

Title	エージェント化プロトタイプとの対話を利用したユーザテスト手法の検討
Author(s)	吉松, 駿平; 高島, 健太郎; 西本, 一志
Citation	情報処理学会研究報告, 2023-HCI-202(39): 1-8
Issue Date	2023-03-15
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/18451
Rights	<p>社団法人 情報処理学会, 吉松駿平, 高島健太郎, 西本一志, 情報処理学会研究報告, Vol.2023-HCI-202, No.39, 2023, 1-8.ここに掲載した著作物の利用に関する注意: 本著作物の著作権は(社)情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 Notice for the use of this material: The copyright of this material is retained by the Information Processing Society of Japan (IP SJ). This material is published on this web site with the agreement of the author (s) and the IP SJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IP SJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.</p>
Description	



エージェント化プロトタイプとの対話を利用した ユーザテスト手法の検討

吉松駿平^{†1} 高島健太郎^{†1} 西本一志^{†1}

概要: デザイン思考をはじめとしたユーザ中心の設計手法において、ユーザテストの段階では、ユーザ体験の品質を高めて付加価値を向上させるために、プロトタイプをユーザが試用した際の利用体験を得ることが必要となる。そこで、ユーザの内観をその場で得るために用いられるのが Think Aloud 法である。しかし、この手法には、認知的負荷の高さから十分な量の発話を得られない可能性がある。また、発話を支援するべく開発者が介入することで、却って率直な発話が阻害される恐れがある。本研究では、ユーザの率直な発話を促し、ユーザビリティ改善のヒントをより多く引き出すため、プロトタイプそのものに身体的特徴を付与しエージェント化する手法を提案し、人間の開発者が介入した場合との比較実験を行った。

キーワード: ユーザテスト, 思考発話法, ヒューマンエージェントインタラクション, 直接擬人化手法

Examination of user testing methods using interaction with agent-based prototypes

SHUNPEI YOSHIMATSU^{†1} KENTARO TAKASHIMA^{†1} KAZUSHI NISHIMOTO^{†1}

Abstract: In user-centered design methods such as design thinking, during the user testing phase, it is necessary to obtain the user experience when the prototype is used in order to improve the quality of the user experience. Therefore, in order to obtain the user's thoughts and emotions, Think Aloud method is used in the field of user testing. However, this method may not be able to obtain a sufficient amount of speech due to the high cognitive load. Additionally, developer intervention to support speech may cause the user to think too much, preventing intuitive speech and the risk of not obtaining honest reactions and impressions. In this study, we propose a method of adding physical characteristics to the prototype and making it an agent to encourage honest user speech and extract more usability improvement hints. We also conducted a comparison experiment with human developer intervention.

Keywords: User Test, Think Aloud Method, Human-Agent Interaction, Direct Anthropomorphization Method

1. はじめに

1.1 背景

近年、アプリケーションやプロダクト等の開発現場では、デザイン思考をはじめとしたユーザ中心の設計手法が多く用いられている。デザイン思考は問題を解決するために、創造的で柔軟なアイデアを生み出すことを目的としており、次のようなステップからなるプロセスを通じて行われる。

まず、観察やインタビューから、問題点やニーズを明確にする。その後、問題解決のため考え出した様々なアイデアを整理し、最も有効なアイデアを選んだ後に、具体的な形にプロトタイプ化する。そして、作成したプロトタイプを使って、実際に問題解決可能かをテストし、その結果を踏まえてプロトタイプを改善する。

その中でも、プロトタイプをユーザが試用し、使用感や満足度を評価する作業は一般にユーザテストと呼ばれる。ユーザに使ってもらうことで、プロトタイプが実際のニーズを満たしているかを確認することができる。また、多くの場合、ユーザテストは繰り返し行われる。その度に改善

点を特定し、修正を重ねることで完成度を高めていく。

1.2 問題提起

ユーザテストの段階では、ユーザ体験の品質を高めるために、ユーザがプロトタイプを試用した際の利用体験の取得が必要となる[1]。特に思考や感情といった、ユーザの内観は観察のみでは記録不可能なため、その取得には発話が多く利用される。その最も基本的な手法として、ユーザテスト後のインタビューが実際の開発現場でも用いられている[2]。しかし、記憶によるバイアスや、状況を回顧する際に理由を後付けする等、課題も存在する。

そのため、ユーザの内観をその場で取得できる手法が求められる。そこで利用されるのが、ユーザに考えていることを逐一発話させる Think Aloud 法（以下 TA 法）である。Priede et al. [3]は、TA 法がインタビューに比べ、回答者の理解を調べることにより適していると述べている。しかし、ユーザは操作と並行して発話することになり、認知的負荷は高まる。よって、操作に真剣になるあまり発話を忘れてしまう恐れがある。これに対して、ユーザテスト中に他者が質問・相槌することで、発話を促す事例も存在する [1][2]。しかし多くの場合、それらの介入は開発者やユーザリサーチの専門家によって行われるため、ユーザの率直な発話が

^{†1} 北陸先端科学技術大学院大学 先端科学技術研究科
Graduate School of Advanced Science and Technology, Japan Advanced
Institute of Science and Technology

得られにくくなる。

そこで、本研究ではエージェントが介入する TA 法に着目し、ユーザテストへの応用の可能性を模索した。人間の代わりにエージェントが質問・相槌を行うことで、開発者には話しにくいと思われるプロトタイプへの不満や、インタフェース上の問題点を聞き出せるようになることが期待できる。ただし、プロトタイプとは独立にエージェントが存在してユーザに介入すると、プロトタイプ本体ではなくエージェントに関心が集まってしまうことが考えられる。そのため、ユーザが抱くエージェントへの関心をプロトタイプに紐づけることが求められる。

1.3 研究目的

本研究では、ユーザの発話増加と率直な感想の取得を目指し、プロトタイプ自体に身体的特徴を付与しエージェント化する。音声で質問・相槌し、アプリが直接語りかけてくるように感じさせることで、通常の TA 法や開発者の介入と比較してよりユーザに話しやすい印象を与えられることが期待できる。また、機能のわかりやすさや操作しやすさ等、ユーザビリティへの言及の増加を第二の目的とした。

2. 関連研究

2.1 ユーザテストにおける TA 法の実施方法と介入の影響に関する研究

タスク中のユーザへの介入について、Boren et al.[4]は、連続的かつ控えめな肯定があると集中を切らさずに発話可能のため、相槌やフォローアップクエスチョンが有効であると述べている。それに加え、予想外の応答で発話が止まった場合に質問することで、ユーザの説明を促せるとしている。一方で、介入があるとユーザは誤認識・誤操作の責任を自らに感じてしまう恐れがあることも指摘している。ユーザテストで発生する問題の多くは、テストされる製品に原因があり、ユーザの責任ではない。したがって Boren et al.は、「誤認識・誤操作は問題を発見したユーザの貢献であり、問題の責任は製品にある」ということを開発者が強調すべきだと結論づけている。

Olmsted-Hawala et al.[5]の実験では web サイトのユーザテストにて、TA 法中ユーザに質問することで、タスクの成功率と web サイトへの満足度が上がることが指摘されている。この結果より、TA 法への開発者の介入によって、ユーザが抱くプロダクトへの印象は向上し、正当な評価でなくなることが考えられる。Krahmer et al.[6]は、厳密に定義された対話のない TA 法と確認応答のある TA 法とでユーザテストを実施し、被験者の評価や検出できた問題点の数に差がないことを示している。これらの研究から、TA 法に開発者が介入することで、タスクの成功率や満足度など、印象は本来の評価より高く評価され、責任の所在についても課題は残るが、ユーザビリティ上の問題点の検出精度への影響はないと考えられる。

しかし、これらはいずれも開発者による介入が TA 法に及ぼす影響についての研究である。すなわち、人間がユーザに対して質問や相槌する形式であり、その役割をエージェントが担った場合については分析や議論は行われてない。

2.2 プロダクトやアプリケーションのエージェント化に関する研究

DiSalvo et al.[7]の研究が示しているように、人間に似た特徴を持つ物体に対して、ユーザはよりインタラクションを行う傾向がある。そこで、プロダクトやアプリケーションに、目や手などの身体的特徴を付与することでエージェント化する先行研究が多く存在する。これらの手法は直接擬人化手法とも呼ばれ、大澤ら[8]は家電製品を題材として説明対象をエージェント化することで、ユーザが説明対象により注視しやすくなり、関心が対象に集まると述べている。Iwamoto et al.[9]は、店頭の商品紹介において、ヒューマノイドロボットが説明する条件と擬人化された商品が自己紹介する条件とで比較実験し、商品説明をユーザが記憶しやすくなると結論づけている。さらに、自己紹介条件がヒューマノイドロボット条件と比べて、質問に対し否定する被験者の割合が多かった点から、被験者の率直な感想が得られる可能性が示唆される。

しかし、これらのエージェント化の事例の多くは店頭での商品説明などを想定しており、擬人化されたエージェントが自己紹介する、いわば情報伝達に特化した研究である。そのため、ユーザテストのように人間の発話から情報を引き出す目的には対応していない。本研究はユーザの発話を支援するために擬人化を用いる点が特徴的である。

3. 提案手法

3.1 提案手法の目的と概要

プロトタイプをエージェント化することで、ユーザテスト中にユーザとプロトタイプの直接対話を可能にすることを目指す。本研究では、プロトタイプのエージェント化研究の探索的な試みとして、対話機能を付加しやすい GUI 上のアプリケーションを対象に手法を設計する。

プロトタイプはTA法に基づいたユーザの発話に対して、音声形式で応答・質問する。また、プロトタイプがユーザの話や操作を理解しようと努めていることをユーザに感じさせるため、画面上にマウスカーソルを追う目を付与する。以後、提案手法を構成する要素である、音声対話と身体的特徴として付与した目について、それぞれ説明する。

3.2 セリフの選定・音声合成

ユーザに対し、プロトタイプそのものが発話していると感じさせることを念頭に手法を設計した。まず、エージェント化したプロトタイプが話すセリフを用意するため、定型文を Coefont[10]で音声合成した。定型文の内容は表 1 に示すように、Boren et al.[4]の議論に基づき、ユーザへの教示から始まり、TA 法によるユーザの発話に対する相槌や、

表 1 実験に用いた定型文の内容

場面	セリフ
教示・自己紹介	1. 私はペイントアプリの（アプリ名）です。よろしくお願いします！
	2. この目であなたの操作をしています。
	3. 私を使ってみて、あなたの感想を教えてください。
相槌	4. はい。
	5. なるほど。
	6. そうですね。
発話を促す質問	7. 今、何をしているところですか。
	8. 何か悩んでいますか？
	9. どこで迷っていますか？
	10. 何かお探しいですか？
意図を問う質問	11. それって何のことですか？
	12. それは何のために使うものですか？

発話が止まった際に促す意図の内容に絞って生成した。音声はシステムらしさを考慮し、Amazon Alexa を参考に女性ナレーターの声を選択した。ユーザの発話内容に応じた定型文を再生し、ユーザとプロトタイプ間の対話を擬似的に実現する。実験中は Wizard of Oz 法（以下 WOZ 法）に基づいて実験者がシステムを装い、ユーザからは見えない PC の裏面に Bluetooth スピーカを配置し、ユーザの背後にいる実験者が、被験者に気づかれないように発話に合わせて、音声を選択・再生した。

3.3 身体的特徴の付与

また、プロトタイプに関心を持たせるために身体的特徴を付与する。本手法では WinEyes[11]を採用し、目をプロトタイプ上の画面上に配置した。ユーザが操作するマウスの動作に、WinEyes の視線が追従するようになっている。実験でテスト対象としたアプリケーションに、WinEyes を配置したスクリーンショットを図 1 に示す。

なお、これらの手法は、将来的にシステムとして自動化することを見据えたものである。そのため発話のタイミングは、秒数で決められている。具体的には、ユーザが 2 秒沈黙した場合、発話完了したとみなして相槌を打ち、10 秒沈黙した場合には現在の状況や悩んでいることを尋ねる。

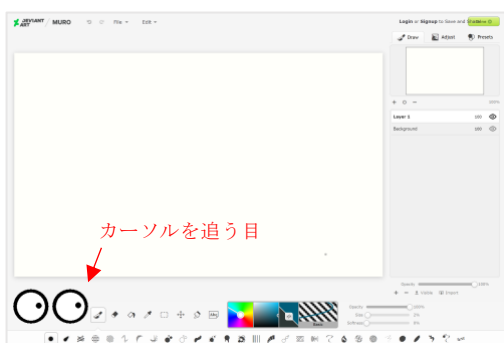


図 1 提案手法の画面イメージ

4. 実験

エージェント化プロトタイプの導入による発話や印象の変化を検証するために実施した実験について説明する。

4.1 実験計画

今回の実験は各条件での印象の違いをユーザ役の被験者からインタビューで聞き出すために、被験者内実験として実施した。また順序効果を避けるため、被験者ごとに各条件の実施する順番を入れ替え、それぞれの順番ごとに同じ人数になるように調整した。

また、ユーザテスト対象にブラウザ上で動作するペイントソフトを用いた。同じ操作感に慣れてしまうと、被験者が使いにくさやわかりにくさを感じても言及しない恐れがあると考え、図 2 に示す通り、muro[12]と 8bit paint[13]の 2 種類を用意した。被験者は同難易度の 2 つのタスクに、それぞれ異なるインターフェースで取り組むことになる。

被験者にはペイントソフトのユーザテストを行うと説明し、20 分間を目安に、図 3 の 2 種類のイラストを条件別に用意し、できるだけ多くの機能を試しながら、イラストを正確に真似て書くよう教示した。また、被験者には、被験者の描く絵を評価するのではなく、ペイントソフトの評価が目的であると実験前に伝えている。

4.2 測定内容

本実験では、タスク中の画面録画と同時に発話を録音し、CLOVA Note[14]でテキスト化した。発話を定量的に扱うた

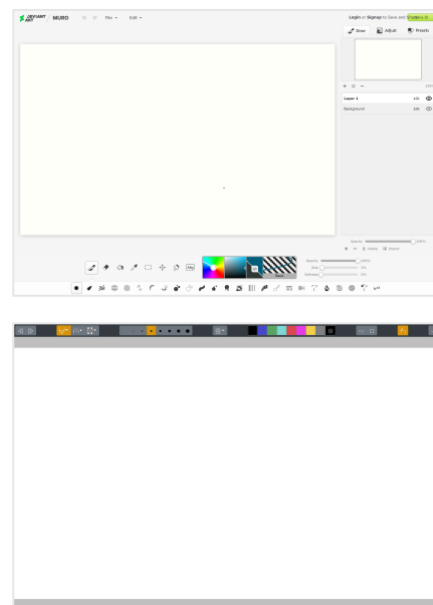


図 2 使用したペイントソフト（上：muro，下：8bit paint）

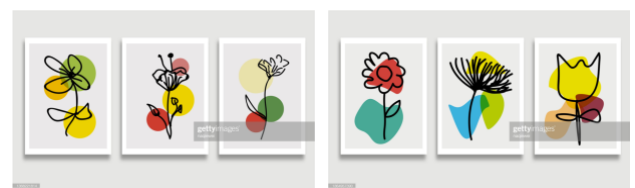


図 3 課題にしたイラスト

めに、海保ら[15]の見解を参考に文末を発話単位の終了と定め、言い淀みで切れ目が不明瞭な場合は2秒以上の間隔を発話終了とみなし、1分あたりの発話数（以下発話数）を集計した。

4.3 予備的実験

エージェント化プロトタイプをユーザテストに用いることは、TA法のセッション中、対話相手が人間からプロトタイプ本体に移行することを意味する。プロトタイプと対話することで、被験者の発話が増加すれば、エージェントのプロトタイプ化は被験者の発話を妨げることなく、ユーザテストを実施できるといえるだろう。しかし被験者の発話が増加する場合、発話を促進するための提案手法は逆効果になってしまっている可能性がある。そこで、エージェント化されたプロトタイプをTA法中に導入することによる、発話数への影響を分析するために、開発者が同席しない一般的なTA法条件と提案手法を用いたAgent条件の比較実験を行った。

TA法条件は予備的実験の統制条件であり、被験者は何の反応も返ってこない状態で、発話しながら操作する。Agent条件は、提案手法に基づき、被験者の発話に合わせて実験者がWOZ法で反応することで対話する。

被験者は著者らの大学の大学院生2名であり、対象のペイントソフトの利用経験がないことを確認した上で、実験に参加してもらった。なお、予備的実験に参加した2名は後述する本実験には参加していない。

この予備的実験の結果、得られた被験者別の発話数を表2に示す。この予備的実験を通して、発話数が増加し、エージェント化したプロトタイプが被験者の発話を阻害しないことを確認した。また、実験後のインタビューでは、「相槌を打ってくれる分だけ多く話したと思う（被験者A）」といった被験者からの感想も得られた。

4.4 本実験

本実験では、開発者が横につくHuman条件を統制条件とし、提案手法であるAgent条件との比較実験を行った。

Human条件では、開発者と名乗る人間が被験者の隣に座り、被験者の発話に反応することで対話する。その際、開発者の発話内容やタイミングは提案手法と同じ条件に揃えている。開発者役は著者らの大学の大学院生1名が行い、全被験者に対して同一の人間がついて実験した。Agent条件では、予備的実験同様、被験者の発話に合わせて実験者がWOZ法で反応することで対話した。実験室内の状況を図4に示す。被験者は著者らの大学の大学院生4名であり、予備的実験と同様、対象のペイントソフトの利用経験がな

表2 予備的実験の1分あたりの発話数

	被験者 A	被験者 B
TA法条件	1.39	2.60
Agent条件	2.96	3.22



図4 本実験中の実験室内

(被：被験者，実：実験者，開：開発者)

いことを確認の上で、実験に参加してもらった。

予備的実験から本実験へ移行するにあたり、発話数に加え、Norman[16]の提唱する行為の7段階に問題が生じている場面を著者の判断で集計し、1分あたりのユーザビリティ上への言及数（以下言及数）として記録している。

行為の7段階とは、ユーザの行為を①ゴールの形成、②行為のプラン、③行為の詳細化、④行為の実行、⑤外界の状況の知覚、⑥外界の状況の解釈、⑦結果の評価の7段階に分類する理論であり、主に②③④が実行過程、⑤⑥⑦が評価過程とされる。本実験では、画面上のどこを操作すべきかわからない場合や、予想外の出力結果が生じた場合に実行・評価に支障をきたしていると筆者が判断し、その場面での被験者の発話を言及数として集計した。

また、Human/Agentの各条件のタスク後に、アンケートによる印象評価を行った。アンケートでは、「話しやすさ」と「関係性」の二つの観点で調査した。「話しやすさ」は、率直な発言ができたかどうか（以下率直性）を表す「考えたことや感じたことを素直に話せた」をはじめ、「開発者の機嫌が悪くなりそうなことも話した」などの5項目からなる。「関係性」はユーザビリティ上の問題発生時における責任の所在を問う「うまく操作できないときでもアプリ側の問題だと思えた」などの5項目を用意した。いずれの項目でも、被験者には7段階評価（1:全くそう思わない～7:非常にそう思う）で回答してもらった。加えて、実験終了後に各条件のアンケート結果を見返しながらインタビューを行い、それぞれの印象とHuman/Agentの2条件の回答の差の理由について尋ねた。

5. 実験結果と考察

実験風景を図5に示す。右のAgent条件では、画面左下に目を配置し、PCの裏に隠したスピーカから音声を再生することでプロトタイプとの対話を擬似的に再現した。

実験では、図6の画面録画および印象評価用アンケート、インタビュー、発話内容などのデータが得られた。まず発話数・言及数やアンケート結果等の定量分析を行い、その後、発話内容に基づいた定性分析を行った。

5.1 アンケート・インタビューに基づく印象評価の分析

印象評価用のアンケート結果とインタビューの回答に

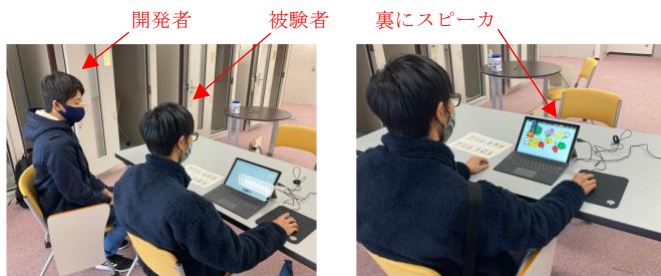


図5 実験風景 (左: Human 条件, 右: Agent 条件)

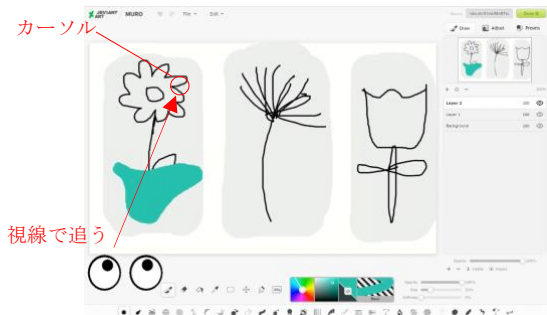


図6 録画された実験中の操作状況 (Agent 条件)

基づいて分析・考察を行う。

5.1.1 ユーザテスト中の「話しやすさ」の評価

「話しやすさ」の評価は、一貫した傾向が見られず、個人差が大きかった。ただし、率直性を表す「考えたことや感じたことを素直に話せた」では、図7のように、被験者 A, D の4人中2人の被験者が Agent 条件を Human 条件より高く評価した。インタビューでは Human 条件について、「悪いところを考えないようにしていた (被験者 C)」、「人間の方が具体的な返答をしてくれるのではと思っていた (被験者 A)」、「こうしたら伝わるかなと考えていた (被験者 B)」という感想が得られ、どのような内容を発話するか考えながら作業していたことが示唆される。一方で Agent 条件では、「同じことしか返さないのわからないだろうなと思って話していた (被験者 B)」、「色々な意味で何を言ってもいいと思った (被験者 C)」との発言があったことから、被験者は何を言ってもいいと感じていたと考えられる。しかし、これには心理的障壁の軽減以上に、反応の種類の少なさが影響していると考えられる。

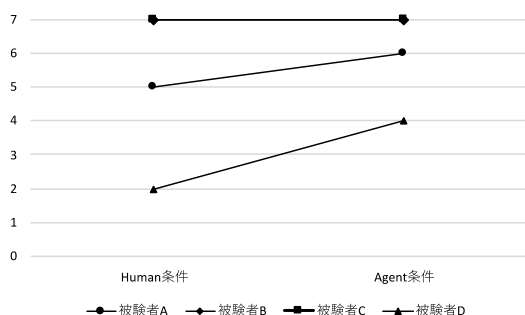


図7 「考えたことや感じたことを素直に話せた」の被験者別評価

5.1.2 被験者との関係性の評価

被験者との「関係性」を比較すると、「話しやすさ」同様、各項目の評価は被験者によって分かれた。しかし、「うまく操作できないときでもアプリ側の問題だと思えた」の項目は、図8の通り、4人中3人の被験者が Agent 条件を Human 条件と比べて高く評価した。これより、被験者がイメージ通りに操作できなかった際に、Agent 条件では Human 条件に比べて、ペイントソフト側にその責任があるとより感じていたことがわかった。インタビューでも、Agent 条件で扱った方のソフトを「アプリの表記に問題があった (被験者 B)」、「必要なものが大体なく、能力の問題でない (被験者 C)」と評価していた。そして、被験者 B と C は Agent 条件下で異なるソフトを操作している。よって、提案手法では、予想外の出力結果になった際、ユーザはその原因を自らの操作ではなく、ペイントソフトのインターフェースに求めるようになることが考えられる。

5.2 発話数・言及数に基づく分析

テキスト化した発話データを元に集計・分析を行った。発話数と言及数は、予備実験同様 4.2 節に示した基準に基づいて算出した。その結果、表3と表4に示すように、発話数と言及数のいずれも全ての被験者で増加がみられた。

この結果より、プロトタイプのエージェント化によって、被験者の発話が増加し、それに伴い、ユーザビリティ上の問題点に対する言及も増加したことが考えられる。ただし、Agent 条件は Human 条件に比べて相槌の頻度が増えたため、これに反応して発話が増加した可能性もある。また、相槌増加の原因として、Agent 条件で WOZ 法を利用したことが挙げられる。教示があるとはいえ、Human 条件ではどうし

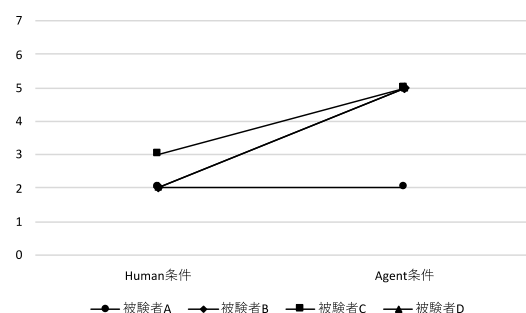


図8 「うまく操作できないときでもアプリ側の問題だと思えた」の被験者別評価

表3 各被験者の1分あたりの発話数

	被験者 A	被験者 B	被験者 C	被験者 D
Human 条件	2.40	3.09	3.89	0.20
Agent 条件	4.75	4.56	4.76	1.48

表4 各被験者の1分あたりの言及数

	被験者 A	被験者 B	被験者 C	被験者 D
Human 条件	0.59	0.39	0.89	0.10
Agent 条件	1.05	1.07	1.35	0.14

でも開発者役が被験者の様子を伺ってしまう様子がみられた。それに比べ、Agent 条件は被験者の機嫌を比較的伺うことなく相槌（音声再生）できるという性質を持つ。その影響を受けて相槌が増加したと推察できる。

ただし、この特性は裏を返せば、被験者の機嫌に関係なく反応するエージェントの相槌が発話数を増加させようとすることもできる。将来的に本手法を自動化した際に有用性が期待できる。

5.3 発話内容に基づく定性分析

テキスト化した発話内容に基づき、各条件間の発話内容や表現の違いから、ケーススタディ的に分析・考察を行う。

5.3.1 責任の所在に関する表現の違い

被験者 B は Human 条件で図 9 のように、「難しい」と 7 回発話していた。思い通りに操作できない原因がインタフェースにあると追及せず、自分の操作にもある程度求めていると筆者は推察した。また、Boren et al.[4]は、開発者の介入によって、ユーザが誤操作を隠す恐れがあると指摘している。そのため、被験者は予想外の応答を誤操作であると感じ、開発者に気を遣って「難しい」という表現でそれを誤魔化していた可能性も考えられる。

それに対し、同じ被験者 B の Agent 条件では「難しい」とは 1 回も発話せず、図 10 のように予想外の出力に対しては、使いたい機能がないペイントソフトに問題があると判断し、そのことを指摘する様子が窺えた。このように Agent 条件では、Human 条件のように婉曲な表現を使ったり誤魔化したりすることなく、アプリケーション側の問題点であると指摘していることがわかる。

5.3.2 ユーザビリティへの言及・指摘の表現の違い

ユーザビリティへの言及・指摘において、Human 条件では図 11 のように頭の中で整理してから発話しているのに

- 被 被験者 07:43**
レイヤーを変えて、後ろから塗ってやれば。
あ、でも色をかぶらないのか。
なるほど、これちょっと難しいです。
- 被 被験者 23:10**
動かしたら描画してないのに出てきます。難しい。

図 9 「難しい」と発話した場面（被験者 B, Human 条件）

- 被 被験者 11:50**
うん、色の淡さとかを変えられたら 1 番いいんですけど、ちょっとなさそうなので
- A Agent 12:07**
はい
- 被 被験者 12:13**
ま、色んな配色をして終わりです

図 10 機能がないことを指摘する（被験者 B, Agent 条件）

- 被 被験者 03:38**
そのまま円という形状が出せるツールがないって
いうのは、ちょっと不便かな。
そうですね、縁は書けない、うまく。

図 11 状況整理して発話する（被験者 C, Human 条件）

- 被 被験者 01:03**
アイコンが多すぎる、何が何なのかわからん。
えっと、スモーク。

図 12 違和感をそのまま発話する（被験者 C, Agent 条件）
対し、Agent 条件では図 12 のように操作中の違和感をそのまま言葉にしている。筆者の推測ではあるが、Human 条件では考えを開発者に伝えることを意識する一方で、Agent 条件ではプロトタイプには考えは伝わらないと感じていたため、表現に違いが生じたと考えられる。

しかしながら、説明しようとする意識が生まれにくいことによる利点もある。ユーザが言語化できない課題に直面した際、Human 条件では説明を諦めてしまう場合でも Agent 条件では感じたままに発話するのではないだろうか。発話データから問題の箇所を特定するのが難しい一方で、その問題の存在を直感的な言及から発見できるかもしれない。ただし、これを活かすには、断片的な発話内容から課題のある機能を特定できるよう、具体的な状況把握が求められる。提案手法を現場で応用する際には、発話データ以外に画面録画や視線情報などを記録する必要がある。

- 被 被験者 12:23**
なんか変わるんですか。
ああああ、なるほど、

- 被 被験者 12:32**
これは超面白いな。

図 13 機能を褒める場面（被験者 C, Human 条件）

- 被 被験者 14:41**
うん、結構用途が特定されそうな筆の形ばかりなので、あの、もうちょっとポピュラーのやつ増やしてもいいかもしれない。

- 開 開発者 15:04**
なるほど、

- 被 被験者 15:13**
うん。使い方は難しい。
速度によって、大きさが変わる、

- 被 被験者 15:37**
使い勝手より面白みを追求してるような筆が多いです。

図 14 指摘後フォローする場面（被験者 B, Human 条件）

5.3.3 Human 条件特有のアプリを褒める発言

Human 条件では「ブラシの種類が多いのは面白い (被験者 B)」、図 13 の「これは超面白いな (被験者 C)」等、Agent 条件ではなかったペイントソフトを褒める発言があった。また被験者 B は Human 条件で図 14 に示す通り、指摘した後に機能に対してポジティブな発言をしており、隣にいる開発者をフォローする意図が示唆される。

5.3.4 各条件の発話内容のまとめ

5.3.1 節から 5.3.3 節で述べた内容も含めた、各条件の発話の特徴を表 5 に示すとともに、ユーザビリティ上の問題発生から発話までの一連の過程を、発話内容と 5.2 節で述べた発話数・言及数の変化に基づき推測する。

まず「問題発生時の責任の所在」から、Human 条件では自らの操作にも責任を感じることや、開発者に気を遣うことで指摘せずに済ませるのに対して、Agent 条件では機能に問題があると判断し、指摘しようとする。次に、「ユーザビリティへの言及・指摘」のための言語化から、Human 条件では、何がどう不便なのか明確にしてから発話していた。一方で Agent 条件では、操作中の違和感をそのまま発話していた。Human 条件では状況整理できずに説明を諦めてしまうような抽象的な違和感も、Agent 条件では感じたままに指摘していたと推察できる。以上を踏まえ、発話までのプロセスを推測し、フロー図にしたものを図 15 に示す。

5.4 提案手法と各条件の使い分けの提案

本実験の結果から提案手法の有用性を改めて述べる。Agent 条件の特徴である、問題の責任がプロダクト側にあ

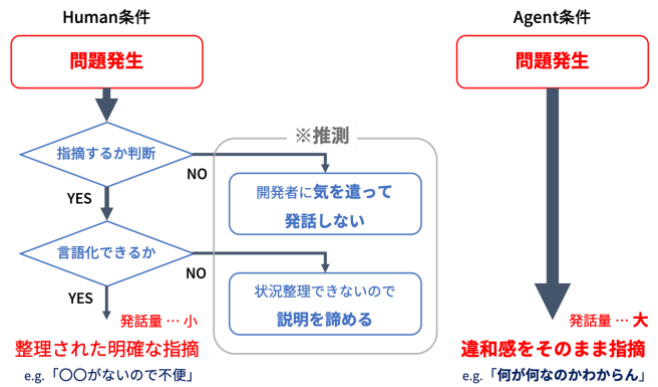


図 15 問題発生から発話までのプロセスの概要
(矢印の太さは発話量を表す)

るとユーザに感じさせる点は、指摘をより多く得たい場合に有効と考えられる。また、操作中の違和感を無理に状況整理せずに率直に発話する点は、使用中の感情推移を考察することに役立つと考えられる。

よって、多少抽象的になったとしても直感的な感想を得たい場面で、エージェント化プロトタイプは特に効果を発揮すると考えられる。ユーザが操作中ストレスを感じる箇所を細かく洗い出したい場合には、違和感を逃さずに発話することから、見落とすことなく改善点を発見できる。

一方で、Human 条件にもメリットは存在する。特に、発話の絶対量が減少してでもユーザの見解が欲しい場合には、対話内容や発話のタイミング等の精度にもよるが、人間の開発者が介入した方がより詳細な発言が得られると考えられる。機能の具体的なアドバイスを引き出し、ユーザと改善の可能性を議論する際は、開発者の介入の下で実施する

表 5 各条件の発話の特徴

	Human 条件	Agent 条件
全体的な口調	<ul style="list-style-type: none"> 常にです・ます調の丁寧語 	<ul style="list-style-type: none"> 最初は丁寧語だが、慣れてくると同時に言い放つような口調に
操作の説明	<ul style="list-style-type: none"> 答えを求めて開発者に質問する 自らの考えを開発者に確認しようとする 開発者に対して伝えることを意識し、より丁寧に説明 	<ul style="list-style-type: none"> 操作内容や、それに付随する印象・感想をそのまま発話 プロトタイプに返答はするが、考えが伝わるとは思っていない
問題発生時の責任の所在	<ul style="list-style-type: none"> 思い通りの結果にならないのは、自らの操作にも責任があると判断 	<ul style="list-style-type: none"> 思い通りの結果にならないのは、プロダクト側に使いたい機能がないためと判断し指摘
ユーザビリティへの言及・指摘	<ul style="list-style-type: none"> 状況や思考を整理してから発話 何がどう不便なのか明確にし、開発者に伝わるように工夫している 	<ul style="list-style-type: none"> 操作中に感じた違和感をそのまま発話 印象や感覚を中心に言及 被験者が上手く説明できない問題も発話を諦めることなく、感じたままに発話している
その他特徴	<ul style="list-style-type: none"> 機能を褒める発言 改善点を意見したあとにポジティブな発言でフォローする 	

ことが望ましい。

6. おわりに

6.1 本研究のまとめ

本研究では、ユーザテストでの TA 法にエージェント化プロトタイプを用いることで、ユーザの発話やユーザビリティへの言及を増加させ、率直な感想を取得することを目的とした。実験の結果、印象評価は個人差の大きい結果となったが、全体の発話とユーザビリティへの言及は増加が確認された。また、Agent 条件の方がより操作中に感じた違和感をそのまま発話していたことから、より率直な感想が得られたといえる。ただし、Human 条件にも、操作の説明をより丁寧に行う場面や、アプリケーションの機能を褒めながらアドバイスする様子が見られたため、今後は、実際の開発現場に応用する際はそれぞれを使い分けることが肝要である。

6.2 今後の展望

本研究は自動化を最終目標としているものの、まだ初期段階にあり、最適なエージェント化手法について十分に検討しきれていない。今後は、直感に基づく感想だけでなく、熟考された意見も開発者の介入と同じ、もしくはそれ以上に引き出したい。そのためには、返答の増加や頷き・瞬きなど、ユーザの発話に反応する動作をエージェントに追加することが必要となる。これらの改良を通して緊張感の軽減、ひいてはユーザに「自分の話を聞いてくれている」と思わせることを目指す。

また、取得したいユーザの反応に応じて、どのような印象を与えると効果的かを明らかにする必要がある。蔵田らの Talking-Ally に代表される弱いロボット[17]のような親近感の持てる設計で、ユーザが辛辣な意見を発話しやすくするなど、エージェントの性格や態度についても同様に模索しながら自動化を進めていく。

謝辞 実験にご協力いただいた協力者の皆様に厚くお礼申し上げます。

参考文献

- [1] 松蘭美帆, 草野孔希: はじめての UX リサーチ ユーザーとともに価値あるサービスを作り続けるために, 翔泳社, 2021
- [2] 木浦幹雄: デザインリサーチの教科書, ビー・エヌ・エヌ新社, 2020.
- [3] Priede, C., and Stephen, F.: Comparing results from different styles of cognitive interviewing: 'verbal probing' vs. 'thinking aloud', *International Journal of Social Research Methodology* 14.4, pp. 271-287, 2011.
- [4] Boren, T., and Judith, R.: Thinking aloud: Reconciling theory and practice, *IEEE transactions on professional communication* 43.3, pp. 261-278, 2000.
- [5] Olmsted-Hawala, L. et al.: Think-aloud protocols: a comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability, *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 2381-2390, 2010.
- [6] Krahmer, E., and Nicole, U.: Thinking about thinking aloud: A comparison of two verbal protocols for usability testing, *IEEE transactions on professional communication* 47.2, pp. 105-117, 2004.
- [7] DiSalvo, C., and Francine, G.: From seduction to fulfillment: the use of anthropomorphic form in design, *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces*, pp. 67-72, 2003.
- [8] 大澤博隆, 大村廉, 今井倫太: 直接擬人化手法を用いた機器からの情報提示の評価, *ヒューマンインターフェース学会論文誌*, Vol.10, No.3, pp.305-314, 2008.
- [9] Iwamoto, T. et al.: The effectiveness of self-recommending agents in advancing purchase behavior steps in retail marketing, *Proceedings of the 9th International Conference on Human-Agent Interaction*, pp. 209-217, 2021.
- [10] Coefont.: <https://coefont.cloud>
- [11] WinEyes.: https://ja.osdn.net/projects/sfnet_wineyes/
- [12] muro.: <https://www.deviantart.com/muro/>
- [13] 8bit paint.: <https://minordaimyo.net/8bitpaintweb/>
- [14] CLOVA Note.: <https://clovanote.line.me>
- [15] 海保博之, 原田悦子: プロトコル分析入門—発話データから何を読むか, 新曜社, 1993.
- [16] Norman, D.: *The Psychology of Everyday Things*, Basic Books, 1988.
- [17] 蔵田洋平, 松下仁美, 小田原雄紀, 大島直樹, DeSilva, P., 岡田美智男: Talking-Ally: 聞き手性をリソースとする発話生成システムの実現にむけて, *ヒューマンインターフェース学会論文誌*, Vol.17, No.2, pp.159-170, 2015.