

# シソーラス辞書を利用した名詞句連体修飾の意味解析

3M-8

平岡丈介 板井由花  
(株)明電舎 基盤技術研究所

## 1. はじめに

日本語の解析において連体修飾節「名詞+の+名詞」の意味解析は困難な問題の一つである。それは意味関係を示す標識が陽には現われていないこと。また、前後の文脈状況によってはさまざまな意味を担うことによっている。現在、この解析には個々の名詞に様々な意味情報を持たせた意味素性を振り、これをもとに解析する方法が考案されている<sup>[1]</sup>。しかし、意味素性を振る作業は、多大な労力が必要であり即座に実行できるという方法ではない。一方、現在では大量の語彙を収録した電子化シソーラス辞書が販売されるようになってきた。そこで、このシソーラス辞書を用いて、連体修飾節の意味解析を行う方法について検討した。

## 2. 連体修飾の意味分類

「名詞Aの名詞B」が担う意味を大きく次の四つに分類した<sup>[2]</sup>。分類の方針は、「(1) 意味関係を《対象》と《関係》で表現する。(2) 名詞の意味内容よりも《対象》と《対象》の関係の仕方を重視する。」である。

- ア. 関係表現：名詞Aまたは名詞Bが関係性的名詞
- イ. 格関係：名詞Aまたは名詞Bがサ変名詞
- ウ. 属性関係：対象、属性、値、に関する表現
- エ. 述語関係：二つの名詞が述語的關係で結合可能

ア：一方の名詞が《対象》を表し、他方が《関係》を表している場合。(図-1)

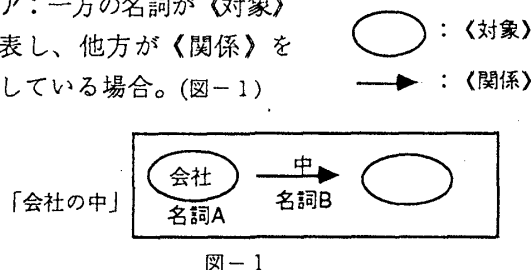


図-1

イ：《関係》が格関係の場合。(図-2)

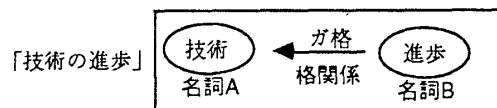


図-2

ウ：名詞が表すものは「対象物」「属性名」「属性値」の三つであり、これら三項の間に生起する関係が《関係》である。(図-3)

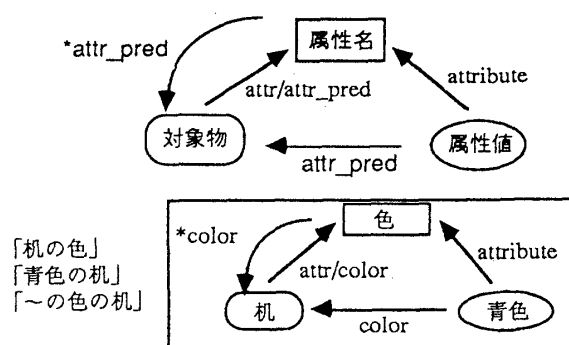


図-3

ウの変形として次の二つがある。

ウ-a 限量的関係：「殆どの」、「多くの」など、対象の量的程度を限定する表現。

ウ-b 形容的關係：形容動詞など、対象の様子を形容する表現。

これらは「属性名」が名詞としてあまり明示されないという特徴がある。

エ：広義の属性関係とも考えられるが、《関係》として動詞的な述語が想定できるもの。名詞Aと名詞Bはこの述語に対しては格関係に立つ。(図-4)

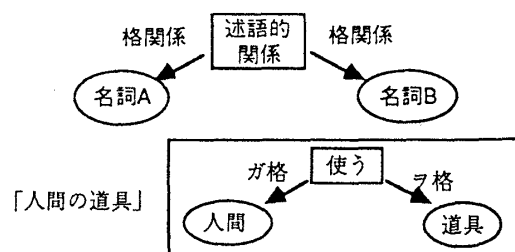


図-4

「人間の道具」のような場合は「人間が道具ヲ使う」と言い直すことができ、述語「使う」が名詞Aをガ格で、名詞Bをヲ格で結び付けている。「さんまの

煙」のような場合は複合的な意味を持つ述語を考える。またエの中で特徴的な次の三つを分離する。

エ-a 含意関係：述語「持つ」で言い換え可能だが、非常に抽象的な内容を表している場合。

エ-b 同格的関係：述語「である」や「という」で言い換えることができる場合であり、名詞Bが名詞Aを数段高い抽象的表現で言い直している。

エ-c 全体部分関係：名詞Bが名詞Aの部分を構成している場合や、集合と要素の関係にある場合。

|                                                                             |                                                               |
|-----------------------------------------------------------------------------|---------------------------------------------------------------|
| ア 私の妻、学校の前、<br>イ 科学の発展、国語の研究<br>ウ 水の温度、三人の女性<br>ウ-a 多くの家、群衆の一部<br>ウ-b 未知の状況 | エ 山の斜面、個人の意見<br>エ-a 経済の構造、研究の目的<br>エ-b 情報処理の分野<br>エ-c パソコンの電源 |
|-----------------------------------------------------------------------------|---------------------------------------------------------------|

表-1 分類例

### 3. シソーラス辞書

我々が用いたシソーラス辞書は学習研究社のものである。これはほぼ国立国語研究所の分類語彙表に準拠した分類法を用いており、記載語数は約3万語で、一つ一つの語義に対して5桁のシソーラスコードが振られている。語義数は約5万語義である。

### 4. シソーラスと意味分類

次の条件を満たすときシソーラスコードによる意味分類が有効となる。

1. ある意味関係で用いられる語が少数のシソーラスコードに集中している。
2. そのシソーラスコードの他の語彙が同じ意味関係で用いられると判断できる。

そこで、約1000例の連体修飾の出現データを上記の意味分類に従って分類し、それぞれの単語のシソーラスコードを調べた。語義が複数ある語については、出現データの文脈に即して人手によりシソーラスコードを選択した。その結果、認められた傾向と処理方法について述べる。

ア【関係表現】上中下左右・時間・人間の意味する語は有効。

イ【格関係】シソーラスは有効ではない。意味素性 act とサ変動詞の結合価を利用。

ウ【属性関係】属性値に相当する語が多く、属性名に相当する語はそれらの分類の中に分散して記載される場合が多い。従って、属性名となりうる語をマーキングすることにより有効となる。「性質」の項には属性名に相当する語が集まっている。

ウ-a【限量的関係】量的割合を限定する語彙はかなり分類されており有効である。

ウ-b【形容的關係】出現度数が低かったため判断保留。

エ【述語関係】《関係》を表す述語は様々なものが考えられる。適当な文脈状況を設定すれば殆どあらゆる解釈が成り立ちうるであろう。しかし、一般的な場合を考えると、解釈も限られてくる。「太郎の本」の場合、{太郎が所有している本} {太郎が書いた本} {太郎について書いた本} といった解釈が考えられる。これは名詞「本」が持っている意味的な特性と考えられるので、以下の方法で解析する。

名詞「本」に対して述語、poss(ガ(hum),ヲ), write(ガ(hum),ヲ), write\_about(ニツイテ(div),ヲ)を与える。これらはそれぞれの解釈の可能性があることを明示している。hum, div は名詞Aに課せられた意味素性の条件である。「歴史の本」の場合は「歴史」の意味素性がhumとは合わず、divのみが残るので {歴史について書いた本} が選択される。

しかし、何万という名詞に対してこのような述語を与えるのは困難である。そこで、シソーラスコードが同一の名詞に対して、同一の述語を与える方法を考えた。

データ数は40例ほどだが有効性があると判断できる場合が約半数であった。

エ-a【含意関係】有効な場合が多い

エ-b【同格的関係】有効でない場合が多いが、この分類に入る語はかなり限られる。

エ-c【全体部分関係】出現度数が低かったため判断保留。

### 5. まとめと課題

シソーラスコードの直接利用が有効となる範囲を確認した。有効でない場合の多くは、語が含み持つニュアンスが意味分類に影響を与えている場合である。意味素性など、他の情報との組合せによりどのような解析ができるか、今後、試みていきたい。

#### 参考文献

- [1] 島津, 内藤, 野村: 助詞「の」が結ぶ名詞の意味関係の解析, 計量国語学Vol.15 No.7 1986
- [2] 草薙, 南, 中野, 吉田: 朝倉日本語新講座「文法と意味2」1 文法形式が担う意味(1985)