

アクセシブルな大学教育  
企業と大学との新しいパートナーシップ  
Saint Mary's University, Canada  
International Manager Mr. Keith Bain

私はカナダのセントメリー大学で障害を持つ学生達をサポートしています。セントメリー大学は、カナダで最も古い大学で1802年に創立されました。セントメリー大学は、歴史の古い大学であり、障害を持つ学生達のサポートも古くから実施してきました。また、大学の図書館は視覚障害を持つ学生でも利用できるように配慮してあります。

今回紹介するプロジェクトは「Liberated Learning Project」です。このプロジェクトは、障害を持つ学生をサポートするために、音声認識の技術を利用して、授業の内容を記録します。音声認識のシステムを利用して「ノートテイク」を行ないます。メモをとるといのは、とても難しいことですし、その人の知識によっては十分なノートが取れていないこともあります。大学での授業の様子を見てください。聴覚障害の生徒がいます。手話通訳がいます。講師はワイヤレスマイクをつけています。教授は、その音声で話しますが、その内容がスクリーン上に表示されます。1人は、全く聞こえない生徒です。文字をみながら授業をうけるということです。素晴らしいことに、すべての学生がスクリーンをみえています。これが健聴者の生徒でも、講義が行われた翌日、このメモが欲しいと要求してきます。一人の生徒は、音声認識により自分のノートでとったものと比較することが大事だということです。

これから、どういったアクセシビリティを提供してきたか、音声認識の機能を話し、デモンストレーションします。また、どのような影響がこれまでもたらされたのか、これから先、どうやっていこうとするのか。最後に一つのモデルということで、世界中に役立つような、大学と企業のパートナーシップについて、説明します。

一般的に言うと、ほとんどの大学、国では、障害を持つ学生達には何らかのサービスを提供しています。基本的なところからお話すると、ほとんどの大学がなんらかの物理的なアクセシビリティ、車イス用のスロープなど提供しています。また、大学ではカウンセリングを行なっているところもあります。生徒の中で学習障害を持っている人などへもテストの便宜をはかっています。点字、デジタルオーディオをつかった試験を実施したり、多くの大学では、手話通訳をいれて授業を行なうことは一般的な利用になっています。最近では、障害者支援技術を利用している大学が多く見受けられます。メモをとるノートテイクのサービスです。このノートテイクですが、非常に大きな課題の一つです。リアルタイムの情報をノートテイクで提供するのは、非常に難しいことです。いろいろなアプローチはありますが、仲介者が入ります。

これは、有料で提供されていたり、ボランティアであったりします。他の方法としては、

字幕、速記などです。また、講義の内容を録音して、視覚障害者が利用することも多く使われています。ただし、録音テープは先進的な技術とは言えません。カナダではオンデマンドで実施します。大学にはお金がないので、やらなければならないとき何とかするのが、オンデマンドです。大学では、障害のある学生にとって一番よい環境を提供できるように努力します。ボランティアのノートテイカーなどがそうです。

現在、大学の中で、速記、テープ起こしなどフルタイムでやっている所はなく、手話通訳も常にいるわけではありません。聴覚障害の生徒は、本人の努力に委ねられるケースもあり、非常に限られた選択肢しかないことも事実です。

カナダの大学の場合、常時の手話通訳はおらず、必要なときだけ手話通訳を呼ぶことになっています。米国では高度な専門的なやり方をしています。アメリカ障害者法 508 条があるからです。リアルタイム字幕とか、速記のタイプで、シープリントやタイプウエルがあります。海外の例としては、常時いるような手話通訳というのは非常に珍しいというのが、オーストラリアの状況です。

イギリスは、二つを組み合わせたモデルです。専門的な者と、その他のボランティア的な人もいます。イギリスの特徴は、障害のある学生は、手当をもらうことができます。

ですから、アメリカ、カナダの場合には、学校に手当が支払われますが、イギリスでは、学生に直接手当が支払われます。学生はそれを自由に使えます。字幕、キャプション、手話通訳を依頼することも出来ます。個々の学生の状況によって、支払われる金額には差があります。この差を埋める大学もあります。

しかし、これらのアプローチには、システム的問題があります。ほとんどの場合、誰かが必要な情報にアクセスするために、何らかの提携が必要です。例えば、誰かがノートを取る必要があるわけです。だいたいの場合には、ボランティアです。また、取ってもらったノートが、常に質の高いものではないという基本的問題が存在します。2番目に、有資格者、いわゆるプロを使った場合、特に田舎の地域では、そういう人材を見つけること自体が難しい場合があります。ほとんどの大学ではコストがかかるという問題に直面してしまいます。

今までお話してきた問題を解決するために、新しいアクセシビリティのプログラムを作りました。これを「Liberated Learning」と呼んでいます。音声認識技術を使ったものです。いくつかの仕様アプリケーションがあります。

1つは、電話です。デジタル化した音声で、ブラウジングするものです。また、バイオメトリックという生物学的、身体の部分をつかったものもあります。音声を使い、文字を作るのが音声認識の技術です。音声認識は、口述筆記、あるいは、E-mail、メモ、手紙と

いう、特にコミュニケーションのために開発されたものではありません。新しい言葉を増やすこともできます。話しかたを学ぶことによって音声を認識するものですが、ベースになっているのは統計です。コンピュータの音声認識のシステムは、いま何を話したかを解析するわけです。私はこの音声認識の専門家ではありませんので、細かい説明はできませんが、どんな仕組みになっているか簡単にご説明します。

何を話したかを文字にするためには3つのモデルが使われます。

最初に、システムは、この音がすでに自分の持っている辞書の中にあるかを検索します。そして次に「音響」。音を認識するわけです。音響認識をただで情報解析することができない場合には、「言語モデル」をつかいます。これは意味論、あるいは構文のモデルにあっているか、あるいは文法にあっているかを見るのではなく、単語のコンビネーション、組み合わせが正しいかどうかを見るわけです。音声認識は非常にむずかしい技術です。まず何が難しいかと言えば、ひとりひとり話しかたは違っています。ゆっくり話す人もいれば、非常に声の小さい人もいます。大きな声の人もいれば速く話す人もいます。また、同じことを言おうとしてもそれぞれ人によって話し方が違います。たとえば、発音されない音もあります。また、子音と母音の並び方で音が違ってくこともあります。また同じ意味のことを言ったとしても、話し方は違うかもしれません。たとえば、英語を第二外国語として学ぶ人にとって「ハロー(こんにちは)」が幾とおりもあるということを理解するのは、なかなか難しいでしょう。

コミュニケーションの多くの割合は、発語されない言葉で伝わります。これは本当に驚くべきことです。しかしながら音声認識は、発語されていない物に対しては理解することができません。内容を理解するということは、できないわけです。

また、音声がきちんと発音されているかというのも問題です。この音声、言葉というのは音の波の続きです。たとえば、私がこれからやりますが、言葉を全部続けてしまってスペースを間に置かない場合は非常に理解がしづらいでしょう。また、特に授業中ではこれが顕著です。たとえばこの音声認識に対してスタートレックのような期待を持たないことが重要です。皆さん、スタートレックというテレビ番組は知っていますか？キャプテン・カークは知っていますか？スタートレックでは、スターシップ・エンタープライズという乗り物があります。そして宇宙人がそれに乗っています。乗組員はコンピュータを介してコミュニケーションを行ないます。キャプテン・ジョン・ルーカ・カークは、この番組のヒーローではありません。まさにコンピュータがヒーローです。彼らが話すときにはコンピュータが反応して、自然なコミュニケーションが図られます。そして失敗は一切ないわけです。

しかしながら、実際のコミュニケーションの中ではそうはいきません。コミュニケーション

の中では、いろいろな言葉、新しい言葉、古い言い回し、人名、地名、難しい専門用語などがあります。これらの言葉をすべて完全に音声認識できるとは言えません。たとえば大学の講義のなかで、特定の言語を扱っている、たとえば異常心理学、有機化学分野など、こういった特別な分野では専門用語が使われます。しかし、音声認識の基本となっている用語の中には、これらの専門用語は入っていません。また、小さな教室では、聞いている人たちに対して教授は質問をし、対話形式で授業が進められます。音声認識のプログラムは、対話というのを学ばせないかぎり理解することはできません。セントメリー大学では、1998年に口述筆記を音声認識で実施する方法を取り入れました。

こういった口述筆記をベースとした技術を使うときには、カンマやピリオドが必要です。それをするためには、言わなければいけないわけです。「ここでカンマ」とか、「次のセンテンス」と言葉で言わなければならない。しかし自発的に話すときにはそういうことは言いません。たとえば家に帰る時、いろいろ考えている時にいちいち句読点を打つでしょうか。普通に話していて、句読点をわざわざ言うのは難しいことです。これがこの商業ベースの音声認識システムを使った場合の口述筆記です。もちろん理解はできますが、授業の中では使えるものではありませんでした。

もう一つの課題ですが、最も大きな課題とっていいでしょう。

「私はフロアに行きました」という言葉です。

このようにしゃべった言葉のなかで、1つだけ間違えているということで、精度 80%となります。この文章はあいまいなわけです。いろいろな理解の仕方があります。

たとえば、お店に行った。ドアのところに行った、あるいは海岸に行った、というふうに考えます。どこに行ったかははっきりわからないわけです。「わたしはお店に行きました」という言葉。お店への to という言葉が間違っています。ほとんどの人がこれを聞いて充分理解できます。もちろんあいまいさということで言えば、前の文章よりもあいまいさは、低いわけです。このように、理解に大きな影響が与えられるわけです。

これは小さな問題ですが、もっと大きな問題の場合には、その影響も大きくなるでしょう。技術の話は少しいたしましょう。これからデモンストレーションを行ないます。IBM の NetScribe というものです。このアプリケーションは、Liberated Learning に関連して開発されたものです。これによって情報へのアクセスが向上できると思っています。特に身障者の方々に対してのアクセシビリティが広がると考えています。

この音声認識のソフトウェアによって、口頭で話した、また大学の講義で話した内容へのアクセスが広がります。この音声認識の技術を介して、わたしが話した言葉がスクリーン上に出てきます。また、精度の高い、教室内で使えるような文字に変換されていきます。たとえば私が息をしたり、止まったりすると、自動的に行が変わります。意味論で

ひとつの意味があるものは、固まって出てきます。この意味の固まりは、例えば、文章になるわけです。実時間で読めるような文字が出てくるわけです。今見ている画面は精度が高いものです。まず、大事なことは教授がスクリーンを見ないということを確認します。

文字をかいている時は、文章は精度が高いですが、自然に話すときは必ずしもそうではないです。繰り返したりしますし。神経質な時は顕著です。この技術は非常にユニークな形で情報を提供でき、例えば、文字と音素を一緒に表すことも可能です。音声認識の信頼性をあげることによって、文字と音声を認識し、今までは正しく理解されなかったことも理解することができます。音素ですが、これを使うことによって、聴覚障害のかたでも口の動きを見ることが可能になります。読唇術を使って、理解を補うことができます。

NetScribe のテストを夏に行ないました。ラップトップコンピュータ、PDA などを使う人は、誰でもインターネットを介してアクセスできます。情報を個人化して、自分たちが学習する内容を自分の好みに変えることができます。すべての人に使ってもらえます。それぞれの人は、学習の仕方が違います。色にも、文字の大きさにも好みがあります。学生達に、この Liberated Learning の理念を理解してもらうことでこのツールを利用してもらっています。使う側からすると、どうしても技術的な難しさゆえに、利用することを躊躇してしまう傾向がありました。これは、できるだけ優しく、ネットワークを介して、教授の音声のプロフィールを理解させる、特別なハードを使わなくても認識できるようにすることに力点を置きました。技術によってそれらを解決し、新たなドアが開かれたといっ

ていいでしょう。

NetScribe に関して少し説明すると、NetScribe は、3つのマルチメディアのファイルを生

成します。

1番目は、文字による転記。

もし、認識が間違っていれば、後で修正可能です。

2番目は、完全なデジタルオーディオファイルが作れます。メディアにアクセスできます。また、文字と音声のファイルによって、可視化できる、ビジュアルなフォーマットをあわせて使えます。

3番目は、講義のあとにオフィスや自分の部屋で、あとからこのファイルにアクセス可能です。

このアプリケーションで情報に対するアクセスが広がります。このように、書かれたテキストと音が連携しています。実用化されているもの、テキストと聴覚情報を一体化するものもあります。煩雑なプロセスを経て、専門の技術がないと使えないではなく、全部自

動的にやってくれるからです。また、視覚的なメディア、プラステキストと聴覚情報を一体化できます。それでは実際のデモをお見せします。

音声認識について話します。

テキストと聴覚情報が同期化し、一体化しますが、プラスPowerPointのスライド、ビデオなど、目で見える視覚情報も入ってきます。プレイバックし、どんな状況かみてみましょう。テキストは小さく下に表れます。話した情報を録音して、再度巻き戻しているような状況です。講義の中で視覚的メディアも使い、講義が終わると、遠隔学習のような形で、テキストと聴覚情報と視覚的メディアを一つのものとして捉え、インターネット上でクリックすると、すぐアクセスできます。他にも、この技術の応用分野があります。言語の翻訳です。今日は皆さんに日本語でお話できず申し訳ありません。日本に着いたら、みんなと話ができないと、バリアを感じました。ですから、この技術を使うことで、最終的に、英語で話したものが、日本語の文字になって、画面に表示されるのも実現できたら素晴らしいと思います。オンライン教育が急速に発展しています。遠隔学習です。データを調べることができます。ビデオテープへ字幕をつけることもできます。ウェブキャスト、ビデオなどで、字幕のない情報がインターネット上にありますが、音声認識のシステムを使えば、エラーがあれば直し、アクセスしやすい媒体に変えられます。

3年間の実績の研究成果をまとめると、学生達は大変役立つと考えており、実用的コメントをもらいました。「綴りがわからない」「ノートの中で聞き落としを補う」「講義が聴けないときは、改めて聞いてみる」とか、大変な可能性を持つ技術だと思います。

教授のほとんどは、たしかにこのソフトウェアを勉強してセットアップしなくてはならないが、手間はかかるが、よりよい授業に役立つ。つまり、言葉の使い方、情報の伝え方など、焦点をもった使い方ができます。問題は、正確性と読解性です。

1時間の講義に対し実際の環境の中で、評価しました。例えば、今やっている環境での評価をしました。早口になれば、正確性が落ちます。気が散っていたり、はっきり言おうとして、かえって正確性が落ちることもあります。もちろん、問題は正確性をあげることですが、編集作業を減らすことです。一万語のなかで、90%が正しくても、10%が間違っているのは、それはそれで大変な作業です。選択の余地が広がるディスプレイをカスタム化する必要もあります。セントメリー大学の教授とIBMは、大学のグループをつくり、この技術をより高度な技術として定着させようとしています。セントメリー大学の実績を踏まえて、アメリカ、オーストラリアに広げてきました。オーストラリア教育省から予算化され、その結果、いくつかの大学がネットワークでつながり、協力できるようになりました。お互いに研究しながら、協調していこうと考えています。

日本IBMを日本での拠点として、日本の大学とつなぎ、同じようなネットワークを作りたいと思っています。そのためには予算が必要ですので、どうなるかは今のところわかり

ませんが、努力していきたいと思っています。今後の計画として、ネットワークのアプリケーションのプロトタイプが出来るように、夏に取り組みを行ないたいと思います。また、事前に話者の声の登録をすることなく、不特定の人声が認識できること、また、音声認識機能が付加されたポータブルな装置で実現することができたら、持ち歩きながら誰とでも話すことができます。聴覚障害者のコミュニケーションの向上が図られると思います。あるいは、英語を第二外国語として学習している人のための、研究プロジェクトがあります。この技術が、第二外国語として学習できるような道具にできたらとも考えています。中心となる日本の大学で、時間、お金、エネルギーをかけて、チームを作ろうという考えのある大学と協力し、基本的知識を持つ関係団体と協力し、日本 IBM の技術を使って開発し、試験し、研究し、日本の市場にあったものを開発していければ素晴らしいと考えています。