

形容詞を含む連体修飾の係り受け特性

2L-1

菊池 浩三
静岡大学

伊東 幸宏

1. はじめに

機械翻訳の日本語解析において長文中の係り処理の精度向上は大きな課題である。重文・複文に対する用言の係り先の解析は、いくつかの研究がなされている（例えば文献(1)）。しかし、用言の係り先を検討するためには個々の用言が支配する節を正確に切り出せることが必要である。一方、連体修飾中に用言が含まれる場合に、それに先行する格助詞がどこに係るかが決定できないケースが少なくない。

そこで、本稿では、イ形容詞・ナ形容詞を含む連体修飾に関して、先行する「が」格・「の」格の形容詞（イ・ナ形）への係り関係の分析を行う。10万文を超えるコーパスから形容詞（イ・ナ形）を含む連体修飾を抽出し、該当する1597件の係り関係を調査した。その結果、形容詞を4つのTYPEに分類し、前後の名詞に分析を加えれば約96%の精度で係りを特定できることが分かった。

2. 分析方法

抽出したのは次の連体修飾形を含む文である。

名詞1 + < *の | *が > + [副詞]

+ < イ形容詞 | ナ形容詞 > + 名詞2

この形式に該当する連体修飾の例を以下に示す。

例1) 経験の少ない利用者

例2) 考え方の著しい変化

例1は、「の」格が「が」格に格シフトを起こし、「経験が少ない」と解釈される。例2は「考え方の変化」のように名詞1は名詞2に係ると解釈される。このように、品詞としては同じ形容詞であっても係り関係から見ると違いが生じる。このような違いを、形容詞の種類と、名詞1、名詞2の性質から分類で

きるか検証した。分析は、大量のコンピュータのマニュアルから例文を抽出し、その係り関係を人手で調べることにより行った。

3. 分類

手始めに、「が」格、「の」格それぞれが形容詞に係るか否かという基準を考える。この基準に基づくと、以下の4タイプを想定することができる。

分類1: 「が」「の」格は形容詞に係る。

分類2: 「が」格は形容詞に係るが、「の」格は形容詞に係らない。

分類3: 「が」格は形容詞に係らないが、「の」格は形容詞に係る。

分類4: 「が」「の」格は形容詞に係らない。

抽出した例文（同じ形容詞が4件以上検出できたものが対象）を事例の出現頻度に従ってこの4つのタイプにマップすると以下ようになる。

(イ形容詞)

分類1: 高い、低い、大きい、小さい等

分類2: 広い等

分類3: 良い、悪い等

分類4: 詳しい、新しい等

(ナ形容詞)

分類1: 可能な、必要な、困難な等

分類2: 豊富な、有効な等

分類3: - (該当なし)

分類4: 詳細な、色々な、主な、~的な等

その他: 異常な等 (傾向が不明なもの)

分類した各単語の係り方の出現回数を分類ごとに累計した結果を表1と表2に示す。表中の○印は分類の基準に該当する部分である。

Dependency analysis of adnominals including adjective and adjective verb

Kozo Kikuchi Yukihiko Itoh

Shizuoka University

イ形容詞	「が」が形容詞に		「の」が形容詞に	
	係る	係らない	係る	係らない
分類1	○ 98	10	○ 217	47
分類2	○ 3	0	1	○ 6
分類3	2	○ 4	○ 52	5
分類4	0	○ 17	0	○ 99
計	103	31	270	157

全体：561件 ○部分の数：496件

表1 イ形容詞の係りの数

ナ形容詞	「が」が形容詞に		「の」が形容詞に	
	係る	係らない	係る	係らない
分類1	○ 134	1	○ 18	15
分類2	○ 9	1	4	○ 24
分類3	0	0	0	0
分類4	21	○ 95	9	○ 695
その他	3	3	1	3
計	167	100	32	737

全体：1036件 ○部分の数：975件

表2 ナ形容詞の係りの数

この分類のみに依存して解析を行っても、以下のような精度は得られる。

イ形容詞： 496/561 (88.4%)

ナ形容詞： 975/1036 (94.1%)

このように、この4分類は係りの決定の標準ルールを設定するのには有効ではあるが、更に解析率を上げるためにはより詳細なルールの設定が必要である。

4. 詳細な分析

これらの係り関係を、前後の名詞（名詞1、名詞2）で分析すると以下の3つのルールが係り関係を規定する上で効果的であることが分かった。

1) 名詞1が、以下のものである場合は、名詞1は名詞2に係る

序数詞（一つの、二つの・・・）

単位（キロの、個の・・・）

連体詞（この、その、・・・）

まとめる語（等の、他の）

時の表現（日常の、将来の・・・）

位置（外の、中の・・・）等

2) 複合助詞は形容詞には係らない

のが、での、からの、への等

3) 名詞2が形式名詞の場合は、係りの観点から次の2種類に分類する

グループ1：場合、時、こと等

グループ2：方、順、側等

グループ1では、名詞1は形容詞に係り、グループ2では、形容詞は名詞2との係り関係が強くなり名詞1は形容詞に係らない。

これらの3つのルールを付加して、分類した形容詞の係り関係を再分析すると、以下のように、高い精度で係り関係を特定できた。

イ形容詞

全体 561件

解釈可能 529件(94.3%)

ナ形容詞

全体 1036件

解釈可能 1004件(96.9%)

5. まとめ

イ形容詞・ナ形容詞に係り関係の特性から4つのタイプに分け、汎用的と思われる3つのルールを適用すると形容詞を含む連体修飾の係り関係を精度良く規定できることが分かった。しかし、単語によっては、これらの方法のみではまだ係り関係を規定するのに十分でないケースもあり、今後更なる特徴抽出を行い係り解釈の精度を向上させる予定である。

これらの分析は、机上で実施した段階であり機械翻訳システムとして実装したわけではないが、この分類が効果的であると判断している。

参考文献

1) 白井、横尾、池原、木村、小見：日本語従属節の依存構造に着目した係り受け解析、情報処理学会研究報告 94-NL-102-9(1994)