

# 声の類似性が信頼に与える効果に関する実証研究

## An Empirical Study on the Effect of Voice Similarity on Trust

高橋 直寛<sup>†</sup>, 能城 沙織

Naohiro Takahashi

<sup>†</sup>木更津工業高等専門学校

National Institute of Technology, Kisarazu College

sdj24b06@kisarazu.kosen-ac.jp

### 概要

人は自身と共通の特徴を持つものを好むことが先行研究より明らかになっており、声に関してもその傾向が示されている。しかし、先行研究では実験参加者の声そのものを実験に使用しており、そのことが評価に影響を与える可能性が否定できない。そこで本研究では、実験参加者の声から作成した自分類似合成音声を用いて実験を行うことを試みた。

実験の結果、信頼の評価には自身の声との類似性は影響を与えず、声自体の信頼性が影響を与える可能性が示唆された。

キーワード：声、類似性、認知科学

### 1. はじめに

介護士やカウンセラーなどの信頼関係が重要な仕事相手を探す際、利用者は自分と相性がいい人を見つけることで仕事相手と円満な関係を結ぶことが期待できる。相性が良い相手とはどのような相手だろうか？

先行研究では実験により、自分類似顔の魅力及び信頼を調べている (DeBruine, 2005)。実験参加者は4~9人のグループに分けられ、グループごとに自身の合成顔が含まれている9枚の画像セット内の2つの顔を比較し、信頼・短期的魅力・長期的魅力をより感じる顔を選択した。この操作を全てのペアについて行い、自身の合成顔を選んだ回数と同じグループ内の他の実験参加者が同じ画像を選んだ平均回数を比較した結果、自分と似た顔の相手は信頼できるということが示された。このような自身と共通の特徴を持つものを好んだり信頼できると感じたりする傾向を以後、同類婚的傾向と呼ぶ。

また、声に関しても自身と共通の特徴を持つものを好むという同類婚的傾向が見られることが実験により分かっている (Peng et al., 2019; Hughes & Harrison, 2013)。Peng et al.(2019)では、平均18種類の録音された音声の魅力度を1~7の7段階で評価してもらった実験を実施している。18種類の音声のうち3種類分は評価した実験参加者自身のものであったが、実験参加者は評価する声自身が自身の声であるかもしれないことは知らされていなかった。実験の結果、実験参加者は自分の声を他の

人が自分の声を評価した時よりも魅力的だと評価した。また、自分の声を他人の声よりも魅力的に聞こえると評価した。Hughes & Harrison(2013)では、男女の実験参加者に同性の実験参加者から録音された3種類の音声(数字、母音、子音)の魅力度を評価してもらった結果、実験参加者は自身の録音した音声を他人の録音した音声よりも魅力的であると評価した。

声に関する同類婚的傾向を扱ったこれらの先行研究はともに実験参加者の声そのものを評価に使用しており、自分自身の声だと認識した状態で評価を行うことで評価に影響を与え、正しい評価ができていない可能性が否定できない。

また、先行研究では魅力についても評価していたが、職業上の相性の良い相手を見つけることを考えた時、魅力ではなく、信頼を評価することの方が適切である可能性が考えられる。

そこで本研究では声に関して、魅力の代わりに信頼を評価することを試みた。自分類似合成音声を利用することで、自分の声そのものを使用せずに信頼を調査する手法を検討した。

本研究で実施した実験は木更津工業高等専門学校の倫理審査委員会の承認を得て実施した。

また、将来的には男女両方の実験を行う予定であるが、今回は途中段階として、実験参加者を男性のみにして実験を行った。

### 2. 研究概要

#### 2.1. 実験方法

本実験では、自分類似合成音声を用いた音声評価実験を行った。概要を以下に示す。

##### 2.1.1. 実験参加者

木更津工業高等専門学校の男子学生21名(平均:18.6歳、標準偏差:1.67歳)を対象に行った。

### 2.1.2. 刺激音声

刺激音声作成のために実験参加者の音声データを収集した。内容は ATR 音素バランス 503 文より、「一週間ばかり ニューヨークを 取材した」を採用した。

この音声データを基に TANDEM-STRAIGHT による音声モーフィング技術を用いて、以下の式(実験参加者が男性の場合)を満たすような、実験参加者の自分類似合成音声(同性)を作成した。

$$V_{sim} = V_{ave} + \alpha(V_i - V_{ave}) \quad (1)$$

ここで、 $V_{sim}$ は実験参加者の自分類似合成音声、 $V_{ave}$ は男性平均音声、 $V_i$ は実験参加者の音声を表す。

音声モーフィングとは、複数の話者の音声を混合する音声加工技術であり、中間的な特徴を持つ音声を生成することができる (Kawahara et al., 2008)。

$\alpha$ の値は木更津工業高等専門学校の学生 5 名(平均: 20.2 歳, 標準偏差: 0.898 歳)を対象に、式(1) $\alpha$ の値が 1/10, 1/5, 1/4 となるような自分類似音声を聞き、誰の自分類似合成音声か予想してもらうという予備実験により決定した。

この予備実験で、類似声の元となる声がわからないような  $\alpha$  の値を求めることで、自身の声が他の参加者の声と比較して高い割合で合成されているが、実験参加者は誰の類似声かわからないという状態で音声評価実験を行うことができる。

実験を行った結果、 $\alpha = 1/5$ となるような自分類似合成音声では、5 名中 5 名全員が誰の自分類似合成音声かわからなかったが、 $\alpha = 1/4$ となるような自分類似合成音声では、5 名中 5 名全員が誰の自分類似合成音声か分かっていた。

以上の結果より  $\alpha$  の値は  $\alpha = 1/5$  に決定した。

### 2.1.3. 実験手順

2.1.2 で作成した自分類似合成音声を用いて音声評価実験を行った。実験参加者(回答者)を 5 名のグループに分け、グループ内の人の自分類似合成音声同士を 1 対 1 で比較したときに、より信頼できると感じた音声を選択してもらう実験をグループ内の全ての組み合わせについて行った(図 1)。

図 1 音声評価実験の概要



信頼の具体的な基準は「2人の店員から、ある高価な商品をあなたに勧めた場合、どちらが勧める商品を買いたいと思うか」とした。

また、実験はデジタル騒音計(CHE-SD1)で 50dB 以下を記録して何度でも音声を聞き直してよい環境で行った。

### 2.1.4. 分析方法

音声評価実験終了後、Wilcoxon の符号付順位検定により回答者が自身の自分類似合成音声を選んだ回数、他者の自分類似合成音声を選ぶ回数の平均よりも有意に多いかを検証した。自身の自分類似合成音声を選んだ回数とグループ内の他者の自分類似合成音声を選んだ回数の平均を比較した。

また、実験に使用した音声そのものの信頼されやすさが音声の選択に影響を及ぼしているかどうかを調査するため、一般化線形混合モデルのロジスティック回帰分析によって分析を行った。

分析には以下のモデル式を用いて、自分自身の声と相手の声そのものが選択の基準となっているか調査した。

$$y = \text{logit}(z_i), z_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + r_i + \epsilon_i \quad (2)$$

ここで、 $y$ は自分類似合成音声を選択したかどうか(0:自分類似合成音声以外を選択, 1:自分類似合成音声を選択),  $x_1$ は自身の自分類似声自体の信頼されやすさ,  $x_2$ は自身とペアになっていた相手の自分類似声自体の信頼されやすさ,  $r_i$ は参加者によるランダム効果,  $\epsilon_i$ は誤差項となっている。

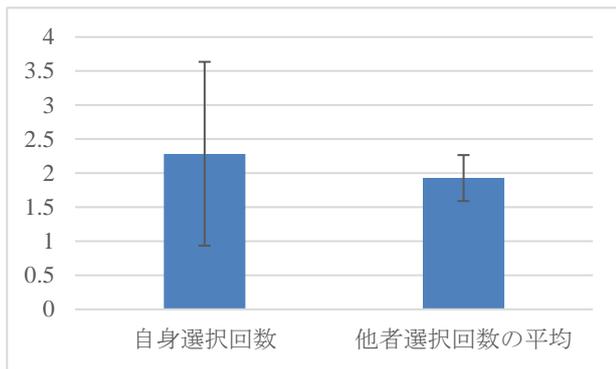
自分類似声自体の信頼されやすさは自分以外の他の実験参加者の回答によって決定した。例えば、実験参加者 A の自分類似合成音声自体の信頼されやすさは他の実験参加者が A の自分類似合成音声を選択した回数

平均によって決定した。ただし、選択した実験参加者自身の自分類似合成音声と A の自分類似合成音声の比較結果は反映しなかった。

## 2.4. 実験結果

Wilcoxon の符号付順位検定を行った結果、自分類似音声と他者の自分類似音声の選択されやすさに有意な差は認められなかった(図 2,  $N=21$ ,  $V=85$ ,  $p=0.4185$ )。

図 2 音声評価実験結果(平均値±標準偏差)



一般化線形混合モデルのロジスティック回帰分析を行った結果、以下ようになった。(表 1)

表 1 一般化線形混合モデル実行結果

変数名	推定値
$\beta_0$	-0.2749
$\beta_1$	1.8076*
$\beta_2$	-1.3093**

※\*は $p < 0.05$ , \*\*は $p < 0.01$ を示す。

自身の自分類似合成音声自体の信頼のされやすさの有意な主効果が認められ( $\beta_1 = 1.81$ ,  $p = 0.025$ )、自身の自分類似音声信頼されやすいほど自身の自分類似音声を選ばれやすいという傾向が示唆された。また、ペアになっていた相手の自分類似合成音声自体の信頼のされやすさの有意な主効果が認められ( $\beta_2 = -1.31$ ,  $p = 0.0035$ )、ペアになっている相手の自分類似音声信頼されやすいほど自身の自分類似音声を選ばれにくいという傾向が示唆された。

## 3. 考察

本研究では、自分類似合成音声を用いて音声の評価

実験を行い、声の類似性が信頼に与える効果について分析を行った。

その結果、自身の音声と類似した音声は他の音声と比較して信頼できるとは言えなかった。

先行研究で示された声の同類婚的傾向は自分自身の声だと認識した状態で行っていることが評価に影響を与えている可能性がある (Peng et al., 2019; Hughes & Harrison, 2013)。また、ロジスティック回帰の結果から、信頼できると感じる声の傾向は、自身の声の類似性によらず、声自体の信頼性が選択に影響を与えているという普遍的な傾向があると考えられる。

声に関する同類婚的傾向が見られなかった原因として、同性では似ている人を好むことによる遺伝的な利益がないからという可能性や、魅力と信頼では違う指標を用いている可能性が考えられる。

今後の展望として、実験対象を男性から女性に変えたり、刺激音声を同性の自分類似合成音声から異性の自分類似合成音声に変えたりすることで、声の類似性が信頼に与える効果を調査できると考えられる。

## 文献

- Lisa M DeBruine(2005). Trustworthy but not lust-worthy: context-specific effects of facial resemblance. PubMed, Vol.272, pp.919-922.
- Susan M Hughes, Marissa A Harrison (2013). I like My Voice Better: Self-Enhancement Bias in Perceptions of Voice Attractiveness. PubMed, Vol.42, pp.941-949.
- Zhikang Peng, Yanran Wang, Linghao Meng, Hongyan Liu, Zhiguo Hu (2019). One's own and similar voices are more attractive than other voices. Australian Journal of Psychology, Vol.71, pp.212-222.
- Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T and Banno, H. (2008) "TANDEM -STRAIGHT: A Temporally Stable Power Spectral Representation for Periodic Signals and Applications to Interference-free Spectrum, FO, and Aperiodicity Estimation,"Proc. f · CASSP2008, Las Vegas, 3933-3936.