

<実践事例>

オンデマンド配信の授業動画への字幕付与による情報保障の実践 ——聴覚障害のある学生への合理的配慮として——

垂門 伸幸¹・吉川 雄也¹・山本 敦也¹・脇坂 紗帆¹・仲兼久 知枝¹

2020年度春学期、本学ではすべての授業をオンラインで実施することとした。これまで障害のある学生に対する授業の情報保障は対面授業を前提としていたため、障害学生教育支援センターでは情報保障の新しい方法を検討した。そのうち、聴覚障害のある学生がオンデマンド配信の授業動画を視聴する際の情報保障として字幕付与を行う方針を立てた。本稿はこの方針のもとに作成した具体的手順を説明し、2020年度春学期の字幕付与の作業結果を報告するものである。2020年度春学期は、106件、動画の長さ合計約124時間の授業動画に字幕を付与した。多数の授業動画を取り扱うことになったが、作業体制やツール等の工夫により、効率よくまた概ね遅滞なく情報保障を行うことができた。一方でいくつかの課題点も浮かび上がってきた。よりよい情報保障のあり方について、今後も模索を続ける必要がある。

キーワード: オンライン授業、動画、聴覚障害、字幕、情報保障

1. はじめに

新型コロナウイルス感染症拡大防止の対応のため、本学では2020年度春学期の授業をすべてオンラインで行うこととした。本学では、聴覚障害のある学生への情報保障を、主として教員等の発話をタイピングにより文字化するパソコンテイクと言われる方法で行ってきた。これは対面授業を前提とした方法であったため、障害学生教育支援センター(以下、「センター」という)では、オンライン授業という授業環境に合わせた情報保障の方法を急遽検討することとなった。筆者らセンターのスタッフ自身も在宅勤務という初めての勤務形態を経験しながら、例年よりひと月ほど遅い5月11日の授業開始日に向けて、オンラインミーティング等で議論を重ねた。

このレポートの目的は、オンライン授業の中でもオンデマンド配信で行われる授業について、聴覚障害のある学生への情報保障として多数の授業動画に効率よく字幕を付与する体制を構築した経緯と結果を報告することである。これによって今後増加することも予想されるオンデマンド配信という、これまで主流ではなかった授業形態に対応した新しい情報保障の方法が普及することを期待している。

2. 授業動画への情報保障の課題と対応方針

先述の通り、聴覚障害のある学生に対する情報保障として本学ではパソコンテイクを導入している。対面授業

業ではDVD等の視聴覚教材(以下、「補助動画教材」という)が用いられる場合があるが、その際にもパソコンテイクで情報保障を行うことも可能ではあるものの、一般に補助動画教材は発話が速い(音声の流れる速度が速い)ため、パソコンテイクのタイピングがその速度に追いつかないといった問題がある。そのため補助動画教材の発話内容を事前に(授業までに)文字化するという方法(以下、「文字起こし」という)が取られることが多い。この補助動画教材の文字起こしは、今回報告する授業動画への字幕付与方法的に似ているが、オンラインという特殊な授業環境のもとでは2つの点を考慮する必要があった。

一つ目は、発話のタイミングの再現である。対面授業を想定した文字起こしの場合、授業中の補助動画教材の発話に合わせてサポーターが文字起こし済の文章をパソコンで一行ずつ表示するか、または紙面上で指差しするという方法で発話と文章提示のタイミングを合わせることができた。一方、オンデマンド配信の場合は、一般的に受講生は自分の好みの時間帯に動画を視聴するものであり、これがオンデマンド配信授業の利点でもある。聴覚障害のある学生がオンデマンド配信の授業動画を視聴する際、彼らの動画視聴の予定に合わせてサポーターを例えばオンライン上で派遣し、動画の発話に合わせて文字起こしされた文章を提示するといった方法をとることは調整に非常に労力がかかることが予想されることから現実的ではない。また、上述のオンデマンド配信授業の利点を損なうことにもなる。

¹ 京都産業大学 教学センター (障害学生教育支援担当)

二つ目は発話の分量である。通常の対面授業では補助動画教材を用いる授業自体、筆者らの経験上それほど多くはなかった。また、補助動画教材が使用される授業でもその分量はひとコマ90分のうちの一部の時間であることが多かった。しかし、2020年度春学期は、全ての授業をオンラインで行うことに伴い、多くの授業でオンデマンド配信の授業形態を取ることが予想された。さらに、授業ひとコマにおける動画の長さも当然長くなることが予想された。

これらの課題に対応するため、音声認識による自動字幕作成機能を活用して動画に字幕を機械的に付与した上で、音声認識の誤りの部分(以下、「誤認識」という)を人手で修正するという方針を立てた。これにより、補助動画教材で行ってきたように単に発話部分の文字起こしをするのとは異なり、発話部分のデータに、発話のタイミングを表す時間データを加えた2次元のデータを作成することができ、また、ゼロから発話データや時間データを作成するよりも、字幕付与の作業を効率的に行うことができると考えられた。

3. 方法

この方針のもとでの字幕付与作業の手順を以下のよう

【手順1】教員が授業動画を動画投稿サイトにアップロードする。

【手順2】教員が動画投稿サイトの機能により自動生成さ

れる字幕データ(以下、「自動字幕データ」という)をダウンロードする。

【手順3】ダウンロードされた字幕データを学生サポーターが修正する。

【手順4】修正された字幕データのデータ構造をセンターがチェックする。

【手順5】教員が修正済み字幕データをもとの動画投稿サイトにアップロードする。

以下、この手順ごとに補足説明する。

【手順1】動画投稿サイトへのアップロード

オンデマンド配信の授業では、一般的に動画は特定の動画投稿サイトにアップロードされ、受講生はサイト内の特定の URL にアクセスすることで受講する人が多いと思われる。本学のオンライン授業では Microsoft Teams (以下、「Teams」という) の使用が推奨され、Teams と連動した動画投稿サイトである Microsoft Stream (以下、「Stream」という) に授業動画をアップロードする教員が多かったと思われる。なお、Stream の他には、YouTube を使用した授業も一部あったことを把握している。

今回報告する字幕付与による情報保障を行うためには、授業動画は Stream や YouTube 等、自動字幕生成機能のある動画投稿サイトに教員が授業動画をアップロードする必要がある。

NOTE Confidence: 0.90593106

fd86110b-a715-4a8e-81cc-3d65762ebb9d

00:00:58.662 --> 00:01:06.614

まずですね どこから とりかかるとのことなんです
今日の配信資料とところどころ穴埋めになっていますので 穴埋めをしながら

NOTE Confidence: 0.90593106

97794edd-2b5c-4f47-9ca2-a9c51a72cb19

00:01:06.614 --> 00:01:14.566

聞いていただければというふうに思います。まずどこからスタートするか
なんです まず 本学の建学の精神ですね。

NOTE Confidence: 0.90593106

760c4f0c-7fc7-49f0-bae9-afb1f33f3a1d

00:01:14.566 --> 00:01:22.518

創設者である荒木先生が掲げられた 建学の精神に照らして
この障害のある学生の支援ということを最初に考えてみたいと

図 1. Stream からダウンロードした vtt データの例

【手順 2】自動字幕データのダウンロード

教員が授業動画を Stream にアップロードした場合、生成された字幕のデータは、「WebVTT」という形式(以下、この形式のことを「vtt 形式」、vtt 形式のデータのことを「vtt データ」という)でダウンロードすることができる。また、YouTube にアップロードした場合でも、字幕データを vtt 形式を含むいくつかの形式でダウンロードすることが可能である。以下では、字幕データとして Stream と YouTube に共通のデータ形式である vtt 形式を扱うことを前提に説明する。

ダウンロードした vtt データのファイルの拡張子は「.vtt」となっている。vtt データはテキストエディタ(Windows であれば「メモ帳」、Mac であれば「テキストエディット」等)で開いて内容を確認することができる。内容は一定の構造に従って構成されており^{註 1}、発話のデータと発話のタイミングを表す時間のデータが含まれている(図 1)。

【手順 3】ダウンロードされた字幕データの修正

Stream 等のサイトによって自動生成された vtt 形式の字幕データの発話部分には誤認識が多数含まれている。字幕を聴覚障害のある学生に対する情報保障として提供するためには誤認識を修正する必要がある。

センターは教員がダウンロードした vtt データの提供を受け、誤認識の修正作業のためのサポーターを募集する。例えば 90 分の動画の場合、30 分ずつに区切って 3 名のサポーターに修正作業を割り当てる。これにより 1 人で 90 分全部の修正を行うよりも、作業を早く終わることができる。一人のサポーターに割り当てる動画の長さは最長 30 分を目安にした。

サポーターが決定したら vtt データを提供し、動画視聴のための URL を知らせる。サポーターは動画を視聴しながら vtt データをテキストエディタで開いて誤認識を修正する。Stream の場合、そもそも誤認識の修正は Web ブラウザ上でも行うことができるようなサイトのインターフェイスとなっているが、ダウンロードした vtt データをテキストエディタで開いて修正する方が作業効率は良い。また、字幕を直接 Web ブラウザ上で修正する場合には、動画自体の編集権限が必要である。教員以外のセンタースタッフやサポーターに編集権限が付与されると動画自体を誤って削除してしまうおそれも考えられるため、編集権限を持つ教員に vtt データを一旦ダウンロードしてもらったものをアップロードした動画データから

切り離して編集する方が安全である。修正作業が終了したらサポーターは修正済みの vtt データをセンターに提出する。

なお、サポーターは有償のアルバイトであり、修正作業等は在宅で行ってもらう。センターでは 30 分の動画を修正するのに 4 倍程度の 2 時間の修正作業時間を要すると見積もった。このように作業にあたる動画の時間の 4 倍を勤務時間として管理することとした。

【手順 4】修正済み字幕データのデータ構造のチェック

vtt データの誤認識修正にあたって、サポーターには事前に vtt データの構造を保ったまま発話部分のみ修正するように指示した。しかし、発話部分を修正するうちに誤って発話部分以外の部分を書き換えてしまうことも当然想定された。vtt データが本来の構造にのっとっていない場合には Stream にアップロードしようとするエラーメッセージが表示され、データを受け付けてもらえない。

そのためセンターでは修正作業を終えてサポーターから返却された vtt データの構造が正しく保たれているかチェックを行うこととした。このチェック作業は膨大な量になるため、いくつかのプログラムを作成し作業の効率化を図った。具体的には、これらのプログラムは本学の事務職員が日常的に使用している Windows パソコンで動作するように、Windows に標準で搭載されている PowerShell (PSVersion 5.1.18362.1110 / PSEdition Desktop) で作成した。参考として図 2 および図 3 にプログラムの一部を示した。

図 2 は、時間表示の構造が保たれていない行を抽出するプログラムである。時間表示行は、

```
00:01:14.566 --> 00:01:22.518
```

のように、数字と「:(コロン)」「.(ピリオド)」「-->」といった文字列から構成されている。例えば、時間部分の末尾、ピリオド以下の 3 桁の数字はミリ秒を示しているが、ここを誤って 2 桁に書き換えてしまった場合や、この行に本来不要な文字(例えばスペース等)を付加してしまった場合には、図 2 のプログラムを実行することでエラーの含まれる行をピンポイントで抽出することができる。

図 3 は、時間表示行の後に誤って空行が入力されている発話部分を抽出するプログラムである。発話部分は、

```
Select-String -Path ./sample.vtt -Encoding UTF8 -Pattern "-->" | `
Select-String -Pattern "^\d\d:\d\d:\d\d\d\d.\d\d\d\d --> \d\d:\d\d:\d\d\d\d.\d\d\d\d$" -NotMatch
# チェックする対象の vtt データのファイル名が「sample.vtt」であった場合。
```

図 2. 時間表示の構造が保たれていない行を抽出するプログラムの例(抜粋)

```

$Line = Get-Content -Path ./sample.vtt -Encoding UTF8
# チェックする対象のvttデータのファイル名が「sample.vtt」であった場合。

$allLineNum = $Line.Length

$j = 1

for ($i=0; $i -lt $allLineNum; $i++){
    if($Line[$i] -match "-->" -and $Line[$i+1] -eq "" -and $Line[$i+2] -match "." -and $Line[$i+2]
-notmatch "^NOTE"){
        Write-Host "##### No." $j " ##### There is 1 empty line "
        Write-Host $Line[$i]
        Write-Host $Line[$i+1]
        Write-Host $Line[$i+2]
        Write-Host ""
        $j++
    }
}

for ($i=0; $i -lt $allLineNum; $i++){
    if($Line[$i] -match "-->" -and $Line[$i+1] -eq "" -and $Line[$i+2] -eq "" -and $Line[$i+3] -match
"." -and $Line[$i+3] -notmatch "^NOTE"){
        Write-Host "##### No." $j " ##### There are 2 empty lines "
        Write-Host $Line[$i]
        Write-Host $Line[$i+1]
        Write-Host $Line[$i+2]
        Write-Host $Line[$i+3]
        Write-Host ""
        $j++
    }
}

```

図 3. 時間表示行の後に 1 行または 2 行の空行が入っている行を抽出するプログラムの例(抜粋)

00:01:06.614 --> 00:01:14.566

聞いていただければというふうに思います。まずどこから
スタートするか
なんですが まず 本学の建学の精神ですね。

のように時間表示行に続けて次の行に配置されて、ひとつの塊り(ブロック)を構成しなければならない。しかし、

00:01:06.614 --> 00:01:14.566

(空行)
聞いていただければというふうに思います。まずどこから
スタートするか
なんですが まず 本学の建学の精神ですね。

のように、時間表示行の後に誤って空行を挿入し、ブロ

ックの構成が崩れていた場合、図 3 のプログラムを実行することで、エラーの含まれるブロックをピンポイントで抽出することができる。

【手順 5】修正済み字幕データのアップロード

サポーターから提出された修正済み vtt データのデータ構造等をセンターがチェックし、必要に応じて再修正したものを教員に提出する。教員はそれを Stream 等にアップロードして字幕付与作業は完了となる。なお、教員がセンターに字幕修正を依頼してから、修正を完了し、センターが教員に修正済み字幕データを提出するまでの期間は 4 日を目安とした。

4. 2020 年度春学期 字幕付与作業結果

2020 年度春学期、上述の流れで字幕付与作業を行

った。以下に作業結果の概要を示す。作業を行った期間は2020年4月27日から8月3日までであった。教員からの字幕修正作業の依頼件数は106件、動画の長さの合計は124時間20分(平均1時間10分)であった。修正作業を割り当てたサポーター数は、延べ266名であった。修正したvtt形式の字幕データは、発話部分のみで107,058行(1件あたり平均1,010行)、文字数(改行文字を除く)は1,882,155文字(1件あたり平均17,756文字)であった。

5. まとめと今後の課題

新型コロナウイルス感染症の拡大防止の一環で2020年度春学期の授業が全てオンライン授業となり、センターは障害のある学生への情報保障についてオンライン授業に即した方法の検討が求められた。聴覚障害のある学生へのオンデマンド配信の授業動画については、字幕付与による情報保障の体制を構築した。

当初から想定はしていたが、教員よりオンデマンド配信の授業動画への字幕付与の依頼を多数受けることになった。これらの依頼について動画投稿サイトが有する音声認識による自動字幕作成機能を活用することや、ひとつの動画教材の字幕データに対し複数名のサポーターを修正作業に割り当てること、vtt形式という特定の字幕データ形式の構造をプログラムにより機械的にチェックすること等、作業体制やツール等を工夫することにより効率的に字幕付与を行うことができた。

一方で、いくつかの課題が浮かび上がってきた。以下に3点を挙げる。まず、理系科目における数式について、音声を文字で表現する際の問題である。例えば、動画で教員が「に えー の にじょう」という発話があった場合、「 $2a^2$ 」なのか「 $(2a)^2$ 」なのか、音声だけでは定まらない。レジュメ等の資料とともに発話されている場合は教員の意図を推測しやすいが、そうでない場合も多い。

次に表現の統一の問題である。これも理系科目に多いが、文字をデータ化する際に表現の仕方が一つに決まらない場合がある。例えば、「 a_1, a_2, a_3, \dots 」のような添え字(下付き文字)については、「 a_1, a_2, a_3, \dots 」とする(添え字が環境依存文字であることから、それを使用せず通常の半角数字を使用する表現)、「 a_1, a_2, a_3, \dots 」とする(TeXの表記ルールに従った表現)等が考えられる。こういった場合に、どのように表記するのか、サポーターには事前に表記のルールをある程度レクチャーしたものの、字幕修正作業にあたる授業動画が多様な分野に渡っているため、サポーターには臨機に対応してもらおう余地を残し、ルールを厳密に守ることまでは求めなかった。サポーターから提出された修正済みのvttデータをセンターが監修する方法も考えられるが、能力と労力の観点から難しい。

最後に、合理的配慮としての字幕は、どこまでの完成度を目指すのかという問題である。情報保障の理想はある形式の情報を、障害のある学生が利用可能な別の形式に100%変換することであると考えられるが、現実的には情報の変換の過程で必ず情報の損失が起る。例えばパソコンテイクの情報保障率は6~8割と言われている(太田2016)。字幕付与による情報保障も情報保障率を100%に近づけることはできるが、時間と労力には限りがあるのも事実である。字幕付与による情報保障の取り組みは始まったばかりであり、よりよい情報保障のあり方について今後も模索を続ける必要がある。

謝辞

字幕付与の方法を検討するにあたって、情報理工学部の先生方には貴重なご助言をいただきました。また学生サポーターの皆さんには、在宅で慣れない中、字幕修正の作業にあたっていただきました。皆さまのご協力に感謝いたします。

注

1)WebVTTの詳細は、
<https://www.w3.org/TR/webvtt1/> や
https://developer.mozilla.org/ja/docs/Web/API/Web_VTT_API 等も参照。

参考文献

太田晴康 (2016)パソコンテイク その特徴と活用 15.
http://www.pepnet-j.org/web/file/tipsheet/2018/15_ohta.pdf (参照 2020.11.11)

Report on Practice of Subtitle-Editing to Access Vocal Information of On-Demand-Class Videos to Accommodate Hearing Impaired Students

Nobuyuki TAREKADO¹, Yuya YOSHIKAWA¹,
 Atsuya YAMAMOTO¹, Saho WAKISAKA¹,
 Chie NAKAGANEKU¹

All the classes in the spring term 2020 were given online at Kyoto Sangyo University. The office for educational support for students with disabilities faced the challenge of developing a plan on how to best accommodate the needs of disabled students, such as hearing impaired students, in an on-line learning context. The

plans included subtitle-editing of on-demand-class videos, which enables hearing impaired students to access vocal information. This paper explains the methods and tools employed, and reports the results of the subtitle-editing from the spring term 2020. A total of 106 videos had subtitles added, and the total time of the videos was about 124 hours. The methods and tools employed facilitated the efficient and effective addition of a large volume of subtitles. In addition, specific areas for improvement of the process were identified and are discussed in this paper.

KEYWORDS: online class, video, hearing impaired, subtitle, accessibility

2020年12月23日受理

1 Educational Support for Students with Disabilities
Office, Kyoto Sangyo University