

ロボット間会話への人間の参入における アイコンタクト同調のゆらぎ構造と文脈維持

Effects of fluctuation structure of eye contact on human participation in dialogue and context maintenance in robot-robot conversations

中澤さくや^{1*} 竹内勇剛¹
Sakuya NAKAZAWA¹ Yugo TAKEUCHI¹

¹ 静岡大学大学院総合科学技術研究科

¹ Graduate School of Integrated Science and Technology, Shizuoka University

Abstract: 従来の人と人工物との円滑な対話コミュニケーションの実現を目指すために研究の多くでは、同調というワードに注目している傾向が強く、実際、話の内容に対する同調や話速・音量と言った韻律情報の同調をはじめとしたコミュニケーション内の同調がコミュニケーションを円滑にすることを示す研究結果が多く報告されている。同時に、2つの話者の発話内容の類似度の変化の仕方においてピンクノイズ (1/f ゆらぎ) があると、話者にとって楽しさや繋がりを感じるものになるという研究がある。この研究から、ゆらぎがあることで対話者の心理状態にプラスの影響を与える可能性があると考えられ、「ゆらぎ」という要素は、従来の円滑な会話か否かという評価指標とは異なる、会話に関わる人間の心理状態に着目した新たな評価指標をもたらす可能性があると推測される。したがって、本研究においては、ロボットの視線移動にゆらぎを適用したときの心理状態、特に会話参入への抵抗感・負担・意欲や対話に対する楽しさに変化が生じるかについて検証を行うために、ゆらぎだけではなく、一定タイミングとランダムタイミングを適用させた条件も追加したうえで、一定・ランダム・ゆらぎの3群の比較実験を行ったところ、3群において有意な差はなかったものの、視線移動にゆらぎを導入することは、エージェントに人間らしさを付与し、対話に対する楽しさやエージェントに対する親しみにプラスの影響を与える効果がある可能性が示唆された。

1 はじめに

現在では、あらゆるコミュニケーションロボットがあるが、これらの設計元となっている人間同士のコミュニケーションに関する先行研究において、よく見られる題材の中に「同調」というものがある。この「同調」に関する研究では、アイコンタクトの同調が円滑なコミュニケーションに寄与しているといった研究 [1] や意見・考えの同調とアイコンタクトとの関係、ならびにアイコンタクトと瞳孔同期との関係について論じた論文 [2]、韻律情報の同調が親しさと盛り上がりのある会話に寄与するといった研究 [3][4] といったものがあり、我々の円滑なコミュニケーションには「同調」が重要な役割を果たしているとする論文が多くあることがわかる。

前段では、同調がコミュニケーションを円滑する要素であることについて触れたが、同時に対話中の「ゆら

ぎ」に注目する研究もいくつかある。コミュニケーションにおける「ゆらぎ」に関する研究として、頭の動きにおける同調がパターン化されて発生しているのではなく、分脈に応じて柔軟に発生する動的な現象であると論じたもの [5] や発話内容の類似度の変化の仕方においてピンクノイズがあると、話者にとって楽しさや繋がりを感じる会話になると論じた論文 [6] がある。このように、コミュニケーションにおいて「ゆらぎ」が存在していること、そして、1/f ゆらぎによって発話者の心理状態にプラスの影響を与えるという側面も報告されている。これはつまり、ゆらぎが本研究で取り上げるエージェントとの対話意欲の向上と言った心理状態に関する新たなコミュニケーションの評価指標をもたらす可能性があると考えられる。したがって、コミュニケーション内の「ゆらぎ」について議論する価値があるといえる。

なお、本研究に先駆けて、身体ねじりによる会話への割り込みの誘発効果を明らかにする研究が行われた [7]。この研究において取り上げたものは「参入」ではなく、

*連絡先: 静岡大学大学院総合科学技術研究科
〒432-8011 静岡県浜松市中央区城北 3-5-1
E-mail: nakazawa.sakuya.21@shizuoka.ac.jp

「割り込み」ではあるものの、話しかける側にとって胴体を向けられていると対話の参加者としてみなされるという結果を本研究では、エージェントの設計において参考にしている。また、本研究では、身体ねじりではなく、ターンテーキングの調整として重要な役割を担うとされている視線外しやアイコンタクトに着目したうえで [8][9][10]、それらにゆらぎを適用させることで、対話の話を遮る「割り込み」ではなく、対話の話を維持した状態での「参入」の機会が生まれ、対話の活性化効果があるのかについて着目している。

以上のことから、本研究は、エージェントの視線そらしのタイミングにゆらぎを適用することで対話参入の機会を作り、会話の活性化を促す効果があるかを明らかにすることが目的である。この研究の結果、ゆらぎを適用させた視線そらしが参入の機会を生み出し、会話の活性化を促す効果があると明らかになった場合、人間と対話するエージェントがより良くなるためには、ただの視線そらしよりも、ゆらぎを適用させた視線そらしのほうが良いと提案することができると考える。また、仮に、逆の結果が得られたとしても、コミュニケーション、特に非言語コミュニケーションにおける 1/f ゆらぎやゆらぎの適用に関する研究はまだ進んでおらず、また非言語コミュニケーションにゆらぎを適用させることが対話参入の機会を生むという視点はまだ未踏であるため、この実験によって、関連研究の発展の一助にもなると考える。

2 1/f ゆらぎ

2.1 定義

1/f ゆらぎとは、スペクトル密度 $S(f)$ が周波数 f に対して、以下の式 (1) のように反比例する信号であり、ゆっくりした変化ほど変動の度合いが大きく、速い変化ほど変動の度合いが小さいという性質を持っている [11]。なお、式 (1) における α は、1 に近いほど 1/f ゆらぎに近く、0 に近いほどホワイトノイズに近いことを意味する。

$$S(f) \propto \frac{1}{f^\alpha} \quad (1)$$

2.2 先行研究

1/f ゆらぎは自然界に多く存在する現象であることが知られているが [11]、人間の心拍や血圧、瞳孔径と言った生体反応 [12] や人間の動作における動から静止へ入る瞬間に 1/f ゆらぎが発生していることも分かっている [13]。

また、数は少ないものの、エージェントの動きに 1/f ゆらぎを適用する研究として、山田 (2022) が行った、フォトリアルな CG キャラクタの歩行・走行動作に 1/f ゆらぎを適用し、適用していないキャラクターを比較対象として用意したうえで、不気味の谷現象が緩和されるか明らかにした研究がある [14]。その結果、適用していない場合に比べて、1/f ゆらぎを適用したキャラクターは、人間に似ているという感想が多くなっただけでなく、親和性が高まっていることもわかっている [14]。ただ、本段落の冒頭でも触れたように、エージェントの動きに 1/f ゆらぎを適用する研究は少ないため、全ての動作に 1/f ゆらぎを適用させても同じ結果になるとは言えない。しかし、O'Neill ら (2024) が行った、CANDOR という大規模な公開会話データセットを用いて、会話の文字情報（意味内容）の時系列的な変動パターンを分析し、発話者に会話の楽しさや継続的なつながりの感覚についてのアンケートを実施した実験では、発話内容の類似度の変化の仕方においてピンクノイズがあると、話者にとって楽しさや繋がりを感じる会話になることを明らかにしたことから [6]、1/f ゆらぎには心理的にプラスの影響を与える可能性は捨てきれないと言える。

2.3 本実験でのゆらぎの実装方法

まずは、1/f ゆらぎを実装するために必要なことを述べる。まず、1/f ゆらぎを実装するにあたって、高速フーリエ変換 (FFT) によって 1/f ゆらぎ信号を生成する必要がある。具体的には、平均 0、分散 1 のホワイトノイズ $w(t)$ を生成したあと、これを FFT により、周波数領域へ変換することで複素スペクトル $W(f)$ を得る。そして複素スペクトル $W(f)$ の各周波数成分に対して、以下の式 (2) のように振幅を周波数 f の平方根に反比例する形でスケールする。なお、直流成分 $f = 0$ は発散を抑えるために除外する。

$$\tilde{W}(f) = \frac{W(f)}{\sqrt{f}} \quad (2)$$

この操作により、パワースペクトル密度は式 (3) を満たすよう調整した後、逆フーリエ変換を行うことで 1/f ゆらぎ信号 $x(t)$ を生成することができる。

$$S(f) \propto \frac{1}{f} \quad (3)$$

1/f ゆらぎ信号の生成の後は、この信号を用いて、発生率に基づく視線移動のイベントの生成を行う。具体的には、時刻 t における視線移動の瞬間発生率 $\lambda(t)$ は、以下の式 (4) のように定義できる。なお、 λ_0 は基準視線移動率、 β はゆらぎの影響強度を表すパラメータで

ある. $\text{clip}(\cdot)$ は, 発生率が設定範囲 $[\lambda_{\min}, \lambda_{\max}]$ を逸脱しないよう制限する関数である.

$$\lambda(t) = \text{clip}(\lambda_0 \exp(\beta x(t)), \lambda_{\min}, \lambda_{\max}) \quad (4)$$

その後, 時間を刻み幅 Δt で離散化し, 区間 $[t, t+\Delta t]$ において視線移動が発生する確率を式 (5) のように定義する.

$$p(t) = 1 - \exp(-\lambda(t)\Delta t) \quad (5)$$

最後に, 各時刻において一様乱数 $u \sim \mathcal{U}(0, 1)$ を生成し, 式 (6) を満たす場合に視線移動イベントを発生させる.

$$u < p(t) \quad (6)$$

以上の手法により, $1/f$ 特性を有する時間変動を含む視線移動イベント列を生成することができる.

ここまで, $1/f$ ゆらぎの実装に必要な手順について触れてきたが, 本実験で用いたエージェントは, 「〇秒~〇秒の間でランダムに1回」という設定しか行えないため, $1/f$ ゆらぎ信号生成や時間的な変動, 発生率の生成ができず, 本エージェントの仕様上, 理論的に示した $1/f$ ゆらぎモデルを直接実装することはできない. したがって, 本実験では, $1/f$ ゆらぎは実装できないものの, $1/f$ ゆらぎに見られる変動構造に着想を得て, 周期性を緩和する目的で一定範囲内のランダム化を行い, 時間間隔にゆらぎを付与する方法を採用した. 具体的には, Web GUI のアイドル機能を使い, 視線そらしのタイミングを 14 秒~16 秒に 1 回の間でランダムに設定することで, 平均 15 秒に 1 回の視線そらしを実現しつつ, 時間間隔にゆらぎを付与した. つまり, 本手法は厳密な $1/f$ ゆらぎを再現するものではないが, 固定周期による強い規則性を回避し, 時間間隔に変動を導入することを目的としたものである.

3 実験方法

3.1 実験の目的

本研究の目的は, エージェントの視線移動にゆらぎを適用することで対話参入を機会を作り, 会話の活性化を促す効果があるかを明らかにすることである.

なお, 上記のようなゆらぎの効果を検証するにあたって, 以下の表 1 の様に一定タイミングの視線移動群 (以降 A 群), ランダムタイミングの視線移動群 (以降 B 群), ゆらぎタイミングの視線移動群 (以降 C 群) の 3 つの群を用意した. 表 1 にあるように, 変化させるのは視線移動のタイミングだけで, 瞬きは 3 群とも「8 秒に 1 回」で固定した.

表 1: それぞれの群の設計

群	瞬き	視線そらし
A 群 (一定)	8 秒に 1 回	15 秒に 1 回
B 群 (ランダム)	"	14~20 秒に 1 回
C 群 (ゆらぎ)	"	14~16 秒に 1 回

3.2 実験の環境

実験は, 主に CommU (図 1) 2 台と CommU を操作する PC を 2 台を用いる. なお, この 2 台の PC のうち, 一台は CommU の操作をしつつ, 同時に ChatGPT を起動し, 実験参加者と音声による会話ができるようにする. また, 1 台のビデオカメラで実験の録画を行い, ChatGPT の音声認識のためにマイクを使用する.



図 1: CommU

上記の機材の配置としては, 図 2 の通りである. なお, 図 2 の A は, 記録するためのビデオカメラ, 図 2 の下部の I は, クイズのお題を映し出し, 参加者に見せるためのモニター, 上部にある II は, 参加者の音声を認識し, CommU と会話できるようにするためのマイクである. また, 2 体の CommU の動作に関しても, 基本的に 2.3 で述べた Web GUI を用いて制御しているが, CommU2 に関しては, それに加えて ChatGPT による参加者の音声認識→返答の出力も行い, 手作業で CommU2 に ChatGPT にて出力された返答を Web GUI におけるコマンドで言わせるようにした. なお, 図 3 と図 4 は, 実際の実験環境であり, それぞれ図 2 の a, b から見た実験環境である.

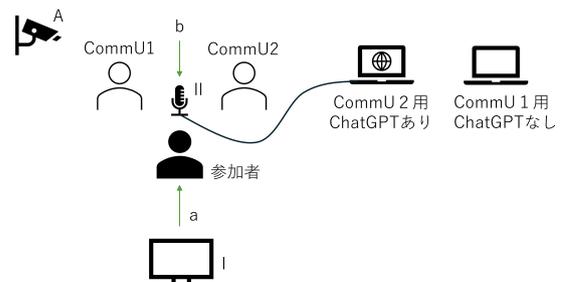


図 2: 実験環境イメージ



図 3: a から見た環境

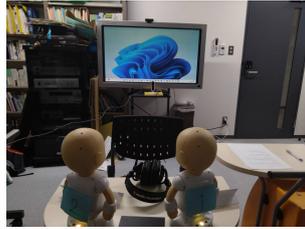


図 4: b から見た環境

3.3 実験の手順

実験の手順は以下のとおりである。

実験手順 1: 取物語の話を要約した映像を実験参加者に見せ、話のあらすじを記憶させる (5分程度)。

実験手順 2: 「CommU の画像認識能力」の性能評価という架空の目的を提示したうえで、物語の内容に関する 4 個のクイズを CommU とともに協力して答えを出してもらおうと説明する。

実験手順 3: 参加者への説明が終わったら、CommU との自己紹介の時間を作る。

実験手順 4: 手順 3 終了後、タスクを開始。

実験手順 5: 全てのクイズに答え終わったら、アンケートへ回答させる。

実験手順 6: アンケート回答に関するインタビューを実施。

実験に関する補足だが、まず、実験を行う前の実験参加者への説明では、CommU の視線移動に関するバイアスがかからないように「CommU の画像認識能力の性能評価」が目的であると伝える。なお、この実験における実験計画は、実験参加者間計画を想定している。したがって、1 人に対して 1 つの群に関する実験を行う。また、それぞれの群における参加者の人数は、数が同じになるように割り振った。また、ビデオカメラによる録画は、手順 3,4,6 で行う。なお、実験手順 1 については手順 2 以降の部屋とは別の部屋において行い、手順 2 以降は全て図 2 に示した場所で行う。

次に、手順 4 における CommU の動作についてだが、実験中の CommU の動作は、ゆらぎを適用させた視線そらしの参入意識の変化と対話の活性化を見るために以下の図 5 のような姿勢と動きに設計している。なお、これ以降、参加者から見て左の CommU を「コー君」、参加者から見て右の CommU を「ミュー君」と呼称する。

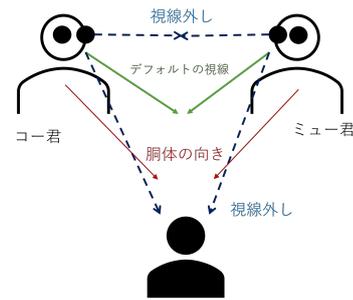


図 5: CommU の動作

図 5 だが、実験参加者に対話の参加者であると認知させるために、本研究の前段研究の結果を踏まえて、胴体は僅かに参加者に向ける。その上で、デフォルトの視線は、胴体の向きより、僅かに反対の CommU 側に向ける。そして、視線外しの向きは、互いの CommU の顔を見る方向と実験参加者の顔を見る方向の 2 パターンである。なお、視線移動の際は、目ではなく、頭を動かして行っている。

なお、いずれのクイズにおいても、参入機会の発生を見るために、お題を出してすぐは、CommU 同士の会話から始めるようにし、後から参加者が対話に参入するように設計した。また、このときの CommU 同士の会話には台本があり、その通りに進行する。また、CommU 同士の会話は、全てのお題に対して、以下の特徴を有している。

- 間違っただけを出そうとする
- 会話中に悩んだり、間を作ったりする
- 台本には終わりがあり、台本の終わりには参加者への答えの確認を行う

このようにすることで、参入する際のきっかけを生じさせるだけでなく、仮に参加者が自主的に対話参入しないことがあっても、対処できるようにした。

なお、参加者がエージェント間の対話に参入したとき、もしくは台本が終わった後は、台本ではなく、ChatGPT による音声認識と自動生成を用いて参加者と対話を行った。

3.4 実験の観測事項と分析方法

実験中に観測するものは、それぞれのクイズにおいて、最初に行われる CommU 同士の会話のどのセリフで参入が行われるのか、参入後の CommU との対話の中でどんな会話が行われるのか、そして CommU に話しかけてからいつ対話を終わらせるのかについて観測する。このとき、意欲に関しては、早い段階で参入が発

生していれば高く、反対に遅ければ低いと判断し、活性化に関しては、対話を終わらせる時間が遅ければ遅いほど活性化していると判断した。

なお、実験後には、CommU 達の会話に参入する際の「参入のしやすさ」や「参入するタイミングに影響を与えたもの」、「参入するにあたっての心理的負担」、「会話の楽しさ」、「CommU への親しみ」、「会話の活発さ」に関するアンケートを5段階リッカート尺度を用いて行いつつ、それぞれの要因についても自由記述にて答えさせた。なお、5段階は「0,1,2,3,4」のように設定し、0と1はマイナス、2は中立、3と4はプラスの意味合いに設定し、数が大きくなるほどプラス、小さくなるほどマイナスとした。

3.5 実験の仮説

仮説は以下のとおりである。

- ゆらぎ群は、他条件群と比較して、対話の盛り上がりや楽しさを感じる。
- ゆらぎ群は、他条件群よりも会話の文脈を維持し、会話時間がより長く持続し、活性化する。
- どの群でも参入行動自体は見られる。

しかし、参入行動開始の早さや会話参入意欲に関しては、先行研究が少なく、仮説を立てることが難しい。したがって、以下では、それぞれの観測事項に関して、どのような違いが3群で生じるのかについて推測を述べる。

- 参入行動開始の早さは群によって異なる（ゆらぎ群が他群と比較して参入が速くなる、もしくは逆に遅くなる）。
- 会話参入意欲は、3群において違いはない。
- 視線移動自体に意欲を増加させる効果がある。
- ゆらぎを適用することで意欲が上昇する。

3.6 実験の参加者

本実験の参加者は、大学生・大学院生の男女合わせて30名であり、A群、B群、C群、それぞれに10人を割り振った。

3.7 実験結果

3.7.1 対話参入行動の有無について

対話参入の有無についてだが、タスク中に一度でも参入した人は、A群は4人、B群は5人、C群は5人であり、3群の中で大きな違いはないという結果になった。また、参入行動自体はどの群でもあるという仮説はあっていたということになる。

3.7.2 参入行動を行った人の特徴について

参入行動を行った人（A1, A4, A6, A7, B3, B5, B6, B9, B10, C1, C2, C4, C8, C9）に注目し、3群における特徴をまとめる。まずは、参入のタイミングについてだが、表2のような結果になった。表2にあるように、A群では、後半のクイズに行くほど参入する人が増える一方で、B群とC群では、前半のクイズで参入する人が多く、後半のクイズでは参入する人が減っていることがわかる。なお、B群とC群を詳しく見ると、一番参入が多いのは、B群はクイズ2、C群はクイズ1であることが分かる。

表 2: 参入のタイミング

参加者	クイズ1	クイズ2	クイズ3	クイズ4
A1	×	×	○	○
A4	×	○	○	○
A6	×	○	○	○
A7	○	×	×	○
B3	○	○	○	○
B5	○	○	×	×
B6	○	○	×	×
B9	○	○	×	○
B10	×	○	×	×
C1	○	○	○	○
C2	○	○	×	×
C4	○	×	×	×
C8	○	○	×	×
C9	○	×	×	×

次に、参入に対する難しさに関してだが、A群では、一回でも「やや簡単だった」と答える人は1人だけであり、参入した全てのクイズにおいて「やや難しかった」と答える人が多かった。一方、B群では、参入した全員が1回でも「やや簡単だった」、「簡単だった」と答えていた。しかし、詳しく見ると、クイズ1では「やや難しかった」もしくは「簡単でも難しくもなかった」と答えた人が3人いたり、最初の参入時には「やや難しかった」、「どちらでもない」と回答したものの、2回目の参入時には「やや簡単だった」と答えるよう

になった人が3人いたりといった特徴が見られた。最後に、C群だが、一回でも「簡単だった」、「やや簡単だった」と答えた人は2人であった。なお、C群に関しては、2回以上の参入が見られた3人のうち2人は、常に難しい寄りの回答をしていた。

次に、参入時の雰囲気についてだが、A群では、4人中3人は常に「参入時に話しかけても問題ない雰囲気を感じた」寄りの回答をしており、残りの一人は、「感じなかった」と常に回答していた。一方、B群では、参入した全員が1回でも「感じた」寄りの回答をしているという結果になった。基本的に「感じた」、もしくは「どちらとも言えない」と回答していたが、2回以上の参入をした4人のうち、2人はマイナスもしくは中立からプラスの回答に変わる傾向が見られた。最後にC群だが、5人中4人は1回でも「感じた」寄りの回答をしており、残りの1人は常に「感じなかった」寄りの回答をしていた。なお、B群とは違い、C群では、クイズの回数に限らず、参入時はプラスの評価をする人と逆に、常にマイナスの評価をする人ではっきりと分かれているという結果になった。

最後に、参入のタイミングを計る際に参考にしたものについてだが、参入した14人のうち、CommUの視線に言及するものは3人だけであった。なお、CommUの会話内容は参入者全員が挙げていた。実際に、参入の難しさや雰囲気でもマイナス寄りの回答をした要因では、「会話の途切れが分からない」、「会話を遮ってしまった」が3群に共通してあげられており、プラス寄りの要因でも「CommUがわからなそうだったから」や「自分の意見と同じだったから」が挙げられていることから、CommUの会話内容に関するものが多いことがわかる。また、参入行動が見られなかった16人のインタビューでも「質問されるまで待っていた」、「CommUの意見を聞いたかった」との発言が多く見られ、CommUの会話内容に注意していたことが分かる。

3.7.3 アンケート結果について

参加者全員に対して行った会話の楽しさやCommUへの親しみ、会話の活発さに関するアンケート結果についてだが、表3から表5、図6から図8のような結果になった。なお、図6から図8の箱ひげ図における×は平均値、点は外れ値となっており、平均値の算出は外れ値を含めたものとなっている。まず、楽しさについてだが、A群が最も高く、次いでC群、B群の順であるものの、分散分析では $p=0.103(p>.05)$ であり、有意差はなかった。次に、親しみについてだが、C群が最も高く、次いでA群、B群の順であったが、こちらも分散分析では $p=0.441(p>.05)$ であり、有意差はなかった。最後に、活発さについてだが、C群が最も高く、次いでA群とB群が同率であった。なお、こ

でも分散分析では $p=0.314(p>.05)$ であり、有意差はなかった。

表 3: 会話の楽しさ

群	平均	分散
A	3.1	0.767
B	2.4	0.844
C	3	1.333

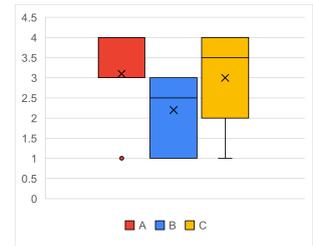


図 6: 会話の楽しさの分布

表 4: CommU への親しみ

群	平均	分散
A	2.5	1.389
B	2.4	2.044
C	3.1	1.656

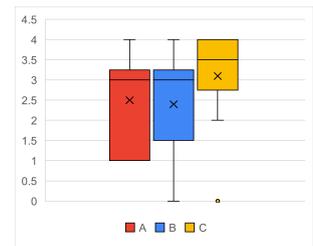


図 7: CommU への親しみの分布

表 5: 会話の活発さ

群	平均	分散
A	2.2	0.844
B	2.2	1.511
C	2.8	0.622

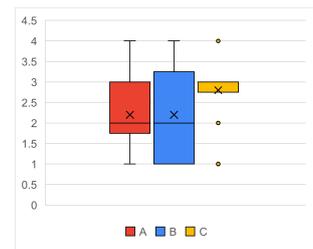


図 8: 会話の活発さの分布

次に、楽しさや親しみ、活発さに関する自由記述についてだが、まず、A群のマイナスの評価を頻出順にまとめると、「機械感(5)」、「テンポの悪さ・ラグ(4)」、「会話の入りづらさ(3)」、「発話量のアンバランス(3)」、「主体性・思考の弱さ(2)」の順で頻出しており、プラスの評価では、「対話感がある(4)」、「教える・導く感覚の楽しさ(3)」、「自分の発言の発展(3)」、「人間らしさ(3)」、「新鮮な体験(2)」の順で頻出していた。次にB群では、「低い応答性・悪いテンポ(7)」、「主体性・意見提示の欠如(6)」、「作業感(3)」、「コー君の消極性(3)」、「ロボット感(3)」の順で頻出しており、プラスの評価では、「体験の珍しさ(4)」、「子供らしさ/人間臭さへの愛着(4)」、「会話の成立(3)」、「キャラクター性・個性(2)」、「ミュー君の積極性(2)」、「視線・首振りなど非言語表現(2)」の順で頻出していた。最後にC群では、マイナスの評価として、「コー君の関与不足(4)」、「応答の遅さ・テンポの悪さ(4)」、「議論にならず説得・説明寄り(3)」、「会話に入りにくい構造(3)」、「見た目・挙

動・話し方の違和感(1)」の順で頻出しており、プラスの評価では、「ミュー君の積極性(5)」、「質疑応答・会話の成立(5)」、「内容・分脈理解(4)」、「親しみやすさ・優しさ(4)」、「思考・理解を深める質問(3)」、「人格・キャラクター性(3)」の順で頻出していた。なお、「」内の()は、出てきた回数である。

3.7.4 タスク中の言動について

タスク中の参加者の言動についてだが、まずA群の特徴をまとめると、3人はずっと敬語だったものの、大半が口調のくだけや子供に話すような口調が見られる。また、大半は話の脱線や試すような行為等がなく、回答の最終確認はあったものの、参加者が主導して淡々と答えに導くような言動が多く、無駄がなかった。また、少ないが、コー君に話しかける人が2人いた。次にB群の特徴だが、こちらも口調のくだけが大半の人に見られるが、A群とは違い、最初敬語だったのが後半にくだける人も確認できる。また、A群では見られない間違っただけの回答への誘導といった回りくどい言動も多く見られる。また、最初の方では説得が中途半端でも切り上げたり、回答の最終確認を行わないといった言動があったのが、後半では説得や説明・確認を行うようになる人も見られた。コー君への話しかけに関しては、大半が行っており、繰り返しコー君に話しかける人に関しては5人いた。最後にC群だが、B群と同様に、口調のくだけが大半の人に見られるが、A群とは違い、最初敬語だったのが後半にくだける人が確認できる。また、B群、C軍と同様に、説得や説明もあるものの、他群と違い、回答をCommUに言わせる行為、回答の最終確認などの行為も多かった。なお、コー君に話しかける人は7人であり、その全員が2回以上話しかけていた。

3.8 考察

結果としては、3群において参入行動をした人数に大きな差はなく、視線行動が参入行動に影響を及ぼしていたという結果や心理的な項目における有意差は見られなかったが、「参入行動のタイミング」や「心理状態」、「言動と自由記述」に関して3群にそれぞれ違いがあったといえる。以下の小節ではそれぞれの考察について述べていく。

3.8.1 参入行動のタイミング

まず、参入行動のタイミングについてだが、表2にあるように、A群では、後半のクイズに行くほど参入する人が増える一方で、B群とC群では、前半のクイ

ズで参入する人が多く、後半のクイズでは参入する人が減っていた。また、B群とC群を詳しく見ると、B群は2問目で一番多いのに対して、C群は1問目で一番多いことが分かる。

以上の結果は、一定リズムによる慣れとランダムなリズムによる不快感、そしてゆらぎのリズムが人の心理に良い影響を与えたことによって行ったと考えられる。まず、一定リズムとランダムなリズムについてだが、松原ら(2026)の実験において、予測可能な一定の速度で動いたロボットといきなり止まるなどの予測不可能な動きをしたロボットとで回避するときの生理反応からストレス反応を調べたり、参加者の主観的な快・不快を調べたとき、一定の方は繰り返し行うことでストレス反応が減少し、比較的快適より感じたのに対して、予測不可能な方は、ストレス反応が高いだけではなく、高いまま持続しており、主観的な評価でも不快に感じるという結果が得られた[15]。次に、ゆらぎに関してだが、2.2の先行研究で述べた、発話内容の類似度の変化の仕方においてピンクノイズがあると、話者にとって楽しさや繋がりを感じる会話になると論じた論文[6]やフォトリアルなCGキャラクタの歩行・走行動作に1/fゆらぎを適用すると、人間に似ているという感想が多くなっただけではなく、親和性が高くなる論じた論文[14]があるほか、人間らしさの異なるロボットや、CGキャラクタの映像を用意し、それぞれの映像に対して感性評価させる実験を行ったところ、人間らしさが不気味さを低下させ、親和性の向上に繋がること[16]もわかっている。以上のことから、A群では、視線移動が一定であるため、参加者がCommUの視線移動に慣れるまでに時間がかかり、その結果、後半になるほど参入しやすくなったと考えられる。そして、C群で、他群と違い一番最初に参入者が多いのは、1/fゆらぎの親和性と似た効果により、初対面でも参入に抵抗がなかった可能性がある。とはいえ、先述のC群のマイナスの評価において「会話に入りづらかった」とあったように、予測が不可能なことで先行研究と同様に無意識的なストレスがあり、B群と同じように後半では参入行動を控えたと考えられる。

3.8.2 心理状態

次に、楽しさや親しみに関するアンケート結果についてだが、有意差はなかったものの、楽しさでは、A群とC群でプラスの意味を持つ3以上の評価となっており、親しみに関してはC群だけが、プラス評価の3以上の評価となっている。

まず、楽しさについてだが、これは前段で述べた通り、一定リズムの慣れとゆらぎが心理的にプラスの影響を与えること、そしてランダムによる不快感が引

たものであり、楽しさや親しみに関してはそれが顕著だったと考えられる。

なお、会話の活発さに関しては、いずれの群でもプラスの評価にならなかったが、これは、自由記述で多く見受けられた「ラグや遅れ」が原因だと言える。実験環境でも述べた通り、ChatGPTの音性認識と文の自動生成を使用したのち、専用のGUIに生成した文を流すことをしていたが、やはりChatGPTの音声認識の時間がかかることと、生成した文をGUIに流すという行為によって発生した返答の遅れが影響を与えてしまったと言える。この遅れによって、タイミングの違いによる効果の有無が正しく検証できなかった可能性があるため、今後の課題と言えるだろう。

3.8.3 タスク中の言動と自由記述

最後に、タスク中の言動、ならびに自由記述について考察する。まず言動についてだが、A群では口調にだけはあるものの、あまり無駄がなく、B群ではずっとくだけた口調の人もいるものの、敬語からくだけた口調に変わったり、不十分な説得をしたり、意地悪の様な事をしたりする言動が見られ、C群では口調に関してはB群と同様ではあるものの、ヒントや質問を投げかけ、CommUに答えを言わせる、回答の最終確認を行うという特徴が見られた。また、コー君に話しかける行動はB群、C群で多く見られた。

次に、自由記述についてだが、ネガティブな意見としてA群では「機械的」という言葉が多く出ており、B群でも多くはないものの「ロボット」という発言があったのに対して、C群では、そういった言葉が一切なかった。また、ポジティブな意見として、どの群でも「人間らしさ」や「性格」に関する言及があるが、B群が一番言及頻度が高い。また、気になった点としてC群のポジティブな意見では他群で見られない「親しみやすさ」や「優しさ」の言及も見られる。

以上の言動と自由記述から、A群は予測可能な動きで不快感はないものの、機械感が強く感じられ、親しみを持つまでに至らず、B群は、予測不可能な動きにより、不快感を抱くものの、機械には感じない状態、そしてC群は、ゆらぎの影響により、人間らしさを感じた、あるいは機械らしさが少なかったために心理的安心感があったのではないかと考察する。この考察を裏付けるために、まずは敬語に関して見ていきたい。心理学において、敬語のような丁寧な言葉遣いは対人関係の距離（社会的距離）を象徴し、同時にそれを創り出す機能があるとわかっている[17]。このことを踏まえると、A群では敬語の人はずっと敬語を使用していたのに対して、B群とC群で前半は敬語だったのに後半から口調がくだける人が現れたのは、最初はCommUに対して、心理的な距離を感じていたのが、時間の経過

か何かしらの要因で心理的距離が無くなっていったからだと考えられる。そして、A群でずっと敬語だった人は変化を与えるものがなく、心理的距離を保ったままだったからであると考えられる。さて、ここでB群とC群の心理的距離に影響を与えたものだが、これが「機械感のなさ」ではないかと考えられる。実際、A群では、ネガティブな意見として「機械感」を多く挙げており、B群とC群では逆に少なかった。したがって、口調に関する違いは、心理的距離と機械感の無さによって発生したと考えられる。しかし、B群とC群は同じだったかと言えば、そうではない。楽しさや親しみに関してはC群はプラス寄りなのに対して、B群は中立寄りであった。この違いに関しては、ランダムの不気味さと不気味の谷現象が関係していると考えられる。不気味の谷現象とは、中途半端に人間に似たデザインにすると、不気味だと感じられてしまう現象のことであるが[18][19]、B群では「機械感はない」が「人間らしさ」が中途半端だったために、不気味だと感じ、楽しさや親しみに関して中立の意見になってしまったと考えられる。そして、1/fゆらぎを導入したキャラクターの動きが人間に似ているだけではなく、親和性も高めたという先行研究と同様に[14]、C群はゆらぎによって「機械感」はなく、「人間らしさ」があったためにB群とは異なって楽しさや親しみを感じる結果になったと考えられる。

3.8.4 明らかになったこと

以上のことから、本実験を通して以下のことが明らかになった。

- 参入行動に関して、3群で大きな違いはなかった。
- 視線行動が参入行動に影響を与えているという結果は得られなかった。
- 3群で心理的な項目における有意差はなかった。
- ゆらぎを視線移動に導入したことで、人間らしさが増加した可能性がある。
- ゆらぎを視線移動に導入したことで、会話の楽しさやエージェントに対する親しみを与える可能性がある。

4 まとめと今後の課題

今回の実験の結果から、2.8.4の考察のまとめでも触れたように、ゆらぎを適用した視線移動には、会話の楽しさやCommUへの親しみといった心理状態にプラスの影響を与える可能性が示唆された。しかし、2.8.2

での考察でも述べた通り、CommU の返答にはラグがあり、本実験での「会話の活発性」に関する観測項目では、正確なデータが得られなかったという課題がある。また、2.3 の実験手順で述べた通り、それぞれのクイズが出題された直後は、CommU 同士の会話から始まるようにしていたが、この CommU 達の会話は Web GUI のボタン操作、つまり実験者の手動で進行していたため、参加者が話しかけるまでの CommU 同士の会話時間や間に違いが発生してしまった可能性もあり、本来観測するはずだった「CommU に話しかけてから、いつ対話を終わらせるのか」についても正確なデータが得られなかった。したがって、今後、CommU に喋らせる実験を行う場合は、より自動化されたシステムを用いることで、ラグを減少させ、参加者によって違いが生じる可能性を減らすことが必要であるといえる。

また、本実験では、ゆらぎを適用した視線移動に加えて、参加者との対話をエージェントに課した上で、ゆらぎの効果の検証を行ったが、ゆらぎを適用した視線移動だけで一切喋らない場合の効果は、対話を加えたときに比べて変化があるのかについても、今後検証する必要があると考える。また、本実験では、CommU の仕様上、単なるゆらぎの効果を検証するだけであったが、別のエージェントを使用することで、1/f ゆらぎを適用した視線移動についても今後検証する必要があると考える。

謝辞

本研究（の成果の一部）は、「内閣府地方大学・地域産業創生交付金事業 駿河湾・海洋 DX 先端拠点化計画 マリンインフォマティクス研究機構 研究助成」の助成を受けたものです。

また、本研究を進めるにあたり、大阪大学の石黒研究室の方々には、CommU の利用にあたって多くの支援を頂きました。心より感謝申し上げます。

参考文献

- [1] Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta psychologica*, 26, 22-63.
- [2] Wohltjen, S., Wheatley, T. (2021). Eye contact marks the rise and fall of shared attention in conversation. *Proceedings of the National Academy of Sciences*, 118(37), e2106645118.
- [3] 西村良太, 北岡教英, 中川聖一. (2009). 音声対話における韻律変化をもたらし要因分析 (<特集> リズムとタイミング). *音声研究*, 13(3), 66-84.
- [4] Kousidis, S., Dorran, D., Wang, Y., Vaughan, B., Cullen, C., Campbell, D., ... Coyle, E. (2008). Towards measuring continuous acoustic feature convergence in unconstrained spoken dialogues.
- [5] Paxton, A., Dale, R. (2017). Interpersonal movement synchrony responds to high-and low-level conversational constraints. *Frontiers in psychology*, 8, 1135.
- [6] O'Neil, K., Finn, E. S. (2024). Pink noise in speakers' semantic synchrony dynamics as a metric of conversation quality. In *Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 46)*.
- [7] 高木カレブ, 坂本孝丈, 市川淳, 竹内勇剛. (2023). 身体ねじりを伴うロボット同士の対話への人の新規参与とその持続. In *人工知能学会全国大会論文集 第 37 回 (2023) (pp. 206OS2b04-206OS2b04)*. 一般社団法人 人工知能学会.
- [8] Rossano, F. (2012). Gaze in conversation. *The handbook of conversation analysis*, 308-329.
- [9] 土屋裕希乃. (2016). 会話場面における視線行動と満足度および印象評価の検討. *国際経営・文化研究*, 21(1), 153-162.
- [10] 深山篤, 大野健彦, 武川直樹, 澤木美奈子, 萩田紀博. (2002). 擬人化エージェントの印象操作のための視線制御方法. *情報処理学会論文誌*, 43(12), 3596-3606.
- [11] 李鹿輝, 加藤千恵子, 齋藤兆古. (2007). 降雪の 1/f ゆらぎ現象. *映像情報 industrial*, 39(3), 57-62.
- [12] 山本光璋. (1994). 生体 1/f ゆらぎ研究の現状. *BME*, 8(10), 1-4.
- [13] 菅井桂子, 齋藤兆古. (2008). 人間の動作に伴う 1/f 揺らぎ周波数の抽出.
- [14] 山田光宏. (2022). フォトリアルな CG キャラクターの動作への 1/f ゆらぎ適用による不気味の谷現象の緩和. *科研費研究成果報告/令和 3 年度終了課題*.
- [15] Matsubara, Y., Tamura, H., Minami, T. et al. (2026). Subjective Emotions and Physiological Responses During Collision Avoidance with a Virtual Autonomous Mobile Robot. *Int J of Soc Robotics*, 18(5). <https://doi.org/10.1007/s12369-025-01341-3>

- [16] Ho, C. C., MacDorman, K. F. (2017). Measuring the uncanny valley effect: Refinements to indices for perceived humanness, attractiveness, and eeriness. *International Journal of Social Robotics*, 9(1), 129-139.
- [17] Trope, Y., Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological review*, 117(2), 440-463. <https://doi.org/10.1037/a0018963>
- [18] Mori, M., MacDorman, K. F., Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & automation magazine*, 19(2), 98-100.
- [19] Katsyri, J., Makarainen, M., Takala, T. (2017). Testing the ‘uncanny valley’ hypothesis in semirealistic computer-animated film characters: An empirical evaluation of natural film stimuli. *International Journal of Human-Computer Studies*, 97, 149-161.