

発声支援デバイスの制御を目的とした手のジェスチャに関する社会的受容性の調査

澤野 令¹ 山田 大志² 佐藤 安理紗 ジェンジエラ¹ 戸原 玄² 矢谷 浩司¹

概要：がんの治療などによって声を失った喉頭摘出者が用いる代用発声法の一つに発声支援デバイスを用いた発声法があるが、一部の発声支援デバイスは声に抑揚がなく不自然に聞こえるという欠点がある。過去の研究において複数の抑揚制御の方法が提案されてきたが、デバイスを持つことにより常に片手が塞がれてしまったり、抑揚制御が呼気量に依存してしまうといった欠点がある。これらを解決するために、本研究では新たにジェスチャを用いて声の抑揚を制御するシステムの設計に向けて、発声障がい者が会話している状況下において人々がジェスチャに対して感じる社会的受容性を調査した。その際ジェスチャを行う側と観測する側の双方の立場からの調査を行い、社会的受容性の観点からどのような特徴を持つジェスチャがシステムの入力に相応しいかを考察した。

Investigating Social Acceptability of Hand Gestures for Voice Generation Support Devices

REI SAWANO¹ TAISHI YAMADA² ARISSA J. SATO¹ HARUKA TOHARA² KOJI YATANI¹

1. はじめに

がんの治療などによって喉頭の全摘出手術を受けた喉頭摘出者は、声帯が無いため通常の方法では声を発することができない。しかし声帯以外の構音器官は正常であることから、小型発音体を使用して音源を口腔内に送り込み、通常の発声を行うように口や舌を動かして声道の形状を変化させることによって会話することができる [1, 2]。これらは代用発声法と呼ばれ、代表的なものに電気式人工喉頭と呼ばれるブザー音を発するデバイスを頸部の皮膚に当てることで発声を可能にする EL 発声法がある。これらの発声支援デバイスには声の抑揚がつけられないという欠点がある [3-5]。先行研究 [6] において、デバイスを用いてピッチを制御するシステムが提案されているが、ユーザが常にデバイスを手に持つ必要があるという欠点がある。

よって本研究では、スマートウォッチに対してジェスチャ入力を行うことでデバイスを手に持たずに声のピッチを制御に用いられる可能性がある手のジェスチャの社会的受

容性を検証した。社会的受容性に影響を与え得る要因として、ジェスチャを行う場所においてプライバシーが保障されているか否かと観測者との関係性、ジェスチャを行う人の立場と観測者の立場の違い、ジェスチャを行う人が障がい者であることを観測者が事前に知っているか否かの三つが考えられるため、本研究においては発声障がい者が会話しているという状況下において、これらの要因を考慮した上でジェスチャの社会的受容性を調査し、どのようなジェスチャの特徴が社会的受容性に影響を与えるのかを分析した。結果として一部のジェスチャの特徴が社会的受容性と有意に相関があることが分かり、その原因を考察した。

2. 関連研究

2.1 電気式人工喉頭の制御に関する研究

電気式人工喉頭を用いた発声法 (EL 発声) は、習得・操作が容易であることから代用発声法の一つとして欠かせないものになっている [2]。EL 発声ではブザー音を発する機器を頸部の皮膚に当て原音の代わりとすることで発声を可能にする。他の代用発声法と比べた場合のこの発声法の利点として、呼気量に左右されないため持続発声が可能であり、

¹ 東京大学 Interactive Intelligent Systems Laboratory

² 東京医科歯科大学

人工喉頭を当てる部位の制御ができればほぼ全ての症例で発声可能であるということが挙げられる。欠点としては、機器を持つため片手が塞がってしまうこと、抑揚がつけられないことが挙げられる [3-5]。

EL 発声は声のピッチが常に一定であり、不自然に聴こえるという問題を解決するために、これまでに様々な研究がなされてきた。Tanaka ら [7] は電気式人工喉頭の周波数を会話内容に応じて機械学習を用いて変化させることで、抑揚を自動的に制御するシステムを作った。しかしこの方法はユーザ自身がピッチを制御するわけではないので、発言が意図しないニュアンスで相手に伝わってしまう可能性がある。上見ら [6] は、電気式人工喉頭のピッチの制御方法として、つまみを回す方法、指で圧力センサを押す方法、喉頭摘出者の前頸部に開けた気管孔から得た呼気圧を用いる方法を実験によって比較し、呼気圧を用いたピッチ制御が最も抑揚の自然性が高くなることを示した。しかし EL 発声には呼気量に左右されないため持続発声が可能であるという利点があるため、呼気圧を用いないピッチ制御の方法を探索する意義があるといえる。また、つまみや圧力センサを用いたピッチ制御では、ユーザが常にデバイスを手を持つ必要があるという欠点があるため、デバイスを手に持たずに抑揚を制御する方法を探索する意義があるといえる。本研究ではこれらの発声支援デバイスとスマートウォッチを連携させ、スマートウォッチに対してジェスチャ入力を行うことで声のピッチを制御するインタフェースを提案する。

2.2 ジェスチャの社会的受容性に関する先行研究

ジェスチャの社会的受容性とは、ジェスチャが社会的にどの程度受け入れられやすいかを意味する。あるジェスチャを入力とするインタフェースを作る場合、社会的受容性が高いジェスチャを選択することが重要である。一般的にプロトタイプ作成には多くのコストと時間を要するため、プロトタイプを作成する前にあらかじめ入力に使用するジェスチャの社会的受容性を調査することが重要である [8,9]。ジェスチャの社会的受容性は様々な要因に影響を受けることが過去の研究によって分かっている。

Rico ら [8] は、ジェスチャを行う場所や観測者との関係性が、ジェスチャに対して感じる社会的受容性にどのように影響を与えるのかを調査し、結果として場所に関してはよりプライバシーが保証された空間であるほど、また観測者に関してはより親しい関係性であればあるほど社会的受容性が高い傾向にあると結論づけた。ここでいう観測者とは、ジェスチャを行なっている本人を周囲から見ている人のことを指す。Fouad ら [10] は Head-Worn Displays(HWDs) の入力ジェスチャに対して感じる社会的受容性を、ユーザと観測者の双方の立場から調査し、一部のジェスチャにおいてユーザと観測者が感じる社会的受容性に差異があることを示した。Halley ら [11] は、HWDs のユーザが障がい者

であるという事前知識が HWDs の入力ジェスチャに対して観測者が感じる社会的受容性にどのように影響するのかを調査し、結果としてユーザが障がい者であると実験参加者が認識している場合の方が実験参加者が感じる社会的受容性が高かったことを示した。

本研究では発声障がい者が会話中に使用する際に社会的受容性が高いとされるジェスチャを考察することを目的としているが、この特殊な状況下におけるジェスチャの社会的受容性を調査した研究は過去に無く、ユーザが発声障がい者であること、また会話中であるという要因がジェスチャの社会的受容性に与える影響は不明である。よってこの特殊な状況下に限定した場合におけるジェスチャの社会的受容性を改めて調査する必要があるといえる。

以上の先行研究をまとめるとジェスチャの社会的受容性に影響を与える要因には以下があると分かる。

- ジェスチャを行う場所においてプライバシーが保障されているか否か、観測者との関係性
- ジェスチャを行う人の立場と観測者の立場の違い
- 観測者がジェスチャを行う人が障がい者であることことを事前に知っているか否か

よって上記の要因がジェスチャの社会的受容性に与える影響を考慮しながら調査を行う必要があるといえる。

3. ジェスチャの社会的受容性の調査方針の検討

2章で述べた通り、ジェスチャの社会的受容性に影響を与える要因の一つとして、「ジェスチャを行う場所においてプライバシーが保障されているか否か」と「観測者との関係性」がある。まず本研究では、「ジェスチャを行う場所においてプライバシーが保障されているか否か」に関して、ジェスチャを行う場所のプライバシーが保証されている、つまり周りに知らない人がいない、という状況を想定した。次に、ジェスチャを行う場所のプライバシーが保証されていない、つまり周りに知らない人がいる、という状況を想定した。これら2つの状況下で、ジェスチャの社会的受容性の調査をそれぞれ行うことにした。観測者との関係性に関して Rico ら [8] の研究では、家族、パートナー、友人、同僚、知らない人の5つを観測者として想定している。しかし、本研究で想定しているシステムのユーザは発声障がい者であり、一般に高齢者の割合が高いと推測される。よって、ユーザにパートナーが存在する場合は家族に含まれる可能性が高いと判断し、家族、友人、同僚、知らない人の4つのみを観測者として想定することにした。以上の考察から7つの状況を会話状況として想定し、それぞれにおけるジェスチャの社会的受容性の評価を個別に行うことにした(表1)。なお、知らない人が会話相手である場合は、すでに周りに知らない人がいる状況であると判断した。

表 1: 7つの会話状況. 周りに知らない人がいるかと誰と会話しているかの二つのパラメータで決定される.

周りに他人/会話相手	家族	友人	同僚	知らない人
いる (Y)	YFa	YFr	YC	YS
いない (N)	NFa	NFr	NC	—

3.1 ジェスチャリストの作成

調査の事前準備として, 社会的受容性を測定するジェスチャのリストを作成した(表 2). 本研究では, スマートウォッチを用いてジェスチャを検出する想定であるため, 指や手首, 腕を使ったジェスチャのみをリストに含めた. 最初に, これらの部位を使った入力ジェスチャに関する過去の HCI の研究 [12–20] を参考に暫定のリストを作成した. その後, 調査するジェスチャの種類が偏らないように, 似ているジェスチャを除いた. また研究室のメンバーと議論する過程で考えたジェスチャを加え, 合計 16 個を最終的にリストに含めた.

表 2: 16 個のジェスチャとラベリング結果

名前	ラベル
指を曲げ伸ばしする動作	FB
手を縦に上下させる動作	HM
両手を近づけたり遠ざけたりする動作	HR
手首から先の上下運動	WM
親指を上下させる動作	TU
拳を上下させる動作	FM
手の甲をさする動作	RB
デバイスをスワイプする動作	DS
手を握ったり開いたりする動作	HG
手を回転させる動作	HRM
肘から先を左右に動かす動作	AR
手首を回転させる動作	WR
つまむ動作	PI
片方の手の平に片手の拳を当てる動作	HP
肘から先を倒す動作	EM
手を体の前で上下させる動作	HV

4. 観測者の立場からの社会的受容性の調査

4.1 ジェスチャを行う人が障がい者であるかどうかの事前知識についての検討

2章で述べた通りジェスチャの社会的受容性に影響を与える要因の一つとして, 「観測者がジェスチャを行う人が障がい者であることことを事前知っているか否か」がある. よって本研究では「ジェスチャを行っている人が発声支援デバイスを用いて会話をしており, このジェスチャを使用して声のピッチを制御している」という条件を実験参加者に提示した場合と, しなかった場合におけるジェスチャの社会的受容性の測定を行うことにした.

4.2 調査方法・手順

一般に調査を行う方法としてクラウドソーシングを用いる方法や Zoom, 対面で行う方法が考えられる. 対面での調査が最も実験参加者への指示がしやすいと考えられるが, Alallah ら [21] は HWDs の入力ジェスチャに対して感じる社会的受容性を, クラウドソーシングを用いた調査方法と研究室で実験参加者と対面で行う調査方法の二つを用いて測定し, その結果クラウドソーシングを用いた場合でも十分に有効な結果が得られることを示した. 本研究においても, 想定しているデバイスは異なるものの, クラウドソーシングを用いた場合でも十分に有効な結果が得られると推測し, またより多くのデータを集めることで, データの信憑性を高めることができると考えたため, クラウドソーシングを選択した. また本調査参加者の条件は 18 歳以上の日本人とし, アンケート調査は以下の流れで実施した.

国内のクラウドソーシングサービスである CrowdWorks に 1 回の所要時間が約 7 分と想定される Microsoft Forms のアンケートをタスクとして投稿した. 調査対象の条件を満たすワーカーがアンケートを回答することができ, 回答を持って調査への参加を同意したものとみなした. 回答が確認できたワーカーには謝金として 100 円を支払った.

4.3 ジェスチャの社会的受容性を問う質問の作成

ジェスチャの社会的受容性を問う質問は Rico ら [8] が行った Web 調査を参考にし以下の要素で構成した.

- ジェスチャの名前とジェスチャの短い説明文
- ジェスチャを行っている動画
- 質問文
- 回答欄

実験参加者はまずジェスチャの名前とジェスチャの短い説明文を読み, その後ジェスチャを行っている動画を視聴し, 最後に質問に回答するという流れになる. ジェスチャの動画に関して Alallah ら [10] の研究によれば, ジェスチャを行う人の性別がジェスチャの社会的受容性に影響を与える可能性があるため, 研究室の女性メンバー 1 人と研究実施者本人のそれぞれが研究室内の同じ場所でジェスチャを行なっている動画を撮影し, 続けて再生されるようにした. Rico ら [8] の Web 調査を参考にしジェスチャは 3 回繰り返して行う様子を撮影した. またジェスチャを行っている人の表情が社会的受容性に与える影響する可能性を考慮し, 動画はマスクをした状態で撮影した. 質問文は「以下の場面・状況において, 相手の方がこの動作をしていたとき, あなたはどの程度違和感を感じますか」とし, 「全く違和感がない, 違和感はない, どちらとも言えない, 違和感がある, 強い違和感がある」の 5 段階のリッカート尺度を用いて評価してもらう形にした. また「ジェスチャを行っている人が発声支援デバイスを用いて会話をしており, このジェスチャを使用して声のピッチを制御している」という条件を

提示した場合としなかった場合におけるジェスチャの社会的受容性の測定を行うために、二種類のフォームを用意して一方のフォームでは全ての質問文の最後にこの条件を説明する文を加えた。

4.4 アンケートフォーム全体の内容

調査で回答してもらったアンケート全体は以下の要素で構成した。

- 性別 (選択式)
- 年齢 (選択式)
- CrowdWorks の表示名 (記述式)
- ジェスチャの社会的受容性を問う質問 × 4 (選択式)
- ダミー用のジェスチャの社会的受容性を問う質問 (選択式)

またクオリティチェックを行うために、明らかに社会的受容性が低いと予想されるダミー用のジェスチャを用意し、そのジェスチャの社会的受容性を問う質問を全てのフォームに含めた。この際ダミー用のジェスチャには、口の中に手を出し入れするジェスチャを使用した。クオリティチェックの基準としては、「知らない人と会話している」という条件下で、ダミー用のジェスチャに対し「全く違和感がない」と評価した実験参加者の回答を全て分析対象から除外することにした。また学習効果の影響を緩和するために、各フォームのジェスチャとダミー用のジェスチャの社会的受容性を問う質問は、ランダム順番で表示されるようにした。

4.5 得られたデータの分布

前述したクラウドソーシングサービスである CrowdWorks を通じて合計 934 件を取得した。その中から、以下の条件を満たすものを抽出し、以降の分析・考察に用いることにした。

- 複数のフォームに重複して回答していないこと。
- 全て同じ選択肢を選んでいないこと。
- クオリティチェックに抵触していないこと。

学習効果の影響を排除するために、「複数のフォームを重複して回答していないこと」を条件に加えた。他のフォームと重複して回答しているもの 205 件、クオリティチェックに抵触したものが 9 件、全て同じ回答のものが 18 件あったのでそれらを除外し、8つのフォーム合わせて合計 702 件の有効データを取得した。

4.6 重回帰分析

まずジェスチャのどのような特徴が社会的受容性に影響を与えるのかを調べるために、各ジェスチャに対して特徴ごとにラベル付けを行った。ラベルは動きの種類に基づき表3の最左列の6つを作成した。その後16個のジェスチャに対してラベル付けを行った(表3)。

このラベルを説明変数、社会的受容性を目的変数として、

重回帰分析を行った。「ジェスチャを行っている人が発声支援デバイスを用いて会話をしており、このジェスチャを使用して声のピッチを制御している」という条件を提示した場合としなかった場合におけるデータが得られたため、それぞれに対して分析を行った。

条件を提示した場合の重回帰分析の結果は表4のようになった。有意水準5%で検定した結果、どの特徴ラベルも社会的受容性に有意に影響を与えていないという結果になった。

条件を提示しなかった場合の重回帰分析の結果は表5のようになった。有意水準5%で検定した結果、手を動かす、手を前に構える、両手を使用する、の3つの特徴が社会的受容性に有意に影響を与えており、手を動かす特徴は正の相関、手を前に構える特徴は負の相関、両手を使用する特徴は正の相関があることが分かった。

4.7 指を使うジェスチャ同士の比較

16個のジェスチャリスト(表2)の中で、指を動かす特徴と手を前に構える特徴の二つを持つジェスチャが4つ(FB, TU, HG, PI)存在し、それらの社会的受容性にどのように差異があるのかを検証した。まず一元分散分析によって、各動作の社会的受容性の平均値間に差異があるのかを有意水準5%で検定した。結果的に条件を提示した場合はp値が0.194となったため有意差は存在せず、条件を提示しなかった場合はp値が0.05を下回ったため、動作間に有意差が存在するという結果になった。

また条件を提示しなかった場合のデータに対して指を使う動作の社会的受容性の平均値の対比較を行った(表6)。

また、FB, TU, HG, PIの社会的受容性の平均値は、それぞれ2.64, 1.83, 2.25, 2.54であった。この値は5段階のリッカート値である。以上の結果から、「指を曲げ伸ばしする動作」は「親指を上下させる動作」と「手を握ったり開いたりする動作」より有意に社会的受容性が高く、「つまむ動作」は「親指を上下させる動作」より有意に社会的受容性が高いことが分かった。

4.8 条件を提示した場合としなかった場合における社会的受容性の比較

次に条件を提示することが社会的受容性に与える影響を検証するために、条件を提示した場合としなかった場合におけるデータの比較を行った。両者の場合における社会的受容性の平均値を、同じジェスチャ同士で比較した(図1)。

各ジェスチャにおいて、条件を提示した場合としなかった場合における社会的受容性の平均値間に有意差が存在するのかを判定するために、条件を提示したか否かと会話状況の種類を因子として、有意水準5%で繰り返しのない二元配置分析を行った。例としてG1, G2の検定結果を示した(表7, 表8)。

表 3: ジェスチャのラベリングの様子. 行カラムがジェスチャのラベルで列カラムがジェスチャを表している.

	FB	HM	HR	WM	TU	FM	RB	DS	HG	HRM	AR	WR	PI	HP	EM	HV
指を動かす	o				o			o	o				o			
手を動かす				o								o				
手首を回転させる										o						
腕を動かす		o	o			o	o			o	o			o	o	o
手を前に構える	o	o	o	o	o	o			o	o	o	o	o	o	o	
両手を使用する			o				o	o						o		

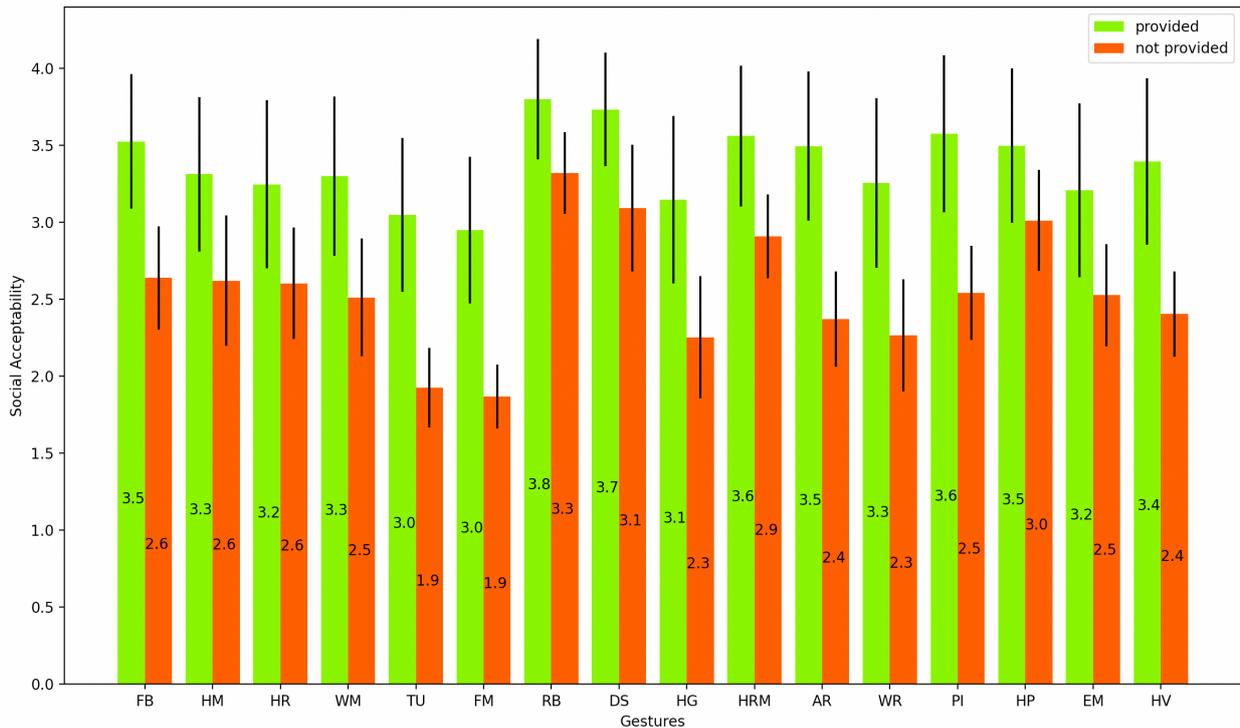


図 1: 各ジェスチャに関して, 条件を提示した場合と提示しなかった場合における社会的受容性の平均値を比較したグラフ. 緑が条件を提示した場合, 赤が条件を提示しなかった場合を示している. エラーバーは標準偏差を表している.

表 4: 条件を提示した場合のデータに対して, 動作の種類を説明変数, 社会的受容性を目的変数として重回帰分析を行った結果

特徴ラベル	係数	std	t 値	p 値
指を動かす	0.2906	0.314	0.927	0.356
手を動かす	0.2791	0.222	1.259	0.211
手首を回転させる	0.2339	0.364	0.643	0.521
腕を動かす	0.2610	0.288	0.905	0.367
手を前に構える	-0.3794	0.213	-1.782	0.078
両手を使用する	0.0907	0.169	0.538	0.592

表 5: 条件を提示しなかった場合のデータに対して, 動作の種類を説明変数, 社会的受容性を目的変数として重回帰分析を行った結果

特徴ラベル	係数	std	t 値	p 値
指を動かす	0.3409	0.238	1.430	0.156
手を動かす	0.5342	0.169	3.169	<0.05
手首を回転させる	0.2888	0.276	1.045	0.299
腕を動かす	0.3975	0.219	1.814	0.072
手を前に構える	-0.4276	0.162	-2.641	<0.05
両手を使用する	0.4329	0.128	3.380	<0.05

結果, 全てのジェスチャにおいて, 一つ目の因子 (条件を提示したか否か) の p 値が閾値を下回っていたため, どのジェスチャにおいても, 条件を提示したか否かによって社会的受容性が有意に異なると結論づけた. また, 条件を提示した場合の方が社会的受容性が高いという結果になった.

5. ユーザの立場からの社会的受容性の調査

5.1 調査方法・手順

本研究で募集した実験参加者は, 高齢者の割合が高いため, オンライン上でフォームに回答することが難しいと判

表 6: 条件を提示しなかった場合のデータに対して、指を使う動作の社会的受容性の平均値間で一対比較を行った結果。

	FB	TU	HG	PI
FB		<0.05	<0.05	0.611
TU			0.115	<0.05
HG				0.183
PI				

表 7: G1 に関して条件を提示したか否かと会話状況を因子とした二元配置分析の結果。右端の列はそれぞれの因子の p 値を示しており、0.05 を下回ったセルを色付けしている。

	sum_sq	df	F 値	p 値
条件提示の有無	250.3	1.0	211.4	<0.001
会話状況	198.7	6.0	27.9	<0.001

表 8: G2 に関して条件を提示したか否かと会話状況を因子とした二元配置分析の結果。右端の列はそれぞれの因子の p 値を示しており、0.05 を下回ったセルを色付けしている。

	sum_sq	df	F 値	p 値
条件提示の有無	152.1	1.0	133.2	<0.001
会話状況	281.8	6.0	41.1	<0.001

断し、調査は対面で行うことにした。また本調査における実験参加者の条件は、発声障がいによって声を発することができない人とし、東京医科歯科大学病院の医師を通じて募集した。アンケート調査は、実験参加者にフォームの内容が表示された PC の画面を見せながら、口頭で説明する形で行った。実験参加者には、選択肢を指差すか、発声支援デバイスをを用いて口頭で回答してもらった。調査を終了した段階で、謝礼として一人目の実験参加者には 500 円分、二人目以降の実験参加者には 1000 円分の図書カードを渡した。また調査中は常に実験参加者の担当医が同席していた。

5.2 ジェスチャの社会的受容性と使いやすさを問う質問の作成

ジェスチャの社会的受容性を問う質問は、クラウドソーシングで使用した質問とほぼ同じ構成にした。実験参加者には、まずジェスチャを行っている動画を視聴してもらい、その後動画内のジェスチャを 3 回程度模倣してもらった。その際、正しくできていない場合は口頭で修正を促した。その後そのジェスチャに関する質問に回答してもらった。ジェスチャの動画は、調査の一貫性を保つためにクラウドソーシングの際に使用したものと同一ものを用いた。社会的受容性を問う質問文は、「以下の場面・状況において、動画内の動作を用いて会話中の声の高さを調整する時、あなたはどの程度抵抗感を感じますか」とし、「全く感じない、感じない、どちらともいえない、感じる、強く感じる」の 5 段階のリッカート尺度を用いて評価してもらった。使い

表 9: 条件を提示した場合のデータに対して、動作の種類のラベルを説明変数、社会的受容性を目的変数として重回帰分析を行った結果

特徴ラベル	係数	std	t 値	p 値
定数	4.2822	0.447	9.573	0.000
指を動かす	0.0296	0.371	0.080	0.936
手を動かす	0.1494	0.262	0.569	0.570
手首を回転させる	0.2447	0.430	0.568	0.571
腕を動かす	-0.6786	0.341	-1.989	<0.05
手を前に構える	-1.2888	0.252	-5.113	<0.05
両手を使用する	-0.7970	0.199	-3.996	<0.05

やすさを問う質問文は、「動画内の動作を用いて会話中の声の高さを調整することを想定します。どの程度使いやすと思いますか」とし、非常に使いやすい、使いやすい、どちらともいえない、使いにくい、非常に使いにくい、の 5 段階のリッカート尺度を用いて評価してもらった。

5.3 アンケートフォーム全体の内容

調査で回答してもらったアンケート全体は、以下の要素で構成した。

- 性別 (選択式)
- 年齢 (選択式)
- ジェスチャの社会的受容性と使いやすさを問う質問 × (6 個か 16 個) (選択式)

一人目の実験参加者に対して使用したフォームには、16 個のジェスチャのうち 6 個のジェスチャに関する質問を含め、二人目以降の実験参加者に対して使用したフォームには、16 個全てのジェスチャに関する質問を含めた。また学習効果の影響を緩和するために、各フォームに含まれるジェスチャはランダム順番で表示されるようにした。

5.4 重回帰分析

観測者の場合と同様に、動作の特徴ラベルを説明変数、社会的受容性を目的変数として重回帰分析を行った (表9)。

有意水準 5% で検定した結果、腕を動かす、手を前に構える、両手を使用する、の 3 つの特徴が社会的受容性に有意に影響を与えており、腕を動かす特徴は負の相関、手を前に構える特徴は負の相関、両手を使用する特徴は負の相関があることが分かった。

5.5 指を使うジェスチャ同士の比較

観測者の立場からのデータ分析と同様に、指を使う 4 つのジェスチャ (FB, TU, HG, PI) 同士の社会的受容性の比較を行った。まず一元分散分析によって、各動作の社会的受容性の平均値間に差異があるのかを有意水準 5% で検定した。結果的に p 値が 0.05 を下回ったため、動作間に有意差が存在するという結果になった。また指を使う動作の社会的受容性の平均値の一対比較を行った (表10)。

表 10: ユーザの立場からのデータに対して指を使う動作の社会的受容性の平均値間で一対比較を行った結果.

	FB	TU	HG	PI
FB		<0.05	<0.05	<0.05
TU			1.0	0.597
HG				0.621
PI				

また FB, TU, HG, PI の社会的受容性の平均値は、それぞれ 3.61, 2.86, 2.86, 2.68 であった。この値は 5 段階のリッカート値である。以上の結果から、「指を曲げ伸ばしする動作」は「親指を上下させる動作」と「手を握ったり開いたりする動作」と「つまむ動作」より有意に社会的受容性が高いことが分かった。

6. 考察

6.1 障がいに関する事前知識がジェスチャの社会的受容性に与える影響

観測者の立場からのジェスチャの社会的受容性の調査結果では、全てのジェスチャにおいて、条件を提示した場合の方が提示しなかった場合に比べて社会的受容性が高いという結果が得られた。これは HWDs のユーザが障がい者であると認識している場合の方が、ジェスチャの社会的受容性が高かったと示している Halley ら [11] の研究と同様の結果になったといえ、原因としてジェスチャの使用理由を知っている方が、観測者がユーザに不信感を抱きにくいという可能性が考えられる。また観測者の立場からのデータの重回帰分析の結果では社会的受容性に有意に影響を与えるパラメータが存在しないという結果になったが、これも同様にユーザが障がい者であると認識していることによって、どのような動きであっても観測者は許容する傾向があったという理由が考えられる。以上を踏まえると、ジェスチャの使用理由が相手に明確に伝わるようなシステム設計にしたり、使用理由を事前に相手に伝える工夫をすることによって観測者が感じる社会的受容性を高めることができると考えられる。

6.2 社会的受容性に影響を与えるジェスチャの特徴

腕を動かす特徴と両手を動かす特徴

本研究において、ユーザの立場からの社会的受容性は腕を動かす特徴と両手を使用する特徴と負の相関があるという結果になった。Ahlstrom らの研究 [22] によれば、動く範囲が大きいジェスチャの方が動く範囲が小さいジェスチャに比べて、ユーザが感じる社会的受容性が低かったという結果を示している。腕を動かす動作や両手を使用する動作は、指先や手首のみを使用するジェスチャに比べてジェスチャが動く範囲が大きいので、ユーザが目立っていると考えた可能性が原因として考えられる。Alallah らの研究 [10] では、

ジェスチャを行う側と観測する側では視点が違うため、動きの小さいジェスチャはジェスチャを行う側にはあまり目立っていないと感じられるが、観測者にとっては目立っていると感じられる可能性があるとしている。同様に考えると、動きの大きいジェスチャはジェスチャを行う側には目立っていると感じられるが、客観的に見るとあまり目立っていないケースも有り得ると考えられ、これも原因の一つとして考えられる。

手を前に構える特徴

また観測者とユーザのいずれの立場においても、手を前に構えるという特徴が社会的受容性と負の相関があった。会話は話している相手の顔を向いていることが多いと考えられるため、ジェスチャが顔の近い位置で行われているほど観測者の視界に入りやすく、目立って感じられたという可能性が原因として考えられる。また今回調査に使用したジェスチャは、手を前に構えるジェスチャの方がジェスチャの動く範囲が多い傾向にあったことも原因として考えられる。

6.3 指を使ったジェスチャ同士の社会的受容性の差異

観測者の立場からのデータにおいては、「親指を上下させる動作」は「指を曲げ伸ばしする動作」と「つまむ動作」より有意に社会的受容性が低いという結果になった。Rico らの研究 [8] では、すでに他の意味を持っているジェスチャは意味が状況の文脈と噛み合わなかったり、相手を困惑させる可能性があるため、モバイルインタフェースの入力には相応しくないと考察しているが、本研究においても「親指を上下させる動作」の一部である「親指を上げる動作」は、日常生活において相手に賛成しているという意味を持ち得るため、会話中にこれを頻繁に繰り返すと会話の文脈に合わず相手を困惑させたり、皮肉に受け取られてしまう可能性があるということが原因として考えられる。またユーザの立場からのデータにおいては、「指を曲げ伸ばしする動作」は「親指を上下させる動作」と「手を握ったり開いたりする動作」と「つまむ動作」より有意に社会的受容性が高いことが分かった。「指を曲げ伸ばしする動作」は他の動作に比べて指を一本しか動かさないの、ユーザは動きが目立ちにくいと感じた可能性が原因として考えられる。

7. 制約

収集データの規模に関する制約

本研究で調査したジェスチャは 16 個のみであったため、より多くのジェスチャに対して調査を行うことでより正確な分析が可能であると考えられる。また本研究ではアンケート調査を行ったが、インタビュー調査によって一部のジェスチャの特徴が社会的受容性に影響を与える理由についてより詳細に議論することができると考えられる。また本研究で調査したユーザは 4 名のみであったため、より多くの

ユーザを募集して調査を行うことで、より正確なデータを得られると考えられる。本研究では、観測者の立場からのジェスチャの社会的受容性の調査の参加者が20代から60代に偏っており、より幅広い年齢層の実験参加者に対して調査することでより正確なデータを得られると考えられる。

調査方法に関する制約

本研究では Rico ら [8] の先行研究を参考に、あえて見えやすいようにジェスチャの動画を撮影したが、実際はジェスチャを見えやすいように行う必然性はないため、体の横や後ろなどの見えにくい位置でジェスチャを行った場合を含めて比較調査することで、より実際の使用時に近いジェスチャの社会的受容性の調査を行うことができると考えられる。また本研究では会話中にジェスチャを行うことを想定しているため、会話音声を含めて動画を撮影することによって、より実際の使用時に近いジェスチャの社会的受容性の調査を行うことができると考えられる。また本研究では、観測者の立場からのジェスチャの社会的受容性の調査に使用したアンケートフォームにおいて、質問項目の読みやすさを重視し、7つの会話状況を全ての実験参加者に対して同じ順序で提示した。これによって得られたデータに順序効果が生じている可能性があり、より正確な調査をするためには、実験参加者によって会話状況の順序を入れ替える必要があると考えられる。

8. まとめと今後の展望

本研究では、ジェスチャによって声のピッチを制御するインタフェースを設計するための事前調査として、発声障がい者が会話しているという状況下におけるジェスチャの社会的受容性を調査した。結果として、一部のジェスチャの特徴が社会的受容性と有意に相関があることが分かった。本研究の知見に基づき、社会的受容性が高いと考えられるジェスチャを入力とするシステムを実装することが今後の展望として考えられる。

謝辞

本研究を進めるにあたり、研究生生活を支えてくださった研究室のメンバーに、心より感謝申し上げます。

参考文献

- [1] 八幡英子, 伊福部達. 音声のピッチパターンを考慮した電気人工喉頭. 音声言語医学, Vol. 30, No. 4, pp. 309–315, 1989.
- [2] 橋場参生, 山本寧, 山口悦範, 須貝保徳, 上見憲弘, 伊福部達. 音声の自然性を備えた電気人工喉頭の開発研究 (第2報). 北海道立工業試験場報告, Vol. 294, pp. 143–150, 1995.
- [3] 富藤雅之. ボイスプロステシス. 日本気管食道科学会会報, Vol. 71, No. 3, pp. 283–284, 2020.
- [4] Mitsuo Hashiba, Yasunori Sugai, Takashi Izumi, Shuichi Ino, and Tohru Ifukube. Development of a wearable electro-larynx for laryngectomees and its evaluation. In *2007 29th Annual*

- International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5267–5270. IEEE, 2007.
- [5] 藪謙一郎, 伊福部達. 拡声器能を備えたウェアラブル電気式人工喉頭の設計. 日本バーチャルリアリティ学会論文誌, Vol. 21, No. 2, pp. 295–301, 2016.
- [6] 上見憲弘, 伊福部達, 高橋誠, 松島純一. ピッチ周波数制御型人工喉頭の提案とその評価. 電子情報通信学会論文誌 D, Vol. 78, No. 3, pp. 571–578, 1995.
- [7] Kou Tanaka, Tomoki Toda, Graham Neubig, Sakriani Sakti, and Satoshi Nakamura. An enhanced electrolarynx with automatic fundamental frequency control based on statistical prediction. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*, pp. 435–436, 2015.
- [8] Julie Rico and Stephen Brewster. Usable gestures for mobile interfaces: evaluating social acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 887–896, 2010.
- [9] Julie Rico. Evaluating the social acceptability of multimodal mobile interactions. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pp. 2887–2890, 2010.
- [10] Fouad Alallah, Ali Neshati, Yumiko Sakamoto, Khalad Hasan, Edward Lank, Andrea Bunt, and Pourang Irani. Performer vs. observer: whose comfort level should we consider when examining the social acceptability of input modalities for head-worn display? In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–9, 2018.
- [11] Halley Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K Kane. The at effect: how disability affects the perceived social acceptability of head-mounted display use. In *proceedings of the 2016 CHI conference on human factors in computing systems*, pp. 4884–4895, 2016.
- [12] Jun Gong, Xing-Dong Yang, and Pourang Irani. Wristwhirl: One-handed continuous smartwatch input using wrist gestures. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 861–872, 2016.
- [13] Yi-Ta Hsieh, Antti Jylhä, Valeria Orso, Luciano Gamberini, and Giulio Jacucci. Designing a willing-to-use-in-public hand gestural interaction technique for smart glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 4203–4215, 2016.
- [14] Jess McIntosh, Charlie McNeill, Mike Fraser, Frederic Kerber, Markus Löchtefeld, and Antonio Krüger. Empress: Practical hand gesture classification with wrist-mounted emg and pressure sensing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 2332–2342, 2016.
- [15] Jess McIntosh, Asier Marzo, Mike Fraser, and Carol Phillips. Echoflex: Hand gesture recognition using ultrasound imaging. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1923–1934, 2017.
- [16] Yasha Iravantchi, Mayank Goel, and Chris Harrison. Beam-band: Hand gesture sensing with ultrasonic beamforming. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–10, 2019.
- [17] Hongyi Wen, Julian Ramos Rojas, and Anind K Dey. Serendipity: Finger gesture recognition using an off-the-shelf smartwatch. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pp. 3847–3851, 2016.
- [18] Laura-Bianca Bilius and Radu-Daniel Vatavu. Demonstration of gesturing, a web tool for ring gesture input. In *The Adjunct Publication of the 34th Annual ACM Symposium on User Interface Software and Technology*, pp. 124–125, 2021.
- [19] Yiqin Lu, Bingjian Huang, Chun Yu, Guahong Liu, and

Yuanchun Shi. Designing and evaluating hand-to-hand gestures with dual commodity wrist-worn devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 4, No. 1, pp. 1–27, 2020.

- [20] William Delamare, Chaklam Silpasuwanchai, Sayan Sarcar, Toshiaki Shiraki, and Xiangshi Ren. On gesture combination: An exploration of a solution to augment gesture interaction. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, pp. 135–146, 2019.
- [21] Fouad Alallah, Ali Neshati, Nima Sheibani, Yumiko Sakamoto, Andrea Bunt, Pourang Irani, and Khalad Hasan. Crowdsourcing vs laboratory-style social acceptability studies? examining the social acceptability of spatial user interactions for head-worn displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–7, 2018.
- [22] David Ahlström, Khalad Hasan, and Pourang Irani. Are you comfortable doing that? acceptance studies of around-device gestures in and for public settings. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, pp. 193–202, 2014.