

## 統語・意味解析コーパス（NPCMJ）チュートリアル

### セッション5

#### Tregex 検索式2

長崎 郁

2021.3.13

### セッション5の内容

1. 単純な関係表現
2. 関係表現における否定
3. いろいろな関係結合子
4. 複雑な関係表現
5. 調べたい表現のアノテーションを知るには
6. 練習問題

1

### セッション5の内容

#### 1. 単純な関係表現

2. 関係表現における否定
3. いろいろな関係結合子
4. 複雑な関係表現
5. 調べたい表現のアノテーションを知るには
6. 練習問題

2

### 単純な関係表現

ノード間の関係は、関係結合子によって指定することができる。

＜ノード＞ ＜関係結合子＞ ＜ノード＞

たとえば、

/<sup>^</sup>IP\b/ < VB

上記の検索式には関係結合子「<」が使われている。

これは、「IP（節）がVB（動詞）を直接支配すること」を表し、動詞述語節を抽出するときに使うことができる。

3

## 単純な関係表現

関係結合子を使った検索の結果では、検索式の左端のノード（マスターノード）がハイライトされる。

<ノード> <関係結合子> <ノード>  
マスターノード

4

## 練習問題1

検索対象ファイルを「396,404p」として行ってください

以下の検索式を使って、表示形式を graphical にして検索し、マスターノードがハイライトされることを確かめなさい。  
次に、表示形式を brackets にして検索し、同様にマスターノードがハイライトされることを確かめなさい。

/^IP\b/ < VB

5

## セッション5の内容

1. 単純な関係表現
- 2. 関係表現における否定**
3. いろいろな関係結合子
4. 複雑な関係表現
5. 調べたい表現のアノテーションを知るには
6. 練習問題

6

## 関係表現における否定

否定を表す「!」は、関係結合子の前に置くこともできる。

/^IP\b/ !< VB

上記の検索式は、「IP（節）がVB（動詞）を直接支配しないこと」を表す。たとえば、形容詞述語節、名詞述語節など、述語が動詞ではない節をまとめて抽出するときに使うことができる。

7

## セッション5の内容

1. 単純な関係表現
2. 関係表現における否定
3. いろいろな関係結合子
4. 複雑な関係表現
5. 調べたい表現のアノテーションを知るには
6. 練習問題

8

## いろいろな関係結合子

関係結合子には、支配関係、前後関係、姉妹関係、イコール関係など、様々な関係を表すものがある。

多くの関係結合子には順行・逆光の対があるため、マスターノードを柔軟に選ぶことができる。

9

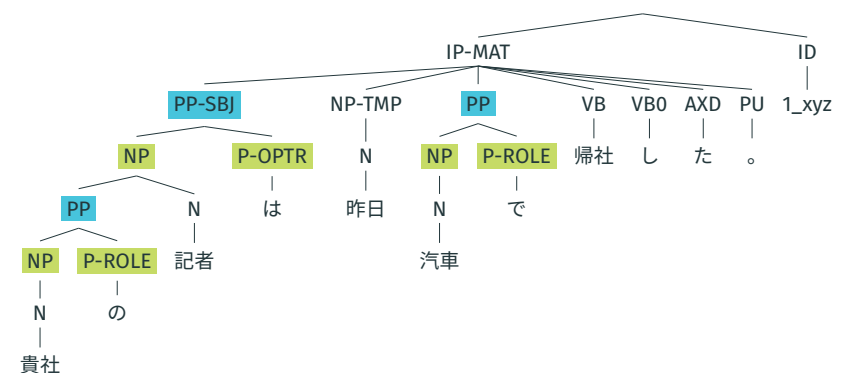
## いろいろな関係結合子 – 支配関係

A < B	AはBを直接支配する (AはBの親である)
A > B	AはBに直接支配される (AはBの子である)
A << B	AはBを支配する (AはBの先祖である)
A >> B	AはBに支配される (AはBの子孫である)

10

## 直接支配

/^PP\b/ < \_\_ (青がマスターノード)  
 \_\_ > /^PP\b/ (緑がマスターノード)

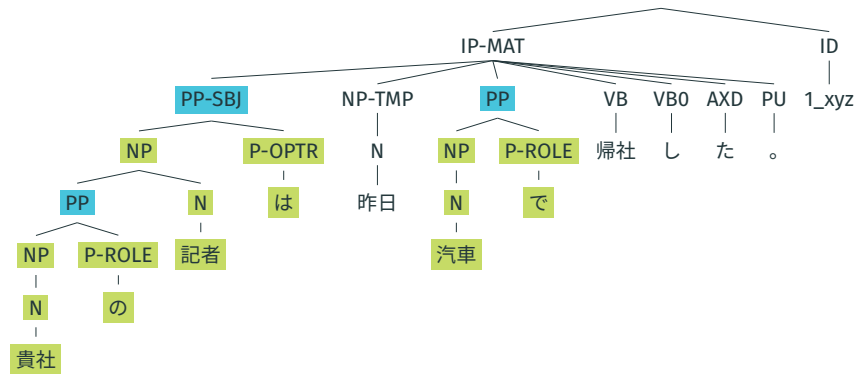


11

## 支配

/^PP\b/ << \_\_ (青がマスターノード)

\_\_ >> /^PP\b/ (緑がマスターノード)



12

## 練習問題 2

検索対象ファイルを 396,404p として行ってください

以下の 2 つの検索式を使って、表示形式を graphical または brackets のどちらかにして検索し、マスターノードが異なることを確かめなさい。

(1) VB < /^食べ/

(2) /^食べ/ > VB

13

## いろいろな関係結合子 – 前後関係

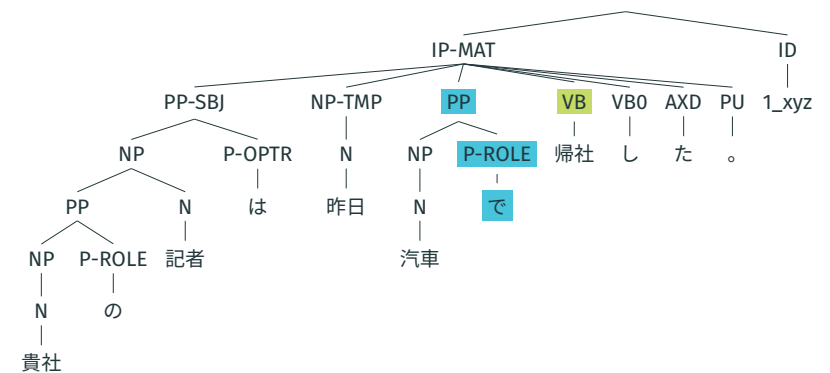
A . B	A は B の直前に置かれる
A , B	A は B の直後に置かれる
A . . B	A は B に先行する
A , , B	A は B に後続する

14

## 直前／直後

\_\_ . VB (青がマスターノード)

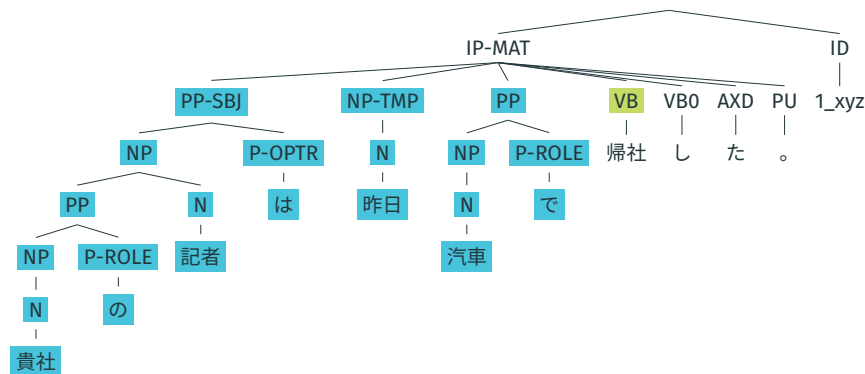
VB , \_\_ (緑がマスターノード)



15

## 先行／後続

\_\_ . . VB (青がマスターノード)  
VB , , \_\_ (緑がマスターノード)



16

## 練習問題3

検索対象ファイルを 396,404p として行ってください

(1) は、「て (or で)」が「もら (or 貰)」の直前に置かれること」を、(2) は、「て (or で)」が「もら (or 貰)」の直後に置かれること」を表し、どちらも、「行ってもらう」「休んで貰いたい」のような、動詞のテ形の直後に補助動詞「もらう」が続く例を抽出するときに使うことができる。

(1) て|で . /^(もら|貰)/

(2) /^(もら|貰)/ , て|で

2つの検索式を使って、表示形式を graphical または brackets のどちらかにして検索し、マスターノードが異なることを確かめなさい。

17

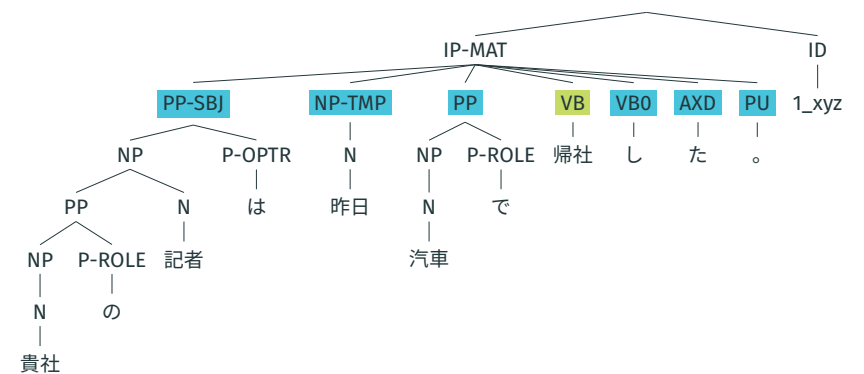
## いろいろな関係結合子 – 姉妹関係

A \$ B	AはBの姉妹である
A \$. B	AはBの直前の姉妹である
B \$, A	AはBの直後の姉妹である
A \$.. B	AはBの先行する姉妹である
B \$,, A	AはBに後続する姉妹である

18

## 姉妹

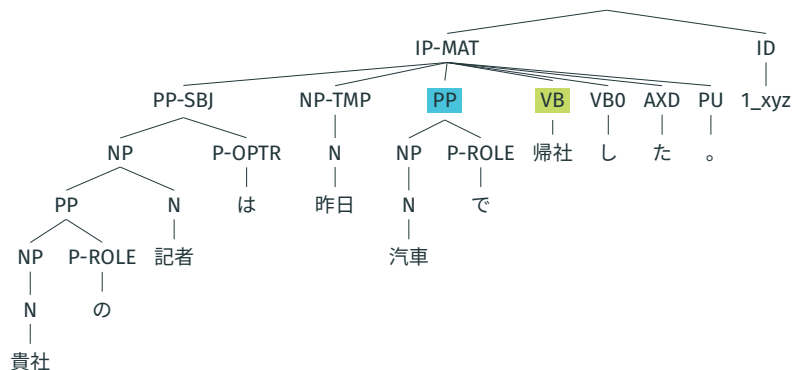
\_\_ \$ VB (青がマスターノード)



19

## 姉妹＋直前・直後

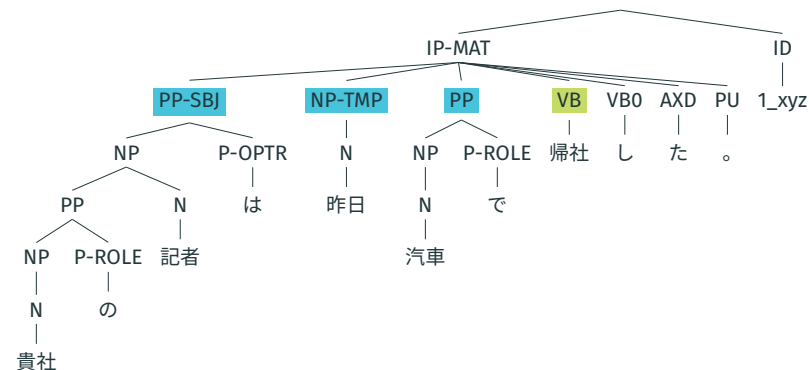
\_\_ \$. VB (青がマスターノード)  
VB \$, \_\_ (緑がマスターノード)



20

## 姉妹＋先行・後続

\_\_ \$. VB (青がマスターノード)  
VB \$, , \_\_ (緑がマスターノード)



21

## 練習問題 4

以下の2つの検索式がどのように異なるか、予想を述べなさい。

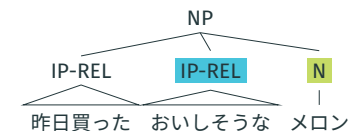
- (1) /<sup>^</sup>IP-REL\b/ \$. /<sup>^</sup>N\b/
- (2) /<sup>^</sup>IP-REL\b/ \$. . /<sup>^</sup>N\b/

/<sup>^</sup>IP-REL\b/ は関係節を、/<sup>^</sup>N\b/ は普通名詞を表す。

22

## 練習問題 4 - 解答

/<sup>^</sup>IP-REL\b/ \$. /<sup>^</sup>N\b/  
⇒「関係節が普通名詞の直前の姉妹である」



/<sup>^</sup>IP-REL\b/ \$. . /<sup>^</sup>N\b/  
⇒「関係節が普通名詞の先行する姉妹である」



23

## いろいろな関係結合子 – その他

A == B	A は B である
A <<, B	B は A の左端の子孫である
B >>, A	A は B の左端の子孫である
A <<- B	B は A の右端の子孫である
B >>- A	A は B の右端の子孫である
A <1 B (A <, B)	B は A の最初の子である
B >1 A (B >, A)	A は B の最初の子である
A <-1 B (A <-, B)	B は A の最後の子である
B >-1 A (B >-, A)	A は B の最後の子である
A <: B	B は A の唯一の子である
B >: A	A は B の唯一の子である
A <<: B	A は B のみを支配する (枝分かれがない)
B >>: A	B は A のみによって支配される (枝分かれがない)

24

## セッション5の内容

1. 単純な関係表現
2. 関係表現における否定
3. いろいろな関係結合子
- 4. 複雑な関係表現**
5. 調べたい表現のアノテーションを知るには
6. 練習問題

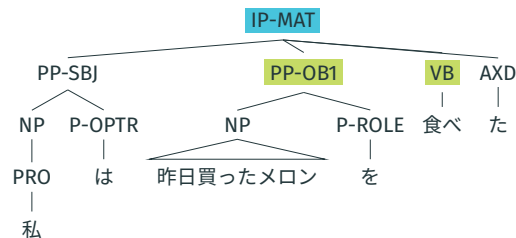
25

## 複雑な関係表現

ノード間の関係は2つ以上指定することもできる。

`/^IP\b/ < /^PP-OB1\b/ < VB`

「IP (節) が PP-OB1 (第一目的語である助詞句) を直接支配し、かつ VB (動詞) を直接支配する」



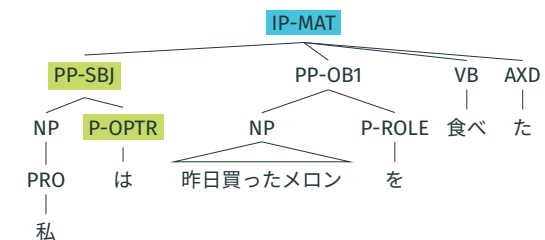
どちらの関係結合子も、マスターノードと、自身の後ろのノードとの関係を表すことに注意。

26

## 複雑な関係表現

`/^IP/ < (/^PP-SBJ/ < /^P-OPTR/)`

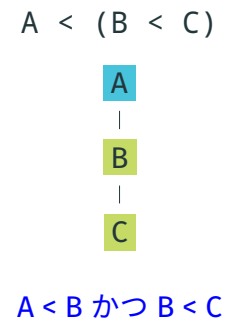
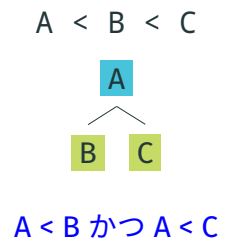
「IP (節) が PP-OB1 (第一目的語である助詞句) を直接支配し、かつ、その PP-OB1 が P-ROLE (格助詞) を直接支配する」



括弧で囲むと、その中の関係結合子は自身の直前のノードと直後のノードとの関係を表すことになる。

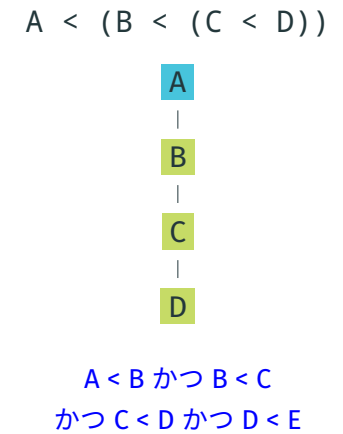
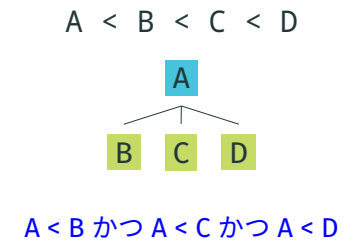
27

## 複雑な関係表現



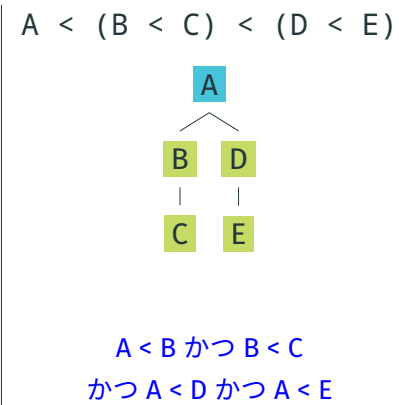
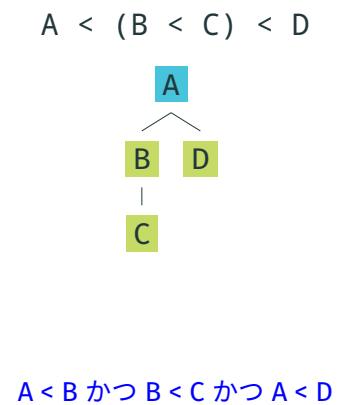
28

## 複雑な関係表現



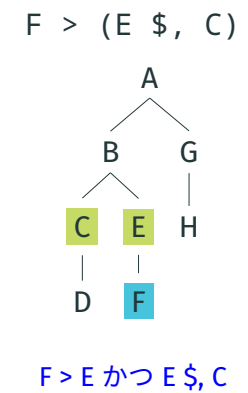
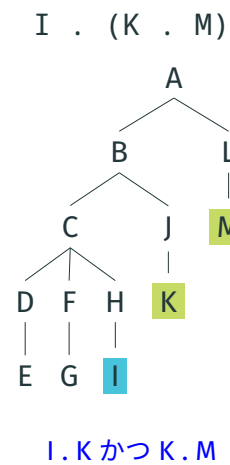
29

## 複雑な関係表現



30

## 複雑な関係表現



31



## 練習問題 5

「理にかなった理由」「私が筋トレをはじめたその理由」のような、「理由」が名詞修飾節を伴った例を抽出するための検索式を書きなさい。以下のツリーを参考にし、「理由」をマスターノードにすること。



32

## 練習問題 5 (解答例)

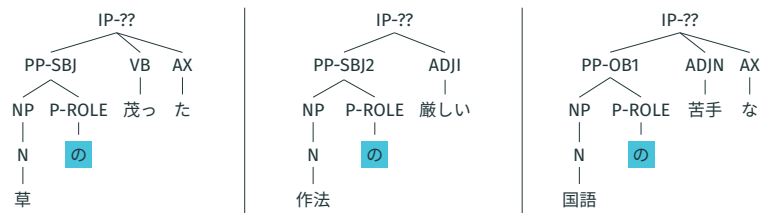
(a) 理由 > (N \$, , / ^ (IP-REL | IP-EMB) \ b / )

(b) 理由 > (N \$, , / ^ IP \ b / )

33

## 練習問題 6

主語や目的語をマークする「の」を含む例を抽出するための検索式を書きなさい。以下のツリーを参考にし、「の」をマスターノードにすること。



注意：主語をマークする拡張タグには「SBJ」と「SBJ2」があり、後者は二重主語文の第二主語を表す。目的語をマークする拡張タグには「OB1」と「OB2」があり、およそ、前者は直接目的語、後者は間接目的語を表す。

34

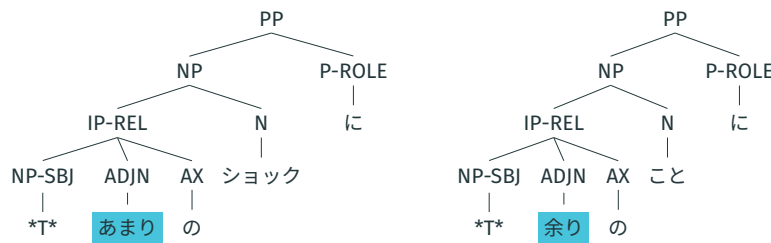
## 練習問題 6 (解答例)

の > (P-ROLE > PP-SBJ | PP-SBJ2 | PP-OB1)

35

## 練習問題7

「あまりのショックに」「余りのことに」に共通する「あまりの～に」を含む例を抽出するための検索式を書きなさい。以下のツリーを参考にし、青いノードをマスターノードにすること。



36

## 練習問題7 (解答例)

(a) あまり|余り . (の . . に)

(b) あまり|余り . の > (ADJN > (/^IP-REL\b/ > (/^NP\b/ \$.. (P-ROLE < に))))

37

## 練習問題7 (解答例)

(a) の検索式はシンプルで、手軽ですが、意図しない例も出てきてしまいます。

... 十人 余り の 見物 が ー かたまり に なって...

また、以下のように1つのツリーの中で、2回ヒットすることがあります。そのうちの一方は意図しない例です。

あまり の ショック に、彼女はその場に倒れ込んだ  
あまり の ショック に、彼女はその場に倒れ込んだ

(b) の検索式では意図しない例は出てきません。よって、検索結果の中に不必要な例が含まれていて欲しくない場合は (b) が適当ですが、どのような例があるかを手軽に見たいだけならば (a) を使う手もあります。

38

## セッション5の内容

1. 単純な関係表現
2. 関係表現における否定
3. いろいろな関係結合子
4. 複雑な関係表現
5. 調べたい表現のアノテーションを知るには
6. 練習問題

39

## 調べたい表現のアノテーションを知るには

調べたい表現の分節のされ方やタグ付けのされ方が分からないと、検索式が書けない！

具体的な表現を文字列検索で検索し、検索結果のツリーを見る

- 文字列検索は、NPCMJ Development Interfaces および NPCMJ Explorer で行うことができる。

NPCMJ Explorer には、TGrep-lite の検索式付きの文法項目がある。

- Tregex と少し書き方が違っているが、参考になるはず。

アノテーションマニュアルで NPCMJ のアノテーション方針の詳細や、具体例を知る。

- NPCMJ ウェブサイトのトップページからダウンロードできる。

40

## セッション5の内容

1. 単純な関係表現
2. 関係表現における否定
3. いろいろな関係結合子
4. 複雑な関係表現
5. 調べたい表現のアノテーションを知るには

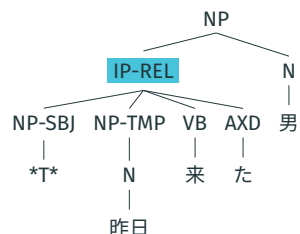
### 6. 練習問題

41

## 練習問題

以下の(8)～(18)を抽出するための検索式を書きなさい。

- (8) 関係節中の主語がトレース (\*T\*) であり、かつ述語が動詞である例



42

## 練習問題

- (9) 主語が空要素である例

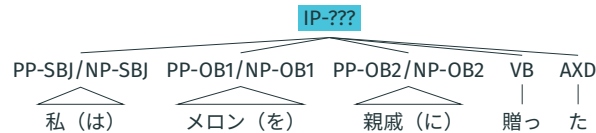


注意：空要素 (\*pro\*、\*speaker\*、\*hearer\*、\*arb\*、\*exp\*、\*T\* など) はすべて \* で始まり、NP に直接支配される。

43

## 練習問題

- (10) 主語 (SBJ) と第一目的語 (OB1)、第二目的語 (OB2) がすべて明示された例

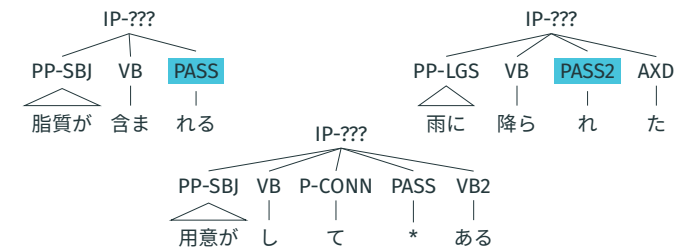


注意：空要素を直接支配する SBJ、OB1、OB2 を排除する必要がある。

44

## 練習問題

- (11) 受動文



注意：直接受身文の助動詞は PASS と、いわゆる間接受身文や持ち主の受身文の助動詞は PASS2 とタグづけされている。PASS はまた、テアル構文で受身文のような項の昇格が起こる場合にも、\* を直接支配する形で使われているので、このようなケースを排除する必要がある。

45

## 練習問題

- (12) 主語がトレース (\*T\*) ではない空要素である例 (主語省略の例)  
 (13) 副詞句 (ADVP) が主語 (SBJ) より前に現れた例  
 (14) 第一目的語 (OB1) が格助詞「が」を伴った例  
 (15) 動詞 (VB) の主語 (SBJ) が格助詞「に」を伴った例  
 (16) 使役の助動詞「せる・させる」を含む例  
 (17) 「死ぬほど」「多ければ多い程」のような、助詞「ほど」を伴った副詞節の例  
 (18) 「来てくれる」「読んでくれた」「教えてはくれない」「手伝っておくれ」のような、補助動詞「くれる」が動詞のテ形を伴う例 (テ形と「くれる」の間に助詞が挟まれる場合のあることに注意)

46

## 練習問題 – 解答例

- (8) a. /<sup>^</sup>IP-REL\b/ < (/<sup>^</sup>NP-SBJ\b/ < /<sup>^</sup>\\*T\\*/ ) < VB  
 b. /<sup>^</sup>IP-REL\b/ < (/<sup>^</sup>NP-SBJ\b/ < /<sup>^</sup>\\*T\\*/ \$ VB)  
 (9) /<sup>^</sup>NP-SBJ\b/ < /<sup>^</sup>\\*/  
 (10) /SBJ/ !< /<sup>^</sup>\\*/ \$ (/OB1/ !< /<sup>^</sup>\\*/ ) \$ (/OB2/ !< /<sup>^</sup>\\*/ )  
 (11) a. PASS|PASS2 !< /<sup>^</sup>\\*/  
 b. /<sup>^</sup>PASS/ !< /<sup>^</sup>\\*/

47

## 練習問題 – 解答例

(12) /<sup>^</sup>NP-SBJ\b/ < /<sup>^</sup>\\*/ !< /\\*T\\*/

(13) /<sup>^</sup>ADVP\b/ \$.. /SBJ/

(14) /<sup>^</sup>PP-OB1\b/ < (P-ROLE < が)

(15) /<sup>^</sup>PP-SBJ\b/ < (P-ROLE < に) \$ VB

(16) /<sup>^</sup>(せ|させ)/ > VB2

(17) ほど|程 , /<sup>^</sup>IP-ADV\b/

(18) VB2 < /<sup>^</sup>(くれ|呉れ|おくれ|お呉れ)/ \$,, (P-CONN < て|で)

(「言ってくれば」のような不必要な例も出てきてしまいます)