

# デュアルディスプレイ型 AI エージェントを使用した 人間らしさの定量的評価手法

## D<sup>2</sup>UI-AIA: Quantitative Analysis of Human Imitation Performance through Dual-display Interface AI Agent Interaction

小林 拓也<sup>1</sup>, 才野 皓平<sup>2</sup>

Takuya Kobayashi<sup>1</sup>, Kohei Saino<sup>2</sup>

<sup>1</sup> HAL 大阪, 大阪芸術大学

<sup>1</sup> Technical School: HAL Osaka, Osaka University of Art

<sup>2</sup> HAL 大阪

<sup>2</sup> Technical School: HAL Osaka

**Abstract:** This study presents a method for measuring how human-like AI agents appear to users, based on our research at the Human Interface Symposium 2024. We developed a dual-display AI agent system that shows a 3D avatar with facial expressions and user tracking capabilities. Our experimental design compares this system with a voice-only interface, focusing on how different levels of agent interaction affect user satisfaction in a practical guidance scenario. Using both survey data and interviews, we measure the impact of AI behavior on user experience, particularly in situations where incorrect information is provided. The study aims to contribute to the development of more effective AI interfaces for digital transformation in service environments.

**Keywords:** AI Agent, Human-Agent Interaction, Nonverbal Communication of AI Agents

## 1 はじめに

### 1.1 AI エージェントの非言語的振る舞い

現在、人間社会は、人工知能 (AI) 技術の急速な進展により、「人間らしさ (Human-likeness)」の概念を見つめ直す必要性に直面している。本研究は AI エージェント (以下 AIA と略す) における「人間らしさ」の実装手法を探求する。

人間らしさの厳密な定義は困難であり、多くの課題が存在する。そこで我々は、チューリングテストの洞察に基づき、「どのような条件下で、人間は AIA を人間として誤認するのか (Human Imitation)」という問いに焦点を当てた。感情パラメータ、デュアルディスプレイ、顔認識などの機能を実装し AIA の統合的な振る舞いを検証した。

実験から得られた興味深い知見として AIA の非言語的な振る舞い - 顔認識による話し手の位置検出と応答、笑顔で相手の名前を讀んでの会話など - が、ユーザーに機械を超えた存在としての印象を与える可能性が観測された。これは従来の音声応答を超える AIA と人間の関係構築可能性を示唆している。

この研究はヒューマンインタフェースシンポジウム 2024 (以下 HI2024) で発表<sup>[2]</sup>され、研究コミュニティから多様な反応を得た。寄せられた意見は研究の深化への示唆となり、本論文ではその結果から特定のテーマに焦点を当て議論を進める。まずは HI2024 の検証から始めたい。

### 1.2 HI2024 発表時の評価

本研究で開発したデュアルディスプレイ型 AI エージェント (Dual-display Interface AI Agent : 以下 D<sup>2</sup>UI-AIA と略す) は、特に以下の点において今後の研究展開の可能性が示唆された。

#### 1.2.1 両面ディスプレイの背景合成技術

向かい合うディスプレイ上において、アバター背景の透過表現手法が新規性の高い取り組みとして評価された。具体的には、透過性のある高価な液晶ではなく、安価な液晶画面とカメラを 2 個ずつ用いて互いの背景を描写するという、シンプルだが独創的なアプローチである。これは、SF でいうところの光学迷彩の概念に近い発想として注目を集めた。

背景合成の精度そのものについては、受け取り方に個人差はあったが、改善の余地がある状況。

### 1.2.2 評価プラットフォームとしての可能性

D<sup>2</sup>UI-AIA が、今後の多様な研究に展開可能な重要なプラットフォームとして高く評価された。特に以下の点で可能性が示唆された。

- ① AIA 研究のための安価な筐体として。
- ② メカニズム制作を伴うヒューマノイドロボット研究において、ユーザーとのインタラクション実験検証時、ヒューマノイドロボットの身長や形状が変わる実験の安価な代替品として活用が期待できる。

## 1.3 HI2024 で明らかになった課題

本研究の過程において、技術的側面と検証に関する重要な課題が明らかとなった。

### 1.3.1 音声認識の環境依存性（技術的側面）

ポスターセッション会場のような背景雑音が大きい環境下において、音声認識の精度が極端に低下する問題が顕在化した。屋外や人が多いなどの騒がしい環境では、システムの音声理解能力が著しく損なわれ、インタラクションの質に重大な影響を与えることが判明した。

### 1.3.2 人間らしさの評価が定性的（検証的側面）

システムの人間らしさに関する評価が、主観的かつ定性的な方法に依存していることが課題として指摘された。客観的で定量的な評価指標の確立が必要不可欠であることが明らかになった。

## 1.4 テーマの決定とフィードバック

以上、HI2024 発表の研究結果への反応について、十分なディスカッションを行った。その上で、「人間らしさ」の探求が本研究において重要であるという認識から、従来定性的な評価にとどまっていた「人間らしさ」の評価を定量的に行う手法の研究を今回のテーマとした。

その他の反応については以下をフィードバックとしたい。

音声認識の環境依存性については、継続的に改善方法を研究し実装のテストを行う。

AIA プラットフォームとしての評価はありがたいことではあるが、使用部材のリストや簡単な組み立て方法、内部プログラムの公開程度にとどめたい。

## 2 人間らしさの評価手法

本章では、AI エージェントの人間らしさを評価するための理論的枠組みを構築する。特に、非言語的振る舞いがユーザインタフェースへの不満を軽減する効果に注目し、その検証方法を提案する。

### 2.1 人間らしさの概念

人間の振る舞いを個別の要素に分解して評価する還元論的アプローチは、今日でも実験手法として一般的であろう。しかしこの手法は、人間らしさという統合的であり全体的な性質を捉えきれないのではないか、という課題を含んでいる。

この認識のもと、本研究では人間らしさの概念について、G. Ryle (1900-1976) の『心の概念 (*The Concept of Mind*<sup>[3]</sup>, 1949)』を基礎として検討する。Ryle は、心的状態を行動傾向として捉える現象学的アプローチを提唱した。この視点は AIA の振る舞いを評価する上で重要な示唆を与える。

また、従来のチューリングテストは主に言語的な応答に焦点を当てていたが、本研究ではこれを非言語的側面に拡張する。視覚的な非言語的要素は、人間らしさの認知に大きな影響を与えるためである。

#### 2.1.1 人間らしさの身体的表現

本研究では人間らしさを単なる言語的応答の正確さではなく、身体的振る舞いを含む全体的な現象として捉える。これは Ryle の行動主義的アプローチを基礎としながらも、視覚的に認知される非言語的コミュニケーションの重要性に注目する立場である。

特に、デュアルディスプレイによる視覚的な身体表現は AIA の存在感を強化し、人間と AI の相互作用をより自然なものとする可能性を持つ。この観点は、従来のチューリングテストが見落としていた身体的次元を補完するものでもある。

Maurice Merleau-Ponty (1908-1961) が『知覚の現象学 (*Phénoménologie de la perception*<sup>[4]</sup>, 1945)』で指摘したように、人間の身体性は単なる物理的な実体ではなく、世界との相互作用の場として機能する。この視点は、デジタル環境における身体的表現の設計にも示唆を与える。本研究で採用する D<sup>2</sup>UI-AIA は、まさに AIA の受肉：身体的存在感を高めること、でありこれにより AIA と人間間の自然な対話環境の構築を目指すものである。

これは、Ryle の行動主義的アプローチを現代のデジタルインタフェース設計に応用する試みとしても位置づけられる。

## 2.2 不満軽減効果の検証アプローチ

実験を行う上では、計測できる程度に明確なユーザーの反応の差異が必要になる。そのため本研究では利用時の情報の正誤に対するユーザーの不満に注目した。具体的には、D<sup>2</sup>UI-AIA の非言語的振る舞いがユーザーの不満を軽減するという仮説を検証する。

検証のため、非言語的振る舞いの強弱と応答の正誤による2×2の実験環境を構築する。この構造により、非言語的振る舞いが強い場合、誤り発生時の不満軽減効果を独立して評価することが可能となる。

評価手法として、十分な数の被験者から得られるアンケートデータを統計的に処理し、補完的にインタビューを実施する。これにより、個別の主観的評価を母集団の特性として定量化し、かつ詳細な質的データによる解釈を可能とする実験設計とした。

## 3 不満軽減効果の実験詳細

### 3.1 実験の概略

本実験ではAIAとの対話における非言語的振る舞いの違いがユーザー体験に与える影響を検証する。今回は試験会場案内というシナリオを用い、D<sup>2</sup>UI-AIA と、音声のみのVUI (Voice User Interface) -AIA との間での反応を比較する。ユーザーの不満要因としては、特定のグループのみ操作時に誤りを人為的に発生させる。また、VUI-AIA は D<sup>2</sup>UI-AIA から視覚的な非言語機能を切ることで反応条件等を揃える。

### 3.2 VUI のユーザインタフェース理論

VUI ユーザインタフェースの理論的基盤として、S. K. Card (1946-) らが提唱した GOMS モデルを検討する。このモデルは、ユーザーの目標、操作、手法、選択ルールを体系的に記述することを可能にする。

GUI 創生を担った Card たちの提唱した GOMS は、本質的に「ユーザーの認知プロセスとタスク構造の分析モデル」であるため VUI への応用が可能である。

現在、会話をを行うのみである AIA に、GOMS モデルに則った実験用の VUI を設計・実装する。

### 3.3 実験の内容と実施手順

#### 3.3.1 実験環境の構成比較

- D<sup>2</sup>UI-AIA (非言語的振る舞いが強い)
  - \* 3D アバターによる視覚的フィードバック
  - \* 感情パラメータに基づく表情変化
  - \* 顔認識機能による対話者の追跡

- VUI-AIA (非言語的振る舞いが弱い)
  - \* 音声による対話のみ
  - \* D<sup>2</sup>UI-AIA と同等の音声応答遅延

#### 3.3.2 参加者とその分類

実験参加者は HAL 大阪に在籍する学生に依頼する。表 1. に示される 4 つのグループで構成される。

	D <sup>2</sup> UI の案内		VUI の案内	
	正	誤	正	誤
グループ A	○			
グループ B		○		
グループ C			○	
グループ D				○

表 1: 実験参加者のグループ分類

#### 3.3.3 実験手順

実験参加者に対しては、次の手順で試験を実施する。また、基本的にはひとり 1. 回のみの参加とする。

- ① 受付で AIA に対して受験番号入力
- ② AIA からの教室番号案内
- ③ 受験教室への移動
- ④ 受験教室の表示から正誤を照合
- ⑤ 誤案内の場合は正しい教室への再移動
- ⑥ アンケートに回答して終了

#### 3.3.4 回答のデータ収集と質問例

HAL 大阪の学生は Microsoft Forms を使用したアンケート入力には慣れており、均質な回答が期待できる。また質問例を表 2 にあげる。

以下の項目について 5 段階で評価  
(1.: 全く感じない ~ 5: 非常に感じた)

苛立ち度	案内を受けて苛立ちを感じましたか?
案内の信頼性	案内が信頼できると感じましたか?
案内の許容度	案内が正しくない場合でも許容できますか?
全体的満足度	案内全体にどれくらい満足しましたか?

表 2: 評価項目と尺度

### 3.4 実験サイズの決定について

本研究では 2×2 の要因計画 (案内方法×正誤) を採用し、各グループ 20 名、総数 80 名を目標サンプルサイズとする。まず 40 名でパイロット実験を実施し、実験手順の妥当性を確認した上で残りの本実験を行う。これにより、統計的な信頼性を確保しつつ、効率的なデータ収集を実現する。

## 4 まとめ

### 4.1 実験の実施に向けて

実際の実験については1月末から2月にかけて実施を予定している。HAI2025のポスターセッションでは実験結果と考察を含めて提案できればと考えている。

今回の実験の結果として、AIAの非言語的振る舞いが、ユーザーに不満を軽減させる効果が認められても、AIAが人間らしい要素を持ち、またその人間らしさを定量的に計測できるようになったと考えるのは早計である。

それでも以下のような有用性を期待しての実験であることをあらためて強調したい。

### 4.2 DX社会における人間とAIの協調

今日の日本社会では、他国との労働効率の比較や少子化など様々な要因から、就労時間の短縮が求められ、DX化というキーワードで従来サービスの転換が図られている。そしてサービス提供側の役務を、デジタルなインタフェースを用いてユーザー側の労務に置き換えるパターンが増えている。具体的には、店舗でのセルフレジや居酒屋でのタブレット注文、お客様相談センターのチャットボット化などである。これらはデジタルデバインドなど多くの問題を孕みつつも、急速に社会へ浸透している。

こうした状況下で、デジタル機器のインタフェースの一部として非言語的振る舞いの可能なAIAを介在させることは、ユーザーの不満を低減する効果が期待できるのではないだろうか。ユーザーがAIAの非言語的振る舞いを認知することで、デジタルインタフェースをより良く活用し、より良いサービスを楽しむ可能性を追求したい。そのために今回の実験では、不満低減効果の測定を指標として選択している。

### 4.3 音声認識の改善

音声認識については、すでにいくつかの音声を分離するAIモジュールを実験しているが、今のところ十分な改善には至っていない。もともと本システムはChatGPTとの間で、話者の音声を変換した文字列をパケット処理している。これに対して、話者発話と背景雑音を分離するために使用するWEBサービス系のAIツールは、音声をストリームで処理するため割り込みが発生し相性が悪い。仮に動作した場合でもひどく遅延して会話が成立しない。

話者の音声を拾うマイクに関しては、多くの場合、ディスプレイの前に設置している機器が多い。この理由を自分たちで実装してみてよくわかった。現在はローカル環境で動作するGPU依存の発話音声分離ツールなどを使用し検証を繰り返している。

WEBサービス型AIの速度が頭打ちの現状で、高性能AIをローカル環境に持つことは、従来と比較にならないレベルのリアルタイム処理を実現する。オンプレミス型のAI機器が安価で運用可能になれば、状況は激変するだろう。先日CES2025ではNVIDIAが、高性能なGrace BlackwellプラットフォームのAIスーパーコンピュータを、価格数十万円で発表した。状況の変化は早いかもしれない。

### 4.4 機材やソフトウェアの公開

D<sup>2</sup>UI-AIAの部材リトや組み立て方、ソフトウェアのブロック図やUnityのパッケージ等については下記に一部公開しており、今後随時更新予定である。

[https://yahaco.xsrv.jp/hal\\_rd/index.html](https://yahaco.xsrv.jp/hal_rd/index.html)

## 謝辞

HI2024での発表時にさまざまなご意見や助言をいただいた研究者のみなさまに感謝を申し上げます。日頃からオープンソースの開示・管理に携わるボランティアのエンジニア諸氏に深く感謝します。OSSがなければこのような研究は成り立ちません。

HAL名古屋、HAL東京のR&Dグループメンバーには、手探りのAI研究において情報交換でずいぶん助けていただき感謝しております。また日本教育財団、HAL教務部には今回このような研究の機会をいただき大変感謝をしております。最後にご多忙な中で試作版の評価や助言を快く引き受けていただいたHAL大阪の同僚教員に深く感謝いたします。

## 参考文献

- [1] Turing, A. M. "Computing Machinery and Intelligence." *Mind*, vol. 59, no. 236 (1950), pp. 433–460.
- [2] Kobayashi, Takuya, and Kohei Saino. "With the Virtual Body of Double-Sided Monitors and the Identification of an Interviewer, an AI Agent Imitates a Human." (2024).
- [3] Ryle, Gilbert. *The Concept of Mind*. (1949).
- [4] Merleau-Ponty, Maurice. *Phénoménologie de la perception*. Gallimard, Paris (1945).
- [5] Card, Stuart K., Thomas P. Moran, and Allen Newell. *The Psychology of Human-Computer Interaction*. L. Erlbaum Associates (1983).