

# 連体節における内の関係・外の関係の自動判別

高瀬 慎也<sup>†</sup> 松本 忠博<sup>†</sup>

<sup>†</sup> 岐阜大学大学院工学研究科

## 1 はじめに

日本語の連体節は英語では、関係節、同格節、前置詞句など様々な表現に対応するため、正しく訳し分けることは日英機械翻訳の課題となっている。それに深く関わる問題の1つが「内と外の関係」の判別である。日本語の連体節は、内の関係・外の関係のどちらも形式上の区別はない。しかし、英語は内の関係と外の関係では異なる形式を用いるため、2つの関係を区別することが訳し分けにおいて重要である。

本研究では、翻訳精度の向上のため、主に動詞の結合価パターンの有無に着目し「内と外の関係」の自動判別プログラムの作成を行い、日英機械翻訳システム jaw/English[1] において実装した。

## 2 連体節について

連体節とは、動詞や形容詞などの用言が連体形で名詞を修飾する文節であり、次のような構造をとる。

述語の連体形 + 被修飾名詞 (以下「底の名詞」とする)

例：昨日あなたが会った人は私の弟です

{昨日あなたが会った} + 人

連体節は、構文的・意味的關係から2つの異なる關係に分類できる。

(1) 廊下を走っている人

(2) 日本語を教える仕事

(1) では、被修飾名詞「人」と連体節中の用言「走る」との間に「人が廊下を走っている」という格關係が成り立つ。このような關係を「内の關係」という。一方(2)では、被修飾名詞「仕事」は連体節中の述語と特定の格關係を持たず、被修飾名詞の内容や原因・理由などを表す。このような關係を「外の關係」という。

## 3 動詞の結合価パターン

動詞の結合価パターンは、NTT 日本語語彙大系 [3] を参考に作成した jaw/English の日本語表現パターン

テーブルに登録されている。図1に動詞「行く」のパターンの例を示す。

## 4 「内と外の関係」の判別方法

本手法の判断方法の流れを、図2に示す。

先行研究 [2] では、判別方法として底の名詞の意味属性が「具体」下であれば「内の關係」とみなしていた。しかし、底の名詞が「具体」下でも「外の關係」となる場合がある。

例：魚を買ったおつり

そのため、このような文にも対応できるように判別方法を考えた。

「内と外の関係」の判別は以下の手順で行う。

1. 底の名詞の意味属性が「具体」か「抽象」か判断
2. 底の名詞が「具体」下の場合、連体節の形式から場合分け
3. 結合価パターンを持つかどうか判断し、「内の關係」「外の關係」の決定

### 4.1 底の名詞の意味属性

底の名詞の意味属性は、前述の日本語表現パターンテーブルを参照し、日本語語彙大系の意味属性体系 (図3) の階層において「具体」下にあるか「抽象」下にあるかを判断する。

### 4.2 連体節のパターンによる場合分け

底の名詞の意味属性が「具体」下である場合と「抽象」下である場合で分けて考える。

#### 4.2.1 底の名詞の意味属性が「具体」下

底の名詞の意味属性が「具体」下である連体節は、以下の4つのパターンに分けられる。

1. 「底の名詞」が「カ格」にマッチする

N1	人	が		
N2	場/場所	に/へ/で		
			行く	物理的移動
N1	主体/乗り物/動物			
N2	道路/鉄道/創作物(その他)/場所/場	に/へ/まで		
N3	道路/鉄道/場所/場	から/より		
			行く	物理的移動
N1	主体/乗り物/動物	が		
N2	場所/場	を		
			行く	身体動作
N1	人	が		
N2	魚	を		
釣り		に		
			行く	身体動作

図 1: 動詞「行く」の結合価パターンの例

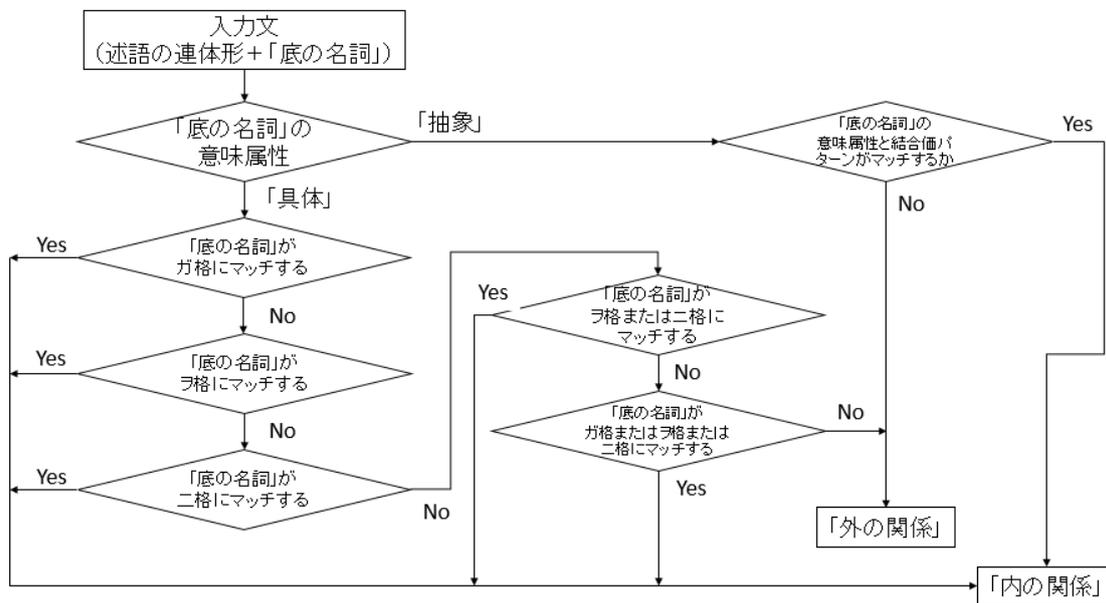


図 2: 判別方法の流れ

2. 「底の名詞」が「ヲ格」にマッチする
  3. 「底の名詞」が「ニ格」にマッチする
  4. 「底の名詞」が「ガ格」または「ヲ」にマッチする
- それぞれの形式について例をあげながら説明する。

1. 「底の名詞」が「ガ格」にマッチする場合  
例：ピアノを弾く男子  
「N1 が N2 を 弾く」の「N2(ヲ格)」には「ピアノ」が入るので、空いている「N1(ガ格)」に底の名詞「男子」がマッチするかを考える。
2. 「底の名詞」が「ヲ格」にマッチする場合  
例：猫が壊したおもちゃ  
「N1 が N2 を 壊す」の「N1(ガ格)」には「猫」

が入るので、空いている「N2(ヲ格)」に底の名詞「おもちゃ」がマッチするかを考える。

3. 「底の名詞」が「ニ格」にマッチする場合  
例：私がクッキーをあげた犬  
「N1 が N2 を N3 にあげる」の「N1(ガ格)」には「私」、「N2(ヲ格)」には「クッキー」が入るので、「N3(ニ格)」に底の名詞「犬」がマッチするかを考える。
4. 「底の名詞」が「ガ格」または「ヲ格」にマッチする場合  
例：日本で作った車、川で泳ぐ子供  
「N1 が N2 を N3 で作る」の「N3(デ格)」には「日本」が入るので、「N1」または「N2」に底の名詞「車」がマッチするかを考える。

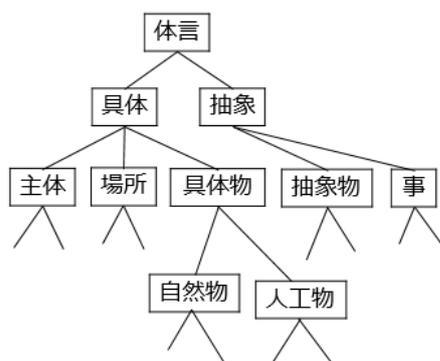


図 3: 意味属性体系

表 1: 例文

番号	例文	種類
(1)	彼が撮った写真	内
(2)	猫を攻撃する癖	外
(3)	公園を走る犬	内
(4)	問題を調査した記憶	外
(5)	日本で作った車	内
(6)	子供を甘やかす傾向	外
(7)	私の猫が壊したおもちゃ	内
(8)	脂肪を溶かす性質	外

以上の4つのパターンに当てはまらない場合は「N1」「N2」「N3」のいずれかにマッチするかを考える。

#### 4.2.2 底の名詞の意味属性が「抽象」下

例文を見たところ、底の名詞の意味属性が「抽象」下であるとき、結合価パターンを持つ場合は「N2」(直接目的語)にマッチする。したがって、「N2」にマッチするかだけを考えればよい。

#### 4.3 「内と外の関係」の決定

底の名詞が結合価パターンにマッチすれば、格関係を持つといえるため「内の関係」となり、マッチしなければ「外の関係」となる。

## 5 判別実験と結果

本研究で提案した「内と外の関係」の判断方法の有効性を検証するために、判別実験を行った。web上の中学生向け英語学習サイトや学習書 [4] から取得した文 135 文(「内の関係」100 文、「外の関係」35 文)に対

表 2: 判別実験結果

	「内の関係」	「外の関係」	合計
例文数	100	35	135
正解数	96	23	119
正解率	96.0%	65.7%	88.1%

表 3: 「内の関係」で誤判別した文

番号	例文	正解	判別結果
(1)	彼女が読む雑誌	内	外
(2)	私たちが読んだ本	内	外
(3)	私たちが図書館で見つけた男性	内	外
(4)	人形を見つけた少年	内	外

し翻訳実験を行った。なお、本実験は判別規則の作成に用いたデータと同一のデータを用いたため、クロウズドテストとなる。

結果を表 2 に示す。

## 6 問題点の考察と改善方法の提案

「内の関係」「外の関係」それぞれの誤判別した文を見て、改善方法を考える。

### 6.1 「内の関係」の誤判別の原因と改善方法の提案

「内の関係」で誤判断した文は表 3 の 4 文である。(1), (2) については、動詞「読む」の結合価パターンが、

- N1(主体) が N2(抽象物(精神)/精神) を読む
- N1(人) が N2(抽象) を N3(詩歌) に読む
- N1(主体) が 票 を読む

の 3 つしかないためである。動詞「読む」の結合価パターンを追加することで、正しく判断できると考えられる。

3, 4 については、動詞「見つける」の結合価パターンが日本語表現パターンテーブルに登録されていないためである。日本語語彙大系に動詞「見つける」の結合価パターンを追加することで解決が可能になると考えられる。

表 4: 「外の関係」で誤判別した文

番号	例文	正解	判別結果
(1)	家に帰る約束	外	内
(2)	彼が来る知らせ	外	内
(3)	彼女の仕事をした褒美	外	内
(4)	あれをする手段	外	内
(5)	工場を売る提案	外	内
(6)	家を失う経験	外	内
(7)	英語を教える方法	外	内
(8)	家に住んでいた跡	外	内
(9)	質問に答える義務	外	内
(10)	新しいことを学ぶ情熱	外	内
(11)	誰かが入った形跡	外	内
(12)	家を売る決心	外	内

表 5: 改善後の判別実験 結果

	「内の関係」	「外の関係」	合計
例文数	100	35	135
正解数	96	33	129
正解率	96.0%	91.4%	94.8%

## 6.2 「外の関係」の誤判別の原因と改善方法の提案

「外の関係」で誤判別した文は表 4 の 12 文である。意味属性には番号がふってあり、意味属性大系の階層が上になるにつれて番号の範囲が大きくなる。

例: 「主体」10003~10387, 「男」10048, 「抽象物」11001~11234

誤判断の原因は、結合価パターンに「N1(「主体」10003~10387)がN2(「抽象物」11001~11234)を売る」のように意味属性の範囲が大きいパターンがあるためである。そのため、意味属性の範囲に制限をかける必要がある。

そこで、底の名詞が「具体」下の場合はその範疇に含まれる意味素の数が 300 以下、「抽象」下の場合には 100 以下という制限を設けた。なお、底の名詞が「主体(10003~10387)」下で、形式が「1」の場合は、N1に「主体」があれば結合価パターンを持つとする。

## 6.3 改善後の判別実験と結果

改善の結果、表 5 のように正解率が向上した。

表 6: 改善後「外の関係」で誤判別した文

番号	例文	正解	判別結果
(1)	彼女の仕事をした褒美	外	内
(2)	家に住んでいた跡	外	内
(3)	誰かが入った形跡	外	内

「内の関係」は結果が変わらなかったので「外の関係」の原因について考える。

「外の関係」で誤判別した文は表 6 の 3 文である。

誤判別の原因は、意味属性の範囲が 100 以下の結合価パターンが存在したためである。意味属性の範囲の制限については検討が必要である。

## 7 おわりに

本研究では、主に結合価パターンの有無に着目し、「内と外の関係」の自動判別プログラムの作成を行った。判別実験を行った結果、クローズドテストではあるが、全体で約 94.8 % の正解率が得られた。今後の課題としては、オープンテストの実施と「内と外の関係」の判別から連体節の翻訳精度の向上があげられる。

## 参考文献

- [1] 池田尚志. 日本語からアジア諸言語への機械翻訳システムの構築奮闘記, 日本語学, 28(12), 62-71, 2009.
- [2] 藤本敬史, 池原悟, 村上仁一, 表克次 (2002). 複文における底の名詞と修飾部の内と外の関係の判断規則, 言語処理学会第 8 回年次大会, pp.679-682.
- [3] NTT コミュニケーション科学研究所: 日本語語彙大系, 岩波書店 (1997)
- [4] 猪野真理枝 佐野洋: 英作文なんかこわくない 連体修飾編 日本語の発想でマスターする英文ライティング, 東京外国語大学出版会 (2014)