

対話的行為の連鎖性を利用した共同注意達成支援

藤原 貴裕^{*1} 川口 一画^{*1} 葛岡 英明^{*1}

Supporting Joint Attention by Sequence Organization of Interactive Actions

Takahiro Fujiwara^{*1}, Ikkaku Kawaguchi^{*1} and Hideaki Kuzuoka^{*1}

Abstract - For a museum guide robot, it is important to achieve joint attention to interact with human visitors effectively. Joint attention is the concept explaining the state in which the speaker and the listener are looking at the same object. However, pointing actions of the robot are often ignored by visitors while they are concentrating on the exhibits, leading to the failure of joint attention. In order to solve this problem, we focused on “restarts and pauses” (RP), which is the strategy to get listeners’ attention used in interpersonal communication. Previous research revealed that the robot can obtain visitors’ attention by RP, but it has not been confirmed whether joint attention is achieved by the robot’s pointing action after getting attention of the visitor. This research conducted an experiment to investigate the effect of RP on achieving joint attention. The experiment compared two conditions, fluent condition and RP condition. The results did not prove the effectiveness of RP on both getting and guiding attention, and questionnaire’s results showed that RP rather deteriorated the impression of the robot.

Keywords: robot agent, communication support, human-robot interaction, and museum guide robot

1. はじめに

人間との対話を目的とするコミュニケーションロボットは、視線や指差しといった非言語的な情報^[1]の伝達が可能であるという特徴を持つ。そのようなロボットの応用例として、博物館等での鑑賞支援ロボットが挙げられる^[2-4]。鑑賞支援ロボットが人間との自然なインタラクションを実現するためには、共同注意の達成が重要である。共同注意とは、話し手と聞き手が同一の対象に視線を向ける状態を指し、両者が注意の対象を共有し円滑なコミュニケーションを行う上で重要な役割を果たす^[5]。

鑑賞支援においては、ロボットが展示物を指し示し、鑑賞者の注意を展示物に誘導することで共同注意が達成される。しかし、鑑賞者の注意が展示物等に集中している場合には、鑑賞者がロボットの指示動作に気付かず、鑑賞者とロボットの間で共同注意が達成されないという問題がある^[6]。本研究ではこの問題に対して、行為の連鎖性^[7]という概念に注目する。人間同士のコミュニケーションにおいて共同注意を達成するには、そこに至るまでに様々な行動が連鎖する。このような連鎖の一つとして、沈黙・言い直しに注目する。

すでに、ロボットが沈黙・言い直しを行うことにより、鑑賞者の注意をロボットに誘導できることは知られている^[8]。しかし、ロボットへの注意誘導の後、ロボットの指示動作により共同注意が達成されるかどうかについては確認されていない。そこで本研究では、沈黙・言い直しの有無が共同注意の達成に与える効果を調査することを目的とする。

2. 関連研究

2.1 沈黙・言い直しに関する社会学的研究

Goodwin^[9-10]は人間同士の対話において、話し手が沈黙・言い直しを用いることにより、聞き手の注意を獲得することを示した。次の例 1、例 2 のトランスクリプトは Goodwin の論文からの引用である。

例 1: 沈黙

A: They've changed- (-----) the China City

B: X_____

例 2: 言い直し

A: Can you bring-? (0.2) Can you bring me here that nylon?

B: X_

このトランスクリプトは、話し手 A の発話の様子と聞き手 B の注意の様子を示している。ここで、A におけるハイフン“-” は 1 つにつき 0.1 秒の時間を表し、話し手が沈黙していることを示す。また、B における空白部分は B が A を見ていない状態を表し、アンダーライン“_” は B が A を見ている状態を表し、“X” は B が A に視線を移動した時点を表している。それぞれの例より、沈黙・言い直しという方略を用いることによって話し手が聞き手の注意を獲得することがわかる。

2.2 鑑賞支援ロボットに関する研究

Pitsch ら^[6]は小型ロボットを鑑賞支援ロボットとして使い、実際に美術館において展示案内を行う実験を実施した。そして、美術館のような環境では他の展示物等に気を取られ鑑賞者の注意がロボットに向いていない状況

^{*1}: 筑波大学

^{*1}: University of Tsukuba

が発生し、そうした場合にはロボットによる指さし等の動作が鑑賞者に気付かれず、注意誘導に失敗するという問題を示した。

この問題に対して川口ら^[8]は、SONY の AIBO を鑑賞支援ロボットとして使用し、通常の発話による説明と、沈黙・言い直しを用いた説明の2条件において、鑑賞者の注意をロボットが獲得できる確率を比較した。その結果、ロボットが沈黙・言い直しをすると、鑑賞者がロボットに注意を向ける確率が高まることを示した。しかし、この研究では沈黙・言い直しによるロボットへの注意獲得のみに着目しており、そこから連鎖的にロボットと人との共同注意が達成できるかどうかはまだ確認されていない。また、この研究ではあらかじめ録音された発話音声を用いており、鑑賞者の状況に応じた沈黙・言い直しの制御は行われていない。

2.3 発話生成に関する研究

Ohshima ら^[11]は聞き手の状態を反映した発話生成システム「Talking-Ally」を実装し、聞き手とともに発話を構築するシステムにより、聞き手の振る舞いがどのように変化するかを検証する実験を行った。このシステムにおいて、話し手の発話開始を示す要素として沈黙・言い直しを用いている。しかし、この研究では話し手と聞き手が相互に会話を組織していくことを目的としており、沈黙・言い直しが注意獲得や注意誘導に与える効果についての検証は行われていない。

2.4 本研究の方針

Pitsch らの示した鑑賞支援ロボットにおける注意誘導の課題に対して、本研究では川口らと同様に Goodwin が示した沈黙・言い直しを用いて注意獲得を行う鑑賞支援ロボットを実装する。沈黙・言い直しの生成に関しては、

Ohshima らと同様に対話者の状況に応じて発話内容を制御する機能を実装する。そしてこのシステムを用いて、沈黙・言い直しによる注意獲得から注意誘導が連鎖的に達成されるかを評価する。

3. システム構成

本研究で実装したシステムの概要を図1に示す。このシステムは鑑賞者の視線を検出し、それに応じて発話内容を制御するものであり、視線センサ、ロボット、音声合成ソフト、C++による制御プログラムから構成される。

3.1 視線センサ

鑑賞者の視線を検出し、鑑賞者がロボットに注意を向けていないときのみ沈黙・言い直しを行うために、視線センサとして OMRON 社の HVC-P2 (広角検出タイプ) を用いた。このセンサは人間の顔を検出すると、センサに対して上下方向 $\pm 90^\circ$ 、左右方向 $\pm 90^\circ$ の範囲で視線角度の推定値を検出することができる。そこで、鑑賞者がロボットを見ていないと判断される視線角度に達したときに沈黙・言い直しを発生させるコマンドを送信する。

3.2 ロボット

本研究で使用したロボットの外観と動作を図2に示す。このロボットは本研究室で開発されたロボットであり、身長51cmで、頭部に2自由度、右腕に2自由度、腰に2自由度を持つ。図2の動画のように、絵画の説明に必要な指差し動作や視線の提示、身体ねじりが可能である。シリアル通信により送信されたコマンドをマイコン (Arduino Pro micro) が受信することで動作する。ロボットの発話を生成する音声合成ソフトとして CeVIO Creative Studio 6 を用いた。C++のプログラムを用い、API を介して音声合成ソフトを制御することで発話を行った。一連の説明は自動で行うものとした。説明用のテキストデータを読み込むことで発話を生成し、説明中あらかじめ設定したタイミングでロボットの動作コマンドが

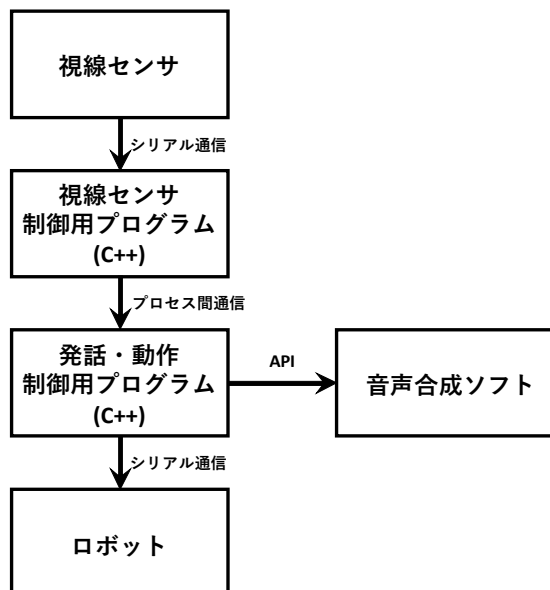


図1 システムの概要

Fig.1 Outline of the System



図2 ロボットの外観と動作

Fig.2 Appearance and Behavior of the Robot

表 1 対象とする連鎖

Table.1 Target Sequence

行為	ロボット	人間
A	沈黙・言い直し	ロボットに注意を向ける
B	指差す	ロボットが指差した方向に注意を向ける

送信されることでロボットが動作する。沈黙・言い直しはロボットの指差し動作の直前に行われる。沈黙・言い直しを行う区間では HVC-P2 による鑑賞者の視線推定値を文節ごとに参照し、ロボットを見ていない場合および視線推定値の検出に失敗した場合に沈黙・言い直しを発生させる。

4. 実験

本研究では、表 1 の行為 A と B の連鎖を対象とし、行為 A が行為 B に与える影響を明らかにするための実験を行った。美術館を想定した環境において、ロボットが絵画の説明を行った。通常の発話による説明（以下、通常発話条件とする）と、沈黙・言い直しを用いた発話（以下、RP 条件とする）の 2 条件で比較を行った。実験に当たっては仮説として以下の 2 点を設定し、これらの確認を行うための評価項目を設定した。

H1. 沈黙・言い直しにより、注意獲得確率が向上する。

H2. 沈黙・言い直しにより、注意誘導時間が短縮される。

4.1 実験環境

Pitsch らの研究で見られたロボットが無視されるといふ状況は、複数の展示物が存在することで、ロボットや説明対象以外に気を取られるような状況であった。そこで本研究では、同様の状況を再現するための環境設定を行った。実験環境を図 3 に示す。

予備実験を行い、ロボット以外に注意が向きやすいかどうかという観点からロボット・絵画の全体図・絵画の部分図の配置を決定した。肖像画は実験参加者の気を逸らすことを意図し、ロボットから離れた位置に配置した。絵画の部分図は縦に 3 つに並べ、ロボットが指差しを行

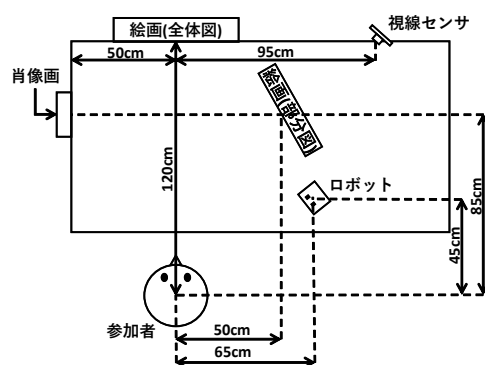


図 3 実験環境

Fig.3 Experiment Environment

うことで実験参加者の注意を誘導できるように配置した。ロボットの発話する音声は、ロボットの動作音を実験参加者に聞こえにくくするためにノイズキャンセリング機能付きヘッドホン SONY MDR-1000X から出力した。

4.2 実験条件

通常発話条件と RP 条件の 2 条件で実験を行った。1 つの絵の説明において、沈黙・言い直しを行う区間を 3 つ設定し、これを評価区間とした。RP 条件では評価区間において実験参加者の視線を検出し、ロボットを見ていない場合に沈黙・言い直しを行った。なお、ロボットの身体動作については各条件で同じものとした。参加者内配置による評価を行うために絵を 2 種類用意した。また、カウンターバランスを取るために絵の説明順を変えて実験を行った。

4.3 実験手順

1 名の参加者に対して絵の説明を 2 回行い、各説明の終了後にアンケートを実施した。説明に用いたスクリプトの例を表 2 に示す。表中の太字下線部分が評価区間であり、RP 条件ではこの区間内で沈黙・言い直しを行った。実験に用いた絵画は、ピーテル・パウル・ルーベンス作「聖母被昇天」と、ミケランジェロ・ブオナローティ作「最後の審判」の 2 作品であった。また、実験参加者は男性 12 名で、いずれも 22~25 歳の大学生と大学院生であった。

4.4 評価方法

仮説 H1, H2 について検証を行うため、沈黙・言い直し判定区間においてロボットが鑑賞者の注意を獲得することに成功した確率（以下、注意獲得確率とする）と、ロ

表 2 スクリプト例

Table.2 Example of Scripts

説明内容	発話	動作
開始	こんにちは。	礼
画家説明	あなたはミケランジェロという芸術家を知っていますか。(中略)	頷き
絵画概要	今回はミケランジェロの「最後の審判」という作品について説明したいと思います。(中略)	身体を絵画に向ける
評価区間 1	まず注目してもらいたいの が、石柱を運んでいる天使たちです。(中略)	指差し
評価区間 2	その中でもひとときわ目を引く のが、人の抜け殻のような皮をもっている聖バルトロマイです。(中略)	指差し
評価区間 3	その中でも特に有名な のが、悪魔の化身とされるヘビをまきつけた地獄の王ミノスです。(中略)	指差し
終了	お聞きいただき、ありがとうございました。	礼

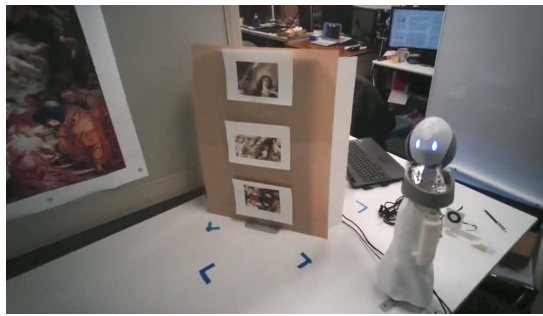


図 4 実験の様子

Fig.4 An Excerpt from the Experiment

ロボットが指差し動作を開始してから鑑賞者がロボットの指差した先を見るまでの時間（以下、注意誘導時間とする）の計測を行った。これらの分析のため、実験中はアイトラッカー (tobii pro glasses 2) を用いて実験参加者の注視方向の記録を行った。また、川口らの論文よりロボットによる沈黙・言い直しによりロボットの印象が悪化する可能性が指摘されていたことから、本研究ではロボットの印象を評価するためのアンケートを実施した。

4.5 実験の様子

実際の実験の様子を図 4に示す。図 4は実験参加者の装着したアイトラッカーにより撮影された動画で、動画中の赤い丸は実験参加者の視線を表示している。

5. 結果と分析

5.1 注意獲得確率

通常発話条件と RP 条件の 2 条件について、注意獲得確率を比較した。各条件における、3 回の評価区間範囲と、3 回分を全て含めた全体についての注意獲得確率を図 5に示す。各回および全体についてそれぞれ χ^2 検定を行ったが結果、いずれの場合も 2 条件の間に有意な差はなかった。

5.2 注意誘導時間

通常発話条件と RP 条件の 2 条件について、注意誘導時間を比較した。各条件における、3 回の評価区間範囲と 3 回分を全て含めた全体についての注意誘導時間の平均を図 6に示す。

発話条件と、評価区間が何回目であるかを要因として二元配置の分散分析（対応あり）を行ったところ、有意差は見られなかった。

5.3 アンケート

各条件終了後に、表 3に示す質問項目についてアンケートを行った。Q1 は、佐久間ら^[12]の研究を参考に、ロボットの印象評価として SD 法を用いて、表 4に示す 10 組の形容詞対について 7 段階で評価するものとした。また Q2~Q5 は、Ohshima ら^[11]の研究を参考に、ロボットのコミュニケーションについて 7 段階で評価する質問項目とした。さらに両方の条件終了後に、Q6~Q8 の 2 条件を比較する質問を行った。

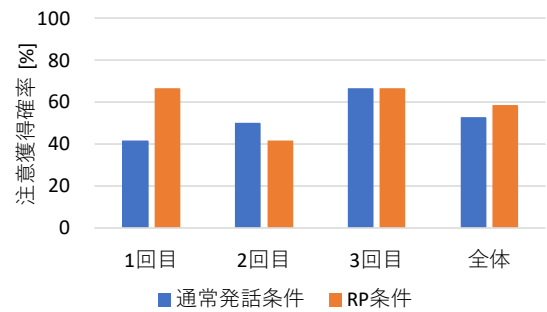


図 5 注意獲得確率

Fig.5 Probability of Getting Attention

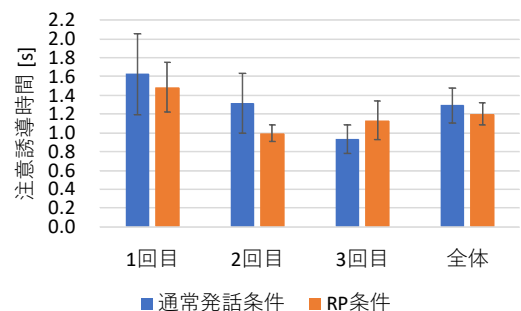


図 6 注意誘導時間

Fig.6 Time of Drawing Attention

表 3 質問項目

Table.3 The Questionnaire

質問	内容
Q1	ロボットに対する印象評価（表 4参照）
Q2	ロボットの説明する内容を理解することができましたか？
Q3	ロボットの説明をもっと聞きたいと感じましたか？
Q4	ロボットの意図を感じましたか？
Q5	ロボットが自分を聞き手として認識していたと感じましたか？
Q6	ロボットがどこを指しているか分かりやすかったのはどちらの条件ですか？
Q7	より理解しやすかったと感じた説明はどちらの条件ですか？
Q8	より好ましいと感じた説明はどちらの条件ですか？

表 4 形容詞対

Table.4 Pair of Adjectives

質問	形容詞対
Q1-1	賢い—愚かな
Q1-2	人間的な—機械的な
Q1-3	敏感な—鈍感な
Q1-4	派手な—地味な
Q1-5	複雑な—単純な
Q1-6	感じのよい—感じの悪い
Q1-7	面白い—つまらない
Q1-8	積極的な—消極的な
Q1-9	好きな—嫌いな
Q1-10	自然な—違和感のある

表 5 因子負荷量

Table.5 Factor Loading

形容詞対	Factor1	Factor2
好きな—嫌いな	0.836	0.029
賢い—愚かな	0.82	0.101
自然な—違和感のある	0.771	0.067
感じのよい—感じの悪い	0.719	-0.017
敏感な—鈍感な	0.651	0.358
積極的な—消極的な	0.375	0.925
面白い—つまらない	0.285	0.609
複雑な—単純な	-0.424	0.545
派手な—地味な	-0.014	0.472
人間的な—機械的な	0.005	0.38

Q1 の結果について因子分析(最尤法、バリマクス回転)を行い、2 つの因子を抽出した。各形容詞対についての因子負荷量を表 5 に示す。因子負荷量の絶対値が 0.5 以上のものをその因子に属するとして形容詞対の分類を行った。各因子に属する形容詞対から、因子 1 は好感度、因子 2 は生物性であると解釈した。また、通常発話条件と RP 条件での因子得点を図 7 に示す。各因子得点について、発話条件 (2 水準) を要因とする t 検定を行ったところ、因子 1 において通常発話条件の因子得点が RP 条件よりも有意に高くなることが示された ($p=0.0051$)。

Q2~Q5 について回答者の回答の平均を図 8 に示す。Q2~Q5 について、発話条件 (2 水準) を要因とした t 検定を行った。その結果、Q3 において通常発話条件の評定の平均が RP 条件よりも高くなる有意な傾向がみられた (Q3: $p=0.096$)。

Q6~Q8 について質問項目ごとの票数を図 9 に示す。Q6~8 について、二項検定 (片側検定) を行った。その結果、Q8 において通常発話条件の票数が RP 条件よりも有意に多い傾向がみられた (Q8: $p=0.073$)。

6. 考察

6.1 注意獲得確率

5.1 節の結果より、2 条件間で注意獲得確率に有意差はなかった。したがって、仮説 H1 は否定された。本実験において 2 条件間で差が見られなかった要因として、実

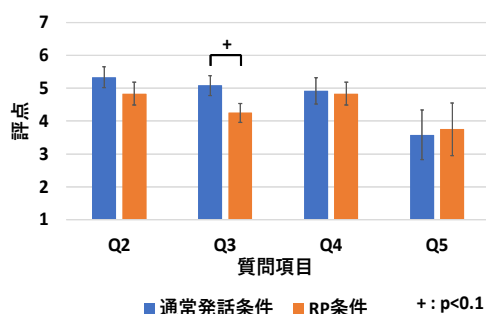


図 8 アンケート結果 Q2-5

Fig.8 Questionnaire results Q2-5

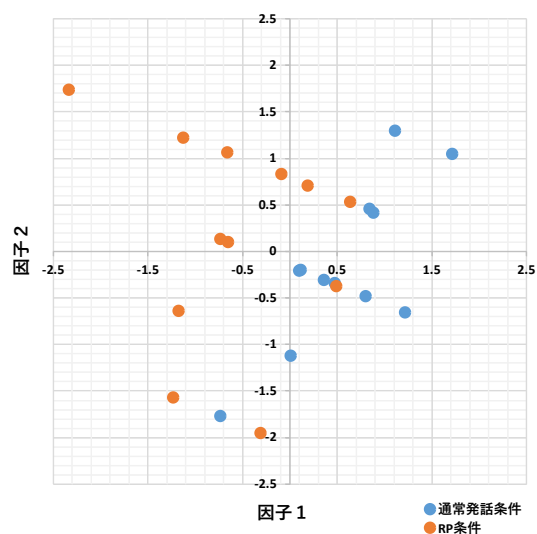


図 7 因子得点

Fig.7 Factor Score

験方法と展示物の種類の 2 つが挙げられる。

まず実験方法について、本実験は実験室実験であったことから、実験参加者にロボットの説明を聞かなければならないという心理が働くことで、ロボットに注意が向きやすかったと考えられる。そのため、評価区間での前提となるロボットが無視されるという状況が作り出せていなかったと考えられる。

次に、展示物の種類について、本研究ではロボットが説明を行う展示物として絵画を選択した。しかし、絵画は受動的に鑑賞するものであるという性質を持つことから、説明を行っている最中にロボットを気にしなくなるほどに絵画に集中するとは考えにくい。よって、実験時にロボットから注意が逸れるという状況が生じにくかったと考えられる。

以上の考察から本実験の改善点として、博物館などでのフィールド実験にすること、展示物を実験参加者が実際に体感するものにする等により、実際に課題として設定したロボットが無視される条件において評価を行う必要がある。

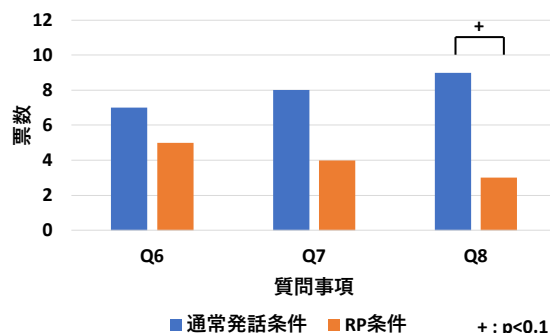


図 9 アンケート結果 Q6-8

Fig.9 Questionnaire results Q6-8

6.2 注意誘導時間

5.2 節の結果より、2 条件間で注意誘導時間に有意差はなかった。したがって、仮説 H2 は否定された。このような結果となった要因として実験配置と、ロボットの動作音が考えられる。

実験配置について、本実験では実験参加者が絵画の全体図を見ている際に、ロボットを周辺視野で捉えることができた。ゆえに、ロボットに注意を向けていない場合でも、ロボットが指示動作を行ったことを知ることができたことで、注意誘導時間が長くなることがなかったと考えられる。

次に、ロボットの動作音について、本実験ではノイズキャンセリング機能付きのヘッドホンから発話音声を出力していたが、ロボットのサーボモータが動作する音を完全に遮断することはできなかった。また、実験室は静かな環境であったことから、実際の美術館等と比較してロボットの動作音が聞こえやすかったといえる。ロボットの指示動作時に動作音が聞こえたことで、ロボットへの注意獲得が失敗していてもロボットの動きを知ることができたため、2 条件で差が生じにくかったと考えられる。

本実験の改善点として、イヤホンと防音用イヤーマフを用いてより遮音性を高めることや、発話音声に加えて実際の美術館等を想定した環境音等を流すことでロボットの動作音をマスキングすることが考えられる。

6.3 アンケート

5.3 節の結果より、沈黙・言い直しを行った際にロボットに対する好感度が悪化することが示唆された。アンケートの自由記述において、沈黙・言い直しによってロボットが故障していると感じたという主旨の記述がみられたことから、ロボットの故障というネガティブな印象を持たせたため、好感度が下がったと考えられる。

また本実験においては、鑑賞者がロボットの方を見えない場合や視線検出に失敗した場合に沈黙・言い直しを行った。その際、視線検出の精度が低く、鑑賞者がロボットを見ているにもかかわらず沈黙・言い直しを行うことがあったため、より印象が悪くなったと考えられる。そこで、より検出精度の良い視線センサを利用した上で再実験を行う必要がある。さらに、本実験では 1 つの評価区間において最大 3 回の沈黙・言い直しを行った。それに対してアンケートにおいて、沈黙・言い直しの回数を多く感じたという主旨の記述がみられたことから、1 つの評価区間での沈黙・言い直しの回数が適切でなかった可能性がある。ロボットに対する好感度が著しく低下しないような沈黙・言い直しの使用回数がどの程度であるかを調査する必要がある。

7. おわりに

本研究では、沈黙・言い直しの有無が共同注意の達成

に与える効果を調査することを目的として、ロボットが絵画を説明するという設定で実験を行った。実験は、通常発話による説明と沈黙・言い直しを用いた説明の 2 条件について行った。実験結果から沈黙・言い直しの有効性を示すことはできなかった。一方で実験参加者に対するアンケートから、沈黙・言い直しによりロボットに対する印象が悪化することが示唆された。本研究では実験における環境設定が不十分であったと考えられることから、今後は博物館等の実環境における実験を行う必要がある。また、視線検出精度を高めた上で、ロボットの印象を著しく悪化させないような沈黙・言い直しの使用回数がどの程度であるかを調査する必要がある。そして、対話的行為の連鎖性を利用した人とロボットのインタラクションのための手法として確立することを目指す。

謝辞

実験に協力してくださったグループウェア研究室の皆様と、お忙しい中実験に参加していただいた皆様に深く感謝いたします。

参考文献

- [1] Vargas, M.: Louder than words: An introduction to nonverbal communication; Iowa State Pr (1986).
- [2] Kuno, Y., Sadazuka, K., Kawashima, M., Yamazaki, K., Yamazaki A., Kuzuoka, H.: Museum guide robot based on sociological interaction analysis; Proceedings of the SIGCHI conference on Human factors in computing systems, p. pp.1191-1194 (2007).
- [3] Shiomi, M., Kanda, T., Ishiguro, H., Hagita, N.: Interactive humanoid robots for a science museum; Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, p. pp.305-312 (2006).
- [4] Yamaoka, F., Kanda, T., Ishiguro, H., Hagita, N.: How close?: model of proximity control for information-presenting robots; Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction, p. pp.137-144 (2008).
- [5] Moore, C., Dunham, P.: Joint attention: Its origins and role in development; Psychology Press (2014).
- [6] Pitsch, K., Wrede, S.: When a robot orients visitors to an exhibit. Referential practices and interactional dynamics in real world HRI; RO-MAN: The 23rd IEEE International Symposium on. IEEE, p. pp.36-42 (2014).
- [7] Schegloff, E.: Sequence organization in interaction: Volume 1: A primer in conversation analysis; Cambridge University Press (2007).
- [8] 川口, 葛岡, 鈴木, 中尾, 山下, 山崎: 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, 14(3), pp.257-263 (2009).

- [9] Goodwin, C.: The Interactive Construction of a Sentence in Natural Conversation; *Everyday Lang. Stud. Ethnomethodology*, vol. 37, p. pp.97-121 (1979).
- [10] Goodwin, C.: *Conversational organization: Interaction between speakers and hearers*; Academic Press (1981).
- [11] Ohshima, N., Ohyama, Y., Odahara, Y., De Silva, P. R. S., Okada, M.: Talking-Ally: The Influence of Robot Utterance Generation Mechanism on Hearer Behaviors; *International Journal of Social Robotics*, Volume 7, pp 51–62 (2015).
- [12] 佐久間, 加藤: ユーザ報酬付与傾向を反映する擬人化エージェントによるボールを使ったやりとり遊び; *人工知能学会論文誌*, 31 巻, 6 号, SP-A (2016).