

聴覚障害児用 発声練習システム 「あいちゃんの手」

新事業開発室	中村 敬和 Yoshikazu Nakamura
新事業開発室	昆 昭彦 Akihiko Kon
山武コントロールプロダクト(株)	青木 功 Isao Aoki
山武テクノシステム(株)	浅輪 晃一 Koichi Asawa

キーワード

マイクユニット, アクティブ・ハンド, インジケータ

聴覚障害のある子どもは、他人の声と自分の声を耳で聞いて比較し、胸/喉/口/舌などの筋肉を運動するフィードバックの成立が困難なため、明瞭な発声や発音がむずかしい。そこでこのフィードバックを補完しながら、発声や発音を練習するシステムを開発した。画面やインジケータなどの視覚と手の平の触覚を利用し、生徒が楽しく学習できる点が従来にない特徴である。

The "Ai-chan no Te" Speech Training System for Aurally Impaired Children
Aurally impaired children have difficulty in comparing their own voices to those of others and in obtaining feedback on how to move the muscles of the chest, throat, mouth and tongue. This makes clear speech and pronunciation difficult to achieve. The authors developed a speech and pronunciation training system that complements such feedback. By utilizing visual aids such as screens and indicators and the sense of touch in the palm of the hand, the system provides greater enjoyment in learning than conventional methods.

1. はじめに

当社の創立90周年にあたる1996年に山武グループでアイデアコンテスト“クリエイト21”が行われ、触覚フィードバック機能を持つ「聴覚障害者用発声練習機」が大賞を受賞した。既存の練習機ではフィードバックと指示は画面上に描かれるデータのみであり、また計測センサーも扱いにくいものであった。これら既存の練習機の改良と受賞した練習機のアイディアを基にして新しい発声練習機の開発を行なったのでそのシステムの概要を紹介する。

2. システムの概要

2.1 システムコンセプト

聴覚や言語に障害のない子供は、幼児期に他人の声と自分の声を聞いて比較するとともに、胸、喉、口、舌などの筋肉を運動させて話すことができるようになる。音声学習

においてフィードバック系を活用している。一方、聴覚障害のある子供は音を(自分の声も含めて)捉えることができにくいため、このフィードバックによる学習が困難になり、明瞭な発声や発音がむずかしくなる。

指導者は早期にこの状態を発見し、聴覚(所有聴力)の最大限の活用と、筋肉運動感覚や教具を利用した言語指導を行い、コミュニケーションが行えるようにする。今回開発したシステムは、総合的な学習環境を目的とするとともに視覚と触覚さらに聴覚のフィードバック系も活用しながら生徒が楽しく学習できるシステムとした。(表.1参照)

最終的にシステムコンセプトならびにユーザ対象練習項目を次のように想定した。

目的	視覚と手の触覚を利用し聴覚も活用しながら、生徒が楽しく学習できるシステム
対象	5歳から12歳を中心に、指導者とともに練習
練習項目	氣息、発声、母音、子音、韻律

表.1 従来の機器との比較

比較項目	あいちゃんの手	従来の機器 / システム
対象 / 目的	生徒と指導者のための 指導 / 定着 / 分かり易い評価	指導者のための 正確な計測・評価
練習プロセス	息・声・母音・韻律・ 子音（開発中）	声のみ、または 息・声・単音節
プラットフォーム	DOS/V-PC + Windows95	専用機、または DOS-PC + 専用ボード
センサ	専用マイクユニットのみ	マイクのみ、または マイク・呼気流センサ・鼻センサ・ 声帯センサ・舌位置センサ
フィードバック	画面：ゲーム インジケータ：4音 ハンド機能（手の触覚）	画面：グラフ表示，ゲーム インジケータ：単独N型，S型
価格帯	未定（数十万円）	100～200万円 単独インジケータは十数万円

2.2 「あいちゃんの手」特徴

開発した練習システムの特徴は次のようになった。

(1) 生徒に負担をかけない

センサ部分はマイクユニットのみとし、生徒に過大な負担がかからない。

(2) 感覚を重視する

視覚的には文字を極力避けた画面，LED表示のインジケータ，手の触覚にタイミングやリズムを伝えるハンド機能（一般的にはタクタイルエイドと呼ばれる）を利用し，多感的に生徒が理解できる。

(3) 楽しく学習できる

アニメーションを用いてゲーム感覚で、生徒が興味を持ち，楽しく練習できる。繰り返し練習ができ，練習毎に何らかのインセンティブを与える。

(4) 生徒に合わせた練習が可能

練習メニューは個々に選択でき，また練習の中でモード変更や調整により，生徒の状況や進捗に合わせた練習ができる。

(5) 容易な操作

操作パッドのみで練習できるようにするなど，パソコンの知識がほとんど無くても使える。また，各練習にお手本をつけ，練習の方法がわかる。

システムの名称は，母音練習の最初に登場する「あ」「い」と，ハンド機能から「手」を取り，聴覚障害児用発声練習システム「あいちゃんの手」とした。（図.1参照）



図.1 システム外形図

2.3 システム構成と各機器の概要

具体的なシステムの構成として，マイクユニット，インジケータ，アクティブ・ハンド，ソフトウェア(パソコン)，操作パッドの5点で構成される。（図.2参照）

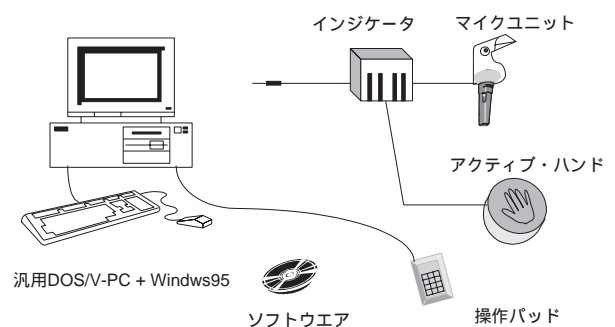


図.2 システム構成図

- (1)マイクユニット：有声音*，無声音*(息)，通鼻音*などが一体で計測できる構造である。
- (2)インジケータ：有声音，無声音(息)，通鼻音，摩擦音*(S音)の4点を同時にLEDレベル表示できる。また，アクティブ・ハンドの状態ランプ表示とマニュアル操作が可能である。
- (3)アクティブ・ハンド：ソフトウェアのゲームと連動して，タイミングやリズム，音素の指示や，ゲーム中の音声処理による応答を，接触子(トントン)や振動子(ブルブル)で手の平に伝える。また，お手本や自分の声を感じ取ることができるための手の平用のスピーカが付いている。
- (4)ソフトウェア：現状で24メニューである。(子音は開発中に付き含まず)基本的に各練習1画面でお手

本と練習が完結する。お手本には音声がついている。また，生徒が親しめるようにするために起動画面(音楽付き)がある。(表.2参照)

- (5)操作パッド：メニュー選択，母音選択，練習操作などの操作を，キー上の絵で選択して操作できる。

(補足説明)

- *有声音：声帯を震わせる母音や濁音などの音
- *無声音：口/舌などの構音器官に息を通して出す音(k, s, tなどの子音)や息(インジケータでは主に息)
- *通鼻音：唇や舌で口腔をふさぎ鼻に声を通す音(m, nなどの子音)
- *摩擦音：無声音の内，唇や舌と硬口蓋の間で息を摩擦させて出す音(インジケータでは「し」の などの子音)

表.2 ソフトウェア練習メニュー

練習項目	練習メニュー	目的
いき(氣息)	シャボン玉	継続して息を出す
	ケーキ	断続的に息を出す
	風船	息の強弱を一定にする
こえ(発声)	UFOキャッチャー	継続して声を出す
	ドーナツ	断続的に声を出す
	うさぎ	一定の大きさで声を出す
	ちょうちょ	一定の高さで声を出す
	自動車	長短のリズムで声を出す
	口形図	口形図での導入練習
	レース	母音を長く発音する
ばいん(母音発音)	ダーツ	断続的に母音を発音する
	魚つり	一定時間内連続で発音する
	自動車	長短のリズムで母音を発音する
	もぐら(あおう)	あおうを連続して発音する
	もぐら(あえい)	あえいを連続して発音する
	もぐら(5母音ランダム)	5母音をランダムに発音する
いんりつ(韻律)	ホルマント	ホルマントグラフ上での発声練習
	高低の練習	アクセントとリズムの基礎練習
	2拍語の練習	アクセントとリズムの練習
	3拍語の練習	アクセントとリズムの練習
	4拍語の練習	アクセントとリズムの練習
	学校で使う言葉	会話の速さとタイミングの練習
	家庭で使う言葉	会話の速さとタイミングの練習
しいん(子音発音) [開発中]	買い物で使う言葉	会話の速さとタイミングの練習
	子音	一部子音単音節の練習

3. 技術内容の説明

3.1 マイクユニット、インジケータ、アクティブ・ハンドについて

3.1.1 マイクユニット

従来の発声発語訓練装置用マイクは、有声音用、摩擦音用、通鼻音用、無声音用と個々に検出マイクが必要であり、操作が複雑、検出位置が不確定、個別音検出は可能だが同時検出がむずかしいなどの問題があった。

有声音/摩擦音、通鼻音、無声音検出用の複数マイクを一体構造のケースに組込んだマイクユニット(図.3参照)を今回開発した。この構成を採ったことにより簡単操作で、同時検出が可能になった。また、マイクユニット外枠前面に鼻下突当て基準部を設けることにより、口から検出マイクまでの検出距離が一定になり、検出条件を安定化できた。

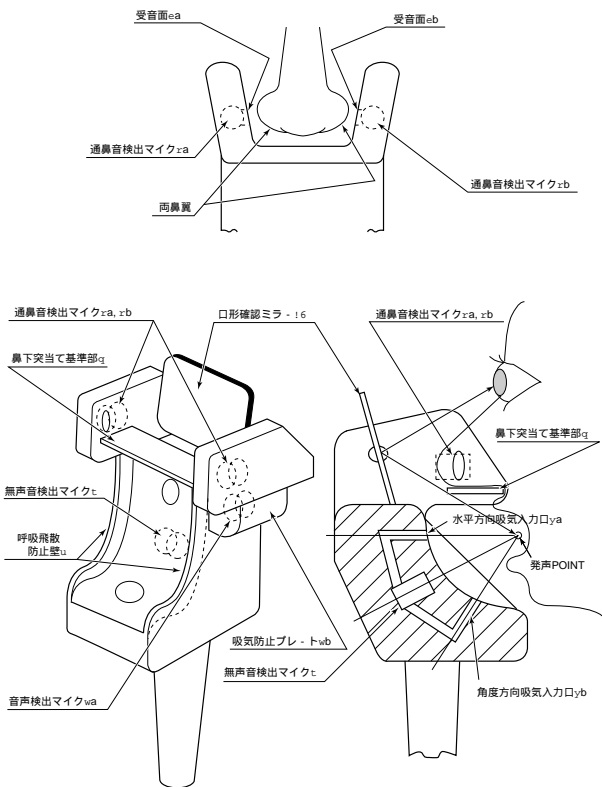


図.3 マイクユニット

有声音/摩擦音検出部は、検出マイク a を口前面ではなく右上に設置し、さらに呼気防止プレート b を設け、呼気の突入を遮断する構造とした。

通鼻音検出部は、鼻下突当て基準部で有声音と通鼻音を分離させ、通鼻音のみを検出する構造とした。両鼻翼振動を非接触で検出する通鼻音受音面 a, b を設けることにより、鼻下突当て基準部と両鼻翼部の検出距離を固定で

きる構造とした。さらに、検出マイク a, b は鼻息が直接突入しないよう両鼻翼の外側に設置することで左右の通鼻音を非接触で検出することが可能になった。

無声音検出部は、水平方向と角度方向の無声音(呼気)をマイクで検出できるように、呼気入力口 a, b を介して呼気流を検出する構造とした。また、鼻下突当て基準部から呼気入力口マイクまでの距離が一定になるよう、呼気受音面を発声点を中心に局面加工し、かつ呼気流が飛散しないよう両サイドに呼気飛散防止壁を設け、呼気入力条件が安定に計測できる構造とした。

発声口形確認ミラーを、マイクと一体にすることにより、発声前に口形を確認できるようにした。

3.1.2 インジケータ

従来製品には摩擦音検出用インジケータと通鼻音検出用インジケータがあり、ろう学校で広く使用されている。しかし高価で、検出位置を固定できず、摩擦音と通鼻音を同時検出ができないなどの問題があった。

今回開発したインジケータは、前記マイクユニットからの信号を使用し、有声音、摩擦音、通鼻音、無声音を同時にLEDレベル表示することにより、4音のレベル状態を一度で判断できる構造とした。(図.4参照)

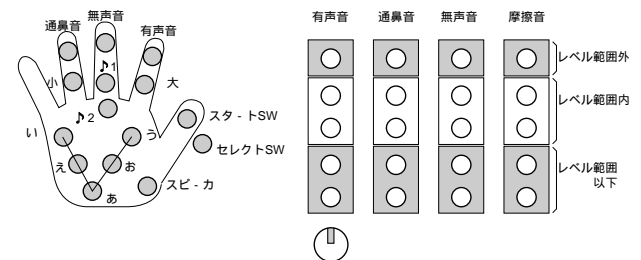


図.4 インジケータ

LEDはろう学校での表示色である赤色(有声音)、青色(摩擦音・無声音)および黄色(通鼻音)を使用して教育現場と矛盾なく理解しやすくした。

また、後記アクティブ・ハンドの状態表示ランプとマニュアル出力操作スイッチも一緒に設け、アクティブ・ハンド併用時の出力状態表示やマニュアル操作指示が簡単にできる構造にした。

このマイクユニット、インジケータそしてアクティブ・ハンドのみ用いてPCなしでも練習が可能である。

3.1.3 アクティブ・ハンド

今回開発したアクティブ・ハンドは、聴覚障害者に音声振動、音声情報(指示)、判定結果(応答)、発声タイミン

リズムを感知させるために、接触子（小型ソレノイド）、振動子（振動モータ）ないしスピーカを介して、左手の指先と手のひらに伝える「触覚利用の聴覚代行機器」であり、従来には無い装置である。（図.5参照）

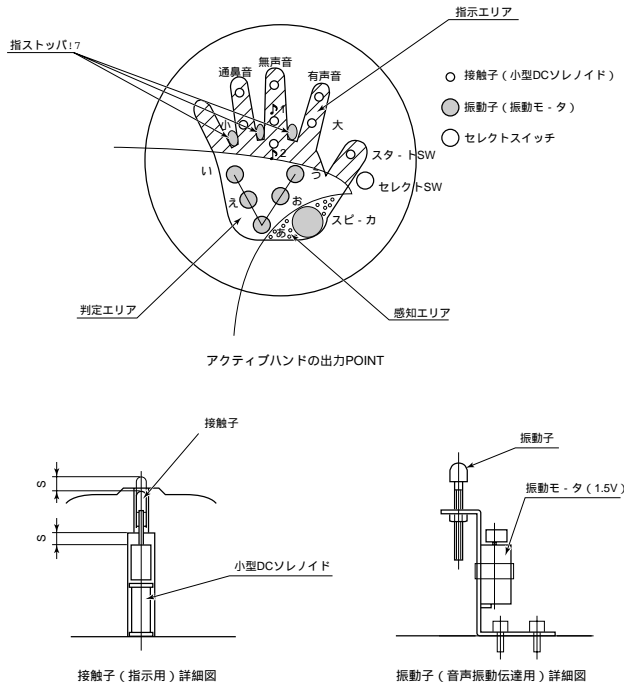


図.5 アクティブ・ハンド

手のひらの皮膚に触知させるために、手のひら部分を「指示エリア」「判定エリア」そして「感知エリア」の3ブロックに分けて、触覚の相互マスキング作用を防ぎ、触覚感覚を100%引き出せる構造とした。

練習対象者と想定した3歳から12歳までの児童の手のひらの寸法を数十人測定し、接触子ないし振動子の出力位置（出力POINT）を決定するとともに、アクティブ・ハンドに正しく手のひらが合わせられるように指ストップを設けている。

接触子と振動子の皮膚圧レベルは、各出力POINTの皮膚感覚受容器（マイスネル触小体とパチニ小体）に、もっとも適切な皮膚圧が加わるよう考慮した構造とした。

アクティブ・ハンドには、母音である「あ」「い」「う」「え」「お」用の5 POINT、有声音・通鼻音・無声音用の3 POINT、大小の2 POINT、韻律用の「1」「2」の2 POINT、スタート用の1 POINTなど合計15個の指示、応答、確認出力POINTとさらにセレクトスイッチ、音声確認スピーカがある。これら15POINTはPC練習ソフトウェアと連動して発声タイミングや発声時間、リズム、音素の種類、音声処理の判定結果などを指示、応答することがで

きる。また、発声音と手本の音声をスピーカを介して皮膚感覚で実感しながら練習できる。

PC練習ソフトウェアとの連動において、短い発声音・指示・リズム・タイミングの皮膚刺激として接触子（トントン）を使用した。長い発声音の指示や練習中の音声処理による応答結果の皮膚刺激には振動子（ブルブル）を使用した。このことにより触覚子と振動子の2つの触覚の違いを明確に区別でき誤りなくアクティブ・ハンドからの指示、応答、確認信号をはっきり判断できる構造となった。

3.2 ソフトウェアについて

3.2.1 ソフトウェアの構成

開発システム環境の選択をするにあたって高性能ゲーム機などの専用機や、Microsoft社「Windows CE」を選択する可能性があったが、ゲーム専用機にしたとき掛かる開発環境の投資額や他製品の開発への活用可能性並びにパーソナル・コンピュータの多くはWindows 95を採用していることなどを考慮して最終的に練習機システムのOSとしてWindows 95を採用するとともにこのOSを基盤とした開発システムを活用している。

開発環境

OS	Windows 95
画面構築	Visual C++ Ver4.2, Director 6J
ロジック処理	Visual C++ Ver4.2, DIO計測制御ライブラリ

3.2.2 ソフトウェア構造

(1) 画面構築

練習システムの起動時に現れるオープニングには、オーサリングツールであるMacromedia社の「Director 6J」を開発環境としている。

Lingoスクリプトを用いて画面のイベント処理や画面全体の構築を行った。

Directorを使うことで音声ファイル/画像ファイルをドラッグ&ドロップで容易にアプリケーションへ組み込むことができた。（図.6参照）

オープニング以外の練習画面や選択画面には、Microsoft社の「Visual C++」を開発環境として採用した。ビットマップファイルをリソースで管理し、それをオブジェクトとしてアニメーション化することで、自由度のあるプログラミングができた。

また、パレットも256色と多彩な色使いでユーザに飽きさせない画面となった。



図.6 オープニング

(2) 音声処理

ウェブフォーム・オーディオ機能として Win32 API には下記の 3 種類のサービスがある。

- 高レベル・オーディオ・サービス
- MCI ウェブフォーム・オーディオ・サービス
- 低レベル・ウェブフォーム・オーディオ・サービス

「高レベル・オーディオ・サービス」は、Wave ファイルの再生のみで録音する機能を持たない。

「MCIウェブフォーム・オーディオ・サービス」は、Wave ファイルの再生 / 録音を行うことができるが、扱うことのできる音声データが Wave 形式のファイルしか扱えないため、マイクからの入力をバッファに取り込む処理とファイルに保存してから読むなどの処理を分けて実行しなければならない。また、「音飛び」を起こす可能性がある。さらに、録音中アニメーション処理部で CPU を長時間占有されてしまうと、Wave データの転送に支障が生じ部分的に音が欠落した Wave データを作成し結果的に音声認識が正確にできなくなる。

これらの事情から、上記 2 つのサービスは使用していない。

「低レベル・ウェブフォーム・サービス」を用いたのでウェブフォーム入出力デバイス (Wave データの入出力を行うドライバ) からの入出力は Wave ファイルを扱わずにメモリ上で処理を行うことができた。さらに、リアルタイム処理にも適しているとともにファイル I/O が発生しないので他のプロセスとの CPU の奪い合いも無くなり「音飛び」の問題もクリアできた。

結果的にアニメーションと音声の同期を実現させることができた。

(3) シーケンサ通信

機能概要

専用ボード (Digital Input/Output ボード) を使用して、各練習内容に合わせ小型シーケンサにコマンドの書き込みを行う。小型シーケンサはこのコマンドに従いアクティブ・ハンドを駆動する。

データ構造

各練習には、それぞれ練習番号が対応付けられ、その練習番号をバイナリ形式としてシーケンサに書き込みを行う。また、タイミングや指示などは、ビットデータ形式でシーケンサに書き込みを行う。

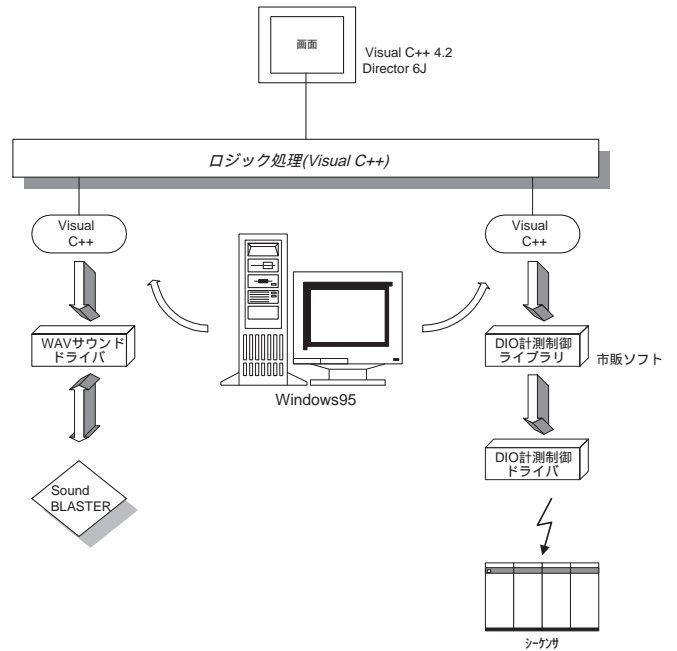


図.7 ソフトウェア構造

3.2.3 画面の構成

図.8にソフトウェア・システム全体の練習構成を示す。

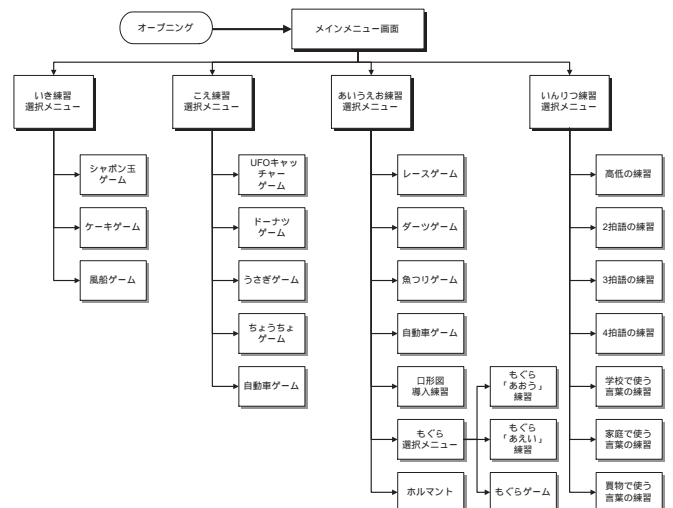


図.8 練習構成

3.3 音声処理技術について

3.3.1 音声処理の概要

マイクを通して音声データはたとえばデータ長 16 ビット, 22kHzのデジタルサンプリングで実時間でコンピュータに取り込み, ノイズカットや高域強調などの各種フィルタリング処理を行った後メモリーバッファに収める。引き続き, バッファ上のデータがある一定領域のデータに区切り, たとえば 30msec の幅のデータを, 前後領域で互いに重なりあうように切り出してこの 30msec の領域に存在する音声パワーや, 音声スペクトルを算出し, 音声時系列データを求める。この時系列データからさらに特徴あるデータを抽出してパターンマッチングなどの処理を行う。(図 .9 参照)

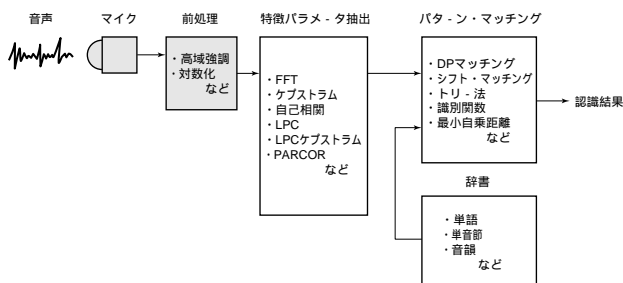


図 .9 一般的な音声処理 (1)

3.3.2 練習と音声処理

(1) LPC ケプストラムとホルマント

音声処理でとくに特徴的なことは, 高速フーリエ変換で音声スペクトルを求めるのではなく LPC ケプストラム(スペクトル周波数包括線のこと)を算出する方法を基本にしていることである。現時点の一点の音声データは過去に観測された数点の音声データの一次結合で現せるものとして音声モデル化している。このモデルに基づいて最終的に LPC ケプストラムを求めるのであるが, 音声上重要な周波数帯であるゼロから 6 kHz前後までの範囲を比較的正確にかつ高速に求めることができる。(図 .10 参照)

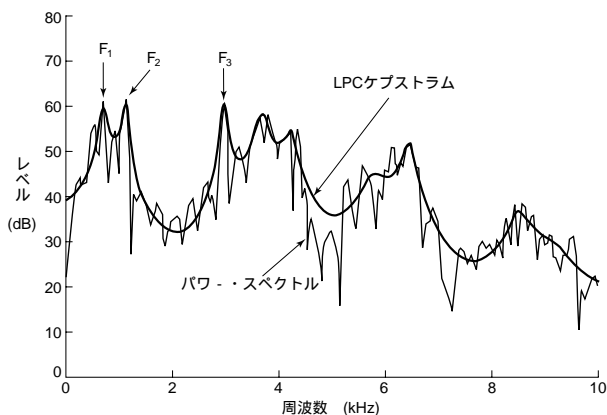


図 .10 LPC ケプストラム (2)

音声とは, 声帯で発生した空気振動に舌の位置と口の開き方などで摩擦音や破裂音の子音や母音に加工して発声するものであるが, スペクトルからこの音声を見ると共振している周波数の値や共振周波数の形が非常に重要なパラメータとなる。たとえば一番周波数の低い共振周波数のことを第一ホルマント F_1 , 次に低い共振周波数のことを第二ホルマント F_2 , その次を第三ホルマント F_3 などと呼びこれらは音声処理ではよく知られた特徴パラメータである。ホルマントは, 基本的には上で述べた LPC ケプストラムから簡単に算出できる。

(2) 母音の練習

母音に関しては第一と第二ホルマントの組み合わせだけでその多くは判別することができる(図 .11 参照)。このホルマント図上で各母音はある領域を占めているがこれは数多くの音声サンプルから母音の領域を求めたからであり, 一人一人の母音は特定な点として現れる。

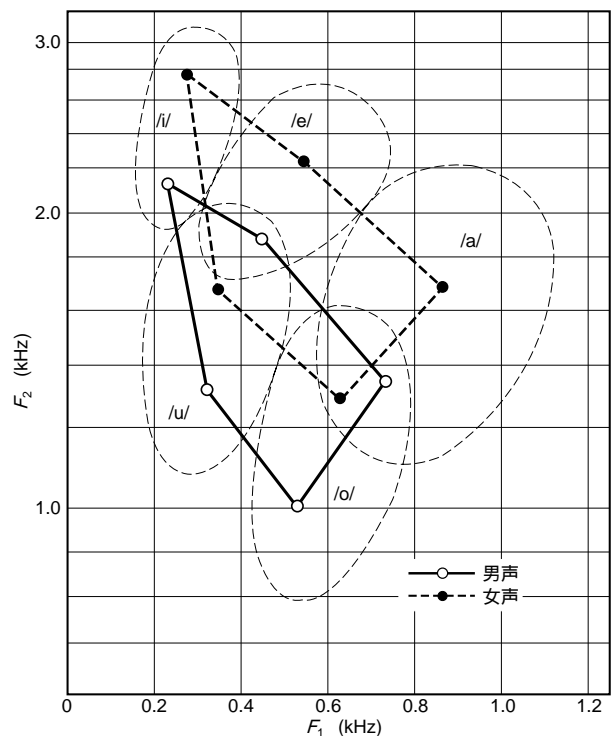


図 .11 ホルマント図 (3)

「あいちゃんの手」の母音の練習ではこのホルマント図を活用して母音の判別を行っている。ホルマント図には, 子ども, 成人男性そして成人女性のものを用意して練習する対象に分けて用いるが, このホルマント図は我々が子ども, 成人男性そして成人女性の母音音声データを独自に収集し解析して最適と思われるものを作成し使っている。

表.2のダーツの練習では、発声した母音がある母音のホルマント領域に入っているか否かを単純に見て判定しているのではなく、正確な発音すなわち各母音領域の中心に近いほど高い得点が得られるように内部処理を工夫している。

(3) 気息の練習

気息(いき)の練習では、マイクが捕らえた音圧レベル、具体的には音声パワーをみて正しく気息を出しているか判定処理を行っている。気息は通常音声処理では周波数成分がないものと仮定して扱っていたが検出することはむずかしい。また気息は周波数成分が含まれしかも音声の周波数帯と重なる領域があるため、単純な音声処理では音声と気息は分離できないこともわかった。

しかし気息と音声は音圧レベルがかなり違うことを音声実験で見出した。この知見から「あいちゃんの手」の内部処理を変更し最終的に気息と音声は容易に判別できるようになった。

(4) 声の高さの練習

声の高さ(ピッチ)の検出は、上で述べたケプストラム処理や音声パワーではなく、ある一定の時間内に計測した周波数、具体的にはゼロ周波数の軸上に何回音声クロスしたか計測するゼロクロス法でピッチ周波数を算出している。

声の高さの練習では、ある定められた範囲の間で正確に発声できるよう訓練する。

最低周波数と最高周波数の組み合わせを複数用意しており声の高い人や声の低い人に合わせて選んで練習できるようにしている。

表.2のドーナツ練習は、気息と、有声ならびに音声パワーの並行処理を行う複雑な内部処理で実現している。

4. 残された課題

現在の練習機「あいちゃんの手」の練習項目で実装していないものに子音の練習がある。子音の判別には単純な処理では対応することがむずかしいためさらに高度な処理である確率統計手法のヒドンマルコフ法⁽⁴⁾や非線型の時間軸変換法のダイナミックプログラミング法⁽⁵⁾その他の音声処理技術の検討を進めている。

我々はさらに使いやすい発声練習機を実現する予定である。

5. おわりに

聴覚障害児用 発声練習システム「あいちゃんの手」は、聴覚、視覚そして触覚などの多くの感覚器へのフィードバックを介して発音発声を楽しく練習するシステムで他に例が無いものとなった。本システムは、ろう学校で実際に授業で使われており教育効果がどの程度向上するのか検証中である。

6. 謝辞

本システム開発にあたり、国立特殊教育総合研究所菅原部長、国立筑波技術短大大沼教授、北海道大学伊福部教授、中部大学梅崎助教授およびろう教育の現場の先生やその他専門家の方々の協力によって初めて完成度の高い発声練習システムを開発できた。この場をかりて感謝申し上げます。

参考文献

- (1)～(3)伊福部 達 「音声タイプライタの設計」 CQ 出版
(4),(5) 中川 聖一 「確率モデルによる音声認識」 電子情報通信学会