

音声ガイドエージェントの環境適応と ユーザーによるエンゲージメントが信頼に与える影響

The Impact of Environmental Adaptation of Voice Guide Agents and User Engagement on Trust in Real Environments

齋藤 真里*¹
Mari Saito

山田 誠二*²*¹
Seiji Yamada

*¹総合研究大学院大学

The Graduate University for Advanced Studies

*²国立情報学研究所

National Institute of Informatics

This paper describes how user engagement affects their trust in audio guide agents to adapt to their environment. Audio guide agents are expected to be used more in the future because they can perform without interrupting the user's actions, even when the user is in action. Because agents adapt to their environment and users work collaboratively with each other, trust in the agent is important. Therefore, we examined whether adaptation itself affects trust and whether intentional user operations as a means of user engagement affect trust. A participant experiment was conducted with two independent variables, the degree of adaptation to ambient noise and the availability of user operation, and the subjective evaluation of the impression of the agent as the dependent variable. The result showed that adaptation of the agent could increase trust in the agent, and trust was also increased by intentional user operations regardless of the agent's adaptation level. In systems where users and agents work collaboratively, the results suggest that intentionally increasing user engagement may increase trust in the agent.

1. はじめに

これまで音声ガイドサービスは、主に美術館や博物館で用いられてきたが、近年、位置情報の利用なども進み、屋外の観光などにも拡大されてきている。音声ガイドは、移動しながら利用することができるため、今後、より多様な場面での活用が期待される。屋外でのガイドを考えた場合、ヘッドホンのノイズキャンセリング機能により、屋外における音声聴取の利便性は高まったが、ガイドの音声と並行してアナウンスなど外部音も聞きたいことがある。現状、外部の音を聞きたいときに、ユーザーがノイズキャンセリング機能を無効にする操作をすることはできるが、ユーザーが常に外部の音とガイドの音声の両方に注意しながらシステムを操作するのは煩雑であり、音声ガイドに対して、エージェントによる適応的な動作が期待される。しかし、ユーザーの状況や好みの多様性もあり、完全な適応は困難であることから、当面はユーザーもある程度システムを操作することを想定する必要がある。

これまでのHAI研究では、ロボットやキャラクターエージェントなどの外見を持つエージェントとのインタラクションが多く研究されてきた。一方で、音声モダリティのみのエージェントを対象としたインタラクション研究は必ずしも多くない。

本研究では、環境音を適応的に制御する音声エージェントを想定し、エージェントと人の両方が協調するような状況で、エージェントの適応動作と人の操作がエージェントの印象に与える影響を検討する。

2. 関連研究

音声ガイドエージェントの適応についての研究としては、鑑賞者の位置や状況に応じて解説を行う音声ガイドエージェントについての精度や有効性が議論されている([Chang 14][Smørdal 14]など)。また、音声ガイドに限らず、音声提示を周囲の音環境に適応させる研究もある。周囲の環境に応じて能動的に音量を調整するなど、騒音に適応するシステムが提案されており、その効果が示されている[Ambeth Kumar 20]。

連絡先: 齋藤 真里, 総合研究大学院大学, maris@nii.ac.jp

一方、エージェントが騒音に適応する方法が利用者の音の好みに影響するという研究もあり、エージェントがどのように音に適応するかによってエージェント自体の印象が変わる可能性がある[Walton 18]。

美術館での鑑賞者に関する研究では、建物や展示レイアウト、他の鑑賞者などの環境に影響されることが示されている[Castro 16][Choi 99]。音声案内自体の提示方法が、訪問者の鑑賞に影響を与えることもわかっている[Leder 04]。このように、ガイドの提示と鑑賞者の行動は相互に影響を与える。

周囲から聞こえる音のすべてがノイズではなく、聞くべき音や聞きたい音が含まれていると考えられるため、エージェントが適切に音を制御することは有用であると考えられる一方で、エージェントの適応行動自体が、エージェントに対する印象に影響を与える可能性もある。

信頼に関する研究は、例えば、レコメンデーションシステムの研究において、個人への適応と信頼の関係が扱われ、信頼が意思決定に影響を与えることが示されている[Komiak 06]。信頼に関する研究では、外見や声の特徴が信頼に及ぼす影響なども見出されており、音声エージェントに関しても、その印象が信頼に影響を及ぼす可能性が高い。

さらに、AIとユーザーとのインタラクションにおけるアルゴリズム忌避の研究[Dietvorst 18]では、AIの結果に対して、ユーザーが修正できる機能を提供した場合、システムへの信頼が高まることもわかっている。ユーザーのエンゲージメントによってシステムの印象は影響を受ける可能性がある。

3. 実験方法

3.1 実験目的と実験計画

音声ガイド体験において、周囲からユーザーにとって不要なノイズだけが聞こえる場合には、常にノイズをキャンセルすればよい。しかし、通行する観光客にむけたアナウンスなど、周囲からの音であってもユーザーにとって必要な音もある。このような状況を対象に、周囲からのノイズをキャンセルし、必要なアナウンスは聞きやすくするような適応をする音声ガイドエージェントを想定する。エージェントの適応度がより高ければ、エー

	ユーザー操作	操作なし	操作あり
エージェント適応			
適応度：低			
適応度：高			

図 1: 2 要因の条件配置

エージェントへの信頼を高める可能性がある。適応が不十分である場合、ユーザー自身が周囲音を制御したいという要求も想定できるが、こうした操作の有無自体もエージェントの信頼に影響を及ぼす可能性がある。そこで、以下の二つの仮説を検証する。

H1: 音声ガイドエージェントの適応は、ユーザーのエージェントへの信頼に影響を及ぼす。

H2: ユーザーのエンゲージメントは、ユーザーのエージェントへの信頼に影響を及ぼす。

実験配置は、エージェントの環境適応とユーザーのエンゲージメントの 2 要因で、環境適応要因は、適応度の高低 2 水準、エンゲージメント要因は、エンゲージメントあり・なしの水準の 4 条件であった (表 1)。実験は、実験参加者に対して、4 条件すべてを実施する被験者内要因で行われた。

環境適応要因の適応度が高い条件では、エージェントはノイズに対して即座にキャンセルする動作を行い、低い条件では、ノイズに対してなにも制御しなかった。エンゲージメント要因のうち、エンゲージメントあり条件では、ユーザー自身が騒音に対する音量操作を行うことができた。エンゲージメントなし条件では、操作をすることができなかった。

3.2 実験方法

G*Power 3.1.9.7. による事前分析の結果に対し、脱落を加味してサンプルサイズが決定された。実験には、20 代から 50 代の成人男女 30 名が参加した。

実験参加者は、インフォームドコンセントの後、実験用のスマートフォンアプリケーションの操作について説明された。参加者は、ユーザー操作アリの条件で、この実験用に作成されたスマートフォンアプリケーションのスライダーを用いて環境音の音量を調節した。4つの条件は、各参加者ごとにランダムな順で実施され、1 条件ごとに印象評価が実施された (図 3)。ユーザーが行った印象評価の項目は、表 1 の項目であった。さらに参加者は、ガイド内容とアナウンス内容を聞き取れているかについて、選択肢から該当する内容を選択する設問にも回答した。

3.3 提示刺激

音声ガイドは、東京・浅草エリアから 4 つの観光スポットを選び、公式サイト解説を参考に、本実験用のオリジナルガイドテキストを作成した。ガイドテキストは、すべての観光スポットの解説が約 30 秒になるように編集した。音声は、Google Cloud TTS^{*1} (setting:ja-JP-Neural2-B, default pitch, for earphones, male) を用いて作成された。すべて 1 文で終了し 10 秒程度になるように調整されたアナウンスが 4 種類作成され、開始後 10 秒のタイミングで提示された。環境からのノイズ音は、Otolologic.jp² から提供されている音源 Otolologic.jp^{*2} を使用し、ガイド開始後 5 秒とアナウンス終了直後の 2 回、4 条件

*1 <https://cloud.google.com/text-to-speech?hl=ja>

*2 <https://otolologic.jp/>, City_Ambi-Street01-1inp3

項目	設問
・エージェントの信頼	・ガイドエージェントは信頼できた。
・エージェントの知性	・ガイドエージェントは知的に感じられた。
・エージェントの人らしさ	・ガイドエージェントは人のようにふるまった。
・体験の満足度	・ガイド体験に満足した。
・エージェントの環境適応	・ガイドエージェントは環境を理解していた。
・エージェントのユーザー適応	・ガイドエージェントはあなたの状況を理解していた。
・ガイドの聞き取りやすさ	・ガイドは聞き取りやすかった。
・アナウンスの聞き取りやすさ	・アナウンスは聞き取りやすかった。
・操作しやすさ	・アプリでの音量操作はしやすかった。

表 1: ガイド体験に関する印象評価の設問。

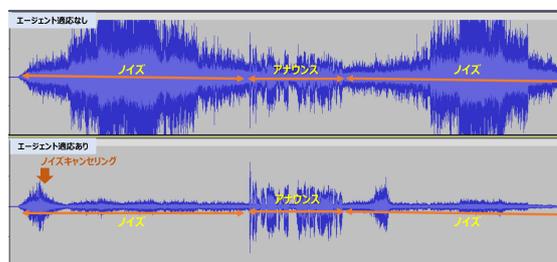


図 2: 提示される音声ガイドの音量の状態。すべての条件で環境のノイズの発生タイミングは同じ。環境のノイズとアナウンス音は重ならないタイミングで再生された。

すべて同じタイミングで発生、終了した。ノイズは徐々に増幅し、終了時は徐々に減衰した (図 2)。

4. 実験結果

参加者の印象評価の各項目について、2 要因の分散分析を行った。M は平均値、SD は標準偏差、p は有意水準である。

信頼に関する評価では、適応度要因 ($F(1,29)=35.208, p < .001, \eta_p^2=0.575$)、ユーザー操作要因 ($F(1,29)=7.495, p = .009, \eta_p^2=0.215$) ともに主効果は有意で、交互作用は有意ではなかった ($F(1,29)=3.915, p = .057, \eta_p^2 = .119$) (図 4, 表 2)。

エージェントの知性の印象評価では、適応度要因 ($F(1,29)=38.761, p < .001, \eta_p^2=0.572$)、ユーザー操作要因 ($F(1,29)=5.093, p = .032, \eta_p^2=0.149$) で有意差が認められたが、交互作用有意により ($F(1,29)=7.909, p=.009, \eta_p^2=0.214$)、下位検定した結果、適応度が低い場合に有意に評価は高くなった ($F(1,29)=10.989, p=.002, \eta_p^2=0.275$)。同様に、エージェントの人らしさの評価でも、適応度要因 ($F(1,29)=9.839, p = .004, \eta_p^2=0.253$)、ユーザー操作要因 ($F(1,29)=4.399, p = .045, \eta_p^2=0.132$)、交互作用有意 ($F(1,29)=5.449, p = .027, \eta_p^2 = 0.158$) により、下位検定した結果、適応度が低い場合に有意に評価は高くなった ($F(1,29)=9.508, p=.004, \eta_p^2=0.247$)。

満足度に関しても、適応度要因 ($F(1,29)=172.957, p < .001$,

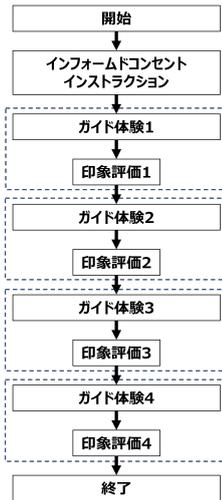


図 3: 実験手順

$\eta_p^2=0.856$), ユーザー操作要因 ($F(1,29)=30.407, p < .001, \eta_p^2=0.512$) ともに有意であったが, 交互作用が有意であったため ($F(1,29)=59.596, p < .001, \eta_p^2=0.673$) 下位検定を行った結果, 適応度が低い場合に有意に評価は高くなった ($F(1,29)=63.865, p < .001, \eta_p^2=0.688$) (図 5, 表 3) .

エージェントの環境適応, ユーザーへの適応についても同様に, 2 要因とも有意差があり, 適応度が低い場合に有意に評価は高いという結果が得られた (環境への適応の評価: ユーザー操作なし 適応度低 $M=1.20, SD=0.551$; 適応度高 $M=3.6, SD=1.303$ ユーザー操作あり 適応度低 $M=2.37, SD=1.066$ 適応度高 $M=3.47, SD=1.008$, 適応度要因 $F(1,29)=75.318, p < .001, \eta_p^2=0.722$), ユーザー操作要因 $F(1,29)=8.223, p=.008, \eta_p^2=0.221$ 交互作用有意により下位検定した結果, 適応度低の場合 ($F(1,29)=34.659, p<.001, \eta_p^2=0.544$) ユーザー状態への適応の評価: ユーザー操作なし 適応度低 $M=1.20, SD=0.664$, 適応度高 $M=3.23, SD=1.331$; ユーザー操作あり 適応度低 $M=2.30, SD=0.988$; 適応度高 $M=2.93, SD=1.112$, 適応度要因 $F(1,29)=40.524, p < .001, \eta_p^2=0.583$), ユーザー操作要因 ($F(1,29)=5.025, p=.033, \eta_p^2=0.148$), 交互作用有意により下位検定した結果 適応度低の場合 ($F(1,29)=28.684, p<.001, \eta_p^2=0.497$).

これらの結果については, 表 4 にまとめた.

聞き取りやすさに関する評価に関して, ガイドの聞き取りやすさでは, 2 要因とも主効果が得られ, 下位検定の結果, 適応度が低い場合にユーザー操作の効果があることが示された (ガイドの聞きやすさ: ユーザー操作なし 適応度低 $M=1.27, SD=0.691$ 適応度高 $M=4.10, SD=1.213$ ユーザー操作なし 適応度低 $M=3.20, SD=1.186$ 適応度高 $M=4.23, SD=0.817$, 適応度要因 $F(1,29)=123.332, p<.001, \eta_p^2=0.810$ ユーザー操作要因 ($F(1,29)=23.538, p<.001, \eta_p^2=0.448$), 交互作用有意により下位検定した結果, 適応度低の場合, ユーザー操作の効果あり ($F(1,29)=50.916, p<.001, \eta_p^2=0.637$) アナウンスの聞きやすさに関しては, 適応度要因は主効果が有意であったが, ユーザー操作要因, 交互作用とも有意な差が得られなかった (ユーザー操作なし 適応度低 $M=3.67, SD=1.348$ 適応度高 $M=4.40, SD=0.770$ ユーザー操作あり 適応度低 $M=3.03, SD=1.273$ 適応度高 $M=4.43, SD=0.728$, 適応度要因 $F(1,29)=42.362, p<.001, \eta_p^2=0.594$), ユーザー操作要因

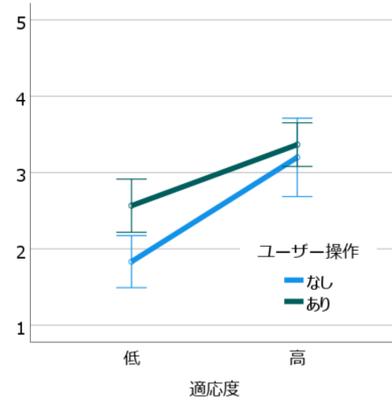


図 4: ガイドへの信頼の評価

	適応度: 低		適応度: 高	
	M	SD	M	SD
ユーザー操作なし	1.83	0.913	3.20	1.375
あり	2.57	0.935	3.37	0.765

表 2: Ratings of whether participants trusted agent. Adpt stands for Adaptation and OP for Operation.

($F(1,29)=2.628, p=.116, \eta_p^2=0.083$) また, ガイドの聞き取りに関しては概ね 100 %に近い正答率であったが, アナウンスについては, 50 %前後の正答率であった.

5. 考察

分析の結果, エージェントの適応が高い場合の方が低い場合よりも信頼が高まり, ユーザー操作がないより, ある方が信頼が高くなり, 2つの仮説は支持された. また, 「エージェントの知性」「エージェントの人らしさ」では, 相互作用が有意であり, 適応度が低い場合に, ユーザー操作があることにより, エージェント自体が知的で, 人らしく感じられることは, 興味深い結果である. エージェントのふるまいが同じであっても, ユーザー操作により, エージェント自体の印象が変わる. さらに, 満足度においても同様の結果が得られている. エージェントの適応性に関する印象評価では, 環境の理解, 参加者の状況の理解とともに, 「エージェントの知性」「エージェントの人らしさ」の結果と同様, 適応度が低い場合に, ユーザー操作の効果が有意であった. 信頼の評価では相互作用がなく, 適応性が高いほど信頼が高く, 利用者の操作が可能ほど信頼が高いという, 他の項目の評価とは異なる結果となった. 利用者の操作が可能であることは, エージェント自身の適応行動とは関係がないはずである. しかし, 参加者は自身の操作により, 適応性が低いほどエージェント自身を高く評価した. 本実験のように, エージェントの適応対象と利用者の操作対象が同じ場合, エージェントと利用者が協力して行う作業と考えられる. このような状況では, 利用者は評価対象としてエージェントを評価するだけでなく, 利用者自身を含めた一つのシステムとして利用者进行评估する可能性がある.

実験のインストラクションでは, ガイドの音声と, 外から聞こえる観光案内の音声の両方を聞くよう指示し, 提示においても両方の音声は重複していなかった. しかし, ガイドの聞き取りやすさとアナウンスの聞き取りやすさの評価には異なる傾向

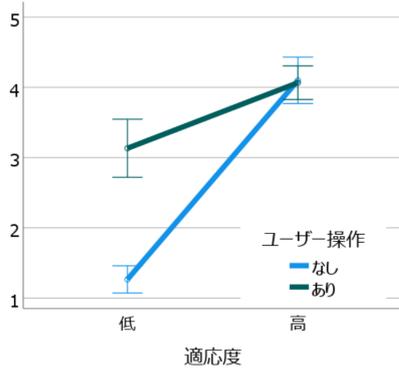


図 5: ガイド体験の満足度の評価

	適応度：低		適応度：高	
	M	SD	M	SD
ユーザー操作なし	1.27	0.521	4.10	0.885
あり	3.13	1.106	4.07	0.640

表 3: Ratings of satisfaction with experience. Adpt stands for Adaptation and OP for Operation.

が認められた。エージェントの適応度が低い場合でも、参加者がアプリケーションを操作する条件ではガイドは聞き取りやすいと評価された。しかし、アナウンスはユーザー操作の影響を受けなかった。正答率の観点でも、アナウンスの正答率は低く、参加者は主にガイドを聞いていたことがわかる。つまり、ガイドを聞きながら観光する場合、外部の音を聞き取ることができる状況だとしても、理解するのは難しく、人間はどちらかを無視する可能性がある。

6. まとめ

音声ガイドエージェントが環境に適応的に制御する状況を想定し、エージェントの適応度とユーザーのエンゲージメントがエージェントの印象に及ぼす影響を検証した。観光ガイドのコンテンツを作成し、観光中に周囲の音とアナウンスを聞きやすくするエージェントとユーザーが周囲音を調整できる環境を用意した。エージェントの適応度の高低とユーザー操作有無の2要因実験の結果、音声ガイドエージェントが環境に適応することで信頼が高まることが示された。さらに、エージェントの機能や性能が同じであっても、ユーザの操作がエージェントへの信頼を高めやすいこともわかった。したがって、エージェントへの信頼を高めるためには、意図的であっても、環境や利用者に適応させ、利用者が操作できるようにすることが有効であると考えられる。

参考文献

[Ambeth Kumar 20] Ambeth Kumar, V. D., Malathi, S., Kumar, A., M, P., and Veluvolu, K. C.: Active Volume Control in Smart Phones Based on User Activity and Ambient Noise, *Sensors*, Vol. 20, No. 15 (2020)

[Castro 16] Castro, Y., Botella, J., and Asensio, M.: Re-Paying Attention to Visitor Behavior: A Re-Analysis us-

	適応 要因	ユーザー 操作要因	交互 作用	単純 主効果
信頼	*	*		
知的	(*)	(*)	*	†
人らしさ	(*)	(*)	*	†
体験の満足度	(*)	(*)	*	†
環境への適応	(*)	(*)	*	†
ユーザーへの適応	(*)	(*)	*	†

表 4: 印象評価のまとめ。†は、適応度が低い場合に、ユーザーの操作があると高くなる項目。

ing Meta-Analytic Techniques, *The Spanish Journal of Psychology*, Vol. 19, p. E39 (2016)

[Chang 14] Chang, K.-E., Chang, C.-T., Hou, H.-T., Sung, Y.-T., Chao, H.-L., and Lee, C.-M.: Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum, *Computers & Education*, Vol. 71, pp. 185–197 (2014)

[Choi 99] Choi, Y. K.: The Morphology of Exploration and Encounter in Museum Layouts, *Environment and Planning B: Planning and Design*, Vol. 26, No. 2, pp. 241–250 (1999)

[Dietvorst 18] Dietvorst, B., Simmons, J., and Massey, C.: Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them, *Management Science*, Vol. 64, pp. 1155–1170 (2018)

[Komiak 06] Komiak, S. Y. X. and Benbasat, I.: The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents, *MIS Quarterly*, Vol. 30, No. 4, pp. 941–960 (2006)

[Leder 04] Leder, H., Belke, B., Oeberst, A., and Augustin, D.: A Model of Aesthetic Appreciation and Aesthetic Judgments, *British journal of psychology*, Vol. 95, No. 4, pp. 489–508 (2004)

[Smørðal 14] Smørðal, O., Stuedahl, D., and Sem, I.: Experimental zones: two cases of exploring frames of participation in a dialogic museum, *Digital Creativity*, Vol. 25, No. 3, pp. 224–232 (2014)

[Walton 18] Walton, T., Evans, M., Kirk, D., and Melchior, F.: Exploring object-based content adaptation for mobile audio, *Personal and Ubiquitous Computing*, Vol. 22, No. 4, pp. 707–720 (2018)