

ネットワーク環境を利用した聴覚障害児用発話訓練支援システム

梅崎 太造[†] 倉谷 和彦[†] 藤吉 弘亘[†]

Speech Training Support System for Hearing Impaired Children Using the Network Environment

Taizo UMEZAKI[†], Kazuhiko KURATANI[†], and Hironobu FUJIYOSHI[†]

あらまし ネットワーク環境下で動作する聴覚障害児用発話訓練支援システムについて述べる。まずはじめに、本システムの概略と設計方針について述べる。DSPを搭載する専用ハードウェアを用いることなく、実時間レベルでの音声分析処理を Pentium CPU を搭載したパソコンのみで実現している。訓練ソフトウェアは、MS-Windows 3.1, 95 および NT 上で動作するように設計されている。そのため、プルダウンメニューやマウスクリック等の優れた操作性により、初心者でも容易に訓練が可能である。次に、声の大きさ（パワー）、リズム、高さ（ピッチ）、音階、母音、無声音/有声音、舌の動き等の訓練ができる発話訓練ソフトウェアについて解説する。最後に、主に指導者向けに作成した訓練履歴の自動管理機能とオンラインヘルプ機能について述べる。これらの機能より、指導者は適時ネットワーク環境下で訓練している児童全員の習熟の度合や訓練ソフトウェアの目的および使い方をすることができる。

キーワード ネットワーク、聴覚障害児、発話訓練、音声分析、MS-Windows

1. まえがき

聴覚障害児教育の重要な課題の一つに、発話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、専門知識をも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、障害児の発話訓練を支援する装置の研究開発が行われてきた [1]~[6]。筆者らの研究室においても、企業と共同で聴覚障害児用の発話訓練機「トーキングトレーナ」を開発してきた。本装置は、現在も各方面で活用されている [7]。

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機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ず、汎用性に欠けている場合が多い。スイッチ等のユーザインタ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ビットパーソナルコンピュータ（以後、パソコンと略す）は演算能力が急激に向上し、その計算速度は、高速な DSP に匹敵するものもある。メモリやハードディスク等の外部記憶装置の容量も十分に確保でき、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。加えて、最近の小・中学校におけるパソコン教室の普及により、32ビット以上の高性能パソコンを用いた汎用性と移植性の高い発話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。

本論文では、MS-Windows 3.1, 95 および NT 上で動作可能な聴覚障害児用発話訓練支援システムについて述べる。本システムでは、アプリケーションソフトウェアや訓練履歴管理をネットワーク上のサーバマシンで一元管理することができる。現在、声の大きさ訓練、リズム訓練、高さ（ピッチ）訓練、および母音訓練を中心とするソフトウェアを利用することが

[†] 中部大学工学部電子工学科, 春日井市
Faculty of Engineering, Chubu University, Kasugai-shi, 487 Japan

できる[8]。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せず、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あるため、機能の拡張や変更が容易に行えるという特徴をもつ。

まず2.で本システムの概要について述べ、3.でシステム設計の基本方針について述べる。4., 5.では発話訓練ソフトウェアのもつ訓練機能と履歴管理機能について解説する。最後に6.で訓練を支援するためのヘルプ機能について述べる。

2. システムの概要

システムは、図1のようにDOS/Vパソコン、サウンドボードおよび、発話訓練ソフトウェアの三つの要素から構成される。DOS/Vパソコン以外でもWindowsが動作するものであれば問題ない。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(互換)を用いているが、Windows対応のサウンドボードであれば、問題なく使用できる。音声信

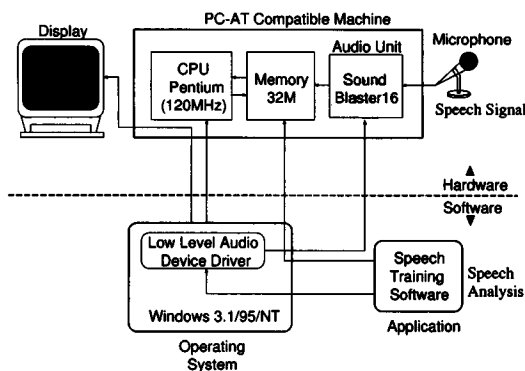


図1 システムの概要
Fig.1 Outline of system.

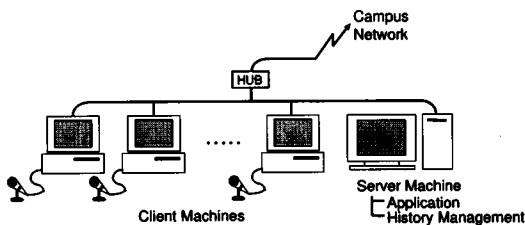


図2 ネットワーク環境
Fig.2 Network environment.

号の入出力には、WindowsがもつWaveFormAudio機能の一つである低レベルWaveFormAudioServiceを利用する。

図2のようなネットワーク環境下での使用を前提としており、発話訓練ソフトウェアと履歴管理ツールはサーバマシンに置かれている。クライアントマシン側でソフトウェアを実行する場合、ネットワークを介してあらかじめコピーしたものを起動する方法とサーバマシン側から起動する方法を選択することができる。訓練履歴は、サーバマシン上のハードディスクに記録される。

発話訓練ソフトウェアの動作は、現在のところMS-Windows 3.1, 95およびNT上で確認している。システム開発には、マイクロソフト社のVisual C++ Version 2.0, 4.0を用いている。

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1に考える必要がある。故に、本システムでは、Windowsの優れた操作性を活用することを基本方針とする。使いやす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している。以下に設計方針として特に考慮した項目を列挙する。

(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、すべ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やすい。訓練ソフトのすべ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よい。アイコンは、ユニークなキャラクタで描かれており、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している。具体例は6.で述べる。

(5) ソフトウェアのメディア

発話訓練ソフトウェアのメディアとしては、最近の

パソコンに CD-ROM ドライバが標準装備されていることを考慮して CD-ROM で提供する。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もつ。

4. 発話訓練ソフトウェア

図 3 の実線で示してある発話訓練ソフトウェアは既に関済済みのものであり、点線で示してあるソフトウェアは現在開発中である。各ソフトウェアは、図 4 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。

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 声の大きさとリズム訓練

(1) 声の誘導訓練 (走れ自動車)

図 5 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。音声のパワー値 P [dB] が、設定されたしきい値以上ならば、自動車が町の中をドライブする。 P は、音声信号を S_n とすると式 (1) で求められる。 N は音声信号のポイント数である。これは、声を恥しめがらずに大きく出す幼児レベルの訓練に用いる。

$$P = \frac{1}{N} \sum_{n=1}^N (10 \log_{10} S_n^2) \quad (1)$$

しきい値はプルダウンメニューにより、5 段階の中から選択できる。

(2) リズム訓練 (うさぎピョンピョン)

図 6 は、声の大きさに合わせてうさぎがジャンプす

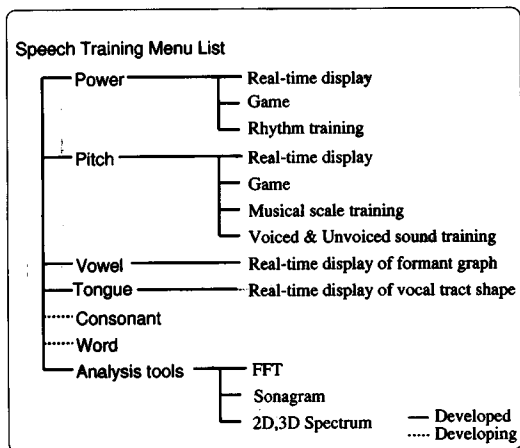


図 3 アプリケーション内容
Fig. 3 Contents of application.

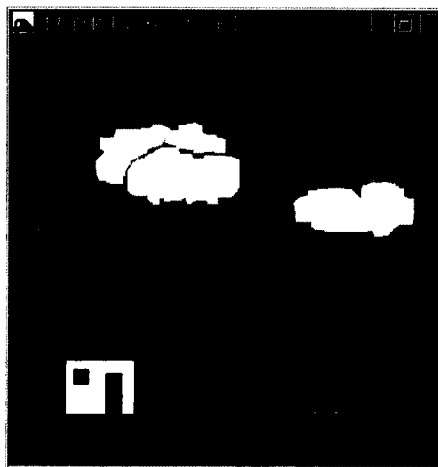


図 5 声の誘導訓練
Fig. 5 Training of speech leading.

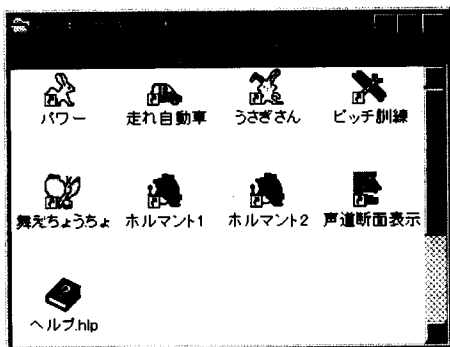


図 4 発話訓練メニュー
Fig. 4 Speech training menu.

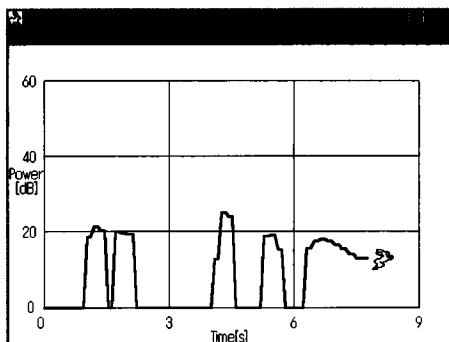
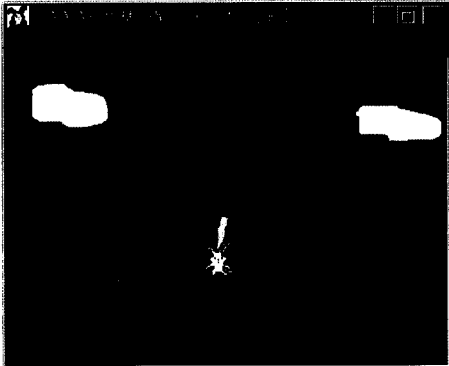
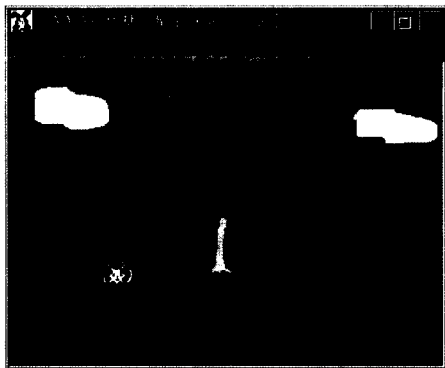


図 6 音声パワーの訓練
Fig. 6 Training of speech power.



(a)Example of inside range.



(b)Example of out of range.

図7 パワーレンジの訓練
Fig.7 Training of power range.

るソフトウェアの出力画面例である。うさぎを任意の間隔でジャンプさせることにより、リズム訓練や声の大きさ訓練を行う。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7(a)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話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(b)のようにうさぎは海の中へ沈んでいく。クリア後若しくは一定時間後に、得点が表示される。採点方法は、初期もち点を100として、目標範囲からパワー値がはずれるごとにもち点から-1する減点方式を用いている。

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

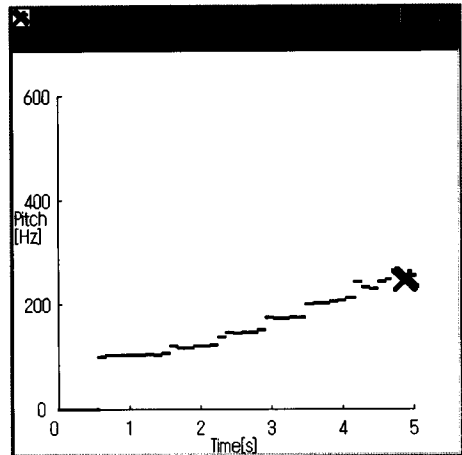


図8 ピッチ訓練例 (ドレミアソラ)
Fig.8 Example of pitch training.

表1 トンボの羽根の色による判別
Table 1 Discrimination by shuttlecock color of dragonfly.

種類	トンボの羽根の色
有声音	赤色
無声音	青色
無音	白色

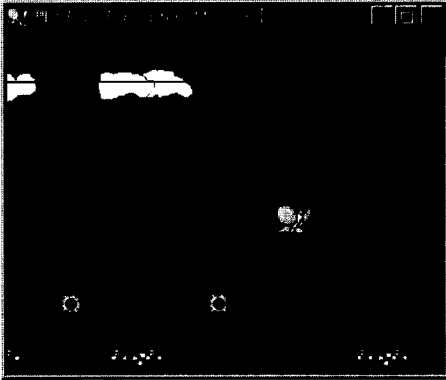
表2 選択可能なピッチレンジ
Table 2 Selectable pitch range.

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

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。聴覚障害児の場合、発話時に不規則なピッチ変動が見られることが多く、一定のピッチで発話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 [9]。ここで、時刻 n の実際の音声信号を S_n 、過去の 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 R_i は、この時系列信号 e_n に対する自己相関であり式 (2) で示される。

$$R_i = \frac{1}{N+1-i} \sum_{i=1}^N e_n \cdot e_{n+i} \quad (2)$$



(a)Example of inside range.



(b)Example of out of range.

図9 ピッチレンジの訓練
Fig.9 Training of pitch range.

変形自己相関関数 R_i が与えられるとき、 R_0 の次にピーク値をとる R_r 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_r/R_0 の比により行う [10]。

(2) ピッチ訓練 II (舞えちょうちょう)

図 9(a), (b) は、訓練者の声の高さに合わせて、蝶が飛んでいくソフトウェアの画面出力例である。蝶は、ピッチが高いほど高く、低いほど低く飛ぶ。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 9(b) のように、発話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。

4.3 ホルマントグラフ

図 10 は、横軸に第 1 ホルマント周波数 F_1 、縦軸に第 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。

表 3 選択可能なピッチレンジ
Table 3 Selectable pitch range.

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

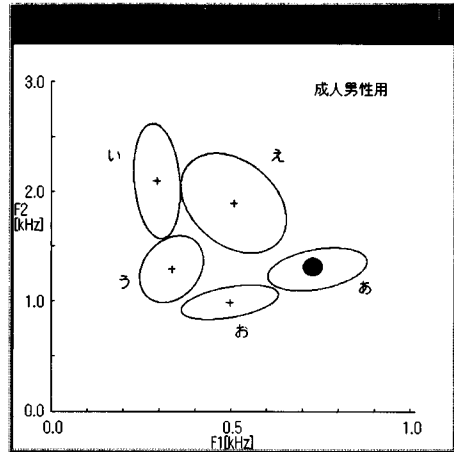


図 10 母音訓練
Fig.10 Training of vowel.

図中のだ円は、各母音における標準的な F_1 - F_2 分布を示している。だ円の大きさにより (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訓練できる。日本語 5 母音の場合、 F_1, F_2 の値のみで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) を係数とする式 (3) で示される高次代数方程式を解くことにより求まる。

$$1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_p z^{-p} = 0 \quad (3)$$

ここで、 $z_i (= z_{iR} + j z_{iI})$ 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、式 (4)、式 (5) により求められる。 T はサンプリング周期とする。

$$F_i = \frac{1}{2\pi T} \tan^{-1} \frac{z_{iI}}{z_{iR}} \quad (4)$$

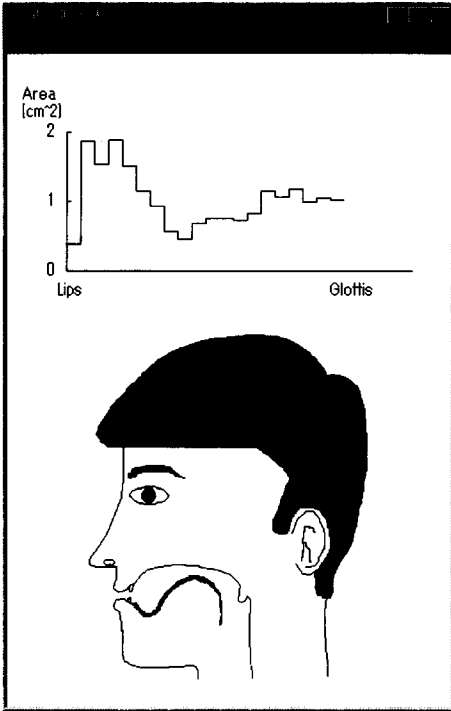


図 11 声道断面図の実時間表示
Fig.11 Real-time display of vocal tract shape.

$$B_i = \frac{1}{2\pi T} \log(z_i R^2 + z_i I^2) \quad (5)$$

帯域幅を制限することで抽出エラーを抑えている [11].

なお、訓練を行う児童だけでなく、指導者も手本を見せることができるように配慮してある。訓練者はブルダウンメニューにより/幼児、小学校低学年、小学校高学年、中学生、成人男性、成人女性/の中から自分に合う訓練環境を選択できる。この分類のために児童 500 人、大人 200 人分の母音音声データベースを用いている。

4.4 声道断面図の実時間表示

図 11 は、声道断面図の実時間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 (図 12 (b)) した声道断面積関数 (図 12 (c)) を求めた後、図 12 (a), (d) に示す実際の声道形状に合わせて表示する [12]。しかし、図 12 (a), (d) 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、ここでは各断面をだ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

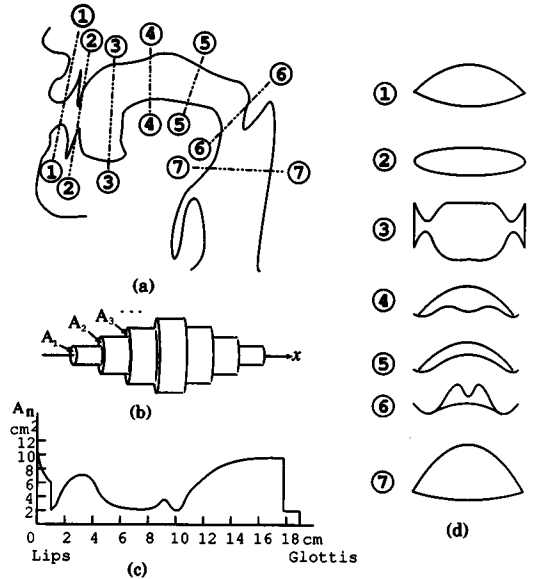


図 12 Fant による声道断面積関数の観測例
Fig.12 Example of vocal tract area function observed by Fant: (a) Vocal tract shape obtained by tracing X-ray photograph, (b) Acoustic cylinder tube, (c) Vocal tract area function, (d) Cross section of vocal tract.

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。

$$A_n = A_{n-1} \cdot \frac{1 - x_n}{1 + x_n} \quad (6)$$

但し、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 [13], [14].

5. 訓練履歴の自動管理機能

本システムでは、訓練を行う児童 1 人ひとりの訓練状況を把握するため、サーバマシン側に図 13 で示されるヒストリマネージャと呼ぶ履歴管理ツールを提供している。この場合、訓練ソフトウェアはそれぞれクライアント側に置かれているとする。クライアントマシン側から書き込まれた履歴情報ファイルにより、訓練者の発話訓練履歴 (訓練回数、平均得点等の評価など) を自動管理することができる。これにより指導者は、ネットワーク環境下で接続された複数のクライアントマシンで別々にトレーニングしている児童達の習熟の度合を確認できる。なお、クライアントマシン側からヒストリマネージャを介して訓練履歴を参照することも可能である。

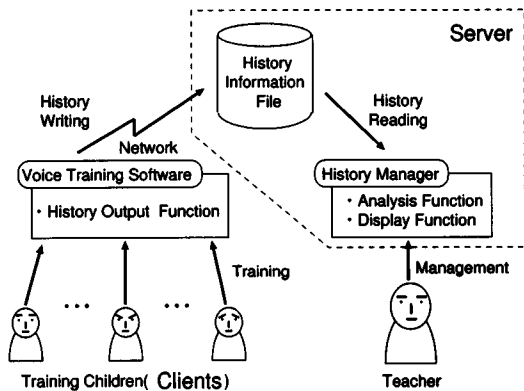


図 13 ネットワーク環境下での履歴管理
Fig. 13 History management in the network environment.

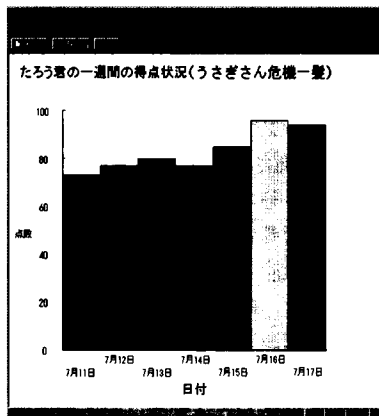


図 15 個人得点の表示例
Fig. 15 Example of displaying the personal score.

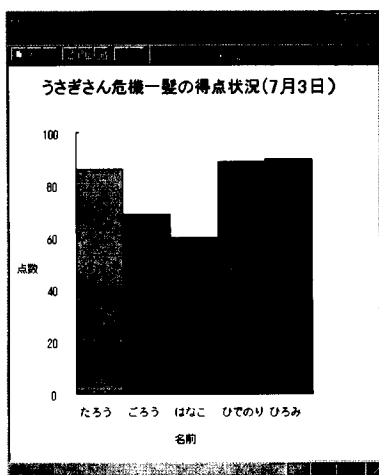


図 14 ヒストリマネージャによる得点の表示例
Fig. 14 Example of displaying the score using history manager.

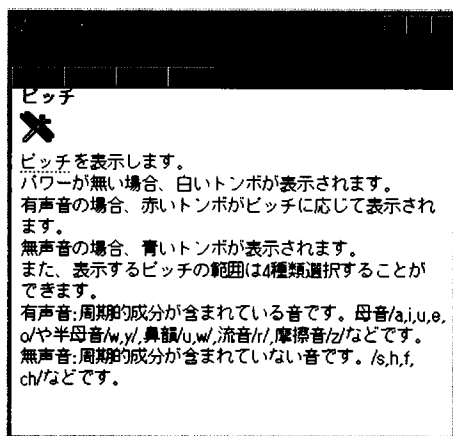


図 16 ピッチ訓練 I のオンラインヘルプマニュアル
Fig. 16 Online help manual of the pitch training I.

図 14 は、5 台のマシンで訓練した 5 人の児童の平均得点結果をヒストリマネージャより表示した例である。図 15 は、個人の 1 週間の訓練結果を表示した例である。

6. オンラインヘルプマニュアル

本システムでは、図 16 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話訓練ソフトウェアの目的、および使用法等をわかりやすく絵や文章で記述している。

7. むすび

本論文では、ネットワーク環境を利用した聴覚障害児用発話訓練支援システムについて述べた。

まずはじめに、本システムの構成と設計方針について述べた。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, 95 および NT を採用したため、プルダウンメニューやマウスクリック等のわかりやすく統一された操作性を提供でき、初心者でも容易に訓練できる。システムのすべ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

に行えるという特徴をもつ。

次に、開発済みの発話訓練ソフトウェアの訓練目的とその内容について述べた。各ソフトウェアの起動はアイコンをクリックするだけでよく、声の大きさ（パワー）、リズム、高さ（ピッチ）、音階、母音、無声音/有声音、舌の動き等の訓練を行える。

最後に、主に指導者の使用を想定して作成した訓練履歴の自動管理機能とオンラインヘルプ機能について述べた。指導者は、離れたネットワーク環境下で訓練している児童の訓練履歴をヒストリマネージャを介して参照することができる。また、ヘルプマニュアルにより、適時訓練ソフトウェアの目的や使い方を知ることができる。これらの二つの機能により指導者側の負担が軽減される。

今後開発済みの発話訓練ソフトウェアを CD-ROM に組み込み、実際に聾学校等の教育施設で評価してもらう予定である。

謝辞 日ごろ、御指導頂く鈴木宣夫教授に深謝致します。

文 献

- [1] 重永 実, “障害児の発話訓練へのマイクロコンピュータの応用,” 聴覚言語障害, vol.12, no.2, pp.79-86, 1983.
- [2] 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), vol.J75-A, no.9, pp.195-203, Sept. 1992.
- [3] 岡 辰夫, 村田程夫, 山田義則, “発声発話訓練装置 CV-3 (聴覚障害児のためのプログラムとその実際),” 信学技報, ET86-10, 1987.
- [4] 小林正幸, “動画を活用した聴覚障害児のための発話訓練システムの研究 (2),” 信学技報, ET95-88, 1995.
- [5] 田中克典, 中村克彦, “聴覚障害児の言語訓練支援システム,” 信学技報, ET93-123, 1994.
- [6] 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,” 信学技報, TL95-11, Aug. 1995.
- [7] 梅崎太造, 桑山清明, “聴覚障害児用発声・発話訓練器の開発,” 信学技報, ET91-40, 1991.
- [8] 倉谷和彦, 梅崎太造, “Windows 上で動作する聴覚障害児用発声訓練支援システム,” 信学技報, ET96-29, 1996.
- [9] M.M.Sondhi, “New methods of pitch extraction,” IEEE Trans. Audio and Electroacoustics, vol.AU-16, no.2, pp.262-266, 1968.
- [10] J.D. Markel, A.H. Gray Jr, “Linear prediction of speech,” Springer-Verlag, 1976.
- [11] 板倉文忠, 高藤収三, “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), vol.J53-A, no.8, pp.35-42, 1972.
- [12] G.C.M. Fant, “Acoustic Theory of Speech Production,” Mouton, 1960.
- [13] 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,” 音響誌, vol.34, 1978.

- [14] 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ーへの応用,” 信学技報, 1994.

(平成 8 年 8 月 30 日受付, 12 月 2 日再受付)



梅崎 太造 (正員)

昭 57 豊橋技科大・工・情報卒。昭 62 名大大学院博士課程了。同年 4 月名大助手。平 2 中部大講師。平 4 より同大助教授。平 5~6 カーネギーメロン大学客員研究員。工博。音声・画像情報処理、聴覚障害児教育の研究に従事。情報処理学会、日本音響学会各会員。

会各会員。



倉谷 和彦 (学生員)

平 7 中部大・工・電子卒。現在、同大大学院博士前期課程在学中。音声情報処理の研究に従事。



藤吉 弘亘 (学生員)

平 4 中部大・工・電子卒。平 6 同大大学院修士課程了。現在、同大学院博士後期課程在学中。画像処理、パターン認識の研究に従事。