Vol.2020-HCI-190 No.15
Vol.2020-UBI-68 No.15
2020/12/9

音声入力使用時の恥ずかしさを軽減する手法の提案

塚本賢治[1]　中島達夫[1]

概要：本論文では，日本人が人前で音声入力を使用するときに感じる恥ずかしさの原因を調査し，その軽減方法について提案する．最近では，スマートフォンなどの身近なデバイスにも音声入力の機能が備わっており，音声入力を使用することのできる機会は多くなっている．一方で，音声入力を使うことができる状況であっても音声入力を使用しない人は少なくない．本研究ではまず，音声入力を使用できる場面であっても音声入力を使用しない原因について調査した．その結果，声を発するという行為自体が恥ずかしいと感じる人が多いことが明らかになった．さらに，その恥ずかしさは音声入力の対象を生物的なキャラクターとすることで軽減できる可能性があることが明らかになった．そこで，音声入力の対象が無機質な機械と生物的なキャラクターの 2 つの場合で，音声入力の恥ずかしさを比較する実験を行った．その結果，生物的なキャラクターを表示した場合の方が音声入力の恥ずかしさが軽減されるという結果が得られた．

1.　はじめに

近年，音声入力はユーザーにとって様々なデバイスと対話する便利で自然な方法であると評価されている[1][2][3]．音声入力はハンズフリーな入力方法であり，それを使用するための特別な訓練も必要としない[4]．我々の身近なデバイスにも音声入力の機能をもつものが増加している．例として，日本で広く普及しているスマートフォンには Google アシスタント[2]や Siri[3]といった音声入力の機能がある．これらの機能を使うと，スマートフォンに話しかけるだけでスマートフォンの操作をすることができたり自身の知りたい情報を得ることができる．また，家庭用のスマートスピーカーも普及が進んでいるほか，公共の場で使用される音声入力もある．既存の研究では，車椅子に音声入力の機能を持たせたもの[5]や，路線バスのバス停を音声入力に対応させたもの[6]が存在する．また，2020 年に開業した東京都内の高輪ゲートウェイ駅の構内には，音声入力によって知りたい情報を得られる端末が複数設置されている[7].

このように，我々が音声入力に触れる機会は増加している．一方で，日本人が実際に音声入力をする機会が増えたとは感じない．実際，2018 年 12 月時点での日本国内におけるスマートスピーカーの認知率は 76.1%であったのに対して，その普及率はわずか 5.9%であった[8]．これは 2019 年 1 月時点でのアメリカでのスマートスピーカーの普及率 26%[9]や，2019 年時点でのイギリスのスマートスピーカーの普及率 19%[10]と比較しても低い水準であることが分かる．このことから，日本人は音声入力をすることに対して何らかの抵抗感を感じていることが窺える．

また，2017 年に行われた日本人の音声操作に対する意識調査[11]では，人前での音声検索は恥ずかしいと回答した人が 71.1%にのぼった．このことから，日本人は人前で音声入力を使用することに恥ずかしさを感じ，それが音声入力の利用を妨げていると推測される．

これらの状況を踏まえて，本研究では日本人が人前で音声入力をするときに感じる恥ずかしさを軽減し，音声入力の使用の敷居を低くすることを目的とする．この目的を達成するため，まず日本人が音声入力をするときに感じる恥ずかしさについての調査を行った．続いて，調査の結果を踏まえて音声入力を使用する時の恥ずかしさを軽減する手法の提案を行う．提案手法では，音声入力の対象を無機質な機械ではなく，図 1 に示すような動物の形をしたスピーカーとした．この手法によって，音声入力を使用する時のユーザーの心情や恥ずかしさの程度が変化するかについて実験を通じて調査した．

図 1　動物の形をしたスピーカー

2.　関連研究

現状，人前での音声入力にはプライバシーや周囲の騒音の影響を受けやすいといった問題が存在し，ユーザーもそれを懸念している[12][13]．本章では，これらの課題を解決する手法を提案した既存の研究を紹介することで，本研究

1 早稲田大学
Waseda University

2 https://assistant.google.com/intl/ja_jp/
3 https://www.apple.com/jp/siri/

ⓒ2020 Information Processing Society of Japan

がこれらの課題を無視して恥ずかしさのみに着目することの妥当性を主張する．

2.1 ジェスチャーによって音声の拡散を抑える研究

Yan ら[14]の研究では，現状の音声入力の課題はプライバシーの問題とトリガーとなる動作を毎回行う必要があることだとしている．この2つの課題を，ジェスチャーをしながら音声入力を行うことで同時に解決した．ジェスチャーは口の片側を覆うようなものである．このジェスチャーをしながら話すことで，声の広がり方が左右で非対称になる．ユーザーが発した音声はユーザーが両耳につけたイヤフォンのマイクで検知し，これが非対称なものであれば音声入力をしていると判断される．

この手法の利点は大きく 2 つ存在する．1 つ目は，音声の拡散を抑えると同時に唇の動きを手で隠すことによって，プライバシーを強化できることである．2 つ目は，音声入力のトリガーとなる動作が不要になり，特にデバイスが手元にない場合は物理的なボタンよりもアクセスが容易になることである．

この手法ではジェスチャーを高精度で認識できたほか，ユーザーにとって分かりやすい方法であったことが示された．

2.2 口の動きによって入力を行う研究

唇の動きによって音声入力と同じような機能を実現しようという研究が存在する[15][16]．Sun ら[15]の研究では，現状の音声入力の課題はプライバシーの問題と周囲の騒音に影響されやすいことだとしている．これらの課題を解決するため，スマートフォンのフロントカメラを利用してユーザーの口の動きをキャプチャすることで音声入力のコマンドを認識する．各コマンドを入力する時の口の動きはディープラーニングを使って学習され，44 個のコマンドのサポートを実現した．

この手法の利点は大きく 3 つ存在する．1 つ目は，1 つのコマンドで効率的に機能にアクセスできることである．2 つ目は，手が利用できない時にも機能を提供できることである．3 つ目は，タッチ操作を支援してデバイスを使用可能な機会を多くすることである．

この手法では，実験で使用したほとんどのタスクにおいて，タッチベースの入力と比較して完了時間が短く入力のしやすさが向上することが示された．

2.3 息を吸いながら発する音で音声入力をする研究

Fukumoto[17]の研究では，現状の音声入力の課題はプライバシーの問題と音声が拡散することだとしている．これらの課題を解決するため，息を吸いながら微かに音声を発することで音声入力をする方法を提案している．ユーザーは手に持った専用のセンサーを唇に当てながら話すことで，微かな音声であっても認識される．センサはマイクを使って空気の流れによる音を検知し，85 個のコマンドのサポートを実現した．

実験によって，この手法では音声の拡散が非常に少ないことが確認された．また，周囲の騒音が大きい環境であっても高い精度で入力を検知することが可能であった．

3. 事前調査

この章では，日本人が音声入力をする時に恥ずかしさを感じる原因について調査する．既存の研究[18][19][20][21]によれば，恥感情が生じる原因は大きく 2 つに分類される．1 つは自己の考え方など自分自身の内面に依存するものである．音声入力をする場合においてこれに当てはまるものは，音声入力そのものに対するイメージや人前で声を出すことについての抵抗感などであると考えられる．もう 1 つは周囲の人間からどのように見られているかなどの環境に依存するものである．音声入力をする場合においてこれに当てはまるものは，話す内容が他人に聞かれることへの懸念や独り言を言っていると思われることを心配する感情であると考えられる．

3.1 調査方法

調査はアンケート形式で行った．音声入力の使用が想定される 7 個の場面のシナリオを設定し，それぞれのシナリオにおいて音声入力を使用することにどのような抵抗を感じるかについて調査した．

表 1　質問項目

質問番号	質問内容
Q1	この状況で音声入力をすると思いますか
Q2	1 つ目の質問で「おそらくしないと思う」「しないと思う」を選んだ方にお聞きします．この状況で音声入力を使用しない理由を教えてください（複数選択可）
Q3	この場面と同じ状況に置かれた場合，あなたの考えに当てはまるものを選んでください（複数選択可）
Q4	この場面で音声入力を使用することについて何か意見や感想があればお書きください

表 2　Q1 の回答選択肢

選択肢番号	選択肢
1	すると思う
2	おそらくすると思う
3	おそらくしないと思う
4	しないと思う

ⓒ2020 Information Processing Society of Japan

表 3　Q2 の回答選択肢

選択肢番号	選択肢
1	喋る内容を周囲に聞かれるのがいやだから
2	この環境で声を出すのが恥ずかしいから
3	独り言を言っていると思われそうで恥ずかしいから
4	音声入力以外の行為（スマホの画面操作など）で目的を達成する方が簡単だから
5	その場の空気を壊しそうだから
6	知らない人の前で声を出したくないから
7	確度が低いから
8	その他（自由記述）

表 4　Q3 の回答選択肢

選択肢番号	選択肢
1	音声入力をする対象として，動物などをモチーフにしたキャラクターが自分だけに見える形で存在する（または表示される）なら，音声入力に抵抗はない
2	音声入力をする対象として，動物などをモチーフにしたキャラクターが自分以外の人にも見える形で存在する（または表示される）なら，音声入力に抵抗はない
3	音声入力をする対象として，人のようなキャラクターが自分だけに見える形で存在する（または表示される）なら，音声入力に抵抗はない
4	音声入力をする対象として，人のようなキャラクターが自分以外の人にも見える形で存在する（または表示される）なら，音声入力に抵抗はない
5	音声入力をする対象が何であっても音声入力に抵抗はあるが，自分のそばにいる人と会話することには抵抗がない
6	上記の全てに抵抗がある

各シナリオは周囲の環境と音声入力の内容の 2 つの項目で区別される．周囲の環境は，音声入力をする時に声の届く範囲に人がいない場合，家族がいる場合，友人がいる場合，知らない人がいる場合の 4 つに分類した．音声入力の内容は，聞き手がいる時の行動に関する内容と聞き手がいない時の行動に関するものの 2 つに分類した．なお，声の届く範囲に人がいない場合では聞き手が存在しないため，音声入力の内容による区別を行っていない．

アンケートでは，各シナリオについて 4 つの質問を行った．質問の内容は表 1 に示す．また，表 1 の Q1 の選択肢を表 2，Q2 の選択肢を表 3，Q3 の選択肢を表 4 に示す．

回答者は 19 人（男性 13 人，女性 6 人）で全員が 20 代（平均 22.7 歳）の日本人であった．被験者の中に海外への長期在住経験のある人はいなかったが，うち 2 人は海外への 6 ヶ月以内の留学経験があった．また，回答者のうち音声入力の使用経験がない人は 2 人で，その他の回答者は少なくとも一度は音声入力の使用経験があった．

調査に使用した 7 個のシナリオの，周囲の環境と音声入力の内容を表 5 に示す．なお，各シナリオでは音声入力以外の方法でも目的の操作を実行できることを明記した．

表 5　各シナリオにおける周囲の環境と音声入力の内容

シナリオ番号	周囲の環境と音声入力の内容
1	聞き手がいないシナリオ
2	聞き手が家族で聞き手のいない場面での行動に関する入力を行うシナリオ
3	聞き手が友人で聞き手のいない場面での行動に関する入力を行うシナリオ
4	聞き手が知らない人で聞き手のいない場面での行動に関する入力を行うシナリオ
5	聞き手が家族で聞き手のいる場面での行動に関する入力を行うシナリオ
6	聞き手が友人で聞き手のいる場面での行動に関する入力を行うシナリオ
7	聞き手が知らない人で聞き手のいる場面での行動に関する入力を行うシナリオ

3.2　調査結果と考察

表 1 の Q1，Q2，および Q3 の結果をそれぞれ図 2，図 3，および図 4 に示す．また，Q1 の回答に対して，「すると思う」の回答を 1，「しないと思う」の回答を 4 として計算した加重平均を図 5 に示す．Q2 の回答の選択肢を，恥ずかしさの原因ごとに分類したものを表 6 に，その分類ごとに結果を集計したものを図 6 に示す．

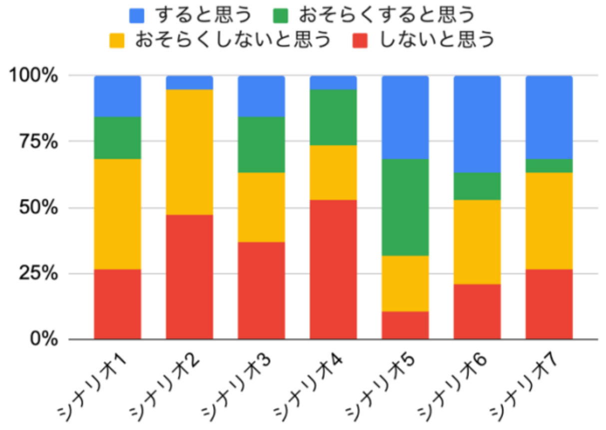

図 2　Q1 の回答

ⓒ2020 Information Processing Society of Japan

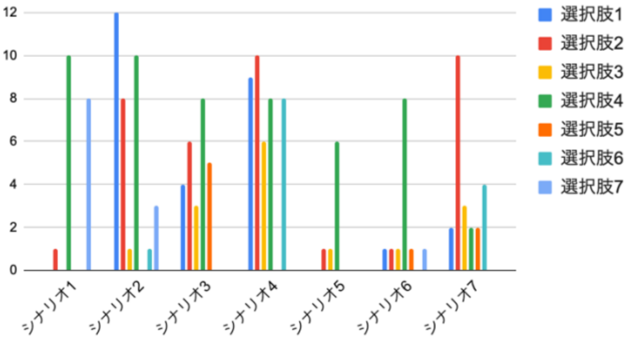

図 3　Q2 の回答

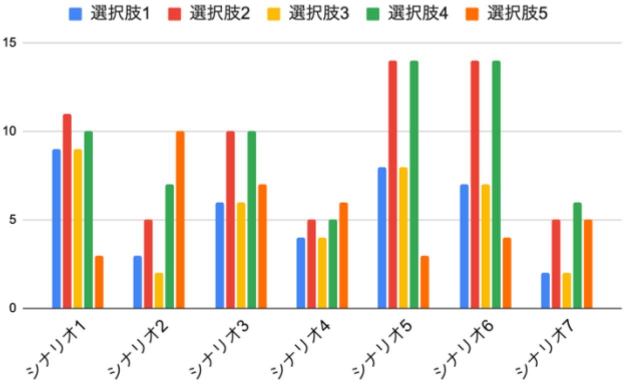

図 4　Q3 の回答

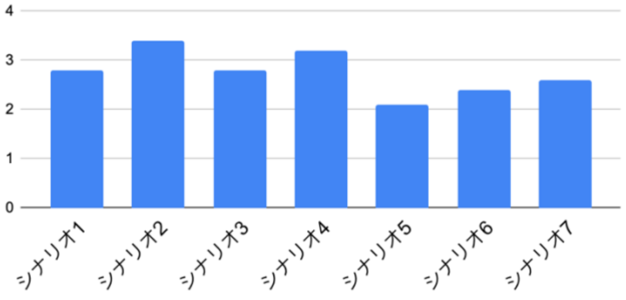

図 5　Q1 の回答の加重平均

表 6　Q2 の選択肢の分類

分類	分類される選択肢
自己の内面に依存する恥ずかしさ	この環境で声を出すのが恥ずかしいから/知らない人の前で声を出したくないから
環境に依存する恥ずかしさ	喋る内容を周囲に聞かれるのがいやだから/独り言を言っていると思われそうで恥ずかしいから/その場の空気を壊しそうだから
その他の要因	音声入力以外の行為（スマホの画面操作など）で目的を達成する方が簡単だから/確度が低いから

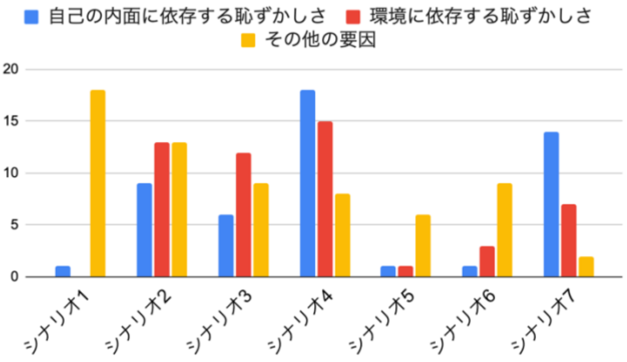

図 6　要因毎に分類した Q2 の回答

図 2 に示す通り，1 つのシナリオを除いて音声入力の使用に否定的な回答が半数以上となった．シナリオの中で唯一音声入力の使用に肯定的な回答が半数より多くなったものは，聞き手が家族で聞き手のいる場面での行動に関する入力を行うシナリオ（シナリオ 5）であった．また，図 6 に示すように音声入力の使用に否定的な回答をした被験者は，自己の内面と環境に依存する恥ずかしさの双方を感じていた．音声入力の使用に否定的な回答をした被験者が，自己の内面または環境に依存する恥ずかしさを感じたとする回答は全てのシナリオの平均で 1 人あたり 1.19 件であった．これを聞き手がいるシナリオ（シナリオ 1 以外）に限定すると，自己の内面または環境に依存する恥ずかしさを感じたとする回答は全てのシナリオの平均で 1 人あたり 1.40 件であった．これは，音声入力の使用に否定的な回答をした多くの被験者が恥ずかしさを感じていたことを示している．

聞き手がいないシナリオ（シナリオ 1）においても，音声入力の使用に否定的な回答が半数を超えた．一方で，図 3 や図 6 に示す通り音声入力の使用に否定的な回答をした原因は，「音声入力以外の行為（スマホの画面操作など）で目的を達成する方が簡単だから」などの恥感情とは関係のないものがほとんどであった．つまり，音声入力をする時に周囲に人がいなければほとんど恥ずかしさを感じないということが分かる．また，このシナリオでは音声入力を使用しない理由として「確度が低いから」をあげる人が全てのシナリオの中で最も多くなった．このシナリオでは音声入力を使用してスマートフォンのアラームをセットする場面としたが，一部の被験者がアラームをセットする行為は確実に実行したいと考え，正しく操作が実行できたかを確認することが難しい音声認識の使用が避けられる結果となった．

聞き手がいるシナリオにおいて，音声入力の内容の違いによる感情の変化に着目して分析する．聞き手が同じシナリオの組（シナリオ 2 とシナリオ 5 の組，シナリオ 3 とシナリオ 6 の組，シナリオ 4 とシナリオ 7 の組）ごとに比較した場合，全ての組において聞き手のいる場面での行動に関する入力を行うシナリオよりも聞き手のいない場面での

ⓒ2020 Information Processing Society of Japan

行動に関する入力を行うシナリオの方が，図 5 に示す加重平均が高くなった．これは，聞き手のいない場面での行動に関する入力を行うシナリオでは，聞き手のいる場面での行動に関する入力を行うシナリオと比較して音声入力を使用に抵抗を感じる傾向があることを示している．また，図 6 を見ると入力内容の違いによって恥ずかしさの感じ方に大きな差があることが分かる．聞き手のいない場面での行動に関する入力を行うシナリオ（シナリオ 2，3，4）では，自己の内面に依存する恥ずかしさを感じたとする回答は平均 11.0 件，環境に依存する恥ずかしさを感じたとする回答は平均 13.3 件であった．一方，聞き手のいる場面での行動に関する入力を行うシナリオ（シナリオ 5，6，7）では，自己の内面に依存する恥ずかしさを感じたとする回答は平均 5.3 件，環境に依存する恥ずかしさを感じたとする回答は平均 3.7 件であり，聞き手のいない場面での行動に関する入力を行うシナリオと比較してどちらも半分以下の回答数となった．これらの結果から，聞き手のいない場面での行動に関する入力を行う場合は，聞き手のいる場面での行動に関する入力を行う場合と比較して恥ずかしさを感じやすい傾向があると言える．

続いて，聞き手の違いによる感情の変化に着目して分析する．聞き手のいる場面での行動に関する入力を行うシナリオ（シナリオ 5，6，7）では，図 2 に示す通り聞き手が家族，友人，知らない人と親密度が下がるにつれて音声入力の使用に否定的な回答が多くなった．一方で，聞き手のいない場面での行動に関する入力を行うシナリオ（シナリオ 2，3，4）では，聞き手が家族の場合に否定的な回答が最も多く，友人の場合に肯定的な内容が最も多くなった．シナリオ 2 で聞き手との親密度が高いにも関わらず音声入力の使用に否定的な回答が多くなった原因は，このシナリオにおける入力内容が他のシナリオと異なる質のものだったため，被験者が入力内容を聞かれたくないと考えたことだと考えられる．この要因を取り除くことができれば，シナリオ 2 で音声入力の使用に肯定的になる人が増え，親密度が高いほど恥ずかしさが減少する傾向がより顕著に表れると推測される．

Q3 の回答では，図 4 に示す通りシナリオ毎の結果に大きな違いは見られなかった．全てのシナリオの平均では，「音声入力をする対象として，人のようなキャラクターが自分以外の人にも見える形で存在する（または表示される）なら，音声入力に抵抗はない」が最も多く，被験者の 52.4%がこの回答をした．次に多かった回答は「音声入力をする対象として，動物などをモチーフにしたキャラクターが自分以外の人にも見える形で存在する（または表示される）なら，音声入力に抵抗はない」であり，被験者の 50.8%の被験者がこれらの回答をした．一方，「音声入力をする対象として，動物などをモチーフにしたキャラクターが自分だけに見える形で存在する（または表示される）なら，音声入力に抵抗はない」と回答した被験者は 31.0%，「音声入力をする対象として，人のようなキャラクターが自分だけに見える形で存在する（または表示される）なら，音声入力に抵抗はない」と回答した被験者は 30.2%であった．このことから，音声入力の抵抗感が少ないのは音声入力の対象が他人の視点からでも見える場合であることが分かる．

4. 実験

事前調査の結果を受けて，音声入力を使用する時のユーザーの恥ずかしさを軽減する手法を提案する．事前調査では，音声入力の使用をしない大きな原因として恥ずかしさを感じていることが挙げられ，音声入力の対象として生物的なキャラクターが他人にも見える形で表示されることで恥ずかしさが軽減される可能性が示唆された．実験では，音声入力の対象がスマートフォンである場合と動物の形をした機械である場合をすることでユーザーの心情にどのような変化が生じるのかを調査する．

4.1 実験方法

実験は主著者と被験者がボイスチャットをしながら行った．被験者はまず音声入力の使用を想定したシナリオを読んだ．続いてそのシナリオの場面を再現した動画を視聴した．動画は，バーチャルリアリティ（以下 VR）空間内でヘッドマウントディスプレイを通して見ている風景を撮影したのである．動画の途中の音声入力をするシーンで音声入力をする言葉が字幕として表示され，被験者はその言葉を実際に声に出して読み上げた．動画の視聴後にアンケートを実施し，音声入力を使用する時の恥ずかしさなどを調査した．動画は音声入力の対象がスマートフォンである場合と動物の形をした機械である場合の 2 つを用意し，各被験者が両方の動画を視聴した．動画のスクリーンショットを図 7 に示す．

図 7　実験に使用した動画のスクリーンショット

アンケートの質問項目のうち，動画毎に行った質問を表 7 に示す．また，実験の最後に表 8 に示す 4 つの項目について，音声入力の対象がスマートフォンである場合と動物

の形をした機械である場合のどちらが当てはまるかを 7 段階で質問した. これを Q5 とする.

被験者は 10 人（男性 8 人，女性 2 人）で全員が 20 代（平均 23.5 歳）の日本人であった．被験者の中に海外への 1 年以上の在住経験や留学経験のある人はいなかった．被験者は全員, 事前調査で実施したアンケートの回答者であった.

表 7　動画毎に行った質問項目

質問番号	質問内容	回答方式
Q1	この状況で音声入力を使用することにどの程度抵抗がありますか	選択式（強い抵抗がある/抵抗がある/少し抵抗がある/抵抗はない）
Q2	この状況で音声入力を使用することにどの程度恥ずかしさを感じますか	選択式（強く感じる/感じる/少し感じる/感じない）
Q3	1 つ目の質問で「抵抗はない」【以外】を選んだ方にお聞きします．抵抗を感じる原因として恥ずかしさ以外の要素があればお書きください	自由記述
Q4	2 つ目の質問で恥ずかしさを「感じない」【以外】を選んだ方にお聞きします．恥ずかしさを感じる原因を選んでください（複数選択可）	選択式（この環境で声を出すのが恥ずかしいから（選択肢 1）/喋る内容を周囲に聞かれるのが恥ずかしいから（選択肢 2）/独り言を言っていると思われそうで恥ずかしいから（選択肢 3）/その他）

表 8　どちらの入力対象がより当てはまるかを聞いた項目

項目番号	項目
1	気軽に使える
2	違和感なく使える
3	知らない人の前でも使える
4	実際に家で使いたい

4.2 シナリオ

実験に使用したシナリオは，聞き手が友人で聞き手のいない場面での行動に関する入力を行うシナリオである．この条件のシナリオを設定した理由は大きく 2 つある．1 つ目は，事前調査の結果，この条件のシナリオでは恥ずかしさを理由として音声入力の使用に否定的な被験者が平均よりも多く，恥ずかしさを軽減する手法の実験に適しているためである．2 つ目は，音声入力の対象として動物などをモチーフにしたキャラクターが他人にも見える形で表示されていれば音声入力をすると回答した被験者が平均よりも多く，提案手法の効果が表れやすいと推定されるためである．また，事前調査では音声入力よりもその代替手段を使って入力をする方が簡単であると考える被験者が一定数いたため，本実験では代替手段を使うことが困難な状況を設定した.

4.3 実験結果と考察

表 7 の Q1，Q2，Q4 の結果をそれぞれ図 8，図 9，図 10 に示す. また, Q1 の結果の「強い抵抗がある」の回答を 3, 「抵抗はない」の回答を 0 として計算した加重平均，中央値，分散を表 9 に示す．Q2 の結果の「強く感じる」の回答を 3，「感じない」の回答を 0 として計算した加重平均，中央値，分散を表 10 に示す．Q5 の結果を図 11 に示し，その結果に対して「スマートフォン（などの無機質な機械）の方が当てはまる」を-3，「動物型のスマートスピーカーの方が当てはまる」を 3 として計算した加重平均, 中央値, 分散を表 11 に示す.

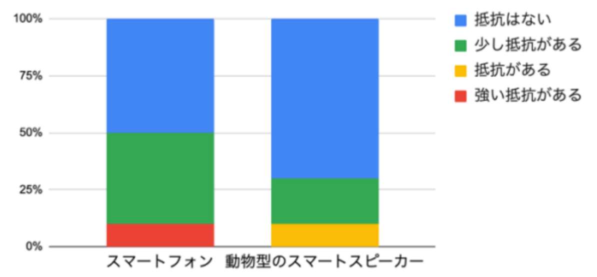

図 8　Q1 の回答

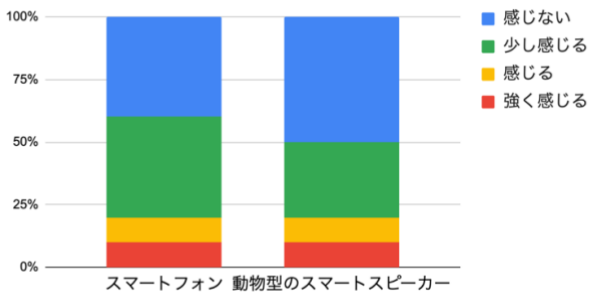

図 9　Q2 の回答

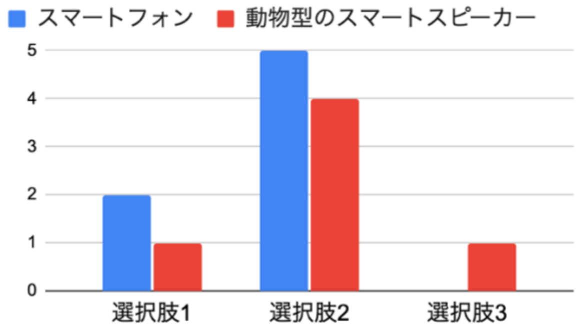

図 10　Q4 の回答

表 9　Q1 の回答の代表値

入力対象	加重平均	中央値	分散
スマートフォン	0.7	0.5	0.90
動物型のスマートスピーカー	0.4	0.0	0.49

表 10　Q2 の回答の代表値

入力対象	加重平均	中央値	分散
スマートフォン	0.9	1.0	0.99
動物型のスマートスピーカー	0.8	0.5	1.07

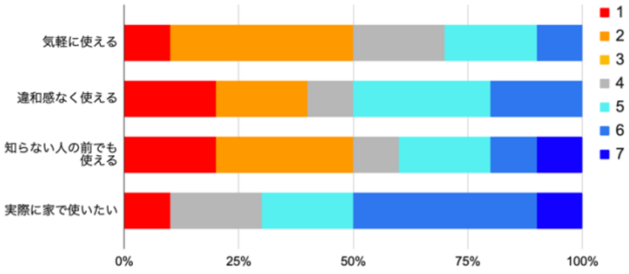

図 11　Q5 の回答

表 11　Q5 の回答の代表値

項目	加重平均	中央値	分散
気軽に使える	-0.7	-1.0	2.90
違和感なく使える	-0.3	0.5	4.01
知らない人の前でも使える	-0.5	-1.0	4.72
実際に家で使いたい	1.0	1.5	2.89

表 9 に示す通り，音声入力の対象がスマートフォンである場合よりも動物型のスマートスピーカーである場合の方が抵抗感の加重平均と中央値が小さくなった．これは，音声入力の対象が動物型のスマートスピーカーである場合の方が音声入力の抵抗感が小さくなったことを表している．また，表 10 に示す通り，音声入力の対象がスマートフォンである場合よりも動物型のスマートスピーカーである場合の方が恥ずかしさの加重平均と中央値が小さくなったが，その差は抵抗感における差よりも小さくなった．音声入力の使用に恥ずかしさを感じるとした被験者が，恥ずかしさの原因として挙げた項目は図 10 の通りである．音声入力の対象がスマートフォンと動物型のスマートスピーカーとで大きな差は見られなかった．このことから，音声入力の恥ずかしさは多少軽減されたが，抵抗感が軽減された原因は恥ずかしさが軽減されたこと以外にも存在すると推定される．

個々の被験者の結果を見ると，音声入力の対象がスマートフォンである場合よりも動物型のスマートスピーカーである場合の方が抵抗感や恥ずかしさが軽減されると感じた被験者は 4 人だった．それに対し，音声入力の対象が動物型のスマートスピーカーである場合よりもスマートフォンである場合の方が抵抗感や恥ずかしさが軽減されると感じた被験者は 1 人，どちらの場合も特に変わらないとする被験者が 5 人であった．この 5 人のうち 3 人は，どの場合でも音声入力を使用することに対して抵抗や恥ずかしさを感じないと回答していた．音声入力の対象が動物型のスマートスピーカーである場合の方が抵抗感や恥ずかしさが軽減されると回答した理由は，声を出すことの恥ずかしさが少なくなったから，自分の話を聞いてもらっている感覚があり話しやすかったから，相手に意志がある印象を受けるので人や動物などと対話している気分になるからなどが挙げられた．音声入力の対象がスマートフォンである場合の方が抵抗感や恥ずかしさが軽減されると回答した理由は，対象の見た目がスマートスピーカーらしくなく周囲から見て独り言を言っていると思われそうだから，無機質な機械なので音声入力という行為はそういうものだと割り切って話すことができるからなどが挙げられた．このことから，動物的な愛嬌のある見た目の中に一般的なスマートスピーカーのような機械的な見た目を取り入れることで，より多くのユーザーにとって有益なものになると考えられる．

図 11，表 11 は，いくつかの項目についてスマートフォンと動物型のスマートスピーカーのどちらが好ましいかを質問した結果である．「気軽に使える」の項目では，加重平均と中央値が共に負になり，スマートフォンの方が好まれる結果となった．理由として，携帯性が高いことなどが挙げられた．「違和感なく使える」と「知らない人の前でも使える」の項目では，加重平均と中央値が 0 に近くなったほか分散も大きく，好みが分かれる結果となった．「実際に家で使いたい」の項目では，加重平均と中央値が共に正になり，動物型のスマートスピーカーの方が好まれる結果となった．理由として，ペットのような愛着が生まれそう，積極的に話しかけたいと感じたなどが挙げられた．また，家で使う場合であっても，一人暮らしであれば動物型のスマートスピーカーが良いが家族がいるのであればスマートフォンの方が好ましいという意見もあった．これらの結果から，動物型のスマートスピーカーは，自宅で使用することが適しており，その場合において音声入力の抵抗感や恥ずかしさを軽減することが期待される．

5. まとめと今後の展望

本論文では，日本人が人前で音声入力を使用する時に感じる恥ずかしさの原因を調査した．また，それを軽減する手法を提案し，実験を通してその効果を確かめた．

調査では，音声入力を使用する状況を音声入力の内容と聞き手の関係性の 2 つに着目して分類し，それぞれの分類

に対応するシナリオを 7 つ用意した．各シナリオについて音声入力の使用にどの程度抵抗があるかについてアンケートを通じて調査した．調査の結果，音声入力の内容が聞き手のいない場所での行動に関するものである場合，特に他人との関係に関わるものである場合は，音声入力の使用に恥ずかしさを感じやすいことが分かった．また，聞き手が知らない人である場合にも音声入力を使用する時の恥ずかしさを感じやすい傾向があった．

実験では，音声入力の対象が機械的な見た目の場合と動物のような見た目の場合でユーザーの心情にどのような変化が生じるのかを確かめた．その結果，動物のような見た目の場合の方が音声入力をする時の抵抗や恥ずかしさが軽減される傾向があったが，人前，特に知らない人の前での使用には適さないことが分かった．

今回の実験では，聞き手が友人で聞き手がいない場所での行動に関する入力を行なったが，実験の結果では聞き手がいない場面など，他の場面でも提案手法が有効であることが示唆された．これらの場面についても本当に提案手法が有効なのかを今後確認する必要がある．また，本実験では VR 空間内で音声入力の使用が想定される場面を再現し，被験者は主著者と通話をしながらその様子を録画した動画を見ながら音声入力をするつもりで声を出す方法で実験を行った．通話越しではなく実際に人と対面して音声入力をする場合にどのような感情の変化が生じるかについて確かめる必要がある．被験者の中には，動物のような見た目を好むグループ，機械的な見た目を好むグループ，どちらの場合も変わらないとするグループの 3 つが存在した．今回の実験では被験者の数が少なく，各グループに対して詳しく考察することは困難であった．今後は被験者を増やしこれらの考察を深める必要がある．また，本論文では研究対象を日本人に絞って調査や実験を行った．本論文で得られた知見が日本以外の国の人に対しても当てはまるのかについても，今後確認しなくてはならない．

参考文献

[1] F. Portet, M. Vacher, C. Golanski, C. Roux, B. Meillon. Design and evaluation of a smart home voice interface for the elderly. Pers Ubiquit Comput (2013) 17, 2013, p. 127-144.
[2] “Google Search by Voice: A case study”. https://static.googleusercontent.com/media/research.google.com/ja//pubs/archive/36340.pdf, (参照 2020-10-31).
[3] D. Wigdor, D. Wixon. Brave NUI World: Designing Natural User Interfaces for Touch and Gesture. Elsevier. 2011.
[4] M.H. Cohen, J.P. Giangola, J. Balogh. Voice User Interface Design. Addison-Wesley Professional. 2004.
[5] S.R. Avutu, D. Bhatia, B.V. Reddy. Voice Control Module for Low Cost Local-Map Navigation Based Intelligent Wheelchair. 2017 IEEE 7th International Advance Computing Conference, 2017, p. 609-613.
[6] 山本 大介, 加藤 りか, 田中 亮佑, 高橋 直久. 公共交換での実利用を想定した「しゃべる」バス路線案内システムの提案とその開発. マルチメディア，分散，協調とモバイル (DICOMO2017)シンポジウム, 2017, p. 747-752.
[7] “高輪ゲートウェイ駅の概要について”. https://www.jreast.co.jp/press/2019/20191203_ho01.pdf, (参照 2020-10-31).
[8] “国内のスマートスピーカー普及率は約 6%，提供機能・サービスの拡大が市場成長のカギ”. https://www.dentsudigital.co.jp/release/2019/0218-000164/, (参照 2020-10-23)
[9] “SMART SPEAKER CONSUMER ADOPTION”. https://voicebot.ai/wp-content/uploads/2019/03/smart_speaker_consumer_adoption_report_2019.pdf, (参照 2020-10-23).
[10] “Amazon Echo Dominates Smart Speaker Market in the UK”. https://www.emarketer.com/content/amazon-echo-dominates-smart-speaker-market-in-the-uk, (参照 20202-10-23).
[11] “文字入力による情報検索『面倒』」は半数 「人前での音声検索は『恥ずかしい』」7 割超，家電などの音声操作，4 割が「『自宅に人がいなければ』利用したい”. https://news.kddi.com/kddi/corporate/newsrelease/2017/10/05/besshi2726.html, (参照 2020-10-23)
[12] A.E. Moorthy, K.L. Vu. Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space. International Journal of Human-Computer Interaction, Volume 31, 2015, p. 307-335.
[13] S. Prabhakar, S. Pankanti, A.K. Jain. Biometric recognition: security and privacy concerns. IEEE Security & Privacy, 2003, vol. 1, no. 2, p. 33-42.
[14] Y. Yan, C. Yu1, Y. Shi, M. Xie. PrivateTalk: Activating Voice Input with Hand-On-Mouth Gesture Detected by Bluetooth Earphones. UIST '19: Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, 2019, p. 1013-1020.
[15] K. Sun, C. Yu, W. Shi, L. Liu, Y. Shi. Lip-Interact: Improving Mobile Device Interaction with Silent Speech Commands. UIST '18: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, 2018, p. 581-593.
[16] K. Mase, A. Pentland. Automatic lipreading by optical-flow analysis. Systems and Computers in Japan, Vol. 22, No. 6, 1991, p. 67-76.
[17] M. Fukumoto. SilentVoice: Unnoticable Voice Input by Ingressive Speech. UIST '18: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, 2018, p. 237-246.
[18] 加藤 祐介, 岡田 顕宏. 準投影法的な恥感情尺度の作成. 日本心理学会第 72 回大会, 2008, p. 1051
[19] P.M. Averill, G.J. Diefenbach, M.A. Stanley, J.K. Breckenridge, B. Lusby. Assessment of shame and guilt in a psychiatric sample: a comparison of two measures. Personality and Individual Differences, Volume 32, Issue 8, 2002, p. 1365-1376.
[20] D.R. Cook EdD. Measuring Shame: The Internalized Shame Scale. Alcoholism Treatment Quarterly, Volume 4, Issue 2, 1988, p. 197-215.
[21] K. Goss, P. Gilbert, S. Allan. An exploration of shame measures—I: The other as Shamer scale. Personality and Individual Differences, Volume 17, Issue 5, 1994, Pages 713-717.