

音声認識を用いた聴覚障がい学生支援 －実環境導入に向けた支援システム構築－

Supporting Hearing Impaired Students Using Speech Recognition Software - Construction of the Support System for Realistic Classroom Environment -

平賀 一馬^{*1}, 皆川 雅章^{*2}

Kazuma HIRAGE^{*1}, Masaaki MINAGAWA^{*2}

^{*1} 札幌学院大学心理学部

^{*1}Department of Psychology, Sapporo Gakuin University

^{*2} 札幌学院大学法学部

^{*2}Department of Law, Sapporo Gakuin University

Email: kazumahiraga22@gmail.com

あらまし：2014年からの施行を得て、2019年度から札幌学院大学では音声認識ソフトを用い、講義中の教員の発話を文字化して聴覚障がい学生に提示し情報保障を行っている。本学では、音声認識ソフトの誤認識を、支援学生がその場で修正する方式を用い、より精度の高い文字情報を提供することを目指している。実際の機器の使用方法、認識精度を上げるための話者の特性分析や修正方法、学生主体の講習会実施などについて報告する。

キーワード：情報保障，音声認識，認識精度，学生スタッフ

1. はじめに

2014年からの施行を得て、2019年度から札幌学院大学では音声認識ソフトを用い、講義中の教員の発話を文字化して聴覚障がい学生に提示し情報保障を行っている。本学では、音声認識ソフトの誤認識を、支援学生がその場で修正する方式を用い、より精度の高い文字情報を提供することを目指している。実際の機器の使用方法、認識精度を上げるための話者の特性分析や修正方法、学生主体の講習会実施などについて報告する。

2. これまでの取り組み

札幌学院大学では、2013年度より音声認識に関するプロジェクトを始めている。はじめは、ノートテイクの代替手段として考えられ、試験段階であった。2013年度は「ドラゴンスピーチ」というアプリを用いて実験を行った。この実験では、話者の特性が「話すのが遅く丁寧」だと認識率が高いこと、講義または話者の話す速度が速い場合は不適であることが判明した。

2015年度から、UDトークを使用した音声認識を始めた。2015年度は支援学生の不足による支援学生の負担が増加していることが特に懸念されていた。このときも実験を行い、その結果、音声認識導入に向けて、誤認識を減らす工夫、想定可能なトラブルの抽出、マニュアルなどの整備が上げられた。この時期から音声認識の基盤が形成され始めてきた。

2017年度では支援学生の養成を目的としたビデオ講習会が研究された。ビデオ講習会は、講習会実施時の講師役の生徒への負担を減らすことを目的として研究された。結果、作業をする際には講義のサポート学生が必要であることと、単調な講習会では

受講生が集中しにくいことが判明した。このビデオ講習会での研究結果は、現在の支援学生の講習会に引き継がれている。

3. 音声認識導入上の課題点

3.1 認識制度

音声認識の精度は、周囲の環境、話者の特性、使用する機器の特徴によって変わってくる。機器に関しては改善することが可能である。しかし、周囲の環境と話者の特性を変えることは、講義に対しての支援としては難しい。そのため、認識制度を改善できない場合、パソコンテイク、ノートテイクなどの他の支援形態との使い分けが必要になる。

3.2 使用する機器

必要な機器には、iPod 3台、パソコン、マイクなどがある。本学では、修正用端末としてiPodとパソコン、表示用端末としてiPadを使用している。この端末には、いずれもUDトークをダウンロードしておく必要がある。そのほか、Wi-Fi環境がない場所でのテイクをするためのポケットWi-Fi、iPod用のモバイルバッテリーがある。

3.3 話者の特性

話者の話し方が聞き取りやすい速度、滑舌であるほど音声認識の精度は上がる。逆に話者の話す速度が極端な速度や、非常に滑舌が悪い場合、音声認識の精度は下がる。また、講義形式によって、話者が一人の場合と、グループワーク形式など話者が複数人の場合がある。前者の場合、認識する音声は1つだけであるため、音声認識の精度は高くなる。また、修正者の負担は比較的小さくなる。しかし後者の場合、複数人が同時にしゃべるケースなど、認識する音声は複数になるため、音声認識の精度は低くなる。

また、修正者の負担も比較的大きい。

3.4 利用する環境の影響

音声認識支援では、教室内の環境に影響されることがある。音声認識支援は、周囲の雑音（複数の声、騒音など）、ネットワークへの接続強度によって、認識制度が低下する。周囲の雑音に対してはマイクの種類を変えることで、認識制度が向上した。ネットワークへの接続強度は、最低でも 3~5Mbps が必要である。これを下回る強度だと音声認識支援は不可能である。接続強度を上げる手段としては、携帯回線（4G 回線）を使う、高性能なポケット Wi-Fi を使うなどが挙げられる。

3.5 誤認識の修正

まず、3-1 で紹介した機器を用意する。

次に、UD コネクトを用いてトークルームを作成する。UD コネクトはホスト用だけ起動し、トークルームの親機とする。UD コネクトでトークルームを作成後、他の UD トークを開いている端末を親機が招待する。これで、音声認識を行うネットワークができ、トークルームに複数端末がつながっている状態となる。すべての機器の接続が完了すると、修正可能な状態となる。次に実際の修正方法である。これはイラストを用いて説明する。



図1 教員からの音声入力



図2 入力音声の誤認識



図3 誤認識の修正

図1~図3の流れで説明する。まず、図1では、講師の「色々ありました」という音声情報が端末に入力される。入力された音声情報は、Ami Voice Cloud と呼ばれるサーバーに送信され、各端末に送信される。図2の中では、「色々ありました」という言葉を、

「イライラしました」と誤って認識している。これを、テイカーは修正したのち、またサーバーに送信している。図3では修正された文が送信され、利用学生はこのときにご認識に気づくことができる。

4. 学生主体の講習会

音声認識の講習会では、テイカーとして支援に携わる際に必要な知識を学んでいく。3-4 で述べた修正方法に加えて、機器の種類と使い方、現状の問題点、そして実践を通して、支援者としての能力を身につけていく。

本学では、支援者の育成を学生主体で行っている。学生主体で講習会を開くことで、講師側を担当する学生は、より支援について深く考えられるメリットがある。また、講師が学生であれば、受講する学生の精神的な負担も軽くなる。講師側を担当する学生は、経験を積んだ学生が担当する。

5. 支援活用の実績と課題

2019 年度から音声認識支援を実践導入していった。結論として、音声認識での支援は講義によって使い分けが非常に難しいことが分かった。

例えば、講義内で複数の、話者がいる状況である場合、認識率が極めて低く、また話者の切り替えが激しいため、利用学生の話者の特定が困難になっていることが多かった。そのため、全 15 回の講義のうち、前半 3 回をのぞいた 12 回は、3 人体制のパソコンテイクでいどむこととなった。また、ネットワーク環境による遅延、および通信切断が起りやすい教室での講義では、ネットワークによる問題の原因がつかめず、パソコンテイクに切り替えた。

現状の音声認識はまだ安定した支援とは言えず、使いどころの限られる支援となっている。今後はネットワーク環境と、学内への音声認識の認知を深めることが課題として上げられる。

6. おわりに

2 音声認識支援は従来のパソコンテイク、ノートテイクに比べて、支援者の負担が減るというメリットがある。しかし、修正者に求められる能力、音声認識の精度が変わる環境を考慮して、支援体制を使い分ける必要がある。今後は、音声認識の精度の向上、修正者の能力向上をはかることで、安定した支援ができるようにすることが求められる。

参考文献

- (1) 札幌学院大学 電子計算機センター,
<https://densan.sgu.ac.jp/>, (参照 2020-02-05)
- (2) 札幌学院大学電子計算機センターサポートデスク,
<http://supportdesk.sgu.ac.jp/>, (参照 2020-02-05)
- (3) ゴーホージャパン株式会社: “Zoho Connect | 機能”,
<https://www.zoho.com/jp/connect/features.html>, (参照 2020-02-05)