文節機能部の確率的書き換えによる言語表現のキャラクター性変換

Probabilistic Conversion of Functional Expressions for Characterization of Dialogue System Utterances

宮崎 千明

Chiaki Miyazaki

NTT メディアインテリジェンス研究所

NTT Media Intelligence Laboratories

miyazaki.chiaki@lab.ntt.co.jp

平野 徹

Toru Hirano

(同 上)

hirano.tohru@lab.ntt.co.jp

東中 竜一郎

Ryuichiro Higashinaka

(同 上)

higashinaka.ryuichiro@lab.ntt.co.jp

牧野 俊朗

Toshiro Makino

(同 上)

makino.toshiro@lab.ntt.co.jp

松尾 義博

Yoshihiro Matsuo

(同 上)

matsuo.yoshihiro@lab.ntt.co.jp

佐藤 理史

Satoshi Sato

名古屋大学大学院工学研究科

Graduate School of Engineering, Nagoya University

ssato@nuee.nagoya-u.ac.jp

keywords: characterization, functional expressions , linguistic expression conversion

Summary

Characterizing dialogue system utterances is important in making human-computer interaction systems more friendly and human-like. A method is described for achieving this by converting functional expressions according to their generation probabilities, which are calculated for specific characters. Experimental results show that the method can add characteristics of the target profiles (i.e., gender, age and closeness with a conversation partner) to dialogue system utterances and in so doing can generate a large variety of linguistic expressions.

1. は じ め に

ユーザの意図に応じて適切な応答を返すことのできる音声対話エージェントは、検索等のための便利なツールとしてだけではなく、コミュニケーションのためのツールとしても期待され始めている。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(キャラクター設定)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。しかし、様々なキャラクターに対応した発話データを全て人手で作成するのはコストが大変高い。よって我々は、特定のキャラクター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、目的のキャラクター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。ここで、本論文において「キャラクター」は「人物像」を指し、「人物像」の構成要素としては「年齢、性別、職業、階層、時代、容姿・風貌、性格等」[金水 03]を想定している。

話者のキャラクター性が伝わる発話の生成は、目的のキャラクターが偏って多く使用する(つまり、キャラクターの特徴が強く表れた)言語表現を抽出し、これを集中的に使用することで実現できる[宮崎 13]。しかしこのような方法では、常に同一の(もしくは類似した)言語表現が生成されることになるため、対話エージェントに適用する場合のように、複数の発話を連続的に生成する場合には、全体として「くどい」印象を与えてしまうという問題がある。さらに、対話エージェントのコミュニケーションをより自然なものにするには、固定的な発話テンプレートを常に使い続けるよりも、バリエーションの豊かな発話ができるほうが望ましいとの主張もある[Higashinaka 05]。

そこで本論文では、対話エージェントの発話に対してキャラクター性を付与し、かつ、バリエーション豊かな言語表現を生成する手法を提案する。具体的には、発話を構成する各文節の機能部を目的のキャラクターに適した確率で適した表現に書き換える。そして、実験により、一

定のキャラクター性がユーザに伝わり、多様な言語表現を含んだ発話が生成できることを示す。また、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どの程度に抑えられるかという側面からも評価を行う。

なお、我々の最終目標は音声対話エージェントを実現することであるが、本論文では、言語表現（テキスト）のみを対象とした技術について述べる。技術の適用先としては、テキストによるチャット対話のアプリケーションを想定している。

2. 関連研究

特定の人物像を想起させる特定の言葉づかいは「役割語」[金水 03] と呼ばれ、「女性語」「男性語」「老人語」から「キャラ語尾」に至るまで、言語的・文化的側面からの研究が行われている。

工学的な側面においては、本研究と同様に、特定の人物像（キャラクタ）に適した発話を生成することを目的とした研究が行われてきた。例えば、話者の性格 [Mairesse 07] や、印象 [沈 12]、人物属性（性別、年代、居住地）らしさ [宮崎 13] や、特定の話者らしさ（話者の個性）[水上 14] を対象とした研究がなされている。これらは全て、人物像（の構成要素）にふさわしい言語表現の生成を目的としている点が共通している。

話者の性格を対象とした研究 [Mairesse 07] では、特に外向性 / 内向性を表現するための言語生成器 (PERSONAGE) の設計を行っている。具体的には、外向性 / 内向性に関わる言語的特徴を表現するためのパラメータ集合（発話内容の評価極性、統語構造の複雑さなど）を用意し、その設定によって外向性 / 内向性が感じられる発話を生成するものである。加えて、映画のスク립トを用いて PERSONAGE のパラメータ設定を自動調整する研究 [Walker 11] や、ロールプレイングゲームの登場人物やストーリー（状況）などに合った発話を生成する研究 [Reed 11] も行われている。

話者の印象を対象とした研究 [沈 12] では、特にまじめさ、かわいさ / かつこ良さ、感情豊かさという観点で、印象形成に有効な言語パターンを音声転記テキストから人手で抽出し、その言語パターンを用いて作られたテキストから目的の印象を感じ取ることができるかを主観評価している。

人物属性らしさを対象とした研究 [宮崎 13] では、特に話者の年齢、性別、居住地の違いを表現することを目的として、文末表現（文末に位置する機能語列）を目的の人物属性らしい表現に書き換えている。具体的には、特定の人物属性に紐付けられたコーパスに偏って多く出現する文末表現を、その人物属性に特徴的な表現として抽出し、これを書き換え先の候補として利用するものである。

特定の話者らしさ（話者の個性）を対象とした研究

図1 書き換えの事前準備

図2 書き換え処理の流れ

[水上 14] では、統計的機械翻訳のアプローチによって言語表現を特定の話者らしく書き換えている。具体的には、目標話者に特徴的な語を抽出したうえで、その言い換え表現を、特徴語との文脈の類似性に基づいて、一般的な話し言葉コーパスの中から獲得し、その際に算出された言い換え確率を翻訳モデルとして代用する。

文献 [宮崎 13, 水上 14] の研究は、本研究と目的が非常に近い。本研究との違いは、これらの関連研究がキャラクター性を表現することに注力しており、何らかの条件下で最良の候補を目的キャラクターらしい言語表現として決定的に選択する点である。これらの研究は、生成される言語表現の多様性には配慮していない。

3. 文節機能部の書き換え手法

本論文で提案する書き換え手法は、発話を構成する各文節の機能部を目的のキャラクターに適した確率で適した表現に書き換えることによって、キャラクター性を表現しつつ、バリエーション豊かな言語表現を生成しようとするものである。

提案手法は、図1に示される事前準備と、図2に示される書き換え処理とで構成される。事前準備では、特定のキャラクターらしさを強調するように人手で書き換えられた発話からなるコーパスから、人間が発話にキャラクターらしさを与えようとする際に行った書き換え操作を抽出し、書き換え規則集を作成する。併せて、複数あるそれぞれのキャラクターについて、コーパスに出現する機能

部の表記の頻度を数えておく。書き換え処理では、文節機能部の認定に必要な文節・形態素情報の取得 (1) を行ったのち、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举 (2) を行い、列举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、本章では、キャラクターらしい発話のコーパスを用いて書き換え規則集を作成する方法 (3.2 節) と、(2) および (3) の具体的な処理内容 (3.3 節, 3.4 節) について述べる (文節・形態素情報の取得 (1) については (2) の中で説明する)。

3.1 キャラクタコーパス

本研究では、日本語において書き換えが可能な言語表現のペアを獲得するために、特定のキャラクター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、キャラクター文と呼ぶ) からなるコーパス [宮崎 14] を利用する。これは、対話システム用に人手で構築された Artificial intelligence markup language (AIML) 形式の対話ルール [目黒 14] からランダムに抽出したシステム応答文をキャラクターらしく書き換えることによって作成されたものである。下記に AIML で書かれた対話ルールの例を示す。ここでは、ユーザの発話が <pattern> にマッチした場合に <template> タグで囲まれた応答文を返すようになっている。この応答文をキャラクターコーパスの書き換え前の文 (以降、元文と呼ぶ) として利用する。

```
<category>
<pattern> アイドル * 憧れ </pattern>
<template> 好きなアイドルは誰ですか? </template>
</category>
<category>
<pattern> 10 年 ぐらい </pattern>
<template> 10 年というと、けっこう長いですね。 </template>
</category>
```

キャラクターコーパスの作成においては、書き換え作業は全て日本語母語話者が行い、キャラクターごとに別の書き換え作業者が担当した。作業者に与えた教示は下記のとおりである。

- 指定されたキャラクターらしさを強調するように発話を書き換えること。
- 元の発話の意味を維持できるのであれば、語順、単語、文字種は変更してもよい。
- 現実世界であまり耳にしないような誇張された表現に書き換えてもよい。

表 1 にキャラクターの一覧と各キャラクターの書き換え結果例を示す。対象とするキャラクターは、年齢、性別、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 種、年齢は 10 歳 (子供)、25 歳 (若年成人)、50 歳 (中年成人)、75 歳

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。

(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 (10 歳 (子供) は会話相手との親密度によって話し方を変えることが想定しづらく、50 歳 (中年成人) と 75 歳 (老年成人) の会話相手として想定されるのは年下が主であろうという観点から、親密度の異なるキャラクタは男女 25 歳のみとした)。

キャラクタを構成する属性は、言語表現との結びつきの強さを考慮して選定した。性別については、現代日本語の実態および役割語として「言葉の性差」が存在すると言われており [金水 03]、言語表現との結びつきが強いことは明らかである。年齢については、日本語の使用実態として「若者語」が観察されていることから [鍵水 13]、言語表現との結びつきが強いものと判断した。会話相手との親密度については、相手との社会的距離 (相手について良く知っているかどうか) が使用される言語表現に影響を与えるとされていることから [Walker 97]、これも言語表現との結びつきが強いと考えた。

本研究で使用するキャラクターコーパスは、12,640 文 (1,264 文 × 10 キャラクタ) からなっている。

3.2 書き換え規則集の作成

キャラクターコーパスの元文と各キャラクター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、どのような環境 (文法的制約条件下) 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。

キャラクターコーパスの元文と各キャラクターのキャラク

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

タ文の対から書き換え規則を抽出するには、まず、元文とキャラクタ文に含まれる文節の対応付けを動的計画法を用いて行う。次に、対応付けられた文節について下記 4 つの要素を抽出する（なお、対応付けられた文節のうち、文節の主辞品詞と係り種別が同一で機能部表記だけが異なるものを規則集への登録対象とする）。

- 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

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 [工藤 02]（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。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（e.g., 「きれいな/色」「横須賀という/市」）、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（e.g., 「すばやく/走る」「高いが/美味しい」）。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（e.g., 「ですか」「ます？」）、そうでないか（e.g., 「ですね」「ますよ」）によって異なる種別とする。書き換え規則の中で、文節の主辞品詞と係り種別は、書き換え元から書き換え先への置換が起こり得る環境を規定するものであり、これは日本語として違和感のある書き換えが生じることを抑制するために導入するものである。なお、文節の主辞品詞と係り種別を用いた絞り込みを行うと、各書き換え元あたりの書き換え先（候補数）は平均 5.5 種類から 4.8 種類に減少する。表 2 に得られた規則の例を示す。

3.3 書き換え先候補の列挙

本節より書き換え処理について説明する。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、表 3 の例では、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、書き換え先候補となる表記は、表中で「書き換え先表記」として示され

ている「よ」「よね」「わ」「わね」などとなる。このとき、日本語として違和感のない書き換え先を列挙できるのは、書き換え規則を適用可能な環境（文節の主辞品詞・係り種別）を適切に規定しているためである。

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。ここでは、各書き換え先候補が生成される頻度を確率的に調整することでキャラクタ性を表現しつつ、特定の候補に偏らず、多様な候補が生成されるようにする。

女性キャラクタの年齢設定が書き換え候補の選択にどう反映されるかを例にとり、候補選択の方法を説明する。「（そうだと）思うな」の「な」を書き換える際は、主辞が動詞の文末文節（平叙文）に出現する「な」の書き換え先の表（表 3）を参照して、各書き換え先のキャラクタ別コーパスにおける出現頻度を取得し、「主辞が動詞の文末文節（平叙文）」という環境における各書き換え先の出現確率を算出する。例えば、書き換え先「わ」の出現確率は、25 歳（F25p）の場合は 8%（3/39）、50 歳（F50p）では 53%（27/51）となる（表中の太字箇所）。各候補が生成される確率がここで算出された出現確率に従うように、最終的に出力する候補の選択を行う。これにより、25 歳や 50 歳キャラクタの発話では、「わ」を一定の頻度で使用することによって女性らしさを印象づけたうえで、残された 92%や 47%では、「わ」の代わりに他の多様な表現を使うことができる。

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、以下の式に従って出現頻度の補正を行う。

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

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

具体的には、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 λ 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 （本論文では全 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

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

図 3 機能部表記の出現頻度分布

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。

図 3 に、キャラクタコーパスにおける機能部表記の出現頻度分布を示す。横軸はキャラクタを表し、縦軸は各キャラクタのコーパスに出現した機能部表記の延べ数を表している。この図から、どのキャラクタのコーパスにも出現頻度が 1 である機能部表記が 500 個程度（全体の 15% 程度）含まれていることが分かる。これらについては、出現頻度が 0 であるものとみなしたうえで頻度補正が行われる。出現頻度 1 の表現を頻度 0 とみなすことの目的は、コーパス中に含まれる誤記（タイプミス等）や、キャラクタコーパス作成時にたまたま見過ごされてしまった不適切な表現が生成されるのを防ぐことである。また、頻度補正を行う理由は、あるキャラクタのコーパスに出現する表現が、他のキャラクタの表現としても利用可能な事例が見受けられ、他キャラクタのコーパスでの出現頻度を借用すれば、書き換え先表記のバリエーションを増やしたり、ゼロ頻度の問題によって書き換えが行えないケースを減らしたりできると考えるためである。ただし、この方法によって頻度を調整しても、書き換え規則が不足している場合や、書き換える前と同じ表記が選択される場合には、結果として書き換えが行われない。

なお本論文では、重み $g_{i,j}$ は人手で指定した。重みの指定にあたって、2 つのキャラクタ i と j が類似している

か否かは、 i と j との間での性別・年齢・親密度という 3 属性の値の重なりを考慮して著者が判断し、重みは著者の直観に基づいて決定した。重みの決定を自動化するには、人間が知覚するキャラクター性の類似度を自動推定することが必要となり、さらなる検討を要すると思われる。よって、重み決定の自動化は今後の課題とする。また、本研究で用いたキャラクタコーパスの量が学習データとして十分であるかどうかは未検証であり、その検証は今後の課題とする。

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）、キャラクター性（目的のキャラクタらしさが表れているか）、言語表現の多様性という 3 つの観点での評価を行った。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 12,139 文（1,214 文 × 9 キャラクタ、1,213 文 × 1 キャラクタ）であり、書き換え規則は 3.2 節の方法で獲得された 7,517 個を使用した。

4.1 実験 1：文法性評価

文に含まれる言語表現を機械的に書き換えると、日本語として違和感のある文が生じる恐れがある。そこで、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。評価に使用したデータは、キャラクタコーパスの作成で使用した対話ルールに含まれるシステム応答文（<template>タグで囲まれたテキスト）からランダムに抽出した 100 文（学習データは含まない）を提案手法によって書き換えたものである。評価者は、言語アノテーション業務に従事する者 2 名（評価者 A：経歴 3 年以上、評価者 B：経歴 5 年以上）である。

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

表 5 文法性判定結果 (キャラクタ別評価値平均)

キャラクタ	評価者 A	評価者 B
F10p	3.97	4.42
M10p	4.01	4.50
F25p	3.75	4.29
F25b	4.11	4.64
M25p	3.97	4.53
M25b	4.19	4.56
F50p	3.92	4.44
M50p	3.92	4.48
F75p	3.94	4.44
M75p	3.73	4.27
平均	3.95	4.46

表 5 に文法性判定結果の平均点を示す。提案手法による書き換え結果は、評価者 A では平均 3.95 点、評価者 B では平均 4.46 点の評価を得た。これは、文末の機能表現のみを書き換え対象とした研究 [宮崎 13] における結果 (文法性スコア 5 段階のうち平均 4.16 点) と同等の値であり、文末以外の文節も書き換え対象とする本研究においては良好な結果であると言える。表 6 には、全 10 キャラクタの評価を点数別に集計した結果を示す。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 10%、評価者 B では 6% であり、提案手法によれば、書き換えを施しても、日本語として違和感のある文は 10% 以下に抑えられることも確認できた。

違和感の原因には、内容語と機能語の意味的不整合 (e.g., 存在動詞「ある」+ 継続「ている」)、機能語の意味・機能的曖昧性 (e.g., 「と」の曖昧性; 「晴れるとうれしい」(順接)、「そうだと思う」(引用)、「朝食と昼食」(並立))、態の変更 (e.g., 「滝に打たれた」「滝に打った」)、提題化の誤り (e.g., 「200 点に落ちる」「200 点って落ちる」) などがあった。提案手法では、書き換え操作を書き換えられる表記と個々の文節の環境 (文節の主辞品詞と係り種別) のみで単純化して捉えているために、こうした問題が起きている。これらの問題を解決するには、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、目的のキャラクタ性が表れているかを評価するために、書き換えら

表 6 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	24(2%)	83(8%)	153(15%)	353(35%)	387(39%)
B	6(1%)	53(5%)	109(11%)	142(14%)	690(69%)

れた文に対するキャラクタ判別実験を実施した。評価者は、性別 (男性 / 女性)、年代 (40 歳未満 / 40 歳以上)、出身地 (東日本 / 西日本) のバランスがとれた 10 名の日本語母語話者が行った。10 名の内訳は表 7 に示す。

この実験では、特定のキャラクタらしく書き換えられた文集合を評価者 10 名それぞれに提示し、属性値を判別させることを行った。例えば、年齢を判別させる場合は、図 4 に示すような文集合を年齢の選択肢 (10 歳, 25 歳, 50 歳, 75 歳) とともに提示し、その文集合の話者の属性値として最も適切だと感じるものを選択させる。

キャラクタ性評価では、本技術を複数発話からなる対話に適用することを念頭に置き、一つの対話を終えた後に伝わるキャラクタ性を評価するため、1 文ではなく、複数文からなる文集合を評価の単位とした。なお、今回の実験では、文集合のそれぞれは 10 文で構成することとした。この 10 文は、キャラクタコーパスの作成で使用した対話ルールに含まれるシステム応答文 (<template> タグで囲まれたテキスト) をランダムに抽出した文 (学習データは含まない) で構成される。

評価対象として評価者に提示したデータは、上記の文集合に対して提案手法で書き換えを施したもの、および下記の 3 種類の方法で作成したものである。

無変換 書き換えられていない文集合 (書き換え前の文集合) を用いる方法。これをキャラクタ性付与性能の下限値を測るためのベースラインとして用いる。
 人手書き換え 人手で書き換えられた文集合 (学習データとして使用した文で構成され、内容語の書き換えも含む) を用いる方法。これをキャラクタ性付与性能の (考え得る中での) 上限値を測るために用いる。
 決定的書き換え キャラクタコーパスにおける出現確率が最も高い候補を常に選択する方法。書き換え対象として用いる文集合、書き換え規則集の作成、および書き換え先候補の列挙については提案手法と同一で、書き換え先候補の選択のみが提案手法と異なる。

評価に使用した文集合は各手法につき 800 個で、内訳は表 8 に示す。性別の判別用に 200 個 (目的キャラクタの性別が男性、女性である文集合が 100 個ずつ)、年齢の判別用に 400 個 (目的キャラクタの年齢が 10 歳, 25 歳, 50 歳, 75 歳である文集合が 100 個ずつ)、親密度の判別用に 200 個 (目的キャラクタの親密度が親密、非親密である文集合が 100 個ずつ) となっている。1 つの文集合につき、提案手法では平均 6.5 か所、決定的書き換えでは平均 6.2 か所、人手書き換えでは平均 18.6 か所が書き換えられていた。標準偏差は順に 2.8, 2.8, 6.2 で

表 7 キャラクター性評価の評価者の内訳

評価者 ID	性別	年代	出身地
C	男性	40 歳以上	東日本
D	男性	40 歳以上	西日本
E	男性	40 歳以上	西日本
F	男性	40 歳未満	東日本
G	男性	40 歳未満	西日本
H	女性	40 歳以上	東日本
I	女性	40 歳以上	西日本
J	女性	40 歳未満	東日本
K	女性	40 歳未満	東日本
L	女性	40 歳未満	西日本

年齢判別の例

文集合	選択肢
1. しかし、最初に食べようと思った人はすごいよね	☐ 10歳
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	☒ 25歳
...	☐ 50歳
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？	☐ 75歳
9. 餌代って結構掛かってる	
10. 残暑が優しいって言い方ってさ、聞かないんだ	

図 4 キャラクター性評価データの例

あった。

図 5 に評価結果を示す。正解率は評価者の判別結果が正解（目的キャラクタ）と一致した文集合の割合（評価者 10 名の平均値）である。提案手法で書き換えられた文集合は、最大で（属性値が男性の場合）0.78 の正解率で正しく属性値が判別された。無変換の場合と比較すると、提案手法の正解率は最大で（属性値が親密の場合）0.27 高かった。また、25 歳の場合を除く全ての属性値の判別において、提案手法の正解率のほうが有意に高かった（両側二項検定で $p < 0.05$ ）。これにより、提案手法には発話にキャラクター性を付与する効果が備わっていることが確認できた。

また、提案手法と決定的書き換えとを比較すると、両側二項検定の結果、75 歳の場合を除く全ての属性値の判別において正解率に有意な差はないことが分かった。よって、提案手法は決定的書き換えと同等のキャラクター性付

表 10 書き換え結果における機能表現の Type-Token Ratio (Type/Token) . p 値は、提案手法と決定的書き換えとの間で両側二項検定の結果。

キャラクタ	文数	人手書き 換え	提案手法	決定的書 き換え	
F10p	820	0.175	0.163	0.151	$p = 0.066$
M10p	820	0.153	0.159	0.146	$p = 0.040$
F25p	590	0.202	0.182	0.156	$p < 0.001$
F25b	950	0.161	0.147	0.137	$p = 0.088$
M25p	590	0.173	0.190	0.180	$p = 0.249$
M25b	950	0.132	0.152	0.139	$p = 0.028$
F50p	820	0.147	0.171	0.151	$p = 0.002$
M50p	820	0.152	0.164	0.157	$p = 0.290$
F75p	820	0.181	0.169	0.148	$p = 0.002$
M75p	820	0.153	0.158	0.143	$p = 0.016$

与性能を持つことが確認できたと言える。ただし、属性値が 10 歳や 75 歳の場合には、提案手法であっても決定的書き換えであっても、判別正解率が著しく低いことが分かった。

決定的書き換えおよび提案手法による書き換え結果を比較するために作成した例を表 9 に示す。書き換えられた表現は太字で示されている。

4.3 実験 3：言語表現の多様性評価

提案手法がバリエーション豊かな言語表現を出力できるものであるかを評価するために、書き換えられた文に含まれる機能表現の多様性を計測した。キャラクター性評価で用いた、決定的書き換えおよび提案手法で書き換えられた全ての文を用いて、それらに含まれる機能表現の多様性を評価した。また、学習データとして使用した人手書き換え（12,139 文）からランダムに抽出した文に対する評価値も参考として示す。

1 つの機能表現は文節を越えないものとして扱い、自立語（名詞、動詞、形容詞、副詞、接続詞、感動詞、連体詞、接頭詞、フィラー）を含まない形態素連続を機能表現として認定した。評価指標としては、機能表現の異なり数（Type）を総数（Token）で除した値である Type-Token Ratio（TTR）を使用した。書き換えられた文の文節・形態素情報は係り受け解析器 CaboCha [工藤 02] の解析結果から取得した。

評価結果を表 10 に示す。想定した通り、全てのキャラクタについて、提案手法のほうが決定的書き換えよりも高い TTR を示しており、M25p と M50p を除く全てのキャラクタにおいて有意差（両側二項検定で $p < 0.05$ ）もしくは有意傾向（両側二項検定で $p < 0.10$ ）があった。よって、提案手法は決定的な手法と比べて、より多様なバリエーションの言語表現を生成可能であることが確認できたと言える。

表 8 キャラクター性評価データの内訳

評価対象	目的属性値	文集合数	計
性別	男性	100	200
	女性	100	
年齢	10 歳	100	400
	25 歳	100	
	50 歳	100	
	75 歳	100	
親密度	親密	100	200
	非親密	100

図5 キャラクタ判別の正解率。n.s., *, ***はそれぞれ、提案手法と無変換または決定的書き換えとの間での両側二項検定の結果が有意差なし、 $p < 0.05$, $p < 0.001$ 未満であることを指す。

表9 書き換え結果の例

	F10p (女性・10歳・親密)	F25p (女性・25歳・親密)
提案手法	<p>一眼レフが欲しいよね ネット上には懸賞サイトが多すぎる気がするよね カップ麺って食べる？ 震災はいつ来るか分からないから、事前の準備って絶対に必要だ [アイドル]のコンサートチケットは[オークション]で5万出してやっと落札できるよ かぼちゃの料理っていったらなんなの？ 野球ってサッカーどっちが好きなの？ 女性剣士は白い防具が似合ったら思う？ 日本三景はいえば？ 沖縄に行ったら、どこへ行きたい？</p>	<p>一眼レフが欲しいね ネット上には懸賞サイトが多すぎる気がするね カップ麺も食べる？ 震災、いつ来るか分からないから、事前の準備は絶対に必要なんだよね [アイドル]のコンサートチケットって、[オークション]で5万出してやっと落札できていたわ かぼちゃの料理っていったらなんなのかしら？ 野球ってサッカーどっちが好き？ 女性剣士は白い防具が似合ってると思う？ 日本三景っていったら？ 沖縄に行ってどこへ行ったら？</p>
決定的書き換え	<p>一眼レフが欲しい ネット上には懸賞サイトが多すぎる気がするよ カップ麺は食べる？ 震災はいつ来るか分からないから、事前の準備は絶対に必要だ [アイドル]のコンサートチケットは[オークション]で5万出してやっと落札できるよ かぼちゃの料理はいったらなんなの？ 野球ってサッカーどっちが好きなの？ 女性剣士は白い防具が似合ってると思う？ 日本三景はいえば？ 沖縄に行ってどこへ行ったら？</p>	<p>一眼レフが欲しいよね ネット上には懸賞サイトが多すぎる気がするよ カップ麺って食べる？ 震災っていつ来るか分からないから、事前の準備は絶対に必要だよ [アイドル]のコンサートチケットって[オークション]で5万出してやっと落札できるよ かぼちゃの料理っていったらなんなの？ 野球ってサッカーどっちが好き？ 女性剣士って白い防具が似合ってると思う？ 日本三景っていったら？ 沖縄に行ってどこへ行ったら？</p>
	M10p (男性・10歳・親密)	M50p (男性・50歳・親密)
提案手法	<p>タイって昔からずっと親日国だね いまや大物歌手の仲間入りだね マラリアになったことある？ 3G、何のことかわかるかな？ 育ててみたら意外と難しいんだよね 地元はどこですか？ 年を取れば痩せにくくなるのが悩みね カードやレターセットってこだわって選ぶことがあるんだ カルーアミルクって意外とアルコール度数があるんだ 料理のチェーン店、みな同じ味だ</p>	<p>タイは昔からずっと親日国だよ いまや大物歌手の仲間入りだ マラリアになったことある？ 3Gは何のことかわかる？ 育ててみたら意外と難しいん 地元はどこですか？ 年を取ると痩せにくくなるのが悩みだね カードやレターセットもこだわって選ぶことがあるんだ カルーアミルクは意外とアルコール度数があるよ 料理のチェーン店はみな同じ味だよ</p>
決定的書き換え	<p>タイは昔からずっと親日国だよ いまや大物歌手の仲間入りだよ マラリアになったことある？ 3Gは何のことかわかる？ 育ててみたら意外と難しいんだよ 地元はどこですか？ 年を取れば痩せにくくなるのが悩みだよ カードやレターセットってこだわったら選ぶことがあるよ カルーアミルクは意外とアルコール度数があるよ 料理のチェーン店はみな同じ味だよ</p>	<p>タイは昔からずっと親日国だ いまや大物歌手の仲間入りだ マラリアになったことある？ 3Gは何のことかわかる？ 育ててみたら意外と難しいんだ 地元はどこですか？ 年を取ると痩せにくくなるのが悩みだ カードやレターセットってこだわって選ぶことがある カルーアミルクは意外とアルコール度数がある 料理のチェーン店はみな同じ味だ</p>

5. ま と め

本論文では、対話エージェントの発話に対してキャラクター性を付与し、かつ、バリエーション豊かな言語表現を生成する手法を提案した。具体的には、発話を構成する各文節の機能部を目的のキャラクターに適した確率で適した表現に書き換えることで、これを実現した。加えて、評価実験により、キャラクター性付与性能（人間が正しくキャラクター性を判別できる割合）を、書き換え前と比較して最大約 30% 向上させられること、従来手法よりも多様な言語表現を用いたバリエーション豊かな発話が生成できること、および、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 10% 以下に抑えられることを明らかにした。

しかしながら、本技術のキャラクター性付与性能は、目的のキャラクター（属性値）によって大きく異なることも分かった。特に、年齢（10 歳，75 歳キャラクター）らしさの付与性能は著しく低いため、原因のさらなる分析が必要である。また、人手で書き換えてもキャラクター性が付与しにくい（他者が正しくキャラクター性を判別できる割合が低い）場合があることも明らかになった。どのような要因でこうした現象が起きるのかについても今後分析していきたい。

◇ 参 考 文 献 ◇

- [Higashinaka 05] Higashinaka, R., Prasad, R., and Walker, M.: Augmenting variation of system utterances using corpora in spoken dialogue systems, in *Proceedings of 2005 IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 262–267 (2005)
- [金水 03] 金水敏：ヴァーチャル日本語役割語の謎，岩波書店 (2003)
- [工藤 02] 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)
- [Mairesse 07] Mairesse, F. and Walker, M.: PERSONAGE: Personality generation for dialogue, in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 496–503 (2007)
- [目黒 14] 目黒豊美，杉山弘晃，東中竜一郎，南泰浩：ルールベース発話生成と統計的発話生成の融合に基づく対話システムの構築，人工知能学会全国大会（第 28 回），2M5-OS-20b-2 (2014)
- [宮崎 13] 宮崎千明，平野徹，東中竜一郎，牧野俊朗，松尾義博：発話にキャラクター性を与えるための文末表現の変換，人工知能学会言語・音声理解と対話処理研究会（第 68 回），pp. 41–46 (2013)
- [宮崎 14] 宮崎千明，平野徹，東中竜一郎，牧野俊朗，松尾義博，佐藤理史：話者のキャラクター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)
- [水上 14] 水上雅博，Neubig, G., Sakti, S., 戸田智基，中村哲：特徴的話者を対象とした言語的個性変換，情報処理学会研究報告，自然言語処理研究会報告，No. 20, pp. 1–7 (2014)
- [Reed 11] Reed, A. A., Samuel, B., Sullivan, A., Grant, R., Grow, A., Lazaro, J., Mahal, J., Kurniawan, S., Walker, M., and Wardrip-Fruin, N.: A step towards the future of role-playing games: The SpyFeet mobile RPG project, in *Proceedings of the 7th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, pp. 182–188 (2011)
- [沈 12] 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ー付与，情報処理学会論文誌，Vol. 53, No. 4, pp. 1269–1276 (2012)
- [Walker 97] Walker, M. A., Cahn, J. E., and Whittaker, S. J.: Improvising linguistic style: Social and affective bases for agent personality, in *Proceedings of the 1st international conference on Autonomous agents*, pp. 96–105 (1997)
- [Walker 11] Walker, M. A., Grant, R., Sawyer, J., Lin, G. I., Wardrip-Fruin, N., and Buell, M.: Perceived or not perceived: Film character models for expressive NLG, in *Proceedings of the 4th Joint Conference on Interactive Digital Storytelling*, pp. 109–121 (2011)
- [鍵水 13] 鍵水兼貴：首都圏の言語の実態と動向に関する研究 全国若者語調査地図集，国立国語研究所 (2013)

〔担当委員：菊池 英明〕

2015 年 6 月 4 日 受理

著 者 紹 介

宮崎 千明

2008 年南山大学外国語学部英米学科卒業，2010 年名古屋大学大学院国際開発研究科博士前期課程修了。2010 年日本電信電話株式会社入社。現在，NTT メディアインテリジェンス研究所に所属。自然言語処理に関する研究開発に従事。言語処理学会会員。

平野 徹

2003 年和歌山大学システム工学部卒業，2005 年奈良先端科学技術大学院大学情報科学研究科博士前期課程，2012 年博士後期課程修了。2005 年日本電信電話株式会社入社。現在，NTT メディアインテリジェンス研究所に所属。自然言語処理に関する研究開発に従事。博士（工学）。言語処理学会会員。

東中 竜一郎 (正会員)

1999 年慶應義塾大学環境情報学部卒業，2001 年同大学大学院政策・メディア研究科修士課程，2008 年博士課程修了。2001 年日本電信電話株式会社入社。現在，NTT メディアインテリジェンス研究所に所属。質問応答システム・音声対話システムの研究開発に従事。博士（学術）。情報処理学会，言語処理学会，電子情報通信学会 各会員。

牧野 俊朗

1987 年東京大学工学部卒業，1992 年同大学大学院工学系研究科博士課程修了。同年，日本電信電話株式会社入社。現在，NTT メディアインテリジェンス研究所主幹研究員。自然言語処理に関する研究開発に従事。博士（工学）。

松尾 義博

1988 年大阪大学理学部卒業，1990 年同大学大学院理学研究科博士前期課程修了。同年，日本電信電話株式会社入社。現在，NTT メディアインテリジェンス研究所，音声言語メディアプロジェクト，音声・言語基盤技術グループリーダー。自然言語処理に関する研究開発に従事。情報処理学会，言語処理学会 各会員。

佐藤 理史 (正会員)

1988 年京都大学大学院工学研究科博士後期課程電気工学第二専攻研究指導認定退学。京都大学工学部助手，北陸先端科学技術大学院大学助教授，京都大学大学院情報科学研究科助教授を経て，2005 年より名古屋大学大学院工学研究科教授。工学博士。言語処理学会，情報処理学会，日本認知科学会，ACM 各会員。