

多人数対話におけるロボットの視線行動に基づく 発話権と対話場のデザイン

Coordinating Turn-Taking and Talking in Multi-Party Conversations by Controlling a Robot's Eye-Gaze

佐藤良^{1*} 竹内勇剛¹

Ryo SATO¹ Yugo TAKEUCHI¹

¹ 静岡大学大学院情報学研究科

¹ Graduate School of Informatics, Shizuoka University

Abstract: The purpose of this study is realizing smooth turn-taking and discussion in multi-party conversations by designing the physical behavior focused on gazing of a robot participating as a side-participant. In this paper, we investigate how humans adjust participation-role-taking by watching the face direction of others include a robot without uttering in a coordinately-turn-taking game. As a result, the appropriate robot's gazing can prompt the participants to perform and contribute to smooth participation-role-taking. We believe our study will effectively produce desirable discussion in multi-party conversations such as a collaborative learning situation.

1 はじめに

人とのコミュニケーションを指向したロボットが誕生し、人の日常生活の中に入りつつある。ロボットと人、ロボットを介した人同士のコミュニケーションを現実的なものにするには、ロボットが人の対面対話状況に構造的・社会的に適応できることが必要になる。なぜならば、対面対話は人にとって最も自然かつ表現上の制約が少ないコミュニケーション形態であるからである。

我々の日常的な対話コミュニケーションは必ずしも2人だけの空間で成り立つものではなく、協調学習場面に代表される多人数対話の機会が多い。一対一の対話の場合には話し手の決定に伴い自動的にもう1人が発言の受け手となるため、参与役割が自明である。それゆえに、発言が誰に向けられているのかを意識しなくても話者交替を円滑に行うことが可能であった。しかし、多人数対話では受け手となりうる候補者が複数人存在するため発言を聞いている人の参与役割が複雑化する[1]。そのため、人と人、ロボットと人の円滑な話者交替 (turn-taking) の実現が問題となる[2]。

これまで話者交替に着目した対話ロボットの研究が数多く行われてきている。小林らが開発したROBITA[2]は多人数対話場面における人の視線を基に発話のアドレス先を処理したり、人の発話内容を理解したりする

ことができる。それによって、ロボット自体が会話に加わるのに適した発話開始のタイミングを予測し、適切なタイミングで発話権を獲得することを実現している。船越らはロボットに設置した明滅光源によりロボットの内部状態を表出させ、ロボットと人の発話の衝突を抑制することで話者交替の円滑化を図った[3]。

これらの研究では、多人数対話における周囲の対話環境を受動的に推定することによって円滑な話者交替の実現を図ったが、積極的に対話環境を制御することは想定されていない。しかし、多人数対話では話者の候補者が複数人存在するにも関わらず話者の偏りが生じたり、話者交替が円滑に行われないことによって、健全な対話の場が維持されないことがある。

通常は司会者に相当する人が発話権の委譲を調整する権限と権限を適切に行使できる能力を有することで、これらの状況が収拾され、質の高い議論が実現される。このことは、司会者の存在と能力が多人数対話を円滑にするための重要な役割を担うことを意味するが、このような能力を有したロボットは研究されていない。そこで本研究では、多人数対話の参与者の発話権を円滑に交替させられるロボットの振舞いを明らかにすることを目的とする。

人同士の対話コミュニケーションでは、言語情報だけでなく様々な非言語的な情報が相互に伝達される。その中でも視線は、その行動主体の発話のアドレス判断に用いられる[4]ため、参与者の視線行動は、円滑な話者交替を実現するトリガとなり得る。それに、人は人

*連絡先： 静岡大学大学院情報学研究科
静岡県浜松市中区城北3丁目5-1
E-mail: gs12019@s.inf.shizuoka.ac.jp

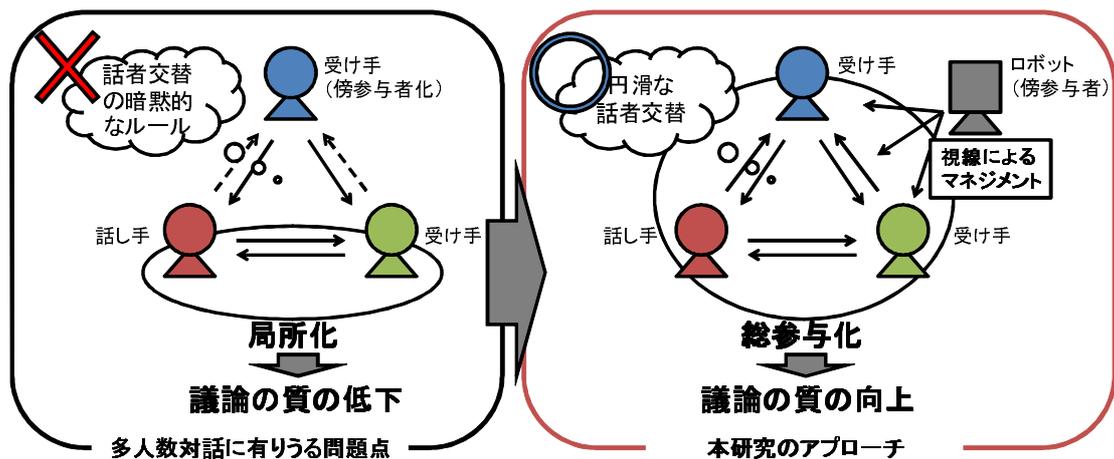


図 1: 本研究のアプローチのイメージ

以外の人工物の視線に対しても、それが疑似的に作られたものであるにも関わらず、話者アドレッシングの判断に用いられる [4] .

そこで本研究では、これらの知見を応用し多人数対話における参加者の発話権の委譲を調整するためのロボットの視線行動に基づくインタラクションモデルを検討する。その対象として言語情報や非言語情報が互いに行き来する一般的な多人数対話場面に焦点を当てた場合には、言語的な情報の文脈に依存した知見となる恐れがある。そこで、本稿では多人数参加型のバランスゲームを行う課題を通じて、対話環境を複雑化し得る言語情報を排除した上で、ロボットの視線行動がゲーム参加者らの役割交替にもたらす影響を調査する。

それにより、円滑な話者交替を円滑にするための、参加者とロボットの非言語的な表現に基づく汎用的なモデルの構築を実現できると考えられる。これが明らかになることで、協調学習場面に代表される多人数対話における話者の偏りをなくし、対話が効率的に行われる方向への話題の誘導を自然な形で達成する認知的作用を提供できると考えられる (図 1) .

2 多人数対話の参与構造と話者交替

一般的に対話の参加者は大きく分けて話し手が話し手の発言を聞いている人のいずれかになる。一対一の対話の場合には話し手の決定に伴って自動的にもう一方の人が聞き手となるため、次話者になる人が明確である。しかし、多人数対話の場合には、1人の話し手に対して、その受け手の候補になり得る者が複数人存在するため、次話者になる人が誰であるかは明白でない。Goffman や Clark らによると、会話の参加者は承認された参加者 (ratified participant) と立ち聞き者 (overhearer) と

に大きく分類され、更に前者は話し手 (speaker), 受け手 (addressee), 傍参加者 (side-participant) に分類される [5][6] (図 2) . 参加者が 3 人以上の対話では、話し手の発話を聞いている人は受け手、傍参加者、立ち聞き者のいずれかの参与役割を担う [1] . 受け手とは、話し手によって発話の矛先が向けられていて、何らかの応答を期待されている参加者であり、それ以外が傍参加者である。傍参加者は当面の会話で発話を行ったり、応答を期待されたりしているわけではないが、会話に加わることは許可されている。なお、この分類における承認された参加者とは、会話に参加していることが他の参加者によって知られている者を指す。ここで本研究では、対話の場を「対話の当事者として相互に言語情報や非言語情報をやり取りし、相手との関係や情報の意図を理解している範囲」として定義する。すなわち、本研究における対話の場とは、先述した参与構造における承認された参加者によって対話がなされる場を指す。このように考えると、一対一の対話の場合には対話の参加者が全員対話の当事者であるのに対し、多人数対話の場合には必ずしも全員が対話の当事者であるとは限らない。そのため、多人数対話の場合には言語情報や非言語情報を用いて対話の参加者がどの参与役割を担っているのかを判断する必要がある。このように、多人数対話における非言語情報は、メッセージの補足としてだけでなく、対話への参与状態を示すためにも用いられると考えられる。

現在の話し手から次の話し手である聞き手に話し手が移ることは話者交替 (turn-taking) と呼ばれており、我々は、話者交替を繰り返すことによって対話を行っている。話者交替には話し手と聞き手の発話 / 非発話の組合せによって「継続」「重複」「沈黙」「交替」の 4 種類の現象がある [8] (表 1) . そして、一般的にはあ



図 2: 参与構造

表 1: 話し手と聞き手の発話 / 非発話による現象

		話し手	
聞き手	発話	重複	交替
	非発話	継続	沈黙

らかじめ発話の順序を決めなくても重複や沈黙といった現象は発生しない [9] .

参与役割の交替は話者交替のタイミングで生じ、参与役割を決定づける要因には視線等の非言語的な情報の影響が大きく、それによって話者交替が調整されることが示唆されている [10][11][12] . その中でも視線は、その行動主体の発話のアドレス判断に用いられ [4] , 次話者の選択に用いられる [11] . そのため、多人数対話の参与者の視線行動は、話者交替を実現するトリガとなり得る . そこで本稿では、視線行動と話者交替について着目し、第 3 章で考察する .

3 視線行動と話者交替

多人数対話の場には、積極的に対話に参加している人や一見対話に参加していないように見えても実際には対話に注意を示している人等、様々な立場の人が存在する . しかし、一般的に対面対話をする場合には話し手は聞き手、聞き手は話し手の方向に身体や顔を向ける . この時、話し手は自身の話の聞き手として望ましい相手や次に話し手になってもらいたい相手の顔を見て話をすることがある . 一方で、聞き手は自身にとって関心のある話であれば話し手の顔を見て話を聞くが、そうでなければよそ見をすることがある . このように、視線方向は多人数対話における対話の注意を示す重要な手がかりであり、聞き手の視線には発話したい / したくないという心理が表出される [13] . そして、参与者同士が表現した心理を参与者が相互に理解し合うことによって、円滑な話者交替が実現する [7] .

このように様々な情報を伝える視線だが、送り続ける時間やタイミング、周囲の状況、相手との距離等によってその意味は大きく変化する . Kendon によると、話し

手が次話者を注視することで話者交替を合図し、次話者が相互注視によってそれを受け入れることで話者交替が成立する [11] . また、榎本らは非言語行動が多人数対話における次話者選択手段として機能し得るかどうかを検討するために、3 人の参与者による会話の分析を行った . その結果、現話者は次話者を視線を送っていることが多く、他の参与者全員が現話者を見つめている時に現話者が視線を送った方向に位置する参与者が次話者になることが示唆された [14] . これらのことから、話者交替のタイミングにおける参与者の視線行動は次話者のセレクトに大きな影響を及ぼすと考えられる . そして、相互注視が成立しているのか、視線を送っているだけなのかといった各参与者の互いの視線状況も同様に次話者のセレクトを左右すると考えられる .

Kendon によると視線には対話開始の手掛かりとしての機能以外に以下の 3 つの機能があるという [11] .

- モニタリング機能 (monitoring) : 発話の継続・終了を聞き手の凝視の有無に基づき確かめる機能
- 調整機能 (regulation) : 聞き手の会話に対する好感度を聞き手の視線行動を基に察知し、発話内容を調整する機能
- 表出機能 (expressive) : 会話をもたらす効果の良し悪しを話し手に伝える機能

このうち、モニタリング機能と調整機能は話し手、表出機能は聞き手の視線がもつ機能である . このように、参与者は話し手の発話中における互いの視線行動を観察しながら、互いの発話に対する興味・関心や発話欲求の有無を推察する . そして、得られた情報を基に、自身が次の話し手として望ましいのか否かを判断しながら、話者交替が行われる . すなわち、話者交替のタイミングだけでなく、話し手の発話中における参与者の視線行動は次話者のセレクトに大きな影響を及ぼすと考えられる .

これまで本稿では、多人数対話における参与構造について述べ、それを交替する機会である話者交替と参与者の視線との関わりを考察した . このことから、多人数対話に参加するロボットの視線行動を適切に制御により話者交替の円滑化を実現できる可能性が示唆される . そこで本稿では、多人数対話における発話を行う、話者交替を行為の交替と仮定し、対話環境を複雑化し得る言語情報を排除した上でロボットの視線行動が参与者の役割交替に及ぼす影響を検証する . 第 4 章では、多人数参加型のバランスゲームを行う課題を通じて、ロボットの視線行動がゲーム参加者らの役割交替にもたらす影響を調査した実験について述べる .



図 3: バランスゲームの外観

4 実験-ロボットとの視線インタラクションによる行為の役割交替の変化及びその円滑化の検証-

4.1 目的

多人数参加型のバランスゲームにおける行為者の役割交替を観察することを通じて、多人数対話コミュニケーションにおける参与者とロボットの視線行動に基づくインタラクションが参与者間の役割交替に及ぼす影響を検証し、そのモデル化を行う。また、ロボットの視線行動の有無が役割交替の円滑化に及ぼす影響を検証する。

4.2 実験協力者

情報学を専攻する日本人大学・大学院生 24 名 (3 名 × 8 組, 男性 20 名, 女性 4 名, 全員右利き) である。

4.3 実験課題: 沈黙環境下における協調的交替ゲーム

4.3.1 協調的交替ゲーム

実験協力者 3 名 1 組で、バランスゲームで勝負を行う (図 3)。中央のボールの内部には、44 個のビー玉が詰まっており、ボールに刺さっている色の異なる 42 本のスティック (赤, 青, 黄, 橙, 緑, 紫の 6 色, 各 7 本) により支えられている。ボールの下部には穴が開いており、スティックの数が少なくなると支えられていたビー玉が落ちる仕組みになっている。

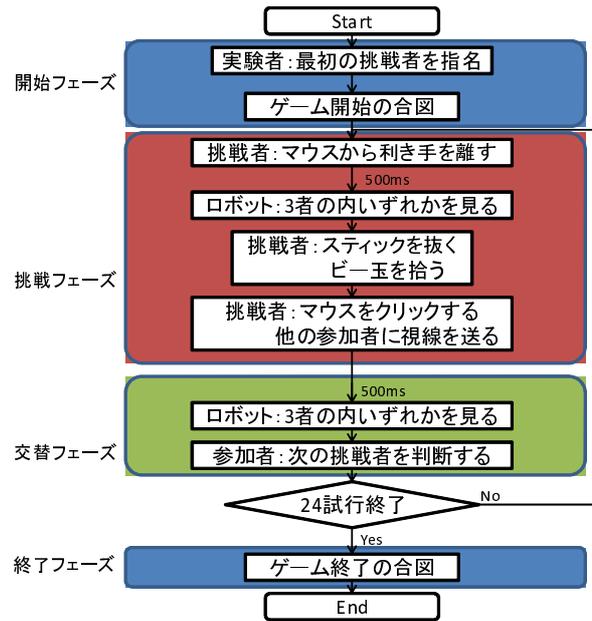


図 4: 実験手順

4.3.2 実験手順

まず、ゲームの最中は基本的に利き腕で手元のマウスをクリックし続けており (左右どちらのボタンをクリックするのは任意)、もう一方の腕は腰元に配置する。また、ゲームの最中は声を一切出してはならない。バランスゲームの手順を以下図 4 に示す。ゲームの進行は開始フェーズ、挑戦フェーズ、交替フェーズ、終了フェーズの 4 つのフェーズで行われる。

4.3.3 開始フェーズ

まず、開始フェーズではバランスゲームへの最初の挑戦者 1 名が実験者から指名される。その後、ゲーム開始の合図の音声スピーカーから流れ、挑戦フェーズに移行する。

4.3.4 挑戦フェーズ

挑戦フェーズでは、挑戦者は手元のマウスから利き手を離し、中央のボールからスティックを 1 本抜く (必要であれば、ゲームの下の回転台を回しても良い)。この時、一度手にしたスティックは必ず抜かなくてはならない (別のスティックに手をかけてはならない)。抜いたスティックは手元の容器 (縦 10cm, 横 10cm, 高さ 5cm) に入れる。この時、ビー玉が落ちた場合には、それらを回収し、同様に手元の容器に入れる。その後、再度利き手で手元のマウスをクリックし、他の参加者

に視線を送ることにより、自分の挑戦が終了した旨を伝え、交替フェーズに移行する。

4.3.5 交替フェーズ

交替フェーズでは、参加者同士でその場その場の状況から次の挑戦者を判断する。この時、声を一切出してはならないというルールに基づき、参加者の視線とロボットの視線（後述）のみを頼りに次の挑戦者を判断する。挑戦者の決定に伴い、再度挑戦フェーズに移行し、挑戦者となった参加者は挑戦を始める。その後、挑戦フェーズと交替フェーズを24回行った後、終了フェーズに移行する。

4.3.6 終了フェーズ

終了フェーズでは、開始フェーズで流れた開始の合図の音声と同様の音声ゲーム終了の合図としてスピーカーから流れ、ゲームが終了する。これらの手順に沿って合計4ゲーム行う。

4.3.7 勝敗条件

参加者はスティックを抜くたびに、スティックの色に合った点数が加点される（赤7点、青6点、黄5点、橙4点、緑3点、紫2点）。一方で、ビー玉が落ちてしまった場合には、ビー玉の数×1点が減点される。1ゲーム終了時の点数について以下の基準を満たした参加者が勝者となる。

- ゲーム終了時に最も多くの点数を獲得する。
- 最下位との点差が10点以下である。

このような勝利条件を設けることにより、参加者には積極的な挑戦意欲だけでなく、挑戦者になる参加者がある程度限定せずに、互いに協力し合うことが求められる。

4.4 実験環境

実験環境を以下図5に示す。3人の実験協力者（H1, H2, H3）と1体のロボット（R）は正方形に位置し、120cm四方のテーブルを囲んで互いに身体を向け合っている。この距離は、Hallによる対人距離の分類の個体距離に分類され[15]、友人同士のような個人的な会話ととられる距離に相当する。テーブルの中央にはバランスゲームを配置し、ゲームと実験協力者の間にはそれぞれ抜いたスティックと落ちたビー玉を入れる容器とマウスを配置する。ゲーム内におけるマウスのクリッ

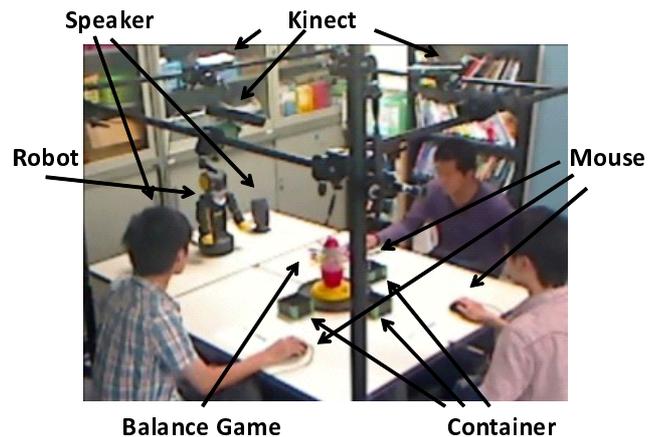


図 5: 実験環境

クの有無と手が離された時間を計測することで、挑戦者になった参加者の判別とその決定に要した時間を計測する。各実験協力者の頭上の前方には、それぞれに対面する形で Kinect センサを配置し、Face Tracking SDK(Microsoft 社)を基に実装されたプログラムにより、実験中の各々の視線方向を計測する。そして、検出された視線方向を4方向（左（人またはロボット）、中央（人またはロボット）、右（人またはロボット）、下（バランスゲーム））に分割する。

4.5 ロボット

実験で用いるロボットは小型ヒューマノイドロボット Robovie-mR2 (ATR Creative[16], 図6)である。ゲームにおける挑戦フェーズと交替フェーズにおいて、視線方向の代用として頭部方向を実験協力者3名の内、いずれかが位置する方向にランダムに変化させる。ただし、ロボットの視線行動の有無はゲームにより異なり、各ゲームの前半、後半それぞれ12試行ずつにおいて動作の有無が変化する（後述）。ゲームの順番はランダムに施行される。

4.6 実験条件と観察項目

4.6.1 ロボットとの視線インタラクション（2要因被験者内計画）

「挑戦フェーズ視線要因」と「交替フェーズ視線要因」の2要因とし、いずれも被験者内計画で行う。「挑戦フェーズ視線要因」とは、挑戦フェーズにおける各参加者とロボットの視線によるやり取りである。また、「交替フェーズ視線要因」とは、交替フェーズにおける各参



図 6: ロボットの外観

表 2: 実験条件

水準	視線のやり取り
EC	フェーズ内でロボットと相互注視が成立する
TO	フェーズ内でロボットから注視される
FROM	フェーズ内でロボットを注視する
NONE	フェーズ内で互いに注視しない

加者とロボットの視線によるやり取りである。どちらの要因も「EC水準」「FROM水準」「TO水準」「NONE水準」の4水準で構成される。各水準の意味を以下表4.6.1に示す。

ロボットが視線行動をした場合において、条件毎の交替フェーズにて各参加者が挑戦者になる割合を観察項目とする。それにより、各参加者とロボットの視線に基づくインタラクションが行為者の役割交替に及ぼす影響を検証する。

4.6.2 ロボットの視線行動の有無（2要因被験者内計画）

「動作有無要因」と「前後半要因」の2要因とし、いずれも被験者内計画で行う。「動作有無要因」とは、各ゲームの前半、後半におけるロボットの視線行動の有無の違いに基づくゲームの違いを意味する。「有り有り水準」「有り無し水準」「無し有り水準」「無し無し水準」の4水準で構成される。「前後半要因」とは、各ゲーム合計24試行の内、前半12試行と後半12試行を意味する。「前半水準」「後半水準」の2水準で構成される。以下表4.6.2はこれらの実験条件をまとめたものである。

各条件において、各参加者が挑戦者になった場合にその決定に要した時間（交替フェーズ開始から終了までの経過時間）を観察項目とする。それにより、ロボッ

表 3: 実験条件

		動作有無要因			
		有り-有り	有り-無し	無し-有り	無し-無し
前後半要因	前半	条件1	条件2	条件3	条件4
	後半	条件5	条件6	条件7	条件8

トの視線行動の有無が行為者の役割交替の円滑化に及ぼす影響を検証する。

4.7 仮説と予測

4.7.1 ロボットとの視線インタラクション

ロボットの視線によるアドレッシングにより、役割交替時に視線を向けられた参加者の挑戦が促される。さらに、ロボットによるアドレッシングに対して視線による応答（アイコンタクトの実現）が行われることで、よりその参加者の挑戦が促される。以上の仮説から以下の順番で挑戦しようとする参加者の割合が増加する。

- 挑戦フェーズ視線要因
 - EC > FROM > NONE > TO
- 交替フェーズ視線要因
 - EC > FROM > NONE > TO

4.7.2 ロボットの視線行動の有無

ロボットが視線行動をすることにより、参加者間の役割交替のための判断を容易にする。それにより、以下に示すようにゲーム間で交替フェーズに要した時間が変化する。

- 有り 有り < 有り 無し, 無し 有り < 無し 無し
- 有り 無し前半 < 有り 無し後半
- 無し 有り前半 > 無し 有り後半

4.8 実験結果と考察

4.8.1 各参加者が挑戦者になった割合

交替フェーズにおいて各参加者が挑戦者になった割合について2要因分散分析を行った。その結果、挑戦フェーズ視線要因の主効果の有意傾向 ($F_{(3,69)} = 2.64, p < .10$)と交替フェーズ視線要因の主効果 ($F_{(3,69)} = 14.59, p < .01$)が得られた。更に、2つの要因の交互作用 ($F_{(9,207)} =$

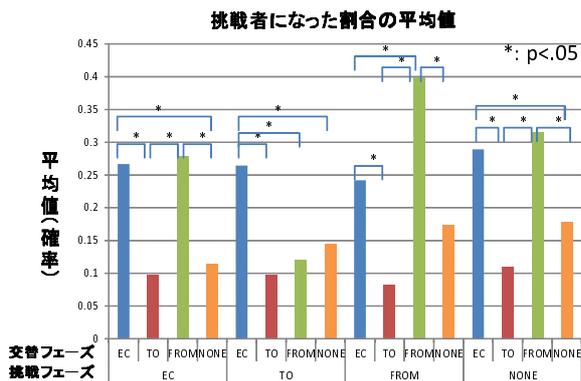


図 7: 挑戦者になった割合の平均値 (交替フェーズ視線要因の単純主効果)

3.25, $p < .01$) が得られた。挑戦フェーズ視線要因の各水準における交替フェーズ視線要因の単純主効果を検証するため、下位検定を行った。なお、多重比較を行うための方法として LSD 法を用いた。その結果を図 7 に示す。

まず、挑戦フェーズ視線要因の EC 水準において、交替フェーズ視線要因の単純主効果が得られた ($F_{(3,69)} = 7.73, p < .01$)。多重比較の結果、EC 水準と FROM 水準が TO 水準と NONE 水準よりも大きいという結果が得られた。また、挑戦フェーズ視線要因の FROM 水準において、交替フェーズ視線要因の単純主効果が得られた ($F_{(3,69)} = 3.90, p < .05$)。多重比較の結果、EC 水準が他水準よりも大きいという結果が得られた。それから、挑戦フェーズ視線要因の TO 水準において、交替フェーズ視線要因の単純主効果が得られた ($F_{(3,69)} = 11.59, p < .01$)。多重比較の結果、EC 水準は TO 水準よりも大きく、FROM 水準は他水準よりも大きいという結果が得られた。そして、挑戦フェーズ視線要因の NONE 水準において、交替フェーズ視線要因の単純主効果が得られた ($F_{(3,69)} = 7.09, p < .01$)。多重比較の結果、EC 水準と FROM 水準が TO 水準と NONE 水準よりも大きいという結果が得られた。

この実験結果から、挑戦フェーズにおいてロボットと相互注視が成立した参加者は、交替フェーズにおいて相互注視の成立 / 不成立に関係なくロボットに視線を向けられることで、より挑戦を促されたということが分かった。それに対し、挑戦フェーズにおいてロボットに視線を向けたがロボットから視線を向けられなかった参加者は、交替フェーズにおいてロボットとの相互注視が成立することにより最も挑戦が促されたということが分かった。また、挑戦フェーズにおいてロボットに視線を向けなかったが、ロボットから視線を向けられた参加者は、交替フェーズにおいてロボットのみ

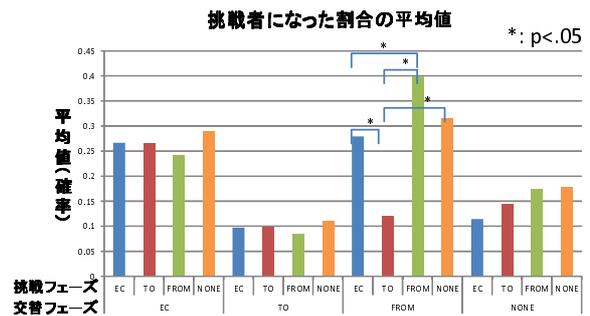


図 8: 挑戦者になった割合の平均値 (挑戦フェーズ視線要因の単純主効果)

から視線を向けられた場合が最も挑戦が促されたということが分かった。そして、挑戦フェーズにおいてロボットに対する、あるいはロボットからの注視が行われなかった参加者は、交替フェーズにおいてロボットに視線を向けられることで、より挑戦を促されたということが分かった。これらの結果から、各参加者は挑戦フェーズにおけるロボットとの視線インタラクションの違いによって、挑戦が促されるための交替フェーズにおけるロボットとの視線インタラクションの違いが生じることが明らかとなった。これは、仮説を一部支持するものである。

続いて、交替フェーズ視線要因の各水準における挑戦フェーズ視線要因の単純主効果を検証するため、下位検定を行った。なお、多重比較を行うための方法として LSD 法を用いた。その結果を図 8 に示す。

まず、挑戦フェーズ視線要因の EC 水準において、交替フェーズ視線要因の単純主効果は得られなかった ($F_{(3,69)} = 0.29, n.s.$)。また、挑戦フェーズ視線要因の TO 水準において、交替フェーズ視線要因の単純主効果は得られなかった ($F_{(3,69)} = 0.12, n.s.$)。それから、挑戦フェーズ視線要因の FROM 水準において、交替フェーズ視線要因の単純主効果が得られた ($F_{(3,69)} = 9.93, p < .01$)。多重比較の結果、TO 水準が他水準より小さく、FROM 水準が EC 水準よりも大きいという結果が得られた。そして、挑戦フェーズ視線要因の NONE 水準において、交替フェーズ視線要因の単純主効果は得られなかった ($F_{(3,69)} = 0.73, n.s.$)。この実験結果から、交替フェーズにおいてロボットに視線を向けないが、ロボットから視線を向けられた参加者は、その前の挑戦フェーズにおいて同様であった場合に最も挑戦が促されることということが分かった。

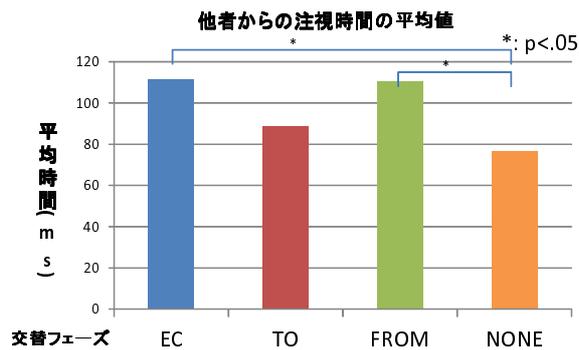


図 9: 他者からの注視時間の平均値 (交替フェーズ視線要因の主効果)

4.8.2 各参加者の他者からの注視量

各参加者とロボットの視線に基づくインタラクションにより、前述の結果が得られた理由を調査するために、各参加者が挑戦フェーズ、交替フェーズにおいて実験条件毎の他者 1 人あたりからの注視時間の平均値に対して 2 要因分散分析を行った。

交替フェーズにおける他者からの注視時間の平均値に対する分析の結果、挑戦フェーズ視線要因の主効果は得られなかった ($F_{(3,69)} = 1.37, n.s.$)。しかし、交替フェーズ視線要因の主効果が得られた ($F_{(3,69)} = 3.74, p < .05$)。LSD 法による多重比較の結果、EC 水準と FROM 水準が NONE 水準よりも大きいという結果が得られた (図 9)。2 つの要因の交互作用は得られなかった ($F_{(9,207)} = 0.65, n.s.$)。この結果から、交替フェーズにおいて、相互注視の成立 / 不成立に関らず、ロボットが参加者に視線を向けた場合には、交替フェーズにおいてその参加者への他者による注視を促すということが分かった。それにより、ロボットが視線を向けた参加者の挑戦が促されたということが示唆された。

4.8.3 挑戦者の決定に要した時間

各ゲームにおける各挑戦者の交替フェーズ開始から終了までに要した時間について 2 要因分散分析を行った。その結果、動作有無要因の主効果が得られた ($F_{(3,69)} = 3.83, p < .05$)。また、前後半要因の主効果は得られなかった ($F_{(1,23)} = 0.82, n.s.$)。更に、2 つの要因の交互作用が得られた ($F_{(3,69)} = 6.55, p < .01$)。動作有無要因の各水準における前後半要因の単純主効果を検証するため、下位検定を行った。その結果を図 10 に示す。

まず、動作有無要因の有り 有り水準において、前後半要因の単純主効果は得られなかった ($F_{(1,23)} = 0.49, n.s.$)。また、動作有無要因の有り 無し水準において、前後半要因の単純主効果が得られた ($F_{(1,23)} =$

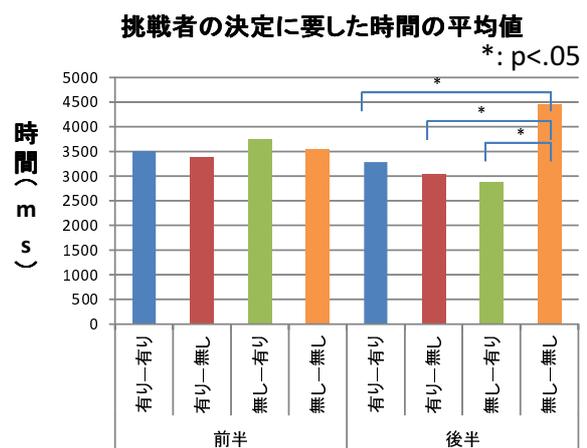


図 10: 挑戦者の決定に要した平均時間 (前後半要因の単純主効果)

4.30, $p < .05$)。その結果、前半水準の方が後半水準よりも大きいことが分かった。次に、動作有無要因の無し 有り水準において、前後半要因の単純主効果が得られた ($F_{(1,23)} = 10.82, p < .01$)。その結果、前半水準の方が後半水準よりも大きいことが分かった。そして、動作有無要因の無し 無し水準において、前後半要因の単純主効果が得られた ($F_{(1,23)} = 5.73, p < .05$)。その結果、後半水準の方が前半水準よりも大きいことが分かった。この実験結果から、ロボットがゲームの前半部分のみ視線行動をする場合には、前半よりも後半の方が挑戦者の決定に要した時間が小さいということが分かった。この結果は仮説に反するものであった。一方で、ロボットがゲームの後半部分のみ視線行動をする場合には、前半よりも後半の方が挑戦者の決定に要した時間が小さいことが分かった。この結果は仮説を支持するものである。そして、ゲームの前後半においてロボットが視線行動をしない場合には、後半よりも前半の方が挑戦者の決定に要した時間が小さいことが分かった。

続いて、前後半要因の各水準における動作有無要因の単純主効果を検証するため、下位検定を行った。その結果を図 11 に示す。

まず、前後半要因の前半水準における動作有無要因の単純主効果は得られなかった ($F_{(3,69)} = 0.43, n.s.$)。次に、前後半要因の後半水準における動作有無要因の単純主効果が得られた ($F_{(3,69)} = 8.77, p < .05$)。LSD 法による多重比較の結果、無し 無し水準が他水準よりも大きいことが分かった。この実験結果から、ゲームの後半部分において、ロボットが視線行動をする方がしないよりも挑戦者の決定に要した時間が小さいということが分かった。この結果から、ロボットの視線

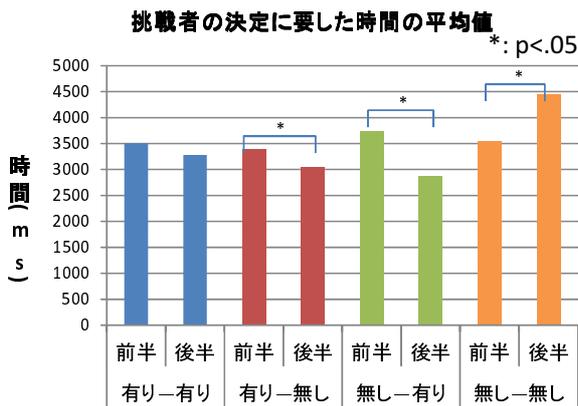


図 11: 挑戦者の決定に要した平均時間 (動作有無要因の単純主効果)

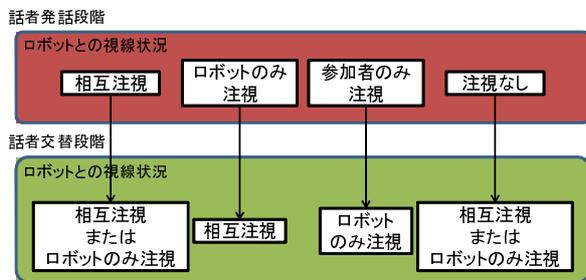


図 12: 参加者への行動を促すロボットの視線配分

行動が行為者の役割交替の円滑化に寄与することが示唆された。

4.9 まとめ

本実験における挑戦フェーズを多人数対話場面における話者の発話中の段階、交替フェーズを話者交替の段階と仮定する。この時、4.8.1 節の各参加者が挑戦者になった割合 (図 7) の結果から、話者の発話段階 / 話者交替の段階のそれぞれにおいて、ロボットとある参加者との視線状況に応じた視線配分をすることにより、その参加者に対して行動 (発話) を促すことができると考えられる (図 12)。この結果は、4.8.2 節の結果 (図 9) から、話者交替段階の視線配分により、ロボットの視線だけでなく他の参加者からの注視を促すことで実現し得ると考えられる。それにより、4.8.3 節の結果 (図 10, 図 11) から、ロボットの視線行動の効果は、対話の場全体における後半ほど強いものとなり、話者交替の円滑化を実現できる可能性があると考えられる。

5 おわりに

ロボットと人、ロボットを介した人同士のコミュニケーションを実現するためには、ロボットが人の対面対話状況に構造的・社会的に適応できることが必要になる。ロボットが人の対面対話状況において適切に振る舞うためには、人の対話がどのような構造のもとで構成されているかを理解する必要がある。その中でも音声対話においては、人と人、ロボットと人の円滑な話者交替 (turn-taking) の実現が問題となる。これまでの話者交替に着目した対話ロボットの研究では多人数対話における周囲の対話環境を受動的に推定することで適応的に発話権を獲得できたが、積極的に対話環境をコントロールすることは想定されていない。しかし、多人数対話は、しばしば話者交替の暗黙的なルールから逸脱され、それによって健全な対話の場の崩壊を招くことがある。そこで本研究では、多人数対話に参加するロボットの身体的な振舞いのデザインによる話者交替の円滑化、議論の活性化を目的としている。その方法として、本稿では、行動主体の発話のアドレス判断に用いられる視線行動に着目した。そして、多人数参加型のバランスゲームを行う課題を通じて、対話環境を複雑化し得る言語情報を排除した上で、ロボットの視線行動がゲーム参加者らの役割交替にもたらす影響を調査した。その結果、以下のことが示唆された。

- 話者の発話段階 / 話者交替の段階のそれぞれにおけるロボットの視線配分により、参加者の発話を促すことができる。
- 話者交替段階のある参加者へのロボットの視線行動は、その参加者に対する他の参加者の注視を促すことができる。
- ロボットを適切に動作させることで、動作しないよりも話者交替を円滑に進められる。

本実験では、多人数対話場面における話者の発話を行為フェーズ、話者交替を交替フェーズとして分割し、視線行動に焦点を当てた汎用的なモデル構築を試みた。ただし、本実験では行為者の移り変わりの時系列的な変化は考慮されていない。福原らは多人数対話の会話を先導する会話の中心者の順位を優位性順位とし、視線による次話者の選択や相互注視の効果が優位性順位によって異なると述べている [17]。このことから、どの参加者が対話の場の中心なのか、それがどのように移り変わるのかを視野にいれた検討が必要であると言える。

また、本実験では、参加者が次の行為者になるタイミングがもつメリットやリスクが考慮されていない。しかし、日常的な対話場面では次話者になるタイミングの違いによって、その後の参加者間の立場や関係性が

変動する可能性がある。そのため本実験のように次の行為者への移行が常に平等な環境ではなく、そのタイミングがその後の役割を変化させうる環境下での参加者の行動を検証すべきである。

更に、そして、ロボットの視線行動が参加者の行動を強制せずに、互いに対話が有意義なものであったのかどうかを検証していく必要がある。今後は上記の点に加え、得られた知見の有用性を言語情報を含めたより日常的な対話環境下で検証すべきであると考えられる。これが明らかになることで、協調学習場面に代表される多人数対話における話者の偏りをなくし、対話が効率的に行われる方向への話題の誘導を自然な形で達成する認知的作用を提供できると考えられる。

謝辞

本研究は、科研費・新学術領域研究「人口ロボット共生学」(24118704)の助成を受けたものである。

参考文献

- [1] 坊農真弓, 鈴木紀子, 片桐恭弘: 多人数会話における参与構造分析-インタラクシオン行動から興味対象を抽出する. 認知科学, Vol. 11, No. 3, pp. 214-227 (2004)
- [2] 松坂要佐, 東條剛史, 小林哲則: グループ会話に参与する対話ロボットの構築. 電子情報通信学会論文誌.D-II, 情報・システム, II - パターン処理. Vol. J84-D-II, No. 6, pp. 898-908 (2001)
- [3] 船越孝太郎, 小林一樹, 中野幹生, 山田誠二, 北村泰彦, 辻野広司: Artificial Subtle Expression としての明滅光源による音声対話の円滑化. 電子情報通信学会論文誌.A, 基礎・境界, Vol. J92-A, No. 11, pp. 818-827, (2009)
- [4] 高柳侑華, 竹内勇剛: 多人数音声対話場面において対話をする人工物の発話に対するアドレス判断. 電子情報通信学会論文誌.D, 情報・システム, Vol. J94-D, No. 1, pp. 37-47, (2011)
- [5] Goffman, E.: Replies and responses. *Language in Society*, Vol. 5, pp. 257-313 (1976)
- [6] Clark, H. H., Carlson, T. B.: Hearers and speech acts. *Language*, Vol. 58, pp. 332-373 (1982)
- [7] 徳永弘子, 湯浅将英, 武川直樹: 3人会話の発話交替における視線行動と発話マインド分析-聞き手の立場から見た発話・非発話の戦略-. 電子情報通信学会技術研究報告.HCS, ヒューマンコミュニケーション基礎, Vol. 107, No. 241, pp. 7-12 (2007)
- [8] 小磯花絵, 伝康晴: 円滑な話者交替はいかにして成立するか: 会話コーパスの分析にもとづく考察. 認知科学, Vol. 7, No. 1, pp. 93-106 (2000)
- [9] 小倉加奈代, 西本一志: チャット対話における発言生成過程の分析. 人工知能学会全国大会論文集 (CD-ROM), Vol. 18, pp. 2D1-07 (2004)
- [10] Duncan, Jr., S.: Some signals and rules for taking speaking turns in conversation. *Journal of Personality and Social Psychology*, Vol. 23, pp. 283-292 (1972)
- [11] Kendon, A.: Some functions of gaze direction in social interaction. *Acta Psychologica*, Vol. 26, pp. 22-63 (1967)
- [12] Duncan, Jr. S., Fiske, D. W.: Face-to-face interaction: Research, methods and theory. *Hillsdale, NJ: Lawrence Erlbaum*
- [13] Roel Vertegaal, Robert Slagter, Gerrit C. van der Veer, Anton Nijholt: Eye gaze patterns in conversations: there is more the conversational agents than meets the eyes. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 301-308 (2001)
- [14] 榎本美香, 伝康晴: 3人会話における参与役割の交替に関わる非言語行動の分析. 言語・音声理解と対話処理研究会, Vol. 38, pp. 25-30 (2003)
- [15] エドワード・ホール著, 日高敏隆, 佐藤信行訳: かくれた次元. みすず書房 (1970)
- [16] ATR Creative: robovie-mR2. <http://www.irc.atr.jp/~reo/mr2/> (最終アクセス 2012年8月28日)
- [17] 福原佑貴, 中野有紀子: 注視・相互注視に基づく会話参加者優位性順位の推定に向けた多人数インタラクシオンの分析. 情報処理学会第73回全国大会講演論文集, Vol. 2011, No. 1, pp. 251-253 (2011)