

# 提題機能を考慮した 日本語文における依存構造の解析

上原 悟      牧野 武則  
東邦大学理学部情報科学科

## 1 はじめに

依存文法 [1] は、導きだす依存構造が、格文法で表現した意味構造に類似している非終端記号のように人工の記号を用いなくて語と語の依存関係を表すなど利点が多い。実用のシステムでも採用され、その有用性が検証されつつある。また、GPSG など単一化文法 [2] では語彙項目に語彙に係わる構文情報を集め、単一化操作で簡素な解析を実現している。

語彙依存文法 [4] は、言語に共通の係り受けのモデルを設定し、語彙に依存する依存情報を語彙項目に集中し、さらに依存構造をネットワークで表現することで、簡素で柔軟な解析、生成の実現を目的としている。しかし、実際には、大局的な構造の把握に適さない [5]、規則の自由度のため係り受けのスコープが限定できない等の問題が起こったため、日本語の持つ基本構造を経験的に得た小数の規則を用いて依存構造を導出する方法を示した [6]。

今回は、提題化された語の特性を考慮した係り受け構造を依存文法をもとに依存構造を導出する方法を提案する。

## 2 係り受けモデル

依存文法では、語と語の依存関係から文の構造が決められる。依存関係にある 2 つの語の一方を主要素、他方を従要素と呼ぶ。主要素と従要素の関係は係り受け情報として与えられる (図 1)。この係り受け情報は各語の情報、つまり語彙項目によって決定される。

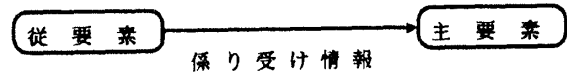


図 1: 依存関係と係り受け情報

## 3 語彙項目

要素から得られる情報を語彙項目といい、大きく 2 つに分けられる。一つはその要素が主要素となる場合、どのような従要素に係り受け可能であるかの情報 (主要素情報)、もう一方は従要素である場合、どのような主要素と係り受け可能であるかの情報 (従要素情報) である。主要素情報は、品詞と格情報であり、それらに対応する従要素情報として、修飾と格がある。

## 4 提題の機能

提題化された要素は、文の表現の中で、ある主題を掲げてそれについて何か述べるという特徴がある [3]。また、文中のある格が提題化されて、名詞+「は」の形をとる。複文においては、いくつもの述語要素が主題を要求していると考えられる。したがって、提題化された要素は、要求されるだけ係り先が存在することを考慮すべきである。文中の全ての述語要素に係り受けるものではなく、いくつかの規則が見受けられた。

- 文中の動詞 (連用) が提題を要求している
- 提題は、「が」格が提題化したものが多い
- 1文1格をもとで「が」格が充足格に存在するなら、その動詞には、係り受けはしない

また、名詞、用言に意味記述を付加することで、提題の係り先チェックをより精妙な仕組みとしている。

## 5 構文解析

日本語の構造は、基本的に前の語が後ろの語へと係り受けていく左下がりの文構造になる。さらに、

Analyzing Dependency Structure with topics  
in Japanese  
Satoru Uehara, Takenori Makino  
Department of Information Science,  
Faculty of Science, Toho University  
2-2-1 Miyama, Funabashi, Chiba 274

No	要素	エントリ	品詞	機能	活用
1	私は	私	名詞	提題	*
2	彼女に	彼女	名詞	に	*
3	会い	会う	動詞	*	連用
4	、	、	特殊	、	*
5	事の	事	名詞	*	連体
6	真相を	真相	名詞	を	*
7	知った	知る	動詞	*	終止
8	。	。	特殊	。	*

表 1: 要素リストの一部

提題は、要求されるすべての要素に係り受け可能である。構文解析の入力としては、形態素解析され、要素に語彙情報が付加されている要素リストを用いる。表1に、例文「私は彼女に会い、事の真相を知った。」の要素リストを示す。

次に、この要素リストに前処理を加える。まず、品詞とエントリから動詞タイプを付加し、次に、機能と活用から修飾と格を決定する。また、動詞タイプに対し、受けることが可能な格を示した要求格情報を与える。同時に、それが対応する充足格情報も持つとする。

構文解析は、まず、日本語の基本構造にしたがって、前の語から後ろの語へと順次係り受けていくように係り受けを仮設定する。次に、要素リストのトップの語から係り受けをチェックして行き、特殊規則に合致する場合は、規則に基づき決定する。最後の語の係り受けが決定された時点で、提題の係り先をチェックする。そして、提題の係り受け構造が決定したら解析終了となる。

特殊変換規則とは、経験的に発見した変換規則に基づくもので、文構造に大きな変更が加わるときに用いられる。

次に、例文を使って、実際の解析の様子を示す。係り受けを仮設定したものが図2、第一段階で提題を考慮してないものが図3、提題を考慮した最終結果が図4になる。第一段階での解析で、「会い」は、「、」に係っているが、「、」に関する特殊処理により、「知った」に係り先が変更され、係り受けがチェックされる。提題以外の要素チェックを終えたら、提題のチェックを開始する。「私は」は、提題の機能を持っているため、文中の動詞（連用）全てに係り先が変更され、係り受けがチェックされる。ここでの係り受けチェックは、動詞の充足格情報と意味記述

情報が考慮される。

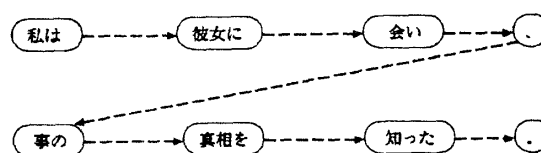


図 2: 係り受けを順次設定したネットワーク

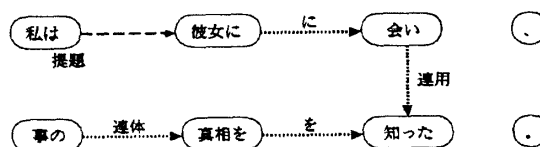


図 3: 提題を考慮してないネットワーク

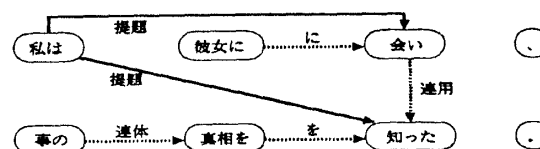


図 4: 構文解析結果を示すネットワーク

## 6 実験

提案した方法を用いて、新聞記事などから得た文に対し解析を行なった。構文上解釈可能な文を正解とすると、正解率は84%であった。これは、主に「が」格によって提題化された要素が多い事を裏付けている。しかし、実際には、「を」格「の」格なども提題化されうるので、柔軟な解析が必要である。

## 7 終わりに

本論では、日本語の基本構造を前提とした語彙依存文法による構文解析方法に新たに提題の機能を考慮した方法を提案した。この方法は、提題の依存関係を明確に示すものである。今後は、今回得た問題点をもとに、規則の見直し、改良を行ない柔軟なシステムにしていく予定である。

## 参考文献

- [1] 児玉「依存文法の研究」研究社出版(1986)
- [2] 黒川「自然言語処理入門」近代科学社(1988)
- [3] 益岡、田窪「基礎日本語文法」くろしお出版(1992)
- [4] 牧野、納富「語彙依存文法について」情報処理学会自然言語処理研究会、82-14(1991)
- [5] 河田、牧野「語彙依存文法における語彙項目の記述について」情報処理学会第49回全国大会 3G-8(1994)
- [6] 酒井、上原、牧野「日本語の基本係り受け構造をもとにした依存構造による解析」情報処理学会第53回全国大会 2H-2(1996)