

# 文節機能部の確率的書き換えによるキャラクタ性変換

宮崎 千明<sup>1</sup> 平野 徹<sup>1</sup> 東中 竜一郎<sup>1</sup> 牧野 俊朗<sup>1</sup> 松尾 義博<sup>1</sup> 佐藤 理史<sup>2</sup>

<sup>1</sup> 日本電信電話株式会社 NTT メディアインテリジェンス研究所

<sup>2</sup> 名古屋大学大学院工学研究科

{miyazaki.chiaki, hirano.tohru, higashinaka.ryuichiro, makino.toshiro, matsuo.yoshihiro}@lab.ntt.co.jp, ssato@nuee.nagoya-u.ac.jp

## 1 はじめに

近年, NTT ドコモのしゃべってコンシェルや Apple の Siri などの, ユーザの発話音声を認識し, ユーザの意図に応じて適切な応答を返す音声対話エージェントが普及してきている. 音声対話エージェントの会話相手としての価値や娯楽性を高めるためには, エージェントに対して何らかの人物像(キャラクタ設定)を与え, より人間らしく親しみやすい存在にすることが重要だと考えられるが, 人手による発話データの作成は大変コストが高い. そこで我々は, 特定のキャラクタらしさを持った発話の自動生成に取り組み, なるべく人手をかけずに, 大量に, 目的のキャラクタに適した発話データを作成できる仕組みの実現を目指している.

キャラクタらしい発話を生成する上で, どのような語彙を選択するかが重要であることは言うまでもないが, その使用頻度にもキャラクタらしさが表れるという知見 [6] がある. 我々はこれに基づき, 言語表現の出現確率の違いによってキャラクタらしさの違いを表現することを考えた. 本稿では, その実現方法として, 発話を構成する各文節の機能部を目的のキャラクタに適した確率で適した表現に書き換える手法を提案するとともに, この手法がどの程度のキャラクタらしさを表現できるものとなっているかを示す.

加えて, 言語表現の書き換えの代償として生じる, 日本語として違和感のある出力を極力減らすことも目標の一つとし, 一定の文法性を担保可能な書き換え規則の構成を提案する.

## 2 関連研究

発話にキャラクタらしさを与えることを目的とした研究は多くはないものの, 話者の性格 [1] や, 印象(まじめさ, かわいさ/かっこ良さ, 感情豊か) [3] などを対象とした研究がなされている. 中でも, 本研究と最も目的に近いのは, 発話の人物属性(年齢, 性別, 居

住地)らしさの変換 [5] や, 特定の話者らしさ(話者の個人性)の変換である [4]. ただし, 文献 [5, 4] の手法は, 目的キャラクタらしい言語表現として, 何らかの条件下で最良の候補を決定的に選択するもので, 確率的な選択でキャラクタらしさの違いを表現するという思想は含まない.

## 3 文節機能部の書き換え手法

本稿で提案する書き換え手法は, 人間が発話にキャラクタらしさを与えようとする際に行う書き換え操作をルール化し, それを模倣するものである. 図 1 に書き換え処理の流れを示す.

入力文を受理したら, まず, 文節機能部の認定に必要な文節・形態素情報の取得 (1) を行う. ここで, 文節機能部は各文節の主辞より後の機能語列を指す. 文節・形態素情報は, 係り受け解析器 CaboCha [2] (形態素解析には MeCab・IPA 辞書を利用) の解析結果を用いて取得する.

次に, 事前に準備しておいた書き換え規則集を参照して書き換え先候補の列挙 (2) を行い, 列挙された候補集合の中から書き換え先候補の選択 (3) を行い, 最後に各文節について選択された候補への置換 (4) を行い, 全文節を結合して出力する.

以降, 本節では (2) および (3) の具体的な処理内容と, キャラクタらしい発話のコーパスを用いて, 人間の書き換え操作を模倣するためのルール集合(書き換え規則集)を作成する方法について述べる.

### 3.1 キャラクタコーパス

本研究では, 人間の書き換え操作をルール化するために, 特定のキャラクタらしさを強調するように人手で書き換えられた発話(以降, キャラクタ文と呼ぶ)からなるコーパス [6] を利用する. 対象キャラクタは, 性別, 年齢, 会話相手との親密度という 3 属性で構成される 10 種類である. 性別は男性 (M) / 女性 (F) の 2

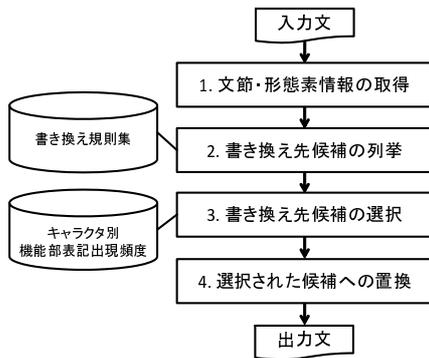


図 1: 書き換え処理の流れ

表 1: キャラクター一覧と書き換え結果の例 (元文:「駅の傍に美味しい寿司屋があります」)

ID	属性値	書き換え結果
F10p	女性・10歳・親密	駅の近くに美味しいお寿司屋さんがあるよ。
M10p	男性・10歳・親密	駅の近くに美味しいお寿司屋さんがあるよ
F25p	女性・25歳・親密	駅の近くに美味しい寿司屋があるよ。
F25b	女性・25歳・非親密	駅の傍に美味しいお寿司屋さんがあるんですよ。
M25p	男性・25歳・親密	駅の傍に美味しい寿司屋があるよ。
M25b	男性・25歳・非親密	駅の傍に美味しいお寿司屋があります。
F50p	女性・50歳・親密	駅の近くに美味しいお寿司屋さんがあるよ。
M50p	男性・50歳・親密	駅の傍に美味しい寿司屋があるよ。
F75p	女性・75歳・親密	駅の傍に美味しいお寿司屋さんがあるのよ。
M75p	男性・75歳・親密	駅の傍にうまい寿司屋があるぞ。

種、年齢は10歳(子供)、25歳(若年成人)、50歳(中年成人)、75歳(老年成人)の4種、親密度は親密(p:会話相手は友人)/非親密(b:会話相手は職場上司)の2種である(親密度の異なるキャラクタは、相手によって話し方を変えることが想定しやすい男女25歳のみ)。表1にキャラクタの一覧と各キャラクタの書き換え結果例を示す。

### 3.2 書き換え規則集の作成

キャラクタコーパスの元文と各キャラクタのキャラクタ文の対から、人間によって施された書き換え操作を抽出し、規則として成形する。書き換え規則には、ある表記から別の表記への書き換えが、どのような生起環境において許容されるのかが記述されている。全10キャラクタ分の書き換え規則の集合を書き換え規則集と呼ぶ。

まず、動的計画法を用いて元文とキャラクタ文に含まれる文節の対応付けを行う。次に、対応付けられた文節のうち、文節の主辞品詞と係り種別(詳細は後述)

表 2: 書き換え規則の抜粋

係り種別	主辞品詞	書き換え元表記	書き換え先表記
文末(平叙)	名詞	だ	だなぁ
文末(疑問)	名詞	かな?	かしら?
連体	名詞	という	っていう
連用	名詞	かと	だなど

が同一で機能部表記だけが異なるものを対象とし、下記4つの要素を抽出する。

- 書き換え元文節の機能部表記
- 書き換え先文節の機能部表記
- 文節の主辞品詞
- 文節の係り種別

文節の主辞品詞は、名詞、動詞、形容詞、形容動詞(名詞、形容動詞語幹)の4種類とする。係り種別は、連体修飾、連用修飾、文末(平叙)、文末(疑問)の4種類からなる。名詞を主辞とする文節に係る文節は連体修飾、動詞・形容詞・形容動詞を主辞とする文節に係る場合は連用修飾と考える。文節が文末に位置する場合は、疑問の意味を持つ形態素を含むか否かによって異なる種別とする。

本研究では、キャラクタコーパスの12,140文(1,214文×10キャラクタ)から抽出した7,517の規則からなる規則集を利用する。表2に抜粋した規則を示す。

### 3.3 書き換え先候補の列挙

任意の入力文に書き換えを施す際には、文中の各文節について、書き換え規則集の中から適用可能な規則を探し、書き換え先表記を取得・列挙する。このとき、処理中の文節と各規則との間で書き換え元表記、および、文節の主辞品詞・係り種別が一致するか否かを照合することで規則の適用可否を判定する。例えば、処理中の文節が「思うな」(機能部表記は「な」、主辞品詞は動詞、係り種別は「文末(平叙)」)である場合、書き換え先候補は表3のようになる(「書き換え先表記」が書き換え先候補となる)。

### 3.4 書き換え先候補の選択

列挙された複数の書き換え先候補の中から、最終的に出力する候補を1つ選択する。ここで、言語表現の使用頻度の違いにキャラクタらしさが表れるという知見[6]に基づき、各書き換え先候補が出力される頻度を調整することでキャラクタらしさを表現するのが本手法の特徴である。

ここで、文献[6]で示されている、女性らしさを印象付ける終助詞「わ」の使用頻度が女性キャラクタの年齢設定が上がるにつれて高くなるという現象を例に

表 3: 書き換え先候補列挙の例

書き換え規則				書き換え先表記のキャラクタコーパスにおける出現頻度										
係り種別	主辞品詞	書き換え元表記	書き換え先表記	F10p	M10p	F25p	F25b	M25p	M25b	F50p	M50p	F75p	M75p	
文末 (平叙)	動詞	な	な	12	0	3	0	0	0	0	0	0	3	
			...											
			よ	15	37	14	0	7	0	8	23	0	0	
			よね	10	11	8	0	7	0	2	9	0	0	
			わ	0	0	3	0	0	0	27	0	21	0	
			わね	0	0	0	0	0	0	0	0	16	0	
			計	55	59	39	54	21	68	51	64	38	48	

とり、これが本手法による書き換えにどう反映されるかを説明する。表3には、各々の書き換え先について、キャラクタ別コーパスにおける出現頻度が併記してある。本手法では、例えば、主辞が動詞の文末文節（平叙文）に出現する「な」を書き換える場合において、終助詞「わ」が選択される確率は、25歳（F25p）の場合は8%（3/39）しかないのに対し、50歳（F50p）では53%（27/51）にまで増える。75歳（F75p）では、「わね」も含め、「わ」を含む表現が選択される確率がさらに高くなる。

ただし上記の方法では、どの候補についても目的キャラクタのコーパスにおける出現頻度が0であった場合に、全候補が棄却され、目的キャラクタには不適切だとしても元文の表記がそのまま出力されてしまう。これを避けるため、類似した別のキャラクタのコーパスにおける出現頻度を利用して、書き換え規則に登録されている全ての機能部表記について、出現頻度の補正を行う。

機能部表記  $w$  の目的キャラクタ  $i$  のコーパスにおける出現頻度  $f_i(w)$  が2以上の場合は、頻度にディスカウント係数  $\lambda$  をかけたものを補正頻度  $r_i(w)$  とし、頻度が1以下の機能部表記には、コーパスに出現した機能部表記の頻度合計  $N$  の  $1-\lambda$  分を分配する（頻度1のものはノイズが多いため頻度0とみなす）。頻度の配分は、キャラクタ  $j$ （本稿では全10種類）の頻度に、目的キャラクタ  $i$  に借用する際の重み  $g_{i,j}$  をかけた値の合計  $c_i(w)$  の、出現頻度が1以下の機能部表記集合  $C$  の全要素についての合計における割合によって決定する。なお本稿では、重み  $g_{i,j}$  は人手で指定した値を使用することとする。

$$r_i(w) = \begin{cases} \lambda f_i(w) & (f_i(w) \geq 2 \text{ の場合}) \\ (1-\lambda)N \frac{c_i(w)}{\sum C} & (f_i(w) \leq 1 \text{ の場合}) \end{cases} \quad (1)$$

$$c_i(w) = \sum_{j=1}^m g_{i,j} f_j(w) \quad (2)$$

表 4: 文法性評価の5段階評価基準

評価	評価基準
5	全く違和感はない
4	違和感はないが、あまり一般的でない表現を含む
3	わずかに違和感があるが、特定の状況における会話文と考えれば許容できる
2	明らかに違和感があるが、おおよその文意は理解できる
1	非常に違和感があり、文意が理解できない

#### 4 評価実験

提案手法によって書き換えを施した文について、文法性（日本語としての違和感が生じていないか）、キャラクタ性（目的のキャラクタらしさが表れているか）という2つの観点での主観評価を行った。評価は著者を含まない2名の日本語母語話者（ともに関東地方在住30代女性）が行った。書き換え規則集の取得および機能部表記の出現確率の算出に用いたデータ（学習データ）は3.1節で示したキャラクタコーパスのうち12,140文（1,214文×10キャラクタ）であり、取得した書き換え規則の数は7,517であった。

##### 4.1 実験1：文法性評価

提案手法によって書き換えを施した文について、日本語として違和感がないかを5段階（1：非常に違和感がある～5：全く違和感はない）で評価した。表4に5段階それぞれについての評価基準を示す。5段階のうち、3以上が日本語として許容可能な文であることを示し、2以下が日本語として明らかに違和感のある文であることを示す基準となっている。評価に使用した文は、キャラクタコーパスの元文と同一の発話データベースからランダムに抽出した100文を提案手法によって書き換えたものである。

表5に女性キャラクタに対する評価の結果を示す。2点（明らかに違和感がある）以下の書き換え結果は、評価者Aで11%、評価者Bでは8%であり、提案手

表 5: 文法性判定結果

評価者	評価値とその評価を得た文の数 (%)				
	1	2	3	4	5
A	13(3%)	39(8%)	78(16%)	161(32%)	209(42%)
B	4(1%)	34(7%)	53(11%)	50(10%)	356(72%)

(A)	(B)
1. しかし、最初に食べようと思った人はすごいよね 2. 社会人であれば家賃相当を家に入れての方がいいと思うんだ …… 8. 金額など法律で規制されていることって知ってたっけ？ 9. 餌代って結構掛かるわ 10. 残暑が優しいって言い方ってさ、聞かないんだ	1. しかし、最初に食べようと思う人ってすごいよね 2. 社会人であれば家賃相当を家に入れての方がいいと思うよ …… 8. 金額など法律で規制されていることは知ってた？ 9. 餌代も結構掛かるわ 10. 残暑が優しいって言い方は、聞かないわね

図 2: キャラクタ性評価データの例

法によれば、書き換えを施しても、日本語として違和感のある文は 10%程度に抑えられることを確認できた。違和感の原因には、内容語と機能語の意味的不整合(例: 存在動詞「ある」+ 継続「ている」)、態の変更(例: 「滝に打たれた」「滝に打った」)、提題化の誤り(例: 「200 点に落ちる」「200 点って落ちる」)等があった。これらを解消するには、語の意味や文節間の関係に踏み込む必要があると考えられる。

#### 4.2 実験 2: キャラクタ性評価

提案手法で書き換えを施した文について、目的のキャラクタらしさが表れているかを評価するために、書き換えられた文に対する二者択一式のキャラクタ判別実験を実施した。

実験では、特定の属性のみが異なる 2 つのキャラクタらしく書き換えられた文集合を評価者に提示し、属性値を判別させることを行った。例えば、図 2 に示すような 2 つの文集合 A, B を提示し、A と B のどちらが 25 歳で、どちらが 50 歳かを判別させる。評価に使用した文集合は 360 個で、内訳としては、男性対女性が 80 個(親密度  $p$  の 4 キャラクタペア  $\times$  20 個)、その他の属性の比較では 40 個(2 キャラクタペア  $\times$  20 個)ずつとなっている。各々の文集合は、キャラクタコーパスの元文と同一の発話データベースからランダムに抽出した、文法性評価で使用した評価データに含まれない 10 文で構成される。性能の比較対象として、人手で書き換えられた文(学習データとして使用した文で、内容語の書き換えも含む)を用いた判別実験も行った。

表 6 に評価結果を示す。正解率は評価者 A の判別結果が正解と一致した文集合の割合で、一致率は評価者 A と B の判定が一致した文集合の割合、 $\kappa$  値は 2 名の評価の偶然によらない一致を表す値である。提案手法で書き換えられた文集合は、平均して 0.87 の正解率で

表 6: キャラクタ性評価結果(正解率は評価者 A の値)

キャラクタ	提案手法			人手書き換え
	正解率	一致率	$\kappa$ 値	正解率
男性対女性	0.83	0.89	0.77	0.76
10 歳対 25 歳	0.73	0.70	0.40	0.78
10 歳対 50 歳	0.75	0.63	0.24	0.88
10 歳対 75 歳	1.00	0.98	0.95	1.00
25 歳対 50 歳	0.83	0.68	0.35	0.83
25 歳対 75 歳	0.95	0.98	0.95	0.98
50 歳対 75 歳	0.90	0.68	0.35	1.00
親密対非親密	1.00	1.00	1.00	1.00
平均	0.87	0.81	0.63	0.90

正しく属性値が判別された(評価者 B の場合は 0.80)。これは、人手で書き換えられた文集合を対象とした場合(0.90)と同等の結果であり、提案手法がキャラクタらしさの表現手段のひとつとして有望であることを示唆するものである。 $\kappa$  値は 0.63 であり、評価者 2 名の判別はかなりの一致(0.61~0.80)を示した。ただし、この値は性別・親密度・75 歳の判別に押し上げられたもので、その他の比較は低い一致(0~0.40)に留まっている。属性値によっては、評価者次第でキャラクタらしさの感じ取られ方が異なるという結果となったため、改善に向けて原因の分析が必要である。

#### 5 まとめ

本稿では、音声対話エージェントの発話への自動的なキャラクタ付けを目的として、発話を構成する各文節の機能部を目的のキャラクタに適した確率で適した表現に書き換える手法を提案した。さらに、提案手法は日本語として違和感のある出力を 10%程度に抑えつつ、人間が 80%以上の正解率で判別可能な程度のキャラクタらしさを付与可能であることを明らかにした。

#### 参考文献

- [1] François Mairesse and Marilyn Walker. PERSONAGE: Personality generation for dialogue. In *Proc.ACL*, pp. 496–503, 2007.
- [2] 工藤拓, 松本裕治. チャンキングの段階適用による日本語係り受け解析. 情報処理学会論文誌, Vol. 43, No. 6, pp. 1834-1842, 2002.
- [3] 沈睿, 菊池英明, 太田克己, 三田村健. 音声生成を前提としたテキストレベルでのキャラクタ付与. 情報処理学会論文誌, Vol. 53, No. 4, pp. 1269–1276, 2012.
- [4] 水上雅博, Graham Neubig, Sakriani Sakti, 戸田智基, 中村哲. 特徴的話者を対象とした言語的個人性変換. 情報処理学会研究報告. 自然言語処理研究会報告, Vol. 20, pp. 1–7, 2014.
- [5] 宮崎千明, 平野徹, 東中竜一郎, 牧野俊朗, 松尾義博. 発話にキャラクタ性を与えるための文末表現の変換. 人工知能学会研究会資料. 言語・音声理解と対話処理研究会, Vol. 68, pp. 41–46, 2013.
- [6] 宮崎千明, 平野徹, 東中竜一郎, 牧野俊朗, 松尾義博, 佐藤理史. 話者のキャラクタ性に寄与する言語表現の基礎的分析. 言語処理学会第 20 回年次大会発表論文集, pp. 232–235, 2014.