

令和 6 年 6 月 6 日現在

機関番号：10102

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K00873

研究課題名（和文）マルチモーダル英語授業コーパスのシステム構築と管理・運用と教師教育への応用

研究課題名（英文）Designing and Managing a Multi-modal Corpus of English Classes to Empower English Teachers

研究代表者

片桐 徳昭 (Katagiri, Noriaki)

北海道教育大学・教育学部・教授

研究者番号：60734829

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究は、新型コロナウイルスのパンデミックにより、システム開発に方針を転換しました。日本語と英語が混在する英語授業を録画し、書き起こしを支援するソフトウェアを開発・公開しました。最終年度にその成果を学会で報告しました。

開発したソフトウェアは、英語授業を録画したものを、インターネット上でAIを使った自動書き起こしを行った後の2種類の言語の書き起こしファイルを1つにまとめる機能があり、複数言語のトランスクリプションの編集作業を可能にします。これにより、50分の授業の書き起こし時間を手動の5～6時間から約2.5時間に短縮できることが確認されました。以上が本研究の成果です。

研究成果の学術的意義や社会的意義

本研究の新規性はMicrosoft Stream上で日本語と英語を別々に書き起こしたトランスクリプションがVideo Text Track (VTT)形式で活用でき、この2種類のVTTファイルを1本にまとめるソフトウェア(VTT2)を開発したことにある。VTT2の活用により、授業のコーパス資料(発話の書き起こしデータ)提供がこれまでより早く提供できることとなること、授業ビデオが映像と音声情報を持つことから、今後マルチモーダルコーパスを開発していく上で、重要なモジュールとして使用される方向性を示すと言う点で学術的意義が存在し、授業データが英語教育の発展に貢献しうることによって社会的意義が存在する。

研究成果の概要（英文）：The developed software records English lessons and uses AI on the internet for automatic transcription. It combines transcription files of two languages into one, enabling the editing of multilingual transcriptions. This development reduced the transcription time for a 50-minute lesson from 5-6 hours manually to approximately 2.5 hours. These are the outcomes of this research.

研究分野：コーパス言語学

キーワード：授業コーパス 自動書き起こし マルチモーダル 英語教育 教師教育

1. 研究開始当初の背景

英語授業データを蓄積するには人手を使用して書き起こす手法が研究開発当初の主流であった。しかし、教室現場で収集する音声には雑音が混在し、書き起こしにどうしても時間がかかってしまう。そのため、この困難さに起因した英語授業データの少なさどうしても課題となり、この少なさを克服し、蓄積した授業データを活かすことが望まれていた。授業データはビデオによる映像と音声を収集することから、これらの情報を統合した英語授業データのマルチモーダル化(発話テキストデータ+授業メタデータ+音声・映像データを統合)してその成果を教師教育(教育実習生・現職教員)や研究者に資することが背景として存在した。

2. 研究の目的

授業のタスクなどの質的な情報の追加とコーパスデータの汎用性の拡張(マルチモーダル化)により、授業データの分析とコミュニケーションな英語授業の応用へ研究がさらに進むことが予想されている中で、マルチモーダル化された英語授業コーパスのデータを公開して有効化するためにも、その蓄積・管理の方法や、データを活用するインターフェイスの深化が必要であった。また、授業データの分析枠は学校種・授業者・タスクなど様々な変数が存在し、同時に学習者に対する効果の評価など授業の質的分析には未解決な分野が多い。そのための分析に必要な研究者自身の手での作業によるタグ付与は多大な時間を要し、ゆらぎの可能性も存在する。キーワードなどで機械的に付与できるタグは、自動化することが可能であり、研究者が自由に利用できるツールを開発する必要性も感じられる。本研究はこのように複雑な授業メタデータを自動タグ付けにより詳細に記述してコーパス化し、加えて、音声・発話・映像情報を有機的に統合形でマルチモーダルコーパスとして成立させ、教師教育や研究者へのデータ提供として公開することを目的とした。

3. 研究の方法

まず第一に明瞭な音声を収集するためのマイクシステムの構築を検討し、プロフェッショナル様の音声収集機器を用いた実験を行った。その後、その機器を用いて、授業データを収集し、書き起こしを進める予定とした。しかし、新型コロナウイルスのパンデミック化の影響で、学校現場を訪問することが不可能となり、加えて授業がほぼ全て遠隔化されたため、教室でのデータ収集が不可能となった。よって研究の方針の転換を余儀なくされ、新型コロナウイルスの沈静化を待つ間、研究の方針をシステム開発に転身することとした。

研究開始から数年で人工知能を利用した自動音声認識技術が格段に進み、インターネット上で自動書き起こしができ、誤認識率も5%程度と低くなったことから、既存の授業データを利用して自動音声認識による書き起こしでコーパス作成をすることとした。

日本語と英語という複数言語が混在する英語授業の録画において、研究遂行段階では、両言語を同時に書き起こすことは不可能であった。そこで、日本語と英語を別々に書き起こしたトランスクリプションを用意して、この2種類のファイルを1本にまとめて編集やファイル出力が可能となるソフトウェアの開発を行なうこととした。

完成したソフトウェア(VTT2)を持ちい新規に授業の書き起こしの実験をして、人間が全て書き起こしを行った場合にかかるであろう時間と、VTT2を用いて書き起こしにかかる時間の比較を行い、研究結果の検証を行うこととした。

4. 研究成果

検証を行った結果、Table 1(次頁)に示すように、1回50分の授業の書き起こし作業が人手だけによる場合(5-6時間程度)と比較して、短い場合で2.5時間程度に短縮できることとなった。

本研究の成果により、英語や日本語などの複数言語が混在する英語の授業のコーパスをより早く作成することが可能となる。同時に、コーパスの作成過程で使用する録画した授業の資料は発話テキストデータ、音声・映像データなどのデータを含むことから、これらに授業メタデータを加えたマルチモーダルコーパスを開発していく上で、重要なモジュールとして使用される方向性を示すものである。

Table 1*Length of Time to Run VTT2 to Create a Bilingual Transcription*

Steps after executing the VTT2 (length of time in HH:MM) [time accumulated in HH: MM]						
1: (00:30) [00:30]	2: (00:30) [01:00]	3: (00:30) [01:30]	4: (00:51) [02:21]	5: (00:25) [02:46]	6: (00:01) [02:47]	7: (00:01) [02:48]
Upload the classroom video to Microsoft OneDrive.	Transcribe the classroom video clip in L1 to obtain WebVTT L1.	Erase WebVTT L1 from the video clip and transcribe in L2 to obtain WebVTT L2.	Delete the unnecessary lines in L1 either on the video transcription or WebVTT.	Delete the unnecessary lines in L2 either on the video transcription or WebVTT.	Combine new-L1 and new-L2 to create mixed WebVTT on the Software VTT2.	Upload the L1 & L2-combined WebVTT transcription to the initial classroom video clip.

[Katagiri, N. (2023). Web Video Text Tracks to Compile an English Classroom Transcript with Bilingual Speech. *Proceedings of the 48th JASELE Annual Convention*, p193 より改変]

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 KATAGIRI Noriaki
2. 発表標題 Web Video Text Tracks to Compile an English Classroom Transcript with Bilingual Speech
3. 学会等名 Japan Society of English Language Education
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
研究分担者	大橋 由紀子 (Ohashi Yukiko) (40589793)	ヤマザキ動物看護大学・動物看護学部・准教授 (32822)	

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
---------	---------