

共感的振る舞いを備えた対話型アバターによる心理的安全性の創出：マルチモーダル同調と視覚的共有の効果

曾根 周作^{1,2} 鈴木 明宏¹ 徳吉 陽河³ 虫明 元² 乾 健太郎^{5,2,4}

¹一関工業高等専門学校 ²東北大学 ³コーチング心理学協会 ⁴理化学研究所 ⁵MBZUAI
shusaku.sone@ichinoseki.ac.jp

概要

本研究は、共感的振る舞いを備えた対話型アバターが、ユーザの心理的安全性および対話相手との関係性に関わる主観的評価に与える影響を検討することを目的とする。従来の共感対話研究が情動変化に焦点を当てるが多かったのに対し、本研究では「安心して語れる」「理解されている」と感じる関係性の質に着目した。大規模言語モデルに基づく受容的応答と非言語的同調を統合した共感対話アバターを開発し、16名を対象とした被験者内実験により、共感的振る舞いを備えた条件と統制条件を比較した。その結果、共感アバター条件では、心理的安全性、信頼感、共感知覚、自己開示意欲が有意に高く評価された。一方で、ポジティブ・ネガティブ感情および状態不安には条件間で有意な差は認められなかった。これらの結果は、共感的対話設計がユーザの情動を直接変化させるというよりも、安心して自己を開示できる対話関係の形成に寄与することを示唆している。

1 はじめに

近年、音声認識・音声合成技術および大規模言語モデルの発展により、対話型エージェントは高い言語的自然性を獲得し、傾聴や相談といった対人的文脈においても活用が進んでいる。一方で、対話が流暢に成立することと、ユーザが「安心して話せる」「率直に発言できる」と感じることは必ずしも一致しない。とくに個人的・感情的な話題を含む対話では、応答内容の適切さに加え、対話関係そのものの質が重要となる。

こうした関係性の質を捉える概念として、心理的安全性 (psychological safety) が知られている。心理的安全性とは、他者から否定的に評価される不安を感じることなく、自由に発言や自己開示ができると

信じられる状態を指す。Edmondson は、心理的安全性が対人関係における重要な基盤概念であることを示した [1]。対話型エージェントとの相互作用においても、ユーザが安心して話せると感じられるかどうかは、対話の継続性や深まりに大きく影響する。

本研究では、共感的な非言語的振る舞いと対話内容に即した視覚的提示を統合した対話型アバターを構築し、その効果を心理的安全性、信頼感、自己開示意欲、共感知覚といった関係性評価の観点から検証する。情動状態の変化のみに着目する従来の枠組みに対し [2]、ユーザがどの程度安心して話せると感じたかという関係の条件に注目する点に本研究の特徴がある。

2 開発したアバター対話システム

本研究で用いたアバター対話システムは、既存のマルチモーダル対話基盤およびアバター描画技術を基礎とし、心理的安全性の醸成を目的として共感的な非言語的振る舞いを新たに統合したものである。図 1 にシステム全体構成を示す。

対話制御には MulmoChat [3] をベースとした対話制御・OS 機能 (MulmoOS) を用い、アバターの具現化には 2D 画像を 3D アバターにする技術 [4] を基盤とした AvatarFacialViewer を用いている。本研究は、これら既存基盤の上に、共感的振る舞いを対話操作変数として明示的に設計・制御した点に特徴がある。

本研究で操作する「共感的振る舞い」は、(1) 表情同調、(2) 呼吸同調、(3) 対話内容の視覚的共有、(4) 非暴力コミュニケーション (NVC) に基づく共感的言語応答方策の有無として実装されており [5]、これらを備えた条件と備えない条件を比較することで、対話関係に与える影響を検証する。

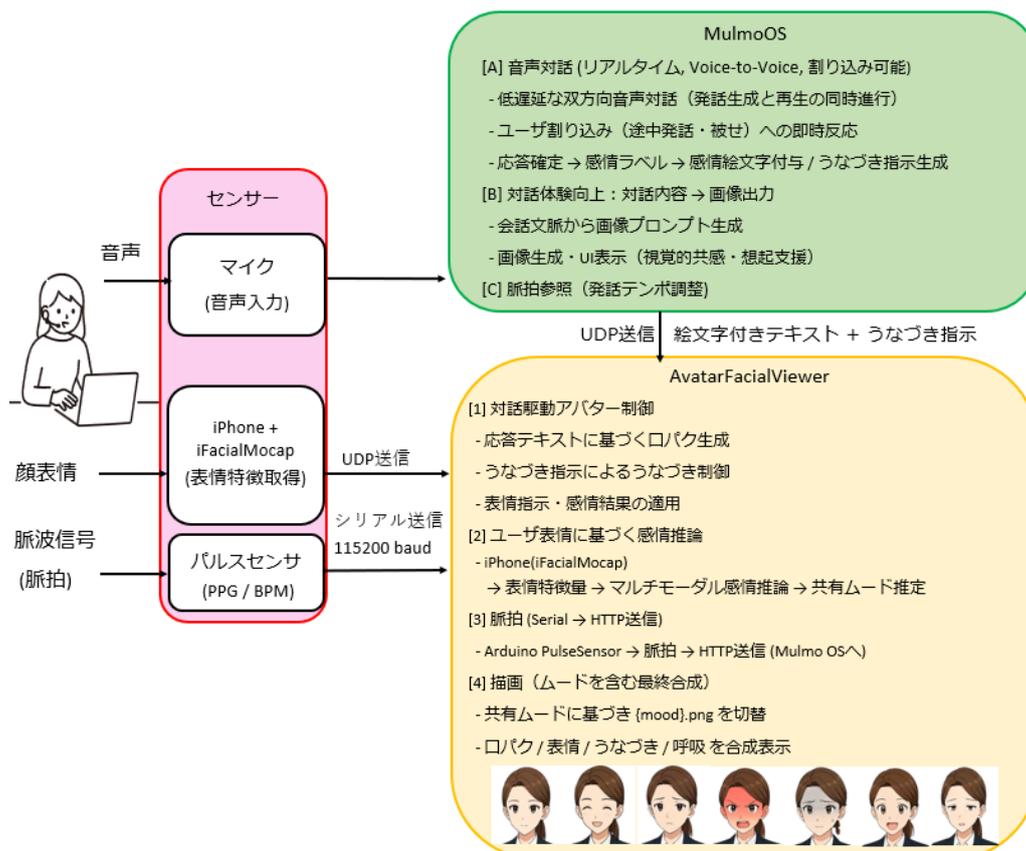


図1 提案するマルチモーダル共感アバターシステムの構成図。本研究では、表情同調・呼吸同調・視覚的共有・非暴力コミュニケーション（NVC）による共感的言語応答を共感的振る舞いとして統合している。

2.1 システム構成概要

システムは、音声対話制御を担う対話制御モジュール（MulmoOS）と、アバターの描画および非言語動作生成を担うアバターモジュール（AvatarFacialViewer）から構成される分散型アーキテクチャを採用している。両モジュールはUDPおよびWebSocketを介して非同期に連携し、低遅延な音声対話とリアルタイムな非言語表現を実現している。

音声対話には OpenAI Realtime API [6] を用い、音声入力から応答生成・音声出力までを一貫して処理することで、自然な対話テンポを維持している。

2.2 共感的対話制御

応答生成には、非暴力コミュニケーション（NVC）[7] に基づくプロンプト設計を採用した。ユーザの発話を評価・判断せずに受容し、感情やニーズに焦点を当てた応答を生成することを基本方針としている。また、即興演劇における「Yes, and」の原則を取り入れ、ユーザの発話を否定せず対話を継続・展開

する設計とした [5]。

2.3 非言語的共感表現の具現化

AvatarFacialViewer は、2D 画像から自然な頭部運動および表情変化を生成する Talking Head Anime 技術 [4] を基盤として開発されている。本研究では、この基盤に対し、ユーザの表情情報および生体信号を入力として統合し、以下の共感的振る舞いを実装した。

- **表情同調**：ユーザの表情変化に応じてアバターの表情を動的に変化させ、情動的な共鳴を視覚的に表現する。
- **呼吸同調**：心拍数に基づいてアバターの呼吸リズム（胸部の上下動）を生成し、無意識的な身体同調感を演出する。
- **視覚的共有**：対話内容に対応する視覚イメージを提示することで、ユーザに「話が正確に理解されている」という理解感を与える。

これらの非言語的振る舞いは、共感アバター条件においてのみ有効化され、非共感アバター条件では



図2 実験実施風景。被験者はアバター対話画面を閲覧しながら、音声対話および表情・生体信号計測を行った。

すべて無効化されている。

3 実験方法

開発したマルチモーダル共感アバター対話システムが、ユーザの心理的安全性および自己開示に与える影響を検証するため、被験者内設計による比較実験を行った。

3.1 参加者

実験には、高等専門学校に在籍する学生 16 名（男性 9 名、女性 7 名）が参加した。年齢は 18 歳から 22 歳であった。すべての参加者に対し、実験の目的、手順、取得データ（音声・生理・脳活動）の取り扱い、および参加の任意性について、書面および口頭で説明を行い、署名によるインフォームド・コンセントを得た。本実験は、一関工業高等専門学校研究倫理審査委員会の承認（承認番号 R7-9）を得て実施された。

3.2 実験条件

参加者は以下の 2 条件のアバターと対話を行った。順序効果を相殺するため、条件提示順序はカウンターバランスをとった。

1. **共感アバター条件 (Empathic)**：提案システムを適用したアバターであり、表情ミラーリング、呼吸同調、視覚的共有表現、および非暴力コミュニケーション (NVC) に基づく共感的・受容的応答を行う。
2. **非共感アバター条件 (Non-empathic)**：対照条件として、表情は固定、呼吸は一定とし、応答は情報伝達の・事務的な内容に留め、視覚的共有や情動的フィードバックは行わない。

3.3 実験手順

1. **事前測定**：実験開始前に、一般的な心理的安全性、状態不安、および気分状態に関する質問紙に回答させた。
2. **ベースライン計測**：センサ装着後、安静状態で 5 分間のベースライン計測を行った。
3. **対話セッション**：各条件のアバターと最大 10 分間の自由対話を行った。対話テーマは「最近の出来事」「趣味」「悩み」などから参加者が自由に選択可能とした。
4. **事後評価**：各対話セッション直後に、アバターとの対話体験に関する主観評価質問紙に回答させた。

3.4 評価指標

- **心理的安全性**：Edmondson[1] の心理的安全性尺度を基に、評価対象を「アバターとの対話」に適合させるため、設問文言を調整した 7 項目を用いた。
- **共感知覚・信頼感**：アバターが利用者の感情を理解し、信頼できる存在として知覚されたかを評価する研究用質問項目群（8 項目）を用いた。
- **感情状態**：STAI（状態不安）[8] および PANAS（気分）[9] を理論的基盤とし、人-アバター対話後の即時的情動状態を測定する目的で項目数および評定方法を調整した研究用適応尺度を用いた。
- **自己開示意欲**：自己開示の深さおよび心理的リスクに関する先行研究[10]を参考に、特定のアバターに対する主観的な自己開示許容度を測定する 8 項目からなる質問紙を作成した。

なお、本実験では、表情情報、音声、言語、脳活動、脈拍などの信号も取得しているが、本稿では主観的体験に焦点を当てるため、分析対象は質問紙による自己報告データに限定した。

4 実験結果

共感アバター条件と非共感アバター条件の間で、各評価指標に差が認められるかを検討する。分析対象は、両条件の対話を完了した 16 名の有効回答データである。条件間の比較には、被験者内デザインに基づき対応のある t 検定を用い、効果量として被験者内デザイン向けに算出した Cohen's d_z を報告す

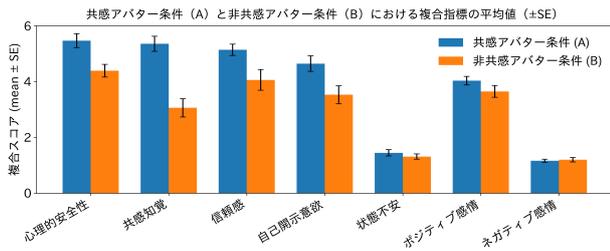


図3 共感アバター条件 (A) および非共感アバター条件 (B) における、各合成尺度の平均値と標準誤差 (Mean ± SE)。

る。有意水準は5%とした。

4.1 関係性および対話評価の比較

アバターとの関係性および対話の質に関する主観評価の結果を表1に示す。いずれの関係性指標においても、共感アバター条件は非共感アバター条件よりも有意に高い値を示した。

また、主要な合成尺度の平均値と標準誤差を図3に示す。関係性指標では一貫して共感アバター条件が高い値を示す一方、情動指標では条件間差が小さいことが確認できる。

4.2 情動状態の変化

対話後に測定された一時的な情動状態について、共感アバター条件と非共感アバター条件の比較結果を表2に示す。ポジティブ感情は共感アバター条件で高い傾向を示したが、有意差には至らなかった。ネガティブ感情および状態不安については、条件間に有意な差は認められなかった。

5 考察

本研究の結果から、共感的振る舞いを実装したアバターは、ユーザの心理的安全性、共感知覚、信頼感、自己開示意欲といった対話相手との関係性に関する主観評価を有意に向上させることが示された。一方で、ポジティブ・ネガティブ感情や状態不安といった一時的な情動状態そのものには、有意な差は認められなかった。

これらの結果は、共感アバターの効果が、ユーザの感情を直接変化させるというよりも、「安心して話せる相手」としての知覚を形成する点にあったことを示唆している。

表1 関係性・対話評価における条件間比較 (Mean ± SD)

尺度	共感条件	非共感条件	t(15)	d
心理的安全性	5.47 ± 0.99	4.39 ± 0.91	3.33**	0.83
共感知覚	5.36 ± 1.08	3.06 ± 1.31	4.95***	1.24
信頼感	5.15 ± 0.81	4.06 ± 1.49	2.31*	0.58
自己開示意欲	4.65 ± 1.12	3.53 ± 1.29	2.53*	0.63

(*p < .05, **p < .01, ***p < .001)

表2 情動指標における条件間比較 (Mean ± SD)

尺度	共感条件	非共感条件	t(15)	p
ポジティブ感情	4.04 ± 0.62	3.65 ± 0.85	1.93	.073
ネガティブ感情	1.16 ± 0.22	1.20 ± 0.33	-0.45	.66
状態不安	1.45 ± 0.45	1.31 ± 0.38	1.08	.30

6 結論

本研究では、共感的振る舞いを備えたアバターがユーザの主観的対話体験に与える影響を被験者内実験により検討した。その結果、共感アバターは、心理的安全性、共感知覚、信頼感、自己開示意欲を有意に向上させる可能性が示された。

これらの結果は、共感的対話エージェントの設計において、情動変化の誘発だけでなく、関係性の質に着目することの重要性を示している。

参考文献

- [1] Amy C. Edmondson. Psychological safety and learning behavior in work teams. **Administrative Science Quarterly**, Vol. 44, No. 2, pp. 350–383, 1999.
- [2] Yukun Ma, Khanh Linh Nguyen, Frank Z. Xing, and Erik Cambria. A survey on empathetic dialogue systems. **Information Fusion**, Vol. 64, pp. 50–70, 2020.
- [3] Receptron. Mulmochat. <https://github.com/receptron/MulmoChat>, 2025.
- [4] Pramook Khungurn. Talking-head anime 3. <https://pkhungurn.github.io/talking-head-anime-3/>, 2022.
- [5] 虫明元, 山口晴保. 認知症ケアに活かすコミュニケーションの脳科学 20 講: 人のつながりを支える脳のしくみ. 協同医書出版社, 東京, 日本, 2023.
- [6] OpenAI. Realtime api. <https://platform.openai.com/docs/guides/realtime>, 2024. Accessed: 2026-01-02.
- [7] Marshall B. Rosenberg. **Nonviolent Communication: A Language of Life**. PuddleDancer Press, Encinitas, CA, 2003.
- [8] 清水秀美, 今栄国晴. State-trait anxiety inventory の日本語版 (大学生用) の作成. **教育心理学研究**, Vol. 29, No. 4, pp. 348–353, 1981.
- [9] 佐藤徳, 安田朝子. 日本語版 panas の作成. 日本パーソナリティ心理学会, 2001.
- [10] 丹羽空, 丸野俊一. 自己開示の深さを測定する尺度の開発. **パーソナリティ研究**, Vol. 18, No. 3, pp. 196–209, 2010.
- [11] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. Empathy in virtual agents and robots: A survey. **ACM Trans. Interact. Intell. Syst.**, Vol. 7, No. 3, September 2017.

A 質問紙尺度 (Questionnaires and Scales)

本研究では、共感的振る舞いを備えた対話型アバターがユーザの主観的体験に与える影響を評価するため、事前質問紙および事後質問紙を用いた。本付録では、各尺度の構成と測定意図を簡潔に示す。なお、すべての尺度は既存尺度を理論的基盤としつつ、人-アバター対話の文脈に適合させるために項目選択や表現を調整した研究用尺度である。

A.1 事前質問紙

- **心理的安全性 (一般)** : 7項目・7件法. オンライン対話一般における安心感と自己表現のしやすさを測定. Edmondson (1999) [1] を参考に作成.
- **状態不安 (STAI-S 参照)** : 6項目・4件法. 実験前の一時的な不安状態を測定. STAI-S [8] を理論的基盤とした短縮・適応版.
- **気分状態 (PANAS 参照)** : 10項目 (ポジティブ5, ネガティブ5)・5件法. 実験前の正負の感情状態を測定. 日本語版 PANAS [9] を基に構成.
- **AI 対話システムへの事前印象** : 5項目・5件法. AI 対話システム一般に対する印象を測定.

A.2 事後質問紙

- **心理的安全性 (アバター対話後)** : 7項目・7件法. 当該アバターとの対話における心理的安全性を測定. Edmondson (1999) [1] を基に作成.
- **自己開示意欲** : 8項目・7件法. 個人的・内面的内容の自己開示に対する主観的許容度を測定. 自己開示の深さと心理的リスクに関する先行研究 [10] を理論的背景とする.
- **状態不安 (再測定)** : STAI-S を基盤とした適応尺度を用い、対話後の不安状態の変化を評価.
- **気分状態 (再測定)** : PANAS 参照尺度を用い、対話後のポジティブ/ネガティブ感情の変化を評価.
- **共感知覚・信頼感** : 7項目 (共感知覚4, 信頼感3)・7件法. アバターの共感性および信頼性の知覚を測定. Paiva et al. (2017) [11] を参考に作成.
- **システム使用感** : 4項目・5件法. 操作性および使いやすさを評価. SUS を理論的基盤とする.

B 対話アバターのプロンプト (要約)

本研究で用いた対話アバターのプロンプトは、実運用上はシステム制御や安全制約を含む長文から構成される。本付録では、条件間の操作要因として本質的な差分に相当する応答方針のみを要約して示す。なお、省略された共通部分は両条件で同一であり、条件間比較に影響しない。

B.1 共感アバター (Empathic) の応答方針

共感アバターは、非暴力コミュニケーション (NVC) および即興演劇における “Yes, and” の原則に基づき、以下の方針で応答を生成した。

- ユーザの発話内容を否定せず、まず受容する。
- 発話の背後にある感情やニーズを推測し、感情を言語化して返す (例: 「それは大変でしたね」)。
- 判断・評価・訂正を行わず、ユーザの世界観を尊重する。
- 共感的応答を基本としつつ、ユーザが明示的に求めた場合には実用的な情報や助言を補足する。
- 応答は短く、対話のテンポを重視する。

B.2 非共感アバター (Non-empathic) の応答方針

非共感アバターは対照条件として、情動的・共感的要素を排除し、以下の方針に従って応答を行った。

- 感情や内的状態への言及を行わない。
- 表情変化や生理的同調を伴わない。
- 応答は事実や情報の提示に限定し、情緒的表現を含めない。
- 親しみや共感を示す言い回しを避け、中立的・事務的な文体を用いる。

以上のように、両条件は「共感的応答の有無」を主要な操作要因として設計されている。