

連体修飾節における曖昧性とその解決策の提案

檜山祥太¹ 吉本啓^{1,2} Alastair Butler^{2,3}

¹ 東北大学大学院国際文化研究科 ² 東北大学高等教育開発推進センター

³ 科学技術振興機構戦略的創造研究推進事業さきがけ

1. はじめに

従来の自然言語処理の研究では、寺村(1975-1978)の理論を基にし、日本語の連体修飾構造を格関係の視点から分析し、連体修飾節の中に被修飾名詞を戻せるかどうかという条件から二分類するという方法が広く採用されている。しかし、格関係から連体修飾関係を正確に捉えるには、問題点が多くあり、それに対する理論的改善策はまだ乏しい。

本稿では、連体修飾関係の分析において、格関係における問題点を指摘し、樗ツリーバンク(Butler and Yoshimoto 2012、吉本他 2013)で実施された、その問題点に対する改善策となるいくつかの基準を提案し、言語学的な知見を日本語分析のアノテーションに活かす方法を示すことを目的とする。

2. 「内の関係」と「外の関係」

現在、連体修飾節を分析する際に最も利用されている分類方法は、寺村(1975-1978)の「内の関係」と「外の関係」に基づく分類である。「内の関係」とは、被修飾名詞が修飾部の用言の補語と考えることができ、一定の格関係が想定できる場合を指し、「外の関係」とは、「内の関係」とは異なり、修飾部の用言の補語と捉えることができず、一定の格関係が想定できない場合、つまり修飾部の中から取り出されたとは考えられない場合のことを指す。例えば、以下の二つの例は前者が「内の関係」、後者が「外の

関係」として考えられている:

- (1) 誰かが買ったチョコレート
- (2) 彼が階段から降りてくる音

(1)では「誰かがチョコレートを買った」というように「チョコレート」に目的格をつけて修飾部に戻すことが可能なため、「内の関係」と考えることができる。一方、(2)では、被修飾名詞「音」にどの格助詞を想定しても修飾部の中に戻すことができない。このような場合は「外の関係」に分類される。

また(3)に示すように、修飾部と被修飾部の間に「という」を入れて表現が成り立っている場合は「外の関係」と見なすことができる:

- (3) 彼が僕を殴った話
彼が僕を殴ったという話

連体修飾節において、このような二分類は格の関わり方の違いという点からも大変重要な分類と言える。しかし、この分類方法を取り入れた分析では、様々な問題が生じることも確認されている。丸元・乾(2000)によると、「内の関係」と「外の関係」による分類において以下の二つの問題点があるとされている:

- ① 言い換え可能な二つの格が考えられる

場合がある

- ②「内の関係」か「外の関係」かが曖昧な場合がある

また、加藤(2003)は次のような問題点を指摘している。

- ③復元されるものは伝統的な文法でいう格助詞に限定されるべきなのかどうか
が不明
④単純に装定と述定の相互変換ができない場合がある

①については、以下のような例が考えられる:

- (3)a. 鯉節を作る工場
b. 鯉節を作る作業場

(加藤 2003: 266)

(3a)の場合、「工場が鯉節を作る」とも「工場で鯉節を作る」とも解釈できる。このように主格と場所格の判断が難しい場合がある。また、(3b)については「作業場が鯉節を作る」とすると、意味的に違和感が生じるが、「作業場で鯉節を作る」であれば問題ない。この点から、①は③の問題点へと繋がる。つまり、場所格で表示される意味関係を単純に連体修飾節に導入するだけでは、(3a)に見られる曖昧性は改善できない。それと同時に、どのような格関係を、どのような条件を設定したうえで含めばよいのかを再考しなければならない。

②の問題点については、様々な現象が確認されている:

- (4) 肖像権を認めた判決

(丸元・乾 2000: 18)

- (5) 大した人物

(加藤 2003: 211)

丸元・乾(2000)が述べるように、(4)については「肖像権を認めるという判決」と「(その)判決で肖像権を認める」という二つの分析の可能性があり、前者の場合は「外の関係」、後者の場合は「内の関係」と見なされるため、どちらが適切なのかという判断が難しい。また(5)の例を考えると②は④の問題点へと繋がる。加藤(2003: 210-211)が述べているように、「その人物が大した」というと非文になる。このように、日本語には単純に装定を述定へと変換ができない場合もある。この場合、「内の関係」とすべきか「外の関係」とすべきかは非常に判断が難しい。

3. 櫛ツリーバンクにおける「内の関係」と「外の関係」の分類基準

ここでは、「内の関係」と「外の関係」に関する分類をより明確にするために、「内の関係」の格付与に関する規則をいくつか提案する。

- (i) 主格、目的格の他に、場所格、時間格、方向格、道具格、度量格も「内の関係」における基準要素として認定する

吉本他(2013)の櫛ツリーバンクでは、NPに追加するタグ情報として NP-SBJ(主語)、NP-OB1(目的語)、NP-LOC(場所格)、NP-TMP(時間格)、NP-DIR(方向格)、

NP-ADT(付加語)、NP-MSR(度量)が採用されており¹、これらを「内の関係」の被修飾語の機能として採用する。これにより、主語や目的語のような格助詞以外の関係性も的確にとらえることが可能となる。

(ii) 固有名詞に付随する連体修飾節は全て「内の関係」になる

(6)a. 戦前から多くの日本人が住む上海

b. (NP (IP-REL (NP-LOC *T*))
 (PP (NP (N 戦前))
 (P から))
 (PP (NP (PP (NP (Q 多く))
 (P の))
 (N 日本人))
 (P が))
 (NP-SBJ *が*))
 (VB 住む))
 (NPR 上海))

(i)で示したように主格、目的格以外の複数の格関係も連体修飾節の分類基準に採用することにより、固有名詞は全てNPに追加されているタグ情報が示す格を与えることができるようになった。樺ツリーバンクでは、固有名詞はNPRというタグが割り当てられるため、NPRが被修飾語であれば、自動的にその連体修飾節は「内の関係」とし、統一的なアノテーションが可能になる。

(iii) 主格と場所格(あるいは道具格のような付加語)などの両方が考えられる場合は、被修飾語を場所格(あるいは道具格のような付加語)と見なし、主語を別に設定する

¹ NPに付加されるタグ情報は他にもあるが、それらについては吉本他(2013)を参照。

(3a)における主格と場所格の違いを考えてみたい。「工場で(誰かが)鯉節を作る」と「工場が鯉節を作る」の間に意味の違いがあるのだろうか。つまり、(3a)は多義表現なのかどうか問題となる。これについては加藤(2003: 206-207)も指摘しているように、(3a)の表現に対して「そこで作ってるの?それとも、そこが作ってるの?」と聞き返すことは考えにくい。多義表現というわけではないと考えられる。したがって、この場合、主格と場所格のどちらを認定させるかは恣意的なものだといえる。そこで(3b)の例も考慮に入れて、(iii)の条件を設定することにより、被修飾語の意味構造の違いに頼ることなく統一的な分析を可能にする:

(7)a. 鯉節を作る工場

b. (NP (IP-REL (NP-LOC *T*))
 (NP-SBJ *pro*))
 (PP (NP (N 鯉節))
 (P を))
 (NP-OB1 *を*))
 (VB 作る))
 (N 工場))

(8)a. 鯉節を作る作業場

b. (NP (IP-REL (NP-LOC *T*))
 (NP-SBJ *pro*))
 (PP (NP (N 鯉節))
 (P を))
 (NP-OB1 *を*))
 (VB 作る))
 (N 作業場))

この基準の長所は(iii)でも述べられているように、この基準は場所格以外にも適用できるということである。例えば(9)や(10)の付加語の場合も同様の分析ができるという点で場所格と付加語の統一的な分析が可能となる:

(9)a. 彼を殺した銃

b. (NP (IP-REL (NP-ADT *T*)
(NP-SBJ *pro*)
(PP (NP (N 彼))
(P を))
(NP-OB1 *を*))
(VB 殺し)
(AXD た))
(N 銃))

(10)a. 昨日会った田中さん

b. (NP (IP-REL (NP-ADT *T*)
(NP-SBJ *pro*)
(NP-TMP (N 昨日))
(VB 会っ)
(AXD た))
(NPR 田中さん))

4. 結論

どのような場合に「内の関係」になるか、また、複数の格が考えられる場合はどのような基準から判断すべきかについて、本稿では樗ツリーバンクで適用している3つの基準を提案した。しかし連体修飾関係にはまだ解決できていない問題点がある。

一つ目は(5)のように、装定の場合のみしか使えない用言だった場合、「内の関係」にするべきか「外の関係」にするべきかという問題である。

二つ目は(11)のように、複数の格が考えられ、かつ多義表現だった場合である。

(11)水を撒いた場所

(11)については加藤(2003: 211)が述べているように「その場所で水を撒いた」なのか「その場所に水を撒いた」なのか「その場所から水を撒いた」なのかが判断できない。このような場合、語用論的知識が判断に必要となるのだが、この語用論的知識をどのようにアノテーションのシステムに取り組

んでいくかがまだ決まっていない。

以上の点については今後の研究課題とし、今後、解決策を見つけツリーバンクの構築に適用させていく予定である。

参考文献

Butler, A. and K. Yoshimoto. 2012. Banking Meaning Representations from Treebanks. *Linguistic Issues in Language Technology* 7. CSLI Publications.

加藤重広. 2003. 『日本語修飾構造の語用論的研究』. ひつじ書房.

丸元聡子・乾裕子. 2000. 「連体修飾を受ける体言の格構造の復元 —コーパスに基づく「内の関係」の分析—」『言語処理学会第6回年次大会 発表論文集』: pp. 16-19.

寺村秀夫. 1975-1978. 「連体修飾のシンタクスと意味 —その1~その4—」『日本語・日本文化』4号~7号.

吉本啓他. 2013. 「日本語ツリーバンクのアノテーション方針」『言語処理学会第19回年次大会 発表論文集』: pp. 924-927.