

Thiết kế tập nhãn cú pháp và hướng dẫn gán nhãn

Nguyễn Phương Thái, Vũ Xuân Lương, Nguyễn Thị Minh Huyền

Đào Minh Thu, Đào Thị Minh Ngọc, Lê Kim Ngân

SP 7.3 – Dự án VLSP

Giới thiệu

Đây là tài liệu hướng dẫn gán nhãn cú pháp cho treebank tiếng Việt. Tập nhãn từ loại và hướng dẫn gán nhãn từ loại được trình bày trong một tài liệu khác. Với mỗi hiện tượng ngữ pháp, chúng tôi trình bày cách nhận diện, cách gán nhãn, cùng với các ví dụ cụ thể để minh họa. Các ví dụ được lấy từ sách ngữ pháp hoặc từ ngữ liệu thực tế. Chúng tôi cố gắng trích dẫn tài liệu tham khảo đầy đủ để người đọc có thể tự tìm hiểu thêm khi cần. Tài liệu này liên tục được chỉnh sửa và bổ sung trong quá trình thực hiện dự án.

Mục lục

1. Toàn bộ tập nhãn.....	3
2. Chú ý chung khi gán nhãn cụm từ.....	6
2.1 Nhãn phần tử trung tâm H.....	6
2.2 Cấu trúc với liên từ đẳng lập.....	8
3. Cụm danh từ.....	8
3.1 Cấu trúc chung.....	9
3.2 Nhập nhằng gán nhãn cụm danh từ cơ sở.....	10
4. Cụm động từ.....	11
5. Cụm tính từ.....	13
6. Cụm phó từ.....	14
7. Cụm giới từ.....	14
8. Cụm từ chỉ số lượng.....	15
9. Ngữ tình thái.....	15
10. Câu trần thuật.....	16
11. Mệnh đề phụ kết (subordinate clause).....	18
12. Câu hỏi.....	19
13. Câu cảm thán.....	22
14. Câu mệnh lệnh.....	23
15. Câu đặc biệt và tit báo.....	23
16. Các nhãn chức năng.....	24
16.1 Nhãn chức năng chủ ngữ.....	24
16.2 Nhãn chức năng tân ngữ trực tiếp.....	25
16.3 Nhãn chức năng tân ngữ gián tiếp.....	26
16.4 Nhãn chức năng khởi ngữ.....	27
16.5 Nhãn chức năng dành cho vị ngữ không phải cụm động từ.....	29
16.6 Nhãn chức năng của chủ ngữ logic.....	29
16.7 Nhãn chức năng bổ ngữ chỉ phạm vi hay tần suất của hành động.....	30
16.8 Thành phần than gọi.....	30
17. Nhãn phân loại phụ ngữ của động từ.....	31
17.1 Phụ ngữ thời gian.....	31
17.2 Phụ ngữ nơi chốn.....	32
17.3 Phụ ngữ chỉ hướng.....	32

17.4	Phụ ngữ chỉ cách thức hay phương tiện	32
17.5	Phụ ngữ chỉ mục đích hay lý do	33
17.6	Phụ ngữ chỉ điều kiện	33
17.7	Trạng ngữ chỉ ý nhượng bộ	34
17.8	Trạng ngữ	34
18.	Nhãn phần tử rỗng	34
18.1	Nhãn *T*	34
18.2	Nhãn *0*	36
18.3	Nhãn *RNR*	36
19.	Các cấu trúc sử dụng liên từ đẳng lập	37
19.1	Các từ trung tâm không có chung bổ ngữ	37
19.2	Các từ (hay cụm từ) có chung bổ ngữ	38
20.	Kết từ đẳng lập (C) và kết từ chính phụ (E)	38
21.	Thành phần chú thích hoặc trích dẫn	39
21.1	Thành phần chú thích	39
21.2	Thành phần trích dẫn	40
22.	Câu ghép	40
22.1	Câu ghép song song (UBKHXX, 1983)	41
22.2	Câu ghép qua lại (UBKHXX, 1983)	41
22.3	Phân biệt câu ghép với câu đơn có thành phần trạng ngữ	42
23.	Tinh lược	43
24.	Câu bị động	50

1. Toàn bộ tập nhãn

Nhãn từ loại:

STT	Tên	Chú thích
	N	Danh từ
	Np	Danh từ riêng

	Nc	Danh từ chỉ loại
	Nu	Danh từ đơn vị ¹
	V	Động từ
	A	Tính từ
	P	Đại từ
	L	Định từ ²
	M	Số từ
	R	Phụ từ
	E	Giới từ ³
	C	Liên từ
	I	Thán từ
	T	Trợ từ, tiểu từ, từ tình thái ⁴
	U	Từ đơn lẻ
	Y	Từ viết tắt
	X	Các từ không phân loại được

Khi gán nhãn, nhãn từ viết tắt sẽ là nhãn kép. Chẳng hạn nếu từ viết tắt là HIV thì nhãn của nó là Ny vì HIV viết đầy đủ thì là danh từ.

Nhãn cụm từ:

STT	Tên	Chú thích
	NP	Cụm danh từ
	VP	Cụm động từ
	AP	Cụm tính từ
	RP	Cụm phụ từ
	PP	Cụm giới từ
	QP	Cụm từ chỉ số lượng
	MDP	Cụm từ tình thái
	UCP	Cụm từ gồm hai hay nhiều thành phần không cùng loại được nối với nhau bằng liên từ đẳng lập
	LST	Cụm từ đánh dấu đầu mục của danh sách

¹ Mới bổ xung vào tháng 5/2008

² Trong tài liệu hướng dẫn cũ là D

³ Trong từ điển (SP7.2) là O

⁴ Trong từ điển (SP7.2) phân ra hai loại là trợ từ (T) và cảm từ (E)

	WHNP	Cụm danh từ nghi vấn (ai, cái gì, con gì, v.v.)
	WHAP	Cụm tính từ nghi vấn (lạnh thế nào, đẹp ra sao, v.v.)
	WHRP	Cụm từ nghi vấn dùng khi hỏi về thời gian, nơi chốn, v.v.
	WHPP	Cụm giới từ nghi vấn (với ai, bằng cách nào, v.v.)

Nhãn mệnh đề:

STT	Tên	Chú thích
	S	Câu trần thuật (khẳng định hoặc phủ định)
	SQ	Câu hỏi
	S-EXC	Câu cảm thán
	S-CMD	Câu mệnh lệnh
	SBAR	Mệnh đề phụ kết (bổ nghĩa cho danh từ, động từ, và tính từ)
	SF	Câu mà chỉ có thể được giải thích hợp lý dưới quan điểm ngữ pháp chức năng

Nhãn chức năng cú pháp:

STT	Tên	Chú thích
	H	Phần tử trung tâm của cụm từ (hay mệnh đề)
	SUB	Nhãn chức năng chủ ngữ
	DOB	Nhãn chức năng tân ngữ trực tiếp
	IOB	Nhãn chức năng tân ngữ gián tiếp
	TPC	Nhãn chức năng chủ đề (khởi ngữ)
	PRD	Nhãn chức năng vị ngữ không phải cụm động từ
	LGS	Nhãn chức năng chủ ngữ logic của câu ở thể bị động
	EXT	Nhãn chức năng bổ ngữ chỉ phạm vi hay tần suất của hành động
	VOC	Nhãn chức năng thành phần than gọi
	TH	Nhãn phần thuyết của câu SF

Nhãn phân loại phụ ngữ của động từ:

STT	Tên	Chú thích
	TMP	Nhãn chức năng phụ ngữ chỉ thời gian
	LOC	Nhãn chức năng phụ ngữ chỉ nơi chốn
	DIR	Nhãn chức năng phụ ngữ chỉ hướng
	MNR	Nhãn chức năng phụ ngữ chỉ cách thức

	PRP	Nhãn chức năng phụ ngữ chỉ mục đích hay lý do
	CND	Nhãn chức năng phụ ngữ chỉ điều kiện
	CNC	Nhãn chức năng phụ ngữ chỉ ý nhượng bộ
	ADV	Nhãn chức năng trạng ngữ (khi không sử dụng được một trong các loại cụ thể trên)

Các nhãn khác:

STT	Tên	Chú thích
	T	Nhãn phần tử rỗng (lưu vết trong phạm vi câu)
	E	Nhãn phần tử rỗng ứng với hiện tượng tỉnh lược
	0	Nhãn phần tử rỗng ở vị trí tác tử phụ ngữ hóa

Các nhãn quy ước trong tài liệu này:

STT	Tên	Chú thích
	.	Nhãn dấu chấm câu, bao gồm: . ? !
	,	Nhãn dấu phẩy
	÷	Nhãn dùng cho cả dấu hai chấm và dấu gạch ngang chú thích

2. Chú ý chung khi gán nhãn cụm từ

2.1 Nhãn phần tử trung tâm H

Mô tả: Phần tử trung tâm của một cụm từ (ngữ đoạn) có các thuộc tính sau (CXHạo, 2007):

- Nó là yếu tố mang tất cả các thuộc tính ngữ pháp của ngữ đoạn
- Nó là yếu tố duy nhất của ngữ đoạn có thể có một quan hệ ngữ pháp và ngữ nghĩa vượt ra ngoài ngữ đoạn
- Các yếu tố khác của ngữ đoạn chỉ có quan hệ phụ thuộc trực tiếp hay gián tiếp với trung tâm ngữ đoạn mà thôi (chứ không có bất cứ quan hệ gì ra ngoài phạm vi ngữ đoạn)

Cụm từ có một từ trung tâm:

(NP (N-H cô gái) (A đẹp))

(NP (N-H võ) (N dân tộc))

(NP (N-H nước) (Np Việt Nam))

Cụm từ có nhiều từ trung tâm:

(NP (NP-H (N-H chén))

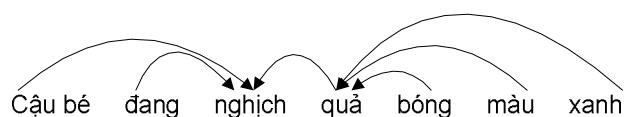
(C và)

(NP-H (N-H đĩa) (A sạch)))

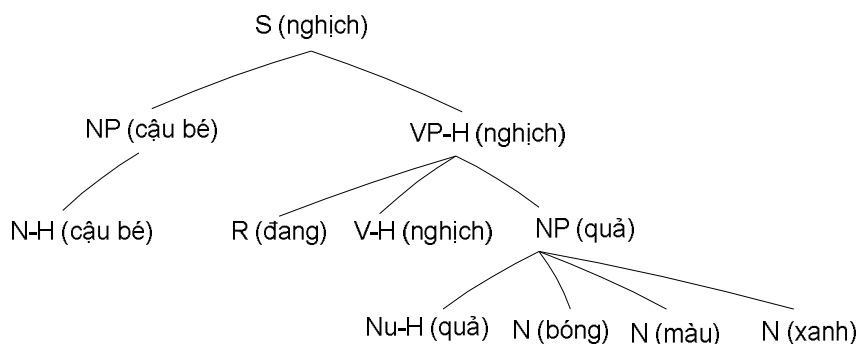
Cụm từ có nhiều từ trung tâm chung bổ ngữ:

(NP (N-H chén) (C và) (N-H đĩa) (A sạch))

Hình 1 thể hiện một cây phụ thuộc trong đó các mũi tên đi từ từ *bổ nghĩa* đến từ *trung tâm*. Xét cụm danh từ “quả bóng màu xanh”, ta dễ dàng thấy là chỉ từ trung tâm “quả” là có liên hệ ra bên ngoài mà cụ thể là với động từ “nghịch”.



Hình 1. Cây phụ thuộc



Hình 2. Cây được từ vệt hóa

Có sự liên hệ chặt chẽ giữa cây phụ thuộc và cây cú pháp thành phần. Hình 2 biểu diễn cây cú pháp thành phần ứng với cây phụ thuộc trong Hình 1.

Thủ thuật xác định phần tử trung tâm (ĐMThu&NPThái, 2008):

- Lược: trung tâm là yếu tố mà nếu lược bỏ nó sẽ làm cho nghĩa và thuộc tính ngữ pháp của ngữ đoạn thay đổi
- Mở rộng văn cảnh: để xác định xem phần tử nào trong ngữ đoạn có quan hệ với văn

cảnh

- Đặt câu hỏi: thay yếu tố phụ bằng từ nghi vấn
- Dựa vào khuôn hình đã được tổng kết: nhận diện nhanh phần từ trung tâm dựa vào kinh nghiệm
- Chèn: để xác định xem có một hay nhiều thành tố trung tâm
- Qui tắc chính trước phụ sau: để nhận diện nhanh ứng cử viên

2.2 Cấu trúc với liên từ đẳng lập

a) Tính đối xứng

Khi đã có liên từ đẳng lập thì một cụm từ hay mệnh đề sẽ có nhiều phần từ trung tâm. Cần chú ý tính đối xứng của một cấu trúc có liên từ nối. Chẳng hạn như:

N C N

NP C NP

Chứ không nên:

N C NP

Hay NP C N

b) Nhân UCP

Trường hợp liên từ đẳng lập được dùng để nối hai từ (hay cụm từ) không cùng loại, ta có thể dùng nhân UCP.

(NP sản phẩm

(UCP (AP rẻ)

và

(NP chất lượng tốt)))

Không nhất thiết phải luôn dùng UCP trong mọi trường hợp, chẳng hạn như khi câu có hai vị ngữ, một cái là cụm tính từ và cái còn lại là cụm động từ. Tuy nhiên với những trường hợp tương tự ví dụ trên, việc dùng UCP là bắt buộc vì nếu không nhóm các thành phần đó lại sẽ gây nhập nhằng cấu trúc.

Ta sẽ quay lại nội dung liên quan đến liên từ đẳng lập trong một phần khác của tài liệu.

3. Cụm danh từ

Ký hiệu: NP

3.1 Cấu trúc chung

Cấu trúc cơ bản của một cụm danh từ như sau [1, trg24]:

<phần phụ trước> <danh từ trung tâm> <phần phụ sau>

Ví dụ: “một mái tóc đẹp” thì danh từ chỉ loại “mái” là trung tâm, “một” là phần phụ trước, còn danh từ “tóc” và tính từ “đẹp” thuộc phần phụ sau.

(NP (M một) (Nc-H mái) (N tóc) (J đẹp))

Một cụm danh từ có thể thiếu phần phụ trước hay phần phụ sau nhưng không thể thiếu phần trung tâm.

Phần phụ trước:

Phần này có tối đa hai thành phần⁵:

<vị trí -2> <vị trí -1>

Ví dụ: “tất cả những chiếc kẹo”

(NP (L tất cả)

(L những)

(Nc-H chiếc)

(N kẹo))

Ở vị trí -2 là định từ chỉ tổng lượng như “tất cả”, “hết thảy”, v.v. Ở vị trí -1 là số từ (hoặc cụm số từ) và định từ. Chi tiết cấu tạo từng thành phần xin tham khảo thêm trong [1, trg45].

Phần phụ sau:

Nói chung phần phụ sau của cụm danh từ có cấu tạo phức tạp hơn phần phụ trước nhiều. Bổ ngữ sau có thể là danh từ, cụm tính từ, cụm động từ, số từ xác định và số từ thứ tự, đại từ chỉ định, cụm giới từ, hay mệnh đề phụ. Đại từ chỉ định, nếu có, thì thường được đặt sau cùng. Sau đây là một số ví dụ:

Ví dụ 1: Cụm danh từ đơn giản (không có bổ ngữ là cụm giới từ, cụm động từ, hay mệnh đề phụ):

quả bóng màu xanh

(NP (Nc-H quả)

(N bóng)

⁵ Cũng có quan điểm coi danh từ chỉ loại thuộc phần phụ trước, như vậy sẽ có tối đa ba thành phần

(N màu)
(N xanh))

Ví dụ 2: Cụm danh từ phức tạp với bổ ngữ sau là cụm giới từ⁶:

cái máy tính của cơ quan

(NP (Nu-H cái)

(N máy tính)

(PP của cơ quan))

Ví dụ 3: Cụm danh từ phức tạp với bổ ngữ sau là mệnh đề phụ:

cái máy tính mà tôi mới mua hôm qua

(NP (Nu-H cái)

(N máy tính)

(SBAR mà tôi mới mua hôm qua))

// Nên trao đổi về nhập nhằng khi gán nhãn NP ở đây

3.2 **Nhập nhằng** gán nhãn cụm danh từ cơ sở

Comment [t1]: Đưa ra “thủ thuật”

Xét hai cụm từ “quả bóng màu xanh” và “bản đồ hàng lậu”:

(NP (Nc-H quả)

(N bóng)

(N màu)

(N xanh))

(NP (N-H bản đồ)

(NP (N-H hàng) (A lậu)))

Tại sao “quả bóng màu xanh” lại có cấu trúc phẳng hơn “bản đồ hàng lậu”? Nói cụ thể hơn, tại sao “màu xanh” không được nhóm lại thành NP trong khi “hàng lậu” lại được? Lý do là vì cả “quả bóng xanh” và “quả bóng màu” đều chấp nhận được, do đó cả “màu” lẫn “xanh” đều có liên hệ với văn cảnh. Trong khi đó ta không thể nói “bản đồ lậu” vì nghĩa bị thay đổi. Như vậy nếu coi “hàng lậu” là cụm danh từ thì đảm bảo yêu cầu là chỉ có từ trung tâm (“hàng”) có liên hệ với văn cảnh.

⁶ Để đơn giản chúng tôi chưa mô tả cấu trúc cụ thể của PP và SBAR, chỉ nêu cụm từ tiếng Việt

4. Cụm động từ

Ký hiệu: VP

Cấu trúc chung:

Giống như cụm danh từ, cấu tạo một cụm động từ về cơ bản như sau:

<phần phụ trước> <động từ trung tâm> <phần phụ sau>

Phần phụ trước:

Phần phụ trước của cụm động từ thường là phụ từ.

Ví dụ:

“đang ăn cơm”

(VP (R đang)

(V-H ăn)

(NP-DOB cơm))

Phần phụ sau:

Động từ có khả năng kết hợp với các từ loại khác một cách rất đa dạng. Mỗi cách kết hợp có thể coi như một mẫu cú pháp của động từ: nội động từ, động từ đi với danh từ, động từ đi với cụm giới từ, động từ đi với mệnh đề, v.v. Ta xét các ví dụ sau:

Nội động từ:

đi

(VP (V-H đi))

Bổ ngữ là cụm danh từ:

yêu cô ấy

(VP (V-H yêu)

(NP-DOB cô ấy))

Bổ ngữ là cụm giới từ:

thanh toán bằng tiền mặt

(VP thanh toán

(PP bằng

(NP tiền mặt)))

Bổ ngữ là hai cụm danh từ :

Nhóm động từ trao nhận (LBiên, 1999) thường có cấu trúc này.

tặng bạn hai quyển sách

(VP (V-H tặng)

(NP-IOB bạn)

(NP-DOB hai quyển sách)))

Bổ ngữ là cụm danh từ và cụm giới từ :

pha cả phê với sữa

(VP (V-H pha)

(NP-DOB cả phê)

(PP với

(NP sữa)))

Bổ ngữ là cụm động từ:

Nhóm động từ tình thái (LBiên, 1999) thường có cấu trúc này.

cần viết thư

(VP (V-H cần)

(VP viết

(NP-DOB thư)))

Bổ ngữ là cụm danh từ và cụm động từ :

Nhóm động từ gây khiến (LBiên, 1999) thường có cấu trúc này.

nhờ bạn chép bài

(VP nhờ

(NP-DOB bạn)

(VP chép

(NP-DOB bài)))

Bổ ngữ là mệnh đề:

Nhóm động từ cảm nghĩ trạng thái (LBiên, 1999) thường có cấu trúc này.

nói rằng cô ấy đẹp

(VP (V-H nói)

(SBAR rằng

(S (NP-SUB cô ấy)

(AP-PRD đẹp))))

Bổ ngữ là cụm tính từ:

Thường áp dụng cho các động từ chỉ quan hệ biến hóa (LBiên, 1999).

Cô ấy trở nên hung dữ.

Ngoài bổ ngữ, góp phần cấu tạo nên cụm động từ còn có phụ ngữ. Phụ ngữ có thể là cụm tính từ, cụm danh từ chỉ thời gian, cụm giới từ, hoặc mệnh đề phụ. Các ví dụ về phụ ngữ là cụm giới từ hoặc mệnh đề phụ xin xem trong phần 9 và 14.

Phụ ngữ là cụm tính từ :

đi nhanh

(VP (V-H đi)

(AP nhanh))

Cô ấy hát hay.

Nó vẽ đẹp.

5. Cụm tính từ

Ký hiệu: AP

Cấu trúc chung:

Cấu tạo một cụm tính từ về cơ bản như sau:

< phần phụ trước> <tính từ trung tâm> <phần phụ sau>

Phần phụ trước:

Phần phụ trước của tính từ thường là phụ từ chỉ mức độ.

Ví dụ:

rất đẹp

(AP (R rất)

(A-H đẹp))

Phần phụ sau:

Bổ ngữ sau có thể là phụ từ chỉ mức độ như trong ví dụ sau:

xinh quá

(AP (A-H xinh)

(R quá))

Bổ ngữ sau có thể là danh từ:

móng cùi

(AP (A-H móng)

(NP cùi))

Bổ ngữ sau có thể là cụm giới từ:

giới về thể thao

(AP (A-H giới)

(PP về

(NP thể thao)))

6. Cụm phó từ

Ký hiệu: RP

Cụm phó từ chủ yếu tạo bởi sự kết hợp giữa các phó từ với nhau. Ví dụ như :

vẫn chưa⁷

(RP (R vẫn) (R chưa))

7. Cụm giới từ

Ký hiệu: PP

Cấu trúc chung :

<giới từ> <cụm danh từ> hoặc <giới từ> <cụm động từ>

Ví dụ :

vào Sài Gòn (trong "đi vào Sài Gòn")

(PP (E-H vào)

(NP Sài Gòn))

⁷ Gán nhãn H ở đầu cũng được, chọn phần từ đầu cho tiện

để tìm cơ hội du học (trong “tôi hay vào mạng để tìm cơ hội du học”)
(PP (E-H để)
(VP tìm
(NP cơ hội du học)))

8. Cụm từ chỉ số lượng

Ký hiệu : QP

Cấu trúc chung :

Thành phần chính của QP là các số từ. Có thể là số từ xác định, số từ không xác định, hay phân số. Ngoài ra còn có thể có phụ từ như "khoảng", "hơn", v.v. QP thường đóng vai trò là thành phần phụ trước của cụm danh từ (vị trí -1).

Ví dụ 1:

năm trăm
(QP (M-H năm) (M trăm))

Ví dụ 2:

hơn 200
(QP (R hơn) (M-H 200))

9. Ngữ tình thái

Ký hiệu: MDP (modal phrase)

Mô tả: Tình thái ngữ là thành phần phụ của câu luôn đứng sau nòng cốt câu, có nhiệm vụ bổ sung những ý nghĩa về tình thái cho câu. Khi gắn nhãn cần chú ý phân biệt tình thái ngữ với: thán từ làm thành một vế của câu ghép, tiểu từ nhấn mạnh, và vị ngữ (NMThuyết và NVHiệp, 1999; tr 235-242).

Cô ta hồi hận thì có⁸.

(S (NP-SUB Cô ta)

⁸ Trong Từ điển Tiếng Việt, “thì có” được phân loại là *khẩu ngữ*.

(VP (V-H hỏi hận))
 (MDP (T thì có))
 (. .))

10. Câu trần thuật

Comment [t2]: Có thể thay bằng thuật ngữ khác không?

Ký hiệu : S

Cấu trúc chung :

Theo quan điểm coi cấu trúc chủ-vị là cấu trúc chủ đạo của câu tiếng Việt [1], câu trần thuật sẽ có cấu trúc sau:

<chủ ngữ> <vị ngữ>

Trong đó chủ ngữ thường là cụm danh từ, còn vị ngữ thường là cụm động từ hoặc cụm tính từ. Với một số ngôn ngữ như tiếng Anh, vị ngữ luôn là cụm động từ.

Ví dụ :

Anh yêu em .
 (S (NP-SUB Anh)
 (VP (V yêu)
 (NP-DOB em))
 (. .))

Nhãn chức năng chủ từ cho ta biết đâu là chủ ngữ của câu (chủ ngữ bề mặt). Cụm động từ theo sau chủ từ sẽ là vị ngữ. Nếu vị ngữ không phải cụm động từ thì sẽ được gán nhãn chức năng PRD.

Chủ ngữ :

Chủ ngữ là cụm danh từ:

“Việc dậy đúng giờ thật khó.”
 (S (NP-SUB Việc
 (VP dậy
 (AP đúng
 (NP-TMP giờ))))
 (AP-PRD (thật khó))
 (. .))

Chủ ngữ là cụm động từ:

“Dậy đúng giờ thật khó .”

(S (VP-SUB dậy
(AP đúng
(NP-TMP giờ)))
(AP-PRD thật khó)
(. .))

Chủ ngữ cũng có thể là cụm chủ vị:

(S (S-SUB (NP-SUB Anh)
(VP nói
(NP-DOB thế)))
(AP-PRD không đúng đâu)
(. .))

Vị ngữ :

Vị ngữ là cụm động từ:

Tôi đi học .

(S (NP-SUB Tôi)
(VP đi
(VP học))
(. .))

Vị ngữ là cụm tính từ:

Nhà anh ấy xa .

(S (NP-SUB nhà anh ấy)
(AP-PRD xa)
(. .))

Vị ngữ cũng có thể là cụm danh từ:

Em bé 7 tuổi.

(S (NP-SUB em bé)
(NP-PRD 7 tuổi)
(. .))

Sự đa dạng trong cấu trúc của cụm động từ và cụm tính từ khiến cho cấu trúc của câu

trần thuật cũng rất đa dạng. Về các khuôn hình câu đơn⁹, DQBan (2005) đã mô tả khá đa dạng.

Câu với động từ “có”:

Có con chuột trong góc nhà.

(S (VP có

(NP-SUB con chuột

(PP-LOC trong

(NP góc nhà))))

(. .))

Động từ này đặc biệt ở chỗ nó đứng đầu câu và danh từ theo sau là chủ ngữ của câu.

Động từ này chỉ sự tồn tại.

// Bổ xung một số động từ tương tự: lấp lánh (?)

// Thêm mẫu câu (những cái chưa được nói đến ở phần cụm động từ)

11. Mệnh đề phụ kết (subordinate clause)

Ký hiệu : SBAR

Cấu trúc và chức năng:

Mệnh đề phụ kết đóng vai trò bổ nghĩa cho danh từ, động từ, hay tính từ. Bản thân nó không thể đứng độc lập làm thành một câu. Về cơ bản cấu trúc của mệnh đề phụ bao gồm một liên từ phụ kết và một mệnh đề (ký hiệu S).

Bổ nghĩa danh từ :

Quyển sách mà anh mượn

(NP (Nu-H Quyển)

(N sách)

(SBAR mà

(S (NP-SUB anh)

(VP mượn))))

Bổ nghĩa động từ:

không đi đá bóng vì bạn gái ốm

⁹ Chú ý là nhãn S không chỉ dùng cho câu đơn.

(VP (R không)
 (V-H đi)
 (VP đá
 (NP-DOB bóng))
 (SBAR-PRP vì
 (S (NP-SUB bạn gái)
 (AP-PRD ôm))))

Trong ví dụ này mệnh đề phụ kết "vì bạn gái ôm" chỉ nguyên nhân của hành động "không đi đá bóng", vì thế có thêm nhãn PRP.

Bổ nghĩa tính từ :

khỏe vì chơi thể thao đều đặn
 (AP (A-H khỏe)
 (SBAR vì
 (S (NP-SUB *T*)
 (VP chơi
 (NP-DOB thể thao)
 (AP đều đặn))))))

Vì nhóm dữ liệu sử dụng nhãn này không nhất quán lắm cho nên phần này sẽ nhấn mạnh các điểm cần chú ý cũng như qui ước cụ thể hơn cách dùng của nó.

- Chú ý 1: Nhãn này chỉ dùng để thể hiện mệnh đề tính ngữ (bổ nghĩa cho danh từ), mệnh đề trạng ngữ (bổ nghĩa cho động từ hoặc câu), hoặc mệnh đề danh ngữ (bổ ngữ của động từ, tính từ).
- Chú ý 2: Không dùng SBAR ngay sau giới từ trung tâm của cụm giới từ. Không gán nhãn SBAR cho câu đơn mà làm thành phần của một câu phức (phân biệt với trường hợp mệnh đề trạng ngữ).

12. Câu hỏi

Ký hiệu : SQ

Khi ta đã thành thạo việc gán nhãn câu trần thuật, việc gán nhãn cho câu hỏi sẽ trở nên đơn giản hơn. Ta xem xét các dạng câu hỏi chính dưới đây :

Câu hỏi chuyên biệt (wh-question):

Loại câu hỏi này được dùng để hỏi về người, vật, địa điểm, thời gian, v.v.

Hỏi người, vật:

Ai đang ở trong nhà ?

(SQ (WHNP Ai)

(VP đang

ở

(PP trong

(NP nhà)))

(. ?))

Cụm danh từ nghi vấn (WHNP) được sử dụng trong loại câu hỏi này. Cụm danh từ nghi vấn có thể là một đại từ nghi vấn (ai) hoặc là một cụm danh từ có đại từ nghi vấn làm bộ ngữ sau (cái gì, con gì).

Hỏi thời gian:

Bao giờ anh đi hội nghị ?

(SQ (WHADV Bao giờ)

(S (NP anh)

(VP đi

(NP hội nghị)))

(. ?))

Hỏi cách thức:

Anh sẽ giải bài toán này bằng cách nào ?

(SQ (S (NP anh)

(VP sẽ

giải

(NP bài toán này)

(WHPP bằng

(WHNP cách nào))))

(. ?))

Cụm giới từ nghi vấn (WHPP) là do giới từ kết hợp với cụm danh từ nghi vấn tạo ra.

Hỏi về trạng thái:

Bàn tay của cô ấy mềm mại ra sao ?

(S (NP bàn tay

(PP của
 (NP cô ấy)))
 (WHAP mềm mại
 (P ra sao))
 (. ?))

Cụm tính từ nghi vấn (WHAP) là do tính từ kết hợp với đại từ nghi vấn tạo ra.

Câu hỏi có-không (yes-no question):

Loại sử dụng cặp phụ từ trái nghĩa “có... không”, “đã... chưa”, v.v.

Ví dụ:

Em có đi chơi không ?
 (SQ (NP-SUB em)
 (VP (R có)
 (V đi chơi)
 (R không))
 (. ?))

Cũng có thể chỉ sử dụng một phụ từ phủ định:

Ví dụ¹⁰:

Mai anh đi chưa ?
 (SQ (NP-TMP Mai)
 (NP-SUB anh)
 (VP (V đi)
 (R chưa))
 (. ?))

Hoặc dùng tiểu từ tình thái:

Ví dụ :

Cô ấy chưa về nhỉ ?
 (SQ (NP-SUB Cô ấy)
 (VP (R chưa)
 (V về))

¹⁰ Trong ví dụ này, TMP là nhân phụ ngữ chỉ thời gian.

(T nhỉ)

(. ?))

// Phụ từ thì ở mức VP còn tiểu từ thì mức câu (?)

13. Câu cảm thán

Ký hiệu : S-EXC

Cấu trúc chung :

Câu cảm thán dùng để thể hiện tình cảm ở một mức độ nhất định. Loại câu này cũng có những đặc trưng về mặt hình thức, chẳng hạn như sử dụng thán từ (ôi, ơi là), tiểu từ (thay), phó từ (lạ, thật), v.v.

Câu cảm thán sử dụng thán từ :

Ồi sức trẻ !

(S-EXC (T Ôi)

(NP sức trẻ)

(. !))

Câu cảm thán sử dụng tiểu từ "thay" :

Vinh quang thay những vị anh hùng dân tộc !

(S-EXC (AP-PRD Vinh quang)

(T thay)

(NP-SUB những vị anh hùng dân tộc)

(. !))

Trong ví dụ này cụm tính từ vị ngữ đứng trước cụm danh từ chủ ngữ. Chúng được nối với nhau bằng tiểu từ "thay".

Câu cảm thán dùng phó từ tình thái :

Con này gớm thật !

(S-EXC (NP-SUB Con này)

(VP (V gớm)

(R thật))

(. !))

14. Câu mệnh lệnh

Ký hiệu : S-CMD

Cấu trúc chung :

Câu mệnh lệnh của tiếng Việt được cấu tạo nhờ những phụ từ tạo ý mệnh lệnh, bằng ngữ điệu mệnh lệnh, và chỉ được chứa những từ liên quan đến nội dung của lệnh (đảm bảo tính ngắn gọn) [1]. Các phụ từ mệnh lệnh hay dùng là : hãy, đừng, chớ, đi, thôi, v.v.

Ví dụ 1 :

Không được làm ồn !
(S-CMD (VP (R không được)
(V làm)
(AP ồn))
(. !))

Ví dụ 2 :

Đi đi, em !
(S-CMD (VP (V đi)
(R đi))
(, ,)
(NP-SUB em))
(. !))

15. Câu đặc biệt và tín báo

15.1 Câu đặc biệt

Nhãn chức năng: SPL

Mô tả : Trong (UBKHXXH, 1983) các tác giả phân biệt nhiều loại câu đặc biệt, trong đó có loại câu xác định trạng thái tồn tại của sự vật. Có một điều cần chú ý là ta không coi câu với động từ “có” là câu đặc biệt.

Tiếng reo.

(S-SPL (NP Tiếng reo)
.)

Giỏi thật!

(S-SPL-EXC (AP Giỏi thật)
!)

15.2 Tít báo

Nhãn chức năng: TTL

Mô tả : Nhãn này được dùng để gán cho cấu trúc cú pháp làm tiêu đề của một bài báo. Tiêu đề có thể là cụm từ hay câu. Sau đây là ví dụ về cụm danh từ và câu làm tiêu đề:

Cuộc đời sau tấm màn nhung
(NP-TTL Cuộc đời
(PP sau
(NP tấm màn nhung)))

Con thi, cha mẹ cũng thi
(S-TTL (S (NP Con)
(VP thi))
(, ,)
(S (NP cha mẹ)
(VP cũng thi)))

16.Các nhãn chức năng

Thông tin cú pháp cơ bản nhất được thể hiện trong cây cú pháp qua các nhãn từ loại, cụm từ, và mệnh đề. Tuy nhiên, trong các ứng dụng của treebank [] nhiều trường hợp cần thông tin cụ thể hơn nữa. Do đó nhãn chức năng được sử dụng để làm giàu thông tin thể hiện trong cây cú pháp.

16.1 Nhãn chức năng chủ ngữ

Ký hiệu : SUB

Mô tả : Nhãn này được dùng để gán cho cụm từ làm chủ ngữ ở trong câu. Xét về vị trí,

chủ ngữ thường đứng trước vị ngữ.

Ví dụ :

Anh này là sinh viên .
(S (NP-SUB Anh này)
(VP là
(NP-DOB sinh viên))
(. .))

Trong ví dụ này nhãn chức năng vị ngữ không được sử dụng vì sau cụm từ có nhãn chức năng chủ ngữ, nếu tồn tại cụm động từ (không phải phụ ngữ) thì cụm đó là vị ngữ. Trong trường hợp vị ngữ không phải cụm động từ, khi đó ta mới dùng nhãn chức năng vị ngữ (phần 13.5).

16.2 Nhãn chức năng tân ngữ trực tiếp

Ký hiệu : DOB

Mô tả : Nhãn này được dùng để gán cho cụm từ làm tân ngữ (object) của động từ trong cụm động từ. DOB ứng với vai bị thể trong [NVHiệp, 2008]. Các dấu hiệu để nhận biết DOB :

- Nó trả lời câu hỏi: ai, cái gì.
- Nó có thể làm chủ ngữ của câu bị động
- Về vị trí, nó thường đứng sau động từ

Ví dụ :

Tôi lái ô tô .
(S (NP-SUB Tôi)
(VP lái
(NP-DOB ô tô))
(. .))

Hỏi : Ai lái ô tô ?

Dạng bị động : Ô tô được tôi lái.

Vị trí : sau động từ "lái"

Các trường hợp đặc biệt:

Có một số động từ mà theo sau là danh từ nhưng danh từ đó không được gán nhãn DOB.

Xét động từ “là”:

Tôi là sinh viên.

(S (NP-SUB Tôi)

(VP là

(NP sinh viên)))

Rõ ràng “sinh viên” là danh từ đi sau động từ “là” nhưng không thể gán cho nó nhãn chức năng DOB, vì nó không phải là đối tượng bị tác động bởi chủ thể “tôi”. Như vậy câu không có dạng bị động.

Tương tự:

bằng

Cái ấm này bằng nhôm.

tại

Việc này tại anh ấy.

của

Cái áo này của tôi.

như

Anh ấy như người ốm.

(Có lẽ câu đầy đủ là “Anh ấy trông như người ốm.”?)

có

Anh ấy có chiếc xe mới.

lên

Em bé này lên 10 tuổi.

Một trường hợp khác, xin xem phần 16.7.

16.3 Nhãn chức năng tân ngữ gián tiếp

Ký hiệu : IOB

Mô tả : Nhãn này được dùng để gán cho cụm từ làm tân ngữ gián tiếp (indirect object) của động từ trong câu. IOB ứng với vai tiếp thể và vai kẻ hưởng lợi trong [NVHiệp, 2008]. Các dấu hiệu để nhận biết :

- Tồn tại giới từ "cho" (cho ai, cho cái gì) trước cụm danh từ đang xét
- Có thể chèn "cho" vào trước cụm danh từ đang xét

Ví dụ:

Tôi tặng bạn quyển sách .
(S (NP-SUB Tôi)
(VP tặng
(NP-DOB bạn)
(NP-IOB quyển sách))
(. .))

Động từ “tặng” trong ví dụ này có hai tân ngữ. Tân ngữ trực tiếp là “sách” thì được gán nhãn chức năng DOB, còn tân ngữ gián tiếp “bạn” thì có nhãn IOB.

16.4 Nhãn chức năng khởi ngữ

Ký hiệu : TPC

Mô tả : Về mặt hình thức, khi một thành phần khác chủ ngữ được đưa lên đầu câu thì thường thành phần đó là khởi ngữ.

Ví dụ:

Vấn đề này chúng tôi đang bàn .
(S-TC (NP-TPC Vấn đề này)
(S (NP-SUB chúng tôi)
(VP đang bàn))
(. .))

Trong ví dụ này, chủ đề (phần đề) của câu được đặt ở đầu câu. Thực chất phần đề này là tân ngữ trực tiếp của động từ “bàn”. Ta sẽ xem xét cách thể hiện thông tin này ở phần 15. Chú ý là nhãn chức năng TC (đề-thuyết hay topic-comment) được dùng để gán cho câu (nhãn S) có thành phần khởi ngữ.

Khởi ngữ còn được gọi là chủ đề, thành phần khởi ý, v.v. Có nhiều nhà ngôn ngữ học đã nghiên cứu về vấn đề này. Chúng tôi coi sách của Nguyễn Minh Thuyết và Nguyễn Văn Hiệp (1999; tr 187-214) và Cao Xuân Hạo (2006) là các tài liệu tham khảo chính. Khi gán nhãn cần chú ý phân biệt khởi ngữ với các thành phần khác của câu như trạng ngữ và vế của câu ghép (Thuyết và Hiệp, 1999; tr 200-204).

Giàu thì tôi đã giàu rồi.

(S-TC (AP-TPC (A Giàu))

(C thì)
 (S (NP-SUB (P tôi))
 (AP-PRD (R đã)
 (A giàu)
 (R rồi))))

Phân thuyết là mệnh đề thiếu chủ ngữ.

Tiền mất 3 triệu đồng, lại cũng mệt mỏi.

(S (S-TC (NP-TPC Tiền)
 (S (NP-SUB *E*)
 (VP mất
 (NP 3 triệu đồng))))

,
 (S (NP-SUB *E*)
 (VP lại cũng mệt mỏi))
 .)

Phân đề là một mệnh đề:

Anh thành công khiến em rất vui mừng.

(S-TC (S-TPC (NP-SUB Anh)
 (VP thành công))
 (VP khiến
 (NP-DOB em)
 (VP rất vui mừng))
 .)

Phân đề là cụm giới từ:

Trên kệ là chiếc tivi.

(S-TC (PP-TPC Trên
 (NP kệ))
 là
 (NP chiếc tivi)
 .)

16.5 Nhãn chức năng dành cho vị ngữ không phải cụm động từ

Ký hiệu : PRD

Mô tả : Nếu vị ngữ của câu không phải là một cụm động từ thì nó được gán nhãn PRD. Nói chung ngoài cụm động từ, cụm tính từ và cụm danh từ cũng có thể làm vị ngữ trong câu. Trong tiếng Việt, cụm tính từ làm vị ngữ là hiện tượng phổ biến.

Ví dụ 1:

Cô gái đẹp .
(S (NP-SUB Cô gái)
(AP-PRD đẹp)
(. .))

Ví dụ 2:

Nhà này 60 mét vuông .
(S (NP-SUB Nhà này)
(NP-PRD 60 mét vuông)
(. .))

16.6 Nhãn chức năng của chủ ngữ logic

Ký hiệu : LGS (logical subject)

Mô tả : Với một câu bị động tiếng Việt được viết đúng ngữ pháp [1, trg149], ta không cần đến nhãn này. Tuy nhiên hiện nay có hiện tượng viết sai ngữ pháp do ảnh hưởng của tiếng Anh. Nếu gặp những câu như vậy thì ta dùng thêm nhãn chức năng LGS. Dấu hiệu nhận biết:

- Tồn tại giới từ “bởi” trước cụm từ đang xét.

Ví dụ 1¹¹:

Yahoo! 360⁰ có thể bị thay thế bởi Yahoo! Mash
(S (NP-TPC Yahoo! 360⁰)
(VP có thể
(VP bị thay thế

¹¹ <http://www.tuoitre.com.vn/Tianyon/Index.aspx?ArticleID=220683&ChannelID=16>

(PP bởi
(NP-LGS Yahoo! Mash))))

Ví dụ này được lấy từ tiêu đề của một bài báo gần đây trên báo Tuổi Trẻ Online.

Ví dụ 2:

Yahoo! 360⁰ có thể bị Yahoo! Mash thay thế
(S (NP-TPC Yahoo! 360⁰)
(VP có thể
(R bị)
(S (NP-SUB Yahoo! Mash)
(VP thay thế))))

Câu trong ví dụ 1 được sửa cho đúng với ngữ pháp tiếng Việt hơn. Khi đó ta không dùng nhãn LGS nữa.

16.7 Nhãn chức năng bổ ngữ chỉ phạm vi hay tần suất của hành động

Ký hiệu: EXT

Mô tả: Nếu cụm danh từ chỉ phạm vi hay tần suất làm bổ ngữ sau cho động từ thì được gán nhãn EXT. Chú ý là trong trường hợp này cụm danh từ không phải tân ngữ (DOB).

Ví dụ:

Anh ấy chạy 5 km .
(S (NP-SUB Anh ấy)
(VP chạy
(NP-EXT 5 km))
(. .))

16.8 Thành phần than gọi

Nhãn chức năng: VOC

Mô tả: Thành phần này nêu lên một lời than hay lời gọi (UBKHXXH, 1983). Nó thường do một cảm từ hay một ngữ có tác dụng như cảm từ đảm nhiệm.

Chao, đường còn xa lắm!
(S-EXC (MDP-VOC Chao)

,

(NP-SUB đường)
 (AP-PRD còn xa lắm)
 !)

Trong ví dụ trên “chao” là cảm từ.

Anh ơi, chờ em với !
 (S-EXC (NP-SUB-VOC Anh ơi)
 ,
 (VP chờ
 (NP em)
 với)
 !)

Trong ví dụ trên “ơi” là cảm từ và được phân tích như thành phần của cụm danh từ. Như vậy “anh ơi” có hai chức năng chủ ngữ và than gọi.

17. Nhãn phân loại phụ ngữ của động từ

Trong ngôn ngữ học hiện đại [2,5], phụ ngữ là thành phần câu đóng vai trò thiết lập tình huống diễn ra hành động hay trạng thái mà động từ chính mô tả. Về hình thức, phụ ngữ có thể là từ, cụm từ, hay mệnh đề. Về ý nghĩa, phụ ngữ thường diễn tả: thời gian, nơi chốn, cách thức, nguyên nhân, mục đích, hay điều kiện.

17.1 Phụ ngữ thời gian

Ký hiệu: TMP

Ví dụ:

Ngày mai tôi đi thi .
 (S (NP-TMP Ngày mai)
 (S (NP-SUB tôi)
 (VP đi
 (VP thi)))
 (. .))

Tôi hay nghe nhạc trong khi làm việc.
 Tôi mới lên Hà Nội được ba tháng.

17.2 Phụ ngữ nơi chốn

Ký hiệu: LOC

Dấu hiệu: Là cụm danh từ chỉ địa điểm hoặc cụm giới từ chỉ địa điểm (giới từ có tác dụng đánh dấu vai địa điểm).

Ví dụ:

Tôi sẽ đi nghỉ ở Tokyo .
(S (NP-SUB Tôi)
(VP sẽ
đi
(VP nghỉ
(PP-LOC ở Tokyo)))
(. .))

17.3 Phụ ngữ chỉ hướng

Ký hiệu: DIR

Mô tả: Vai này cho biết chuyển động diễn ra theo đường nào hay hướng nào. Dấu hiệu nhận biết:

- Tồn tại các giới từ như “từ”, “ra”, “vào”, “lên”, “xuống”, “đến”, “tới”, “sang”, “qua”, “lại”, “về” trước cụm từ đang xét

Ví dụ:

Anh ấy sẽ bay từ Sài Gòn ra Hà Nội .
(S (NP-SUB Anh ấy)
(VP sẽ
bay
(PP-DIR (PP từ
(NP Sài Gòn))
(PP ra
(NP Hà Nội))))
(. .))

17.4 Phụ ngữ chỉ cách thức hay phương tiện

Ký hiệu: MNR

Mô tả: Vai này cho biết một hành động được thực hiện như thế nào. Dấu hiệu nhận biết:

- Đặt câu hỏi *thế nào*
- Tồn tại các giới từ như “bằng” (chỉ phương tiện) ở trước cụm từ đang xét

Ví dụ:

Cô gái ăn chè bằng thìa .
 (S (NP-SUB Cô gái)
 (VP ăn
 (NP-DOB chè)
 (PP-MNR bằng
 (NP thìa))))
 (. .))

17.5 Phụ ngữ chỉ mục đích hay lý do

Ký hiệu: PRP

Mô tả: Có một số dấu hiệu nhận biết sau

- Tồn tại các giới từ như “để”, “cho”, “mà” (mục đích), “vì”, “do”, “tại”, “bởi” (lý do, nguyên nhân) ở trước cụm từ hay mệnh đề đang xét
- Có thể chen các giới từ đó vào trước mệnh đề đang xét
- Đặt câu hỏi *để làm gì* hoặc câu hỏi *tại sao* cho cụm chủ vị chính

Ví dụ:

Nó không đi làm được vì ốm .
 (S (NP-SUB-1 Nó)
 (VP không
 đi
 (VP làm)
 được
 (SBAR-PRP vì
 (S (NP-SUB-1 *T*)
 (AP-PRD ốm))))
 (. .))

// “đi làm” nên được ghép lại?

Con nên mượn sách của bạn mà học tiếng Anh.

17.6 Phụ ngữ chỉ điều kiện

Ký hiệu: CND

Dấu hiệu nhận biết: Tồn tại các từ/cụm từ “nếu”, “giá mà”, “miễn là”, “hễ” đi trước cụm từ hay mệnh đề đang xét.

Nếu thời tiết đẹp thì lớp chúng tôi sẽ thăm rừng Cúc Phương vào chủ nhật này.
Hễ ông Hòa đến thì anh gọi tôi đây.

17.7 Trạng ngữ chỉ ý nhượng bộ

Ký hiệu: CNC¹²

Dấu hiệu nhận biết: Tồn tại các từ/cụm từ “tuy”, “dù”, “mặc dù” đi trước cụm từ hay mệnh đề đang xét.

Tuy anh ấy bị hỏng mắt nhưng anh ấy vẫn sống rất lạc quan.¹³

17.8 Trạng ngữ

Nhãn chức năng: ADV

Mô tả¹⁴: Nhãn này được sử dụng khi thành phần trạng ngữ không thuộc một trong các loại cụ thể đã được mô tả ở phần trên. Nếu một trạng ngữ được phân loại cụ thể hơn (chẳng hạn TMP) thì cũng được hiểu là có loại ADV.

// ví dụ?

18. Nhãn phần tử rộng

18.1 Nhãn *T*

Mô tả: Nhãn phần tử rộng *T* được dùng để thể hiện cấu trúc sâu của một số hiện tượng ngữ pháp¹⁵ như bị động, khởi ngữ, mệnh đề phụ kết, **cụm động từ làm bổ ngữ**.

Comment [t3]: Thêm ví dụ

Ví dụ 1:

Thuyền được đẩy ra xa .
(S-PV (NP-SUB-1 Thuyền)
(VP được
(SBAR (NP *E*))

¹² Viết tắt của concession

¹³ Ví dụ này được lấy từ (LBiên, 1999; tr166)

¹⁴ Tài liệu hướng dẫn gán nhãn cú pháp phần 1 đã phân biệt các loại trạng ngữ cụ thể.

¹⁵ Nhãn *T* dùng cho A-movement

(VP đẩy
(NP-DOB-1 *T*)
ra xa)))

(. .))

Đây là một câu bị động trong đó tân ngữ của động từ “đẩy” được đưa lên đầu làm chủ ngữ của câu.

Ví dụ 2:

Vấn đề này chúng tôi đang bàn .

(S-TC (NP-TPC-1 Vấn đề này)

(S (NP-SUB chúng tôi)

(VP đang bàn

(NP *T*-1)))

(. .))

Ví dụ 3:

Tôi đã mua quyển sách mà thầy giáo giới thiệu .

(S (NP-SUB Tôi)

(VP đã mua

(NP-DOB (NP-1 quyển sách)

(SBAR mà

(S (NP-SUB thầy giáo)

(VP giới thiệu

(NP-DOB *T*-1))))))

(. .))

Câu này có hai mệnh đề, trong đó mệnh đề phụ kết bổ nghĩa cho từ “quyển sách”. Ở mệnh đề này, tuy tân ngữ không trực tiếp xuất hiện sau động từ “giới thiệu” nhưng ta ngầm hiểu đó là cụm từ “quyển sách”. Do đó ta cần đưa vào ký hiệu cụm danh từ rỗng giá trị chỉ số là 1, đồng chỉ số với cụm danh từ “quyển sách”¹⁶. Một cụm danh từ rỗng vẫn được gán nhãn chức năng như bình thường, trong trường hợp này là DOB.

Ví dụ 4:

Anh ấy khỏe vì chơi tennis đều đặn .

(S (NP-SUB-1 Anh ấy)

¹⁶ Nếu cụm từ không có phần từ rỗng tương ứng thì không cần gán chỉ số.

(AP-PRD khỏe
 (SBAR vì
 (S (NP-SUB *T*-1)
 (VP chơi
 (NP tennis)
 đều đặn))))

(. .))

Ở câu này thì phần từ rỗng lại là chủ ngữ của mệnh đề phụ kết bỏ nghĩa cho tính từ vị ngữ của mệnh đề chính.

18.2 Nhãn *0*

Mô tả: Nhãn này dùng để gán cho những mệnh đề phụ kết mà thiếu tác từ phụ ngữ hóa (complementizer). Dấu hiệu nhận biết:

- Có thể chèn “rằng”, “mà” vào trước mệnh đề đang xét (mệnh đề này có thể là bỏ ngữ của động từ hoặc định ngữ của danh từ)

(S (NP-SUB Chúng)
 (VP dẫn
 (AP thắng)
 (PP-DIR lên
 (NP núi
 ,
 (NP-LOC nơi
 (SBAR *0*
 (S (NP-SUB hàng lậu)
 (VP được
 (VP chuyển đến))))))))

(. .))

18.3 Nhãn *RNR*

Mô tả:

Anh ấy vừa ăn vừa nói trong bữa tiệc .
 (S (NP-SUB Anh ấy)
 (VP (VP vừa ăn

(PP-TMP *RNR*-1))
 (VP vừa nói
 (PP-TMP *RNR*-1))
 (PP-TMP-1 trong
 (NP bữa tiệc)))
 (. .))

19. Các cấu trúc sử dụng liên từ đẳng lập

Liên từ đẳng lập được dùng để nối (hay thể hiện quan hệ đẳng lập giữa) hai từ hay cụm từ.

19.1 Các từ **trung tâm** không có chung bổ ngữ

Nếu hai (hay nhiều) từ được nối với nhau bằng liên từ độc lập thì ta có thể chỉ cần gán như sau:

Bố, mẹ, và con
 (NP (N-H Bố) (, ,) (N-H mẹ) (, ,) (C và) (N-H con))

Chú ý là ở ví dụ này cụm danh từ có 3 danh từ trung tâm.

Trường hợp có ít nhất một thành phần là cụm từ thì ta gán cho chúng nhãn cụm từ:

Hai bút chì và một quyển sách
 (NP (NP hai bút chì)
 và
 (NP một quyển sách))

Các ví dụ khác:

Cấu trúc cú pháp và ngữ nghĩa
 (NP Cấu trúc
 (NP cú pháp và ngữ nghĩa))
 Đã, đang và sẽ thực hiện mua sách, giấy và bút
 (VP (RP đã, đang và sẽ)
 (VP thực hiện
 (VP mua
 (NP sách, giấy và bút))))

Comment [t4]: Có nhập nhằng ko? Nên giải thích cho nhóm.

19.2 Các từ (hay cụm từ) có chung bổ ngữ

Họ thường chỉ nghĩ tới nghĩa vụ đóng góp của Việt kiều, chú trọng tới khía cạnh “khai thác”.

20. Kết từ đẳng lập (C) và kết từ chính phụ (E)

Qui ước:

- Dùng nhãn C với kết từ đẳng lập (và, hoặc, v.v.)
- Dùng nhãn E với kết từ chính phụ mà theo sau là mệnh đề (vì, do, v.v.), khi đó nút cha sẽ là SBAR
- Dùng nhãn E với kết từ chính phụ mà theo sau là cụm danh từ hoặc cụm động từ (của, cho, để, v.v.), khi đó nút cha sẽ là PP

Một tài liệu tham khảo tốt về các loại kết từ là (DQ Ban, 2007; Tập 1, tr132-).

Ví dụ:

Tôi yêu anh vì những nguyên nhân sâu xa hơn.

```
(S (NP Tôi)
  (VP yêu
    (NP anh)
    (PP-PRP vì
      (NP những nguyên nhân
        (AP sâu xa hơn))))
  (. .))
```

Người ta nhìn anh vì anh giống Rô-mê-ô của họ.

```
(S (NP Người ta)
  (VP nhìn
    (NP anh)
    (SBAR-PRP vì
      (S (NP anh)
        (VP giống
          (NP Rô-mê-ô
            (PP của
              (NP họ))))))
      (. .))
    (. .))
  (. .))
```

Trong cả hai câu “vì” đều có từ loại E. Trong câu thứ nhất, “vì” là phần tử trung tâm của cụm giới từ PP. Trong câu thứ hai, “vì” lại là phần tử trung tâm của mệnh đề phụ SBAR.

21. Thành phần chú thích hoặc trích dẫn

21.1 Thành phần chú thích

Về hình thức, thành phần chú thích có thể được đặt trong cặp dấu gạch ngang, cặp dấu ngoặc, hay dấu phẩy. Về chức năng, phần chú thích có thể bổ nghĩa cho cụm từ hoặc cho cả câu. Ta xét các ví dụ sau:

Phần chú thích bổ nghĩa cho cụm danh từ:

Bà Hiền, giám đốc công ty X, cho biết ...

(S (NP (NP-H Bà Hiền)
(, ,)
(NP giám đốc công ty X)
(, ,))
(VP cho biết ...))

Phần chú thích bổ nghĩa cho cả câu:

Em học sinh ấy – thật là gan dạ – đã nổ mìn diệt cả tổp địch.

(S (NP Em học sinh ấy)
(- -)
(AP thật là gan dạ)
(- -)
(VP đã
(VP nổ
(NP mìn))
(VP diệt
(NP cả tổp địch))))

Nếu thành phần chú thích được bao trong cặp dấu ngoặc thì dấu ngoặc trái ‘(’ sẽ được chuyển thành LBKT còn dấu ngoặc phải ‘)’ thành RBKT. Chú ý là cặp dấu chú thích nên được đặt ngang bậc với nhau trong cây cú pháp. Chẳng hạn như câu trên không nên được gán nhãn như sau :

(S (NP Em học sinh ấy)
 (- -)
 (AP thật là gan dạ -)
 (VP đã
 (VP nỏ
 (NP mìn))
 (VP diệt
 (NP cả tốp địch))))

Trong đó dấu gạch kết thúc chú thích được đặt ở mức sâu hơn dấu gạch bắt đầu chú thích.

21.2 Thành phần trích dẫn

Thành phần này thường là câu nói, được đặt trong cặp dấu nháy kép, đi sau các động từ như "nói", "cho biết", v.v.

(S (NP-SUB Cô Lan)
 (VP nói thêm
 :
 (SBAR *0*
 "
 (S (NP-SUB chúng tôi)
 (VP lên
 (NP-LOC đây)
 (VP-TMP được
 (NP-TMP 6 năm)
 rồi)))
 "))
 (. .))

Chú ý cặp dấu nháy kép cũng được đặt ngang bậc nhau trong cây cú pháp.

22. Câu ghép

Trong các tài liệu về ngữ pháp tiếng Việt (UBKHXXH, 1983; DQBan, 2007) câu ghép được nhận biết và phân loại dựa vào các dấu hiệu: số thành phần chủ-vị, cách dùng phụ từ, và cách dùng kết từ. Thêm vào đó quan hệ ngữ nghĩa giữa các thành phần câu cũng là căn cứ phân loại câu ghép và phân biệt câu ghép với câu đơn.

22.1 Câu ghép song song¹⁷ (UBKHXXH, 1983)

Dấu hiệu hình thức là câu ghép này có từ hai cụm chủ vị trở lên. Các cụm chủ vị này được nối với nhau bằng dấu phẩy hoặc kết từ đẳng lập.

Chim kêu, vượn hú.

```
(S (S (NP Chim)
      (VP kêu))
  ,
  (S (NP vượn)
      (VP hú))
  .)
```

Tôi chưa làm kịp, hay anh làm giúp tôi vậy?

```
(SQ (S (NP Tôi)
        (VP chưa làm kịp))
    ,
    hay
    (S (NP anh)
        (VP làm giúp
            (NP tôi)
            vậy))
    ?)
```

22.2 Câu ghép qua lại (UBKHXXH, 1983)

Phụ từ được dùng trong vị ngữ của các vế: vừa...đã, chưa...đã, mới...đã, v.v.

Thầy giáo vừa dạy xong cậu đã quên rồi à.

```
(S (S (NP Thầy giáo)
      (VP vừa dạy xong))
  (S (NP cậu)
      (VP đã quên rồi à))
  .)
```

¹⁷ Còn gọi là câu ghép đẳng lập

Kết từ được dùng ở đầu mỗi vế: vì...nên, vì...mà, nếu...thì, v.v.

Vì trời mưa *nên* tôi không đi chơi nữa.

```
(S [Vi]
  (S (NP trời)
    (VP mưa))
  nên
  (S (NP tôi)
    (VP không đi chơi nữa))
  .)
```

Comment [t5]: PRP

22.3 Phân biệt câu ghép với câu đơn có thành phần **trạng ngữ**

Comment [t6]: Chỉ bị nhập nhằng 2 loại là CND và PRP

Sau khi tham khảo các tài liệu (NXBKHXH, 1983; Thuyết và Hiệp, 1999; DQBan, 2007) chúng tôi thấy rằng chưa có một cách phân biệt thống nhất giữa các tác giả. Tuy nhiên chúng tôi thấy rằng (Thuyết và Hiệp, 1999, tr314-316) có đề xuất một cách giải quyết vấn đề thiên về mặt cú pháp hình thức và khá rõ ràng, thuận tiện cho việc áp dụng. Do đó chúng tôi lựa chọn cách đó cho gán nhãn Treebank. Tiêu chuẩn như sau:

- Nếu liên từ thứ hai vắng mặt thì ta sẽ có một câu đơn có trạng ngữ chỉ nguyên nhân

```
(S (SBAR-PRP Vì
  (S (NP trời)
    (VP mưa)))
  (S (NP tôi)
    (VP không đi chơi nữa))
  .)
```

- Các trường hợp còn lại ta có câu ghép (3 trường hợp)

```
(S [Vi]
  (S (NP trời)
    (VP mưa))
  nên
  (S (NP tôi)
    (VP không đi chơi nữa))
  .)
```

Comment [t7]: PRP

(S (S (NP trời)
 (VP mưa))
 nên
 (S (NP tôi)
 (VP không đi chơi nữa))
 .)

Comment [t8]: PRP thiếu “vì”

(S (S (NP trời)
 (VP mưa))
 (S (NP tôi)
 (VP không đi chơi nữa))
 .)

Comment [t9]: PRP thiếu “vì”

23. Tỉnh lược

Kí hiệu *E*. Mô tả: Trong nhiều trường hợp cần phải dùng nhãn phần tử tỉnh lược để mô tả đầy đủ hơn cấu trúc ngữ pháp của 1 câu.

1. Tỉnh lược chủ ngữ (TLCN)

Dạng tỉnh lược này xuất hiện trong các trường hợp sau:

- TLCN trong câu mệnh lệnh

Ví dụ:

Ông chủ lại nói to:

-A! Nó đói! Ø Đi bắt cho nó vài con châu chấu!

(S-CMD(NP-SUB *E*))

(VP (VP Đi))

(VP(V bắt

(PP-IOB cho nó

(NP-DOB vài con châu chấu)))

(. .))

- Câu có dạng lời cầu chúc / cầu mong / lời chào...

Ví dụ:

- Mời ông vào trong này. Chúng tôi đợi mãi. Ø Mời ông vào thưởng trăng.

(S(NP-SUB *E*))

(VP(V Mời)

(NP-DOB ông)
 (VP vào
 (VP-PRP thường
 (NP trống))))

(. .))

- Câu chứa các từ tình thái “cần, nên, phải, hãy”.

Ví dụ:

Nó chết phải đem chôn. Ø Phải mua cỗ gỗ. Ø Phải mời xóm, mời làng.

(S(NP-SUB *E*))
 (VP(V phải)
 (VP mua)
 (NP-DOB cỗ gỗ)))

(. .))

(*) kết ngôn: phát ngôn được liên kết với phát ngôn bị tình lược

Ø: kí hiệu cho yếu tố bị tình lược

- Câu có chủ ngữ mang ý nghĩa khái quát

Ví dụ:

Cụ Tiên chỉ làng Vũ Đại nhận ra rằng: Ø Hãy ngấm ngấm đẩy người ta xuống sông nhưng rồi dắt nó lên để đền ơn.

(Chí Phèo – Nam Cao)

Xin giải thích thêm đối với trường hợp này: chủ ngữ ở đây có thể tự do tùy vào hoàn cảnh sử dụng (tình lược trong trường hợp này thường gặp trong các câu thành ngữ, các câu nêu kinh nghiệm: Ăn quả nhớ kẻ trồng cây, uống nước nhớ nguồn...)

(S(NP-SUB Cụ Tiên chỉ làng Vũ Đại)
 (VP (V nhận ra)
 (SBAR rằng :
 (S(NP-SUB *E*))
 (VP (V Hãy)
 (AP (A ngấm ngấm)
 (VP (V đẩy)
 (NP-DOB người ta)

Comment [t10]: Hay đây là Head?

(VP xuống
(NP-LOC sông))))
(C rồi)
(VP (V đất)
(NP-DOB-1 nó)
(R lên)
(SBAR-PRP (E để)
(S (NP *T*-1)
(VP (V đèn on)))))))))
(. .))

- Chủ ngữ trong câu bị tính lược đồng chức năng với chủ ngữ ở các kết ngôn (*)
Ví dụ này cho thấy yếu tố đóng vai trò là chủ ngữ trong phát ngôn bị tính lược “anh”
cũng đồng thời là chủ ngữ bị tính lược trong các phát ngôn đằng sau.

Ví dụ:

Anh cứ hát. Gò ngực mà hát.

(S (NP-SUB *E*)
(VP gò
(NP-DOB ngực)
(C mà)
(VP hát))
(. .))

Comment [t11]: Chưa ổn! có lẽ gần là
VP-PRP ?

- Chủ ngữ trong câu bị tính lược không cùng chức năng trong các kết ngôn

Ví dụ:

(1) Sau cùng thì y gạ THAI cổ vườn. (2) Ø Không muốn cổ thì Ø xoay tiền mà
trả y.

Trong ví dụ trên, “Thai” là bổ ngữ trong phát ngôn (1) nhưng sang đến (2) có
chức năng là chủ ngữ của các phát ngôn đó.

(S
(S(NP-SUB *E*)
(VP không muốn cổ)))
(C thì)
(S(NP-SUB *E*)
(VP xoay tiền)
(C mà)

(VP trả)
(NP y))))

(. .))

- Chủ ngữ là yếu tố được ngầm định
 - Chủ ngữ là lời của chính tác giả

Ví dụ:

Cái đầu ông ngoẹo xuống, như đầu một thằng bé khi nó cúi. Trông thật là thiếu
não.

(S(NP-SUB *E*))
(VP Trông)
(VP(T thật)(T là)
(AP thiếu não)))

(. .))

- Chủ ngữ là nhân vật được nói đến trong truyện

Ví dụ:

Nguyên là bà ấy béo quá – gớm! Ø Béo đến nỗi bụng sẽ xuống.

(S(NP-SUB *E*))
(VP Béo)
(X đến nỗi)
(SBAR(S Bụng sẽ xuống)))

(. .))

- Chủ ngữ là một trong những người đang đối thoại

Ví dụ:

Rồi đổi giọng, cụ thân mật hỏi:

- Ø Về bao giờ thế?

(S(NP-SUB *E*))
(VP Về)
(WHRP bao giờ)
(T thế))

(. .))

- Chủ ngữ ẩn

Ví dụ:

Ông Bình cho biết sắp tới sẽ đề nghị trung ương phân cấp cho TP, phân cấp cho
tổng công ty để đẩy nhanh hơn tốc độ cổ phần hoá.

(S (NP-SUB Ông Bình)
 (VP cho biết
 (SBAR (RP-TMP sắp tới)
 (S (NP-SUB *E*)
 (VP sẽ đề nghị
 (SBAR (S (NP-SUB trung ương)
 (VP (VP phân cấp
 (PP cho
 (NP TP)))
 (, ,)
 (VP phân cấp
 (PP cho
 (NP tổng công ty)))
 (PP-PRP để
 (VP đẩy
 (AP nhanh hơn)
 (NP-DOB tốc độ cổ phần
 hóa))))))))))

(. .))

Câu tính lược trên đã lược bỏ thành phần chủ ngữ trong mệnh đề thứ nhất. Mà khó có thể khôi phục chính xác thành phần chủ ngữ. Do đó gán nhãn *E* nhằm ngầm thông báo rằng đã có 1 thành phần bị tính lược.

- Chủ ngữ bị tính lược khi câu có mô hình : Là + danh từ
- Mô hình hiển ngôn từ “là”

Cấu trúc “Danh là Danh” biểu thị quan hệ đồng nhất giữa các sự vật, hiện tượng. Ví dụ: Ông là người cha nghiêm khắc của lũ con ích kỉ, đần độn. Ø Là người chồng đáng kính của các bà vợ tầm thường...

(S (NP-SUB*E*)
 (VP Là
 (NP người chồng
 (VP đáng kính)
 (PP của
 (NP các bà vợ
 (AP tầm thường))))))

(. .))

- Mô hình ẩn từ “là”

Ví dụ:

Mà bọn cô đầu thì ác quá. Họ cậ có quần áo đẹp, tóc uốn quăn. Họ cứ nhìn cái đầu thợ nhà quê xén vụng của tôi mà cười. Huống chi lại còn có bao nhiêu khách của ô. H nữa. Ø Toàn những phú thương cả.

Nếu khôi phục câu này sẽ là: Họ toàn là những phú thương cả.

(S (NP-SUB *E*))

(VP (R toàn)

(NP những phú thương)

(NP (P cả)))

(. .))

2. Tính lược vị ngữ

Khi tính lược vị ngữ thì nòng cốt còn lại trong câu là chủ ngữ. Ngoài ra còn có thể có những thành phần phụ khác nếu chúng thuộc phần báo. Khi đó, giữa chủ ngữ với các thành phần phụ này thường có dấu phẩy (hoặc dấu ngang nối...) ngăn cách – đây là một hình thức đánh dấu vị trí tính lược.

Ví dụ:

Vấn Viện kiểm sát: Bộ có chủ trương cho thứ trưởng ký xác nhận vào công văn vào công văn Công ty tiếp thị không?

(S

(NP-SUB(R Vấn)

(Np Viện Kiểm sát))

(VP *E*))

(: :)(“ “)

(SQ(NP-SUB Bộ))

(VP(R có)

(V chủ trương)

(PP cho)

(NP-DOB thứ trưởng))

(VP-IOB ký)

(NP xác nhận))

(PP

vào)

(NP(N

công văn)

(NPCông ty)

(VP tiếp thi))))))

(R không))))

(. ?)(" ")

(. .))

Nếu các thành phần phụ cũng thuộc phần nêu thì chúng có thể tỉnh lược cùng với vị ngữ. Khi đó chủ ngữ sẽ là thành phần duy nhất còn lại trong câu bị tỉnh lược. Ví dụ:

Lan vừa bước vào nhà. Ø Cả bố và mẹ.

(S (NP-SUB cả bố và mẹ)

(VP *E*))

(. .))

3. Tỉnh lược C-V

Ví dụ:

Thoáng chốc, Quyên nhớ đến mọi nét mọi vẻ của Cà My. Ø1 Ø2 Cả cái cử chỉ khi Cà My ôm cô mà hôn thiết là kêu.

Ø= Quyên nhớ đến...

(S (NP-SUB *E*))

(VP (V *E*))

(NP (P Cà)

cái cử chỉ

(NP-TMP khi

(SBAR (S Cà My ôm cô mà hôn cô thiết là kêu))))))

(. .))

Trong tiếng Việt, một cấu trúc khá quan trọng là C-V-B. Nên khi lược bỏ C-V thì vẫn còn một thành phần khá quan trọng là bổ ngữ. Điều này thể hiện rõ trong các câu tồn tại.

Ví dụ:

Nhìn lại đằng sau, Dũng có cả một khu gang thép. Và Ø1 Ø2 một gia đình sau bao nhiêu năm tan tác đã dần dần đoàn tụ.

Ø= Dũng có cả...

(S (C Và)

Comment [t12]: Là gì vậy?

Comment [t13]: Hay thành phần kết nối câu?

(NP-SUB *E*)

(VP (V *E*))

(NP-SUB một gia đình

(VP (NP-TMP sau bao nhiêu năm tan tác)

đã dần dần đoàn tụ)))

(. .))

Đây là những thành phần dễ bị tính lược nhất. Còn các thành phần khác như: V-B, C-V-B thương ít gặp. Khi khôi phục sẽ khó khăn và mang tính chủ quan. Khi làm mọi người nên gán nhãn cho các thành phần đã được mô tả trên đây.

Comment [t14]: Phân tích thể nào?

24. Câu bị động

I. Cấu trúc chung của câu bị động tiếng Việt

Tiếng Việt không biến hình từ, nên động từ không có dạng chủ động và bị động. Tuy nhiên tiếng Việt cũng có cách diễn đạt ý bị động một cách đều đặn như các quy tắc ngữ pháp, bằng hai phương thức ngữ pháp hư từ và trật tự từ. Với hai phương thức này dạng bị động của câu tiếng Việt được xác định bằng một số yếu tố hữu hạn có quan hệ cấu trúc khá chặt chẽ làm thành ba điều kiện cần và đủ cho việc tạo nên một kiến trúc bị động ổn định.

- Chủ ngữ bị động, về mặt nghĩa chịu ảnh hưởng của động từ ngoại động trong câu bị bao (điều kiện cần để phân biệt nó với chủ ngữ chủ động).
- Có mặt trợ động từ bị động bị hay được (điều kiện cần để phân biệt câu bị động với câu trung tính).
- Vị tổ là một câu bị bao (trong đó chủ ngữ chủ động có thể vắng mặt, vị tổ là động từ ngoại động, thực thể nêu ở chủ ngữ chủ động của câu bị bao không trùng với thực thể nêu ở chủ ngữ bị động của câu. Đây là điều kiện cần để phân biệt bị, được là trợ động từ bị động với bị, được là động từ tình thái).

Do đó ta có mô hình cấu trúc cú pháp chung của câu bị động tiếng Việt

CN1	+	TĐT bị động	+	Vị tổ1(câu bị bao)
(bị động)		(bị, được)		(NC2 + Vị tổ2(ĐT ngoại động) + Bổ ngữ)

Ví dụ: Thuyền được (người lái) đẩy ra xa.

Em bé được mẹ rửa chân cho.

II. Gán nhãn cú pháp cho câu bị động tiếng Việt

Việc gán nhãn cho câu bị động tiếng Việt dựa trên vai trò của từng thành phần trong câu.

Nó cũng giống như việc gán nhãn cho các câu chủ động trong tiếng Việt, cụ thể như sau:

- Nhãn chức năng bị động cho câu là: PV (passive voice)
- CN1 (bị động): đóng vai trò là chủ ngữ của câu, gán nhãn SUB
- Trợ động từ bị động “bị” , “được”: đóng vai trò là động từ trong câu, nên gán nhãn V cho “bị” và “được”
- Vị tố 1(câu bị bao): là một mệnh đề đóng vai trò bổ ngữ của câu, nên gán nhãn là SBAR

Ví dụ 1: Thư được Giáp gửi cho Tị.

(S-PV (NP-SUB-1(N-H Thư))
 (VP (V-H được)
 (SBAR(S(NP-SUB(Np-H Giáp))
 (VP(V-H gửi)
 (NP-DOB-1 *T*)
 (PP-IOB (E-H cho)
 (NP(Np-H Tị))))))))
 (. .))

Ví dụ 2: Nó bị cảnh sát bắt.

(S-PV (NP-SUB-1(P-H Nó))
 (VP (V-H bị)
 (SBAR (S (NP-SUB (N-H cảnh sát))
 (VP (V-H bắt)
 (NP-DOB-1 *T*)))))
 (. .))

Ví dụ 3: Mái nhà bị gió lật.

(S-PV (NP-SUB(N-H Mái nhà))
 (VP (V-H bị)
 (SBAR (S (NP-SUB(N-H gió))
 (VP (V-H lật)
 (NP-DOB-1 *T*)))))
 (. .))

III. Phân biệt trợ động từ bị động (bị, được) với động từ thực và động từ tình thái
 Trong tiếng Việt thường có sự nhập nhằng và nhầm lẫn giữa trợ động từ bị động "bị", "được" với động từ thực "bị", "được" và động từ tình thái "bị", "được". Vì vậy để gán nhãn chính xác điều đầu tiên là cần xác định chính xác đó có phải là câu bị động hay không? Sau đây là những điều kiện dùng của động từ thực bị, được và của động từ tình

thái bị, được, và trợ động từ bị động bị, được thể hiện ở chu cảnh cú pháp và chức năng cú pháp riêng biệt trong cách dùng hai từ này.

1. Chức năng và chu cảnh cú pháp của trợ động từ bị động “bị”, “được”
- Làm tác tố bị động, không tham gia vào vị tổ
- Đứng trước vị tổ là câu bị bao (câu này có thể vắng chủ ngữ), vị tổ của câu bị bao là động từ ngoại động tác động lên thực thể nêu ở chủ ngữ của toàn câu
- Chủ ngữ của câu bị bao và của “bị”, “được” không trùng nhau.

Ví dụ 1: Họ bị kẻ gian lấy mất tiền.

(S-PSS(NP-SUB(P-H Họ))
(VP(V-H bị)
(SBAR(S(NP-SUB(N-H kẻ gian))
(VP(V-H lấy)(V mất)
(NP-DOB(N-H
tiền))))))
(. .))

Ví dụ 2: Tường được treo tranh.

(S-PSS(NP-SUB(N-H Tường))
(VP(V-H được)
(SBAR(S(NP-SUB *E*)
(VP(V-H treo)
(NP-DOB(N-H
tranh))))))
(. .))

2. Chức năng và chu cảnh cú pháp của động từ thực bị, được
- Với tư cách một thực từ, tức là từ mang ý nghĩa từ vựng đầy đủ các từ bị, được dễ dàng làm vị ngữ và có chu cảnh cú pháp sau đây:

+ Bỏ ngữ là một danh từ hay cụm danh từ

Ví dụ 1 : Con thỏ bị đạn .

(S (NP-SUB(Nc-H Con)(N thỏ))
(VP(V-H bị)
(NP-DOB(N-H đạn)))
(. .))

Ví dụ 2: Cậu bé được cái bút rất đẹp.

(S (NP-SUB(N-H Cậu bé))
(VP (V-H được)
(NP-DOB (N-H cái bút)

(AP(R rất)(A-H đẹp))))

(. .))

+ Bổ ngữ là một câu bị bao với hai điều kiện

- Chủ ngữ 1(của toàn câu) không chịu tác động của vị tổ 2 trong câu bị bao. Vị tổ 2 của câu bị bao có thể là động từ nội động hay ngoại động.
- Thực thể của chủ ngữ 2(của câu bị bao) không trùng với thực thể ở chủ ngữ 1 (chủ ngữ của toàn câu)

Ví dụ: Em này bị bố mẹ mất sớm.

(S (NP-SUB(N-H Em)(P này))

(VP (V-H bị)

(SBAR (S (NP-SUB (N-H bố mẹ))

(VP (V-H mất)

(AP(A-H sớm))))))

(. .))

Kết luận: Chức năng và chu cảnh cú pháp của 2 động từ thực “bị”, “được”:

- Làm vị ngữ; đứng trước bổ ngữ do danh từ hoặc cụm danh từ đảm nhiệm.
 - Làm vị tổ; đứng trước bổ ngữ do một câu đảm nhiệm với điều kiện:
- + Chủ ngữ của toàn câu không chịu tác động của vị tổ trong câu bị bao.
- + Thực thể ở chủ ngữ của toàn câu không trùng với thực thể ở chủ ngữ của câu bị bao.
3. Chức năng và chu cảnh cú pháp của động từ tình thái bị, được
- Các động từ tình thái đích thực có nét chung là chúng đứng trước một động từ khác có chủ ngữ là một thực thể trùng với thực thể ở chủ ngữ của câu.
 - Chức năng của 2 động từ tình thái bị, được trong câu diễn đạt tính tình thái được phép, bắt buộc, nằm trong phần tình thái của câu, chứ không giữ vai trò vị tổ diễn đạt sự thể như động từ thực từ. Vì vậy mặc dù động từ tình thái khi gắn nhãn treebank được gán là V nhưng không phải là động từ chính của vị tổ trong câu.
 - Mặt khác hai động từ tình thái này không thực hiện chức năng biến câu thành câu bị động như trợ động từ bị động.

Ví dụ 1 : Nó được đi xem kịch.

(S (NP-SUB (N-H Nó))

(VP (V được)

(VP (V-H đi)

(VP (V-H xem)

(NP-DOB (N-H kịch))))))

(. .))

Ví dụ 2: Bạn ấy bị ốm.

(S (NP-SUB (N-H Bạn) (P ấy))
 (VP (V bị) (VP (V-H ốm)))
 (. .))

Kết luận: Chức năng và chu cảnh cú pháp của 2 động từ tình thái “bị”, “được”.

- Làm yếu tố tình thái, không tham gia vào vị tổ.
- Đứng trước vị tổ là động từ nội động, ngoại động, tính từ hay quan hệ từ; các từ này có chủ ngữ trùng với chủ ngữ của bị, được

IV. Phân tích một vài cách dùng bị, được có thể gây lẫn lộn

Ví dụ 1: (A) Cầu thủ X bị phạm lỗi.(câu này dùng trong thuyết minh bóng đá hiện nay có quan hệ nghĩa với câu sau đây: Cầu thủ Y phạm lỗi đối với cầu thủ X.)

(B) Em này bị phạm lỗi chính tả trong bài viết.

Trong hai câu (A) và (B) thì câu (A) là câu bị động vì chủ ngữ của toàn câu khác với chủ ngữ của “phạm lỗi”. Nếu diễn đạt khác đi sẽ là “ Cầu thủ X bị cầu thủ Y phạm lỗi”. Về mặt nghĩa cầu thủ X là người bị hại. Ở câu (B) chủ ngữ của “bị” và “phạm lỗi” đều là “em này”. “Bị” là động từ tình thái, nó không giữ chức năng vị tổ của câu, có thể bỏ “bị” mà không làm thay đổi nghĩa sự việc của câu .

- Gán nhãn cho hai câu trên sẽ là
- (A) Cầu thủ X bị phạm lỗi.
 (S-PV (NP-SUB (N-H Cầu thủ X))
 (VP (V-H bị)
 (SBAR (S (NP-SUB *E*)
 (VP (V-H phạm)
 (NP-DOB (N-H lỗi))))))
 (. .))

- (B) Em này bị phạm lỗi chính tả trong bài viết.
 (S (NP-SUB (N-H Em này))
 (VP (V-H bị)
 (VP (V-H phạm)
 (NP-DOB (N-H lỗi) (N chính tả)
 (PP (E-H trong)
 (NP (N-H bài viết))))))
 (. .))

-Ví dụ 2 : (C) Các nhà báo được chất vấn.

Được đặt trong mối quan hệ nghĩa với câu

(D) Ông cố vấn bị chất vấn.

Trong hai câu trên thì câu (C) không phải là câu bị động. “Bị” là động từ tình thái, nó

không tham gia vào vị tổ của câu. chủ ngữ của toàn câu và của “chất vấn” là một “các nhà báo”. Còn câu (D) là câu bị động, chủ ngữ của câu là “ông cố vấn”, còn chủ ngữ của “chất vấn” là “các nhà báo”.

- Gán nhãn cho hai câu trên sẽ là
- (C) Các nhà báo được chất vấn.
 (S (NP-SUB (L Các) (N-H nhà báo))
 (VP (V được)
 (VP (V-H chất vấn)))
 (. .))
- (D) Ông cố vấn bị chất vấn.
 (S-PV (NP-SUB-1 (N-H Ông) (N cố vấn))
 (VP (V-H bị)
 (SBAR (S (NP-SUB *E*)
 (VP (V-H chất vấn)
 (NP-DOB-1 *T*))))))
 (. .))

Tài liệu tham khảo:

Tiếng Việt

- [1] Diệp Quang Ban. Ngữ pháp tiếng Việt. 2005. *NXB Giáo dục*.
- [2] Lê Biên. Từ loại tiếng Việt hiện đại. 1999. *NXB Giáo dục*.
- [3] Vũ Tiến Dũng. Tiếng Việt và ngôn ngữ học hiện đại sơ khảo về cú pháp. 2003. *VIET Stuttgart – Germany*.
- [4] Cao Xuân Hạo. Tiếng Việt sơ thảo ngữ pháp chức năng. 2006. *NXB Khoa học xã hội*.
- [5] Cao Xuân Hạo. Tiếng Việt mấy vấn đề ngữ âm ngữ pháp ngữ nghĩa. 2007. *NXB Giáo dục*.
- [6] Nguyễn Văn Hiệp. *Vài nét về lịch sử nghiên cứu cú pháp tiếng Việt. Tạp chí Ngôn ngữ, Hà Nội, số 10/2002*.
- [7] Nguyễn Văn Hiệp. Cơ sở ngữ nghĩa phân tích cú pháp. 2008. *NXB Giáo dục*.
- [8] Đào Minh Thu, Nguyễn Phương Thái. Thủ thuật phân tích câu và cụm từ. 2008. *Tài liệu nội bộ nhóm VTB*.
- [9] Nguyễn Minh Thuyết, Nguyễn Văn Hiệp. Thành phần câu tiếng Việt. 1999. *NXB Đại học Quốc gia Hà Nội*.

Tiếng Anh

- [1] Peter Sells. Lectures on Contemporary Syntactic Theories. 1987. *CSLI*.

- [2] Mitchell P. Marcus et al. Building a Large Annotated Corpus of English: The Penn Treebank. 1993. Computational Linguistics.
- [3] Fei Xia et al. Developing Guidelines and Ensuring Consistency for Chinese Text Annotation. 2000. COLING.
- [4] Nianwen Xue et al. Building a Large-Scale Annotated Chinese Corpus. 2002. COLING.
- [5] Chung-hye Han et al. Development and Evaluation of a Korean Treebank and its Application to NLP. 2002. LREC.
- [6] Sabine Brants et al. The TIGER Treebank. 2003. COLING.