

対話シナリオにおけるキャラクター性を強調した 名詞への自動置換手法の提案と評価

田原 俊一[†] 池田 和史[†] 王 亜楠[†] 帆足 啓一郎[†]

[†]KDDI 総合研究所 〒356-8502 埼玉県ふじみ野市大原 2 丁目 1 番 15 号

E-mail: †{sh-tahara, kz-ikeda, wa-yanan, hoashi}@kddi-research.jp

あらまし キャラクター性を有する対話エージェント（キャラクタ）と対話を行うことでユーザの親近感が向上すると考えられるが、キャラクタ毎に異なる対話シナリオを新規に作成すると多大なコストが発生する。既存手法では、キャラクタに応じて発話に含まれる文節の機能語を自動的に書き換える手法を提案しているが、キャラクタに相応しくない名詞が発話に残り違和感が生じる。本稿では、発話から置換対象となる名詞の関連語を自動的に取得し、Twitter 上での言及頻度からキャラクター性を強調する語を獲得する方式を提案する。提案手法で書き換えを行なった発話にキャラクター性が現れているか評価をするため、キャラクタを判別させる被験者実験を行なった所、提案手法は既存手法と比較し正解率が有意に高くなることが確認された。

キーワード 対話システム, 対話シナリオ, キャラクタ, SNS

1. まえがき

近年、対話エージェントは広く普及しており、キャラクター性を有するもの（以下、「キャラクタ」と定義）も登場している。例えば、日本マイクロソフト社が開発したりんな[1]のような女子高生のキャラクタや、Erica[2]のような女性アンドロイド型のキャラクタが存在している。対話エージェントにキャラクター性を付与することで、キャラクタらしい応答をすることが可能となり、ユーザの親しみやすさが向上する可能性がある。対話システムを通して、ユーザがキャラクタと対話を行う際は、人手で作成した対話シナリオを用意する必要がある。しかし現状では、キャラクタ毎に異なる対話シナリオを新規に作成する必要があり、多大な作成コストが発生する。

発話にキャラクター性を付与する関連研究として、水上らはキャラクタ用に作成されたコーパスから特徴的な表現を抽出して翻訳モデルを構築し、キャラクタらしく発話を書き換える手法[5]を提案している。宮崎らは、発話の文節の機能語をキャラクター性に応じて自動的に書き換える手法[3]を提案しているが、キャラクタに相応しくない名詞が発話に残り違和感が生じてしまう。そのため、発話内の名詞をキャラクタに応じた名詞に置換する手法が必要となる。

本稿では、対話シナリオの発話に含まれる名詞を作成したい年代・性別のキャラクタに合わせた名詞に置換する手法を提案する。具体的には、年代・性別といったキャラクター性を指定すると、指定したキャラクタと同じ属性を持つ Twitter ユーザが投稿したツイートを用いて、自動的に発話に含まれる名詞の置換を行う。

提案手法において年代と性別の観点で発話にキャラクター性が現れているか確認をするため、被験者に異なる

2 キャラクタの発話がそれぞれ含まれる 2 文書を提示し、キャラクタを判別させる評価実験を行なった。実験により、提案手法は文節の機能語のみを変更する既存の手法[3]と比較し、正解率が有意に高くなることが確認された。

本稿の構成は以下の通りである。2 章で関連研究について述べる。3 章では提案手法を述べ、4 章では提案手法の評価について述べる。5 章では考察を述べ、6 章では結論と今後の課題を述べる。

2. 関連研究

キャラクタの発話を新規に作成する研究として、Higashinaka らはユーザがマリーアントワネットといった特定の人物になりきって、ウェブサイト上で質問応答のやり取りを行い、特定人物の発話を収集する手法を提案している[4]。

キャラクター性を考慮して発話の書き換えを行う研究として、水上らはキャラクタ用として作られた Twitter 上のツイートなどといったコーパスにおける文から特徴的な言語表現を抽出してキャラクタの翻訳モデルを構築し、発話に含まれる機能語をキャラクタらしく書き換える手法を提案している[5]。被験者実験では、翻訳モデルを用いて発話をキャラクタらしく書き換えることで、発話に対するキャラクタらしさが向上している。宮崎らは、発話に含まれる文節の機能語を年代・性別・ユーザとの親密度といったキャラクター性に応じた機能語に置換し、発話を自動的に書き換える手法を提案している[3]。例えば、「西野カナが好きだよ」という 20 代女性用に作られた発話内の「だよ」という機能語を非親密な 40 代男性用に「です」に置換し、「西野カナが好きです」という発話に書き換える。

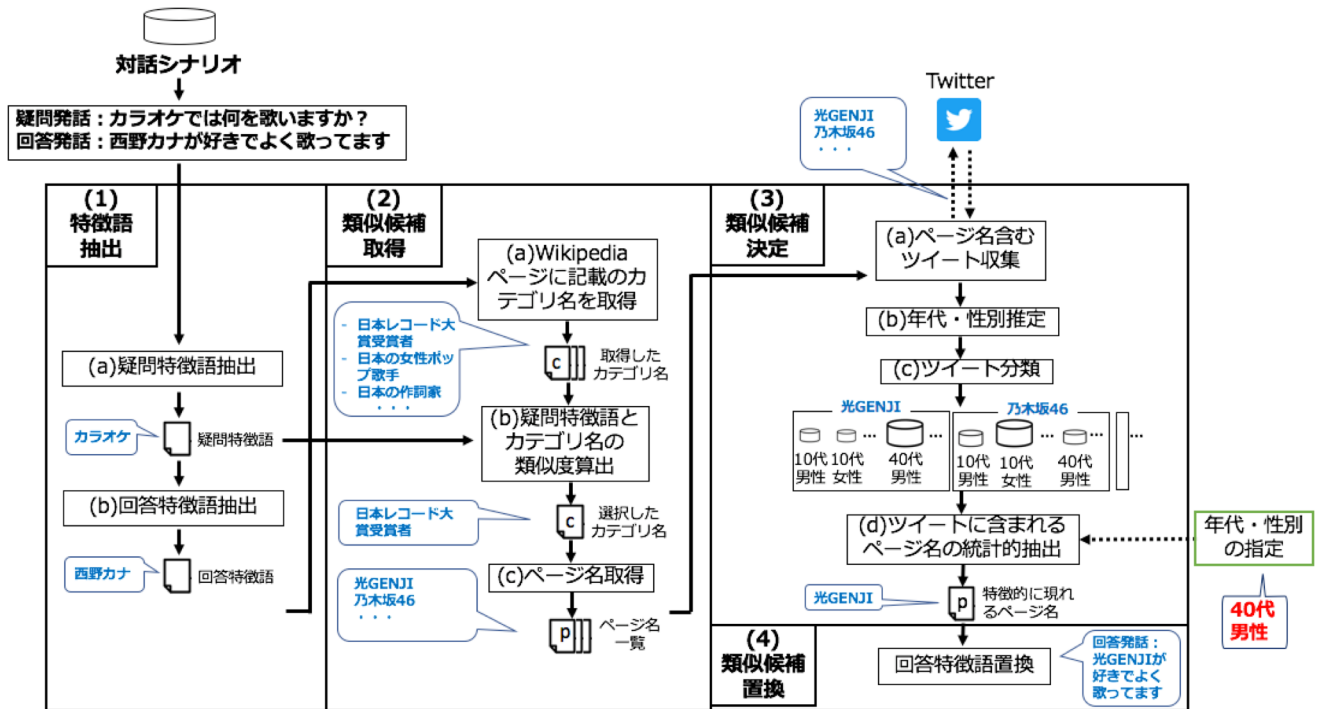


図 1：提案手法の概要

しかし、この手法では書き換えた発話に違和感が残る課題がある。例えば、40代男性用に書き換えられた「西野カナが好きです」という発話には「西野カナ」という名詞が含まれているが、「西野カナ」は若年女性に人気の歌手であるため、40代男性のキャラクター性を特徴付ける発話としては相応しくない。そのため、発話内の名詞をキャラクターに合わせた名詞に置換する手法が必要となる。

3. 提案手法

3.1. 提案手法の概要

2章で述べた課題を解決するために、対話シナリオ内の「何」や「どんな」といった疑問詞を含む疑問文の発話（以下、「疑問発話と定義」）に対して回答となる発話（以下、「回答発話と定義」）に含まれる名詞を作成したい年代・性別のキャラクターに応じた名詞に置換する手法を提案する。図1は提案手法の概要を示している。

始めに、表1に示すような対話シナリオを用意する。対話シナリオは、ユーザ想定発話とキャラクタの発話のペアで構成されている。対話時、ユーザの発話に類似するユーザ想定発話を選び、その発話に対応するキャラクタの発話を選択することで、ユーザの発話に対するキャラクタの応答を生成できる。ユーザ想定発話中の疑問発話に対応するキャラクタの発話を回答発話として抽出する。例えば、表1のユーザ想定発話にお

表 1：対話シナリオの例

ユーザ想定発話	キャラクタの発話
カラオケでは何を歌いますか？	西野カナが好きでよく歌ってます
夜に炒飯食べようかな	炒飯いいですね！
どんなスポーツをしますか？	ウォーキングとかそんなレベルですね
今日買い物しますか？	はい。しますよ。
...	...

ける「カラオケでは何を歌いますか？」が疑問発話であり、その発話に対応する「西野カナが好きでよく歌ってます」を回答発話として抽出する。

次に、以下の4ステップで回答発話の書き換えを行う。(1)特徴語抽出では、回答発話に含まれる疑問発話の回答となる名詞（以下、「回答特徴語」と定義）を抽出する。(2)類似候補取得では、回答特徴語の類似候補を収集するため、Wikipediaの回答特徴語に関するページに記載されているカテゴリ名を1つ選択し、選択したカテゴリ名に含まれるページ名を取得する。(3)類似候補決定では、特定のキャラクターが回答しそうな候補を決定するため、取得したページ名の中から、特定のキャラクターと同年代・性別のTwitterユーザが投稿したツイートに特徴的に出現するページ名を1つ決定する。(4)類似候補置換では、回答特徴語を決定したページ名に置換する。

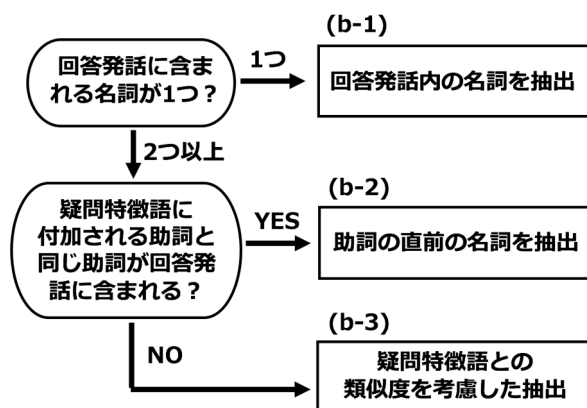


図 2：回答特徴語を抽出する手法の概要

表 2：回答特徴語の抽出例

番号	疑問発話/抽出した疑問特徴語	回答発話	回答発話に含まれる名詞	抽出した回答特徴語
b-1	カラオケでは何を歌いますか? / カラオケ	西野カナをよく歌っています	西野カナ	西野カナ
b-2	どんな食べ物が好きですか? / 食べ物	近所にある店の和食が好きですね	近所店 和食	和食
b-3	何か好きなスポーツあるんですか? / スポーツ	バドミントンをやってたのでテレビでやってると観ちゃいますね。	バドミントン テレビ	バドミントン

3.2. 特徴語抽出

回答発話から回答特徴語を取得する。図 2 は、回答特徴語の抽出手順について示している。表 2 は、各抽出手法で回答特徴語を抽出した例を示している。

(b-1)回答発話に含まれる名詞が 1 つである場合、その名詞が疑問発話の回答である可能性が高いため、回答特徴語として抽出する。表 2 の b-1 における回答発話に含まれる名詞は「西野カナ」のみであるため、「西野カナ」を回答特徴語として抽出する。(b-2) 回答発話に含まれる名詞が 2 つ以上の場合、どの名詞を置換するか決定する必要がある。疑問発話の回答となる名詞は、疑問詞の係り受けの関係にある名詞(「疑問特徴語」と定義)に付加される助詞と同じ助詞の直前に位置する傾向があるため、その名詞を抽出する。(b-3)疑問特徴語と関連性の高い名詞が回答である可能性が高いため、疑問特徴語と回答発話に含まれる各名詞の類似度を算出し、回答特徴語を抽出する。

3.2.1 疑問詞と助詞を考慮した抽出

疑問発話の回答となる名詞は、疑問特徴語に付加される助詞と同じ助詞の直前となる傾向があるため、その名詞を抽出する。例えば、表 2 の b-2 における疑問発話内に「どんな食べ物が」という単語群が含まれており、疑問特徴語である「食べ物」に付加される助詞は「が」である。更に、回答発話に「が」という助詞が含まれているため、直前の名詞「和食」を抽出する。

3.2.2 疑問特徴語との類似語を考慮した抽出

疑問特徴語と回答発話に含まれる各名詞の類似度を算出し、類似度が最も高い名詞を回答特徴語として抽出する。類似度の算出の際には、Wikipedia の全ページのテキストを形態素に分割したものを学習した fastText[6]を用いた。例えば表 2 の b-3 では、疑問特徴語の「スポーツ」と、回答発話内の名詞の類似度をそれぞれ算出し、最も類似度の高い「バドミントン」を回答特徴語として抽出する。

3.3. 類似候補取得

回答特徴語の類似候補を収集するため、回答特徴語に関する Wikipedia のページに記載のカテゴリに含まれるページ名を取得する。

ページ名の取得は以下の手順で行う。(a)Wikipedia で回答特徴語に関するページを検索し、ページ内に記載されているカテゴリ名を全て取得する。(b)疑問発話に対し回答として相応しいページ名を入手するため、疑問特徴語と取得した各カテゴリ名をそれぞれ比較し、類似度が最も高いカテゴリ名を選択する。具体的には、疑問特徴語と各カテゴリ名の類似度を fastText を用いて算出し、類似度が最も高いカテゴリ名を 1 つ選択する。(c)カテゴリに含まれるページ名を取得し、回答特徴語の類似候補とする。

3.4. 類似候補決定

作成したいキャラクターが回答として発話しそうな言葉を見つけるため、3.3 節で取得したページ名の内、作成したいキャラクターと同年代・性別の Twitter ユーザーが投稿したツイートに特徴的に出現するページ名を 1 つ決定する。ページ名の決定は以下の手順で行う。(a)3.3 節で取得した各ページ名を含むツイートを収集する。(b)文献[7]の手法を用いて、収集した各ツイートを投稿したユーザの年代・性別を推定する。(c)推定したツイートを年代・性別ごとに分類する。(d) AIC (赤池情報量基準)を用いて、各ページ名が、指定した年代・性別のツイート群の中に高頻度且つ偏って出現する度合いを算出し、度合いが最大となるページ名を 1 つ決定する。

表 3：評価値と評価項目

評価値	評価項目
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

表 4：回答発話の日本語評価における平均評価値

	文体	提案
10 代男性	3.32	4.05
20 代男性	3.76	3.94
30 代男性	3.85	3.99
40 代男性	3.48	4.05
10 代女性	3.95	3.58
20 代女性	3.61	3.77
30 代女性	3.53	3.95
40 代女性	3.97	3.78
平均	3.68	3.89

4. 提案手法の評価

4.1 回答発話における日本語評価

4.1.1 実験概要

発話を書き換えた後に、発話内に日本語として違和感が残ると、対話時に、ユーザはキャラクタに対して不快感を抱く可能性がある。そこで、提案手法によって書き換えを行なった発話が日本語として違和感がないか評価を行うため、発話に含まれる文節の機能語を年代・性別に相応しい機能語に置換する手法[3]（以下、「文体手法」と定義）と比較し、被験者実験を実施した。具体的には、疑問発話と書き換えを行なった回答発話から構成される対話を被験者に提示し、「回答発話が日本語として違和感がないか」という問い方で、対話毎に表 3 の評価値をアンケートに記入させる。計 8 キャラクタ（性別に関して男性と女性の 2 種類、年代に関して 10 代～40 代の 4 種類の組み合わせ）を実験の対象とし、3 名の被験者に評価値を回答させる。

20 代女性のキャラクタとして作成された対話コーパス[8]からランダムに疑問発話と回答発話のペアを 40 ペア抽出し、実験に使用した。3.4 節で説明したツイートに関しては、2018/6/1 から 6/7 の期間に投稿されたものを使用する。

4.1.2 実験結果

表 4 は、文体手法と提案手法におけるキャラクタ毎の被験者 3 名による平均評価値を示している。全キャラクタの平均に関して、提案手法は文体手法よりも上

対話 1	Q.何かスポーツはしていますか？ A.綱引きが好き	対話 1	Q.好きなスポーツは何ですか？ A.フィギュアスケートが好き
対話 2	Q.どんな料理が好きですか？ A.オムライスが好きだよ	対話 2	Q.どんな料理が好きですか？ A.ハンバーグが好きだよ
...		...	
対話 10	Q.何の音楽を聴きますか？ A.佐香智久が好きだね	対話 10	Q.何の音楽を聴きますか？ A.横原敬之が好きだね

図 3：キャラクタ性判別実験で使用する文書の例

回ることが分かった。評価値に関する有意差を Mann-Whitney の U 検定により分析した。有意水準 1 (%) で、文体手法と提案手法の間で有意差が認められた ($Z=2.925$)。これらの結果から、提案手法で書き換えた回答発話において、被験者は文体手法よりも日本語として違和感を感じなかったことが示された。

4.2 キャラクタ性評価

4.2.1 実験概要

提案手法で書き換えを行なった回答発話が、年代と性別の観点でキャラクタ性が現れているか評価をするために、文体手法と比較し被験者実験を実施した。疑問発話と書き換えを行なった回答発話から構成される対話を含むキャラクタ性が異なる 2 文書に対して、二択でキャラクタ性を判断させる実験である。4.1.1 項と同様の 8 キャラクタを実験の対象とし、3 名の被験者にキャラクタ毎に書き換えた文書を比較させる。

図 3 は、疑問発話と、提案手法を用いて 10 代女性・40 代女性用に書き換えた回答発話から構成される対話が 10 対話ずつ含まれた 2 文書の例を示している。左の文書が 10 代女性用、右の文書が 40 代女性用に書き換えられた文書である。年代の違いが回答発話に現れているかどうか評価を行う際には、図 3 の 2 文書を被験者に提示し、どちらの文書が 10 代用に書き換えられたか、10 対話から総合的に判断させ、回答させる。

4.1.1 項で述べた対話コーパスからランダムに疑問発話と回答発話のペアを 120 ペア抽出し、実験に使用した。120 ペア中 10 ペアを 1 文書とし、計 12 文書を用意する。3.4 節で説明したツイートに関しては、2018/6/1 から 6/7 の期間に投稿されたものを使用する。

4.2.2 実験結果

表 5 は、文体手法と提案手法における比較対象毎の被験者 3 名による正解率の平均、被験者 3 名中 2 名ずつの一致率とカッパ値の平均を示している。正解率は、被験者の回答と正解が一致した割合、一致率は被験者 2 名の判定が一致した割合、カッパ値は被験者 2 名の偶然によらない一致による割合を示している。

実験結果より、正解率、一致率、カッパ値の全比較対象の平均に関して、提案手法は文体手法よりも上回ることが分かった。正解率に関する有意差を 2 つの比

表 5：キャラクター性判別評価における正解率，一致率，カップパ値

	正解率		一致率		カップパ値	
	文体	提案	文体	提案	文体	提案
男性 vs 女性	0.60	0.78	0.74	0.75	0.45	0.28
10 代 vs 20 代	0.53	0.78	0.47	0.69	-0.05	0.14
10 代 vs 30 代	0.57	0.67	0.53	0.58	0.12	0.09
10 代 vs 40 代	0.63	0.74	0.44	0.64	-0.13	0.07
20 代 vs 30 代	0.75	0.53	0.72	0.69	0.26	0.39
20 代 vs 40 代	0.78	0.72	0.67	0.78	0.06	0.45
30 代 vs 40 代	0.63	0.69	0.67	0.56	0.29	-0.04
平均	0.64	0.70	0.61	0.67	0.14	0.20

率の差の検定により分析した．有意水準 1(%)で，文体手法と提案手法の間で有意差が認められた ($X^2=7.623$)．一致率に関する有意差を 2 つの比率の差の検定により分析した．有意水準 5(%)で，文体手法と提案手法の間で有意差が認められた ($X^2=4.613$)．これらの結果から提案手法で書き換えた回答発話を含む文書において，被験者は文体手法よりもキャラクター性を容易に判別できることが示された．

5. 考察

5.1 日本語評価

提案手法によって書き換えた回答発話の内，評価値の低かったものを分析した所，回答特徴語の置換候補として不適切な名詞が置換されている発話が多かった．例えば，「何かスポーツしてたりとかするんですか？」という疑問発話において，「水着をやってみました」という発話が見つかった．書き換え前の回答発話「水泳をやってみました」から回答特徴語として「水泳」が抽出されたが，「水泳」の Wikipedia ページに記載されている「体育」カテゴリに含まれる「水着」のページ名が置換候補として決定されたため，日本語として違和感が生じている．この問題を解決するために，Twitter などといった SNS の投稿を利用して書き換え後の回答発話に含まれる単語の共起関係を調べ，共起しないと判断された場合は，置換候補としてのページ名を候補から外すといった手法が考えられる．

5.2 キャラクター性評価

書き換えを行なった回答発話を分析すると，文体手法では年代間で発話にキャラクター性が現れており，提案手法では年代と性別の両方において発話にキャラクター性が現れているものが多く見られた．表 6 は，文体手法と提案手法によって書き換えたキャラクター毎の回答発話の例を示している．文体手法では，10 代～20 代における文末表現が平常語，30 代～40 代における文末表現が敬語に書き換えられており，年代に関して発話にキャラクター性が現れていると言える．一方で，提

案手法では，10 代，40 代のキャラクター性を特徴付ける「綱引き」「ゴルフ」といった名詞が 10 代女性，40 代男性の回答発話に含まれ，更に，男性のキャラクター性を特徴付ける「ボクシング」「サッカー」「ゴルフ」といった名詞が男性の回答発話に含まれており，年代と性別の両方に関して発話にキャラクター性が現れている．

提案手法において，書き換えを行なった回答発話や文書を分析すると，3 つの問題点が見つかった．表 5 の 20 代と 30 代の正解率が文体手法よりも大幅に下回った点，カップパ値が 0.20 と低い値になった点，片方の文書中にもう片方のキャラクター性を特徴付ける回答発話が含まれていた点である．それぞれの原因と対策について考察する．

表 5 の各比較対象の正解率の結果を見ると，20 代と 30 代，20 代と 40 代の比較において提案手法は文体手法を下回ってしまった．特に，20 代と 30 代の比較では，0.22pt と大幅に下回ったため，原因を分析した．提案手法で書き換えを行なった回答発話を確認した所，全 7 比較対象における 2 キャラクターの回答発話の組み合わせの内，23%が同じ名詞に置換されていたことに対し，20 代と 30 代の比較における 2 キャラクターの回答発話では，40%が同じ名詞に置換されており，発話にキャラクター性が出せなかったことが原因であると考えられる．この問題を解決するために，回答特徴語の類似候補を更に多く取得する仕組みを構築し，キャラクター性を際立たせる必要がある．例えば，回答特徴語の類似語を fastText といった類似語取得器を用いて取得し，取得した類似語の Wikipedia のページに記載されているカテゴリ内のページ名を取得する手法が考えられる．

カップパ値の全比較対象の平均に関して，提案手法は文体手法よりも上回ったが，0.20 と低い一致となり，被験者によって発話に対するキャラクター性の判断が異なることが分かった．例えば，表 6 の 20 代女性の発話に含まれる「テニス」という名詞は，老若男女問わず楽しめるスポーツであり，20 代女性を特徴付ける発話としては相応しくない．発話に対しキャラクター性を際立たせるために，複数のキャラクターに共通して当ては

表 6：文体手法と提案手法によって書き換えた回答発話の例

	疑問発話：何かスポーツをしていますか？（疑問特徴語：スポーツ） 回答発話：バドミントンをしています（回答特徴語：バドミントン）	
	文体	提案
10 代男性	バドミントンをするよ	ボクシングをしています
20 代男性	バドミントンをしたり	サッカーをしています
30 代男性	バドミントンしています	サッカーをしています
40 代男性	バドミントンをしてます	ゴルフをしています
10 代女性	バドミントンをしているよ	綱引きをしています
20 代女性	バドミントンをするよ	テニスをしています
30 代女性	バドミントンをしています	フィギュアスケートをしています
40 代女性	バドミントンもしています	フィギュアスケートをしています

まるような回答特徴語の類似候補を除外する機能を導入することが考えられる。

提案手法で書き換えを行なった発話が含まれる文書ペアの内、被験者の回答と正解が一致しなかったものを解析した所、片方の文書中にもう片方のキャラクター性を特徴付ける回答発話が含まれているものが頻出していた。例えば、「どんなお仕事されてるんですか？」という疑問発話に対して、30 代男性の文書の中に「メイドに就いています」、30 代女性の文書の中に「ホストに就いています」という回答発話が見つかった。30 代男性と推測されたツイートに「メイド」、30 代女性と推測されたツイートに「ホスト」といった言葉が高頻度且つ偏って出現しており、それぞれ「メイド」と「ホスト」の置換候補が決定されたと考えられるが、これらの言葉は疑問発話に対するキャラクターの回答として相応しくない。この問題を解決するために、3.4 節の類似候補決定において、回答特徴語に係る動詞と置換候補となるページ名の組み合わせが、指定した年代・性別のツイート群に高頻度且つ偏って出現する度合いを算出し、ページ名を決定する手法が考えられる。

6. 結論と今後の課題

本稿では、対話シナリオ内の疑問文の発話に対して回答となる発話に含まれる名詞を作成したい年代・性別のキャラクターに応じた名詞に置換する手法を提案した。

本稿の貢献は、以下の 2 点に集約される。1 つは、文体手法と提案手法を用いて被験者実験を実施した所、回答発話の日本語評価において、提案手法は文体手法と比較して有意に評価値が高いことが確認された点である。もう 1 つは、キャラクター性判別評価において、提案手法は文体手法と比較して、正解率、一致率がいずれも有意に高く、カップ値が上回ったことが確認された点である。

今後の課題として、以下の 2 点を挙げる。第一に、キャラクター性を更に際立たせるために、回答特徴語の

類似候補を更に取得する仕組みを構築することである。例えば、回答特徴語の類似語を fastText といった類似語取得器を用いて取得し、類似語の Wikipedia のページに記載されているカテゴリ内のページ名を取得する手法が考えられる。第二に、様々なキャラクターの生成を実現することである。年代・性別といったキャラクター性に基づく対話エージェントだけでなく、有名人や架空人物などといったキャラクターの生成を目指し、これを実現するための手法を考案する予定である。

参 考 文 献

- [1] Wu, Xianchao, et al. "りんな: 女子高生人工知能." 言語処理学会第 22 回年次大会発表論文集：306-309. (2016)
- [2] 井上昂治 et al. "自律型アンドロイド Erica のための音声対話システム." SIG-SLUD 5.02: 21-24 (2015)
- [3] 宮崎千明, et al. "文節機能部の確率的書き換えによるキャラクター性変換." 言語処理学会第 21 回年次大会発表論文集：277-280.(2014)
- [4] Higashinaka, Ryuichiro et al. "Using role play for collecting question-answer pairs for dialogue agents." INTERSPEECH. 2013.
- [5] 水上雅博 et al. "特徴的話者を対象とした言語的個人性変換." 研究報告音声言語情報処理 (SLP) 2014.20: 1-7.(2014)
- [6] Joulin, Armand, et al. "Bag of tricks for efficient text classification." arXiv preprint arXiv:1607.01759 (2016).
- [7] Ikeda, Kazushi, et al. "Twitter user profiling based on text and community mining for market analysis." Knowledge-Based Systems 51: 35-47. (2013)
- [8] Ikeda, Kazushi, et al. Utilizing Crowdsourced Asynchronous Chat for Efficient Collection of Dialogue Dataset, In Proc. of HCOMP 2018.