

遠隔地リアルタイム字幕支援のためのキーワード入力ソフトウェアの研究開発

† 筑波技術大学産業技術学部産業情報学科

‡ 筑波技術大学障害者高等教育研究支援センター

若月大輔 †, 加藤伸子 †, 河野純大 †, 塩野目剛亮 †, 村上裕史 †, 皆川洋喜 †, 西岡知之 †,
三好茂樹 ‡, 黒木速人 ‡, 内藤一郎 †

要旨 : 聴覚障害者の社会参加を支援するための遠隔コミュニケーションに関する研究開発を進めている。本研究では、専門的な内容に対する遠隔地リアルタイム字幕を支援するために、字幕入力者向けに提示するキーワードの入力方法について検討を行った。これまでのキーボードからの直接入力だけでなく、入力履歴や、講義資料を活用してキーワード入力を支援するソフトウェアを試作した。試作ソフトウェアを用いて大学の講義を対象に主観評価実験を行い、各入力方法の特徴と有効性について調査した。

キーワード : 遠隔地リアルタイム字幕, 字幕入力者支援, 遠隔情報保障, 聴覚障害者支援

1 はじめに

社会参加する聴覚障害者を支援するために、図1に示すような遠隔情報保障システムを構築し、遠隔から手話通訳やリアルタイム字幕を提供する遠隔コミュニケーション技術に関する研究や実験的支援を進めてきた [1, 2, 3, 4, 5]

遠隔地リアルタイム字幕では、会場の講師の発話音声とカメラで撮影した映像をネットワークを介して遠隔のスタジオへ送る。スタジオでは送られてきた音声とカメラの映像をもとに、字幕入力者が字幕を作成し、リアルタイムに会場へ送り返して字幕が必要な受講者（被情報保障者）へ提示している。しかし、大学の講義などの専門性の高い内容が対象の場合、字幕入力者が専門用語や専門的内容への対応ができず字幕が不正確になったり途切れたりすることがある。そこで、専門知識を持った健聴の支援者を会場に配置し、講師が発した専門用語や内容の解説等を入力して遠隔のスタジオに送っている。また、専門用語だけでなく聞き取りにくい単語、講師の状況、および「これ」、「それ」などの指示語の対応等の情報も遠隔の字幕入力者に伝える場合もある。このように、遠隔からの字幕による通訳をより正確に行うことができるように、専門知識を持った支援者（キーワード入力者）が入力したキーワードによる字幕入力者の支援に関する研究を進めている。

字幕入力者向けのキーワードを入力する方法として、キーワード入力者がキーボードで直接入力する方法を試行してきた。キーワードの導入によって専門的な内容についても、より正確かつ安定な遠隔地リアルタイム字幕

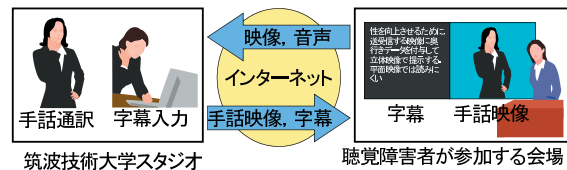


図1 遠隔情報保障システム

が可能となった。しかし、一般にキーワード入力者のキーボードを打つ速度には個人差があり、キーワードが字幕入力者に提示されるまでの時間のばらつきがある。また、タイピング速度だけでなくミスタイプや漢字の誤変換などによって、間違ったキーワードを提示したり、著しく遅れたりすることもある。このため、字幕入力者が求めるキーワードが間に合わない、必要なキーワードが表示されないなどの問題が発生している。より正確で安定な字幕提供するためには、多様な分野のキーワード入力者がキーワードを入力しやすい環境、およびインターフェースについて検討する必要がある。

本研究では、より正確で安定な字幕を被情報保障者に提供するために、字幕入力者向けのキーワード入力を支援する方法について検討を行う。キーボードからの直接入力だけでなく、入力履歴から参照したり、講義資料からキーワードを抜き出したりすることができるソフトウェアを試作した。試作ソフトウェアを用いて大学の講義を対象にキーワード入力を行い、よりキーワードを入力しやすい方法について入力者に対する主観評価実験を実施した。

遠隔地リアルタイム字幕システムの構成を図2に示す。サーバはウェブサーバとフラッシュメディアサーバ(adobe社)で構成され、ウェブ上で映像の送受信、字幕データの送受信、ならびにキーワードの送受信が可能なソフトウェアが動作している。これらのデータは一般的なウェブブラウザ、および専用ソフトウェアでの閲覧、参照が可能である。

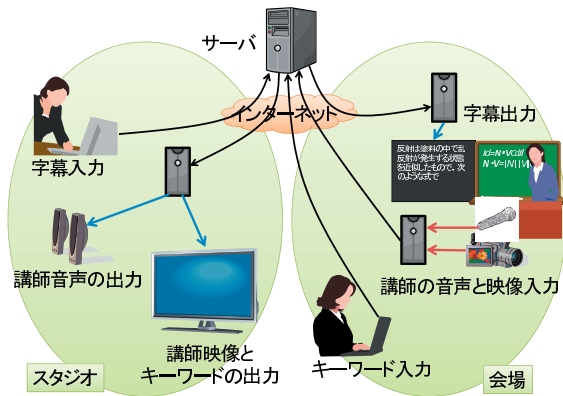


図2 遠隔地リアルタイム字幕システムの構成

これまで我々が研究開発を行ってきた遠隔地リアルタイム字幕システムでは、会場側の講師の発話や様子をマイクとビデオカメラでPCに取り込み、インターネットを介してサーバへ送信する。スタジオ側ではサーバにアクセスして会場から送信されてくる講師の音声聞き、映像を参照しながらリアルタイムに字幕を作成してサーバに送り返す。会場側ではサーバ上の字幕データを逐次受信し、会場にいる被情報保障者へ字幕を提示する。このように、インターネット接続環境をもつ会場やスタジオであれば原理的にどこでも支援を受けたり提供したりすることが可能である。

大学の講義などの専門性が高い内容の字幕を提供するためには、字幕入力者に専門的知識が要求される。字幕の正確さを向上させるために事前に講義資料等を提供して予習してもらおう場合もあるが、資料に掲載されていない内容や用語については対応が難しい。そこで、専門用語や解説、ならびに教員の様子や指示語の対応などの遠隔から字幕入力するために必要な情報をキーワードとして字幕入力者に提示する。会場側に専門知識を持った支援者が入力したキーワードをサーバへ送信し、スタジオ側の字幕入力者がそのキーワードを音声や映像とともに参照できるシステムを構築した。

3.1 従来のキーワード入力

第2節で紹介した遠隔地リアルタイム字幕システムにおいて、キーワード入力者はキーボードから直接入力のみでキーワード入力を行う。しかし、タイピング速度の個人差、ミスタイプや漢字の誤変換によって必要なキーワードの提示が遅れたり、表示されなかったりする場合がある。入力者の訓練や変換辞書の調整も考えられるが、多様な分野に対応するのは難しい。

また、キーワードを参照する字幕入力者から、専門用語がたびたび繰り返される場合はそのつどキーワードとして提示してほしい、講義資料などに書かれている単語や解説等についても提示してほしいという要望がある。繰り返し同じキーワードを入力したり、講義資料の内容を抜粋して入力する必要がある場合、入力するキーワードが増加するためミスタイプや誤変換のリスクが高まる。そこで本研究では、入力履歴や講義資料を参照してキーワードとして再利用できるソフトウェアを試作した。

3.2 入力履歴と講義資料を参照できるキーワード入力法

本研究で試作した入力履歴と講義資料を参照することができるキーワード入力ソフトウェアを図3, 4に示す。キーワード入力者が操作する入力側(図3)と、字幕入力者がキーワードを参照する提示側(図4)に分かれており、入力部で入力されたキーワードがネットワークを介して即時に提示側に表示される。

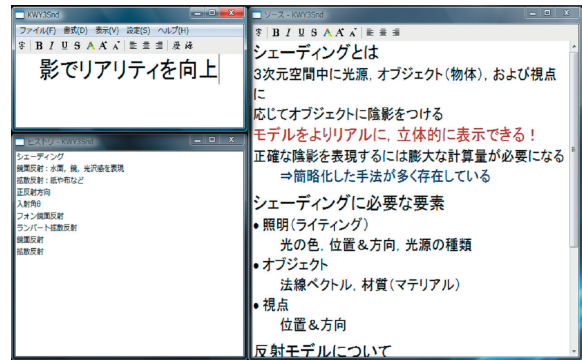


図3 キーワード入力ソフトウェア(入力側)

3.2.1 キーワード入力側

キーワード入力側は直接入力、入力履歴、資料の3つのウィンドウで構成される。直接入力ウィンドウ(図3左上)では、キーワードを入力しエンターキーを押下すると即時に提示側へ送られる。単に文字データを送るだけでなく、フォント、スタイル(強調、斜体、アンダーライン、打ち消し線)、サイズ、カラー、およびアライメント(左寄せ、中央寄せ、右寄せ)を設定したキーワー

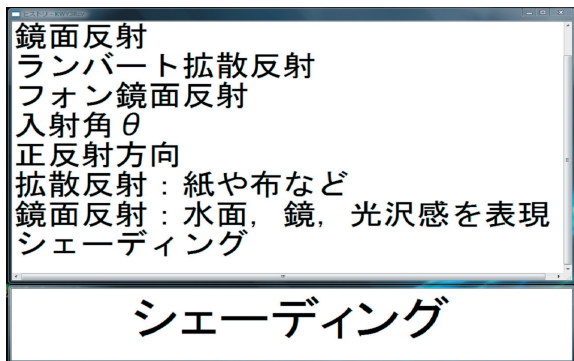


図4 キーワード入力ソフトウェア（提示側）

ども送ることができる。入力履歴ウインドウ（図3 左下）では、送信したキーワードの履歴がリストとして表示される。リストのキーワードを選択してエンターキーを押下すると即時にそのキーワードが改めて提示側へ送られる。マウスのダブルクリックでも同様に動作する。資料ウインドウ（図3右）では、講義資料などを読み込み表示する。資料上の単語や文字をドラッグして選択してエンターキーを押下することで即時に選択部分をキーワードとして提示側へ送ることができる。なお、対応するファイル形式はリッチテキストフォーマットの文書ファイル(*.rtf) かテキストファイル(*.txt) であるため、資料等はいずれかに変換して読み込ませる。

このように、必要に応じて入力履歴や資料に容易にアクセスしてキーワードとして再利用することができる。なお、3つのウインドウはそれぞれ独立であり、キーワード入力者が利用しやすいように自由にレイアウトを変更したり、表示／非表示を切り替えることが可能である。

3.2.2 キーワード提示側

キーワード提示側はカレントキーワード、表示履歴の2つのウインドウで構成される。カレントキーワードウインドウ(図4下)では、入力側から送られてきたキーワードを即時に表示する。ウインドウの最大内接サイズまでフォントが自動的にリサイズされて、ウインドウいっぱいにキーワードが表示される。表示履歴ウインドウ(図4上)では、カレントキーワードウインドウに表示されたキーワードの履歴がリストとして表示される。フォントの種類やサイズも設定することができる。なお、それぞれのウインドウは独立しているため自由にレイアウトを変更することが可能である。

4 実験・結果

4.1 実験環境

試作したキーワード入力用ソフトウェアの特徴を確認するために、本ソフトウェア適用した遠隔地リアルタイ

ム字幕システムを疑似的に構築してキーワード入力者と字幕入力者に対する主観評価を実施した。構築した実験環境を図5に示す。今回はあらかじめ撮影しておいた講義のビデオ映像を対象にした。講義内容は静岡福祉大学で開講された法学のうちの14回分の講義である。講師と講義資料が合成されたビデオ映像に対してキーワード入力を行い、講師、講義資料、およびキーワードを合成した映像を生成してサーバへ送る。スタジオ側ではその合成された映像をもとに字幕を作成しリアルタイムにサーバへ送り返す。会場側でその字幕を確認する。

キーワード入力者は、以前に同じ授業でキーワード入力を実施した経験がある健聴者6名である。字幕入力者は、普段から放送用の字幕入力を行っているプロの入力者に依頼した。

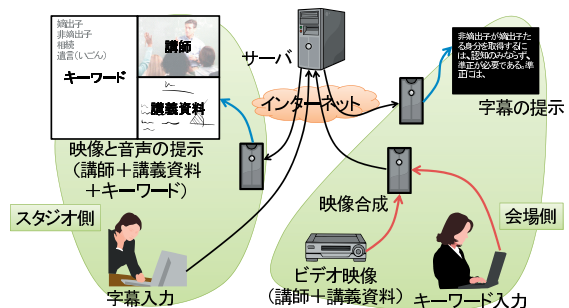


図5 実験環境

キーワード入力方法として、次のA~Dの4つの入力条件について調査を行った。

- A: 直接入力のみ
- B: 直接入力+入力履歴
- C: 直接入力+資料
- D: 直接入力+入力履歴+資料

1回の授業で4人のキーワード入力者がキーワードを入力する。90分の講義を時間で区間I~IVの4つに区切って、それぞれの区間で4人のキーワードの条件が重複しないように担当させた。その例を図6に示す。

90分の講義				
	区間I	区間II	区間III	区間IV
被験者1	A	B	C	D
被験者2	B	A	D	C
被験者3	C	D	A	B
被験者4	D	C	B	A

図6 1回の講義に対する各被験者のキーワード入力条件の変更例

字幕入力者には講義の前半と後半で異なる条件で入力されたキーワードを提示した。前半と後半で同一のキーワード入力者が担当するようにした。つまり、図6の例では、前半は条件Aで被験者1の区間Iと被験者2の区間IIの入力、後半は条件Cで被験者1の区間IIIと被験者2の区間IVを字幕入力者へ提示した。14回すべての授業において上記のように入力者や条件の偏りがないうようにキーワード入力者のシフト組んで実験を行った。

4.2 キーワード入力者に関する評価

各講義の終了後にキーワード入力者に回答してもらった主観評価結果について述べる。第4.1節で設定した入力条件A~Dについて入力しやすさについて7段階(1:「かなり入力しにくい」~4:「どちらともいえない」~7:「かなり入力しやすい」)で評価してもらった結果を図7に示す。なお、図中の棒グラフは平均値、エラーバーは標準偏差を表す。

条件D(直接入力+入力履歴+資料)によるキーワードがキーワード入力者にとって入力しやすい結果となった。条件B(直接入力+入力履歴)と条件C(直接入力+資料)を比較すると、条件Cの方が1ポイント以上平均値が高い。また、図8に示した直接入力、入力履歴、資料ウィンドウそれぞれについての入力しやすさについて回答してもらった結果でも、資料によるキーワード入力は、他の2つの入力方法に比べて1ポイント以上高く、資料から抜粋してキーワードを入力する方法が最も入力しやすい方法であると考えられる。

条件A(直接入力)は標準偏差が約2ポイントと大きく入力者によってばらつきが大きい。図8の結果からも直接入力は標準偏差が2ポイント程度と大きい。これは入力者のキーボードのタイピング速度の個人差が影響しており、タイピング速度の速い入力者は評価が高く、遅い入力者は評価が低かったためだと考えられる。条件CやDにおいては標準偏差が1ポイント程度であり、ほとんどの入力者が入力しやすいと回答している。

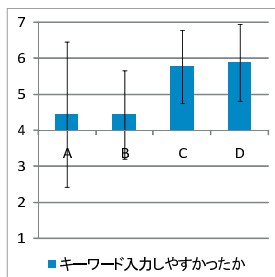


図7 入力条件A~Dに対する入力しやすさ

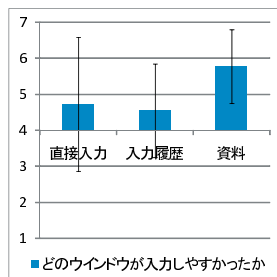


図8 各ウィンドウに対する入力しやすさ

次に、キーワード入力中に返信されてきた字幕を確認

できたか、講義内容を把握できたかについて7段階(1:「まったくできない」~4:「どちらともいえない」~7:「かなりできた」)で評価してもらった結果について図9、図10に示す。

ここでも条件CとDが条件AやBと比べて0.5ポイント程度高く、字幕入力者からの字幕の遅れ具合や用語等の正確さを確認したり、講義内容を把握しやすいことが示唆された。これは、資料ウィンドウの単語や解説の選択領域をキーワードとして送信できる手軽さから入力に余裕が生まれたためだと考えられる。また、必要なキーワードを拾うために講義の進行に合わせて資料ウィンドウを常に参照しているのも理由の一つとなっている可能性が高い。

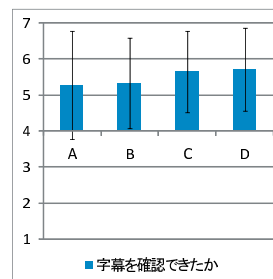


図9 入力中に字幕を確認できたか

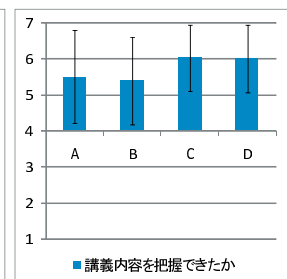


図10 入力中に講義内容を把握できたか

4.3 字幕入力者に関する評価

前節の実験で入力されたキーワードを提示して字幕をリアルタイムに作成してもらった。その字幕入力者に回答してもらった主観評価結果について述べる。第4.1節で設定した条件A~Dについて字幕入力の参考になったかについて7段階(1:「まったく参考にならない」~4:「どちらともいえない」~7:「かなり参考になる」)で評価してもらった結果を図11に示す。なお、図中の棒グラフは平均値、エラーバーは標準偏差を表す。

平均値をみるとすべての条件で、4の「どちらともいえない」よりも0.5~1ポイント程度高く、キーワードは字幕入力の役に立っていることがわかる。条件CとDの標準偏差は約1.5ポイントであり条件AやBと比較しても高めである。また条件Dの平均値が約0.5ポイントでほかの条件と比べて若干低い。これは、資料ウィンドウからのキーワード入力が原因となっていると考えられる。キーワード入力者の中には長い文章をそのままキーワードとして送る入力者も見受けられた。このため、これまで提示してきた専門用語や短めの解説等を中心としたキーワードと異なって見えたのが1つの原因であると考えられる。

キーワードが表示される速度について評価してもらっ

た結果について図 12 に示す。平均値をみるとすべての条件で、「どちらともいえない」という結果となっており、どのキーワード入力方法を利用してもキーワードの遅れについてはほとんど差がみられないことがわかった。

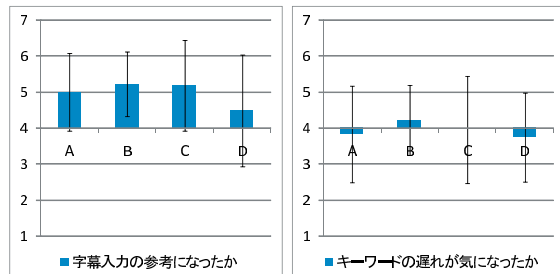


図 11 字幕入力の参考になったか

全体を通してみると、どの入力方法を選択した場合でも字幕入力者にとってはほとんど差がみられないことがわかった。一方、キーワード入力者にとっては、条件 C と D の結果から資料ウインドウを利用した場合に全体的に評価が高く、字幕の確認や授業の進行を把握できる余裕が出ることがわかった。一方、条件 B の結果から直接入力に入力履歴ウインドウを追加しただけでは入力しやすさにはつながらないことが示唆された。これは、入力キーワード数が多いため過去に入力した内容を探すのに時間がかかることが原因になっており、履歴の格納方法や参照方法を改良すれば評価が上がる可能性が高い。

5 まとめ

本研究では、大学の講義などの専門分野においても、より正確で安定な遠隔地リアルタイム字幕を被情報保障者に提供するために、字幕入力者向けのキーワード入力を支援する入力方法についての検討を行った。これまでのキーボードからの直接入力だけでなく、入力履歴や講義資料を参照してキーワードの入力を支援するソフトウェアを試作した。

試作ソフトウェアの有効性を確認するために大学の講義を対象にキーワード入力実験を実施した。直接入力、入力履歴、および講義資料を組み合わせたキーワード入力方法について主観評価を行った。その結果、キーワードは字幕入力者にとって役立っており、いずれの入力方法でもほとんど差がみられないことがわかった。また、キーワード入力者にとっては講義資料から用語や解説を拾ってキーワードとして入力する方法の評価が高く、字幕の確認や授業の進行状況の把握に有利であることが示唆された。

今後は、入力履歴の参照方法の改善、ならびに新たな入力方法の考案を行う。また、講師の発話に対する各入力方法のキーワード提示のタイミングと、遠隔地リアルタイム字幕への効果に関する調査を進めていきたい。

謝辞

実験の実施にあたり多大なご協力とアドバイスをいただいた字幕入力者の皆様、講義をご担当された宮下修一先生、ならびにスタッフの皆様へ感謝いたします。

参考文献

- [1] 加藤, 他 10 名: 講義資料とキーワードを画面合成した遠隔手話通訳システムー通訳スタジオにおける検討ー, ヒューマンインタフェース学会研究報告集 Vol.9, No.1, pp.23-28, 2007.
- [2] 河野, 他 10 名: 講義資料とキーワードを画面合成した遠隔手話通訳システムにおける聴覚障害学生への提示方法, ヒューマンインタフェース学会研究報告集, Vol.9, No.1, pp.29-32, 2007.
- [3] 加藤, 他 9 名: 聴覚障害者の情報保障におけるパソコン要約筆記入力者に対するキーワード提示, ヒューマンインタフェース学会論文誌, Vol.9, No.2, pp.125-134, 2007.
- [4] 西岡, 他 9 名: 遠隔地リアルタイム字幕提示システムにおける字幕作成者に対するキーワード提示について, WIT2005-88, pp.81-86, 2006.
- [5] 村上, 他 8 名: 就職面接場面での遠隔情報保障に関する一考察, WIT2006-57, pp.41-46, 2007

Keyword Input System to Support Captionists of Real-Time Remote Stenocaptioning System

Daisuke WAKATSUKI¹⁾, Nobuko KATO,¹⁾ Sumihiro KAWANO,¹⁾ Takeaki SHIONOME,¹⁾
Hiroshi MURAKAM,¹⁾ Hiroki MINAGAWA,¹⁾ Tomoyuki NISHIOKA,¹⁾ Shigeki MIYOSHI,²⁾
Hayato KUROKI,²⁾ and Ichiro NAITO¹⁾

¹⁾ Department of Industrial Information, Faculty of Industrial Technology,
Tsukuba University of Technology

²⁾ Research and Support Center on Higher Education for the Hearing and Visually Impaired,
Tsukuba University of Technology

Abstract: We have been developing a remote communication support system to assist the social participation of hearing impaired people. In this paper, we described a keyword input system to support captionists using a real-time remote stenocaptioning system for higher or professional education. The keywords have been input directly via the keyboard in our previous work. We made a prototype keyword input system that can pick up key words from input history or lecture documents. We performed experiments to highlight the features of each input method in order to support the captionists.

Key words: Real-time remote stenocaptioning system, support for captionist, remote communication support, hearing impaired