

放射線検査における音声認識入力システムを用いた患者誘導の検討

Investigation of patient guidance using a voice recognition input system in radiological examinations

中西 優太^{1)*}, 鶴野 慶太郎²⁾, 高島 悅²⁾, 高田 麻朗²⁾, 安田 有佑²⁾, 小松 裕司³⁾, 阿部 修司⁴⁾, 吉田 昌裕³⁾

1) (学生) 大阪ハイテクノロジー専門学校 診療放射線技師学科 (現 国立循環器病研究センター)

2) (学生) 大阪ハイテクノロジー専門学校 診療放射線技師学科

3) 修士 (医療安全管理学) 大阪ハイテクノロジー専門学校 診療放射線技師学科

4) 博士 (保健学) 大阪ハイテクノロジー専門学校 診療放射線技師学科

Key words: Radiology, Hearing impairment, Patient Guidance, Voice Recognition Input Systems

【Abstract】

In the past, radiological examinations have been performed remotely using voice instructions for patient positioning and other procedures. However, there is a risk that patients with hearing impairments or hearing loss may not fully understand the explanations given by the radiology technologist during the examination. The purpose of this study was to investigate the usefulness of a simple instruction transmission method using a voice recognition input system, assuming a gastric X-ray fluoroscopy examination.

The results of this study suggest that a simple support transmission method using a voice recognition input system is useful for gastrofluoroscopy for the hearing-impaired, because the use of an appropriate display format made it possible to visually transmit real-time instructions to the subject.

【要旨】

これまで、放射線検査においては患者への体位変換などの指示を音声を用いて遠隔で行ってきた。しかし、聴覚障害や難聴を自覚している患者に対して、診療放射線技師による指示が十分に理解できていない状況で検査が行われている恐れがある。そこで本研究は胃X線透視検査を想定し、音声認識入力システムを用いた簡便な指示伝達方法の有用性を検討することを目的として実施した。

今回の結果より、適切な表示形式を用いることで、被験者に対してリアルタイム性の高い指示を視覚的に伝達することが可能であったことから、聴覚障害者の胃X線透視検査において、音声認識入力システムを用いた簡便な指示伝達方法は有用であることが示唆された。

緒言

放射線検査において、呼吸停止や体位変換などの指示は、検査精度を向上させる上で重要な要因となる。一般的には、検査担当者による口頭指示や装置からの自動音声などといった、聴覚を活用するコミュニケーションが行われている。しかし、難聴などの聴覚障害

NAKANISHI Yuta^{1)*}, HIBARINO Keitarou²⁾, TAKASHIMA Jyun²⁾, TAKADA Mao²⁾, YASUDA Yusuke²⁾, KOMATSU Yuji, M.S.³⁾, ABE Shuji, Ph.D.⁴⁾, YOSHIDA Masahiro, M.S.³⁾

1) Student, Department of Radiological Technology, Osaka College of High Technology (Currently National Cerebral and Cardiovascular Center)

2) Student, Department of Radiological Technology, Osaka College of High Technology

3) Department of Radiological Technology, Osaka College of High Technology

4) Department of Radiological Technology, Osaka College of High Technology

* E-mail: nakanishi.yuta@ncvc.go.jp

Received January 21, 2025; accepted August 5, 2025

のある被験者にとって、音声による情報伝達は困難である。2022年に日本補聴器工業会が実施した全国調査 (JapanTrak 2022) によると、18歳以上の日本人のうち約11.6%が自身の聴力に問題があると回答している¹⁾。さらに聴覚障害者では触覚の感受性が高まる場合があり、突然の身体接触に対して不安や不快感、過敏な身体反応を示すことがあることが報告されている²⁾。また視覚的注意の強化により、視野外からの視覚刺激に対して敏感に反応する傾向も示されている³⁾。これらの特性は、放射線検査における誘導や指示伝達に影響を与える可能性がある。こうした背景を踏まえ、聴覚障害のある被験者に対する効果的なコミュニケーション手段の開発や評価が求められている。

そのような中、日本診療放射線技師会は、2024年に聴覚障害者のための放射線部門におけるガイドライン改訂版⁴⁾を発行した。そのガイドラインによると、聴覚障害者とのコミュニケーション手段として、筆談、文字や絵による伝達、ジェスチャーまたは体の一部に触れるといった方法、手話、読唇術、室内の電気のON・OFFといったライトによる伝達、電光文字表示器など

の特殊機器の使用、遠隔手話通訳といった8種類を挙げている⁴⁾。また浅野らによると、胃X線透視検査時における高度難聴の受診者には、検査前に体位変換を含めた検査風景と撮影方法についてまとめた動画を見てもらい、検査中は、あらかじめ用意しておいた動作や呼吸の指示を記載したシートを、撮影中に受診者へ見せながら検査を実施している⁵⁾。これらの手段は簡便で伝わりやすいが、電気が消灯したら吸気で止める指示が記載されたシートを見せるなど、あらかじめ伝達内容が決まっているものや、手話や読唇術といった技術習得が必要なものとなっている。また筆談は筆記する時間が必要であり、伝達の遅延が生じる。従って技術が必要な手話や読唇術以外は、自由度とリアルタイム性が相反する課題を抱えている。

現在、音声認識技術が急速に発展しており、音声を認識して文字化するソフトウェアが多く開発されている。河原らは、聴覚障害者を対象として、講演での音声をリアルタイムにテキスト化し、字幕付与する手段を開発している⁶⁾。また音声認識システムの中には無償で提供されているものも多く存在する。そこでわれわれは胃X線透視検査を想定して、音声認識入力システムを用いた簡便な指示伝達方法の有用性を検討することを目的として、本研究を実施した。

1. 方 法

1.1 使用機器

本研究では、Intel Core i7-11700 2.5GHz, Operating SystemにWindows 11、オーディオインターフェース (Scarlett 2i2) を搭載したパーソナルコンピューター (NEC, 東京, 以下, PC) を用いた。また音声認識入力による文字表示を可能にするため、市販の音声認識入力ソフトウェアを用いた。入力された文字の表示には、視認性を考慮して50インチの液晶モニター (SHARP, 大阪, 以下, モニター) を用いた。

1.2 音声認識入力ソフトウェアの選定

本研究で使用するソフトウェアの選定のために、音声入力の正確性および表示時間の計測による性能評価を行った。対象のソフトウェアは無償で公開されており、汎用性や利便性が高い、プレゼンテーションソフトウェア (Microsoft, PowerPoint, Windows 11, Microsoft 365版, オフライン, 以下, M社), ワープロソフトウェア (Google, Google ドキュメント, Windows 11, Chrome ブラウザー, オンラ

イン, 以下, G社), メモ作成ソフトウェア (Apple, iPhone メモ, iPhone 13, iOS 17.1, オフライン, 以下, A社), 文字起こしソフトウェア (WellSource, Speechnote, Windows 11, Chrome ブラウザー, オンライン, 以下, W社) の4種類とした。

性能評価では、「右を向いて」という指示を100回行い、正確に表示された回数と、指示出しから文字表示完了までの時間を計測した。「右を向いて」は、胃X線透視検査で頻繁に用いられる基本的な体動指示の一つであり、簡潔かつ誤認の少ない文言であることから、音声認識による表示の正確性を評価する基準として採用した。100回の音声入力は、各ソフトウェアの認識安定性や反応時間のバラツキを把握するのに十分な回数と判断した。なお音声入力には、A社を除く3ソフトウェアは、コンデンサーマイク (AT4040) を使用してPCへ、A社はスマートフォンに対して直接入力を行った。また結果に影響を与える声質の変化を防ぐため、全ての音声入力は同一の研究員が行った。

1.3 音声認識入力システムを用いた体動調査

1.2で選定したソフトウェアを用いて実際に音声入力をを行い、モニターに表示された文字に基づいて被験者が正確に体動できるかを調査した。体動の正確性は、①表示された指示と一致する動作が行われたか②誤動作や遅れがなかったかの2点で評価した。これらの評価は、検査室のカメラ映像を操作室のモニターで確認しながら実施した。評価者は本研究に関与していない2人とし、互いに独立して評価を行った。判定が一致しない場合には協議により評価を確定し、客観性と再現性の確保に努めた。

被験者は30代から60代の健常な一般成人11人とし、いずれも医療従事者ではなく、胃X線透視検査の受診経験があるボランティアであった。参加に当たっては、既往歴や身体的障害がないことを事前に確認した上で依頼した。

実験室はX線透視検査室とし、検査室側に被験者、操作室側に研究員を配置することで音声入力が被験者に聞こえないように施した (Fig.1)。また指示を与えるモニターは被験者の右前方に配置した。さらに表示された内容が正確であるかを研究員が確認できるように、被験者と同様の画面が表示できるPCを研究者の前面に配置した。

指示内容は、Fig.2に示す5種類である。体動の再現性が確認しやすく、音声認識による表示の評価に適したもの5つを選定した。指示を表示する際の文字の

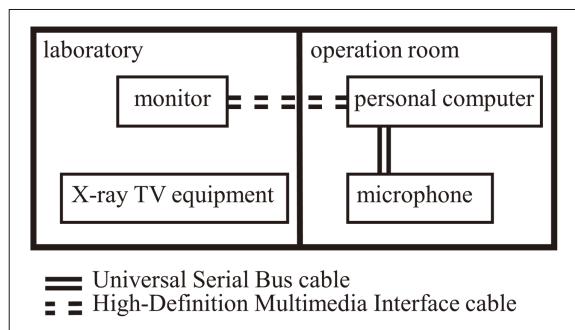


Fig.1 Location of each device

- 左手を挙げてキープ
- 左手を下げる
- 体ごと右を向いて
- 時計回りに1回転して
- 体ごと左を向いて

Fig.2 Instruction content

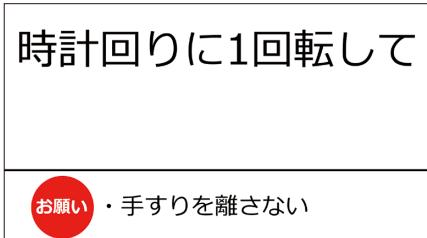


Fig.3 Monitor Screen Configuration

大きさは、文字数が最も多い「時計回りに1回転して」が1行で表示できる最大の大きさである60ポイントとした。また指示文を消去するタイミングは、被験者が指示を受けて体動を開始したことを確認したタイミングとした。文字の書体はメイリオを使用し、文字の色は黒色(HEX値:#000000)を選定した。またモニターに表示する画面構成は、指示が表示される上部エリアと、実験中の注意喚起が表示される下部エリアに分けた(Fig.3)。

1.4 被検者を対象とした体動実験の質問紙調査

音声認識入力システムを用いた体動実験の被験者に、体動実験でのモニターに表示された文字の分かりやすさを調査するために、被験者11人を対象とした無記名による自記式質問紙法を実施し、内容を分析した。質問項目と選択肢は、以下に示す6項目について

行い、個別の意見は自由記述とした。

- ・問1 文の長さは適切か (はい・いいえ)
- ・問2 文字の大きさ (小さい・ちょうどよい・大きい)
- ・問3 文字の表示スピード (遅い・ちょうどよい・早い)
- ・問4 文の消去のスピード (遅い・ちょうどよい・早い)
- ・問5 文字の書体は適切か (はい・いいえ)
- ・問6 文字の色は適切か (はい・いいえ)

1.5 倫理的配慮

被験者に対して、研究への参加および質問紙への回答は自由意志であり、研究協力拒否による不利益はないことを口頭で説明を行い、書面で同意を得た。また質問紙へは無記名とし、得られた内容は研究以外には使用しない他、個人が特定されないように配慮した。なお、本研究は日本放射線カウンセリング学会倫理審査委員会の承認を得て実施した(JSRC倫第2023-1号)。

06

2. 結 果

2.1 音声認識入力ソフトウェアの性能評価

音声認識入力ソフトウェアの正確性の結果をFig.4に、表示時間の計測結果をFig.5に示す。各社のソフトウェアの正確性は、A社が74%で最も低く、M社は95%、G社97%、W社92%であった。この結果から、A社以外の3社の正確性に大きな差は確認できなかった。また文字の表示時間はM社が1.164秒で最も速く、G社が2.528秒で最も遅いという結果であった。以上の結果から、高い正確性と表示時間の早さを考慮して、体動調査ではM社を採用した。

2.2 被験者による体動調査

体動実験での被験者の体動の正確性では、被験者全員がモニターに表示された指示に応じて、5種類全ての項目で正確に体動することが可能であった。

2.3 質問紙調査

体動実験でのモニターに表示された、文字の分かりやすさに対する被験者の質問紙調査の結果をFig.6に示す。文字色と書体では、1人が不適切と回答した。自由記述において、文字色については「黒背景に白文字が良い」、書体については「丸ゴシック M-PRO が好き」との回答を得た。表示スピードでは2人が速い

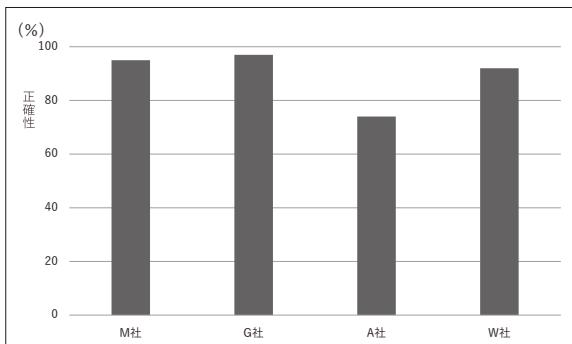


Fig.4 Accuracy of voice recognition input software

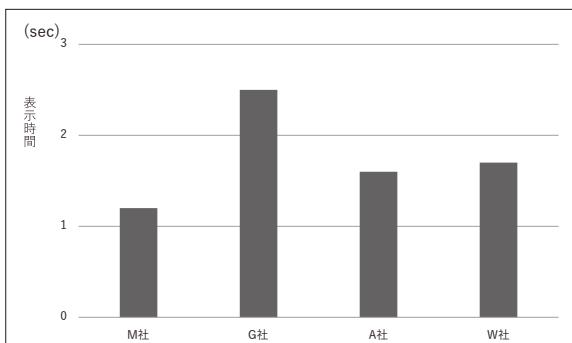


Fig.5 Display time for voice recognition input software

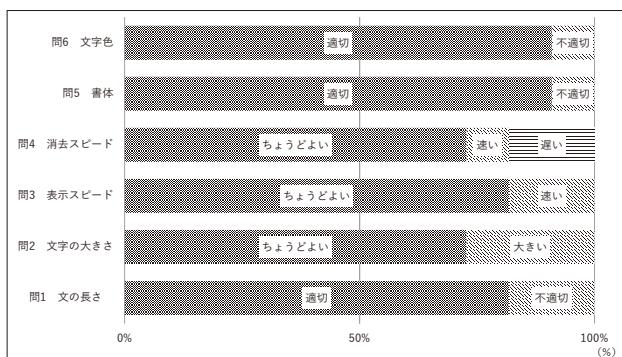


Fig.6 Questionnaire survey results

と回答した。文の長さでは、2人が不適切と回答した。この2人の自由記述では、「基準がないため」「長さがバラバラ」との回答を得た。消去スピードでは、1人が速い、2人が遅いと回答した。

3. 考 察

音声認識入力ソフトウェアの選定において、音声認識の正確性についてはA社のソフトウェアが74%と、他社のソフトウェアがいずれも90%を超えていたのに対して低い値となった。これは、A社以外は音声入力

にコンデンサーマイクを使用したが、A社はスマートフォンに直接入力をしているため、使用するデバイスの性能の影響のためであると考えられる。ただし、マイクの種類や使用機器の違いが表示速度や認識精度に与える影響については、本研究では詳細な検証を行っておらず、ソフトウェアの処理性能やネットワーク遅延などの要因も関与している可能性がある。従って今後はマイクの違いによる影響だけではなく、複数要因を統制した条件での比較検討が必要である。表示時間はM社のソフトウェアが1.164秒と最も速く、最も遅いG社のソフトウェアに対して2.17倍の差があった。これは、ソフトウェアの性能に依存するものであると考えられる。指示のリアルタイム性の観点からも、表示速度が優れていることが重要である。このことから、表示の正確性が高く表示速度の速いM社のソフトウェアが優れていると考える。

体動実験では、被検者11人全員が全ての項目で正確に体動することが可能であった。このことから、音声認識入力システムを用いてモニターに表示された文字の指示を、被験者は理解できたことが示唆される。従って体動の指示が必要な検査において、本システムを用いた指示により検査を実施することが可能であると考えられた。聴覚障害者に対する胃X線透視検査において、視覚的手段を用いたコミュニケーション支援の先行研究として、あらかじめ指示文を用意して提示する「指示シート型」の支援ツールの開発が報告されている⁷⁾。さらに平成12~13年度の科研費研究では、電光掲示板を用いた固定式の指示表示システムが検討されていた⁸⁾。これらの研究では、表示内容が事前に設定されたものであるため、即時的な柔軟性に課題があった。一方で、本研究では音声認識入力に、リアルタイムかつ個別対応が可能な指示を汎用デバイス上に視覚表示する点に特徴がある。つまり検査者の発話に応じてその場で内容を替えられるという点で、事前の表示設定を必要としない柔軟性と即時性を有しており、新たな実用性と応用可能性があると考える。なお本研究では、体動の正確性を評価するに当たり、事前に定義した基準に基づき、研究に関与しない複数の評価者が独立して判定を行った。しかしながら、この評価方法は観察による質的判断であるため、完全な客觀性を担保するものではない。今後は動作解析センサーなどを活用し、定量的かつ客觀的な評価手法の導入が望まれる。また今回の体動実験は健常者のボランティアを対象としており、全ての指示において正確な体動が可能であったが、体動に制限がある被検者に対して同様の

結果が得られるかについては、今後の検討課題である。質問紙調査において、文字の大きさや表示・消去のスピード、文の長さや書体、文字色に関して、「ちょうどよい」との回答が7割以上を占め、おおむね満足度の高い評価が得られた。一方で、一部の被検者からは、文字の大きさや表示・消去のスピード、書体、文字色に関して否定的な回答も見られた。これらの評価のバラツキは、表示形式に対する個人の感じ方の違いが関与している可能性があるが、本研究では11人という小規模なサンプルを対象としており、参加者の背景と評価傾向の関連性については十分に検討されていない。従って今後はより多様な属性の参加者を対象とし、評価項目の精緻化や自由記述を含む調査によって、表示設定の最適化に資する知見を得る必要があると考える。今回の体動実験では、健常者のボランティアに対して限られた指示について検討を行った。しかし、実際の検査では、検査の状況に応じたさまざまな内容の指示をする必要がある。また被験者によっては表示された指示に対して、その読み取りや理解および指示された体動に制限が生じている場合もある。これらのことから、体動の正確性を向上させるためには、状況や被験者に合わせて表示方法や内容を個別に調整するなど、最適な表示についてはさらに検討を行う必要があると考える。

結論

本研究では、音声認識入力システムを用いた検査指示伝達の有用性について、ソフトウェアの選定、体動実験、および質問紙調査を通して検討した。その結果、適切な表示形式を用いることで、被検者に対してリアルタイム性の高い指示を視覚的に伝達することが可能であったことから、聴覚障害者の胃X線透視検査において、音声認識入力システムを用いた簡便な指示伝達方法は有用であることが示唆された。一方で、使用ソフトウェアやデバイス環境による認識精度・表示速度の違いや、表示内容に対する評価には一部バラツキも見られた。質問紙調査では肯定的な評価が多かったものの、否定的な意見の背景までは明らかにできず、小規模なサンプルであることや選択式の評価形式による限界も存在する。さらに実際の検査現場では、体位の微調整、不正確な動作への再指示など、多様で複雑な指示が求められる場面も多く、1行表示に限定された本研究の設計では、そうしたニーズに対応できているとは言えない。今後は、より多様な表示形式や内容、

さらに複雑な指示伝達への拡張性について検討とともに、参加者属性の多様化や自由記述による評価を取り入れた多面的な検証が求められる。

利益相反

筆頭著者および共著者全員に開示すべき利益相反はない。

謝辞

なお、本論文の要旨は第33回大阪府診療放射線技師会学術大会（2023年）で発表し、最優秀研究発表賞を受賞した。

本研究に当たり、株式会社アイエスゲートの宮田充さまより、研究に関する貴重な資料を多数ご提供いただきました。心より感謝申し上げます。

06

図の説明

- Fig.1 各装置の設置位置
- Fig.2 指示内容
- Fig.3 モニター画面構成
- Fig.4 音声認識入力ソフトの正確性
- Fig.5 音声認識入力ソフトの表示時間
- Fig.6 質問紙調査の結果

参考文献

- 1) 日本補聴器工業会：JapanTrak 2022 調査報告書、https://hochouki.com/files/2023_JAPAN_Truk_2022_report.pdf (参照 2025年5月1日)
- 2) Heidenreich M, et al.: KCNQ4 K⁺ channels tune mechanoreceptors for normal touch sensation in mouse and man. Nat Neurosci, 15(1), 138-145, 2012.
- 3) Bavelier D, et al.: Do deaf individuals see better?. Trends Cogn Sci, 10(11), 512-518, 2006.
- 4) 公益社団法人日本診療放射線技師会 業務改善推進委員会 編：聴覚障害者のための放射線部門におけるガイドライン. 9-11, 公益社団法人日本診療放射線技師会, 2024.
- 5) 浅野みず江, 他：胃X線検診の撮影現場における検査困難症例への対応～放射線技師の立場から～. 日本消化器がん検診学会雑誌, 58(1), 35-41, 2020.
- 6) 河原達也, 他：聴覚障害者のための講演・講義の音声認識による字幕付与. 日本音響学会誌, 74(3), 156-162, 2018.
- 7) 宮田 充, 他：聴覚障害者や難聴の方々の胃部X線検査受診における胃部X線検査支援システムの有用性について. 日放技誌, 69(832), 134-141, 2022.
- 8) 西山勝夫, 他：聴覚障害者の胃レントゲン検査時のコミュニケーション支援システムの開発と評価. 平成12年度～13年度科学研究助成金（基盤研究(C)(1))研究成果報告書. 2002.