

聴覚障がい者向け複数人会話支援システム Multi-Person Conversation Support System for the Hearing Impaired

後藤颯斗[†] 宮治裕[†]
Hayato Goto[†] Yutaka Miyazi[†]

[†] 青山学院大学 社会情報学部

[†] School of Social Informatics, Aoyama Gakuin University.

要旨

本研究では、複数人会話における聴覚障がい者への情報保証と対話支援を目的としたシステムを提案した。本システムはチャットアプリをベースとし、カメラで手を挙げている人を検知し、その音声のみを文字起こしすることで話者分離を実現した。また、迅速な情報伝達手段として合成音声による発話機能と頻出表現のショートカットボタン等も実装した。実験では既存サービスとの比較を行い、複数の項目で本システムの優位性が示された。一方、音声認識の結果を聴覚障がい者に見やすい形で提示できていない課題も判明した。

1. はじめに

聴覚障がい者にとって他人とのコミュニケーションは難度が高い。特に「議論型」が聴覚障がい者にとって難度が高いとされている[1]。議論型とは複数人で情報を受発信する形式で、会議や普段の会話がこれに該当する。難度が高い理由は主に2つあり、1つ目は「会話内容の理解が困難」という点である。会話内容を理解できないタイミングとして93.4%の聴覚障がい者が「打ち合わせや会議」を挙げている[1]。2つ目は「情報の発信が困難」という点である。81.3%が「意見を言うタイミングが掴めない」、44.7%が「自分の出した情報が他の人に間違っって伝わる」と感じている[1]。

健聴者にとっても聴覚障がい者とのコミュニケーションは難度が高く、50.8%の健聴者が「自分の情報が間違っって伝わる」、51.9%が「会話が噛み合わない」と感じている[1]。

本研究の目的は、複数人会話における聴覚障がい者支援システムの構築を行うことである。聴覚障がい者と健聴者が相互に情報の送受信ができることを目指す。

2. 関連研究・サービス

鈴木[2]は、複数音声をリアルタイムで字幕化し、聴覚障がい者に話者を区別しやすい状態で提示するWEBアプリケーションを提案している。音声認識はボタンで制御しており、話者は発話の際録音ボタンを押下し、発話終了後決定ボタンを押下することで文字起こしが行われる。字幕提示の手法はチャット形式のUIを採用しており、話者ごとの色分け機能、メッセージ間の余白量調節機能等を用いることで複数話者の音声を認識しやすいUIを実現している。

UDトーク[3]はスマートフォンで利用可能なコミュニケーション支援アプリである。チャット形式のUIで、ボタン制御の音声認識、過去メッセージの編集、手書きメッセージ等の機能がある。(無料プラン) 広く普及しているアプリで、多くの聴覚障がい者向けシステムの論文内で引用されている。

3. 提案手法

本システムはチャットアプリがベースのWEBアプリケーションであり、最大5名(内1名が聴覚障がい者)で利用できる。会話参加者は全員PCでシステムを利用する。システムはチャット機能に加え、カメラから手を挙げている人を検知してその人の音声のみを文字起こしする機能、会話参加者の顔映像配信機能、合成音声での発話機能等を有している。

3.1. システム構成

システム構成図を図2.1に示す。画像認識はTensorFlow.js[4]のhand-pose-detectionを利用して手掌部を

検知している。音声認識は react-speech-recognition[5]を利用している。音声認識結果はサーバーへ送信され、WebSocket を用いてクライアント側に送信される。メッセージも同様に送受信を行なっている。顔映像配信機能に関しては SkyWay[6]の P2P 通信を利用している。

3.2. 手掌・音声認識部

手掌の検知には TensorFlow.js の hand-pose-detection を、音声認識には react-speech-recognition を用いている。PC のカメラに手を写すと音声認識が行われ、認識結果はリアルタイムで会話参加者に共有される。システム利用時、発話者は手を挙げると言うルールを設け、話者分離を実現している。

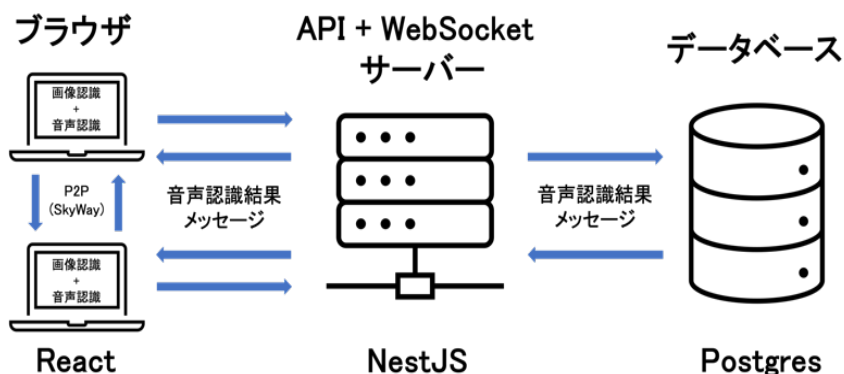


図 2.1 システム構成図

3.3. メッセージ部

会話参加者はタイピングしたメッセージを送信することができる。聴覚障がい者モードにはショートカットボタンが用意されており、頻出する表現（ありがとうございます等）に関しては迅速にメッセージを送信できる。合成音声（Web Speech API）による発話機能も用意されており、健聴者との差を埋める工夫をしている。聴覚障がい者のメッセージは黄色くハイライトされ、メッセージ送信時に通知音を鳴らすことで目立つようにしている。各メッセージに対して、スタンプでリアクションをすることもできる。

3.4. 顔映像配信部

聴覚障がい者の画面には健聴者全員の顔映像が常時配信され、画面から目を話さずに会話参加者の表情を確認することができる。映像配信には SkyWay の P2P 通信を用いている。

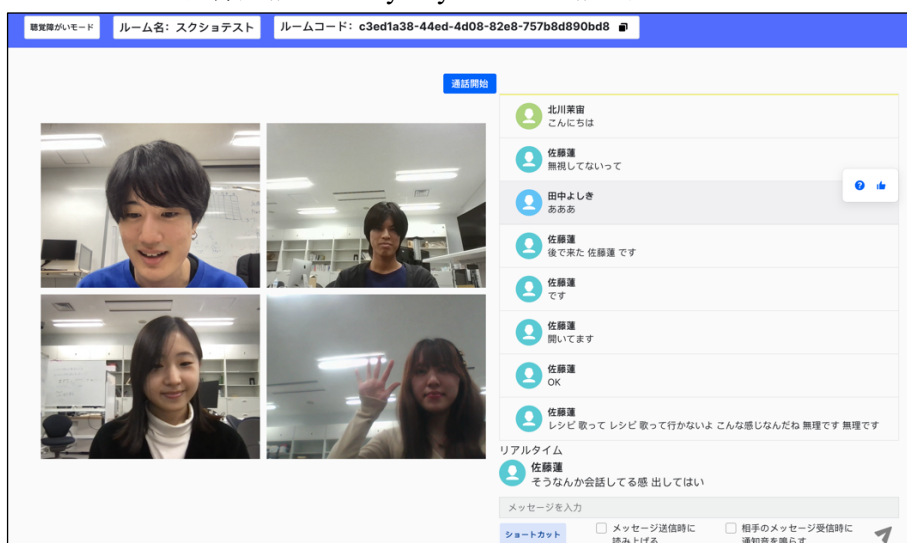


図 2.2 聴覚障がい者モード画面

4. 評価実験

評価実験では既存サービスである UD トークとの比較を行なった。4人1組のグループを作成し、本システムと UD トークを用いて計2回決められたテーマについて話し合った。(5グループ, 計20名) テーマは「好きな食べ物」または「将来行ってみたい場所」で、各グループ1名は聴覚障がい者役として耳栓を装着し、発話を禁止した。各会話が終了後、聴覚障がい者役と健聴者役で別々のアンケート調査(4段階, 自由記述)を行なった。

健聴者アンケートの同項目における本システム, UD トークの比較では, Wilcoxon の符号付き順位和検定を用い, 有意水準は5%未満とした。聴覚障がい者アンケートに関しては, 標本数が少ないため各種統計量を用いて分析を行なった。アンケート調査の結果を表3.1, 表3.2に示す。

表3.1 健聴者アンケート中央値比較 (n=15)

番号	設問	本システム	UDトーク	有意差
1	聴覚障がい者へ素早く情報発信ができた	3	2	あり
2	聴覚障がい者へ容易に情報発信ができた	4	2	あり
3	聴覚障がい者の意見を素早く理解できた	4	2	あり
4	聴覚障がい者の意見を容易に理解できた	4	2	あり
5	会話が円滑に進んだ	3	2	あり

表3.2 聴覚障がい者アンケート比較 (n=5)

番号	設問	統計量	本システム	UDトーク
1	会話の最中疎外感を感じなかった	中央値	3	3
		分散	0.5	1.5
2	意見を言うタイミングが掴めた	中央値	3	2
		分散	0.8	0.3
3	意見を素早く伝えられた	中央値	3	1
		分散	0.7	0.2
4	発信した情報が相手に正しく伝わった	中央値	4	3
		分散	0.3	1.7
5	言いたいことを全て言えた	中央値	3	2
		分散	1.2	0.7
6	健聴者同士の会話が理解できた	中央値	4	2
		分散	0.3	0.7

健聴者アンケートでは全ての設問で中央値が向上しており, 本システムの方が優れているという結果であった。自由記述欄では, 「合成音声発話機能によって普通に会話している感覚があった」「メッセージのハイライトで発言に気づけた」「リアクションボタンが良かった」等の記述が多く見られた。

聴覚障がい者アンケートに関しては, 設問1以外で中央値が向上した。自由記述欄では「顔が見られ

て良かった」「会話が円滑に進んだ」等の記述があった。システムの問題点としては、「手をおろすのを忘れてしまう」「誰が入力中か見たい」「タイピングが遅いと大変かも」「文字を追うのに目が疲れる」等の意見があった。

健聴者にとって本システムは有用であると考えられる。特に「合成音声発話機能」「ハイライト機能」「リアクションボタン」の3つが高評価であり、情報の送受信を容易に行えることによって健聴者の負担を減らすことができたと考えられる。しかし、「手上げ文字起こし」に関しては意見が分かれており、より優れた文字起こし方法を検討する必要があると考えられる。

一方、聴覚障がい者にとっては「情報の受信」で問題があると判明した。音声認識機能によって素早く文字起こしが行われるが、そのスピード感に合った情報提示ができていないと考えられる。実際、「文字を追うのが疲れる」という意見が多く見られ、チャット形式以外の情報提示方法を検討する必要があると考えられる。

5. まとめ

聴覚障がい者にとって複数人会話への参加は難度が高いものになっている。情報の送受信で問題が起きており、本研究では複数人会話で聴覚障がい者と健聴者が相互に情報の送受信を行うことを目的とするシステムを提案した。システムはチャットアプリがベースのWEBアプリケーションであり、「手上げ文字起こし機能」「合成音声発話機能」「メッセージ機能」「顔映像配信機能」を有している。評価実験では既存サービスとの比較を行ない、本システムの優位性が示された。しかし、聴覚障がい者にとっては「情報の受信」で問題があったと判明した。音声認識機能による文字起こし情報の提示が間に合っておらず、今後はチャット形式以外の情報提示方法を検討する必要があると考えられる。

参考文献

- [1] 水野映子, “聴覚障害者が働く職場でのコミュニケーションの問題”, 第一生命経済研究所, 2014.
- [2] 鈴木拓弥, “複数音声同時字幕提示ユーザインタフェースの開発”, 情報処理学会研究報告, Vol.2020-AAC-14 No.2.
- [3] UD トーク, <https://udtalk.jp/>, (参照 2023-11-07).
- [4] TensorFlow.js, <https://www.tensorflow.org/js>, (参照 2023-11-07).
- [5] react-speech-recognition, <https://webspeechrecognition.com/>, (参照 2023-11-07).
- [6] SkyWay, <https://skyway.ntt.com/>, (参照 2023-11-07).