

# 連用・連体接続に基づいた 係り受け構文解析\*

1 Q-3

細井 貴晴 小島 丈幸 山口 昌也 乾 伸雄 小谷 善行 西村 恕彦  
(東京農工大学 工学部 電子情報工学科)

## 1 はじめに

構文解析をおこなうさいの構文規則は、句構造文法や係り受け文法で記述される。英語文の構文解析の場合、多くは、句構造文法での記述がなされるが、日本語文の構文解析は、係り受け文法の考え方でおこなわれることがある[1]。そして、本研究でも、日本語文における係り受けの考え方での構文規則の提案をおこなう。

日本語文の文節は、二つの属性をもつ。一つ目として、そのものの性質を考える。ある文節は体言的か用言的な働きをになう。二つ目は、接続の仕方を考える。係り方には、連用接続か、連体接続、もしくは終止接続がある。接続詞、感動詞、副詞、連体詞などのような係る文節がないものもある。これらの日本語文の構造を考慮した構文規則の提案である。この提案で、もっとも主張する点は、文法規則が体系的で、一つにまとまる点である。

## 2 構文規則

### 2.1 文節の種類

1節で述べた考えの下で、文節の種類を定義する。日本語の特性として、句の単位には、二つの属性がつけられる。

- 句そのものの性質
- どのような句に接続するか

まず、句そのものの性質としては、用言的な働きをするもの、体言的な働きをするもの、終止的な働きをするものがある。動詞や形容詞は用言であり、名詞は体言であることは自明である。副詞や連体詞、接続詞、

感動詞は、多くの場合係られることがないので、終止的であると考えられる。

そして、どのような句に接続するかを考えると、接続のしかたは、用言的であり、体言的であり、終止的である。例をあげて考えると、連体詞は体言に接続し、副詞のほとんどは用言に接続する。どの品詞にも接続のすることのない感動詞などは、終止的であるといえる。

以上により、表1に示す九つの文節の種類が考えられる。文節の種類にはどのようなものがあるか、その一部を表1に示す。

表 1: 文節の種類

文節の種類	文節の形の例	文節の例
体言終止句	普通名詞 + 助動詞	僕だ
体言連用句	普通名詞 + 格助詞	僕が
体言連体句	普通名詞 + 接続助詞	僕や
用言終止句	動詞の終止形	眠る
用言連用句	動詞の連用形 + 格助詞	眠りは
用言連体句	形容詞の連体形	きれいな
終止終止句	接続詞, 感動詞	はい
終止連用句	副詞	とても
終止連体句	連体詞, 指示詞	あの

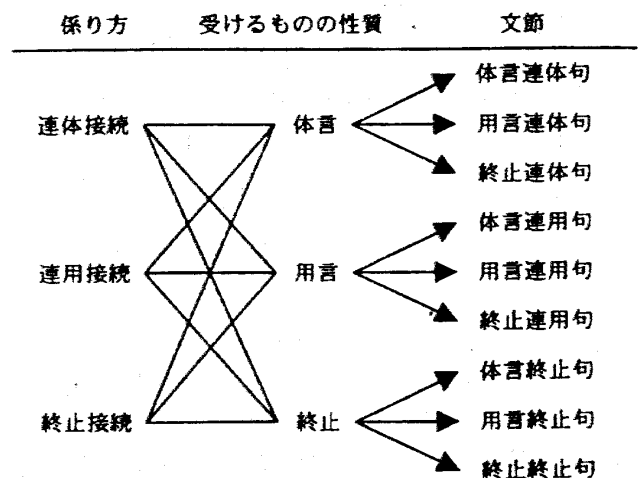


図 1: 文節を成す係り受け関係図

\*Syntactic analysis based on link of declinable words and link of uninflected words, Takaharu HOSOI, Takeyuki KOJIMA, Masaya YAMAGUCHI, Nobuo INUI, Yoshiyuki KOTANI, Hirohiko NISHIMURA, Tokyo University of Agric. and Tech., Dept. of Computer Science

## 2.2 構文規則

2.1節での文節の種類をDCGの形式で記述する。また、図1に文節の種類における係り受け関係を示す。

### 制約規則

- $A, B, C \in \{\text{体言, 用言, 終止}\}$
  - $s(A, B) = \text{文節}$
  - $t(A, B) = \text{品詞の列}$
- $s(A, B) \rightarrow s(C, A), s(A, B), \{A \neq \text{終止}\}.$   
 $s(A, B) \rightarrow t(A, B).$

## 2.3 解析例

単純な文を用いて、解析の流れを示す。  
 “僕の本は、おもしろい。”

### 2.3.1 解析の手順

- (a) ある一つの文に対して、品詞情報から文節の認識をおこなう
- (b) 文節の係り先を調べる
- (c) 規則に当てはめる

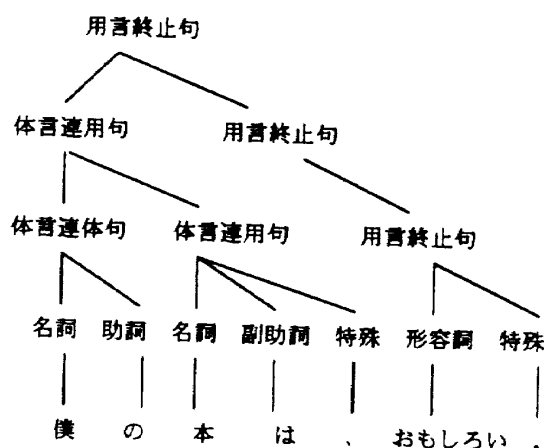


図2: “僕の本は、おもしろい”の解析結果

(a)において、“僕の”、“本は”、“おもしろい”をそれぞれ体言連体句か体言連用句、体言連用句、用言終止句であると認識する。(b)において、読点の情報を吟味して“僕の”は“本は”に係ることが決まり、“僕の”の文節を体言連体句とする。(c)において出される結果は、図2のようになる。

## 3 文節の認識

表1に示したような文節の種類を作成を用例に基づいておこなう。品詞の情報や係り受けの関係を獲得するためのデータは、次の3.1節に示すものを用いた。

### 3.1 文節の形を決めるデータ

文節の形を決めるデータは、[2][3]のものを用いた。それらは、毎日新聞の記事を京都大学の形態素解析システム(JUMAN)、構文解析システム(KNP)で自動解析を行い、その結果を人で修正されたものである。よって、ここでの品詞の区別は、JUMANの品詞体系に基づいている。

### 3.2 品詞の分類

2節の考え方に基づいた品詞の分類をおこなう。

係る側のタイプは用言的なもの、体言的なもの、終止的なものの三種類のものに分ける。用言的なものは動詞・形容詞を含んだ文節、体言的なものは名詞を含んだ文節、終止的なものは、それ以外の品詞とする。接頭辞・接尾辞は、動詞・形容詞・名詞につくかで、それらのつくものをそれぞれ、動詞・形容詞・名詞で扱う。

受ける側のタイプも係る側の考え方と同じように考え、分類をおこなう。

## 4 おわりに

体系的な構文規則を提案した。

この規則を用いて構文解析をおこなうことの利点は、文法規則が非常に体系的でわかりやすいので、確率情報などの情報が取り入れやすいことなどある。

## 参考文献

- [1] 秋山典文, 藤井敦, 徳永健伸, 田中穂積: 形態素で残る曖昧性を考慮した日本語文の係り受け解析, 言語処理学会第1回年次大会発表論文集, pp.129-132, 1995.
- [2] 京都大学テキストコーパス Version1.0, 1997.
- [3] コーパス作成の作業基準 Version 1.4, 黒橋禎夫, 斎藤由衣子, 坂口昌子, 1997.