

音声グループアシスタントを用いた 対面協調 Web 検索環境の構築と評価

山本 卓嗣^{1,a)} 高田 秀志^{†1,b)}

概要：

音声認識技術の向上に伴い、スマートフォン、タブレット端末、テレビ、スピーカ、PC など、様々なデバイスに音声インタフェースが搭載されるようになってきている。しかし、これらの音声インタフェースの多くは、パーソナルアシスタントであり、複数人で用いるような機能は搭載されていない。そこで、本研究では、複数人での利用を想定した機能を搭載する音声グループアシスタントの構築を目指す。また、対象環境を、人々が日常的に行う作業である対面モバイル協調 Web 検索作業環境とし、この作業環境に向けた機能を搭載する音声グループアシスタントを構築する。対面モバイル協調 Web 検索において音声グループアシスタントを用いることが作業者に与える効果を評価した実験の結果、Engagement について高い評価が得られた。一方で、音声を用いることから Usability は下がる傾向にあり、Awareness については、音声を用いた方が全体的に評価が高くなったものの、有意差は見られなかった。

1. はじめに

音声認識技術の向上に伴い、スマートフォン、タブレット端末、テレビ、スピーカ、PC など、様々なデバイスに音声インタフェースが搭載されるようになってきている。これらのデバイス^{*1*2*3*4}の登場によって、人々は様々な入力操作を音声で実行できるようになった。さらに、これら多くの音声インタフェースは、名前を付与されることで、擬人化され、人間のように振る舞う。しかし、これらの音声インタフェースの多くは、パーソナルアシスタントであり、複数人で用いるような機能は搭載されていない。

そこで、本研究では、複数人での利用を想定した音声グループアシスタントの構築を目指す。対象とする作業は、人々が日常的に行っている対面モバイル協調 Web 検索作業とする。この作業で利用される機能を搭載した音声グループアシスタントを構築し、作業者に与える効果を検証・評価する。

2. 研究方針

本節では、対象とする協調 Web 検索作業、関連研究について述べる。

2.1 モバイル協調 Web 検索

モバイル協調 Web 検索作業は、複数人の作業者によって、図 1 に示すような流れで行われる [1]。単独検索フェーズとは、作業者がそれぞれ持ち寄ったモバイル端末を用いた Web 検索を行う場面であり、個人で作業が行われる。意見交換フェーズとは、Web 検索によって集まった情報について作業者同士で意見を交換する場面であり、共同で作業が行われる。候補比較フェーズとは、作業者同士の話し合いによって生まれたいくつかの候補を作業者同士で選定していく場面であり、共同で作業が行われる。

このように、モバイル協調 Web 検索作業は個人作業と共同作業を往復することによって行われる。しかし、個人作業と共同作業を往復することはユーザによっては困難であり、モバイル協調 Web 検索作業で重要とされている「作業の共有」「意識の共有」[2] が図りにくいという問題がある。この問題に対して、個人作業を行いながら、「作業の共有」「意識の共有」を図りやすい対面モバイル協調 Web 検索環境を構築する。この環境により、作業者同士のより協調的な作業が実現できると考える。既存環境よりも協調的に作業が行えるようにするために、本研究では、検索クエ

¹ 立命館大学情報理工学研究科

^{†1} 現在、立命館大学情報理工学部

^{a)} is0240hx@ed.ritsumeai.ac.jp

^{b)} htakada@cs.ritsumeai.ac.jp

^{*1} Siri <https://www.apple.com/jp/ios/siri/>

^{*2} Google Assistant https://assistant.google.com/intl/ja/_jp/

^{*3} Google Home https://store.google.com/product/google_home?hl=ja

^{*4} Amazon Echo https://www.amazon.co.jp/dp/B07V1KHDSC?tag=googhydr-22&ref=pd_sl_32ia8287yw_b

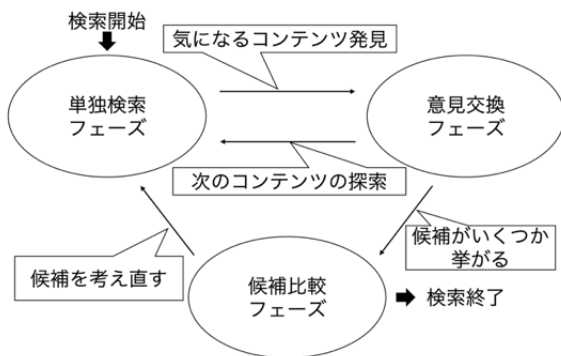


図 1 モバイル協調 Web 検索作業の流れ

りの入力と作業者同士の Web ページの共有を音声操作で実現する環境を提案する。これにより、作業者同士で他者の状態が把握しやすくなり、「作業の共有」「意識の共有」[2] がよりしやすい環境が構築できると考える。

2.2 音声アシスタントを用いた協調作業支援

Google Home^{*3}等の製品化されているスマートスピーカは、ユーザがコマンド会話によって、ユーザのやりたいことを、音声インタフェースを介して実現するデバイスである。音声認識の認識精度を向上させる研究成果は日々向上している。しかし、音声インタフェースをどのようにして日常の環境に組み合わせるかについては十分な研究結果は得られていない [3]。スマートスピーカは、対面環境下におけるマルチユーザでの利用も十分に考えられる。従って、グループワーク等での使用を目的とした機能を検討することで、グループワークを支援できる可能性があると考えられる。

実際に、オフィス会議にスマートスピーカを用いることによる作業効率化を目的とした研究 [4] の研究成果は、ユーザ同士の会話ログの解析にとどまっており実際にシステムの構築には至っていない。そこで、本研究では、日常生活で行われる対面協調 Web 検索でのスマートスピーカの利用方法を検討し、実際にその環境で必要な機能を搭載した音声グループアシスタントを構築する。

グループチャット環境へ検索 Bot を導入し、ユーザの作業を支援することを目的とした研究 [5] では、「ユーザの互いの活動に対する意識が高まる」、「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」といった有用性をユーザに与えることが示されている。しかし、この研究では Bot のアクションを人間が実行する Wizard of OZ [6] 実験を実施しており、実際に Bot の構築には至っていない。本研究では、このような「マルチユーザ+Bot」という環境を対面モバイル協調 Web 検索作業で再現することを目指す。

マルチユーザの使用を想定したスマートスピーカについての研究 [3][4][7] や、スマートスピーカの日常利用について検討した研究 [3][4][8] は多く行われている。特に、マル

チユーザの会議をスマートスピーカを用いて支援することを目的とした研究 [4] では、ユーザ同士の会議で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を行っている。しかし、この研究は、マルチユーザの会話ログの解析にとどまっている。

グループワークでのスマートスピーカの利用を目的とした研究 [3] は、スマートスピーカの既存製品を用いたワークショップを行っている。しかし、既存製品が搭載する機能は、シングルユーザ向けの機能であり、モバイル協調 Web 検索時の作業者の機能要求を満たせないと考える。そこで、本研究では、日常的に人々が行っている対面モバイル協調 Web 検索 [9] で利用されるスマートスピーカを想定し、その場面でのユーザの機能要求を満たすスピーカの構築を行う。さらに、構築する音声グループアシスタントを用いた対面モバイル協調 Web 検索環境を構築し、音声グループアシスタントがユーザに与える効果を検証・評価する。

3. 音声グループアシスタント

本節では、対面モバイル協調 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) の機能と実装について述べる。

3.1 機能

構築する VGA は、音声により与えられるクエリに基づいて Web 上の情報を検索する機能 (共同検索機能) と、ユーザが手元のモバイル端末で開いているページを他のユーザと共有する機能 (共有機能) を有する。また、これらの機能は起動トリガとなる発話によって起動される。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こととした。

検索機能の説明を図 2 に示す。検索機能は、マルチユーザの一人が VGA に「イザナギ 検索 (起動トリガ+検索機能の選定)」、「(検索クエリ)」の 2 ステップで VGA に話しかけると、Web 検索結果がそれぞれのユーザのモバイル端末に表示される機能である。



図 2 共同検索機能

共有機能の説明を図3に示す。共有機能は、マルチユーザの一人がVGAに「イザナギ 共有 (起動トリガ+共有機能の選定)」, 「(送信者名)」の2ステップでVGAに話しかけると、送信者がそのとき開いているWebページを他のユーザに共有する機能である。ここで、iPadには端末名としてユーザの名前が設定してあるものとし、上記の(送信者名)により、どのiPadで開いているWebページを共有するかを特定する。



図3 モバイル端末間のWebページ共有機能

3.2 音声認識機能の実装

VGAの構築には、Raspberry Piにマイク付きスピーカをUSBインターフェースにより接続し、Google Speech to Text API⁵を用いる。このAPIを用いることで、図4に示すように、スピーカへの音声入力をテキスト変換し、テキストに応じて、検索、共有機能を実現する。

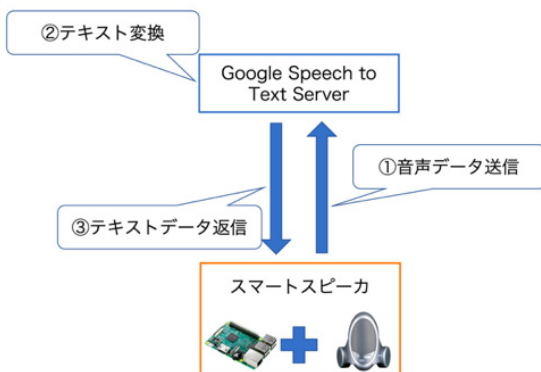


図4 音声認識の実装概要

また、VGAからの応答には、Open Jtalk ライブラリ⁶を用いて、テキストファイルから作成した音声合成のwavファイルを用いる。

3.3 共同検索機能・共有機能の実装

共同検索機能および共有機能の実装には、VGAに加えて、ユーザがモバイル端末として用いるiPadと、iPadやVGAの間でデータ共有を行うサーバとしてMacbook Proを用いる。

サーバでは、CGIライブラリを用いたWebアプリケーションが動作しており、VGAの共同検索機能によって見つかったWebページのURLをiPadに配信したり、共有元のiPadで開いているWebページのURLを他のiPadに配信したりする処理を行う。

4. 評価実験

本節では、音声グループアシスタントを用いた対面モバイル協調Web検索作業のユーザ評価を得るために行った実験について述べる。

4.1 実験内容

本実験は、3節で述べた音声グループアシスタントを用いた対面モバイル協調Web検索作業と、音声グループアシスタントを用いない対面モバイル協調Web検索作業の比較実験である。被験者は、情報系の学生18名(3名一組×6グループ)である。被験者はそれぞれのモバイル協調Web検索環境下で、北海道への旅行計画と、沖縄県への旅行計画を立案し、最終的な旅行計画は、旅行計画シートに記入した。本実験の比較システムとして構築したWeb検索アプリケーションの画面を図5に示す。

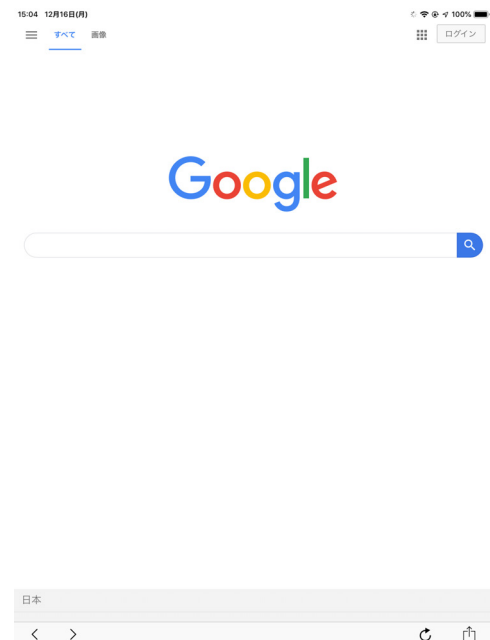


図5 比較システムのアプリケーション画面

音声グループアシスタントを用いたモバイル協調Web検索用のアプリケーションと比較システムのWeb検索アプリケーションはともに、ツールバーにBackボタン、Forward

ボタン, Reload ボタンを設けた。

一方で, 音声グループアシスタントを用いない全体共有機能を実現するため, 右下の Capture ボタンを設けた。被験者は, このボタンをタップすることで他者への Web ページ共有を行うことができる。

また, 順序効果を考慮し, 音声グループアシスタントを用いた旅行計画の立案作業と, 音声グループアシスタントを用いない旅行計画の立案作業は, 全 6 グループのうち, 3 グループは音声グループアシスタントを用いた作業を先に, 残りの 3 グループは音声グループアシスタントを用いない作業を先に行った。

さらに, 作業中の制限として, 作業中に他のアプリケーションを開かないこと (Safari や Google Map アプリケーションに移動すると, ログが取れないため), 他者と iPad を直接見せ合う Web ページ共有を避けること (音声操作による共有と音声操作によらない共有を比較するため) 等を被験者にあらかじめ説明した。

4.2 評価方針

各システムを用いた対面モバイル協調 Web 検索作業についてのユーザ評価を得るため, 各作業後に, Usability[10], Engagement[10], Awareness[10] をそれぞれ実施した。アンケートには, 図 6 に示すリッカート尺度による 7 段階評価を用いた。

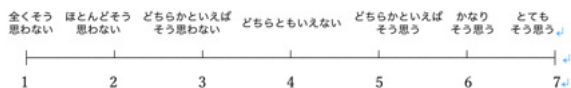


図 6 リッカート尺度

また, 各設問に対して, 自由記述欄を設け, 被験者がなぜその値で評価したかを被験者の任意で回答できるようにした。Usability に関するアンケートの質問内容を表 1, Engagement に関する内容を表 2, Awareness に関する内容を表 3 にそれぞれ記す。

4.3 結果と考察

Usability に関するアンケート結果を図 7, Engagement の結果を図 8, Awareness の結果を図 9 にそれぞれ示す。

はじめに, Usability について述べる。Usability について, 音声グループアシスタントを用いた作業のユーザ評価の平均値は全体的に低くなった。t 検定 ($p<0.05$) を行ったところ, Q.2($p=0.0176$), Q.3($p=0.0111$), Q.6($p=0.0068$) の結果に有意差が見られた。

記述内容を踏まえて考察すると, 音声グループアシスタントを用いた作業の方が低い評価になった原因として, 「音声インタフェースを普段使用しないため」, 音声入力に抵抗があることがあげられる。また, 入力時に, 音声グルー

表 1 Usability のアンケート内容

設問番号	設問内容
Q.1	作業を行うためにこのシステムは使いやすいかった
Q.2	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
Q.3	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
Q.4	このシステムを快適に使用することができた
Q.5	このシステムの使い方を覚えることは簡単だった
Q.6	このシステムを使えば, 作業の生産性は向上する
Q.7	このシステムのレスポンスから何をすればよいかわかりやすかった
Q.8	このシステムを使うと楽しかった
Q.9	このシステムに満足している

表 2 Engagement のアンケート内容

設問番号	設問内容
	このシステムを使った作業は
Q.1	興味深くなかった ----- 興味深かった
Q.2	楽しくなかった ----- 楽しかった
Q.3	退屈だった ----- 熱中した
Q.4	面白くなかった ----- 面白かった
	このシステムを用いた今回の作業についてどのように感じましたか?
Q.5	全く夢中にならなかった ----- とても夢中になった
Q.6	注意を向けられなかった ----- 注意をむけた
Q.7	全く集中できなかった ----- 完全に集中できた
Q.8	全く没頭できなかった ----- 深く没頭できた

表 3 Awareness のアンケート内容

設問番号	設問内容
Q.1	作業中に, 私は作業環境のすべてのことを認識できた
Q.2	作業中に, 私の周りで何が起きているか意識できた
Q.3	作業中に, ほかの作業者が私をどう思うか心配だった
Q.4	作業中に, ほかの作業者が何をしていたかしっかりと把握できた
Q.5	作業中に, 他人は私が何をしているかしっかり認識できたと思う
Q.6	作業中に, チームでどれだけ上手に作業を行っているか認識できた

プアシスタントと会話するため, タップ操作での入力よりも時間がかかったことも原因として考えられる。実際に, 6 グループの作業平均時間は, 音声グループアシスタントを用いた作業が 23 分 36 秒であり, 音声グループアシスタントを用いていない作業は 21 分 01 秒であった。さらに, 音声認識精度の低さも原因として考えられる。実際に, 作業中に音声入力の誤認識が多く発生した, グループ 2, 6 では, 音声グループアシスタントを用いていない作業よりも著しく評価が低くなっている。

次に, Engagement について述べる。Engagement について, 音声グループアシスタントを用いた作業のユーザ評価の平均値は全体的に高くなった。t 検定 ($p<0.05$) を行ったところ, Q.2($p=0.0392$) に有意差が見られた。また, 自由記述回答として, 「音声を用いた Web ページ共有に新しさを感じて楽しかった」という趣旨の回答を被験者 7 名から得た。さらに, 作業への集中度合いに関する質問項目 (Q.5-Q.8) についても, 音声グループアシスタントを用いていない作業と同等の評価を得たことから, 被験者は音声グループアシスタントを用いた作業により関心を持って取り組めたと考えられる。

次に, Awareness について述べる。Awareness について, 音声グループアシスタントを用いた作業の平均値は, 全体的に高くなった。しかし, t 検定 ($p<0.05$) を行ったところ, 有意差は見られなかった。音声グループアシスタントを用いた作業の平均値が高くなった原因として, 音声インタフェースを採用したことがあげられる。モバイル協調

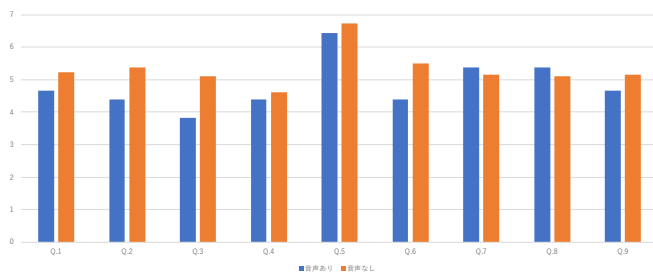


図 7 Usability

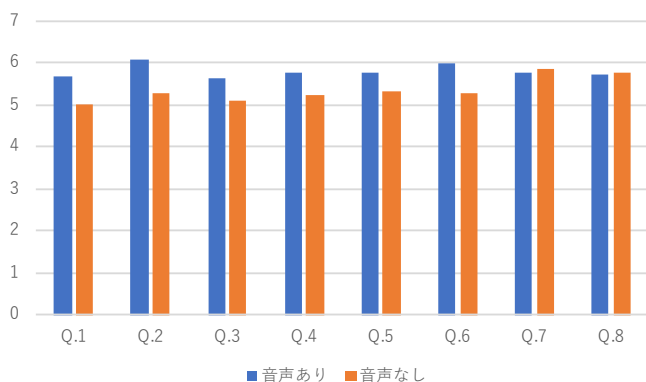


図 8 Engagement

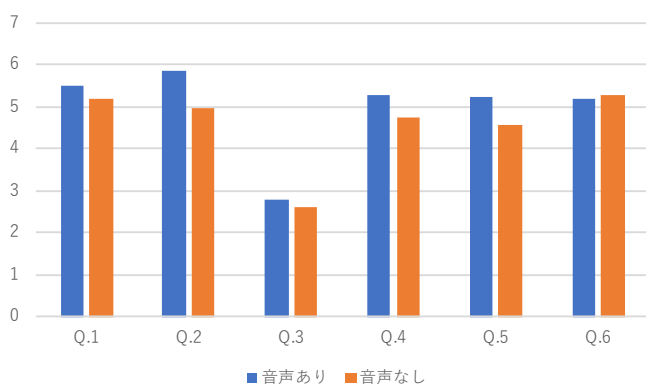


図 9 Awareness

Web 検索では、「意識の共有 [2]」を行うことによって、作業仲間間のコミュニケーションが促進される。したがって、発話しないと Web ページ共有ができない仕様にするによって、普段あまり発言しない人等も発言するようになる。これにより、そのような被験者の意識も全体に共有されるため Awareness の評価値が高くなったと考えられる。一方、有意差が見られなかった原因として、グループによっては共同検索機能があまり使われず、回答のばらつきが多くなったことが考えられる。今回の実験では、音声グループアシスタントを用いた作業用の Web 検索アプリケーション、音声グループアシスタントを用いない作業用の Web アプリケーションともに、タブ機能を導入しなかった。そのため、被験者は、他者から共有された Web ページによって、自身の検索した Web ページの閲覧が一時的に

できなくなるという問題が発生した。本システムは、Back ボタンを設け、Web ページを遷移できるように設計した。しかし、他者からの共有ページが各被験者の作業を中断してしまい、認知的負荷を与えてしまったと考えられる。これにより、観光地を決める人、食事処を決める人等の役割分担を各タスクのはじめに行ったグループ (グループ 3, 4) では、共同検索機能は一度も使われなかった。したがって、タブ機能等を追加し、自身の Web ページ閲覧作業の中断によるストレスを軽減すれば、共通の検索クエリで共同検索機能がより多く利用されるようになり、Awareness の評価値がより高くなると考えられる。

5. おわりに

本研究では、音声グループアシスタントを用いた対面モバイル協調 Web 検索環境の構築を行い、音声グループアシスタントが作業者に与える効果を検証・評価した。音声グループアシスタントの機能として共同検索機能と、モバイル端末間の Web ページ共有機能を設計した。

音声グループアシスタントを用いた対面協調 Web 検索作業と音声グループアシスタントを用いていない対面協調 Web 検索作業の比較実験を行った結果、Engagement について高い評価が得られた。一方、Usability, Awareness に関しては高い評価は得られなかった。今後は、音声グループアシスタントによる Web ページ推薦機能、Web ページのブックマーク機能等を設計・構築し、より実用的な音声グループアシスタントの構築を目指す。

参考文献

- [1] 奥梓, 小牧大次郎, 荒瀬由紀, 原隆皓: 携帯端末を用いた協調 Web 検索におけるコンテンツ比較インタフェース, DEIM Forum 2010(2010).
- [2] 岡田謙一: 協調作業におけるコミュニケーション支援, 電子情報通信学会誌 (2006): 213-217.
- [3] Martin Porcheron, et al: Talking with Conversational Agents in Collaborative Action, companion of the 2017 ACM conference on computer supported cooperative work and social computing ACM, 2017.
- [4] McGregor, Moira, and John C. Tang: More to Meetings: Challenges in Using Speech-Based Technology to Support Meetings, Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. ACM, 2017.
- [5] Avula, Sandeep, et al: SearchBots: User Engagement with ChatBots During Collaborative Search, Proceedings of the 2018 Conference on Human Information Interaction & Retrieval. ACM, 2018.
- [6] Dahlback, Nils, Arne Jonsson, and Lars Ahrenberg: Wizard of Oz studies — why and how, Knowledge-based systems 6.4 (1993) 258-266.
- [7] Luger, Ewa, and Abigail Sellen: Like having a really bad PA: the gulf between user expectation and experience of conversational agents, Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. ACM, 2016.
- [8] Porcheron, Martin, Joel E. Fischer, and Sarah Sharples:

“Do Animals Have Accents?” : Talking with Agents in Multi-Party Conversation, Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. ACM, 2017.

- [9] Morris, Meredith Ringel, Jarrod Lombardo, and Daniel Wigdor : WeSearch: supporting collaborative search and sensemaking on a tabletop display, Proceedings of the 2010 ACM conference on Computer supported cooperative work. ACM, 2010.
- [10] Shah, Chirag : Collaborative information seeking: The art and science of making the whole greater than the sum of all, Vol. 34. Springer Science & Business Media, 2012.