

## 係り受け解析のための日本語述語句の細分類

白井 諭<sup>\*1</sup> 畑山満美子<sup>\*1</sup> 木村 淳子<sup>\*2</sup> 十河 則子<sup>\*2</sup> 横尾 昭男<sup>\*3</sup> 池原 悟<sup>\*4</sup>

<sup>\*1</sup>NTT コミュニケーション科学研究所

<sup>\*2</sup>NTT アドバンステクノロジー

<sup>\*3</sup>ATR 音声翻訳通信研究所

<sup>\*4</sup>鳥取大学 工学部

### 1 はじめに

日本語の構文解析の精度向上には、体言と用言の関係を精密に分類することが必要である。このうち用言の相互関係に関して、従属節の3分類とその相互制約に関する言語学の成果が知られている[南 64,74]が、分類には意味的分析を必要とするなど、日本語の構文解析への適用は困難であった。これに対して、筆者らは、先に、表層的な情報から述語句を3分類するとともに、係り受けの経験則等を加味することにより、述語句を都合 52 種類に分類する方法を提案し、用言の相互関係の高精度な解析が可能であることを示した[白井 95]。本稿では、先の分類において述語句の関係認定を誤るものについて重点的に検討する。そして、一部の動詞の振舞いを考慮した述語句の細分類を提案する。

### 2 述語句の基本分類

まず、南不二男が指摘した従属節の性質を紹介し、それを応用した係り受け解析のための従属節の分類法について述べる。

#### 2.1 従属節の性質

南は、述語表現に現れる助動詞や助詞の種類の違いに着目して、従属節を次の 3 種類に分類した[南 64,74]。

- A 類：「～しながら」等。
- B 類：「～たら、～と、～なら、～ので、～のに、～ば、～て (従属的用法)」等。
- C 類：「～が (順接)、～から、～けれど、～し、～て (独立した用法)」等。

南は、さらに、上記の3種類の従属節間に、下記

の強い傾向があることを指摘した。

- A 類は、他の A 類、B 類、C 類の一部となることができる。
- B 類は、他の B 類、C 類の一部となれるが、A 類の一部とはなれない。
- C 類は、他の C 類の一部となれるが、A 類、B 類の一部とはなれない。

#### 2.2 係り受け解析向きの分類法[白井 95]

##### (1)従属節の種類

言語処理の立場から節を分類すると、主節のほか、連用節、連体節、引用節の3種の従属節に分けられる。このうち主節は、他の節の係り先にはなるが係り元にはなり得ない。連体節は述語の活用形などの文法的性質によって係り先が明確に決まる。引用節は引用の助詞等を伴うことが多く、係られる側の動詞のタイプが限定されるなど、形態的にその係り先がほぼ明確である。従って、述語間の係り受け解析で問題となるのは、係り元が連用節である場合、すなわち、連用節から連用節へ、連用節から連体節へ、連用節から引用節への3つの係り受け関係である。

##### (2)従属節の再分類

南が指摘した関係を、従属節の中心的要素である述語間の係り受け決定に利用することを考える。南の分類では、接続が順接か逆接かなど、意味的な判断を必要とするが、構文解析段階でそれを判断し、分類するのは困難である。そこで、南の分類の主旨を生かしながら、語尾表現をより長単位で分類した「述語句」を分類の基本単位とし、意味的判断の困難な表記はデフォルトの解釈で分類することなど

により、従属節の述語句を次のように再分類した。

- A: 「～しつつ」, 「～することに加え」など, 「同時」の表現。
- B: 「～するので」などの「原因」, 「～して」, 「～することで」等の「中止」の表現。
- C: 「～するし」などの「独立」の表現。

この分類では、表層的に明らかに A または C と判定できるもの以外は B に分類した。従って、南の分類に比べ B の範囲が広がっているが、分類相互の包含関係の傾向は保存されていると期待される。なお、以下では特に必要ない限り「述語」を「述語句」の意味で使用する。

### (3) 読点の有無による分類

述語間の係り受けに曖昧さが生じるのは、2 つ以上の従属節がある場合である。そこで、新聞記事からそのような文を抽出して、各従属節の出現頻度を調べたところ、A が 10% 余り、B が 80% 余り、C が 5% 余りで、圧倒的に B が多いことが分かった。

ところで、書き手が解釈しにくいと思った文には読点を入れる傾向があり、係り受け解析では「読点を伴う文節は直後には係りにくい」という経験則として利用されている。そこで、従属節を含む長文ではその傾向が強いのに着目して、従属節 A, B, C の分類に、読点の有無を加味した 6 種類の分類を考える。書き手が読点を付与した従属節は、それだけ遠くに係る可能性が強い、すなわち、独立性が強いと考えられるから、それぞれの従属節の包含関係は、

A C 「A+読点」 C B C 「B+読点」 C C C 「C+読点」

となることが予想される。

新聞記事の場合、B の従属節のうち、読点を持つものと持たないものは、ほぼ同数である。従って、読点を考慮したことにより、B 相互の係り受けの約半数は、曖昧なく決定できると期待される。

### (4) 連用節の中止性

上記の分類では、B 同士, 「B+読点」同士の係り受け関係は決定できない。これらの語尾の表現は、

新聞記事を対象にした調査では、ほぼ下記の形式に限定される。

～し (用言連用形), ～して, ～ (する/した) もので, ～ (する/した) ことで, ～しており, (名詞) で, その他

これらのうち, 「し」「して」のタイプは中止性が弱く, 行為の連続, 並行, 順序等の意味に使用される傾向が強い。それに対して, 「もので」「ことで」「しており」のタイプは前置き的な内容を表し, 主題の変わり目に使用される中止性が強い表現であると考えられる。中止性の強い述語ほど遠くに係る傾向があると推定されるから, 以下のヒューリスティックスを導入する。

- 連用節述語のうち, 中止性の強い B 類の述語は, 他の連用節述語を飛び越える。逆に, 他の述語は, 次に現れた連用節述語に係る。
- 連用形の単独型は, 他の連用形の単独型を飛び越えないで, それに係る。

### (5) 述語の状態性と動作性

連用節のうち, B 類同士, 「B+読点」類同士の係り受けを決めるため, 次に, 述語の状態性もしくは動作性に着目する。述語は, その動作性から見ると, 動作性の強い順 (状態性の弱い順) に, 他動詞性, 自動詞性, 形容詞性, 名詞性の 4 種の述語に分類することができる。但し, 使役系の表現は他動詞性, 受身系の表現は自動詞性とする。

読者に分かりやすくするため, 書き手は主題や動作主体を統一的に捉えて表現する傾向があるため, 動作性の述語と状態性の述語が同一レベルで表現されることは少ない。また, 状態性の強い述語は, 動作性の強い述語に包み込まれる傾向を持つ。これらの傾向に着目して, 以下のヒューリスティックスを設ける。

- 動作性の強い述語は, 動作性の弱い述語を飛び越し, 動作性の弱い述語は, 動作性の強い述語に係る。

ところで, 連用節の中には, 用言が格助詞相当語の一部として使用される「～と比べ (～より) 高い

値段」のようなものがある。このタイプの連用節は、他の述語の格要素として使用されたものである。そのため、述語間の係り受けとしてではなく、格関係に準じて扱う必要がある。

#### (6)引用節と連体節

連用節と引用節の関係を見ると、「～すると（発表する）」などの引用節は独立性が高いため、連用節が引用節を飛び越えて他の節に係ることは考えにくい。しかし、「～するよう（依頼する）」など、様態化し独立性が弱められた引用相当節の場合は、飛び越えられる可能性もある。そこで、次のルールを設けた。

- 引用節述語は、連用節述語の「C+読点」に準じる。
- 引用相当節述語は、「B+読点」に準じる。

次に、連用節と連体節の関係を見る。連体節を形式名詞「もの、こと、の」を伴うもの（～したもので、～することで、など）とその他の名詞性のものに分けて考えると、形式名詞を伴う連体節は、対象を捉え直す[三浦 75]ために形式名詞が使用されたとも考えられるので、連用節がこれを越えて他の述語に係ることは考えにくい。これに対して、普通の名詞性の連体節の場合は、必ずしもそれが係り先になるとはいえない。以上から、連体節述語の扱いは、以下の通りとする。

- 形式名詞に係る連体節述語は、連用節の「B+読点」に準じる。
- 通常の連体節述語は、Bに準じる。

#### 2.3 述語句の分類

前節の分類をまとめると、表1ようになる。

### 3 述語句の細分類

本章では、2章で述べた語尾表現に着目した述語句の分類に対して、問題点を整理するとともに、個別の動詞の性質を考慮することにより分類の詳細化を図る。

#### 3.1 述語句の分類の問題点

2章の分類は、2.2節(2)でも指摘したように、南の分類と比較すると、Bの範囲が広く取られている。このため、「～と比べ」のように格関係に準じて扱う必要があるものや南がA類に分類している一部の「～て」など、Aに分類した方が良いものがBに分類されている。

新聞記事を対象とした実験では、AやCの出現比率が少ないため、Bの範囲を広くとったことにより余分な多義を生じるものの、結果的には解析誤りとなったケースは少ない。一般的には、解析処理の品質の向上を図る上で、余分な多義を削減する方が有利である。

#### 3.2 細分類の基本方針

2章の分類では、Bの下位分類として中止性の強いものを分離することにより、好結果を得た。そこで、逆に、依存性が高いものを従属型として分離することを考える。南がA類に分類した「～て」など、助詞相当表現の分離が主なターゲットになる。

しかし、2章の分類は係り受け解析のための、すなわち、形態素解析で得られる述語句の語尾表現に着目した分類である。語尾表現を見るだけでは分類を詳細化することは困難である。

助詞相当表現に使われる動詞は比較的限定され

表1 述語句の分類

節の種類		分 類		動作性	
従 属 節	連用節	A（読点なし）			
		A+読点			
		B（読点なし）	通常		他動詞性
			強中止		自動詞性
		B+読点	通常		形容詞性
			強中止		名詞性
	引用節	C（読点なし）			
		C+読点			
		引用相当節（→B+読点）			
		通常の引用節（→C+読点）			
連体節	一般名詞型（→B）				
	形式名詞型（→B+読点）				
主節					

ていることに着目して、動詞の字面そのものと、述語句の前後、特に直前の格要素を見ることにより、問題点の解決を図る。

### 3.3 細分類の検討

#### (1)「～て」単独形

副詞的な性質を有する。例えば、「8月相場入りで大証修正は反発して始まった。」（日本経済新聞社の市況速報記事）のような格助詞を伴わない「動詞+て」の表現がある。このように、「～て」単独形の用法しかない動詞をAの従属型とした。市況速報記事では、このほか「先駆して」が該当する。

#### (2)「格助詞+～て」形

助詞相当の表現である。例えば、「特に事故現場など各地の移動局から直接、衛星を通じてニュース映像を送るSNGの用途が伸びると見ている」（日経産業新聞）や「任天堂が後場に入って年初来高値を更新し、兼松日産農、小野薬も買われた」（市況速報記事）である。このうち、前者は直後の述語に係るのでAの通常型に分類した。これに対し、後者はいくぶん前置きの式であると考えて、Bの従属型に分類した。このほかには、後者として、「に加えて」、「に対して」、「に関して」などが属する。

#### (3)「格助詞+動詞連用形単独」形

前節の「格助詞+～て」型に似た助詞相当表現である。しかし、[森田 89]で指摘されているように、本動詞の用法に近いものも少なくなく、係り受け解析の段階でこれらを固定的に扱うのは危険である。そこで、市況速報を含む新聞記事に限定するという条件で、「～を含め」、「～に続き」、「～を控え」などをBの従属型に分類した。例えば、「加盟国の経済改革を後押しする狙いで、特に融資の拡大を求めているロシアなど旧ソ連諸国に対し、新たな支援策になる。」では読点を伴っているが、従属型であるので、前側の述語句（狙いで）は係らない。

### 3.4 述語句の細分類

前節の分析結果をまとめると、表2のようになる。AタイプとBタイプ、および、それらの読点ありと読点なしの4種類それぞれに対して、従属型を設けた体系となっている。

表2 述語句の細分類（\*新設）

節の種類		分類		動作性
従属節	連用節	A（読点なし）	*従属型 通常	
		A+読点	*従属型 通常	
		B（読点なし）	*従属型 通常 強中止	他動詞性 自動詞性
		B+読点	*従属型 通常 強中止	形容詞性 名詞性
	引用節	引用相当節（→B+読点） 通常の引用節（→C+読点）		
	連体節	一般名詞型（→B） 形式名詞型（→B+読点）		
主節				

## 4 おわりに

本稿では、係り受け解析を行なうための日本語の述語句の基本分類と問題点について述べ、それを解決するための細分類試案について報告した。具体的には、AタイプとBタイプの述語の下位分類として従属型を設けた。本稿で報告した試案に基づく係り受け解析は実験中であり、その効果については次の機会に報告する。

### 参考文献

- [南 64] 南不二男：複文，講座現代語6（時枝・遠藤監修），明治書院（1964）
- [南 74] 南不二男：現代日本語の構造，大修館書店（1974）
- [三浦 75] 三浦つとむ：日本語の文法，勁草書房（1975）
- [森田 89] 森田，松木：日本語表現文型，アルク（1989）
- [白井 95] 白井，池原，横尾，木村：階層的認識構造に着目した日本語従属節間の係り受け解析の方法とその精度，情報処理学会論文誌，Vol.36，No.10，pp.2353-2361（1995）