

## 音声認識ソフトの講義活用のための導入実験

皆川雅章\*

Email:minagawa@sgu.ac.jp

\*1: 札幌学院大学社会情報学部社会情報学科

### 1. はじめに

音声認識ソフトを使って教員の口頭説明を文字化した結果を講義で活用するための導入実験を行った。

本学では、聴覚障がいを持つ学生(以下、対象学生)に対して、講義中に本学学生による「ノートテイク」あるいは「パソコン通訳」の支援<sup>(1)</sup>が行われている。これによって講義中、対象学生の横に支援を行う学生(以下、支援学生)が着席し、教員が話す言葉をノートに筆記、あるいはタイプングにより PC 画面上に表示し、対象学生はそれを見ながら講義を受けることができる。

一昨年度、この支援と並行して、PC を用いた演習型講義での説明や指示などを、音声認識ソフトを用いて文字化し、対象学生に伝えた。これらの講義においても支援学生が配置されているが、実質的にテイクを必要とするのは、講義の始めに行う、演習に関する説明や指示、演習中の補足説明の部分に限定されているので、教員だけで対応できないかと考えて作業代替方法を探った。タブレット型端末やスマートフォン上の音声認識ソフトを使用して、この支援を2つの講義で実施し、一定の成果を得たが、音声の変換結果、変換スピード、対象学生への文字情報提供方法など、実際の教室で円滑な活用を行うためには、次のような点で検討・解決すべき課題が見つかった。

- ①リアルタイム性の問題
- ②誤認識を起こさないような発話の工夫
- ③誤変換結果の速やかな訂正
- ④対象学生への文字情報の提示方法
- ⑤講義全体の運営の問題
- ⑥機器とソフトウェアの性能・利用環境

これまでに、これらの課題に取り組み<sup>(2)</sup>このアプローチが適用可能な条件を探ってきた。今回は、上記のような演習型ではない通常講義を行う教室での適用可能性を探るための導入実験を行っている。

### 2. これまでの取り組み

上述の課題に対するアプローチとして、講義中に行った説明を実験用例文(文字数 1256)として用い、下記のように、デバイスやアプリケーションソフトが異なる環境における、音声認識ソフトの性能、デバイスの操作性、誤変換の修正法、リアルタイム性等を比較実験し、ノートテイクによる情報保障の代替支援として音声認識ソフトを導入する際の課題を考察した。

装置 1 : スマートフォン+音声認識ソフト(DD)

装置 2 : タブレット型端末+音声認識ソフト(DD)

装置 3 : PC+音声認識ソフト A(AV)

(DD:Dragon Dictation, AV:Ami Voice)

考察結果は次の通りである。

①装置 1, 2 において変換のためのタップ操作を介する影響は、教室内でのやりとり時間に時間的な余裕がある場合、例えば、ゼミナールなどにおいて、意見交換を行うような場合には、許容範囲である。

②誤変換の訂正は、いずれの装置でも同様の操作(訂正箇所の確認、訂正箇所へのカーソル移動、キーによる修正入力)を行う必要があり、操作上の優位性に大きな差は見いだせない。ただし、装置 1, 2 は片手に持って操作できるので、講義中の発話時に教員が教室内を移動するような場合に有利で、教壇等の定位置にいる場合には装置 3 が向いている。

③変換の特性を把握し、正しく認識されるように発話の工夫で誤変換を減らす余地があり、例文の文字数程度の文章で 10 個以内の誤変換数であれば、訂正時間が発生しても講義の流れへの影響は小さいと予想される。

③教室外での対象学生の個別の修学相談等において、担当教職員に手話のスキルがなくても発話を文字情報として伝達できるので、より正確な意思疎通のために活用できる可能性がある。

### 3. 今回の取り組み

上記の取り組みは継続的な検証と並行して、通常講義における音声認識ソフトの活用の可能性を探る。ここでの通常講義とは、教員による説明が主体の講義で履修者が 100 名程度を想定している。また、講義運営上の想定として、特に聴覚障がい学生を意識するのではなく、履修者全体を対象として発話することを前提としている。

実験を実施した講義の内容は次の通りである。

①1 年次前期選択科目(社会科学系学生を対象)

②履修者数 85 名の中教室

③聴覚障害学生 1 名(パソコン通訳による情報保障あり)

④授業のポイントを書画カメラで提示

⑤教科書の説明にはタブレット型端末を使用

⑥説明時に教室用マイクを使用

音声認識のための装置は、PC 上で動作する

DragonSpeech<sup>(3)</sup>である。ヘッドセットは有線(教壇上の可動範囲を確保するために延長ケーブル使用)で PC に接続されている。発話に変換される間、誤変換の修正は行わず、そのままの結果を収集することとした。

なお、本学ではノートテイクやパソコンテイクが行われている場合、テイカーの存在を意識した発話を考慮するように指示されているので、それに従った。講終了後にテイカーに発話の状態を確認し、これまでのところ、特に問題はなかったとの回答を得ている。

#### 4. 実験結果

8回の講義で、音声認識結果データの収集を行った。その1例を以下に示す(文字数1205)。文中、下線を引いた部分が誤変換(話者の意図した言葉とは異なるという意味で)と考えられる部分である。説明内容はメール送信時の留意点についてである。「後期」→「講義」、「県名」→「件名」、「わい」→「場合」、「追記」→「追加」、「速攻」→「すぐに」、「渡して」→「はたして」、「要領」→「要領」、「明るい」→「わかる」、「ゆとぎに」→「いうときに」であることは推測でき、講義内容もほぼ再現できる。

##### 変換例1

それではこれから後期を始めます。9まず最初に今日の予定を説明しますでは教科書34ページを開いてくださいまずメールを送るときに大事なことがあります1つ目は適切な県名を書くと言うことです県名を見ただけでメールの中身がわかるように簡潔に適切な県名を描きましょう2番目これは相手の名前を本文に書くということです他の受信者には分からないようにBCCで送ったのに自分のメールだと勘違いして遠因にメールを返してしまいういもあります3番目自分の名前を本文に書きましょう送信先に応じて必要な情報を追記することが望ましいといえます。水はメールコミュニケーションについて話をしますではメールコミュニケーションについて話をします。目に換算。同教科書34ページを開いてください今日のお題はメールによるコミュニケーションです皆さんは普段メールを送っていると思いますけれどもいくつか注意してほしいことがありますまず1つ目ですここにあるように適切な県名を書くということですただ単に赤みだけを書くのではなくってメールを受け取った人がどういうことに関するものなのかということ速攻見ただけでわかるようにしておきましょう速攻を見ただけでわかるようにしておきましょうそこを見ただけでわかるようにしておきましょうプさて次に3歩目です3番目の話題に入りますそれは相手の名前を本文に書くということです受け取る側が渡して誰にきたのかということがわからないときはあります要件のだけを書かれてもわからないということですでは次に来ますよ添付ファイルについての説明です皆さんは本文以外にwordのファイルとかエクセルのファイルとか添付することがあると思いますその時に注意しておかなくてはならないことがありますまず要領の大きすぎるファイルは使わない要領というのはコンピューターの中で表現されるデータのオブションですデータの大きさですここでCCとBCCこれらの違いについて説明しておきますまずCCですけれども直接の下早春の相手以外にも同じ内容のものを送りたい場合がでできますその場合にはメールのオプションで親子というのがありますのでCCというのがありますからこれを使いましょう頭に描いていますこれが自分ですここに直接の送信相手がいますツールを使う場合はこのように直接メールが送られますいまお話ししたCCというのは内容のことを例えば君リークンピクンでこれら2人にも伝えるということになりますCCというのは直接の送信相手以外にも例えば太郎くん花子さんにも伝えるということになります。あお互いに他の人にメールが送られているということが明るいようになっていきましたところがBCCの場合これは違いますBCCで送られた人は他の誰に送られたかが分かりますがそれ以外の人にとってはこのECCで送られた相手が誰かということがわかりません。このことは実際にメールを送る場合にどのような目的で誰に送るかということを明確に決めて送るとゆとぎに大事なことになります

上記の例は比較的、発話の再現が容易と思われるものであるが、以下の例のように、再現が困難な場合も生じる。説明内容は2進数における基数の考え方についてである。変換例1と比較して、もとの発話内容を推測しにくい部分が多いことがわかる。講義中の教室内は比較的静穏であるが、ヘッドセットの装着方法、教室内マイクの音声の影響などを受けていることも考えられる。

##### 変換例2

我われがこの手の指を使って数えたという話をいもしたんですけどジョンソンをドラえもんがね同じように自分のなんか体のパーツを使って数えるって時に神戸のオラオラ超えて123030ひとかたまりでできますねさてドラえもんをどうする我々は熟考ひとかたまりで考えや自分たちのこの手法でドラえもんの場合には猫をひとかたまりとして上ってこれが実は2進数の考え方それであと少しだけ説明して終わりますね2進数の場合は基ことを住んでしまった。井関すねと汚い気が深いので買え先の365頭に寝るでねさて鉛筆が目の前にご本であると思つてますさてドラえもん式の数え方日本集まったら一言があまり効率をよく脱いで日本集まったら一方、最後一方だった。さて今こうやって円を付けたけどもう一つ。つけれますねこれが一言もいいのです2進数の世界では101と書きます酒の受信子思い出してくださいいちばんしたかねいちばん腹の少ないです2進数値の夫婦両手いっぱいなきやと足1本後それでこれは熟日本原が残りましたそれから9桁というのは日本東ねたものがありますかということなんだけどそれはなくてさらに日本東ねたものを入れ2つ集めてeriko東にしたのが本屋で101と読みますがこれが2進数でのあらわしかたです

#### 5. おわりに

通常講義で音声認識ソフトを使って教員の発話を文字化する実験を行った。今回の実験を通じて、以下の点を確認することができた。

- ①変換例1のように、比較的認識結果が良好な場合には、講義後に教員がテキストを修正・編集して配布することは可能である。
- ②対象学生向けにリアルタイムでノートテイク/パソコン筆記の代替手段とすることは困難である。そのため、安定して誤変換数を減らす工夫が必要となる。なお、講義中にソフトの動作が不安定になる場合もあり、教員が講義に集中しつつ安定した状態で音声認識を行うためには、講義のサポート態勢の整備、あるいはこのような講義方法を考慮した教室環境の整備が必要になると考えられる。

#### 参考文献・URL

- (1) 本学で実施されているノートテイクに関して  
札幌学院大学バリアフリー委員会  
<http://bfc.sgu.ac.jp/>
- (2) 皆川：”音声認識ソフトを用いたノートテイク代替支援の可能性に関する1考察”，2013PCカンファレンス講演論文集，2013年8月
- (3) Dragon Speechに関して  
NUANCE <http://japan.nuance.com/dragonspeech/>