

修士論文

ろう・難聴者に最適な 会話型ナチュラルユーザインタフェースの検討

令和3年度

筑波技術大学大学院修士課程技術科学研究科

産業技術学専攻

加藤 隆司

目次

第 1 章 序論	1 -
1.1 研究背景	1 -
1.1.1 ユーザインタフェースの進化とユーザ体験	1 -
1.1.2 ろう・難聴者ユーザ視点に立った NUI 設計	3 -
1.1.3 会話型ユーザインタフェース	5 -
1.1.4 手話インタフェース関連の行動喚起	7 -
1.1.5 ユーザインタフェースのガイドライン	9 -
1.2 関連研究	12 -
1.2.1 スマートスピーカーの普及状況と課題	12 -
1.2.2 ジェスチャインタフェースの活用事例と課題	15 -
1.2.3 Wizard of Oz 法	17 -
1.3 研究目的	18 -
RQ1:光を用いた応答のユーザビリティ	19 -
RQ2:手話と文字の最適な表示方法	20 -
RQ3:ウェイクアップにおける注視の有効性	20 -
RQ4:フィードバックの必要性	21 -
RQ5:ウェイクアップのより良い方法	21 -
RQ6:フィードバックのより良い提示方法	21 -
1.4 本論文の構成	22 -
第 2 章 光通知と注視を用いた手話会話型ユーザインタフェースの検討	23 -
2.1 検証する研究課題	23 -
2.2 実験 1 の方法	24 -
2.2.1 参加者	24 -
2.2.2 システムの構築	26 -
2.2.3 実験 1 の環境	29 -
2.2.4 評価項目	31 -
2.3 結果	33 -
2.3.1 System Usability Scale	33 -
2.3.2 反応時間	34 -
2.3.3 手話と文字に対する必要性	35 -
2.3.4 注視入力	36 -
2.4 考察	38 -
RQ1:光を用いた応答のユーザビリティ	38 -
RQ2:手話と文字の最適な表示方法	39 -
RQ3:ウェイクアップにおける注視の有効性	39 -

第 3 章 ウェイクアップ方法とフィードバック方法の検討	40 -
3.1 検証する研究課題	40 -
3.2 実験 2 の方法	40 -
3.2.1 参加者	40 -
3.2.2 実験条件	42 -
3.2.3 システムの構築	44 -
3.2.4 実験 2 の環境	47 -
3.2.5 評価項目	49 -
3.3 結果	54 -
3.3.1 フィードバックの有無が及ぼす影響の解析結果	54 -
3.3.2 満足度の解析結果	56 -
3.3.3 順位付け解析結果	62 -
3.3.4 参加者の挙動の解析結果	67 -
3.3.5 システムの挙動の解析結果	73 -
3.3.6 主観評価の解析結果	75 -
3.4 考察	78 -
RQ4:フィードバックに対する意識	78 -
RQ5:ウェイクアップのより良い方法	78 -
RQ6:フィードバックのより良い提示方法	80 -
第 4 章 ろう・難聴者に最適な会話型 NUI 設計ガイドラインの提案	81 -
第 5 章 結論	83 -
5.1 まとめ	83 -
5.2 今後の課題	84 -
参考文献	85 -
研究業績	90 -
謝辞	91 -
付録 A 実験 1 の参加者マニュアル	94 -
付録 B 実験 1 のアンケート	94 -
付録 C 実験 1 の結果	98 -
付録 D 実験 2 の参加者マニュアル	94 -
付録 E 実験 2 のアンケート	104 -
付録 F 実験 2 の結果	106 -

図目次

図 1 : ユーザインタフェースの進化	- 1 -
図 2 : ユーザインタフェース設計工程	- 3 -
図 3 : CUI, VUI とその他のシステムとの関係	- 5 -
図 4 : スマートスピーカーの国別出荷台数予測	- 12 -
図 5 : スマートスピーカーの主要 4 社	- 13 -
図 6 : Echo Show シリーズ	- 14 -
図 7 : Pixel4 のジェスチャ機能「Motion Sense」	- 15 -
図 8 : Wizard of Oz 法を用いた実験配置	- 17 -
図 9 : 聴者ユーザとシステムの会話プロセスの様式	- 18 -
図 10 : 実験 1 の参加者(N=12)の特性	- 25 -
図 11 : スマートスピーカーの利用機能	- 26 -
図 12 : 実験 1 で構成するシステムの外観および挙動内容	- 27 -
図 13 : 実験 1 で構築するシステムが実行する入出力内容 (タスク別)	- 28 -
図 14 : 実験 1 の配置図	- 29 -
図 15 : 光提示あり条件における反応時間の特定方法	- 32 -
図 16 : 各参加者(N=12)の SUS 得点	- 33 -
図 17 : システムからの応答に対する反応時間	- 34 -
図 18 : 手話と文字それぞれに対する必要性のアンケート結果	- 35 -
図 19 : ニュース確認のコマンドを入力する時の各参加者の動作継続時間	- 37 -
図 20 : 実験 2 の参加者(N=24)の特性	- 41 -
図 21 : 実験 2 で構築するシステムの外観	- 44 -
図 22 : 実験 2 で構築するシステムの挙動内容	- 45 -
図 23 : 実験 2 で構築するシステムが実行する入出力内容	- 46 -
図 24 : 実験 2 の配置図	- 47 -
図 25 : I3×O2 のときの参加者とシステムの会話時間	- 49 -
図 26 : I2×O3 条件の時の参加者挙動に関する解析時間	- 51 -
図 27 : I1×O1 条件の時の参加者挙動に関する解析時間	- 51 -
図 28 : I2×O3 条件の時のシステム挙動に関する解析時間	- 52 -
図 29 : I1×O1 条件の時のシステム挙動に関する解析時間	- 52 -
図 30 : フィードバックの個別満足度の平均値	- 54 -
図 31 : フィードバックの個別ウェイクアップ終了からコマンド入力開始までの時間平均値	- 55 -
図 32 : 各 9 条件の満足度平均値	- 56 -
図 33 : 補聴器/人工内耳の装用に関する特性別の満足度の平均値	- 57 -
図 34 : 手話レベルに関する特性別の満足度の平均値	- 57 -
図 35 : 会話時の声使用に関する特性別の満足度の平均値	- 58 -
図 36 : アイディンティティに関する特性別の満足度の平均値	- 58 -

図 37 : ウェイクアップ方法の順位結果	- 63 -
図 38 : フィードバック方法の順位結果	- 64 -
図 39 : 手話レベル特性別のウェイクアップ条件の満足度順位結果	- 65 -
図 40 : 手話レベル特性別のフィードバック条件の満足度順位結果	- 66 -
図 41 : フィードバックそれぞれの手振り動作時間の平均値	- 67 -
図 42 : 参加者の各属性における手振り動作時間の平均値	- 68 -
図 43 : 9 条件毎のウェイクアップ終了-コマンド入力開始の時間平均値	- 70 -
図 44 : 9 条件毎のシステムの反応開始 (変化終了) からコマンド入力開始までの時間平均値 ...	- 71 -
図 45 : 9 条件毎のシステムの反応開始からコマンド入力開始までの時間平均値	- 72 -
図 46 : 9 条件毎のウェイクアップ開始からシステムの反応開始(変化開始)までの時間平均値.....	- 73 -
図 47 : 9 条件毎のウェイクアップ終了からシステムの反応開始(変化開始)までの時間平均値.....	- 74 -
図 48 : フィードバックに対する意識の調査結果 (N=24).....	- 76 -
図 49 : サインネームおよび手振りの動作中における注視の意識調査結果(N=24).....	- 77 -
図 50 : 本研究で提案するろう・難聴者ユーザとシステムとの会話プロセスの様式	- 81 -

表目次

表 1 : NUI の主な操作手段とその活用事例	- 2 -
表 2 : 学際的なワークショップで促された行動喚起の概要	- 7 -
表 3 : ヤコブの 10 ユーザビリティヒューリスティックス	- 9 -
表 4 : ユーザー中心デザインの 7 つの原則	- 9 -
表 5 : インタフェースデザインの 8 つの黄金律	- 10 -
表 6 : Alexa デザインガイド	- 10 -
表 7 : Google アシスタントの会話デザインに関するガイドライン	- 10 -
表 8 : 本研究で示す研究課題	- 19 -
表 9 : ユーザとシステムが会話する時の相互入出力モダリティ	- 23 -
表 10 : 色光それぞれの特徴	- 27 -
表 11 : ユーザが実験 1 中に実行可能なコマンド	- 30 -
表 12 : 実験 1 のプログラム	- 30 -
表 13 : System Usability Scale のアンケート項目	- 31 -
表 14 : コマンド入力前に注視した割合と注視開始からコマンド入力開始までの時間	- 36 -
表 15 : ウェイクアップ条件	- 43 -
表 16 : フィードバック条件	- 43 -
表 17 : 実験 2 プログラム	- 48 -
表 18 : ウェイクアップとフィードバックの満足度を要因とした分散分析結果（特性別）	- 59 -
表 19 : 手振り条件の満足度が他の条件に比べて有意に高くなった参加者特性	- 61 -
表 20 : 挙動順番の構成比と挙動開始時刻の差	- 69 -
表 21 : 注視と手振り動作の両方を行った時間	- 69 -

筑波技術大学

修士（工学）学位論文

第1章 序論

1.1 研究背景

1.1.1 ユーザインタフェースの進化とユーザ体験

ユーザインタフェース（User Interface：以下，UI と記す）とは，コンピュータを含む製品やサービスとそれを使う人間（以下，ユーザと記す）の間にあって，ユーザの指示を製品やサービスに伝えたり，製品やサービスからの出力結果を人間に伝えたりするためのソフトウェアやハードウェアの総称である．コンピュータの普及に伴い，コンピュータとの会話を実現するため UI は進化を遂げてきた（図1）．

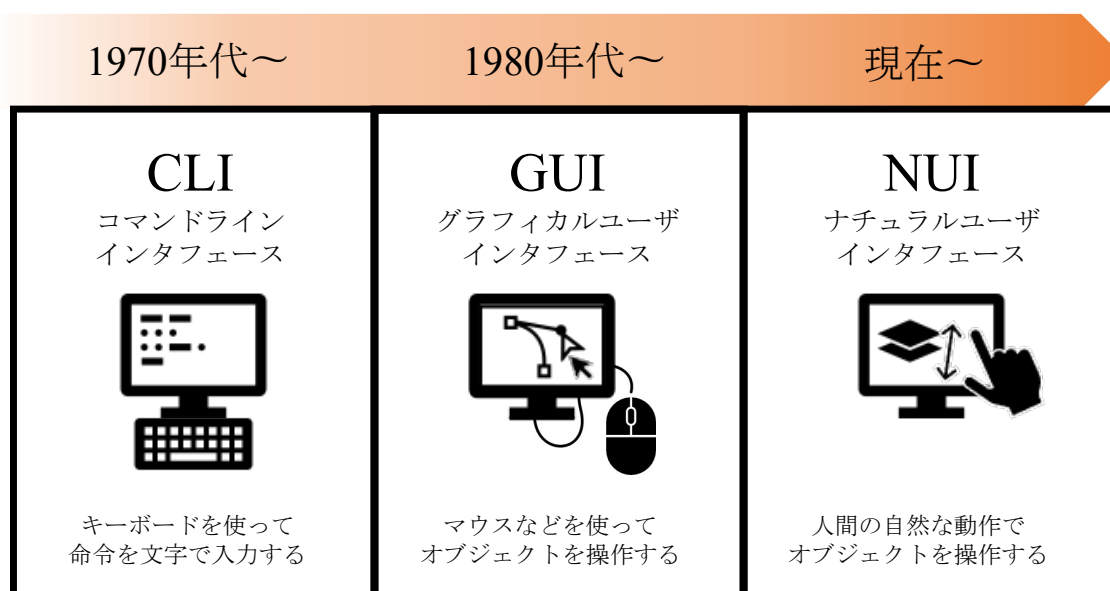


図1：ユーザインタフェースの進化

1970年代以降，パーソナルコンピュータ（以下，パソコンと記す）が飛躍的な普及を遂げ，ますます一般ユーザとの接点としての UI 研究の重要性が叫ばれるようになった[1]．特に，Apple 社のパソコン Macintosh では，UI 設計に力が注がれ，直接操作の重要性が説かれた．そして，ゼロックス Alto というコンピュータにヒントを得て，グラフィカルユーザインタフェース(Graphical User Interface：以下，GUI と記す)の概念が生み出され，それまでのコマンド言語入力方式から，パソコンのインタラクティブ性を大きく飛躍させた．

コマンド言語入力方式は，キーボードによる文字入力が基本で，コマンドラインインタフェース(Command Line Interface：以下，CLI と記す)と呼ばれる．これに対し，GUI では，画面上にグラフィカルに描かれたデータ（オブジェクト）をマウスなどのポインティングデバイスで操作することにより，コマンド入力と等価な作業を直感的に行えるようにした．

また 2008 年のカンファレンスで、Microsoft の Surface を担当するリサーチマネージャである Dennis Wixon は、未来の UI は GUI からナチュラルユーザインタフェース(Natural User Interface:以下, NUI と記す),そしてオーガニックユーザインタフェース(Organic User Interface:以下, OUI と記す)に進化していくと講演した[3]. NUI とは、人間が操作を覚えるのではなく、人間の自然で直観的な行動・感覚をコンピュータが判断し、その結果を返す UI である. NUI の活用事例として、ジェスチャ認識や、視線認識、音声認識、タッチパネルで操作できるデバイスが挙げられる. NUI の具体例を表 1 に示す. そして、OUI は、物理的な入力によって変化する、平面ではない表示装置を持っている UI である.

表 1 : NUI の主な操作手段とその活用事例

操作手段	主な活用事例
タッチ	スマートフォンやタブレット端末、券売機など
音声	音声アシスタントやスマートスピーカー
ジェスチャ	Play Station Move ¹ や Wii ² , Pixel4 のモーションセンス ³
視線	Tobii Eye Tracker 4C ⁴ , GazeChat ⁵

ユーザビリティは国際標準化機構の規格である ISO 9241-11 で「特定の利用状況において、特定のユーザによって、ある製品が、指定された目標を達成するために用いられる際の、有効さ、効率、ユーザの満足度の度合い」とされている[2]. 昨今、UI の上位概念として、ユーザエクスペリエンス (User Experience, UX : 以下、ユーザ体験) の重要性が叫ばれている[1]. ユーザ体験とは、提供された製品やサービスによってユーザにもたされる体験や経験を指す. ユーザビリティの高い機器や作業効率の高い製品やサービスを提供するだけでなく、ユーザに満足感や快適感、感動なども含む価値のある体験をもたらすことが重要であるという考え方に基づくものである.

1 SONY, PlayStation Move モーションコントローラー,
<https://www.playstation.com/ja-jp/accessories/playstation-move-motion-controller/>

2 任天堂ホームページ, Wii TOP, <https://www.nintendo.co.jp/wii/index.html>

3 Pixel Phone ヘルプ, 手を振れずに Pixel を操作する, <https://support.google.com/pixelphone/answer/9517454?hl=ja>

4 Tobii, 製品一覧, <https://www.tobii.com/ja/product-listing/>

5 Zhenyi, H., et al., GazeChat: Enhancing Virtual Conferences with Gaze-aware 3D Photos. UIST '20, 769–782, 2021.

1.1.2 ろう・難聴者ユーザ視点に立った NUI 設計

ユーザビリティの高い UI を開発するには、ユーザの視点に立って、ユーザがそのシステムのメンタルモデルを構築しやすく、あるいは、ユーザの抱いているメンタルモデルと実際のシステムとのギャップが少ないように設計することが重要である。このように、ユーザの視点に立った製品の設計は人間中心設計と呼ばれ、そのプロセスが ISO 9241-210 [4]として規格化されている。この考えを盛り込んだ UI の設計工程が、「目標を設定し、ユーザやヒューマンファクタを分析、システムを設計、プロトタイプを開発、ユーザビリティを評価、その結果をまた仕様にフィードバックさせる」というものである（図 2）。

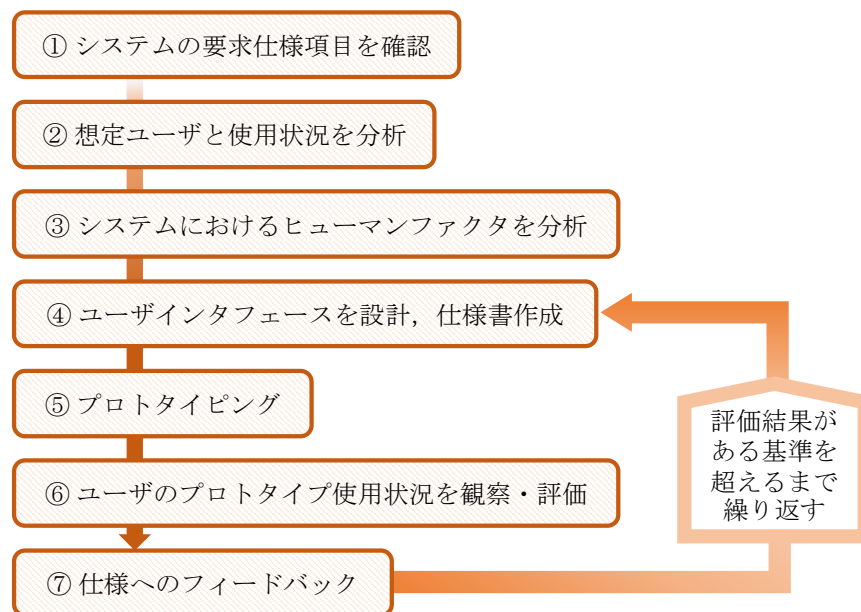


図 2：ユーザインターフェース設計工程

[1]を参考に著者作成

よって、ろう・難聴者ユーザの NUI を設計するにあたり、図 2 に記載されている「② 想定ユーザと使用状況を分析」に重点を置く必要がある。ろう・難聴者がどのような人なのか、能力、特性などを分析する必要がある。また、能力や特性については、実際に使用するろう・難聴者ユーザもしくは想定ユーザに近いモデルユーザを対象に、観察したり、測定したり、意見をもらったりする必要がある。また、ろう・難聴者の使用状況や取り巻く環境についても分析する必要がある。

ここで、ろう・難聴者ユーザ向けでジェスチャを操作手段とした NUI 研究に、飯島らは、音楽を楽しめるようにスマートフォンでドラムのような振動触覚を提示するアプリを開発した[5]。アプリ開発の前段階で、難聴者 1 名 にスマートフォンで楽器を楽しむために必要な情報をヒアリングした。そこから得られた知見を元に、アプリを開発し、ろう・難聴者 6 名が体験・評価を実施した。実験でろう・難聴者ユーザのニーズ分析から、システムの構築に対して再検討を行う必要があることを主張している。

また，ろう・難聴者ユーザ向けで音声を操作手段とした NUI 研究として，Jeffrey らは，ろう・難聴者がスマートデバイスのような音声で操作できる UI にアクセスできるための初期研究を実施した[6]．現在の音声認識技術で，ろう・難聴者の音声認識されないことをスマートデバイスで実証し，ろう・難聴者が音声で操作できる UI にアクセスできるための代替的なアプローチを 2 つ提示している．1 つ目は，ヒューマンコンピューテーション（計算機技術のみでは自動的に処理することが困難な問題を，人の処理能力によって解決しようとする手法⁶）による認識手法を用いることである．2 つ目は文字入力から音声を生成するアプリを用いることである．

しかし，ろう・難聴者の視点に立ち，設計した NUI の研究は極めて少ない．そこで，ろう・難聴者によるアクセシビリティの向上を狙いとした NUI の検討が望まれている．

6 Von Ahn, L.: Human Computation, Doctoral Thesis. UMI Order Number: AAI3205378, CMU, (2005).

1.1.3 会話型ユーザインタフェース

会話型ユーザインタフェース（Conversational User Interface：以下，CUI と記す）とは，図 3 に示す通り，人間と機械やプログラムとのインタラクションに音声ユーザインタフェース（Voice User Interface：以下，VUI と記す）を中心に，GUI，テキスト，車載システムや IoT など，いくつかの手段を使った支援における，システムとユーザの接面のこと[7]である．CUI を作る意義として，大きく挙げられるのが，「人間を精神的に支えることができる」ことである．人間は，たとえ相手がコンピュータであっても，やり取りを行うものであれば，あたかもそれが人間であるかのように扱ってしまうという性質がある．この性質をメディアの等式[8]と呼ぶ．人間は相手に話を聞いてもらったり励ましてもらったりすることで気分が良くなったり，気力が湧いてくる．そのような，会話を含む UI の実現によって，家庭や病院など多くの領域に様々な新しいサービスが生まれることが予想される．そして，現在は音声識別，音声合成，自然言語処理など技術の進歩により，人間が UI と自然に会話できるようになってきている[9]．

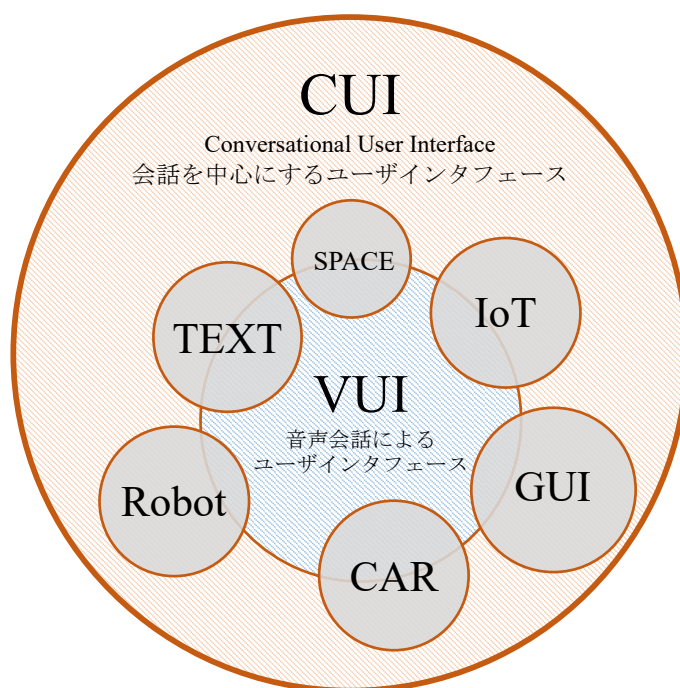


図 3：CUI，VUI とその他のシステムとの関係

[6]を参考に著者作成

しかし，ろう・難聴者が音声モダリティで会話する UI を活用することは極めて困難である．この理由として，ろう・難聴者は音声情報を獲得できないことの他に，「ワードエラー率(Word Error Rate：以下 WER，と記す)」の高さが挙げられる．WER は，音声認識技術を評価する基準として，人間が発話した単語の総数に対して聞き取りができなかった単語の確率である．Google が開発した音声認識システムの WER は 2017 年に平均で 5 %未満を達成した[10]．しかし，Bigham らは，ろう・難聴者の WER が 40 %と音声認識がうまく機能しなかったことを

報告[6]した。また、Abraham は、今の音声認識技術で、ろう・難聴者はまだ使えないことを示したが同時に、ろう・難聴者が発話した音声データを多く収集できれば低い WER を達成できる可能性があることを報告[11]をした。しかし、ろう・難聴者の音声データを多く収集するには多大な時間を要することの問題が述べられている。音声操作はユビキタスなインターフェース技術になりつつあり、この傾向が続くにつれ、この技術におけるアクセシビリティの課題に対処する必要性が高まっている。

ここで、ろう・難聴者のための CUI のアクセシビリティ先行研究に、音声に代わる入力手段として、テキストやジェスチャより手話での会話を希望していることの報告[12]があり、また入力手段にタッチスクリーンよりも手話を用いる方が望ましいことの報告[13]があった。そして、ろう・難聴者ユーザはシステムと手話でインタラクションを行えることに関心を示していることの報告[14]があった。近年、ヒューマンコンピュータインタフェース分野の研究者たちは、手話でインタラクションが行える CUI の設計を検討し始めている[15]。

1.1.4 手話インタフェース関連の行動喚起

Bragg らは、手話インタフェースの設計・構築が、近年、取り組む価値がある研究課題として認識されるようになったことを掲げている[16]. そのために必要な技術として挙げられるのが、手話の認識(recognition)・生成(generation)・翻訳(translation)の 3 つである. それらを総称した技術を、Bragg らは"sign language processing"と呼んでいる. "sign language processing"の技術は学際的な視点からも評価されるようになり、更なる発展のために Bragg らは多様な専門家を集めてワークショップを開催した[17]. そこで、「データセット」、「認識とコンピュータビジョン」、「モデリングと自然言語処理」、「アバターとグラフィックス」、「UI/UX デザイン」の 5 つのトピックに分け、それぞれの現状と課題について議論を行った. さらに、「各分野を超えて共鳴する、行動喚起は何か？」の問題についても議論され、結果、5 つの行動喚起が示された(表 2).

表 2：学際的なワークショップで促された行動喚起の概要

No.	行動喚起(Call to Action)
1	ろう者コミュニティとの連携
2	リアルワールドにおける活用の重視
3	手話 UI に関するガイドラインの作成
4	より大規模で代表的な公開ビデオデータセットの作成
5	アノテーションの標準化とその支援ソフトウェアの開発

1 つ目は、ろう者チームのメンバーを"sign language processing"の技術の研究に参加させることは不可欠であることを示している. これは、手話を母語とするろう当事者を尊重し、ろう者の生活経験から得られる重要な洞察を取り入れるためである. そして、"sign language processing"のシステム開発がろう者コミュニティの要望やニーズに沿ったものであるか確認できるようにするためである.

2 つ目は、リアルワールドの活用事例に対応し、かつ影響力のあるシステムを構築するためには、リアルワールドで利用できるかを考えることは不可欠であることを示している. "sign language processing"は領域特化に適しており、技術には限界があるため、リアルワールドで受ける制約を考慮してシステムを構築する必要がある.

3 つ目は、"sign language processing"のシステムを搭載した UI 設計のガイドラインがあれば、研究者や開発者が独自に設計パターンを作る必要はなく、常に効果的なインターフェースの設計・構築・評価ができることを示している. "sign language processing"の技術はまだ発展途上であり、人々がどのように手話を使うかについて系統的に理解できてないのが現状である.

4 つ目は、多様な手話使用者を含む大規模なデータセットは、ソフトウェアを多様なユーザに適したものにするための学習に不可欠であることを示している. また、そのデータセットが公開されていることは、システムの開発を促進するだけでなく、ろう者が手話を用いるのが可能であることを保証するために重要である.

5つ目は、アノテーションを標準化することで、データの共有、ソフトウェアの互換性の拡大、品質管理が可能になることを示している。アノテーションは、認識システムの学習、自然言語処理や機械学習ソフトウェアへの入力、手話アバターの生成などに不可欠である。また、アノテーションを支援することは、精度、信頼性の向上、コストの改善に貢献できる。そして、アノテーション標準化は手話表記法の標準化にも貢献でき、そこから手話データセットの収集、ラベル付け、結合、さらには手話での一般的な読み書きがより効率的に行えるようになる。

これらの行動喚起の中で、著者が特に着目したのは、3つ目の「手話 UI に関するガイドラインの作成」である。現状で、手話ユーザがシステムとどのようなインタラクションが行われるかについての基礎研究が不足している。手話ユーザのために開発されたシステムは数多くあるが、これらの研究は、システム単体の評価に焦点を当てており、ユーザとシステムの間で交わされるインタラクションの原則については概説されていない。その結果、新しいシステムを開発する各チームは、研究に基づく一般的なデザインガイドラインの恩恵を得られずに、ほぼゼロからインターフェースをデザインしなければならない。

以上から、研究が進んでいる NUI の中でも、ろう・難聴者に最適な会話型 NUI の設計ガイドラインの提案が望まれている。

1.1.5 ユーザインタフェースのガイドライン

ろう・難聴者に最適な会話型 NUI の設計ガイドラインを策定する際に、既存の GUI と VUI の設計ガイドラインは適用できるものか調査する。

GUI 設計ガイドラインとして、主に Nielsen による「ヤコブの 10 ユーザビリティヒューリスティックス」[18][19], Norman による「ユーザー中心デザインの 7 つの原則」[20], Shneiderman による「インタフェースデザインの 8 つの黄金律」[21]があげられる。それぞれのガイドラインを表 3, 表 4, 表 5 に示す。

表 3 : ヤコブの 10 ユーザビリティヒューリスティックス

No.	原則
1	システム状態の視認性を高める
2	実環境に合ったシステムを構築する
3	ユーザにコントロールの主導権と自由度を与える
4	一貫性と標準化を保持する
5	エラーの発生を事前に防止する
6	記憶しなくても、見ればわかるようなデザインを行う
7	柔軟性と効率性を持たせる
8	最小限で美しいデザインを施す
9	ユーザによるエラー認識, 診断, 回復をサポートする
10	ヘルプとマニュアルを用意する

表 4 : ユーザー中心デザインの 7 つの原則

No.	原則
1	外界にある知識と頭の中にある知識の両方を利用する
2	作業の構造を単純化する
3	対象を目に見えるようにして, 実行のへだたりと評価のへだたりに橋をかける
4	対応づけを正しくする
5	自然の制約や人工的な制約などの制約の力を活用する
6	エラーに備えたデザインをする
7	以上の全てがうまくいかない時には標準化をする

表 5：インタフェースデザインの 8 つの黄金律

No.	原則
1	一貫性を追求しよう
2	よく使うユーザにはショートカットを提供しよう
3	有意義なフィードバックを提供しよう
4	処理の完了が明確に分かるようにしよう
5	エラーに簡潔に対処できるようにしよう
6	アクションを簡単にやり直せるようにしよう
7	ユーザに主導権を持たせるようにしよう
8	人間の短期記憶に負担を減らそう

GUI は古くから、設計を成功させ、ユーザビリティの問題を回避するための鍵としてヒューリスティックを用いてきた。GUI 設計ガイドラインは VUI に直接適用することはできない [22]。ここで、VUI と手話会話を含めた NUI はどちらもノータッチインタラクション、自然言語で対話をするものの共通点がある。よって、GUI 設計ガイドラインをベースにろう・難聴者に最適な NUI ガイドラインを考案することは困難と考える。

VUI 設計ガイドラインとして、主に Alexa⁷[23]と Google アシスタント⁸[24]があげられる。それぞれのガイドラインを表 6、表 7 に示す。

表 6：Alexa デザインガイド

No.	原則
1	柔軟性を持たせる
2	パーソナライズする
3	分かりやすくする
4	自然な会話にする

表 7：Google アシスタントの会話デザインに関するガイドライン

No.	原則
1	会話に適したアクション体験にする
2	ブランドペルソナを作成する
3	会話中のエラーに対処する
4	画面での見え方を考える前に 音声のみの体験を試す

7 Alexa とできること。], <https://www.amazon.co.jp/meet-alexa/b?ie=UTF8&node=5485773051>

8 「それ、Google にやらせよう。」, <https://www.apple.com/jp/siri/>

音声会話システムの開発において、将来のデザイナーの指針となる VUI 原則の基盤を構築することが不可欠であることの報告[25]がある。また、VUI 設計ガイドライン（表 6, 表 7）は、音声ベースでインタラクションを行うことを前提に設計されている。そして、VUI と会話する時に主に使われる音声情報から手話情報に置き換えるだけでは、ろう・難聴者の身体性を考慮した、快適なユーザ体験を作り出すことができないことが考えられる。ろう・難聴者の対人コミュニケーションにおいては手話だけでなく、他のモダリティ（詳細は 1.3 節にて説明）が活用されている。ゆえに、ろう・難聴者に最適な会話型 NUI ガイドラインの作成方法として、VUI との会話の各ステップで、代替可能なろう・難聴者によるアクセス方法を検討する。以上から、本研究では既存の VUI との会話ステップをベースにした、ろう・難聴者に最適な会話型 NUI のガイドラインの提案を目指す。

1.2 関連研究

会話型 NUI の活用事例として挙げられるスマートスピーカーとジェスチャインタフェースそれぞれの現状と課題を示し、著者が提案するろう・難聴者に最適な会話型 NUI はその課題を克服できるものか示す。また、実験デザインにおける Wizard of Oz 法についても、今節で説明する。

1.2.1 スマートスピーカーの普及状況と課題

2010 年代後半になると、音声認識性能や AI 技術の向上を背景に、スマートスピーカーが各社から発売され、実際に家庭の中で利用され始めている[26]。これは、会話形式 AI エージェントであり、ユーザの音声による問いかけに応じて、音楽やニュースや天気予報を音声で再生するものである。ここでは、ユーザが、近くに置いたマイク/スピーカを内蔵した本体に、「ハローAI」などのウェイクワードを発話してから質問や要求を続けることによって、タスクを実行することができる。スマートスピーカーの普及理由について、音声認識・自然会話技術の向上、作業を中断せずハンズフリーで使える[27]ことが挙げられる。Canalys 調査[28]では、全世界で 2020 年にスマートスピーカーの出荷台数は 3 億 2,000 万台を達成し、今後 2024 年に 6 億 4,000 万台に達すると予測されている（図 4）。

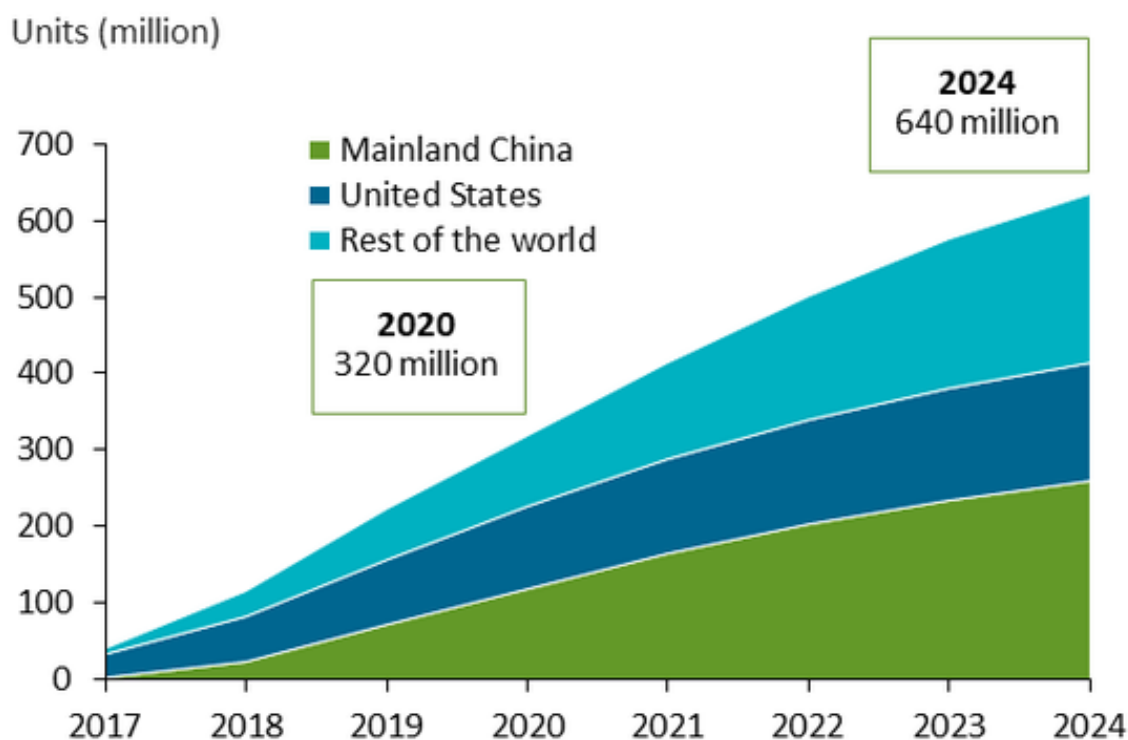


図 4：スマートスピーカーの国別出荷台数予測

文献[26]より引用

スマートスピーカーに搭載されている AI アシスタントの主な活用事例として、Amazon の「Alexa」、Apple の「Siri⁹」、Google の「Google アシスタント」や、LINE の「Clova¹⁰」の 4 つが挙げられる。これらのスマートスピーカーの外観を図 5 に示す。Alexa は連携できるスマート家電の種類が豊富であり、Amazon のショッピングや、Kindle 本の読み上げが可能である。Siri は、iPhone と初期設定や音楽再生に関する連携ができ、iPhone からアラームを設定するなど遠隔操作も可能である。Google アシスタントは、「Google Play Music」や、「Google Map」、「Google Photo」など、日常でも使うシーンの多い各種サービスと連携でき、Google 検索で検索可能な範囲で回答が可能である。Clova は、LINE で繋がっている相手と音声だけで LINE メッセージの送受、LINE 通話が可能である。



図 5：スマートスピーカーの主要 4 社

タッチ入力や視覚的フィードバック獲得ができる画面付きスマートスピーカー（以下、スマートディスプレイと記す）として Alexa の Echo Show、Google の Nest Hub などが導入された。そして、そのスマートディスプレイの形状・機能も発展を遂げてきた。その代表事例を Echo Show で説明する（図 6）。

Echo Show¹¹は、スマートスピーカーに画面が付いたデバイスであり、Prime ビデオ視聴の他に、ビデオ通話、写真の表示が可能になった。また、音声による出力情報が多すぎるとユーザの認知負荷が大きくなる問題も、ディスプレイでビジュアルな表現を利用することで解消できる。Echo Show¹²は Echo Show5 より画面が大きく、動画を視聴しやすくなった。また、音質も向上し音楽が聞こえやすくなった。Echo Show¹³では、350 度回転の首振り運動が可能になった。ヒトの動きや声に追従して、適切な方向へ画面の向きを自動で調整してくれる。これにより、ビデオ通話中は常にフレームインできる他に、内蔵カメラに安全にアクセスして、外出先からも部屋の様子がいつでも確認できるようになった。これまでの Echo Show シ

9 「Siri がさらに活躍します。あなたが頼む前から。」、<https://www.apple.com/jp/siri/>

10 「CLOVA は人に寄り添う「ひとにやさしい AI」です。」<https://clova.line.me/>

11 Echo Show5 第 2 世代-スマートディスプレイ、www.amazon.co.jp/dp/B08KGY97DT

12 Echo Show8 第 2 世代-HD スマートディスプレイ、www.amazon.co.jp/dp/B084TNP2B4

13 Echo Show10 第 3 世代-モーション機能付きスマートディスプレイ、www.amazon.co.jp/dp/B084PVPV27Z

リーズはテーブルやデスクに置くタイプだが、最新モデルの「Echo Show15¹⁴」は薄型で壁掛けできる他に、カウンターなどに置くことができる。また、15.6 インチの大きな画面を活かし、スケジュールやメモ、付箋紙など、複数の情報を表示することができる。なお、これらの Echo Show はアクセシビリティ機能として、字幕の付与が可能である[29]。

そして、スマートスピーカーの課題に、オフィスや公共スペースで使うことで会話内容が周囲の人たちに聞かれることのプライバシー問題、騒音環境下で音声の認識困難な問題が挙げられる。前者の問題については、聴覚は視覚と比べて受動的かつ無意識下の刺激であるため、他者が近距離にいる限り、音声を出すとどの方向であっても気づかれやすいからである[30]。そのため、ろう・難聴者に最適な会話型 NUI は、音声情報を利用しないため、オフィスや公共スペースでプライバシー問題を解消できることが期待される。後者の問題については、音声認識を必要としないことから、騒音環境下での認識問題が生じない。



Echo Show5



Echo Show8



Echo Show10



Echo Show15



図 6 : Echo Show シリーズ

14 Echo Show15 15.6 インチフル HD スマートディスプレイ, www.amazon.co.jp/dp/B08MQNJC9Z

1.2.2 ジェスチャインタフェースの活用事例と課題

ジェスチャインタフェースは、視覚機能と腕、手指、顔の表情などの身体機能を利用するインタフェースである。応用例としては、Google Pixel 4 の Motion Sense (図 7) や、ドローンなどがある。モーションセンスとは、スマホに直接触れず、手をかざすことで音楽再生時の進む・戻る操作や、着信やアラームの停止などの操作が可能な機能である。また、モーションセンスによって、スマホに手を近づけただけで、画面を点灯させる機能も搭載されている。

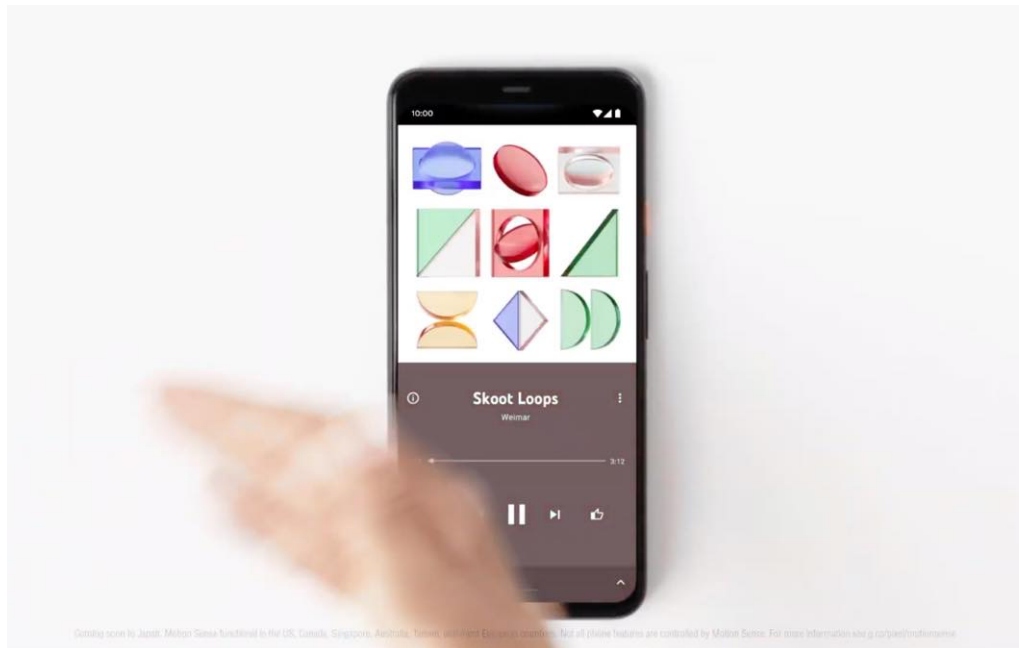


図 7 : Pixel4 のジェスチャ機能「Motion Sense」¹⁵

そしてジェスチャインタフェースの課題には、「記憶負荷」[31][32][33]，「意図の区別」[34]，「肉体疲労」[35]の 3 つが挙げられる。

記憶負荷

ジェスチャ入力、言語性がないため、CLI のようにある程度コマンドを覚えておく必要がある。CLI に比べればある程度は対象や状況がヒントになる可能性はあるものの、ジェスチャ操作に依存すればするほど、ユーザが覚えるべきジェスチャの種類が多くなる。そのため、コマンドの表現能力は低下し、コマンドとして用いるジェスチャに一貫性を持たせるのは困難である。

意図の区別

ジェスチャ操作は、人間が頻繁に行う直感的で自然な動作で操作できることが一つの強みである。しかし、その直感と自然さが、人間の日常的な動作とオーバーラップし、本来の操作と日常的な動作の区別がつかず誤ってコマンドと認識されて実行されてしまうという問題が発生する。

肉体疲労

15 Pixel Phone ヘルプ、手を振れずに Pixel を操作する、<https://support.google.com/pixelphone/answer/9517454?hl=ja>

ジェスチャでは、腕を大きく上下左右に動かす、腕を長時間上げているような動作をすることがあり、人間の負担が大きくなる。

ろう・難聴者に最適な会話型 NUI は、手話をメインとしインタラクションを行うための UI であり、「記憶負荷」や「意図の区別」といったジェスチャインタフェースの課題を克服する手段として期待される。また、ろう・難聴者に最適な会話型 NUI すなわち手話を用いた UI であれば、手話を母語とするろう・難聴者にとって自然な手段であり、「肉体疲労」といったジェスチャインタフェース特有の問題は生じえないと考えられる。

1.2.3 Wizard of Oz 法

手話を用いて UI とのインタラクションに関わる実験の実施において、連続手話認識システムが必要である。しかし、現時点で開発された実生活での連続手話認識システムにおける認識率は 39.6%にとどまる[36]など、実験で使用することはできない。この問題を解決する手法として、Wizard of Oz 法[37][38]がある。これは、Wizard と呼ばれる人間が、システムのふりをしてユーザと会話する手法である。Wizard of Oz 法では、会話型 UI を搭載したシステム全体が完成していない状態でも、Wizard が未開発の箇所を補完し、システムを動作させることが可能になる。本研究では、このアプローチを使って、図 8 のように、Wizard がろう・難聴者に最適な会話型 UI を搭載したシステムを振る舞い、ろう・難聴者ユーザと会話する場面に生じるデータを抽出する実験を行う。

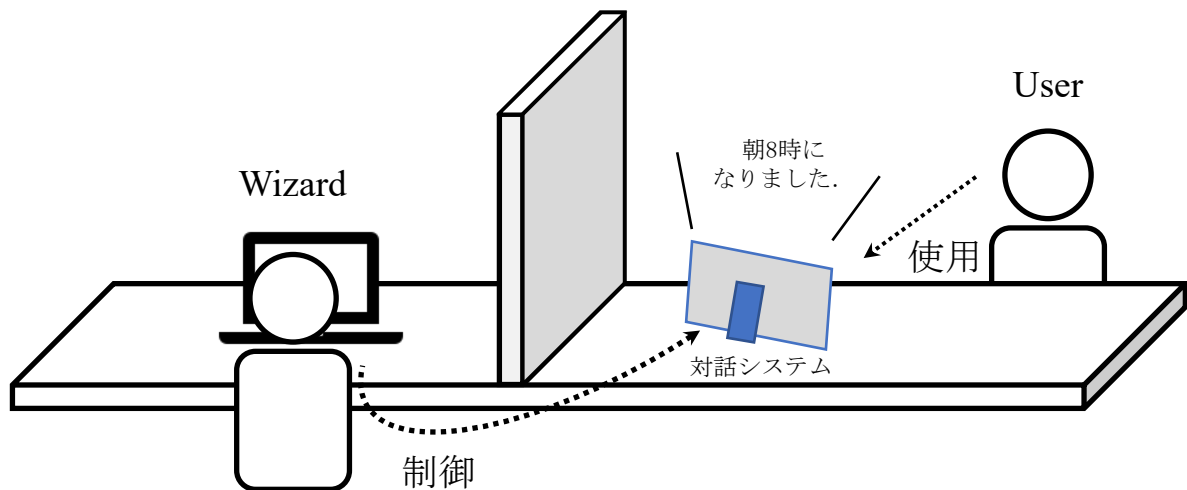


図 8 : Wizard of Oz 法を用いた実験配置

1.3 研究目的

本研究では、ろう・難聴者ユーザ視点に立った NUI の設計ガイドラインを提案することを本論文の目的とする。

目的を果たすために、会話型 NUI を搭載したシステム（以下、システムと記す）と聴者の会話プロセスから、ガイドラインを作成するために必要な要素を解明する。聴者とシステムの会話プロセスの様式についてまとめたものを図 9 に示す。会話の最初に、ユーザからウェイクワードでシステムを呼び起こす必要がある。ウェイクワードは、AI アシスタントとして、音声を受け付けるためのトリガーとしての簡単な音声フレーズである。Alexa の場合は"Alexa", Google アシスタントの場合は"OK, Google.", そして Siri の場合は"Hey, Siri."などのウェイクワードを用いる。次に、システムはユーザのウェイクアップを検出し、コマンド受付可能状態のフィードバックを提示する。例として、Alexa の場合は画面下部に青いバーが表示され、Google アシスタントの場合はユーザの発言内容のフィードバックが図れるよう上部にバーが表示される。続いて、ユーザは天気予報やニュース、アラームなどのタスクを実行するためにシステムに音声でコマンドを入力する。システムはユーザのコマンドに従い、音声読み上げでタスクを実行する。ほかの会話シチュエーションとして、「アラーム通知」や相手からの「ビデオ着信」でシステムからユーザに音声で呼びかける場合がある。

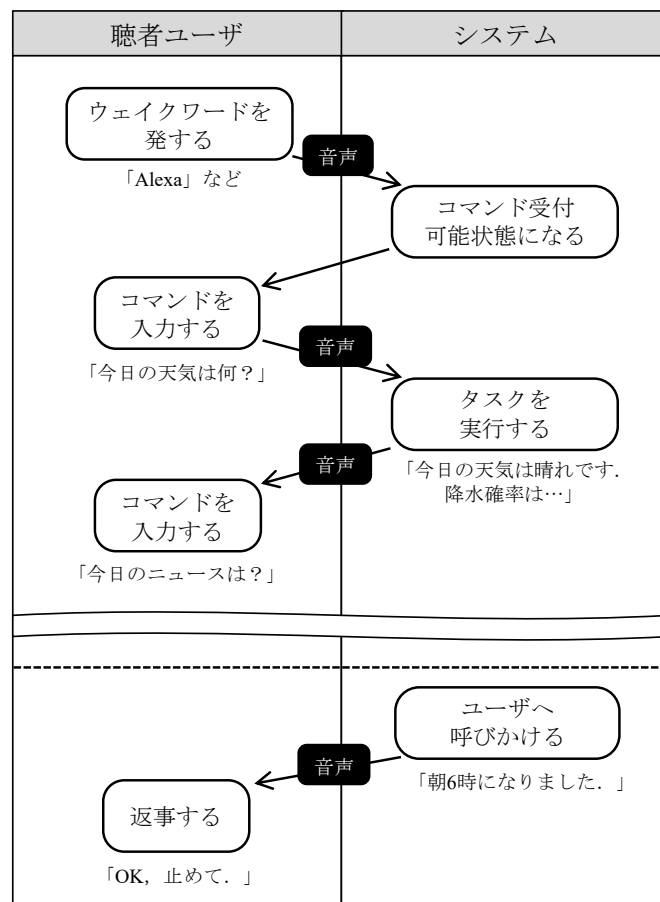


図 9：聴者ユーザとシステムの会話プロセスの様式

聴者ユーザとシステムの会話ステップで、代替可能なろう・難聴者によるアクセス方法を検討する。そのために考案した6つの研究課題（表8）を検証し、その結果に基づいたガイドラインを提案する。

また、これらの研究課題を検証するにあたって、調査対象者は筑波技術大学[39]の学生とした。筑波技術大学は、視覚障害者および聴覚障害者のための大学であり、聴覚障害者学生の両耳の聴力レベルはおおむね60デシベル以上である。そのため、筑波技術大学の学生はろう者だけでなく・難聴者も含む。

表8：本研究で示す研究課題

No.	研究課題(Research Question)
1	システムが光を用いた応答を示すことで、ろう・難聴者にとってユーザビリティは向上するか
2	ろう・難聴者にとってシステムに適した手話/文字の表示方法は何か
3	ろう・難聴者がシステムに行うウェイクアップ方法に注視は有効か
4	ろう・難聴者がコマンド指示を行う前に、システムよりコマンド受付可能であることを提示を確認したいか
5	ろう・難聴者がシステムをウェイクアップするより良い方法は何か
6	ろう・難聴者のウェイクアップを検出したとき、システムがコマンド受付可能であることを示すより良い方法は何か

上記の研究課題について概要を説明する。

RQ1:光を用いた応答のユーザビリティ

システムが出力する音声情報を、聴者はアイズフリーで（常にシステムを見ずに）獲得することができる。音声情報を獲得できない、ろう・難聴者がシステムを用いる場合、ディスプレイ付きのデバイス（Echo Show や Nest Hub）を用いても、システムを視界から外すときにシステムからの応答を捉えにくいことがある。すなわち、システムの音声出力に代わり、ろう・難聴者ユーザに最適なシステムの出力方法を調査することが重要である。

一方で、ろう文化において、光による通知は馴染みの深いものとなっている。例えば、ろう・難聴者の日常生活で一般的に活用されているものとして、光通知機能が搭載されたインターフォン、アラーム、火災警報器[40]などがある。また、周囲で鳴るアラームの音源の方向を光通知で、ろう・難聴者に伝達する発光装置が開発されている[41]。そこで、本論文では、光通知による出力を用いることで、ろう・難聴者によるユーザビリティは向上できるか調査を行う。

RQ2:手話と文字の最適な表示方法

ろう・難聴者ユーザがシステムの字幕機能を用いたとしても、ユーザからの手話入力を可能とする場合は、会話で用いる言語は手話がメインになることの理由から、システムから字幕が出力されると、ユーザビリティが低くなることが懸念される。Alexaのようなデバイスが家庭に入ること、キッチンやリビングルームにデバイスを配置し、ハンズフリーでインタラクションを行うのが増えたことの報告[42]がある。そこで、他のことをする片手間にインタラクションを行うことが多くなることが考えられる。ろう・難聴者も同様に、他のことをしながら、システムからの応答情報を捉えたい。ここで、ろう・難聴者による流れる手話の表示、一定時間残る文字の表示の必要性を明らかにすることは重要である。

中等教育、高等教育の環境下でのろう学生に対する支援について、手話通訳、字幕、またその両方それぞれの条件からの影響を調査したもの[43]がある。ここでは、字幕のみの場合、他の2つの条件に比べてろう・難聴者の成績が有意に高かったことの報告がなされている。テレビのコンテンツを理解する上で、字幕と手話の複合効果を評価したもの[44]がある。ここでは、字幕と手話の両方を提供することが、単独で提供するより多くのろう・難聴者がテレビにアクセスできるようにするのに非常に役に立ったことの報告がなされている。また、文字言語が主体であるWebページに、手話ビデオを組み入れることでろう・難聴者によるアクセシビリティは改善されるか調査した研究[45]がある。ここでは、手話ビデオを通して提示された情報で、ろう・難聴者ユーザーのWebに対する関心が高まったことの報告がなされている。しかし、これらの研究は、一方通行のメディアを対象としている。

双方向のインタラクションを行うシステムは、ユーザから入力を行う必要があり、その場合に適した手話/文字の表示に対する必要性を明らかにすることが重要である。そして、テレビやWebの場合では、話し手の言語に合わせて翻訳しなければならないが、システムはコンピュータが発信するので、言語は受け手に合わせて調整することができる。

すなわち、ろう・難聴者ユーザによる必要性を考慮した上で、手話/文字の表示方法を検討しなければならない。以上から、字幕だけでなく手話も付与する場合に、並行作業を行う状態で、ろう・難聴者による手話および字幕に対する必要性を調べる。

RQ3:ウェイクアップにおける注視の有効性

ろう・難聴者同士で会話を始める場合に、話したい相手とアイコンタクトを取る必要があり、そのためには肩を叩く、または手を振る動作を行う[46][47]。また、会話の開始時点では発話だけでなく非言語的情報も重要な役割を果たす場合がある。その中でも「注視」は相手の行動を喚起・促進し、インタラクションの開始および進行に関わる[48]。

Mandeらは、AIアシスタントのウェイクアップに関して、Wizard of Oz法によってろう・難聴者ユーザが好む方法を明らかにした研究[49]を実施した。その結果、ろう・難聴者にとって、デバイスのサインネームを使うことがベストであり、次に手振り、拍手、リモコン使用、スマホアプリ使用、デバイスの指文字の順番に好んだ。しかし、ろう・難聴者との対人コミュニケーションで会話のスタートに必須のアイコンタクトの活用については検討されていない。

以上から、注視を用いることで、明示的なウェイクアップを用いない自然な会話ができ、ユーザの満足度は高くなることが考えられる。以上を踏まえ、システムのウェイクアップにおける注視活用の有効性を調査する。

RQ4:フィードバックの必要性

ろう・難聴者の対人コミュニケーションの際に、聴者と比較して呼びかけるときに相手の注意獲得は必須であるとの報告[50]がある。これより、聴者がシステムとインタラクションを行うとき、ろう・難聴者と比較して、システムが提示するフィードバックの確認を行わないことが推測される。聴者は「常に自分が発する音声は聞かれている」と思う一方、ろう・難聴者は「見ているときだけ伝わる」と思うといった違いがある。そのため、ろう・難聴者によるウェイクアップへのシステムのフィードバックに対する必要性を明らかにする。

RQ5:ウェイクアップのより良い方法

Mande らは、ろう・難聴者ユーザのウェイクアップ方法に関する研究で、意図しない手振りを行う場合、システムにウェイクアップとして認識されてしまう懸念があることを指摘した[49]。ここで、注視は呼びかけ対象に視点を向けるといった意図的に行えるウェイクアップ動作と考える。そのため、サインネーム条件と手振り条件に注視条件を加えて、ろう・難聴者ユーザのウェイクアップ方法それぞれの満足度を比較し調査する。

RQ6:フィードバックのより良い提示方法

ろう・難聴者との対人コミュニケーションで会話のスタートに必須のアイコンタクト[46][47]について、システムをウェイクアップした際に、システムからのより良いフィードバック方法は何か知る必要がある。前節で述べたように、Echo Show5 の青いバー表示や、Echo Show10 の首振り運動だけで、ろう・難聴者のユーザ体験の向上に繋がるかまだ明らかになっていない。そのため、システムがコマンド受付可能であること（フィードバック）を示す方法それぞれの満足度を比較して調査する。

1.4 本論文の構成

本論文は全5章からなる。第2章では、RQ1-RQ3を検証するために、光通知と注視を用いた手話会話システムがろう・難聴者によるユーザビリティの向上に繋がるか調査する。次に、第3章では、RQ4-RQ6を検証するために、ろう・難聴者に最適なウェイクアップ方法とフィードバック方法を検討する。第4章では、6個の研究課題への検証結果に基づき、ろう・難聴者に最適な会話型NUIの設計ガイドラインを提案する。最後に、第5章で本論文のまとめと今後の課題について述べる。

第2章 光通知と注視を用いた手話会話型ユーザインタフェースの検討

2.1 検証する研究課題

今章で検証する研究課題は RQ1-RQ3 である。これらは、表 9 に示す通り、ろう・難聴者ユーザがシステムと会話するときに実現したい相互入出力モダリティを網羅した研究課題である。聴者ユーザがシステムと会話するときの相互入出力モダリティは音声のみで行える。一方、ろう・難聴者ユーザの場合、音声情報を手話情報に置き換えるだけではユーザ体験の向上に繋がらないと考える。ろう・難聴者ユーザがシステムと会話をするときに手話/文字モダリティを使用する(RQ2)。システムから、ろう・難聴者ユーザに呼びかけるときには光通知による出力モダリティを使用する(RQ1)。それに対し、ろう・難聴者ユーザからシステムにウェイクアップを行う（呼びかける）ときには注視モダリティを使用する(RQ3)。こういったシステムとの相互入出力モダリティの使用が、ろう・難聴者のユーザ体験の向上に繋がるか調査する。RQ1 から RQ3 を検証するために、Wizard of Oz 法に基づいた、光通知を用いた手話会話システムを構築し、評価実験（以下、実験 1 と記す）を行う。なお、実験 1 は筑波技術大学研究倫理委員会の承認を受けている（承認番号：2020-13）。

表 9：ユーザとシステムが会話する時の相互入出力モダリティ

パターン ユーザ	会話中		呼びかけ	
	システム→ユーザ	ユーザ→システム	システム→ユーザ	ユーザ→システム
聴者	音声		音声	音声
ろう・難聴者	手話/文字(RQ2)		光(RQ1)	注視(RQ3)

2.2 実験 1 の方法

2.2.1 参加者

実験参加者は本学学生のろう・難聴者 12 名（20～24 歳）である。また、参加者の属性による実験結果への影響の解析も同時に行うため、特性を調査した。性別や補聴器/人工内耳の装用の有無、手話歴、手話レベル、日常会話における声の使用の有無、視覚情報/聴覚情報の依存状況、AI アシスタントの使用経験それぞれの結果をまとめたものを図 10 に示す。

なお、調査項目の中にある手話レベルについて、長南は、「日本手話」を「日本語とは異なる独自の統語構造を持つ日本手話」，「中間手話」を「日本語の統語構造に従って手話単語が表現される中間型手話」と定義している[51]。そして、手話を読み取る際に、音声情報は不要であるかによって手話レベルに影響すると考える。以上から、手話レベルが高い順に「声なし日本手話を読み取れる」，「声なし中間手話を読み取れる」，「声あり中間手話を読み取れる」の 3 つのレベルを用いた。

性別については、男性 7 名、女性 5 名であった。また、補聴器/人工内耳の装用の有無について、「装用している」と回答した参加者が 8 名、「装用していない」と回答した参加者が 4 名であった。次に、手話歴について、5 年単位で刻み分類した結果、生まれてからずっと手話を使っている参加者が 4 名、15 年以上 20 年未満手話を使っている参加者が 4 名、10 年以上 15 年未満手話を使っている参加者が 2 名、5 年以上 10 年未満手話を使っている参加者が 1 名、そして 5 年未満手話を使っている参加者が 1 名であった。続いて、手話のレベルについて、「声なし日本手話を読み取れる」と回答した参加者が 6 名、「声なし中間手話を読み取れる」と回答した参加者が 3 名、「声あり中間手話を読み取れる」と回答した参加者が 3 名であった。また、日常生活で交わされるコミュニケーションに声を使うかどうか尋ねた結果、「使う」参加者が 9 名、「使わない」参加者が 3 名であった。最後にコミュニケーションを図る際に、視覚情報と聴覚情報をどのように活用しているかについて、「視覚情報と聴覚情報の両方」を活用する参加者が 6 名、「視覚情報のみ」活用する参加者が 6 名であった。

なお、AI アシスタント使用経験については、4 段階評価（1.よく使う，2.たまに使う，3.あまり使わない，4.使わない）で回答してもらった。その結果、「3.あまり使わない」および「4.使わない」と回答した参加者が、11 名であり、全体の 92%を占めていた。そう回答した理由として、「自分の声を認識してもらえない」「音声を用いて話さない生活を送っている」とコメントしている人が大半であった。ろう・難聴者ユーザのうち、AI アシスタントを使用している人は非常に少数であることの報告[52]があり、実験 1 の参加者と同様の傾向があることを示している。他に、「1.よく使う」と回答した参加者は 0 名、「2.たまに使う」と回答した参加者は 1 名であった。

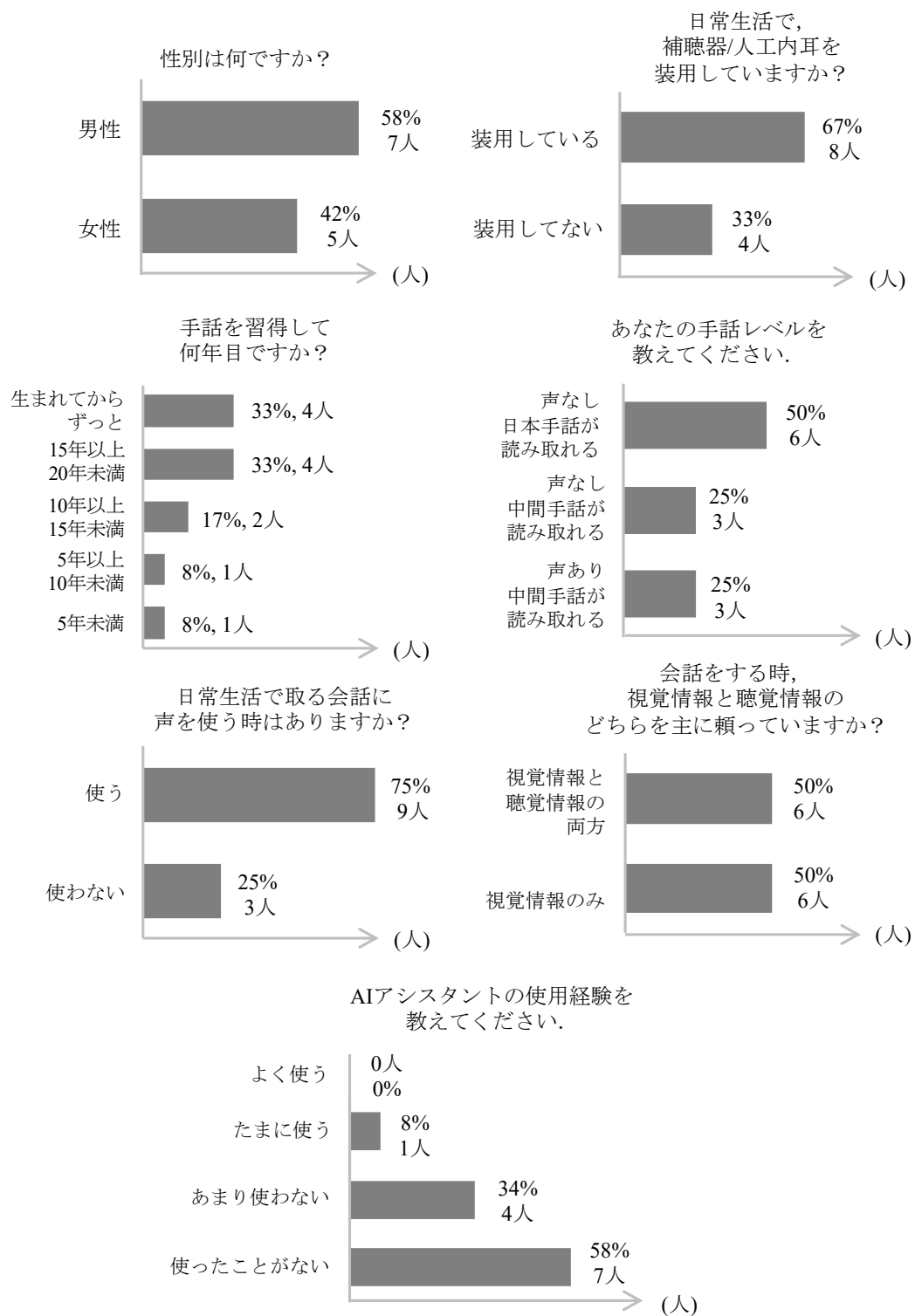


図 10：実験 1 の参加者(N=12)の特性

2.2.2 システムの構築

システムが実行するタスクの選定

システムを構成するにあたって最初に、システムが実行するタスクの選定を行った。株式会社電通デジタルのスマートスピーカーの利用機能に対する調査結果を図 11 に示す。これらのタスクにおけるユーザへの通知方法で、ろう・難聴者が抱える問題について、グルーピングを行った。その結果、システムから呼びかけてくることに気づかない問題、ユーザの質問に対し回答が終了するタイミングに気づかない問題の 2 つに分類した。そのため、本実験でシステムが実行するタスクの数は、「呼びかけ」のタスクを 2 つ、「回答」のタスクを 2 つ、合計 4 つとした。タスクの選定において、図 11 に掲載されているランキング上位の「音楽を聞く（1 位）」と「ラジオ（radiko など）聞く（5 位）」は、特に聴覚情報に依存するタスクのため、対象外とした。他に上位である、2 位の「天気予報」と 4 位の「ニュース」は、ユーザの質問に対し回答を行うタスクである。そして、3 位の「アラーム/タイマー機能」は、システムは、アラームやタイマーが鳴ったことをユーザに呼びかけるタスクである。そして、ランキング圏外のタスクであるが、「電話着信」もシステムからユーザに呼びかける。以上から、「呼びかけ」のタスクは「アラーム設定」と「電話着信」、そして「回答」のタスクは「天気予報」と「ニュース」とした。

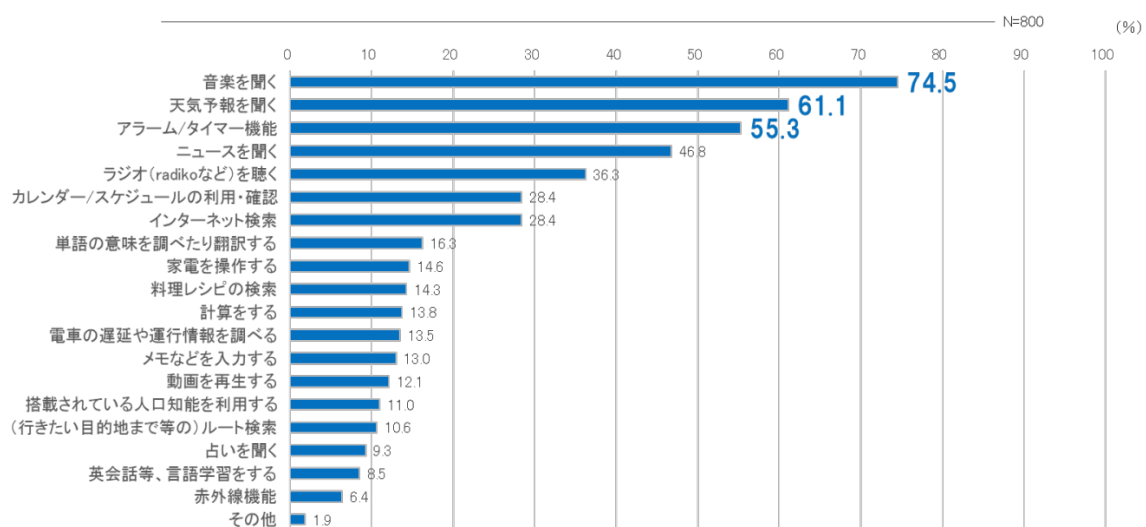


図 11：スマートスピーカーの利用機能¹⁶

¹⁶ DENTSU DIGITAL NEWS RELEASE, 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, 「電通デジタル, スマートスピーカーの国内利用実態調査を実施」

システムの構成

実験 1 で構築するシステムの基本構成は、iPad、Meross Wifi スマート LED 電球（以下、LED 電球と記す）、GoPro9 カメラとした。システムの外観と動作の様子を図 12 に示す。iPad は、Echo Show8 のディスプレイ用コンテンツを模擬し、Microsoft 社製 Microsoft PowerPoint 2019 でプレゼンテーションを作成し表示したものである。また、システムからの応答に手話も表示するため、ディスプレイに手話者の動画と合成を行った。



図 12：実験 1 で構成するシステムの外観および挙動内容

Wizard 側から遠隔でシステムを制御するために、iPhone と MacBook Pro を用いた。iPhone は、LED 電球の色（RGB1600 万色）と点滅周期を設定でき、Wi-Fi 経由で遠隔から LED 電球の制御が可能である。MacBook Pro は、Sidecar を用いることでデスクトップを拡張するディスプレイとして iPad の遠隔制御が可能である。

実験 1 で用いる LED 電球での光呈示には、システムからの呼びかけの時に黄色の点滅、システムの応答中に緑色の点灯と 2 つのパターンを設定した。表 10 に、色度輝度計 CS-100A にて計測した、各パターンの色光の色度と輝度を示す。なお、GoPro9 カメラは、利用者からの手話入力を正面で見えるように設置した。

表 10：色光それぞれの特徴

パターン	タスク	色	輝度 (cd/m ²)	色度		点滅周期 (s)
				x	y	
呼びかけ	アラーム通知 /電話着信	黄色	241	0.54	0.41	1.21
回答	天気予報 /ニュース	緑色	131	0.32	0.56	

システムが実行する入出力

システムが実験中に、どのような入力 がされ、どのような出力を示すかフローチャート形式でまとめたものを図 13 に示す。

アラーム設定のタスクについて、ユーザがシステムに時間指定でアラーム設定のコマンドを出してきたら、システムよりセットしたことを出力する。そして、時間になったらシステムからユーザに光で通知を行い、時間になったことと LED 電球による黄色光の点滅を一緒に出力する。ユーザが止めることのコマンドを出してきたら、システムの LED 電球を消灯してホーム画面に戻る。

電話着信のタスクについて、システムよりユーザに電話が着信したことを LED 電球による黄色光点滅と一緒に出力する。次に、ユーザが着信を拒否することのコマンドを出してきたら、システムの LED 電球を消灯し、着信を拒否したことを出力する。そして、ホーム画面に戻る。

天気予報とニュースのタスクについて、ユーザがシステムに天気およびニュースのコマンドを出してきたら、それに対応した情報を LED 電球による緑色光の点灯と一緒に出力する。システムの情報の出力が終わったら LED 電球を消灯し、ホーム画面に戻る。

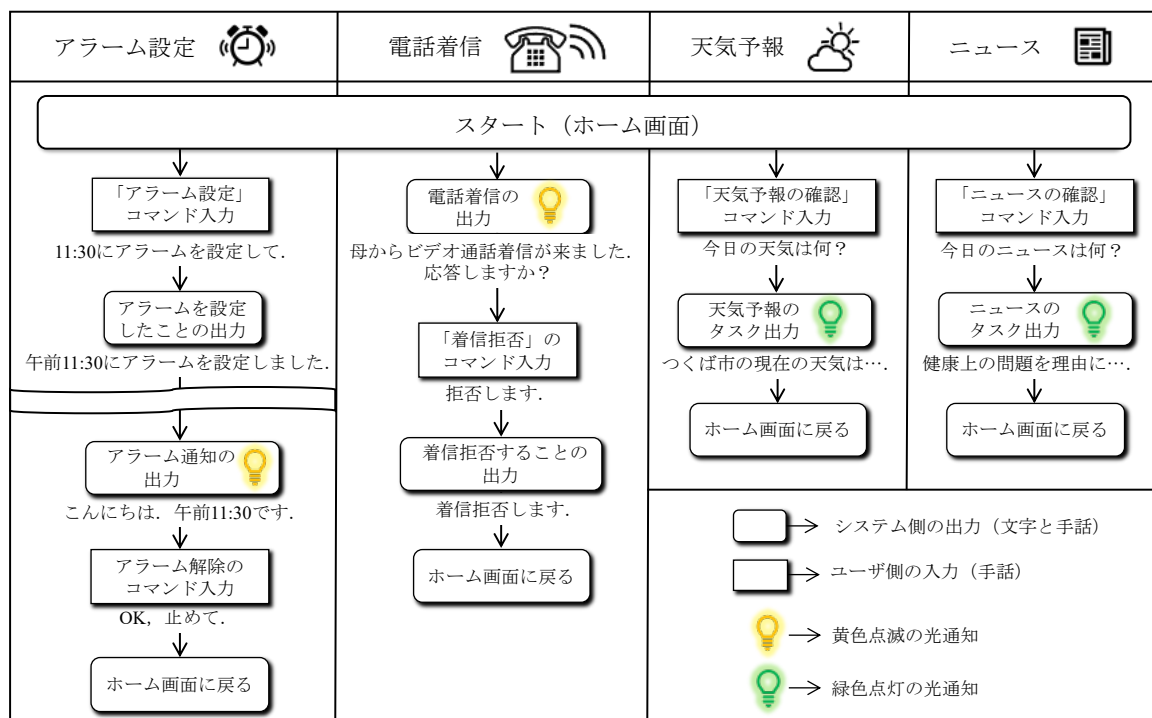


図 13：実験 1 で構築するシステムが実行する入出力内容 (タスク別)

2.2.3 実験 1 の環境

実験 1 の環境では、仕事をしながらシステムとインタラクションを行う場面を想定し、参加者正面の机上左側 45 度方向にシステム、正面方向に仕事用 PC を配置した（図 14）。これは、仕事をしつつシステムからの出力に気付くことを可能とし、常にシステムを視界に入れないようにしたものである。指示機器は、あるタイミングで、参加者のコマンドを促すものである。仕事用 PC には、画面のランダムな位置に数字/英語をランダムに表示させ、ユーザの文字キーボード入力に応じて、他のランダムな位置に別の数字/英語をランダムに表示させることを繰り返すプログラムを PsychoPy (v2021.1)[53]で組み入れた。また、設置したカメラのフレームレートは 50fps である。

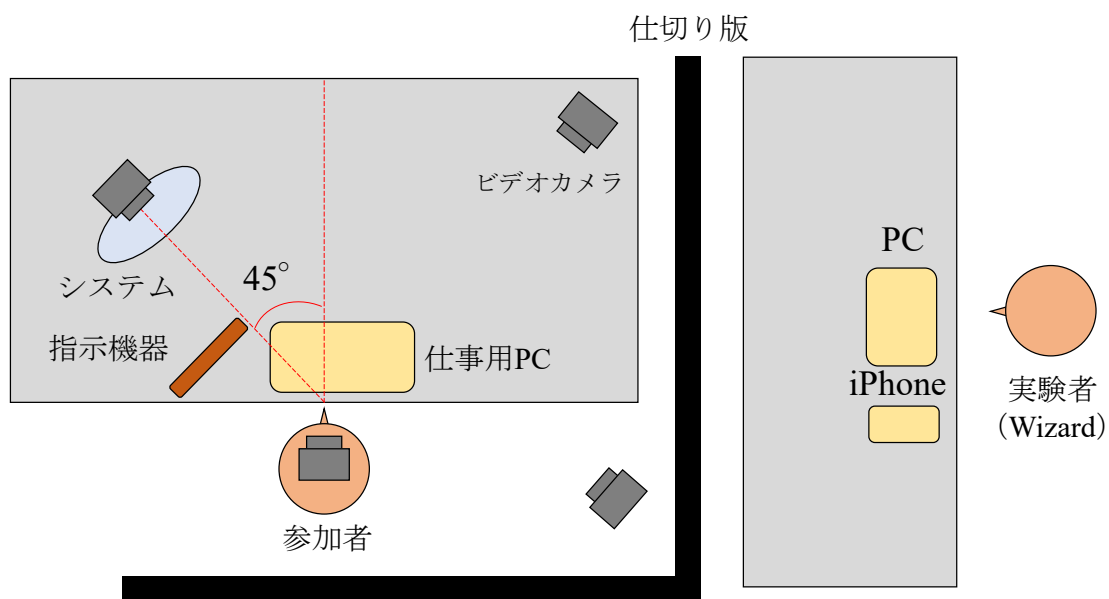


図 14：実験 1 の配置図

この環境で参加者と実験者(Wizard)それぞれにとって重要なポイントを以下に述べる。

参加者

- Wizard of Oz 法の特性上、参加者は手話認識システムがあると想定しており、人 (Wizard) がシステムを制御しているとは知らない。
- 実験中は、「仕事用 PC 画面のランダムな位置に表示される数字/英語をそのままキーボードで入力する」仕事に常に取り組みなければならない。
- 「アラーム設定」・「天気予報の確認」・「ニュースの確認」のコマンドを手話でシステムに実行させることができる。システム側からの天気予報/ニュースの回答が終了したことに気づいたら直ちに Space キーを押す。
- 工作中、任意のタイミングでシステムから「アラーム通知」、「電話着信」の出力があった場合、手話で「止める」ことのコマンドを出す。
- 実験中は、ヘッドストラップマウントに装着した GoPro カメラを頭に装着する。

実験者（Wizard）

- Wizard of Oz 法の特性上、実験者は連続手話認識システムを演じる際、参加者にシステムを制御していることを知られてはならない。
- 実験中は、iPhone と MacBook Pro を用いてシステムと LED 電球の制御を行う。
- 指示機器の操作について、任意のタイミングで参加者がコマンドを出すとき、実施するタスクの内容ならびにコマンド例を出し、参加者が把握したことを確認したら即画面を消すようにする。

また、参加者が実験中、システムに対して手話で実行可能なコマンドを表 11 に示す。

表 11：ユーザが実験 1 中に実行可能なコマンド

No.	コマンド内容
1	アラームを設定する 例「11:30 にアラームを設定して」
2	天気予報を確認する 例「今日の天気は何？」
3	ニュースを聞く 例「今日のニュースを教えて」
4	かかってきた電話着信を拒否する 例「拒否します」
5	鳴ったアラームを止める 例「OK, 止めて」

本番前、参加者にシステムの使い方、並びに 4 つのタスクそれぞれのシステムの挙動について説明を行った。さらに、参加者が手話を用いたコマンド入力に慣れるよう、本番同等のタスク実行の練習を行った。本番では、1 回の実験につき、図 13 に記載した 4 つのタスクを各 1 回実施した。そして、その実験を 2 回（光提示あり/なし）繰り返して行った。実験 1 の手順および、それぞれの所要時間を表 12 に示す。さらに、順序効果をなくすため、参加者ごとのタスク、光提示あり/なしの 2 条件の順序はカウンターバランスを取った。

表 12：実験 1 のプログラム

No.	概要	時間
1	教示・練習	約 15 分間
2	本番（1 回目）	約 10 分間
3	アンケート	約 2.5 分間
4	本番（2 回目）	約 10 分間
5	アンケート	約 2.5 分間
6	インタビュー	約 20 分間

2.2.4 評価項目

RQ1 の検証については、まず、「光呈示あり」と「光呈示なし」それぞれの条件に対するシステムのユーザビリティを調べるために、ユーザビリティの定量的評価に広く用いられている評価指標 System Usability Scale（以下、SUS と記す）[54]を用いて比較した。SUS は 10 項目からなるアンケートであり、5 段階のリッカートスケールによって回答させた。また、SUS の質問項目を表 13 に示す。集計方法について、奇数項目（肯定的な質問）では、回答番号から 1 を引いて、偶数項目（否定的な質問）では、5 から回答番号を引く。そして、すべての項目は 0（強くそう思わない）から 4（強くそう思う）で評価し、各項目の評価値の合計を 2.5 倍することで、0 から 100 のスケールへ変換できるようにした。

表 13：System Usability Scale のアンケート項目

No.	質問項目
1	このシステムを今後も繰り返し使いたい。
2	このシステムは必要以上に複雑だった。
3	このシステムは簡単に使えた。
4	このシステムを使えるようになるには、技術に詳しい人のサポートが必要だ。
5	このシステムは機能が良くまとまっていた。
6	このシステムは一貫性に欠けるところが多かった。
7	たいていの人は、このシステムをすぐ使えるようになるだろう。
8	このシステムはとても使いづらかった。
9	このシステムを使っていて不安を感じることはなかった。
10	このシステムを使い始める前に、多くのことを学ぶ必要があった。

また、システムからの通知に対する気づきを評価するために、通知から参加者が気づき反応するまでの時間を図 15 の通りにビデオを用いて計測を行う。

「呼びかけ」のときの反応時間について、「光呈示なし」の条件で反応開始ポイントはディスプレイの画面の変化、反応終了ポイントはユーザが画面に視点を向けたときとし、その間の時間と定義する。「光呈示あり」の条件でも同様のポイントであるが、画面変化より先に光が点灯した場合の反応開始ポイントは光点灯とする。

「回答終了」のときの反応時間について、「光呈示なし」の条件で反応開始ポイントはディスプレイの画面の変化、反応終了ポイントはユーザがボタンを押下したときとし、その間の時間と定義する。「光呈示あり」の条件でも同様のポイントであるが、画面変化より先に光が消灯した場合の反応開始ポイントは光消灯とする。

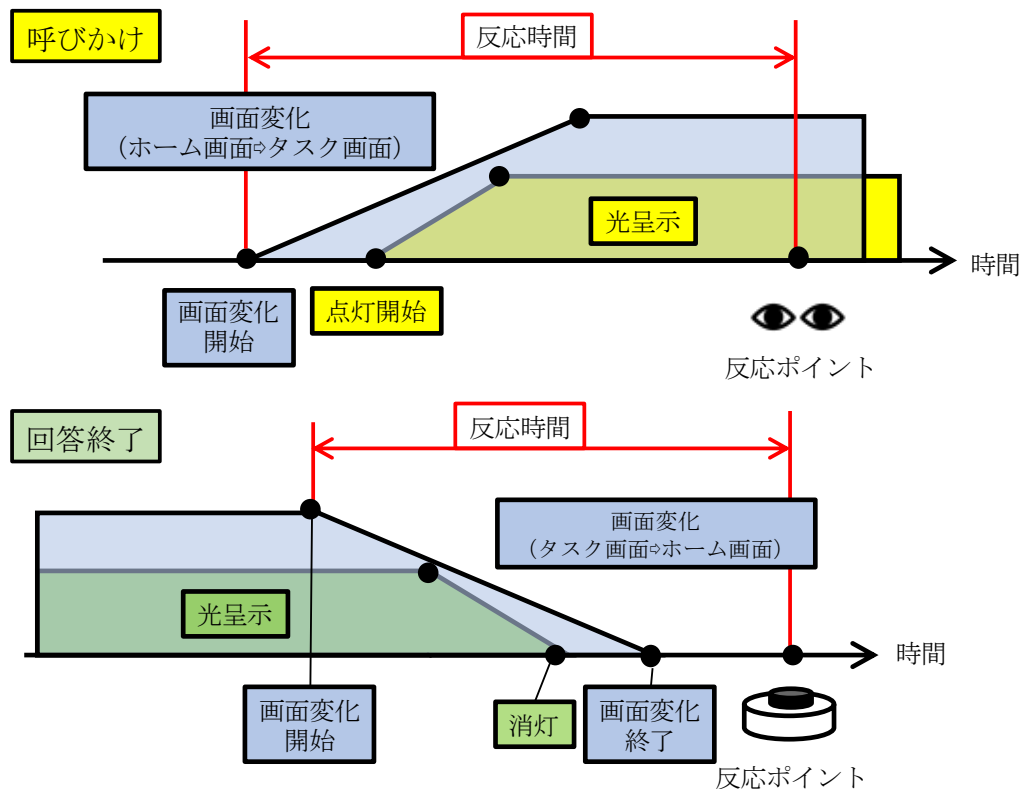


図 15：光呈示あり条件における反応時間の特定方法

RQ2 の検証については、参加者による手話/文字の必要性を調べる。実験終了後に手話/文字の必要性の有無を（1.必要，2.やや必要，3.どちらでもない，4.やや必要でない，5.必要でない）に分けた 5 段階のリッカートスケールによるアンケートを実施した。

RQ3 の検証については、参加者が手話でコマンドを入力する前にシステムに注視したかどうか調べる。そのために、全体のうちコマンドを入力する前の 5 秒間でデバイスに対し 1 回でも注視した回数の割合，ならびに注視開始から手話開始までの時間についてビデオを用いて計測を行った。なお、解析の対象となるデータについて、実験中に、システムから「アラーム通知」や「電話着信」の出力があり、ユーザは「止める」ことのコマンドを入力させた場面がある。ユーザが応答画面を見てからコマンドを入力することは、コマンドを入力する前に注視をしたかどうかを調査するための解析データにはならない。解析の対象データになるものは、ユーザが能動的にコマンドを入力するタスク、「アラーム設定」、「天気予報の確認」、「ニュースの確認」の 3 つである。

最後に、ビデオを用いた時間の解析については、ELAN¹⁷のツールを用いた。

17 The Language Archive, (June 5, 2021), ELAN, <https://archive.mpi.nl/tla/elan>

2.3 結果

2.3.1 System Usability Scale

実験終了後に調査した SUS 得点の結果についてまとめたものを図 16 に示す．SUS 得点の平均値は，「光呈示あり」条件で 80.67 (S.D.: 7.62), 「光呈示なし」条件は 68.96 (S.D.: 14.6) であった．SUS 得点における光呈示の有無に伴う有意差を調べるために，ウィルコクソンの符号付き順位検定を行った結果，「光呈示あり」条件の方が有意に高いことが認められた ($p < 0.05$).

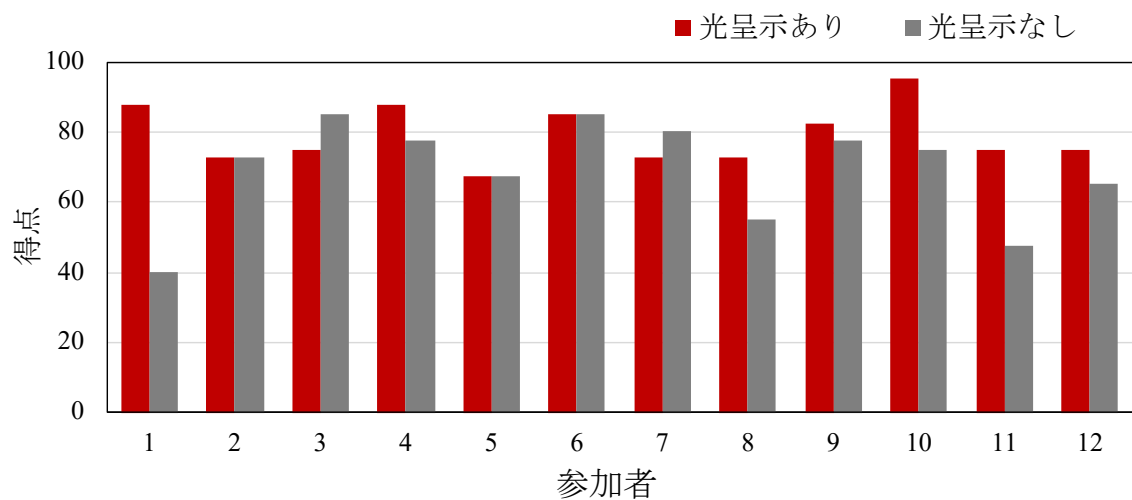


図 16：各参加者(N=12)の SUS 得点

2.3.2 反応時間

反応時間の結果についてまとめたものを図 17 に示す。電話着信・アラーム通知の「呼びかけ」に対する反応時間の平均値で「光呈示あり」条件は 0.91 s (S.D.: 0.35), 「光呈示なし」条件は 1.19 s (S.D.: 0.57) であった。呼びかけに対する反応時間における光呈示の有無に伴う有意差を調べるために、ウィルコクソンの符号順位検定を行った結果、「光呈示あり」条件の方が反応時間は有意に短くなることが認められた ($p < 0.01$)。一方、天気予報・ニュースの回答終了に対する反応時間の平均値で「光呈示あり」条件は 1.37 s (S.D.: 0.50), 「光呈示なし」条件は 1.91 s (S.D.: 1.22) であった。「回答終了」に対する反応時間における光呈示の有無に伴う有意差を調べるために、ウィルコクソンの符号順位検定を行った結果、「光呈示あり」と「光呈示なし」の間に有意差は認められなかった ($p > 0.05$)。

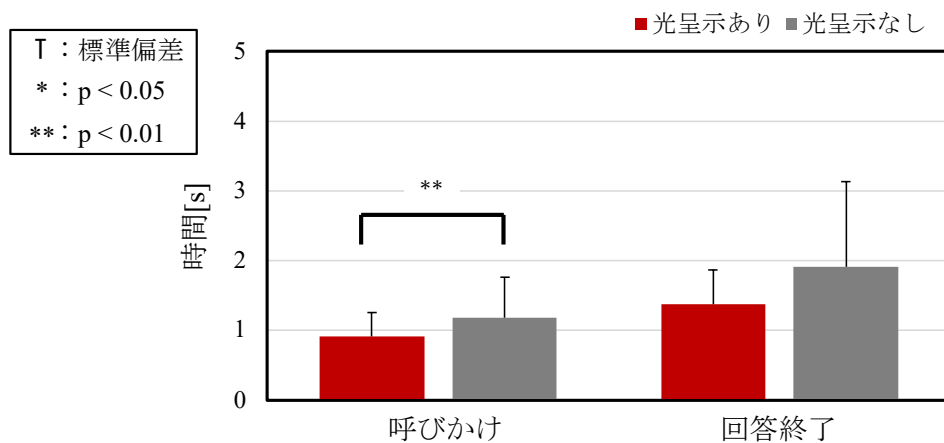


図 17：システムからの応答に対する反応時間

2.3.3 手話と文字に対する必要性

参加者 12 名に手話と文字それぞれに対する必要性を 5 段階のリッカートスケールによるアンケートで問い合わせた結果についてまとめたものを図 18 に示す。手話について、「1:必要」は 4 名、「2:やや必要」は 2 名、「3:どちらでもない」は 3 名、「4:やや必要でない」は 2 名、「5:必要でない」は 1 名であった。手話を必要としない(4,5 と回答した)参加者は 3 名いて、手話歴が短い人から順番に、3 年(1 位)、5 年(2 位)、15 年(5 位)であった。一方、文字について、「1:必要」は 9 名、「2:やや必要」は 3 名であった。

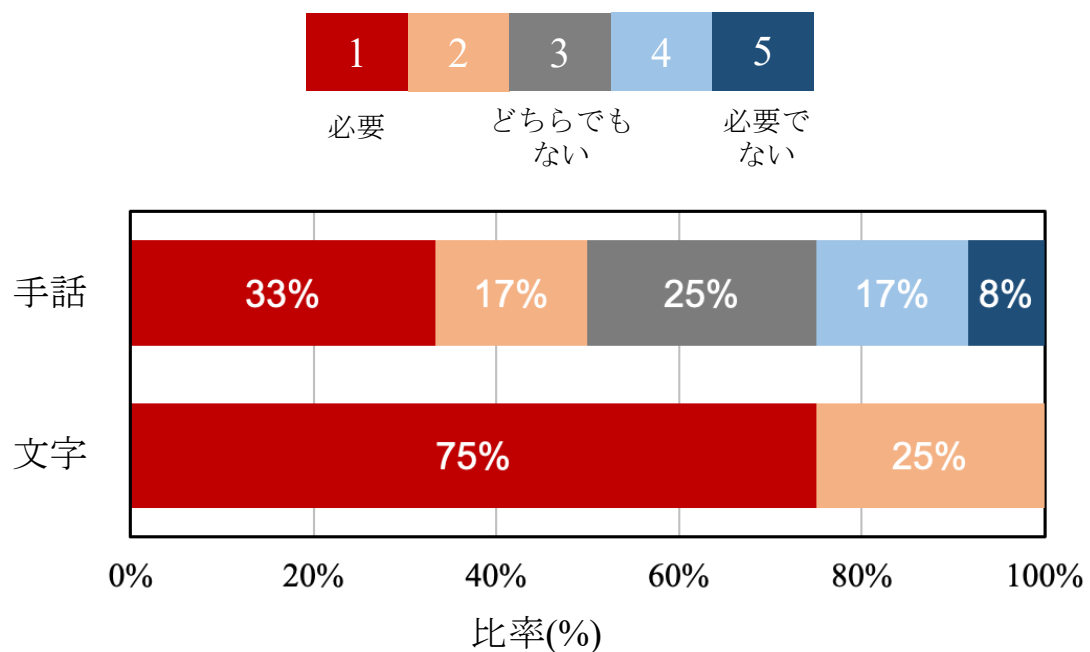


図 18：手話と文字それぞれに対する必要性のアンケート結果

2.3.4 注視入力

実験中、実施するタスクを示す指示機器の画面を消すのが遅れてしまい、参加者が、ながら読みしながらコマンドを入力するパターンがあった。そのデータはシステムに対して注視したかどうか調べるのに不適なデータとなるため、解析の対象となるデータから削除した。参加者(N=12)がシステムにコマンドの入力を行った回数について、「アラーム設定」は 23 回、「天気予報の確認」は 24 回、「ニュースの確認」の 22 回、合わせて 69 回である。コマンドを入力する前の 5 秒間でシステムに対し 1 回でも注視した回数の割合、ならびに注視開始から手話開始までの時間についてまとめたものを表 14 に示す。高い割合でコマンドを入力する前にシステムに注視を行っていたことは明らかである。

表 14：コマンド入力前に注視した割合と注視開始からコマンド入力開始までの時間

タスク	割合(%)	平均値±標準偏差(s) (最小値, 最大値)
アラーム設定	100	0.76 ± 0.61 (0.20, 3.18)
天気情報の確認	100	0.43 ± 0.23 (0.10, 1.08)
ニュースの確認	86.4	0.59 ± 0.44 (0.20, 2.08)
総計	93.4	0.59 ± 0.47 (0.10, 3.18)

本番 1 回目と本番 2 回目それぞれ 1 回、合わせて 2 回ユーザが「ニュースの確認」のコマンドを出す時の挙動についてまとめたものを図 19 に示す。挙動について、手振り、手話、システムへの注視の 3 つに分類した。また、手話開始の時間を 0 s とし、全体的に参加者のシステムへの注視は手話開始より早く始まっていることを示した。P1 と P5 については、1 回不適なデータがあったため除外した。図 19 より、手話開始前に手振りの動作を行っていたのは P3, P8, P9 の 3 人であった。この 3 人は、スマートスピーカーの使用経験がありウェイクワードを使うべきであることを知っていた。また、今回の実験においてシステムに話しかける前に明示的なアクションを起こす必要があると考えていたことはインタビュー調査でわかった。

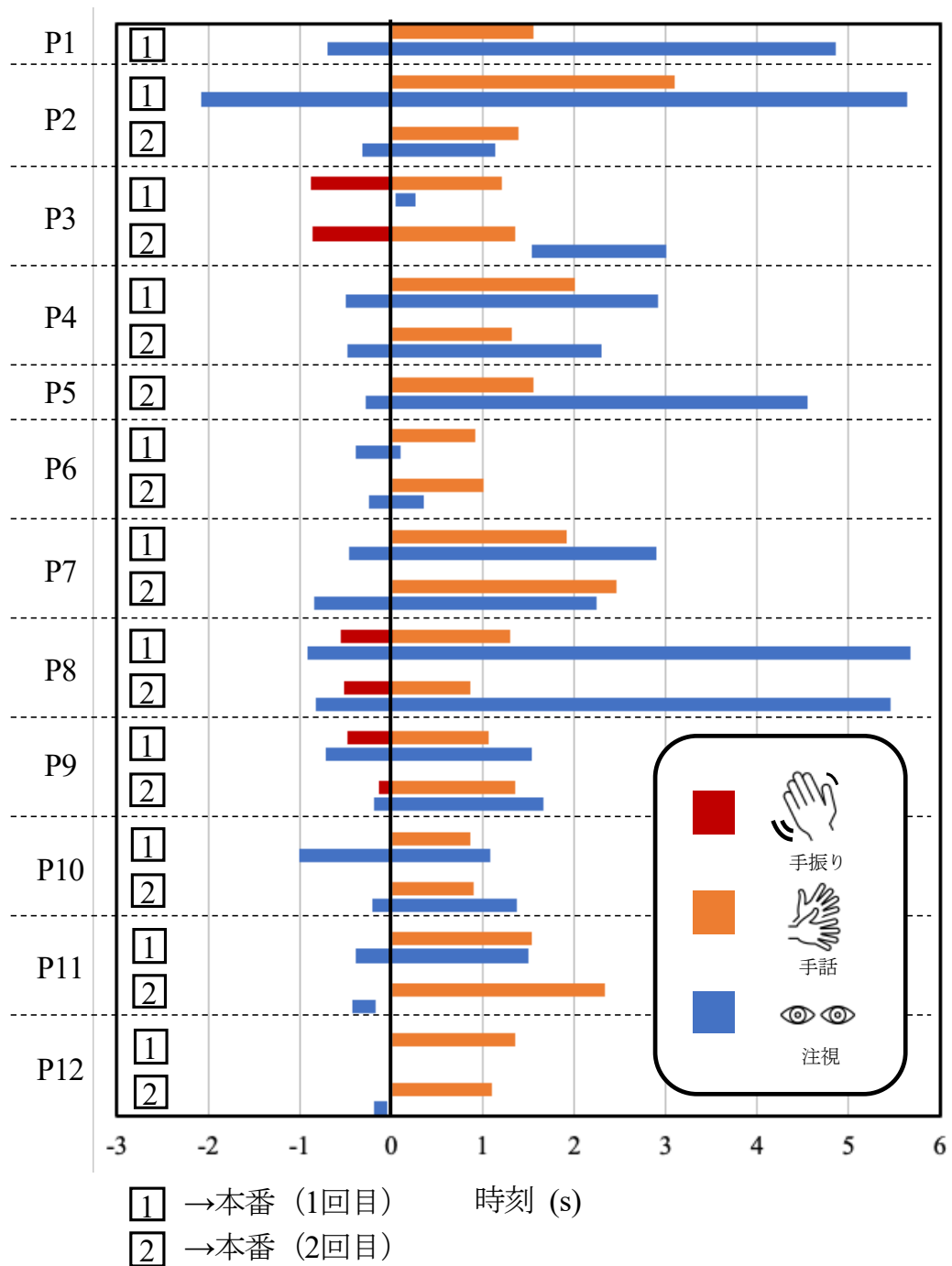


図 19 : ニュース確認のコマンドを入力する時の各参加者の動作継続時間

2.4 考察

RQ1:光を用いた応答のユーザビリティ

2.3.1 節で述べた結果から、ろう・難聴者にとって、システムからの通知に気づく方法として、光を用いる方がユーザビリティは向上することが示唆される。また、2.3.2 節より、呼びかけ（電話着信・アラーム通知）に対する反応について、光を用いた方が有意に反応時間は短くなったことから、システムからの通知に気づきやすくなったことがいえる。

参加者からは、「家のインターフォン通知も光で知らせてくれるなど、光を使った通知方法には馴染みがあるのでシステムにも光を付与した方が印象良い。」「作業に集中していても光通知なら気づくことができる。」のコメントがあった。しかし、「音を頼りに生活しているため、光手段に違和感があった。」とのコメントもあり、聴覚機能を活用し生活を送っている人には不向きな場合もある。

図 16 より、P3, P7 は光を用いた方がユーザビリティは低下した。そういった参加者からは、「システムの画面変化だけで気づいたため、光を用いる必要性は感じられなかった。」のコメントがあった。これは、光を用いずとも、ろう・難聴者だから反応できる場合が生じたためと考えられる[55]。本実験環境では、システムをユーザ正面の机上左側 45 度と周辺視野内に設置した。すなわち、周辺視野の範囲内にシステムが設置されていたため、システムを見なくとも画面変化に気づけたため光を用いてもユーザビリティは向上しない参加者がいることを考慮しなければならない。

一方、回答終了（天気予報・ニュース）に対する反応について、2.3.2 節で述べた通り、光を用いても反応時間に関して有意差はなかった。また、「回答がいつ終了したかシステムを見なくてもわかったので便利に感じた」と肯定的なコメントがある一方、「自分から求める情報（天気予報・ニュース）で光呈示は必要ない」、「光が明るすぎる」と否定的なコメントをする参加者がいた。以上から、気づきやすさは変わらないが、光呈示の方法について光露出を控えめにするなど改善することで、ユーザビリティは向上する可能性がある。

LED 電球について「システムと一体化になってないので、光通知に煩わしさを感じた。」とコメントする参加者がいた。光通知のために LED 電球を用いたが、本来は照明器具として使われるものである。それゆえに、光装置の形状はシステムの形状に合わせることを検討する必要がある。

RQ2:手話と文字の最適な表示方法

2.3.3 節から、文字の表示はユーザ属性問わず全員が必要とするのに対し、手話の表示についての必要性は参加者特性による影響を受けることがわかる。また、手話歴が浅い人ほど手話は必要と考えない傾向にあることがわかる。

手話を必要としない参加者からは、「手話がわからない」、「手話と文字の両方が同時に映し出されると情報処理に苦しむ」とのコメントがあった。一方で、手話を必要とする参加者からは、「手話があることでシステムからの応答内容が頭に残りやすくなる」とのコメントがあった。また、参加者から「仕事をしながらだとタスク画面を見る場合、流れる手話より残る文字の方がいい」とのコメントがあった。

聴者にとって、スマートスピーカーのような VUI とのインタラクションは、アイズフリーで行える利点がある[56]。そのため、システムを使う上で、他の作業をしながらインタラクションが交わされることが頻繁に発生する。しかし、ろう・難聴者の場合、音声情報を獲得できず画面を注視してしまうことからアイズフリーの利点は失われる。それを補完するために、ろう・難聴者が少し目を離しても認知できる文字情報が必要になることが予想される。この課題の解決方法として、ユーザの視線を外したときに手話動画は止まり、視線を戻したときに再び手話動画を再生する仕様にする方法などが考えられ、今後検討の必要がある。

RQ3:ウェイクアップにおける注視の有効性

2.3.4 節の結果より、参加者は手話で話しかける前に画面を注視する傾向にある。本実験において、ウェイクアップ方法についての指示は特に行っていない。それにも関わらず、参加者は高い確率で自然にシステムを注視した。

一方で、図 19 より、手振りを行う参加者は 12 人のうち 3 人(P3, P8, P9)いた。しかし、手振りをウェイクアップの方法として用いる場合、電話着信やアラーム通知のようにユーザが予測できないタイミングで、他人と話しているときの手話が手振りとして認識されることが懸念される。また、コマンドを出すたびに手振りを行うより、注視を用いる方が自然な会話ができることが考えられる。

システムを注視してからコマンド入力開始までの最大時間は 3.18 秒であった(表 14)。ここで、Alexa がユーザからの応答を待つ時間制限は 8.0 秒間である[57]。したがって、ろう・難聴者ユーザが注視してから、システムの待ち時間内にコマンドを入力することができる。以上から、注視はウェイクアップ方法に有効であることが示唆される。

第3章 ウェイクアップ方法とフィードバック方法の検討

3.1 検証する研究課題

今章で検証する研究課題は、RQ4-RQ6である。これらは、システムとの会話開始時における、ろう・難聴者ユーザにとって最適な入出力方法を網羅した研究課題である。最初に、ろう・難聴者ユーザはシステムのフィードバックが必要であるか調査する(RQ4)。続いて、ろう・難聴者ユーザからシステムを起動させるために最適なウェイクアップ方法を検討する(RQ5)。そして、システムがユーザのウェイクアップを検出し、コマンド受付可能であることを示す、最適なフィードバック方法を検討する(RQ6)。RQ4 から RQ6 を検証するために、Wizard of Oz 法に基づいた、様々なフィードバックを提示する会話システムを構築し、評価実験（以下、実験2と記す）を行う。なお、実験2は筑波技術大学研究倫理委員会の承認を受けている（承認番号: 2021-12）。

3.2 実験2の方法

3.2.1 参加者

実験2の参加者は本学学生のろう・難聴者24名（20～23歳）である。また、参加者の属性による実験結果への影響の解析も同時に行うため、特性を調査した。性別や補聴器/人工内耳の装用の有無、手話歴、手話レベル、日常会話における声の使用、アイデンティティ、AIアシスタントの使用経験それぞれの結果をまとめたものを図20に示す。

性別については、24名のうち、男性14名、女性10名であった。また、日常生活における補聴器/人工内耳の装用の有無について、「装用している」と回答した参加者が18名、「装用していない」と回答した参加者が6名であった。次に、手話歴について、5年単位で刻み分類した結果、生まれてからずっと手話を使っている参加者が6名、15年以上20年未満手話を使っている参加者が8名、10年以上15年未満手話を使っている参加者が3名、5年以上10年未満手話を使っている参加者が3名、そして5年未満手話を使っている参加者が4名であった。続いて、手話のレベル（特性の決め方は2.2.1節を参照）について、「声なし日本手話を読み取れる」と回答した参加者が11名、「声なし中間手話を読み取れる」と回答した参加者が8名、「声あり中間手話を読み取れる」と回答した参加者が5名であった。また、日常生活で取るコミュニケーションに声を使うかどうか尋ねた結果、「使う」参加者が20名、「使わない」参加者が4名であった。最後に、アイデンティティを尋ねた結果、「ろう者」と回答した参加者が8名、「難聴者」と回答した参加者が6名、「聴覚障害者」と回答した参加者が9名、そして「考えたことがない」参加者が1名であった。

なお、AIアシスタント使用経験については、（1.毎日、2.週に数回、3.月に数回、4.月1回以下、5.使わない）の5段階評価で回答してもらった。その結果、「5.使わない」と回答した参加者が、20名と一番多く、全体の83%を占めていた。実験2の参加者は、ろう・難聴者ユ

ーザのうち、AI アシスタントを使用している人は非常に少数であることの報告[50]と同様の傾向を示している。そして、「1.毎日」と回答した参加者は2名であり、使用場面については、「料理するときのタイマー設定で Siri を使う」、「アプリを開くとき、また運転するときに車載音声アシスタントを用いる」であった。「3.月に数回」と回答した参加者は1名、「4.月1回以下」と回答した参加者は1名であり、使用場面については、「何か調べたいとき Siri を使う」、「今日の天気や気温を知るために使う」であった。

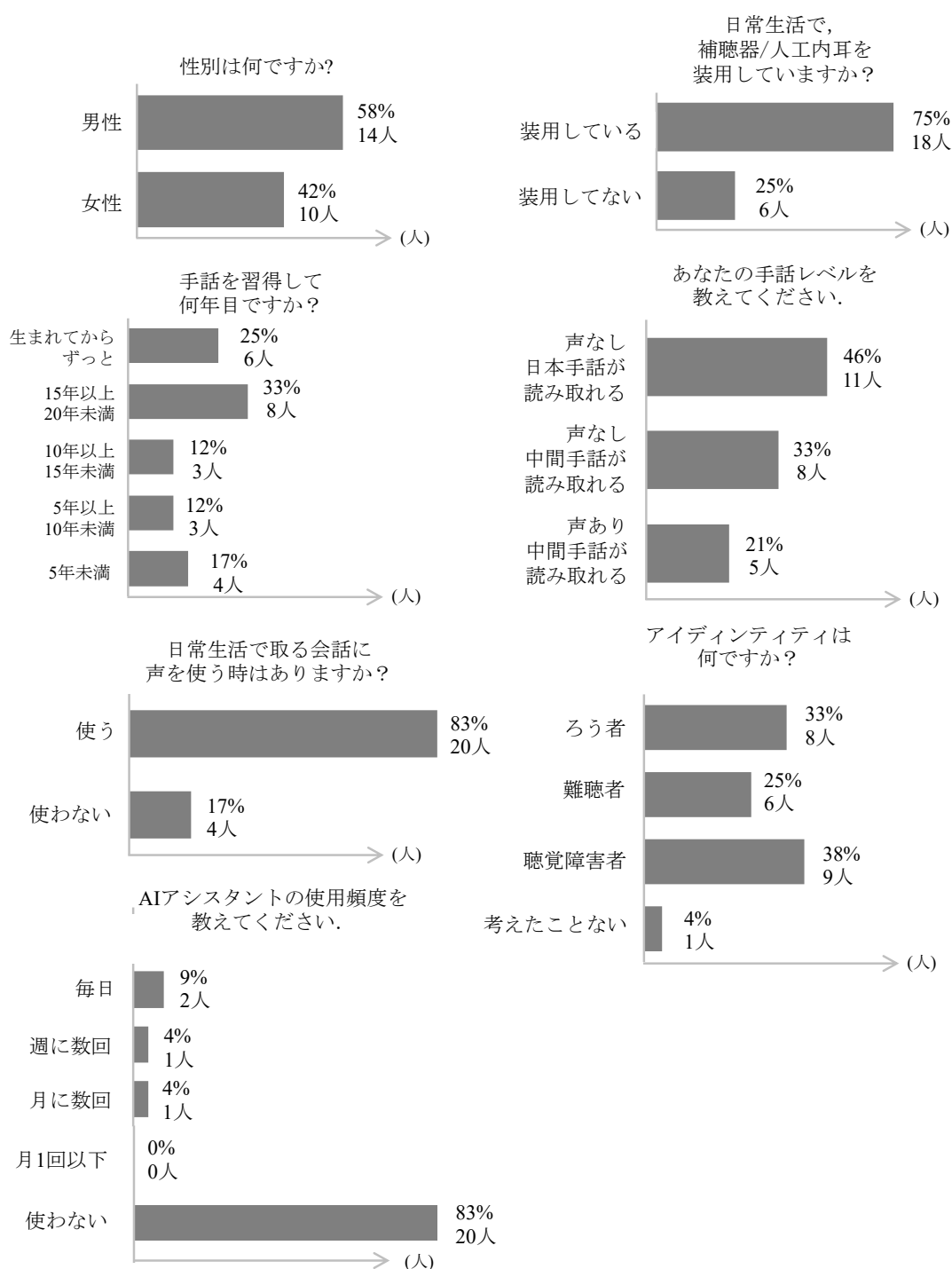


図 20：実験 2 の参加者(N=24)の特性

3.2.2 実験条件

研究課題について検証を行うため、ウェイクアップ方法とフィードバック方法のそれぞれに設定した条件それぞれを表 15, 表 16 に示す。なお、それぞれに振り当てた番号について、ウェイクアップ方法は「ユーザ側の入力 = Input」の頭文字の"I", フィードバック方法は「システム側の出力 = Output」の頭文字の"O"とした。

ウェイクアップ方法

I1.「注視」条件

前章で有効性が示唆された注視条件に、他のウェイクアップ条件と比較し、より良い方法を明らかにする。ここで、ろう・難聴者のためのウェイクアップに関する研究の報告[49]によれば、一番評価の高いウェイクアップ方法は、サインネーム条件であった。しかし、その比較に注視条件は検討されてなかった。そのため、実験 2 においてウェイクアップ方法に注視条件を加えることにより、ろう・難聴者の満足度は変わるか調査を行う。

I2.「サインネーム」条件

システムには Alexa が搭載されるものとし、そのサインネームとして、[12]の研究を参考に、サインネームの方法は、「システムに向けて"A"指文字を挙げ、その指文字を"X"の形で空描きする」ことである。

I3.「手振り」条件

システムの方に手のひらを出し、左右に往復運動させる動作である。研究[49]の報告によれば、意図しない手振り動作について、ウェイクアップとしてシステムに認識されてしまうことの懸念が生じる。そこで、相手の行動を喚起・促進し、インタラクションの開始および進行に関わる注視条件と比較することで、新たな知見が得られると考えた。

フィードバック方法

O1.「青いバー表示+明度低」条件

Echo Show は、ウェイクワードを認識し、ユーザからのリクエストを処理する状態として、青いバーを画面の下に表示する¹⁸。また、青いバー表示を目立たせるよう、画面の明るさを下げている。こういった、従来の反応提示方法が、ろう・難聴者ユーザにとって最適なフィードバックであるか探るために、ベースラインとして条件に加えた。

O2.「O1 に手話表示を追加」条件

ろう・難聴者ユーザが手話をメインにシステムと会話する場合、システムからも手話で出力することは、ろう・難聴者にとって最適であることが推測される。そのため、「何？」の手話で反応を提示することの条件を加えた。

O3.「O1 に首振り運動を追加」条件

Echo Show10 のような首振り運動はろう・難聴者にとって最適なフィードバックであるか調べるために、首振り運動で反応を提示することの条件を加えた。

18 Alexa および Alexa 対応端末に関する FAQ, https://www.amazon.co.jp/gp/help/customer/display.html?ref_=help_search_1-1&nodeId=201602230&qid=1641309401742&sr=1-1

O4.「変化なし」条件

ろう・難聴者ユーザのフィードバックに対する必要性を調べるために、O1-O3 と対照的にフィードバックを提示しない、すなわちシステムに変化しないことの条件を加えた。

表 15：ウェイクアップ条件

No.	条件
I1	注視
I2	サインネーム
I3	手振り

表 16：フィードバック条件

No.	条件
O1	青いバー表示+明度低
O2	O1 に手話表示追加
O3	O1 に首振り運動追加
O4	変化なし

3.2.3 システムの構築

実験 2 で構築するシステムの外観を図 21 に示す。システムの基本構成は、iPad（第 9 世代）、BAOSHISHAN 社の直径 20 cm 耐荷重 20 kg の電動回転台の 2 つとした。iPad はシステムのディスプレイとして用いた。なお、iPad は、Echo Show10 のディスプレイ用コンテンツを模擬し、Microsoft 社製 Microsoft PowerPoint 2019 でプレゼンテーションを作成し表示したものである。iPad（第 9 世代）の画面サイズは 10.2 インチ、Echo Show10 の画面サイズは 10.1 インチであり、画面サイズはほぼ一致している。iPad のディスプレイ周りのカバーは、FLASHFORGE 社の FFA-103 型 3D プリンタ Adventure 3 で出力し、組み立てを行った上で使用した。また、システムからの応答に手話も表示するため、ディスプレイに手話者の動画と合成を行った。



図 21：実験 2 で構築するシステムの外観

Wizard 側から遠隔でシステムを制御するために、MacBook Pro と電動回転台リモコンを用いた。MacBook Pro は、Sidecar を用いることでデスクトップを拡張するディスプレイとして遠隔で iPad の制御が可能である。電動回転台リモコンは、遠隔で電動回転台の回転速度、方向角度の遠隔制御が可能である。

システムの挙動内容を図 22 に示す。フィードバック（O1-O3）を提示するためのシステムの仕様を次に説明する。「O1：青いバー表示+明度低」の出力機能については、ホーム画面の明るさを -25% に設定し、画面下に Echo Show8 のディスプレイに表示される青いバーを模擬表示した。「O2：O1 に手話表示追加」の出力機能については、O1 の画面変化に加えて、手

話者が表現する「何？」の手話動画を合成し表示した。なお、手話者はホーム画面には登場せず、参加者のウェイクアップ動作に反応して手話者がフェードイン登場（継続時間：1.5 秒間）しながら「何？」の手話を行い、その後は静止するようにした。「O3：O1 に首振り運動追加」の出力機能については、O1 の画面変化に加えて、電動回転台の動作を行った。なお、システムは 11.25 度、0.62 秒間の単方向の首振り運動を行い、その後は静止するようにした。タスク画面について、天気予報とニュースの情報をグラフィックに文字と合わせて 6 秒間表示し、ホーム画面に戻るようにした。



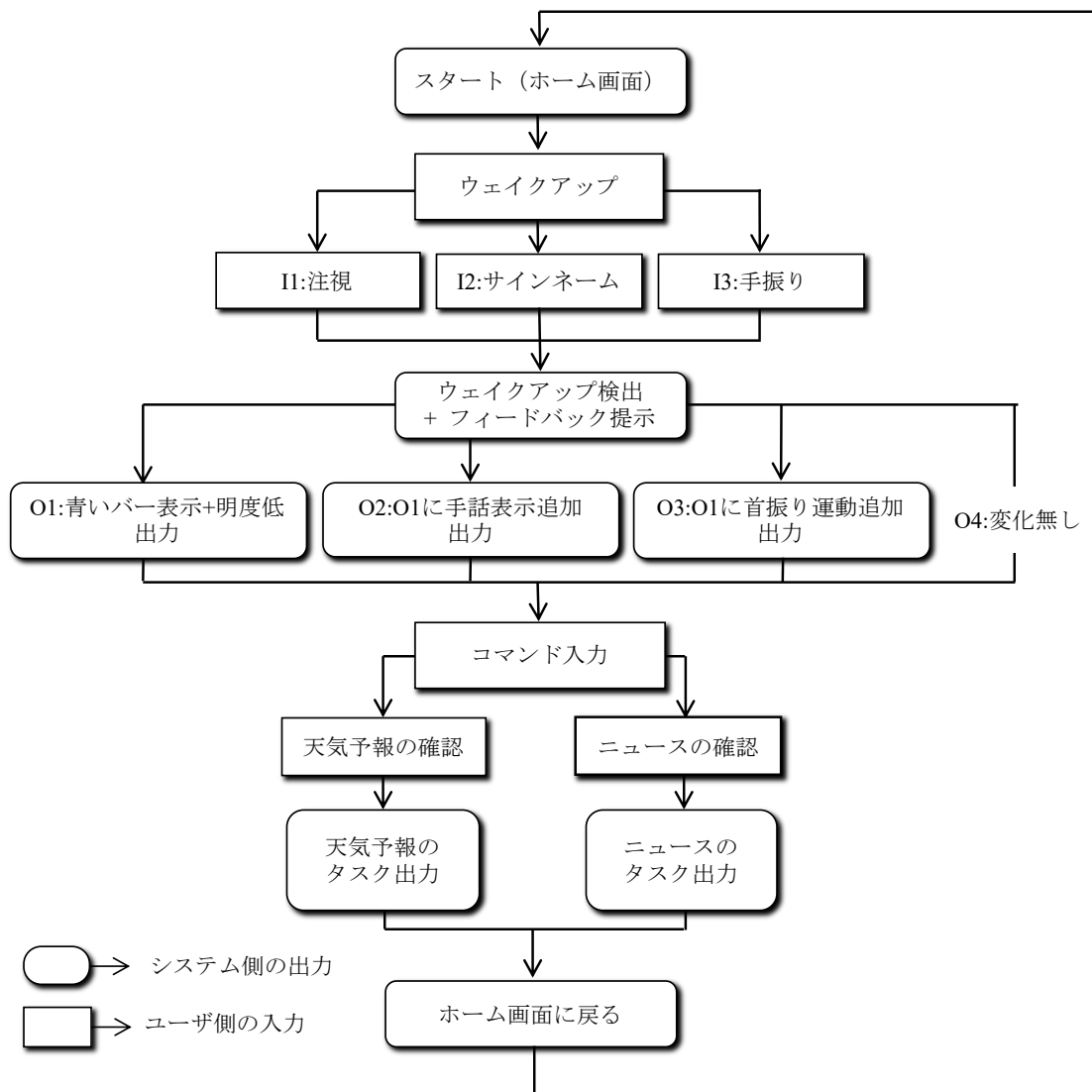


図 23：実験 2 で構築するシステムが実行する入出力内容

3.2.4 実験 2 の環境

実験 2 の環境では、仕事をしながらシステムとインタラクションを行う場面を想定し、参加者正面の机上右側 45 度方向にシステム、正面方向に仕事用 iPad、机上右側 45 度左側方向にアンケート画面と指示画面の 2 画面を表示させる iPad を配置した（図 24）。なお、アンケートについて、Microsoft Forms で各条件におけるウェイクアップ条件とフィードバック条件の両方に対する満足度質問、それぞれの満足度順位の質問、そしてフィードバックに対する必要性の質問を 7 段階リッカートスケールで作成した。指示画面は、次に参加者に行ってもらうウェイクアップ方法、システムから提示されるフィードバックを示すものとし、Keynote で作成した。この配置は、参加者が指示画面を確認し（左側）、仕事をしつつ（正面）、システムと会話し（右側）、アンケートに満足度を回答し（左側）、そして指示画面を確認する（左側）ことに戻るといったサイクルにした。仕事内容はタイピングゲーム¹⁹のプレイとした。また、設置したビデオカメラのフレームレートは 50fps である。

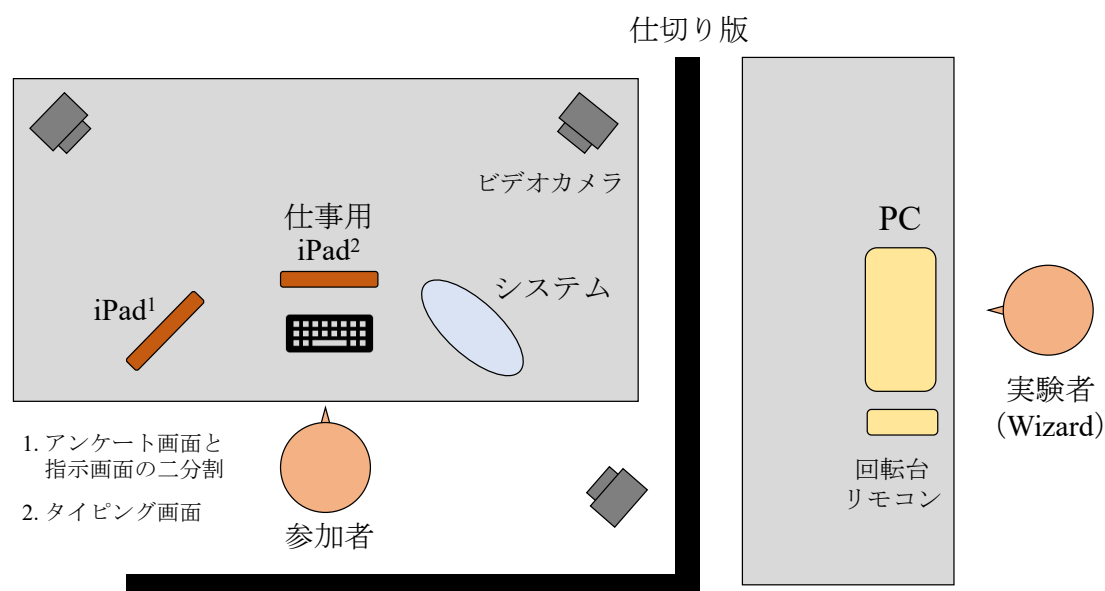


図 24：実験 2 の配置図

この環境で参加者と実験者(Wizard)それぞれにとって重要なポイントを以下に述べる。

参加者

- Wizard of Oz 法の特性上、参加者は手話認識システムがあると想定しており、人 (Wizard) がシステムを制御しているとは知らない。
- 実験中は、タイピング画面に表示される数字/英語をそのままキーボードで入力することの仕事を常に取り組みなければならない。
- 指示画面に表示される、次に行うウェイクアップ方法と、システムから提示されるフィードバックを確認する。
- 3 字程度、タイピングをしてから、システムをウェイクアップし、システムのフィードバ

19 Moku Uchi 無料オンラインタイピングゲーム, <http://www13.plala.or.jp/hydrogen/piano.html>

ックが提示され次第、「天気予報の確認」と「ニュースの確認」のコマンドから一つ選択し、システムに手話で入力する。

- システムの「天気予報」や「ニュース」の回答が終わったことを確認したら、「ウェイクアップ方法とシステムのフィードバックの両方に対する満足度」のアンケートに回答する。

実験者(Wizard)

- Wizard of Oz 法の特性上、実験者は手話認識システムを演じる際、参加者にシステムを制御していることを知られてはならない。
- 実験中は、システムの制御を行う。
- 指示機器の操作について、任意のタイミングで参加者にあるタスクを出させるときに、実施するタスクの内容ならびにコマンド例を出し、参加者が把握したことを確認したら即画面を消すようにする。

そして、本番前にシステムの使い方、システムの挙動について説明を行った。さらに、ウェイクアップおよび手話を用いたコマンド入力に慣れるよう、本番同等のタスクを行う練習を行った。参加者には、ウェイクアップの3条件とフィードバックの4条件を組み合わせた計12条件をそれぞれ1回、すなわち繰り返し12回で行ってもらった。実験2のプログラムおよび、所要時間を表17に示す。また、順序効果をなくすため、12条件の順序はラテン方格法を用いて決めた。

表 17：実験2 プログラム

No.	概要	時間
1	教示・練習	約 15 分間
2	本番※	約 36 分間
3	アンケート	約 10 分間
4	インタビュー	約 10 分間

※{(実践：2分間)+(評価：1分間)}×12条件

ウェイクアップ方法の3つの条件(I1-I3)、フィードバック方法の4つの条件(O1-O4)から最適な組み合わせ条件を調べるために、それぞれの7段階満足度(3.非常に満足, 2.満足, 1.やや満足, 0.どちらでもない, -1.やや不満, -2.不満, -3.非常に不満)のリッカートスケールによるアンケートを実施した。その回答データ(N=24)から、より良いウェイクアップ条件とフィードバック条件を明らかにする。

そして、システムとの会話における有効性を評価するために、参加者とシステムの挙動について ELAN²⁰のツールでビデオ解析を行った。ここで、参加者とシステムのやりとりの挙動タイミングの詳細について、図 25 に示す。



参加者の挙動ポイント

- 20 The Language Archive, (June 5, 2021), ELAN, <https://archive.mpi.nl/tla/elan>

システムの挙動ポイント

- S_1 : ホーム画面からフィードバックまでの変化が始まる時刻 (O1 と O2 と O3)
 - O1 (青いバー表示+明度低) → 青いバーの表示が始まる時刻とした.
 - O2 (O1 に手話表示追加) → 手話者が登場し始まる時刻とした.
 - O3 (O1 に首振り運動追加) → 首振り運動が始まる時刻とした.
- S_2 : ホーム画面からフィードバックまでの変化が終わる時刻 (O1 と O2 と O3)
 - O1 (青いバー表示+明度低) → 青いバー表示と明度が低くなるための変化を要する時間は 0 s に近似すると見なす ($S_1 = S_2$).
 - O2 (O1 に手話表示追加) → 手話者が表現する「何?」の手話が終わり静止し始めた時刻とした. これは, 参加者とシステムが手話で会話することにより, 参加者はシステム側からの手話を確認してからコマンド入力を行ったかどうか明らかにするためである.
 - O3 (O1 に首振り運動追加) → 首振り運動が終わり静止する時刻とした.

これらの挙動ポイントをもとに, 参加者の挙動に関する解析時間は以下の 3 つの通りとする. また, I2×O3 条件の時の参加者の挙動に関する解析時間を図 26 に, I1×O1 条件の時の参加者の挙動に関する解析時間を図 27 に示す. なお, 図 26, 図 27 の時間軸について, 手話開始の時間を 0 s とした.

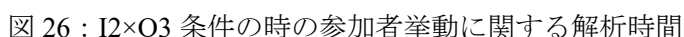
1. 手振りの動作時間 ($T_2 - T_1$)

I1 (注視) 条件は, 参加者がシステムを注視し始めた瞬間がウェイクアップ動作と見なしたため, 時間 ($T_3 - T_2$) は 0 s である. そして, I2 (サインネーム) 条件は, "A" の指文字を "X" の形で空描きする動作である一方, I3 (手振り) 条件は, 手のひらを前方に出し左右に往復運動するといった, 終わりのタイミングは特に決まってない動作である. そこで, 著者は, I3 (手振り) 動作の時間がシステムの提示するフィードバック条件(O1-O3), または, 参加者の属性によって変わる可能性を懸念し, 明らかにするために調査する.

また, 2.3.4 節より実験 1 の参加者がシステムに手話でコマンド入力する前に, 注視する確率が高いことが明らかになっている. また, 1.3 節より, 手振りは, 意図しない手振り動作が認識されてしまうここで, 参加者がウェイクアップの I3 (手振り) 条件を実践するとき, 画面を見るタイミングはウェイクアップ動作の前後のどちらの割合が高いのか調査する. そして, 注視と手振り動作の順番について, 注視が先の場合は, 注視開始から手振り動作開始までの時間, 手振り動作が先の場合は, 手振り動作開始から注視開始までの時間それぞれの平均値と標準偏差, 最大値と最小値を解析する. また, 注視と手振り動作の両方を行った時間の平均値と標準偏差, 最大値と最小値を解析する. これらの解析は, システムを開発する際に, 注視と手振り動作の両方をウェイクアップ動作として認識する場合, どちらが先の認識対象とすべきか, また, その許容時間は何秒以上とすべきかを設計ガイドラインとして提示できるためである.

1.3 節で述べたように、参加者がシステムとインタラクションを行う際に、ウェイクアップとコマンド入力を行う必要がある。インタラクションの効率性が上がるほどユーザビリティが向上することを踏まえ、最も時間が短いウェイクアップ条件は何か、そして時間はフィードバック条件(O1-O3)に影響されるか調査する。また、フィードバックがなされない時の効率性も知るため、O4（変化なし）条件の時間も合わせて解析を行った。

システムが提示するフィードバックそれぞれの条件(O1-O3)に対する参加者の挙動を明らかにする。また、フィードバック条件(O1-O3)の違いが時間 ($T_3 - S_2(S_1)$) に及ぼす影響も明らかにする。システムが O2 (O1 に手話表示追加) 条件のフィードバックを提示した場合、参加者は手話者が登場してからか、もしくは「何？」手話の表現が終わってからコマンド入力されるか明らかにする。そして、システムが O3 (O1 に首振り運動追加) 条件のフィードバックを提示した場合、参加者はシステムの首振り運動が開始してからか、もしくはシステムの首振り運動が終わってからコマンドを入力するか明らかにする。



また、システムの挙動に関する解析時間は以下の 2 つの通りとする．そして、I2×O3 条件の時のシステムの挙動に関する解析時間を図 28 に、I1×O1 条件の時のシステムの挙動に関する解析時間を図 29 に示す．

1. ウェイクアップの動作開始からシステムの反応開始（変化開始）までの時間（ $S_1 - T_1$ ）

システムの挙動が、参加者の行うウェイクアップ方法、また Wizard が制御するフィードバック方法によって、どのように変わるか明らかにする．

2. ウェイクアップの動作終了からシステムの反応開始（変化開始）までの時間（ $S_1 - T_2$ ）

時間（ $S_1 - T_2$ ）は、参加者がコマンドを入力するまでの待ち時間である．その時間の長さが、システムの挙動が、参加者の行うウェイクアップ方法、また Wizard が制御するフィードバック方法によって、どのように変わるか明らかにし、また、参加者の満足度の結果にどう影響するか調査する．

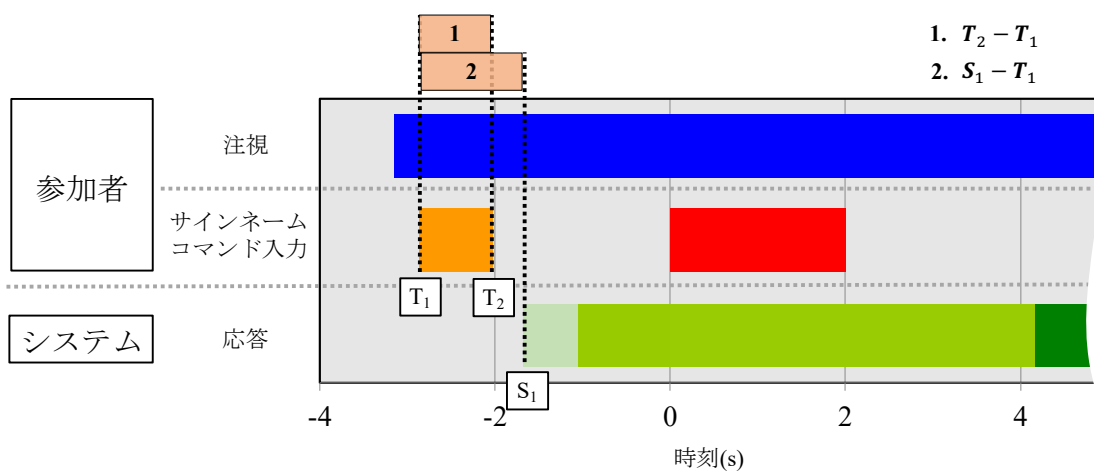


図 28 : I2×O3 条件の時のシステム挙動に関する解析時間

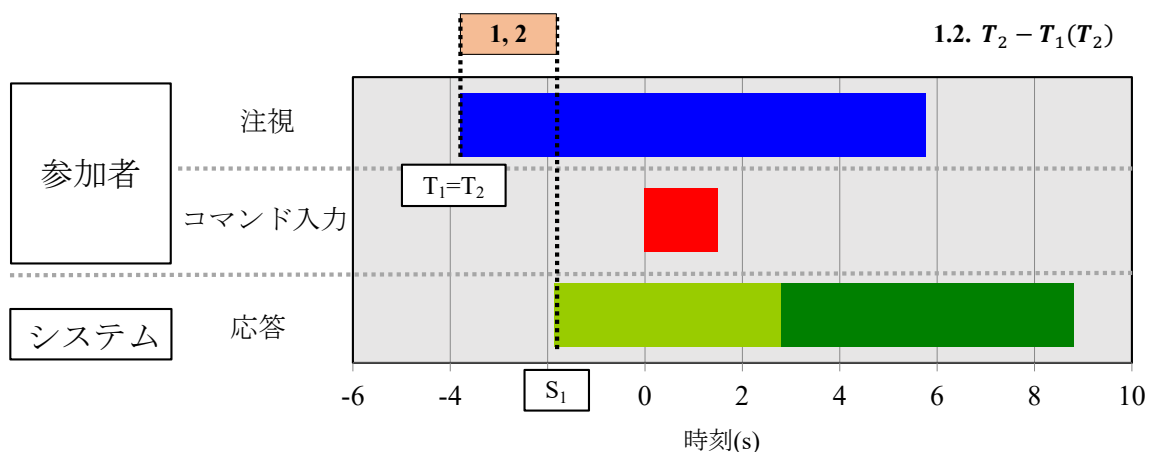


図 29 : I1×O1 条件の時のシステム挙動に関する解析時間

そして、RQ4 のフィードバックの提示に対する必要性について調査するために、ウェイクアップとコマンド入力を連続で行うより、間を空けてシステムからの反応を確認したいか、必要性（1.とてもそう思う、2.思う、3.やや思う、4.どちらでもない、5.あまり思わない、6.思わない、7.全くそう思わない）の7段階リッカートスケールによるアンケートを実施した。そして、RQ3 で検証したウェイクアップにおける注視の有効性を高められるか調査するために、I2（サインネーム）条件と I3（手振り）条件のウェイクアップ方法で行うときに、I1（注視）条件も合わせて行いたいかどうかをインタビューで尋ね、参加者の回答を3つのパターン（1.加えたい、2.限定的に加えたい、3.加えなくても良い）で分類した。

3.3 結果

3.3.1 フィードバックの有無が及ぼす影響の解析結果

フィードバック(O1-O4)それぞれの満足度の平均値および標準偏差を図 30 に示す。満足度平均値は、O1 条件で 1.90 (S.D.:1.16), O2 条件で 1.67 (S.D.:1.21), O3 条件で 1.67 (S.D.:1.55), O4 条件で -0.18 (S.D.:1.64)であった。4 水準間(O1-O4)で満足度に有意差は認められるか、Tukey 法の多重比較(群の数=4, $v=72$, $\alpha=0.05$ および $\alpha=0.01$)を行った。その結果、O4 (変化なし) 条件での満足度が他のフィードバック(O1-O3, 変化あり)条件より有意に低くなったことが認められた ($p < 0.01$)。

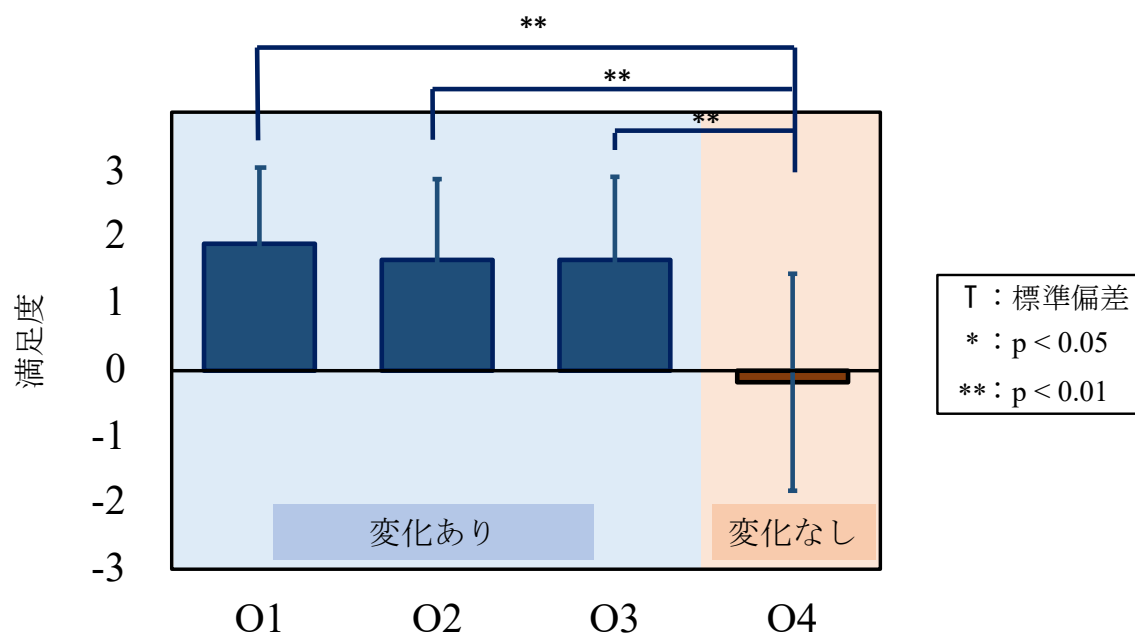


図 30 : フィードバックの個別満足度の平均値

フィードバック条件(O1-O4)それぞれのウェイクアップ動作終了からコマンド入力開始までの時間 ($T_3 - T_2$) の平均値および標準偏差を図 31 に示す。時間平均値は、O1 条件で 2.60 s (S.D.:1.06), O2 条件で 2.92 s (S.D.:1.31), O3 条件で 2.57 s (S.D.:1.17), O4 条件で 1.79 s (S.D.:1.14)であった。4 水準間(O1-O4)で時間に有意差は認められるか Tukey 法の多重比較(群の数=4, $v=72$, $\alpha=0.05$ および $\alpha=0.01$)を行った。その結果、O4 (変化なし) 条件の時間が他のフィードバック(O1-O3, 変化あり)条件の時間より有意に短くなったことが認められた ($p < 0.01$)。

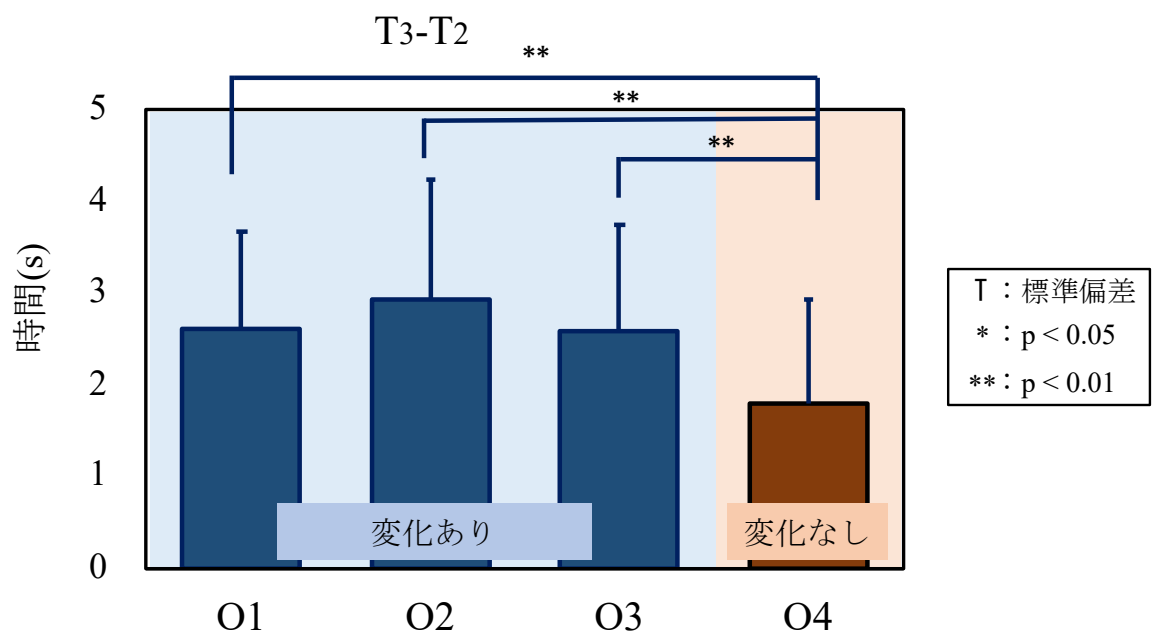


図 31 : フィードバックの個別ウェイクアップ終了からコマンド入力開始までの時間平均値

3.3.2 満足度の解析結果

ウェイクアップ方法とフィードバック方法の各条件を組み合わせた計 9 条件それぞれの満足度の平均値および標準偏差についてまとめたものを図 32 に示す。なお、図 30 よりフィードバックあり条件(O1-O3)に比べて有意にフィードバックなし条件(O4)の満足度が有意に低くなることを認められたため、より良いフィードバック方法を検討する際に比較条件から O4 (変化なし) 条件を省いた。

入出力方法の満足度平均値について、最も高かったのは、I3 (手振り) 条件と O1 (青いバー表示+明度低) 条件の組み合わせ方法の 2.13 (S.D. :1.08)であり、最も低かったのは、I1 (注視) 条件と O2 (O1 に手話表示追加) 条件の組み合わせ方法の 1.21 (S.D. :1.41)であった。

また、ウェイクアップ方法 (I1-I3, 3 水準) とフィードバック方法 (O1-O3, 3 水準) を要因として、2 要因分散分析 (繰り返しあり : 1 標本あたり 24 行数) を行った。その結果、ウェイクアップ方法の満足度について主効果が有意であった ($p < 0.05$)が、フィードバック方法の満足度について主効果は有意でなかった ($p > 0.05$)。また、ウェイクアップ方法とフィードバック方法それぞれの満足度の交互作用は有意でなかった ($p > 0.05$)。満足度の主効果が有意に認められたウェイクアップ方法について、満足度平均値は、I1 (注視) 条件で 1.49 (S.D. :1.32), I2 (サインネーム) 条件で 1.79 (S.D. :1.13), I3 (手振り) 条件で 1.96 (S.D. :1.13)であった。ウェイクアップ方法 3 水準間の満足度平均値の差を調べるために、Tukey 法の多重比較 (群の数=3, $v=72$, $\alpha=0.05$ および $\alpha=0.01$) を行った結果、I3 (手振り) 条件が I1 (注視) 条件より満足度が有意に高くなったことが認められた ($p < 0.05$)。

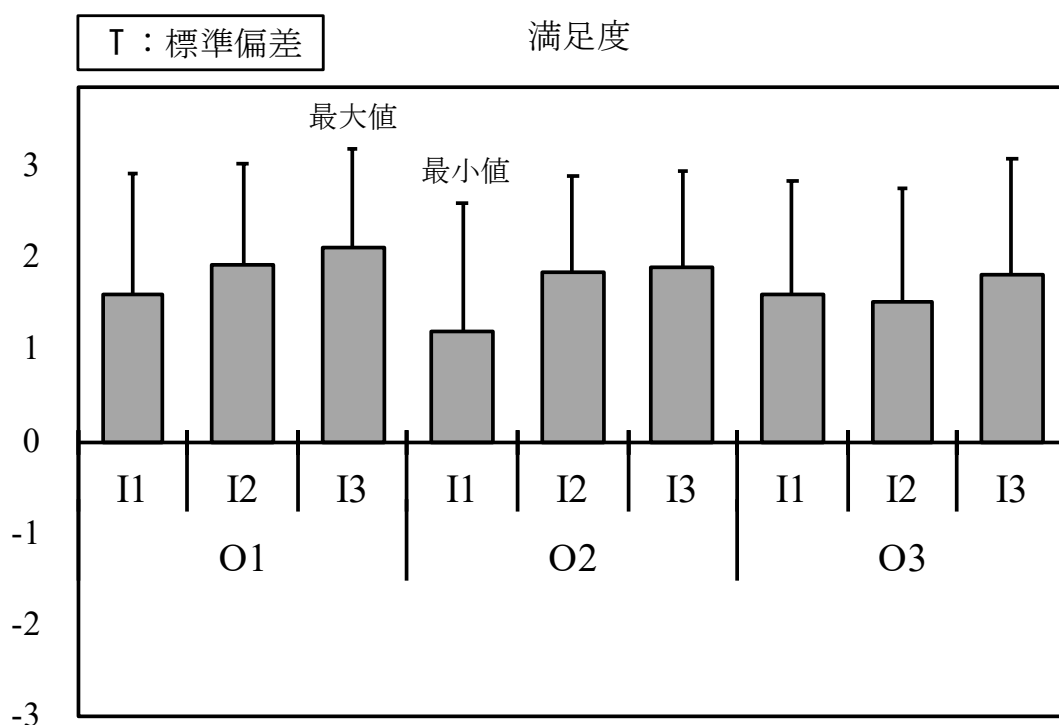


図 32 : 各 9 条件の満足度平均値

参加者属性による特性の違いが及ぼす影響の解析結果

参加者の各属性における満足度の平均値および標準偏差についてまとめたものを図 33 から図 36 に示す。属性について、図 33 を日常生活における補聴器/人工内耳の装用の有無、図 34 を手話レベル、図 35 を会話時における声の使用の有無、図 36 をアイディンティティとして示す。

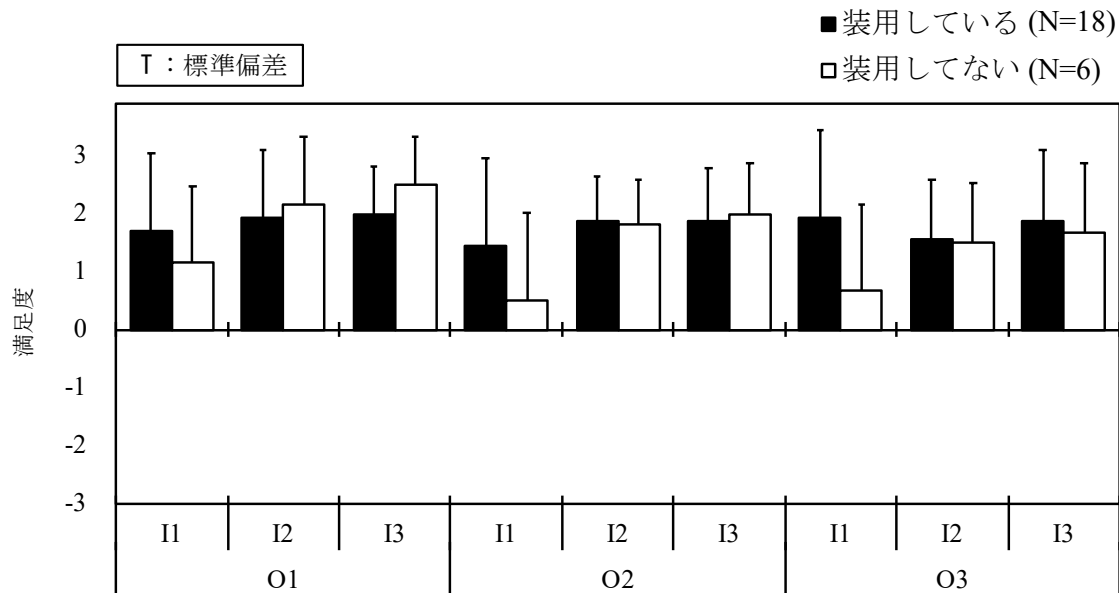


図 33 : 補聴器/人工内耳の装用に関する特性別の満足度の平均値

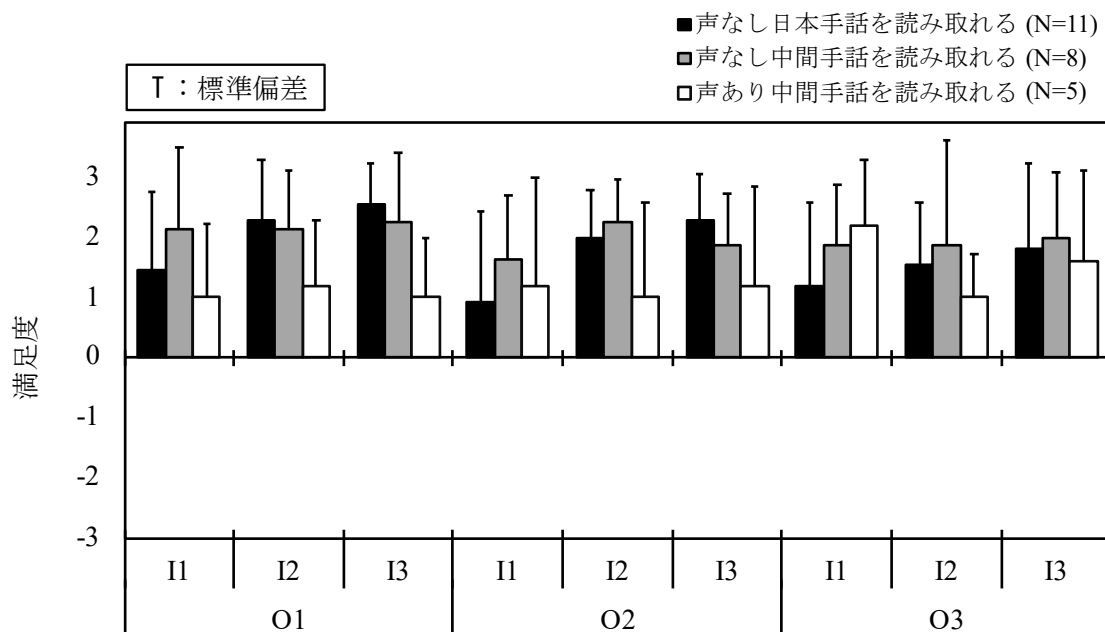


図 34 : 手話レベルに関する特性別の満足度の平均値

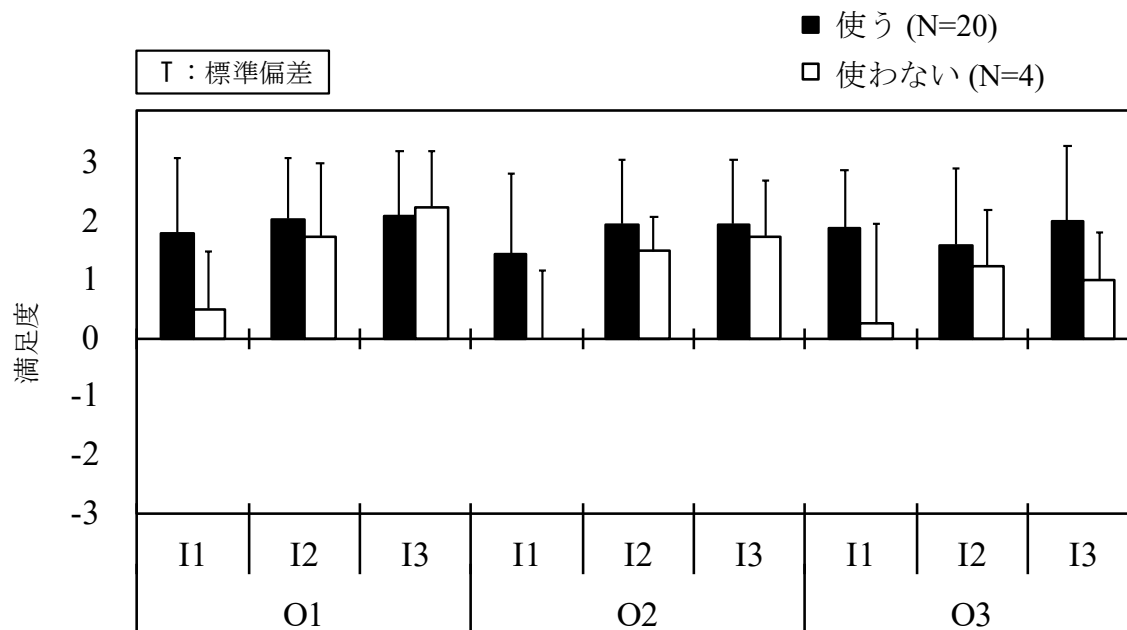


図 35 : 会話時の声使用に関する特性別の満足度の平均値

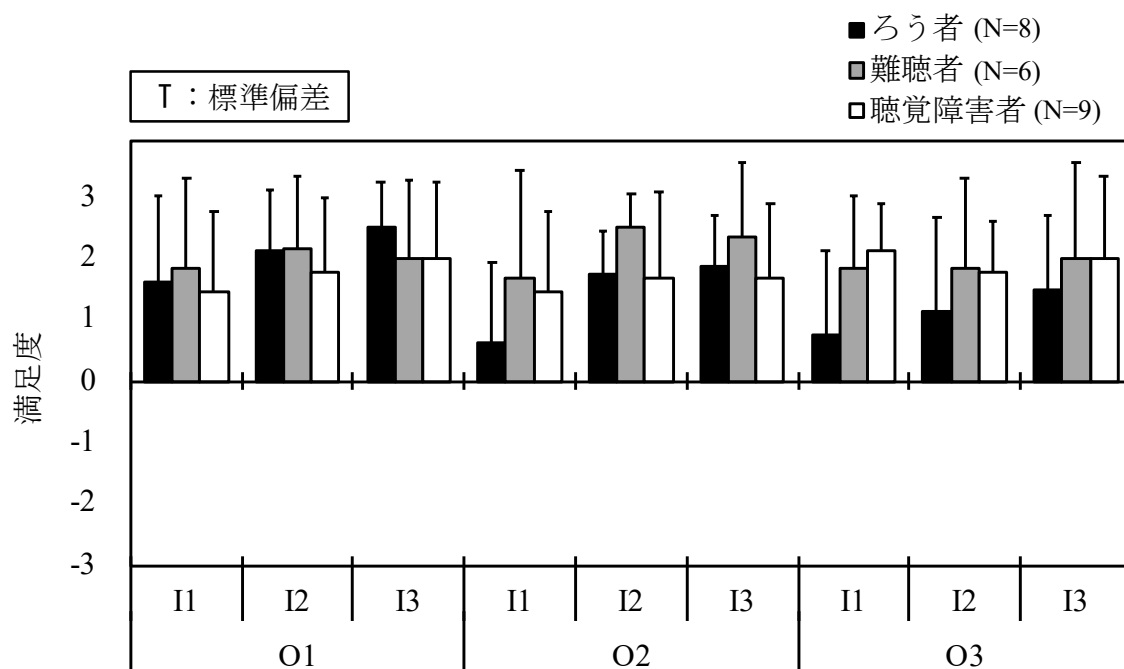


図 36 : アイデンティティに関する特性別の満足度の平均値

そして、特性ごとにウェイクアップ方法（I1-I3, 3 水準）とフィードバック方法（O1-O3, 3 水準）を独立変数とし、2 要因分散分析（繰り返しあり：1 標本あたり 24 行数）を行った。その結果、特性別にウェイクアップ方法とフィードバック方法、交互作用それぞれの有意確率についてまとめたものを表 18 に示す。

表 18：ウェイクアップとフィードバックの満足度を要因とした分散分析結果（特性別）

		有意確率 (p 値)		
参加者属性	特性	ウェイクアップ (I1-I3, 3 水準)	フィードバック (O1-O3, 3 水準)	交互作用
補聴器/人工内耳の 装用の有無	装用している (N=14)	0.63	0.81	0.67
	装用していない (N=10)	$4.45 \times 10^{-3}^{**}$	0.22	0.98
読み取り 手話レベル	声なし日本手話 (N=11)	$1.20 \times 10^{-3}^{**}$	0.12	0.82
	声なし中間手話 (N=8)	0.80	0.68	0.92
	声あり中間手話 (N=5)	0.71	0.50	0.83
会話時の 声使用の有無	使う (N=20)	0.38	0.63	0.61
	使わない (N=4)	$6.43 \times 10^{-3}^{**}$	0.33	0.87
アイディンティティ	ろう者 (N=8)	$1.82 \times 10^{-2}^{*}$	$1.82 \times 10^{-2}^{*}$	0.91
	難聴者 (N=6)	0.64	0.82	0.95
	聴覚障害者 (N=9)	0.53	0.79	0.90

*: $p < 0.05$, **: $p < 0.01$

最初に、日常生活における補聴器/人工内耳の装用の有無の属性について、2 要因それぞれに主効果が認められた特性は、「装用していない」と回答した参加者のウェイクアップ方法であった ($p < 0.01$)。また、ウェイクアップ方法とフィードバック方法それぞれの時間の交互作用は有意でなかった ($p > 0.05$)。「装用していない」と回答した参加者内のウェイクアップ方法について、満足度平均値は、I1（注視）条件で 0.78 (S.D.: 1.40), I2（サインネーム）条件で 1.83 (S.D.: 0.99), I3（手振り）条件で 2.06 (S.D.: 1.00)であった。ウェイクアップ方法 3 水準間の満足度平均値の差を調べるために、Tukey 法の多重比較（群の数=3, $v=18$, $\alpha=0.05$ および $\alpha=0.01$ ）を行った結果、I2（サインネーム）条件が I1（注視）条件より満足度が有意に高くなったことが認められ ($p < 0.05$)、また、I3（手振り）条件が I1（注視）条件より満足度が有意に高くなったことが認められた ($p < 0.01$)。

次に、手話レベルの属性について、2 要因それぞれに主効果が認められた特性は、「声なし日本手話を読み取れる」と回答した参加者のウェイクアップ方法であった ($p < 0.01$)。また、

ウェイクアップ方法とフィードバック方法それぞれの時間の交互作用は有意でなかった ($p > 0.05$). 「声なし日本手話を読み取れる」と回答した参加者内のウェイクアップ方法について、満足度平均値は、I1 (注視) 条件で 1.18 (S.D.: 1.38), I2 (サインネーム) 条件で 1.94 (S.D.: 0.97), I3 (手振り) 条件で 2.21 (S.D.: 1.02)であった. ウェイクアップ方法 3 水準間の満足度平均値の差を調べるために、Tukey 法の多重比較 (群の数=3, $v=33$, $\alpha=0.05$ および $\alpha=0.01$) を行った結果、I2 (サインネーム) 条件が I1 (注視) 条件より満足度が有意に高くなったことが認められ ($p < 0.05$), また、I3 (手振り) 条件が I1 (注視) 条件より満足度が有意に高くなったことが認められた ($p < 0.01$).

そして、会話時における声の使用の有無の属性について、2 要因それぞれに主効果が認められた特性は、「使わない」と回答した参加者のウェイクアップ方法であった ($p < 0.01$). また、ウェイクアップ方法とフィードバック方法それぞれの時間の交互作用は有意でなかった ($p > 0.05$). 「使わない」と回答した参加者内のウェイクアップ方法について、満足度平均値は、I1 (注視) 条件で 0.25 (S.D.: 1.22), I2 (サインネーム) 条件で 1.50 (S.D.: 0.90), I3 (手振り) 条件で 1.67 (S.D.: 0.98)であった. ウェイクアップ方法 3 水準間の満足度平均値の差を調べるために、Tukey 法の多重比較 (群の数=3, $v=12$, $\alpha=0.05$ および $\alpha=0.01$) を行った結果、I2 (サインネーム) 条件と I3 (手振り) 条件が I1 (注視) 条件より満足度が有意に高くなったことが認められた ($p < 0.05$).

最後に、アイデンティティの属性について、2 要因それぞれに主効果が認められた特性は、「ろう者」と回答した参加者のウェイクアップ方法とフィードバック方法であった ($p < 0.05$). また、ウェイクアップ方法とフィードバック方法それぞれの時間の交互作用は有意でなかった ($p > 0.05$). 「ろう者」と回答した参加者内のウェイクアップ方法について、満足度平均値は、I1 (注視) 条件で 1.00 (S.D.: 1.38), I2 (サインネーム) 条件で 1.67 (S.D.: 1.17), I3 (手振り) 条件で 1.96 (S.D.: 1.00)であった. ウェイクアップ方法 3 水準間の満足度平均値の差を調べるために、Tukey 法の多重比較 (群の数=3, $v=24$, $\alpha=0.05$ および $\alpha=0.01$) を行った. その結果、I3 (手振り) 条件が I1 (注視) 条件より満足度が有意に高くなったことが認められた ($p < 0.05$). また、フィードバック方法について、満足度平均値は、O1 (青いバー表示+明度低) 条件で 2.08 (S.D.: 1.10), O2 (O1 に手話表示追加) 条件で 1.42 (S.D.: 1.10), O3 (O1 に首振り運動追加) 条件で 1.13 (S.D.: 1.36)であった. フィードバック方法 3 水準間の満足度平均値の差を調べるために、Tukey 法の多重比較 (群の数=3, $v=24$, $\alpha=0.05$ および $\alpha=0.01$) を行った. その結果、O1 (青いバー表示+明度低) 条件が O3 (O1 に首振り運動追加) 条件より満足度が有意に高くなったことが認められた ($p < 0.05$).

以上の結果から、ウェイクアップ方法について手振り条件の満足度が有意に高いことが認められた参加者の特性、その備考についてまとめたものを表 19 に示す.

表 19：手振り条件の満足度が他の条件に比べて有意に高くなった参加者特性

No.	手振り条件の満足度が他の条件に比べて有意に高くなった特性	備考
1	日常生活で補聴器/人工内耳を装用しない	聴覚情報に頼らない
2	日本手話を読み取れる	手話レベルが高い
3	コミュニケーションで声を使わない	音声情報を活用しない
4	アイディンティティはろう者である	ろう学校育ち，デフファミリー出身， ろうコミュニティに属しているなど

3.3.3 順位付け解析結果

全参加者の順位結果

各ウェイクアップ条件(I1-I3)に対して、参加者(N=24)が行った順位付けの結果を図 37 に示す。I1（注視）条件を 1 位に回答した参加者は 5 名、2 位に回答した参加者は 3 名、3 位に回答した参加者は 16 名であった。I2（サインネーム）条件を 1 位に回答した参加者は 6 名、2 位に回答した参加者は 14 名、3 位に回答した参加者は 4 名であった。I3（手振り）条件を 1 位に回答した参加者は 13 名、2 位に回答した参加者は 7 名、3 位に回答した参加者は 4 名であった。また、平均順位について、I1（注視）条件で 2.46 位 (S.D.: 0.82), I2（サインネーム）条件で 1.92 位 (S.D.: 0.64), I3（手振り）条件で 1.63 位 (S.D.: 0.75)であった。ウェイクアップ方法の違いによって平均順位は変化するか、フリードマン検定を用いて調査した結果、有意に変化することが認められた ($p < 0.05$)。

I1（注視）条件について、1 位（最上位）に回答した参加者からは、主に「手を使わないといった、楽な呼びかけ動作だった」、「スマートフォンの顔認証を用いるのと同じ感覚で使用できた。」、「手を使わない、楽な動作だったから。しかし、意図なしでちらりとシステムの方を見ただけで認識される懸念があった。なので、注視に 2 秒間ルールとか設けてシステムに反応してもらうようにすればもっといいかも。」のコメントがあった。一方、3 位（最下位）に回答した参加者からは、主に「呼ぶ動作において、見るだけでは違和感があった。」、「手が離せないときには便利だと思うが、その後手話でコマンドを出すなら最初から手を使った呼びかけをする方がいいと思った。注視だけでは認識してもらえないか不安だったため、手を使う方が確実に呼び出せると思った。」、「見るだけだと、いつ反応してもらえないかわからない。」のコメントがあった。

I2（サインネーム）条件について、1 位（最上位）に回答した参加者からは、主に「システム特有の呼び方のため、システムからもサインネーム動作が認識されやすいと思ったから。」、「システム特有の呼び方であり周りに誤解されることもないので、この呼び方にしっくりきた。」、「名前を使った呼び方なので、システムに対して親近感が沸いた。」のコメントがあった。一方、3 位（最下位）に回答した参加者からは、主に「日常生活で人を呼ぶときに、名前で呼ぶことはほぼないため、違和感あった。」、「手話に慣れてないのでサインネームで呼ぶことに慣れてない。」、「注視や手振りと比べると、サインネームは認知しづらい呼び動作であった。」のコメントがあった。

I3（手振り）条件について、1 位（最上位）に回答した参加者からは、主に「すぐ動作に移せるから。」、「日常で会話する時に呼びかけるのと同じ動作だから。」、「相手問わず、フレンドリーな呼び方ができる。初めてでも呼べる。」のコメントがあった。一方、3 位（最下位）に回答した参加者からは、主に「サインネームと比べたら呼ぶ対象がはっきりしないため。」、「システムに手振りで呼びかけると、周りにいる友人に誤解される可能性があったため。」のコメントがあった。

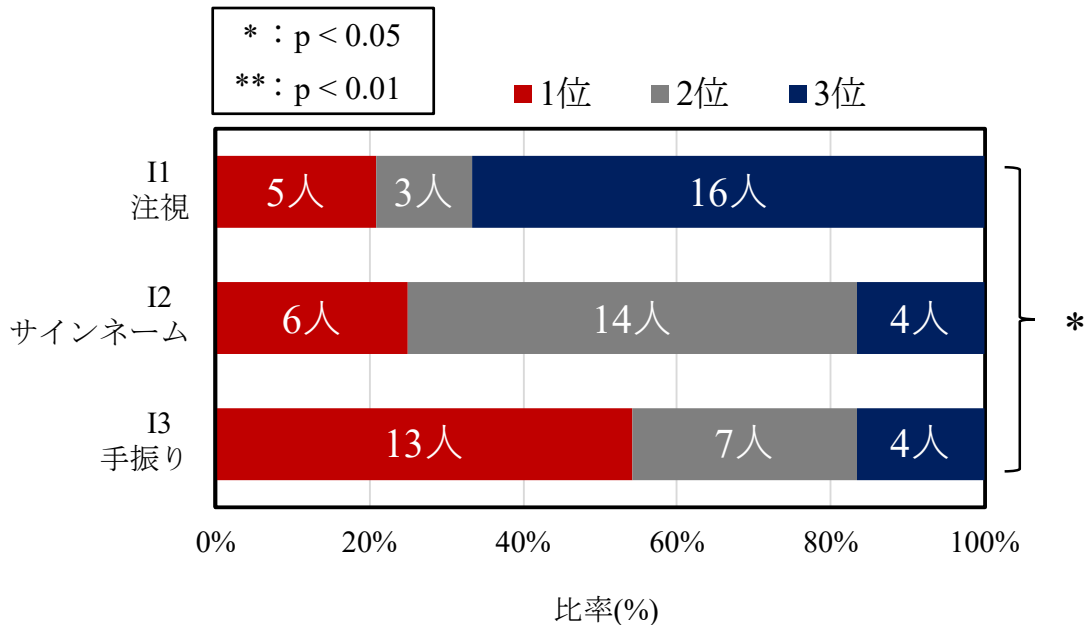


図 37 : ウェイクアップ方法の順位結果

各フィードバック条件(O1-O4)に対して、参加者(N=24)が行った順位付けの結果を図 38 に示す。O1（青いバー表示+明度低）条件を 1 位に回答した参加者は 14 名、2 位に回答した参加者は 3 名、3 位に回答した参加者は 7 名で、4 位に回答した参加者はいなかった。O2（O1 に手話表示追加）条件を 1 位に回答した参加者は 4 名、2 位に回答した参加者は 13 名、3 位に回答した参加者は 6 名、4 位に回答した参加者は 1 名であった。O3（O1 に首振り運動追加）条件を 1 位に回答した参加者は 5 名、2 位に回答した参加者は 8 名、3 位に回答した参加者は 10 名、4 位に回答した参加者は 1 名であった。O4（変化なし）条件を 1 位に回答した参加者は 1 名、2 位に回答した参加者は 1 名、3 位に回答した参加者は 22 名で、4 位に回答した参加者はいなかった。また、平均順位について、O1（青いバー表示+明度低）条件で 1.71 位 (S.D.: 0.91), O2（O1 に手話表示追加）条件で 2.17 位 (S.D.: 0.76), O3（O1 に首振り運動追加）条件で 2.29 位 (S.D.: 0.86), O4（変化なし）条件で 3.83 位 (S.D.: 0.64)であった。フィードバックの違いによって平均順位は変化するか、フリードマン検定を用いて調査した結果、有意に変化することが認められた ($p < 0.01$)。

O1（青いバー表示+明度低）条件について、1 位（最上位）に回答した参加者からは、主に「私の呼びかけに対してシステムが反応してくれたことが目で見てすぐ分かったから。」、「スマートフォンの Siri と同じような反応のため馴染みやすかった。」、「他の反応は過剰に感じ、青いバーの表示だけで十分だと思った。」のコメントがあった。また、「手話表示や首振り運動と比べるとシンプルすぎてわかりにくかった。」と評価が低いコメントがあった。

O2（O1 に手話表示追加）条件について、1 位（最上位）に回答した参加者からは、主に「手話で話しかけるので、相手も手話で反応して欲しいと思ったから。」、「手話で反応してくれたことに安心感があった。」、「手話で反応してくれることで表情などが読み取れ、用件を聞かれていることもわかりやすかった。」のコメントがあった。一方、「システムからは人間で

はなくアバターで表示してくれた方がいい。」，「手話でなく，文字情報の表示とかで十分と感じた。」，「呼びかけるときにだけ，登場してくるのでなく，ホーム画面などデフォルト状態のときから手話者が表示されていたら，評価はもっと上がった。」など，手話表示方法の改善に関するコメントがあった。

O3（O1に首振り運動追加）条件について，1位（最上位）に回答した参加者からは，主に「はっきりとシステムが動作し反応してくれたことが見て分かり，可愛い印象も受けた。」，「首振り反応してくれるのを見ていて，とても分かりやすかった。」のコメントがあった。他の参加者からは，主に「外方を向かれたことに違和感あった。もし，注視しない，サインネームと手振りでウェイクアップするなら首振りは横じゃなく縦振り運動がよかった。」，「見やすさの面では良かったが，これにさらに多くの文字情報が付与されると見にくくなるのではと懸念があった。」，「首振り運動には時間が少しかかるので評価は下がる。」のコメントがあった。

O4（変化なし）条件について，1位（最上位）に回答した参加者からは，「変化なしの方が，呼びかけ動作から間要らず用件言えるため。」のコメントがあった。また，「練習の時では，分かりにくい反応と思ったが，次第に使っていくことで，変化がないということは呼びかけからすぐに用件を言えるので案外良いかも？と感じた。」のコメントもあった。他の参加者からは，主に「いつコマンドを出すべきか，タイミングがわからなかったため不安に感じた」，「いつコマンドを出すべきかタイミングがわからなかったため不安に感じた」，「目に見える変化が欲しい。」と評価が低いコメントが多々あった。

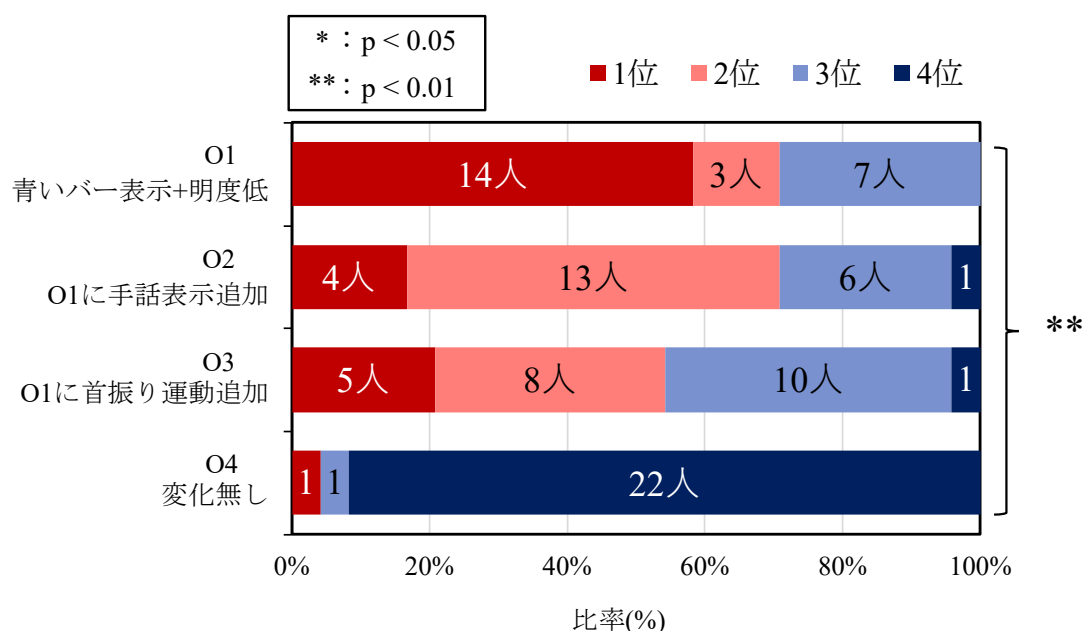


図 38：フィードバック方法の順位結果

手話レベルの特性の違いが順位に及ぼす影響の解析結果

参加者の手話レベル属性におけるウェイクアップ方法の順位データについてまとめたものを図 39 に示す。そして、参加者特性ごとにウェイクアップ条件 (I1-I3) の違いによって平均順位は変化するか、フリードマン検定を用いて調査した。その結果、有意に変化することが認められた参加者特性は、「声なし日本手話が読み取れる」と回答した参加者 ($p < 0.01$)のみであった。

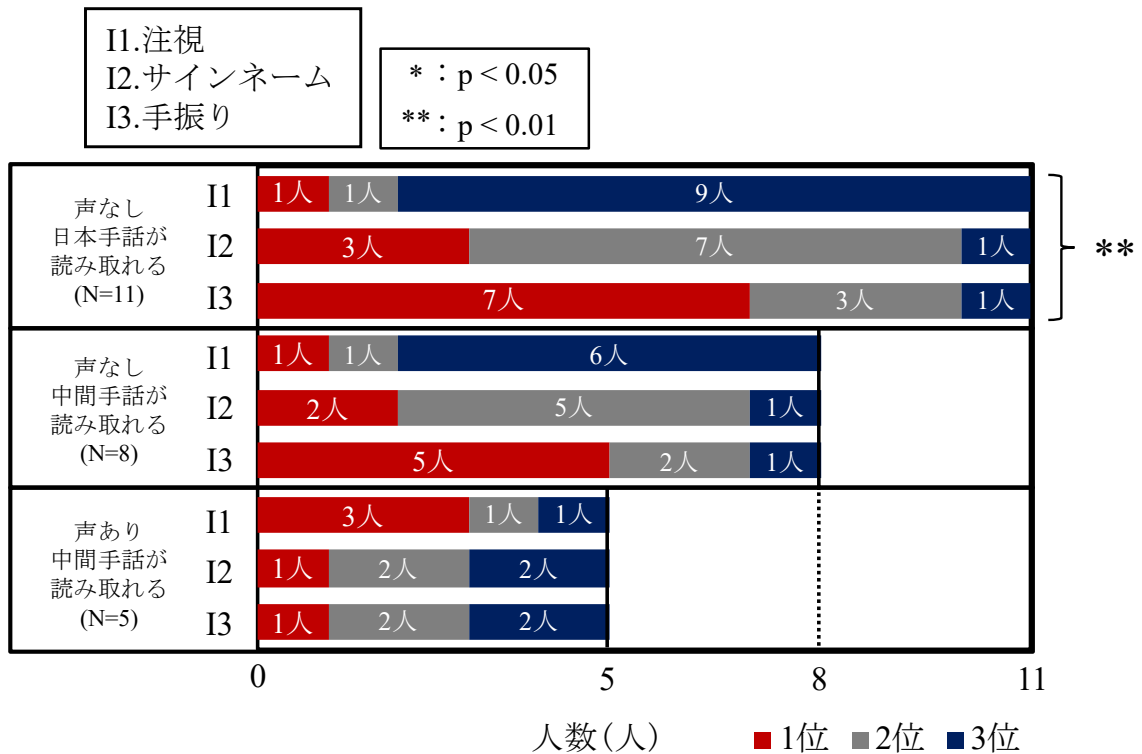


図 39 : 手話レベル特性別のウェイクアップ条件の満足度順位結果

参加者の手話レベル属性におけるフィードバック方法の順位データについてまとめたものを図 40 に示す。そして、参加者特性ごとにウェイクアップ条件 (I1-I3) の違いによって平均順位は変化するか、フリードマン検定を用いて調査した。その結果、有意に変化することが認められた参加者特性は、「声なし日本手話が読み取れる」と回答した参加者 ($p < 0.01$)と、「声なし中間手話が読み取れる」と回答した参加者 ($p < 0.01$)であった。

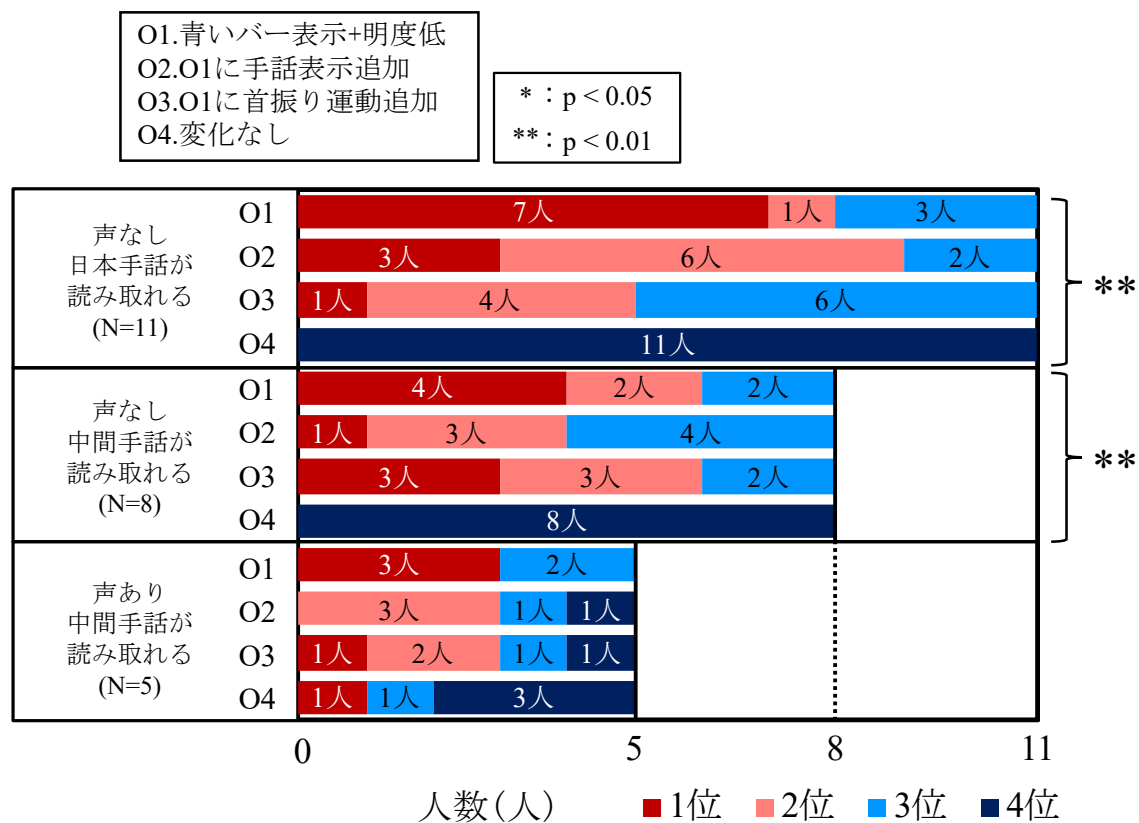


図 40：手話レベル特性別のフィードバック条件の満足度順位結果

3.3.4 参加者の挙動の解析結果

1. 手振り動作時間 ($T_2 - T_1$)

フィードバック条件の違いが手振り動作時間に及ぼす影響の解析結果

フィードバック条件(O1-O3)ごとの手振り動作時間の平均値および標準偏差についてまとめたものを図 41 に示す。手振り動作時間平均値は、O1（青いバー表示+明度低）条件で 1.04 s (S.D.:0.50), O2（O1 に手話表示追加）条件で 1.21 s (S.D.:0.60), O3（O1 に首振り運動追加）条件で 1.11 s (S.D.:0.42)であった。3 水準間(O1-O3)で時間に有意差は認められるか Tukey 法の多重比較（群の数=3, $v=72$, $\alpha=0.05$ および $\alpha=0.01$ ）を行った。その結果、有意差が認められるものはなかった ($p > 0.05$)。

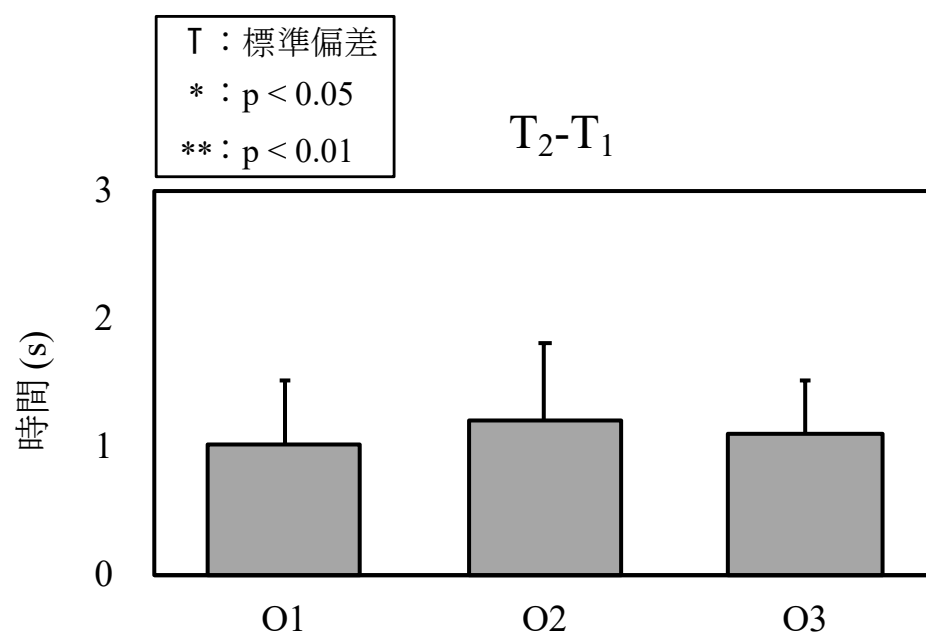


図 41：フィードバックそれぞれの手振り動作時間の平均値

参加者特性の違いが手振り動作時間に及ぼす影響の解析結果

参加者の各属性における手振り動作時間の平均値および標準偏差についてまとめたものを図 42 に示す。属性について、手話レベル、会話時における声の使用の有無、アイディンティティの 3 つを示す。そして、水準間に有意な差があるか比較するために、手話レベル属性（3 水準）とアイディンティティ属性（3 水準）では、Tukey 法の多重比較を行い、会話時における声の使用の有無の属性（2 水準）では、Welch の t 検定を行った。

手話レベルの属性について、手振り動作時間は「声なしで日本手話が読み取れる」と回答した参加者で 0.99 s (S.D.:0.12), 「声ありで中間手話が読み取れる」と回答した参加者で 1.43 s (S.D.:0.37)であり、手話レベルが高い方が手振り動作にかかる時間は有意に短くなったことが認められた ($p < 0.01$)。

会話時における声の使用の有無の属性について、手振り動作時間は「使わない」と回答した参加者で 1.17 s (S.D. :0.52), 「使う」と回答した参加者で 0.87 s (S.D. :0.41)であり、声を使わない方が手振り動作にかかる時間は有意に短くなったことが認められた ($p < 0.05$).

アイディンティティの属性について、手振り動作時間は「ろう者」と回答した参加者で 0.93 s (S.D. :0.35), 「聴覚障害者」と回答した参加者で 1.26 s (S.D. :0.38)であり, 「ろう者」のアイディンティティを持つ参加者が「聴覚障害者」のアイディンティティを持つ参加者より時間は有意に短くなったことが認められた ($p < 0.01$).

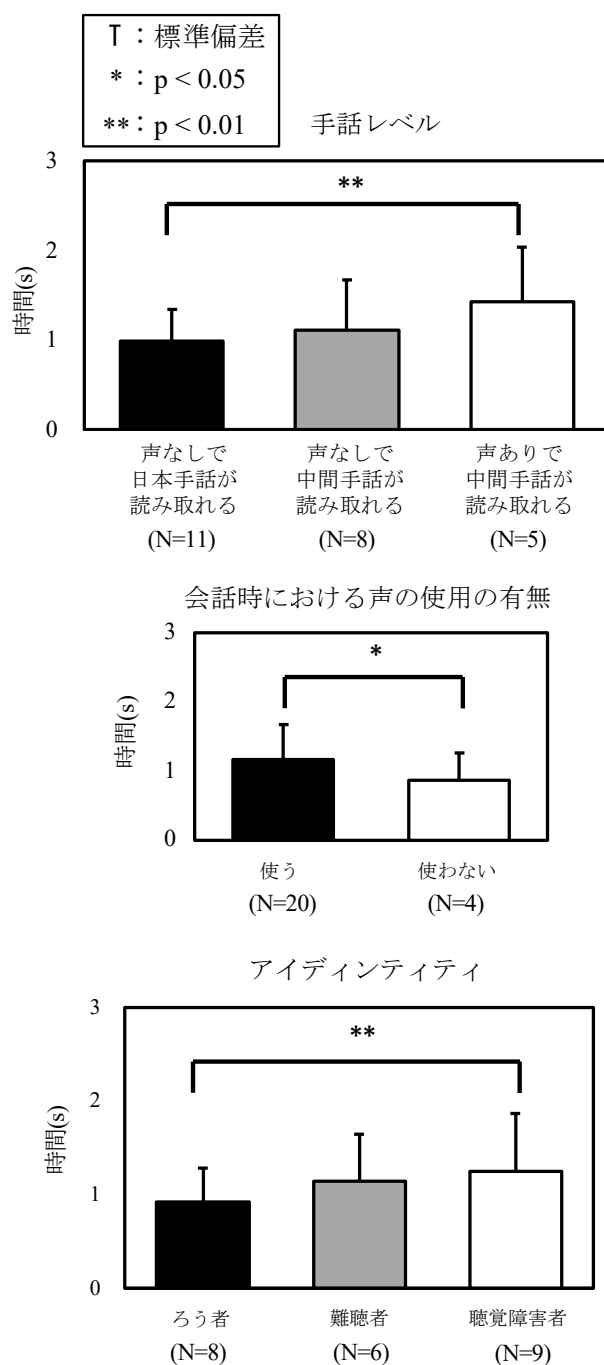


図 42 : 参加者の各属性における手振り動作時間の平均値

手振り条件における注視動作

注視と手振り動作の開始時刻について早かった方のデータ数の構成比と挙動開始時刻の差についてまとめたものを表 20 に示す．なお，注視と手振りそれぞれの開始時刻の差が 0.02 s 以外であったデータ数は全体 72 個のうち，2 個あり同時に開始したと見なした．また，注視と手振りの両方行った時間を表 21 に示す．注視と手振りの両方行った時間がフィードバック方法 (O1-O3) の違いによる影響を受けるか Tukey 法の多重比較 (群の数=3, $v=24$, $\alpha=0.05$ および $\alpha=0.01$) を行った．その結果，有意差が認められたものはなかった．

表 20：挙動順番の構成比と挙動開始時刻の差

条件	挙動順番	構成比(%)	挙動開始時刻の差 (s)
			平均値 \pm 標準偏差 (最小値, 最大値)
I3 手振り	注視→手振り	81.9	0.37 ± 0.35 (0.04, 1.60)
	手振り→注視	15.3	0.30 ± 0.32 (0.04, 1.12)
	同時	2.8	

表 21：注視と手振り動作の両方を行った時間

挙動	平均値 \pm 標準偏差(s) (最小値, 最大値)
注視 \cap 手振り	1.07 ± 0.52 (0.20, 2.46)

2. ウェイクアップ動作終了からコマンド入力開始までの時間($T_3 - T_2$)

ウェイクアップ方法とフィードバック方法の各条件を組み合わせた計 9 条件それぞれのウェイクアップ動作終了からコマンド入力開始までの時間の平均値および標準偏差を図 43 に示す。ウェイクアップ方法 (I1-I3, 3 水準) とフィードバック方法 (O1-O3, 3 水準) を要因として、2 要因分散分析 (繰返しあり: 1 標本あたり 24 行数) を行った。その結果、ウェイクアップ方法について主効果は有意であった ($p < 0.01$) が、フィードバック方法について主効果は有意でなかった ($p > 0.05$)。また、ウェイクアップ方法とフィードバック方法の交互作用は有意でなかった ($p > 0.05$)。

時間の主効果が有意に認められたウェイクアップ方法について、時間平均値は、I1 (注視) 条件で 3.45 s (S.D.: 1.28), I2 (サインネーム) 条件で 2.52 s (S.D.: 1.20), I3 (手振り) 条件で 2.12 s (S.D.: 0.87) であった。ウェイクアップ方法 3 水準間の満足度平均値の差を調べるために、Tukey 法の多重比較 (群の数=3, $v=72$, $\alpha=0.05$ および $\alpha=0.01$) を行った結果、I2 (サインネーム) 条件は I1 (注視) 条件より有意に短くなったことが認められ ($p < 0.01$)、I3 (手振り) 条件は I1 (注視) 条件より有意に短くなったことが認められ ($p < 0.01$)、そして I3 (手振り) 条件は I2 (サインネーム) 条件より有意に短くなった ($p < 0.05$) が認められた。

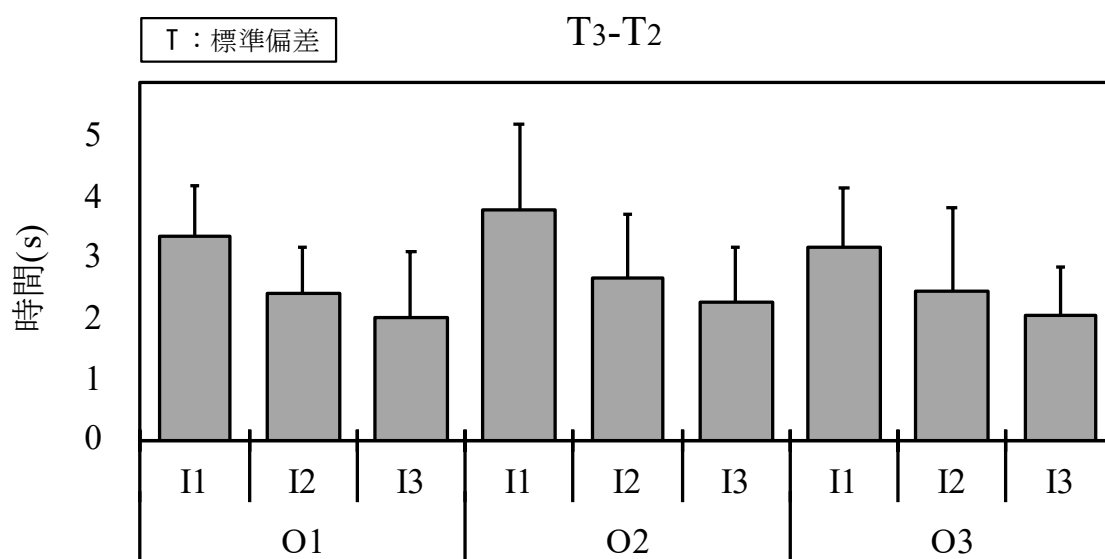


図 43 : 9 条件毎のウェイクアップ終了からコマンド入力開始までの時間平均値

システムの反応開始（変化終了）からコマンド入力開始までの時間 ($T_3 - S_2(S_1)$)

ウェイクアップ方法とフィードバック方法の各条件を組み合わせた計 9 条件それぞれのシステムの反応開始（変化終了）からコマンド入力開始までの時間の平均値および標準偏差を図 44 に示す。ウェイクアップ方法（I1-I3, 3 水準）とフィードバック方法（O1-O3, 3 水準）を要因として、2 要因分散分析（繰り返しあり：1 標本あたり 24 行数）を行った。その結果、ウェイクアップ方法について主効果が有意でなかった ($p > 0.05$) が、フィードバック方法について主効果は有意であった ($p < 0.01$)。また、ウェイクアップ方法とフィードバック方法それぞれの時間の交互作用は有意でなかった ($p > 0.05$)。

時間の主効果が有意に認められたフィードバック方法について、時間平均値は、O1（青いバー表示+明度低）条件で 1.51 s (S.D.: 0.71), O2 (O1 に手話表示追加) 条件で 0.27 s (S.D.: 1.07), O3 (O1 に首振り運動追加) 条件で 1.83 s (S.D.: 1.11) であった。フィードバック方法 3 水準間の時間平均値の差を調べるために、Tukey 法の多重比較 (群の数=3, $v=72$, $\alpha=0.05$ および $\alpha=0.01$) を行った結果、O2 (O1 に手話表示追加) 条件が、O1（青いバー表示+明度低）条件と O3 (O1 に首振り運動追加) 条件より時間が有意に短くなったことが認められた ($p < 0.01$)。

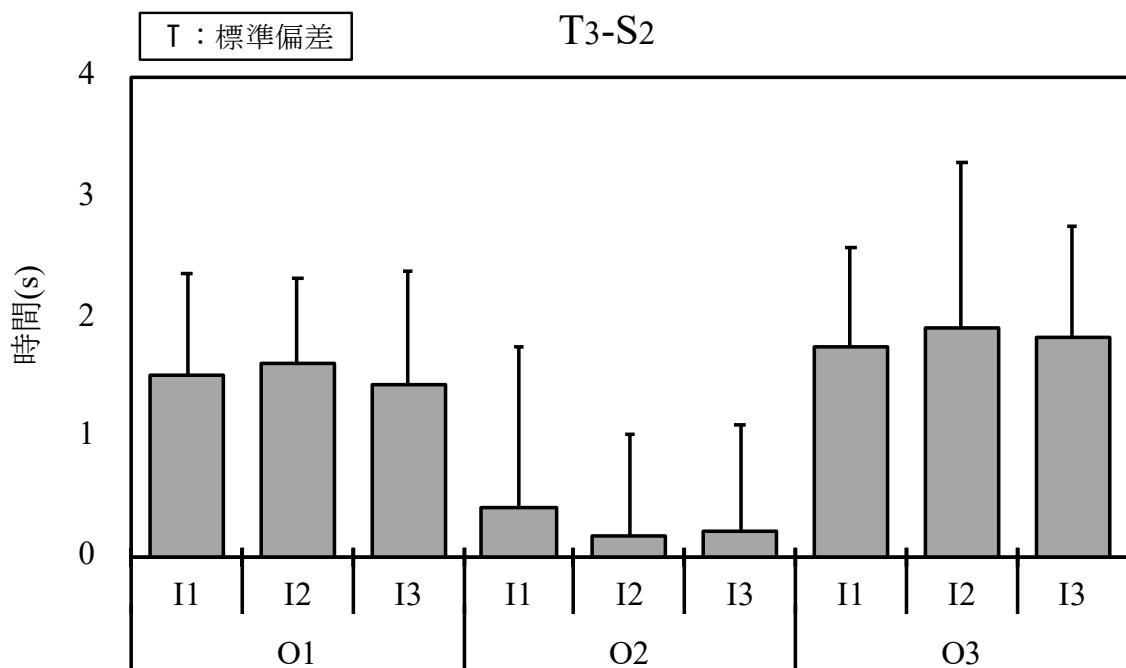


図 44 : 9 条件毎のシステムの反応開始（変化終了）からコマンド入力開始までの時間平均値

ここで、システムの反応開始のタイミングについて、有意に短くなった O2 (O1 に手話表示追加) 条件はシステム反応開始(変化開始, S_1)とし、システムの反応開始からコマンド入力開始までの時間の平均値および標準偏差を図 45 に示す。ウェイクアップ方法 (I1-I3, 3 水準) とフィードバック方法 (O1-O3, 3 水準) を要因として、2 要因分散分析(繰り返しあり: 1 標本あたり 24 行数)を行った。その結果、ウェイクアップ方法とフィードバック方法それぞれについて主効果が有意でなかった ($p > 0.05$)。また、ウェイクアップ方法とフィードバック方法それぞれの時間の交互作用は有意でなかった ($p > 0.05$)。

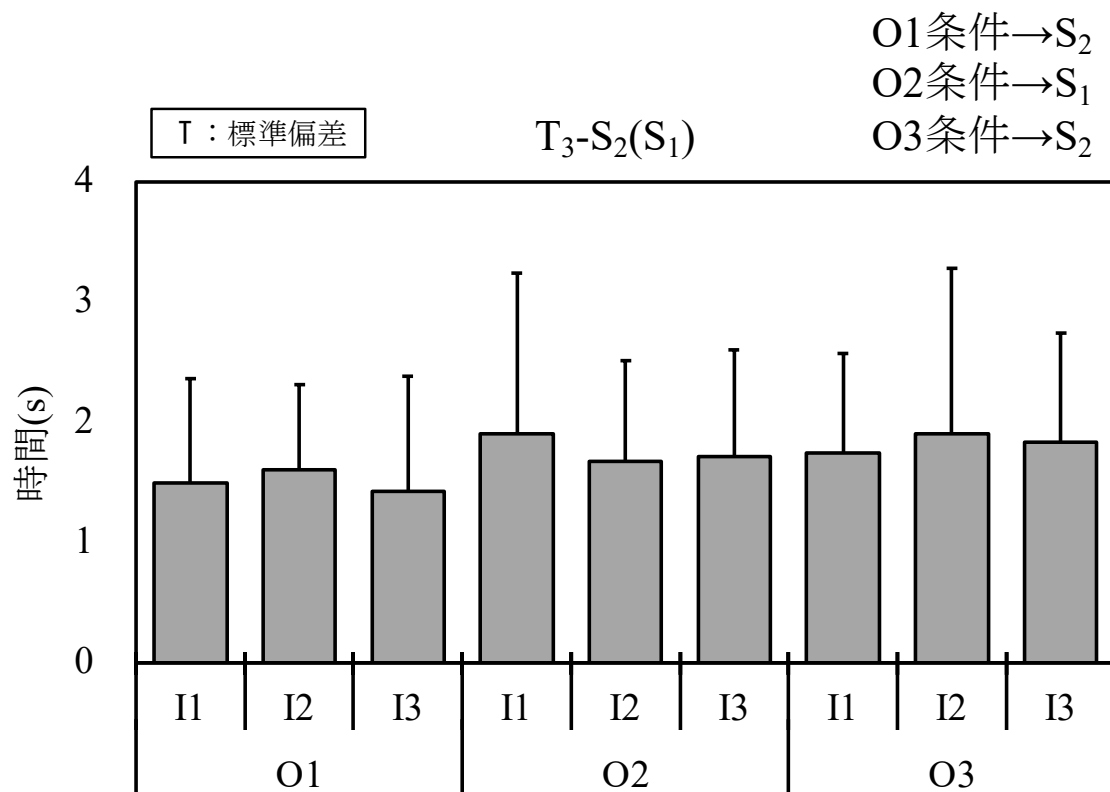


図 45 : 9 条件毎のシステムの反応開始からコマンド入力開始までの時間平均値

3.3.5 システムの挙動の解析結果

1. ウェイクアップの動作開始からシステムの反応開始（変化開始）までの時間（ $S_1 - T_1$ ）

ウェイクアップ方法とフィードバック方法の各条件を組み合わせた計 9 条件それぞれのウェイクアップ動作開始からシステムの反応開始（変化開始）までの時間の平均値および標準偏差を図 46 に示す。ウェイクアップ方法（I1-I3, 3 水準）とフィードバック方法（O1-O3, 3 水準）を要因として、2 要因分散分析（繰り返しあり：1 標本あたり 24 行数）を行った。その結果、ウェイクアップ方法について主効果は有意であり（ $p < 0.05$ ）、フィードバック方法についても主効果は有意であった（ $p < 0.01$ ）。また、ウェイクアップ方法とフィードバック方法の交互作用は有意でなかった（ $p > 0.05$ ）。

時間の主効果が有意に認められたウェイクアップ方法について、時間平均値は、I1（注視）条件で 1.73 s (S.D.: 0.26), I2（サインネーム）条件で 1.78 s (S.D.: 0.15), I3（手振り）条件で 1.58 s (S.D.: 0.21)であった。ウェイクアップ方法 3 水準間の時間平均値の差を調べるために、Tukey 法の多重比較（群の数=3, $v=72$, $\alpha=0.05$ および $\alpha=0.01$ ）を行った結果、I3（手振り）条件は I2（サインネーム）条件より有意に短くなったことが認められた（ $p < 0.05$ ）。また、時間の主効果が有意に認められたフィードバック方法について、時間平均値は O1（青いバー表示+明度低）条件で 1.78 s (S.D.: 0.18), O2（O1 に手話表示追加）条件で 1.87 s (S.D.: 0.24), O3（O1 に首振り運動追加）条件で 1.44 s (S.D.: 0.11)であった。フィードバック方法 3 水準間の時間平均値の差を調べるために、Tukey 法の多重比較（群の数=3, $v=72$, $\alpha=0.05$ および $\alpha=0.01$ ）を行った結果、O3（O1 に首振り運動追加）条件は O1（青いバー表示+明度低）条件と O2（O1 に手話表示追加）条件より有意に短くなったことが認められた（ $p < 0.01$ ）。

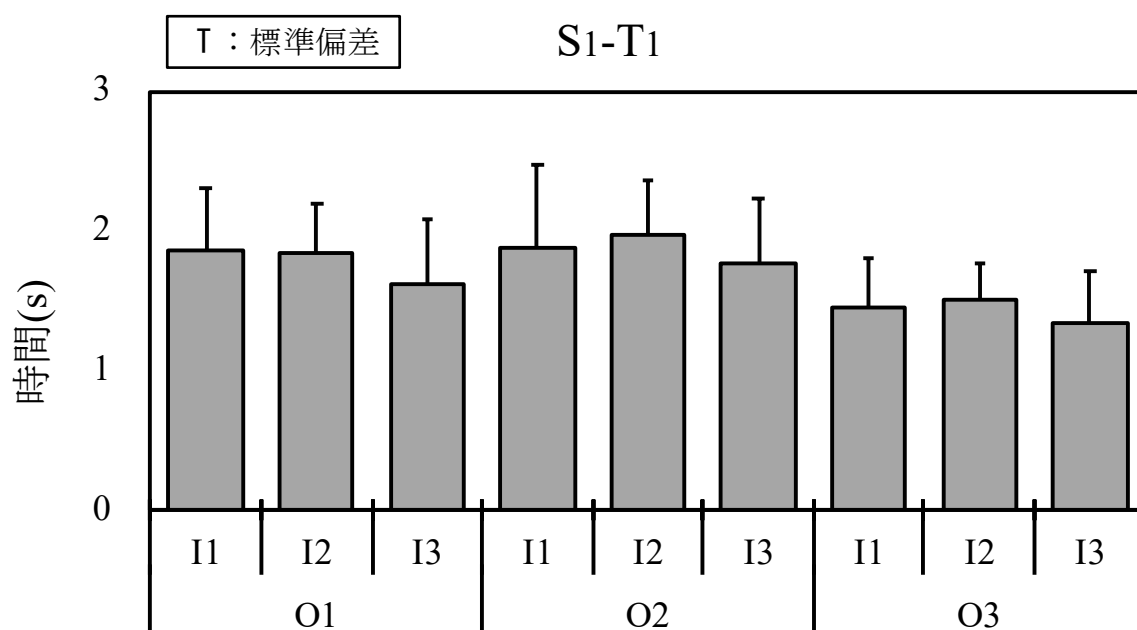


図 46：9 条件毎のウェイクアップ開始からシステムの反応開始(変化開始)までの時間平均値

2. ウェイクアップの動作終了からシステムの反応開始（変化開始）までの時間（ $S_1 - T_2$ ）

ウェイクアップ方法とフィードバック方法の各条件を組み合わせた計 9 条件それぞれのウェイクアップ動作終了からシステムの反応開始（変化開始）までの時間の平均値および標準偏差を図 47 に示す。ウェイクアップ方法（I1-I3, 3 水準）とフィードバック方法（O1-O3, 3 水準）を要因として、2 要因分散分析（繰り返しあり：1 標本あたり 24 行数）を行った。その結果、ウェイクアップ方法について主効果は有意であり（ $p < 0.01$ ）、フィードバック方法についても主効果は有意であった（ $p < 0.01$ ）。また、ウェイクアップ方法とフィードバック方法の交互作用は有意でなかった（ $p > 0.05$ ）。

時間の主効果が有意に認められたウェイクアップ方法について、時間平均値は、I1（注視）条件で 1.73 s (S.D.: 0.26), I2（サインネーム）条件で 0.79 s (S.D.: 0.20), I3（手振り）条件で 0.46 s (S.D.: 0.29)であった。ウェイクアップ方法 3 水準間の時間平均値の差を調べるために、Tukey 法の多重比較（群の数=3, $v=72$, $\alpha=0.05$ および $\alpha=0.01$ ）を行った結果、I1（注視）条件は I2（サインネーム）条件と I3（手振り）条件より有意に長くなったことが認められた（ $p < 0.01$ ）。また、時間の主効果が有意に認められたフィードバック方法について、時間平均値は O1（青いバー表示+明度低）条件で 1.09 s (S.D.: 0.51), O2（O1 に手話表示追加）条件で 1.15 s (S.D.: 0.63), O3（O1 に首振り運動追加）条件で 0.74 s (S.D.: 0.39)であった。フィードバック方法 3 水準間の時間平均値の差を調べるために、Tukey 法の多重比較（群の数=3, $v=72$, $\alpha=0.05$ および $\alpha=0.01$ ）を行った結果、O3（O1 に首振り運動追加）条件は O1（青いバー表示+明度低）条件より有意に短くなったことが認められ（ $p < 0.05$ ）、また、O3（O1 に首振り運動追加）条件は O2（O1 に手話表示追加）条件より有意に短くなったことが認められた（ $p < 0.01$ ）。

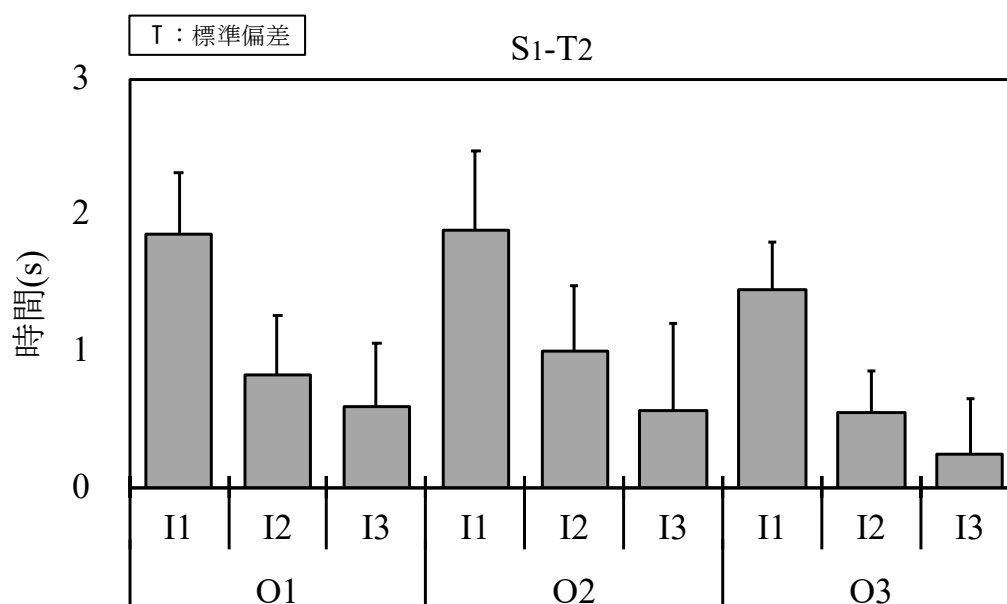


図 47：9 条件毎のウェイクアップ終了からシステムの反応開始(変化開始)までの時間平均値

3.3.6 主観評価の解析結果

フィードバックに対する意識調査

参加者は、ウェイクアップとコマンド入力を連続で行うより間を空けてシステムからの反応を確認したいかどうかの結果を図 48 に示す。その結果、「1.とてもそう思う」と回答した参加者は7名、「2.思う」と回答した参加者は4名、「3.やや思う」と回答した参加者は8名、「4.どちらでもない」と回答した参加者は3名、「5.あまり思わない」と回答した参加者は2名であった。

「1.とてもそう思う」と回答した参加者からは、主に「相手（システム）の反応が伺えると安心して用件が言える。」、「こっちがシステムからの反応を確認したいのもあり、間を置かないとシステムの方が僕の手話を認識することが困難になることの懸念があった。」、「普段の生活で手話を使って会話するときに、逐一間を置くことは自然であり、相手がシステムでもそうしたい。」のコメントがあった。

「2.思う」と回答した参加者からは、主に「システムが私の呼びかけに対する反応を示してくれることを確認できてからコマンドを出したい。」、「聴者同士では間はそこまで必要ないと思う。しかし、私は機械が相手でもきちんと反応してくれるか確認してから用件を言いたい。」、「自分は聴学校育ちだったが本大学に入学してから友人を呼ぶとき相手の反応を確認するようになった。そのため今回でも『思う』の結果になった。」のコメントがあった。

「3.やや思う」と回答した参加者からは、主に「Alexa のような音声操作デバイスでは、間は要らなくとも会話できると思う。しかし、私は、呼びかけてシステムから示された反応、首振り運動や手話表示といったフィードバックを確認してから、用件を話したい。」、「普段の生活ではそうだが、忙しい場面、例えば料理する時に間は必要ないかなと考え、『やや思う』にしました。」、「間を空けた方がシステムの私に対する手話の読み取りが安定すると考えている。もし、読み取り精度が人間並みに優れているならば間はなくてもいいかと思いました。」のコメントがあった。

「4.どちらでもない」と回答した参加者からは、主に「システムの手話認識の精度次第。呼びかけ動作に対してきちんと認識してもらえるのであれば間は必要なく、十分でなければ間は欲しいため、『どちらでもない』と回答しました。」、「忙しい、また何か作業する場面と考えると呼びかけてからすぐコマンド出したいと思った。」のコメントがあった。

「5.あまり思わない」と回答した参加者からは、主に「対人コミュニケーションを取るときでも、呼びかけから用件を聞くまでの間はない。むしろ、今回のシステムは最初から私を見ていることがわかるので、私の呼びかけにすぐ気づいてくれるはず。だから、間はあまり要らないと考えた。」、「私はせっかちなところがあり、呼びかけてからもすぐに用件を言いたくなったためです。それでも、システムから反応してくれることのフィードバックは必要であると感じた。」のコメントがあった。

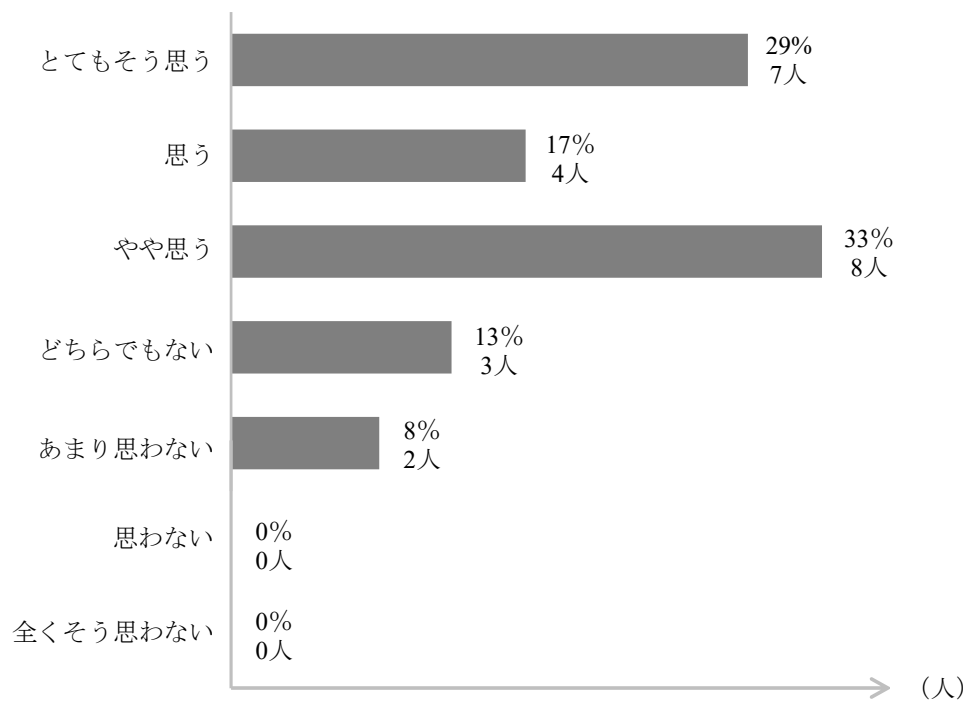


図 48：フィードバックに対する意識の調査結果 (N=24)

手を使ったウェイクアップ動作を行うとき注視を行うことの意識調査

参加者はサインネームや手振りといった、手を用いたウェイクアップ動作でも注視は合わせて行いたいかを確認したいかどうかの結果を図 49 に示す。その結果、「加えたい」の回答パターンに該当する参加者は 19 名、「限定的に加えたい」の回答パターンに該当する参加者は 2 名、「加えなくても良い」の回答パターンに該当する参加者は 3 名であった。

「加えたい」の回答パターンに該当する参加者からは、主に「他の人と会話するとき周囲にあるシステムを見ないで呼ぶとき間違って認識される可能性が懸念されたので、注視も合わせて行いたい。」、「目で見ないで呼びかけることには慣れてない。日常生活の対人コミュニケーションでは、目で見て呼びかけるため。」、「視線を合わせないで呼びかけることは相手に対して失礼と感じたから。」、「人と話すとき、目で見ると癖があるので、注視を加えた方が安心できる。」、「注視を加えることで呼びかけ対象がシステムであることが双方向的に知れるため。」、「システム相手の反応が確認できるためにも注視は加えたい。」のコメントがあった。

「限定的に加えたい」の回答パターンに該当する参加者からは、主に「システム特有の呼び方であるサインネーだけでも、呼びかけ対象ははっきりできるので注視は行わなくてもいいと考えた。周りにいる人に呼びかけるときに使う手振り動作なら、呼びかけ対象を明確にするために注視も合わせて行いたい。」、「仕事で忙しいときなら、注視はしないで呼ぶ方法もあるかも。」のコメントがあった。

「加えなくても良い」の回答パターンに該当する参加者からは、主に「サインネームと手振りで注視は必要ないと思う。システムは私の手話が読み取れるならそれで見なくても問題ない。」、「サインネーム、手振りの場合見る必要はないと思った。家に帰って使用する場面を想定すると、画面を見ながら呼びかける気持ちが湧かない。注視することに縛られたくない。呼びかけの時は見ず、システムが応答した時に見るのがいい。」のコメントがあった。

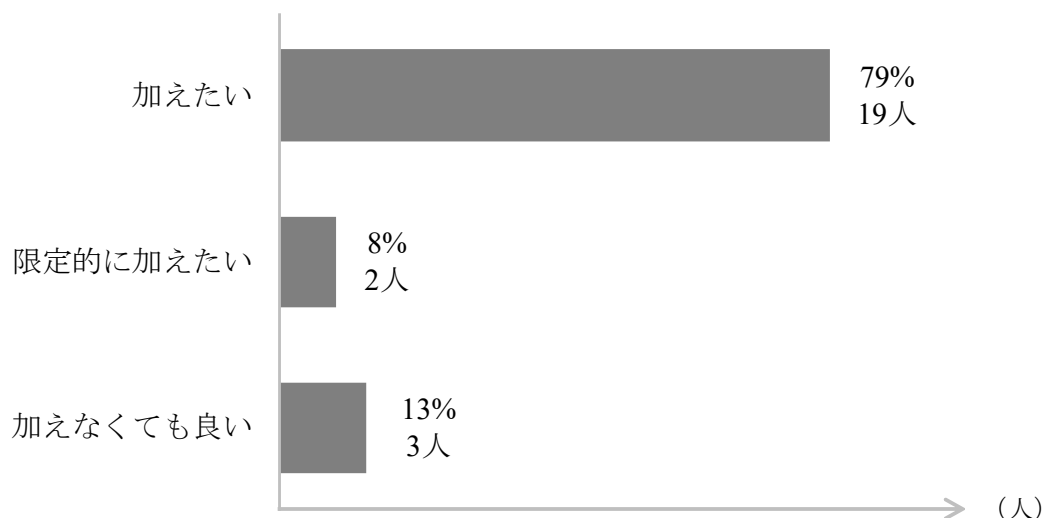


図 49：サインネームおよび手振りの動作中における注視の意識調査結果(N=24)

3.4 考察

RQ4:フィードバックに対する意識

ユーザビリティの評価指標である「有効さ」，「効率」，「ユーザの満足度」の3つの度合い[2]をもとに，フィードバックを提示することは参加者のユーザビリティ向上に寄与したか考察を行う．システムがフィードバックを提示しないことと比べて，提示する方が，参加者の満足度が有意に高くなった（図 30）ことから，「ユーザの満足度の度合い」におけるユーザビリティは向上したことが考えられる．しかし，システムがフィードバックを提示しないことと比べて，提示する方が，システムに呼びかけてからコマンド入力するまでの時間が有意に長くなった（図 31）ことから，「効率」におけるユーザビリティは低下したことが考えられる．また，フィードバックを提示するより，提示しない方が，満足度の順位は下がった（図 38）．そして，不安に感じることのコメントが多かったことから安全性は確保できず，ゆえに「有効さ」におけるユーザビリティは低下したことが考えられる．よって，ろう・難聴者ユーザとシステムの会話において，コマンド入力を早くするより，システムからの反応をきちんと獲得できる状態にすることがユーザビリティ向上のために重要であることが示唆される．また，ろう・難聴者ユーザは，ウェイクアップ動作とコマンド入力の間に時間を空けてシステムからの反応を確認したい傾向が見られた（図 48）．

以上から，ろう・難聴者ユーザのウェイクアップ動作に対してシステムよりフィードバックを提示することは必要であることが示唆される．

RQ5:ウェイクアップのより良い方法

図 32 および図 37 より，全体のろう・難聴者ユーザで，注視やサインネームより手振りでウェイクアップをするのが好ましいことが明らかになった．ここで，3.3.3 節より，参加者「手振りは，対人コミュニケーションにおいて，呼びかける時によく使うので楽だった．」のコメントがあった．これは，ろう者が物理的距離のある相手に接触を図る場合，主に手を振ることが報告[47]されている．このような知見と比較して考えると，システム相手にウェイクアップする場面でも，ろう者同士の会話と同様な行動が見られたと言える．そして，3.3.3 節より，参加者の「手振り動作はすぐ移せるものだから」のコメントがあった．また，手振り動作は，他のウェイクアップ条件より，ウェイクアップ動作終了からコマンド入力開始までの時間が有意に短くなった（図 43）．よって，手振りの動作は，次に行うコマンド入力動作まで滑らかに移行するのに有効な動作であることが示唆される．また，手振り条件に対する満足度が高くなった理由として，サインネームより手振りの方が，ウェイクアップ動作開始からシステムの反応開始までの時間は有意に短くなった（図 46）．サインネーム動作は，"A"の指文字を"X"の形で空描きする動作である一方，手振り動作は，手のひらをシステム方向に出し左右に往復運動するといった，終わりのタイミングが特に決まってない動作である．なので，Wizard 側としても手振り条件で早く画面変化させてしまったことが示唆される．そのため，手振りはサインネームより，システムに早く認識され，結果満足度は上がったといえる．

ここで、表 20 より、参加者が手振り動作を行うとき、必ず注視を行っていた。そして、図 49 より、サインネームや手振りといった手を用いるウェイクアップ動作を行うときでも、注視と合わせて行いたいことの傾向が見られた。以上から、注視と手振りの両方を用いることは、ろう・難聴者に最適なウェイクアップ方法であることが示唆される。

ろう・難聴者ユーザの特性について、「聴覚情報に頼らない」、「手話レベルが高い」、「音声情報を用いない」、といった身体特性およびアイディンティティがろう者である参加者は、他のウェイクアップ方法に比べて、手振り動作の満足度は有意に高くなり（表 19、図 39）。そして、「手話レベルが高い」、「音声情報を用いない」といった身体特性およびアイディンティティがろう者である参加者は、他の特性を持つ参加者より手振り動作にかかる時間は有意に短くなった（図 42）。ろう者は、小さい頃から手話を習得しており、また、ろうコミュニティに属していることから手話を含め、手を使ったコミュニケーションに慣れている。ここで、ろう者や難聴者は健常者に比べて視覚的感覚が優れており、また周囲への注意意識があるときにより明らかになることの報告[58]がある。よって、呼びかける相手がろう者の場合、手振り動作にかかる時間は短くとも相手は気づきやすくなる。ろう者の身体特性を持つ参加者は、そういった状況に慣れているため、手振り動作時間が短く、そして手振りのような手を用いる動作を行うことに対して抵抗感はないことが示唆される。

一方、「聴覚情報に頼る」、「手話レベルが低い」、「音声情報を用いる」といった、身体特性およびアイディンティティが難聴者（聴覚障害者）である参加者は、他のウェイクアップ方法に比較して手振りの満足度が高くなる傾向は見られず（図 33、図 34、図 35、図 36）、身体特性およびアイディンティティがろう者である参加者より、手振りにかかる動作時間は有意に長くなった（図 42）。このことから、難聴者（聴覚障害者）の身体特性を持つ参加者は、手振りのような手を用いる動作は時間がかかってしまうため、ウェイクアップ動作として用いることに対して抵抗感があることが示唆される。ここで、「手話レベルが低い」参加者 5 名のうち、3 名（全体の 66 %）が注視を 1 位に選んだ（図 39）。その参加者のコメントに「手を使わない」、「楽な動作だった」が挙げられたことから、難聴者（聴覚障害者）の身体特性およびアイディンティティを持つユーザは、手を用いない、注視を用いることが望ましい可能性が示唆される。また、Vtyuria らは、Wizard of Oz 法によって柔軟な応答を行うスマートスピーカーシステムを想定した実験を行い、ウェイクワードを用いない自然な対話ができる場合に、ユーザの満足度が高くなることを示した[59]。この研究対象者は聴者である。難聴者（聴覚障害者）の身体特性およびアイディンティティを持つユーザは、日常生活のコミュニケーションにおける声の使用や聴覚機能を活用するため、聴者と同様に手振りやサインネームといった明示的なウェイクアップ方法を好まない可能性がある。

RQ6:フィードバックのより良い提示方法

フィードバック方法の違いがシステムの反応開始（変化終了）からコマンド入力開始までの時間に及ぼす影響について、手話表示が他のフィードバックに比べて有意に短くなった（図 44）。ここで、手話表示についての時間で、開始ポイントをシステムの反応開始（変化終了）から反応開始（変化開始）に変えて再度比較を行った結果、時間に有意差は見られなくなった（図 45）。よって、参加者は、システムの青いバー表示、手話表示するといった画面が変化する場合には変化開始直後から「システムはコマンド認識中の状態」と見なすのに対し、システムが首振り運動する場合には首振り運動終了直後から「システムはコマンド認識中の状態」と見なし、コマンド入力を行なっていることが示唆される。

また、フィードバック方法の違いがウェイクアップ動作開始からシステムの反応開始までの時間に及ぼす影響について、システム首振り運動が他のフィードバックと比べて有意に短くなった（図 46）。また、ウェイクアップ動作終了からシステムの反応開始までの時間で、システム首振り運動が他のフィードバックと比べて有意に短くなった（図 47）。これは、Wizard 側が他のフィードバック（青いバー表示、および手話表示）は PowerPoint のみで制御するのに対し、システム首振り運動は PowerPoint と同時に回転台リモコンで制御するため、システムをきちんと作動させようと早く押してしまったことが理由と考えられる。

図 32 より、全体のろう・難聴者ユーザで、満足度が有意に高くなったシステムのフィードバックの提示方法はなかった。しかし、図 38 より、全体のろう・難聴者ユーザで、他のフィードバックより青いバー表示の状態を提示するのが好ましいことが明らかになった。ろう・難聴者ユーザの特性について、アイディンティティがろう者である参加者は、システムの首振り運動と比べて、青いバー表示のみの満足度は有意に高くなり（図 36）、手話レベルが高い参加者は満足度順位付けで、有意に青いバー表示のみを好んだ（図 40）。また、3.3.3 節より、参加者のコメントに、「青いバー表示だけで十分」が挙げられたことから、システムの反応が示すのに要する時間は短かったことがシステムのユーザビリティ向上に繋がったことが示唆される。しかし、参加者の入力方法は手話であるが、出力方法は文字や青いバー表示と非対称になっている。本来は、手話表示の方が自然なインタラクションが行えるはずと考える。満足度順位結果より、「声あり中間手話が読み取れる」といった手話レベルの低い参加者(N=5)全員が、手話表示を 1 位に選ばなかった（図 40）。しかし、「声なし日本手話が読み取れる」といった手話レベルの高い参加者(N=11)について、手話表示を満足度 1 位と 2 位を合わせて一番多かった。ここで手話表示が青いバーのみ表示より満足度が下がった理由として、3.3.3 節の手話表示に対する参加者のコメントから、「手話者はアバターで表示する」、「ホーム画面といったデフォルト状態から表示させ、ユーザの入力に反応するようにする」など、手話表示の仕様を改善することで、満足度は向上する可能性がある。以上から、システムのフィードバックには更なる検討の必要性が示唆される。

第4章 ろう・難聴者に最適な会話型 NUI 設計ガイドラインの提案

第2章，第3章における研究課題の検証を踏まえ，聴者と比較した，ろう・難聴者が会話型 NUI と会話を行うときに実現すべきであると考え，プロセスの様式を図 50 に示す。

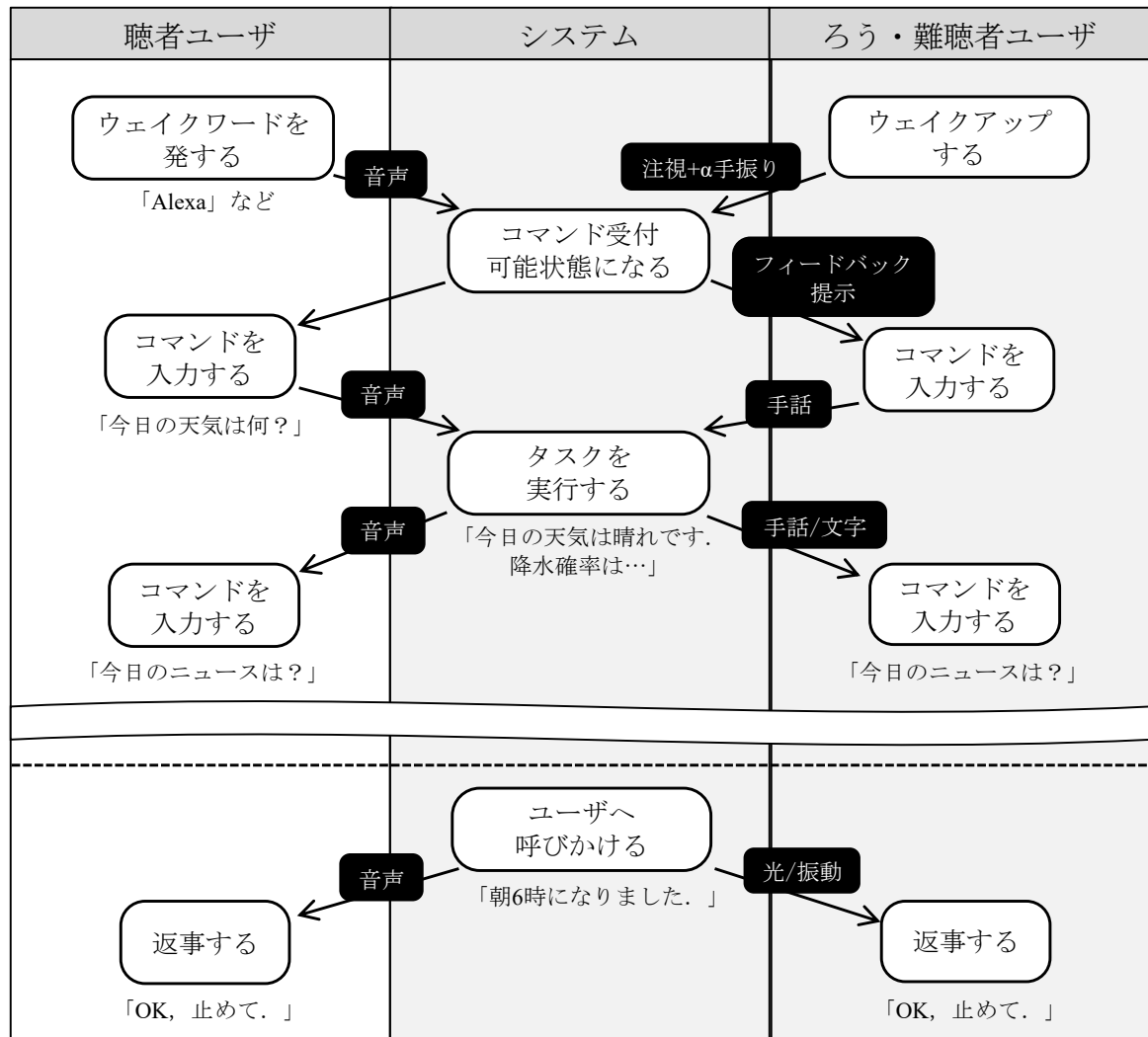


図 50：本研究で提案するろう・難聴者ユーザとシステムとの会話プロセスの様式

図 50 をもとに，ろう・難聴者に最適な会話型 NUI を設計するための 5 つのガイドラインを提案する。

1. ウェイクアップは注視ベースで行えるようにする

ろう・難聴者ユーザにとって，システムを注視してからコマンド入力することは自然である (RQ3)．しかし，ユーザ特性によって好みのウェイクアップ方法は異なることに注意する必要がある．身体特性およびアイデンティティが「ろう」であるユーザには，手振りもオプ

ションとして使えるなどパーソナライズする (RQ5) . また, 設計者が手振りと注視の両方を認識するシステムを作る場合, 表 21 のデータを参考にできる.

2. ウェイクアップに対するフィードバックを提示する

ろう・難聴者ユーザは, ウェイクアップした後, システムにコマンドを入力しても良いか確認したい (RQ4) . そのため, ユーザに安心感を持たせるために, システムから, 手話表示などのフィードバックを提示するようにする.

3. コマンド入力の手話で行う

ろう・難聴者によるコマンド入力, 音声に代わる入力手段の研究に, 手話が望ましい [12][13]. また, ろう・難聴者ユーザはシステムと手話でインタラクションを行えることに関心を示している[14]ため, 手話で入力を行えるようにする.

4. システムからは手話と文字を出力する

文字出力は, 片手間にシステムとインタラクションを行う場合に最適であり, 手話入力をメインに行うユーザには, システムの出力モダリティとして手話が最適である (RQ2) . しかし, 手話に慣れていないユーザにとって手話を出力することは, 必要でない可能性があるため, 表示方法についてもパーソナライズをする必要がある.

5. ユーザに呼びかける時は光か振動を用いて出力する

ろう・難聴者がシステムを視界から外す時に, システムから呼びかけられ, それに気づかない場面がある. それを避けるために, 光による通知を用いた出力を行う (RQ1) . なお, モバイルアプリ (スマートフォン, タブレットなどのデバイス上で直接動いているアプリ) について, ろう・難聴者のユーザ体験を考慮し作成されたガイドラインにおいても, 通知に振動を用いることが検討されていた[60]. そのため, ろう・難聴者に最適な会話型 NUI において, 振動による通知方法を活用することが期待できる.

第 5 章 結論

5.1 まとめ

本研究では，ろう・難聴者に最適な会話型 NUI の設計ガイドラインの提案を目指し，音声ユーザインタフェースとの会話の各ステップで，ろう・難聴者に最適なアクセス方法を検討した．そのために，検証実験を 2 回に分け，行った．1 回目の実験では，ろう・難聴者ユーザ (N=12) を対象とし，光通知と注視を用いた手話会話システムがユーザビリティの向上に寄与するか調査した．2 回目の実験では，ろう・難聴者ユーザ (N=24) を対象とし，ろう・難聴者に最適なウェイクアップ方法とそれに対するフィードバック方法を検討した．実験結果，ろう・難聴者に最適な通知手段，情報伝達方法，ウェイクアップ方法，フィードバック方法の 4 つについて以下の通りに明らかにした．

1. 光通知を用いた出力はろう・難聴者ユーザの満足度を高めるものだと示した．
2. ろう・難聴者による手話・文字の出力に対する必要性を明らかにした．
3. ろう・難聴者に最適なウェイクアップ方法は自然に行える注視であり，また手振り動作を好むユーザがいることを示した．
4. ろう・難聴者ユーザによるウェイクアップへのフィードバックに対する意識を明らかにした．

以上より，ろう・難聴者によるアクセシビリティが向上するための，将来の会話型 NUI の設計ガイドラインを提案した．

5.2 今後の課題

実験 1 と実験 2 両方の参加者の年齢層は低く、全員が大学生という小規模なサンプルであった。そのため、幅広い年齢層の者の好みや行動パターンの調査を行うことはできなかった。ここで、AI アシスタントを搭載した会話型ユーザインタフェース(CUI)に対する、若年層と高齢者の好みの違いが明らかになったことの報告[61]があった。よって、今後の研究では、より多様な人々を対象とし、システムの評価実験を行なっていく必要がある。

また、実験 1 の参加者の 92 %、実験 2 の参加者の 83 % は、AI アシスタントの使用経験がほとんどない。すなわち、参加者はシステムの使用に慣れていないため、システムの挙動はあまり予想できず、システムに対する評価が変化する可能性がある。よって、参加者にはシステムの使用に十分に慣れてから評価実験を行う必要がある。また、日常生活における課題やユーザビリティの評価も視野に入れていく必要がある。

システムの出力機能についても、今回の研究では、ろう・難聴者ユーザへの応答手段に、光通知を組み入れた。しかし、一部の参加者からは「光だけでなく振動を用いた通知方法もあったほうが気づきやすくなると思う。」とのコメントがあった。今後は、振動通知を組み入れた検証実験も必要であると考ええる。さらに、本実験は参加者前方にシステムが設置された上での結果であり、後方からの通知にも気づけるような方法を模索していく必要がある。

参考文献

- [1] 北原 義典, イラストで学ぶヒューマンインタフェース, 講談社, 2019 年
- [2] Online Browsing Platform(OBP), ISO 9241-11:2018(en),
<https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en> (2022 年 2 月 24 日アクセス)
- [3] UX Week 2008 | Dennis Wixon | The Challenge of Emotional Innovation, Vimeo,
<https://vimeo.com/2893051> (2022 年 2 月 10 日アクセス)
- [4] Online Browsing Platform(OBP), ISO 9241-210:2019(en),
<https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en> (2022 年 2 月 24 日アクセス)
- [5] Iijima, R., Shitara, A., Sayan. S., and Ochiai, Y., Smartphone Drum: Gesture-based Digital Musical Instruments Application for Deaf and Hard of Hearing People, In Symposium on Spatial User Interaction, Association for Computing Machinery, New York, USA, Vol. 25, pp. 1–2, 2021.
- [6] Jeffrey. P. B., Raja. K., Ting-Hao. K. H., Juan. P. F., and Saiph. S., On How Deaf People Might Use Speech to Control Devices, In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, Association for Computing Machinery, New York, USA, pp. 383–384, 2017.
- [7] 梁 一誠, 植村 朋弘, 会話型ユーザインタフェースのためのデザイン方法に関する考察—乳児の子育てを支援するサービス「メティス」を事例として—, 日本デザイン学会研究発表大会概要集, Vol. 28, No. 424, 2019.
- [8] Reeves. B., Nass. C., The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places, Bibliovault OAI Repository, the University of Chicago Press, 1996.
- [9] Candello. M., et al., CUI@CHI: Mapping Grand Challenges for the Conversational User Interface Community, In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, pp. 1–8, 2020.
- [10] TECH DRIVERS, Making sense of Google CEO Sundar Pichai’s plan to move every direction at once,
<https://www.cnbc.com/2017/05/18/google-ceo-sundar-pichai-machine-learning-big-data.html>
(2022 年 2 月 24 日アクセス)
- [11] Abraham. G., Automatic Speech Recognition Services: Deaf and Hard-of-Hearing Usability, In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, USA, pp. 1–6, 2019.
- [12] Jason. R., Evan. G., Brittany. W., Christian. V., and Raja. K., Accessibility of voice-activated agents for people who are deaf or hard of hearing, Journal on Technology and Persons with Disabilities, San Diego, Vol. 7, No. 16, 2019.
- [13] Gabriella. W., Colleen. G., Barbra. S., Kristen. F., Christian. V., and Raja. K., “Alexa, Can You See Me?” Making Individual Personal Assistants for the Home Accessible to Deaf Consumers, Journal

- on Technology and Persons with Disabilities, San Diego, Vol. 8, No. 19, 2020.
- [14] Abraham. G., Vaishnavi. M., and Matt. H., Understanding deaf and hard-of-hearing users' interest in sign-language interaction with personal-assistant devices, In Proceedings of the 18th International Web for All Conference, Association for Computing Machinery, New York, USA, Vol. 24, pp. 1–11, 2021.
 - [15] Abraham. G., Vaishnavi. M., and Matt. H., Accessibility for Deaf and Hard of Hearing Users: Sign Language Conversational User Interfaces, In Proceedings of the 2nd Conference on Conversational User Interfaces, Association for Computing Machinery, New York, USA, Vol. 55, pp. 1–3, 2020.
 - [16] Danielle. B., et al., Sign Language Interfaces: Discussing the Field's Biggest Challenges, In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, USA, pp. 1–5, 2020.
 - [17] Danielle. B., et al., Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective, In The 21st International ACM SIGACCESS Conference on Computers and Accessibility, Association for Computing Machinery, New York, USA, pp. 16–31, 2019.
 - [18] Nielsen Norman Group. 1994. 10 Usability Heuristics for User Interface Design. <https://www.nngroup.com/articles/ten-usability-heuristics/> (2022 年 2 月 24 日アクセス)
 - [19] Jakob. N., Enhancing the explanatory power of usability heuristics, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, USA, pp. 152–158, 1994.
 - [20] Donald. N., The design of everyday things. Doubled Currency, 1988.
 - [21] Ben. S., Catherine. P., Designing the User Interface: Strategies for Effective Human-Computer Interaction, 2010.
 - [22] Christine. M., Cosmin. M., Leigh. C., and Benjamin R. C., Design guidelines for hands-free speech interaction, In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, Association for Computing Machinery, New York, USA, pp. 269–276, 2018.
 - [23] Alexa developer documentation, Alexa Skills Kit, はじめに, <https://developer.amazon.com/ja-JP/docs/alexa/alexa-design/get-started.html> (2022 年 2 月 24 日アクセス)
 - [24] Google Assistant, Conversational Actions, Guides Design guidelines, <https://developers.google.com/assistant/interactivecanvas/design> (2022 年 2 月 24 日アクセス)
 - [25] Christine. M., Cosmin. M., "I don't know what you're talking about, HALexa": the case for voice user interface guidelines, In Proceedings of the 1st International Conference on Conversational User Interfaces, Association for Computing Machinery, New York, USA, Vol. 9, pp. 1–3, 2019.
 - [26] Lopatovska. I., et al., Talk to me: Exploring user interactions with the Amazon Alexa, Journal of Librarianship and Information Science, Vol. 51, 2018.
 - [27] Shashank. A., Aaron. P., Oguz. Ö., Michael. R., and Wolfgang. N., How Compatible is Alexa with Dual Tasking? — Towards Intelligent Personal Assistants for Dual-Task Situations, In Proceedings

- of the 9th International Conference on Human-Agent Interaction, Association for Computing Machinery, New York, USA, pp. 103–111, 2021.
- [28] Canalys, Global smart speaker market 2021 forecast, <https://canalys.com/newsroom/canalys-global-smart-speaker-market-2021-forecast> (2022 年 2 月 24 日アクセス)
- [29] Amazon, ヘルプ&カスタマーサービス, スクリーン付き Echo 端末のユーザ補助機能を使用する, <https://www.amazon.co.jp/gp/help/customer/display.html?nodeId=202158200> (2022 年 2 月 24 日アクセス)
- [30] 寺西立年, 音の聞こえと認識, 日本音響学会誌, Vol. 44, No. 10, pp. 759–767, 1988.
- [31] Frick-Horbury. D., The Use of Hand Gestures as Self-Generated Cues for Recall of Verbally Associated Targets, The American journal of psychology, Vol. 115, No. 1, pp. 1–20, 2002.
- [32] Frick-Horbury. D., Robert. E. G., The Effects of Restricting Hand Gesture Production on Lexical Retrieval and Free Recall, The American Journal of Psychology, University of Illinois Press, Vol. 111, No. 1, pp. 43–62, 1998.
- [33] Martin. H., Tom. G., and Richard. J., Touchless Gestural Interaction with Wizard-of-Oz: Analysing User Behaviour. In Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction, Association for Computing Machinery, New York, USA, pp. 207–211, 2015.
- [34] Julia. S., Charles. C. M., Tommer. L., Scott. E. H., and Jennifer. M., Combining body pose, gaze, and gesture to determine intention to interact in vision-based interfaces, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, USA, pp. 3443–3452, 2014.
- [35] Yee. W., Potential Limitations of Multi-touch Gesture Vocabulary: Differentiation, Adoption, Fatigue, In: Jacko J.A. (eds) Human-Computer Interaction, Novel Interaction Methods and Techniques, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, Vol. 5611, 2009.
- [36] Runpeng. C., Hu. L., and Changshui, Z., A Deep Neural Framework for Continuous Sign Language Recognition by Iterative Training, IEEE Transactions on Multimedia, Vol. 21, No. 7, pp. 1880–1891, 2019.
- [37] Flona. H., Wizard of Oz testing – a method of testing a system that does not yet exist, <https://www.simpleusability.com/inspiration/2018/08/wizard-of-oz-testing-a-method-of-testing-a-system-that-does-not-yet-exist/> (2022 年 2 月 24 日アクセス)
- [38] Fraser, N.M., Gilbert, G.N. Simulating speech systems. Computer Speech and Language, Vol. 5, No. 1, pp. 81–99, 1991.
- [39] 筑波技術大学, トップページ, <https://www.tsukuba-tech.ac.jp/> (2022 年 2 月 24 日アクセス)
- [40] Berke. J., verywellhealth, Hearing Aids and Technology, Assistive Listening Devices for the Deaf and HOH, <https://www.verywellhealth.com/assistive-listening-devices-1046105> (2022 年 2 月 24 日アクセス)
- [41] Matsuda. A, Sugaya. M., and Nakamura. H., Luminous device for the deaf and hard of hearing

- people, In Proceedings of the second international conference on Human-agent interaction, Association for Computing Machinery, New York, USA, pp. 201–204, 2014.
- [42] Frank. B., et al., Understanding the Long-Term Use of Smart Speaker Assistants. In Proceeding the ACM on Interactive, Mobile. Wearable Ubiquitous Technologies. Vol. 2, No. 3, pp. 1–24, 2018.
- [43] Marc. M., et al., Benefits of Sign Language Interpreting and Text Alternatives for Deaf Students' Classroom Learning, The Journal of Deaf Studies and Deaf Education, Volume 11, No. 4, pp. 421–437, 2006.
- [44] Sharma. D., Rao. R., The Combined Effect of Captioning and Sign Language in Understanding Television Content in Deaf, Journal of Communication Disorders, Deaf Studies & Hearing Aids, Vol. 6, No. 1, pp. 1–7, 2018.
- [45] Debevc. M., Kosec. P., Holzinger. A., Improving multimodal web accessibility for deaf people: sign language interpreter module. Multimed Tools Appl, Vol. 54, pp. 181–199, 2010.
- [46] SignGenius, Do's & Don'ts - Getting Attention in the Deaf Community, <https://www.signgenius.com/info-do's&don'ts.shtml> (2022 年 2 月 24 日アクセス)
- [47] Bartnikowska. U., Significance of touch and eye contact in the Polish Deaf community during conversations in Polish Sign Language: ethnographic observations, Hrvatska Revija za Rehabilitacijska Istraživanja, Vol. 53, pp. 175-185, 2017.
- [48] Heath. C., Nicholls. K., Body Movement and Speech in Medical Interaction, Studies in Emotion and Social Interaction, Cambridge University Press, 1986.
- [49] Vaishnavi. M., Abraham. G., Becca. D., and Matt. H., Deaf Users' Preferences Among Wake-Up Approaches during Sign-Language Interaction with Personal Assistant Devices, Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, USA, pp. 1–6, 2021.
- [50] Lieberman A. M., Attention-getting skills of deaf children using American Sign Language in a preschool classroom, Applied psycholinguistics, Vol. 36, No. 4, pp. 855–873, 2015.
- [51] 長南 浩人, 日本手話, 中間型手話, 日本語対応手話の構造の違いが聴覚障害者の手話の理解に与える影響, 教育心理学研究, Vol. 49, No. 4, pp. 417–426, 2001.
- [52] Alisha. P., Kanika. M., and Leah. F., "Accessibility Came by Accident": Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities, In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, USA, No. 459, pp. 1–13, 2018.
- [53] Jonathan. P., et al., PsychoPy2: Experiments in behavior made easy, Behavior Research Methods, Vol. 51, pp. 195–203, February 2019.
- [54] Brooke. J., SUS: A quick and dirty usability scale, Usability Eval. Ind. 189. pp. 1–7, 1995.
- [55] Bavelier. D., et al., Visual attention to the periphery is enhanced in congenitally deaf individuals, The Journal of neuroscience : the official journal of the Society for Neuroscience, Vol. 20, No. 17, pp. 1–8, 2000.
- [56] Chathy. P., Designing Voice User Interfaces: Principles of Conversational Experiences, O'Reilly

Media, 2017.

- [57] Developer documentation, amazon alexa, Alexa Design Guide. Be Available, <https://developer.amazon.com/en-GB/docs/alexa/alexa-design/available.html>
(2022 年 2 月 24 日アクセス)
- [58] Daphne. B., Matthew. W. G. D., Peter. C. H., Do deaf individuals see better?, Trends in Cognitive Sciences, Vol. 10, No. 11, pp. 512–518, 2006.
- [59] Alexandra. V., Adam. F., Exploring the Role of Conversational Cues in Guided Task Support with Virtual Assistants, in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, USA, pp. 1–7, 2018.
- [60] Ricardo. P. S., Matheus. S. B., and Luciana. A. M. Z., Supporting the Development of Social Networking Mobile Apps for Deaf Users: Guidelines Based on User Experience Issues. In Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, Association for Computing Machinery, New York, USA, pp. 278–285, 2018.
- [61] David. G., Gerhard. W., Age-Related Differences in Preferences for Using Voice Assistants. In Mensch und Computer 2021, Association for Computing Machinery, New York, USA, pp. 156–167, 2021.

研究業績

- [1] Kato. T., Shitara. A., Kato. N., and Shiraishi. Y., Sign Language Conversational User Interfaces Using Luminous Notification and Eye Gaze for the Deaf and Hard of Hearing, Proceedings of ACHI'21: The 14th International Conference on Advances in Computer-Human Interactions, pp. 30–36, Nice, France, 2021.

謝辞

本論文は、国立大学法人 筑波技術大学 大学院 技術科学研究科 産業技術学専攻 情報科学コース修士課程在学中に行った研究をまとめたものです。本論文をまとめるまでに多くの方々のご指導、ご支援、ご協力をいただきました。

本論文の作成にあたり、同大学 大学院 技術科学研究科 産業技術学専攻の白石優旗准教授には、指導教員として、院入学半年前より研究テーマについての考え方から研究の遂行まで、丁寧かつ熱心なご指導をいただきました。ここに深謝の意を表します。

同大学 大学院 技術科学研究科 産業技術学専攻の加藤伸子教授には、副指導教員として、本研究の遂行にあたって、ヒューマンインタフェースの知見に基づくご指導を数多くいただきました。ここに深謝の意を表します。

筑波大学 大学院 博士後期課程 図書館情報メディア研究科 図書館情報メディア専攻の設楽明寿氏には、本学客員研究員として、研究の進め方や実験環境の構築、データ解析に対して様々な示唆をいただきました。ここに深謝の意を表します。

また、本論文をご精読いただき、数々のご経験を基に有用なコメントをいただきました同大学 大学院 産業技術学専攻の河野純大先生、副査の井上正之先生に深く感謝申し上げます。

第2章の研究の元になった論文の国際会議発表において、国際手話をご指導いただきました、同大障害者高等教育研究支援センター 障害者基礎教育研究部（聴覚障害系）の 小林洋子講師には、深く感謝いたします。

また、同大学 大学院 技術科学研究科 産業技術学専攻 情報科学コース修士課程の先輩である土屋智彦氏には、日常の議論、実験デザインの構築のサポートなど、多岐にわたりお世話になりました。ここに感謝の意を表します。

ろう当事者視点に立ったご意見をいただきました同大学 産業技術学専攻の米山文雄講師、そしてそれぞれの学問分野を超えた先生方からのご助言やご助力をいただける、このような素晴らしい環境を用意していただいた筑波技術大学へもお礼申し上げます。

また、本研究における第2章と第3章の実験のシステムに合成する手話動画に関して、手話通訳士の宮原二三弥氏には、手話動画の撮影にご協力いただきました。そして、実験の実施に関して、同大学の在學生に参加者としてご協力をいただきました。皆様の協力のもと、本研究を遂行することができました。ここに感謝の意を表します。

最後に、何不自由なく本研究ならびに学業全般の環境を与えていただきました家族には、深く感謝し、お礼を申し上げます。

付録 A 実験 1 の参加者マニュアル

実験する際に行ってもらう内容

1. 実験全般についての説明内容

実験時の作業と注意事項

これから行う実験は手話会話型 UI を用いたインタラクションを解明するためのものです。実験対象者には、以下のことを行ってもらいます。

- 1) 手話会話型 UI を介してタスクを実行
- 2) 仕事用 PC で課題に取り組む
- 3) ニュース/天気予報の応答終わりに素早く巨大 Enter キー叩きで反応

実験時の注意事項

- (1) 実験時間中は他の人々と話をするのは一切禁止します。質問やトラブルの際には、手を挙げて知らせて下さい。また、こちらから指示した入力以外に、コンピュータのキーを叩いてはいけません。
- (2) 実験中、携帯電話の電源は切っておいてください。
- (3) 2 種類の実験に参加してもらいます。以下ではそれぞれを実験 1, 実験 2 と呼びます。
- (4) これからすることをまとめますと、次のようになります。
 - 1) 実験の説明・ユーザ属性アンケート回答…5 分間
 - 2) 練習…15 分間
 - 3) 実験 1 本番…10 分間
 - 4) アンケート記入…2.5 分間
 - 5) 休憩…5 分間
 - 6) 実験 2 本番…10 分間
 - 7) アンケート記入…2.5 分間
 - 8) 総括アンケート記入…20 分間
 - 9) インタビューに回答…20 分間

2. 実験前の確認

- ・ライトは明るすぎないか確認をします。
- ・椅子の高さ、シートポジションを確認してください。

3. 留意事項

- ・手話会話型ディスプレイは常にユーザを見続けています。
- ・課題用 PC に数字を入力するときには、姿勢を正しくしてください。

4. 使用機材

(1) 手話会話型ディスプレイ

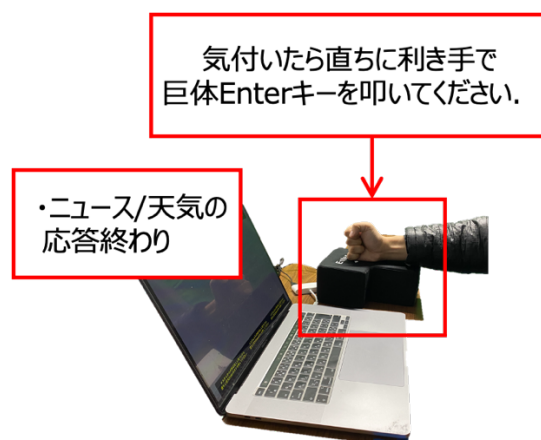


(2) 仕事用 PC



仕事には常に取り組むようにしてください。

(3) 反応



付録 B 実験 1 のアンケート

ユーザ属性アンケート

令和 年 月 日

氏名： _____

性別： 男 ・ 女 ・ その他 _____

——音声会話ユーザインタフェースの使用状況について教えてください。——

現在、普及している音声会話型ユーザインタフェースには次のような例があります。

Apple Siri, Amazon Alexa, Google アシスタント, Google Home, Google Nest Hub,
喋ってコンシェル, Contana など

1. 音声会話型ユーザインタフェースの使用経験を教えてください。

① よく使う ② たまに使う ③ 使ったことある ④ 使わない

④と回答した方は3. の質問に飛んでください。

2. 使用経験のある音声会話型ユーザインタフェースを教えてください。

(回答例：Siri, Google アシスタント)

()

3. 上記のように回答した理由を記入してください。

()

4. 音声会話型ユーザインタフェースに求める改良点があれば記入してください。

()

システム満足度アンケート（光呈示あり）

令和 年 月 日

氏名： _____

-----システムに対する満足度をお聞かせください。-----

	強く 思わない	1	2	3	4	5	強く 思う
Q1. このシステムを今後も繰り返し使いたい							
Q2. このシステムは必要以上に複雑だった							
Q3. このシステムは簡単に使えた							
Q4. このシステムを使えるようになるには、技術に詳しい人のサポートが必要だ							
Q5. このシステムは機能が良くまとまっていた							
Q6. このシステムは一貫性に欠けるところが多かった							
Q7. たいていの人は、このシステムをすぐ使えるようになるだろう							
Q8. このシステムはとても使いづかった							
Q9. このシステムを使っていて不安を感じることはなかった							
Q10. このシステムを使い始める前に、多くのことを学ぶ必要があった							

-----システムからの呼び出し、および応答に対する満足度を教えてください。-----

	強く 思わない	1	2	3	4	5	強く 思う
1. システムからの「呼び出し（電話/アラーム）」…							
Q1. 快適であった							
Q2. 気づきやすかった							
Q3. 光呈示の色に黄色は適切であった							
Q4. 光呈示の明るさは適切であった							
Q5. 黄色での光呈示の点滅周期は適切であった							
2. システムからの「応答（ニュース/天気）」…							
Q1. 快適であった							
Q2. 終わりに気づきやすかった							
Q3. 光呈示の色に緑色は適切であった							
Q4. 緑色での光呈示の明るさは良かった							

システム満足度アンケート（光呈示なし）

令和 年 月 日

氏名：

-----システムに対する満足度をお聞かせください。-----

強くそう思わない 強く思う
1 2 3 4 5

- | | | | | | |
|--|--|--|--|--|--|
| Q1. このシステムを今後も繰り返し使いたい | | | | | |
| Q2. このシステムは必要以上に複雑だった | | | | | |
| Q3. このシステムは簡単に使えた | | | | | |
| Q4. このシステムを使えるようになるには、技術に詳しい人のサポートが必要だ | | | | | |
| Q5. このシステムは機能が良くまとまっていた | | | | | |
| Q6. このシステムは一貫性に欠けるところが多かった | | | | | |
| Q7. たいいていの人は、このシステムをすぐ使えるようになるだろう | | | | | |
| Q8. このシステムはとても使いづらかった | | | | | |
| Q9. このシステムを使っていて不安を感じることはなかった | | | | | |
| Q10. このシステムを使い始める前に、多くのことを学ぶ必要があった | | | | | |

——システムからの呼び出し、および応答に対する満足度を教えてください。——

1	2	3	4	5
---	---	---	---	---

- | | | | | |
|-----------------------------|--|--|--|--|
| Q1. システムからの呼び出しは快適であった | | | | |
| Q2. システムからの呼び出しには気づきやすかった | | | | |
| Q3. システムからの応答は快適であった | | | | |
| Q4. システムからの応答の終わりには気づきやすかった | | | | |

-----システムを使って気づいたことがあれば、記入してください。-----

$$\left(\begin{array}{c} \text{ } \end{array} \right)$$

-----より改良すべき点があれば、記入してください。-----

$$\left(\begin{array}{c} \text{ } \end{array} \right)$$

アンケート

令和 年 月 日

氏名： _____

1. あなたは、光呈示はあったほうが、システムとの会話に有効だと感じますか。
以下の①～⑤に、○を付けて回答してください。

1「当てはまる」 2「やや当てはまる」 3「どちらともいえない」 4「やや当てはまらない」 5「当てはまらない」

2. 1.の質問で、そのように選択した理由を記入してください。

(_____)

3. あなたは、それぞれについてどの程度当てはまりますか？
以下の①～⑤に、○を付けて回答してください。

【ディスプレイ表示に「手話」は必要である】

1「当てはまる」 2「やや当てはまる」 3「どちらともいえない」 4「やや当てはまらない」 5「当てはまらない」

【ディスプレイ表示に「文字」は必要である】

1「当てはまる」 2「やや当てはまる」 3「どちらともいえない」 4「やや当てはまらない」 5「当てはまらない」

4. 3.の質問で、そのように選択した理由を記入してください。

(_____)

5. 光呈示について改良できるところがありましたら、内容を記入してください。

(_____)

6. 今回の通知方法に、光手段を採用しましたが、他に良いと思う通知方法（振動・風など）
がありましたら、内容を記入してください。

(_____)

7. 手話会話型ユーザインタフェースについて、改良できるところがありましたら、
内容を記入してください。

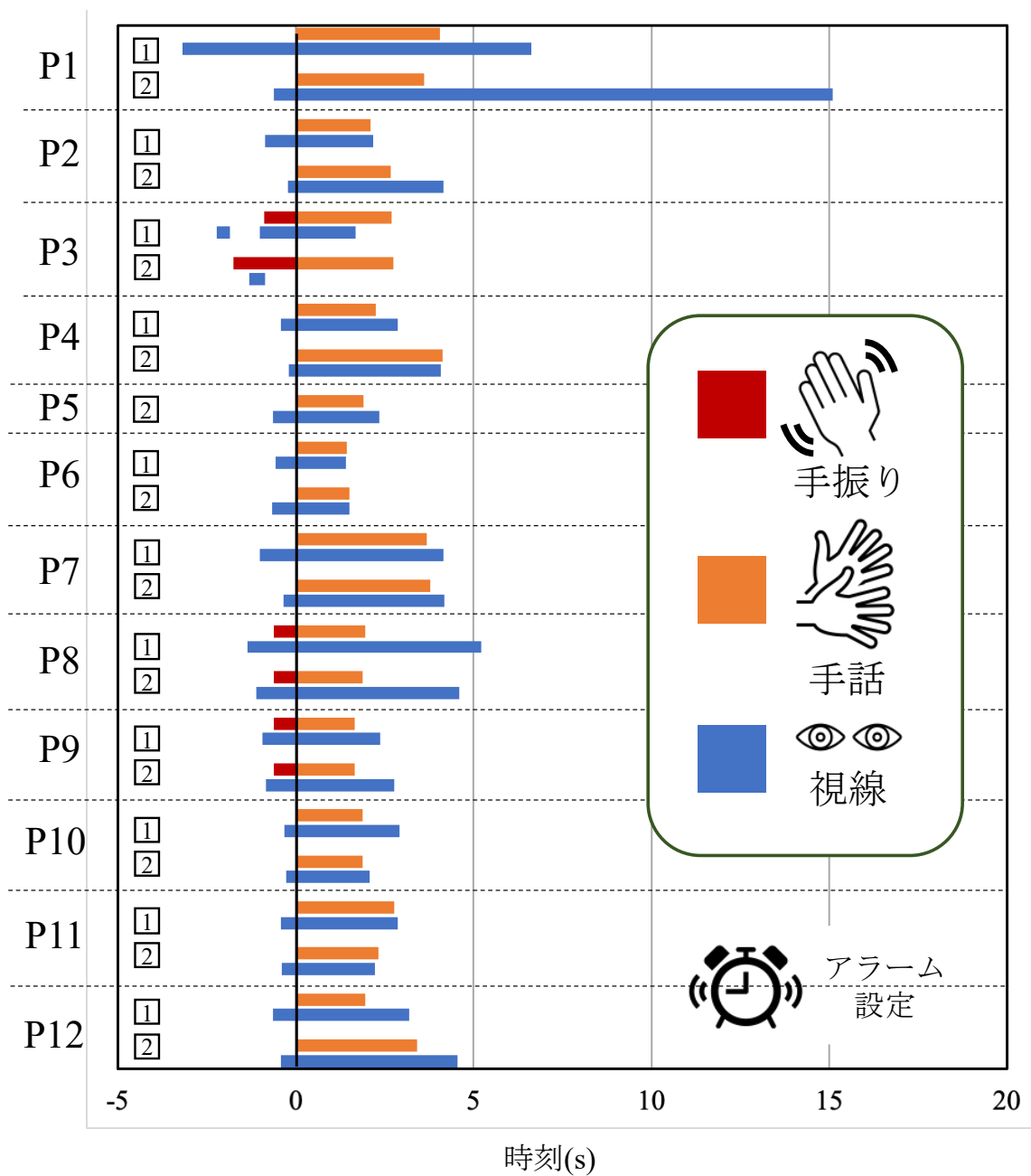
(_____)

アンケートのご協力ありがとうございました。

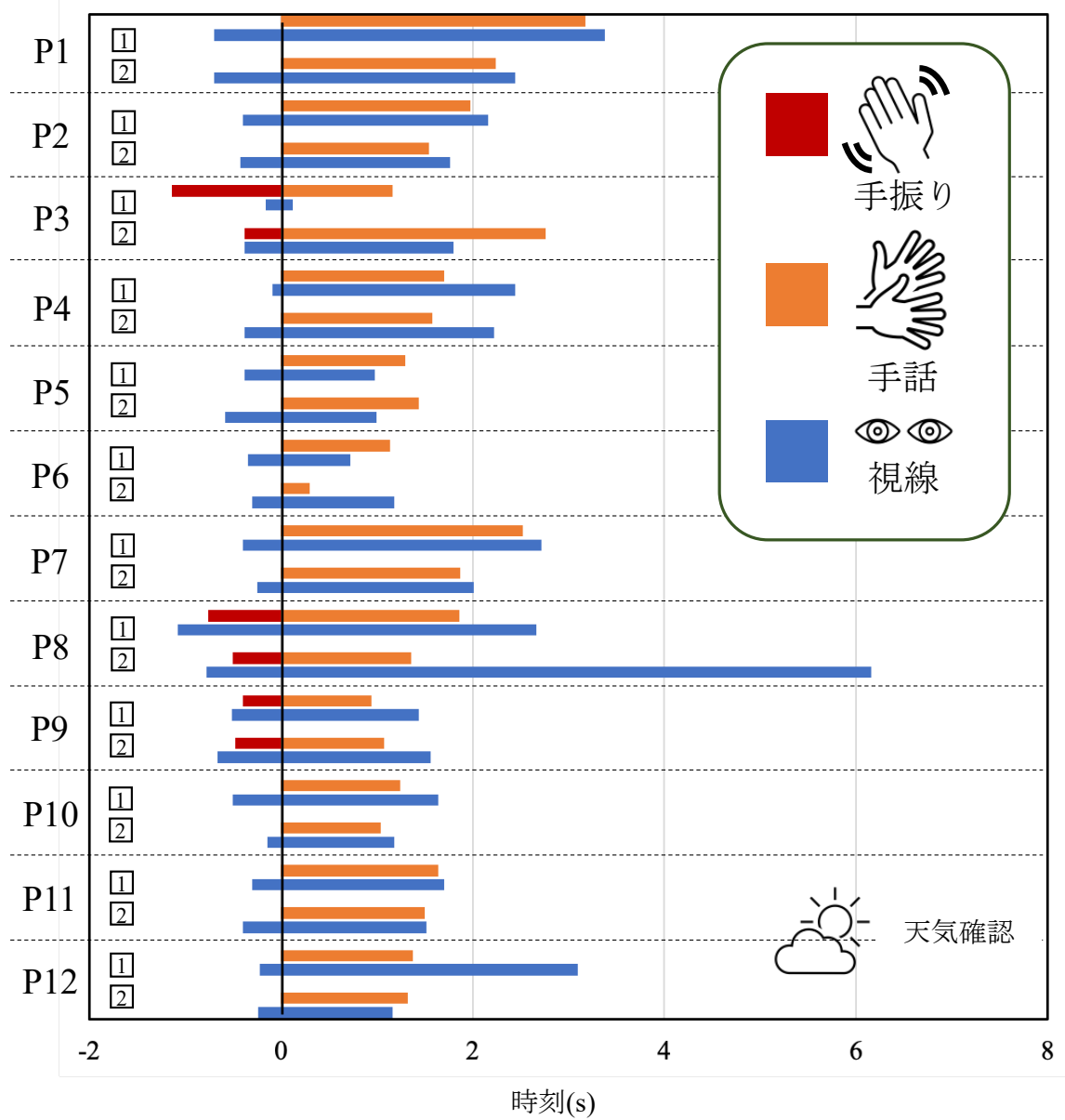
付録 C 実験 1 の結果

参加者およびシステムの挙動時間表

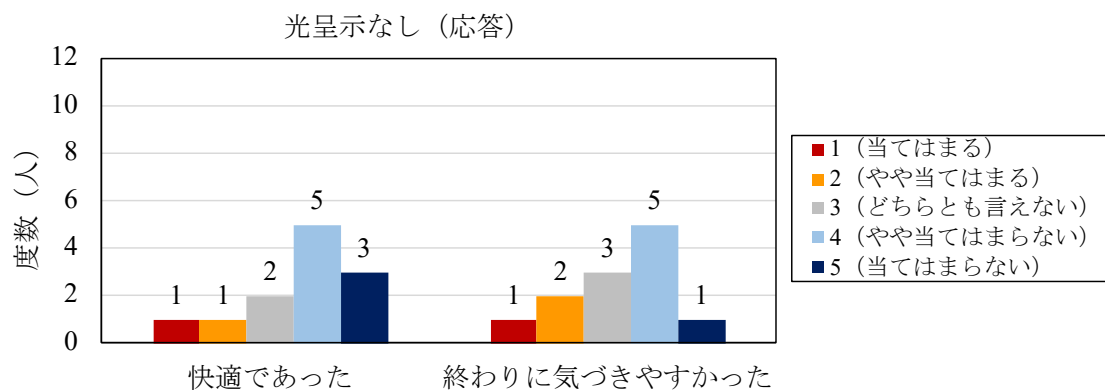
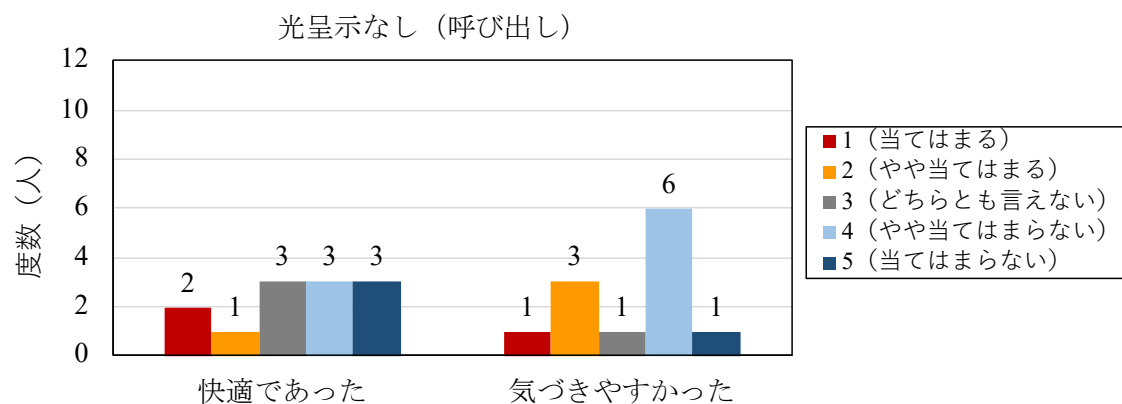
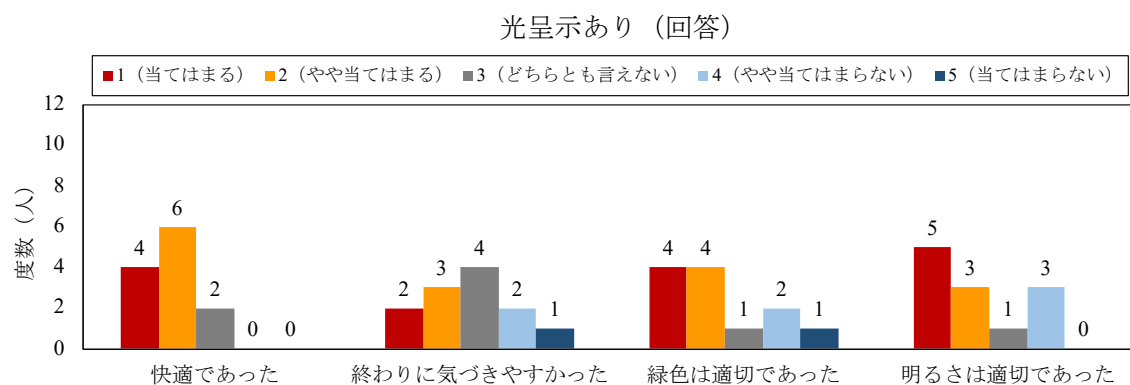
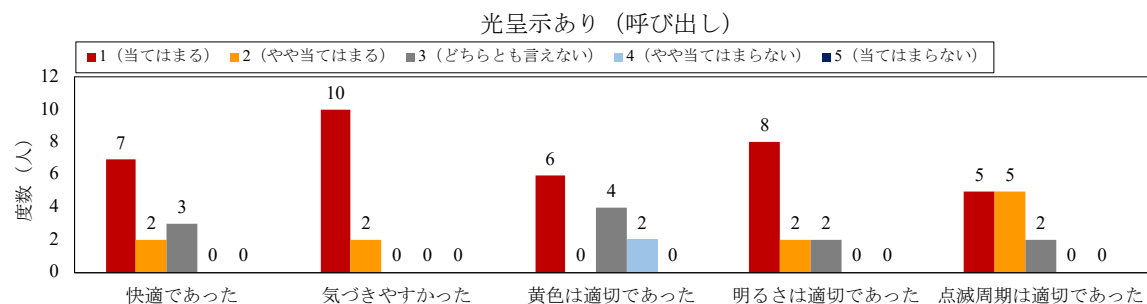
1. アラーム設定

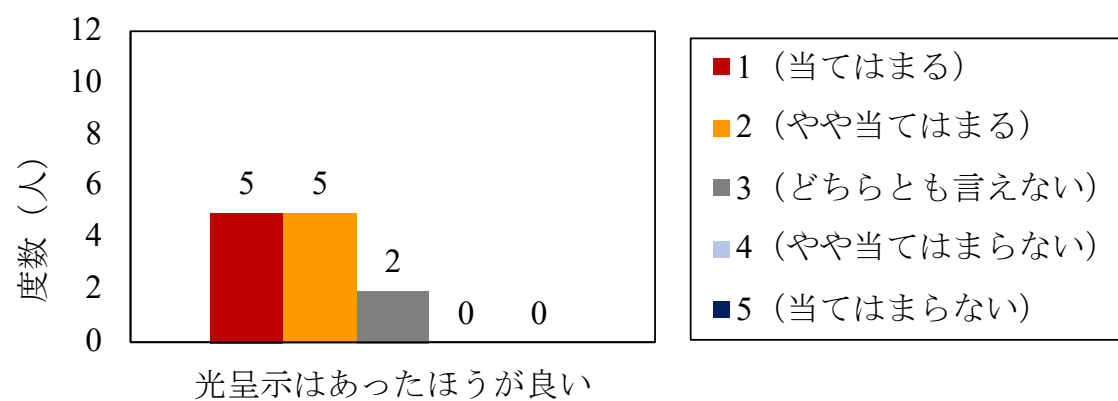


2. 天気確認



光呈示の有無それぞれに対するアンケート結果





付録 D 実験 2 の参加者マニュアル

実験手順

5. 実験全般についての説明内容

実験時の作業

これから行う実験は会話型 UI を用いたインタラクションを解明するためのものです。

実験対象者には、以下のことを行ってもらいます。

- 4) 会話型 UI に対する入力条件(3 つ)と出力条件(4 つ)で組み合わせ入出力方法 12 条件を実践する
※入出力方法の内容については次頁参考
- 5) 各々の条件に対し使用満足度評価(7 段階)とランク付け評価を行う
- 6) 常にアルファベット・数字を一字ずつタイピングする

実験手順(合計 80 分間)

これからすることをまとめますと、次のようになります。

- ① 練習:10 分間
- ② 本番:51 分間
 1. [(実践:2 分間)+(評価:1 分間)]×12 条件:36 分間
 2. 条件ランク付け:5 分間
 3. 途中休憩:10 分間
- ③ アンケート調査:8 分間
- ④ インタビュー調査:11 分間

6. 使用機器



(左)アンケート画面

1つ1つのパターン実施後、
アンケートに回答する

(左)指示画面

次に入力する方法、
システムからの反応方法を
知らせてくれる



(中央)タイピング画面

アルファベットまたは数字が
表示されるのでタイプする



(右)会話型UI

手話で天気・ニュースを
知ることができる

7. フローチャート



【留意事項】

- ・参加者はタイピングしながら会話型 UI と対話する環境にいます。
- ・指示画面の指示を確認したら、アルファベットを3字程度タイプした後に好きなタイミングで会話型 UI に入力を行ってください。
- ・逐一「③応答画面」を確認したら、アンケート(7段階評価)にご回答ください。終わりましたら指示画面を確認してください。それを繰り返し12回で行います。

付録 E 実験 2 のアンケート

1

氏名 *

回答を入力してください

今より実験が始まります。
1回ずつ実践・評価（7段階）を行なっていただきます。
全て12回あります。頑張ってください！！

今から指示画面にあるウェイクアップ動作を確認し、アルファベット3字程度タイプした後に好きなタイミングで実践してください。
それが終わりましたら「次へ」を選択してください。

2

【1回目】ウェイクアップ・システム反応に対する満足度を教えてください（1/12） *

	非常に満足	満足	やや満足	どちらでもない	やや不満	不満	非常に不満
1回目	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

評価ありがとうございます！
続いて2回目の実践に入りましょう。

残り11回です！
今から指示画面にあるウェイクアップ動作を確認し、アルファベット3字程度タイプした後に好きなタイミングで実践してください。
それが終わりましたら「次へ」を選択してください。

（中略...以後 12 回まで繰り返し）

1つ1つの実践・評価はこれでおしまいです！
お疲れ様でした！

まだ続き（3ページ）にアンケートがございます。
ご回答頂きましたら、加藤（研究実施者）が後に参りますのでしばらくお待ちください。

満足度ランク評価

14

3個のウェイクアップ条件を満足した順に順位をつけてください。 *

A. サインネーム

B. 手振り

C. 注視

15

4個のシステムからの反応条件を満足した順に順位をつけてください。 *

A. 青いバー表示+明度低

B. Aに手話表示を追加

C. Aに首振り運動を追加

D. 変化無し

17

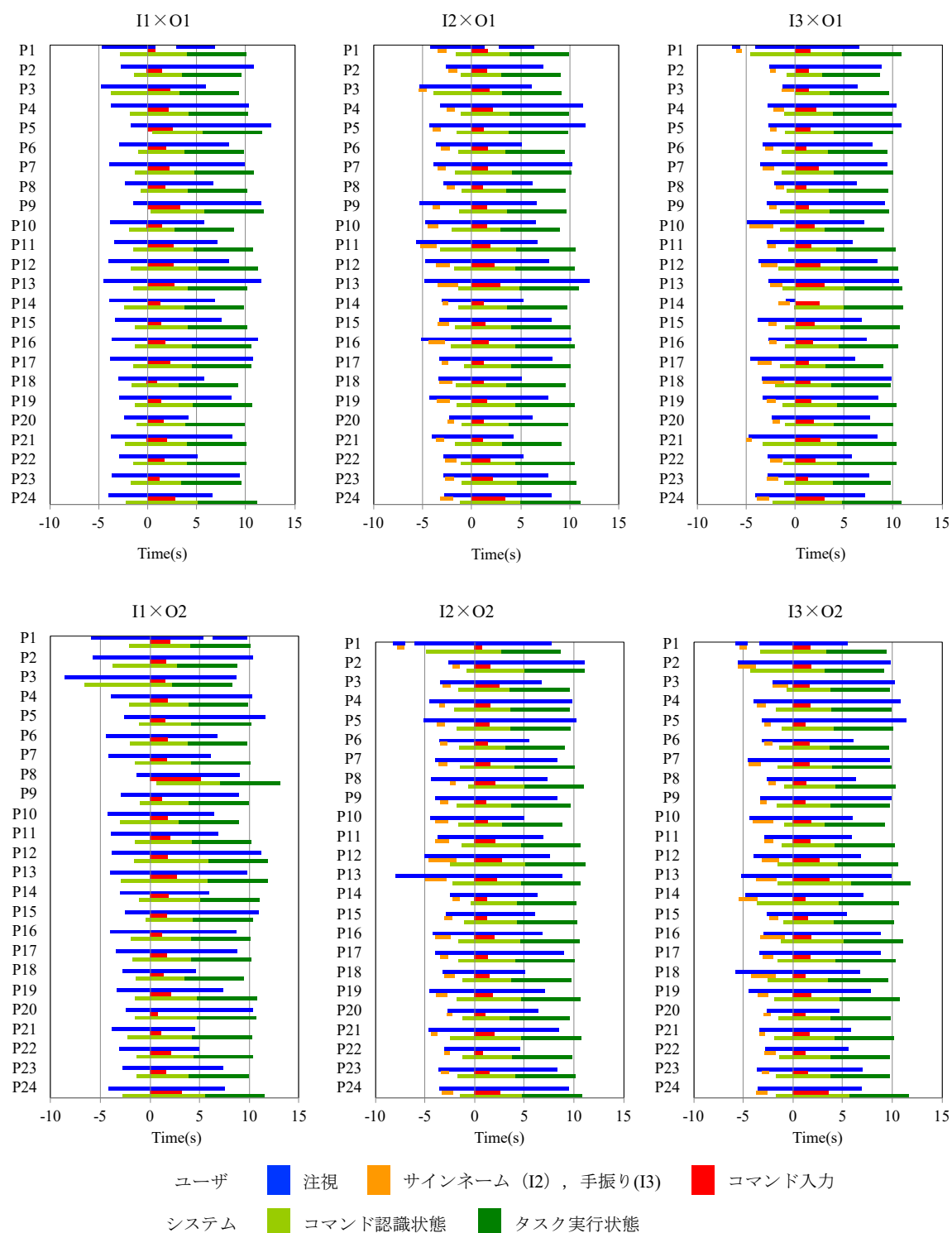
ウェイクアップ動作とコマンド指示を連続で行うより、間を空けてシステムからの反応を確認したいと思いますか？ *

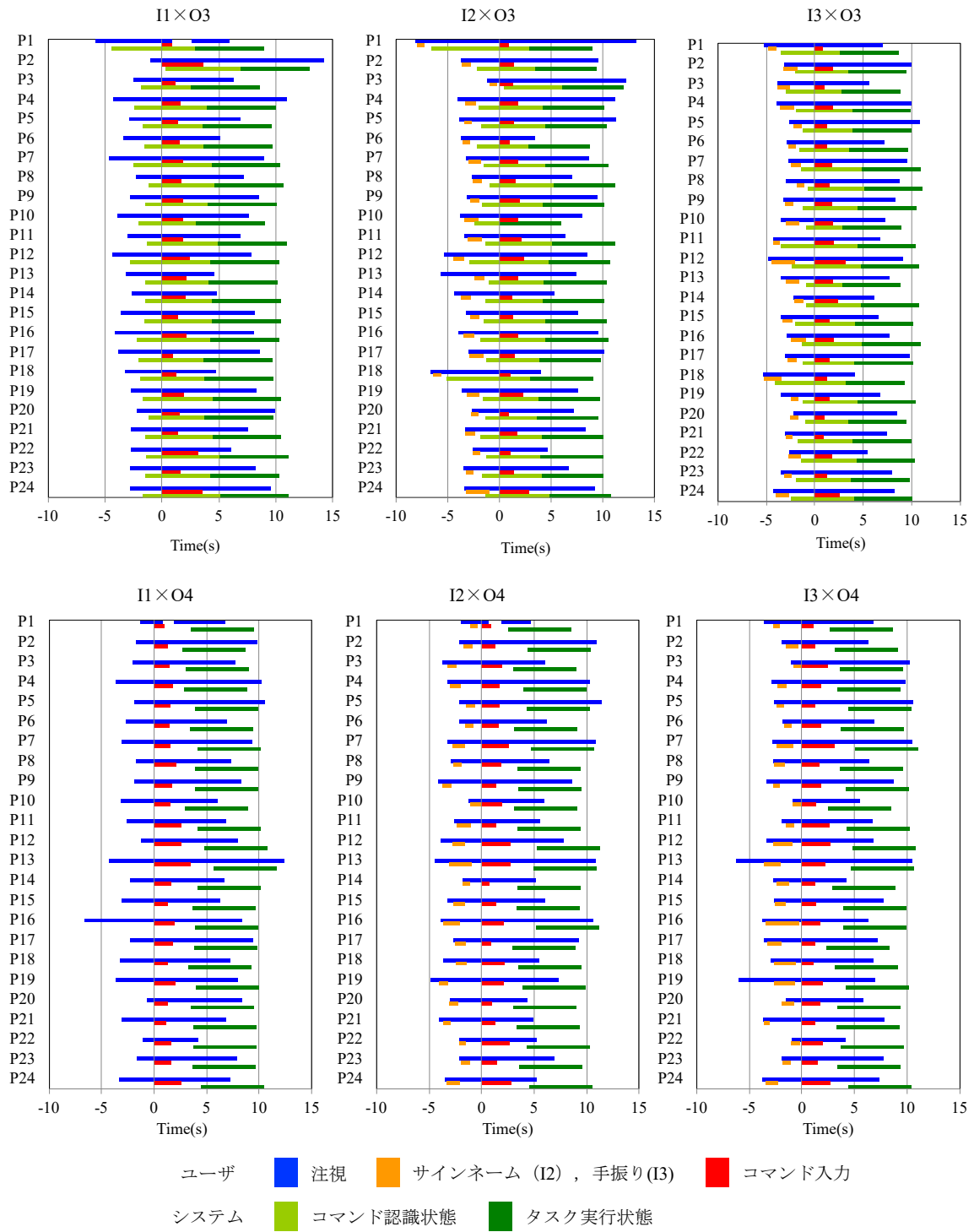
ウェイクアップは、システムにコマンド指示する前に起動させるための動作です。
システムの反応は、ユーザーにコマンド入力可能であることを知らせるためのものです。

- ☐ とてもそう思う
- ☐ 思う
- ☐ やや思う
- ☐ どちらでもない
- ☐ あまり思わない
- ☐ 思わない
- ☐ 全くそう思わない

付録 F 実験 2 の結果

参加者およびシステムの挙動時間表





ウェイクアップ動作にかかった時間のヒストグラム

