

## 音声入力システムの教育的利用

一般教育等（聴覚障害系） 細谷美代子  
教育方法開発センター（聴覚障害系）石原保志

**要旨：**リアルタイム字幕提示システムが聴覚障害者への情報保障手段としてきわめて有意義なものであることは既に明らかにされ、技術的にも完成の域に達していると言える。しかし、従来のキーボード入力方式は専門的技術を必要とするため、聴覚障害者を対象とする学校教育における利用・普及は低調であった。近年開発された音声入力方式は専門的技術を必要としないことから、その教育的利用に期待が寄せられている。音声入力方式には、授業者の発声を話者以外の第三者が復唱し音声入力する方法と授業者の発声を直接入力する方法とがある。本研究は後者の方法で、本学の授業改善を試みようとするものである。本稿では予備的研究と今後の課題について報告する。

**キーワード：**リアルタイム字幕提示・聴覚障害者・情報保障・音声認識・授業

### 1. はじめに

共同研究者の一人が本学聴覚部で担当する「国語」（一年次必修）では、教科目の特性から言語表現に関することがらが指導の中心的内容となる。したがって聴覚に障害を有する学生にいかに豊かな言語体験を与えようか、ということが教育成果をあげることに緊密に結びついている。乏しい情報量の下での貧しい言語体験という環境では有意なコミュニケーション能力の育成が困難であるということはいままでの「教授場面での情報量の増加」を図ることが教育成果を挙げるためには必須である。

従前の手話・口話・板書・プリント等に加えてリアルタイム文字提示システムを新たな情報保障手段として教場に導入することが教育効果を挙げるのに有効であろうと考えられた。

### 2. 先行研究

リアルタイム文字提示システムには大別してキーボード入力方式と音声入力方式とがある。

前者に関しては本学創設当初から、教育方法開発センターを中心に先端的研究が鋭意進められてきた。すでに数々の機器開発等所期の研究成果を挙げていることは周知の通りであり、聴覚障害者に対する情報保障手段としては機能面で実用段階に達している。しかし、入力担当者の人件費・処遇、システムの設営・運用における専門技術者の確保などの諸点から、いつでも手軽に学内授業に使うことができるわけではない。全国の聾学校への普及という面でもこの点が懸念されるところである。

後者は近年音声認識ソフトが実用化され、市販のもの

が安価に入手できるようになったことを背景に関心が高まっている。音声認識ソフトを用いた入力方式はさらに、次の二種に分けられる。

- ・ 話者の音声を取録しマイクに吹き込む担当者や字幕修正担当者を配置する
- ・ 話者自身がマイクを通して入力する

本学教育方法開発センターではすでに前者の研究が進行中である。当該研究の目的は、話者以外に二人の要員を配置した上で、より高い変換効率能力を持つ機器・ソフト開発を目指すものであることが報告されている。

後者は人件費等の負担を考慮せずにすむという点で本学の教育環境の中でも実現可能な方策と考えられた。情報収集の結果、こうした音声認識文字提示システムが聴覚障害者の教育場面に有効に働く可能性が高いにも関わらず、その教育方法や効果等に関する研究はまだ蓄積されていないということがわかった。それは安価で実用に耐えるシステムがつい先頃まで未開発であったことが最大の理由であるが、同時に企業における技術開発の主たる動機が文章作成の省力化という点にあったことも影響していると思われる。さらに音声による機械操作が可能であるという点から視覚障害者や肢体不自由者の利用を目指すという発想に傾きがちであったことが聴覚障害者に対する利用に関する研究成果の少ない原因と思われる。

### 3. 本研究の位置づけと目的

このような研究動向を踏まえ、本研究では話者自身が

マイクを通して入力するという方法を採用し、聴覚障害教育における教育的利用可能性を追究する。

本研究が採用する入力方法は最も簡便、安価で、導入しやすいという条件を持つ反面、限定された厳しい変換効率という条件を伴うものでもある。字幕の誤変換修正者を置かないという条件の下で、情報保障手段としての音声入力システムの機能・使い勝手、支援機器としての効果的な教育利用の方法等を明らかにすることが本研究の目的である。

本研究は広く聴覚障害児童・生徒・学生に対する教育的利用を探ろうとするものであるが、当面の資料収集、及び検証の場は本学における「国語」「総合文化（文学）」などの授業を主とする。

#### 4. 研究環境の設定

##### 4.1 機器類

実際の講義場面での利用については、216講義室の大型モニターに字幕を提示することを目指す。予備的研究として、現段階ではノートパソコン、モニターに提示して調整中である。

ハード：ノートパソコン

Gateway Solo 2500（日本ゲートウェイ）

モニター

- ・液晶ディスプレイ TV LC - 150 F 1（シャープ）
- ・KV - 34 ST 70（ソニー）

マイク

ヘッドホン型 ANTI-NOISE NC61  
（アンドレア）

ソフト：ウインドウズ98（マイクロソフト）

ViaVoice98（日本アイ・ビー・エム）  
一太郎9（ジャストシステム）

市販の音声認識ソフトは他に次のようなものがあるが、今回はViaVoice98を用いた。

- ・SmartVoice Ver.2.0（NEC）
- ・DragonSpeech Select（ドラゴンシステムズ）

##### 4.2 音声認識ソフトの準備

音声認識ソフトに授業者の声や話し方の特徴を記録して認識効率を挙げる「エンロール」を約2時間実施した。さらに、昨年度の講義プリントより「文学」の10コマ分をテキストファイルの形にして学習させ、語彙拡張を計った。

#### 5. 検証

##### 5.1 マイク装着の問題点

音声認識システムを授業者本人が使う際には授業者がマイクを装着しなければならない。このマイク装着が実際の講義場面でどのような問題をもたらすかを探った。

##### 5.1.1 学生の側から

マイク装着が学生の口形の読みとりに影響があるか否かは大きな問題である。授業者は顔の正面が受講生の方に向くような位置を確保するのが聴覚障害教育の基本であるが、授業のはじめから終わりまで授業者が一定の場所に止まり、一定の方向を向いているわけにはいかない。実際にはさまざまな方向、姿勢をとることになる。

そこで実験的に216講義室における授業のはじめから終わりまでマイクを装着し、学生に評価を依頼した。途中でマイクの位置を少し変えた。前半は左頬のあたり、後半は顎のあたりに置いた。調査時はマイクを装着するだけで、字幕をモニターに流すことはしなかった。本学聴覚部一年次学生を対象に、5クラスで資料1のアンケートを実施し、48人分の回答を得た。

##### 資料1 アンケート

1から4の質問について、該当する記号に をつけてください。

1：「左ほお」のあたりにマイクがある時、

- A：見えないことが多かった
- B：見にくい時があった
- C：気にならない

2：「あご」のあたりにマイクがある時、

- A：見えないことが多かった
- B：見にくい時があった
- C：気にならない

3：あなたの席は白板に向かってどのあたりですか。

- A：右手の方
- B：正面に近い
- C：左手の方

4：あなたは、人とコミュニケーションを取るのに口形を重視していますか。

- A：非常に重視している
- B：どちらかといえば重視している
- C：あまり重視していない
- D：ほとんど頼らない・口形を読めない
- E：なんとも言えない

##### 結果

結果を表1に示した。口形を重視している学生が多いなかで、マイクが口形読みとりに決定的な障害にはならないであろうことが示唆された。

表1 マイク装着に関する学生評価

学生の位置 マイク位置	左手		正面		右手	
	左頬	顎	左頬	顎	左頬	顎
見えないことが多かった	0	0	0	0	0	0
見にくい時があった	3	1	7	0	3	0
内訳(学生の口形に対する意識)	A-1 E-2	A-1	A-3 B-3 C-1		B-3	
気にならない	15	17	10	17	10	13
内訳(学生の口形に対する意識)	A-7 B-7 C-1	A-7 B-9 C-1	A-7 B-2 C-1	A-10 B-5 C-2	A-1 B-6 C-1 D-1 E-1	A-1 B-9 C-1 D-1 E-1
小計	18		17		13	
総計			48			

### 5.1.2 授業者の側から

マイクを装着して講義をするときの不都合として、顔の近くにマイクがあると、手話を使いにくいということがあった。具体的には「本当」「英語」「紹介」「発表」など、顔の近くで表現する手話単語の使用に支障が生じた。マイクに手があたると、マイクの位置がずれたり、それを避けようとするとき正確な表現ができないという類の問題も発生した。

パソコンとマイクをつなぐコードにも問題がある。授業者は講義室内の二方の白板を使用するためにその間を行き来するが、コードが短いときは移動が困難であり、十分な長さを確保しようとするとき、時には足元が危うくなるおそれがある。伸縮するコードまたはワイヤレスマイクなどが使えるとよい。

### 5.2 仮名での読み提示

短歌や俳句を授業で取り上げる際には聴者に対する時とは異なる困難がある。授業プリントを与えるだけでは、意味を捉えるのは難しいものもある。一般には一度朗読することで、息継ぎ、強弱を通じて、意味の句切れのヒントをつかむことができる。聴覚障害者の場合、それが難しい。文字提示の方法を工夫することで単語の読み、意味の切れ目などをわかりやすく伝達できないかという

観点から検証した。

新聞の投稿歌壇からの作品を取り上げ、わかりやすい提示方法を探ってみた。

a：余所目には細し貧しと見えしかもほしいまま来し一本の道

(朝日新聞 平成11年11月28日(日)東京版朝刊)

これを学生に朗唱させると「余所目には細し貧しと見え、しかも、ほしいまま来し一本の道」

と読むものが少なくはないであろう。しかし、仮名で

よそめには ほそしまずしと みえし かも  
ほしいまま こし いっぽんのみち

と提示できれば、鑑賞のレベルはさておき、一応の歌意をつかむことは容易になると予想された。そこで、この短歌を音声入力してみたところ、変換効率も極めて低かった。

原因は次のような点にある。

- ・「余所目」という語彙は入っているが、冒頭の語であることから、認識しにくいようである。

- ・「細し」「貧し」が文語であるため認識しにくい。「細く」「貧しく」であれば認識するが、「ho」の「h」は認識しにくい。
- ・「見えし かも」も「し」「かも」が文語であるため認識しにくい。

b：番号で名前を呼ぶ暮しにも慣れて郵便看守より受く

投稿者はアメリカで刑務所に入っている男性である。「番号で名前を呼ぶ、暮しにも慣れて、郵便看守より、受く」と読むと「何を受け取るのか？」ということになりかねない。「番号で名前を呼ぶ暮し、にも慣れて、郵便、看守より受く」と読まねばならない。しかし、ここでも「呼ぶる」に受け身の助動詞「る」が出現するために認識しにくかった。

同日の朝日歌壇の他の作品30余首についても、やはり文語は苦手という同じような結果が出た。

次に、口語短歌の代表ともいえる俵万智の作品を検証してみた。講義で紹介した、『チョコレート革命』よりの17首はいずれも高い認識率を示した。歌集名にもなった「男ではなくて大人の返事する君にチョコレート革命起こす」をはじめ、「無理をしてほしいと言えば会いにくる深夜かなしく薔薇を抱えて / 父として君を見上げる少女あり深く鋭く我と関わる」などは認識率100%であった。17首の総音数540 に対して誤認識総音数は42であり、認識率は92%と高い数値を示した。

### 5.3. 効果的な利用方法

授業中のコミュニケーションは授業者と学生、学生と学生など、多角的な方向の発信・受信が飛び交う。授業者からの一方通行的な講義になるのを避けるため、また、学生の発表能力を高めたいという意図から学生の発表の場を意識的に設けて授業を構成している。発信方法としては、口頭発表(手話等を含む)以外に、学生に板書させる、講義プリントに各自が記入したものを即、張り出し、全員で見合うなどの形をとってきた。互いの考え・意見の確実な交信という目的は達するが、時間がかかりすぎること、授業の進度が遅くなることなどのマイナス面があった。こうした点の改善を目指した利用方法を探った。

モニター画面の文字提示行数をクラス人数分に一致させる。一行文字数は20字に設定した。一人一人の意見を聞いたり、学生に発表させるとき、学生の発言内容を授業者が復唱し、音声入力していくと、発表が一巡りすると同時に全員分の答えを提示できる。授業者が要約筆記的に復唱し音声入力することで、学生発言のうちの無駄

を省くことも可能である。

発言させただけでは情報が伝わりにくく、記憶にも残りにくい。さらに比較検討する際の障害となる。しかし、板書していると時間をとってしまうという矛盾を解決する使い方として有望な方法である。リアルタイム字幕提示はその情報量と臨場感に一つの価値があるものであろうが、発言を完全に復元することに主眼を置かず、タイムリーな利用方法を探ることが、聴覚障害学生の教育場面においては重要である。

### 6. 今後の課題

音声認識ソフトの現在の性能・特性を踏まえて、聴覚障害教育の日常の教育場面における実用的な情報保障手段の一つとして、使いやすい条件、効果的な利用方法を検証・開発することが今後の課題である。

### 参考文献

- 1) 石原保志 他：聾学校生徒のテレビ視聴のための字幕挿入の研究 - 文字量・呈示時間の番組内容理解に及ぼす影響 - . 特殊教育学研究 27(2) : 25-37,1989 .
- 2) 石原保志 他：漢字仮名混じり字幕における漢字の読み表示に関する研究 - 聴覚障害者を対象とした音声文字変換による情報保障において - . 日本特殊教育学会第33回大会発表論文集 : 218,1995 .
- 3) 兼子次生：教育字幕の技術的可能性についての基礎的研究 リアルタイム提示型教育字幕の技術的報告 . 第33回全日本聾教育研究大会研究集録 : 120-121,1999 .
- 4) 小林正幸 他：音声認識を利用したリアルタイム字幕提示システム . 筑波技術短期大学テクノレポート6 : 71-74,1999 .
- 5) 小林正幸 他：聴覚障害学生のためのリアルタイム字幕提示システムの改良・研究(6) . 日本特殊教育学会第37回大会発表論文集 : 46 , 1999 .
- 6) 坂井美恵子：聾学校高等部における、リアルタイム字幕付き授業の有効性について . 第33回全日本聾教育研究大会研究集録 : 118-119 , 1999 .
- 7) 内野権次：音声認識技術を用いた音声文字変換提示システムの開発 - 聴覚障害教育のための情報保障支援機器としての実用化を目指して - . 現代聴覚障害教育研究 : 89-92 , 1999 .

本研究の環境設定については、小林正幸氏(本学教育方法開発センター)にご協力いただいた。記して謝意を表する。

本稿は平成11年度文部省科学研究費補助金の交付を受けて行われた研究の一部である。課題番号：11680290

## Educational Use of Speech Recognition Systems

Miyoko HOSOYA Yasushi ISHIHARA

Real-time video caption systems are extremely valuable for providing information to the Deaf and Hearing Impaired, and are considered state-of-the-art technology. However, as manual data input requires specialist technicians, it has been difficult to develop and use the system in education for the Deaf and Hearing Impaired. Recently developed speech recognition systems do not require specialist technicians and thereby increase the potential of employing video caption systems in the education of the Deaf and Hearing Impaired. Speech recognition systems consist of two modes; the lecturer's utterances are input by a third person; or the lecturer inputs directly. The current research focused on the latter to improve college lessons. This paper reports the preliminary research and topics for future study.

Keywords: real-time video caption systems, the Deaf and Hearing Impaired, providing information, speech recognition systems, lessons