

国立国語研究所学術情報リポジトリ

Analyzing Utterance Intentions in Entertainment Novels

| | |
|-------|--|
| メタデータ | 言語: jpn 出版者: 公開日: 2023-03-24 キーワード (Ja): キーワード (En): 作成者: 夏目, 和子, 佐藤, 理史, Natsume, Kazuko, Sato, Satoshi メールアドレス: 所属: |
| URL | https://doi.org/10.15084/00003740 |

エンタメ小説における会話文の発話意図分析 —発話文表現文型辞書の改良に向けて—

夏目 和子 (名古屋大学)

佐藤 理史 (名古屋大学)

Analyzing Utterance Intentions in Entertainment Novels

Kazuko Natsume (Nagoya University)

Satoshi Sato (Nagoya University)

要旨

小説の会話文を生成するためのツールとして『発話文表現文型辞書』を編纂し改訂を重ねている。本辞書は、「発話意図」と「表現文型」と「話し方の特徴」で構成されており、「発話意図」の項目の抜本的な見直しが目下の課題である。本研究では、エンタメ小説の会話文を発話意図分析することによって、現在の辞書の発話意図（60種）で対応できない発話文はどれか、追加すべき発話意図は何か、どのような発話意図のラベルや体系を採用すれば、小説発話文生成時に使いやすいかなどを検討する。

1. はじめに

小説の会話文を生成するためのツールとして『発話文表現文型辞書』を編纂し改訂を重ねている (夏目 2017, 2019, 2022)。本辞書は、発話意図 (=発話の目的)、表現文型 (=発話意図を表現する言語形式)、話し方の特徴ベクトル (6軸 20次元) および例文で構成されている。2022年8月現在の辞書のサイズは、発話意図が60種、話し方の特徴を付与した表現文型のエントリ数が8,055、例文数が23,489である。発話文生成および変換における本辞書の役割について2節で紹介する。

『発話文表現文型辞書』の改良のために取り組むべき多くの課題のうち、本研究では「発話意図」項目の抜本の見直しに取り組む。まずは現在の辞書の発話意図の課題を3節で確認する。それらの課題の解決策をさぐるために、実在のエンタメ小説の会話文に対して、実際に発話意図を付与することを行なった。本稿ではその内容を報告する。

エンタメ小説の会話に発話意図を付与する際に特に留意した4点を以下に示す。

1. 発話意図をどのくらい細かく区別するか
2. 語用論的な意図をどう扱うか
3. 感情をどう扱うか
4. 発話意図がひとつに定まらない場合はどのように記述するか

2. 発話文生成における『発話文表現文型辞書』の役割

2.1 発話文表現文型辞書

『発話文表現文型辞書』は、小説の発話文生成を念頭に、「ある目的(発話意図)で発話する時、ある話し方で表すならば、この文型を使う」という情報を提供する辞書である。

発話文表現文型辞書の各エントリは、主に下記の3種類の情報で構成されている。

発話意図＝発話の目的
 表現文型＝発話意図を表現する際に使用する言語形式
 話し方の特徴ベクトル (6 軸 20 次元)

2.2 発話文生成

話文生成システム (刀山 2016) (木村 2018) (米田 2022) における『発話文表現文型辞書』の役割は、「発話意図」と「話し方の特徴ベクトル」が入力されると、それに適合した「表現文型」を提案することである。以下に例を示す。「話し方の特徴ベクトル」は、実際は6軸20次元のベクトルであるが、ここでは便宜上文字で表現する。

| | | |
|-------------------|---|--------------|
| 入力 | | 提案 |
| 発話意図＝非難-実行 | → | 表現文型＝Vt なんて！ |
| 話し方の特徴＝女性的、強い、下扱い | | |

発話文生成システムは、受け取った表現文型 (Vt なんて!) と発話内容 (チョコレートを食べる) を HaoriBricks3 (佐藤 2020) で表層化し、発話文「チョコレートを食べたなんて!」を出力する。

2.3 発話文変換

『発話文表現文型辞書』は、「話し方の特徴」を変更すると、別の「表現文型」を提案する。以下に例を示す。

| | | |
|----------------------------------|---|--------------|
| 入力 | | 提案 |
| 発話意図＝非難-実行 | → | 表現文型＝Vb やがって |
| 話し方の特徴＝男性的 2、強い、 下扱い、くだけ、ぞんざい | | |

この場合、発話文変換システム (木村 2019) は、発話文「チョコレートを食べたなんて!」を「チョコレートを食べやがって」に変換する。

『発話文表現文型辞書』は、「発話意図」を変更すると、別の「表現文型」を提案する。以下に例を示す。

| | | |
|----------------------------------|---|----------------|
| 入力 | | 提案 |
| 発話意図＝禁止 | → | 表現文型＝Vb んじゃねえぞ |
| 話し方の特徴＝男性的 2、強い、 下扱い、くだけ、ぞんざい | | |

この場合、発話文変換システムは、発話文「チョコレートを食べやがって」を「チョコレートを食べんじゃねえぞ」に変換する。

3. 発話文表現文型辞書の発話意図の課題

本辞書の発話意図を選択し体系化するために参考にした資料とその内容を以下に示す。下線資料は、この度の発話意図見直しのために新たに参照したものである。

- a. 日本語学：表現意図およびそれに応ずる文表現の分類 (国立国語研究所 1960, 1963)
 モダリティに関する項目 (日本語記述文法研究会 2003) (益岡 1992)
 談話に関する項目 (日本語記述文法研究会 2009)

評価のモダリティ (高梨 2010)
 発話機能 (山岡 2008)
 伝達的な機能による発話の分類 (宮島 1995)

- b. 日本語教育： 発話内行為およびモダリティ (グループ・ジャマシイ 1998)
 表現意図 (市川 2019)
- c. 談話分析： 発話単位タグ (荒木 1999)
 談話行為タグ (伝 2015) (居關 2018)
 発話意図タグ (入江 2005)
 発話機能の連鎖組織 (筒井 2012)

いずれも有益な情報である。しかしながら、小説発話文を生成する際に発話意図として使用しにくいものも含まれている。例えば、モダリティ (またはムード) を表す「確言」「当為」「概言」「比況」や「示唆」など日常使わない専門的用語はラベルとして使いにくい。また、主に談話分析で用いられる「情報要求」「情報提供」などの機能を表すラベルや、「評価」「確認」などやや抽象的で、それ単体で発話意図として用いるのはやや困難である。

現在の辞書の発話意図のリストを表1に示す。現状は、上記のように扱いづらいものは保留にしているため、本当に必要な発話意図が欠落している可能性がある。加えて、「推量」「断定」など情報の確からしさのモダリティに関しては、既に多くの表現文型を本辞書に収録しているが、小説発話文生成時に用いる発話意図としては問題がある。

表1 発話文表現文型辞書のグループ別発話意図

| グループ | 発話意図 | 発話意図の数 | エントリの数 |
|----------|--|--------|--------|
| 行為要求 | 命令、依頼-実行、依頼-非実行、勧誘、忠告、勧告、許可、禁止、勧め | 9 | 378 |
| 意向伺い | 申し出 | 1 | 26 |
| 態度表明 | 褒める、貶す、非難-過失、非難-実行、非難-非実行 | 5 | 212 |
| 意志表示 | 願望-実行、願望-非実行、願望-事態、願望-他者の行為、後悔-実行、後悔-非実行、決心-実行、決心-非実行 | 8 | 261 |
| 感情表出 | — | 0 | 0 |
| 情報伝達 | 説明-物事、説明-理由、説明-原因、説明-結論、説明-換言、伝聞 | 6 | 1,132 |
| 情報の確からしさ | 推量-だろう、推量-のだろう、推量-ようだ、推量-みたいだ、推量-らしい、推量-かもしれない、推量-と思う、断定-言い切り、断定-にちがいない、確信-はずだ、否定的確信-はずがない、否定的確信-わけがない、不確定 | 13 | 5,714 |
| 質問 | 確認-肯否要求、確認-未知情報要求、確認-念押し、許可要求 | 4 | 103 |
| 応答 | 肯定、承諾-依頼、依頼-勧誘、受諾-申し出、否定、断り-依頼、断り-勧誘、断り-申し出、不明、未定 | 10 | 172 |
| 感動詞 | よびかけ、別れ、感謝、謝罪 | 4 | 57 |
| | | 60 | 8,055 |

現在の発話意図の主な課題を以下にまとめる。

- 全体の課題
小説発話文生成のための発話意図リストを作成するために
 1. 保留中の発話意図を解決する
 2. 足りない発話意図を見つける

- グループ別の課題
情報の確らしさ（モダリティ）：確らしさは発話意図ではない。表現文型エントリーを適切な発話意図に振り分ける。グループの見直しも必要である。
質問：質問の中心的形式は、yes-no 質問と wh 質問であるが、辞書の表現文型が文末述語（述部）を対象としているので、とりあえず保留にしてきた。
感情表出：どのような発話意図が必要か。現状はゼロ。
情報伝達：既存の発話意図の見直し。追加でどのような発話意図が必要か。

4. エンタメ小説の発話文への発話意図付与

前節で示した発話意図の課題に解決の糸口をさぐるために、実際の小説での会話文一つ一つに対して、発話意図を付与することを行なった。

4.1 対象データ

分析の対象としては、エンタメ小説を選択した。その理由を以下に示す。

1. プロットが明解で、会話の状況や場面設定がわかりやすい
2. 登場人物のキャラクタがはっきりしている
3. 会話が多い。作品がシリーズ化されることも多いので、特定の人物の会話文を大量に収集できる

これらのことは、今後、特定の話者の表現文型を多く収集し、将来、発話意図を自動で付与するという目標に役立つかもしれないと考えてのことである。

具体的には、以下の3小説の会話文を抽出して利用した。

結城浩『数学ガール秘密のノート：微分を追いかけて』（以下、【結城】と略記）

有川浩『図書館戦争』（【有川】）

三浦しをん『風が強く吹いている』（【三浦】）

4.2 方法

3冊の小説の巻頭から各 100 発話文を抽出し、それぞれに文型と発話意図を付与した。文型は、述語を含む文には HB3 の仕様の接続形式と主要部と末尾部を記述した。まず、発話意図を付与したのち、発話意図を 3 種類に区分した。

文型辞書の発話意図を付与

新規の発話意図を付与

不明

次に、不明に区分した発話文を対象に、その理由を分析しながら、複数の発話意図を付与するか、発話意図を付与しないかを決定した。すなわち、最終的に以下の 4 種類に区分した。

1. 辞書の発話意図を付与
2. 新規の発話意図を付与
3. 複数の発話意図を付与
4. 付与しない

5. 結果

表2に、エンタメ小説3作品の各100発話文、合計300発話文に対する発話意図の付与結果を集計した表を示す。種類の計は3作品の重複を除いた数を()内に示す。

表2 小説作品の各発話文に対する発話意図の付与結果

| | 作品 | 発話 文数 | 種類 | 発話意図 (回数) |
|----------------------|------|----------|-------|--|
| 1 辞書の 発話意図 を付与 | 【結城】 | 15 | 12 | 説明-(6)、依頼-(3)、確認-(2)、肯定-(2)、感謝、命令 |
| | 【有川】 | 20 | 13 | 説明-(5)、依頼-(2)、確認(2)、禁止(2)、忠告(2)、命令(2) 他 |
| | 【三浦】 | 30 | 21 | よびかけ(4)、説明-(4)、命令(3)、確認-(3)、依頼-(2)、否定 |
| | 計 | 81 | (30) | (2) 他 |
| 2 新規の 発話意図 を付与 | 【結城】 | 63 | 32 | 説明-(20)、wh 質問-(7)、了解-(5)、聞き返し(5)、意見-(3)、助言(3)、yes-no 質問(2)、指摘(2)、申し出-説明(2)、予告(2) 他 |
| | 【有川】 | 43 | 23 | 意見-(12)、評価-(5)、主張-(4)、報告-(2)、説明-(2)、同意要求(2) 他 |
| | 【三浦】 | 53 | 21 | 説明-(13)、wh 質問-(9)、紹介(9)、挨拶(4)、評価(4)、yes-no 質問(3) 他 |
| 計 | 154 | (62) | | |
| 3 複数の 発話意図 を付与 | 【結城】 | 20 | 17 | 意見/不満(2)、説明-理由/助言(2)、了解/不本意(2)、wh 質問-何/反感、wh 質問/意見要求、wh 質問/説明要求 他 |
| | 【有川】 | 32 | 29 | 気遣い/嫌味(2)、評価/開き直り(2) 他 |
| | 【三浦】 | 15 | 12 | 説明-事情/言い訳(4)、wh 質問/怒り、wh 質問/非難 他 |
| 計 | 67 | (57) | | |
| 4 付与し ない | 【結城】 | 2 | | |
| | 【有川】 | 5 | | |
| | 【三浦】 | 2 | | |
| | 計 | 9 | | |
| 作品別 計 | 【結城】 | 100 | 62 | |
| | 【有川】 | 100 | 67 | |
| | 【三浦】 | 100 | 35 | |
| 合計 | | 300 | (151) | |

5.1 発話意図の階層化

発話意図付与では、(-)を使用した階層化ラベルを用いた。発話意図の階層化について説明する。(-)を用いた発話意図の階層化は、発話意図を細分化するために、発話文表現文型辞書の初期から採用していた方法である。(-)を挟んだ左右の関係は様々ある。以下に例を示す。

1. 説明-理由：理由を説明する。例：「等しい速度で運動してるから」【結城】
2. 依頼-非実行：ある行為をしないように依頼する。例：「まあ、そう言わないで」【有川】
3. wh 質問-場所：質問の対象が場所。例：「どこで野宿してたんだ？」【三浦】
4. 肯定-確認：確認に対して肯定の応答をする。例：「できるよね？」に対して「できるよ、もちろん」【結城】

このように階層化を採用した理由は、その発話意図を表現する文型を特定するために、意図を細分化して、意味を限定する必要があったからである。例えば、「説明」の下位分析には「理由」の他に、「結果」、「前提」、「事情」、「状況」、「物事」などがあり、それぞれ表現する文型は異なる。

5.2 辞書の発話意図を付与

表2のグループ1は、既に辞書に収録されている発話意図を付与した発話文である。対象となった発話文数と発話意図の種類、および、その内訳を示している。なお、内訳では、発話意図の最上位の階層が同一のものは、まとめて集計している。

5.3 新規の発話意図を付与

表2のグループ2は、辞書に収録されていない新しい発話意図を付与した発話文である。新しい発話意図は、合計で62種類となった。

5.4 複数の発話意図を付与

表2のグループ3は、複数の発話意図を (I) で併記して付与した発話文である。付与した発話意図は、既存の発話意図か新しい発話意図かは問わない。どのような場合に複数の発話意図を付与するかを以下に示す。

1. 語用論的な意図を併記
例：wh 質問/非難「なんなんだ、その穴は」【三浦】
例：wh 質問/説明要求「どゆこと？」【結城】
2. 感情を併記
例：yes-no 質問/警戒「俺を脅そうっての？」【三浦】
感情のラベルは、基本的に（Ackerman 2020）を参照している。
3. 発話意図がひとつに定まらない場合、全て併記
例：拒絶-説明/取り消し-依頼「やっぱり説明はいいや」【結城】

これらのうち、前者の2つは、発話文の文型を定めるために併記が不可欠である。

5.5 発話意図を付与しない

表2のグループ4は、発話意図を付与しなかった発話文の数を示している。どのような場合に付与しなかったかを以下に示す。

1. 発話文が倒置 例：「あたしだけよこんな腕立て食らってんの！」【有川】
2. 述語がない 例：「…ハイジさん、ここが？」【三浦】
3. 記号のみの文 例：「？」【結城】

6. 考察

6.1 3作品に付与した発話意図の特徴

対象とした3作品のそれぞれの物語の特徴と、その作品に特徴的な発話意図を表3に示す。

表3 物語の特徴と特徴的な発話意図

| 作品 | 物語の特徴 | 特徴的な発話意図と「発話文」 |
|------|------------------------------------|--|
| 【結城】 | 高校生の僕が中学生の従姉妹のユーリに数学を教える | <ul style="list-style-type: none"> ・説明-前提 「たとえば、直線の上を一つの点が動くとしよう」「つまり、点Pの位置xが1から4に変わったとする」 ・説明-手順 「ここまでで、点Pとその位置の話が終わり」「次は時刻だ」 ・意見/確信-はずだ 「『なーんだ、それだけのことなの?』って言うはずだよ」 ・意見/断定-言い切り 「ユーリにもちゃんとわかる」 |
| 【有川】 | 新人図書隊員の女の子が、厳しい教官や仲間たちと一緒に図書館の敵と戦う | <ul style="list-style-type: none"> ・指摘/怒り 「しかもコケたのゴールの後じゃないよっ!」 ・意見 「絞られてへこたれる程度なら辞めりゃいいんだ」「女子から採るとすれば彼女しかいないと思うけど」 |
| 【三浦】 | 箱根駅伝にチャレンジする大学生たちの話 | <ul style="list-style-type: none"> ・説明-理由/言い訳 「だって、これしかないんだもん」 ・非難-過失 「気をつけろって言っただろ」 ・説明-事情/言い訳 「気をつけてはいたんだ」 ・wh 質問/怒り 「ホントにあんた、なんなんですか?」 ・wh 質問/非難 「なんなんだ、その穴は」 |

【結城】に特徴的な発話意図は、「説明」と「意見」のバリエーションである。これらには、バリエーションがあり、適切な発話文を生成するためには、細分化が必要である。発話意図の階層化は、このための手段を提供する。

【有川】に特徴的な発話意図は、感情を伴う強めの「意見」交換である。感情は、発話文の文型に影響するため、感情の併記が不可欠である。

【三浦】に特徴的な発話意図は、「wh 質問」や語用論的な意図を伴うものである。「wh 質問」に伴う感情や語用論的意図も、併記が必要である。

6.2 発話意図の体系化に向けて

今回の調査により、不足している発話意図が62種類明らかになった。これらの多くは保留にしていたものが多いが、まったく想定していなかった発話意図もある。小説の発話文に実際に発話意図を付与してみることは、発話意図の体系の妥当性を評価するためには不可欠である。

それぞれの小説で付与した発話意図の分布には、大きな偏りがあった。発話意図の体系は、拡張可能とすることが望ましく、階層化が不可欠である。階層的なラベル体系を用いることにより、必要に応じて発話意図の細分化が可能となる。

発話文によっては、発話意図をひとつだけ付与することは現実的ではなく、複数の発話意図を併記することが必要となる。特に、語用論的意味や感情は、併記することにより、適切な文型を選択できるようになると考えられる。

7. おわりに

本稿では、『発話文表現文型辞書』の「発話意図」の抜本的な見直しのために実施した、小説の発話文への発話意図付与について述べた。今後は、さらに、対象とする発話文を拡大し、辞書の「発話意図」の改訂案の策定を行う予定である。

エンタメ小説の会話文への発話意図付与は、付与するラベルの体系が定まれば、それほど難しくなくという感触を得ている。考えられる理由2点を以下に示す。

1. 小説の会話文の発話意図は、話者の発話意図というよりも、話者の口を借りた小説の筆者の意図なので、ストーリー展開がわかりやすいエンタメ小説は、発話意図もわかりやすい。
2. 小説の会話文の発話意図は、音声や表情などのマルチモーダルに頼ることができないので、熟練の作家ほど、文字情報だけで誤解なく意図が伝わるように書いている。

以上の理由から、適切な発話意図が付与された小説発話文 DB が実現すれば、小説発話文の生成以外にも、日本語教育や会話生成にも、役に立つ言語資源になることが期待できる。

謝 辞

本研究は、JSPS 科研費 JP18H03285、JP21H03497 の助成を受けたものです。

文 献

- 夏目和子・刀山将大・佐藤理史 (2017) 「発話文自動生成のための日本語表現文型辞書の作成」言語資源活用ワークショップ 2016 表論文集, pp.126-136.
- 夏目和子・佐藤理史 (2019) 「発話文表現文型辞書の設計と編纂」言語資源活用ワークショップ 2019 発表論文集, pp.295-312.
- 夏目和子・佐藤理史 (2022) 「話し言葉生成のための発話文表現文型辞書」言語処理学会第 28 回年次大会発表論文集, pp.1612-1617.
- 刀山将大・夏目和子・佐藤理史他 (2017) 「話者の特徴を反映した発話文生成器の作成」言語処理学会第 23 回年次大会発表論文集, pp.28-31.
- 木村遼・夏目和子・佐藤理史・松崎拓也 (2018) 「発話表現文型辞書を利用した多様な発話文生成機構」第 32 回人工知能学会全国大会, 2E2-02.
- 米田智美・佐藤理史・夏目和子他 (2022) 「話者の性格を反映した発話文の生成」言語処理学会第 28 回年次大会発表論文集, pp.1744-1748.
- 佐藤理史 (2020) 「HaoriBricks3: 日本語文を合成するためのドメイン特化言語」自然言語処理, 27:2, pp.411-444.
- 木村遼 (2019) 「表現文型を用いた発話文の解析と生成」名古屋大学大学院工学研究科修士論文 (未公開)
- 国立国語研究所編 (1960, 1963) 『話しことばの文型』国立国語研究所.

- 益岡隆志・田窪行規 (1992) 『基礎日本語文法』 改訂版, くろしお出版.
- 日本語記述文法研究会 (2003) 『現代日本語文法』 4 モダリティ, くろしお出版.
- 日本語記述文法研究会 (2009) 『現代日本語文法』 7 談話・待遇表現, くろしお出版.
- 高梨信乃 (2010) 『評価のモダリティ: 現代日本語における記述的研究』 くろしお出版.
- 山岡政紀 (2008) 『発話機能論』 くろしお出版.
- 宮島達夫・仁田義雄編 (1995) 『日本語類義表現の文法』 上: 単文編, くろしお出版.
- グループ・ジャマシイ編著 (1998) 『教師と学習者のための日本語文型辞典』 くろしお出版.
- 市川保子 (2019) 『日本語類義表現と使い方のポイント: 表現意図から考える』 スリーエーネットワーク.
- 荒木雅弘・伊藤敏彦・熊谷智子・石橋雅人 (1999) 「発話単位タグ標準化案の作成」 人工知能誌, 14:2, pp. 251-260.
- 伝康晴 (2015) 「対話への情報付与」 小磯花絵 (編), 『講座日本語コーパス3: 話しことばコーパス-設計と構築-』, 朝倉書店, pp.101-130.
- 居關友里子・門田圭佑・伝康晴 (2018) 「『日本語日常会話コーパス』への談話行為アノテーションの試み: タグ選択が困難な事例に焦点を当てて」 言語資源活用ワークショップ 2018 発表論文集, pp.47-56.
- 入江友紀・松原茂樹・河口信夫・山口由紀子・稲垣康善 (2005) 「音声対話コーパスにおける発話意図タグの設計と評価」 電子情報通信学会論文誌 D, 88:10, pp. 2169-2173.
- 筒井佐代 (2012) 『雑談の構造分析』 くろしお出版.
- 結城浩 (2015) 『数学ガールの秘密のノート: 微分を追いかけて』 SB クリエイティブ.
- 有川浩 (2011) 『図書館戦争』 角川文庫.
- 三浦しをん (2009) 『風が強く吹いている』 新潮文庫.
- Angela Ackerman, Becca Puglisi・滝本杏奈 (2020) 『感情類語辞典』 増補改訂版, フィルムアート社.