

BGMのジャンルが音声エージェントとの対話印象に与える影響

飯田 隆太郎¹

概要：音声エージェントとの対話において、抵抗感を軽減し好印象を感じさせるためにエージェントの振る舞いや感情表出を操作する手法が研究がされてきている。このような手法は、より人間同士のコミュニケーションに近い対話を可能にする。その中でも対話環境情報をエージェントに操作させる振る舞いが検証されている。本稿では、この対話環境情報である BGM の効果を詳細に検証するため、BGM のジャンルに着目した。具体的には、音声エージェントとの雑談を想定したやり取りにおいて、提示される BGM のジャンルによってエージェントへ感じる印象が異なるかどうかを検証した。その結果、ロックはクラシック、ジャズ、ヒーリング音楽に比べて、好感因子に強く影響を与えていることが示された。また、実験結果を踏まえて、気分にあった音楽を提供してくれる音声エージェント制作への応用も検討した。

1. はじめに

1.1 研究の背景

近年、音声を媒介としたインターフェースが普及し、駅やレストランなど様々な場所で用いられている。特に、Google home や Alexa などの音声エージェント搭載のスマートスピーカーは AI 技術の発展や IoT の観点から様々な研究がされてきた。しかし、その活発な研究活動とは反して、国内でのスマートスピーカーの所有率はアメリカと比べてとても低い [1][2]。その原因として、やはり対音声エージェントとの会話における抵抗感が大きいことが挙げられる [3]。また、対話はタスク指向型対話（駅の案内など）と非タスク指向型対話（雑談）に区分され、近年のスマートスピーカーや音声エージェントの普及により後者の性能に期待するユーザーが増えてきた。これにより、ユーザーは音声エージェントに対し人間らしさを強く求めるようになってきた。

この問題を解決するためには、音声認識精度・音声合成精度・エージェントの振る舞い・エージェントの感情表出の4つの要素を最適化し、より人間らしい存在にすることが重要だと考える。特に、人間同士の会話では表面上の言葉のやり取りだけではなく、多くの非言語的情報をやり取りしコミュニケーションを取っている。これこそが対人間の対話との最大の違いである。

これに関連する研究として、目良らはユーザーの応答によって気分が変化し異なる反応を返す手法を提案し、従来

の音声エージェントよりも好ましい印象を与えることに成功している [4]。また、福田らはユーザーの嗜好情報を動的に更新する手法を提案し、好ましい印象を与えることに成功している [5]。さらに、佐々木らは音声エージェントの呼びかけ後の丁寧さを操作することによって、主従関係の理解が変化することを報告している [6]。このように音声エージェントの振る舞い・感情表出を操作することによってユーザーの音声エージェントへの好感度を向上させたり、特定の印象を与えることが可能だと言える。

ここで筆者は、音声エージェントの振る舞い・感情表出の操作として、対話環境を取り巻く情報の一つである音楽（BGM）に着目した。関連する研究として、小林らはカフェでの会話行動に光と音楽が与える影響を検証している [7]。さらに、高橋らは BGM を付加することによって音声エージェントの社会的存在感の強化に成功している [8]。しかし、音楽（BGM）と音声エージェントの印象に関する研究はまだ少なく、より詳細な検討をする必要があると考えられる。

1.2 研究の目的

このような背景を踏まえ、本研究では音声エージェントとの対話における BGM の役割に着目し、BGM のジャンルごとに印象形成にどのような効果があるかを検討した。具体的には、スマートスピーカーを介した音声エージェントとの雑談対話時に、エージェントが自主的に BGM を提示することによって、エージェントの人間らしさが BGM 無の対話と比べてどの程度向上したかを複数ジャンルの BGM

¹ 早稲田大学基幹理工学研究科

で検討した。なお、本研究の位置付けは以下の通りである。

- (1) スマートスピーカを用いた音声エージェントとの対話において、
- (2) 雑談場面を想定した時に、
- (3) 音声エージェントの印象に影響を与える、
- (4) 音楽ジャンルの比較検証

1.3 本論文の構成

はじめに、本研究の背景と目的、本論文の構成について述べる。

第2章では、音声エージェントとの対話におけるBGMの影響に関する実験について述べる。

第3章では、第2章の実験結果を提示する。

第4章では、音声エージェントとの対話におけるBGMの影響に関する実験結果について考察する。

第5章では、本研究を総括し、結論を述べる。

2. 実験

本章では、スマートスピーカを用いた音声エージェントとの対話において、提示される音楽のジャンルが音声エージェントの印象に与える影響について検証した実験の詳細を述べる。

2.1 実験概要

実験は、音声エージェントとの雑談する場面を想定して行った。実験参加者には、スマートスピーカを通じて音声エージェントと台本に沿って対話をさせた。対話時には聴覚刺激を提示した。その後、それぞれについて5件法(0:感じない, 1:やや感じない, 2:普通である, 3:やや感じる, 4:感じる)で印象評価を行った。質問項目は表1に示す17項目である。項目の選定基準としては、高吉ら[9]がロボットの評価に用いた15項目を参考に、「知性がある」「対話が成立する」の2項目を加えた。

さらに、全ての手続きの終了後、対話時の音声エージェントの印象及び聴覚刺激の印象についてインタビューを行った。

なお、聴覚刺激はランダムに提示することでカウンターバランスをとった。また、それぞれの条件の前に聴覚刺激なしで会話させることによってエージェントに対する印象のリセットを行った。

会話の話題は、参加者による興味にばらつきが出ないように実験当時の気温・季節の話とした。また、全ての条件において同一の台本を用いた。また、できる限り雑談に近い心理状態にさせるために2箇所2択の選択肢を入れることにした。ただし、選択肢の組み合わせによる印象のばらつきを防ぐため、1回目の対話で選択した選択肢を毎回選ぶように指示した。使用した台本を付録として図A-1, A-2に示す。

実験は、20代の男女5名に対して行なった。

2.2 実験装置

実験環境は図1に示す通りである。参加者は卓上のスマートスピーカ(Google home)に真っ直ぐ向き合って座り、台本に沿って対話した。音声エージェントの音声と聴覚刺激は共にスマートスピーカを用いて提示した。音声エージェントの対話システムの起動はノートPC(MacBook Pro)で行なった。

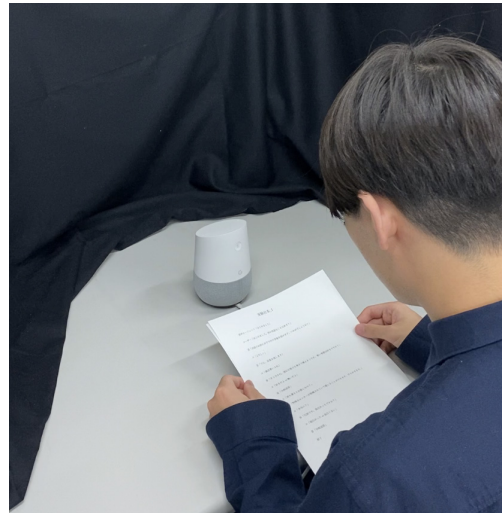


図1 実験環境

2.3 提示刺激

音声エージェントとの対話システムには、Google Assistant*1とDialogflow*2を用いて実装した。

聴覚刺激として、ロック/クラシック/ジャズ/ヒーリング音楽の4ジャンル×2曲の8種類を選定した。曲による好みで偏りが出ないように各ジャンル2曲を選曲した。さらに、調性による影響も排除するためにヒーリング音楽以外の調性をはっきりしている音楽に関してはkeyをEmで統一した。このようにして選んだ聴覚刺激を表2に示す。

3. 実験結果

本章では、スマートスピーカを用いた音声エージェントとの対話において、提示される音楽のジャンルが音声エージェントの印象に与える影響について検証した印象評価実験、自由回答結果を示す。

3.1 音楽ジャンルと音声エージェントに感じる印象の関係

今回提示した8種類の聴覚刺激における17項目の評定結果(5名×8条件×17項目=680点)を用いて、因子分析(最小二乗法, プロマックス回転)を実施した。因子

*1 <https://developers.google.com/assistant>

*2 <https://dialogflow.cloud.google.com/>

表 1 実験で使用した質問項目

外交的である	友人にしたい
格好いい	感じがいい
仲良くなれそう	また会いたい
優しそう	共感が持てる
気が利く	真面目そう
おとなしい	やんちゃ
不真面目	機械らしい
人間らしい	対話が成立する
知性がある	

表 2 聴覚刺激

No.	ジャンル	キー	曲名	アーティスト名
1	ロック	Em	サウダーズ	ポルノグラフィティ
2	ロック	Em	勝手にシンドバッド	サザンオールスターズ
3	クラシック	Em	交響曲第 9 番「新世界より」- 第 4 楽章	ステイーヴン・ガンゼンハウザー
4	クラシック	Em	禁じられた遊び (ロマンス)	ナルシソ・イエベス
5	ジャズ	Em	ジャニー・ギター「大砂塵」	ベギー・リー
6	ジャズ	Em	シャレード	フィルム・スタジオ・オーケストラ
7	ヒーリング音楽	—	水辺にて	菅野由弘
8	ヒーリング音楽	—	森の営み	菅野由弘

数は 4 因子を仮定した。

因子分析結果を表 3 に示す。第 1 因子は「外交的である」「感じがいい」といった項目が高い因子負荷を示しており、音声エージェントの積極的な印象を表す項目と言える。そこでこの因子を「積極性因子」と命名する。

第 2 因子は、「また会いたい」「友人になりたい」といった項目が高い因子負荷を示しており、音声エージェントとの友好関係を印象つける項目だと言える。そこでこの因子を「友好因子」と命名する。

第 3 因子は、「おとなしい」「やんちゃ」といった項目が高い因子負荷を示しており、音声エージェントの消極的な印象を表す項目と言える。そこで、この因子を「消極的因子」と命名する。

第 4 因子は、「人間らしい」「機械らしい」といった項目が高い因子負荷を示しており、音声エージェントの人間らしい印象を表す項目と言える。そこでこの因子を「人間性因子」と命名する。

各因子間相関を表 4 に示す。表 4 より「積極性因子」と「友好因子」、「積極性因子」と「人間性因子」にやや強い正の相関があった。

さらに因子分析によって得られた各因子得点を 8 種類の条件に分類し、実験参加者 5 名の各データに対して平均値と標準誤差を算出した。

図 2 は「積極性」因子の因子得点の各条件における標本平均値と標準誤差を示したものである。全条件において大きな影響の差は見られなかった。条件 5, 6 が同ジャンルの中で差が生じているように見える。これらの点に関して、楽ジャンルの種類を独立変数、因子得点を従属変数として 1 要因分散分析を行なった。その結果、音楽ジャンルの主

表 3 音声エージェントの印象に対する評定値の因子分析結果

項目	f1	f2	f3	f4	共通性
外交的である	.811	-.115	-.262	-.027	.607
感じがいい	.723	.034	-.256	.039	.648
気が利く	.654	.006	.393	.117	.654
対話が成立する	.555	-.021	-.105	-.007	.302
優しそう	.539	-.178	-.220	.007	.258
格好いい	.444	.245	.162	.004	.371
知性がある	.441	.260	.341	.120	.516
共感が持てる	.367	.215	.240	.150	.372
また会いたい	-.144	.894	-.090	-.075	.705
友人にしたい	.012	.853	-.084	.107	.871
仲良くなれそう	.237	.693	-.224	-.179	.696
おとなしい	-.249	-.082	.847	.054	.798
やんちゃ	.097	-.089	-.779	.123	.646
真面目そう	-.080	-.182	.656	.073	.508
不真面目	-.409	.217	-.623	.304	.717
人間らしい	.136	-.092	-.122	.868	.883
機械らしい	-.068	.030	.073	-.786	.687
寄与	4.219	3.809	3.280	3.203	

表 4 因子間相関行列

Table 4 Factor correlation matrix.

因子	f1	f2	f3
f1	1.000	.730	-.021
f2	.730	1.000	.080
f3	-.021	.080	1.000

効果に有意な差 ($F(7,32) = 1.33$) は見られなかった。

図 3 は「友好」因子の因子得点の各条件における標本平均値と標準誤差を示したものである。条件 1, 2 のロックが最も影響を及ぼしていることがわかる。条件 3, 4 のクラシックと条件 7, 8 のヒーリング音楽が同程度の影響と

なっている。条件 5, 6 のジャズに関しては、曲によって大きく差が出ている。これらの点に関して、音楽ジャンルの種類を独立変数、因子得点を従属変数として 1 要因分散分析を行なった。その結果、音楽ジャンルの主効果が有意 ($F(7,32) = 6.37, p < .01$) であった。さらに、各水準間の差について Holm 法による多重比較を行ったところ、条件 1 と条件 5, 条件 1 と条件 8, 条件 2 と条件 5, 条件 2 と条件 8, 条件 5 と条件 6 との間有意差があった ($p < .01$)

図 4 は「消極性」因子の因子得点の各条件における標本平均値と標準誤差を示したものである。全条件において大きな差はなかった。これらの点に関して、音楽ジャンルの種類を独立変数、因子得点を従属変数として 1 要因分散分析を行なった。その結果、音楽ジャンルの主効果に有意な差 ($F(7,32) = 0.27$) は見られなかった。

図 5 は「人間性」因子の因子得点の各条件における標本平均値と標準誤差を示したものである。消極性因子の場合と同様に全条件において大きな差はなかった。これらの点に関して、音楽ジャンルの種類を独立変数、因子得点を従属変数として 1 要因分散分析を行なった。その結果、音楽ジャンルの主効果に有意な差 ($F(7,32) = 1.25$) は見られなかった。

3.2 音声エージェント及び聴覚刺激の印象についての自由回答結果

全ての手続き終了後に実施したインタビューで得られた主な回答を以下にまとめる。

「今回聞いた音楽に好き嫌いがありましたか？」という質問に対しては、全員一致でないと答えた。

「音楽がある場合とそうでない場合で印象に違いはありましたか?」「違いがあった場合、どのように感じましたか?」という質問に対しては、全員一致であると答え、「明るい曲だと会話をしている感じが増した」「日本語の歌詞がついている曲だと機械と喋っている感じが薄れる」「エージェントの雰囲気が曲の雰囲気に合わせて変化するように感じた」といった回答が得られた。

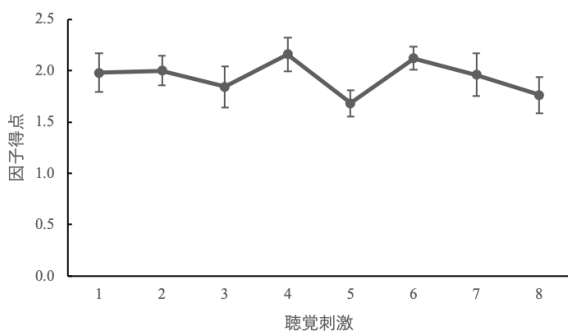


図 2 「積極性」因子得点の平均値

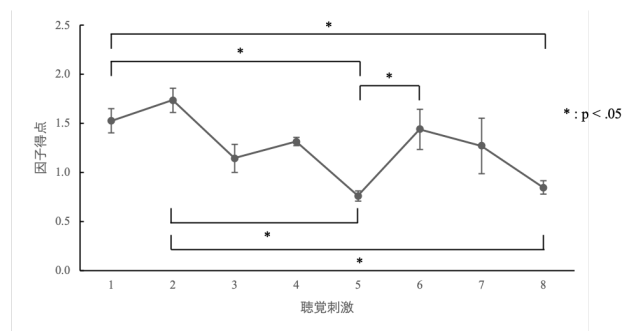


図 3 「友好」因子得点の平均値

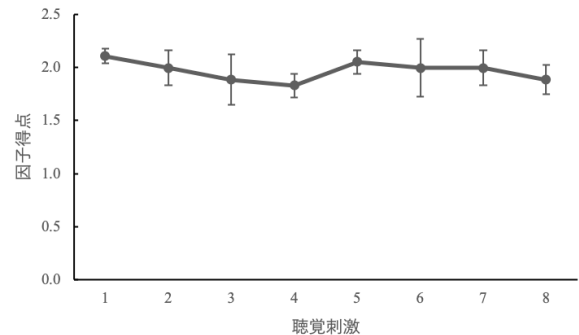


図 4 「消極性」因子得点の平均値

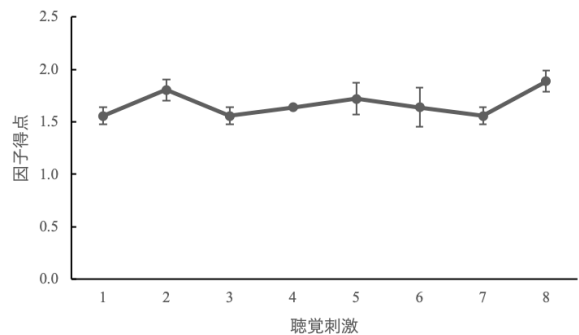


図 5 「人間性」因子得点の平均値

4. 考察

本章では、スマートスピーカを用いた音声エージェントとの対話において、提示される音楽のジャンルが音声エージェントの印象に与える影響について検証した実験について考察する。

今回の実験結果より音声エージェントとの対話において、音楽ジャンルによって「友好因子」及ぼす影響に差が生じることがわかった。特に、今回分析したジャンルではロックが最も影響が強く、ヒーリング音楽やクラシックは影響が低かった。これはロックというアップテンポの曲調がローテンポの音声エージェントの音声と組み合わせることにより、擬似的な会話の緩急のようなものが生じたからではないかと推察する。高杉らは、ロボットとの対話における交代潜時長と頷きのタイミングの影響を調査し、「丁

寧さ」以外の全ての項目において交代潜時間が短く領きタイミングが発話開始タイミングに近いほど評価が高くなることを明らかにした [10]. このことから、音声エージェントとの会話において音楽が会話テンポの感じ方に影響を与え、それが音声エージェントの印象につながっていると考えられる。また、この現象は日本語の歌詞も要因になっている可能性がある。音声エージェントの音声だけではなく歌詞も言語情報として頭に入ってくるため、認知負荷がかかり結果的に会話のテンポが速く感じたのではないかと考えられる。

さらに、「友好因子」の因子得点において、条件5と条件6が同ジャンル内で有意差が生じた。これに関しては、ロックが「友好因子」に及ぼす影響の原因と同様に、条件5は条件6に比べてテンポが低いことと、ビート（ドラムなどの打楽器音）が少ないことに要因があると考えられる。

また、本実験の参加者は5名だったが、音楽の感じ方や音声エージェントへの印象は個人差が大きい要素である。したがって、今回の実験結果の信頼性には議論の余地がある。さらにより多くの実験参加者からデータを集め詳細な検討を重ねる必要がある。

5. コンテンツへの応用

本章では実験結果を踏まえ、コンテンツ制作への応用に関して述べる。

本研究では音声エージェントとの対話を想定した場面でのエージェントによる音楽提供を行うシステムを実装し、実験を行った。実験用では会話内容の統制のために1つのパターンでの会話しか用意されていなかった。しかし、対話状況が変化したとしても、音声エージェントの印象を音楽で操作できるかを詳細に検討する必要がある。

そこで、筆者は、ユーザの気分合う音楽を流し雑談することで、音楽療法のような振る舞いをする音声エージェントへの応用を着想した。これは、同質の原理 [11] から着想を得た。同質の原理とは、落ち込んでいるときには暗い音楽を興奮しているときには明るい音楽を聴くというように、気分やテンポに合った曲を聴くことで聴取者は精神的に良い方向へと誘導されるという原理である。

具体的には、最初にディスプレイ等に気分を測定するためのマトリックスを用意し、ユーザはそこをタッチする。タッチされた座標の x 軸/y 軸と音楽の明るさ/テンポが対応しており、気分に合わせて曲を再生する。曲を再生した後は、時間経過によって再生する音楽を変更する。また、音声エージェントの話すテンポや明るさも曲に対応して変わるといえるものである。体験者の気分のできる限り寄り添った振る舞いをするので、音声エージェントとの間に親近感や愛着といった感情を醸成をさせられるのではないかと考えている。また、今回の実験結果を踏まえて、提供可能な曲を増やし歌詞やテンポによる偏りを少なくする。

6. 結論

本稿では、スマートスピーカーを用いた音声エージェントとの対話において、提示される音楽のジャンルが音声エージェントの印象に与える影響について検証した。その結果、ロックはクラシック、ジャズ、ヒーリング音楽に比べて、好感因子に強く影響を与えていることが示された。また、実験結果を踏まえて、気分合った音楽提供による音楽療法のような振る舞いをする音声エージェントの制作という応用例も検討した。

参考文献

- [1] "The Smart Audio Report", April 2020, NATIONALPUBLICMEDIA · Edison Research, <https://www.nationalpublicmedia.com/insights/reports/smart-audio-report/>
- [2] "スマートスピーカーの日本における利用実態についてインターネット調査", 2019年2月18日, 株式会社電通デジタル, <https://www.dentsudigital.co.jp/release/2019/0218-000164/> (最終閲覧: 2020年12月1日)
- [3] "日本人の音声操作に対する意識調査 2017", 2017年10月5日, KDDI 株式会社, <https://news.kddi.com/kddi/corporate/newsrelease/2017/10/05/2726.html>
- [4] 目良 和也, 市村 匠, 黒澤 義明, 竹澤 寿幸: 情緒計算手法と心的状態遷移ネットワークを用いた音声対話エージェントの気分変化手法, 日本知能情報ファジィ学会誌, Vol.22, No.1, pp.10-24 (2010)
- [5] 福田 和彰, 加藤 昇平: 親しみやすい感性対話エージェントの実現に向けた動的嗜好更新手法, 電気学会論文誌, Vol.138, No.12, pp.1485-1492 (2018)
- [6] 佐々木 健, 北村 美穂, 倉田 啓一, 渡邊 克巳: 呼びかけ語の丁寧さが音声エージェントへの愛着に与える影響, 第33回人工知能学会全国大会, (2019)
- [7] 小林 茂雄, 小口 尚子: 光色とBGMの種類がカフェでの会話行動に与える影響, 日本建築学会環境系論文集, 第599号, 143-150, (2006)
- [8] 高橋ともみ, 田中一晶, 小林賢一郎, 岡夏樹: 音声対話エージェントの社会的存在感向上のための感情表現の音楽的強調, HAI シンポジウム 2020, (2020)
- [9] 高吉幸治, 田中俊也: ロボットの振る舞いと知性・性格の印象の関係, 情報処理学会研究報告 (CVIM), コンピュータビジョンとイメージメディア, Vol.2007, No.87, pp.43-48, (2007)
- [10] 高杉 将司, 吉田 祥平, 沖津 健吾, 横山 正典, 山本 知仁, 三宅 美博: コミュニケーションロボットとの対話における交替潜時間と領き先行時間長の影響評価, 計測自動制御学会論文集 Vol.46, No.1, 72/81 (2010)
- [11] ALTSHULER, I. M., "The past, present and future of musical therapy.", Music therapy, Philosophical Library, pp. 24-35, (1954)

付 録

音声エージェント「はじめまして」

ユーザー「最近寒いよね」

音「そうですね。流石に私でも寒さで凍えそうです。寒い季節は好きですか？」

ユ「好きだよor嫌いだよ」

音「分岐返答」

ユ「君も暑さとか感じるの？」

音「特殊なセンサーが搭載されていて感じるができます。汗もかきます。」

ユ「本当に？」

音「冗談です。面白かったですか？」

ユ「面白かったorつまらない」

音「分岐返答」

終了

図 A.1 音声エージェントとの対話台本（聴覚刺激なし）

音声エージェント「はじめまして」

ユーザー「はじめまして。何か得意なことはある？」

音「会話のお供におすすめの音楽を流せます。いかがでしょうか？」

ユ「よろしく」

音「わかりました。音楽を流します。」

——聴覚刺激提示——

ユ「最近寒いよね」

音「そうですね。流石に私でも寒さで凍えそうです。寒い季節は好きですか？」

ユ「好きだよor嫌いだよ」

——分岐応答：好きだよ——

音「そうなんですね。冬はスキーなどのウィンタースポーツがとても楽しいですよ。ちなみに私は夏が好きです。とても元気な気分になります。」

——分岐応答：嫌いだよ——

音「そうですね。寒いのが苦手なので。私も苦手です。ちなみに私は夏が好きです。とても元気な気分になります。」

ユ「君も暑さとか感じるの？」

音「特殊なセンサーが搭載されていて感じるができます。汗もかきます。」

ユ「本当に？」

音「冗談です。面白かったですか？」

ユ「面白かったorつまらない」

——分岐応答：面白かった——

音「それは良かったです。また会話したくなったらいつでも声をかけて下さいね。それではまた。」

——分岐応答：つまらなかった——

音「そうですね。次は面白い冗談を考えてきますね。また会話したくなったらいつでも声をかけて下さいね。それではまた。」

終了

図 A.2 音声エージェントとの対話台本（聴覚刺激あり）