

# 対面協調 Web 検索支援のための音声グループアシスタントの構築

立命館大学情報理工学部協調メディア研究室

2021 年 9 月 11 日

本研究では、例えば「一緒にいる 3 人でレストランを探す」などの共通の目的を持って Web 検索を行う「対面協調 Web 検索」という作業を支援するために、ユーザの発話によって Web ページを検索および共有する機能を持った音声グループアシスタントを構築し、その有用性について評価を行う。これまでに、離れたユーザ同士の会話を支援するグループチャット環境に対して、ユーザの会話に基づいておすすめの Web ページを提示する検索ボットを導入するような研究が行われ、ユーザの検索作業への意識が高まることなどの効果が得られているが、対面しているユーザが行っている協調 Web 検索にこのような検索ボットを導入するには、文字入力ではなく、ユーザの発話によって動作することが望ましい。本論文で提案する音声グループアシスタントは、各ユーザが利用する携帯端末と組み合わせて動作し、ユーザの発話による検索機能、および、ページ共有機能を提供する。協調 Web 検索にこのような音声グループアシスタントを用いる場合と、通常の画面操作により検索および共有を行う場合とを比較した結果、音声認識精度が十分ではないことにより使いやすさは低下するが、作業への関わり度や周囲への意識（アウェアネス）は高まる傾向が見られた。

## A Voice Group Assistant for Supporting Co-located Collaborative Web Search

Distributed and Collaborative Systems Laboratory

Faculty of Information Science and Engineering, Ritsumeikan University

September 11, 2021

This paper presents a voice group assistant with functions for searching and sharing Web pages, followed by the results of evaluation on its usefulness in the co-located collaborative Web search tasks, in which multiple people search the Web with a common goal such as "finding a restaurant to go together with three of us." In the past, a search bot has been proposed to assist the Web search in a group chat environment for remote users and its effectiveness such as increased awareness has been obtained. However, such a search bot introduced to the collaborative Web search in a face-to-face manner is expected to operate according to the user's speech instead of text inputs. The voice group assistant proposed in this paper works in combination with mobile terminals used by each user, and provides search and page sharing functions triggered by the user's speech. As a result of evaluation by comparing the use of the voice group assistant for collaborative web search with the use of normal screen operation for search and sharing, we found that the usability of the voice group assistant was reduced due to insufficient speech recognition accuracy, but the degree of engagement with the task and awareness of the surroundings tended to increase.

### 1 はじめに

近年の携帯端末の普及に伴い、複数人で、ある共通の目的を達成するために Web 上の情報の検索や共有を行う「協調 Web 検索」という作業が日常的に行われている [1][2]。協調 Web 検索では、例えば、「一緒にいる 3 人で夕食を食べに行くレストランを決める」という共通の目的にしたがって、各人が Web 検索を行って候補を抽出し、見つけた候補をお互いに共有しながら、最終的に皆が納得するような結論を導く、というような作業が行われる。これまでに、このような作業を支援するために検索結果やコメントなどを共有するシステムやツール

が多く開発されてきたが、実際にはユーザが日常的に利用するには至らず、既存のメッセージングツールなどを Web ブラウザと組み合わせて利用することが多いことから、ユーザが日常的に利用している既存の環境に協調 Web 検索を支援する機能をどのように導入していくかが課題となっている [3]。

このような中、Slack など、ユーザが日常的に利用しているグループチャットツールにユーザ以外の第 3 者的な立場のアシスタントとして検索ボットを導入する研究がある [4]。この研究では、ユーザ間のチャットの中から検索要求を抜き出し、その結果を発言するようなチャットボットを構築することを想定した上で、チャット

トボットの動作を人間が実行する Wizard of Oz[5] 実験を実施することにより、「作業者の互いの活動に対する意識の高まり」「共有、コミュニケーション、合意形成に達する容易さ」「より高いレベルでの楽しみ」といった有用性を作業者に与えたことが示されている。一方、上記の例のように「一緒にいる3人」が対面して協調 Web 検索を行うような場面においては、このような検索ボットを導入する研究は行われておらず、どのような有効性を持つのかを検証する意義は高いと考えられるが、対面環境においてはユーザ間では文字入力ではなく直接の会話によりコミュニケーションが行われることから、検索ボットへの要求を文字入力で行うのではなく、ユーザの発話を理解して動作を行うような音声グループアシスタントを導入することが適切であると考えられる。

本論文では、このような考えに基づいて構築した対面協調 Web 検索支援のための音声グループアシスタントについて述べる。音声によって操作できる第三者的な立場の音声グループアシスタントが協調 Web 検索に加わることで、グループチャットに導入された検索ボットのような有用性が生まれるのではないかと考える。また、音声で操作できるようにすることで、端末上のタップ操作などに比べて使いやすさは低下するものの、作業への関与において一体感が生まれ、他のユーザが何を行っているかに対するアウェアネスが向上することが期待される。本研究では、協調 Web 検索を支援するための音声グループアシスタントが持つべき機能について検討し、実装を行った上で、その有用性について「使いやすさ」「作業への関与度合い」「アウェアネス」の観点で評価を行う。

## 2 関連研究

協調 Web 検索作業を支援するための環境としては、離れたユーザ同士の作業を支援するために、検索したページにコメントなどを付けた上で共有できるようにしたもの [6]、チャットツールと Web 検索ツールを統合することによってコミュニケーションの取りやすさを向上させたもの [7] がある他、対面しているユーザ同士の作業を支援するために、ユーザごとの携帯端末と 1 台の共有ディスプレイの間で検索ページのやり取りをできるようにしたもの [8]、テーブルトップディスプレイを用いたもの [9]、候補の比較を行いやすくするためのインタフェースを提供するもの [10]、携帯端末を把持している方向によって個人作業と共同作業の切り替わりに対するアウェアネスを向上させたもの [11] などが存在している。しかし、先に述べたように、これらのシステムやツールが多くの人に日常的に利用されるには至っておらず、通常のチャットツールのように、ユーザが使い慣れ

た環境に協調 Web 検索を支援する機能を組み込んでいくことが必要とされている。本研究は、その一つとして提案されているグループチャットツール上の検索ボットに相当するものを、対面しているユーザが行っている協調 Web 検索に適用するべく、ユーザの発話を認識して動作する音声グループエージェントを構築しようとするものである。

今日、Apple 社の Siri や Goole 社の Google Home のように、コマンド形式の会話による音声アシスタントが多く存在している。このような音声アシスタントをユーザが日常的にどのように受け止めているかに関する研究 [12] や、グループでの会話中に音声によるパーソナルアシスタントをどのように利用しているかを調査した研究 [13] が存在しているが、これらは音声アシスタントを個人で利用する場合を対象としており、グループで共有して使うことを想定しているわけではない。

また、複数人の会議における参加者同士の自然言語による会話からスマートスピーカの適切な起動トリガとなるワードを選定することにより、オフィス会議における作業を効率化することを試みている研究 [14] がある。この研究では、参加者同士の会話ログを解析することにより、スマートスピーカが有効に機能する場面がどのようなものかについて検討が行われているが、実際にシステムを構築するには至っていない。

これらに対して、本研究では、対面して行われる協調 Web 検索を対象とした上で、グループワークへ導入可能な音声アシスタントを実際に構築した上で、評価を行う。

## 3 音声グループアシスタント

本章では、対面協調 Web 検索環境に導入する音声グループアシスタント (VGA: Voice Group Assistant) の設計方針、機能、実装について述べる。

### 3.1 設計方針

先に挙げた、候補の比較を行いやすくするためのインタフェースを提供することを目的とした研究 [10] では、ユーザの協調 Web 検索作業を分析することにより、作業は「個人検索フェーズ」「意見交換フェーズ」「比較フェーズ」を遷移しながら行われることを示している。具体的には、個人検索で見つけた Web ページに基づいて意見交換を行い、これを何度か繰り返した後、収集した Web ページを比較することにより、最終的な合意に至る。

このような作業を対面して行う場合、通常、各自が持っている携帯端末上の Web ブラウザを用いて個人検索および Web ページの閲覧を行い、必要に応じて画面を見せ合いながら会話するなどの方法で意見交換や比

較を行うことが想定される。このような場面を支援するVGAを導入する上で、VGAがどのような機能を持つべきかについて検討すると、まず、情報を収集する段階である「検索機能」が必要であると考えられる。現在、音声によってクエリを入力することによって検索を行う機能はスマートフォンなどでも実用化されており、有用性や実現可能性が高い。次に、収集した情報を比較するために必要となる「共有機能」が必要であると考えられる。現状、Apple社のスマートフォン等に搭載されているAirDropのように、相手端末を指定した上で現在開いているWebページを共有する機能が実現されているが、これを音声により行えるようにすることが考えられる。

VGAの機能を利用するに当たっては、Apple社のSiriにおける”Hey, Siri”のような起動トリガを用いることとする。起動トリガを用いない方法としては、作業者の会話の中から言葉を認識し、それに適した機能を自動で実行するという方法も考えられるが、自然な会話の中から機能要求を抽出して実行する機能を実現するには、複数人作業者の会話を高い精度で認識する必要があり、今回、音声認識システムとして利用することを想定しているGoogle Speech to Text API[15]では十分な認識精度を得られなかった。したがって、今回はVGAを「イザナギ」と名付け、この名前を起動トリガとして用いる。

また、作業者とVGAの対話においては、コンテキスト認識は行わないこととする。コンテキスト認識を行うことで、作業者は、音声インターフェースを使用する際に、例えば「さっきのレストラン」のように、以前の会話に基づいた要求を行うことができるようになるが、複数人作業者の会話に対する認識精度が十分でない状況においては、誤認識による使い勝手の低下を招く可能性が高いと考える。

さらに、VGAとの会話形式としては、VGAによる質問と作業者による応答を組み合わせたコマンド会話形式を採用する。VGAとの対話を自然会話で行うことも考えられるが、語彙のばらつきや発話区間の区切りなどへ対処するのが困難であると判断した。

### 3.2 機能

3.1で述べたように、構築するVGAは、ユーザが音声でクエリを入力することによってWeb上の情報を検索する機能と、自分の手元の端末で閲覧しているWebページを他者の端末にも表示する共有機能を有する。

検索機能の使用方法を図1に示す。ユーザの一人が(1)「イザナギ 検索」と発話すると、VGAから(2)「検索ワードを教えてください。」という質問が返ってくる。この質問に対して、ユーザが例えば(3)「北海道 観光」という検索クエリを発話すると、VGAが(4)「お調べ

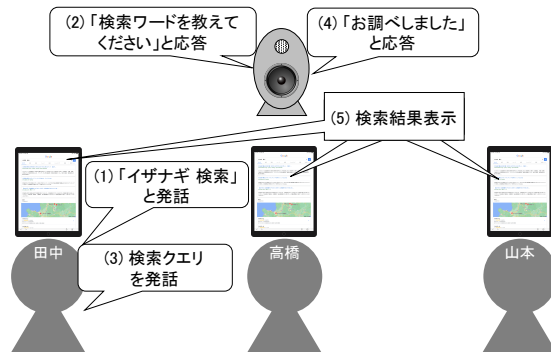


図1 検索機能

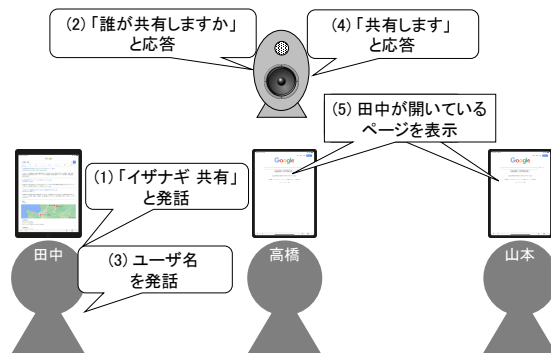


図2 共有機能

しました」と返答し、(5) Web 検索の結果がユーザそれぞれの端末に表示される。

これに加えて、使い勝手を向上させるため、以下のような対話も可能としている。これにより、少ない対話数で検索結果を取得することができる。

- ユーザが(1)「イザナギ 北海道 観光で検索」と発話することで、VGAが(2)「お調べしました」と返答し、(3) 検索結果が端末に表示される。
- 「～で検索」の代わりに、「調べて」「探して」「教えて」という口語調の語句を使う(「イザナギ 北海道 観光で調べて」など)。

次に、共有機能の使用方法を図2に示す。ユーザの一人が(1)「イザナギ 共有」と発話すると、VGAから(2)「誰が共有しますか」という質問が返ってくる。この質問に対して、ユーザが例えば(3)「田中」という送信者名を発話すると、VGAが(4)「共有します」と返答し、(5) 送信者が閲覧しているWebページが他のユーザに共有される。ここで、各ユーザの名前は、ユーザが利用している端末上のアプリケーションの設定画面により設

定できるようにしておき、発話された名前からの端末に表示されているページを共有するのが特定できるようにする。

共有機能においても、使い勝手を向上させるため、ユーザが(1)「イザナギ 田中が共有」と発話することで、VGAが(2)「共有します」と返答し、(3)送信者が閲覧しているWebページが他のユーザに共有される、というような対話も可能としている。

なお、VGAが「イザナギ」という言葉を認識したものの、上記の対話パターン以外の状態に陥った場合には、「認識できませんでした」と返答し、次の発話待ちの状態になる。

### 3.3 実装

本研究で実装した協調Web検索環境のシステム構成を図3に示す。各ユーザが利用する端末としては、Apple社のiPadを用いる。iPad上にはブラウザ画面が表示されており、VGAを通じて検索あるいは共有されたWebページが表示される他、ハイパーリンクのナビゲート、前後のページへの移動、ページの再読み込みができるようになっている。

VGAはPythonで記述されたプログラムによりRaspberry Pi上で動作している。VGAに対する音声入力は、Raspberry Piにマイク付きスピーカをUSBインタフェースにより接続し、Google Speech to Text APIにより音声データをテキストに変換することで行っている。また、検索機能および共有機能におけるVGAの音声出力による応答は、IBM Watson Text to Speechサービスのデモ版を用いて音声合成データを作成することで実現している。

Raspberry Pi上ではVGAとユーザ端末間で情報をやり取りするためのサーバ機能も動作しており、VGAによる検索結果のページ、および、各モバイル端末が開いているページのURLを受信し、保持している。音声による検索が行われた場合には、VGAからサーバ機能に対して検索結果ページのURLが送信され、そのURLが各端末に送信される。また、音声によるページ共有が行われた場合には、VGAからサーバ機能に対して発話されたユーザ名が送信され、そのユーザ名に対応するモバイル端末が開いているページのURLが各モバイル端末に送信される。各モバイル端末がRaspberry PiからURLを受け取ったときには、そのURLに対応するWebページを画面上に表示する。

## 4 評価実験

本章では、音声グループアシスタントの有用性を評価するために行った実験について述べる。

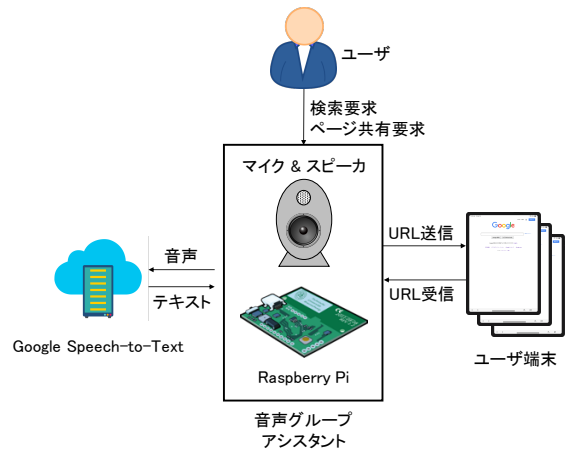


図3 システム構成

### 4.1 実験内容

本実験では、対面協調Web検索作業に音声グループアシスタントを用いる場合と用いない場合とを比較し、音声インタフェースを用いた検索機能および共有機能がユーザにどのような有用性をもたらすかを明らかにする。実験には、3.で述べた音声グループアシスタントを備えたシステム（提案システム）に加えて、音声を用いずに検索や共有ができる機能を実現したシステム（比較システム）を用意した。比較システムでは、検索はGoogleの検索ページを利用して行い、共有は画面上のボタンを押すことで行うことができる。

被験者は情報系の学生6名であり、3名一組の計2グループを形成した。また、各グループ3名の被験者それぞれは、音声の誤認識が生じにくい名前として、実験中に「田中」「高橋」「山本」と名乗ることとし、これらをアプリケーション上のユーザ名として設定した。

次に、被験者が行う作業内容について説明する。被験者は、「研究室旅行で行う宴会の場所を決める」という作業に取り組む。作業に取り組む上で、被験者間でいくつかの候補を比較しながら結論を導くことを促すため、「25人程度が入れること」「4,500円程度までで飲み放題付きのプランを注文できること」「主要駅の近くであること」「全国チェーンのお店はできるだけ避けること」という制約を設けた。対象となる主要駅としては「富山駅」「金沢駅」「高知駅」の中から指定した。結論を出すまでの制限時間は10分とした。

実験手順は表1に示すとおりである。被験者は、比較システムと提案システムのそれぞれについて、使い方について説明を受けた後、作業を実施した。

表1 実験手順

手順	内容
1	実験概要説明
2	比較システムの使い方説明
3	比較システムを用いた作業
4	作業についてのアンケート
5	提案システムの使い方説明
6	提案システムを用いた作業
7	作業についてのアンケート

#### 4.2 評価方針

システムの有用性を検証するために、システムの「使いやすさ」、作業への「関与度合い」、作業における「アウェアネス」の観点でアンケートを行う。使いやすさの評価には、主観的ユーザビリティ測定手法のひとつである IBM Computer Usability Satisfaction Questionnaires[16][17]に基づき、本研究で対象とするシステムに関連する質問で構成したものを用いる。関与度合いの評価には、人々のフロー状態を評価するための指標 [18] を用いる。アウェアネスの評価には、「状況的自己認識尺度」(Situational Self-Awareness Scale)として提案されている質問 [19] を用いる。これら3種のアンケートを「標準設問によるアンケート」と呼ぶ。加えて、協調 Web 検索に特化した独自の質問を用意する(これを「独自設問によるアンケート」と呼ぶ)。

標準設問、独自設問ともに、回答形式は7段階のリッカート尺度であり、「全くそう思わない(1)」「ほとんどそう思わない(2)」「どちらかといえばそう思わない(3)」「どちらとも言えない(4)」「どちらかといえばそう思う(5)」「かなりそう思う(6)」「とてもそう思う(7)」のいずれかを選択する。

さらに、作業風景を録画した映像の観察やシステムに記録されたログの解析を行い、検索機能や共有機能の使用回数、および、音声認識における誤認識回数を計測する。

#### 4.3 結果

まず、提案システム、および、比較システムの機能が何回利用されたかの結果を表2に示す。提案システムにおける数値には、音声の誤認識により利用し直した回数も含まれており、括弧内の数値は誤認識の回数を表示している。比較システムにおいては、検索はそれぞれの端末で行うことから、すべての被験者の合計回数を示している。この結果によれば、グループ2の共有回数を除いて、比較システムの方が利用回数としては多い傾向が見られた。

表2 各機能の利用回数

	グループ1		グループ2	
	提案	比較	提案	比較
検索	3 (1)	9	3 (1)	8
共有	1 (0)	2	5 (0)	3

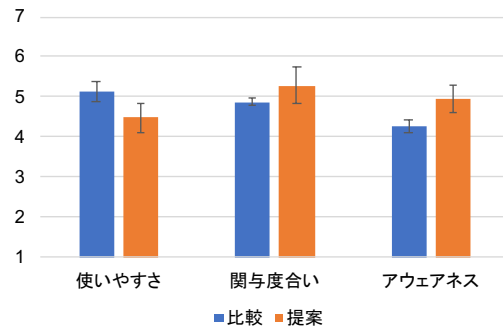


図4 標準設問によるアンケート結果の平均値

次に、使いやすさ、関与度合い、アウェアネスに関する標準設問によるアンケート結果を表3に示す。表では、各被験者のそれぞれの質問に対する回答に加えて、使いやすさ、関与度合い、アウェアネスのそれぞれの観点での回答の平均値を提案システムと比較システムに分けて示している。また、3つの観点ですべてのユーザの回答を平均したものを図4に示す(エラーバーは標準誤差を示す)。

この結果を見ると、提案システムは比較システムに比べて、使いやすさの観点では評価が低くなる傾向が見られる。また、提案システムの使いやすさと比較システムの使いやすさに差があるのかをウィルコクソンの符号付順位和検定で検証したところ、有意差が見られた( $p = 0.027$ )。一方、関与度合いやアウェアネスの観点では提案システムの方が評価が高くなる傾向が見られるが、提案システムが比較システムに比べて関与度合いを上げるかについては有意差が見られず( $p = 0.345$ )、提案システムが比較システムに比べてアウェアネスを向上させるかについても有意差は見られなかった( $p = 0.072$ )。

さらに、協調 Web 検索に特化した独自設問によるアンケートの結果を表4に示す。提案システムで実現した音声による検索機能が使いやすかったかどうかに対する評価は、中央値が4.5、最頻値が7となり、評価が分かれていることが分かる。また、共有機能については、提案システムの音声による共有よりも、比較システムのボタンを押すことによる共有の方が使いやすさの点でやや評価が高くなった。音声認識の精度が十分であったかど

表3 標準設問によるアンケート結果

設問	システム	被験者					
		A	B	C	D	E	F
<b>使いやすさ</b>							
作業を行うために、このシステムは使いやすかった。	提案	3	4	4	5	6	3
	比較	5	5	5	5	6	5
このシステムを使って効果的に作業を行うことができた。	提案	3	4	5	4	5	4
	比較	6	5	5	5	5	5
このシステムを使って効率良く作業を行うことができた。	提案	3	3	4	3	5	4
	比較	6	4	5	5	5	5
このシステムを快適に使用することができた。	提案	2	3	4	4	5	5
	比較	5	4	4	5	6	6
このシステムの使い方を覚えることは簡単だった。	提案	5	5	5	5	7	7
	比較	7	5	6	6	7	7
このシステムを使えば、作業の生産性は向上する。	提案	4	3	4	4	5	4
	比較	6	4	5	5	6	5
このシステムの応答から何をすればよいかわかりやすかった。	提案	4	3	5	5	7	6
	比較	5	3	4	5	7	4
このシステムを使うと楽しかった。	提案	5	3	5	6	7	7
	比較	5	4	5	4	5	4
このシステムに満足している。	提案	3	3	4	5	6	5
	比較	5	4	5	5	7	5
平均	提案	3.56	3.44	4.44	4.56	5.89	5.00
	比較	5.56	4.22	4.89	5.00	6.00	5.11
<b>関与度合い</b>							
このシステムを使った作業は興味深かった。	提案	5	4	5	6	6	6
	比較	5	5	5	5	5	6
このシステムを使った作業は楽しかった。	提案	5	4	5	6	7	7
	比較	5	5	5	5	5	5
このシステムを使った作業に熱中した。	提案	5	4	5	5	7	5
	比較	5	5	5	4	6	4
この作業を使った作業は面白かった。	提案	5	4	5	6	7	7
	比較	6	5	5	5	5	5
このシステムを使った作業にとっても夢中になった。	提案	4	4	4	5	7	7
	比較	5	5	4	4	6	4
このシステムを使った作業に注意を向けることができた。	提案	5	4	5	5	7	6
	比較	5	5	5	5	5	5
このシステムを使った作業に集中できた。	提案	5	4	4	4	7	6
	比較	5	4	4	5	4	5
このシステムを使った作業に没頭できた。	提案	5	4	4	4	7	6
	比較	5	4	4	5	5	5
平均	提案	4.88	4.00	4.63	5.13	6.88	6.25
	比較	5.13	4.75	4.63	4.75	5.13	4.88
<b>アウェアネス</b>							
作業中に、私は作業環境のすべてのことを認識できた。	提案	4	4	5	5	6	5
	比較	5	5	4	5	5	4
作業中に、私の周りで何が起きているか意識できた。	提案	5	4	6	6	7	5
	比較	5	5	5	5	3	4
作業中に、ほかの作業者が私をどう思うか心配だった。	提案	4	4	5	4	7	2
	比較	2	4	4	3	7	2
作業中に、ほかの作業者が何をしていたか把握できた。	提案	4	4	6	5	6	5
	比較	4	5	6	4	1	3
作業中に、他の人は私が何をしているか認識できた。	提案	5	4	6	5	6	5
	比較	5	4	5	4	3	6
作業中に、チームでどれだけ上手に作業を行えているか認識できた。	提案	4	4	5	5	6	5
	比較	4	4	5	5	5	4
平均	提案	4.33	4.00	5.50	5.00	6.33	4.50
	比較	4.17	4.50	4.83	4.33	4.00	3.83

表4 独自設問によるアンケート結果

設問	システム	被験者						中央値	最頻値
		A	B	C	D	E	F		
共同検索機能は使いやすかった.	提案	2	3	4	5	7	7	4.5	7
Web ページ共有機能は使いやすかった.	提案	5	3	4	5	7	6	5	5
	比較	6	5	5	5	6	6	5.5	6
音声認識の精度は十分だった.	提案	4	4	6	5	4	5	4.5	4
結果に全員の意見が反映されている.	提案	5	5	5	5	6	5	5	5
	比較	5	5	6	5	5	4	5	5
他の人と多くの会話をするのができた.	提案	3	4	5	4	5	6	4.5	4
	比較	5	5	6	4	7	5	5	5

うかには、どちらとも言えないという回答が多く、精度の向上が必要であると言える。作業の結果に全員の意見が反映されているかどうかについては、提案システムの場合と比較システムの場合で差は見られなかった。多くの人と会話することができたかどうかについては、提案システムを利用した場合の方がやや低くなる傾向が見られた。

#### 4.4 考察

まず、音声を用いることによる効果について、アンケートの自由記述の結果も示しながら考察を行う。

システムの使いやすさに関して、音声を用いる方が使いやすさが低下するという結果が得られた。この要因としては、自由記述において「音声を使うと、特に誤認識の際に言い直さなければならないので時間がかかる」「音声認識がうまくいかないときがあって使いにくい」などの意見があり、音声認識の精度が十分でなかったことが考えられる。今回利用した Google Speech to Text API では、発話の区切れ目を認識し、発話区間ごとにテキスト化された発話内容が得られるようになっているが、被験者間の通常の会話も連続して音声認識が行われていることから、被験者が検索機能あるいは共有機能を利用する際の発話のタイミングと、通常の会話に対する音声認識の区切れ目のタイミングがうまく合わず、検索要求や共有要求がうまく認識されない場面が見られた。今後、より自然な会話の中から音声グループアシスタントへの要求を聞き分けることができるようにするなど、音声認識技術の進歩と合わせた使いやすさの向上が必要である。

作業への関わり度合いについては、有意差は見られなかったものの、音声を用いる方が向上する傾向が見られた。この要因としては、自由記述において、「作業への熱中度が上がり、周りの雰囲気も和やかになる」「音声認識がうまくいくように集中するため退屈せず取り組めた」などの意見があり、音声によるインタフェースによってグループで行っている作業への意識が高まる効果

を持つことが示唆された。

作業へのアウェアネスに関しては、有意差は見られなかったものの、音声を用いる方が向上する傾向が見られた。この要因としては、自由記述において、「声を出すので周りの人の状況を把握することができた」「声を出すことにより自然と会話が増えたように思う」などの意見があり、音声によるインタフェースは他者が行っていることへの意識を高め、コミュニケーションを行いやすくする効果を持つことが示唆された。

次に、システムが提供している機能の妥当性について考察する。

検索機能について、比較システムでは各々のユーザが個別に検索を行うが、提案システムでは一人のユーザの発話による検索の結果をユーザ全員で共有するようになっている。この点について、自由記述では、「全員で検索結果を共有すると非効率」「検索すると必ず検索結果が画面に表示されるので作業が中断される」という否定的な意見があった反面、「全員で必然的に作業に参加することにつながる」という肯定的な意見もあった。今回の評価実験では、提案システムで検索を行う際にはVGAを用いることを促していたが、個人で検索する場面では音声を使わずに自身の端末で、全員でまとまって検索する場面ではVGAで検索をするような使い分けをすることが妥当と考えられる。

共有機能については、比較システムでは個人で検索した後ページ共有を行うのに対して、提案システムでは全員で検索を行い、同じ検索結果を共有した後に個別のページを共有するようになっている。この点について、自由記述では、比較システムでは「相手が何をしているのか共有するまで分からず心配になった」「共有をして初めてその人の作業が明らかになるので、それまで複数人が同じことを調べていると効率が悪い」などの否定的な意見があったが、提案システムでは、先に述べたように「周りの人の状況を把握することができた」という肯定的な意見が見られた。音声認識の精度によって使いや

すさは低下するものの、音声を用いることによって、グループで一体感を持って作業を行うことにつながるものが伺える。

以上をまとめると、本研究により以下のような知見が得られたと考える。

- 対面協調 Web 検索において、音声による検索機能および共有機能を搭載した音声グループアシスタントを導入することにより、作業への関わり度合いやアウェアネスが向上する傾向が見られ、グループ作業への一体感につながる可能性がある。
- 音声認識の精度への不安、および、実際に誤認識が発生することにより、使いやすさが低下する。
- 検索をする場面において、音声グループアシスタントを用いた検索はメンバーの一致した考えに基づいたクエリを用いる場合に有効であるが、個別に検索したい場合も存在することから、全体での検索と個人での検索を使い分けることが必要である。

## 5 おわりに

本研究では、対面して行われる協調 Web 検索において、検索および Web ページ共有の機能を持った音声グループアシスタントを構築し、その有用性について評価を行った。その結果、通常の画面操作により検索や共有を行う場合と比べて、音声認識精度が十分でないことから使いやすさは低下するが、作業への関わり度やアウェアネスについては向上する傾向が見られた。

今後は、使いやすさの改善に向けて、音声認識技術の発展による精度向上に期待しつつ、より自然な対話の中からユーザの要求を認識できるような方法について検討を進める。また、検索や共有以外の機能の必要性についても検討し、様々な要求に応えられる音声グループアシスタントの実現をしていく。

## 参考文献

[1] Morris, M. R.: A survey of collaborative web search practices, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1657–1660 (2008).

[2] Shah, C.: *Collaborative information seeking: The art and science of making the whole greater than the sum of all*, Vol. 34, Springer Science & Business Media (2012).

[3] Morris, M. R.: Collaborative search revisited, in *Proceedings of the 2013 conference on Computer supported cooperative work*, pp. 1181–1192 (2013).

[4] Avula, S., Chadwick, G., Arguello, J. and Capra, R.: SearchBots: User Engagement with ChatBots During Collaborative Search, in *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*, pp. 52–61 (2018).

[5] Dahlbäck, N., Jönsson, A. and Ahrenberg, L.: Wizard of Oz studies—why and how, *Knowledge-based systems*, Vol. 6, No. 4, pp. 258–266 (1993).

[6] Morris, M. R. and Horvitz, E.: SearchTogether: an interface for collaborative web search, in *Proceedings of the 20th annual ACM symposium on User interface software and technology*, pp. 3–12 (2007).

[7] Bentley, F. R. and Peesapati, S. T.: SearchMessenger: Exploring the use of search and card sharing in a messaging application, in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp. 1946–1956 (2017).

[8] Amershi, S. and Morris, M. R.: CoSearch: a system for co-located collaborative web search, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1647–1656 (2008).

[9] Morris, M. R., Lombardo, J. and Wigdor, D.: WeSearch: supporting collaborative search and sensemaking on a tabletop display, in *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pp. 401–410 (2010).

[10] Komaki, D., Oku, A., Arase, Y., Hara, T., Uemukai, T., Hattori, G. and Nishio, S.: Content comparison functions for mobile co-located collaborative web search, *Journal of Ambient Intelligence and Humanized Computing*, Vol. 2, No. 3, pp. 239–248 (2011).

[11] Teevan, J., Morris, M. R. and Azenkot, S.: Supporting interpersonal interaction during collaborative mobile search, *Computer*, Vol. 47, No. 3, pp. 54–57 (2014).

[12] Luger, E. and Sellen, A.: Like having a really bad PA: the gulf between user expectation and experience of conversational agents, in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 5286–5297 (2016).

[13] Porcheron, M., Fischer, J. E. and Sharples, S.: Do animals have accents?: talking with agents



- in multi-party conversation, in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp. 207–219 (2017).
- [14] McGregor, M. and Tang, J. C.: More to meetings: challenges in using speech-based technology to support meetings, in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp. 2208–2220 (2017).
- [15] Google, : Cloud Speech-to-Text - 音声認識: 入手先 <<https://cloud.google.com/speech-to-text>> (参照 2020-03-31) .
- [16] Lewis, J. R.: IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use, *International Journal of Human-Computer Interaction*, Vol. 7, No. 1, pp. 57–78 (1995).
- [17] Computer System Usability Questionnaire 入手先 <<https://garyperlman.com/quest/quest.cgi?>> (参照 2021-03-29) .
- [18] Ghani, J. A., Supnick, R. and Rooney, P.: The Experience of Flow in Computer-mediated and in Face-to-face Groups., in *ICIS*, Vol. 91, pp. 229–237 (1991).
- [19] Govern, J. M. and Marsch, L. A.: Development and validation of the situational self-awareness scale, *Consciousness and cognition*, Vol. 10, No. 3, pp. 366–378 (2001).