

## …平成8年度助成研究より…

# 聴覚障害者のための音声情報の 色呈示方式の開発

東北大学工学部  
助教授 工学博士 木 幡 稔

### 1. 聴覚障害者の現状と発話訓練の必要性

障害者や老人に対する福祉の問題は、今日マスコミ等でも大きく取り上げられるようになり、研究分野として福祉工学という新しい分野を生み出すに至っています。日本では障害者福祉や高齢者福祉が欧米に比べて立ち遅れているといわれますが、本研究で聴覚障害者の方たちとお話をする（実際は手話）機会を得て、現在の日本では聴覚障害者の方が不便を感じる事が極めて多いことを知ることができました。

さて、聴覚障害者は障害の重さにもよりますが、一般に発声にも障害を生じます。これは、聴覚によるフィードバックにより調音運動を制御し、発声することができなくなるからです。このことは、意思伝達の手段を失うことになり、心理的に大きな負担となるとともに、緊急時には大きな問題となります。

重度の聴覚障害者の方々は、意志伝達の手段としては手話や筆談を用います。手話は健聴者側にも訓練が必要ですが、ところが日本では手話を理解できる人はまだわずかであり、その結果、手話は聴覚障害者同士での意志疎通の手段に限られてしまいます。また、筆談は伝送速度が遅いことと、筆記用具等を要することから、限られた状況下でしか用いられることはないようです。

このような不便さを何とか解消することで聴覚障害者の方々の意志疎通の幅を拡大し、多くの人々と交流できるようにすることはできない

か、と考えたことが本研究の発端でした。聾学校には、実際に発話訓練の課程がありますが、舌や下顎の位置、唇の形に関する視覚的な指示をたよりに調音するため、最適な調音方法を会得することは非常に困難です。すなわち、健聴者と同じように何らかのフィードバックが発話訓練には必要と考えられます。そこで、まず視覚によるフィードバックを利用することに着目しました。そして、パーソナルコンピュータのディスプレイ上に発声した音声記号の特徴量をリアルタイムで表示し、視覚的なフィードバックを行うことで調音の修正を行える発話訓練装置を開発することを目的として本研究はスタートしました。

### 2. パーソナルコンピュータを用いた発話訓練装置

パーソナルコンピュータを用いて聴覚障害者の発話訓練や視覚・触覚などによる音声認識、また健聴者が英語などの言語訓練を行うシステムは、実はずいぶん前から研究されてきました。特に視覚・触覚などにより音声知覚を行う方式は「聴覚代行」と呼ばれ、さまざまな方式が試みられてきました。しかし、聴覚と他の感覚との決定的な性質の違いによって、使い勝手のよいシステムは実現しませんでした。その性質というのは、人間の五感の中では聴覚は最も時間的変化の激しい信号を取り扱わなければならないということです。例えば日本語の場合、1秒間に数モーラの速度で発声されますが、そうす

ると1モーラ当たりの持続時間は100ms 余りとなり、さらにそのなかでも子音の持続時間は10ms程度である場合があります。こうした素早い変化をする信号を画面に呈示したり、振動や電気刺激により触覚的に呈示したりすることはもちろん可能ですが、聴覚障害者が発話を訓練するための適切なフィードバックにすることは難しいというわけです。

従来までの方式の多くは、音声信号を帯域フィルタやFFTを用いて周波数分析し、ホルマント周波数やピッチ周波数を抽出して視覚や触覚に呈示するものでした。これは、人間の蝸牛での周波数分析のメカニズムをそのまま他の感覚に置き換えたものと考えることができます。しかし、こういった音声信号のスペクトル表現は健聴者の聴覚神経系において処理される場合には最適なフィードバック効果であると考えられますが、視覚や触覚を用いる場合にはもっと他の適当な（発話訓練に向けた）表現方法があるのではないか、という疑問を持ったわけです。

さて、以上のような経緯でいろいろな方式を

検討した結果、統計モデルによる音声の表現方法の1つであるHMM（隠れマルコフモデル）を用いた方式を提案するに至りました。HMMは近年飛躍的に向上している音声認識の基礎となる技術です。

HMMは図1に示すように、単語、音節、音素などを単位として、その音響特徴パラメータ（例えばケプストラム）の時間変化を、マルコフモデルにおける状態遷移で表すモデルです。音声認識においては、この複数個のHMMに対する尤度を計算し、最大になるものを認識結果とするわけですが、そのためただ1つのHMMの尤度が支配的に大きくなることが望ましいわけで、またそうなるようにHMMを設計しなければならないわけです。しかし不特定話者に対して十分な認識率が得られるようにHMMを設計することは非常に難しい、というのが現状です。本研究では、このHMMの出力尤度がいまいになって認識率を低下させるという性質を逆に発話訓練の情報として利用することに着目しました。つまり、似通った音響的特徴を有する音声に対応するHMMの出力尤度を訓練目標の音声

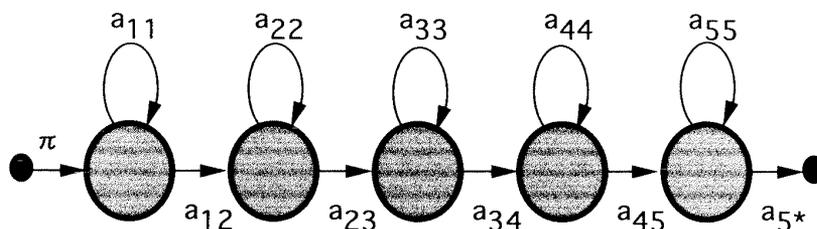


図1 HMM (Hidden Markov Model, 隠れマルコフモデル)

に対応する HMM の出力尤度とならべて表示することにより、調音方式の類似性や差異を対比的に会得することができると考えました。(例えば、/a/, /pa/, /ta/, /ka/, /ha/) また、訓練目標の音声に対する尤度を向上させるという単純な目標にすることによって、ホルマント等を直接呈示する方式よりも訓練効率が向上するのではないかと考えたわけです。

HMM のプログラムの実行は処理速度の要求から、従来までは研究室のワークステーションで行うのが一般的でした。しかし、近年パーソナルコンピュータの性能が著しく向上し、ワークステーションと大差ない程度になり、HMM のプログラムの実行も十分に可能になりました。また、聴覚障害者がいつでも訓練できるためには、自宅にシステムを設置する必要があります。そこで、この訓練システムはパーソナルコンピュータ上で動作可能なアプリケーションとして開発することにしました。

### 3. 色情報のコンフィギュレーション

HMM を設計するプログラムと、音節を訓練の単位として、入力音声の尤度を算出するプログラムはすでに研究室のワークステーション上で実現していましたが、難しい課題が残っていました。それは、パーソナルコンピュータへのプログラムの移植と、HMM から出力される尤度をどのように画面上に呈示するかという問題です。まず、移植の問題に関しては Windows95 上でのアプリケーションとして移植することにしまし

た。また、HMM の設計は計算量とデータ量が非常に大きな処理となることから、ワークステーション上で行い、HMM のデータだけをパーソナルコンピュータに転送して用いることにしました。

第 2 の問題は非常に難しい問題で、この訓練システムの性能を左右する問題であるともいえます。HMM から出力される尤度は単なる数値データで、これを 2 次元の制約された画面上にどのように呈示するのが効果的かということについて色々と我々の研究グループで議論しましたが、色情報を効果的に用いることが有効ではないかということになりました。具体的には、尤度をバークラフで画面表示するとともに、音節を色を用いて識別する方式です。ただ、110 個の音節すべてに適当に色を割り当てたのでは、訓練効果は期待できませんから、先に述べたように音節の特徴の類似性(調音様式、後続母音など)を考慮して、特徴の近い音節には同系色を配色する必要があります。この問題に関しては

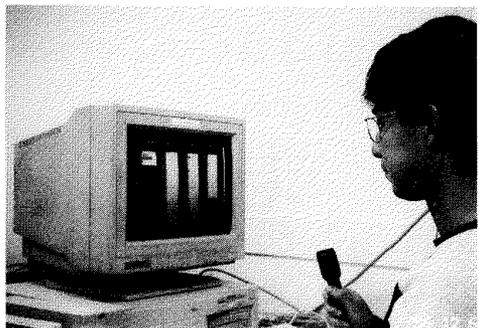


図 2 開発中の発話訓練システム

今後、実際に聴覚障害者の方に訓練システムを使っていたがき、実験的に検討し、改善していく予定です。(図2は開発中のシステム)

#### 4. 障害者と福祉、我が国でのありかたについて

我国の障害者福祉は欧米諸国に比べ立ち遅れているということを冒頭で述べました。このことは施設や介護器具、料金や税制上の優遇など物質的な側面だけについて評価されがちですが、実際に大事なことは障害者に対する我々の接し

方ではないかと思います。例えば、本研究の目的とは矛盾するかもしれませんが、多くの健聴者が手話を理解できるようになれば聴覚障害者の交流の場は大きく拡大するはずです。現在の日本人にはそういった時間的、精神的余裕がないことも理解できますが、将来は一人一人が障害者のために自分ができることを実行できるようになって欲しいと思います。最後になりましたが、本研究に対しサウンド技術振興財団から多大な助成を頂きましたことを感謝致します。

