

# 用例による換喩の解析

山本 専 村田 真樹 長尾 眞

京都大学工学部 電気工学第二学科

## 1 はじめに

本研究では、用例を利用して換喩を解析することを試みる。換喩解析についての従来の手法では換喩に関する特殊な知識や意味ネットワークなどによって換喩の解析がなされてきたが[1][2]、本研究では、換喩を解析するために新聞コーパスから「名詞 A の名詞 B」の形をした用例を収集し、それによって換喩を解析することを試みる。この手法では、用意したコーパスから用例を取り出して換喩の解析に使用するので換喩に関する特別な知識は必要ない。また、比較的新しい換喩表現でもコーパスを新しいものにすることによって解析できるようになると期待できる。

## 2 比喩について

比喩表現は文章中でよく使われる表現技法の一つである。その種類としては、主に 直喩、隠喩、換喩などがある。直喩とは「のような」「みたいな」などの語を伴って明示的にある対象をそれに似ている別のものに喩える比喩表現である。隠喩は、直喩と同じようにあるものをそれに類似した別のもので表現するものであるが、直喩とは異なり、明示的に比喩を示す語はない。換喩は、あるものをそれと関係のある他のもので表わす比喩である。換喩が持つ関係には 作者-作品、主体-手段、容器-中身、主体-付属物などがある。表1にその関係と換喩の例を示す。

以下では本研究の対象である換喩とはどのようなものかについて例を挙げて考えてみる。

漱石を読む (1)

これは作者-作品の関係による換喩である。この「漱石を読む」という文を字義通り解釈しようとする「漱石」という人物自体を読むことになり、意味がとれない。この文では作家名である「漱石」という語によってその作品にあたる「(漱石が書いた) 小説」という語を表わしており、「漱石の小説を読む」という意味で解釈される。

本研究では換喩のみを対象とし、隠喩や直喩は対象としていない。それは、以下のような理由からであ

表 1: 換喩における関係とその例

換喩の関係	換喩の例	解釈
作者と作品	漱石を読む	漱石の小説
メーカーと製品	フェラーリに乗る	フェラーリの車
原料と製品	アルコールを飲む	酒
容器と中身	鍋を食べる	鍋料理
主体と付属物	詰め襟が歩いて来る	詰め襟の学生
主体と手段	白バイに捕まる	白バイ警官
主体と近接物	ハムサンドが勘定を払う	ハムサンドの客
地名と産地	大島を仕立てる	大島紬

る。直喩は、比喩の意味を持つ語を文中に含むためそのままの解釈で意味がとおり他の文と同様に扱っても差し支えない。隠喩などの類似性による比喩は文脈の影響を受けやすく、文章の流れによって解釈が変化することが多く、解釈を決定することが難しい。それに対し、関連性の比喩である換喩は語の関連性による寄与が大きく、狭い範囲の解析で解釈の決定が可能となると考えられる。

## 3 用例による換喩解析

### 3.1 換喩と用例「名詞 A の名詞 B」

換喩は、いわば言い替えの比喩である。前述の「漱石を読む」は、「小説」を「漱石」で言い替えている。しかし、それは単なる言い替えではない。「小説」は小説全般を指すのではなく「漱石の小説」を意味している。「漱石を読む」は「作家-作品」の関係が暗黙のうちに認識され「漱石の小説を読む」の「小説」が省略されたものであるとも考えられる。

このように考えると換喩の文の多くは文中で省略されている語を補ってやることで解釈できる。換喩において喩える語と喩えられる語の間には、前述のような「作家-作品」等の関係が存在しているので、その関係を踏まえた上で文中の喩える語を手がかりにして喩えられる語を見つけ出せば解釈が可能になると思われる。

本研究ではこの考え方に基いて、換喩の文の単語と

単語との関係から省略された語を補い、これを解釈とする。ここで、省略されている語を補うために用例を利用する。換喩の解析に用いる用例として以下の表現を利用する。

- 「名詞 A の 名詞 B」
- 「名詞 A 名詞 B」

「名詞 A の名詞 B」の名詞 A と名詞 B がもつ関係の中には「漱石の小説」(作者-作品)のように換喩の持つ関係のほとんどが含まれる。このため換喩の解析に「A の B」の用例が利用できると考えられる。従って換喩の解析においては「名詞 A の名詞 B」の形をしたものを用例として集め、この中の「名詞 B」のうち最適なものを選びだし、これを換喩の文に補って解釈とすればよいことになる。

また、換喩の文「鍋を食べる」が「鍋料理を食べる」と解釈されるように、複合名詞で「鍋料理」のように表現することの方が一般的になっているものもある。このようなものも解釈の候補とするため、複合名詞「名詞 A 名詞 B」も用例として集めて解析に用いる。

### 3.2 用例による換喩解析方法

本研究では、換喩を解析するために用例を収集し、それによって換喩を解析する。

換喩の検出は、比喩の検出をした上でその中から換喩の文を特定しなければならず、隠喩などの他の比喩解析と切り離して換喩のみの検出を行なうのは困難である。そこで本研究では換喩の検出は扱わず、換喩の文が入力文として与えられたものとして換喩の解析を行なった。

本研究では以下のような手順で解析を行なう。

1. 換喩の入力文を格フレームと照合し、換喩で表わされている語を特定する。
2. 新聞のコーパスから換喩の語(名詞 A)を含む用例「名詞 A の名詞 B」、「名詞 A 名詞 B」を集め、名詞 B を解釈の候補とする。
3. 候補の中から格フレーム、動詞との関連、頻度を考慮して解釈とする語を一つに決定する。この語を入力文に補って解釈とする。

#### 3.2.1 換喩で表わされている語の特定

換喩で表わされている語の特定には格フレームを使う。換喩の入力文を形態素解析、構文解析し、以下の

形のものを取り出して格フレーム辞書と照合する。

「名詞+格助詞」+「名詞+格助詞」…+「述語動詞」  
(2)

それぞれの格についてその名詞が格フレーム辞書の意味属性に合致するか否かを調べ、合致しない格を換喩であると判定した。

例として以下の換喩の文を解析することを考える。

僕が漱石を読む (3)

まず、換喩の語を特定する。「読む」の格フレームは以下のようなものとする。

{人}が{本、新聞、小説}を [読む] (4)

入力文をこの格フレームに照らし合わせる。ガ格の「僕」は「人」であるのでこの格は格フレームの条件を満たす。ヲ格の「漱石」は「本」、「新聞」、「小説」のどれにも属さず、この語は格フレームの条件を満たさない。このことによりヲ格の「漱石」が換喩の語であると判定される。

#### 3.2.2 解釈候補の抽出

次に換喩の解釈候補の用例を集める。用例としては新聞のコーパスにある文を利用する。

前節で換喩であると判定された語(名詞 A)について用例を収集する。新聞のコーパスから「名詞 A の名詞 B」、「名詞 A 名詞 B」の用例を収集し、換喩解釈の候補とする。

まず、名詞 A を含む文を新聞のコーパスから集め、その文を形態素解析、構文解析し、それを用例とする。その用例の中で「名詞 A の名詞 B」「名詞 A 名詞 B」の形をしたものを解釈の候補とする。この時それぞれの「名詞 A の名詞 B」「名詞 A 名詞 B」がコーパス中に出現した回数を名詞 B の頻度とする。

同時に、動詞との関係による絞り込みのために換喩の文の中で名詞 A と同じ格を受けている動詞についても同様に新聞のコーパスからその動詞を含む文を全て集め、形態素解析、構文解析を行なって用例を作る。この中から「名詞 C を ~ する」の形をしたものを全て集めその名詞 C を次の候補の絞り込み用いる。

例として前述の例文(3)の場合を考える。この解釈候補を用例から集める。解析の対象となる語は「漱石」であるので、新聞のコーパスから「漱石の」を含む文を全て集め、これを形態素解析、構文解析したものを用例とする。この用例の中に「漱石の小説は…」

という文があったとすると、この中から「漱石の」が係っている名詞「小説」を解釈の候補として抽出する。このようにして「漱石の名詞 B」の形の用例から名詞 B を全てとりだし解釈の候補とする。

### 3.2.3 解釈候補の絞り込み

最後にこの名詞の用例から集めた候補を格フレームや動詞とのつながりによって絞り込み、頻度から解釈を決定する。この絞り込みと決定は、名詞の用例から集めた候補に対して行なう。なぜなら、換喩は名詞と名詞との関連性に由来するものであり、それを考えると、名詞から集めた候補を換喩解析の候補とすることが適当であると思われるからである。

まず、解析した結果として得られる解釈を表わす文はリテラルでなくてはならないから、候補となる語を格フレームの格の意味的制限を満たす語に絞り込む。次に、残った候補の中の語で、動詞の用例から集めた名詞 C に同じ語があるものだけを候補として残す。このような語は名詞とのつながり、動詞とのつながりがそれぞれ実際の文章にあるということがいえるから、その単語を解釈として採用する根拠が強くなる。最後に、この候補の中からさらに一つの解釈に決定するために頻度の情報を使う。

しかし、単純に候補の語の頻度だけによる決定では、換喩である語と意味的な結び付きの強い語が候補としてあっても、その語がコーパス全体にあまり出現しない場合は解釈として採用されない。このことによる誤りを減らすために、シソーラスにおける上位語の頻度を利用する。換喩解析の候補の中で、ある語の上位語があった場合、上位語の頻度の何分の 1 かの頻度でその語も出現したと換算する。これによって出現頻度の低い下位語の頻度を補強できると考えられる。このような考え方から、解釈候補の語群にあるその語の上位語全ての頻度を足したものを上位語頻度とした。本研究においては、その語の頻度とこの上位語頻度とを [1.5 : 1] の割合で加算して、これが最も高いものを解釈と決定する。

こうして最も値の高いものを解釈として採用し、この単語を入力文に補い換喩解析の結果得られた解釈とする。

例として例文 (3) の場合を考える。前節の候補の抽出によって集めた候補が { 小説-11、作品-3、新書-1、遺族-1、影響-1 } であったとする。それぞれの単語にふられた数字はその単語の頻度である。解釈候補

表 2: 本研究の実験結果

●換喩ではないと判定された文 6/41

●換喩の文 35/41

正しく換喩を解析できた文	66%(23/35)
解析できなかった文	17%( 6/35)
単語辞書中に単語がない	50%(3/6)
動詞が格フレーム辞書にない	50%(3/6)
誤った解釈をされた文	17%( 6/35)
用例の中に正解がない	67%(4/6)
頻度情報により不正解	17%(1/6)
格フレームの選択失敗	17%(1/6)

表 3: 実験における正解例

換喩の文	解釈
ユーミンを聞く	ユーミンの <u>歌</u> を聞く
フォードに乗る	フォードの <u>車</u> に乗る
今、平安神宮 が満開だ	平安神宮の <u>桜</u> が満開だ
鍋を食べる	鍋 <u>料理</u> を食べる
白バイ が違反者を逮捕する	白バイ <u>警官</u> が逮捕する

の絞り込みははじめに「遺族」、「影響」などの格フレームに合致しない語や、動詞から集めた名詞 C の語群にない語を取り除かれる。その後、上位語頻度を求め、頻度と上位語頻度を 1.5:1 の割合で足した値が最も大きいものを解とする。ここで上位下位の関係があるのは、小説と作品だけであるので作品の下位語である小説の上位語頻度は 3 となり、その他は 0 である。その語の頻度と上位語頻度を加算して最も大きい「小説」が解釈の語となり、解釈は「漱石の小説を読む」となる。

## 4 実験と考察

### 4.1 実験

前述した手法を用いて実際にどの程度換喩の解析が行なえるかを調べるために換喩の文 (41 文) について実験を行なった。これらの文はあらかじめ文法書 [3][4] などから用意した。格フレーム辞書として、NTT の日本語語彙大系 [5] を用いた。入力文の形態素解析には JUMAN [6] を使い、構文解析には KNP [7] を用いた。コーパスは新聞 3 年分を使用した。また、上位下位関係を求める際のシソーラスとして NTT の日本語語彙大系の意味体系、単語体系を用いた。この実験の結果と正解の例をそれぞれ表 2, 表 3 に示す。

## 4.2 考察

実験において41文中換喩ではないと判定された文が6文あった。これらの文は「庭を掃く」などのように通常換喩であることを意識せずに使われているもので、比喩的な意味合いの薄いものであった。それ以外の35文の換喩の中で、23文が正しく解析できた。

精度について考察する。入力41文中正しい解釈が得られたものは23文であった。これを正解率としてみると56%である。しかし、解析を誤った18文のうち6文は格フレームによってリテラルであると解釈され、また実際そのままでも意味の通る文であった。これを換喩でないと考えて入力文から外すと、正解率は66% (23 / 35) となり、まずまずの値であるといえる。しかし、入力文に用いた換喩の文は一般の文章からとってきたのではなく、文法書などから集めたため、この結果だけを見て実際の文章においても同様の精度で解析できるとは言えない。

以下で実験において明らかになった問題点を考察する。

格フレーム辞書にない動詞「ゲンナリする」や名詞辞書にない「吉本ばなな」などの語が解析する文に含まれているため解析できないという問題があった。全ての語を格フレームで記述することは不可能なので、これを解決するためには、何らかの方法で未知語を格フレーム辞書、単語辞書で扱えるようにする必要がある。

用例が少なく、正解が用例の中にないため、解析を誤ったものとして「一升瓶を飲む」があった。この文の解析で用例として得られたのは「ラッパ飲み」、「ラベル」、「栓」がそれぞれ1つずつで、正解である「酒」は用例にはなかった。この問題は用例を取り出すために用いた新聞のコーパスを増やすことである程度改善できると考えられる。

動詞による絞り込みにおいて、実際の文章中に解釈候補と換喩の文の動詞との組合せが必ずしも出現するとは限らないため、正解が棄却されてしまう問題があった。実験では「霞が関が動揺する」の解析において、正解の語「官僚」は候補に挙がったが、新聞記事中に「官僚が動揺する」の用例がなかったため、動詞とのつながりからの絞り込みによって正解である「官僚」が棄却されてしまった。これを防ぐために、出現しなかった候補を棄却するのではなく、その頻度を減らすなどの方法で動詞とのつながりの情報を反映すればよいと思われる。

頻度による絞り込みにおいて、新聞のコーパス中に出現する頻度により誤った解析をした例もあった。

「鷗外を読む」の解析において解釈候補の頻度が「小説」1に対し、「研究」5で「鷗外の研究を読む」と解析された。

## 5 おわりに

本研究では、換喩の文を用例によって解釈する手法を示した。これは、換喩を喩えられる語が省略されている比喩であるとみなし、換喩における語の関係を接続助詞「の」の持つ働きで表わせることに着目して、用例を用いて省略されている語を補うことによって換喩を解析することを目的としたものである。

この手法を用いて換喩の解析の実験を行なったところ、35個の換喩の文において66%の精度で解析できた。しかし、入力文は一般の文章からとってきたのではなく、文法書などから集めた換喩の例文であったため、一般の文章にはこの例文にないパターンの換喩も存在することが考えられ、実際の文章中に現れる換喩について実験をし、どの程度の精度で解析できるかを確かめる必要がある。

## 参考文献

- [1] Eric Iverson and Stephen Helmreich, Metallel: An integrated approach to non-literal phrase interpretation, *Computational Intelligence*, Vol. 8, No. 3, (1992), pp. 477-493.
- [2] D. Fass, Collative semantics: A semantics for natural language, *PhD thesis, New Mexico State University*, (1988).
- [3] 国立国語研究所 中村明, 比喩表現の理論と分類, (秀英出版, 1977).
- [4] 山梨正明, 比喩と理解, (東京大学出版会, 1988).
- [5] 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), 日本語語彙大系, (岩波書店, 1997).
- [6] 黒橋禎夫, 長尾真, 日本語形態素解析システム JUMAN 使用説明書 version 3.4, (京都大学大学院工学研究科 長尾研究室, 1997).
- [7] S. Kurohashi and M. Nagao, A method of case structure analysis for japanese sentences based on examples in case frame dictionary, *IEICE Transactions on Information and Systems*, Vol. E77-D, No. 2, (1994), pp. 227-239.