

## 3 F - 1

## 日本語発話からの照応情報の抽出

吉本啓 島津明

NTT 基礎研究所

## 1 はじめに

照応情報(主題、前提、時制)は各発話行為について一定であることにもとづいて、これらの情報を発話から抽出して解析結果を得る方法について述べる。複文が1つの発話行為を表す場合には照応情報が従属節・主節間で一致するよう解析される。より独立性の強い従属節については、主節とは異なる発話行為として扱われる。その際、主節と従属節の関係は別に明示される。これにもとづいた HPSG/JPSG による解析結果は部分 DRS (談話表示構造)として与えられ、談話全体の DRS を構成するための入力となる。

## 2 文の4階層

日本語の文は4つの階層から構成される [4]。

A ::= 様態副詞 + 補語 + 述語 + (否定)

B ::= 制限的修飾句 + 主格 + A + (否定) + 時

制

C ::= 非制限的修飾句 + 主題 + B + モーダル

D ::= 呼掛け + C + 終助詞

A < B < C < D の順で、上位の述語構成語(補助動詞・助動詞・終助詞)は下位のそれをサブカテゴリー化するが、逆はありえない。従属節にもこれに関係する。A~Cの各々の節をサブカテゴリー化してできる従属節はそれ自体各節の構成要素となる。例えばB類をサブカテゴリー化する「カラ」をヘッドとする従属節はB類の構成要素である制限的修飾句と同様に機能する。

他に、上の階層はさまざまな統辞的現象と関連することが指摘されている。

## (1) 主題・主語

B類従属節(B類の節をサブカテゴリー化する接続助詞「ト」「カラ」等をヘッドとする節)を伴う文において、主語が主節・従属節ともに非主題の場合は2つの主

語は一致しなくてよいが(例1)、主題の場合は一致する(例2)。C類従属節の主語および主題は主節から独立している。

1. [春子<sub>i</sub>が卒業した]<sub>B</sub> から  $\phi_j$  結婚した。
2. 春子<sub>i</sub>は [ $\phi_i$  卒業した]<sub>B</sub> から結婚した。

## (2) 前提

質問・否定・推量の作用域(焦点の起こりうる範囲)にA類従属節は含まれる。B類従属節はそのままでは作用域(< >)外だが(例3)、主節述語が「ノカ」「ノデハナイ」「ノダロウ」で終わると作用域内となる(例4)。C類従属節はつねに作用域外である。

3. 春子は [結婚した]<sub>B</sub> から < 就職した > なかった。
4. 春子は < [結婚した]<sub>B</sub> から就職した > のではない。

## (3) 時制

B類従属節の時制はその基準を主節の表す出来事の発生時に負っている。例1で、B節中の「タ」は主節の「結婚した」の表す時点を基準時としてその時点で完了していることを表す。C類従属節は概ね時制に関して主節から独立している。

## 3 照応情報の抽出

以上の諸点が統辞構造の4階層と関連することはこれまで指摘されてきたが、なぜそうなのかについて理論的に説明されたことはなかった。これは、階層のうちC類までが判断伝達の発話行為の単位をなし、その中で(1)~(3)の情報が一定でなければならぬと考えることによって解決される。B類従属節はC類の節に含まれるので、これらの情報に関して全体と一致しなければならない。C類従属節は主節とは独立して1つの発話行為を表す。

(1)~(3)の情報は、従属節・主節間の一致・不一致が問題になる、談話全体との一致が問題になる、との2点

[[HEAD [[POS ADV][FORM KARA]	
COH [[HEAD [[POS V]	
MODL [[BODY -][CAUS -][DEAC -][ASPC -]	
[DONT -][OPTT -][PAST -][EVID -]	(2-1)
[TENT -][POLT -][SPAA -]]]]]	
SUBCAT {[[HEAD [[POS P][GRF SUBJ]]][SEM ?X1]]}	(1-1)
PRAG [[ANAP [[TOP ?X1	
[RTIME ?X2]	(3-1)
[ETIME ?X3]]]]]]]]]	
SUBCAT {[[HEAD [[POS V]]][SUBCAT {}]	
[SLASH {[[HEAD [[POS P][GRF SUBJ]]][SEM ?X1]]}]	(1-2)
[SEM [[CONT ?X4]]]	
PRAG [[ANAP [[PRSP ?X5	(2-2)
[RTIME ?X3]]]]]]]]]	(3-2)
[SEM [[CONT ?X4]]]	
PRAG [[ANAP [[PRSP ?X4 U ?X5]]]]]]]	(2-3)

図1 接続助詞「カラ」の語義記述

が代名詞等照応表現と共通なので、「照応情報」として一括する。

Head-Driven Phrase Structure Grammar[5]/Japanese Phrase Structure Grammar[1]にもとづく、(1)~(3)の一致を規定したB類接続助詞「カラ」の語義記述を図1に挙げる。(1-1),(1-2)で主節・従属節間での主語・主題の一致を扱う。(1-1)の主題化された主節主語が(1-2)の省略された従属節主語と一致することが示されている。主題が文中にあらわれない場合については「カラ」の別の語義記述で扱う。

(2-1~3)で前提の一致を扱う。(2-1)で係り先の主節を制限し、「カラ」従属節が作用域外となる。そのため、従属節の意味内容が文全体の前提に含まれる。一般にCOH素性でコンプリメントのヘッド(ここでは主文述語句)に対する制約を表す。(2-2)で示された、サブカテゴリーライズされる節の意味である?X4、その前提である?X5の和が(2-3)で従属節としての前提であるとされ、文全体の前提を計算するために利用される。

(3-1),(3-2)は時制の一致を扱う。主節の発生時である?X3が従属節の基準時と一致することが示される。

以上を踏まえて、各文の解析結果は通常の Universe (discourse referentsの集合) および Conditions (それに対する条件の集合) の他に照応情報をも構成要素とする一種の部分DRS[2]として与えられる。C類複合文と引用文を伴う文については、主節と従属節にそれぞれ対応するDRSが関係づけられた複合DRSとして扱う。

#### 4 DRSの構成

以上の方法による各文の解析結果を入力としてDRSを構成する。構成規則の詳細は他に譲る。各文の解析結果から、発話行為を単位として、ユニフィケーションにもとづいてDRSを形成[3]する。照応情報は、主題の導入と継続、それによるゼロ代名詞の同定や、文の前提が談話モデル内の知識に適合するかどうかによる発話行為としての適切性のチェックに利用される。

#### 5 むすび

日本語の文脈処理の一環として、発話行為を単位として照応情報を抽出し、DRSの構成に利用する方法について述べた。

#### 参考文献

- [1] Gunji, *Japanese Phrase Structure Grammar*. 1987.
- [2] Kamp, H. *A Theory of Truth and Semantic Representation*.
- [3] Maeda, H. et al. *Parsing Japanese Honorifics in Unification-Based Grammar*. 26th ACL, 1988.
- [4] 南不二男. 『現代日本語の構造』大修館. 1974.
- [5] Pollard, C. and I. Sag. *Information-Based Syntax and Semantics*. vol.1. CSLI. 1987.