

# 人の振る舞いへの気付きの表出による ロボットのソーシャルプレゼンスの強化

岩崎雅矢<sup>†1</sup> 周剣<sup>†1</sup> 池田瑞<sup>†1</sup> 小池祐輝<sup>†1</sup>  
大西裕也<sup>†1</sup> 河村竜幸<sup>†2</sup> 中西英之<sup>†1</sup>

**概要**：日常生活において、接客ロボットやガイドロボットなどのロボットが普及してきているが、ロボットの発言はそのソーシャルプレゼンスの弱さのために、人々に無視されることが多い。この問題を解決するには、ロボットの振る舞いは人々の注意を引き付け、傾きや返答をしてもらえるようなものである必要がある。本研究では、このような振る舞いを調査するために2つの実験を行った。まず、聞き手がロボットを見ている時に見返すロボットの動作によって、聞き手が自分の動作をロボットに認知されているという感覚になるのかを定量的に調べる実験室実験を行った。次に、実際の店舗においてロボットと訪問者の自然なインタラクションを調べる実験を実施し、さらにそのデータをもとに開発した社会的応答モデルの実環境での有効性を検証した。その結果、人とロボットとのインタラクションの初めにおいて、ロボットが人の行動を認識して行動できることを表現することが、聞き手の感じるロボットに見られている感覚を向上させ、ロボットの発言に対する返答の促進に有効であると分かった。

## 1. はじめに

コミュニケーションロボットと呼ばれる会話を行うことができるロボットは多くの場面で用いられている。例えば、博物館のガイドロボット[1][2]や教育[3]、店舗において訪問客を応対することを目的としたロボットも存在している。そんな中で、ロボットのある特定の行動がロボットの印象に対して効果的であるという研究[4][5][6]や、ある特定の行動や発話方法によって人の注意をロボットに向けることができるというような研究がなされてきた[7][8][9]。しかし、これらのロボットにおいて、共通の行動がすべての目的に対して必ずしも効果的というわけではなく、それぞれの目的に沿って行動する必要がある。本研究ではこれらのコミュニケーションロボットの中で接客ロボットに着目する。近年外国人観光客は増えているが、店員が外国語を話せずコミュニケーションがとれない店舗がある。このような店舗に対して接客ロボットを置くことは、外国語での客を呼び込みや商品説明を行うことを可能にさせる。また、顔認識機能や購入情報などの顧客データをもとに、商品の購入の頻度や嗜好などに応じてサービスを提供することができる。さらに、人の販売員は人であるために警戒されてしまうかもしれないが、ロボットは気を使う必要がなく、客から関心を持って近づく可能性もある。このような理由から、接客ロボットは様々な店舗に急速に配備されてきた。しかし、どのような行動が販売に有効かはまだわかっていない。接客ロボットの発言は、ロボットの社会的プレゼンスの弱さが原因となり人に無視される可能性があると考えられ、これを解決するために様々な研究がされてきた[8][9][19]。ロボットは発言を無視されると、商品紹介などを行うことができず、ロボット本来の目的を果たせないと

考えられる。これらのロボットの振る舞いは人々の注意をロボットへ誘導し、その人々にロボットの発言を注意深く聞いてもらえるものであるべきである。過去の我々の論文で、ロボットと訪問客とのインタラクションが双方向になることによって、接客ロボットが訪問客により人のように扱われるということを示した。そして、そのような双方向の会話を行うためには、訪問客にロボットの最初の発言に返答してもらうことが重要であり、それらのインタラクションの初期におけるロボットの印象がその後の会話に大きく影響していると結論付けた[10]。そこで本研究では、その最初の発言に返答してもらうためには具体的にどのようなタイミングでロボットがどのような行動をするべきなのかを調査することを目的とする。まず本論文では、実験室実験において、聞き手がロボットを見た瞬間にロボットが聞き手を見返すことによって、ロボットが自分の行動を認識しているという感覚を向上させるかどうかを調べる。しかし、実験室実験のみではそこで得られた統計的な結論が本当に実際の環境においても同じであるかはわからない。したがって、この実験室実験の評価に加えて実際の店舗におけるロボットを訪問客と会話させる実験を行う。その訪問客の言語情報だけでなく非言語情報も含めたマルチモーダル会話分析を行うことで、訪問客がロボットの話を聞き、それに対して返事をするために必要な接客ロボットの設計指針を得る。そして、我々はこのフィールド実験のデータに基づいて、ロボットが社会的に反応するモデルを開発する。

## 2. 関連研究

### 2.1 フィールドにおける実験

ロボットと人間の自然な相互作用を研究するために、実

<sup>†1</sup> 大阪大学大学院 工学研究科 知能・機能創成工学専攻

<sup>†2</sup> 京都イノベーション株式会社

際に多くの実験が既に行われている。たとえば、博物館[2][8][11][12][13]と教室[3]で行われた実験がある。また、ロボットをさまざまな目的でロボット販売員として使用する実験もある[14][15][16]。このロボット販売員の実験の中には、ショッピングモールで実施されたものもある[17][18]。これら2つの実験のうち、前者はロボットを用いて顧客との関係を構築するものであり、後者はロボットを用いて商品の売り上げを向上させるクーポンを提供したものである。しかし、これらの実験では、ロボットは顧客に直接商品を紹介していない。また、商業施設で直接商品を紹介するロボットの研究開発、店舗のオーナーに対する社会的受容性を調査した研究も存在している[28]。しかし、具体的にどのような行動が販売促進に有効であるかは未だに不明である。そこで、本研究では、実際の店舗でロボットが商品を紹介する際のロボット販売員の有効な行動を調査する。

## 2.2 顧客の注意を引くロボット

客の注意を引く実験も多く存在する。博物館でのロボットの頭部動作が客の注目を集めるという研究がある[8][9]。しかし、ここでロボットは自律行動を行っているため、客と自然に会話を行うことはできていない。さらに、インフォメーションセンターにロボットを設置し、客をロボットとの会話に引き付けるようにした研究もある[19]。ただ、そのデータ数は十分なものではない。また、街角で移動しつつ顧客に声をかけ、ロボットの発言に回答してもらうための、話しかけ行動モデルの開発に取り組んだ研究がある[27]。これには、人の移動方向を予測してロボットが人に気付かれやすいように近づいて声をかける手法が使われている。しかし、小さい店舗などでロボットが使用される場合、安全にロボットが移動を行うことができない場合も存在する。本研究では、そのような場合においても使用できる方法を調査する。さらに、ロボットの視線制御による対話相手の行動変化に関する研究も行われてきた[29][30]。しかし、実際の環境ではどのように働くのかは明確になっていない。

## 2.3 人の注意を誘導するデジタルサイネージ

デジタルサイネージにおいても、人々の注意を誘導する試みが数多くなされている。Kinectを用いて、ディスプレイの前を通る歩行者の視点位置を検出し、その位置に応じて、情報を正対表示するシステムを提案したもの[20]や、ディスプレイ前の人の位置移動に基づいて、提示する情報を変化させることで、人々に提示情報を認知させることができるというものがある[21]。しかし、これらのシステムは、そもそも音声を発さないシステムである上に、視覚的な情報提示がメインであるため、人間とシステム間の言語的、非言語的なやり取りは見られていない。

## 3. 実験室における実験

先行研究で、接客ロボットが人のように扱われ、適切にその役目を果たすためには、ロボットと訪問客とのインタ

クションが双方向になることが必要であり、そのためには、訪問客にロボットの最初の発言に返答してもらうことが重要であるということが示された [10]。そこで本章では、人とロボットのインタラクションの初期に着目し、より詳細に分析を行うため、まずは実験室実験を行う。

### 3.1 仮説

従来研究に、博物館のガイドロボットが、ロボットの発話の切れ目において、観客の方へ振り返ることで、頷きなどの反応を観客に促すことを示したものが存在する[8][11]。これらの研究では、ロボットの発話の適切なタイミングで振り返りを入れている。しかし、これらはガイドロボットであるためにロボットが一方向的に話す場合が多く、ロボットの発言内容に合わせてロボットを行動させている。しかし、接客ロボットの場合は客がロボットに対して話す場合も多く存在し、このような方法では対応できない可能性が考えられる。そこで、我々はロボットの発言を含む初期動作を被験者の動きに反応して行わせる。具体的には、被験者がロボットの方を振り返った瞬間に、ロボットもそれに応じて被験者の方に振り返り話し始めるという動作を行わせる。ロボットがこのような行動をすることで、ロボットの動作をプログラムであらかじめ時間を指定して行わせているのではなく、被験者の動きをロボットが理解して行い始めたという印象を聞き手に与えることができるのではないかと考えた。そこで以下の2つの仮説を立てる。

**仮説 1** 被験者がロボットに視線を向けた瞬間に、ロボットの言動を開始することで、被験者の感じるロボットに見られている感覚が向上する。

**仮説 2** 被験者がロボットに視線を向けた瞬間に、ロボットの言動を開始することで、聞き手はロボットの発言に対して相槌や返答を行うようになる。

### 3.2 実験条件

前述した仮説を検証するため、発話開始時における聞き手の振り返り動作に応じたロボットの見返し動作を要因とし、以下に示す実験条件を設定した。

見返し動作なし条件：被験者が実験室に入室した時から、絶え間なく、ロボットは、あらかじめ決められた原稿を一定の間を取りながら、繰り返し発言する。

見返し動作あり条件：ロボットは、被験者が実験室に入室した時には、「見返し動作なし条件」と同様に、発言を行っているが、被験者が席に着くタイミングで、発言を止め、被験者がロボットの方に視線を向けたタイミングで、ロボットも被験者を見返し、発言を再開する。

今回は条件順序によって被験者のロボットに対する印象に影響が出ることが予想されるため被験者間実験を行った。

### 3.3 実験環境

図 1 に実験中の様子を、図 2 に被験者側の部屋の実験環境を示す。被験者との距離は 100cm 離れており、ロボットは被験者がロボットを確認する際に振り返りを行わなければならないように被験者の左斜め後ろに設置した。



図 1 実験室における実験中の様子

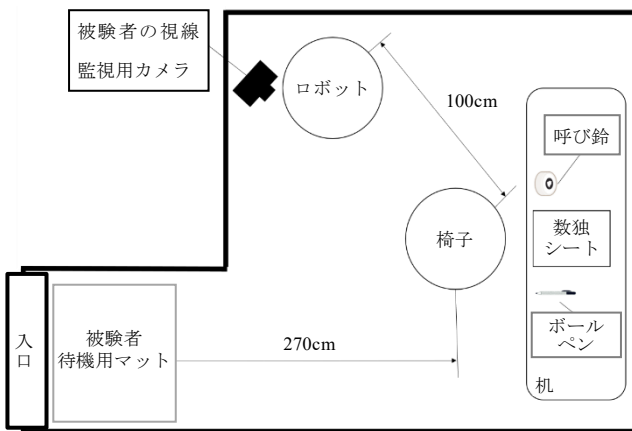
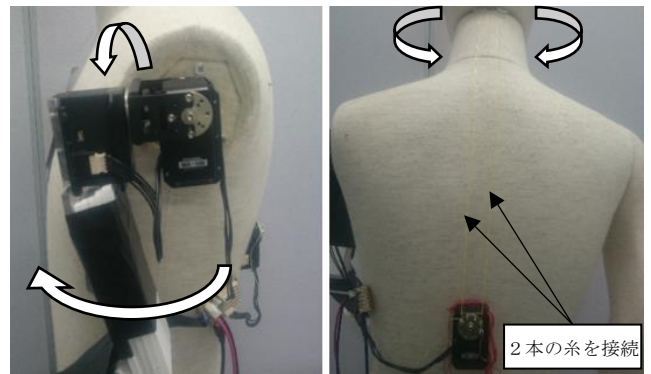


図 2 実験環境

図 1 からわかるように、このロボットは顔に目や鼻、口のような感覚器を有していない。このようなロボットを採用した理由は、ロボットの表情が構成され、聞き手のロボットに対する印象に影響を与えることを防止するためである。また、ロボットは図 5 のような方向を向かせて設置した理由は、見返し動作あり条件の話し始めで、被験者の方を振り返るロボットの動作以外の場面で、被験者とロボットとの間で顔が向かい合う状況を作らないようにするためである。このロボットの肩関節の自由度は図 3(a)に示すように、2 自由度となっており、前方及び後方の挙上動作と側方の外転及び内転の動作を行うことができる。図 3(b)に示すように、ロボットの首関節の自由度は 1 自由度となっており、頸部の左右方向の旋回運動を行うことができる。ロボットの背中部分にサーボモーターを取り付け、そこから 2 本の糸を伸ばして、ロボットの首元と接続する。これら 2 本の糸をサーボモーターで引っ張ることによって、頸部の左右方向の旋回運動を実現した。次に、ロボットの発話を遠隔操作するためのインターフェースについて説明する。ロボットの発話内容は、あらかじめ決められており、PC 上でボタンをクリックすることで、所定の発話が開始されるようになっている。発話が開始されると、一文ごとに一定の間を取りながら、発話を続けるように設定を行った。実験時

にはロボットの近くにカメラを設置し、実験者が被験者の様子を観察し遠隔操作を行えるようにした。また、音声ロボットから発せられていると被験者に示すため、ロボットは発言中頭部と腕を用いたランダムな動作を行う。



(a) 腕用モーター

(b) 首用モーター

図 3 ロボットの構造

### 3.4 タスク

本実験では、被験者のロボットに対する自然な振る舞いを観察するため、被験者にロボットの話聞くことがタスクであると思われないように、被験者には  $9 \times 9$  の数独を解いてもらうこと、数独の開始の合図はロボットから行われることの 2 点を事前案内で伝えた。以下に、実験の流れを示す。被験者が案内人と共に実験室に入室する時、既にロボットは発言を始めている。その発言中に、被験者は図 2 に示す待機マットの上で立って待機し、案内人からの事前説明を受ける。案内人からの説明が終わると、被験者は席に座り、ロボットから数独の解答を開始する合図を待つ。このとき、見返し動作なし条件では、事前説明終了後もロボットの発言は絶え間なく続いていく。一方で、見返し動作あり条件では、事前説明終了後、被験者が椅子に座るタイミングでロボットは一度発言を止め、被験者がロボットに視線を移した瞬間にロボットが被験者の方を見返し、発言を再開する。この見返し動作あり条件において、被験者が椅子に座るタイミングでロボットが一度発言を止めた理由としては、被験者にロボットの方へ振り返らせるためであり、被験者がロボットの方へ振り返りを行わない限り、見返し動作を行うことができずこの条件が成り立たないからである。以降はロボットが数独に関する話を約 4 分間話し続けた後、ロボットが数独の解答開始の合図を行い、被験者は数独の解き始める。数独を解き終えた被験者は机の上に置かれているベルを鳴らし、案内人を呼んだところで実験は終了となる。図 4 に各条件での実験の流れを示す。

### 3.5 参加者

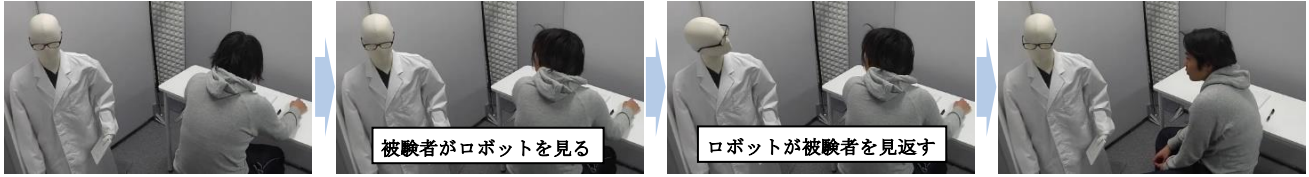
関西圏在住の 18 歳から 24 歳の大学学部生 20 名 (男性 10 名、女性 10 名) に対して実験を行った。

### 3.6 評価方法

#### (1) 行動評価

本実験では、まず、ロボットの発言に対して、相槌を行

### 見返し動作あり条件



### 見返し動作なし条件

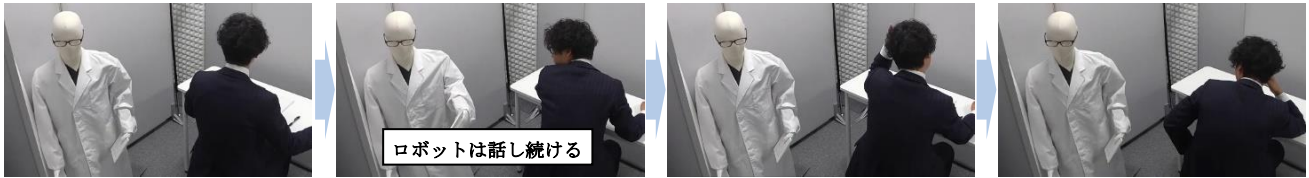


図4 各条件における実験の流れ

ったかどうかをカウントすることで評価した。これは仮説2に対応する。相槌をカウントするルールに関して、被験者がロボットの発話に対して、「はい」、「そうですね」のような発声を伴う返答を行った場合、もしくは、声には出さないが頷きを入れた場合に相槌行動を1回としてカウントした。会話初期における「こんにちは」や「よろしくお願いします」は「見返し動作あり条件」では、実験者が被験者の動きを観察しながら、被験者が振り向いたタイミングや、返事をしたタイミングを見計らって発言がされる。一方で、「見返し動作なし条件」では、一文の発言ごとに一定の間隔をおいて発言がされる。発言と発言の間が開くほど、相槌を打ちやすくなる可能性が考えられるため、分析は両条件間で各発言後の間隔が両条件で統制されている区間において、相槌回数やロボットに視線を向けた秒数の行動評価を行った。被験者がロボットに視線を向けている時間は、図2に示す視線監視用カメラで撮影した映像を確認して行った。「見返し動作あり条件」は被験者がロボットの方を振り向かないと条件として成り立たない。そこで、どうしてもロボットの方を振り向かない被験者の場合は、「こっちを向いてください」とロボットに発話させることにした。しかし、この言葉をかけることで、意図的に被験者をロボットの方へ向かせてしまっている可能性が考えられる。そこで、実験結果を示す際には、全ての被験者データを含めたものに加えて、「こっちを向いてください」と被験者に声をかけたデータを除いた場合のものも示すこととする。

#### (2) アンケート評価

ロボットに対する主観的な印象を尋ねるために、アンケートを採用した。実験後、被験者に対して7段階のリッカート尺度を用いたアンケートを取った。全尺度をそれぞれ1:全くあてはまらない, 4:どちらともいえない, 7:非常によくあてはまる, に対応させた。アンケートの質問項目に加えて、自由記述欄を設け、アンケート記述後にスコアをつけた理由、実験中取った行動の根拠を尋ねるインタビューを行った。

### 3.7 結果

アンケートによる印象評価の結果を図5に、行動評価の

結果を図6に示す。図の箱は各項目のスコアの平均値を表し、棒は標準誤差を表す。分析は対応のないt検定を用いた。まず、図5に示すアンケートによる印象評価の結果では、音声の質を問う項目では、各条件間で有意な差は見られなかったため、ロボットの発話の質による被験者の行動及び、ロボットの発話に対する印象の結果には影響がないことを確認した。また、Q2のロボットに観察されている感覚では、見返し動作あり条件が見返し動作なし条件より有意にスコアが高くなった( $t(18)=2.11, p<.05$ )。したがって、仮説1は支持され、ロボットがタイミングよく見返すことによって、被験者に対してロボットに見られている感覚を与えられているということが確認できた。アンケートのその他の項目に関しては、条件間でスコアの差は見られなかった。次に、図6に示す行動評価の結果について、相槌の回数に関しては、見返し動作あり条件の方が見返し動作なし条件に比べて、有意に相槌の回数が増える傾向が得られた( $t(18)=1.86, p<.1$ )。また、声に出して返答を行った回数に関しては、条件間で有意差は見られなかった。しかし、ロボットに視線を向けた時間に関しては、見返し動作あり条件が見返し動作なし条件よりも有意に時間が長くなった( $t(11)=2.92, p<.05$ )。さらに、「こっちを向いてください。」と被験者に声をかけたデータを除いた場合においても、見返し動作あり条件の方が見返し動作なし条件よりも有意に時間が長くなった( $t(11)=2.49, p<.05$ )。図7に両条件における、ロボットの各発言に対して、相槌を打たれた回数を表したヒストグラムを示す。図の箱は、ロボットの各発言に対して打たれた相槌の数を表しており、薄い灰色で示されたグラフは、ロボットが話す原稿の文の数を4分割にし、各パートで打たれた相槌の合計を示したものである。4分割とした理由としては、全原稿を発言するためにかかる時間がおおよそ4分間であったため、4分割にすることで、おおよそ1分毎の相槌の総数を区切りよく集計することができるからである。以降、ロボットの発言の3~13番目の発言を第1パート、14~24番目の発言を第2パート、25~35番目の発言を第3パート、36~45番目の発言を第4パートと呼ぶことにする。図7の縦軸は、話し始めから話し終わりま



でのロボットの発言を示す。また、グラフに示された横軸は、ロボットの各発言に対して打たれた相槌の平均回数とパート毎の相槌の平均回数を示している。図7を見ると、どちらの条件においても、ロボットの話が進むにつれて、相槌の回数が減っている傾向が読み取れる。また、第1パートが最も各条件間で、相槌の回数に差があると分かる。そこで、各パートで行われた相槌の回数に条件間で差が見られるか、分析を行った。その結果、図8に示すように、第1パートにおける相槌の回数は、見返し動作あり条件が見返し動作なし条件に比べて、有意に相槌の回数が多くなる結果となった ( $t(18)=2.47, p<.05$ )。また、第2パートにおける相槌の回数は、見返し動作あり条件が見返し動作なし条件に比べて、有意に相槌の回数が増える傾向が得られた ( $t(18)=1.77, p<.1$ )。第3パート、第4パートでは、条件間に相槌の回数で有意な差は見られなかった。したがって、仮説2はロボットの発言の序盤部分(第1パート)に着目すると支持された。

### 3.8 考察

まずはアンケート評価について考える。Q2のロボットに観察されている感覚において、見返し動作あり条件が見返し動作なし条件よりも有意にスコアが高くなる結果となった。これは、ロボットが被験者の振り返りに対して見返す行為が直接的に働いたためであると考えられ、被験者に対してしっかりと観察されている感覚を与えることができたということが確認できる。実験後のインタビューにおいて、「はじめに後ろを振り向くと、ロボットと目が合った」という意見や、「振り向くと、こっちを見返してきた。」という意見が得られた。したがって、ロボットが被験者の行動や態度を理解して行動できるのではないかと被験者が感じていると考えられる。また、ここでのロボットの行動が不自然だというコメントは得られなかった。

次に行動評価について考える。ロボットに視線を向けた時間で差が見られた理由としては、被験者のロボットに見られている感覚が向上することで、行動や態度を理解できるロボットなのではないかと被験者が感じるようになり、そのロボットに対して話をしっかり聞いていることを態度で示そうとしたためであると考えられる。実験後のインタビューにおいて、「ロボットに話を聞いているアピールをしたかった。」という意見が多く得られた。さらに、見返し動作あり条件では、顔だけロボットの方へ向けるのではなく、体ごとロボットの方へ向いて話を聞く様子が多く見られた。

また、相槌行動に関しては、ロボットの発言全てに対する相槌の回数で有意な傾向が示された。有意差までは得られなかった理由としては、話し始めに被験者の振り返り動作に合わせて行ったロボットの見返し動作の印象が時間の経過とともに薄れていき、行動や態度を理解できないロボットであるという印象を与えたためであると考えられる。これは、図7に示されるように、ロボットが話す原稿の文

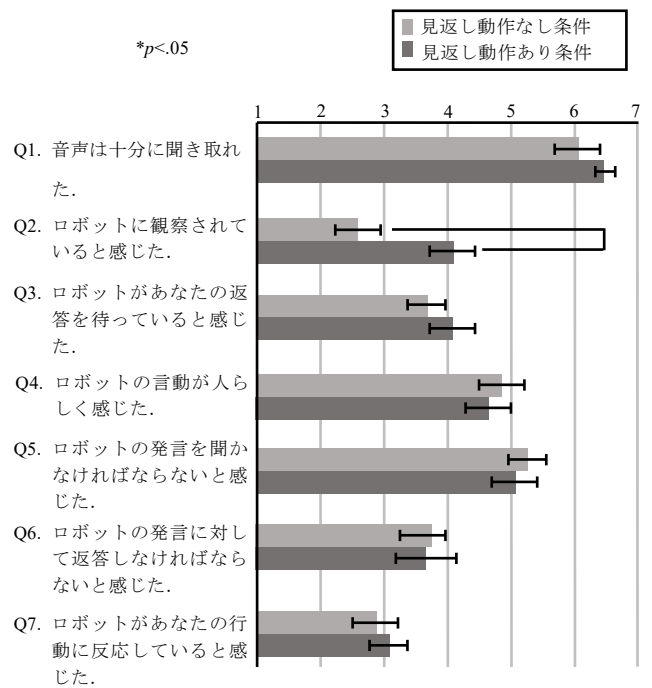


図5 アンケート結果

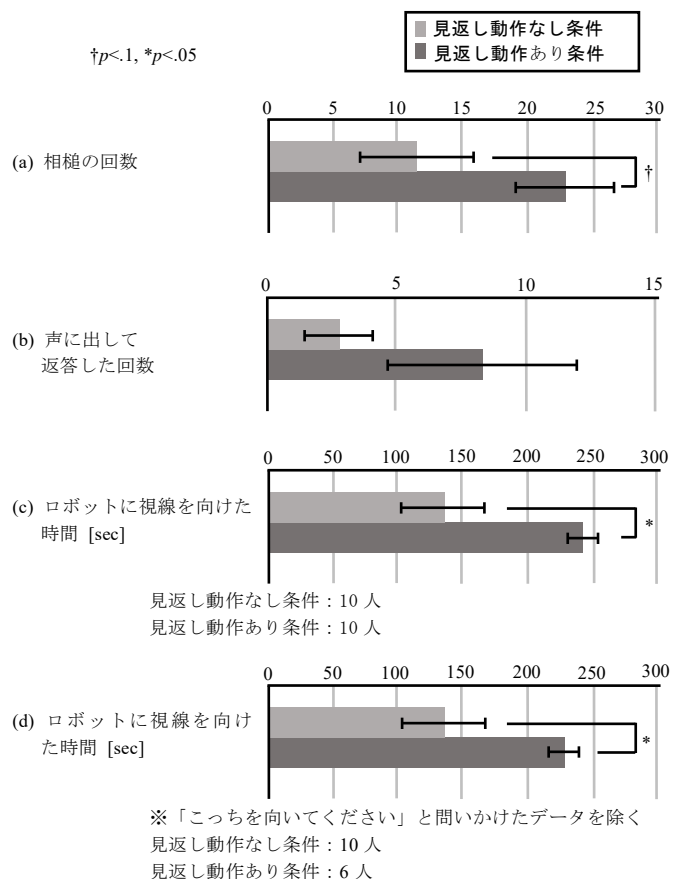
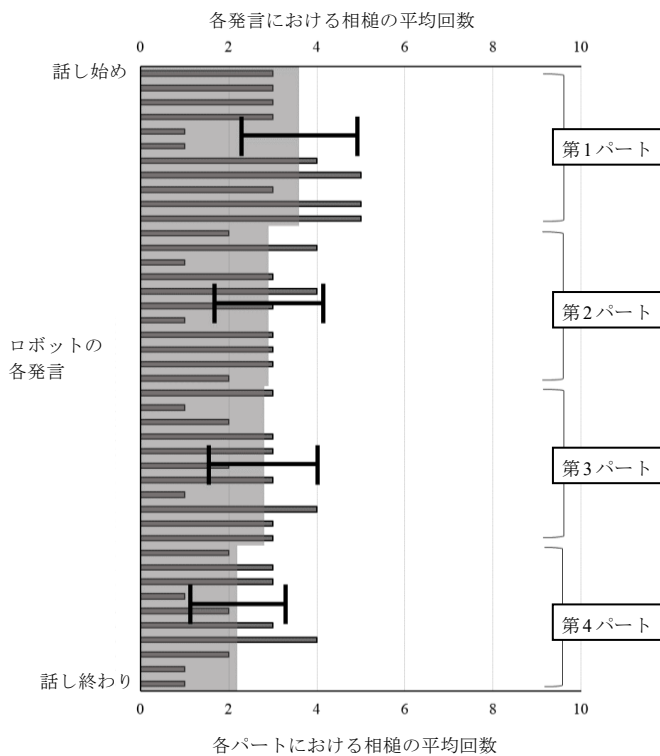
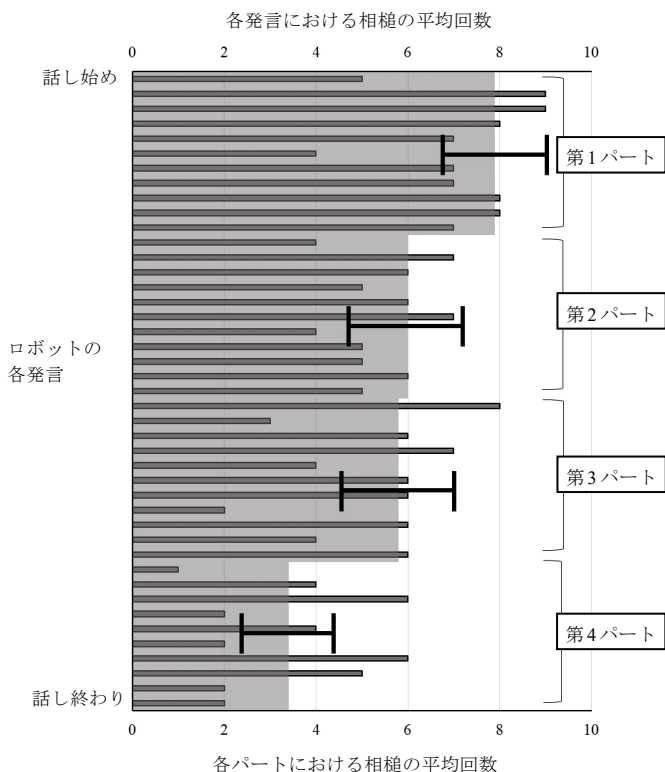


図6 行動評価結果

の数を4分割したグラフにおける相槌の回数がロボットの話が進むにつれて減少傾向にあることから読み取ることができる。また、行動分析においては前半のみ条件間で相槌の回数に有意な差が見られたが、アンケート評価においてQ5やQ6には見られなかった原因としても、アンケー



(a) 見返し動作なし条件



(b) 見返し動作あり条件

図7 ロボットの各発言に対する相槌回数

トは実験後に行われるため、実験後半の印象が強かったためであると考えられる。実験後のインタビューで「話が進むにつれて印象が薄れていったが、始めのうちは見返されたこともあり、こっちの行動が見られていると感じた。」という意見が得られた。また、見返し動作あり条件で相槌を

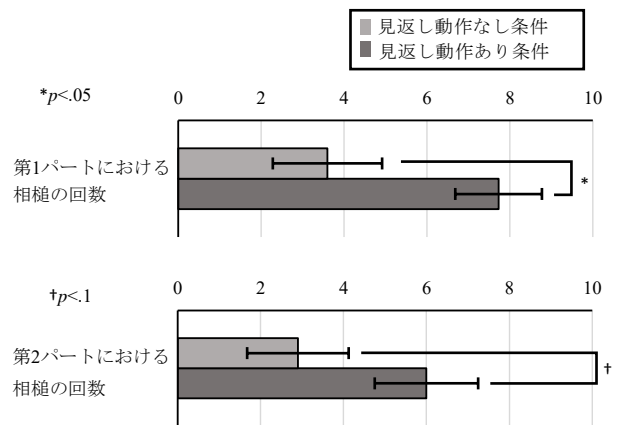


図8 第1, 第2パートにおける相槌の平均回数

返した被験者にその理由を聞くと、「返事をしないとロボットに申し訳ないと思った」という意見が多く、多くの被験者から得られた。また、「相槌を返さないとロボットが悲しみそう」だと答えた被験者も、見受けられた。したがって、「無視された」ということをロボットに感じさせてしまうことに、被験者は罪悪感のようなものを覚え、相槌を返してくれるようになったのではないかと考えられる。以上のことから、人の行動や状態をロボットがわかっていることを表出することで、そのロボットに見られている感覚が向上し、人はロボットに対して話をしっかり聞いていることを態度で示そうとすると考えられる。

本実験ではロボットが聞き手の行動を理解していることを示す動作として見返しを採用したが、どのような動作が適切であるかは調査していない。また、本実験では見返る動作は会話の冒頭のみで行われていたが、この動作の印象が後半で薄れたのではなく、単に被験者が飽きた可能性もある。さらに、会話の最中に同様の振る舞いを行うことでソーシャルプレゼンスを高い状態のまま維持できる可能性がある。これらの調査は今後の課題である。

#### 4. 実際の店舗における実験

前章の実験室実験の結果より、実際の店舗においてもこの結論が成り立つのかを調査する。

##### 4.1 実験環境

この実験は、本大学の研究倫理委員会の承認を得て行われている。本研究では、京都の商店街にある七味専門店「ちんたら」において実験を行った。図9はその実験の様子を示している。この店内に接客ロボットとして Pepper を置き、訪問客と会話をさせる実験を行った。Pepper を使用した理由は、Pepper には多くのセンサーが搭載されているため、顧客からのリアルタイムデータを簡単に取得でき、さらにそれらを用いた安全装置が搭載されているからである。この実験では、ロボットを遠隔操作するために Wizard of Oz 法[22]を用いた。そのため、我々はリモートコントローラを開発し、あらかじめ Pepper に様々な言動をインストールした。ロボットの主要な行動には、会話の他に、握手、試食

の勧誘、店員を呼ぶなどがある。ロボットの背後に会話用カメラを設置し、実験者は店舗の前にある商店街の通路の向かい側に立ち、カメラと肉眼の両方を用いて店内の状況を観察し、ロボットが次に行うべき行動を判断し選択した。訪問客の顔を見えないようにピントをずらした3つの不鮮明カメラを正面側、側面側、背面側に設置し使用した。また、顔を認識可能な映像を映した鮮明カメラも側面側に設置し、訪問客から撮影の同意が与えられた場合のみ映像を取得した。さらに、実験の情報を記載したハンドアウトをロボットの体に設置した。

#### 4.2 実験結果

実験は2017年4月10日 15:15-16:22, 13日 15:21-16:22, 14日 16:01-17:21, 2017年8月14日 16:21-17:58, 15日 15:20-16:22, 16日 14:15-15:58 の2017年11月27日 16:08-17:34, 12月02日 14:07-16:58, 04日 14:05-17:13, 05日 14:13-16:45 の10日間行った。164組, 360人の訪問客が実験中に店舗を訪れ、各組の平均人数は2.2人であった。



図9 実際の店舗における実験の様子

#### 4.3 分析と考察

ここでは、ロボットと訪問客とのインタラクションの初期部分に着目する。まず訪問客がロボットと会話を開始しなかった組と会話を開始した組の比較を行う。ただし、トランスクリプト中の括弧内の数字は沈黙の時間の秒数を示している。トランスクリプト1はロボットと会話をしなかった組の会話の例を示している。この組は、ロボットは訪問客に試食の提案を行った(3行目)。しかし、訪問客は商品を見ており、ロボットに反応しなかった。この理由として、訪問客はロボットが誰に話しかけているかを理解していなかったことが考えられる。次に、トランスクリプト2はロボットと会話をした組の会話の例を示している。この組は、図10のように訪問客がロボットの方に振り返り、見ているときに(4行目)、ロボットは「こんにちは」と発言した(5行目)。すると訪問客はロボットに返答した(6行目)。訪問客がロボットに返答したのは、訪問客がロボットを見たときにロボットが挨拶をしたからであると考えられる。つまり、訪問客はロボットが自分に話しかけてい

(P=ロボット, C1=男性, C2=女性)

- |   |   |
|---|---|
| 1 | ((入店))  |
| 2 | ((商品を見る))   |
| 3 | P Would you like to try a sample? You can taste here. |
| 4 | (3.3)   |
| 5 | P Welcome. Please feel free to watch the products.    |
| 6 | P Are you troubled to select?                         |
| 7 | P May I shake hand with you?                          |

トランスクリプト 1: ロボットと会話をしなかった組  
(8月16日 00:16:07-00:18:54) .

(P=ロボット, C1=男性, C2=女性)

- |   |                          |
|---|--------------------------|
| 1 | ((入店))                   |
| 2 | ((商品を見る))                |
| 3 | P May I help you?        |
| 4 | ((C1は振り向いてロボットを見る))(0.5) |
| 5 | P Hello!                 |
| 6 | C2 Hello!                |
| 7 | C1 Hi::!                 |
| 8 | P My name is Pepper.     |

トランスクリプト 2: ロボットと会話をした組  
(4月10日 00:37:45-00:42:30) .

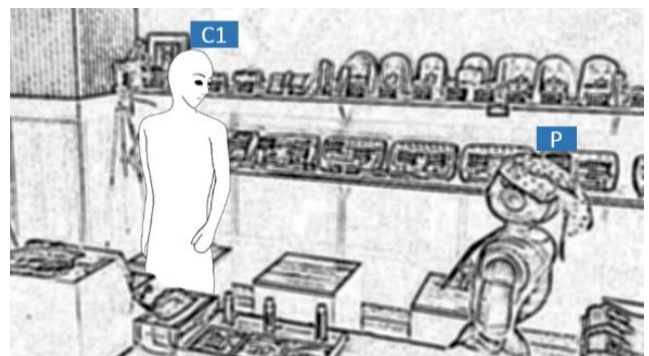


図10 ロボットの方に振り返り向く訪問客

ることを理解し、相互に認識しあった状態になったため返答を得ることができたと考えられる。

以上を踏まえて、訪問客がロボットを見ている時であればいつ挨拶をしても効果があるのかを調べる必要がある。トランスクリプト3は、訪問客がロボットを見ている時に挨拶をしたが返答しなかった組の会話の例である。この組では、訪問客がロボットの方に振り返ったとき、ロボットはすぐに挨拶をしなかった。訪問客はロボットの方に振り返った後、訪問客はロボットを3.6秒間見ている(4行目)。しかし、ロボットはこの間、何も行動をしなかった。その結果、訪問客は製品を見始めた(5行目)。したがって、訪問客がロボットを見たときにロボットはすぐに挨拶をする必要があると考えられる。そうしなければ、訪問客はす

(P=ロボット, C1=女性)

1	((入店))
2	P May I help you?
3	((C1 は振り向いてロボットを見る))
4	(3.6)
5	P ((C1 は商品を見始める))
6	P Nice to meet you!

トランスクリプト 3: ロボットを見るが返答しない訪問客 (8月16日 01:32:27-01:35:49) .

表1 ロボットの発言時における訪問客の状態と反応

	ロボットに返答した組	ロボットに返答しなかった組
ロボットを見ていた組	78	20
ロボットを見ていなかった組	8	58

ぐにロボットへの興味を失ってしまう。

以上から、訪問客がロボットを見たときに、ロボットが挨拶をする最適なタイミングである。このタイミングでロボットが挨拶をすることで、訪問客に、ロボットは自分が振り向いたという行動を理解して話しかけていると感じさせることができる。すべての訪問客について、ロボットの発言に返答したかどうかを調査した結果を表1に示す。表1から、ロボットに返答した訪問客の大部分は、ロボットが挨拶したときにロボットを見ていたことがわかる。

## 5. 社会的応答モデル

これまでの章から、訪問客に対してロボットが自分の振り向いたという行動を理解して話しかけていると感じさせることで返答してもらえんということが明らかとなった。したがって、ロボットは訪問客の非言語的な行動から判断できる相手の状況に応じて適切な行動をすることが必要である。本章ではロボットが訪問客にそのような印象を持たせることができるような社会的応答モデルを開発する。

### 5.1 エンゲージメントの識別と非言語要因

人がどれほどロボットに興味を持っているかを示す値として、エンゲージメントがある。本章ではロボットはこの値を用いて次の行動を選択する応答モデルを開発する。エンゲージメントの広く認知された明確な定義は未だ提案されていないが、先行研究からエンゲージメントをモデル化するのに有用な非言語要因の候補を収集することは可能である。先行研究によると、一般的な要因の1つはアイコンタクトであり、これは相手に注意を向けていることを表す強力なサインである[7][23]。さらに、アイコンタクトをすることによって相手が笑顔なのか顔をしかめているのかなどという表情も分析することが可能である。また、ロボッ

トに対する反応や姿勢などからも重要な情報を得ることができ[7][24]、ロボットと人との間の距離も重要な要因である[25][26]。

以上から、エンゲージメントの値を計算するのに用いる要因として、アイコンタクト、アイコンタクトの継続時間、距離、接近、笑いという5つの非言語要因を選択した。

### 5.2 社会的応答モデル開発のためのデータ分析

この応答モデルの開発には前章で説明した実際の店舗における実験のデータを用いる。この実験では164組の訪問客のデータを得た。しかし、このデータの中には訪問客から話しかけたものとロボットから話しかけたものが混在する。ここではロボットの最初の発言に対する訪問客の反応の有無について調査する必要があるため、164組の内、ロボットが最初に発言を行った153人の訪問客について、ロボットが最初の発言を行うまでの訪問客の行動を記録した動画をデータとして使用する。そして、実験者がこの動画からロボットが最初の発言を行う前の5つの非言語要因を取得する。また、以下の変数を定義した。

**Y: エンゲージメント**, 訪問客がロボットの発言に反応すれば1, しなければ0となる。

**X1: アイコンタクト**, ロボットが訪問客に挨拶をしたときに訪問客がロボットを見ていた場合1, 見ていなければ0となる。

**X2: アイコンタクトの継続時間**, 訪問客が入店してからロボットが挨拶するまでに測定されたアイコンタクトの時間(秒)。

**X3: 距離**, ロボットが訪問客に挨拶をしたときのロボットと訪問客との距離(メートル)。

**X4: 接近**, ロボットが挨拶をしたときにその訪問客がロボットに近づいてきていれば1, そうでなければ0。

**X5: 笑い**, ロボットが挨拶をしたときに訪問客が笑っていれば1, そうでなければ0。

表2 Yと各変数の相関

	X1	X2	X3	X4	X5
Yとの相関	0.65	0.69	0.54	0.36	0.38

表2はこの実験データから得られたYと各変数の相関を示している。各相関係数の値は0から1であり、値が1に近い場合相関が強いことを示す。この表2からアイコンタクトおよびアイコンタクトの継続時間はエンゲージメントの最も重要な要因であることが分かる。そして、距離もまた重要な要因であると言える。

### 5.3 社会的応答モデルの開発

データを分析するため、機械学習のモデルの1つであるロジスティック回帰を用いる。ロジスティック回帰を使用する理由は二値分類問題を解くための簡単なモデルであるからである。以下の式5.1と5.2にロジスティック回帰モデルを示す。ここでのWは重み、bはバイアスである。これらの式を用いた実験データの分析結果を表3に示す。



$$Y = \frac{1}{1+e^{-z}} \quad (5.1)$$

$$Z = W^T X + b \quad (5.2)$$

コスト関数を最小化することにより、最適化された重みを得ることができる。また、バイアスのパラメータは-0.20である。我々はモデルを検証するために5分割交差検証を用いた。5分割交差検証はデータ全体を5つのパートに分け、4つのパートを使用してモデルをトレーニングし、1つをテストするための有効な方法である。交差検証から、訪問客のエンゲージメントを予測するのに十分であると考えられる88.9%の精度が得られた。

表3 重みのパラメータ

要因	重み
アイコンタクト	0.53
アイコンタクトの継続時間	1.35
距離	-0.99
接近	0.17
笑い	1.20

#### 5.4 実環境における社会的応答モデルの有効性

ここまではモデル作成のためのデータ収集について述べたが、本章ではこのモデルを実環境において用いる場合について述べる。ロジスティック回帰の方程式と重みをPepperにインストールし、Pepperの頭のカメラで訪問客の顔を認識し、ロボットとアイコンタクトを行うかどうか、笑っているかどうかを判断する。また、Pepperの足に搭載されたレーザセンサを使用し、訪問客とロボットとの距離を測定し、その変化量から接近中かどうかを検知する。このようにロボットが訪問客の非言語情報を0.5秒ごとに測定し、自動的に出力を計算するように設定した。ロジスティック関数は0から1の値を取り、単調増加であるため、今回はロジスティック関数の値が0.33以上0.66未満の場合、ロボットは「ねえねえ」あるいは「Hey」と訪問客に呼びかけ、0.66以上の場合「こんにちは」あるいは「Hello」と挨拶をするという社会的応答モデルを開発した。例えば、2m離れて0.6秒間ロボットを見て接近している訪問客の場合は $x_1=1, x_2=0.6, x_3=2, x_4=1, x_5=0, y=0.33$ となり、1.8m離れて0.7秒間ロボットを見ながら笑顔で立っている訪問客は $x_1=1, x_2=0.7, x_3=1.8, x_4=0, x_5=1, y=0.66$ となる。

今回作成したモデルの有効性を検証するため、実店舗において再び実験を行った。実験の設定は前述した実店舗における実験と同様である。実験は2018年9月10日15:21-17:28、9月11日14:20-17:04、9月20日14:17-17:17、9月21日13:52-17:08の4日間行った。44組の訪問客が実験中に店舗を訪れた。本実験ではこのモデルの有効性を検証するため、この応答モデルを用いてロボットから会話を開始した組において、ロボットの最初の発言に対して訪問客が返答を行ったかどうかについて調査する。具体的には、この社会的応答モデルを用いてロボットが挨拶をする組と、

本店舗の入り口がロボットから約2mであることから、約2m以内に近づいた時にロボットが挨拶する組における訪問客の返答率を比較して調査した。その結果、このモデルを用いて最初の挨拶に反応した組は20組中15組で75.0%、ロボットから約2m以内に近づいた時のロボットの挨拶に反応した組は19組中11組で57.9%であった。ここでは、訪問客から会話を開始した5組は除いた。これらを比較するとこのモデルを用いた方が訪問客の返答率が高いことが分かる。また、前章の2017年に行った実験における訪問客の返答率は153組中75組で49.0%であった。したがって、これと比較してもモデルを用いた方が訪問客の返答率が高いことが分かった。このように、実環境において75.0%の訪問客がロボットの挨拶に対して返答を行ったことや、これまでの実験や距離のみを制限した場合よりも高い返答率が得られたことからこの応答モデルは実環境においても有効であると推測される。今回高い返答率が得られた理由として以下のことが考えられる。手動で人が遠隔操作する場合、離れた場所をカメラや肉眼で観察することになるが、その場合、距離はある程度判断できても、訪問客の視線の方向までは正確に判断できない。しかし、このモデルを用いることで、訪問客がロボットに視線を向けたかどうかを適切に判別でき、さらには挨拶を行うべきタイミングにおいて、遠隔操作による遅延がなく、すぐに反応をすることができる。今回は七味専門店内の狭いスペースでの実験のため、Pepperの移動機能は使用しなかった。しかし、自ら対話相手へ近づく動作を行いながら話しかけ動作を行うことで、より高い返答を得られる可能性もある。そのような検討は今後の課題である。

## 6. おわりに

本論文では、インタラクションの初期時のロボットの行動を調査することによって、接客ロボットの最初の発言に対して訪問客に返答してもらうためには具体的にどのようなタイミングでロボットがどのような行動をするべきなのかを調査することを目的とし、フィールド実験と実験室実験の二つの実験を行った。まず、聞き手がロボットを見た瞬間にロボットが聞き手を見返すことによって、ロボットが自分の行動を認識しているという感覚を向上させるかどうかを調査する実験室実験を行った。次に、実店舗でのロボットと訪問者の自然な相互作用を調べるためのフィールド実験を行った。そして、実際の店舗における実験のデータをもとに社会的応答モデルを開発した。その結果、接客ロボットと訪問客との間のインタラクションの初めにおいて、ロボットが人の行動を認識して行動できることを表現することが、聞き手の感じるロボットに見られている感覚を向上させ、ロボットの発言に対する返答の促進に有効であると分かった。本研究から得た最も重要な知見は、人の状態をロボットがわかっているということを表出すること

が人とロボットの間の会話において重要であるということである。このような行動をロボットが行うことで、人はそのロボットが自分の行動を理解していると思い、話を聞いているというアピールをするために返答しうなずくようになる。我々はこの研究によってロボットと人の会話が促進され、ロボットがより効果的に使われることを期待する。

**謝辞** 本研究を遂行するにあたり、京七味専門店ちんちらの皆様には実験場所の提供および実験における多大な協力を頂きました。Michelle Geilenberg氏、金延儒氏には実験にご協力頂きました。心より厚く御礼申し上げます。

## 参考文献

- [1] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita, 2006. Interactive humanoid robots for a science museum. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 305–312.
- [2] R. Gehle, K. Pitsch, and S. Wrede, 2014. Signaling trouble in robot-togroup interaction. emerging visitor dynamics with a museum guide robot. In *Proceedings of the second international conference on Human-agent interaction*. ACM, 361–368.
- [3] F. Tanaka, K. Isshiki, F. Takahashi, M. Uekusa, R. Sei, and K. Hayashi, 2015. Pepper learns together with children: Development of an educational application. In *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*. IEEE, 270–275.
- [4] C. L. Sidner, C. D. Kidd, C. Lee, and N. Lesh, 2004. Where to look: a study of human-robot engagement. In *Proceedings of the 9th international conference on Intelligent user interfaces*. ACM, 78–84.
- [5] T. Kanda, H. Ishiguro, and T. Ishida, 2011. Psychological analysis on human-robot interaction. In *Robotics and Automation, 2001. Proceedings 2001 ICRA*. IEEE International Conference on, vol. 4. IEEE, 4166–4173.
- [6] C. L. Sidner, C. Lee, C. D. Kidd, N. Lesh, and C. Rich, 2005. Explorations in engagement for humans and robots. *Artificial Intelligence*, vol. 166, no. 1-2, 140–164.
- [7] M. L. Knapp, J. A. Hall, and T. G. Horgan, 2003. Nonverbal communication in human interaction. *Cengage Learning*.
- [8] A. Yamazaki, K. Yamazaki, Y. Kuno, M. Burdelski, M. Kawashima, and H. Kuzuoka, 2008. Precision timing in human-robot interaction: coordination of head movement and utterance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 131–140.
- [9] K. Yamazaki, A. Yamazaki, M. Okada, Y. Kuno, Y. Kobayashi, Y. Hoshi, K. Pitsch, P. Luff, D. vom Lehn, and C. Heath, 2009. Revealing gauguin: engaging visitors in robot guide's explanation in an art museum. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1437–1446.
- [10] M. Iwasaki, J. Zhou, M. Ikeda, T. Kawamura, H. Nakanishi, 2018. A Customer's Attitude to a Robotic Salesperson Depends on Their Initial Interaction. In *Proceedings of the 27th IEEE International Conference on Robot and Human Interactive Communication*. IEEE.
- [11] Y. Kuno, K. Sadazuka, M. Kawashima, K. Yamazaki, A. Yamazaki, and H. Kuzuoka, 2007. Museum guide robot based on sociological interaction analysis. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1191–1194.
- [12] M. Bennewitz, F. Faber, D. Joho, M. Schreiber, and S. Behnke, 2005. Towards a humanoid museum guide robot that interacts with multiple persons. In *Humanoid Robots, 2005 5th IEEE-RAS International Conference on*. IEEE, 418–423.
- [13] 川口 一画, 葛岡 英明, 鈴木 祐也, 中尾 誉, 山下 淳, カラーラピッチ, 山崎 敬一: ロボットの発話途中の沈黙と言い直しによる人の注意誘導, 日本バーチャルリアリティ学会論文誌, Vol. 14 No. 3. pp. 257-264, 2009.
- [14] K. Nakagawa, M. Shiomi, K. Shinozawa, R. Matsumura, H. Ishiguro, and N. Hagita, 2013. Effect of robots whispering behavior on peoples' motivation. *International Journal of Social Robotics*, vol. 5, no. 1, 5–16.
- [15] M. Niemela, A. Arvola, and I. Aaltonen, 2017. Monitoring the acceptance of a social service robot in a shopping mall: first results. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 225–226.
- [16] M. Niemela, P. Heikkilä, and H. Lammi, 2017. A social service robot in a shopping mall: Expectations of the management, retailers and consumers. In *The Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 227–228.
- [17] T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, and N. Hagita, 2010. A communication robot in a shopping mall. *IEEE Transactions on Robotics*, vol. 26, no. 5, 897–913.
- [18] M. Shiomi, K. Shinozawa, Y. Nakagawa, T. Miyashita, T. Sakamoto, T. Terakubo, H. Ishiguro, and N. Hagita, 2013. Recommendation effects of a social robot for advertisement-use context in a shopping mall. *International Journal of Social Robotics*, vol. 5, no. 2, 251–262.
- [19] M. K. Lee, S. Kiesler, and J. Forlizzi, 2010. Receptionist or information kiosk: how do people talk with a robot? In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. ACM, 31–40.
- [20] 小玉駿, 須藤翔太, 渋谷進. デジタルサイネージに向けた情報を正対表示させ歩行者を引き付けるシステム, 研究報告エンタテインメントコンピューティング (EC), 2014, 2014-EC-31, no.4, pp.1-8.
- [21] 木原民雄, 横山正典, 渡辺浩志. 人の位置移動による状況即応型デジタルサイネージの構成法, 情報処理学会論文誌. 2012, vol.53, no.2, pp.868-878
- [22] J. F. Kelley, "An iterative design methodology for user-friendly natural language office information applications." *ACM Transactions on Information Systems (TOIS)*, Vol. 2 No. 1, pp. 26-41, 1984
- [23] G. Castellano, I. Leite, A. Pereira et al, 2014. Context-Sensitive Affect Recognition for a Robotic Game Companion. *ACM Transactions on Interactive Intelligent Systems*, vol. 4, no. 2, 1–25.
- [24] H. Yanco, 2012. Personalization in HRI: A Longitudinal Field Experiment. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 319–326.
- [25] M. P. Michalowski, S. Sabanovic, and R. Simmons, 2006. A Spatial Model of Engagement for a Social Robot. *Advanced Motion Control, 9th IEEE International Workshop on*, 762–767.
- [26] 神田 崇行, "ロボットメディアによる人間情報処理研究: 4. コミュニケーションロボットと人間との距離." 情報処理, Vol. 49 No.1, pp. 24-29, 2008.
- [27] Satake, Satoru, et al. (2013) "A robot that approaches pedestrians." *IEEE Transactions on Robotics* 29.2: 508-524.
- [28] 石超, 佐竹聡, 神田崇行, 石黒浩. (2017) "客引きロボット導入に向けた社会実験." 日本ロボット学会誌 35.4: 334-345.
- [29] Andrist, Sean, et al. (2014) "Conversational gaze aversion for humanlike robots." *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM.
- [30] Admoni, Henny, and Brian Scassellati. (2017) "Social eye gaze in human-robot interaction: a review." *Journal of Human-Robot Interaction* 6.1: 25-63.