

# 機械処理のための方言結合価辞書の作成

## —置賜方言を対象に

横山 晶一 我妻 信博

山形大学 工学部

### 1 はじめに

方言は、日本の各地で、その地方独自の文化、交通、地理などさまざまな条件のもとで独自の発展をとげている。これまで、方言は、主として民俗学や言語地理学的な観点から研究されてきた。

本研究は、方言を機械处理的な観点からとらえようとする試みの一つである。音声的な面からは、方言の地域によるアクセントやイントネーションの違いを明らかにする研究がすでに行われている(たとえば[1])。

本研究では、自然言語処理や機械翻訳的なアプローチの準備段階として、方言に関する結合価と意味素性の辞書を作る試みについて報告する[2]。方言を機械処理の観点から扱うと次のような利点がある[3]。

1. 方言は、語彙や文法の面で、古い日本語の形態を残していると言われている[4]。もしそうであれば、機械処理に適した形での系統的な記述も可能と考えられる。また、その記述から、逆に共通語の記述への示唆が得られる可能性がある。
2. 方言は主として話し言葉として用いられる。したがって、話し言葉の特徴をよく表している。方言の文法を記述することは、話し言葉の研究につながる。
3. 方言と共通語の間の機械翻訳的なシステムを作ることによって、両者の関係、また各々の特徴が明確になる。

本研究では、対象として置賜方言を取り上げる。その理由は次の通りである。

1. 方言の研究にはネイティブの存在が必須である。筆者の一人(我妻)はネイティブであり、しかも3世代同居しているために曖昧なケースでの確認が容易にできる。
2. 資料が比較的揃っており[5, 6]、それらを参照できる。

3. 大学のキャンパスがこの方言の区域内であり、継続的に研究できる可能性がある。

4. すでに山形県内の別の方言である庄内方言に関して調査を行ったノウハウがある[3, 7]。

以下では、置賜方言について、その特徴を述べるとともに、資料[5, 6]に基づいて、現在も使われている動詞を182語抽出し、その各々に対して、意味の違いにより、結合価と、付随する名詞の意味素性を調査した。いわば「IPAL[8]置賜方言版」を作成した。その結果について述べる。

### 2 置賜方言の特徴

山形県は村山、置賜、庄内、最上の4つの地方に大きく分かれ、それぞれの地方に特有の方言がある。そのうち、特に庄内と他の3地方とは、地理的な要件から来る著しい差があり、一般的には庄内が北東北方言、他の3地方が南東北方言に属すると言われている[6]。

置賜方言は、他の方言と同じように、話し言葉の特徴を顕著に持っており、文の中で助詞が省略される傾向があったり、独特の敬語法があったりする。以下に簡単にその特徴を述べる。

#### 2.1 動詞の特徴と活用

置賜方言では、他の方言と同様に、サ変動詞の活用に一段化の傾向が顕著に見られるほか、活用形にも、共通語にない特徴がある。表1に、主な動詞とその活用形を示す。表1で、括弧に入れたものは、稀にしか用いられないものである。

活用形における特徴は、主として、後続の助動詞による。未然形には、ネ(打ち消し)、レル・ラレル(受身、可能)、セル・ラセル<sup>1</sup>(使役)などがつく。「ネ」、「ラセル」を除いては、共通語と同じである。活用形は、カ変、サ変で形が少し変わる。

<sup>1</sup>「見ラセル」といった形で用いられる

表1 動詞の活用形

語例	活用	未然形	連用形	終止形	連体形	仮定形	命令形
いく	五段	イカ	イキ イツ	イク	イク	イケ	イケ
みる	上一	ミ	ミ	ミル	ミル ミツ ミン	ミレ	ミロ
ねる	下一	ネ	ネ	ネル	ネル ネ ネン	ネレ	ネロ
くる	カ変	コ ク	キ	クル	クル クッ クン	クレ	コエ
する	サ変	サ シ	シ	シル (スル)	シル (スル)	シレ	シロ

連用形は、中止法はほとんど用いられない。後続の語としては、ダエ・チェ（希望）、タ・ダ（完了、過去）、テ（接続助詞）などがある。

終止形、連体形は通常同形で、共通語と同じように文の終りや、連体修飾形として用いられる。後続の語としては、ベ（推量、意志、勧誘）、ス・シ（丁寧）、ナ（禁止）、ド（伝聞、仮定）、ハ（意志）、ケ（回想）、コンダラ（仮定）などがある。

仮定形は共通語と同じようにバ（仮定）がつき、命令形は言い切りの形である。

## 2.2 共通語との助詞の違い

結合価に関連する主な格助詞の用法を表2に示す。表で括弧に入っているのは省略可能であることを意味する。

表2 置賜方言の格助詞

主格	-
対格	(バ)
方向	サ
受け手	サ
場所	サ
目的	サ

### (1) 主格を示す助詞

主格を示す助詞は、多くの場合には省略される。しかしながら、話し言葉としての語調を整えるために、次の例のように、文としてのリズムの形で残る場合がある。この場合の「ア」は、完全な助詞というよりは語調に近い。

(例) アメアフル（雨が降る）

### (2) 対格を示す助詞

対格を示す助詞もほとんど省略されるが、強調などの場合に「バ」が用いられる。

(例) キンナ マヂデ ホンバ カッテキタ  
(昨日町で本を買ってきた)

### (3) 方向を示す助詞

方向を示す助詞は「サ」が用いられる。

(例) キシャデ トウキョウサ イグ  
(汽車で東京に行く)

### (4) 受け手を示す助詞

受け手を示す助詞も「サ」が用いられる。

(例) コノホンバ オメサ カシヤル  
(この本をお前に貸してやる)

### (5) 場所を示す助詞

場所を示す助詞は、「サ」が一般的であるが、「ニ」、「エ」も用いられる。しかし、これらは現在「サ」によって駆逐されつつある。

(例) ホンソゴサ アル（本はそこにある）

表3 名詞の意味素性と例

略号	Full Name	素性名	例
DIV	Diverse		
CON	Concrete	具体名詞	
ANI	Animal	動物	ベゴ(牛)、ワンコ(犬)、ビッキ(蛙)
HUM	Human	人間	オボゴ(子供)、シャデ(弟)、シェンシェ(先生)、カガ(嫁)
ORG	Organization	組織、機関	ガッコ(学校)
PLA	Plant	植物	
PAR	Parts	生物の部分	オッポ(尾)
NAT	Natural	自然物	
PRO	Products	生産物、道具	オチャオギ(茶菓子)、ゴツォ(ご馳走)、シャッポ(帽子)
PHE	Phenomenon	現象名詞	
ABS	Abstract	抽象名詞	ボダユギ(あられ)
ACT	Action	動作、作用	ツラダシ(訪問)、ネブカゲ(居眠り)
MEN	Mental	精神	
LIN	Linguistic Products	言語作品	テンツ(嘘)
CHA	Character	性質	
REL	Relation	関係	
LOC	Location	空間、方角	グルリ(周囲)、カマバ(火葬場)
TIM	Time	時間	キンナ(昨日)、ユンベ(昨夜)
QUA	Quantity	数量	ヒシテ(一日中)

(6) 目的を示す助詞

目的を示す助詞も主として「サ」が用いられるが、ごくまれに省略したり、「ニ」を用いる場合もある。場所を示す助詞と同様に、若年層に「サ」を用いる傾向がある。「サ」はこのように、次第に用法を拡大しつつある。

(例) ミンナデアソビサ イグ  
(みんなで遊びに行く)

(7) その他

少数ではあるが、必ず数量詞を伴う動詞がある。この場合には、格助詞の省略と区別するために、数量詞の入るべきところに $\phi$ という記号をつける(例: 3日イキル)。

### 3 結合価と意味素性

IPAL にならって、抽出した置賜方言の動詞 182 語(音便形も含む)について、それらの結合価と、付随する名詞の意味素性を調査した。選択した語は、資料 [5, 6] に掲載されていて現在も使用されているものに対して、

主として内省と聞き取りに基づいて、結合価と意味素性を付与した。

予備的な調査から、意味素性は、最も粗い動詞辞書の意味素性を用いれば十分であることが推測されたので、今回はこの意味素性を用いた。用いた意味素性と、方言名詞の例を表 3 に示す。例の欄が空なのは、共通語の名詞とほとんど同じで、方言独特の表現が見当たらなかった箇所である。以下に、動詞「あがる」の記述例を示す。この例において、助詞が省略される部分は下線で示し、任意の部分は括弧に入れて示す。

・あがる(ラ行五段)

(1) 田畑の仕事を終わって家に帰る

ミンナ シゴトカラ アガルコロダベ

(みんなが仕事から帰る頃だろう)

N[HUM].+N[ACT] から +V

(2) 学校から帰る

モウスグ ガッコウカラ アガッテクル

(もうすぐ学校から帰ってくる)

- N[HUM/ANI/PRO/ORG]-(+N[LOG] (ば) )  
+N[LOC] から+N[LOC] さ+V
- (3) 学校を卒業する  
チュウガッコウ (バ) アガル  
(中学校を卒業する)  
N[HUM]-+N[ORG] (ば) +V
- (4) 入学する  
ハルニ コウコウサ アガッタバツカリダ  
(春に高校に入学したばかりだ)  
N[HUM]-+N[ORG] さ+V
- (5) 食べる (丁寧)  
サクランボ アガットゴエ  
(サクランボを召し上げて下さい)  
N[HUM]-+N[CON] (ば) +V
- (6) 寄る、家にはいる  
ウヂノナガサ アガル  
(家の中に入る)  
N[HUM]-+N[LOC] さ+V

上記の例から分かるように、「あがる」には、(3) のように、「学校を卒業する」という意味と、(4) のように、「入学する」という一見相反する意味があるが、格助詞が異なることから、二つの意味を区別することができる。また、その他の多義性も、結合価と意味素性ですべて分類できる。

置賜方言には、数は少ないが、一つの動詞で能動と受動の両方の意味を持つものもある。以下に例を示す。

- ・おぼる (=おぶる) (ラ行五段)
- (1) おんぶする  
バッチャ オボゴバ オバル  
(おばあさんが子供をおんぶする)  
N[HUM/ANI]-+N[HUM/ANI] (ば) +V
- (2) おんぶされる  
バッチャ オボゴサ オバル  
(おばあさんが子供におんぶされる)  
N[HUM/ANI]-+N[HUM/ANI] さ+V

この場合でも、結合価と意味素性を用いれば意味を区別できる。

#### 4 おわりに

本研究では、置賜方言の動詞について、その結合価と付随する名詞の意味素性を調査し、この手法が動詞の意味の区別に有効であることを確認した。また、結合価に

用いる格助詞の用法についても調査して、その働きを明らかにすることができた。

しかしながら、格助詞の省略の多い方言では、対格を示す助詞まで省略されると、主格と対格が同じ意味素性を持つ場合に、どちらが主格かが曖昧になる場合がある。たとえば、次のような例がある。

- ・せめる (=いじめる) (マ行下一)  
オメ コネゴバ セメンナ  
(お前は子猫をいじめるな)  
N[HUM/ANI]-+N[HUM/ANI] (ば) +V

この例で、助詞「バ」が省略されると、語順で判断するしかなくなる。

また、本研究では触れなかったが、終助詞についてもさまざまな用法が知られている。これらについても今後調査する予定である。

方言は現在急速に変貌しつつある。世代によって言葉は異なるし、全体としては衰退の方向にあると言われていいる。逆に、若者言葉として、新しい方言を作る場合も少数ある。ここでは、そこまでは扱わなかった。これらの結果に基づいて、方言解析システムを構築することや、方言と共通語との機械翻訳システムを作成することが今後の目標である。

#### 参考文献

- [1] 板橋 秀一代表：音声の方言的特徴の抽出と方言音声の判別に関する研究、平成5年度科研費報告書(1993)
- [2] 我妻 信博：結合価と意味素性を考慮した置賜方言の機械処理用辞書、山形大学卒業論文(1998)
- [3] 横山 晶一、安野 克彦：方言の機械処理に関する予備的考察—庄内方言を対象として、信学技報 NLC95-45(1995)
- [4] 藤原 与一：方言学の方法、大修館書店(1977)
- [5] 上村 良作監修：米沢方言辞典、桜楓社(1969)
- [6] 平山 輝男編：山形県のことば、日本のことばシリーズ9、明治書院(1997)
- [7] 安野 克彦：庄内方言の解析システムに関する予備的研究、山形大学卒業論文(1995)
- [8] IPA 技術センター：計算機用日本語基本動詞辞書 IPAL、解説編、辞書編(1987)