

# アバターの視線と口の動きが視聴者の記憶と印象に及ぼす影響 Effect of the gaze and mouth movements of an avatar on viewers' memory perception and impression

瀬島 章仁<sup>†</sup>, 安田 哲也<sup>‡</sup>, 小林 春美<sup>§\*</sup>  
Akihito Sejima, Tetsuya Yasuda, Harumi Kobayashi

<sup>†</sup>東京電機大学大学院, <sup>‡</sup>東京大学, <sup>§</sup>東京電機大学  
Graduate School of Tokyo Denki University, The University of Tokyo, Tokyo Denki University  
h-koba@mail.dendai.ac.jp

## 概要

本研究は、ニュース映像でアナウンサーを務めるアバターの口の動きと顔の向きが、視聴者に与える印象やニュース内容の認知に及ぼす効果を調べた。異なる4つの条件でアバターが映った刺激動画が提示され、理解度と好感度に関する質問紙調査を行った。結果、アバターの口の動きがある場合、参加者は動画をより理解したと考えやすかった。さらに、アバターの顔方向がニュース映像の方に動くと、参加者はアバターにより注意を向けていた。

キーワード：選択行動, アバター

## 1. 目的

アバターを使用した動画は YouTube などでも多くみられ、全国各地で人気コンテンツとなっている。ニュース動画は現代社会において情報伝達の主要な手段となり、その中で AI がニュースの読み上げに利用される例が増加している。しかし、AI によるニュースの読み上げのような純粋な音声情報だけでは、視聴者が視聴するためのモチベーションの維持・向上が問題となることが考えられる。アバターを導入して視覚的な情報を提供することにより、視聴者が情報を効果的に受容しやすくなる可能性がある。本研究では、YouTube でよく見られる情報発信動画を模し、アバターを利用して、その動画の印象等を調べた。

Aseri & Interrante (2021)は、仮想現実における3つの異なるタイプのアバター表現がもたらす視覚的および非言語の手がかりの影響を調べた。比較されたアバタータイプは、HMD とコントローラーのみでアバターの体は表示されない「No Avatar」、3D モデルを利用しアバターの体及びコントローラー等が表示される「Scanned Avatar」、およびクロマキー処理を利用し自身の体が表示される「Real Avatar」(コントローラー等も表示される)の3種類であった。なお、参加者は必ず HMD(Head

Mounted Display)を装着していた。タスクは、参加者同士(一方はモデレータ)と作業を行うものであり、その後、7件法により心的な項目が含まれる質問紙に答えた。その結果、Real Avatar は、他のアバター表現よりも、高い対人信頼(Interpersonal Trust)をもたらしたが、Scanned Avatar と No Avatar には違いがなかった。また、Scanned Avatar は No Avatar や Real Avatar よりも体の姿勢(posture)への気づきが少なかった。これに関しては、Scanned Avatar は表情(facial expression)に関する気づきが少なく、その結果姿勢をよく見ている可能性を挙げていた。共存度(Co presence)<sup>1</sup>に関しては、Real Avatar や Scanned Avatar は No Avatar よりも高かった。他にもメタバースにおけるアバターを調べている研究もあり、アバターを介したコミュニケーションの研究は仮想空間等で調べられていることが多い。

Mizuho et al. (2023)では、バーチャル・オムニバス講義と呼ばれる新しい遠隔講義形態が提案された。この講義は、多数でオムニバス講義を行うというわけではなく、1人の講師だが自身のアバターを変えることで複数の講師で疑似的にオムニバス講義を行うというものである。この研究では、異なる環境での学習が記憶力を向上させることが示されていることを基に、複数の講師から仮想的に学習することでも同様の効果があるか調べるものであった。90分間の講義を4つのセッションに分け、セッションごとに異なるアバターを使って教える「4アバター状態」と、すべてのセッションで同一のアバターを使用する「1アバター状態」を比較する実験を行った。結果、講義後の理解度テストの点数は、1アバターの状態よりも4アバターの状態の方が有意に高いことが示された。これは、複数のアバターで講義を行うことで、学生が講義に注意しやすくなる効果が得られる可能性を示唆した。

情報の伝達において非言語的な情報は重要な役割を

<sup>1</sup> Schroeder (2002)によれば、共存感とはユーザーが他の人間などの主体と一緒にいるという感覚のこ

ととされる。

担っている。なかでも人の視線は相手との会話促進などに影響を与えることが知られている(大坊,1998)。他者の瞬間的な視線移動は、その方向に対しての注意を促すことがわかっている(Driver et al.,1999)。また、顔が提示された場合に直視の方が逸視よりもその顔を再生しやすいことがわかっている(Mason et al.,2004)。学習効果に関しては、Pi et al. (2020)がビデオ教示場面を調べており、教材に視線を向けた方がアイコンタクト(直視)よりも学習効果が増加したという知見を示している。また、

近年、早朝のニュース番組でAIがニュースを伝えている。このニュースはアバター等の利用はなく、Mizuho et al.(2023)の観点を考慮すると、視聴者の注意継続が困難になる可能性を示唆する。本研究では、予備的に、アバターの助けが、参加者の注意を促すという仮定を基に実験を行った。ここでのアバターの助けとは、ナレーション時に音声だけではなく、アバターがあたかも伝えているというような助けである。具体的には、ニュース映像において、現場にいるアナウンサーが何か伝えている場合のみ、アバターがそのニュース映像を見て、ナレーション時には参加者の方向を向いているか、そうでないか(つまり常に参加者の方向を向いている)という2水準の視線・顔条件があった。加えて、アバターの口がともに動く、または動かない、という2水準の口の動き条件があった。

本研究の予測では、アバターの口の動きがある場合、視聴者はニュース動画の内容をより理解しやすくなり、アバターに対して好感を持つと予測される。口の動きがあることにより、視覚的な手がかりが増え、言葉の意味やニュース内容がより明確に伝わると考えられる。また、アバターの顔方向がニュース動画に向く場合は、視聴者は動画の内容により注目しやすくなると考えられる。顔方向の動きがあることで、視聴者の注意がニュース動画に惹きつけられ、動画の重要な情報に集中しやすくなると予測される。

本研究は、視線と口の動きが視聴者のアバターに対する印象やニュース内容の記憶に及ぼす効果について調べた。

## 2. 方法

### 2.1. 参加者

理工系大学に通う18歳から24歳の学生32人(男性27人、女性5人:平均年齢20.2歳)であった。

### 2.2. 条件

顔方向の動きと口の動きの条件をそれぞれ2水準で設けた。顔方向の動き条件は、ニュースを読み上げる際に、ニュース映像を見るように向ける、または、ニュース映像は見ずに参加者を見るようにまっすぐ前を向いている、というものであった。口の動き条件は、ニュースを読み上げる際に、口を動かす、または口を動かさず閉じている、というものであった。よって条件は、顔方向の向き条件(アバターの顔の向きが動く、アバターの顔の向きが動かない)と口の動き条件(口が動く、口が動かない)の2条件であった。

### 2.3. 刺激

アバターは、メッセージアプリの機能にあるミー文字(開発会社名:Apple社,iOS12から導入された機能)を使用して作成した。その後、CGで作成した人の上半身に、作成したアバターの顔を合成した。このアバターに対して、各条件から構成される動きを映像編集した。なお、これらの映像は、うなずくなどの動きを考慮し相互作用の要素が含まれていた。全ての刺激は、左側にアバターが配置され右側にニュースが配置された(図1)。ニュース動画は、4種類作成し、いずれも天気に関する人々へのインタビュー動画であった。条件ごとにアバターの動きが異なっていた。また、アバターの口の動き条件がある場合、ニュース動画の解説音声とアバターの口の動きが一致している状態を提示した。これは、視覚的な手がかりとして口の動きと聴覚情報との統合が、情報の理解度に寄与することができると考えられたからである。

図1 実験の一場面



### 2.4. 手続き

参加者は実験同意書が明記されたGoogle formのQRコードをスマホで読み取り回答した。同意のあった参加者に対し、実験参加用のQRコードを提示し、参加者はそのQRコードを読み取り実験に参加した。なお、参加者には必要に応じてイヤホンやヘッドフォンなどをつけて動画を再生するよう求めた。

参加者は自身でGoogle Form上の実験動画を流した

後に、その下にある5つの質問項目が含まれた質問紙を答えた。なお、この順で4回繰り返した。実験順序をランダム化することはなく、実験動画を見た順番は、視線・顔方向・口の動きがある刺激、口は動くが顔は動かない刺激、視線・顔方向は動くが口は動かない刺激、顔も口も動かない刺激であった。

質問内容は、アバターへの好感度、動画とアバターのどちらを見ていたか、参加者のニュースに関する理解度の自己評価、他にフィラーとして、動画の内容に関する簡単な2つの質問(例えば誰が雪かきをしていたか)であった。

質問紙調査は6段階のリッカートスケールで回答するものであった。設問は、アバターの好感度はどうだったか(1. とても悪い・6. とても良い)、参加者は動画とアバターのどちらを多く見たと思うか(1. 動画・6. アバター)、提示されたニュース動画に関しての参加者自身の理解度はどうだったか(1. 全く理解できなかった・6. とても理解できた)であった。実験の終わりに、「どの動画のアバターの印象が良かったか」や「実験自体の内容を理解できたか」に関する質問紙調査を実施した。

## 2.5. 分析方法

データの分析には統計ソフトウェア R, `ordinal`, `ggeffect` を利用した。得られたデータは順序尺度に該当するため、累積ロジット混合モデル(`clmm`: Cumulative

Link Mixed Model)を用いた。独立変数は、顔方向の動き(Direction)と口の動き(Face)であった。なお、累積ロジットモデルは最大モデルとし、モデルが収束したことが確認できたモデルを採用した。

要因に関しては、エフェクトコーディングを採用した。なお、ランダム効果に関しては、個人差においてランダム切片とランダム傾きを考慮した。なお、提示順序もランダム効果に含めた。例えば、好感度を調べたモデルは、`clmm(popularity ~ Face * Direction + (1+Face + Direction | participants) + (1|order))`となった。

## 3. 結果

### 3.1. 好感度

`Clmm` の結果、主効果、交互作用ともに有意ではなかった。参加者の評定する好感度について、今回の実験では顔方向の動きや口の動きは有意な効果はもたらさないことがわかった。

### 3.2. 動画とアバターのどちらを見たか

`Clmm` の結果、顔方向の動きの主効果のみが有意であった( $z = 1.969, p = .049$ , 図2)。顔方向の動きがあれば、参加者は顔の動きがないときと比べ動画をよりみなくなると評定したことがわかった( $\beta = 0.8728$ )。ま口の動きがどちらを見ていたかについての評定に影響を及ぼすことはなかった。

図2 顔方向の動き条件を基準とした「動画とアバターのどちらをみていたか(に関する参加者自身の主観的評価)」についての CLMM の結果

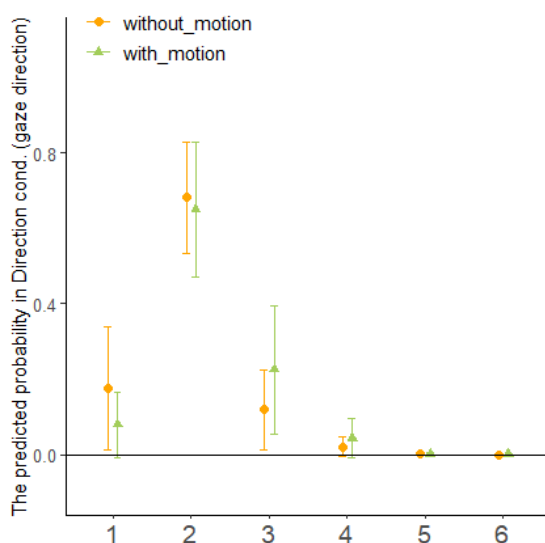
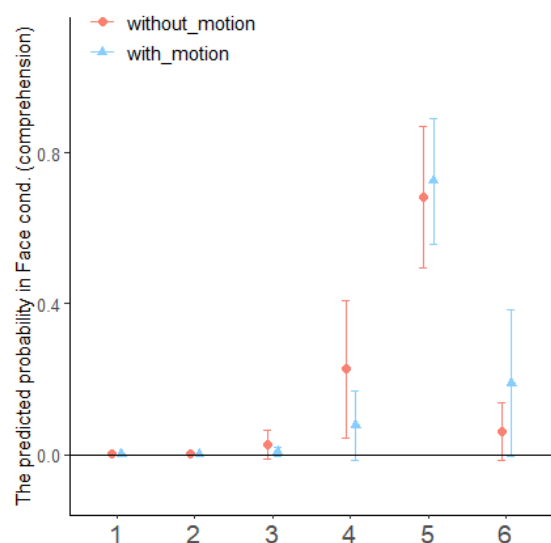


図3 口の動き条件を基準とした「ニュース動画に関しての参加者自身の主観的理解度」についての CLMM の結果



### 3.3. 動画の理解度

Clmmの結果、口の動きの主効果のみが有意であった( $z = 2.413, p = .01589$ , 図 3)。口の動きがあれば、参加者は口の動きがないときと比べ動画を理解したと評定することがわかった( $\beta = 1.3190$ )。顔方向の動きが評定に影響を及ぼすことはなかった。

## 4. 考察

好感度に関して、顔方向の動きや口の動きが評定に有意な影響を与えなかった。これは、参加者がアバターの動作に対して特定の方向性や口の表情に対して感じる好みや嗜好が一樣でなかった可能性があった。アバターの好感度は主観的な評価であり、個々の参加者の主観的な感じ方が大きく影響する要素と考える。次に、動画とアバターの注視に関して、他人の視線の方向に注意を向けることを視線手がかりと呼び(Frischen et al., 2007)、社会的相互作用を成功させる重要な能力である。

顔方向がニュース映像に向いている時の方が、映像の方を向いていない時よりも参加者が動画をあまり注視しないという結果が得られた。これは、アバターの動作が参加者の視線を惹きつけ、動画の視聴に影響を与えた可能性があった。アバターの動作がより注目されると、動画への注意が削がれることが示唆された。また、口の動きによる注視に関しては、結果に影響を与えなかったことから、口の動きが注目を集める要因ではなかった可能性が示された。最後に、動画の理解度においては口の動きがある場合に参加者が動画を理解しやすいと評価するという結果が得られた。これは口の動きが意思疎通に寄与する可能性があり、参加者が口の動きがあると情報を取得しやすくなると考えていることを示唆する。しかしながら、このことを明確に示唆するためには、様々なバリエーションで行う等、追試する必要がある。

本研究では、アバターの口の動きとニュース動画の解説音声が一致していた。この一致は視聴者にとって重要な視覚の手がかりとなり、聴覚情報との統合により理解が促進されると考えられる。Johnson & Mayer (2009)の「Multimedia Learning」は、視覚と聴覚の統合が学習者の理解を促進することが示されており、口の動きは視覚の手がかりとして機能し、聴覚情報との一致により参加者の理解を助けた可能性がある。

## 5. 結論

アバターの動作が視覚的な注意を惹きつけ、その

結果として動画の注視や理解度に影響を与えることが示唆された。ただし、好感度に関しては個人差が大きく、主観的な感じ方が多様であることがわかった。また、アバターの口の動きがある場合、参加者は口の動きがない時と比べ動画をより理解する効果を与える。さらに、アバターの顔方向が動くと、参加者はアバターに惹きつけられることがわかった。これらの知見は、アバター動画におけるコミュニケーションや情報伝達において、アバターのデザインや動作が重要な要素となることを示唆した。そのため、何かをアバターに発言させたい際に、アバターの顔方向を動かすことで、参加者に注目されやすくなり、さらに口を動かすことでスムーズに内容を理解しやすくなると考える。しかし、本実験ではアバターの各条件での提示順序の影響、視線の行動データや記憶課題を使った記憶の測定は行っていないため、それらを行うことで今回の結果を検証し、さらに明らかにしていく必要がある。また、本実験は動画内にアバターが登場するものであるため、今後の研究では、異なる実験条件やコンテキストにおけるアバターの影響をより詳細に調査する必要がある。

## 文献

- Aseeri, S., & Interrante, V. (2021). The influence of avatar representation on interpersonal communication in virtual social environments. *IEEE transactions on visualization and computer graphics*, 27(5), 2608-2617.
- Schroeder, R. (2002). Social interaction in virtual environments: Key issues, common themes, and a framework for research. In *The social life of avatars: Presence and interaction in shared virtual environments* (pp. 1-18). London: Springer London.
- Driver IV, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienti
- Mizuho, T., Amemiya, T., Narumi, T., & Kuzuoka, H. (2023, March). Virtual omnibus lecture: investigating the effects of varying lecturer avatars as environmental context on audience memory. In *Proceedings of the Augmented Humans International Conference 2023* (pp. 55-65).
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin*, 133(4), 694.
- Johnson, C. I., & Mayer, R. E. (2009). A testing effect with multimedia learning. *Journal of Educational Psychology*, 101(3), 621-629.