

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

宮崎 千明¹ 平野 徹¹ 東中 竜一郎¹ 牧野 俊朗¹ 松尾 義博¹ 佐藤 理史²

¹ 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

² 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は, 音声検索のツールとしてだけでなく, 会話や反応を楽しむための娯楽ツールとしても多くのユーザの関心を集めている.

我々は, 音声対話エージェントの持つ, 会話相手としての価値や娯楽性をより高めるためには, エージェントに対するキャラクター付けが非常に重要であると考えている. キャラクター付けとは, エージェントに対して何らかの人物像を与えることであり, エージェントをより人間らしく, 親しみやすい存在にすることを可能にするものである. 音声対話エージェントに多様なキャラクターを用意し, ユーザが好きな時に好きなキャラクターを選んで話せるようにすれば, 会話から得られる満足感はより大きくなると考えられる.

しかし, 対話エージェントのキャラクター付けには, 非常に大きなコストが発生するという問題がある. なぜなら, エージェントの発話に特定のキャラクターらしさを持たせるには, キャラクター設定に応じて, 内容と言葉遣いが共に適切であることを満たす発話データを作成する必要があり, 大規模な発話データを用意する必要性が生じた場合や, 何種類もの異なるキャラクターを用意したい場合に, すべての作業を人手で実施すると非常に手間がかかってしまうからである.

そこで我々は, 特定のキャラクターらしさを持った発話を自動生成することにより, コストを削減し, 簡単に多様なキャラクターを持った対話エージェントを作り上げることを目指す. 本稿では, 部分的に発話を書き換えることによって話者のキャラクター性を表現する技術の実現を目指し, 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を行う.

2 関連研究

発話にキャラクター性を与えることを目的とした研究には, 英語を対象に, 話者の性格 (Big Five; 外向性, 協調性, 情緒安定性など) を考慮した発話を生成する研究があり [1], 性格に応じた発話生成のためのパラメータを映画の SCRIPT を用いて学習する研究もある [2]. 他には, キャラクターに対する印象 (まじめさ, かawaiiさ/かっこ良さ, 感情豊か) の形成に貢献する言語パターンを手作業で抽出し, 抽出されたパターンに基づいて文書を書き換えを行う研究もある [3]. 本研究は日本語を対象とし, キャラクター性の構成要素としては, 性別, 年齢, 会話相手との親密度の 3 つに着目してキャラクターらしさに寄与する言語表現の分析を行う.

3 キャラクター性を強調した発話データ

話者のキャラクター性を分析する対象として, 特定のキャラクター性を強調するように人手で書き換えた発話データを用意した. 書き換えの元となる発話データには, 対話システムの発話用に人手で構築された発話データベースからランダムに抽出したものを利用した. すべての書き換え作業は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, 文字種は変更することができる.
- 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構わない.

指定するキャラクターとしては, 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 種類を設定する. 表 2 に全キャラクターの一覧と各キャラクターの書き換え発話例を示す. 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

表 1: 書き換え文字数ごとの発話数の分布

書き換え文字数	キャラクタ									
	F10p	M10p	F25p	F25b	M25p	M25b	F50p	M50p	F75p	M75p
5 未満	130	48	128	195	153	198	110	202	53	27
5 以上 10 未満	134	77	121	82	132	76	130	97	132	144
10 以上 15 未満	36	75	43	22	18	24	48	7	83	76
15 以上 20 未満	6	37	13	7	4	7	10	1	31	39
20 以上 25 未満	1	32	2	1	0	2	6	0	4	16
25 以上 30 未満	0	20	0	0	0	0	2	0	3	4
30 以上	0	18	0	0	0	0	1	0	1	1

表 2: 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」）

キャラクタ	属性	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。

(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 種とした。なお, 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 歳のみを対象とした。

表 3 に, 書き換え作業によって収集したデータの概観を示す。収集したデータは, 632 発話 (平叙文 307 発話, 質問文 325 発話) の元文と, その元文を各キャラクタらしさを強調するように書き換えた発話データ (平叙文 3070 発話, 質問文 3250 発話) で構成されている。表 3 の書き換え異なり数は, 元文と 10 種のキャラクタを合わせた最大 11 種類の発話の中に, 異なる発話が平均していくつ含まれているかを示したものである。書き換え文字数は, 置換または削除された文字の合計 (一発話平均) を示す。

表 1 に, 各キャラクタに対する書き換え文字数ごとの発話数の分布を示す。表中の数値は, 左端に記載の書き換え文字数の範囲に該当する発話の数となっている。どのキャラクタにおいても, 書き換え文字数 5 や 10 を境に発話数が急激に減少しており, 10 文字未満

表 3: 書き換えデータの概観

元文 (発話平均)				書き換え結果 (発話平均)	
発話種別	発話数	形態素数	文字数	異なり数	文字数
平叙文	307	12.6	21.5	9.4	6.73
質問文	325	10.8	17.6	8.0	6.33

の書き換えが大半を占めていることが分かる。ただし, キャラクタ M10p には, 書き換え文字数が 30 以上の発話も少なからず存在しており, 書き換えを施すべき, または, 施せる箇所の量はキャラクタ間で異なっていることが示唆されている。

4 キャラクタ性認知実験

前節で示した書き換えデータが話者のキャラクタ性に寄与する言語表現の分析対象として妥当であるかを検証するために, 日本語母語話者の主観評価によるキャラクタ性の認知実験を行った。

実験では, 特定の属性のみが異なる 2 つのキャラクタの発話ペアを提示し, 属性値を判別させることを行った。例えば, 「(A) 辛い回鍋肉が好き」 「(B) 辛い回鍋肉が好きなの」という発話ペアを提示し, A と B のどちらが男性で, どちらが女性かを判別させる。実験では, 作業員 A, B (筆頭著者) の 2 名が同一のデータに対して個別に評価作業を行い, 2 名の判定の一致度を測定した。評価に使用する発話数は, 平叙文, 質問文ともに 360 ペアずつとした。内訳としては, 男性対女性が 80 ペア (親密度 p の 4 キャラクタペア × 20 発話), その他の属性の比較では 40 ペア (2 キャラクタペア × 20 発話) ずつとなっている。

表 4 に平叙文に対する実験結果を, 表 5 に質問文に対する実験結果を示す (正解率の値は作業員 A の結果)。平叙文, 質問文ともに, 7 割以上の正解率で正しく属性値を判別できた。さらに, 作業員間の判別結果は 7 割以上が一致し, 平叙文の 25 歳対 50 歳を除いては, カップ値は 0.41 以上 (中程度の一致) となった。よって, 本書き換えデータには第三者が認知できる程度の話者のキャラクタらしさが備わっていると言

表 4: キャラクタ判別結果 (平叙文)

キャラクタ	正解率	一致率	カッパ値
男性 対 女性	0.75	0.91	0.82
10 歳 対 25 歳	0.83	0.85	0.70
10 歳 対 50 歳	0.80	0.83	0.65
10 歳 対 75 歳	1.00	1.00	1.00
25 歳 対 50 歳	0.73	0.68	0.34
25 歳 対 75 歳	0.98	0.98	0.94
50 歳 対 75 歳	0.98	1.00	1.00
親密 対 非親密	0.98	0.98	0.95

表 5: キャラクタ判別結果 (質問文)

キャラクタ	正解率	一致率	カッパ値
男性 対 女性	0.75	0.81	0.63
10 歳 対 25 歳	0.78	0.78	0.55
10 歳 対 50 歳	0.73	0.80	0.60
10 歳 対 75 歳	1.00	1.00	1.00
25 歳 対 50 歳	0.70	0.70	0.41
25 歳 対 75 歳	0.98	1.00	1.00
50 歳 対 75 歳	1.00	1.00	1.00
親密 対 非親密	1.00	1.00	1.00

える。

5 キャラクタらしさを強調するための書き換え操作の分析

本研究の大目標である、発話を部分的に書き換える技術を実現するためには、それぞれのキャラクタらしさを強調するために、発話の「どのような箇所」を「どのような表現」に書き換えればよいかを把握する必要がある。本稿ではまず、全てのキャラクタに共通して多く書き換えられる品詞に分析対象を絞り、その品詞の書き換えられ方に表れるキャラクタ間の違いを分析する。

分析の下準備として、元文と書き換え文のペアを比較し、元文から削除された要素、書き換えによって挿入された要素、および置換された要素同士の対応付けを行う。対応付けは動的計画法を利用して文字単位で行い、連続した置換(削除・挿入)箇所は一つの書き換え箇所として扱う。その後、元文と書き換え文のそれぞれに対して形態素解析を行い、書き換え箇所の範囲内に左端が位置する形態素の品詞を、その書き換え箇所の品詞として付与する。

元文 駅の傍に美味しい寿司屋があります。
書き換え文 駅の近くに美味しいお寿司屋さんがあるよ。

対応付け 駅の【近く／傍】に美味しい【お／】寿司屋【さん／】があ【るよ／ります】。

対応付け+品詞 駅の【近く 名詞／傍 名詞】に美味しい【お 接頭詞／】寿司屋【さん 接尾辞／】があ

表 7: 書き換えパターンとその出現回数 (書き換え元品詞ごとの上位 3 項目)

書き換え元	書き換え先	出現回数
助詞	助詞	531
助詞	φ	438
助詞	読点	67
助動詞	φ	470
助動詞+終助詞	φ	428
助動詞	終助詞	165
判定詞	判定詞	217
判定詞	φ	206
判定詞+終助詞	φ	204

【るよ 終助詞／ります 助動詞】。

上記のように書き換え箇所が対応づけられたデータを用いて、書き換え箇所数を品詞別に集計した。書き換えられた箇所の品詞別集計を表 6 に示す。集計結果から、助詞、助動詞、判定詞が全てのキャラクタに共通して多く書き換えられていることが分かる。よって、これら 3 品詞はキャラクタ性を表現するうえで重要な書き換え箇所であると言える。それぞれの品詞の書き換え先となる頻度が高い品詞上位 3 項目を表 7 に示す。

表 7 の書き換えパターンについて、具体例を観察したところ、下記の 4 点にキャラクタ間の違いが表れていることが分かった。

- 提題助詞「って」の使用
- 助詞の省略
- 丁寧表現「です」「ます」の削除
- 丁寧表現「です」「ます」と終助詞との置換

格助詞(が, を, に)や提題助詞(は)と「って」との置換は、どのキャラクタでも共通して多く出現した。ただし、キャラクタによって書き換えの向きが異なっていることが分かった。表 8 に示される引用助詞「って」とその他の助詞との書き換え回数の分布を見ると、非親密な相手との会話(M25b)や高齢のキャラクタ(F75p, M75p)の発話を想定した場合にのみ、その他のキャラクタでは積極的に使用されている「って」が逆に避けられていることが分かる。よって、提題助詞(および提題化が可能な助詞)は着目すべき書き換え箇所であり、キャラクタによって「って」とその他助詞の使い分けを行うことでキャラクタ性を表現できる箇所であると言える。

助詞の省略については、省略が起こる文脈がキャラクタらしさを表現する特徴となっている。表 9 に文脈ごとの助詞削除の頻度分布を示す。助詞省略が行われる文脈を観察してみると、「～ことがある」という、自己の経験や習慣について語る文脈での「が」の省略「～

表 6: 書き換えられた箇所の品詞別集計 (上位 5 項目. 括弧内は頻度.)

キャラクタ	1	2	3	4	5
F10p 女性・10歳・親密	助動詞 (373)	助詞 (320)	判定詞 (206)	名詞 (49)	終助詞 (42)
M10p 男性・10歳・親密	助詞 (639)	助動詞 (415)	名詞 (344)	判定詞 (223)	動詞 (183)
F25p 女性・25歳・親密	助詞 (498)	助動詞 (397)	判定詞 (222)	名詞 (87)	動詞 (75)
F25b 女性・25歳・非親密	助詞 (181)	助動詞 (130)	名詞 (44)	判定詞 (35)	終助詞 (33)
M25p 男性・25歳・親密	助詞 (474)	助動詞 (386)	判定詞 (223)	名詞 (121)	動詞 (88)
M25b 男性・25歳・非親密	終助詞 (127)	助詞 (116)	動詞 (115)	助動詞 (83)	名詞 (50)
F50p 女性・50歳・親密	助詞 (530)	助動詞 (386)	判定詞 (218)	名詞 (144)	動詞 (83)
M50p 男性・50歳・親密	助動詞 (349)	助詞 (224)	判定詞 (197)	動詞 (33)	終助詞 (22)
F75p 女性・75歳・親密	助動詞 (328)	助詞 (321)	判定詞 (199)	動詞 (86)	終助詞 (84)
M75b 男性・75歳・親密	助動詞 (397)	判定詞 (240)	名詞 (201)	助詞 (201)	終助詞 (176)

表 8: 助詞同士の置換の分布

置換前/後	キャラクタ									
	F10p	M10p	F25p	F25b	M25p	M25b	F50p	M50p	F75p	M75p
って/*	41	24	79	67	11	0	14	0	3	0
*/って	0	1	0	1	0	29	3	0	18	29
その他	9	33	18	25	26	13	35	3	36	13

表 9: 文脈ごとの助詞削除の頻度分布

文脈	キャラクタ									
	F10p	M10p	F25p	F25b	M25p	M25b	F50p	M50p	F75p	M75p
ことがある	16	28	20	10	29	0	29	5	1	0
その他	10	32	36	13	107	1	76	0	8	9
計	26	60	56	23	136	1	105	5	9	9

ことある」が大半を占めていることが分かった。ただし、キャラクタ M25p とキャラクタ F50p では、その他の文脈における助詞の省略も多く発生した。表 10 の例文は、キャラクタ M25p, F50p の助詞省略の例であるが、助詞を抜き去ることで、よりくだけた言葉遣いの印象が付加されていることが分かる。よって、助詞省略が起こりにくい箇所をでの助詞省略は、キャラクタ性を表現する書き換えとして効果的であると言える。

助動詞と判定詞の書き換えでは、丁寧表現「です」および「ます」の書き換えが大半であった。「です」「ます」の削除は、非親密な相手との会話を想定したキャラクタ F25b, M25b 以外の全キャラクタで多く観察されており、発話に親密さを与えるために行われた書き換え操作であると考えられる。

丁寧表現「です」「ます」の書き換え操作としては、

表 10: 「～ことがある」以外の文脈での助詞省略の例

キャラクタ	書き換え文の例
M25p	ホラー映画でも年齢制限【/が】あるんだ
M25p	両親とも仕事【/を】してる
F50p	中学校にバスケット部【/は】あった
F50p	吉野家に一人で入るの【/に】抵抗ない

表 11: 助動詞「ます」と置換された終助詞の分布

終助詞	キャラクタ									
	F10p	M10p	F25p	F25b	M25p	M25b	F50p	M50p	F75p	M75p
な	5	0	2	0	0	0	0	0	0	1
よ	13	22	13	0	11	0	12	19	0	0
の	2	1	2	0	0	0	0	0	2	0
わ	0	0	5	0	0	0	13	0	22	0
ね	0	0	0	0	2	0	0	0	1	0
その他	0	1	0	0	0	0	0	0	1	15

単純な削除に次いで、終助詞との置換も頻繁に発生した。表 11 に助動詞「ます」と置換された終助詞の分布を示す。女性のキャラクタに着目すると、キャラクタ F25p (25 歳), キャラクタ F50p (50 歳), キャラクタ F75p (75 歳) とキャラクタの年齢設定が上がるほど、「ます」と終助詞「わ」との置換が増加していくことが分かる。発話の女性らしさを表現するために、女性ことばの定番である終助詞「わ」が利用され、キャラクタの年齢設定が上がるるとともにその傾向が強くなったものと思われる。判定詞「です」の削除箇所についても、「ます」と同様の傾向が観察された。キャラクタ M75p では、「ぞ」「のう」等の役割語への書き換えが多く発生した。よって、丁寧表現「です」「ます」は、置換先の終助詞を変えることで異なるキャラクタ性を表現できる書き換え箇所だと言える。

参考文献

- [1] François Mairesse and Marilyn Walker. PERSONAGE: Personality generation for dialogue. In *Proc.ACL*, pp. 496–503, 2007.
- [2] Marilyn A. Walker, Grace I. Lin, and Jennifer E. Sawyer. An annotated corpus of film dialogue for learning and characterizing character style. In *LREC*, pp. 1373–1378, 2012.
- [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.