarrow V (NP) (PP)$

hoặc sơ đồ X-gạch:

- (12) $VP \rightarrow V'$
 (13) $V' \rightarrow V' (PP)$
 (14) $V' \rightarrow V (NP)$

người ta có thể giải thích được các kiểu nói trên. Thế nhưng nếu đặt câu hỏi, trường hợp nào dùng bổ ngữ cho hợp lý, trường hợp nào không, thì các quy tắc ấy không giải thích được. Vài ví dụ:

- | | | |
|------|--------------------------------|--------------|
| (15) | * She sings a man. | (tiếng Anh) |
| (16) | * Sie singt einen Mann. | (tiếng Đức) |
| (17) | * Cô ấy hát một người đàn ông. | (tiếng Việt) |

Mặc dù những câu trên đúng ngữ pháp, nghĩa là thỏa quy tắc chủ-vị $S \rightarrow NP VP$, nhưng lại không nói được. Rõ ràng là không thể dùng động từ “*hát*” (*sings, singt*) cho “*một người đàn ông*” (*a man, einen Mann*).

Nghĩ ra một bổ ngữ cho thích hợp với vị từ là khả năng thẩm định ý nghĩa từ ngữ của con người. Mỗi con người đều biết phân biệt vị từ nào cần bổ ngữ, vị từ nào không, và khi cần, thì cần loại nào. Những thông tin ngữ nghĩa ấy được chứa trong kho **từ vựng (lexical)** khổng lồ của bộ óc chúng ta.

Trong chương này, chúng ta sẽ bàn về từ vựng và việc hạn chế sử dụng sơ đồ X-gạch, tạm gọi là *X-gạch hạn chế (constrain X-bar)*. Hạn chế có thể hiểu nôm na là sự đòi hỏi vị từ hoặc những phần tử có liên quan đến nó phải được sử dụng

sao cho hợp lý. Nói ngắn gọn, tính hạn chế có nhiệm vụ bắt buộc cách dùng quy tắc X-gạch sao cho câu đúng ngữ pháp và ngữ nghĩa. Những trường hợp phi lý sẽ bị loại trừ.

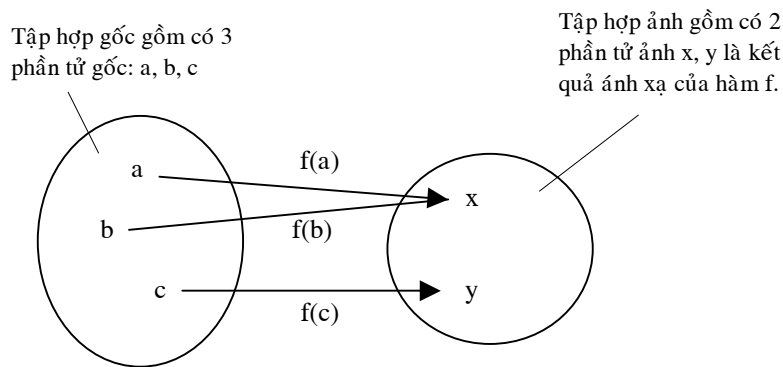
TỪ VỰNG

1. Vị từ và tham số

Vị ngữ (*predicate phrase*), hiểu theo định nghĩa của ngôn ngữ học Âu châu, là phần *động từ* cộng với *bổ ngữ* của một câu hoặc mệnh đề. Động từ giữ chức năng như một **vị từ (*predicate*)**¹⁾.

Xét về mặt cấu trúc thì vị từ được xem như là *đầu* của vị ngữ. Ngoài ra, còn một khía cạnh cũng cần được xét đến, đó là tính *quan hệ* (*relation*) của vị từ. Tính quan hệ ở đây được hiểu theo nghĩa toán.

Hẳn chúng ta ai cũng hiểu ý nghĩa của **hàm số (*function*)**. Giả sử có một hợp tập gốc, nếu ứng dụng một hàm số lên một phần tử thuộc tập hợp này, ta sẽ được một ảnh tương ứng và chỉ một ảnh duy nhất. Đây là một loại quan hệ 1-1.



Ví dụ, hàm số f ứng dụng lên phần tử gốc a , sinh ra ảnh $f(a) = x$.

¹⁾ “*Predicate*” có lúc có nghĩa là vị ngữ, có lúc có nghĩa là vị từ. Đây là một điểm khá lộn xộn của thuật ngữ “*predicate*” mà giới ngôn ngữ học Tây phương thường vướng phải. Để tránh hiểu nhầm, chúng tôi gọi vị từ là “*predicate*”, vị ngữ là “*predicate phrase*”.

Khác với tính “**1 đối 1**” của hàm số, một **quan hệ (relation)** tổng quát cho phép một phần tử có quan hệ với *nhều* phần tử. (Chỉ cần hiểu đơn giản vậy cũng đủ hiểu phần lý thuyết sau này).

Động từ (vị từ) diễn tả mối quan hệ của những phần tử có mặt trong câu. Những phần tử tham dự được coi là những **tham số (argument)**. Ví dụ khi nói:

(18) She sings a song.

thì “*she*” và “*a song*” là hai tham số. Hai tham số này nối nhau thành một câu. Vị từ “*sing*” là cầu nối.

Như vậy, động từ cần một số lượng tham số bắt buộc để bảo đảm câu đúng ngữ pháp. Số lượng tham số được gọi là **hóa trị (valence)**²⁾. Có ít nhất 3 trường hợp:

a. Một tham số: Đây là trường hợp **động từ bất xuyên tính (intransitive verb)**, tức loại động từ chỉ có tác dụng riêng cho chủ ngữ chứ không hướng đến bổ ngữ bên ngoài. Chủ ngữ là một tham số.

[She]₁ sings.

(“*She*” là 1 tham số)

b. Hai tham số: Đây là trường hợp **động từ xuyên tính (transitive verb)**, tức loại động từ có tác dụng xuyên ra khỏi chủ ngữ và hướng đến bổ ngữ bên ngoài. Chủ ngữ, bổ ngữ là hai tham số.

[She]₁ sings [a song]₂.

(“*She*”, “*a song*” là 2 tham số)

c. Ba tham số: Đây là trường hợp **động từ nhị xuyên (ditransitive verb)**, tức loại động từ có tác dụng hướng đến hai cá thể thuộc phần bổ ngữ. Chủ ngữ cộng với 2 cá thể trong phần bổ ngữ là 3 tham số.

[She]₁ gives [me]₂ [a book]₃.

²⁾ một thuật ngữ vay mượn từ ngành hóa học.

(“*She*”, “*me*”, “*a book*” là 3 tham số)

Có trường hợp, một động từ mang hai tính chất: xuyên tính và bất xuyên tính. Ví dụ “*sing*” (*singen, hát*) :

Bất xuyên tính:

- | | | |
|------|------------|--------------|
| (19) | She sings. | (tiếng Anh) |
| (20) | Sie singt. | (tiếng Đức) |
| (21) | Cô ấy hát. | (tiếng Việt) |

Xuyên tính:

- | | | |
|------|------------------------|--------------|
| (22) | She sings a song. | (tiếng Anh) |
| (23) | Sie singt ein Lied. | (tiếng Đức) |
| (24) | Cô ấy hát một bài hát. | (tiếng Việt) |

Khi xem một động từ có bao nhiêu tham số, người ta chỉ xét đến các yếu tố *biệt định* (*specify*) và *bổ nghĩa* (*complement*) (bổ túc cho đầy đủ ý nghĩa). Yếu tố *phụ nghĩa* (*adjunct*) không dùng được ở đây.

Thế nhưng hai yếu tố *biệt định* và *bổ nghĩa* cũng bị giới hạn; có trường hợp dùng được, có trường hợp không. Điển hình là trường hợp một ngữ đoạn danh từ NP hoặc một ngữ đoạn tác tử bổ ngữ hóa CP giữ nhiệm vụ thêm nghĩa cho động từ:

- | | |
|------|----------------------------|
| (25) | I know [NP that]. |
| (26) | I know [CP that you know]. |

Động từ “*know*” dùng được cho hai trường hợp NP và CP, trong khi đó “*love*” chỉ dùng được cho NP:

- | | |
|------|------------------------------|
| (27) | I love [NP that]. |
| (28) | * I love [CP that you know]. |

Như vậy mỗi động từ còn đòi hỏi mỗi tham số phải có ngữ nghĩa hợp lý. Ta ghi nhớ điểm này trước khi vào chi tiết.

2. Tiểu phạm trù

Lý thuyết truyền thống đã liệt động từ vào phạm trù từ vựng tiêu chuẩn *động từ* V, song trên thực tế, động từ có nhiều loại khác nhau (bất xuyên tính, xuyên tính, ...). để bảo đảm cách dùng sao cho hợp với số lượng tham số, người ta phải chia nhỏ loại động từ thành những *tiểu phạm trù (subcategory)*.

Tạm ngưng dùng sơ đồ X-gạch, chúng ta xem lại quy tắc VP truyền thống của tiếng Anh:

$$(29) \quad VP \rightarrow V \text{ NP}$$

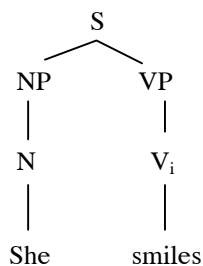
Quy tắc này có thể được mở rộng thành:

$$(30) \quad \begin{aligned} VP &\rightarrow V_i \\ VP &\rightarrow V_t \text{ NP} \\ VP &\rightarrow V_p \text{ PP} \\ VP &\rightarrow V_{tp} \text{ NP PP} \\ VP &\rightarrow V_{t2} \text{ NP NP} \end{aligned}$$

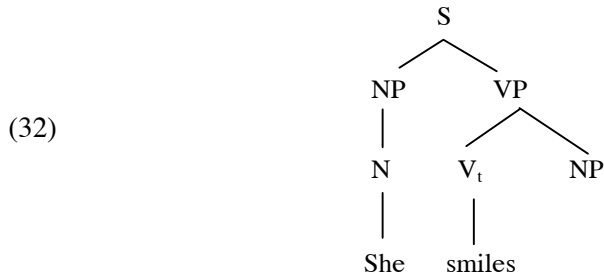
(i = intransitive, t = transitive, p = preposition, tp = transitive with preposition, t2 = ditransitive)

Với cách phân loại ấy, phạm trù động từ không còn đơn thuần là động từ nữa. Những loại động từ V_i , V_t , V_p , V_{tp} , V_{t2} được coi như tiểu phạm trù của V.

$$(31) \quad \text{She smiles.}$$



Câu “*she smiles*” không thể dùng quy tắc $VP \rightarrow V_t \text{ NP}$, vì “*smiles*” thuộc tiểu phạm trù V_i (động từ bất xuyên tính).



Ý tưởng chia nhỏ phạm trù thành những tiểu phạm trù đã làm mất tính tổng quát của quy tắc $VP \rightarrow V NP$. Người ta nói, tiểu phạm trù mang tính **cục bộ nghiêm ngặt (strikt local)**, mỗi tiểu phạm trù chỉ còn chấp nhận một quy tắc cục bộ đúng cho mỗi trường hợp nhất định đã được lựa chọn từ tập hợp phạm trù lớn. Ví dụ, gặp động từ bất xuyên tính, bắt buộc phải dùng quy tắc $VP \rightarrow V_i$; gặp động từ xuyên tính, bắt buộc phải dùng quy tắc $VP \rightarrow V_t NP, \dots$

Sự kiện mất tính tổng quát dẫn đến vấn đề là làm sao có thể mô tả được cái *cấu trúc toàn thể* của ngữ đoạn động từ VP.

Hãy xem một chút lý thuyết và cách giải quyết vấn đề:

a. Tôn ty Chomsky

Quy tắc ngữ pháp tạo sinh của Chomsky phản ánh mô hình *ngữ pháp cấu trúc ngữ đoạn (Phrase Structure Grammar (PSG))*. Mỗi quy tắc bao gồm hai vế: *vế trái* và *vế phải*. *Vế trái* sinh ra *vế phải* (ký hiệu mũi tên $\rightarrow$)³⁾.

(33) $X \rightarrow yZ$

(34) $y \rightarrow \text{value}$

Mỗi ký hiệu X, Z (viết hoa) là một **biến số (variable)** hoặc phần tử **không tận cùng (non-terminal)**, tức còn khai triển được. Ngược lại, nếu đã tận cùng bằng một trị số “value” nào đó, ví dụ y (viết nhỏ), nó sẽ được coi như một **nguyên tử (atom)** hoặc phần tử **tận cùng (terminal)**.

³⁾ Có người đọc là “vế trái được thay thế bằng vế phải”, “vế trái gồm ...”, ...

Theo **tôn ty Chomsky** (*Chomsky's hierarchy*) (S.s. Chomsky 1965), (Hopcroft, Ullman 1979), ngữ pháp được chia ra làm 4 loại:

- **Loại 0 (type 0)**: Ngữ pháp vũ trụ (universal grammar)
- **Loại 1 (type 1)**: Ngữ pháp cảm ngữ cảnh (context sensitive grammar)
- **Loại 2 (type 2)**: Ngữ pháp phi ngữ cảnh (context free grammar)
- **Loại 3 (type 3)**: Ngữ pháp chính quy (regular grammar)

Ngữ pháp chính quy (regular grammar) là loại ngữ pháp đơn giản nhất, có dạng:

$$(34) \quad X \rightarrow t$$

X là một phần tử tận cùng bằng trị “t”

Ví dụ:

Danh từ $N \rightarrow \text{house} \mid \text{car} \mid \text{Frau} \mid \text{Mann} \mid \dots$

Chỉ định từ $D \rightarrow \text{the} \mid a \mid \text{der} \mid \text{die} \mid \text{das}$

($\mid$ là ký hiệu của phép “hoặc” (OR) trong logic).

Ngữ pháp phi ngữ cảnh (context free grammar) có dạng:

$$(35) \quad X \rightarrow yZ$$

X, Z là những phần tử không tận cùng; y là phần tử tận cùng, có thể nằm bên phải hoặc bên trái Z.

Ví dụ:

$$VP \rightarrow V \, NP$$

Ngữ pháp phi ngữ cảnh là tập hợp chứa ngữ pháp chính quy.

Ngữ pháp cảm ngữ cảnh (context sensitive grammar) có dạng:

$$(36) \quad AXB \rightarrow AYB$$

A, B, X, Y là những phần tử không tận cùng. Vế bên trái không được chứa nhiều phần tử hơn vế bên phải. X có thể được thế bằng Y , nghĩa là tùy ngữ cảnh, X có thể sinh ra một hiển trị nhất định. X và Y có thể trống (NULL).

Ngữ pháp cảm ngữ cảnh là tập hợp chứa ngữ pháp phi ngữ cảnh.

Ngữ pháp vũ trụ (*universal grammar*) bao gồm tất cả các loại ngữ pháp.

Để giải quyết vấn đề ngữ đoạn động từ $VP \rightarrow V NP$, người ta có thể biến quy tắc này thành loại *ngữ pháp cảm ngữ cảnh* bằng cách xem động từ V như một phần tử không tận cùng và tiếp tục khai triển nó:

- (37) $V \rightarrow V_i$
 $V \rightarrow V_t NP$
 $V \rightarrow V_p PP$
 $V \rightarrow V_{tp} NP PP$
 $V \rightarrow V_{t2} NP NP$

b. Tiểu phạm trù và đặc điểm cú pháp

Giải quyết vấn đề bằng *ngữ pháp cảm ngữ cảnh* không phải là cách tối ưu. Loại ngữ pháp này rất phức tạp mà người ta cố tránh. Thay vì thế, người ta khai thác những đặc điểm của *tiểu phạm trù*.

Cụ thể, ta có thể xem *phạm trù từ vựng* là một *thuộc tính* chứa những trị *danh từ*, *động từ*, *tính từ*, ... Cũng với tinh thần ấy, thuộc tính là phạm trù ngữ nghĩa *bậc hai* (*secondary*) chứa các thuộc tính như *phái tính*, *số lượng*, *thể cách*, ... Phái tính có các trị như giống đực, giống cái, giống trung. Số lượng có các trị như số nhiều, số ít. Thể cách có các trị như danh cách, đối cách, tặng cách.

Xem một ví dụ về những đặc điểm của danh từ “*Kind*” (trẻ con) trong một cụm từ tiếng Đức dưới đây:

- (38) *liebt die Weisenkinder.*
 (hãy thương trẻ mồ côi):

Danh từ nguyên mẫu “*Kind*” nằm dưới dạng số nhiều (*Kinder*), nói chung gồm các thuộc tính:

Thuộc tính	Trị
Phạm trù	Danh từ
Phái tính	giống trung
Số lượng	số nhiều
Thể cách	đối cách (accusative) đóng vai bổ ngữ trực tiếp

Ở trường hợp khác:

- (39) *helft den Weisenkindern*
(hãy giúp trẻ mồ côi)

ta thấy động từ “*helft*” (nguyên mẫu: *helfen* = *giúp*) biến “*Kind*” thành *bổ ngữ gián tiếp*. Bổ ngữ gián tiếp được chia dưới dạng *tặng cách* (*dative*). Trong khi đó, tiếng Anh hoặc tiếng Việt không có cách biến thể này.

Nếu một thuộc tính mang một đặc điểm thực sự, thì đặc điểm ấy được xem như có tính *duang*, ngược lại, có tính *âm*. Đặc điểm này được ghi chú bằng dấu + hoặc – trong ngoặc []. Vài ví dụ:

- (40) Danh từ có đặc điểm phạm trù danh từ: $N \rightarrow [+N]$
 (41) Động từ có đặc điểm phạm trù động từ: $V \rightarrow [+V]$
 (42) Tính từ có đặc điểm phạm trù tính từ: $A \rightarrow [+A]$
 (43) “*Cái bàn*” có đặc điểm danh từ số ít: $[+N, +\text{số ít}]$. “*Cái bàn*” không chứa đặc điểm động từ và số nhiều: $[-V, -\text{số nhiều}]$.
 (44) Đặc điểm “*Kind*” trong “*liebt die Weisenkinder*”: $[+N, +\text{trung}, +\text{số nhiều}, +\text{đối cách}]$.

c. Tiểu phạm trù phi ngữ cảnh

Để dễ hiểu ***tiểu phạm trù phi ngữ cảnh*** (*context free category*), hãy lấy một đặc tính tiêu biểu của danh từ làm ví dụ: ***danh từ riêng*** (*proper noun*) và ***danh từ chung*** (*common noun*).

Danh từ riêng là những tên riêng như Tí, Sửu, Huế, Sài Gòn, Hà Nội, ... Trái lại, *danh từ chung* là những danh từ chỉ những thứ chung chung đại loại như bàn ghế, xe cộ, nhà cửa, ... nói chung là mọi thứ không chứa tính chất gì riêng tư cả. Hai đặc điểm này đối lập nhau trong phạm trù từ vựng danh từ.

Một danh từ có thể được đánh giá hoặc là danh từ chung, hoặc là danh từ riêng qua ký hiệu âm/dương:

$$(45) \quad [+N] \rightarrow \begin{cases} [+danh \text{ từ chung}] \\ [-danh \text{ từ chung}] \end{cases}$$

hoặc viết gọn:

$$(46) \quad [+N] \rightarrow [\pm danh \text{ từ chung}]$$

Loại quy tắc này được gọi là ***quy tắc tiểu phạm trù phi ngữ cảnh (context free subcategory rule)***.

Danh từ chung còn chứa những tính chất khác như *đếm được* hay *không đếm được*, *cụ thể* hay *không cụ thể*. Từ đó mới có thêm quy tắc thường thấy ở ngôn ngữ Ấn-Âu:

$$(47) \quad [+danh \text{ từ chung}] \rightarrow [\pm \text{đếm được}, \pm \text{cụ thể}]$$

Ví dụ tiếng Đức:

$$(48) \quad [+danh \text{ từ chung}, +\text{đếm được}, +\text{cụ thể}], \text{ như:} \\ \text{Haus (nhà), Auto (xe hơi), Kuh (bò), Tisch (bàn)}$$

$$(49) \quad [+danh \text{ từ chung}, +\text{đếm được}, -\text{cụ thể}], \text{ như:} \\ \text{Idee (sáng kiến), Gedanke (suy nghĩ)}$$

$$(50) \quad [+danh \text{ từ chung}, -\text{đếm được}, +\text{cụ thể}], \text{ như:} \\ \text{Wasser (nước), Sand (cát), Möbel (bàn ghế)}$$

$$(51) \quad [+danh \text{ từ chung}, -\text{đếm được}, -\text{cụ thể}], \text{ như:} \\ \text{Liebe (tình yêu), Glauben (niềm tin)}$$

Tính “*cụ thể*” còn được xem có “*sống*” (alive) thật không:

(52) [+cụ thể] → [±sống]

Nếu sống, có mang tính “*người*” không:

(53) [+sống] → [±người]

Sau khi phân tích từ tổng quát đến chi tiết, ta thấy, một danh từ có thể làm chủ một tập hợp chứa những đặc tính nhất định, ví dụ như:

(54) N --> [+N]

(55) N --> [+N, +danh từ chung]

(56) N --> [+N, +danh từ chung, +đếm được, +cụ thể]

(57) N --> [+N, +danh từ chung, +đếm được, +cụ thể, +sống]

(58) N --> [+N, +danh từ chung, +đếm được, +cụ thể, +sống, -người]

Đối với tiếng Đức, tiếng Anh, những danh từ thú vật “*Kuh*”(bò), “*cow*”, “*Pferd*”(ngựa), “*mustang*”, “*Hund*”(chó), “*dog*”, “*Katze*”(mèo), “*cat*”, ... là những danh từ chung, đếm được, cụ thể, sống, không mang tính người:

(59) N --> [+N, +danh từ chung, +đếm được, +cụ thể, +sống, -người]

Nhân đây xin lưu ý. Khi bàn đến tính đếm được trong tiếng Việt, chúng ta nên cẩn thận. Trong tiếng Việt có nhiều danh từ cụ thể tưởng chừng đếm được nhưng lại không. Người Việt không nói “*một bò*”, “*hai ngựa*”, “*ba chó*”, “*bốn mèo*”, ... mà nói “*một con bò*”, “*hai con ngựa*”, “*ba con chó*”, “*bốn con mèo*”, ... tức là phải dùng thêm *danh từ hình thức* (chẳng hạn “*con*”) trước danh từ chính.

d. Tính nghiêm ngặt của tiểu phạm trù

Một từ trong cấu trúc cú pháp được công nhận là một phần tử tận cùng nếu đặc điểm của nó giữ tính nghiêm ngặt khắp nơi. Dựa vào nguyên tắc này, người ta có thể giải quyết vấn đề ngữ đoạn động từ VP → V NP tốt hơn. Mỗi loại động từ sẽ được đánh dấu bằng **SUBCAT_x**; *x* là số đặc trưng cho mỗi tiểu phạm trù. Thay vì dùng các quy tắc:

- (60) $VP \rightarrow V_i$
- (61) $VP \rightarrow V_t NP$
- (62) $VP \rightarrow V_p PP$
- (63) $VP \rightarrow V_{tp} NP PP$
- (64) $VP \rightarrow V_{t2} NP NP$

ta dùng:

- (65) $VP \rightarrow [+V, SUBCAT0]$
- (66) $VP \rightarrow [+V, SUBCAT1] NP$
- (67) $VP \rightarrow [+V, SUBCAT2] PP$
- (68) $VP \rightarrow [+V, SUBCAT3] NP PP$
- (69) $VP \rightarrow [+V, SUBCAT4] NP NP$

- SUBCAT0 biểu thị động từ bất xuyên tính
- SUBCAT1 biểu thị động từ xuyên tính
- SUBCAT2 biểu thị động từ đi với ngữ đoạn giới từ PP
- SUBCAT3 biểu thị động từ xuyên tính đi với ngữ đoạn giới từ PP
- SUBCAT4 biểu thị động từ nhị xuyên với hóa trị 2

Một ví dụ ứng dụng kỹ thuật lập trình: Thử thành lập ***bảng từ vựng (lexical bank)*** với một bảng từ chứa những thông tin về đặc điểm của từng từ như:

Tiếng Anh:

- (70) play: [+V, SUBCAT0]
- (71) kick: [+V, SUBCAT1]

Tiếng Đức:

- (72) spielen: [+V, SUBCAT0]
- (73) kicken: [+V, SUBCAT1]

Tiếng Việt:

- (74) chơi: [+V, SUBCAT0]
- (75) đá: [+V, SUBCAT1]

v.v.

3. Quan hệ chủ đề và vai trò “theta”

Việc dùng bổ ngữ còn lệ thuộc vào tiểu phạm trù; nói chính xác là còn lệ thuộc vào sự **giới hạn tiểu phạm trù hóa** (*subcategorization restriction*). Không những vậy, một bổ ngữ hợp lý còn tùy thuộc vào sự giới hạn của ngữ nghĩa. Chuẩn tắc ngữ nghĩa nói lên sự **giới hạn chọn lựa** (*selectional restriction*). Sau đây, thử xem các nhà ngôn ngữ học Tây phương giải thích thế nào về mối quan hệ lệ thuộc giữa các phần tử tham gia qua lý thuyết **quan hệ chủ đề** (*thematic relation*) và **vai trò theta** (*theta role*).

Một trong những cách giải thích sự giới hạn chọn lựa là làm sáng tỏ mối *quan hệ chủ đề*. Mối quan hệ này bao hàm các vế ngữ nghĩa được dùng để mô tả sự hòa hợp của các tham số đối với *vị từ*.

Khi nói đến hành động, người ta có thể tưởng tượng như sau:

Tác nhân (*agent, actor*) là người gây ra hành động, *người khởi động* (*initiator*). Tác nhân phải là một vật thể *sống* (*alive*) và *có khả năng làm theo ý muốn*.

(76) Bill kicks the ball.

(77) Bill đá trái banh.

Bill là một tác nhân, một vật thể sống, có khả năng hành động theo ý muốn là đá trái banh. Tuy nhiên, không phải một câu bao giờ cũng có một tác nhân theo tinh thần vừa kể.

(78) “The rain is falling on the leaves”.

“*The rain*” (mưa) không được kể là tác nhân vì mưa không phải là vật thể sống. Mưa chỉ là hiện tượng thiên nhiên.

Dù tác nhân có khả năng làm theo ý muốn, song nhiều khi một hành động không hẳn phải là sự cố ý:

(79) Bill accidentally broke the glass.

Bill vẫn được kể là tác nhân theo định nghĩa ở trên. Tác nhân thường là chủ ngữ, nhưng vẫn có thể xuất hiện ở vị trí khác.

Cảm nhân (experiencer) là người có kinh nghiệm, hoặc cảm nhận, hoặc hiểu được một cái gì đó. Cảm nhân có thể xuất hiện như một tham số nằm trong phần chủ ngữ hoặc phần bổ ngữ:

(80) *Monica* loved Bill.

(81) Bill kissed *Monica*.

Monica là cảm nhân. *Monica* có thể đóng vai chủ ngữ và bổ ngữ.

Những gì xuyên qua hành động được nói tới, được cảm nhận, được hiểu như ý chính, thì được coi như **chủ đề (theme)**.

(82) Monica talked about *her love story*.

“*Her love story*” là chủ đề, là câu chuyện tình được nói tới.

Điểm cần hướng tới được hiểu là **mục tiêu (goal)**.

(83) Monica went to *Bill*.

Bill là mục tiêu mà Monica muốn tiến tới.

Có một loại mục tiêu đặc biệt: **thụ nhân (recipient)**, tức người nhận. Thụ nhân đi với động từ sẽ làm thay đổi sở hữu chủ:

(84) Monica gave *Bill* her hand.

(85) *Bill* was kissed by Monica.

Bill là mục tiêu. Ngược với mục tiêu là **nguồn gốc (source)**, tức khởi điểm chuyển động. *Monica* ở hai ví dụ trên là nguồn gốc.

Hiện trường xảy ra biến cố được gọi là **vị trí (location)**:

(86) Monica met Bill in *Washington*.

Những gì được dùng để gây ra hành động được hiểu là **dụng cụ (instrument)**:

(87) *Monica* took a *Taxis* to visit Bill.

Taxis là dụng cụ.

Còn một vị trí đặc biệt nữa được gọi là cá thể được phục vụ, ***cống hiến (benefactive)***.

(88) Monica bought the flowers for *Bill*.

(89) She cooked *him* dinner.

Cần để ý, một chỉ định ngữ DP không phải bao giờ cũng chỉ có một quan hệ chủ đề. Chẳng hạn chỉ định ngữ Monica sau đây chứa một quan hệ chủ đề vừa là tác nhân, vừa là nguồn gốc:

(90) *Monica* wrote a letter to Bill.

Đã gọi là *quan hệ (relation)* được hiểu theo toán học thì không thể bảo đảm sự liên hệ của nó giống như phép áp " $1 \rightarrow 1$ ", nghĩa là từ điểm gốc, nó chỉ chiếu qua một điểm ảnh theo nguyên tắc " 1 nối 1 ". Quan hệ chủ đề và tham số tương tự vậy. Vì lẽ đó mà các nhà ngôn ngữ Tây phương đã đề nghị một cấu trúc đặc biệt được gọi là ***vai trò theta (θ role)***. Vai trò này cho phép thành lập quan hệ " $1 \rightarrow 1$ " dùng cho tham số (giống hình thức hàm số). Muốn vậy phải xem vai trò θ tương đương với một tham số nằm dưới dạng một tập hợp các quan hệ chủ đề dính chùm thành một. Ví dụ tham số Monica ở trên vừa là tác nhân, vừa là nguồn gốc, tức một tập hợp gồm hai tham số {tác nhân, nguồn gốc} nhưng chỉ giữ một vai trò θ duy nhất.

Hãy xem vai trò θ mô tả cấu trúc tham số của động từ. Thử xét một động từ ví dụ như "*put*". Động từ "*put*" đòi hỏi 3 tham số:

1. Một chủ ngữ làm tác nhân.
2. Một bổ ngữ trực tiếp làm chủ đề.
3. Một bổ ngữ gián tiếp làm mục tiêu hay vị trí.

(91) Bill put the Samsonite on the table.

Ngoài cách sử dụng này ra, tất cả mọi cách dưới đây đều sai ngữ pháp do thiếu hoặc dư tham số:

(92) * Put the Samsonite on the table.

(93) * Bill put the Samsonite.

(94) * Bill put on the table.

(95) * Bill put the Samsonite on the table the hut.

Câu sau cũng sai ngữ pháp:

(96) * The rain put the mountain with the car.

Câu này vô nghĩa vì các chỉ định ngữ DP không đóng đúng vai trò theta. “*The rain*” không thể đóng vai trò tác nhân, “*the mountain*” không thể đóng vai trò chủ đề và cũng không thể lấy *xe hơi* (*with the car*) chở cả quả núi đi chỗ khác.

Để làm sáng tỏ sự đòi hỏi tham số của động từ, người ta thường dùng **lưới theta** (**theta grid**) (Haegeman 1991). Một ví dụ cho động từ “*put*”:

Trước nhất, vẽ ra 3 cột tương ứng với 3 tham số mà vị từ “*put*” đòi hỏi. Mỗi cột có hai dòng. Dòng trên ghi tên vai trò theta (có thể có nhiều vai trò khác nhau). Dòng dưới xác định chỉ số của từng vai trò:

(97) [Bill]_i put [the Samsonite]_j on [the table]_k.

Chỉ số *i* biểu thị vai trò tác nhân “*Bill*”, chỉ số *j* biểu thị chủ đề “*the Samsonite*”, ...

(98) “*put*”:

ngoại chủ đề	nội chủ đề	
<u>gốc/tác nhân</u>	chủ đề	mục tiêu
i	j	k

Vai trò theta có hai loại (Marantz 1984):

- **Vai trò theta ngoài (external theta role)** chỉ chủ ngữ. Vai trò này thường được gạch dưới (trong lưới).
- **Vai trò theta trong (internal theta role)** chỉ bổ ngữ trực tiếp (và gián tiếp).

TIỂU PHẠM TRÙ HÓA

Đã xuyên qua tiểu phạm trù và vai trò theta, nay chúng ta tiếp tục xét chi tiết hơn về vai trò cú pháp của từ vựng.

Một ngữ pháp tạo sinh gồm có ba thành phần:

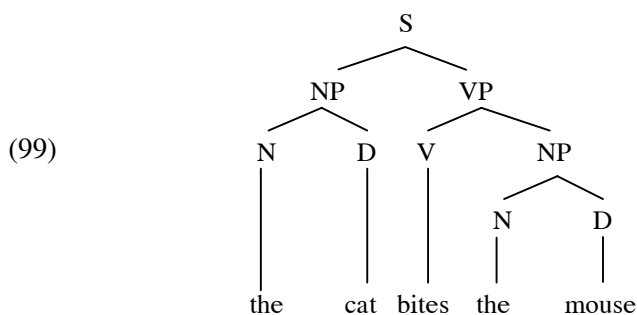
1. Một tập hợp quy tắc ngữ đoạn, còn được gọi là *quy tắc phạm trù* (*categorial rule*).
2. Một kho từ vựng chứa từ ngữ của từng ngôn ngữ, trong đó mỗi từ đều có những thuộc tính cú pháp.
3. Một *quy tắc lồng từ vựng* (*lexical insertion rule*) chỉ cách gắn từ vào đúng vị trí trong ngữ đoạn.

Ví dụ:

1. Quy tắc ngữ đoạn:
 - (i) $S \rightarrow NP VP$
 - (ii) $NP \rightarrow D N$
 - (iii) $VP \rightarrow V NP$
2. Từ vựng:

“bite” = V, “chase” = V
 “dog” = N, “cat” = N, “mouse” = N
 “the” = D, ...
3. Quy tắc lồng từ vựng:

Từ có thể được gắn vào nút tận cùng có cùng phạm trù từ vựng.



Trong ví dụ trên, “*bite*” là động từ xuyên tính. Nó cần một bổ ngữ NP. Ta ký hiệu:

(100) *bite*: V, +[- NP]

Ký hiệu muốn nói, “*bite*” là một động từ và có thể gắn vào vị trí “-” trước NP. (“-” là chỗ dành sẵn cho V). Trong khi đó:

(101) *fall*: V, -[- NP]

“*fall*” là động từ bất xuyên tính, không thể đứng trước bổ ngữ NP (được ký hiệu là - [- NP]). Không thể nói:

(102) * the cat falls the mouse

Theo cách trình bày trên, ta nói, “*bite*” đã được *tiểu phạm trù hóa* (*subcategorised*) và cho phép một ngữ đoạn danh từ NP theo sau. Những thông tin đóng trong ngoặc [...] được gọi là *khung tiểu phạm trù hóa* (*subcategorisation frame*).

Một động từ vừa xuyên tính, vừa bất xuyên tính cũng có thể được ký hiệu tương tự, ví dụ động từ “*eat*”:

(103) *eat*: V, ±[- NP]

Theo mô hình vừa kể, một phần tử từ vựng không những chỉ chứa thông tin về phạm trù mà còn chứa thông tin về tiểu phạm trù hóa. Mở rộng quy tắc lồng từ vựng:

Quy tắc lồng từ vựng (lexical insertion rule) cho phép gắn một từ đã được tiểu phạm trù hóa vào một nút tận cùng có cùng phạm trù từ vựng.

Một ví dụ khác:

(104) *put*: V, +[- NP PP]

Cách tiểu phạm trù hóa “*put*” cắt nghĩa được tính đúng/sai ngữ pháp với bổ ngữ “*on the table*” là ngữ đoạn giới từ PP :

(105) Bill put the Samsonite on the table.

(106) * Bill put the Samsonite.

(107) * Bill put on the table.

Cũng là ngữ đoạn giới từ PP, song nếu dùng các giới từ khác như “for”, “from”, “to”, câu sẽ sai ngay:

(108) *Bill put the Samsonite for the table.

(109) *Bill put the Samsonite from the table.

(110) *Bill put the Samsonite to the table.

Như vậy, dùng bổ ngữ PP của “put” sao cho thích hợp còn tùy vào vị trí (location). Ta thêm ký hiệu LOC (viết tắt của *locative*) để trở nét này.

(111) put: V, + $\left[\begin{array}{cc} \text{NP} & \text{PP} \\ & +\text{LOC} \end{array} \right]$

X-GẠCH HẠN CHẾ

Lý thuyết X-gạch tổng quát có thể đề ra câu sai ngữ pháp. Tính hạn chế có nhiệm vụ lọc trường hợp sai. Nó cần một chuẩn tắc được gọi là:

Chuẩn tắc theta (theta criterion):

- i. Mỗi tham số phải nhận một và chỉ một vai trò theta.
- ii. Mỗi vai trò theta phải nhận một và chỉ một tham số.

Cụ thể, thử phân tích động từ “love” bằng lưới theta.

Đối với một câu chứa vị từ “love”, ta xem từng tham số đóng vai trò gì và cho nó một chỉ số tương ứng. Chẳng hạn câu:

(112) [Monica]_i loves [Bill]_j.

Câu này đúng ngữ pháp vì dùng đúng số lượng tham số và vai trò. Monica là cảm nhân và Bill là chủ đề.

(113) “love”:

<u>cảm nhân</u>	chủ đề
i	j

Ngược lại, một câu sai ngữ pháp vì thiếu chủ đề:

(114) [Monica]_i loves.

“love”:

<u>cảm nhân</u>	chủ đề
i	

Điểm này không thỏa chuẩn tắc theta, bởi *mỗi vai trò theta phải nhận một và chỉ một tham số*.

Câu sau cũng sai ngữ pháp vì dư tham số:

(115) [Monica]_i loves [Bill]_j . [Tom]_k

“love”:

<u>cảm nhân</u>	chủ đề	
i	j	k

Chỉ có 2 vai trò theta nhưng có đến 3 tham số. “Tom” không nhận vai trò theta nào cả. Điểm này không thỏa chuẩn tắc: *Mỗi tham số phải nhận một và chỉ một vai trò theta*.

MÔ HÌNH NGỮ PHÁP

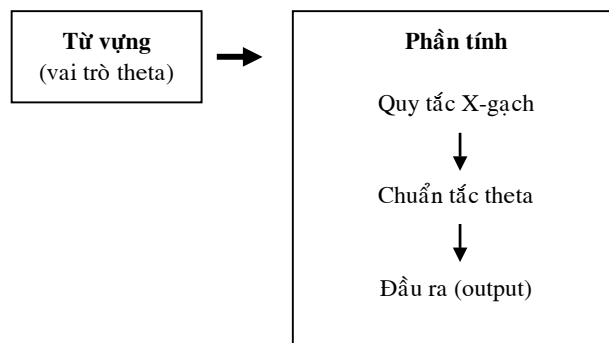
Cú pháp học là một *khoa học nhận thức (cognitive science)*. Theo Chomsky, sự hiểu biết ngôn ngữ của con người dựa trên hai thành phần căn bản: **từ vựng (lexical)** và **phần tính (computational component)**.

Phần tính chứa đựng mọi quy tắc và tính hạn chế. Nhờ quy tắc, con người có thể tạo sinh ra câu cú, và cũng qua đó, nhờ hạn chế, những ngữ nghĩa vô lý sẽ bị loại trừ. Phần tính không thể làm việc với “chân không” mà cần từ vựng.

Từ vựng là một cuốn từ điển sống khổng lồ của con người. Nó chứa đựng tất cả các thông tin về từ. Đối với mỗi vị từ, mọi vai trò theta đều được định nghĩa rõ ràng, tinh vi và chỉ riêng cho vị từ ấy. Từ vựng còn bao gồm cả những trường hợp ngoại lệ. Như một ***đơn vị từ vựng (lexical item)***, mỗi từ chứa ít nhất những thông tin:

- Ý nghĩa
- Cách phát âm
- Những thông tin ngoại lệ
- Lưới theta

Học một từ mới nghĩa là phải học và nhớ những thông tin kể trên. Ta có thể tưởng tượng một mô hình của hệ thống ngữ pháp trừu tượng như sau:



Từ vựng cung cấp cho *phần tính* mọi thông tin tương ứng. Phần tính sẽ ráp từ vào quy tắc X-gạch sau khi lọc lựa từ theo những chuẩn tắc theta nhất định. Sản phẩm cuối cùng là câu ở đầu ra. Trong quá trình này, mọi dữ liệu từ vựng được sử dụng phải có khả năng mô tả mọi bình diện sao cho đúng ngữ pháp để tạo ra hình thức câu.

NGUYÊN TẮC CHIẾU MỞ RỘNG

Theo nguyên tắc vai trò theta, mỗi vị từ cần sự góp mặt của các vai trò theta. Song dĩ nhiên cũng có những trường hợp ngoại lệ, tỉ dụ như trong tiếng Anh, tiếng Đức. Vài ví dụ:

a. Chủ ngữ giả như “it”, “es”:

- (116) It rains. (tiếng Anh)
- (117) Es regnet. (tiếng Đức)
- (118) Trời mưa.
- (119) It snows. (tiếng Anh)
- (120) Es schneit. (tiếng Đức)
- (121) Trời tuyết.

Rõ ràng các vị từ diễn tả thời tiết ở đây không chấp nhận các đại danh từ “it”, “es” như các vai trò theta. Thế nhưng vì như đã nói, các ngôn ngữ Âu châu thiên về cấu trúc chủ-vị, tức vị từ bắt buộc phải có chủ ngữ, cho nên người ta phải dùng các đại danh từ “it”, “es” như một hình thức **bổ thể (expletive)** để trám vào chỗ trống đó. Ngoài ra, còn có trường hợp:

b. Chủ ngữ là một mệnh đề:

- (122) [_{CP} that Monica loves Bill] is likely.

Mệnh đề “that Monica loves Bill” là một bổ ngữ. Câu trên có thể hiểu cùng một nghĩa mặc dù bị đảo ngược. “It” lại đóng vai bổ thể:

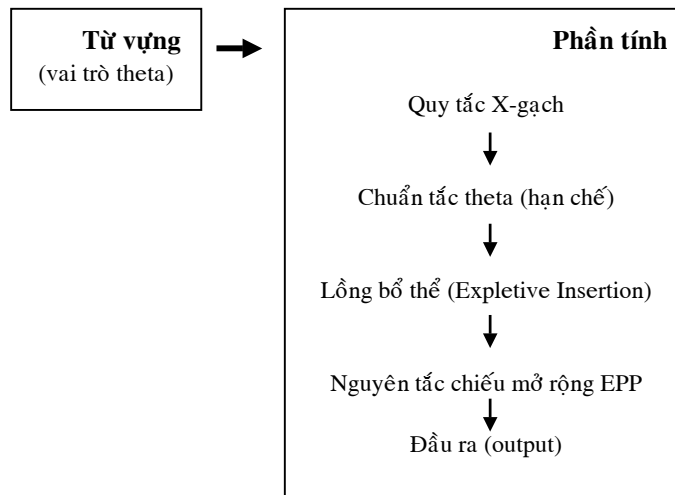
- (123) It is likely [_{CP} that Monica loves Bill].

Để giải thích các trường hợp trên, vai trò theta được mở rộng:

Nguyên tắc chiếu mở rộng (Extended Projection Principle (EPP)): Tất cả các mệnh đề phải có chủ ngữ. Dữ liệu từ vựng phải được mô tả rõ ràng.

Nếu sơ đồ X-gạch không chọn lựa được vai trò theta thích hợp, nó sẽ loại bỏ câu này. Gặp trường hợp hơi khắc nghiệt này, các nhà ngôn ngữ học đã đề nghị thêm quy tắc **lồng bổ thể (expletive insertion)**. Nguyên tắc này có thể hiểu đơn giản là phải lồng các đại danh từ “it”, “es” vào câu để chúng giữ chức năng chủ ngữ.

Mô hình của hệ thống ngữ pháp cũ được mở rộng như sau:



TÓM LƯỢC

Vị từ diễn tả mối quan hệ của những phần tử có mặt trong câu. Những phần tử này được gọi là *tham số*. Mỗi vị từ có một số tham số nhất định tùy loại bất xuyên tính, xuyên tính, ...

Lý thuyết truyền thống liệt động từ vào phạm trù từ vựng tiêu chuẩn *động từ V*, song trên thực tế, động từ có nhiều loại khác nhau. Chúng được chia nhỏ thành những *tiểu phạm trù* nhằm bảo đảm cách dùng động từ sao cho hợp với số lượng tham số.

Ý tưởng chia nhỏ phạm trù thành những tiểu phạm trù làm mất tính tổng quát của quy tắc $VP \rightarrow V NP$. Tiểu phạm trù mang tính ***cục bộ nghiêm ngặt***. Điều này dẫn đến vấn đề là làm sao có thể mô tả được cái *cấu trúc toàn thể* của ngữ đoạn động từ VP. Một trong những giải pháp là dựa vào loại *ngữ pháp cảm ngữ cảnh*. Người ta xem động từ V như một phần tử không tận cùng và tiếp tục khai triển nó:

- $V \rightarrow V_i$
- $V \rightarrow V_t \text{ NP}$
- $V \rightarrow V_p \text{ PP}$
- $V \rightarrow V_{tp} \text{ NP PP}$
- $V \rightarrow V_{t2} \text{ NP NP}$

Giải quyết vấn đề bằng *ngữ pháp cảm ngữ cảnh* không phải là phương cách tối ưu. Có một sáng kiến khác là khai thác những đặc điểm của tiểu phạm trù: thuộc tính của từ. Tốt nữa là dùng *tiểu phạm trù phi ngữ cảnh*. Ví dụ:

- $\text{VP} \rightarrow [+V, \text{SUBCAT0}]$
- $\text{VP} \rightarrow [+V, \text{SUBCAT1}] \text{ NP}$
- $\text{VP} \rightarrow [+V, \text{SUBCAT2}] \text{ PP}$
- $\text{VP} \rightarrow [+V, \text{SUBCAT3}] \text{ NP PP}$
- $\text{VP} \rightarrow [+V, \text{SUBCAT4}] \text{ NP NP}$

- SUBCAT0 biểu thị động từ bất xuyên tính.
- SUBCAT1 biểu thị động từ xuyên tính.
- SUBCAT2 biểu thị động từ đi với ngữ đoạn giới từ AP.
- SUBCAT3 biểu thị động từ xuyên tính đi với ngữ đoạn giới từ AP.
- SUBCAT4 biểu thị động từ nhị xuyên với hai hóa trị.

Một bổ ngữ hợp lý còn tùy thuộc vào sự giới hạn của ngữ nghĩa. Chuẩn tắc ngữ nghĩa nói lên sự *giới hạn chọn lựa* (*selectional restriction*).

Một trong những cách giải thích sự giới hạn chọn lựa là làm sáng tỏ mối *quan hệ chủ đề*. Mối quan hệ này bao hàm các vế ngữ nghĩa được dùng để mô tả sự hòa hợp của các tham số đối với *vị từ*.

Nói đến hành động, người ta có thể tưởng tượng đến những vai trò như: *tác nhân, cảm nhân, chủ đề, mục tiêu, thụ nhân, nguồn gốc, vị trí, dụng cụ, ...* Đây là những *vai trò θ (theta)*. Vai trò này cho phép thành lập quan hệ “1→1” dùng cho tham số (giống quan hệ 1→1 của hàm số). Vai trò theta tương đương với *một tham số* nằm dưới dạng một *tập hợp các quan hệ chủ đề* dính chùm thành một. Ví dụ, động từ “*put*” cần 3 tham số:

- Một chủ ngữ làm *tác nhân*.
- Một bổ ngữ trực tiếp làm *chủ đề*.
- Một bổ ngữ gián tiếp làm *mục tiêu* hay *vị trí*.

Vai trò cú pháp của từ vựng được xác định qua khung tiểu phạm trù hóa.

Lý thuyết X-gạch tổng quát có thể đề ra câu sai ngữ pháp, cho nên cần hạn chế. Cụ thể bằng *chuẩn tắc theta*:

- Mỗi tham số phải nhận một và chỉ một vai trò theta.
- Mỗi vai trò theta phải nhận một và chỉ một tham số.

Dù vậy vẫn không thể tránh những trường hợp ngoại lệ (tiếng Anh, tiếng Đức) như *chủ ngữ giả* (“it”, “es”), *chủ ngữ là một mệnh đề*. Vai trò theta được mở rộng bằng *nguyên tắc chiếu mở rộng*: Tất cả các mệnh đề phải có chủ ngữ. Thông tin từ vựng phải được mô tả rõ ràng. Ngoài ra, còn có thêm quy tắc *lồng bổ thể*. Quy tắc này cho phép lồng các đại danh từ “it”, “es” vào câu để giữ chức năng chủ ngữ.

NHẬN XÉT VÀ ĐIỂM MỞ

Ngôn ngữ Âu châu thiên về cấu trúc *chủ-vị* (*subject-predicate*). Đây là một trong những đặc điểm cơ sở được dùng trong lý thuyết cú pháp. Theo cấu trúc này, vị từ trong câu lệ thuộc vào chủ ngữ trong câu. Tiếng Việt cũng có cấu trúc này nhưng cũng có trường hợp chủ thể không nằm trong câu.

Vài đề nghị cho điểm mở:

- 1) Thử tìm thêm những vai trò theta khác mà tiếng Việt có thể có.
- 2) Thử ứng dụng khung tiểu phạm trù hóa để thành lập một bảng dữ liệu chứa vài ví dụ động từ, tính từ, ... đặc biệt là các từ “hả”, “đâu”, “à”, “nhá”, “kia”, “đi”, “nhỉ”, ... Thử xác định các trường hợp câu/mệnh đề/thành tố thường bị tỉnh lược.

Chương 12

PHÉP BIẾN HÌNH

Ý NGHĨA CỦA PHÉP BIẾN HÌNH

Mỗi câu nói trong thực tế đều có một cấu trúc cú pháp: *cấu trúc bề mặt (surface structure)*. Theo Radford (1981:146), có thể giả định rằng bình diện này miêu tả cái cấu trúc cú pháp thô thiển nhất của câu. Công việc chúng ta đã và đang gắng làm là đưa ra một cú pháp hợp lý nhằm xác định mọi cấu trúc bề mặt chuẩn hình của một ngôn ngữ. Cụ thể là dựa vào một tập hợp *quy tắc phạm trù* (ví dụ: cấu trúc ngữ đoạn), *từ vựng* và *quy tắc lồng*. Tuy nhiên, không phải là không tồn tại những trường hợp mà các loại quy tắc ấy không đủ khả năng giải thích cho hợp lý. Điển hình là *câu hỏi- wh (wh-question)*. Muốn giải thích, Chomsky đề nghị dùng *quy tắc biến hình (transformation rule)*.

Trước hết, hãy xem vài cách hỏi “*wh*” (*who, what, why, when, which, where, how*) trong tiếng Anh.

Ví dụ một người A nói:

(1) I bought a car.

Khi nghe vậy, một người B có thể ngạc nhiên, hỏi:

(2) You bought a car ?

Cách hỏi này là cách lặp lại cấu trúc câu nói của người A. Loại câu hỏi này (tạm) gọi là *câu hỏi hồi ngôn (echo question)*. Nó giống như câu hỏi có/không (yes/no question) bởi người A có thể trả lời bằng yes/no. Tương tự, người B cũng có thể hỏi:

(3) You bought a what ?

Cũng là một câu hỏi hồi ngôn nhưng người ta sẽ không trả lời bằng yes/no mà là “*a car*” chẳng hạn.

Đó là trường hợp lặp lại câu xác định. Ngoài ra còn có trường hợp “hồi ngôn” trong câu hỏi thực sự:

- (4) A: Did you buy a house ?
 (5) B: Did I what ?

Hoặc có thể hỏi hồi ngôn với câu cùng loại, ví dụ câu mệnh lệnh:

- (6) A: Don't touch my hand.
 (7) B: Don't touch your what ?

Những cách hỏi hồi ngôn vừa nêu rất phổ biến trong tiếng Việt.

- (8) A: Tôi mới mua cái nhà.
 (9) B: Anh mới mua cái gì ?

Ngược với câu hỏi hồi ngôn trong tiếng Anh là câu hỏi *phi hồi ngôn* (*nonecho question*), điển hình là câu hỏi-wh (s.s. Radford 1981:149). Thử xét câu sau đây:

- (10) Which car will your father put in the garage ?

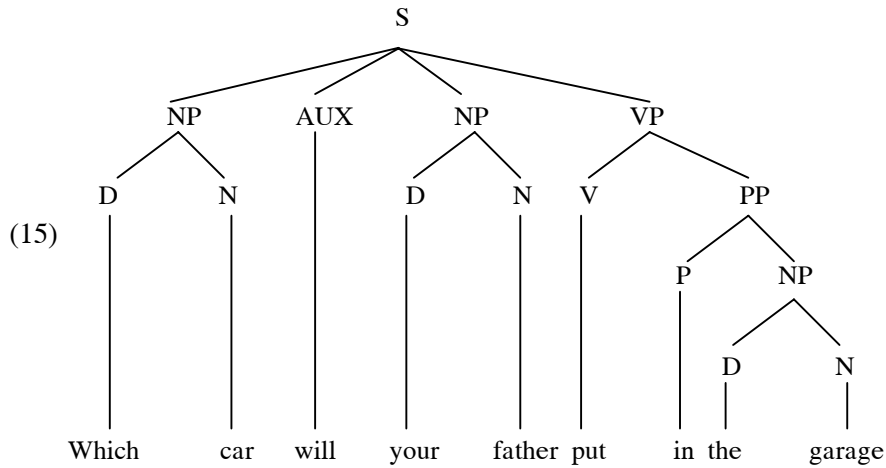
Nếu xem:

- “Which car”, “your father” là ngữ đoạn danh từ NP.
- “Will” là động từ phụ (trợ động từ) AUX.
- “Put in the garage” là ngữ đoạn động từ VP, với “put” là động từ V và “in the garage” là ngữ đoạn giới từ PP,

chiếu theo quy tắc cấu trúc ngữ đoạn:

- (11) $S \rightarrow NP\ AUX\ NP\ VP$
 (12) $VP \rightarrow V\ PP$
 (13) $PP \rightarrow P\ NP$
 (14) $NP \rightarrow D\ N$

ta có cây cú pháp:



“Put” nằm ở vị trí V là hợp lý. Song không phải bao giờ cách lồng này cũng đúng cho mọi trường hợp.

Chúng ta còn nhớ ở phần từ vựng, “put” được tiểu phạm trù hóa như sau:

$$(18) \text{ put: V, + } \left[\begin{array}{l} - \text{ NP PP} \\ + \text{ LOC} \end{array} \right]$$

“Put” cần một bổ ngữ NP và một bổ ngữ PP chỉ vị trí. Ví dụ một câu nói được:

(19) John put the car in the garage.

trong khi đó, những trường hợp sau không nói được vì thiếu bổ ngữ:

(20) *John put the car.

(21) *John put in the garage.

Chuỗi từ “John put in the garage” không nói được vì thiếu bổ ngữ NP, (put cái gì). Vậy, có thể suy ra như thế này được không, “Which car will your father put in the garage” không nói được vì trong câu cũng xuất hiện một cấu trúc thiếu bổ ngữ NP tương tự “your father put in the garage” ?

Mặt khác, ở một câu nói được:

(22) Will your father go in the garage ?

ta không cần bổ ngữ NP bởi vì “go” được tiểu phạm trù hóa:

(23) $go: V, + \left[\begin{array}{c} - PP \\ +LOC \end{array} \right]$

Tuy nhiên, không phải vì thế mà có thể nói:

(24) *Which car will your father go in the garage ?

Câu này sai ngữ pháp.

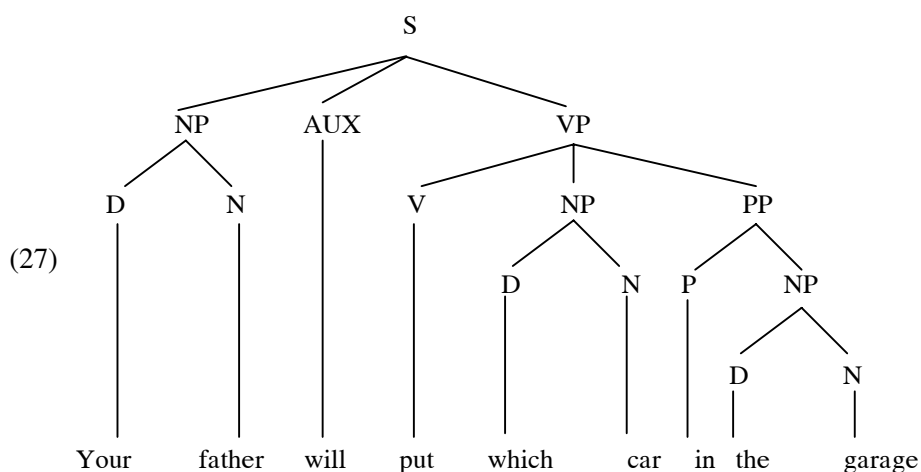
Nói tóm lại, rõ ràng, những gì ta đã biết trước nay không thể giải thích được trường hợp câu hỏi-wh, nếu wh nằm đầu câu (v.d. “which car”).

Quay về trường hợp câu hỏi hỏi ngôn.

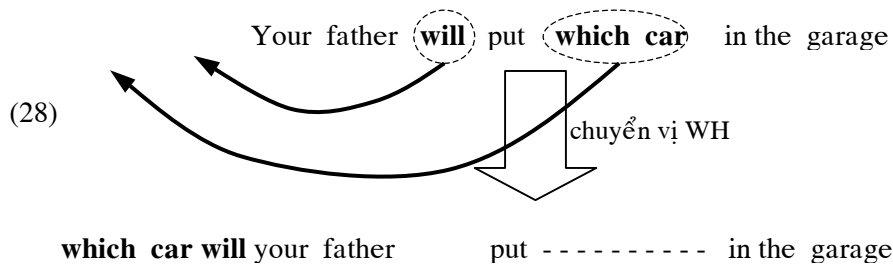
(25) A: My father will put the Mercedes in the garage.

(26) B: Your father will put which car in the garage ?

Cấu trúc câu xác định và câu nghi vấn giống nhau:



Nếu “*which car*” nằm sau động từ, thì không có vấn đề. Còn không, để giải thích câu “*Which car will your father put in the garage*” trong cách hành ngôn ngoài đời, ta phải dùng đến quy tắc **chuyển vị-wh** (*wh-movement*), nghĩa là chuyển vị trí “*wh*” (“*which car*”).



Vị trí “-----” là chỗ của “*which car*” lúc chưa bị chuyển vị. Song song đó động từ phụ “*will*” cũng bị chuyển vị.

Qua phân tích, người ta có thể tiên giả định: có đến hai bình diện cấu trúc cú pháp: Một là cấu trúc bề mặt liên quan, hai là thêm một bình diện cấu trúc trừu tượng (Chomsky gọi là *cấu trúc sâu* (*deep structure*) (*D-Structure*, *Base Structure*, *Underlying Structure*)). Ngay trường hợp này, *Underlying Structure* có liên quan đến cấu trúc bề mặt trong quá trình dùng quy tắc biến hình (cụ thể: chuyển vị-wh).

Đó là ý nghĩa của phép biến hình. Tiếp theo, thử xem lại các quy tắc X-gạch (Carnie 2000).

Lý thuyết X-gạch nguyên thủy đã dựa trên nền tảng ngữ pháp tiếng Anh. Các quy tắc ngữ pháp tiếng Anh có thể được tóm gọn bằng 3 quy tắc chính:

- (29) Quy tắc biệt định ngữ: $X' \rightarrow (YP) X'$
 (30) Quy tắc phụ ngữ: $X' \rightarrow X' (ZP) \mid (ZP) X'$
 (31) Quy tắc bổ ngữ: $X' \rightarrow X (WP)$

Vì lý do, các quy tắc trên hoặc không giải thích được cấu trúc các ngôn ngữ khác, hoặc chỉ từng phần, chúng đã được tổng quát hóa:

- (32) Quy tắc biệt định ngữ: $X' \rightarrow X' (YP) \mid (YP) X'$
 (33) Quy tắc phụ ngữ: $X' \rightarrow X' (ZP) \mid (ZP) X'$
 (34) Quy tắc bổ ngữ: $X' \rightarrow X (WP) \mid (WP) X'$

Mặc dù đã giải quyết xong khúc mắc này nhưng vẫn chưa hết vấn đề. Chính tiếng Anh cũng có những cấu trúc không thỏa các quy tắc tổng quát vừa nêu, điển hình là trường hợp mệnh đề phụ và câu có thì:

(35) *Mệnh đề phụ:* $S' \rightarrow (C) S$

(36) *Câu có thì:* $S \rightarrow NP (T) VP$

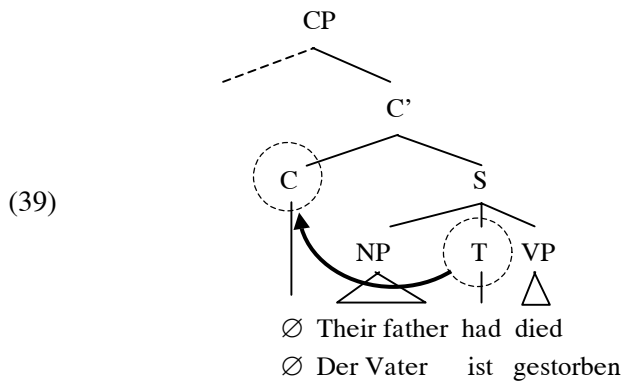
Để giải thích các trường hợp trên, giới ngôn ngữ tạo sinh đã đưa ra thêm 2 loại ngữ đoạn mới: **ngữ đoạn tác tử bổ ngữ hóa** (*complementizer phrase*) **CP** và **ngữ đoạn thì** (*tense phrase*) **TP**. Thế S' bằng CP, được:

(37) *Ngữ đoạn tác tử bổ ngữ hóa:* $CP \rightarrow C'$
 $C' \rightarrow (C) S$

Thế S bằng TP, ta được:

(38) *Ngữ đoạn thì:* $TP \rightarrow NP T'$
 $T' \rightarrow T VP$

Cuối cùng, cách giải thích ấy cũng dựa vào tính *chuyển vị*. Một ví dụ:



Ngôn ngữ trùng trùng điệp điệp. Chúng ta thừa biết, không phải ngôn ngữ nào cũng giống tiếng Anh, kể cả các trong nội bộ ngôn ngữ Tây phương (Pháp, Ý, Tây Ban Nha, Bồ Đào Nha, ...). Tuy nhiều ngôn ngữ Âu châu có dạng căn bản *Subject-Verb-Object* (SVO), song không phải bao giờ cũng tuyệt đối thế. Tiếng Đức là một ví dụ:

- (40) Sie mag Rosen sehr gern.
Subject verb Object (adv)
 Cô ta thích hoa hồng lắm
- (41) Rosen mag sie sehr gern.
Object verb Subject (adv)
 Hoa hồng thích cô ta lắm
 (Câu này có cấu trúc *Object-Verb-Subject (OVS)*)

Các ngôn ngữ Á châu như tiếng Trung Hoa, Nhật, ... lại càng không giống tiếng Âu châu.

Tiếng Thổ Nhĩ Kỳ:

- (42) [Can] [kitab-i] [oku-du].
Subject Object verb
 (Can (đã) đọc cuốn sách)

Tiếng Nhật:

- (43) [Phong-ga] [Hung-o] [mita].
Subject Object verb
 (Phong (đã) thấy Hùng)

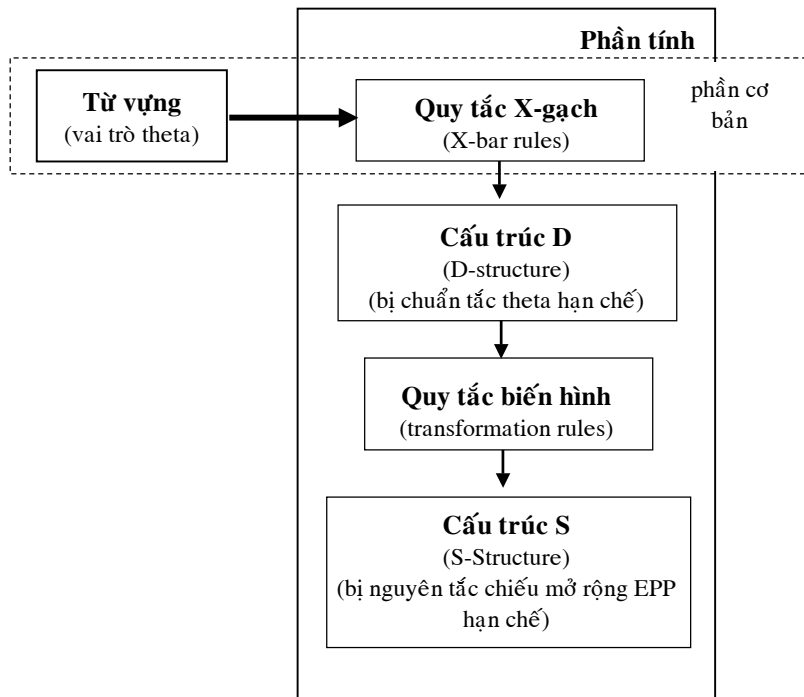
Cấu trúc với thứ tự *SVO* không dùng được cho những ngôn ngữ kể trên. Ngay tiếng Pháp mặc dù mang cấu trúc *SVO* nhưng không phải là không có vấn đề:

- (44) Je mange *souvent* des pommes.
 (I) (eat) (often) (of. the) (apples)
- (45) I often eat apples.
- (46) Tôi hay ăn táo.

Ngay ví dụ này, ta thấy trạng từ “*souvent*” của tiếng Pháp nằm giữa động từ “*mange*” (ăn) và “*pommes*” (táo), trong khi đó tiếng Anh không phải vậy và tiếng Việt cũng không.

Nói ngắn gọn, cấu trúc từng ngôn ngữ có thể giống mà cũng có thể khác nhau. Điển hình là hình thức câu hỏi của tiếng Việt giống như câu hỏi hỏi ngôn của tiếng Anh song những từ nghi vấn “*wh*” của tiếng Việt không bị chuyển vị như câu hỏi-*wh* của tiếng Anh.

Sự biến hình từ cấu trúc sâu sang cấu trúc bề mặt được giới ngôn ngữ tạo sinh giải thích qua mô hình dưới đây:



Phần từ vựng và phần quy tắc X-gạch được coi như phần cơ bản. Hai phần này giữ nhiệm vụ tạo sinh cây cú pháp. Kết quả đầu tiên là **cấu trúc D (D-structure)** hoặc **cấu trúc sâu (deep structure)**. Theo Chomsky, cấu trúc này miêu tả ngữ nghĩa nguyên thủy và có giá trị chung cho mọi ngôn ngữ. Sự kiện mỗi ngôn ngữ có cách nói khác nhau có thể giải thích được bằng quy tắc *biến hình*. Dựa vào quy tắc biến hình, thứ tự của từ trong cấu trúc D sẽ được chuyển vị sao cho thích hợp với ngữ pháp của ngôn ngữ ấy trong thực tế. Kết quả này là **cấu trúc S (S-structure)**, **cấu trúc bề mặt (surface structure)** sau khi đã được gạn lọc bằng *nguyên tắc chiếu mở rộng EPP*.

Đó là nguyên tắc, sau đây chúng ta đi sâu vào quy tắc biến hình.

Quy tắc biến hình có hai loại: **quy tắc chuyển vị (movement rules)** và **quy tắc lồng (insertion rules)**. *Quy tắc chuyển vị* có nhiệm vụ chuyển vị trí từ trong câu.

Quy tắc lồng thêm những từ mới vào câu. Xin lưu ý, quy tắc lồng ở đây không hoàn toàn đồng nghĩa với quy tắc lồng ở chương trước (sẽ được giải thích sau).

Trước nhất, hãy xem một loại quy tắc chuyển vị tạm gọi là ***chuyển vị đầu tới đầu*** (*head-to-head movement*).

CHUYỂN VỊ ĐẦU TỚI ĐẦU

Ý nghĩa của loại quy tắc biến hình này là chuyển vị trí của đầu ngữ đoạn này sang vị trí của đầu ngữ đoạn khác. Có hai trường hợp: *chuyển vị động từ* và *chuyển vị thì*.

1. Chuyển vị động từ V ($V \rightarrow T$)

Dùng lại ví dụ câu tiếng Pháp bên trên:

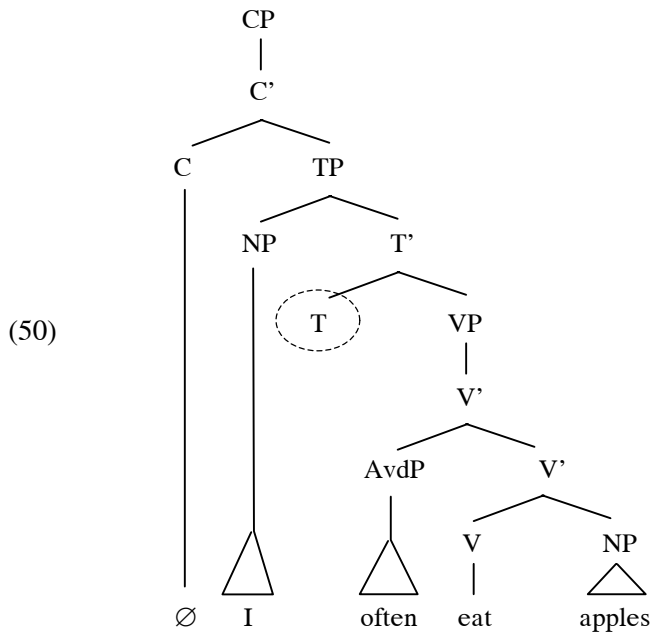
- (48) Je mange souvent des pommes.
(I) (eat) (often) (of the) (apples)

Giới ngôn ngữ học Tây phương cho đây là hiện tượng một *phụ ngữ* xuất hiện giữa *đầu ngữ đoạn động từ VP* và *bổ ngữ*. “Mange” là động từ, “des pommes” là bổ ngữ, đứng giữa đó là phụ ngữ “souvent”. Trường hợp này khác trường hợp phụ nghĩa bên trái động từ trong tiếng Anh, ví dụ “often”:

- (49) I often eat apples.

Người Việt cũng có thể nói tương tự: “tôi hay ăn táo”, “cô ta hay ăn táo”, “nó hay ăn táo”, v.v..

Theo nguyên tắc X-gạch, câu trên có cây cú pháp:



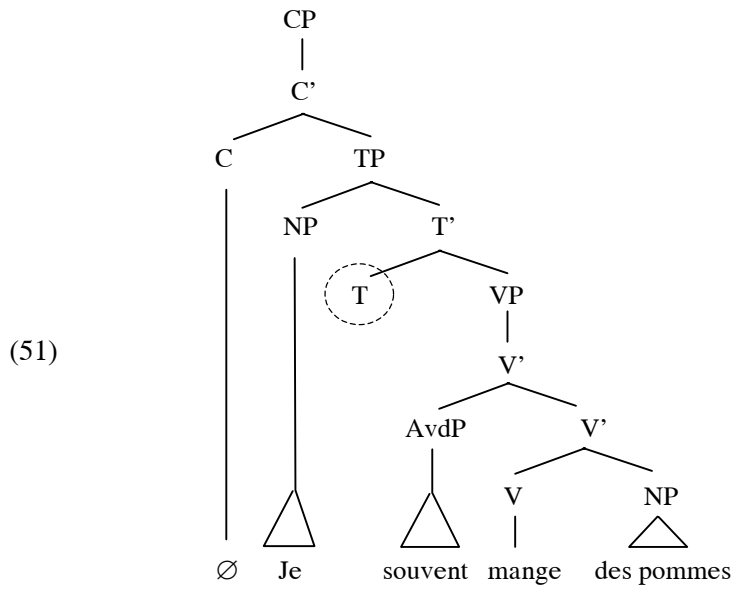
Hãy để ý, cấu trúc này có một đầu ngữ đoạn T đứng giữa NP và trạng từ (often). Nút T là chỗ dành riêng cho động từ phụ, nhưng ở đây không dùng tới. Bảng bên dưới cho thấy từng vị trí từ trong trường hợp không dùng động từ phụ và trường hợp dùng đến:

1	2	3	4	5
I		often	eat	apples
Je	mange	souvent		des pommes
I	have	often	eaten	apples
Je	ai	souvent	mangé	des pommes

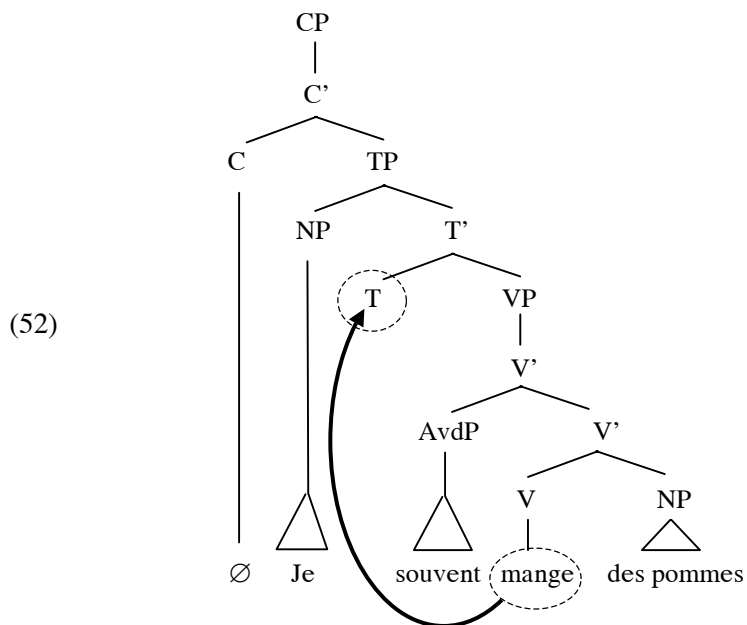
Muốn cây cú pháp trên cũng dùng được cho tiếng Pháp, bắt buộc phải thành lập một quy tắc mới. Quy tắc này có công dụng chuyển động từ “*mange*” ra khỏi ngữ đoạn VP và đưa nó vào vị trí T. Đó là ý nghĩa của quy tắc biến hình $V \rightarrow T$ (tức **verb raising**, tạm gọi là **lên cấp động từ**). Định nghĩa:

Quy tắc lên cấp động từ (verb raising) $V \rightarrow T$ là quy tắc đưa đầu động từ V lên đầu thì T.

Cấu trúc D của câu “*Je mange souvent des pommes*” có hình dạng:



Chuyển vị động từ “*mange*” lên vị trí T, cấu trúc D sẽ đổi thành cấu trúc S:

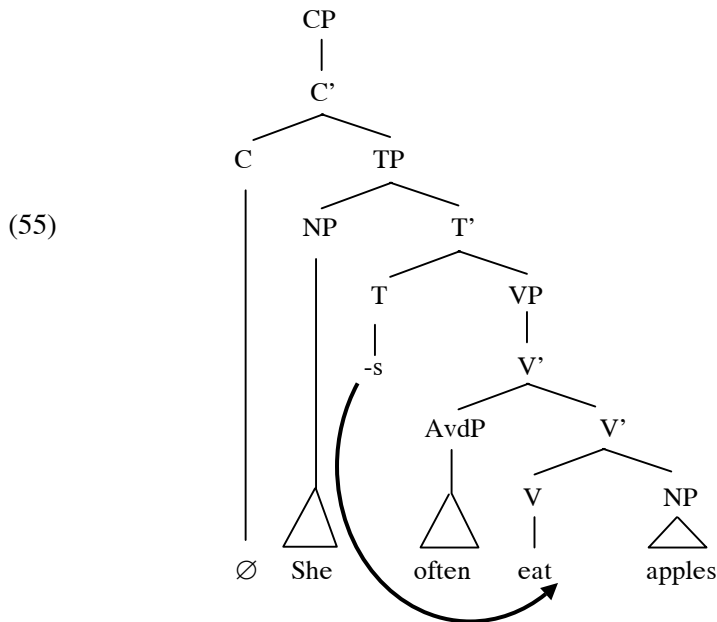


Cần để ý, cấu trúc D không phải là câu đúng ngữ pháp của tiếng Pháp mà là của tiếng Anh (hoặc tiếng Việt). Nó phải sinh ra cấu trúc S đúng với câu tiếng Pháp trong thực tế:

(53) Je mange_i souvent t_i des pommes.

Trong cấu trúc trên, t_i được coi như một *dấu vết (trace)* nằm ở vị trí động từ của cấu trúc D. Cả tiếng Anh cũng cần yếu tố này khi muốn giải thích hình vị động từ ở ngôi ba số ít. Diễn hình là trường hợp thêm *s* sau động từ:

(54) She *often* eats apples.



Ngay đây, ta cần một động tác chuyển vị từ trên xuống dưới: $T \rightarrow V$. Tạm gọi là hình thức ***xuống cấp động từ (verb lowering)***.

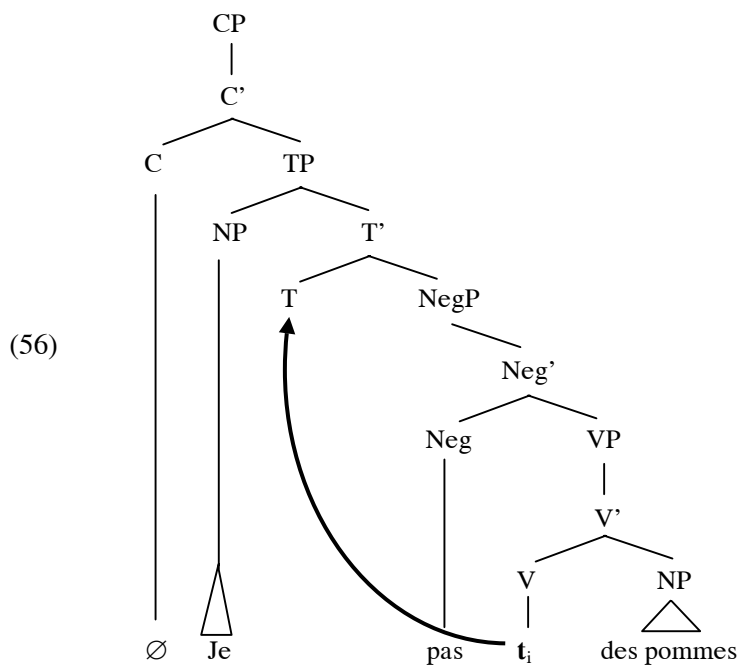
Đó là trường hợp câu xác định, kế tiếp, thử xem trường hợp câu phủ định.

Lập bảng giống như trên, ta có:

1	2	3	4	5
I	do	not	eat	apples
Je	ne-mange	pas		des pommes
I	have	not	eaten	apples
Je	n'ai	pas	mangé	des pommes

Trước hết, tạm thời đừng để ý đến hình vị “*ne*” trong tiếng Pháp mà hãy tập trung vào những vị trí phủ định “*pas*”, “*not*”. Ta thấy trường hợp này na ná như câu có trạng từ. Tất cả các động từ phụ của tiếng Anh và tiếng Pháp đều đứng trước vị trí phủ định. Hãy quan sát sự chuyển vị xung quanh động từ trong cây cú pháp dưới đây.

Giả sử “*not*” là đầu ngữ đoạn của một phép chiếu đặc biệt, tạm gọi là **phép chiếu phủ định (negation projection)**, ký hiệu là **NegP** (Pollock 1989). Phép chiếu này làm bổ ngữ cho ngữ đoạn thì TP và chi phối ngữ đoạn động từ VP.



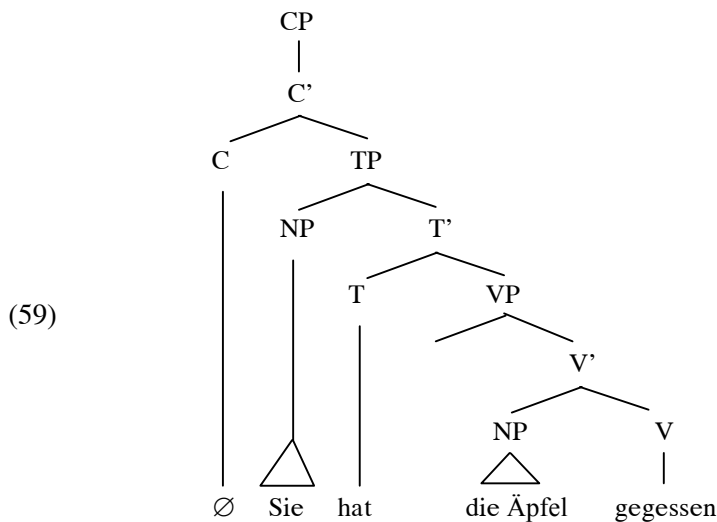
Mũi tên trong cây chỉ hướng chuyển vị động từ xoay quanh “*pas*”. Bằng cách này, người ta có thể giải thích được nhiều ngôn ngữ, thế nhưng không phải là tất cả. Xem thêm một ngôn ngữ khác, chẳng hạn tiếng Đức:

- (57) Sie ißt die Äpfel.
 (She) (eats) (apples)
 She eats apples.
 Cô ta ăn táo.

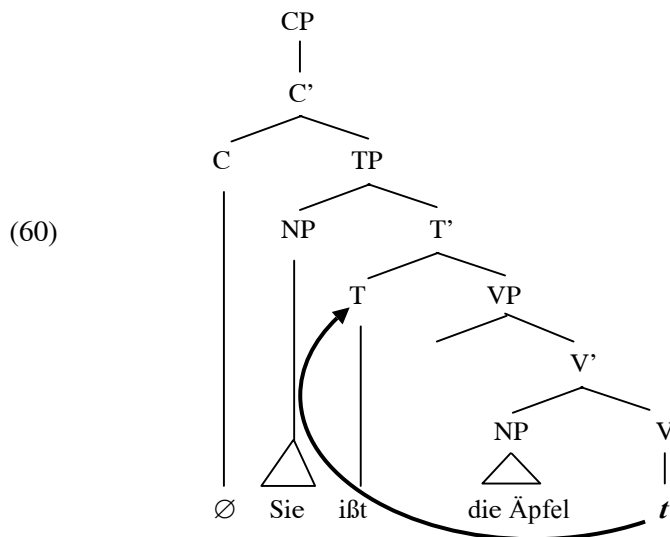
Trong trường hợp câu chỉ có động từ chính, cấu trúc câu tiếng Đức sẽ giống tiếng Anh (hoặc tiếng Việt ngay ví dụ này). “*ißt*” xuất phát từ động từ nguyên mẫu “*essen*”, được chia cho đại danh từ nhân vật ngôi 3 số ít. “*Essen*” là một động từ bất quy tắc. Nếu trong câu có động từ phụ, động từ chính sẽ không nằm sau động từ phụ mà nằm ở cuối câu. Ví dụ ngay nay: động từ phụ = “*hat*”, động từ chính = “*gegessen*”.

- (58) Sie hat die Äpfel gegessen.
 (She) (has) (apples) (eaten)
 She has eaten apples.

Cấu trúc này khác tiếng Anh hoặc tiếng Việt:



Ở một loại thì quá khứ như trên, T không cần sự trợ giúp của động từ chính, nhưng ở trường hợp khác, chẳng hạn ở thì hiện tại, thì lại cần. Trường hợp này có thể giải thích được nhờ thao tác chuyển vị:



Tóm lại, chuyển vị động từ là hình thức chuyển vị trí động từ đối với vị trí của thì T. Có lúc chuyển lên bằng quy tắc lên cấp động từ: $V \rightarrow T$, có lúc chuyển xuống bằng quy tắc xuống cấp động từ $T \rightarrow V$.

2. Chuyển vị thì T ($T \rightarrow C$)

Ở chương trước, chúng ta đã biết qua loại *câu hỏi có/không* (yes/no question) của những ngôn ngữ chứa cấu trúc *ngịch đảo chủ ngữ/động từ phụ* (subject/aux inversion)¹⁾ (tiếng Anh, tiếng Đức, tiếng Pháp). Ví dụ:

Câu xác định:

(61) The sons looked for the treasure.

(62) Their father had died.

¹⁾ Thuật ngữ *subject/aux inversion*, theo ý chúng tôi, có vẻ hơi hạn hẹp. Dùng *subject/verb inversion* có lẽ tổng quát hơn. Bởi vì trong tiếng Đức, có những trường hợp nghịch đảo mà không nhất thiết phải là động từ phụ.

(63) Die Kinder wollte die Erbe.

(64) Der Vater ist gestorben.

Câu hỏi:

(65) Did the sons look for the treasure ?

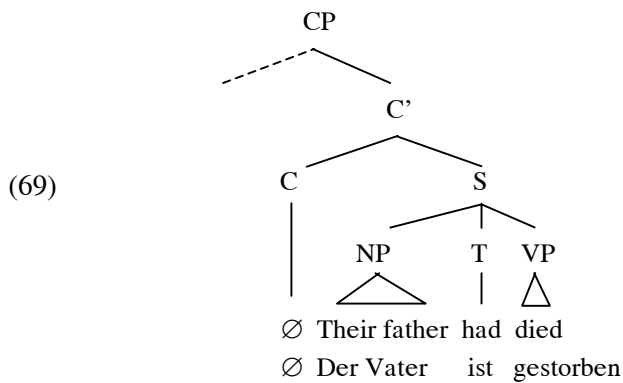
(66) Had their father died ?

(67) Wollten die Kinder die Erbe ?

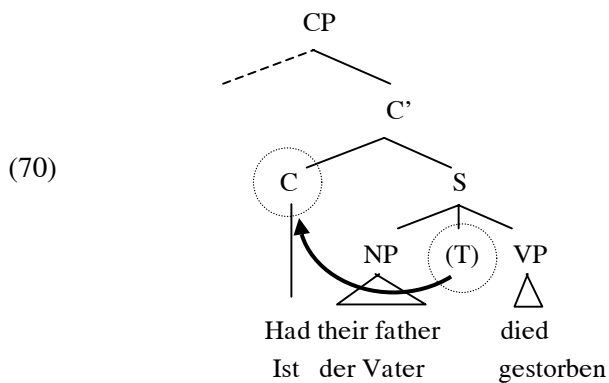
(68) Ist der Vater gestorben ?

Muốn biến câu trả lời trở thành câu hỏi, những ngôn ngữ nghịch đảo hoặc phải thêm động từ phụ ở đầu câu (trường hợp tiếng Anh) hoặc cũng có thể chuyển vị động từ chính lên đầu câu (trường hợp tiếng Đức). Đây là cách **chuyển vị thì T** (*T movement*) tới đầu một tác tử bổ ngữ hóa C: $T \rightarrow C$.

Cây cú pháp câu xác định:



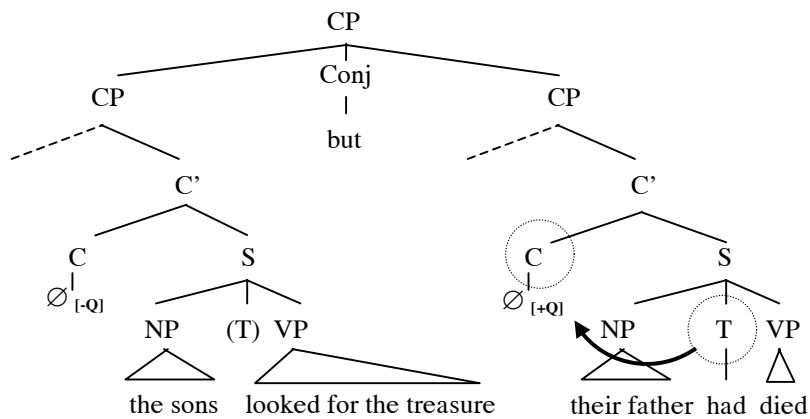
Cây cú pháp câu hỏi yes/no (ja/nein):



Bằng cách chuyển vị động từ phụ “*had*”, “*ist*” từ vị trí T lên vị trí C, câu xác định biến thành câu hỏi. Việc giải thích cấu trúc câu hỏi tương đối dễ dàng. Trong cấu trúc câu hỏi, vị trí T trống. Điều này chấp nhận được, bởi T là một lựa chọn, không bắt buộc phải có.

Không giống như tiếng Việt, tiếng Ba Lan, Irish, ... khi lập câu hỏi, người ta chỉ cần thêm một từ đặc biệt như “*chuta*”, “*phải không*”, “*Czy*”, “*An*”, v.v., tiếng Anh, tiếng Đức cần gắn thêm tính chất câu hỏi [+Q]²⁾ vào vị trí tác tử bổ ngữ hóa C. Trong trường hợp câu xác định, C có trị zero, tức một tác tử bổ ngữ hóa trống $\emptyset_{[-Q]}$. Ví dụ:

(71) The sons looked for the treasure but had their father died ?



Cùng một cấu trúc $C' \rightarrow C S$, tác tử bổ ngữ hóa C vừa mang tính chất của một câu xác định [-Q], vừa mang tính chất của một câu hỏi [+Q]. Ở phần bên trái của cây, C có trị trống $\emptyset_{[-Q]}$, trong khi đó phần bên phải chứa C có trị trống $\emptyset_{[+Q]}$. Nếu chuyển vị “*had*” vào vị trí C, mệnh đề “*their father had died*” sẽ trở thành câu hỏi: “*had their father died ?*”.

Tựu trung, những ngôn ngữ có cấu trúc *ngịch đảo chủ ngữ/động từ phụ* đều cần quy tắc chuyển vị đầu tới đầu. Ngay trường hợp này, ta có phép chuyển vị từ đầu T tới đầu C ($T \rightarrow C$). Cần lưu ý, đối với câu hỏi, trong tiếng Anh, chỉ có động từ phụ mới chiếm chỗ đầu ngữ đoạn T. Động từ chính không được. Nói cách khác là chỉ có động từ phụ mới áp dụng được quy tắc chuyển vị $T \rightarrow C$, trong khi đó,

²⁾ Q viết tắt của Question. +Q có nghĩa có tính chất câu hỏi, ngược lại là -Q.

ví dụ như tiếng Đức, động từ chính lại cần quy tắc chuyển vị $T \rightarrow C$. Khi đặt câu hỏi, nếu trong câu không có động từ phụ, động từ chính phải được chuyển lên đầu câu:

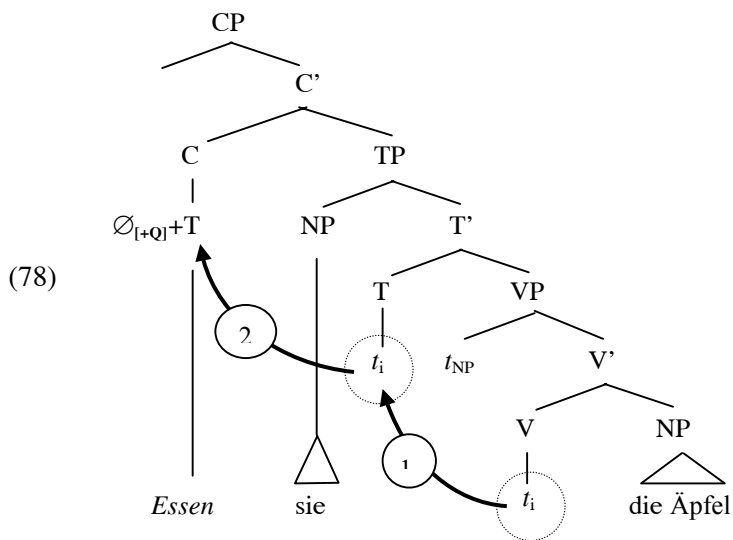
Câu xác định:

- (72) Sie essen die Äpfel. (tiếng Đức)
 (73) They eat the apples. (tiếng Anh)
 (74) Họ ăn táo.

Câu hỏi:

- (75) Essen sie die Äpfel ? (tiếng Đức)
 (76) *Eat they the apples ? (tiếng Anh)
 (77) Họ ăn táo à ?

Động từ chính “essen” (ăn) của tiếng Đức trong câu xác định được đảo ngược để biến thành câu hỏi. Tiếng Anh không nói được mà phải cần động từ phụ “do”. Để thực hiện phép biến hình, tiếng Đức phải chuyển vị hai lần. Trước nhất chuyển dấu vết t_i vào vị trí T (tức $V \rightarrow T$) và sau đó vào vị trí C (tức $T \rightarrow C$) :



QUY TẮC LỒNG

Hình thức biến hình “chuyển vị đầu tới đầu” có thể giải thích nhiều ngôn ngữ. Ngoài hình thức này còn có *quy tắc lồng* (*insertion rules*). Xin lưu ý, quy tắc lồng ở đây khác với quy tắc lồng bình thường. Ý nghĩa quy tắc lồng bình thường là lồng từng từ cho đúng vị trí nút tận cùng (terminal node) trong cây cú pháp. Quy tắc lồng ở đây có nghĩa là lồng thêm một phần tử “từ”, ví dụ phần tử động từ phụ “do” trong câu hỏi tiếng Anh.

Ở thể xác định, động từ “do” là một *động từ chính* (*main verb*) bình thường:

- (79) The boy does the homework. (tiếng Anh)
 (80) Der Junge macht die Hausaufgabe. (tiếng Đức)
 (81) Thằng bé làm bài tập nhà.

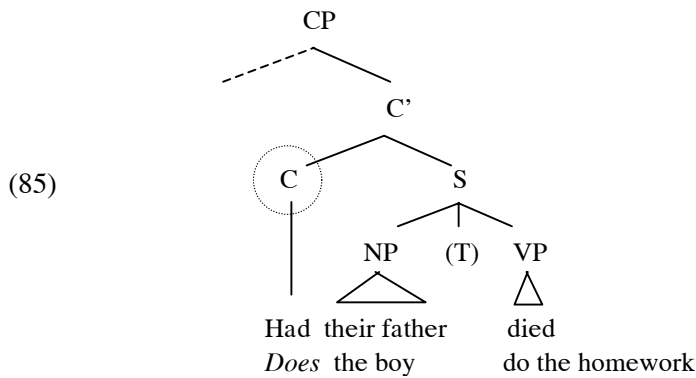
Khi hỏi, người Đức sẽ đặt động từ “macht” (“làm”) (nguyên mẫu: “machen”) ở đầu câu; tiếng Anh không được:

- (82) Macht der Junge die Hausaufgabe ? (tiếng Đức)
 (83) * Does the boy the homework ? (tiếng Anh)

Câu hỏi tiếng Anh cần lồng thêm động từ phụ “do” (“do” chia thành “does” cho nhân vật ngôi ba, số ít “the boy”):

- (84) Does the boy do the homework ? (tiếng Anh)

Giống trường hợp dùng động từ phụ “have” cho thì quá khứ, trong câu “Had their father died ? ” bên trên, động từ “do” được coi như một *hư từ* (*dummy word*) hao hao như “have” và được lồng vào vị trí C:



Nếu không xét đến hình vị động từ, thì chuỗi từ [*the boy **do** the homework*] trong cây cú pháp tương đương với câu xác định trong tiếng Việt “*Thằng bé làm bài tập nhà*”. So với câu hỏi tiếng Anh “*Does the boy do the homework*”, có lẽ câu hỏi sau đây của tiếng Việt:

(86) *Có phải* thằng bé làm bài tập nhà *không* ?

gần giống nhất. Ở đây, “*có phải*” na ná như “*does*”. Chỉ có khác là trong câu hỏi trên của tiếng Việt còn có thêm “*không*” ở cuối câu. Ngoài ra, tiếng Việt còn có nhiều cách hỏi khác nhau:

(87) Thằng bé *có* làm bài tập nhà *không* ?

(88) Thằng bé *có* làm bài tập nhà *hay không* ?

(89) Thằng bé làm bài tập nhà *phải không* ?

(90) Thằng bé làm bài tập nhà, *chứ gì* ?

(91) Thằng bé làm bài tập nhà *à* ?

(92) Thằng bé làm bài tập nhà *ư* ?

(93) Thằng bé làm bài tập nhà *hả* ?

(94) Thằng bé làm bài tập nhà *hở* ?

... v.v.

Sơ lược về quy tắc lồng là vậy. Nay chúng ta sẽ xem tiếp một loại biến hình khác: **chuyển vị NP (NP movement)**.

CHUYỂN VỊ NP

Chuyển vị NP là một loại chuyển đổi vị trí của ngữ đoạn danh từ NP và biệt định ngữ CP (hãy xem như NP cho gọn).

Không giống như loại chuyển vị đầu tới đầu, chuyển vị NP không nhằm mục đích chuyển đổi thứ tự của từ sao cho thích hợp với các quy tắc cấu trúc ngữ đoạn mà là nhằm giải quyết trường hợp một ngữ đoạn danh từ NP nằm ở vị trí không thích hợp đối với vai trò theta, nghĩa là một ngữ đoạn danh từ NP không thỏa *chủ đề (theme)*. Thử xem một ví dụ với động từ “*sing*” của tiếng Anh.

Động từ “*sing*” cần một tham số bắt buộc, đó là tác nhân (actor, agent), tức người gây ra hành động “*sing*” (hát).

(95) Mary sings.

(96) Mary sings a song.

Vai trò tác nhân phải nằm trong mệnh đề chứa “sing” (nếu ngoài là sai ngữ pháp):

(97) * I want Mary [that sings].

Nguyên tắc vai trò theta hạn chế nói rằng, một tham số nhận một vai trò theta từ vị từ (*predicate*) bắt buộc phải có tính cục bộ đối với vị từ, nghĩa là phải nằm gần vị từ đó. Nói rõ hơn:

Điều kiện về tính cục bộ (locality condition) đối với vai trò theta: Vai trò theta phải được đặt trong cùng một mệnh đề như vị từ.

Ví dụ một câu đúng ngữ pháp tiếng Anh:

(98) I want [that Mary sings].

Một câu khác:

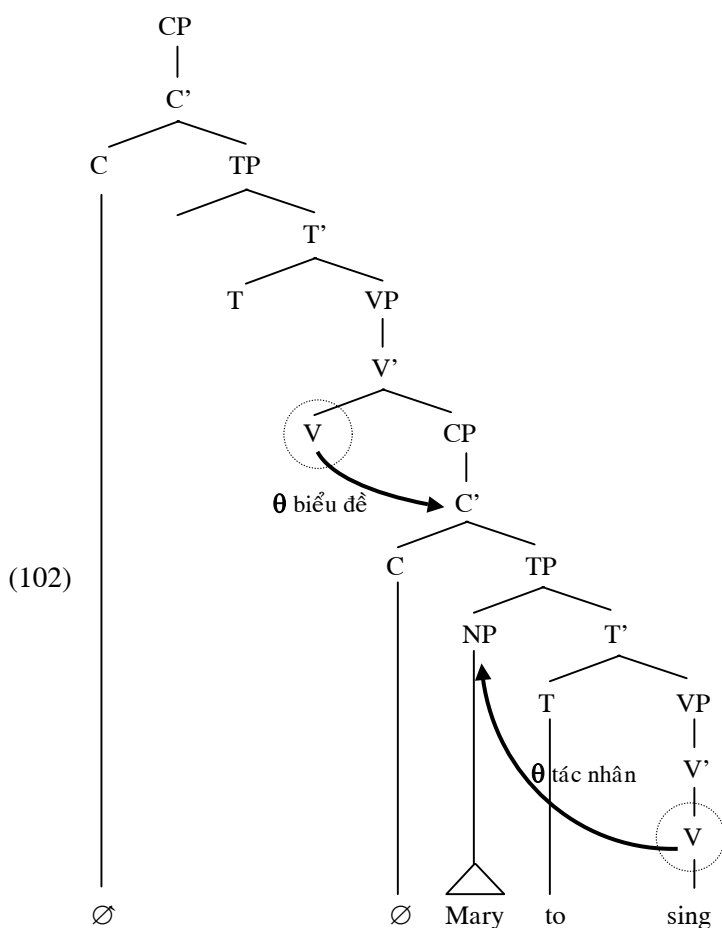
(99) Mary is likely [to sing].

Câu này nói được. Mary là tác nhân của “sing” nhưng lại nằm trong mệnh đề chính thay vì phải nằm trong mệnh đề [to sing] cho gần vị từ. So sánh cấu trúc trên với hai cấu trúc:

(100) It is likely [that Mary sings].

(101) [That Mary sings] is likely.

ta thấy, mệnh đề [that Mary sings] là một *biểu đề (proposition)* (tạm dịch) diễn tả sự việc Mary gây ra hành động. Trường hợp này khác với trường hợp “Mary is likely”, ở đó, “Mary” không nhận bất cứ vai trò theta nào của “is likely”. Trường hợp này không thỏa *điều kiện tính cục bộ đối với vai trò theta* vừa nói. Do đó công việc phải làm trước khi biến hình cấu trúc ngữ pháp là phải xác định vai trò theta. Cấu trúc D của câu “Mary is likely [to sing]” trông giống cây cú pháp dưới đây:



Sau khi đã lập xong một ngữ đoạn danh từ NP ngay vị trí chủ ngữ của mệnh đề lồng (đang giữ vai trò tác nhân), công việc kế tiếp là đưa nó về dạng cấu trúc bề mặt. Để thực hiện việc này, cần có một phép biến hình có khả năng chuyển vị ngữ đoạn danh từ NP (Mary) lên bình diện *biệt định ngữ* của mệnh đề chính. Phép biến hình này là phép *chuyển vị NP* (*NP movement*). Đây là hình thức chuyển vị lên cấp NP, tức NP sẽ được chuyển từ vị trí thấp lên vị trí cao hơn. Nhãn t_i trong cấu trúc D là dấu vết để lại:

Chuyển vị NP là một hình thức làm rõ vai trò theta (tác nhân, thụ nhân, ...) trong một câu tưởng chừng sai ngữ pháp. Ở nhiều ngôn ngữ, người ta thường dùng **cách (case)** để phân biệt vai trò theta:

- **Danh cách (nominative)** đóng vai chủ ngữ.
- **Đối cách (accusative)** đóng vai bổ ngữ trực tiếp.
- **Tặng cách (dative)** đóng vai bổ ngữ gián tiếp.

Hình vị từ có thể diễn tả *cách*. Cách mô tả mối quan hệ ngữ pháp của các đối tượng đang giữ từng vai trò trong câu. Chẳng hạn trong tiếng Nhật, người ta đánh dấu “-ga” cho danh cách, “-o” cho đối cách, “-ni” cho tặng cách:

- (106) Hanakata-ga ronbun-o kai-ta.
 (Hanakata) (đề tài) (đã viết)
 (danh cách) (đối cách)
- (107) Hanakata-ga heya-ni haitte-kita.
 (Hanakata) (phòng) (đã đi vào)
 (danh cách) (tặng cách)

Cách trong tiếng Anh không được đánh dấu như trong tiếng Nhật mà chỉ ở dạng đại danh từ (không kể câu hỏi):

- **Danh cách:** I, you, he, she, it, we, you, they
- **Đối cách:** me, you, him, her, it, us, you, them

Tiếng Đức tương tự:

- **Danh cách:** ich, du, er, sie, es, wir, ihr, sie
- **Đối cách:** mich, dich, ihn, sie, es, uns, euch, sie

Tiếng Đức còn có thêm:

- **Tặng cách:** mir, dir, ihm, ihr, ihm, uns, euch, ihnen

Cách của tiếng Đức không những chỉ được biểu thị xuyên qua đại danh từ mà còn qua mạo từ, danh từ, tính từ, sở hữu từ, giới từ, ... rất phức tạp. Tiếng Đức là một ngôn ngữ giàu hình vị. Trái lại, có ngôn ngữ nghèo hình vị, tiếng Anh chẳng hạn. Tất nhiên cũng có ngôn ngữ không có hình vị.

Không đặt nặng về hình vị, các nhà ngôn ngữ để ý đến ***cách trừu tượng (abstract case)*** của những ngữ đoạn danh từ NP. Theo Chomsky (1981), một ngữ đoạn danh từ NP có thể mang tính chất của một trong những cách trừu tượng sau đây:

- Danh cách (nominative, viết tắt NOM): Cách này diễn tả một biệt định ngữ của thì *T hữu hạn (finite T)*.
- Đối cách (accusative, viết tắt ACC): Cách này là chị em của một động từ xuyên tính.
- Tặng cách (dative, viết tắt DAT): Cách này gắn với một giới từ.

Việc đầu tiên phải làm trước khi chuyển vị một ngữ đoạn danh từ NP là lọc lựa cách. Nguyên tắc:

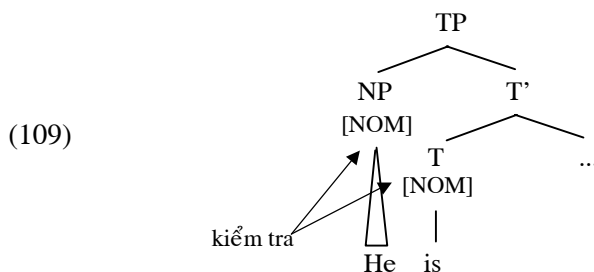
Lọc cách (Case Filter): Tất cả ngữ đoạn danh từ NP đều phải được ghi chú bằng một dấu hiệu chỉ *cách (case)*. Còn không, không thể phát sinh.

Một ví dụ tiếng Anh:

(108) She

$$\left(\begin{array}{c} \text{feminine} \\ 3^{\text{rd}} \text{ person} \\ \text{singular} \\ \text{nominative} \end{array} \right)$$

Ma trận trên chứa các tính chất của “*she*”. Đặc biệt, “*she*” là một đại danh từ chỉ được dùng cho danh cách. Nguyên tắc lọc cách chỉ cho phép một danh cách đi với một động từ ở dạng danh cách tương ứng, ví dụ “*is*”. Ngoài ra, như đã biết, danh cách diễn tả một biệt định ngữ của thì *T hữu hạn (finite T)*, và lọc cách có nhiệm vụ *kiểm tra* đặc tính này. Ví dụ trong cây cú pháp dưới đây, “*she*” và “*is*” được ghi chú bằng dấu hiệu [NOM] (nominative):

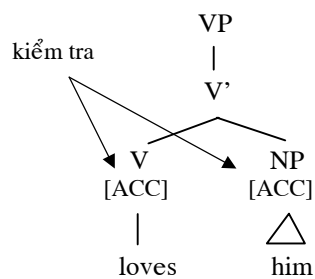


Trường hợp đối cách tương tự:

(110) Him

$\left(\begin{array}{l} \text{masculine} \\ 3^{\text{rd}} \text{ person} \\ \text{singular} \\ \text{accusative} \end{array} \right)$

(111) Mary loves him.
[ACC] [ACC]



Cây cú pháp cho thấy đối cách tồn tại trong trường hợp một ngữ đoạn danh từ NP đóng vai chị em của động từ V. “Him” là chị em của “loves”.

Xem lại các câu ví dụ của tiếng Anh chứa mệnh đề với thì T hữu hạn:

(112) It is likely [that Mary sings].

(113) [That Mary sings] is likely.

ta thấy các ngữ đoạn danh từ NP đều là danh cách và là biệt định ngữ, trong khi đó, thì phi hạn T (*non-finite T*), nghĩa là chứa “to”, như:

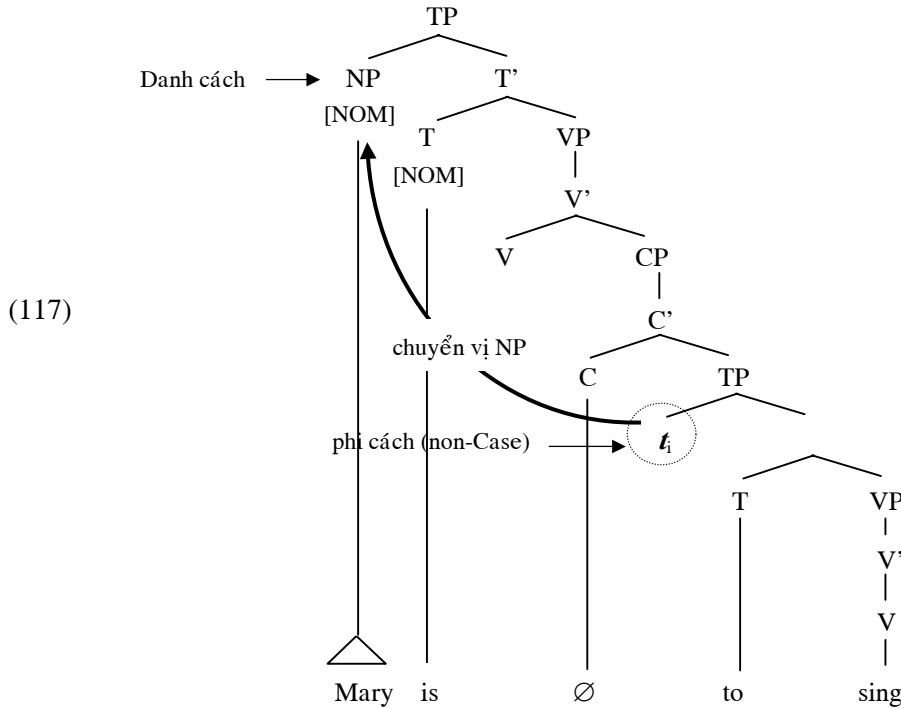
(114) * It is likely Mary to sing.

(115) * Mary to sing is likely.

không chứa tính danh cách [NOM]. Nguyên tắc lọc cách bị tổn thương là sự cắt nghĩa, tại sao câu không nói được. Nhưng câu dưới đây lại nói được:

(116) Mary is likely to sing.

Trong câu này, mệnh đề chính chứa T hữu hạn. Ngữ đoạn danh từ NP “*Mary*” sẽ được chuyển tới vị trí biệt định ngữ để nhận vai trò danh cách:



Quan sát kỹ, ta sẽ thấy, từ một vị trí không có cách (phi cách), một ngữ đoạn danh từ NP có thể được chuyển vị lên vị trí có cách. Mặc dù vị trí ấy không có cách nhưng vẫn chứa vai trò theta.

Ngoài trường hợp câu chủ động (active) vừa nêu, cả trường hợp câu bị động (passive) cũng giải thích được (có thể xem chi tiết ở phần *Generalization* của Burzio (1986)). Ngoài ra cũng cần lưu ý, cách nhận xét như trên về tặng cách chỉ tương đối đúng. Không như tiếng Anh, tặng cách trong tiếng Đức không bắt buộc phải đi với giới từ.

Câu có tặng cách đi sau giới từ:

- (118) Ich habe das Buch von Dir bekommen.
 (Tôi) (đã) (cuốn sách) (từ/của) (anh) ((đã) nhận)
 [ACC] [prep] [DAT]

Câu có tặng cách không có giới từ:

- (119) Gib mir das Buch.
 (đưa) (tôi) (cuốn sách)
 [DAT]

Vị trí cách trừu tượng của tiếng Anh thường theo thứ tự từ trái sang phải: danh cách + đối cách. Tiếng Đức, tiếng Việt tương tự, song cũng có trường hợp đảo ngược: đối cách + danh cách.

- (120) Ich lese das Buch.
 (Tôi) (đọc) (cuốn sách)
 [NOM] [ACC]

- (121) Tôi đọc cuốn sách.
 [NOM] [ACC]

Cùng nội dung ấy, người Đức có thể nói:

- (122) Dieses Buch habe ich schon gelesen.
 (cuốn sách này) (đã) (tôi) (rồi) ((đã)đọc)
 [ACC] [NOM]

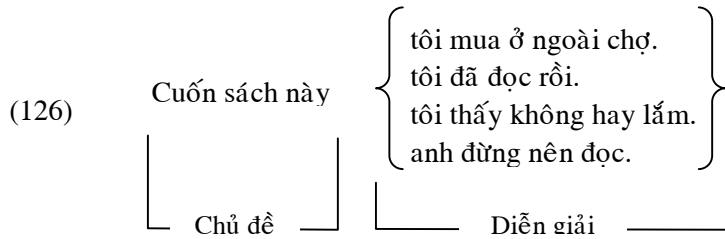
Người Việt cũng thường hay nói:

- (123) Cuốn sách này tôi đã đọc rồi.
 [ACC] [NOM]

Ở cách nói trên, “*cuốn sách này*” được hiểu là đang đóng vai chủ đề. Thử so sánh hai câu tiếng Anh và tiếng Việt:

- (124) I read the book.
 (125) Tôi đọc cuốn sách.

“*The book*” và “*cuốn sách*” là cái đang được nói đến, tức *chủ đề (theme)*. Nếu như người Việt nói “*Tôi đọc cuốn sách*”, thì đây là một cách diễn tả rất bình thường, nghĩa là một nhân vật đang làm gì đó. Cách lập câu này không khác với các ngôn ngữ khác. Song khi nói “*Cuốn sách này, tôi đã đọc rồi*” thì có nghĩa chủ đề “*cuốn sách này*” đi trước rồi sẽ được diễn giải, bổ sung sau, ví dụ:



Cấu trúc *đề diễn* (*chủ đề – diễn giải, theme – rheme*³⁾) là một cấu trúc ngôn ngữ mà người Việt thường dùng.

Cách nói với thành tố được đặt phía trước cũng thường thấy trong tiếng Đức. Đó cũng là một hình thức nhấn mạnh:

- (127) *Von ihm* will ich nicht mehr hören.
(từ) (hắn) (muốn) (tôi) (không) (nữa) (nghe)

Hiểu theo tiếng Việt đại để là “*tôi không muốn nghe nói gì về hắn nữa*”. Cụm chữ “*von ihm*” nhấn mạnh nhân vật “*hắn*”. Ngay đây, “*von ihm*” không phải là ngữ đoạn danh từ NP mà là một ngữ đoạn giới từ PP (có quan hệ lệ thuộc với động từ “*hören*”). Câu không đảo nghịch có dạng:

- (128) Ich will nicht mehr *von ihm* hören.

Một ví dụ tương tự:

- (129) Ich mag *diese Musik* nicht so sehr.
(tôi) (thích) (nhạc này) (không) (lắm)

Hiểu theo tiếng Việt là “*Tôi không thích nhạc này lắm*”. Thay vì nói xuôi, người Đức cũng thường nói ngược bằng cách đẩy cụm từ “*Diese Musik*” lên phía trước:

- (130) *Diese Musik* mag ich nicht so sehr.
(nhạc này) (thích) (tôi) (không) (lắm)

Câu trên có trật tự na ná như người Việt thường nói:

- (131) Nhạc này tôi không thích lắm.

³⁾ cách gọi của Halliday 1994. Gọi interpretation (diễn giải) có lẽ tổng quát hơn.

“*Diese Musik*” và “*nhạc này*” là hai ngữ đoạn danh từ NP tương đương và là danh cách nằm ở đầu câu. Cách lập câu vừa làm rõ tính chất chủ đề vừa cho thấy cấu trúc đề diễn với

- phần đề là “Diese Musik” và “nhạc này”,
- phần diễn là “mag ich nicht so sehr”, “tôi không thích lắm”

Giả sử đó là cách đề diễn trong tiếng Đức, thì ranh giới giữa hai phần có thể nhận diện được nhờ vị trí động từ. Động từ đứng sau chủ ngữ trong một câu xuôi sẽ được đưa lên phía trước chủ ngữ trong câu đảo:

Câu xuôi:

(132) Ich mag diese Musik nicht so sehr.
(chủ ngữ) (động từ) (bổ ngữ)

Câu đảo:

(133) Diese Musik *mag* ich nicht so sehr.
 (bổ ngữ) (động từ) (chủ ngữ)
 (phần đề) (----- phần diễn -----)

Bình thường trong cấu trúc đề diễn của tiếng Việt, người ta thường dùng “*thì*”, “*là*”, “*mà*” để phân ranh giữa phần đề và phần diễn:

(134) Nhạc này thì tôi không thích lắm.
(phần đề) (----- phần diễn -----)

Không dùng “*thì*”, câu nói có vẻ gọn gàng, văn vẻ hơn, song biên giới giữa phần đề và phần diễn sẽ trở nên lu mờ, hoặc có khi còn dẫn đến hiện tượng đa nghĩa ngoài ý muốn:

(135) Nhà tôi ở không được sạch lắm.

Nếu hiểu “*nhà tôi*” theo tinh thần “*cái nhà mà tôi ở*”:

(136) Nhà tôi ở không được sạch lắm.
(cái nhà mà) (tôi ở ...)

thì không sao, nhưng hiểu “*nhà tôi*” như “*ông/bà xã tôi*” thì hơi nguy. Do đó, muốn bỏ “*thì*”, “*là*”, “*mà*”, thiết nghĩ cũng nên thêm cái dấu phẩy vào cho chắc. Có hai trường hợp:

(137) Nhà, tôi ở không được sạch lắm.

(138) Nhà tôi ở không được sạch lắm.

Xét cho cùng, dấu phẩy cũng không phải là một giải pháp tuyệt đối. Viết còn phẩy được, nói thì không thể. Cấu trúc mơ hồ thường là một trở ngại lớn trong công đoạn phân tích cú pháp của hệ nhận diện ngôn ngữ mà chắc hẳn người nghiên cứu ngôn ngữ nào cũng đã từng gặp phải. Dĩ nhiên còn nhiều vấn đề khác, nhưng theo kinh nghiệm chung, máy không thể nhận diện tiếng Việt dễ dàng như tiếng Anh, Đức cũng bởi vì cấu trúc tiếng Việt khá đa dạng và hay chứa những trường hợp không sắc bén hoặc chứa nhiều yếu tố/thành tố zero.

CHUYỂN VỊ - WH

Như đã biết, trong các ngôn ngữ nghịch đảo (Anh, Pháp, Đức, ...), cấu trúc câu xác định ngược hẳn câu hỏi. Giới ngôn ngữ tạo sinh giải thích điểm này bằng sáng kiến *chuyển vị-wh*. Chuyển vị-wh chẳng qua là một trong những hình thức của quy tắc biến hình. Câu hỏi có/không là một trường hợp tiêu biểu.

Ví dụ ứng với một câu xác định của tiếng Anh, tiếng Đức:

(139) Kiều had died.

(140) Kiều ist gestorben.

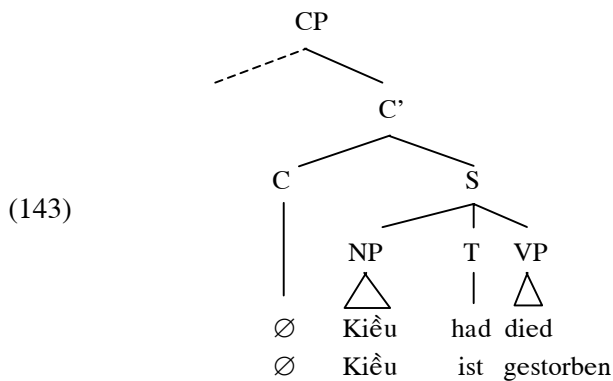
(Kiều đã chết)

Thì câu hỏi sẽ là:

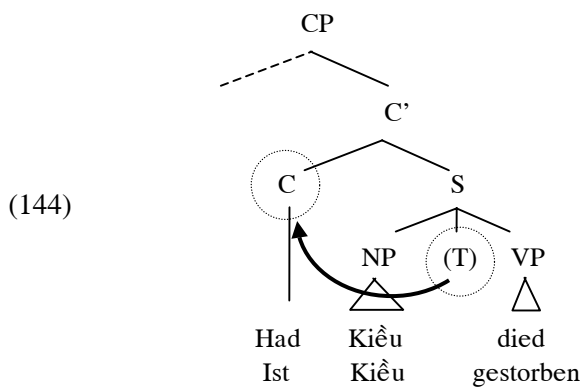
(141) Had Kiều died ?

(142) Ist Kiều gestorben ?

Cây cú pháp câu xác định:



Cây cú pháp câu hỏi yes/no (ja/nein):



Bằng cách chuyển vị động từ phụ “*had*”, “*ist*” từ vị trí T lên vị trí C, câu xác định biến thành câu hỏi.

Tương tự câu hỏi có/không, các loại câu hỏi khác cũng có cấu trúc ngược với câu xác định. Cụ thể, với “*what*”, “*who*”, “*where*”, ... (tiếng Anh), “*was*”, “*wer*”, “*wo*”, ... (tiếng Đức):

(145) Who was Nguyễn Du ?

(146) Wer war Nguyễn Du ?

(147) Nguyễn Du là ai ?

Giống câu hỏi có/không, ở loại câu hỏi wh, cấu trúc câu nghi vấn của tiếng Việt cũng không ngược với cấu trúc câu xác định. Nếu thay thế “*ai*” bằng một giá trị

hữu lý, ví dụ như “*một đại thi hào*”, ta sẽ được câu trả lời. Cấu trúc vẫn giữ nguyên. Đây là trường hợp câu hỏi hồi ngôn mà tiếng Anh cũng có.

Câu hỏi hồi ngôn thường chỉ theo sau một câu vừa được nói ra bằng cách sử dụng lại cấu trúc của câu nói ra ấy. Có trường hợp câu hỏi hồi ngôn đi từ một câu xác định hoặc mệnh lệnh:

- (148) A: I bought a car.
 (149) B: You bought a car ?
 (150) A: Don't touch my hand.
 (151) B: Don't touch your what ?

có trường hợp câu hỏi hồi ngôn đi từ một câu hỏi:

- (152) A: Did you buy a house ?
 (153) B: Did I what ?

nhưng cũng có trường hợp, người Anh không thể đổi một câu hỏi hồi ngôn thành câu trả lời bằng cách đưa một giá trị xác định vào vị trí nghi vấn-wh. Ví dụ không thể thế “*whom*” trong câu nghi vấn:

- (154) **Whom** did Kiều love ?

bằng “*Kim Trọng*” để được câu xác định:

- (155) ***Kim Trọng** did Kiều love.

theo tinh thần người Việt khi hỏi:

- (156) Kiều thương **ai** ?

và được trả lời là:

- (157) Kiều thương **Kim Trọng**.

Đó là trường hợp hồi ngôn, nhưng nếu là một câu hỏi-wh thực sự, nghĩa là từ nghi vấn-wh nằm đầu câu, thì ở câu hỏi của tiếng Anh, từ nghi vấn “*whom*” lại nằm xa “*love*”, hoặc ở tiếng Đức, “*wen*” lại nằm trước “*liebte*” (Wen liebte Kiều ?).

Muốn được một câu xác định có nghĩa tương ứng với cấu trúc không bị đổi như tiếng Việt, người Anh, người Đức không thể thế “*whom*”, “*wen*” bằng “*Kim Trọng*” một cách dễ dàng:

(158) * Kim Trọng did Kiều love.

(159) Kim Trọng liebte Kiều.

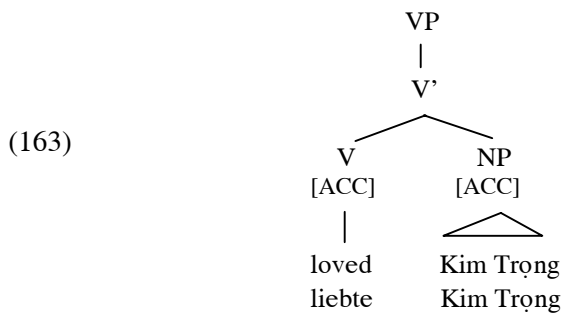
(160) Kiều đã thương Kim Trọng.

Ta thấy, câu tiếng Việt nói được, câu tiếng Anh không nói được, còn câu tiếng Đức lại có nghĩa khác hoàn toàn: “*Kim Trọng đã thương Kiều*” chứ không phải “*Kiều đã thương Kim Trọng*”. Muốn đúng, bắt buộc phải đảo ngược cấu trúc câu tiếng Anh và tiếng Đức.

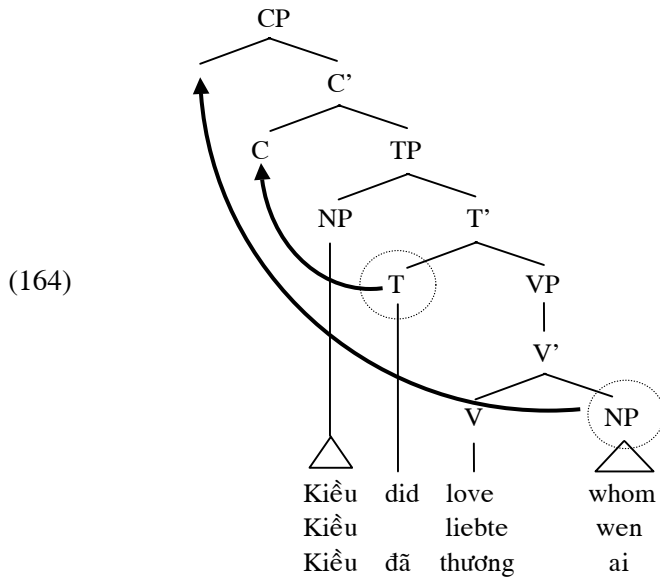
(161) Kiều loved Kim Trọng.

(162) Kiều liebte Kim Trọng.

Trong câu trả lời, “*Kim Trọng*” là đối cách. Xin nhắc lại, đối cách chỉ tồn tại ở cấu trúc một ngữ đoạn danh từ NP là chị em của một động từ V:



Nhưng ở câu hỏi-wh, đối cách của tiếng Anh và tiếng tiếng Đức không phải là chị em của động từ V, nói cách khác là không nằm đúng vị trí vai trò theta và thậm chí cũng không có cách (case) gì cả. Thử nhìn cấu trúc sâu (cấu trúc D), tức câu trả lời, dưới đây:



Để đưa cấu trúc D về cấu trúc S, tiếng Anh và tiếng Đức phải chuyển vị “whom”, “wen” lên nút biệt định ngữ CP. Không những vậy, động từ phụ diễn tả thì “did” cũng được chuyển đi. Tiếng Anh chuyển được nhưng tiếng Đức thì không vì không có động từ phụ trong trường hợp này. Tiếng Việt vẫn giữ một cấu trúc “tịnh” không thay đổi: cấu trúc D và S giống nhau.

Ở đây, tiếng Anh còn bị một điểm lấn cấn nữa là câu hỏi của mệnh đề lồng (embedded clause) không được phép chuyển vị. Đó là trường hợp [*whom Kiều has loved*] của câu sau:

(165) Nguyễn Du knew [*whom Kiều has loved*] ?

(166) * Nguyễn Du knew [*whom has Kiều loved*] ?

Dịch y hệt nội dung câu này sang tiếng Đức, ta có:

(167) Nguyễn Du wusste [*wen Kiều geliebt hat*].

Nếu câu hỏi là mệnh đề chính, nó sẽ có dạng:

(168) Wen hat Kiều geliebt ?

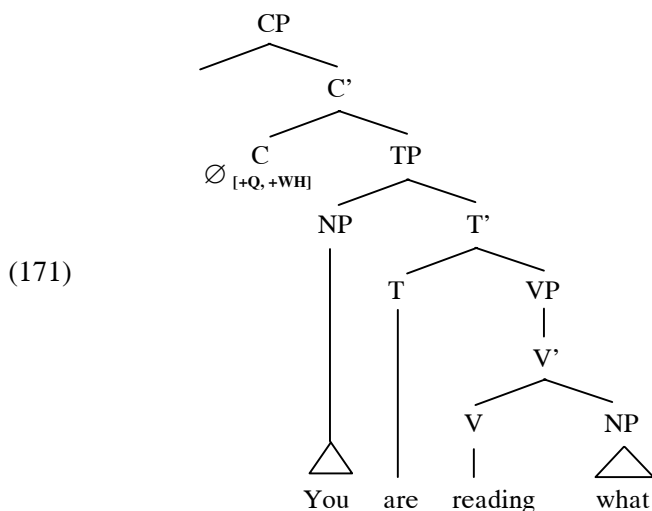
(169) Kiều đã yêu ai ?

Chúng ta sẽ nhận ra ngay, nếu câu hỏi tiếng Đức là một mệnh đề lồng, thì động từ phụ sẽ nhảy xuống cuối câu. Cụ thể, động từ phụ “*hat*” bị xuống cấp, ngược với động tác lên cấp $T \rightarrow C$ trong trường hợp câu hỏi là mệnh đề chính.

Một động từ cần có thêm đặc điểm *WH* để phân biệt giữa câu hỏi có/không và câu hỏi-wh.

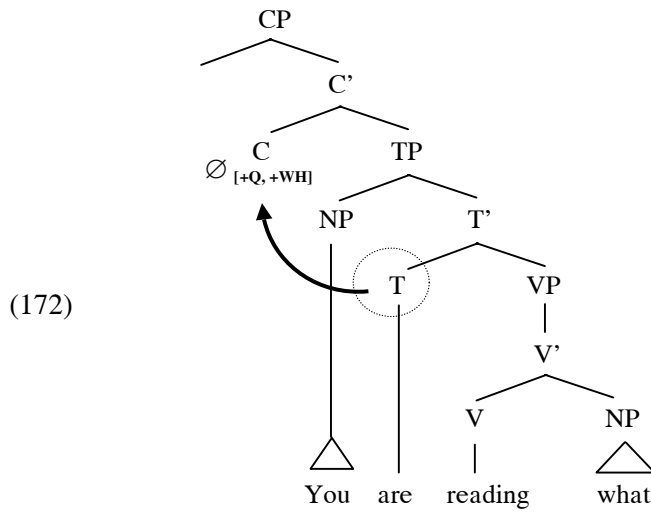
(170) What are you reading ?

Cấu trúc D có hình dạng:

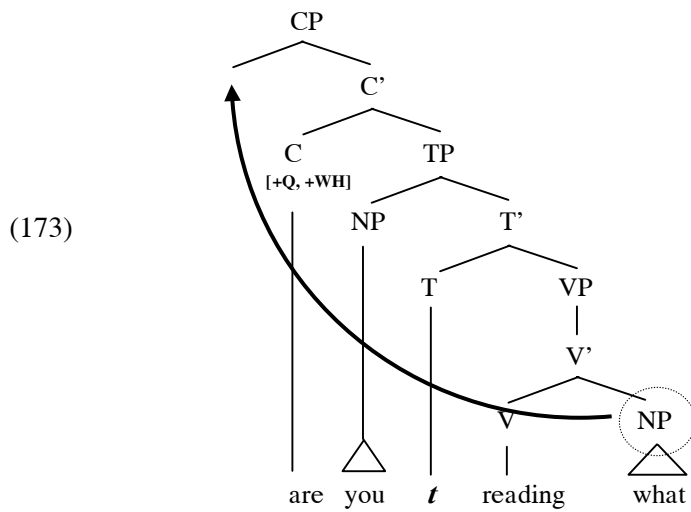


Cấu trúc D của cây cú pháp diễn tả, đây là câu hỏi được đánh dấu bằng [+Q] và là loại *tác tử bổ ngữ hóa C* chứa giá trị [+WH]. Trong câu hỏi, “*you*” và “*what*” nhận vai trò theta và cách. “*You*” là danh cách, “*what*” là đối cách. Quá trình chuyển vị sẽ trải qua hai bước, một, chuyển vị $T \rightarrow C$ để điền vào tác tử bổ ngữ hóa C mang dấu hiệu câu hỏi [+Q], hai, ứng dụng chuyển vị wh để cho biết đây là loại câu hỏi WH.

Chuyển vị $T \rightarrow C$:



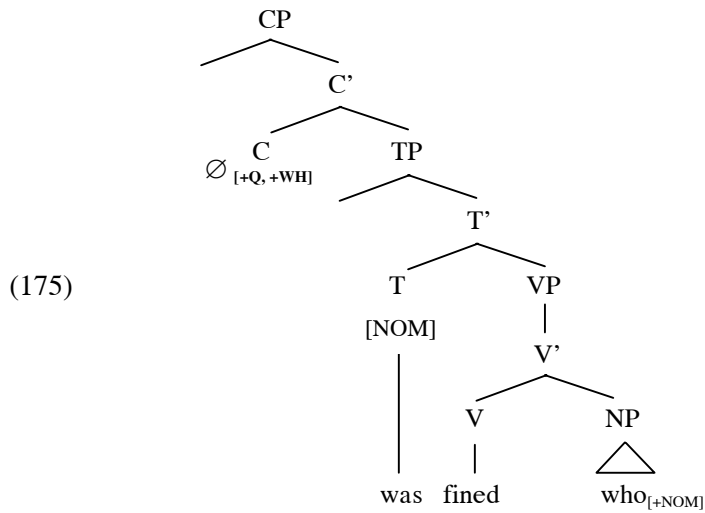
Chuyển vị wh:



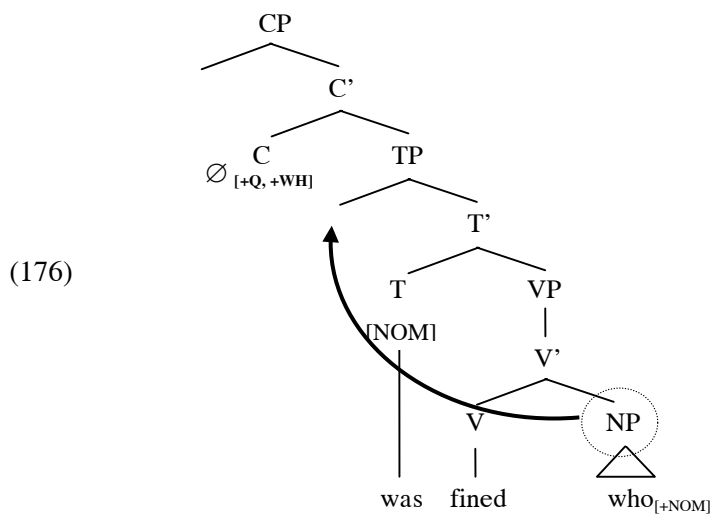
Bên trên là trường hợp câu chủ động. Sau đây là một ví dụ câu bị động:

(174) Who was fined ?

Cấu trúc D có dạng:

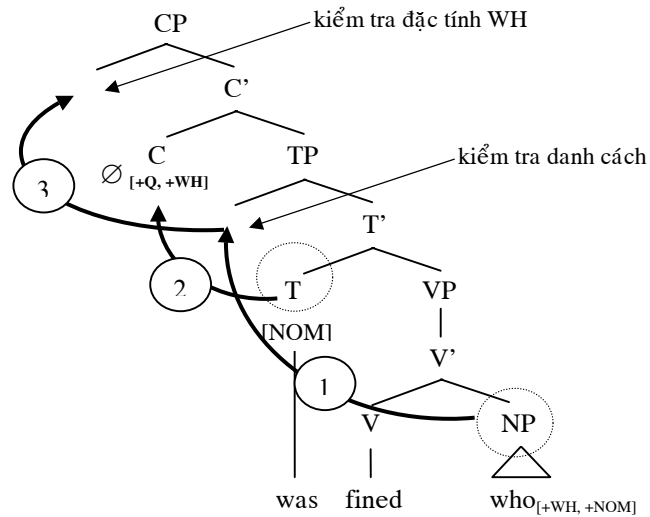


“Who” chỉ là một trong những tham số của câu và đóng vai như một bổ ngữ của động từ. Vì ở dạng bị động, nó không thể nhận cách ở vị trí này mà phải được chuyển lên phần biệt định ngữ của TP để được kiểm tra danh cách ở đó:



Sau khi kiểm tra cách và nếu nhận được danh cách, nó sẽ tiếp tục được chuyển lên biệt định ngữ của CP để được kiểm tra đặc tính WH. Động từ phụ chuyển vị với quy tắc $T \rightarrow C$ và nhận đặc tính [+Q]:

(177)



Quá trình chuyển vị gồm ba giai đoạn. Một, chuyển vị NP, hai, chuyển vị $T \rightarrow C$, và ba, chuyển vị wh. Xin lưu ý, chuyển vị wh chỉ chuyển những NP đã được kiểm tra cách.

Ngoài những hình thức chuyển vị kể trên còn có những hình thức khác, v.d. như biến hình cấu trúc *Verb-Subject-Object (VSO)* sang dạng *Subject-Verb-Object (SVO)*, tiêu biểu là tiếng Irish, hoặc chuyển vị $N \rightarrow D$ trong tiếng Hebrew, Ý (Carnie 2000, Ritter 1988).

TÓM LƯỢC

Quy tắc biến hình là một quy tắc đặc trưng của hướng ngữ pháp tạo sinh. Mục đích của quy tắc biến hình là nhằm giải thích cấu trúc ngôn ngữ trong quá biến hóa từ nhận thức-nhận diện sang cách diễn đạt thực tế. Quy tắc biến hình bao gồm nhiều hình thức chuyển vị: chuyển vị "wh", chuyển vị NP, chuyển vị đầu tới đầu (thì, động từ), ...

Chuyển vị "wh" là một cách biến hình tiêu biểu dành cho các câu hỏi "wh" thuộc nhóm ngôn ngữ nghịch đảo. Ở những ngôn ngữ này, cấu trúc câu hỏi ngược với cấu trúc câu xác định.

Ở trường hợp khác, muốn biến câu trả lời trở thành câu hỏi, những ngôn ngữ nghịch đảo hoặc phải thêm động từ phụ ở đầu câu (trường hợp tiếng Anh) hoặc cũng có thể chuyển vị động từ chính lên đầu câu (trường hợp tiếng Đức). Đó là hình thức chuyển vị đầu tới đầu.

Chuyển vị đầu tới đầu có hai trường hợp: *chuyển vị động từ* và *chuyển vị thì*. Riêng trường hợp chuyển vị động từ, có lúc động từ được chuyển lên bằng quy tắc lên cấp động từ: $V \rightarrow T$, có lúc được chuyển xuống bằng quy tắc xuống cấp động từ $T \rightarrow V$.

Chuyển vị NP là một loại chuyển đổi vị trí của ngữ đoạn danh từ NP và biệt định ngữ CP.

NHẬN XÉT VÀ ĐIỂM MỞ

Hình thức câu hỏi của tiếng Việt giống như câu hỏi hồi ngôn. Những từ nghi vấn “wh” của tiếng Việt không bị chuyển vị như câu hỏi-wh của tiếng Anh. Hãy thử xem:

- 1) Cấu trúc nào trong tiếng Việt có thể ứng dụng được quy tắc biến hình ?
- 2) Quy tắc biến hình có ứng dụng được cho các cấu trúc đề diễn không ?
- 3) Cấu trúc câu hỏi tiếng Việt thường giống cấu trúc câu trả lời, nghĩa là theo nguyên tắc tịnh (tịnh tiến). Có trường hợp nào có hai cấu trúc khác nhau không ?

PHỤ LỤC

BẢNG ĐỐI CHIẾU THUẬT NGỮ VIỆT - ANH**A**

A-gạch	A-bar
ăn khớp	matching
âm tiết	syllable
âm tố	speech sound
âm vị	phoneme
âm vị học	phonology
ẩn số	variable

B

bao gồm	involve
bảng	list
bảng mẫu trùng hợp	pattern matching list
bảng dữ liệu	database
bảng từ vựng	lexical bank
bất đối xứng	asymmetric
biến hình	transform
biến số	variable
biến thể	inflection
biệt định ngữ	specifier
biểu đạt	utterance
biểu thức	expression
biểu thức dẫn chỉ	referring expression
bình diện	level
bình diện trung gian	intermediate level
bị động	passive
bóng chiếu bình diện gạch	bar-level projection
bóng chiếu X-gạch	X-bar projection
bổ ngữ	complement
bổ thể	expletive

C

cách	case
------	------

cách trừu tượng	abstract case
cảm đề	subject/theme sensitive
cảm ngữ cảnh	context sensitive
cảm nhân	experiencer
câu	sentence
câu bị động	passive sentence
câu chẻ	cleft sentence
câu chẻ giả	pseudo-cleft sentence
câu chủ động	active sentence
câu hỏi có/ không	yes/no question
câu hỏi hồi ngôn	echo question
câu hỏi phi hồi ngôn	nonecho question
cấu trúc bề mặt	surface structure
cấu trúc Chủ từ-Động từ-Túc từ (SVO)	Subject-Verb-Object structure (SVO)
cấu trúc Chủ từ-Túc từ-Động từ (SOV)	Subject-Object-Verb structure (SOV)
cấu trúc cú pháp	syntax structure
cấu trúc cú pháp tối thiểu	minimal syntax structure
cấu trúc phụ	substructure
cấu trúc sâu	deep structure
cấu trúc-D	D-structure
cấu trúc-S	S-structure
cây	tree
cây cú pháp	syntax tree
chân lý	axiom
chỉ định ngữ	x. ngữ đoạn chỉ định từ
chỉ định từ	determiner
chỉ số	index
chỉ số phụ	coindex
chi phối	dominance, government
chi phối tận mức	exhaustive domination
chi phối trực tiếp	immediately dominate
chiếu	project
chính quy	regular
chữ cái	x. mẫu tự
chức năng	function
chuẩn hình	well-formed
chuẩn tắc theta	theta criterion

chủ	possessor
chủ đề	theme
chủ đề – diễn giải	theme - rheme
chủ ngữ	subject
chủ ngữ ngữ pháp	grammatical subject
chủ động	active
chủ thể	subject
chủ thể logic	logical subject
chủ thể ngữ pháp	grammatical subject
chủ thể tâm lý	psychological subject
chủ từ	x. chủ ngữ
chuyển vị	move, movement
chuyển vị-alpha	alpha-movement
chuyển vị lên cấp	raising movement
chuyển vị đầu tới đầu	head-to-head movement
chuyển vị T (thì T)	T-movement
chuyển vị-wh	wh-movement
cộng hiến	benefactive
cộng đồng tiếng nói	speech-community
cục bộ nghiêm ngặt	strikt local
cụm ký tự	group of characters
cú pháp	syntax
cú pháp học	syntax

D

danh cách	nominative
danh ngữ	x. ngữ đoạn danh từ
danh từ	noun
danh từ chung	common noun
danh từ hình thức	formal noun
danh từ riêng	proper noun
dẫn chỉ	refer
dấu hiệu	symbol
dấu vết	trace
diễn giải	interpretation
diễn đạt	express
diễn thức	expression
dữ liệu	data

dữ liệu mẫu	pattern
dữ liệu muốn tìm	sample
dụng cụ	instrument
dụng năng ngôn ngữ	language performance
ngữ dụng	pragmatic
dụng pháp	pragmatic
dụng pháp học	pragmatics

D

đa nghĩa	ambiguous, ambiguity
đa ngôn ngữ	multi-lingual
đầu	head
đầu từ vựng	lexical head
đầu tới đầu	head-to-head
đại danh từ	pronoun
đại danh từ chỉ thị	demonstrative pronoun
đại danh từ hỗ tương	reciprocal pronoun
đại danh từ nghi vấn	interrogative pronoun
đại danh từ nhân vật	personal pronoun
đại danh từ quan hệ	relative pronoun
đại danh từ sở hữu	possessive pronoun
đại danh từ tái quy	reflexive pronoun
đề diễn	Theme-Rheme
đi trước	precede
điều kiện	condition
điều kiện về tính cục bộ	locality condition
đối cách	accusative
đối lập	opposition
đối tượng	object
đối tượng chất liệu	material object
đối tượng hình thức	formal object
đối vị	paradigmatic
đối xứng	symetric
đồng nghĩa	synonymous
động ngữ	x. ngữ đoạn động từ
động từ	verb
động từ chính	main verb
động từ nguyên mẫu	infinitive verb

động từ nhị xuyên
động từ phụ
đơn vị cú pháp
đơn vị từ vựng
đứng trước

ditransitive verb
auxiliary verb
syntax unit
lexical item
precede

G

giả thuyết
giống tính
giọng
giới hạn chọn lựa
giới hạn của ngôn ngữ
giới hạn tiểu phạm trù hóa
giới ngữ
giới từ

hypothesis
gender
voice
selectional restriction
delimitation of a language
subcategorization restriction
x. ngữ đoạn giới từ
preposition

H

hạn chế cục bộ
hàm chọn lựa
hàm đặc tính
hàm đặc trưng
hàm số
hậu tố
hệ chuyên gia
hệ hình
hệ nhận diện ngôn ngữ
hệ thống ký hiệu
hiện tại
hình thái học
hình thức
hình thức liên từ (dạng liên từ)
hình tự
hình vị
hình vị phái sinh
hư từ
hồi quy
hoán vị

locality constrain
selecting function
characteristic function
characteristic function
function
postfix
expert system
paradigm
language recognizing system
system of symbols
present
morphology
formal
subjunctive form
ideo-graph
morpheme
derivational morpheme
dummy word
recursive
permutation

hữu hạn
hóa trị

finite
valence

K

kênh
kết thúc
khẳng định
kho từ vựng
khoa học nhận thức
không
không tận cùng
khung tiểu phạm trù hóa
kiểm soát
ký hiệu
ký tự

channel
x. tận cùng
maintain
vocabulary
cognitive science
zero
non-terminal
subcategorisation frame
control
symbol
character

L

lệ thuộc ngữ nghĩa
lên cấp (giáng cấp) động từ
lệnh-C
lệnh-C bất đối xứng
lệnh-C đối xứng
liên đới
liên từ
loại từ vựng
logic mờ
logic vị ngữ
lọc cách
lồng bỏ thể
lưới theta
lớp chỉ định từ
lớp danh từ
lớp động từ
lớp tính từ
lớp từ
lớp tương đương
Lý thuyết tiêu chuẩn

according to the context
verb raising
C-command
asymetric C-command
symetric C-command
interdependency
conjunction
lexical type
fuzzy logic
predicate logic
case filter
expletive insertion
theta grid
determiner class
noun class
verb class
adjective class
word class
equivalent class
Standard Theory

Lý thuyết tiêu chuẩn mở rộng
Lý thuyết chi phối và ràng buộc
Lý thuyết nguyên tắc và thông số

Lý thuyết ràng buộc
Lý thuyết X-gạch

Extended Standard Theory
Government and Binding Theory
Principles and Parameters
Theory
Binding theory
X-bar theory

M

ma trận
mã nhị phân
mạo từ
mạo từ bất định
mạo từ xác định
mẫu cấu trúc
mẫu phát âm
mẫu tự
mẫu-trùng hợp
mệnh lệnh ngữ
mệnh đề
mệnh đề chính
mệnh đề lồng
mệnh đề phụ
mô tả cấu trúc
mục tiêu

matrix
binary code
article
non-specified article
specified article
structure pattern
accentual pattern
alphabet
pattern-matching
imperative phrase
clause
main clause
embedded clause
subordinate clause
structure description
goal

N

N-gạch
nghỉ
nghị vấn
nghịch đảo chủ ngữ/động từ phụ
ngoại chủ thể
ngoại đề
ngôn ngữ chết
ngôn ngữ cơ thể
ngôn ngữ con người
ngôn ngữ hình thức
ngôn ngữ học điện toán

N-bar
break
interrogative
subject/aux inversion
external subject
external theme
dead language
body language
human language
formal language
computational linguistics

ngôn ngữ lập đồ	projecting language
ngôn ngữ lập trình	programming language
ngôn ngữ lập trình cao	high programming language
ngôn ngữ máy	machine language
ngôn ngữ nghèo hình vị	morphologically poor language
ngôn ngữ nghịch đảo	inversion language
ngôn ngữ ngoại hóa	externalized language
ngôn ngữ nội hóa	internalized language
ngôn ngữ thứ nhất (ngôn ngữ chính)	primary speech
ngôn ngữ tự nhiên	natural language
ngôn ngữ vũ trụ	universal language
nguồn gốc	source
nguyên mẫu	infinitive
nguyên tắc chiếu mở rộng	Extended Projection Principle (EPP)
nguyên thể	infinitive
nguyên tử	atom
ngưng	break, stop
ngữ âm học	phonetics
ngữ cảnh	context
ngữ nghĩa	semantic
ngữ nghĩa học	semantics
ngữ đoạn	phrase
ngữ đoạn chỉ định từ (DP)	determiner phrase (DP)
ngữ đoạn danh từ (NP)	noun phrase (NP)
ngữ đoạn giới từ (PP)	prepositional phrase (PP)
ngữ đoạn động từ (VP)	verb phrase (VP)
ngữ đoạn tác tử bổ ngữ hóa (CP)	complementizer phrase (CP)
ngữ đoạn thì (TP)	tense phrase (TP)
ngữ đoạn tính từ (AP)	adjective phrase (AP)
ngữ pháp	grammar
Ngữ pháp biến hình tạo sinh	Generative Transformation Grammar
Ngữ pháp cấu trúc ngữ đoạn	Phrase Structure Grammar
Ngữ pháp cấu trúc ngữ đoạn tổng quát hóa	Generalized Phrase Structure Grammar
ngữ pháp cảm ngữ cảnh	context sensitive grammar
ngữ pháp chính quy	regular grammar
Ngữ pháp chức năng	Functional Grammar

Ngữ pháp chức năng từ vựng	Lexical Functional Grammar
ngữ pháp phi ngữ cảnh	context free grammar
Ngữ pháp tạo sinh	Generative Grammar
ngữ pháp truyền thống	traditional grammar
người khởi động	initiator
người đối thoại	colocutor
người nói tiếng mẹ đẻ	native speaker
nhãn	label
nhãn hiệu	x. nhãn
nhánh	branch
nhận dạng	recognition
nhận diện	x. nhận dạng
nhấn mạnh	stress
nhận thức	recognition
nhân vật	person
nhìn trước	look ahead
nở (lan) rộng	expansion
nút	node
nút chị em	sister node
nút con	daughter node
nút không tận cùng	non-terminal node
nút mẹ	mother node
nút đỉnh	root node
nút tận cùng	terminal node

P

P-gạch	P-bar
phạm trù cú pháp	syntax category
phạm trù ngữ pháp	grammatical category
phạm trù rỗng	empty category
phạm trù từ vựng	lexical category
phạm trù trống	empty category
phạm vi ràng buộc	binding domain
phái sinh	derivation
phái tính	gender
phát âm	pronunciation
phân bố	distribution
phân bố bổ sung	complementary distribution

phân loại	taxonomy
phân đoạn	segment
phân tích cú pháp	syntax analyse
phân tích ngữ nghĩa	semantic analyse
phân tích tín hiệu	signal analyse
phần tính	computational component
phép biến hình	transformation
phép bố trí	coordination
phép chiếu bình diện gạch	bar-level projection
phép chiếu phủ định	negation projection
phép tính lược	reduction
phép thế	substitution
phép thế-“do so”	do-so-replacement
phép thế-“one”	one-replacement
phép thế-“so”	so-replacement
phi ngữ cảnh	context free
phi tự nhiên	non-natural
phụ ngữ	adjunct
phương ngữ	dialect
phương pháp truyền thông	method of communication
phương tiện giao lưu	communicative devices

Q

quan hệ	relation
quan hệ (logic)	coherence
quan hệ chủ đề	thematic relation
quan hệ chủ động-bị động	active-passive relation
quan hệ lệ thuộc	dependency
quan hệ lệ thuộc cấu trúc	structure dependence
quan hệ ngữ pháp	grammatical relation
quá khứ	past
quét	scan
quy nạp	induction
quy tắc	rule
quy tắc biến hình	transformation rule
quy tắc cấu trúc ngữ đoạn	phrase structure rule
quy tắc lồng	insertion rule
quy tắc lồng từ vựng	lexical insertion rule

quy tắc logic	logic rule
quy tắc ngữ pháp	grammatical rule
quy tắc phạm trù	categorial rule
quy tắc sản sinh	production rule
quy tắc tiểu phạm trù phi ngữ cảnh	context free subcategory rule

R

ràng buộc	bind
rỗng	null, empty

S

siêu ngôn ngữ	metalanguage
sinh cách	genitive
sinh cách tự do	free genitive
sinh cách-“s”	's genitive
sinh cách-“of”	of genitive
so sánh	comparison
số (lượng)	number
số ít	singular
số nhiều	plural
sống	alive

T

tác nhân	agent, actor
tác tử bổ ngữ hóa	complementizer
tán thán từ	interjection
tận cùng	terminal
tân ngữ (thường gọi: túc từ)	object
tặng cách	dative
tham số	argument
thao tác	procedure
thành tố	constituent
thành tố trực tiếp	immediate constituent
thái	voice
thái bị động	passive voice
thái chủ động	active voice
thẩm năng ngữ dụng	pragmatical competence

thảm năng giao lưu	communicative competence
thảm năng ngôn ngữ	language competence
thảm năng ngữ pháp	grammatical competence
thế tiền vị	precedence
thế tiền vị trực tiếp	immediate precedence
thì	tense
thói quen ngôn ngữ	language habit
thông báo	convey
thông số	parameter
thông tin	information
thụ nhân	recipient
thuộc chủ	possessed
tiền ngữ	antecedent
tiền tố	prefix
tiếng	sound
tiếng địa phương	dialect
tiếp vĩ ngữ	suffix
tiểu phạm trù	subcategory
tiểu phạm trù hóa	subcategorised
tiểu phạm trù phi ngữ cảnh	context free subcategory
tìm	search
tín hiệu đầu vào	input signal
tín học	informatics
tính ngữ	x. ngữ đoạn tính từ
tính từ	adjective
tính từ sở hữu	possessive adjective
Tối thiểu luận	Minimalism
tôn ti	hierarchy
tôn ty Chomsky	Chomsky's hierarchy
trạng từ	adverb
trí thông minh nhân tạo	artificial intelligence
trống	empty, NULL
trợ động từ	auxiliary verb
trùng hợp	matching
trùng ngữ	anaphor
túc từ	x. tân ngữ, bổ ngữ
túc từ gián tiếp	indirect object
túc từ trực tiếp	direct object
tuyến chiếu	projection line

tuyến tính	linear
từ	word
từ dưới lên trên	bottom up
từ ghép	compound word
từ điển	lexicon
từ đơn	single word
từ trên xuống dưới	top down
từ vựng	lexical
tương lai	future
U	
ưu thế	dominance
V	
V-gạch	V-bar
vai trò theta	theta role
vai trò theta ngoài	external theta role
vai trò theta trong	internal theta role
vết	trace
vị ngữ	predicate (phrase)
vị từ	predicate
vị trí	location, position
vị trí ưu thế	dominant position
vô hạn	infinite
vòng vô tận	endless loop
vũ trụ	universal
X	
X-gạch hạn chế	constrain X-bar
xuống cấp (nâng cấp) động từ	verb lowering

BẢNG ĐỐI CHIẾU THUẬT NGỮ VIỆT - ANH

BẢNG ĐỐI CHIẾU THUẬT NGỮ ANH - VIỆT**A**

A-bar	A-gạch
abstract case	cách trừu tượng
accentual pattern	mẫu phát âm
according to the context	lệ thuộc ngữ nghĩa
accusative	đối cách
active sentence	câu chủ động
active	chủ động
active voice	thái chủ động
active-passive relation	quan hệ chủ động-bị động
adjective	tính từ
adjective class	lớp tính từ
adjective phrase (AP)	ngữ đoạn tính từ (AP)
adjunct	phụ ngữ
adverb	trạng từ
agent	tác nhân
alive	sống
alphabet	mẫu tự, chữ cái
alpha-movement	chuyển vị-alpha
ambiguity	đa nghĩa
ambiguous	đa nghĩa
anaphor	trùng ngữ
antecedent	tiền ngữ
argument	tham số
article	mạo từ
artificial intelligence	trí thông minh nhân tạo
asymmetric	bất đối xứng
asymmetric C-command	lệnh-C bất đối xứng
atom	nguyên tử
auxiliary verb	động từ phụ, trợ động từ
axiom	chân lý

B

bar-level projection	bóng chiếu bình diện gạch
----------------------	---------------------------

benefactive	cống hiến
binary code	mã nhị phân
bind	ràng buộc
binding domain	phạm vi ràng buộc
Binding Theory	Lý thuyết ràng buộc
body language	ngôn ngữ cơ thể
bottom up	từ dưới lên trên
branch	nhánh
break	ngừng, nghỉ

C

case	cách
case filter	lọc cách
categorial rule	quy tắc phạm trù
C-command	lệnh-C
channel	kênh
character	ký tự
characteristic function	hàm đặc trưng, hàm đặc tính
Chomsky's hierarchy	tôn ty Chomsky
clause	mệnh đề
cleft sentence	câu chẻ
cognitive science	khoa học nhận thức
coherence	quan hệ (logic)
coindex	chỉ số phụ
colocutor	người đối thoại
common noun	danh từ chung
communicative competence	thẩm năng giao lưu
communicative devices	phương tiện giao lưu
comparison	so sánh
complement	bổ ngữ
complementary distribution	phân bố bổ sung
complementizer	tác tử bổ ngữ hóa
complementizer phrase (CP)	ngữ đoạn tác tử bổ ngữ hóa (CP)
compound word	từ ghép
computational component	phần tính
computational linguistics	ngôn ngữ học điện toán
condition	điều kiện
conjunction	liên từ

constituent	thành tố
constrain X-bar	X-gạch hạn chế
context	ngữ cảnh
context free	phi ngữ cảnh
context free grammar	ngữ pháp phi ngữ cảnh
context free subcategory	tiểu phạm trù phi ngữ cảnh
context free subcategory rule	quy tắc tiểu phạm trù phi ngữ cảnh
context free subcategory rule	quy tắc tiểu phạm trù phi ngữ cảnh
context sensitive	cảm ngữ cảnh
context sensitive grammar	ngữ pháp cảm ngữ cảnh
control	kiểm soát
convey	thông báo
coordination	phép bố trí

D

data	dữ liệu
database	bảng dữ liệu
dative	tặng cách
daughter node	nút con
dead language	ngôn ngữ chết
deep structure	cấu trúc sâu
definite article	mạo từ xác định
delimitation of a language	giới hạn của ngôn ngữ
demonstrative pronoun	đại danh từ chỉ thị
dependency	quan hệ lệ thuộc
derivation	phái sinh
derivational morpheme	hình vị phái sinh
determiner	chỉ định từ
determiner class	lớp chỉ định từ
determiner phrase (DP)	ngữ đoạn chỉ định từ (DP)
dialect	tiếng địa phương, phương ngữ
direct object	túc từ trực tiếp
distribution	phân bố
ditransitive verb	động từ nhị xuyên
dominance	ưu thế, chi phối
dominant position	vị trí ưu thế

do-so-replacement
D-structure
dummy word

phép thế-“do so”
cấu trúc-D
hư từ

E

echo question
embedded clause
empty
empty category
endless loop
equivalent class
exhaustive domination
expansion
experiencer
expert system
expletive
expletive insertion
express
expression
Extended Projection Principle (EPP)
Extended Standard Theory
external subject
external theme
external theta role
externalized language

câu hỏi hỏi ngôn
mệnh đề lồng
rỗng, trống
phạm trù trống, phạm trù rỗng
vòng vô tận
lớp tương đương
chi phối tận mức
nở (lan) rộng
cảm nhân
hệ chuyên gia
bổ thể
lồng bổ thể
diễn đạt
diễn thức, biểu thức
Nguyên tắc chiếu mở rộng
Lý thuyết tiêu chuẩn mở rộng
ngoại chủ thể
ngoại đề
vai trò theta ngoài
ngôn ngữ ngoại hóa

F

finite
formal
formal language
formal noun
formal object
free genitive
function
Functional Grammar
future
fuzzy logic

hữu hạn
hình thức
ngôn ngữ hình thức
danh từ hình thức
đối tượng hình thức
sinh cách tự do
chức năng, hàm số
Ngữ pháp chức năng
tương lai
logic mờ

G

gender	giống tính, phái tính
Generalized Phrase Structure Grammar	Ngữ pháp cấu trúc ngữ đoạn tổng quát hóa
Generative Grammar	Ngữ pháp tạo sinh
Generative Transformation Grammar	Ngữ pháp biến hình tạo sinh
genitive	sinh cách
goal	mục tiêu
government	chi phối
Government and Binding Theory	Lý thuyết chi phối và ràng buộc
grammar	ngữ pháp
grammatical category	phạm trù ngữ pháp
grammatical competence	thẩm năng ngữ pháp
grammatical relation	quan hệ ngữ pháp
grammatical rule	quy tắc ngữ pháp
grammatical subject	chủ ngữ ngữ pháp, chủ thể ngữ pháp
group of characters	cụm ký tự

H

head	đầu
head-to-head	đầu tới đầu
head-to-head movement	chuyển vị đầu tới đầu
hierarchy	tôn ti
high programming language	ngôn ngữ lập trình cao
human language	ngôn ngữ con người
hypothesis	giả thuyết

I

ideo-graph	hình tự
immediate constituent	thành tố trực tiếp
immediate precedence	thế tiên vị trực tiếp
immediately dominate	chi phối trực tiếp
imperative phrase	mệnh lệnh ngữ
index	chỉ số

indirect object	túc từ gián tiếp
induction	quy nạp
infinite	vô hạn
infinitive	nguyên mẫu, nguyên thể
infinitive verb	động từ nguyên mẫu
inflection	biến thể
informatics	tin học
information	thông tin
initiator	người khởi động
input signal	tín hiệu đầu vào
insertion rule	quy tắc lồng
instrument	dụng cụ
interjection	tán thán từ
intermediate level	bình diện trung gian
internal theta role	vai trò theta trong
internalized language	ngôn ngữ nội hóa
interdependency	liên đới
interpretation	diễn giải
interrogative	ngghi vấn
interrogative pronoun	đại danh từ nghi vấn
inversion language	ngôn ngữ nghịch đảo
involve	bao gồm
L	
label	nhãn hiệu, nhãn
language competence	thẩm năng ngôn ngữ
language habit	thói quen ngôn ngữ
language performance	dụng năng ngôn ngữ
language recognizing system	hệ nhận diện ngôn ngữ
level	bình diện
lexical	từ vựng
lexical bank	băng từ vựng
lexical category	phạm trù từ vựng
Lexical Functional Grammar	Ngữ pháp chức năng từ vựng
lexical head	đầu từ vựng
lexical insertion rule	quy tắc lồng từ vựng
lexical item	đơn vị từ vựng
lexical type	loại từ vựng

lexicon	từ điển
linear	tuyến tính
list	bảng
locality condition	điều kiện về tính cục bộ
locality constrain	hạn chế cục bộ
location	vị trí
logic rule	quy tắc logic
logical subject	chủ thể logic
look ahead	nhìn trước

M

machine language	ngôn ngữ máy
main clause	mệnh đề chính
main verb	động từ chính
maintain	khẳng định
matching	trùng hợp , ăn khớp
material object	đối tượng chất liệu
matrix	ma trận
metalanguage	siêu ngôn ngữ
method of communication	phương pháp truyền thông
minimal syntax structure	cấu trúc cú pháp tối thiểu
Minimalism	Tối thiểu luận
morpheme	hình vị
morphologically poor language	ngôn ngữ nghèo hình vị
morphology	hình thái học
mother node	nút mẹ
move	chuyển vị
movement	chuyển vị
multi-lingual	đa ngôn ngữ

N

native speaker	người nói tiếng mẹ đẻ
natural language	ngôn ngữ tự nhiên
N-bar	N-gạch
negation projection	phép chiếu phủ định
node	nút
nominative	danh cách

non-echo question	câu hỏi phi hồi ngôn
non-natural	phi tự nhiên
non-specified article	mạo từ bất định
non-terminal	không tận cùng
non-terminal node	nút không tận cùng
noun	danh từ
noun class	lớp danh từ
noun phrase (NP)	ngữ đoạn danh từ (NP)
null	rỗng, trống
number	số (lượng)

O

object	đối tượng, tân ngữ (thường gọi: túc từ)
of genitive	sinh cách-“of”
one-replacement	phép thế-“one”
opposition	đối lập

P

paradigm	hệ hình
paradigmatic	đối vị
parameter	thông số
passive	bị động
passive sentence	câu bị động
passive voice	thái bị động
past	quá khứ
pattern	dữ liệu mẫu
pattern matching list	bảng mẫu trùng hợp
pattern-matching	mẫu-trùng hợp
P-bar	P-gạch
permutation	hoán vị
person	nhân vật
personal pronoun	đại danh từ nhân vật
phoneme	âm vị
phonetics	ngữ âm học
phonology	âm vị học
phrase	ngữ đoạn

Phrase Structure Grammar

phrase structure rule

plural

possessed

possessive adjective

possessive pronoun

possessor

position

postfix

pragmatic

pragmatical competence

pragmatics

precede

precedence

predicate

predicate (phrase)

predicate logic

prefix

preposition

prepositional phrase (PP)

present

primary speech

Principles and Parameters Theory

procedure

production rule

programming language

project

projecting language

projection line

pronoun

pronunciation

proper noun

pseudo-cleft sentence

psychological subject

Ngữ pháp cấu trúc ngữ đoạn

quy tắc cấu trúc ngữ đoạn

số nhiều

thuộc chủ

tính từ sở hữu

đại danh từ sở hữu

chủ

vị trí

hậu tố

dụng pháp, ngữ dụng

thẩm năng ngữ dụng

dụng pháp học

đi trước, đứng trước

thế tiên vị

vị từ

vị ngữ

logic vị ngữ

tiền tố

giới từ

ngữ đoạn giới từ (PP)

hiện tại

ngôn ngữ thứ nhất (ngôn ngữ

chính)

Lý thuyết nguyên tắc và thông số

thao tác

quy tắc sản sinh

ngôn ngữ lập trình

chiếu

ngôn ngữ lập đồ

tuyến chiếu

đại danh từ

phát âm

danh từ riêng

câu chẻ giả

chủ thể tâm lý

R

raising movement

chuyển vị lên cấp

recipient	thụ nhân
reciprocal pronoun	đại danh từ hỗ tương
recognition	nhận dạng, nhận diện, nhận thức
recursive	hồi quy
reduction	phép tính lược
refer	dẫn chỉ
referring expression	biểu thức dẫn chỉ
reflexive pronoun	đại danh từ tái quy
regular	chính quy
regular grammar	ngữ pháp chính quy
relation	quan hệ
relative pronoun	đại danh từ quan hệ
root node	nút đỉnh
rule	quy tắc

S

's genitive	sinh cách-“'s”
sample	dữ liệu muốn tìm
scan	quét
search	tìm
segment	phân đoạn
selecting function	hàm chọn lựa
selectional restriction	giới hạn chọn lựa
semantic	ngữ nghĩa
semantic analyse	phân tích ngữ nghĩa
semantics	ngữ nghĩa học
sentence	câu
signal analyse	phân tích tín hiệu
single word	từ đơn
singular	số ít
sister node	nút chị em
so-replacement	phép thế-“so”
sound	tiếng
source	nguồn gốc
specified article	mạo từ xác định
specifier	biệt định ngữ
speech-community	cộng đồng tiếng nói
speech sound	âm tố

S-structure	cấu trúc-S
Standard Theory	Lý thuyết tiêu chuẩn
stop	ngưng
stress	nhấn mạnh
strikt local	cục bộ nghiêm ngặt
structure dependence	quan hệ lệ thuộc cấu trúc
structure description	mô tả cấu trúc
structure pattern	mẫu cấu trúc
subcategorisation frame	khung tiểu phạm trù hóa
subcategorised	tiểu phạm trù hóa
subcategorization restriction	giới hạn tiểu phạm trù hóa
subcategory	tiểu phạm trù
subject	chủ thể, chủ ngữ (thường gọi: chủ từ)
subject/aux inversion	ngịch đảo chủ ngữ/động từ phụ
Theme-Rheme	đề diễn
subject/theme sensitive	cảm đề
Subject-Object-Verb structure (SOV)	cấu trúc Chủ từ-Túc từ-Động từ (SOV)
Subject-Verb-Object structure (SVO)	cấu trúc Chủ từ-Động từ-Túc từ (SVO)
subjunctive form	hình thức liên từ (dạng liên từ)
subordinate clause	mệnh đề phụ
substitution	phép thế
substructure	cấu trúc phụ
suffix	tiếp vĩ ngữ
surface structure	cấu trúc bề mặt
syllable	âm tiết
symbol	dấu hiệu, ký hiệu
symetric	đối xứng
symetric C-command	lệnh-C đối xứng
synonymous	đồng nghĩa
syntax	cú pháp, cú pháp học
syntax analyse	phân tích cú pháp
syntax category	phạm trù cú pháp
syntax structure	cấu trúc cú pháp
syntax tree	cây cú pháp
syntax unit	đơn vị cú pháp
system of symbols	hệ thống ký hiệu

T

taxonomy	phân loại
tense	thì
tense phrase (TP)	ngữ đoạn thì (TP)
terminal	tận cùng, kết thúc
terminal node	nút tận cùng, nút kết thúc
thematic relation	quan hệ chủ đề
theme	chủ đề
theme - interpretation	chủ đề – diễn giải
theta criterion	chuẩn tắc theta
theta grid	lưới theta
theta role	vai trò theta
T-movement	chuyển vị T (thì T)
top down	từ trên xuống dưới
trace	vết
traditional grammar	ngữ pháp truyền thống
transform	biến hình
transformation	phép biến hình
transformation rule	quy tắc biến hình
tree	cây
universal	vũ trụ
universal language	ngôn ngữ vũ trụ
utterance	biểu đạt

V

valence	hóa trị
variable	ẩn số, biến số
V-bar	V-gạch
verb	động từ
verb class	lớp động từ
verb lowering	xuống cấp (nâng cấp) động từ
verb phrase (VP)	ngữ đoạn động từ (VP)
verb raising	lên cấp (giáng cấp) động từ
vocabulary	kho từ vựng
voice	giọng; thái

W

well-formed	chuẩn hình
wh-movement	chuyển vị-wh
word	từ
word class	lớp từ

X

X-bar projection	bóng chiếu X-gạch
X-bar theory	lý thuyết X-gạch
yes/no question	câu hỏi có/ không

Z

zero	không
------	-------

BẢNG ĐỐI CHIẾU THUẬT NGỮ ANH - VIỆT

TÀI LIỆU THAM KHẢO

- Abney, Steven (1987) *The English Noun Phrase in its Sentential Aspect*. Ph.D. Dissertation. MIT.
- Abraham, Werner (Hrsg.) (1978) *Valence, Semantic Case and Grammatical Relations*. Amsterdam: John Benjamins.
- Aoun, Joseph (1985) *A grammar of Anaphora*. Cambridge: MIT Press.
- Arens, Hans (1969) *Sprachwissenschaft. Der Gang ihrer Entwicklung von der Antike bis zur Gegenwart*. 2 Bde. Frankfurt: Athenäum Fischer Taschenbuch Verlag.
- Baker, Mark, Kyle Johnson, and Ian Roberts (1989) *Passive Arguments Raised*. *Linguistic Inquiry* 20: 219-251.
- Baltin, Mark (1981) *Strict Bounding*. In Baker C. L. and John McCarthy (eds) *The Logical Problem of Language Acquisition*. Cambridge: MIT Press: 257-295.
- Baltin, Mark and Anthony Kroch (1989) *Alternative Conceptions of Phrase Structure*. Chicago: Chicago University Press.
- Bauer, Laurie (1983) *English Word-formation*. Cambridge: Cambridge University Press.
- Bierwisch, Manfred (1966) *Strukturalismus. Geschichte, Probleme und Methoden*. In: *Kursbuch 5* (1966), 77-152. Verbesselter Abdruck in: Ihwe (1971), 17-90.
- Bloomfield, Leonard (1926) *A Set of Postulates for the Science of Language*. In: *Language* 2, 153-64.
- Bloomfield, Leonard (1933) *Language*. London: Allen & Unwin, 1935.
- Borsley, Robert (1996) *Modern Phrase Structure Grammar*. Cambridge Mass: Blackwell.
- Bresnan, Joan (ed.) (1982). *The Mental Representation of Grammatical Relations*. Cambridge: MIT Press
- Bresnan, Joan (forthcoming) *Lexical Functional Grammar*. Oxford: Blackwell.
- Bùi Kỳ, Trần Trọng Kim (1925) *Kiểu*. Hà Nội: Vĩnh Hưng long thư quán.
- Burzio, Luigi (1986) *Italian Syntax*. Dordrecht: Reidel.
- Bühler, Karl (1934) *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Frankfurt: Ullstein Verlag, 1978 (Ullstein Buch Nr. 3392)
- Bünting, Karl-Dieter (1987) *Einführung in die Linguistik*. 12. Auflage, Frankfurt: Athenäum.
- Carnie, Andrew (1995) *Head Movement and Non-Verbal Predication*. Ph.D. Dissertation, MIT.
- Carnie, Andrew (2000) *Syntax*. Arizona: University of Arizona Press.
- Chafe, Wallace L. (1970) *Meaning and the structure of language*. Chicago: The University of Chicago Press.
- Chametzky, Robert (1996) *A Theory of Phrase Markers and the Extended Base*. Albany: SUNY Press.
- Cheng, Lisa (1997) *On the Typology of Wh-Questions*. New York: Garland Press.
- Chomsky, Noam (1957) *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam (1964) *Current Issues in Linguistic Theory*. The Hague: Mouton.
- Chomsky, Noam (1965) *Aspects of the Theory of Syntax*. Cambridge (Mass.): The M.I.T. Press (Dt. *Aspekte der Syntax-Theorie*. Frankfurt/M: Suhrkamp, 1969)
- Chomsky, Noam/Halle, Morris (1968) *The Sound Pattern of English*.: New York; Evanston; London: Harper & Row, Publishers.

- Chomsky, Noam (1970) *Remarks on Nominalization* in R. Jacobs and P. Rosenbaum (eds.) *Readings in English Transformational Grammar*. Waltham: Ginn. pp 184-221.
- Chomsky, Noam (1975) *The Logical Structure of Linguistic Theory*. New York: Plenum
- Chomsky, Noam (1977) *On Wh-movement*. In Peter Culicover, Thomas Wasow and Adrian Akmajian (eds). *Formal Syntax*. New York: Academic Press. pp. 71-132
- Chomsky, Noam (1979) *Language and Responsibility*. New York: Pantheon Books.
- Chomsky, Noam (1980) *On Binding*. *Linguistic Inquiry* 11, 1-46
- Chomsky, Noam (1981) *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam (1993) *A Minimalist Program for Linguistic Theory*. In Kenneth L. Hale and Samuel J. Keyser (eds) *The View from Building 20*. Cambridge: MIT Press.
- Chomsky, Noam (1995) *The Minimalist Program*. Cambridge: MIT Press.
- Dalrymple, Mary, Ronald Kaplan, John Maxwell & Annie Zaenen (eds) (1995) *Formal Issues in Lexical-Functional Grammar*. Stanford: CSLI.
- Duden (1981) *Das große Wörterbuch der deutschen Sprache in sechs Bänden*. Hrsg. und bearb. v. Wiss. Rat und den Mitarbeitern der Dudenredaktion unter Leitung von Günther Drosdowski; Mannheim: Bibliographisches Institut.
- Đỗ Quang Chính (1972) *Lịch sử chữ Quốc Ngữ 1620-1659*. Saigon: Nxb. Ra Khơi.
- Edmondson, Jerold and Donald A. Burquest (1998) *A Survey of Linguistic Theories*. (3 rd ed.) Dallas: Summer Institute of Linguistics.
- Esser, Wilhelm K. (1970) *Wissenschaftstheorie I. Definition und Reduktion*. Alber: Freiburg; München.
- Fabb, Nigel (1994) *Sentence Structure*. London: Routledge.
- Falk, Yehuda N. (forthcoming) *Lexical Functional Grammar: An Introduction to Parallel Constraint-Based Syntax*. Stanford: CSLI Publications.
- Fillmore, Charles J. (1968) *The case for case*. In: *Universals in linguistic theory*. New York: Ed. by E. Bach & R.J. Harms.
- Fleischer, Wolfgang (1972) *Wortbildung der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Gazdar et al. (1985) *Generalized Phrase Structure Grammar*. Cambridge: Harvard University Press
- Gleason, H. A. (1961) *An Introduction to Descriptive Linguistics*. Revised edition, New York: Holt, Rinehart and Winston.
- Glinz, Hans (1965) *Grundbegriffe und Methoden inhaltbezogener Text- und Sprachanalyse*.
- Grewendorf, Günther/Hamm, Fritz/Sternefeld, Wolfgang (1989) *Sprachliches Wissen. Eine Einführung in moderne Theorien der grammatischen Beschreibung*. 3. Aufl.: Frankfurt/M: Suhrkamp (suhrkamp taschenbuch wissenschaft 695)
- Grimshaw, Jane (1990) *Argument Structure*. Cambridge: MIT Press.
- Haegeman, Liliane. (1991) *Introduction to Government and Binding Theory*. Oxford: Basil Blackwell.
- Habermas, Jürgen (1971) *Vorbereitende Bemerkungen zu einer Theorie der kommunikativen Kompetenz*. In: Jürgen Habermas/ Niklas Luhmann (Hrsg.), *Theorie der Gesellschaft oder Sozialtechnologie — Was leistet die Systemforschung?* Frankfurt/M: Suhrkamp, 101–141
- Halliday, M.A.K, Angus McIntosh, Peter Stevens (1964) *The Linguistic Sciences and Language Teaching*. London: Longmans.
- Halliday, M.A.K (1994) *An introduction to Functional Grammar*. London: Edward Arnold.
- Heim, Irene and Kratzer, Angelika (1998) *Semantics in Generative Grammar*. Malden: Blackwell.

- Higginbotham, James (1980) *Pronouns and Bound Variables*. Linguistic Inquiry 11, 697-708.
- Higginbotham, James (1985) *A Note on Phrase Markers*. MIT Working Papers in Linguistics 6: 87-101.
- Hopcroft, J.E. and J.D. Ullman (1979) *Introduction to automata theory, languages and computation*. New York: Addison-Wesley.
- Holzman, Mathilda (1997) *The Language of Children* (2nd Edition). Cambridge: Blackwell.
- Hymes, Dell (1972) *Competence and performance in linguistic theory*. In: R. Huxley & E. Ingram (eds.), *Language Acquisition: Models and Methods*. New York: Academic Press, 3-28.
- Ihwe, Jens (Hrsg.) (1971) *Literaturwissenschaft und Linguistik. Ergebnisse und Perspektiven. Bd. 1: Grundlagen und Voraussetzungen*. Frankfurt: Athenäum.
- Jackendoff, Ray (1977) *X-bar Syntax: A Theory of Phrase Structure*. Cambridge: MIT Press.
- Jackendoff, Ray (1993) *Patterns in the Mind*. London: Harvester-Wheatsheaf.
- Jackendoff, Ray (1983) *Semantics and Cognition*. Cambridge (Mass.): MIT Press
- Jaeggli, Osvaldo (1986) *Passive*. Linguistic Inquiry 17. 4: 587-622.
- Jaeggli, Osvaldo and Kenneth Safir (eds) (1989) *The Null Subject Parameter*. Dordrecht: Kluwer Academic Press.
- Jelinek, Eloise (1984) *Empty Categories, Case, and Configurationality*. Natural language and Linguistic Theory 2: 39-76.
- Kaplan, Ronald (1995) *The Formal Architecture of Lexical Functional Grammar*. in Dalrymple, Mary et al. (eds). *Formal Issues in Lexical Functional Grammar*. Stanford: CSLI.
- Kaplan, Ronald and Annie Zaenen (1982) *Lexical Functional Grammar: A Formal System for Grammatical Representation*. In Joan Bresnan (ed.) (1982). *The Mental Representation of Grammatical Relations*. Cambridge: MIT Press.
- Kastovsky, Dieter (1982) *Wortbildung und Semantik*. Düsseldorf: Pädagogischer Verlag Schwann; Bagel, Bern & München: Francke Verlag. (Studienreihe Englisch 14)
- Lasnik, Howard (1989) *Essays on Anaphora*. Dordrecht: Kluwer Academic Publishers.
- Lê Công Liệu (1973) *Đại Việt Thông Sử của Lê Quý Đôn (1759)*. Hà Nội: Viện Sử Học.
- Lê Mạnh Liêu (1973) *Lê Quý Đôn: Đại Việt Thông Sử*. Saigon: Bộ Văn Hóa Giáo Dục và Thanh Niên.
- Lê Quý Đôn (1977) *Đại Việt Thông Sử*. Viện Sử Học Hà Nội, Nxb. Khoa học Xã hội.
- Lê Văn Lý (1948) *Le Parler Vietnamiens*. Paris: Huong Anh.
- Lightfoot (1976) *Trace Theory and Twice Moved NPs*. Linguistic Inquiry 7: 559-582
- Lightfoot, David (1991) *How to Set Parameters: Evidence from Language Change*. Cambridge: MIT Press.
- Lightfoot, David and Norbert Hornstein (eds) (1994) *Verb Movement*. Cambridge: Cambridge University Press.
- Leech, Geoffrey (1981) *Semantics. The Study of Meaning*. Second Edition, London: Penguin Books.
- Leont'ev, A.A. (1971) *Sprache - Sprechen - Sprechtaetigkeit*. Stuttgart: Kohlhammer
- Lewandowski, Theodor (1990) *Linguistisches Wörterbuch*. 3 Bde., Heidelberg; Wiesbaden: Quelle & Meyer (UTB 518)

- Linke, Angelika; Nussbaumer, Markus; Portmann Paul R. (1991) *Studienbuch Linguistik*. Tübingen: Niemeyer (Reihe Germanistische Linguistik; 121: Kollegbuch)
- Lyons, John (1977) *Semantics*. 2 vols. Cambridge etc.: Cambridge University Press
- Lyons, John (1981) *Language and Linguistics. An Introduction*. Cambridge etc.: Cambridge University Press.
- Maas, Utz (1973) *Grundkurs Sprachwissenschaft I: Die herrschende Lehre*. München: List Verlag (List Taschenbücher der Wissenschaft 1424)
- Manzini, H. Rita (1983) *On Control and Control Theory*. Linguistic Inquiry 14: 421-446.
- Marantz, Alec (1984) *On Grammatical Relations*. Cambridge: MIT Press.
- May, Robert (1985) *Logical Form*. Cambridge: MIT Press.
- McCloskey, James (1983) *A VP in a VSO language* in Gazdar, G., G. Pullam, and I. Sag eds. *Order Concord and Constituency*. Foris, Dordrecht. 9-55.
- Nguyễn Kim Thản (1997) *Nghiên cứu ngữ pháp tiếng Việt*. Hà Nội: Nxb. Giáo Dục.
- Nguyễn Kim Thản (1999) *Động từ trong tiếng Việt*. Hà Nội: Nxb. Khoa học Xã Hội.
- Nguyễn Tài Cẩn (1999) *Ngữ pháp tiếng Việt*. Hà Nội: Nxb. ĐHQG Hà Nội.
- Perlmutter, David (1979) *Syntactic Argumentation and the Structure of English*. Berkeley: University of California Press.
- Petter, Marga (1998) *Getting PRO under Control*. The Hague: Holland Academic Graphics.
- Phan Khôi (1997) *Việt Ngữ Nghiên Cứu*. Đà Nẵng: Nxb. Đà Nẵng.
- Pinker, Steven (1995) *The Language Instinct*. New York: Harper Perennial.
- Pollard, Carl and Ivan Sag (1994) *Head-Driven Phrase Structure Grammar*. Stanford: CSLI Publications and Chigago: The University of Chicago Press.
- Pollock, Jean-Yves (1989) *Verb-movement, Universal Grammar, and the Structure of IP* Linguistic Inquiry 20, 365-424.
- Postal, Paul (1974) *On Raising*. Cambridge: MIT Press
- Radford, Andrew (1981) *Transformational Grammar: A student's guide to Chomsky's Extended Standard Theory*. Cambridge: Cambridge University Press.
- Radford, Andrew (1988) *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.
- Radford, Andrew (1997) *Syntactic Theory and The Structure of English: A Minimalist Approach*. Cambridge: Cambridge University Press.
- Radford, Andrew (1997) *Syntax: A Minimalist Introduction*. Cambridge: Cambridge University Press.
- Rauh, Gisa (1988) *Tiefenkasus, thematische Relationen und Thetarollen. Die Entwicklung einer Theorie von semantischen Rollen*. Tübingen: Narr.
- Reclam (Hrsg.) (1986) *Friedrich Schiller: Maria Stuart*. Stuttgart: Reclam.
- Redslob, Edwin (1967) *Goethes Leben*. Reclam, Stuttgart.
- Ritter, Elizabeth (1988) *A Head Movement Approach to Construct State Noun Phrases*. Linguistics 26, 909-929.
- Ries, J. (1931) *Was ist ein Satz?*
- Robins, R.H. (1964) *General Linguistics. An Introductory Survey*. London: Longmans.
- Robins, R.H. (1966) *The development of the word class system of the European grammatical tradition*. Foundations of Language, 2. 3-19.
- Robins, R.H. (1973) *Ideen- und Problemgeschichte der Sprachwissenschaft*. Frankfurt/M: Athenäum.
- Rosenbaum, P. S. (1967) *The Grammar of English Predicate Complement Constructions*. Cambridge: MIT Press.

- Saito, Mamoru and Howard Lasnik (1994) *Move Alpha: Conditions on its application and output*. Cambridge: MIT Press.
- Sapir, Edward (1921) *Language*. New York: Harcourt, Brace & World, Inc.
- Saussure, Ferdinand de (1916) *Cours de linguistique générale*. Publié par Charles Bally et Albert Sechehaye, Paris: Payot, 1968.
- Seidel, E. (1935) *Geschichte und Kritik der wichtigsten Satzdefinitionen*.
- Sells, Peter (1985) *Lectures on Contemporary Syntactic Theories*. Stanford: CSLI.
- Soames, Scott and David M. Perlmutter (1979) *Syntactic Argumentation and the Structure of English*. Berkeley: University of California Press.
- Somers, H.L. (1987) *Valency and Case in Computational Linguistics*. Edinburgh: University of Edinburgh Press.
- Suppes, Patrick (1957) *Introduction to Logic*. Princeton: van Nostrand
- Speas, Margaret (1990) *Phrase Structure in Natural Language*. Dordrecht: Kluwer Academic Publishers.
- Stechow (1988) *Bausteine syntaktischen Wissens*. Wesdeuscher Verlag, Opladen,
- Stowell, Tim (1981) *Origins of Phrase Structure*. Ph.D. Dissertation, MIT.
- Tạ Hùng – Bích Hà (1996), *Từ Điển Toán học*. Nxb. Đồng Nai.
- Thái Doãn Hiểu, Hoàng Liên (1994) *Ca dao dân ca – Tình yêu*. TP.HCM: Nxb. TP. HCM
- Travis, Lisa (1984) *Parameters and Effects of Word Order Derivation*. Ph.D. Dissertation, MIT.
- Uriagereka, Juan (1998) *Rhyme and Reason: An Introduction to Minimalist Syntax*. Cambridge Mass: MIT Press.
- Wagner (1999) *Syntax*. In *Linguistik Grundlage*. Hamburg: Uni Manuskript.
- Williams, Edwin (1994) *Thematic Structure in Syntax*. Cambridge: MIT Press.

TÀI LIỆU THAM KHẢO

CHỈ LỤC

“

“à 63

“chứ gì 63

“để 63

“đúng chứ gì 64

“đúng không 64

“thì, là, mà 29, 49, 56

“to 63

A

Abney 217, 341

A-gạch 185

A-gạch (A-bar) 178

Alexandre de Rhodes 6

Amiral 6

ASCII 73

Assembler 7

ảnh trống 18

âm 67

âm thanh 67

âm tiết (syllable) 75

âm tố (speech sound) 75

âm vị (phoneme) 3

B

Ba tham số 248

Barbosa 6

bảng (list) 72

bảng mẫu trùng hợp (pattern matching list)
74

Bảng dữ liệu (database) 71

bảng từ vựng (lexical bank) 70

bài bản 12

biến hình 116

biến số 21, 112

biến thể (inflection) 240

biệt định ngữ (specifier) 52

Biệt định ngữ (specifier) 216

Biểu thức dẫn chỉ (referring expression)

163

bình diện (level) 178

Bình diện chính 188

bình diện gạch (bar-level) 178

Bình diện tận cùng 189

bình diện trung gian 172, 188

bị động 32

Bit 73

Borri 6

**bóng chiếu bình diện gạch (bar-level
projection)** 178

bóng chiếu X-gạch (X-bar projection) 173

bổ nghĩa/phụ nghĩa 38

bổ ngữ (complement) 52, 85, 193

bổ ngữ gián tiếp 294

bổ ngữ trực tiếp 294

bộ phân tích 115

bổ thể (expletive) 267

Bloomfield 2, 89

Bresnan 16

Bùi Kỳ 48

Burzio 297

Byte 72

C

Cathéchismus (Phép giảng tám ngày) 7

cách trừu tượng (abstract case) 295

cảm đề (subject/theme sensitive) 34

Cảm nhân (experiencer) 259

câu 2, 67, 76, 132

câu bị động 297, 307

câu bị động (passive sentence) 59

câu chẻ (cleft sentence) 39, 62, 95

Câu chẻ giả (pseudo-cleft sentence) 96

câu có thì 222

câu hỏi “phải không 63

câu hỏi có/không (yes/no question) 63, 230

câu hỏi có/không (yes/no question) 230

câu hỏi hồi ngôn (echo question) 271, 274,
277, 303, 310

câu hỏi phi hồi ngôn (nonecho question) 272
 câu hỏi- *wh* (*wh-question*) 271, 272, 274, 277, 303, 304, 306, 310
 câu nghi vấn 274, 303
 câu phụ 223
 câu phủ định 282
 cấu trúc bề mặt (*surface structure*) 271, 278
 cấu trúc cú pháp (*syntax structure*) 100, 143
 cấu trúc cú pháp tối thiểu (*minimal syntax structure*) 93
 cấu trúc *D* 278, 281, 282, 292, 304, 305
 cấu trúc *D* (*D-structure*) 23
 cấu trúc luận 3
 cấu trúc ngữ pháp 57
 cấu trúc phụ (*substructure*) 176
 cấu trúc *S* 278, 281, 282, 305
 cấu trúc *S* (*S-structure*) 24
 cấu trúc sâu (*deep structure*) 278
 câu xác định 272, 274, 282, 286, 287, 288, 290, 301, 302, 303, 304, 310
 cây (*tree*) 110
 cây cú pháp (*syntax tree*) 23, 143
 cây trừu tượng 143
 Chân lý thế tiền vị (*Axioms of Precedence*) 149
 Chân lý tính chi phối (*Axioms of Dominance*) 145
 chất liệu 4
 chỉ định từ 103, 117
 chỉ số (*index*) 165
 chỉ số phụ (*coindex*) 165
 chi phối (*government*) 151
 chi phối tận mức (*exhaustive domination*) 146
 Chi phối trực tiếp (*immediately dominate*) 145
 chia động từ 70
 Chomsky 2, 10, 11, 12, 77, 113, 143, 173, 271, 275, 278, 295
 chuẩn hình (*well-formed*) 75
 Chuẩn tắc *theta* (*theta criterion*) 264
 chuỗi từ 67, 70
 chủ (*possessor*) 218
 chủ đề 17
 chủ đề (*theme*) 36, 259
 chủ ngữ 132

chủ ngữ (*subject*) 223
 chủ ngữ ngữ pháp (*grammatical subject*) 33, 245
 chủ động 32
 chủ thể 20, 32
 chủ thể logic (*logic subject*) 36
 chủ thể ngữ pháp (*grammatical subject*) 36
 chủ thể tâm lý (*psychological subject*) 36
 Chủ từ-Động từ-Túc từ 22
 chủ-vị (*subject-predicate*) 245, 270
 chuyển vị (*movement*) 134
 chuyển vị lên cấp (*raising movement*) 25
 chuyển vị đầu tới đầu (*head-to-head movement*) 279
 Chuyển vị động từ *V* (*V* → *T*) 279
 chuyển vị *NP* (*NP movement*) 290, 292
 chuyển vị thì *T* (*T movement*) 285, 286
 chuyển vị *wh* (*wh-movement*) 22
 chuyển vị-*wh* 275, 301
 chữ 70
 chữ Hán 6
 chữ Nôm 6
 chữ viết 67
 chức năng truyền thông 3
 communicative competence 11
 competence 10
 Cours de linguistique générale 4
 công hiến (*benefactive*) 260
 cộng đồng tiếng nói (*speech-community*) 2
 cộng đồng tiếng nói đồng nhất 10, 77
 cục bộ nghiêm ngặt (*strikt local*) 251
 cụm ký tự (*group of characters*) 67
 cụm từ 67, 70
 Cú pháp 14
 cú pháp cấu trúc ngữ đoạn 171
 cú pháp học (*syntax*) 75

D

danh cách 253
 danh cách (*nominative*) 86, 294
 danh ngữ 113
 danh từ 68, 69, 74, 78, 79, 80, 82, 83, 84, 85, 86, 88, 96, 97, 99, 100, 101, 102, 103, 105, 106, 107, 108, 109, 110, 111, 112, 113,

115, 117, 118, 119, 120, 121, 123, 124,
127, 131, 135, 136, 137
danh từ chung 102, 103, 136
danh từ chung (common noun) 254
danh từ hình thức 103, 107, 108, 119, 120
danh từ hình thức (formal noun) 107
Danh từ kiểm soát danh từ 85
Danh từ kiểm soát tính từ 84, 86
danh từ riêng 102, 103
danh từ riêng (proper noun) 254
dấu (symbol) 67
dấu phẩy 53
dấu vết (trace) 282
diễn giải 20
discovery procedures 89
dụng năng ngôn ngữ 58
dụng năng ngôn ngữ (language performance) 11
dụng pháp học (pragmatics) 75
dữ liệu mẫu (pattern) 72
dữ liệu muốn tìm (sample) 72

Đ

đa nghĩa 11
đa nghĩa (ambiguity) 138
Đại danh từ (pronoun) 164
đại danh từ hồ tương 101, 102, 103, 136
đại danh từ nghi vấn 101, 102, 103, 136
đại danh từ nhân vật 101, 102, 103, 136
đại danh từ nhân vật (personal pronoun) 27
Đại danh từ nhân vật kiểm soát động từ 88
đại danh từ quan hệ 101, 102, 103, 136
đại danh từ sở hữu 101, 102, 103, 136
đại danh từ tái quy 101, 103, 136
Đại Việt Thông Sử 47
đầu 188
đầu động từ V 280
đầu thành tố 109
đầu thì T 280
Đầu từ vựng (lexical head) 110
Đề Diễn (Theme-Rheme) 36
đếm được 118
điệp ngữ 28
Điều kiện về tính cục bộ 291

Độ phức tạp của bình diện (level complexity) 190
Đỗ Quang Chính 6
đối cách 176, 253, 254
đối cách (accusative) 86, 294
đối tượng 4, 17
Đối tượng chất liệu 4
Đối tượng của ngôn ngữ học 4, 10
đối tượng hình thức 4, 5
đối tượng nghiên cứu 4
đối tượng vật chất 5
đối vị 90
đồng nghĩa 11
động ngữ 113
động từ 68, 69, 70, 74, 78, 79, 84, 85, 87, 88,
94, 95, 97, 98, 99, 100, 101, 102, 103, 104,
106, 110, 113, 115, 128, 129, 131, 132,
135, 137
động từ bất xuyên tính 248, 250, 251, 257,
263, 269
động từ bất xuyên tính (intransitive verb) 128
Động từ kiểm soát bổ ngữ 87
Động từ kiểm soát danh từ 85
Động từ kiểm soát tính từ 85
Động từ kiểm soát trạng từ 84
động từ nguyên mẫu (infinitive verb) 69
động từ nhị xuyên 248, 257, 269
động từ phụ 25
động từ phụ (auxiliary verb) 132
động từ xuyên tính 248, 251, 257, 263, 269
động từ xuyên tính (transitive verb) 128
đơn vị cú pháp 76
đơn vị cú pháp (syntax unit) nhỏ nhất 68
đơn vị cú pháp nhỏ nhất 99
đơn vị ký tự có ý nghĩa 75
đơn vị từ vựng (lexical item) 266
đúng (true) 19

F

Ferdinand de Saussure 4, 9

G

giới hạn chọn lựa (selectional restriction) 258, 269

giới hạn tiểu phạm trù hóa
(*subcategorization restriction*) 258
giới ngữ 113
giới từ 68, 101, 102, 110, 113, 115, 136
Giới từ 103, 136
Gleason 4
Glinz 89
Goethe 39, 41
grammatical competence 11

H

Habermas 11
Haegeman 261
Hai tham số 248
hạn chế cục bộ (locality constrain) 169
Halliday 3
hàm chọn lựa (*selecting function*) 19
hàm đặc tính (*characteristic function*) 19
hàm số 18, 247, 248, 260, 269
hành động ngôn ngữ 52
Hậu tố (postfix) 68
Head-Driven Phrase Structure Grammar
(*HPSG*) 16
hệ chuyên gia (*expert system*) 8
hệ hình 90
hệ nhận diện ngôn ngữ 70
hệ nhận diện ngôn ngữ (*language*
recognizing system) 8
hệ thống ký hiệu tự tạo 2
hệ thống siêu cá thể 10
hình học của cây (*geometry of tree*) 143
hình thái học (morphology) 75
hình tự (ideo-graph) 67
hình vị 20, 69
Hjelmslev 10
Hoa Đào Thi 45, 51
hóa trị (valence) 248
Hồ Xuân Hương 6, 42, 46, 51, 54
hồi quy 177
hồi quy (recursive) 137
Hymes 11

I

Internet 8

J

Jackendoff 173

K

Kaplan 16
khả năng ngôn ngữ 10
khoa học nhận thức (cognitive science)
15, 265
không tận cùng (*non-terminal*) 111
khung tiểu phạm trù hóa (*subcategorisation*
frame) 263
kiểm soát 82, 83
kiểm tra danh cách 308
Kiều 6
Kim Vân Kiều 48, 56
Koiné 6
ký tự 67
ký tự (*character*) 10, 67

L

La Dictionarium Annamaticum Lusitium et
Latinum 7
langage 10
langue 10
Leont'ev 5, 9
Lewandowski 4, 76, 102
Lê Công Liệu 47
Lê Lợi 48
Lê Quý Đôn 42, 47, 52, 54
lệ thuộc cấu trúc (*structure dependence*) 3
Lê Văn Lý 13
Lê Văn Đặng 44, 51
lệnh-C (c-command) 150
Lệnh-C bất đối xứng (asymetric c-
command) 151
Lệnh-C đối xứng (symetric c-command)
151
liên từ 53, 68, 91, 103, 136
loại từ vựng (*lexical type*) 100
Lọc cách (Case Filter) 295
logic mờ (*fuzzy logic*) 15
lồng bỏ thể (expletive insertion) 267
lựa chọn 121
lược từ 53

lượng 86
lưới theta (theta grid) 261
 Lớp chỉ định từ (*determiner class*) 102
 Lớp danh từ (*noun class*) 102
 Lớp động từ (*verb class*) 102
 Lớp tính từ (*adjective class*) 102
 lớp từ 110
Lớp từ (word class) 103
 lớp từ tương đương 102
Lớp tương đương (equivalent class) 102
 Lục Vân Tiên 6
 Lyons 2, 3
 Lý Thường Kiệt 42, 43
 Lý thuyết chi phối và ràng buộc
 (*Government and Binding Theory (GB)*)
 16
Lý thuyết chi phối và ràng buộc
 (*Government and Binding Theory*) 151
Lý thuyết ràng buộc (binding theory) 150
 Lý thuyết tiêu chuẩn (*Standard Theory*) 16
 Lý thuyết tiêu chuẩn mở rộng (*Extended*
Standard Theory) 16
Lý thuyết X-gạch (X-Bar Theory) 173

M

Marantz 261
 Maurice Durand 46
 mã nhị phân (*binary code*) 72
 mạo từ 83, 101, 102, 103, 104, 105, 106,
 107, 117, 118, 120, 136
mạo từ bất định (non-specified article)
 104
mạo từ xác định (specified article) 104
 mẫu (*pattern*) 75
 mẫu cấu trúc 12
mẫu tự (alphabet) 67
 mẫu-trùng hợp (*pattern-matching*) 72
 mệnh đề (*clause*) 223
 mệnh đề chính (*main clause*) 61, 223
mệnh đề lồng (embedded clause) 223
 mệnh đề phụ 95, 222
 mệnh đề phụ (*subordinate clause*) 39, 61,
 223
Mô tả cấu trúc (structure description)
 113
Một tham số 248

mục tiêu (goal) 259

N

Nam Quốc Sơn Hà 43
NegP 283
N-gạch (N-bar) 178
ngịch đảo chủ ngữ/động từ phụ
 (*subject/aux inversion*) 230
 ngoại chủ thể 54
 ngoại chủ thể (*external subject*) 33
 ngôn ngữ 1
 ngôn ngữ cơ thể (*body language*) 1
 ngôn ngữ con người 2
 ngôn ngữ con người (*human language*) 67
 ngôn ngữ hình thức (*formal language*) 1,
 143
 ngôn ngữ học điện toán 7
ngôn ngữ học điện toán (computational
linguistics) 75
 ngôn ngữ học sinh học 9
 ngôn ngữ học tâm lý 9
 ngôn ngữ học xã hội 9
 ngôn ngữ lập trình (*programming language*)
 1
 ngôn ngữ lập trình cao 7
 ngôn ngữ máy (*machine language*) 7
ngôn ngữ nghịch đảo (inversion language)
 23
ngôn ngữ ngoại hóa 12
ngôn ngữ nội hóa 12
 ngôn ngữ toán 2
 ngôn ngữ tự nhiên (*natural language*) 1, 67
 ngôn ngữ tuyến tính 21
 ngôn ngữ vũ trụ (*universal language*) 135
nguồn gốc (source) 259
 Nguyễn Công Trứ 42, 49
 Nguyễn Du 6, 42, 47, 56
 Nguyễn Kim Thản 14
 Nguyễn Đình Chiểu 6
Nguyên tắc bình diện 176
Nguyên tắc chiếu mở rộng (Extended
Projection Principle (EPP)) 267
Nguyên tắc di truyền của đầu trong ngữ
đoạn 176
Nguyên tắc đầu (head) X-gạch 174
Nguyên tắc ngữ đoạn 176

Nguyên tắc ràng buộc A (binding principle A) 167
Nguyên tắc ràng buộc A (mở rộng) 169
 Nguyên tắc và thông số (Principles and Parameters (P&P)) 16
 nguyên tử 112
 Nguyễn Trãi 6, 42, 44, 45, 51
 ngữ âm 6
ngữ âm học (phonetics) 75
 ngữ cảnh 11
 ngữ cảnh (context) 80
 ngữ nghĩa 6
ngữ nghĩa học (semantics) 8, 75
Ngữ đoạn (phrase) 110
ngữ đoạn chỉ định từ (determiner phrase) 219
Ngữ đoạn danh từ 113
 ngữ đoạn danh từ (noun phrase) 111
 ngữ đoạn danh từ NP 112, 116, 123, 128, 130, 132, 178, 272, 290, 292, 295, 296, 297, 299, 300, 304, 310
Ngữ đoạn giới từ 113
 ngữ đoạn giới từ (prepositional phrase) 40
 ngữ đoạn giới từ PP 186, 246, 263, 264
 Ngữ đoạn giới từ PP 127
Ngữ đoạn động từ 113
 ngữ đoạn động từ VP 128, 132, 182
 ngữ đoạn tác tử bổ ngữ hóa 222
ngữ đoạn tác tử bổ ngữ hóa (complementizer phrase) 228
ngữ đoạn tác tử bổ ngữ hóa (complementizer phrase) CP 276
 Ngữ đoạn tác tử bổ ngữ hóa CP 228
 Ngữ đoạn thì 222
ngữ đoạn thì (tense phrase) 240
ngữ đoạn thì (tense phrase) TP 276
Ngữ đoạn tính từ 113
 ngữ đoạn tính từ AP 124, 185
 Ngữ đoạn tính từ AP 124
 ngữ pháp 6, 76
ngữ pháp biến hình tạo sinh (Generative Transformation Grammar) 10
 Ngữ pháp biến hình tạo sinh (Generative Transformation Grammar) 16
Ngữ pháp cấu trúc ngữ đoạn (phrase structure grammar) 113

Ngữ pháp cấu trúc ngữ đoạn tổng quát hóa (Generalized Phrase Structure Grammar (GPSG)) 16
Ngữ pháp cảm ngữ cảnh (context sensitive grammar) 252
Ngữ pháp chính quy (regular grammar) 252
 Ngữ pháp chức năng từ vựng (Lexical Functional Grammar) 16
 ngữ pháp chuẩn hình (well-formed grammar) 31
 ngữ pháp học 15
Ngữ pháp phi ngữ cảnh (context free grammar) 252
 ngữ pháp tạo sinh (Generative Grammar) 16, 113
Ngữ pháp vũ trụ (universal grammar) 252, 253
 người nói/nghe lý tưởng 10, 65
 nhãn hiệu 123
Nhãn hiệu (label) 144
 nhánh 111
Nhánh (branch) 144
 nhìn trước (look ahead) 74
 nhân cách hóa 80
 nhận dạng từ 75
 nhận diện → nhận thức 18
 nói tắt 30
nút (node) 110
Nút (node) 144
Nút chị em (sisters) 146
Nút con (daughter node) 145
Nút không tận cùng (non-terminal) 144
Nút mẹ (mother node) 145
Nút đỉnh (root node) 144
Nút tận cùng (terminal node) 144

O

Object-Verb-Subject (OVS) 277

P

performance 10
 P-gạch 186
P-gạch (P-bar) 178
 phạm trù cú pháp 109

phạm trù cú pháp (syntax category) 101, 110
phạm trù gạch (bar-category) 173
phạm trù ngữ pháp (grammatical category) 101
phạm trù từ vựng (lexical category) 101
phạm vi ràng buộc (binding domain) 169
 Phan Khôi 70, 107, 224
 phái sinh 115, 282, 295
 phái tính 20, 86, 253
phân bố bổ sung (complementary distribution) 218
 phần diễn 38
phân loại (taxonomy) 12
 phần đề 38
phân đoạn (segment) 89
phân tích cú pháp (syntax analyse) 70
Phân tích cú pháp (syntax analyse) 71
Phân tích ngữ nghĩa (semantic analyse) 71
phân tích tín hiệu 71
phân tích tín hiệu (signal analyse) 70
Phân tích tín hiệu (signal analyse) 71
phần tính (computational component) 265
 phần tử ảnh 18
 phần tử gốc 18
 phần tử **đối lập** 90
phép áp 18
 Phép biến hình 95
 Phép biến hình (transformation) 89
 Phép bố trí 91
 Phép bố trí (coordination) 89
phép chiếu (projection) 178
phép chiếu phủ định (negation projection) 283
 Phép hoán vị 94
 Phép hoán vị (permutation) 89
 Phép tính lược 93
 Phép tính lược (reduction) 89
phép thế (replacement, substitution) 178
 Phép thế (substitution) 89
phép thế-“do so 183
phép thế-“one 179
phép thế-“so 186
 Phép thử câu hỏi 91
 Phép thử câu hỏi (question test) 89

phi bản năng 2
 phi tự nhiên (non-natural) 3
phụ ngữ (adjunct) 52, 85, 193
phương pháp truyền thông 2
P-marker 143
 Pollard 16
 Pollock 242, 283
pragmatic 11
pragmatical competence 11
 PROLOG 72

Q

quan hệ 247, 248, 258, 260, 268, 269
quan hệ chủ đề (thematic relation) 258
Quan hệ chủ động-bị động (active-passive relation) 96
quan hệ lệ thuộc 97, 110
 QUAN HỆ LỆ THUỘC 81
Quan hệ lệ thuộc (dependency) 82
quan hệ lệ thuộc của từ 82
quét (scan) 74
 Quốc Âm Thi Tập 44
 quy tắc biến hình 271, 275, 278, 279, 280, 301, 309, 310
quy tắc biến hình (transformation rule) 22
quy tắc cấu trúc ngữ đoạn (phrase structure rule) 112
quy tắc chuyển vị (movement rules) 278
quy tắc suy diễn (deduction rule) 115
Quy tắc lên cấp động từ (verb raising) V → T 280
quy tắc lồng (insertion rules) 278, 289
quy tắc lồng từ vựng (lexical insertion rule) 262
quy tắc phạm trù (categorical rule) 262
quy tắc sản sinh (production rule) 112, 143
quy tắc tiểu phạm trù phi ngữ cảnh (context free subcategory rule) 255

R

Radford 175, 271, 272
Ràng buộc (bind) 167
 Ries 76
 Robins 4, 6

S

Sag 16
 sai (false) 19
 Sanskrit 6
 Sapir 2
 Schiller 41
 Schützenberger 113
 Seidel 76
 SELECT 72
 siêu ngôn ngữ (metalanguage) 80
 sinh cách (genitive) 86
 sinh cách tự do (free genitive) 216
sinh cách- "s" 216
sinh cách- "of" 216
 sinh hoạt ngôn ngữ 57
 sinh học 8
 số lượng 253
 sở hữu 20
 SQL 72
 Stechow 12, 176
 Subject-Object-Verb (SOV) 222
 Subject-Verb-Object (SVO) 222, 276, 309
sự liên đới (interdependency) 88

T

tác nhân 60
tác nhân (actor) 36
Tác nhân (agent, actor) 258
tác tử bổ ngữ hóa (complementizer) 134
 tác tử bổ ngữ hóa C 286, 287, 306
 tặng cách 253, 254
 tặng cách (dative) 86
Tặng cách (dative) 294
 tâm lý học 7, 8
 tận cùng (terminal) 111
 tập hợp gốc 18
 thành thạo ngôn ngữ 77
 thành tố 82, 90, 95, 99
Thành tố 109
thành tố (constituent) 75, 147
 thành tố cùng loại 90
Thành tố của (constituent of) 147
Thành tố trực tiếp của (immediate constituent of) 147
 thành tố zero 54

thái bị động (passive voice) 32, 40
 thái chủ động (active voice) 32, 40
thẩm năng ngữ dụng 11
thẩm năng giao lưu 11
 thẩm năng ngôn ngữ 1, 77
thẩm năng ngôn ngữ (language competence) 10
 thể cách 86, 253
Thế tiền vị (precedence) 148
Thế tiền vị trực tiếp (immediate precedence) 149
 thì (tense) 25
 thì phi hạn T (non-finite T) 296
 thì T hữu hạn (finite T) 295
 thiết lập trình biên dịch (compiler design) 143
 thiếu chủ ngữ 30
 thông số 12
thông số (parameter) 222
 Thông số và thứ tự của Từ 221
 thực thể đếm được 108
 thuần cấu trúc 3
 thuần cảm giác 78
 thuần trực giác 14
 thuần túy người 1
thụ nhân (recipient) 259
thuyết ràng buộc (binding theory) 163
 tiền danh từ 107
Tiền ngữ (antecedent) 164
Tiền tố (prefix) 68
 tiếng 70
 tiếng (sound) 72
 tiếng La tinh 42
 tiếng địa phương 10
 tiếng Phạn 6, 42
 Tiếng Việt ngày nay 57
 Tiếng Việt xưa 43
 tiếng xã hội 10
tiểu phạm trù (subcategory) 250
 tiểu phạm trù hóa 273, 274
 tiểu phạm trù hóa (subcategorised) 263
tiểu phạm trù phi ngữ cảnh (context free category) 254
 tiểu tập hợp hóa 28
 tìm 72
 tín hiệu đầu vào (input signal) 71
 tin học 7

Tính cảm đề 32, 54
tính chi phối (dominance) 144
 Tính giàu thông tin cụ thể 52
 Tính hồi quy 137
Tính nghiêm ngặt của tiểu phạm trù 256
tính ngữ 113
tính ưu thế 144
 tính tỉnh lược 29, 30, 53
 Tính tịnh tiến 21, 50
 tính từ 68, 69, 78, 79, 80, 82, 83, 84, 86, 87, 88, 98, 101, 102, 103, 104, 106, 107, 109, 110, 112, 113, 115, 121, 123, 124, 125, 126, 127, 128, 136, 142
Tính từ kiểm soát tính từ 84
Tính từ kiểm soát trạng từ 84
 tính tuyến tính (linearity) 82
 Tối thiểu luận (Minimalism (MP)) 16
 tôn ti (hierarchy) 82
tôn ty Chomsky (Chomsky's hierarchy) 252
 tự 70
 tự nhiên 3
tịnh tiến 26, 127
 Từ 41, 67
từ (word) 68, 101
 từ dưới lên trên (bottom up) 151
 từ ghép 28
 từ láy 28
 từ nghi vấn 21, 277, 303, 310
từ đơn (single word) 68
 từ trên xuống dưới (top down) 156
 từ vựng (lexical) 68, 246, 265
 Trần Tế Xương 42, 49
 Trần Trọng Kim 42, 48
 trạng từ 82, 101, 102, 103, 113, 116, 124, 125, 126, 128, 129, 130, 131, 136, 142
 Travis 222, 345
 trí thông minh nhân tạo (artificial intelligence) 8, 155
 triết học 8
 trình biên dịch 115
trống 228
 trống (NULL) 72
 trung tâm ý nghĩa 82, 109
 trùng hợp 72

Trùng ngữ (anaphor) 164
tuyến chiếu (projection line) 175
tuyến tính 18

U

UNICODE 73

V

vai trò theta 290, 291, 294, 297, 304, 306
vai trò theta (θ role) 260
vai trò theta (theta role) 258
Vai trò theta ngoài 261
Vai trò theta trong 261
 V-gạch 182
V-gạch (V-bar) 178
 văn bản thánh 6
 văn nói 42, 65
 văn phạm 76
 văn viết 42, 65
 vận dụng thẩm năng 58
 vật lý 7
 vị ngữ 53, 132
vị ngữ (predicate phrase) 223
 vị từ 32
vị trí (location) 259, 264
 vị trí từ 82, 97
 Vịnh Bánh Trôi 46

W

Wagner 76

X

xã hội học 8
 X-gạch hạn chế (constrain X-bar) 246
xuống cấp động từ (verb lowering) 282

Z

zero 228, 240