

親しみやすいチャットボットを目指したエンタメ対話技術

サービスイノベーション部 つのもり 角森 ゆいこ 唯子 しらみず 白水 ゆうたろう 優太郎

近年、さまざまな用途でのチャットボットの需要が高まっている。コミュニケーションの提供を目的としたチャットボットとの対話においては、ユーザに「親しみやすい」「楽しい」と思ってもらうことが重要である。そこで、親しみやすい対話を実現するために、特定のキャラクターを再現する技術と、なぞかけなどの言葉遊び技術を開発した。これにより、チャットボットにエンターテインメント性を付与することが可能になった。さらに、エンタメ対話を提供するツールとして、チャットツール上で動作する対話アプリを開発した。

1. まえがき

近年、スマートスピーカーや対話エージェントアプリなどのチャットボットが普及している。これらは「アラームをかける」や「今日の天気は？」などのような、何かしらのタスク依頼を主に受け付けているものがほとんどである。これらのチャットボットはタスク依頼を受け付けることを念頭に設計されているが、ドコモが2012年にリリースした「しゃべってコンシェル」では、タスク依頼ではないコミュニケーションを目的としたユーザ発話が多く入力されていた。このようなユーザの入力に答えるために、

ドコモではNTTメディアインテリジェンス研究所の成果技術に基づき、雑談対話API (Application Programming Interface)*1を開発し、2013年よりdDs (docomo Developers support) [1]にて公開してきた。さらに、2018年には、雑談対話APIを基に開発した「かたらい [2]」をインターメディアプランニング株式会社と共同でリリースした。これらは、タスクを受け付けるのではなく、ユーザとコミュニケーションをとることを目的としている。このようなコミュニケーションを目的とした対話は、ユーザに「親しみやすい」「楽しい」と思ってもらうことが重要である。

©2021 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

*1 API: OSやミドルウェアなどが提供する機能を、他のソフトウェアが利用するためのインタフェース。

そこで、親しみやすい対話を実現するために、特定のキャラクターを再現するキャラクター雑談技術と、なぞかけなどチャットボットとのゲーム性のある対話を実現できる言葉遊び技術を開発した。これにより、ユーザが親しみやすいと感じる、エンターテインメント性をチャットボットに付与することが可能になった。さらに、ユーザが手軽にエンターテインメント性のある対話（以下、エンタメ対話）を楽しめるように、チャットツール上で動作するエンタメ対話アプリを開発した。

本稿では、キャラクター雑談技術と言葉遊び技術、対話アプリについて解説し、活用事例についても紹介する。

2. キャラクター雑談技術

2.1 概要

チャットボットにキャラクター性を付与することで、ユーザに親しみやすいと思ってもらえる対話を実現できる可能性が高い。例えば、ユーザの好きなキャラクターや、オリジナルのキャラクターを設定することで、ユーザの興味を引くことができる。

キャラクター雑談技術は、NTTメディアインテリジェンス研究所の成果技術 [3] を基に開発した、任意のキャラクターとの対話を再現する技術である。作成したいキャラクター風の発話ペアデータ（質問、回答）を作成し、それを基に対話システムを構築する。新規に対話システムを構築する場合だけでなく、組み合わせることで既存の対話システムにもキャラクター性を付与することが可能である。

本技術では、低コストでリアルタイム性のある、キャラクターを再現した高精度な対話を実現している。まず発話ペアデータベースを検索することで候補を抽出し、次にさまざまな手法で算出したスコアを用いて最終的な発話を決定している。このように、最初に検索で回答候補を絞ることで、高スペックな

サーバ以外でも、リアルタイム性のある対話を実現することができる。

2.2 処理の流れ

キャラクター雑談技術のシステム構成図を図1に示す。ユーザ発話が入力されると、次の処理を実行する。

まず、発話ペア検索部では、入力されたユーザ発話に対する複数の回答候補を、発話ペアデータベースから検索・抽出する。発話ペアデータベースには、任意のキャラクターの発話ペア（質問と回答の対）をあらかじめ格納している。発話ペアデータベースにおいて、入力されたユーザ発話に類似した質問を検索し、質問とそれに対応する回答候補のペア集合を取り出す。

次に、発話スコア計算部では、翻訳モデルや word2vec^{*2}モデルを含むさまざまな手法を用いて、複数のスコアを計算する。翻訳モデルは、一般には両言語（例：日本語と英語）での文間の対応が与えられた対訳コーパス^{*3}を用いて学習するが、本技術では質問と回答の対応が与えられた大量の発話ペアを用いて学習している。この翻訳モデルに、入力されたユーザ発話と回答候補の対を入力することで、そのユーザ発話と回答候補の対がどのくらい自然かに関するスコアを算出する。word2vecモデルでは、入力されたユーザ発話と回答候補との類似度に関するスコアを算出する。

最後に、発話ランキング部では、複数のスコアを統合して発話のランキングを行う。最もスコアが高い発話を、最終的なシステム発話として出力する。

2.3 活用事例

「携帯人間AIジミー」（以下、AIジミー）は、キャラクター雑談技術を初めて活用した商用サービスで、株式会社ラフ&ピースマザーから提供予定の、芸人のジミー大西氏を再現したAIとの音声対話が

*2 word2vec：テキストデータを解析し、各単語の意味をベクトル表現化する手法。

*3 コーパス：テキストや発話を大量に収集してデータベース化した言語資料。

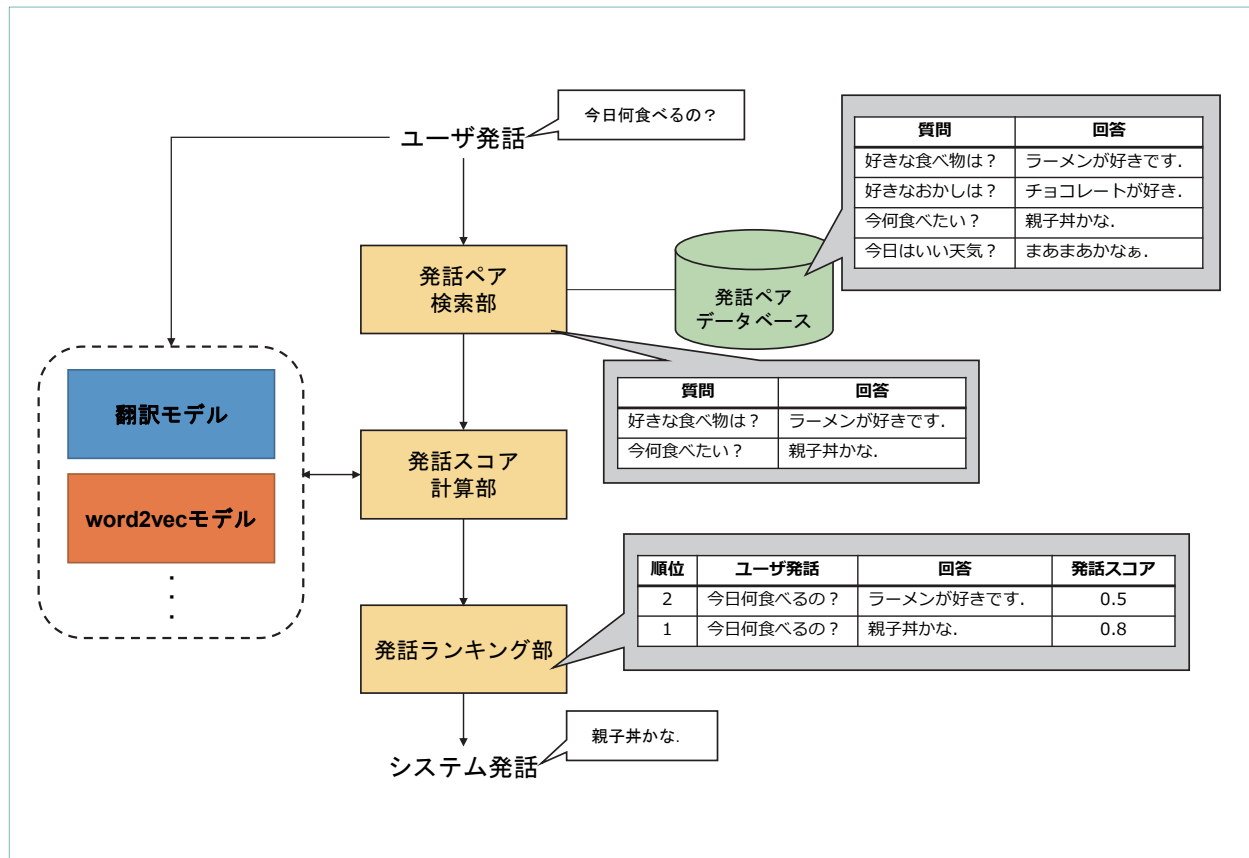


図1 キャラクター雑談技術のシステム構成図

可能なアプリケーションである (図2)。本アプリでは、ジミー大西氏が実際に質問に回答することで質問回答ペアを作成し、それらを学習させることでジミー大西氏を再現した対話を実現している。例えば、ユーザが「元気?」と発話すると、「僕ですか?もう至って元気です。もう生まれへんぐらい元気です。」のように、ジミー大西氏風の応答を行う。キャラクター雑談技術を用いることで、よりジミー大西氏らしい応答を可能にしている。

AIジミーでは、音声対話を実現するために、本技術に加えてAIエージェントと、NTTメディアインテリジェンス研究所が開発した音声合成モデルが使用されている。AIエージェントは、音声認識・音声合成機能を有しており、ユーザと音声での対話

が可能となる。AIエージェントの音声合成には、NTTメディアインテリジェンス研究所の協力のもとで開発した、ジミー大西氏の声を学習したオリジナル音声合成モデルを使用している。

3. 言葉遊び技術

3.1 概要

ゲーム性のある対話機能をチャットボットに付与することで、ユーザの興味度を高め、ユーザに「もっと話したい」と思ってもらえる可能性が高くなる。本技術では、山手線ゲーム、百人一首、なぞかけ、クイズを実現している。



図2 携帯人間AIジミー

(1)山手線ゲーム

山手線ゲームとは、例えば「山手線の駅名」のような特定のお題に対して、「新宿」「渋谷」と回答していく対話型ゲームである。回答が重複したり、誤った回答をしたりしてしまうと負けとなる。

回答をユーザが入力する際の課題として、「表記ゆれ」が挙げられる。例えば、ユーザが回答しようとしている語が「新宿三丁目駅」である場合、「新宿三丁目」のように語末の「駅」を省略したり、「新宿3丁目駅」のように漢数字をアラビア数字に置換したりして入力するケースが想定される。このような問題に対処するため、Wikipediaから自動で表記ゆれ情報を収集し、定期的に辞書を構築している。また、ゲーム性を高めるための工夫として、Wikipediaのページ閲覧数やリンク頻度から算出した「回答の知名度」に従って、システム回答順序を決定している。例えばお題が「山手線の駅名」であれば、「新宿」「渋谷」のような比較的知名度の高い駅名を優先的に回答し、「田町」のような比較的知名度の低い駅名を後に回答する。

(2)百人一首

小倉百人一首を利用した競技かるたは一般的に、詠手が上の句を詠み上げ、下の句のみが書かれた札（字札）の中から対応する札を競技者が取る形式で行われる。

本機能も同様に、まずシステムは、1件の上の句と、その上の句に対応する1件を含む複数件の下の句を、それぞれ詠み札・取り札としてユーザに提示する。ユーザは、システムから提示された下の句の中から任意の句を選択する。システムは、ユーザが選択した下の句が上の句に対応しているかどうか（正誤）をユーザに返却する。

後述のmy daizメンバー^{*4}「ことばあそび」は、上の句の表示と同時に、音声合成により句を詠み上げるが、上の句の表示テキスト（歴史的仮名遣）をそのまま発話音声とするのではなく、詠上げ文字列として別途発話文字列に調整を加えることで、既存の音声合成機能を利用した自然な札詠みを実現している。

(3)なぞかけ

なぞかけとは「XとかけてYととく。その心はど

^{*4} my daizメンバー：my daiz（^{*5}参照）に追加できる拡張機能（サービス）のこと。

ちらもZでしょう」といった形式の言葉遊びである。「かたらい」では、システムからランダムに提示される4つの単語から選択するか任意の単語を入力することで、それをお題Xとしたなぞかけを楽しむ。なぞかけはSNSから収集したテキストデータを基に構築しており、投稿文を（X，Y，Z）形式のトリプルデータに変換したものを利用している。

(4)クイズ

継続して気軽に遊べる対話型ゲームとして、クイズ機能を開発した。システムはランダムにクイズ本文1件と、回答候補として正答・誤答を併せて4つの選択肢をユーザに提示する。ユーザは必要に応じて、クイズ本文に対応するヒントを参照できる。問題に正解すると、問題の難易度に応じた点数が付与される。

後述の「ことばあそび」メンバーでは、一定期間

中の出題問題を固定し、獲得点数をユーザごとに蓄積することで、ユーザ間で点数の比較（ランキング化）を可能としている。

3.2 活用事例

my daiz^{*5}において、本技術を活用した対話型ゲームを「ことばあそび」として提供している（図3）。

my daizは、電車の運行状況や天気などの情報提供を主な機能としているが、対話型ゲーム機能を追加することで、エージェントに対するユーザの心的障壁を取り除き、親しみをもってもらうことをねらいとしている。

my daizメンバー「ことばあそび」独自の工夫として、音声・テキストによる操作だけでなく、GUI（Graphical User Interface）^{*6}によるボタン操作を可能としている。これにより、音声認識や誤字



図3 ことばあそび

*5 my daiz：ユーザに合わせた幅広い情報を提供する、スマートフォンやタブレット上で動作する音声対話エージェント。

*6 GUI：操作や表示の対象がボタンやアイコンで表現され、直感的な操作や視認性に優れたインタフェース。

脱字による入力ミスの低減を図っている。また、「クイズ」機能では、ユーザの利用継続率向上を狙って、日ごとに回答数の制限をつける、難易度に応じて点数を付与するなどの仕組みを追加しており、さらに「百人一首」機能では、my daizの音声合成を利用して上の句を詠み上げることで、競技かるたの自己練習としての活用や学習効果の定着を狙っている。

4. かたらいSlackアプリ

4.1 背景

昨今のリモートワークの全国的な普及に伴い、Slack[®]*7などのチャットツールが脚光を浴びている。一方で、リアルでの対話機会が減少したことで、「コミュニケーションが減った」など、リモートワーク特有の課題が浮き彫りになっている。エンターテインメント性の高い対話は、人と人、人とモノのコミュニケーションを活性化する効果が期待できる。

そこで、ユーザが手軽にエンタメ対話を楽しめるように、主要なチャットツールの1つであるSlack上で動作する対話アプリを開発した。Slackにインストールできるアプリケーションとしては初の、チャンネル内での雑談活性化機能をもつ。

4.2 機能

本Slackアプリでは、インストールするだけで以下に述べる機能を利用できる。

①投稿内容の返答

チャンネル^{*8}へ投稿されたメッセージの内容に応じて、自然な返答を投稿する。応答文は、かたらいの雑談対話エンジンからの出力を利用している。本アプリでは、チャンネルのログが不用意に流れることを抑止するために、ユーザ発言に対して毎回返答するのではなく、一定確率での応答としている。

②投稿内容に含まれるキーワードを拾った言葉遊び

前述したなぞかけのデータベースに格納されている任意の語が、チャンネルへ投稿されたメッセージに含まれているとき、一定の確率で言葉遊びを返答する。例えば、「ラーメン食べたいなあ」という投稿に対して、「ラーメンとかけて祇園祭ととく。その心は、どちらも出汁(山車)が魅力でしょう」と返答する。

③なぞかけやクイズの自動投稿

チャンネルを盛り上げるために、なぞかけやクイズなどのコンテンツを自動で投稿する。これにより、雑談専用チャンネルの埋没・沈滞を防いだり、雑談専用チャンネルへの投稿を利用者に喚起したりする。

本アプリのトライアル利用者9名を対象に8項目からなるアンケート調査を実施した(図4)。各質問事項の評定平均値に対して、7段階リッカートスケール^{*9}の「4. どちらともいえない」を基準とした1標本のt検定^{*10}を実施したところ、「メッセージを送りたくなった」「業務の息抜きになった」などの項目において5%水準で有意差が見られた。このことから、本アプリを利用することで、チャンネルへの投稿に対する心理的ハードルが下がり、組織内のコミュニケーションを促進させるとともに、対面の対話が減ることによる孤独感の解消が期待できる。なお、図中のエラーバーは標準偏差(収集したデータのばらつき具合)を示す。また、グラフ上部の「*」は有意水準を示す(* $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$)

5. あとがき

本稿では、キャラクター雑談技術と言葉遊び技術、対話アプリについて解説した。キャラクター雑談技

*7 Slack[®]: Slack Technologies, Inc.の登録商標。

*8 チャンネル: Slackにおいて、メッセージをプロジェクトやチーム単位で整理・共有する場所。

*9 リッカートスケール: アンケートなどで使われる心理検査的評価尺度の一種であり、提示された文に対して回答者がどの程度同意できるかを回答する。一般的に5段階の尺度が用いられる

が、7, 9段階の尺度を用いることもある。

*10 1標本のt検定: 1つの母集団の母平均が特定の値と等しいかの検定(統計学的な検証手法)のこと。本稿では「アンケート結果の平均値が4.0(=特定の値)である」と仮説(帰無仮説)を立てて検定を実施した。

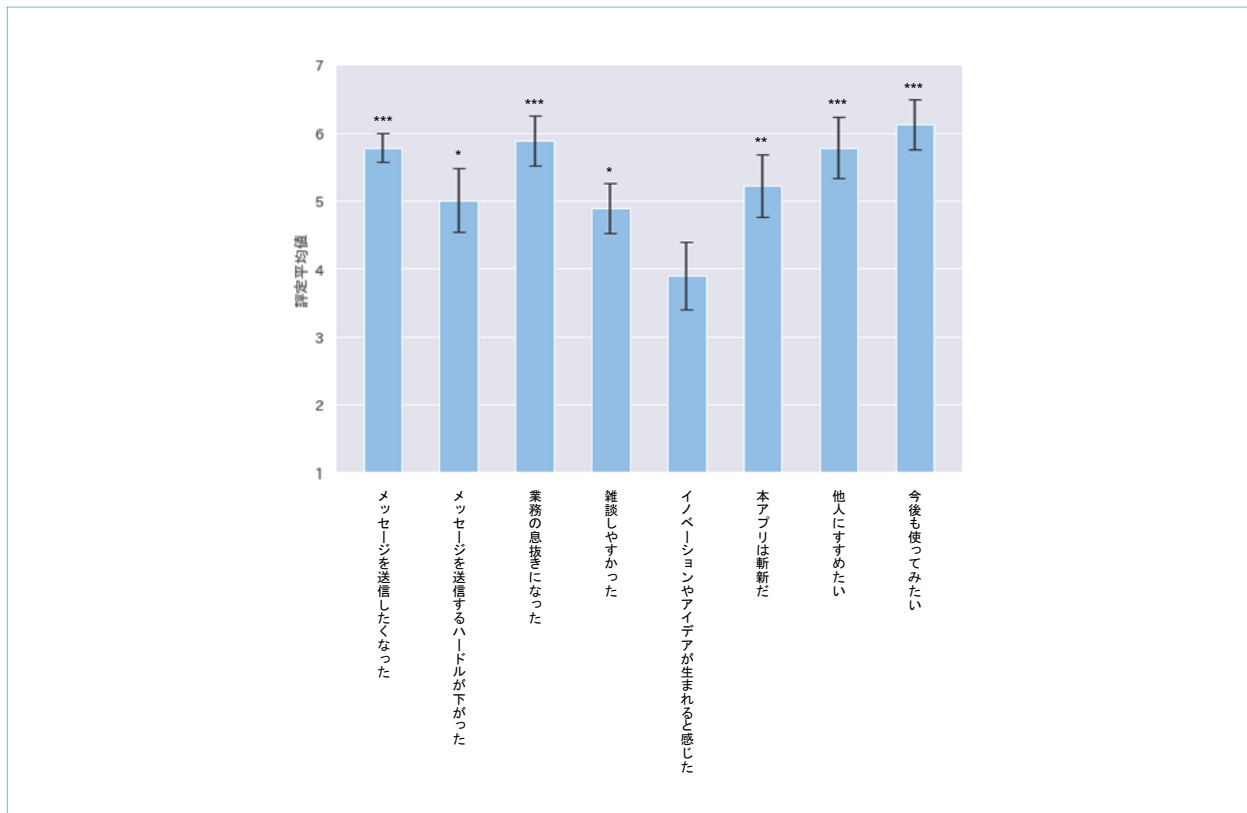


図4 かたらいアプリに関するアンケート結果

術は、作成したいキャラクター風の発話ペアデータを基に対話システムを構築することで、任意のキャラクターとの対話を再現することができる。言葉遊び技術は、Wikipediaなどの知識ベースを基に構築した「山手線ゲーム」や「なぞかけ」など、4種類のゲームを提供している。これらの技術により、チャットボットにユーザが親しみやすいと感じる、エンターテインメント性を付与することが可能になった。さらに、エンタメ対話を提供するSlackアプリを開発し、ユーザが手軽にエンタメ対話を楽しむことができるようになった。

今後は、実サービスで使用される中で課題を抽出

していき、より親しみやすいエンターテインメント性の高い対話を提供するサービスへと改良していきたい。

文献

- [1] docomo Developer support Webサイト.
<https://dev.smt.docomo.ne.jp/>
- [2] katarai Webサイト.
<https://www.katarai/>
- [3] R. Higashinaka, M. Mizukami, H. Kawabata, E. Yamaguchi, N. Adachi and J. Tomita: "Role play-based question-answering by real users for building chatbots with consistent personalities," Proc. of SIGDIAL, Jul. 2018.