

令和 5 年 6 月 6 日現在

機関番号：12103

研究種目：基盤研究(C)（一般）

研究期間：2019～2022

課題番号：19K02996

研究課題名（和文）聴覚障害者のためのモバイル端末文字通訳システムに関する研究

研究課題名（英文）Research on Web-Based Manual Speech-to-Text Captioning using Mobile Devices for Deaf or Hard of Hearing

研究代表者

若月 大輔（Wakatsuki, Daisuke）

筑波技術大学・産業技術学部・教授

研究者番号：50361887

交付決定額（研究期間全体）：（直接経費） 2,600,000円

研究成果の概要（和文）：聴覚障害者の情報保障の1つであるPC文字通訳をインターネット上で行えるシステムを構築した。専門的なアプリケーションを導入することなく、利用者と文字通訳者がモバイル端末のウェブブラウザだけで、容易に文字通訳を生成、閲覧できるようになった。また、音声認識を活用して、通訳者の技能をサポートする機能を実現した。さらに、文字通訳に画像を挿入できるようにしたハイブリッド文字通訳を実現し、利用者の学習行動や学習効果を明らかにした。当初の予定に含まれていなかった成果として、専門家以外が文字通訳を担う方法や、オンデマンドの動画コンテンツに関する新たな文字通訳方法について提案し、有効性を検討した。

研究成果の学術的意義や社会的意義

本研究は、文字通訳の専門性や、人材確保とその養成の課題を解決するために、モバイル端末の特性を活かした次世代の文字通訳を提案するものである。提案法をcaptiOnline4として実装し、ウェブブラウザのみでの文字通訳を実現した。ネットワークや情報機器の専門知識が不要となり、誰でもオンライン文字通訳を容易に利用できるようになった。音声認識を活用した文字通訳者間の連絡や、聞き溜めのサポートも可能にした。これらの成果によって、文字通訳者に求められる技能の敷居を下げることができ、先に挙げた課題解決の一助になったと考える。captiOnline4は、令和4年度末で320を超える団体に利用するまでに至った。

研究成果の概要（英文）：I have developed a Web-based system for manual speech-to-text captioning, which is one of information support methods for deaf or hard of hearing (DHH). Without the need for specialized applications, users and captioners can easily view and generate text caption using only a web browser on a mobile terminal. In addition, by leveraging speech recognition, I have implemented features to support the skills of captioners. Furthermore, a hybrid caption that can insert images into the text caption was implemented, and the user's learning behavior and learning effects were clarified. As results that were not included in the original plan, we proposed a method for non-professional captioners and a new captioning method for on-demand video content, and examined its effectiveness.

研究分野：ヒューマンインタラクション，知能情報学，教育工学，学習支援システム，通信工学

キーワード：遠隔情報保障 聴覚障害 文字通訳 モバイル端末 ウェブアプリケーション 音声認識

## 1. 研究開始当初の背景

聞こえに障害がある聴覚障害者に対して、本来聴覚から得られる情報を、視覚などの他の感覚で代替して、情報獲得を保障しようという考え方を情報保障と呼ぶ。聴覚障害者の情報保障は音声を手話へ変換する手話通訳や、音声を文字に変換する文字通訳<sup>1</sup>といった方法が一般的に普及している。日本における大学などの高等教育の現場においては、手話において専門用語などの表現が困難なことから、主に文字通訳が利用されることが多い。大学における文字通訳はノートテイクと呼ばれる手書きによる方法や、PCを利用して専用ソフトで文を入力して提示するPCテイクと呼ばれる方法もある。また、学会などを聴講する場合は、発話された文を、複数人で担当を分けて入力する連係入力<sup>2</sup>と呼ばれる日本独自の方法で文字通訳が行われている。

研究開始当初の日本語のクラウド型の音声認識は、一般的な環境下で80%程度と言われており、高等教育や学術的会議の場面で修正することなくそのまま利用するのは現実的ではなかった。このため、日本における文字通訳は、専門技能をもった人が全文を打ち込んだり、誤認識を修正したりなど、人が介在しなければ成立しない。この状況は、音声認識技術の進歩により改善が見込めるが、同音異義語が多い日本語において、その見通しは明らかではない。一方、高等教育に耐えうる質の文字通訳を提供できる通訳者に要求される技能は多く、人材の確保やその養成も課題となっている。

1990年代から開発が進められてきたIPtalkなどに代表される専用ソフトは、PCを利用した文字通訳が前提とされている。一方、近年ではスマートフォンやタブレットなどのモバイル端末が広く普及し、PCを利用しなくても文字通訳ができる下地が整いつつある。大学生を対象にした文字入力方法についても、スマートフォンによるフリック入力の方が好まれ、習熟度が高いことが報告されている。つまり、フリック入力に熟練した人を文字通訳の入力者として取り込むことができれば、人材確保の課題の解決につなげることができる。しかし、現状ではモバイル端末での文字通訳に対応したシステムは皆無である。適した入力および連係インタフェースについて検討を行い、モバイル端末の特性を活かした新たな文字通訳手法を考案する必要がある。

## 2. 研究の目的

本研究では、現状のPCを利用した文字通訳の高い専門性に起因する課題と、人材確保と養成の課題を解決するために、近年広く普及し、多くの人々が扱いに慣れているモバイル端末の特性を活かした次世代の文字通訳手法として、モバイル端末文字通訳を提案する。さらに、提案手法を実装したシステムを構築し、普及と発展の礎を築くことを目指す。

具体的には、PCと文字通訳用ソフトを扱う知識や技能、話の聞き溜め・要約の技能、文を入力する技能をサポートする方法を検討する。また、従来の文字通訳の指示語などに対して、画像を併せて提示するハイブリッド文字通訳も検討する。

## 3. 研究の方法

本研究では、聴覚障害者のためのモバイル端末文字通訳システムを実現するために、次の4つのテーマに分けて研究を進めた。

### 【PCと文字通訳用ソフトを扱う知識や技能】

既存の専用ソフトはダウンロード、インストールおよびネットワーク設定など、文字通訳以外に要求される知識や技能が多く敷居が高い。そこで、モバイル端末からウェブブラウザでアクセスするだけで文字通訳が可能になる環境構築を目指す。

サーバをNode.jsを使用して構築し、WebSocketによるリアルタイム通信により、ウェブベース遠隔文字通訳システムを実装した。研究代表者は、既に同様の構成でcaptiOnline(キャプションライン)と呼ばれるシステムを構築していたが、本研究用に文字入力・修正機能、アカウント管理機能、QRコードによるアクセス機能など、より利便性を強化した研究用の新たなシステムを構築する。

### 【話の聞き溜め・要約の技能】

文字通訳者が音声をきき取って文字化する文字通訳は、通訳者の技能に頼るところが大きく、

<sup>1</sup> 本応募では音声の全文や要約を文字で表出する方法全般を文字通訳、文字通訳を担う者を(文字)通訳者と記述する。

<sup>2</sup> パートナーと分担して文を入力する方法である。例えば、文の前半を自分が入力し、後半をパートナーが入力するといった分担をする。単独入力と比較して1.5倍から2倍の速さで文を入力できるようになる。

数や長い固有名詞などの聞き逃しで誤った文を表出してしまうこともある。そこで、音声認識を併用して聞き落としした内容を参照できる通訳者支援インタフェースを提案し、その実用性を評価する。

方法としては、音声認識結果を文字通訳結果として直接出力するだけでなく、原稿やメモ等に一時的に出力し、通訳者がそれらを参照して文字入力に役立てたり、修正して出力したりできるインタフェースを構築する。

#### 【文を入力する技能】

PC文字通訳ではタッチタイピングが前提だが、【PCと文字通訳用ソフトを扱う知識や技能】で述べたように、タイピング以前にPCの扱いが敷居を高めている。本研究ではモバイル端末に特化した文字通訳方法について検討する。フリック入力や音声入力に適した入力インタフェースを考案する。この際、PC文字通訳と同様に他者との連係入力が可能であることに配慮する。また、連係入力では、通訳者同士が分担して入力した文を自然かつ連続的に接続していくための特殊技能が求められる。本研究では通訳者同士が、正確に文が接続されるよう配慮しなくても、自動で適切に文が接続されるスマート連係入力を提案し、その実現を目指す。

#### 【モバイル端末ならではの工夫】

現在の文字通訳では、「この式」や「あの図」のような指示語が何を指すのか伝わりにくいため、「画面右中央の式/図」といった書き換えを行っている。モバイル端末を利用できる場合、写真撮影で指示語の部分を直接画像で示すことが可能になる。文字と画像のハイブリッド文字通訳を提案し、その実現を目指す。

ハイブリッド文字通訳では、従来の文字通訳に画像を挿入することになるが、挿入しやすい方法や、適切な配置方法などを検討する。

### 4. 研究成果

ここでは、3方法で述べたそれぞれのテーマを中心に、それらの研究成果について報告する。2020年度から2023年度の研究期間の2年～4年目の間は、新型コロナウイルス感染症の影響を大きく受け、中には実験や検討が困難なテーマがあった。評価や検討が不十分なテーマについては、今後も継続して研究を進めていきたい。一方、本研究を通して得られた知見もあり、当初の研究目的には含まれていなかった成果についても併せて報告する。

#### 【PCと文字通訳用ソフトを扱う知識や技能】

以前から研究開発を重ねてきた、ウェブブラウザでアクセスするだけで、遠隔文字通訳を実現可能なシステムであるcaptiOnlineを更新し、captiOnline ver.4を構築しリリースした。同システムは、日本語の文字通訳に必要な不可欠な連係入力のための機能を始め、文字通訳に必要な機能はすべて採用した。また、音声認識を活用する機能を充実させた。音声認識結果を文字通訳として使用して、通訳者がそれを即時に修正できる機能を基本として、音声認識結果を通訳者間のチャットで利用したり、直接出力せずに原稿としてためておいたりする機能も実装した。さらに、アカウント管理機能やQRコード出力機能等を実装したことで、通訳者や利用者のアクセス管理や、モバイル端末でのアクセスが容易になった。

本研究で開発したcaptiOnlineは、既に多くの団体や企業に活用していただいている。特に、コロナ禍にともない、対面での文字通訳が困難になったことから、ビデオ会議アプリを併用して実施している団体が急増した。本研究開始前の2018年度末では約190だった利用団体・企業数が、2022年度末では約320となった。

#### 【話の聞き溜め・要約の技能】

先に述べたように、音声認識結果を直接出力せずに、原稿へ出力してためておける機能を実装した。これによって、話の聞き溜めが難しい経験の浅い文字通訳者が、それらを参照して入力に役立てたり、聞き落としした内容を参照したりできるようになった。また、音声認識結果を原稿に出力できるようになったことで、音声認識で逐次生成される原稿を裏で修正して、文字通訳結果として送出することが可能になった。これにより、音声認識の誤認識を含んだ結果が、利用者の目に触れる機会が減り、通訳者の肉体的、心理的負担の軽減につながることを期待できる。これらの機能は、captiOnline ver.4に実装済みであり好評を得ているが、評価実験を実施するまでには至らなかった。今後、有効性や実用性を明らかにしたい。

#### 【文を入力する技能】

当初の予定では、フリック入力などのモバイル端末の文字入力法を活かした、文字通訳の方法を提案する予定であったが、今回の研究期間では十分な検討を行なうことができなかった。一方、先に述べたように、音声認識による文字入力機能や修正機能を充実させた。これによって、通訳者からはモバイル端末による文字入力の最適化よりも、音声認識をうまく活用する方法についての声が多く寄せられるようになった。本テーマについても、十分な評価実験を実施するまでに至らなかった。音声認識結果をモバイル端末で修正することも有効だと考えられるので、引き続き

き検討を進めて行く予定である。

文字通訳は、訓練を積んだ専門家が担うことが一般的である。ここでは、専門家以外が文字通訳の入力を担うことができる方法の検討も行った。具体的には、タイムライン形式で聞こえたことを投稿して貰う方法と、ライブ動画をリアルタイムで分割して、クラウドソーシングで入力して文字化してもらう方法について検討した。これらについては、【当初の研究計画に含まれていなかった成果】で述べる。

#### 【モバイル端末ならではの工夫】

ハイブリッド文字通訳を実現するために、PC などのローカル環境に保存された画像や、インターネット上の画像などを、ドラッグアンドドロップで挿入できるインタフェースを構築した。原稿に対しても画像を挿入できるようにし、あらかじめ送出する文字通訳コンテンツを作り込めるようになり、一部の通訳者から好評を得ている。一方、文字通訳の文に対する画像の配置については、従分な検討を行うことができなかったため、今後も継続して検討を進めたい。

文字通訳の利用者が、ハイブリッド文字通訳を利用した際の効果や、学習行動について、大学の受業を想定した実験を行い、その有効性を明らかにした。当該実験を通して得られた研究成果は、国内外の学会で公表した。

#### 【当初の予定に含まれていなかった成果】

文字通訳の専門家ではない人が、文字通訳を担う方法として、タイムライン形式の方法と、クラウドソーシングを利用して文字化をする方法について検討した。タイムライン形式の方法については、イベント等でその場に居合わせた人たちに、聞こえることを入力してもらう形式で文字通訳を行なう方法である。聞こえること以外に、見える、感じる、専門的な情報も投稿しあうことで、障害の有無に関係なく情報を相補的に共有できる新たな情報保障のかたちを模索した。この方法は、ISee TimeLine として実装して、実用化を目指して様々な場面で実験を重ね、国内外の学会等で成果を公表した。

また、クラウドソーシングによる文字化の方法は、ライブ動画を適切な長さに分割して文字化を担当する人(ワーカ)に配り、ワーカが文字化したものを適切に連結して、文字通訳の文を生成する方法である。ワーカの文字化する能力にみあった長さの動画を配ることで、専門家ではない複数のワーカで文字化を担うことができる方法を提案し、システムを試作した。同システムで実験を行い、国内外の学会等で成果の公表を行った。

本研究では、ライブの文字通訳を対象に、リアルタイムで文字通訳を行なうことを中心に検討を行ってきたが、オンデマンドの映像コンテンツに対する文字通訳やその提示方法についても検討した。字幕を主体として映像コンテンツを閲覧できる方法を新たに提案した。captionNovel として実装したシステムで実験を行い、国内で研究成果の公表を行った。なお、本提案は、聴覚障害者の映像視聴環境の改善に大きく寄与できる見込みがあるため、継続して研究を進める計画である。

## 5. 主な発表論文等

〔雑誌論文〕 計9件（うち査読付論文 6件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 新井達也, 若月大輔, 塩野目剛亮	4. 巻 121(142)
2. 論文標題 聴覚障害学生向けノベルゲーム型オンデマンド教材に関する基礎的検討	5. 発行年 2021年
3. 雑誌名 電子情報通信学会技術研究報告	6. 最初と最後の頁 11-16
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 新井達也, 若月大輔, 塩野目剛亮	4. 巻 63(2)
2. 論文標題 字幕付き数学講義場面における聴覚障害学生の受講行動 図表や数式を挿入した字幕に対する視線行動に着目して	5. 発行年 2021年
3. 雑誌名 ろう教育科学	6. 最初と最後の頁 51-62
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Wakatsuki Daisuke, Arai Tatsuya, Shionome Takeaki, Tsukuba University of Technology 4-3-15 Amakubo, Tsukuba, Ibaraki 305-8520, Japan, Teikyo University 1-1 Toyosatodai, Utsunomiya, Tochigi 320-8551, Japan	4. 巻 25
2. 論文標題 Hybrid Caption Including Formula or Figure for Deaf and Hard-of-Hearing Students	5. 発行年 2021年
3. 雑誌名 Journal of Advanced Computational Intelligence and Intelligent Informatics	6. 最初と最後の頁 187 ~ 194
掲載論文のDOI (デジタルオブジェクト識別子) 10.20965/jaciii.2021.p0187	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 張健偉, 白石優旗, 若月大輔, 塩野目剛亮, 熊井克仁, 森嶋厚行	4. 巻 J104-A
2. 論文標題 クラウドソーシングに基づく聴覚障害者によるリアルタイム文字情報保障システムの開発	5. 発行年 2021年
3. 雑誌名 電子情報通信学会論文誌A	6. 最初と最後の頁 64 ~ 74
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Tanaka Kohei, Wakatsuki Daisuke, Minagawa Hiroki	4. 巻 2
2. 論文標題 A Study Examining a Real-Time Sign Language-to-Text Interpretation System Using Crowdsourcing	5. 発行年 2020年
3. 雑誌名 Computers Helping People with Special Needs(ICCHP 2020)	6. 最初と最後の頁 186 ~ 194
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-58805-2_22	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Wakatsuki Daisuke, Arai Tatsuya, Shionome Takeaki	4. 巻 2
2. 論文標題 Analysis of the Gaze Behavior of Deaf and Hard-of-Hearing Students During a Captioned Lecture	5. 発行年 2020年
3. 雑誌名 Computers Helping People with Special Needs(ICCHP 2020)	6. 最初と最後の頁 202 ~ 208
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-58805-2_24	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Rumi Hiraga, Daisuke Wakatsuki, Yuhki Shiraishi, Masayuki Inoue, Yuka Kogo, Yoshiki Fukunaga, Makoto Kobayashi, Manabi Miyagi, Takeaki Shionome, Jianwei Zhang and Atsuyuki Morishima	4. 巻 -
2. 論文標題 First Evaluation of Information Support of everyone by everyone for everyone TimeLine (ISeeTL) applied to Deaf and Hard of Hearing People Watching Sport	5. 発行年 2019年
3. 雑誌名 AAATE2019: The 15 th International Conference of the Association for the Advancement of Assistive Technology in Europe	6. 最初と最後の頁 S135-S136
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 田中康平, 若月大輔, 皆川洋喜	4. 巻 119(322)
2. 論文標題 クラウドソーシングを用いた手話文字通訳 ~ ライブ動画配信を活用したシステムの試作 ~	5. 発行年 2019年
3. 雑誌名 電子情報通信学会技術研究報告	6. 最初と最後の頁 95-100
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 新井達也, 若月大輔, 塩野目剛亮	4. 巻 119(165)
2. 論文標題 字幕付き講義場面における聴覚障害学生の視線行動の分析	5. 発行年 2019年
3. 雑誌名 電子情報通信学会技術研究報告	6. 最初と最後の頁 35-39
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

ウェブベース遠隔文字通訳システムcaptionline <a href="https://captionline.org">https://captionline.org</a>
--

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------