

# 深層学習を用いた対話文への「かわいさ」の付与

## Attachment of “Cuteness” to Dialogue Sentences Using Deep Learning

後藤 優介<sup>\*1</sup>

Yusuke Goto

砂山 渡<sup>\*2</sup>

Wataru Sunayama

畑中 裕司<sup>\*2</sup>

Yuji Hatanaka

小郷原 一智<sup>\*2</sup>

Kazunori Ogohara

<sup>\*1</sup>滋賀県立大学大学院工学研究科

Graduate School of Engineering, The University of Shiga Prefecture

<sup>\*2</sup>滋賀県立大学工学部

School of Engineering, The University of Shiga Prefecture

Nowadays, automatic generation of dialogue sentences using deep learning has been conducted. Automatic generation systems of dialogue sentences require not only sequences of conversation but also characteristics of conversation partner. Because, it makes users feel charms and personalities in conversation. However, previous studies often transformed the sentence's characteristics to another one which easy to define transformed vocabularies, for example, the dialect and the style of samurai. Therefore, this study will transform the sentence's characteristics to another one which hard to define transformed vocabularies. Especially, the purpose of this study is that give the cuteness to sentences. We considered how to give the cuteness to sentences, and evaluate sentences that can give the cuteness or not.

### 1. はじめに

近年, Apple 社の Siri や Amazon 社の Alexa を皮切りに, 対話型システムの注目度が増している. これらのように, 特定のタスクを実行する事を目的とした対話型システムの事をタスク指向型対話システムと呼ぶ. 一方で, SoftBank 社の Pepper や Microsoft 社の りんなに代表される, ユーザとコミュニケーションを取る事を目的とした対話型システムの事を非タスク指向型対話システムと呼ぶ. これは主に, コミュニケーションロボットでの活躍が期待され, AI と対話できる事で, より AI に対して親しみを持ってもらう事が期待できる [1]. こういった非タスク指向型対話システム固有の問題として, 無個性な対話による対話の質の低下が挙げられる. 対話の応答文が無個性で機械的なものであると, ユーザに AI との間に壁を感じさせる. それにより, ユーザの対話に対する意欲が失われる事や, AI に対して親しみを感じる事が難しくなる事が懸念される.

そこで本研究では, 付与するキャラクター性として, 具体的なキャラクターではなく, 人間の感性の 1 つである「かわいさ」を挙げた. そして, コーパスの調査に加え, 深層学習によって学習されたモデルの学習内容を分析する事で, かわいさが現れている文にどのような特徴があるかを網羅的に調査した. その後, 調査で得られた要素を自動で付与する事で, 文にかわいさを付与するアルゴリズムを作成した. また, そのアルゴリズムに, 深層学習による判定を取り入れる事で, ユーザによってどれくらい強くかわいさを付与するかを調整できるようにした. これにより, 従来手法と比べて, ユーザの好みに沿った強さのかわいさを, 文中の幅広い範囲で付与できる事が期待される.

### 2. 関連研究

#### 2.1 文に現れるキャラクター性とそれが現れる場所を調べた研究

文のどの部分にキャラクター性を感じさせる要素を埋め込む自由度があるかを調べた研究 [2] がある. この研究では, 内容表現 (名詞, 動詞, 形容詞, 副詞など) と機能表現 (助詞, 助動詞, 接続詞など) の両方にキャラクター性を埋め込む自由度が

あると結論付けている. しかし, 内容表現は非常に自由度が高く, 単純な置換は難しいとしている.

一方で, 機能表現は同義表現への置き換えが比較的簡単であり, キャラクター性を容易に埋め込む事ができるとしている. 本研究でも, 単語の同義表現への置換を扱う事になるので, 比較的置換が容易とされていた機能表現は優先的に取り入れた. 同時に, 内容表現内でも比較的自由度が低いとされた形容表現にも着目した.

#### 2.2 文にキャラクタ付けを行う研究

文節機能部に着目し, 人手で書き換えたコーパスを分析する事で見つけ出した置換ルールを, 確率に乗っ取り自動適用する研究 [3] がある. これにより, 変換により感じる違和感を 10% 以下に抑えつつ, 人間が 80% の確率で判別できるキャラクター性変換を行う事に成功している. また, よりコストを抑えた変換を行うため, 変分オートエンコーダと注意機構を用いてキャラクター性を付与する研究 [4] がある. この研究では, 注意機構を用いる事で, 生成モデルを用いる場合でも変換する単語を制御しつつ, 文全体の変換を行う事ができるという事を示している.

しかし, これらの研究における変換は, 機能表現のみを対象とした置換に限られており, キャラクター性を感じさせる要素を網羅的に適用できているとは言い難い. また, これらの研究では「関西弁」「侍風」といった具体的なキャラクター性の付与に留まっており, キャラクター性の付与が直接魅力的な対話につながっているとは言えない. そこで, 本研究では, より人間の感性に近いキャラクター性である「かわいさ」を付与する場合について調査を行った. 調査するコーパスとして人手による変換を用いていない既存のものを選び, 要素を網羅的に調査する事で, 自動適用できる変換ルールを調査した.

### 3. 「かわいさ」を付与する変換ルール

#### 3.1 「かわいさ」の定義

本節では, 本研究で付与するキャラクター性である「かわいさ」について定義する. 「かわいい」という単語は, 辞書的には「愛らしい, 小さい」という意味がある. しかし, 現代日本において, 「かわいい」という単語は様々な意味で用いられるようになっており, 物理的・数値的特徴からは明確に分類する

事ができない [7]. その人個人の感性によって、こういったものが「かわいい」のかは変わってくる。

「かわいい」の意味について調べた研究 [5][6][7] によると、心理学的に人が他者に対して「かわいい」と感じる要素は、大きく「共感性」「不完全性」「親和性」の3つの要素に集約できると言われている。それぞれの用語の定義は以下の通りである。

- 「共感性」: 他者の状況や感情に共感し、他者志向の温かい気持ちを持たせる性質。
- 「不完全性」: 幼さ・弱さから、守ってあげたいという気持ち (保護欲) を感じさせる性質。
- 「親和性」: 相手と親密な関係を維持し、仲良くなりたいと感じさせる性質。

このうち、共感性と親和性については、自分と他者の関係性を評価して生まれる性質であり、相対的なものであるため、統一した基準での定義が難しい。そのため、本研究では不完全性を取り上げ、「文全体から読み手に不完全性を感じさせる性質」を「かわいさ」と定義する。また、外見の要素を含まない不完全性の要素として、「幼さ (未熟さ)」と「精神的な弱さ (気の弱さ)」の2つを挙げ、それにつながる要素を調査していく。

本研究では、「かわいさ」を含む学習用のテキストデータとして、ソーシャルゲーム「アイドルマスターシンデレラガールズ」の登場人物 183 人のうち、「正統派かわいい」キャラクターとして位置づけられている 67 名のセリフを合計 50,113 件 [12] を用いることとした。

また、「かわいさ」を含まない学習用のテキストデータとして、老若男女 199 人の日常会話を録音し、それを文字起こしたデータセット「名大会話コーパス」[9][10][11] に収録された発話文のうち、20 文字以上のものを 55,054 件を用いることとした。

### 3.2 「かわいさ」を付与する変換ルール

「かわいさ」を付与する変換ルールの作成に際して、前節で用意したデータをもとに、深層学習により「かわいさ」を含む文と含まない文とに分類する学習を行った。その結果を考察することで、変換ルールを用意した。

#### 3.2.1 「幼さ, 未成熟さ」についての変換ルール

幼さや未成熟さを感じさせる表現として、以下の要素がある事がわかった。

- 幼さや柔らかさを感じさせる表記や表現を使う。
- 相手との距離を縮め、相手を頼ろうとする表現を使う。
- 感情を表に出し、裏表を感じさせない表現を使う。
- 敬語の使用は最低限にする。

これらの要素をもとに、分類に寄与する単語を調べた上で、表 1 に示す「幼さ, 未成熟さ」を付与する変換ルールを作成した。

#### 3.2.2 「精神的な気の弱さ」についての変換ルール

精神的な気の弱さを感じさせる表現として、以下の要素がある事がわかった。

- 自分の意見を強く主張できず、断定出来ない事を感じさせる表現を使う。
- 自分の考えを相手に強制できない事を感じさせる表現を使う。

表 1: 「幼さ, 未成熟さ」を付与する変換ルール

変換ルール	変換対象の単語	変換後の単語
名詞をひらがなに置換	漢字のみ名詞 (10%の確率で適用)	その単語をひらがなに置換
副詞をオノマトペに置換	体言か名詞の前にある一般副詞 (30%の確率で適用)	評価の高い副詞 27 件のうち 1 つと置換
相触・同意・終助詞長音化	相触や同意を示す表現 (うん, なるほど等) または終助詞	対象の単語の後に「-」を挿入
語尾表現置換	です, ます等の丁寧な文末表現とその前の単語 文末記号の前の, 名詞や副詞では無い単語	よ, ねに置換もしくは挿入 必要に応じて前の単語も活用させる
感動詞の挿入・置換	感動詞 文頭・文末 (50%の確率で適用)	評価の高い感動詞 (5 件) のうちランダムで 1 つと置換もしくは挿入
名詞にさんをつける	一般名詞もしくは固有名詞 (15%の確率で適用)	対象の単語の後に「さん」を挿入

表 2: 「精神的な気の弱さ」を付与する変換ルール

変換ルール	変換対象の単語	変換後の単語
否定表現を婉曲表現に置換	言い切る表現 (~ない等)	対象の単語の後に「かも」を挿入
曖昧な表現を追加	文末の動詞 (50%の確率で適用)	対象の単語の後に「かも」を挿入
命令をお願いする表現に置換	命令形である動詞	対象の単語を基本形に置換し、「~てほしいな」「~よ」を挿入
逆接の接続詞を削除	逆接の接続詞 (しかし, だけど等)	対象の単語を削除, 必要に応じて句読点を挿入
強い否定を緩和する表現の追加	強い否定を表す副詞や形容動詞 (ちっとも, 全く等)	「多分」「できれば」と置換

- 相手の意見に対して、強く否定できない事を感じさせる表現を使う。

これらの要素をもとに、分類に寄与する単語を調べた上で、表 2 に示す「精神的な気の弱さ」を付与する変換ルールを作成した。

## 4. 深層学習を用いて「かわいさ」を付与するシステム

本章では、前節で用意した変換ルールをもとに、対話文に「かわいさ」を付与するシステムについて述べる。「かわいさ」を付与するシステムのフローチャートを図 1 に示す。以下で、フローチャート内の「かわいさ」の評価システムについて述べる。

### 4.1 深層学習による「かわいさ」の評価

本システムにおける「かわいさ」の評価 (以降スコアと呼ぶ) は、対話文を 3.2 節で構築した深層学習に入力して分類した時の、SoftMax 関数により出力される値  $E$  (0 から 1 の実数値) を 100 点満点に換算した値に、変換ルールが適用された回数  $n$  の 5 倍の数値を加算した式 1 の値を用いて行われる。ただし、100 点を超えた時には 100 点として扱う。

$$S = 100E + 5n \quad (0 \leq S \leq 100) \quad (1)$$

深層学習においては、この値が 50 点を超えると「かわいい」文として判定されるが、本研究では文の分類が目的ではなく、文に「かわいさ」を付与することを目的とするため、目標スコアを与え、このスコアを超えるまで前章で用意した変換ルールの適用を試みる。

### 4.2 「かわいさ」を付与する変換ルールの適用

表 1 と表 2 に示した 11 個の変換ルールのうち、まだチェックを行っていないルールについて、変換対象となる単語があるかをチェックする。すなわち、文中の各単語について、変換対象の単語かどうかをチェックし、変換対象の単語が存在する場合、その変換ルールの適用処理を行う。変換後に目標となるかわいさスコアを超えていた場合には変換ルールの適用を終了し、目標スコアを超えていない場合は、次のルールを適用する。なお、チェックの順番は、名詞をひらがなに置換、感動詞の挿入、曖昧な表現を追加以外の 8 個のルールについて乱数を用いてランダムに決定する。その 8 つのルールについての

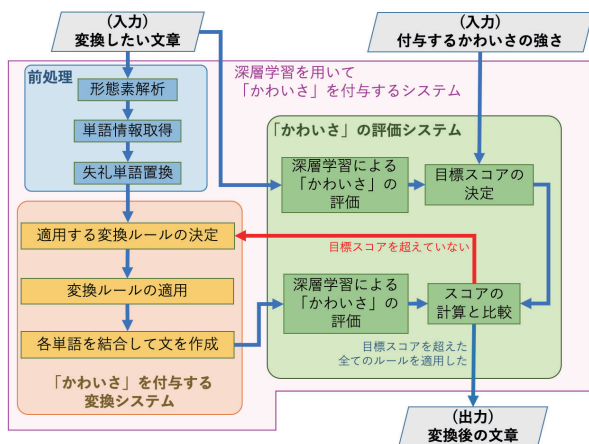


図 1: 「かわいさ」を付与するシステムのフローチャート

表 3: 変換ルールの適用で「かわいい」と答えた人数の増減

	文数 (/944)
「かわいい」と答えた人が増えた文	474
「かわいい」と答えた人が減った文	111

チェックが終了した後、曖昧な表現を追加→名詞をひらがなに置換→感動詞の挿入の順にルールを適用する。

## 5. 「かわいさ」を付与するシステムの検証実験

### 5.1 実験設定

大学生・大学院生 16 名に 3 章で用意したデータの一部（「かわいさ」を含まない文 1000 件を学習データとは別に用意して、全く変換が行われない文を除いた 944 件）を用いて、無作為に分けた変換ルールの適用前と適用後のいずれかの文について「各文がそれぞれ自分に対する発話であった場合、その文を発した人に対してかわいいと感じるかどうかが」を回答してもらった。

### 5.2 実験結果と考察

#### 5.2.1 変換ルールの適用前後で「かわいい」と答えた人数の比較

変換ルールの適用により「かわいい」と答えた人数の増減表 3 に示す。約 50% の文で「かわいい」と回答する人の数が増えており、一定の効果があったことが確認できる。

変換後の文のうち、特にかawaiiと判定した人数が増加した文の例を表 4 に示す。これらの文は、変換前の文はかわいいと答えた人数は 0 人だが、変換後は 6 人もしくは 7 人まで増加していた。一方で、変換後にかawaiiと判定した人数が減少した文の例を表 5 に示す。これらの文は、変換前の文はかわいいと答えた人数は 3 人または 4 人であるが、変換後は 0 人であった。

変換によってかわいくならなかった 470 件の文について、その原因を著者が分類した結果を表 6 に示す。

もとの文がネガティブな内容の話、説教や教訓を含む話、若い人があまりしなさそうな内容を含む話という 3 つの要因で、かわいくならなかった文全体の約 42% を占めている。そのため、文の内容によっては、どのような変換を行ってもかわいさを付与できない可能性が考えられる。

表 4: 変換後の文のうち、特にかawaiiと判定した人数が増加した文の例

変換前の文	変換後の文
結構、周りがワイワイガヤガヤしても寝れるんです。	ふふふ、結構、周りがワイワイガヤガヤしても寝れるね。
来年までいる気なわけだから、だから、ぎりぎりでしょう？	うふふ、来年までいる気なわけだからね。だから、ぎりぎりよね？
何とかっていうバンドのボーカルだったんだよ。	何とかっていうバンドのボーカルさんだったんだよー。

表 5: 変換後の文のうち、特にかawaiiと判定した人数が減少した文の例

変換前の文	変換後の文
私は自分のを取ったら、手をひぎの上に乗せて待ってます。	わたしは自分のを取ったら、手をひぎの上に乗せて待ってよ。
それもほんとに昔行っただけども、高校ぐらいのときに。	それもつやつしたむかし行っただけども、高校ぐらいのときにね。
いやでも、ヒジキとか、ちゃんとしてあげようと思った。	いやでも、ヒジキとか、ほかほか作ってあげようと思った。

#### 5.2.2 変換ルールごとのかわいさと判定した人数の変化

各変換ルールの適用が、読み手に与える「かわいさ」という印象にどのように寄与しているかを調べるため、3.2 節で述べた各変換ルールごとに、かわいさと判定した人数の変化をまとめた。各変換ルールごとの「かわいさ」と答えた人数の変化を表 7 に示す。全ての変換ルールについて、変換ルール適用後の文のほうがかわいさと判定した人数は増加していた。

特にかawaiiと答えた人が増加した文が多いルールとして、感動詞の追加や置換をするルールと、命令をお願いする表現に置換するルールがある。これらのルールは、適用した文の約 6 割の文でかわいさと答えた人数が増加しており、比較的有用なルールであるといえる。かわいさと答えた人数が大きく増加した理由として、これらのルールが文のイメージを大きく変える効果を持っているという事が考えられる。感動詞を追加したり置換するルールの場合、本システムで付与する感動詞はにこやかな笑いや微笑みを感じさせるものであり、文全体に柔らかな印象を与えることができる。また、命令をお願いする表現に置換するルールの場合、命令口調というわかりやすくきつい印象を与える表現を無くし、お願いするという柔らかい印象を付与することで、きつい印象の文を一気に柔らかい印象の文に変えることができる。これらのように、文の印象に明確な変化を与えることができるため、柔らかさや気の弱さといった面から「かわいさ」につながったと考えられる。

一方で、否定表現を緩和するというルールと、断定表現を婉曲表現に変換するルールについては、ルールを適用後にかawaiiと答えた人数が増加した文は約 34%~37% であり、約 45% の文ではかわいさと答えた人数が変わらなかった。これは、強い否定表現を弱い否定表現に置換しても結局否定している事には変わらず、文全体が読み手に与える印象としてはあまり変化がなかったからだと考えられる。同様に、断定する表現を遠回しな婉曲表現に置き換えても、結局内容を肯定している事には変わらないため、文の印象に変化がなかったといえる。しかし、これらのルールでもかわいさと答えた人が増加している文は存在しているので、印象が大きく変わらないからといって不要なルールであるとは一概には言い切れない。

#### 5.2.3 事後アンケートから見るかわいさ文の傾向

事後アンケートとして「どういった文がかawaiiと感じたか」「どういった文がかawaiiと感ぜなかったか」について回答してもらった。特に用意した変換ルールとは異なる要素について、かわいさと感じた文についての回答を以下に挙げる。

- ゆるい内容の世間話的な会話
- 食べ物の話が含まれる文
- 全体的に柔らかく、ポジティブな印象を与える文
- 日常のふとした感じの会話

表 6: 変換によってかわいくならなかった原因の分類

かわいくならなかった原因	該当する文数 (/470)
ネガティブな内容の文	105
前後の会話が無いと理解し辛く、違和感を感じる文	90
難しい話題や説明的な内容を含む文	76
昔話やお酒の話など、若い人があまりしなさそうな話題の文	25
お金の話題についての文	14
変換前後であまり印象が変わらない文	14
教訓や説教と言った、相手に言い聞かせるような内容を含む文	7
その他(上記のどれにも当てはまらないもの, 原因不明のもの)	139

表 7: 変換ルールごとの「かわいい」と答えた人数の増減 (%)

ジャンル	ルール	増加	減少
幼さ, 未成熟さ	名詞をひらがなに置換	50.88	12.87
	副詞をオノマトベに置換	48.28	12.07
	相槌・同意・終助詞長音化	48.94	8.51
	語尾表現置換	53.93	9.63
	感動詞の挿入・置換	64.40	8.20
精神的な気の弱さ	名詞にざんを付ける	43.40	11.32
	断定表現を婉曲表現に置換	37.84	20.27
	曖昧な表現を追加	55.56	16.67
	命令をお願いする表現に置換	61.29	9.68
	逆接の接続詞を削除	47.44	12.82
	強い否定を緩和する表現の追加	34.78	19.57

- 「うふふ」などの笑い声に準ずる言葉が含まれた文

また、かわいくないと感じた文についての回答を以下に挙げる。

- 内容が硬く、難しく感じる会話
- 会話の前後がよくわからず、内容が意味不明な文
- 否定的な文や、同調・共感とは逆の文
- きっぱりした断定表現や、誰かを突き放すような棘のある表現を含む文
- 「うふふ」や「ふふふ」が使われている文

これらを比較すると、5.2.1 項での考察と同様、文の内容から受ける印象がかわいさの判定に大きく寄与しているという事がわかる。また、被験者にかわいくないと判定した文について理由を聞くと、「表現はかわいいと思ったが、文の内容的にかわいさとは感じなかった」という解答が一定数得られた。

#### 5.2.4 考察のまとめ

文におけるかわいさには、「表現から感じるかわいさ」と「内容から感じるかわいさ」の2つの要素があり、それら2つの要素が同時に存在する場合に、読み手はその文から「かわいい」という印象を感じていたと考えられる。本システムでは、「表現から感じるかわいさ」を文に対して付与する事はできるが、「内容から感じるかわいさ」については十分に付与する事ができない。そのため、元々の文の内容が難しい場合やネガティブな内容に触れている場合は、十分なかわいさを付与する事ができないといえる。

「内容から感じるかわいさ」を付与するためには、難しい内容の文やネガティブな内容の文を、ゆるい内容の文やポジティブな内容の文に置き換える必要があり、既存のルールによる変換では対応できない。また、文の内容を書き換えると対話の内容も変わってしまうという問題点も存在する。そのため、本システムの手法とはまた別のアプローチを試みる必要がある。

## 6. おわりに

対話文に対して「かわいさ」を付与するシステムについて述べ、提案システムがかわいさをどの程度付与できるかを実験により検証した。実験の結果、約50%の文に対して一定のかわいさを付与できる事を確認した。かわいさを付与できなかった文については、ネガティブな内容の話や、説教や教訓を含む話、若い人があまりしなさそうな内容を含む話など、表現を変換しても内容からかわいさとは感じない文が多かった。本システムでは「かわいい表現」を付与する事はできるが、「かわいい内容」を付与する事はできないため、難しい文やネガティブな文に対しては十分な「かわいさ」を付与できなかったと考えられる。

今後は、かわいさを付与できる文と付与できない文を区別するとともに、より長い一連の対話文に対してかわいさを付与する量とタイミングを検討していきたいと考えている。

## 参考文献

- [1] 狩野芳伸:「コンピューターに話を通じるか-対話システムの現在」、情報管理学会誌, Vol.59, No.10, pp.658-665 (2017)
- [2] 刀山将大, 佐藤理史, 松崎拓也, 宮崎千明, 平野徹, 松尾義博:「文のどこにキャラクター性を埋め込む自由度があるか」、言語処理学会第22回年次大会発表論文集, pp.721-724 (2016)
- [3] 宮崎千明, 平野徹, 東中竜一郎, 牧野俊朗, 松尾義博, 佐藤理史:「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」、人工知能学会誌, Vol.31, No.1, pp.1-9 (2016)
- [4] 谷川晃大, 藤田寛泰, 壹岐太一:「変分オートエンコーダと注意機構を用いた発話文のキャラクター性変換」、第32回人工知能学会全国大会, 4G2-01 (2018)
- [5] 石川なつ美:「『かわいい』の意味について」、東京女子大学言語文化研究 24, pp.21-35 (2015)
- [6] 家内嘉宏, 入野野宏:「共感性と親和動機による“かわいい”感情の予測モデル構築」、パーソナリティ研究, Vol.23, No.3, pp.131-141 (2015)
- [7] 入野野宏:「“かわいい”感情の心理学モデル」、情報処理学会誌, Vol.57, No.2, pp.128-131 (2016)
- [8] kuromoji: (URL)<https://www.atilika.com/ja/kuromoji/>
- [9] 名大会話コーパス: (URL)<https://mmsrv.ninjal.ac.jp/nucc/>
- [10] 藤村逸子, 大曾美恵子, 大島ディヴィッド義和:「会話コーパスの構築によるコミュニケーション研究」(2011)
- [11] 藤村逸子, 滝沢直宏編:「言語研究の技法:データの収集と分析」、ひつじ書房, pp.43-72 (2011)
- [12] iM@S-CGWordsSearch( $\beta$ ): (URL)<https://icws.indigo-bell.com/>