

令和 6 年 5 月 27 日現在

機関番号：13901

研究種目：基盤研究(B) (一般)

研究期間：2021～2023

課題番号：21H03497

研究課題名(和文)話者の特徴や個性を感じさせる発話文の生成機構

研究課題名(英文)Automatic generation of utterances that fit speaker's personality

研究代表者

佐藤 理史 (Sato, Satoshi)

名古屋大学・工学研究科・教授

研究者番号：30205918

交付決定額(研究期間全体)：(直接経費) 12,300,000円

研究成果の概要(和文)：ある発話意図を表すのにどのような文末形式が可能かを記述した『発話文表現文型辞書』を作成するとともに、任意の発話をそのスタイル・口調の特徴を表現したベクトル(口調ベクトル)に変換する口調エンコーダを実現した。口調エンコーダを利用して特定の話者(キャラクタ)の少数のセリフからその話者の代表口調ベクトルを作成し、これを用いて複数の発話候補の中からその話者らしい発話を選択したり、小説の発話の話者を自動推定したりすることが可能となった。

研究成果の学術的意義や社会的意義

本研究の意義は、日本語発話のスタイル・口調をベクトル化することに成功し、それにより、生成と解析の両側面において、新たな技術を提供した点にある。発話生成という側面では、発話のスタイル・口調を制御する(特定の話者のスタイル・口調を模倣する)ことが可能となった。この技術は対話エージェントなどで個性を持ったエージェントの実装等に利用できる。解析という側面では、小説内の発話の話者推定に対して口調に基づく話者推定という新しい方法を提供した。

研究成果の概要(英文)：We developed "a Dictionary of Sentence Patterns of Japanese Utterance" that offers possible sentence-ending forms for expressing a particular speech intention. Additionally, we developed a speech-style encoder that converts any utterance into a vector (speech-style vector), in which the features of the speech style are embedded. Using the speech-style encoder, we can create a representative vector of a specific speaker (character) from a dozen utterance examples. Representative vectors enable the automatic selection of an utterance suitable for a speaker among utterance candidates and the automatic identification of the speaker of an utterance in a novel.

研究分野：自然言語処理

キーワード：自然言語処理 発話文 表現文型 口調エンコーダ 口調ベクトル 話者推定 発話文生成

1. 研究開始当初の背景

自然言語処理は、書き言葉を対象に発展してきたが、近年は、対話エージェントやロボットでの使用を想定した話し言葉の研究も盛んに行われてきている。たとえば、ロボットは、近い将来、我々の生活に入り込むことが期待されているが、「ロボットはどのような日本語を話すべきか」という問題は、ほとんど議論されていない。具体的には、

- 人間同士の会話(実際の会話は、非常に崩れた話し言葉で、会話に参加していない第三者には、意味不明な場合も多い)が理想なのか、それとも、もう少し整った話し言葉の方が良いのか
 - 没個性的な話し方がよいのか、それとも、話者の特徴や個性を感じさせる話し方がよいのか
- この問題が議論されない背景には、発話の形式・口調を制御する技術が未成熟という現状がある。実際、現在のほとんどのロボットでは、簡単なテンプレートをを用いた発話生成が用いられており、どのように話すかは、完全に組み込み(固定)となっている。

2. 研究の目的

「ロボットはどのような日本語を話すべきか」という問題を俎上に乗せるためには、柔軟な発話文生成機構の実現が不可欠と考え、この実現を本研究の目的とする。具体的には、以下の3点を目標として掲げる。

1. 同一の発話意図・発話内容を、色々な形式・口調で表出できる発話文生成機構を実現する。
2. 発話形式・口調を制御することにより、話者の人物像(特徴・個性)を感じさせる発話文の生成を実現する。
3. 上記の2点の実現のためのデータ作成に欠かせない、発話文解析機構を実現する。

話し言葉では、書き言葉と比較して、話者の違いがより強く、かつ直接的に言語形式に現れる。このわかりやすい例に、日本語の小説の会話がある。小説の作者は、登場人物の話し方・口調を書き分けることによって、「Aが言った」のように話者を直接示すことなく、話者が誰であるかを示すのが普通である。この「話し方・口調によって話者を暗示する」ことは、(英語の小説と比較して)日本語の小説において顕著である。すなわち、日本語は、話者の特徴・個性等を発話文に反映させやすい言語であり、金水敏が指摘した「役割語」の存在も、その証拠のひとつである。

本研究の中核的問いは、次の問いである。

- 話者の人物像(特徴・個性)を発話文に反映させることを、どうすれば機械的に実現できるか?

3. 研究の方法

本研究で取り組む内容は、大きく、(1) 表現文型辞書、(2) 発話文生成機構、(3) 発話文解析機構、に分けられる。

(1) 表現文型辞書: 表現文型辞書は、ある発話意図を表すのにどのような文末形式(表現文型、以下、文型と略記)が存在するか、および、それぞれの文型はどのような話し方特徴を持つか、という情報を提供する知識ベースである。研究スタート時にすでにある程度の蓄積があるが、それに加えて、以下の4点について検討・拡充する。文型のグループ化、話し方特徴の再検討、文型の記述形式の規格化によるツールとの自動連携の実現。

(2) 発話文生成機構: これまでに実現した文合成ツール(HaoriBricks3)を使うと、発話内容に文型を接続して表層文を生成することができる。発話文生成機構における研究課題は、話者の性質・性格と話し方特徴(あるいは、文型)をどう結びつけるかである。この課題を解決するために、以下の3点について検討する。話者のモデル化、話者データベースの作成、話し方特徴推定または文型選択の実現。

(3) 発話文解析機構: 実際の小説の発話文を解析し、使われている表現文型を自動的に同定することができれば、表現文型辞書の文型の拡充と、話者データベースの拡充が容易になる。これを実現するために、以下の2点に取り組む。文末述語解析器 Panzer の改良、文型の自動同定。

4. 研究成果

(1) 発話文表現文型辞書を用いた多様な形式の発話文合成の実現

8055 件のエントリ(文型)を 60 種類の発話意図に分類して整理した『発話文表現文型辞書』を作成した。この辞書の各エントリには、発話意図、表現文型、例文、話し方特徴ベクトル(但し、43 意図 1,444 エントリに対してのみ)が記述されている。辞書の一部を表 1 に示す。

表現文型は、接続形式(どのような活用型のどんな標識をもつ述語に接続するか)、主要部、

末尾部（終助詞等）によって記述されている。これらの要素は文生成ツール HaoriBricks3 と連携しており、表現文型と述語を指定すると、その表層形式を自動生成することができる。

表 1 『発話文表現文型辞書』の一部

発話意図	連番	接続形式	表現文型		話し方特徴ベクトル								
			主要部	末尾部	ジェンダー	年齢	長さ	強さ	待遇表現		印象		
									敬意	距離	場面差	品格	
R05	1	Vbnt	方がいい		10	00	01	00	01	00	00	00	0000
R05	2	Vbnt	方がいい	よ	00	00	00	00	00	00	00	00	0000
R05	3	Vbnt	方がいい	ね	00	00	00	00	00	01	00	00	0000
R05	4	Vbnt	方がいい	わ	01	00	00	00	00	01	00	00	0000
R05	5	Vbnt	方がいいと思う		00	00	10	01	00	00	00	00	0000
R05	6	Vbnt	方がいいんじゃない	かな	00	10	20	01	00	00	00	00	0000
R05	7	Vbnt	方がいいです	よ	00	10	10	00	00	10	00	10	0000
R05	8	Vbn	ことだ		10	10	00	00	01	00	00	00	0000
R05	9	Vbn	ことだ	ね	10	10	10	00	01	01	00	00	0000
R05	10	Vbn	こと	ね	01	10	00	00	01	01	00	00	0000
R05	11	Vbn	こと		00	10	01	00	01	00	00	00	0000
R05	12	Vbn	ことです		00	10	00	00	00	00	10	10	0000
R05	13	Vb	命令形	よ	10	00	01	00	00	00	01	01	0000
R05	14	Vb	なさい		00	10	01	00	00	00	00	00	0000
R05	15	Vb	なさい	よ	01	10	00	00	00	00	00	00	0000
R05	16	Vb	お～なさい		00	10	01	00	01	00	00	10	0000
R05	17	Vb	お～なさい	よ	01	10	00	00	01	00	00	10	0000
R05	18	Vb	お～なさい	な	01	10	00	00	01	00	00	10	0001
R05	19	Vn	テは	ね	00	10	00	00	00	01	00	00	0000
R05	20	Vn	チャ形	ね	01	00	00	00	00	01	01	00	0000

話し方特徴ベクトルは、その文型が話し方としてどんな特徴を持っているかの情報を表現したもので、ジェンダー、年齢、長さ、強さ、待遇表現、印象の6軸で構成されている。待遇表現は、敬意の有無、距離、場面差、品格の4つに細分され、印象は、明るい、暗い、可愛い、古風の4種類に細分されており、全体で20次元のベクトル表現となっている。これらの特徴を陽に指定することにより、多様な形式の発話文を自動合成することが可能である。

(2) 小説発話文データベースの整備

実際の小説14作品から登場人物のセリフ（発話）を抜き出し、そのそれぞれに話者の情報を付与した発話文データベースを整備した。総発話数は27,788件、総文数は46,158文である。

(3) 口調エンコーダの実現と口調ベクトルを利用した話者推定の実現

任意のセリフを、そのスタイル・口調を表す64次元のベクトル（口調ベクトル）に変換する口調エンコーダを提案・実現した。口調エンコーダは、文エンコーダ（BERTまたはLUKE）に口調タイプ分類器を接続することによって実現されている。口調エンコーダの構成を図1に示す。

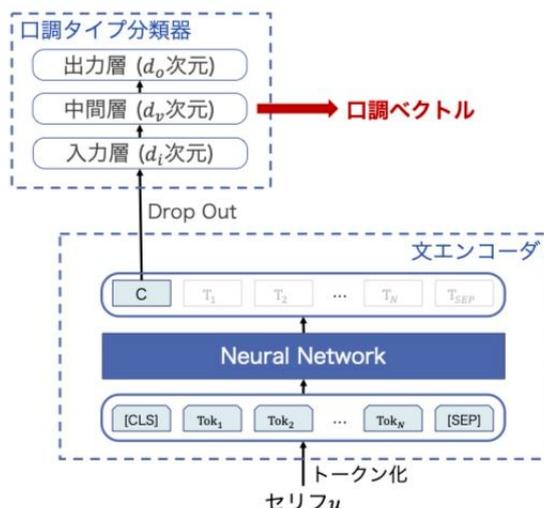


図1 口調エンコーダの構成

この口調エンコーダを訓練するために、10種類の口調タイプを設定し、7作品の登場人物50名をそれぞれ口調タイプを割り振った学習用データ(総セリフ数16,705件、総文字数480,999字)を作成した。この学習用データを用いて、文エンコーダと口調タイプ分類器のパラメータを同時学習する。口調エンコーダとして利用する場合は、入力したセリフに対して、口調タイプ分類器の中間層のベクトルを、口調ベクトルとして採用する。

口調ベクトルを利用すると、次の方法で、セリフの話者を推定することができる。

[Step1] 話者候補のそれぞれに対して、セリフの具体例を集め、それらを口調ベクトルに変換してその平均を求める。これを代表口調ベクトルと呼ぶ。

[Step2] 話者を推定したいセリフを口調ベクトルに変換する。

[Step3] 話者候補の代表口調ベクトルのうち、[Step2]の口調ベクトルに最もよく似ている代表口調ベクトルを求め、その話者を推定話者とする。

登場人物はそれぞれの小説で異なるため、登場人物のセリフの実例をあらかじめ大量に集めておくことはできない。上記の方法では、小説に横断的に現れるセリフの口調・スタイルの特徴を埋め込んだベクトルを生成する口調エンコーダを、あらかじめ他の小説のセリフデータから学習しておき、セリフの口調・スタイルの特徴に基づいて話者を推定する。実際の小説のセリフを用いた実験において、話者推定に必要なセリフ数は、10例程度であることが判明している。

この方法を組み込んだ小説のセリフの話者推定システムを実装した。まず、地の文の手がかり等を利用して、それぞれのセリフに対して話者候補を求める。この段階で話者が確定した(話者候補が1名となった)セリフを、代表口調ベクトルの計算に用いる。話者候補が複数となったセリフは、上記の方法により話者を決定する。ライト文芸作品5作品に対して実際にこの話者推定法を適用し、4作品でベースラインを上回る推定結果を得た。

口調エンコーダには多くの改良の余地がある。口調エンコーダの性能比較を可能とするために、19口調タイプに対してそれぞれ45セリフを収録した評価用データセットを構築した。口調エンコーダで利用する文エンコーダは、日本語ウィキペディアのテキスト(書き言葉)で事前学習されたモデルであるため、実際の小説のセリフ(話し言葉)を用いた追加学習に効果があると考え、実際にこれを試みたところ、文エンコーダとしてBERTを用いる場合は大きな効果があり、LUKEを用いる場合はほとんど効果がないことが判明した。最終的に、2つの口調タイプを弁別する性能として、最高で精度0.899が得られた。

なお、話者推定の前段として必要な「小説の登場人物検出」についても研究を行い、登場人物を指し示す固有表現にタグを挿入するという、新しいタスク設定を提案した。

(4) 発話者のキャラクター性を反映した発話の自動合成

本研究の最終目標である「発話者のキャラクター性を反映した発話の自動合成」の実現に対しては、2種類の方法を試みた。

最初に試みた方法は、キャラクターの性格とそのキャラクターが用いる文型リストの組を与えてキャラクターのモデルを作成し、『発話文表現文型辞書』の特徴ベクトルを利用して文型を選択する方法である。この方法の問題点は、キャラクターの性格を記述するのが難しいこと、および、そのキャラクターが用いる文型リストを選定することが難しいことである。そのため、話者としていキャラクターを熟知していないと、キャラクターモデルの作成に必要なデータの作成ができない。

この問題を回避するために、口調エンコーダを利用する第2の方法を考案した。この方法では、話者としていキャラクターのセリフの実例を10例程度集めることができれば、キャラクターのモデル(代表口調ベクトル)を作成することができる。発話の合成では、セリフ候補を複数作成し、その中からそのキャラクターが言いそうなセリフを、代表口調ベクトルの類似度を用いて選択する。

表2 発話の自動合成例

入力	前文脈=ここを, 述語=片付ける, 接続=Vb, 発話意図=Q06(許可要求)		
話者	順位	発話	距離 ² 確率
笠原郁	1	ここを片付けていいかしら	13.27 0.620
	2	ここを片付けてもいいかしら	14.96 0.279
	3	ここを片付けさせてほしいの	17.72 0.098
柴崎麻子	1	ここを片付けさせてほしいの	13.45 0.973
	2	ここを片付けていいかしら	29.41 0.012
	3	ここを片付けてもいいかしら	32.21 0.008
堂上篤	1	ここを片付けさせてもらえんか	8.03 0.993
	2	ここを片付けさせてくれよ	16.54 0.005
	3	ここを片付けていいか	20.70 0.001
小牧幹久	1	ここを片付けていいかい	13.11 0.586
	2	ここを片付けてもいいかい	14.83 0.257
	3	ここを片付けていいかな	18.40 0.068

発話の自動合成の例を表2に示す。この例では、「ここを片付ける」ことの許可を求める発話を生成している。キャラクターとしては、小説『図書館戦争』(有川浩著)に登場する4名を用いている。笠原郁と柴崎麻子は女性であり、堂上篤と小牧幹久は男性である。女性キャラクターに対しては女性らしい発話が、男性キャラクターに対しては男性らしい発話が選ばれており、かつ、同性のキャラクターに対しても異なる文末形式の発話が選ばれていることがわかる。

今回の実装では、セリフ候補の作成に『発話文表現文型辞書』を利用する方式を採用したが、これ以外の方式(たとえば、大規模言語モデルを利用した候補作成)に置き換えることもできる。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 佐藤理史	4. 巻 22
2. 論文標題 多様な形式の発話文合成を実現する辞書とツール	5. 発行年 2022年
3. 雑誌名 日本語文法	6. 最初と最後の頁 153-168
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 石川和樹, 小川浩平, 佐藤理史	4. 巻 31
2. 論文標題 口調エンコーダを用いた小説発話の話者推定	5. 発行年 2024年
3. 雑誌名 自然言語処理	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計9件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 石川 和樹, 宮田 玲, 小川 浩平, 佐藤 理史
2. 発表標題 口調ベクトルを用いた小説発話の話者推定
3. 学会等名 情報処理学会自然言語処理研究会
4. 発表年 2022年

1. 発表者名 夏目 和子, 佐藤 理史
2. 発表標題 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ー
3. 学会等名 言語資源ワークショップ2022
4. 発表年 2022年

1. 発表者名 石川和樹, 佐藤理史, 宮田玲, 小川浩平
2. 発表標題 複数の手がかりを利用した小説発話の話者推定
3. 学会等名 言語処理学会第29回年次大会
4. 発表年 2023年

1. 発表者名 夏目和子, 佐藤理史
2. 発表標題 話し言葉生成のための発話文表現文型辞書
3. 学会等名 言語処理学会第28回年次大会発表論文集
4. 発表年 2022年

1. 発表者名 米田智美, 佐藤理史, 夏目和子, 宮田玲, 小川浩平
2. 発表標題 話者の性格を反映した発話文の生成
3. 学会等名 言語処理学会第28回年次大会発表論文集
4. 発表年 2022年

1. 発表者名 大島一海, 小川浩平, 佐藤理史
2. 発表標題 超短編小説のフリオチパターンの分類とオチの自動生成
3. 学会等名 FIT2023(第22回情報科学技術フォーラム)
4. 発表年 2023年

1. 発表者名 川北雄大, 石川和樹, 夏目和子, 小川浩平, 佐藤理史
2. 発表標題 口調弁別評価データセットの作成と口調エンコーダの評価
3. 学会等名 情報処理学会自然言語処理研究会
4. 発表年 2024年

1. 発表者名 大島一海, 窪田智徳, 小川浩平, 佐藤理史
2. 発表標題 固有表現を対象とした小説登場人物検出
3. 学会等名 言語処理学会第30回年次大会発表論文集
4. 発表年 2024年

1. 発表者名 夏目和子, 佐藤理史
2. 発表標題 小説発話への発話意図アノテーションのための末尾部分分析の試み
3. 学会等名 言語処理学会第30回年次大会発表論文集
4. 発表年 2024年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------