

GUI を用いた聴覚障害児用発声訓練システムの開発

倉谷 和彦*, 梅崎 太造**

Development of the voice training system for auditory handicapped children using GUI

Kazuhiko KURATANI* and Taizo UMEZAKI**

(Received May 31, 1996)

We describe the voice training system for auditory handicapped children running on the MS-Windows 3.1 and 95. Because the calculations of speech analysis are done only using the Pentium CPU in this system, extra hardware is not needed. The mouse is used in all operations of the software. Therefore even beginners are able to play easily. Users are able to train the power control, the pitch control and the vowel utterance in the training of basic utterance. In this paper the features of this software are explained in detail.

Keywords : auditory handicapped children, voice training, MS-Windows, speech analysis

1. まえがき

聴覚障害児教育の重要な課題の一つに、発声訓練を可能な限り早期に実施することが挙げられる。その際、各個人の能力に合わせて個別指導することが望ましいが、実際には、専門知識を持つ熟練した教師の絶対数が不足しているという問題や教育予算の問題があり、実現することは不可能に近い。

これまで、上記の問題を解決する手段として、教育機関や企業等により、障害児の発声・発語訓練を支援する装置の研究開発が行われている⁽¹⁾⁻⁽⁴⁾。筆者らの研究室においても、学研(株)及び中央発條(株)と共同で聴覚障害児用の発声・発語訓練器「トーキングトレーナ」を開発して来た。本装置は、現在も各方面で活用されている⁽⁵⁾。

しかし、従来用いられている訓練装置は、演算速度の問題から、高価な DSP (Digital Signal Processor) を搭載する専用のハードウェアで構成されているものが多

い。8ビットや16ビットの安価なマイクロチップ (CPU) のみで、DSP を使わずに実時間の音声分析処理が可能な訓練器を設計することは、不可能である。メモリの容量も少ないため、特定の訓練しか行えない装置が多く、汎用性に欠けている。スイッチ等のユーザーインターフェースの悪さから、操作が極端に複雑なものさえある。

これに対し、最近の32ビットパソコンは演算能力が急激に向上し、その計算速度は、これまでの DSP に匹敵するものもある。メモリや外部記憶装置 (ハードディスク等) の容量も十分に確保できる。音声入出力等のマルチメディア機能も充実している。また、MS-Windows に代表される GUI (Graphical User Interface) 環境の操作性も著しく向上していることから、32ビット以上のパソコンを用いた汎用性と移植性の高い発声訓練システムを開発することが関係機関より強く望まれている。

本稿では、MS-Windows3.1 及び 95 上で動作可能な聴覚障害児用発声訓練支援システムについて報告する。

*電気工学専攻大学院生 (Graduate Student, Electrical Engineering)

**電子工学科 (Department of Electronic Engineering)

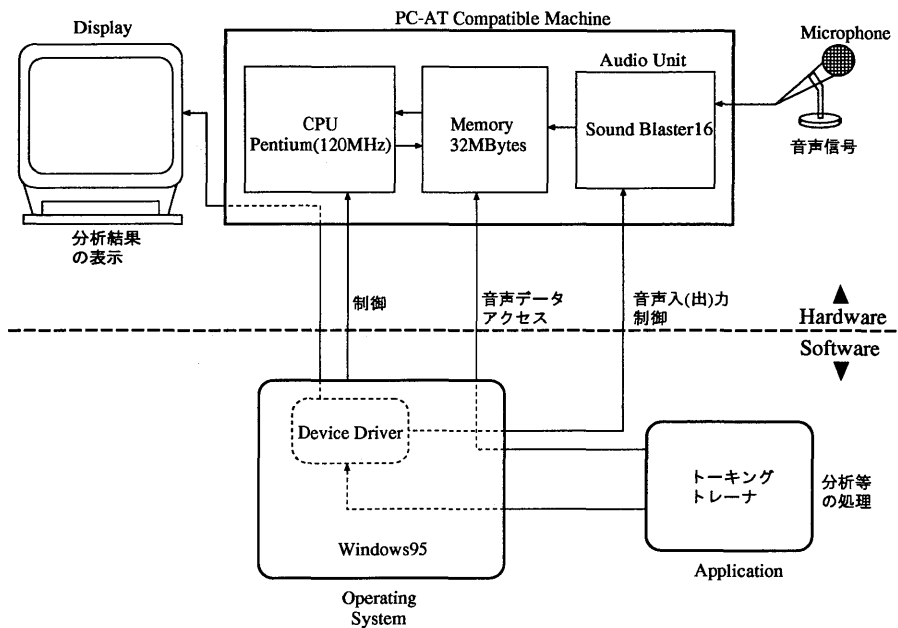


図1 システム構成の概要

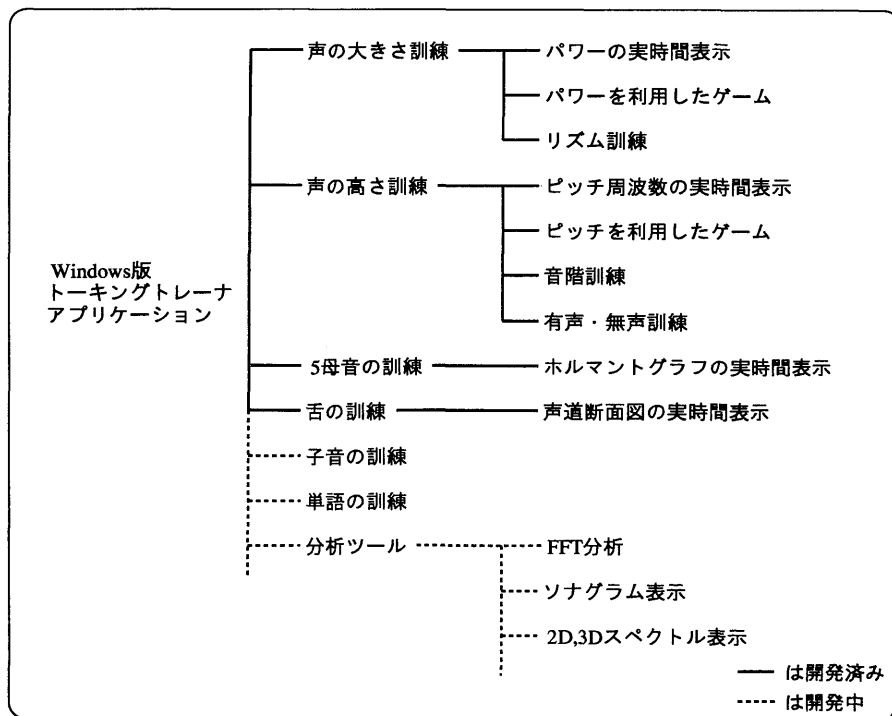


図2 アプリケーション内容

現在、声の大きさ訓練、リズム訓練、高さ訓練、及び母音訓練を中心とするソフトウェアを開発している。システムは、サウンドボードを搭載した DOS/V パソコン以外に特別な装置を必要としない。従来のシステムに比べてハードウェアに依存する部分が少ない。システム全体がソフトウェアのみで構成されているため、機能の拡張や変更が容易に行えるという特徴を持つ。

2. システムの概要

本システムは、図1のように DOS/V パソコン、サウンドボード及び、発声訓練ソフトウェアの3つの要素から構成される。訓練用パソコンとしては、Pentium CPU を搭載しているマシンを想定している。サウンドボードは、DOS/V マシンのサウンドボードとして最も広く利用されている Sound Blaster16 を用いるが、Windows 対応のサウンドボードであれば、問題なく使用できる。発声訓練ソフトウェアの動作は、現在のところ MS-Windows 3.1 及び 95 上のみで確認しているが、システム開発にマイクロソフト社の Visual C++ Version 4.0 を用いているため、Windows NT での動作も可能である⁽⁶⁾⁽⁷⁾。

3. システムの設計方針

訓練対象者が聴覚障害児（幼児、小、中学生）であるため、コンピュータにまだ慣れていない初心者の使用を第一に考える必要がある。故に、Windows の優れた操作性を活用することがシステム設計の基本方針である。

3.1 GUI 機能の活用

使い易さに重点を置くため、可能な限り GUI 機能を活用し、マニュアルレスでも利用できることを目指す。

1. フルマウスオペレーション

Windows 上の操作は、基本的にキーボード操作を必要とせず、全てマウスで行えるため、初心者でも簡単に操作できる。

2. 統一された操作性

プルダウンメニューによる選択や、ダブルクリック等の操作は、MS-Windows における基本操作であり、初心者にも受け入れられ易い。訓練ソフトの全てが同じ操作で扱える。

3. 訓練内容の起動

訓練内容の選択は、アイコンをダブルクリックするだけで良い。アイコンは、ユニークなキャラクターで描かれており、訓練者が一目で訓練内容を理解できるように配慮してある。

4. ヘルプ機能

主として指導者向けに、オンラインヘルプマニュアル

を提供する。アプリケーションの特徴や使用法を分かり易く絵や文章で表現している。

3.2 ソフトウェアのメディア

発声訓練ソフトウェアは、家庭での使用も考えており、将来的に CD-ROM で提供する。CD-ROM は、最近のパソコンに標準装備されており、情報媒体として、容量の大きいソフトウェアを提供できるということと、製作費が安価であるという特徴を持つ。

4. 発声訓練ソフトウェア

図2に、現在開発済みの発声訓練ソフトウェアを示す。各ソフトウェアは、図3で示されるアイコンをクリックすることにより起動できる。



図3 発声訓練メニュー

声の大きさや高さ等の基礎的な訓練を行うものと母音訓練を中心としたものである。以下、各ソフトウェアに

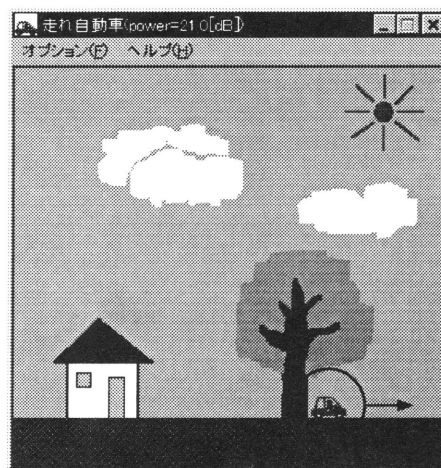


図4 声の誘導訓練

ついて説明する。なお、音声のサンプリング周波数は11[kHz]、量子化ビット数は16ビットである。

4.1 声のリズムと大きさ訓練

① 声の誘導訓練 (走れ自動車)

図4は、声の誘導訓練を行うソフトウェアの出力画面例である。音声のパワーが、設定された閾値以上ならば、自動車が町の中をドライブする。声を恥しがらずに大きく出す幼児レベルの訓練に用いる。

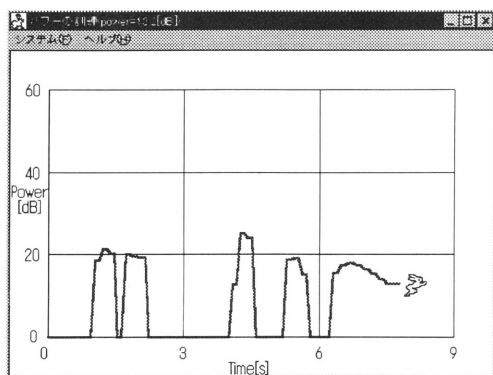


図5 音声パワーの訓練

② リズム訓練 (うさぎピョンピョン)

図5は、声の大きさに合わせてうさぎがジャンプする。うさぎを任意の間隔でジャンプさせることにより、リズム訓練を行う。スケール表示を選択して、音声パワーの具体的な数値を知ることができる。

③ 声の大小訓練 (うさぎさん危機一髪)

図6に示すように、音声のパワーが目標範囲内に入れば、うさぎは孤島に上陸し、一定時間範囲内の発声を続けるとヘリコプターに救助される。目標範囲にパワー値が入らないときは、図7のようにうさぎは海の中へ沈んで行く。

4.2 声の高さ (ピッチ) 訓練

① ピッチ訓練Ⅰ (トンボ)

図8は、実時間でピッチ周波数の変動を表示するソフトウェアの出力画面例である。表1に示すように、トンボの羽根の色により、/有声音、無声音、無音/の判別ができる。時間軸を長く取ることにより、音階の訓練も可能である。表2に、選択できるピッチの範囲を示す。聴覚障害者の場合、発声時に不規則なピッチ変動が見られることが多く、一定のピッチで発声訓練することは重要な訓練項目の一つである。

ピッチ周波数は、音声信号の変形自己相関関数より推定する。ここで、 n 時刻の実際の音声信号を S_n 、過去の

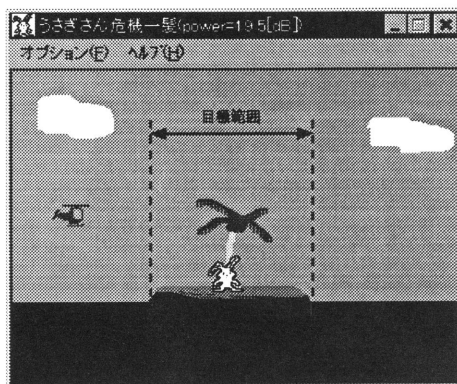


図6 声の大きさが範囲内の場合

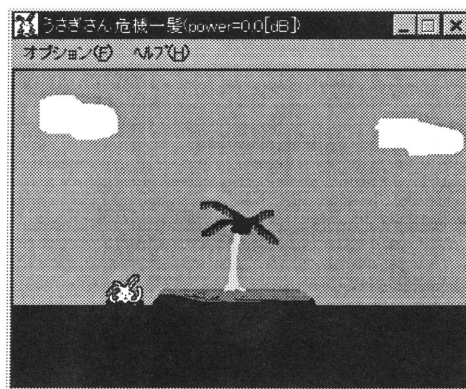


図7 声の大きさが範囲外の場合

表1 トンボの羽根の色による判別

種類	トンボの羽根の色
有声音	赤色
無声音	青色
無音	白色

表2 選択できる縦軸の範囲

種類	ピッチの範囲 [Hz]
高い	250 ~ 450
やや高い	150 ~ 350
ふつう	50 ~ 250
その他	0 ~ 600

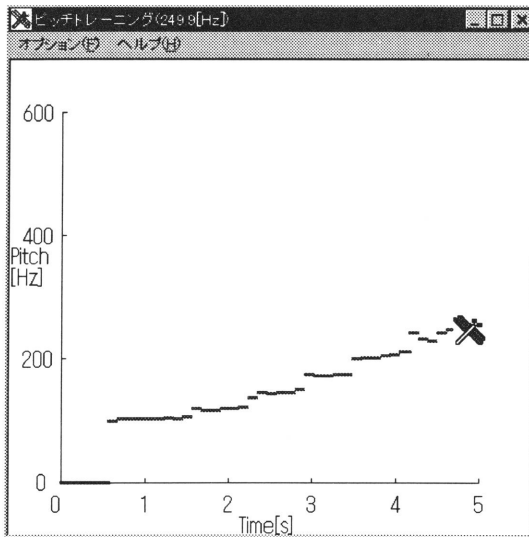


図8 ピッチ周波数の表示例（ドレミファソラ）

p 点 ($S_{n-p}, S_{n-p+1}, \dots, S_{n-1}$) の多項式近似より S_n を予測した値を $\hat{S}_n$ とするとき, $e_n = S_n - \hat{S}_n$ を残差信号と呼ぶ。変形自己相関関数は, この時系列信号 e_n に対する自己相関である。

変形自己相関関数 $R(n)$ が与えられるとき, $R(0)$ の次にピーク値を取る $R(\tau)$ を探索する。このときの τ の値をピッチ周期 T とする。あらかじめ, ピッチの存在範囲を定めておき, その区間内で探索を行う。

有声/無声の判別は $R(\tau)/R(0)$ の比により行う⁽⁸⁾。

② ピッチ訓練II（舞えちょうちょう）

図9, 図10は, 訓練者の発声した声の高さに合わせて, 蝶が飛んで行くソフトウェアの画面出力例である。蝶は, ピッチが高い程高く, 低い程低く飛ぶ。訓練者は, 表3に示すように4種類の訓練範囲を選択できる。図10のように, 発声音のピッチ周波数が, 目標範囲から外れると, 画面のスクロールが止まり一定時間先に進めない。

表3 選択できる範囲

種類	ピッチの範囲 [Hz]
高い	225 ~ 275
やや高い	175 ~ 225
ふつう	85 ~ 135
低い	75 ~ 125

4.3 ホルマントグラフ

図11は, 横軸に第1ホルマント周波数 F_1 , 縦軸に第

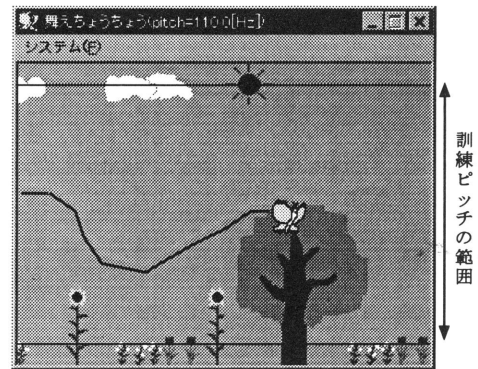


図9 声の高さが範囲内の場合

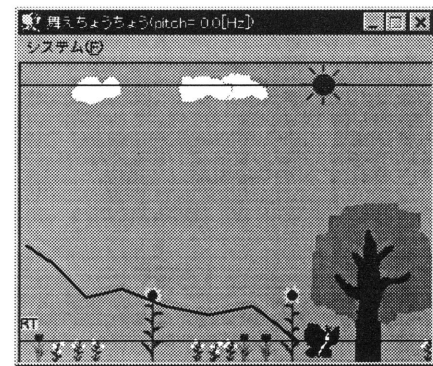


図10 声の高さが範囲外の場合

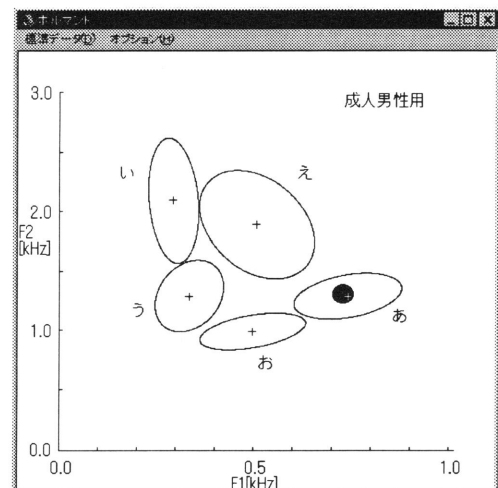


図11 ホルマントグラフの表示例

2ホルマント周波数 F_2 の値を表示したもので, ホルマントグラフと呼ばれている。

図中の楕円は, 各母音における標準的な F_1 - F_2 分布を

示している。楕円の大きさにより (F_1 - F_2 分布の分散値で決定される) 難易度を3段階に設定できる。訓練者は、自分の母音が標準的な母音からどの程度ずれているかを、実時間で逐次確認しながら高精度に判別できる。

ホルマント周波数は、線形予測係数(時系列信号 $y_{n-p}, y_{n-p+1}, \dots, y_{n-1}, y_n$ において y_n の値を過去の p 点の多項式近似で推定することを線形予測法と呼び、全極型の伝達関数 $1/(1+\sum_{i=1}^p a_i z^{-i})$ で短時間のスペクトル推定を行うことに等しい。このときの a_i を指す。) を係数とする(1)式で示される高次代数方程式を解くことにより求まる。

$$1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_p z^{-p} = 0 \quad (1)$$

ここで、 $z_i (=z_{iR} + jz_{iI})$ が方程式の根となる。ホルマント周波数を F_i 、ホルマント帯域幅を B_i とすると、次式により求められる。 T はサンプリング周期とする。

$$F_i = \frac{1}{2\pi T} \tan^{-1} \frac{z_{iI}}{z_{iR}}$$

$$B_i = \frac{1}{2\pi T} \log (z_{iR}^2 + z_{iI}^2)$$

帯域幅を制限することで抽出エラーを抑えている⁽⁹⁾。

4.4 声道断面図のリアルタイム表示

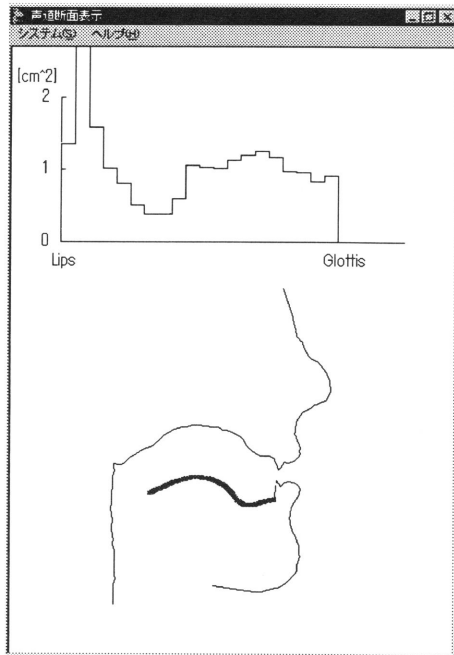


図12 声道断面図の表示例

図12は、声道断面図のリアルタイム表示を行うソフトウェアの出力画面例である。訓練者の舌やあごの動きを具体的に確認できる。通常、声道の断面を円形に近似した声道断面積関数を求めた後、図13に示す実際の声道形状に合わせて表示する。しかし、図13に示すような複雑な形状に近づけるのは困難であるため、ここでは各断面を楕円形に近似して表示している。

声道断面積関数 A_n は、PARCOR 係数 x_n (偏自己相関関数: 時系列信号におけるある点の値を未来と過去から推定したとき、それぞれの誤差の相互相関関数) から次式で求められる。

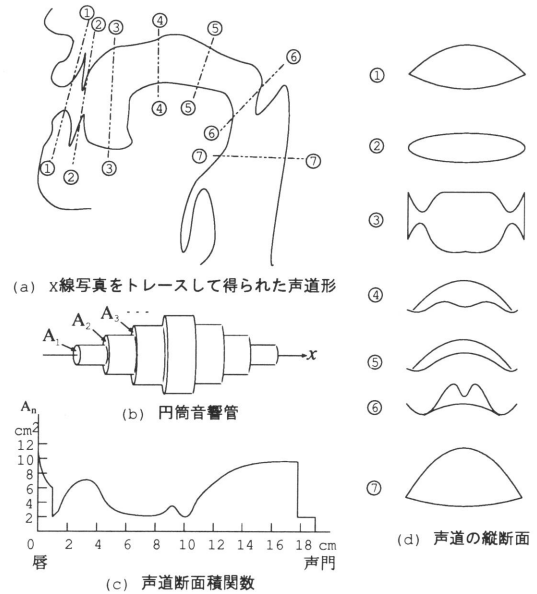


図13 Fant による声道断面積関数の観測例

$$A_n = A_{n-1} \cdot \frac{1-x_n}{1+x_n} \quad (2)$$

ただし、音声には、音源特性と口からの放射特性も含まれているので、声道の特性を分離するために適応逆フィルタを通して⁽¹⁰⁾⁽¹¹⁾。

5. オンラインヘルプマニュアル

本システムでは、図14、図15のようにオンラインヘルプマニュアルを提供しており、訓練中であっても適時呼び出すことが可能である。システムの全体的な説明や各発声訓練内容の目的、及び使用法等を分かり易く記述している。

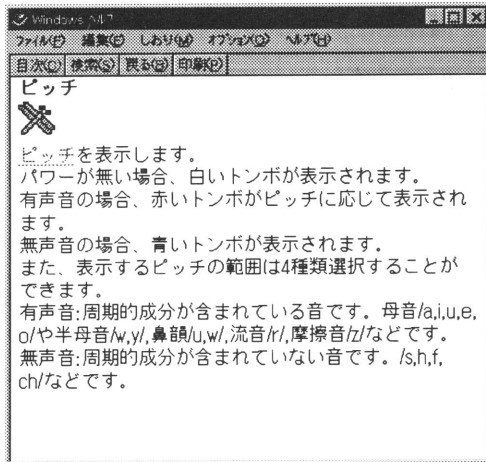


図14 ピッチ訓練 I のオンラインヘルプ

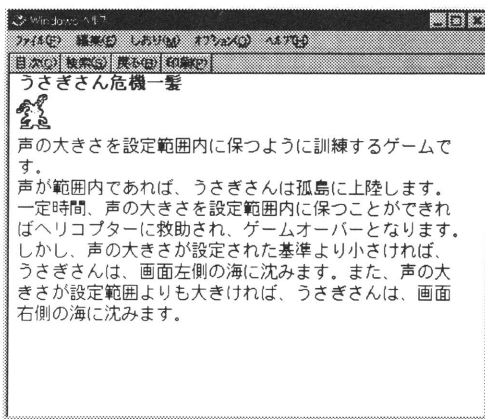


図15 声の大小訓練のオンラインヘルプ

6. 開発予定の機能

開発予定のソフトウェアを以下に示す。

1. 子音訓練

子音 (/s, h, z/, /n, m/, /b, d, g/, /p, t, k/, /r/, /w, y/) を含めた日本語単音節の発声訓練ソフトウェアを開発する。また、聾学校等の教育で用いられている /p, d, k/, /p, b, m/... 等の子音組についても訓練できるソフトウェアを開発する。

2. 単語訓練

日常生活で使用頻度の高い単語の発声訓練ソフトウェアを開発する。

3. 音声分析ツール

指導者向けに、訓練者の詳細な音声分析が可能なツールを提供することを考えている⁽¹²⁾。音声データのファイル化、波形表示、FFT 分析、ソナグラム表示等を行う機能を用意する。

なお、本システムは、聴覚障害児教育以外の他分野(外国人の日本語教育等)にも応用していく予定である。

7. むすび

従来、DSP を搭載した専用ハードウェアを用いなければ、実時間レベルでの音声分析処理は不可能であったが、今回、Pentium CPU を搭載したパソコン上で実現できることを示した。訓練ソフトウェアの動作環境として MS-Windows 3.1 及び 95 を採用したため、分かり易く統一された操作性を提供でき、初心者でも容易に操作できる。システムの全てをソフトウェアで構築しているため、仕様の変更や拡張が容易に行え、コスト面でも貢献できる。最後に、今後の課題として、以下の項目を挙げておく。

1. 訓練履歴の自動管理

訓練者個人毎に発声訓練した履歴を自動管理し、指導者に児童の訓練状況と習熟の度合を示すようにする。

2. 訓練レベルの自動設定

訓練者個人の発声能力に合わせて練習コースを自動選択する機能を付加する。

3. 試用と評価

開発済みのソフトウェアを CD-ROM で提供し、実際に聾学校等の教育施設で評価してもらう。

参 考 文 献

- 1) 村松広丘, 重永実: “聴覚障害者用文音声発話練習システム”, 信学論(A), **J75-A**, 9, pp. 195-203 (1992-2)
- 2) 小林正幸: “動画を活用した聴覚障害児のための読話訓練システムの研究(2)”, ET95-88, 5, pp. 35-39 (1995-11)
- 3) 田中克典, 中村克彦: “聴覚障害児の言語訓練支援システム”, ET93-123, 8, pp. 127-123 (1994-1)
- 4) 塚本榮一, 伊藤智之: “特殊教育における言語獲得を支援するコンピュータシステム”, TL95-11, 8, pp. 17-24 (1995-01)
- 5) 梅崎太造, 桑山清明: “聴覚障害児用発声・発語訓練器の開発”, ET91-40, 6, pp. 13-18 (1991-6)
- 6) 日向駿馳: “入門 Visual C++”, 翔泳社
- 7) “Microsoft Windowsソフトウェア開発キット”, 翔泳社
- 8) 安居院猛, 中嶋正之: “コンピュータ音声処理”, 秋葉出版
- 9) 板倉文忠, 斎藤収三: “統計的手法による音声スペクトル密度とホルマント周波数の推定”, 信学論(A), **J53-A**, 8, pp. 35-42 (1972-1)
- 10) 中嶋隆之, 鈴木虎三, 大村浩, 石崎俊, 田中和世:

- “デコンボリューションによる声道形の推定と適応型音声分析システム”, 日本音響学会誌, 34(1978)
- 11) 菅野文武, 鎌田弘之, 石田義久: “音声認識技術を利用した単音節・単語の発話練習とスピーチトレーナへの応用”, ET93-122, 8. pp. 119-126(1994-1)
- 12) 山田賢治, 梅崎太造: “Motifベースの音声分析表示システム”, 日本音響学会平成7年春季研究発表会, No. 1-8-8, pp. 357-258(1995-3)