

同格的な節を補部とする名詞について

金城 由美子* 柏岡 秀紀 田中 英輝

ATR 音声言語コミュニケーション研究所

1 はじめに

日本語の連体修飾構造は、[[太郎が_{t_i}買った]本_i]のように被修飾名詞が修飾節中の要素である関係節によるものと、[[全員が救助された]事実]のように被修飾名詞が修飾節の要素ではない同格節によるものに大別される。

翻訳などのタスクにおいて、統語的・意味的に依存関係の異なる二種類の連体修飾構造を区別することは重要であるが、両者を統語構造のみで区別することは難しい。このような連体修飾構造の理解には、同格節を補部とする名詞の体系的記述が欠かせないが、動詞などの述語に関する記述に比べ、統語的・意味的記述はすすんでいない。

同格的な節を補部とする名詞の記述と分類の試みは既に行われているが、比較的少数の語に対して詳細な統語的・意味的記述を行うというものであった[1, 2]。本発表では、大規模なコーパスから連体形式「という」の存在を手がかりに同格的な節を補部とする名詞の収集とその分類を行うことを目標とする。

2 関係節と同格節

日本語の連体修飾構造は、(1)のように被修飾名詞が修飾節中の要素であると考えられる関係節によるもの(いわゆる「内の関係」)と、(2)のように修飾節が被修飾名詞を含まない同格節によるもの(「外の関係」)の二種類に大別される[3]。

- (1) [太郎が_{t_i}買った]本_i;
- (2) [全員が救助された]事実

関係節による修飾はその被修飾名詞に特に制限がないのに対し、同格節による修飾は限られた意味内容をもつ被修飾名詞にのみ起こり、同格節は被修飾名詞の補部として被修飾名詞の内容を表す役割を果たしている[4]。

「という」「との」「とする」「といった」などの連体形式が同格節と共起する場合があります、同格節は、その共

起関係により、連体形式が必須のもの(3a)と任意のもの(3b)の2種類に分けられる。

- (3) a. [全員救助された]という報告
- b. [外国人に日本語を教える](という)仕事

「外の関係」には「京都に引っ越す前」のような相対名詞修飾節が含まれるが、相対名詞修飾節は連体形式の介在を許さないので、本稿では考察の対象としない。

同格節を補部とする名詞は、発言・思考に関わる名詞や、「事実」のように内容の説明を要するものに限られるとされている。

関係節が関係代名詞といった機能語を伴わない日本語では、修飾節が連体形式を伴う場合は同格節であることが明らかであるが、連体形式を伴わない場合は同格節であることを明示する形式的な特徴はない¹。

また、同格節を補部とする名詞は、関係節の被修飾語となることも可能である。

- (4) [外国人に教える]仕事

日本語の場合、修飾節が空代名詞を含む場合もある。(4)の例では、修飾節の述語「教える」の対象が明示されていないが、被修飾名詞である「仕事」をその対象と考える関係節の読み[[外国人に_{t_i}教える]仕事_i]の外に、修飾節が空代名詞を含むという同格節の読み[[外国人に(何かを)教える]仕事]も可能である。つまり、述語の格関係と修飾節に現れる名詞の数という統語的情報のみでは、修飾節と被修飾名詞の関係を決定することはできない。関係節か同格節かを決定するためには、統語的な情報に加えて、どのような名詞が同格節を補部とするかという語彙的な情報に着目する必要がある。

日本語の連体修飾構造には、語用論的な関係に基づくものもある。

- (5) トイレに行けないコマーシャル

*E-mail: yumiko.kinjo@atr.co.jp

¹統語的要因により、関係節が連体形式を伴う場合もある[4]。

(5)は、「コマーシャルがとても面白くてできていて、トイレに行くことができなくなってしまう」状況についての表現だとされている[5]。この例では、修飾節と被修飾名詞が臨時的な修飾関係にあると考えられるが、そのような解釈を行うためには、関係節ではないという統語的な情報と、同格節ではないという語彙的な情報が必要である。

つまり、連体修飾構造の決定という問題に対処するためには、同格節を補部とする名詞に関する網羅的な語彙記述が必要ということになる。そこで、体系的記述への手始めとして、大規模なコーパスからの名詞の収集を試みることにした。

3 同格節を補部とする名詞の収集

先行研究[1, 2]では、連体修飾構造のうち、次の条件を満たす名詞について詳細な記述がなされている。

- (6) i. 「外の関係」(当該見出し語とそれに先行する述語との意味的な関係がガランなどの格助詞によって復元不可能)
- ii. 「という」の介在が可能

ここではより多くの語を得るために、毎日新聞[6]の5年分の記事データ(7,245,799文、延べ語数130,743,305語、異なり語数367,733語)から、同格節をとる名詞の収集を試みた。

3.1 候補となる名詞の収集

修飾節が連体形式を伴わない場合、関係節と同格節を形式的に判定することは難しい。そこで、連体形式「という」を伴う同格節を選び、被修飾名詞を収集することにした。

連体形式「という」が出現する文について、KNP[7, 8]で構文解析を行い「節+という+名詞」という連体修飾構造を持つ140,068文を選んだ。その被修飾名詞11,011語のうち、日本語語彙体系[9]の一般名詞属性体系(2,710属性、最大12段の深さ)により、意味属性番号を付与できたもの8,834語を候補語とした。候補語の意味属性の分布を図1に示す²。下位の意味属性として、言動、思考を含む「事」の比率が高く、同格節を補部とする名詞は発言・思考に関わる名詞や、「事実」のように内容の説明を要するものである、という既存の説明と一致する分布を示している。

²各項共通である3段目までの意味属性を示した。複数の語義を持つものを含む。

意味属性		(語義数)
具体	主体	2,784
	場所	1,750
	具体物	1,224
抽象	抽象物	1,988
	事	3,172
	抽象関係	2,270

図1: 候補語の意味属性の分布

3.2 名詞の選択

「節+という+名詞」という構造を持つものでも、同格節を補部とする名詞として不適切なものや、伝聞など同格節ではない節に修飾を受ける名詞も候補から除外した。具体的には下記のようなものがあった。

主体を表す名詞

人や組織など、主体を表す具体名詞が被修飾語となる場合、「という」は本動詞として使われるか、伝聞を表し、修飾節は同格節とはいえない。

- (7) 「反撃はすべき」という市民だが「かえってアラブ全体を敵にしよう」と複雑な反応だった。
- (8) また、契約書らしい紙に名前を書かされたという女性もいる。

また主体を表す名詞は、同格節を補部とする名詞の修飾語として現れる場合も多い。

- (9) 「NATO空爆に参加するが、地上軍派遣には反対」という米国の主張はどうしても説得力に欠ける。

意味属性として、主体または主体と場所のみを持つ名詞を候補語から除外した。

形式名詞

連体形式「という」と共起する「こと、もの、方」などの形式名詞は、慣用的な用法の一部となったり、照応的に使用されるなど、用法が非常に多岐に渡るため、候補語から外した。

- (10) 「お休みいただくしかない」ということになった。
- (11) この大会は、シリーズの成績で世界ランキングを決め、男女の上位8選手が世界一を争うというもの。

- (12) そんな短い間に産業構造を転換をしるという方が無理

副詞的名詞など

「時、具合、ところ」など格助詞を伴わずに副詞節を作る名詞や、「従来」のように述語の項として独立して名詞の機能を果たさないもの。

- (13) これ以上乗ったら沈むという時、ボートにすがりついてきた人を、あなたならどうしますか。
- (14) 吽形は運慶、阿形は快慶の作」という従来の定説がこれでぐらついたことになる。

その他、候補語の中から、解析誤りなどで「節+という+名詞」と判定された語などを除外し、6,000語あまりの語を同格節を補部とする名詞として収集した。

3.3 収集方法について

「という」の存在を手がかりにして同格節を補部とする名詞の収集を行ったが、この方法に問題がないかどうかについて検討しておく。

「節+という+名詞」という構造を持つ文を対象にしたため、「という」の介在が任意の同格節を補部とする名詞については収集できなかった可能性がある。また同格節に「という」の介在が任意である名詞は、「という」を伴わずに現れることの方が多という傾向がある³。しかし、収集された名詞には、「という」の介在を任意とする名詞が多数含まれており、大きな問題はないと思われる。

参考のため、IPAL [2] の名詞辞書に「外の関係」を持つものとして記述されている語が、今回収集した中に含まれているかどうかを調査した。

コーパスから収集した約 6,000 語の中には、IPAL の名詞辞書⁴ に「外の関係」を持つとして記述されている名詞 116 語のうち 110 語が含まれており、含まれていなかったのは「窮状、騒動、手品、値打ち、民話、ルポ」の 6 語であった。名詞を収集したコーパスから、これらの名詞を被修飾語とする「という」を伴わない同格節が出現する例を探したところ、窮状 6、騒動 17、手品 1、値打ち 8、民話 1、ルポ 11 の計 46 例があった。以上の結果からも、「という」を伴う同格節から名詞を収集したことに大きな問題はないといえる。

³例えば、「事実」という語が同格節に現れる回数は、「という」を伴う含む場合が 108 回であるのに対し、「という」を伴わない場合が 452 回 (95 年度分のデータ)。

⁴1,082 語が基本語として収録されている。

4 名詞の意味属性と分類

同格節を補部とする名詞の意味的な性質と統語的な性質の関係について、連体形式の必須性と、同格節による修飾の可能性を取り上げて考察する。

4.1 意味属性と統語的性質

同格節を補部とする名詞は、連体形式を必須とするものと、そうでないものに分けられるが、統語的に異なる性質を持つ 2 つのクラスの名詞の間に意味の違いがあるかどうか、収集した名詞の意味属性の比較を行った。

収集した語から高頻度語 100 語を選び、同コーパス中に (15) のような連体形式を伴わない同格節の例が現れるかどうかを調査した。

- (15) 少しでも値切りたい思惑がある。

調査した 100 語のうち、連体形を必須とするものが 40 語、任意とするものが 60 語あった。

(連体形を必須とするもの)

言葉、声、見方、意見、内容、思い、認識、理由、批判、表現、指摘、議論、論旨、趣旨、事情、説、質問、結果、疑問、判断、主張、危機感、報告、期待、数字...

(連体形を任意とするもの)

気持ち、意味、点、話、感じ、考え、問題、気、意識、姿勢、印象、狙い、形、事実、情報、発想、動き、立場、状況、イメージ、時代、現実、不安、条件...

調査した 100 語全体、連体形を必須とするもの、任意のものについて、それぞれ意味属性を語義数の割合で示した。

意味属性	全体	必須	任意
主体	2.9	2.9	2.9
場所	1.8	2.9	1.0
具体物	4.7	2.9	5.7
抽象物	32.9	36.8	30.5
事	32.9	38.2	30.5
抽象関係	24.7	16.2	29.5

図 2: 分類した名詞の意味属性の分布 (%)

図 2 が示す通り、全体的に似た意味属性の分布を見せるが、「抽象関係」という意味属性は、連体形式を必須とするものに少ないという特徴があった。100 語のうち、「抽象関係」の意味属性を持つものは 40 語あり、そのうちの 37 語は連体形式が任意の語であった。「抽象関係」

の下位の意味属性についても調べたが、顕著な特徴を示すものはなかった。

先行研究 [1] では、下記の 2 つの組合せにより名詞の組合せにより下位分類を行っている⁵。

- (16) i. 主語が入れるか否かについての制約による分類
- ii. 修飾節と主名詞の間の意味関係による分類

今回は (16) のように修飾節に主語が入れるかどうかという観点からの分類を行っておらず、この点も含めて、意味的な観点から統語的性質が決定できるかどうかについては、今後さらに検討していく必要がある。

4.2 意味の階層性と統語的性質

収集した名詞には「話、笑い話、思い出話、作り話」「評、風評、悪評、世評」など形態素を共有する語のグループが見受けられた。これらの語が意味的なつながりを持ち、階層的な関係にあることは明らかであるが、同格節をとるかどう点で異った振舞いを見せるものがある。

「話」という語は、出来事などを表す同格節を補部としてとるが、「結婚話」や「恋愛話」など下位に属する語でも出来事を特定したものは、同格節の補部がとりにくい。

- (17) ?* 親友が結婚するという結婚話

出来事を特定するのではない要素を含む下位語は、同格節をとることが可能である。

- (18) 熊五郎がまんじゅうを騙しとるといふ笑い話

同格節をとることができない原因を、冗長性を避けるという意味的なものとして考えることもできる。しかし、そういった制限がない語もあり、同格節の補部の意味的な役割を考える上で興味深い例だと思われる。同格節の補部の意味的な役割は、出来事であったり、発言・思考の内容であったりと個々の名詞によって異り、それほど明確でないことを考えると、このような例に注目していくことが重要だと思われる。

また、同格節をとるかどう点という統語的性質を視野に入れることによって、よりきめの細かい意味階層が得ることが可能になるとと思われる。

⁵166 語を 17 種類の名詞クラスに分類している。

5 おわりに

同格節を補部とする名詞について、大規模なコーパスからの収集を試みた。多数の語彙を収集するために、構文解析システムを利用して「節 + という + 名詞」という構造から候補語を収集し、シソーラスの意味属性を利用して選択を行った。また、収集した名詞の一部を統語的性質によって分類し、意味属性の分布を比較した。収集した語彙について、統語的・意味的により詳細な下位分類が可能かどうかは、今後の検討課題としたい。

謝辞

本研究は通信・放送機構の研究委託「大規模コーパスベース音声対話翻訳技術の研究開発」により実施したものである。

参考文献

- [1] 大島資生. 名詞の統語的・意味的分類の試み—いわゆる「同格連体名詞」について. 計量国語学, Vol. 18, No. 1, pp. 9-25, 1991.
- [2] 本多啓. 連体修飾語としての用法 2. 計算機用日本語基本名詞辞書 IPAL (Basic Nouns)—解説編—, pp. 149-162. 情報処理信仰事業教会技術センター, 1996.
- [3] 寺村秀夫. 連体修飾のシンタクスと意味—その 1—. 日本語・日本文化, Vol. 4, , 1975. 【寺村秀夫論文集 I】(くろしお出版) 所収.
- [4] 金城由美子. 連体修飾におけるトイウの機能について. 日本語学会第 123 回大会予稿集, pp. 134-139. 日本語学会, 2001.
- [5] 松本善子. 日本語名詞句構造の語用論的考察. 日本語学, Vol. 12, pp. 101-114, 11 1993.
- [6] 毎日新聞社 (編). CD-毎日新聞. 1991~1995 年度版.
- [7] 黒橋禎夫, 長尾真. 日本語構文素解析システム KNP version 2.06b6. Technical report, 京都大学, 1998.
- [8] 黒橋禎夫, 長尾真. 日本語形態素解析システム JUMAN version 3.61. Technical report, 京都大学, 1998.
- [9] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦 (編). 日本語語彙体系 1—意味体系—. 岩波書店, 1997.