

研究題名	聴覚障害者のための音声情報の色呈示方式の開発	報告書作成者	木幡 稔
研究従事者	木幡 稔		
研究目的	<p>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る。聴覚障害者は障害の重さにもよるが、一般に発声にも障害を生じる。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ある。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になるとともに、緊急時には大きな問題となる。</p> <p>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を開始した動機である。聾学校には、実際に発話訓練の課程があるが、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ある。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る。そこで、視覚によるフィードバックを利用することに着目し、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信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本研究の目的とする。</p>		

研究内容	<p>従来までの発話訓練方式の多くは、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ある。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る。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るが、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考えられる。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った。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ある。</p> <p>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ある。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。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、というのが音声認識における現状である。本研究では、このHMMの出力尤度があ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た。</p>
------	---

研究内容

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。

HMMのプログラムの実行は処理速度の要求から、従来まではワークステーションで行うのが一般的であった。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った。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る。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た。

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たが、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課題が残されていた。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る。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は、色情報を効果的に用いることが有効ではないかと考えた。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

研究内容

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音節間の類似度と非類似度という独自の3つのパラメータに圧縮して表示する方法を提案した。（詳細は報告書に述べる）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だき、実験的に検討し、改善していく予定である。（図2は開発中のシステム）

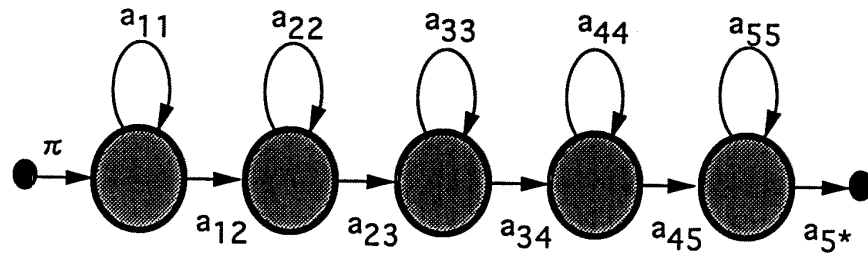


図1 HMM (Hidden Marcov Model)

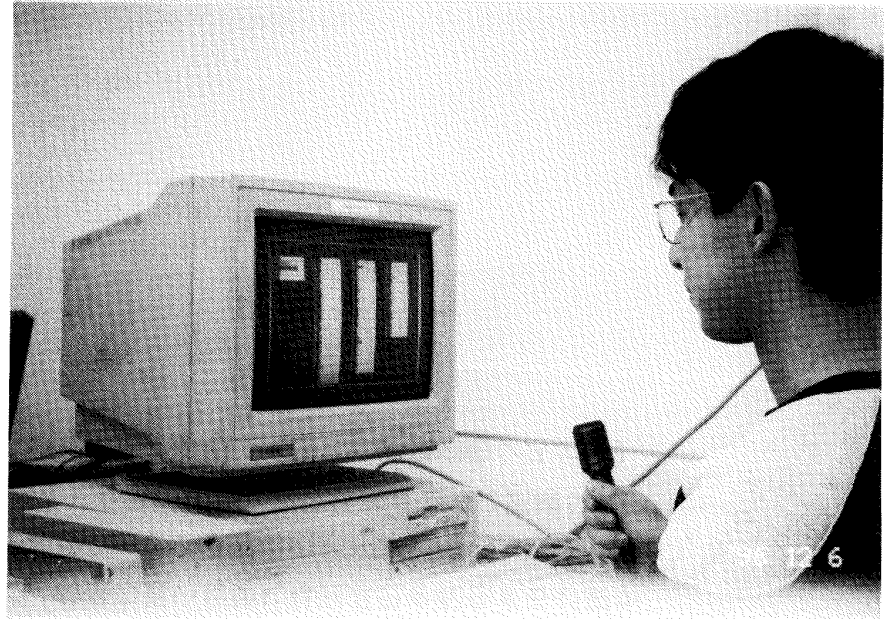


図2 開発中の発話訓練システム

(注： フローチャート図，ブロック図，構成図，写真，データ表，グラフ等 研究内容の補足説明に御使用下さい)