



# 声質変化に伴う自己認識の変化についてのアンケート調査

國見友亮<sup>1,2)</sup>, 畑田裕二<sup>1)</sup>, 木村健太<sup>1,2)</sup>, 鳴海拓志<sup>1)</sup>, 持丸正明<sup>1,2)</sup>

1) 東京大学 (〒 113-0033 東京都文京区本郷 7-3-1, y.kunimi,hatada,narumi@cyber.t.u-tokyo.ac.jp)

2) 産業技術総合研究所 (〒 277-0882 千葉県柏市柏の葉 6-2-3, kenta.kimura,m-mochimaru@aist.go.jp)

**概要:** バーチャル空間における自己表象 (アバター) の特性は自己認識を含む認知に影響を与えることが知られている。従来研究では、自己表象としてのアバター外見の効果に主眼が置かれ、アバター外見や外見と声質の変化の組み合わせがソーシャル VR 上での新たなアイデンティティの形成にどのように貢献するかが調査されてきた。他方、声質の変化がどの程度まで自己認識や自己連続性に影響を与えているかは十分に調査されていない。本研究では、ソーシャル VR 上でボイスチェンジャーや発声の工夫によって別性の声質を利用しているユーザーを対象としたアンケート調査を通じて、声質の変化が自己認識に影響を与える可能性を確認した。また、声質を変化させる方法の違いが自己連続性の捉え方に影響を与える可能性が示唆された。

**キーワード:** 声質、聴覚提示、自己認識、ソーシャル VR

## 1. はじめに

VR ヘッドマウントディスプレイを用いてアクセスした多数のユーザーがアバターを介したコミュニケーションを行えるバーチャル空間をソーシャル VR という。ソーシャル VR では実身体とは異なる特徴を持つアバターを用いて他者との社会的インタラクションを行うことが可能である。自らが表現したい自己の要素を反映したアバターを用いた社会的インタラクションは、アバター使用者のアイデンティティにも影響することが報告されている [1,2]。

他方、ソーシャル VR における理想的な自己の表現にはアバター外見だけではなく、ボイスチェンジャー等を利用して声質を変える方法もよく用いられている。VRChat を始めとするソーシャル VR には、実身体の性とは異なる性のアバターを使用し、声の性別もアバターの外見に合わせて変化させているユーザーも一定数存在する。実際、美少女アバターを利用しているユーザーへの調査では、アバターと一致した声質を利用することが自らのアイデンティティを理想的な自己に近づけていくために重要であることが示唆されている [3]。他方、声質変化の効果そのものに注目し、ソーシャル VR における声質変化がアイデンティティ形成の土台にある自己認識や社会的インタラクションにどのような影響を与えているかを調査した研究はない。

そこで本研究では、主要なソーシャル VR プラットフォームの一つである VRChat<sup>1)</sup>において声質変化を利用しているユーザーを対象に、声質変化の前後で自己認識や自己連続性、社会的インタラクションの変化を自覚するかを調査した。ソーシャル VR で声質変化を利用しているユーザーには、声質変化の方法として、図 1 に示すようにボイスチェンジャーでリアルタイムに音声を変換する方法を用いているユーザーのほか、自らの肉体的な発声方法の工夫によ

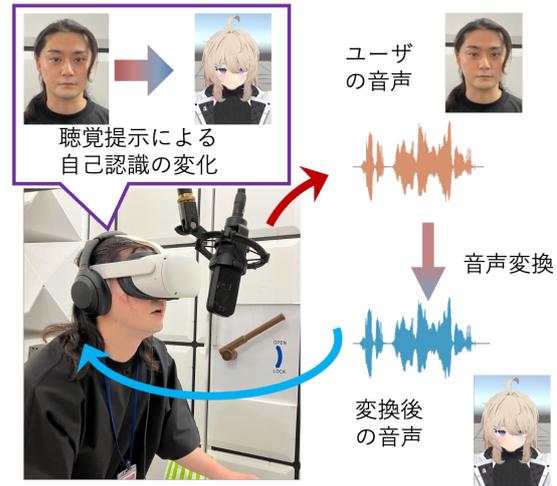


図 1: 声質の変化による自己認識の変化の概念。

て声質を変化させる「両声類」と呼ばれるユーザーがいる。本研究では、こうした声質を変化させる方法の違いが自己認識の変化に影響を与えているかについても調査した。

## 2. アンケート調査

### 2.1 調査内容

本研究では、自身の声質の変化に伴い自己認識が変化したように感じているか、また声質を変化させる方法の違いによってそうした回答の傾向が異なるかを明らかにするため、以下の問いに取り組む。

- RQ1: 声質の変化の前後で自己認識が変化したように感じるか?
- RQ2: 声質の変化方法によって回答の傾向に違いが生じるか?

本アンケート調査では、基本情報として、現実世界での

<sup>1)</sup><https://hello.vrchat.com/>

身体と心の性別、VRChat 内での身体と心の性別及び声の性別、声質を変化させる方法、声質を変化させている期間、声質を変化させている時の聴覚フィードバックの有無、声質を変化させるタイミングについて回答を行ってもらった。また、声質を変える前後の自己認識についての質問を3項目、自身の声質に対するイメージについて6項目、声質を変える前後での社会的インタラクションの違いについて5項目、VRChat をプレイする前後での自己連続性について4項目用意した。これらのアンケートでは、5段階のリッカート尺度(1: 強く同意しない, 3: どちらでもない, 5: 強く同意する)を利用した項目と自由記述の項目について回答を行ってもらった。本稿では表1に記載するリッカート尺度を用いたアンケート調査結果について触れる。

## 2.2 参加者

本アンケート調査はソーシャル VR(VRChat) 上で使用している声質が現実の身体とは別性であると自認しているユーザーを調査対象とした。参加者の募集は SNS への投稿を通じて行い、インフォームドコンセントに同意した参加者のみに回答を許諾した。本アンケートに参加したユーザーの中で、希望者はアンケート回答後に500円のAmazonギフト券を受け取った。

## 3. 結果

本調査では、Cisgender 全体の回答を、全ての回答が「どちらでもない」だった場合(平均値が3)と対応のある t 検定(両側)を行うことで RQ の検証を行った。VC と両声類の回答については、正規性が棄却され、等分散性が棄却されなかったため、ウィルコクソンの順位和検定(両側)を用いて検証を行った。

### 3.1 基本情報

本調査では、現実世界での身体と心の性別が一致しているユーザーを Cisgender, その他のユーザーを Others とし分類を行った。以下では声質の変化方法について、ハードウェア及びソフトウェアのボイスチェンジャーを VC, 自身の発声方法で声質を変化させているユーザーを両声類と表記する。アンケート回答者の声質変化方法を表2に示す。今回の調査では Others のアンケートの回答者が少なかったため、本稿では Cisgender ユーザーのみのデータについて報告する。Cisgender のユーザーの実身体の性別の内訳は、男性81人、女性5人であった。

### 3.2 自己認識についての調査の結果

Q1 の Cisgender 全体の回答(平均=3.44, 標準偏差=1.37)に対して対応のある t 検定を実施した結果、平均値は3よりも有意に高かった( $t = 2.99, p < .005$ )。Q1 の VC(平均=3.59, 標準偏差=1.31)と両声類(平均=3.24, 標準偏差=1.44)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった( $p \geq .100$ )。

Q2 の Cisgender 全体の回答(平均=3.60, 標準偏差=1.32)に対して対応のある t 検定を実施した結果、平均値は3よりも有意に高かった( $t(X) = 4.24, p < .005$ )。Q2 の VC(平

均=3.76, 標準偏差=1.18)と両声類(平均=3.41, 標準偏差=1.48)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった( $p \geq .100$ )。

Q3 の Cisgender の回答(平均=3.98, 標準偏差=1.11)に対して対応のある t 検定を実施した結果、平均値は3よりも有意に高かった( $t(X) = 8.19, p < .005$ )。Q3 の VC(平均=3.80, 標準偏差=1.15)と両声類(平均=4.22, 標準偏差=1.00)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意傾向が確認された( $p < .100$ )。

### 3.3 声の印象についての調査の結果

Q4 の Cisgender の回答(平均=3.03, 標準偏差=1.24)に対して対応のある t 検定を実施した結果、有意差は確認されず、平均値は3よりも有意に大きいとは言えなかった( $t(X) = 0.26, p > .100$ )。Q4 の VC(平均=3.00, 標準偏差=1.15)と両声類(平均=3.08, 標準偏差=1.36)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった( $p \geq .100$ )。

Q5 の Cisgender の回答(平均=4.06, 標準偏差=0.82)に対して対応のある t 検定を実施した結果、平均値は3よりも有意に高かった( $t(X) = 12.01, p < .005$ )。Q5 の VC(平均=4.06, 標準偏差=0.80)と両声類(平均=4.05, 標準偏差=0.85)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった( $p \geq .100$ )。

### 3.4 社会的インタラクションに関する調査の結果

Q6 の Cisgender の回答(平均=1.52, 標準偏差=1.01)に対して対応のある t 検定を実施した結果、平均値は3よりも有意に低かった( $t(X) = -13.50, p < .005$ )。Q6 の VC(平均=1.45, 標準偏差=1.02)と両声類(平均=, 標準偏差=)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった( $p \geq .100$ )。

Q7 の Cisgender の回答(平均=3.57, 標準偏差=1.31)に対して対応のある t 検定を実施した結果、平均値は3よりも有意に高かった( $t(X) = 4.04, p < .005$ )。Q7 の VC(平均=3.45, 標準偏差=1.26)と両声類(平均=3.73, 標準偏差=1.37)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった( $p \geq .100$ )。

Q8 の Cisgender の回答(平均=3.88, 標準偏差=1.34)に対して対応のある t 検定を実施した結果、平均値は3よりも有意に高かった( $t(X) = 6.11, p < .005$ )。Q8 の VC(平均=3.57, 標準偏差=1.38)と両声類(平均=4.30, 標準偏差=1.18)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差が見られた( $p < .005$ )。

Q9 の Cisgender の回答(平均=3.34, 標準偏差=1.52)に対して対応のある t 検定を実施した結果、平均値は3よりも有意に高かった( $t(X) = 2.06, p < .050$ )。Q9 の VC(平均=3.47, 標準偏差=1.49)と両声類(平均=3.16, 標準偏差=1.55)の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった( $p \geq .100$ )。

Q10 の Cisgender の回答(平均=2.85, 標準偏差=1.51)に対して対応のある t 検定を実施した結果、有意差は確認さ

表 1: 声質変化と自己認識に関係するアンケート調査の内容.Q は 5 段階のリッカート尺度を用いた項目.FD は自由記述も含む項目.

アンケート	質問内容
自己認識	Q1 声の性別を変えて喋った際に、別人になったような感覚を感じますか？
	Q2 声の性別を変える前後で、自分自身のイメージが変化したように感じますか？
	Q3 声の性別を変える前後で、他者から見た自分のイメージが変化したと思いますか？
声の印象	Q4 声の性別を変える前の自分の声は好きですか？
	Q5 声の性別を変えた後の自分の声は好きですか？
	FD1 声の性別を変えたいと思った理由を教えてください
	FD2 声の性別を変化させる前の自分の声にどのような印象を持っているか教えてください
	FD3 声の性別を変化させている際の自分の声にどのような印象を持っているか教えてください
社会的インタラクティブ	Q6 現実世界での自分の身体と、VRChat での身体の見た目は似通っていますか？
	Q7 声の性別を変化させる前後で、言葉選びが変わりますか？
	Q8 声の性別を変化させる前後で、喋り方が変わりますか？
	Q9 声の性別を変化させる前後で、他者への触れ合い方が変化しますか？
自己連続性	Q10 声の性別を変化させる前後では、違うタイプの人たちと交流していると感じますか？
	Q11 VRChat 内の身体を使用した際に、現実世界での声質を使用すると違和感を感じますか？
	Q12 現実世界での自分と VRChat 内での自分に繋がりがあのように感じますか？
	Q13 VRChat 内での経験が、現実世界での自分に影響を与えていると感じますか？
FD5 VRChat 内で喋る内容を一度頭の中で考える際に、声の性別を変化させた後の声で考えますか？	

表 2: 声質の変化方法についての分類

	ボイスチェンジャー (VC)	両声類
Cisgender	49	37
Others	5	19

れず、平均値は 3 と有意に異なるとは言えなかった ( $t(X) = -0.93, p \geq .100$ ). Q10 の VC (平均=3.06, 標準偏差=1.48) と両声類 (平均=2.57, 標準偏差=1.54) の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった ( $p \geq .100$ ).

### 3.5 自己連続性に関する調査の結果

Q11 の Cisgender の回答 (平均=3.35, 標準偏差=1.56) に対して対応のある t 検定を実施した結果、平均値は 3 よりも有意に高かった ( $t(X) = 2.07, p < .050$ ). Q11 の VC (平均=3.61, 標準偏差=1.47) と両声類 (平均=3.00, 標準偏差=1.63) の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差傾向が見られた ( $p < .100$ ).

Q12 の Cisgender の回答 (平均=3.51, 標準偏差=1.22) に対して対応のある t 検定を実施した結果、平均値は 3 よりも有意に高かった ( $t(X) = 3.87, p < .005$ ). Q12 の VC (平

均=3.29, 標準偏差=1.21) と両声類 (平均=3.81, 標準偏差=1.20) の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差が見られた ( $p < .050$ ).

Q13 の Cisgender の回答 (平均=4.08, 標準偏差=1.10) に対して対応のある t 検定を実施した結果、平均値は 3 よりも有意に高かった ( $t(X) = 9.14, p < .005$ ). Q13 の V (平均=3.96, 標準偏差=1.12) C と両声類 (平均=4.24, 標準偏差=1.06) の回答に関してウィルコクソンの順位和検定を実施した結果、両者の間に有意差は見られなかった ( $p \geq .100$ ).

## 4. 考察

### 4.1 RQ1

以上の結果から、声の性別の変化前後で自己認識が変化したように感じていることを確認できたともに、声の性別変化後の VRChat での自己と現実の自己の間に自己連続性もあるという可能性が示唆された。

#### 4.1.1 自己認識の変化

Q1-3 の結果から、声の性別の変化前後で主観的及び客観的な自己イメージが変化していることを確認でき、声の性別は自己認識を切り替える一助となっていることが示唆された。Q6 と Q11 の結果を併せて考えると、現実の身体とは乖離した特徴のアバターを使用する際にアバターのイメージに一致する声の性別を利用することで、アバター及

び声の性別に対する没入感が向上し、自己認識の変化が促進される可能性がある。

#### 4.1.2 自己連続性

Q1, Q11 が示唆しているように、声の性別の変化が自己認識に影響を与えている。ここからユーザーは、現実世界と VRChat のそれぞれにおいて異なる自己が存在していると認知しているのだと考えられる。他方で Q12, 13 の結果から、ユーザーは現実世界と VRChat における自己の間に繋がりを感じており、VRChat 上での経験が現実の自己に影響を与えていると感じていることも示唆された。現実と VRChat の自己は別々の自己であると認識しながらも自己連続性を感じている可能性があるのは興味深い点である。

#### 4.1.3 社会的インタラクション

Q7, Q8 の結果から、ユーザーは社会的インタラクションにおいて声の性別だけではなく、自身の言葉選びだけや発話スタイルも変化していると自覚していることがわかる。また Q9 の結果から、VRChat 上では他者との触れ合い方が変化していると感じるユーザーが多くいることも確認できた。他方で Q10 の結果からは、多くのユーザーが現実世界と VRChat のそれぞれで違うタイプの人たちと交流をしているわけではないと捉えていることが示唆されているため、VRChat 上では交流相手ではなく社会的インタラクションの内容が主に変化している可能性がある。

### 4.2 RQ2

Q3, Q8, Q10, Q11 の結果から、VC ユーザーと両声類ユーザーには回答の傾向が異なる設問も存在することが確認された。ただし、これらの回答には異なる声質を体得するための準備時間・練習時間や努力量なども影響すると考えられる。VC ユーザーがボイスチェンジャーを設定するための時間や、両声類ユーザーが異なる性別の声を体得するための時間を考慮できていない点は本調査の限界である。

#### 4.2.1 行為主体感と記憶による自己連続性

Q11, Q12 の結果から、両声類ユーザーは VC ユーザーよりも現実世界と VRChat 上での自己の連続性が高いと感じていることが示唆された。両声類ユーザーは別性の声質の発声方法を努力の末に体得しているため、別性の声質に対しても高い所有感があると考えられる。このような声の所有感が自己連続性に影響している可能性がある。行為主体感と記憶の研究では、行為に対する主体感が高いほど結果が自己に帰属し、記憶課題の成績が高くなることが分かっている [4]。行為主体感が記憶に影響を与えることを考えると、行為主体感に影響を与える発話と聴覚提示の間に遅延が存在しない両声類の方が記憶を保持しやすく、発話と聴覚提示の間に遅延が存在する VC ユーザーよりも自己連続性が高いと感じる可能性がある。今後は、発話と聴覚提示の間に存在する遅延、聴覚提示される音声の声質が行為主体感や声の所有感に影響を与えるか、また行為主体感や声の所有感が記憶や自己連続性について影響を与えるか調べる必要が存在する。

### 5. 結論

本研究では、声質の変化が自己認識の変化に影響を与えるか、また声質の変化手法の違いによって自己に関する回答の傾向に違いが生じるか検証した。アンケート調査の結果、声質の変化が現実とソーシャル VR 上での自己の認識に影響を与え、声質の変化前後で自己認識が切り替わっていると感じていることを確認した。これは、声質の変化が自己認識の変化のスイッチとして利用できる可能性を示唆している。また、ボイスチェンジャーを利用して声質を変化させているボイチェンユーザーおよび発声方法を変化させ声質を変化させている両声類ユーザーでは、両声類ユーザーの方が現実とソーシャル VR 上での自己に自己連続性を感じていることを確認した。しかしながら、本研究では声質変化による影響か、それともアバターの外見の変化による影響か、介入条件が統制されていないため分からないことも多い。今後は、両声類ユーザーのように遅延が少なく変換精度の良いボイスチェンジャーを開発し、実験条件が統制された介入実験を行うことで、遅延が少ない声質変化が自己認識や認知能力の変化にどのような影響を与えるか調べる必要がある。

謝辞本研究の一部は、JST ムーンショット型研究開発事業 (JPMJMS2013) ,JST さきがけ (JPMJPR22S9) および JST 次世代研究者挑戦的研究プログラム (JPMJSP2108) の支援を受けて行われた。

### 参考文献

- [1] Hatada, Y., Barbareschi, G., Takeuchi, K., Kato, H., Yoshifuji, K., Minamizawa, K., & Narumi, T. (2024, May). People with Disabilities Redefining Identity through Robotic and Virtual Avatars: A Case Study in Avatar Robot Cafe. In Proceedings of the CHI Conference on Human Factors in Computing Systems, 1–13.
- [2] Freeman, G., & Maloney, D. (2021). Body, avatar, and me: The presentation and perception of self in social virtual reality. Proceedings of the ACM on human-computer interaction, 4(CSCW3), 1–27.
- [3] Bredikhina, L., & Giard, A. (2022). Becoming a virtual cutie: digital cross-dressing in Japan. Convergence, 28(6), 1643–1661.
- [4] Hon, N., & Yeo, N. (2021). Having a sense of agency can improve memory. Psychonomic Bulletin & Review, 28, 946–952.