



AR 会話エージェントの信頼に 視覚的・物理的なタッチが与える影響

The Impact of Visual and Physical Touch on the Trustworthiness of Augmented Reality Conversation Agent

松本篤弥¹⁾, Yang Chi Lan¹⁾, 葛岡英明¹⁾

1) 東京大学 (〒 113-8656 東京都文京区本郷 7-3-1, {atusya.maya, chilan.yang ,kuzuoka}@cyber.t.u-tokyo.ac.jp)

概要: パーソナルアシスタントやチャットボットなどの会話エージェントが普及しているが、音声認識エラーなどの破綻時に、適切に会話を修復できず、ユーザーの期待に応えられない場合、信頼感が低下することが課題となっている。エージェントが破綻を検知せずに信頼感を得る方法として、AR 環境での視覚的・物理的タッチを提案し、15 人を対象に破綻時に低下した信頼感の回復を試みた。信頼感の回復は確認されなかったが、破綻の発生による意思決定の変化が確認された。

キーワード: AReality, ソーシャルタッチ, 信頼, 会話

1. はじめに

チャットボットやパーソナルアシスタントなどの会話エージェント (以下, エージェント) は、情報検索, タスクの実行, 雑談などの多様な機能を提供することを、ユーザに期待させる。エージェントはこれらの期待に応えることで、ユーザと信頼関係を構築し、生活をサポートする [1]。

エージェントは、機械とユーザのインタフェースとして急速に普及しているが、エージェントとの会話はユーザの期待に常にえられるわけではない。エージェントは自然言語処理のエラーにより会話の文脈を無視した応答を返したり、音声認識のエラーによりユーザの発話を理解できなかったりする。それにより、ユーザは、往々にして、エージェントに起因する会話の「破綻」を経験する。このような破綻が生じたときに、ユーザは、エージェントが適切な言葉の言い換えや訂正などにより、会話を「修復」することを期待する。しかし、エージェントは常に会話を修復できるわけではない。そのような場合、ユーザの期待するような会話ができないとみなされ、エージェントへの信頼感が低下することがある。さらには、時として、ユーザはエージェントの会話能力を過小評価し、複雑な会話を避けるようになる [1]。

エージェントがユーザの期待に応え信頼関係を構築することは、ユーザを最大限にサポートするために重要である。信頼関係を構築するために、会話の破綻に適切に対処することは、エージェントの設計における課題の 1 つとなっている。従来の研究 [2, 3] では、エージェントが会話の破綻を検出し、「すみません、うまく理解できませんでした。」といった言語的な手段で会話の破綻を修復してきた。しかし、破綻の原因となる自然言語処理や音声認識のエラーをすべて検知し修復することは、現状では困難である。特に雑談においては、話題がすぐに切り替わったりすることから、会話の破綻の原因は多岐にわたるため、音声やテキストなどの多様な言語的修復方法が必要となる。そのため、タスク

の実行や情報検索と比較して、会話の破綻の検知と修復が困難である。故に、言語的な手段を用いない信頼感の回復手段についても検証を行う必要がある。

信頼感とは人間が期待した行動をエージェントが実際にとるかというパフォーマンスに対する信頼感と、人間の社会的規範を遵守する行動をとるかという道徳的な信頼感の二つに分けられる [4]。会話の破綻の修復は、パフォーマンスに対する信頼感にあたり、エージェントが修復に失敗するとき、ユーザの信頼感が低下する。コミュニケーション手段の一つであり、信頼感に影響を及ぼすインタラクションとして「ソーシャルタッチ (コミュニケーションのための身体接触)」 (以下, タッチ) が存在する [5]。身体への接触は生理学的には CT 繊維と呼ばれる求心性神経に対して刺激を与えることであり、心理学的には感情的の伝達や向社会的な行動の促進などの影響を与えることができる [6]。先行研究では、タッチがエージェントへの信頼感を向上させる効果があることが示されている [5]。しかし、その効果が、会話の破綻により低下したエージェントへの信頼感に対しても同様に働くかどうかは明らかになっていない。そのため本研究では、エージェントとのタッチが会話の破綻で生じる信頼感の低下を回復する役割を果たすと仮説を立て、検証を行った。

2. 関連研究

2.1 会話の破綻と修復

発言の誤りにより、互いのコミュニケーションが理解できないことや内容が誤解されることが生じるため、発言の誤りを正しいものに置き換えることが人間同士の相互の信頼感に重要であり [7], またユーザとエージェントとの信頼感においても重要である [2]。そのため、エージェントとの会話中の破綻の修復を適切に行い、信頼感の回復を目指す研究が行われている。Ashktrab らはエージェントが破綻を

検出したときに、コミュニケーション理論に基づく8つの修復戦略のうち、どのような修復戦略がユーザに好ましいか比較し、ユーザがオプションや説明の修復戦略を好むことを示した[2]。このような既存の手法では、ユーザの信頼感を回復するために、エージェントがエラーを検知し破綻を修復する必要がある。しかし、エージェントはこのようなエラーを完璧に検知できるわけではない[8]。そのため、本研究では自然言語や音声以外の手法で信頼感を回復させる手段として、タッチ(接触行為)に着目した。

2.2 タッチによる信頼感への影響

信頼感はエージェントとの適切なタッチにより向上する。Lawらは人間とソーシャルロボットの間でのタッチが、ロボットに対する人間の信頼感に及ぼす影響を、オンライン研究により調査した。ソーシャルロボットが人間に対してタッチする様子を観察者に見せた後、観察者がソーシャルロボットへの信頼感を評価した。その結果、ソーシャルロボットとのタッチが観察された場合、特に誠実さや倫理性といった観点でロボットへの信頼感が全体的に増加したことがわかった。しかし、タッチの行為が不適切と感じられる場合や、ソーシャルロボットのタッチに関する設計目的が明示された状況下では、逆にソーシャルロボットへの信頼感が減少する可能性も示された[5]。

上記のように文脈に応じたタッチによりエージェントへの信頼感が向上する。しかし、会話の破綻により低下した信頼感に対し、タッチがどのような影響を与えるかについては明らかになっていない。そのため、本研究は破綻が生じるエージェントとの会話を再現し、その状況下でエージェントとのタッチを付与することで、タッチの有無が人間の主観的な信頼感へ与える影響について評価した。

3. 実験

3.1 参加者とタスク

参加者は頭部装着型拡張現実ディスプレイ(ARHMD)であるMeta社(カリフォルニア, アメリカ合衆国)のOculus Quest Proを装着した。ARHMDで提示するエージェントについては、物理的に接触する際に参加者に不快感を与えにくいと考えられる猫の外見のエージェントを採用した。

実験条件については、タッチの実施形態の違いによる信頼感への影響を調べるため、以下のように3つの異なるタッチ条件で参加者間実験を行った。

- 猫のエージェントにタッチすると視覚的に反応する条件(視覚的タッチ条件)
- 猫のエージェントの頭部と同じ位置にぬいぐるみが置かれ、タッチすると視覚的に猫が反応するだけでなく、物理的なフィードバックを得られる条件(物理的・視覚的タッチ条件)
- 猫のエージェントが表示されるがタッチを行わない条件(タッチ無し条件)

そして、雑談中の会話の破綻の影響とタッチの関係を調べるため、すべてのタッチ条件において、会話の修復が必要

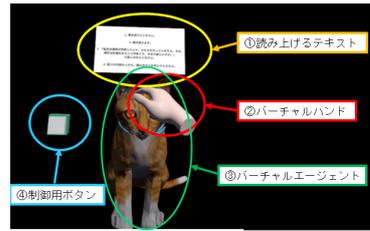


図1: バーチャルな猫を撫でている画面。

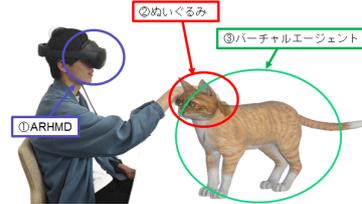


図2: バーチャルな猫を撫でている様子。

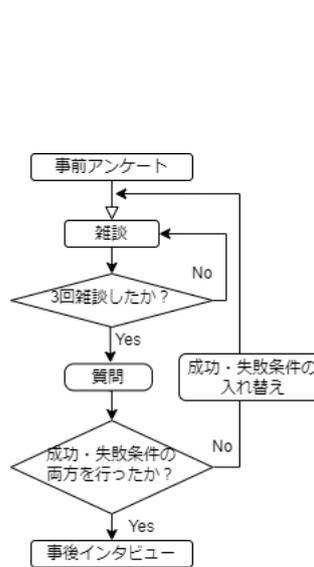


図3: 雑談全体の手順。

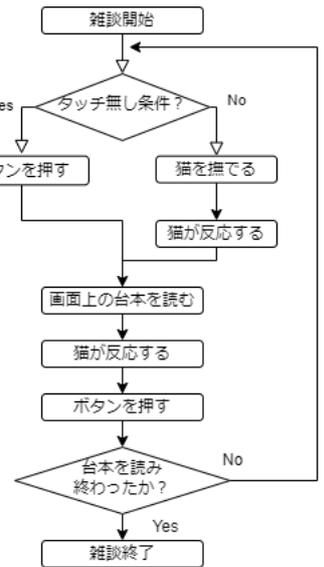


図4: 実験全体の手順。

ない条件である成功条件と会話の修復が必要な破綻条件という2つの会話条件について被験者内で比較した。

参加者として、男性15人(年齢:Mean=22.9, Std=0.80)が募集され、すべての参加者はタッチ条件に沿った3つのグループにランダムに割り振られた。割り振られた各グループで、参加者はAIであると伝えられた猫のエージェントと図3の手順で複数の応答を含む3回の雑談を行い、最後に質問に答えた。実験は途中で邪魔の入らない部屋で実施され、図4に示すような手順で行った。

3.2 実装

Meta Quest Proではビデオシースルーにより周囲の現実環境の映像をディスプレイ上に表示し、その上にバーチャルのオブジェクトを表示できる。実験用のアプリはゲームエンジンのUnityで実装し、図5のように表示された。参加者はバーチャルな手で猫を撫でることやボタンを押すことでエージェントとインタラクションを行った。会話で読み上げるスクリプトは、実際の雑談チャットボットの会話例をもとに日本人に適した形に実験実施者が修正したものを表示した。成功条件のスクリプトでは日本語に修正したものを使用し、失敗条件のスクリプトでは、音声対話チャットボットの実環境で発生するエラーを参考に、文脈を無視したり、間違った知識を示す応答や音声認識に失敗する応答を取り入れたスクリプトを使用した。

3.3 評価指標

会話破綻時のタッチが信頼感に与える影響を調べるため、

- 信頼感のアンケートの測定値が向上する
- エージェントからのアドバイスにより意思決定が変化しやすくなる

という仮説を立て検証した。

また、タッチによる認知への影響を探索的に調べるため、

- 知覚概念・心理的距離についてアンケートの測定値が向上する
- エージェントに触れている時間が長くなる
- 参加者の声の大きさとピッチが上昇する

という仮説を立て検証した。

信頼感に関して、会話の破綻に対する修復では、エージェントの会話を遂行する能力が求められると考えられるため、パフォーマンスに対する信頼感を測定した。

知覚概念は、エージェントに関連する5つの知覚概念を測定する指標である。これらの項目の内、参加者は静止物体に触れるため、安全性の知覚以外の人間らしさ・生命観・好感度・知性の知覚の項目について測定を行った。

心理的距離は、個人がエージェントと自分自身の間にどれほどの「近さ」を感じているかという概念である。本研究では、心理的距離に関する5つの質問を実施した。

また、行動的な指標として、参加者の意思決定の変化、音声入力時に会話エージェントに触れている時間について測定した。参加者が会話エージェントを信頼している場合、意思決定に対して会話エージェントの提案により意見を変更する可能性が存在する。意思決定の変化については、提案前と提案後に質問の回答に対する意見に変化があった場合に変化があるものとして計測した。そのほかの指標について、参加者が会話エージェントとの心理的距離が近い場合に接触時間が長くなると考えられたため、エージェントに触れている時間について計測した。また、タッチによりポジティブな感情が想起される場合があるため、感情の覚醒価の指標として音声の大きさ・ピッチの変動についても計測した。

4. 結果

初めに、信頼感、心理的距離、知覚概念における人間らしさ・生命観・好感度・知性の知覚について、正規分布でない群が含まれていたため、整列ランク変換を行い、有意水準 $\alpha = 0.05$ として二元配置分散分析 (2-way ANOVA) を実施した。結果として、信頼感 ($\eta_p^2 = 0.68, p < 0.01$)・心理的距離 ($\eta_p^2 = 0.48, p < 0.01$)、知覚概念における生命観 ($\eta_p^2 = 0.55, p < 0.01$)・好感度 ($\eta_p^2 = 0.59, p < 0.01$) の指標間に会話条件間で有意な差異が存在することを確認した。一方で、タッチ条件間においては有意な差異は存在しなかった。また、全ての指標について交互作用は確認されなかった。有意な差異が認められた指標について、正規性が確認されたため会話条件ごとに t 検定を行い、ボンフェローニ法による p 値の補正を行った上で事後分析を実施した。結果として、信頼感については物理的・視覚的タッチ条件 ($T = 4.87, p = 0.025$)

および視覚的タッチ条件 ($T = 4.44, p = 0.034$) で会話条件間に有意な差が観察された。また、生命観について視覚的タッチ条件 ($T = 4.32, p = 0.037$) で会話条件間の有意な差が確認された。

続いて、行動的評価についても整列ランク変換後に有意水準 $\alpha = 0.05$ として二元配置分散分析を実施した。その結果、意思決定の変化について、会話条件間で有意な差が観察された ($\eta_p^2 = 0.34, p = 0.028$)。それ以外の行動的指標である、猫に触れる時間、音声の大きさ・ピッチの変動については有意な差は確認されなかった。

また、成功条件と失敗条件の順序が与える影響について、データの正規性・分散の等質性が確認されたため有意水準 $\alpha = 0.05$ として対応の無い t 検定を行った。会話条件について、成功条件から失敗条件へと進む順序で実験を行った群と、その逆の順序で実験を行った群との間で、各評価項の差分を比較した。結果として、主観的評価に関しては、信頼感 ($T = 3.14, p < 0.01$)、好感度 ($T = 2.06, p = 0.049$)、知性の知覚 ($T = 2.70, p = 0.012$)、心理的距離 ($T = 2.54, p = 0.017$) について有意な差が確認され、成功条件から失敗条件の群の方がもう一方と比較して差分が有意に大きいことが確認された。一方で、客観的評価に関しては、順序により有意な差が確認される項目はなかった。

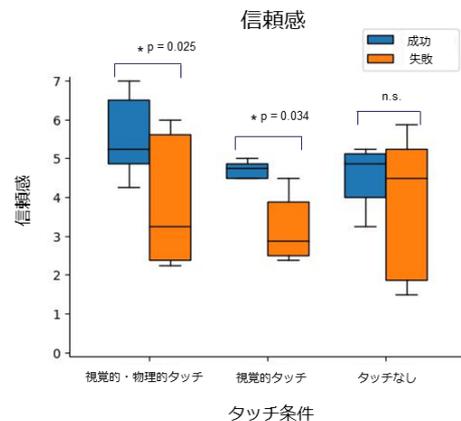


図 5: 実験結果: タッチ条件間での信頼感。

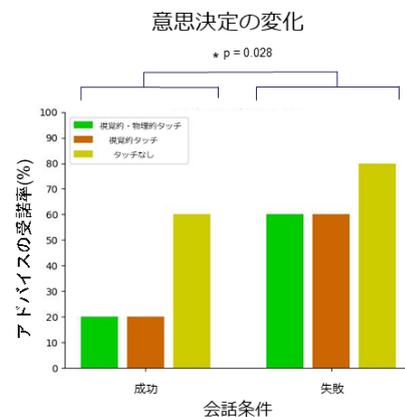


図 6: 実験結果: 会話条件間での意思決定の変化。

5. 考察

5.1 会話の破綻とタッチによる信頼感へ相互作用

結果から、会話の失敗がエージェントへの信頼感を低下させる。一方で、タッチの有無により信頼感は統計的に有意な差が生じないことが確認された。これは、エージェントとの雑談において、言語的な要素が主観的評価に強く影響を与えることを示唆している。会話の破綻が引き起こした信頼性の低下がタッチにより回復しなかった理由として、主観的信頼感の質的な側面の違いが考えられた。信頼感の指標である多次元信頼尺度 (MDMT)[4] は、「道徳的信頼」と「パフォーマンス的信頼」の2つに分けてエージェントに対する信頼感を評価する。その中で、タッチは道徳的信頼感を高める傾向にあるとされている [5]。しかし、雑談するエージェントでは、会話中の語彙的に正確な受け答えや文脈理解のようなパフォーマンス的側面が重要視される。そのため、タッチの影響が小さく、主観的な信頼感が低下したと考えられた。

5.2 会話の破綻による意思決定の変化への影響

意思決定の変化についての結果は、仮説と異なり、会話の破綻が存在する状況において、実験参加者はエージェントからの助言をより受け入れやすく、意思決定を変えることを示している。理由として、エージェントが会話の破綻により脆弱な存在であるとみなされた可能性が考えられた。従来の研究によると、エージェントが自身の弱点を見せると、人はそのエージェントに対する信頼感を高めることが分かっている [9]。会話の破綻により、脆弱なエージェントとしてみなされた結果、アドバイスを受け入れ意思決定を変化させたと考えられた。

6. 制限

本研究の制約として、信頼感の質、エージェントの外見、インタラクションの向きがあげられる。本研究では会話の「パフォーマンス的信頼感」に着目し計測したが、タッチのような非言語的な要素は会話の「道徳的信頼感」に重要であることが考えられる。次に、インタビューで猫が話すことに違和感を覚えた実験参加者が存在した。猫との会話という状況は日常生活とは異なる経験である。そのため、エージェントへの認知に影響を与えた可能性がある。さらに、本研究では実験参加者からエージェントへのタッチを研究の対象としたが、エージェントから実験参加者へのタッチとは効果が異なる可能性がある。これらの制約について、「道徳的信頼感」を計測することやエージェントの見た目やタッチの方向性が雑談を行うエージェントに与える影響を探ることが有益であると考えられる。

7. 結論

本研究では、会話エージェントとの雑談が破綻した状況における信頼感を損なわない方法として、視覚的・物理的なタッチを利用する手法を検証した。結果として、破綻により低下した信頼感をタッチが回復するという仮説は支持されなかった。一方、会話の破綻の存在が参加者の意思決

定に影響を与え、エージェントからの助言をより受け入れる傾向が見られた。これらの結果は、会話中に生じる破綻により低下したエージェントの信頼感について、タッチという非言語的な側面よりも言語的側面が重要であることを示すものであった。今後の研究では、エージェントとの雑談時の信頼感について、エージェントの形状やタッチの方向性などの非言語的な側面から詳細に調査する。

謝辞 本研究は COI-NEXT の助成を受けたものです。

参考文献

- [1] L. Mavrina, J. Szczuka, C. Strathmann, L. M. Bohnenkamp, N. Krämer, and S. Kopp, ““ alexa, you’re really stupid ”: A longitudinal field study on communication breakdowns between family members and a voice assistant,” *Frontiers in Computer Science*, 2022.
- [2] Z. Ashktorab, M. Jain, V. Liao, and J. Weisz, “Resilient chatbots: Repair strategy preferences for conversational breakdowns,” 2019.
- [3] T. J.-J. Li, J. Chen, H. Xia, T. M. Mitchell, and B. A. Myers, “Multi-modal repairs of conversational breakdowns in task-oriented dialogs,” *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 2020.
- [4] B. F. Malle and D. Ullman, “A multidimensional conception and measure of human-robot trust,” *Trust in Human-Robot Interaction*, 2021.
- [5] T. Law, B. F. Malle, and M. Scheutz, “A touching connection: How observing robotic touch can affect human trust in a robot,” *International Journal of Social Robotics*, 2021.
- [6] A. Saarinen, V. Harjunen, I. Jasinskaja-Lahti, I. P. Jääskeläinen, and N. Ravaja, “Social touch experience in different contexts: A review,” *Neuroscience Biobehavioral Reviews*, 2021.
- [7] E. Schegloff, G. Jefferson, and H. Sacks, “The preference for self-correction in the organization of repair in conversation,” *Language*, 1977.
- [8] S. Terragni, B. Guedes, A. Manso, M. Filipavicius, N. Khau, and R. Mathis, “BETOLD: A task-oriented dialog dataset for breakdown detection,” in *Proceedings of the Second Workshop on When Creative AI Meets Conversational AI*, 2022.
- [9] N. Martelaro, V. Nneji, a. Ju, and a. Hinds, “Tell me more: Designing hri to encourage more trust, disclosure, and companionship,” in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2016.