

# 対話行為と情緒を解析するための文末表現パターンの作成

徳久 雅人 前田 浩佑 村上 仁一 池原 悟  
鳥取大学工学部知能情報工学科 〒 680-8552 鳥取県鳥取市湖山町南 4-110  
E-mail: { tokuhisa, s032049, murakami, ikehara } @ike.tottori-u.ac.jp

**あらまし** 本稿では、テキスト対話において、話し手の意図と情緒を解析するための文末表現パターンを作成する方法を示す。日本語の文末表現は、話し手の主体性が表れやすいと言われているが、助詞や助動詞を単独に解析したのでは、その主体性が捉え難い。そこで、本稿では、文末表現を分解せずに意味解析することを狙い、文末表現パターンの作成を行う。このパターンは、タグ付きテキスト対話コーパスを手作業で分析することで作成する。具体的には、まず、対話行為を表す単位で文末表現を台詞から抽出し、その表現を基礎にパターンを記述する。次に、その文末表現と共に生じやすい話し手の情緒を、その文末表現の持つ感情的なニュアンスと見なして集計する。この方法を行った結果、「行動を表す台詞」および「属性を伝達する台詞」(合計 8,935 件)から、3,371 件のパターンが作成された。

**キーワード** 対話行為, 文末表現, 情緒, 感情, パターン言語処理, 非線形言語モデル

## Constructing Suffix Expression Patterns to Analyze Dialogue Acts and Emotions

Masato TOKUHISA Kousuke MAETA Jin'ichi MURAKAMI Satoru IKEHARA  
Department of Information and Knowledge Engineering, Faculty of Engineering, Tottori University  
4-110 Koyama-Machi-Minami, Tottori, 680-8552, Japan  
E-mail: { tokuhisa, s032049, murakami, ikehara } @ike.tottori-u.ac.jp

**Abstract** In this paper, we explain how to construct suffix patterns for analyzing dialogue acts and emotions from text-dialogue. It is said that Japanese suffix expressions express the speaker's subjectiveness. It is, however, difficult to analyze the subjectiveness by the compositional method of the auxiliary words. Therefore, we aim to realize the non-compositional method and try to construct the suffix patterns from the previously constructed text-dialogue corpus with dialogue act tags and emotional tags. For instance, suffix expressions are extracted from the utterances of the corpus in multi-words just expressing the dialogue act, and then new patterns are constructed on the suffix expression basis. As the results of the pattern generation from “utterances for plan” and “utterances to inform attribution” (8,935 utterances), the 3,371 suffix expression patterns are newly obtained.

**Keyword** dialogue act, suffix expression, emotion, affect, pattern language processing, non-compositional language model

### 1 はじめに

本稿では、テキスト対話において、話し手の意図と感情を解析することを目指し、その解析に必要な発話文の文末表現について、パターン形式の知識ベースを構築する。

日本語表現は、主体的表現と客体的表現で構成されており、主体的表現に話し手の意図や感情が表されると言われている[1]。従来研究において、主体的表現は、助詞・助動詞・形式名詞などが複数組み合わせ合わせた文型を単位として意味分類がなされている[2]、[3]。こうしたまとめ方がなされたのは、助詞・助動詞の個々の単語の意味を合成することにより主体的表現の意味を導くことができ

ないことによるものと考えられる。

ところで、これは、言語表現と意味の間に存在する非線形性の現れである。言語の非線形性を扱うための言語モデルとして、「非線形言語モデル」が提案されている[4]。このモデルでは、単語の表す概念を「単一概念」、複数の単語で表す概念を「複合概念」、そして、単語や句などの表現を結び付ける表現構造の表す概念を「抽象的概念」と呼んでいる<sup>\*1</sup>。1つの文は、一般に複合概念であり、文のどの一部でも変更するとその概念は変化してしまうのだが、文の表現構造とその抽象的概念を固定しておけば、文の複合概念は、表現構造により結び付けられた部分的な表現の表す概念と抽象的概念とに分離することができる。この分離のための言語

\*1 たとえば、「富士山は大山より高い。」という文は「～は～より～」という表現構造があり「富士山」、「大山」、「高い」という3つの単語が結び付けられている。この表現構造は、「2者の比較(ただし、ここでは説明のためあえてラベル名を与えている)」という抽象的概念を表している。

知識ベースの形式として、文型パターンの形式が採用され、その記述仕様が示された[5]。なお、表現構造のことを「非線形要素」、表現構造で結び付けられている部分表現のことを「線形要素」と呼ばれている。

さて、テキスト対話の理解において、発話文から対話行為を解析することが第一の課題となるのだが、その解決に非線形言語モデルを適用するには、発話文の表現構造の表す抽象的概念を、対話行為に対応させ、その表現構造をとらえる文型パターンを作成する必要がある。しかし、従来研究[2],[3]で示された文末表現とその解説は、文型パターンの作成に参考になるものであるが、対話行為を解析することを中心としておらず、抽象的概念の捉え方が本稿で目指すものと異なるため、そのまま採用することができない。

そこで、本稿では、タグ付きテキスト対話コーパスを分析することにより、対話行為を解析することのできる文末表現パターンの作成を行う。そして、文末表現と感情との共起関係を集計し、文末表現における感情的なニュアンスについて分析するとともに、文末表現の周辺の構造について検討する。

## 2 タグ付きテキスト対話コーパス

本稿では、既に構築されているテキスト対話コーパスを改良して使用する。このコーパスには、対話行為タグ、表情タグ、および、情緒タグが付与されている[6],[7]。

収録されているテキスト対話は、漫画「ちびまる子ちゃん」の第1巻から第10巻までである[8]。漫画に描かれる出来事は架空の事であるが、使われる日本語は、一部の登場人物を除いて、自然なものであると言えるので、言語表現の分析対象となりうる。

まず、付与されている対話行為タグは、次のとおり設計された：

{行為タイプ, 対象タイプ, 補助属性1, 補助属性2}

各項目は、表1のとおりに要素が定義された。行為タイプは、比較的常識的な分類である。対象タイプは対話内容として、「対話者の行動が描かれること」を想定して設けられたものである。具体的には、「勧誘」など「対話相手を誘うための対話」である。対話から行動を理解することを研究する上で、行動の原因となる生理的・心理的状況から欲求や情緒の発生と、プランや行動過程との関連を捉える必要がある。それらについての断片的な情報は発話に表されると想定されることから、対象タイプが定められた。

ここで、項目要素の組み合わせに次の関係が定められている：

- 「はい」、「いいえ」、「その他」は対象タイプをとらない。
- 対象タイプには補助属性が付随する。
- 「質問」、「要求」には、「実在」の補助属性は付随しない。
- 「要求」には、「過去」の補助属性は付随しない。

したがって、111通りのタグの使用が想定される。たとえば、「うん、いいよ。」という台詞に対しては{はい}というタグが付与され、「明日、スキーに行くよ。」という台詞に対しては{伝達, プラン, 非過去, 想像}というタグが付与される。

対話行為タグは、台詞の1文ごとに付与されている。

表1 対話行為タグの構成要素

項目	要素	説明
行為タイプ	質問	仮説が述べられていない疑問文
	確認	仮説が述べられている疑問文
	伝達	事態を述べる文
	要求	聞き手の行動を定める文
	はい	「はい」、「うん」など承諾を表す文
	いいえ	「いいえ」、「ちがう」など否定を表す文
対象タイプ	その他	挨拶や相づちなど上記以外のもの
	生理	身体の内外部や表面で感じる事
	欲求	生理や社会的に起因して生じる願望
	情緒	喜び、悲しみなどを感じる事
	プラン	人物の行動の1つ以上の系列
	属性	物事に対する価値や性質
補助属性1	その他	上記以外のもの
	過去	対象は過去のものである
補助属性2	非過去	対象は過去のものでない
	否定	対象は想像上であり、否定する
	想像	対象は想像上であり、否定しない
	実在	対象は実在した、あるいは、実在する

一方、表情タグおよび情緒タグは、漫画のコマに登場する人物の単位に付与されている。人物が登場せず、台詞だけの部分には付与されていない。

表情タグは、次の7種類である：

{幸福}, {悲しみ}, {嫌悪}, {驚き}, {恐れ}, {怒り}, {背後}, {普通}

エックマンの表情分類[9]に従い5種類が採用され、漫画の描き方の特徴に着目して{背後}のタグが設けられた。

情緒タグは、次の9種類である：

{喜び}, {悲しみ}, {好ましい}, {嫌だ}, {驚き}, {期待}, {恐れ}, {怒り}, {なし}

この分類は、Plutchikの8つの基本情緒を参考にした[10]。Plutchikの分類は感情語彙の分類によるものであるため、言語処理に相性が良いものと思われる。たとえば、言語表現には、時制や感嘆という特徴があるが、この分類には、予測に関する情緒として「期待」や「恐れ」が、感嘆に関する情緒として「驚き」が含まれている。また、対話において対話者の「怒り」は円滑なコミュニケーションの障害となるため「悲しみ」や「嫌だ」とは異なる扱いが必要になると予想されるため、この分類には「怒り」が含まれている。

使用するタグ付きコーパスの規模を表2にまとめ、一部を表3に例示する。

表2 使用するタグ付きコーパスの規模

項目	規模
ナレータを除く台詞	26,738 文・315,240 文字
付与された対話行為タグ	26,603 個
付与された表情タグ	17,676 個
付与された情緒タグ	20,895 個

※ 表情{普通}に対して本稿で新たに情緒タグを追加した。

表 3 タグ付きコーパスの例（台詞は文献[8]第3巻より引用）

#	頁	コマ	話し手	台詞	対話行為タグ	表情タグ	情緒タグ
1	87	7	お母さん	あれ、まる子、何してるの？	[質問, プラン, 非過去, 想像]	〈驚き〉	《驚き》
2			まる子	あ、まる子ね、今から静岡のお婆ちゃんのところに行こうと思 って、支度しているんだよ。	[伝達, プラン, 非過去, 実在]	〈幸福〉	《喜び》, 《期待》
3	88	1	お母さん	何言ってるの、この子はっ。	[その他]	〈怒り〉	《怒り》
4				誰がお婆ちゃんちに連れて行ってってくれるって言ったのよ？	[質問, プラン, 過去, 想像]		
5			まる子	誰も言っていないよ。	[伝達, プラン, 非過去, 否定]	〈普通〉	《なし》
6		2	お母さん	じゃあダメでしょ。	[伝達, 属性, 非過去, 実在]	〈怒り〉	《怒り》
7				行けないじゃないの。	[伝達, プラン, 非過去, 否定]		
8			まる子	まる子一人で行くもん。	[伝達, プラン, 非過去, 想像]	〈普通〉	《なし》
9		3	お母さん	お馬鹿ッ。	[伝達, 属性, 非過去, 実在]	〈怒り〉	《怒り》
10				お婆ちゃんの家は遠いのよ。	[伝達, 属性, 非過去, 実在]		
11				電車とバスに乗るのよ。	[伝達, プラン, 非過去, 想像]		
12		4	まる子	電車は電車の運転手さんが、バスはバスの運転手さんが、 みんな運転してくれるもん。	[伝達, その他, 非過去, 想像]	〈幸福〉	《期待》
13				まる子は乗ってるだけだもん。	[伝達, プラン, 非過去, 想像]		
14				一人でできるもん。	[伝達, プラン, 非過去, 想像]		
15			お母さん	まあナマイキ言って。	[伝達, プラン, 非過去, 実在]	-	-
16		5	まる子	だから、今から行くよ。	[伝達, プラン, 非過去, 想像]	〈幸福〉	《期待》
17			お母さん	コラ、待ちなさい。	[要求, プラン, 非過去, 想像]	〈怒り〉	《怒り》

※ #はこの表における通番。#15 では、コマ内に「お母さん」の表情が描かれていないので、表情タグと情緒タグは付与されていない。

### 3 文末表現パターンの作成方法

タグ付きコーパスを利用して文末表現パターンを作成する。作成の手順は次の通りである：

**手順1. 文末表現の認定：** 対話行為に注意しながら、台詞から文末表現とする部分を抽出する。

**手順2. パターン化：** 文末表現を非線形要素として字面で記述し、それ以前の部分を線形要素として変数で記述する。そして、制約的な記述子を追加して、パターン化する。

**手順3. 対話解析用の情報付与：** パターンが文に適合する際にその文の意味解析の結果となる情報をパターンに付与する。

**手順4. 感情的ニュアンスの付与：** 台詞と共に存在する情緒タグから、感情的なニュアンスの解析できることを期待して、作成したパターンに情緒タグの共起頻度を付与する。1つのパターンには複数の台詞が対応するので、それらの台詞の情緒タグを集計に使用する。

#### 3.1 文末表現の認定

台詞の部分表現のうち文末表現とみなすための条件は次のとおりである：

- 対話行為を表す部分を文末表現とみなす。
- 対話行為の内容を表す部分は文末表現とみなさない。

たとえば、「一緒に魚釣りに行こうよ。」という発話文に、[要求, プラン, 非過去, 想像]という対話行為タグが付与されていたとする。「一緒に魚釣りに行こうよ。」までが、対象タイプの内容、すなわち、プランの内容を表しており、「うよ。」の部分が「プランの実行を要求する」という対話行為を識別するための鍵となる表現である。このことから、文末表現と認める部分は、「うよ。」となる。

他に、幾つかの例を示そう：

- 台詞:「お腹が痛いの？」  
対話行為:[確認, 生理, 非過去, 想像]  
→ 文末表現:「～の？」  
生理内容:「お腹が痛い」
- 台詞:「当時は、地動説でも良かった。」  
対話行為:[伝達, 属性, 過去, 実在]  
→ 文末表現:「～た。」  
属性内容:「当時は、地動説でも良い」

#### 3.2 文末表現のパターン化

対話行為を解析するためには、文末表現の字面だけでなく、文末表現より先行する表現部分における統語的・意味的な情報が必要である。

たとえば、「～の？」だけでは「今から食事に行くの？」という表現も適合するため、[確認, 生理]という対話行為の解析は不適切である。適切に適合させるには、パターンが、統語的に「〈形容詞節〉の？」であり、意味的に「〈身体状態〉の？」でなければならない。

したがって、前節の例に対しては表 4 のようにパターン化する。まず、統語的な制約を課すために、変数と関数を使用する。CLV は、動詞述語の節を表す変数である。CLAJ は、形容詞述語の節を表す変数である。ここには記載していないが、名詞述語の節として CLN も使用できる。次に、<sup>^</sup>mizen, <sup>^</sup>renyou, <sup>^</sup>kihon は関数である。関数の左側の記述子に対して活用形の制約を課す。

次に、意味的な制約を課すために、変数に意味属性制約を付与することが一つの方法である(たとえば、日本語語彙大系[11])。また、変数に対応する表現の用例を保持しておき、運用時に参照する方法もある。本稿では、初期の取組として、意味的な制約は課さずにパターン化を行う。

表 4 解析用情報を付与したパターン化の例

<ul style="list-style-type: none"> <li>パターン: CLV^mizen うよ。</li> <li>対話行為: [要求, プラン, 非過去, 想像]</li> <li>内容: CLV</li> <li>パターン: CLAJ^kison の ?</li> <li>対話行為: [確認, 生理, 非過去, 想像]</li> <li>内容: CLAJ</li> <li>パターン: CLAJ^renyou た。</li> <li>対話行為: [伝達, 属性, 過去, 実在]</li> <li>内容: CLAJ</li> </ul>
---

### 3.3 対話行為解析用の情報の付与

パターンの適合により対話行為の解析ができるように、解析用の情報を付与する。「対話行為」のスロットによりその種類を、「内容」のスロットに「プラン」、「生理」、「属性」の内容を記入しておく。内容のスロットには、変数を記述することが多いが、字面による表記などでも構わない。

### 3.4 感情的なニュアンス

同一のパターンとなる台詞に付与されている情緒タグを集計して、その比率より感情的なニュアンスを調べる。

たとえば、対話行為 [要求, プラン] を表すパターン「CLV^mizen うよ。」の元になった台詞と情緒タグとして、以下のものがある:

- 「ねえお父さん花札しようよ。」〈幸福〉, 《喜び》, 《期待》
- 「買おうよ。」(表情タグ・情緒タグなし)
- 「ねエ、文鳥飼おうよ。」〈普通〉, 《期待》
- 

その結果、次の頻度情報が得られる。

パターン: CLV^mizen うよ。 [要求, プラン, 非過去, 想像]

喜	好	期	悲	恐	嫌	怒	驚	無	?	タ計	セ計
20	1	31	0	2	4	0	1	1	2	62	44

※?は情緒タグなし, “タ計”はタグの合計, “セ計”は台詞の合計

この結果より、この文末表現で構成される台詞の 70% (31/44) は《期待》とともに使われていたことが分かる。

## 4 文末表現パターンの作成結果

### 4.1 対象

本稿では、行動に関する発話を重視しているので、[要求, プラン], [伝達, プラン], [確認, プラン]について、ならびに、使用頻度の高かった[伝達, 属性]についての対話行為に焦点を当て、パターン化を進める。

### 4.2 プランについての対話行為

#### 4.2.1 [要求, プラン]の場合

本コーパスには、2,166 件の台詞に[要求, プラン]のタグが付与されていた。そこからパターンを作成したところ、598 パターンが得られた。頻度の高いパターン(上位 10 件)とその感情的ニュアンス(上位3件)を表 5 にまとめる。この表について次のことが言える:

- 《期待》の占める割合が高い。

- 「なさい。」で《怒り》の割合が高いのは、親が子供に使うための発話だからである。
- 「CLN。」は主に「頂戴。」から作られたパターンである。
- 「ないでよ。」は「否定」の言い方である。他と異なり、《嫌だ》の割合が高い。

「否定」かどうかにより、情緒の内訳に違いが生じ易いと思われるので、表に記載していない[要求, プラン]のパターンについても調べた。「否定」の言い方の場合、《嫌だ》, 《怒り》の割合がそれぞれ 37%, 31%であり、《喜び》, 《期待》の割合が 9%, 10%であった。これに対して否定の対極となる「想像」の言い方の場合、《嫌だ》, 《怒り》の割合が 15%, 14%であり、《喜び》, 《期待》の割合が 28%, 38%であった。

表 5 使用頻度の高い[要求, プラン]の文末表現パターン

#	パターン	頻度	頻度の高い情緒
1	CLV^renyou てよ。	125	期待(65), 嫌だ(38), 喜び(23)
2	CLV^renyou なさい。	114	怒り(44), なし(16), 喜び(15)
3	CLV^mizen う。	110	喜び(62), 期待(59), 好ま(12)
4	CLV^meirei 。	103	期待(34), 怒り(27), 嫌だ(17)
5	CLV^renyou て。	93	期待(40), 喜び(27), ? (14)
6	CLN 。	50	期待(35), 喜び(14), 怒/? (5)
7	CLV^mizen うよ。	44	期待(31), 喜び(20), 嫌だ(4)
8	CLV^renyou てね。	40	喜び(16), 期待(13), 恐れ(9)
9	CLV^renyou て下さい。	38	期待(12), なし/? (8), 喜び(5)
10	CLV^mizen ないでよ。	35	嫌だ(20), 怒り(8), 期待(4)

#### 4.2.2 [伝達, プラン]の場合

本コーパスには、2,767 件の台詞に[伝達, プラン]のタグが付与されていた。そこからパターンを作成したところ、1,231 パターンが得られた。頻度の高いパターンを表 6 にまとめる。この表について次のことが言える:

- 《喜び》の占める割合が高い。
- 「〜。」は、[要求, プラン](表 5)にも存在する。たとえば、「一緒に帰ろう。」は「要求」であり、「これは、いつかは使おう。」は「伝達」である。

補助属性の組み合わせにより、時制的な区別が可能である。[非過去, 想像]は未来, [非過去, 実在]は現在, [過去, 実在]は過去をそれぞれ表すと言える。ここで、《期待》の割合を調べてみると、それぞれ 36%, 16%, 14%であった。しかし、《恐れ》も同様に調べたが、それは 9%~10%で大差が無かった。

表 6 使用頻度の高い[伝達, プラン]の文末表現パターン

#	パターン	頻度	頻度の高い情緒
1	CLV^kison よ。	108	喜び(41), 期待(40), 嫌だ(13)
2	CLV^renyou たよ。	71	喜び(39), 期待(14), 好/恐(7)
3	CLV^mizen う。	66	期待(33), 喜び(27), 恐れ(8)
4	CLV^kison よ。	37	喜び(19), 期待(14), 嫌だ(5)
5	CLV^renyou てる。	31	驚き(12), 喜び(6), 恐れ(5)
6	CLV^kison ぞ。	27	期待(16), 喜び(10), 嫌だ(5)
7	CLV^renyou ます。	26	期待(8), 喜び(6), ? (4)
8	CLV^mizen ないよ。	24	恐れ/嫌だ/怒り(5)
9	CLV^renyou たんだよ。	24	喜び(12), 嫌だ(4), 好/期(3)
10	CLV^kison 。	23	喜び/期待(7), 嫌だ(5)

### 4.2.3 [確認, プラン]の場合

本コーパスには、615 件の台詞に[確認, プラン]のタグが付与されていた。そこから 385 パターンが得られた。頻度の高いパターンを表 7 にまとめる。この表について次のことが言える：

- 《驚き》の割合が高いことがある。「確認」を行う状況として、思いがけない情報、すなわち、驚くべき情報の聞き直しが考えられる。
- 「CLN ?」は、たとえば、「こう?」や「今、帰り?」という台詞から作られたパターンである。

表 7 使用頻度の高い[確認, プラン]の文末表現パターン

#	パターン	頻度	頻度の高い情緒
1	CLV^kohon の?	31	驚き(16), 嫌だ(6), 恐れ(5)
2	CLV^renyou たの?	18	驚き(8), 喜/嫌/怒(3)
3	CLV^renyou た?	16	喜び/期待(7), 恐れ(3)
4	CLV^kohon ?	14	期待(8), 喜び(4), なし(2)
5	CLV^mizen うか。	9	喜び(7), 期待(4), 嫌/驚/無(1)
6	CLV^renyou てるの?	7	期待(3), なし(2), 喜/怒/?(1)
7	CLV^mizen ないの?	7	期待(4), 喜/悲/恐/驚(1)
8	CLN ?	7	喜/驚(2), 期/怒/無/?(1)
9	CLV^renyou てるんでしょ。	6	嫌だ(3), 好/期/恐/怒/驚/無(1)
10	CLV^renyou てみようか。	6	期待(5), 喜び(3), 恐れ(1)

### 4.3 属性についての対話行為

#### 4.3.1 [伝達, 属性]の場合

本コーパスには、3,387 件の台詞に[伝達, 属性]のタグが付与されていた。そこから 1,157 パターンが得られた。頻度の高いパターンを表 8 にまとめる。この表について次のことが言える：

- 《喜び》と《嫌だ》の割合が高い。
- 「CLN。」は、「バーカ。」や「ダメ。」などの台詞から作られたパターンである。
- 感情的なニュアンスを伴う文末表現は、この表においては見られない。それにも関わらず偏った情緒との共起が目立つのは、属性の内容に依るものと考えられる。
- 属性を伝達する場合、形容詞述語による台詞(たとえば、「危ないぞ。」)が予想されるのだが、名詞述語が目立つ。

表に記載していない[伝達, 属性]のパターンについても調べたところ、述語の品詞として、名詞、形容詞、動詞、副詞、その他があり、それぞれの頻度は、1,454 件、1,217 件、581 件、86 件、49 件であった。

表 8 使用頻度の高い[伝達, 属性]の文末表現パターン

#	パターン	頻度	頻度の高い情緒
1	CLN。	230	喜び(72), 怒り(59), 嫌だ(53)
2	CLAJ^kohon。	95	喜び(36), 嫌だ(23), ?(16)
3	CLN だよ。	92	喜び(32), 嫌だ(26), 期待(12)
4	CLAJ^kohon よ。	66	嫌だ(27), 喜び(18), 悲しみ(12)
5	CLN だね。	64	喜び(24), 嫌だ(16), 怒り(11)
6	CLAJ^kohon ね。	54	喜び(27), 驚き(10), 好ましい(9)
7	CLN だ。	46	嫌だ(15), 喜び(10), 怒り(9)
8	CLN よ。	41	嫌だ(19), 怒り(10), 恐れ(6)
9	CLN ッ。	38	怒り(27), 嫌だ(13), 悲しみ(4)
10	CLAJ^kohon じゃん。	38	喜び(15), 期待(9), 怒り(8)

## 5 検討

### 5.1 発話文の構造

文末表現と対話行為を中心に発話文の構造を考察する。文末表現パターンの作成を通じて、発話文の構造を図 1 のように捉えることができた。対話行為の表現として文末表現を対応させ、それより前の部分を「内容の表現」とここでは呼んでいる。「内容の表現」を見てみると、主節だけでなく従属節や引用節が見られ、また、主節を構成する述語には文末表現に含まれなかった付属語的な表現が見られることがある(ここでは「付属表現1」と呼ぶ)。一方、文末表現の後には、対話行為を表すとは言いにくい付属表現が見られることがある(ここでは「付属表現2」と呼ぶ)。

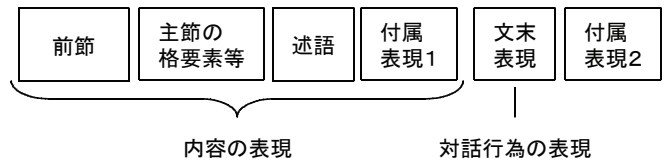


図 1 対話行為と文末表現を基準とした発話文の構造

これらの構造に対して、パターン化の際に判断を要した点を以下に挙げる：

- 重文複文は「内容の表現」であるのか？
- 内容の表現に使われる述部は、述語のみであるのか、それとも、「付属表現1」を認めるか？
- 文末が同じ台詞に、異なる対話行為の対応を認めるか？
- 文末表現よりも後ろ側に、対話行為の表現とは言えない表現のあるとき、どのようにするか？

本節では、上記について例を交えて対応の仕方を説明する。

#### 5.1.1 内容の表現における構造

本稿では、台詞が重文複文の場合、主節以外の節を内容の表現に含めてパターン化することにした。内容の表現に含めることが不適当と思われる台詞は、複雑な対話行為を表すものであるため、文末表現パターンではなく、文中の構造にも着目した文型パターンでなければならないと考えた。本稿で作成したパターンは、複雑な場合は主節のみにしか考慮していないことになる。

内容の表現に含めることの利点として、内容の表現の表す概念が分解されずに得られることである。そして、内容の表現においてさらに表現構造を解析することが重要になってくる。以下に台詞とパターンの例を示す：

- 「それにしても、その作文へたっぴいだから、もういっぺん書き直したほうがいいぞ。」(CLAJ^kohon ぞ。)
- 「「子供Vロート」じゃなきゃ駄目だよ。」(CLN だよ。)

前者は、「～だから～たほうが～」という構造により「理由付きプランの評価」という属性の内容を表し、後者は、「～じゃなきゃ～」という構造により「指定外への評価」という属性の内容を表している。これらは属性表現パターン集として構築することが考えられる。

### 5.1.2 内容の表現における述語後続の表現

述語に後続する助詞や助動詞などが、対話行為を表すためではなく、内容を表すために使われる場合がある。

- 「ステージまで僕にエスコートさせておくれよ。」

プランの内容として「(相手が)ステージまで僕にエスコートさせる」と解釈すべきところである。「せる」を内容の表現から外すと、プランの内容の格解析に支障をきたす。

対策は、「CLV」に適合するのは、動詞のみではなく、「せる」や「られる」の付随した動詞も認めることである。したがって、上記の台詞は、「CLV^renyou せておくれよ。」ではなく「CLV^renyou おくれよ。」とパターン化することができる。

また、「られる」も同様の問題が生じる。格解析だけでなく、可能動詞の判断とも関連する。たとえば、[伝達, 属性]においては、可能性という属性を扱うので、「CLVK」という可能動詞と「られる」付き動詞を認める変数を使う。以下の台詞は「CLVK^kohon ぞ。」と「CLVK^kohon よ。」とそれぞれパターン化することができる：

- 「こうでザリガニの子が捕れるぞ。」(可能動詞)
- 「これで安心して新学期を迎えられるよ。」(動詞+られる)

しかし、その他の表現において、判断が不明確なものがある。以下の例では、「てくれる」はプランの内容として必須ではないし、[確認, プラン]を解析する上でも必須ではない：

- 「ほんとにしっかり持ってくれてるでしょうね。」

今後、判断の仕方を再検討する必要がある。

### 5.1.3 文末表現の多義性

多義の発生の仕方には、大きく次の2つがある：

- 1つの台詞に複数の対話行為が見られる場合
- 異なる対話行為である複数の台詞から同一の文末表現パターンが作られる場合

前者の例を示す：

- 「まるちゃんも、自転車乗れるようになりなさいよ。」
- 「早く行った方がいいわよ。」

[要求, プラン]として上記2つの台詞を解釈すると、二重線部がプランの内容の表現でありそれ以降が文末表現となる。[要求, 属性]や[伝達, 属性]として解釈すると、下線部全てが属性の内容の表現であると言える。この判断の違いは、対話行為タグの付与の段階から生じる。

対策は、第 5.1.1 節のように内容表現における構造解析を前提として、下線部全てを内容の表現と見なすことである。

後者は、パターンを運用する問題である。運用時に曖昧性解消に必要な情報(たとえば、意味属性制約)をあらかじめ準備しておくことが、パターンの作成時での対策の一つである。

### 5.1.4 対話行為の文末表現よりも後の表現

倒置の表現などでは、通常では対話行為を表す文末表現の後に、個別的な表現が存在する。そのような台詞の例を示す：

- 「下手だな、俺って。」

本稿の段階では、倒置の対策ができておらず、「CLN だな、俺

って。」と文末表現の一部としてパターン化した。汎用性を高めるために「俺」という個別の部分を変数化しなければならないが、節変数 CLN に「俺」の部分の変数 N を合成するためのパターン記述子が設計できていない。

## 5.2 文末表現からの情緒推定の可能性

表 5 ~ 表 8 を見る範囲では、文末表現と情緒の間に緩やかな傾向は見られるものの、文末表現が情緒推定を行う上での決定的な判断材料になっているとは言い難い。

その理由として、第一に、文末表現を抽出する基準として対話行為を用いたため情緒的な特徴をパターンにあまり強く取り込められなかったこと、第二に、対話からの情緒推定では単語が語義として持つ感情的ニュアンスよりも対話文脈の方が強く影響する可能性があることが考えられる。

今後、文末表現を分解して感情的ニュアンスの強い単語を抽出することや、対話内容と情緒タグとの考察を深める必要がある。

## 6 おわりに

本稿では、対話行為と情緒を表すタグ付きテキスト対話コーパスから、「プランを表す台詞」と「属性を伝達する台詞」(合計 8,935 件)を対象に、対話行為を表す文末表現のパターンを作成し、情緒タグとの共起頻度をパターンに対応付けた。対話行為ごとに情緒的な傾向は見られたが、情緒推定のための根拠として使用するには改良を要すると言える。一方、対話行為を解析するための文末表現は適切に得られた。文末表現とその周辺の表現における構造の検討を行うことができた。

**謝辞** 本研究は、博報『ことばと文化・教育』研究助成の下で行いました。台詞の解析において、奈良先端科学技術大学院大学(松本研究室)の chasen を使用させて頂きました。

## 参考文献

- [1] 時枝誠記：“国語学原論。”岩波書店，1941。
- [2] 森田良行，松木正恵：“日本語表現文型。”アルク，1989。
- [3] 砂川有里子，駒田聡，下田美津子，鈴木睦，筒井佐代，蓮沼昭子，ベケシュ・アンドレイ，森本順子：“日本語文型辞典。”くろしお出版，1998。
- [4] 池原悟：“言語で表現される概念と翻訳の原理。”電子情報通信学会技術研究報告，思考と言語，TL2003-25，pp.7-12，2003。
- [5] 池原悟，阿部さつき，徳久雅人，村上仁一：“非線形な表現構造に着目した重文と複文の日英型パターン化。”自然言語処理，Vol.11，No.3，pp.69-95，2004。
- [6] 徳久雅人，村上仁一，池原悟：“漫画における表情に着目した情緒タグ付きテキスト対話コーパスの構築。”自然言語処理，Vol.14，No.3，pp.192-217，2007。
- [7] 徳久雅人，前田浩佑，村上仁一，池原悟：“心的状態を表す対話行為タグ付きテキスト対話コーパスの構築。”電子情報通信学会技術研究報告，思考と言語，TL2007-45，pp.25-30，2007。
- [8] さくらももこ：“ちびまる子ちゃん。”Vol.1-10，集英社，1987-1993。
- [9] P.エックマン，W.V.フリーゼン：“表情分析入門。”工藤力(訳)，誠信書房，1987。
- [10] Plutchik,R.：“Emotions and Life: Perspectives from Psychology, Biology, and Evolution.” American Psychological Association, 2002.
- [11] 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：“日本語語彙大系。”岩波書店，1997。