

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 4 月 1 日現在

機関番号：23803

研究種目：基盤研究（C）

研究期間：2010～2012

課題番号：22500513

研究課題名（和文） 音声認識による聴覚障害者の発音訓練法の開発  
 —音声セグメント技術の導入—

研究課題名（英文） Development of pronunciation training method using speech  
 recognition for the hearing impaired.

研究代表者

松浦 博（MATSUURA HIROSHI）

静岡県立大学・経営情報学部・教授

研究者番号：60451085

研究成果の概要（和文）：

音声分析システムVoiceAnalyzerに従来から整備されている子音領域推定機能に加えて母音領域推定機能を整備し、単語における音素の構成を把握しやすくした。また、音声セグメントをアルファベットで表示することに加え、平仮名で表示することによって分かりやすくした。さらに、Windows8において新たに加わったタッチ入力への対応を行い、使い勝手を向上させた。聴覚障害者によって発話された235単語×10人分のデータを収集し、一般に聴覚障害者は濁音の発声が難しいことが分った。また、本システムによって発音を可視化することによって、音素の省略された発話やあいまいだった発話が改善した。

研究成果の概要（英文）：

We added a vowel region estimation function to the existing consonant region estimation function in the VoiceAnalyzer speech analysis system, which made it easier to understand the configuration of phonemes in words. We further facilitated understanding by adding a hiragana display to the alphabetical display of the phonetic segment labels. Moreover, usability was enhanced through compatibility with the touch-input function newly added in Windows 8. We collected data of 10 persons × 235 words spoken by hearing-impaired people and found that the utterance of voiced speech is generally difficult for hearing-impaired people. We also understood that speech omitting phonemes or vague speech was improved through this system's visualization of sound.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	1,500,000	450,000	1,950,000
2011年度	900,000	270,000	1,170,000
2012年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,300,000	990,000	4,290,000

研究分野：総合領域

科研費の分科・細目：人間医工学・リハビリテーション科学・福祉工学

キーワード：ヒューマンインターフェイス、健康・福祉工学、パターン認識、音声情報処理、障害学、発音評価、音声認識、聴覚障害

### 1. 研究開始当初の背景

言語とりわけ音声言語は人間の本質と言っても良い。なぜなら、人間を上回る様々な能力を備える動物は数多く存在するが、音声言語をあやつることはない。幼児は簡単な言葉をかけるだけで特別な教育なしに、最低限の音声言語を獲得する。この獲得機会を奪われる聴覚障害は重大な障害であると言える。

人間は自らの発声を耳でとらえ脳に信号としてフィードバックする。この信号と発声した音韻ごとに期待される信号との誤差を時々刻々、補正しながら正しく発音する。幼児期から人の声を聞くことによって、音韻モデルが構築されると考えられるが、聴覚障害者はフィードバック情報が少ないため、音韻モデルの構築が難しい。聾学校等の教師の指導によって発音訓練が行われる場合においても、十分なフィードバック情報を与えることは難しい。本研究においては、音声セグメント技術によってフィードバック情報を提供し、正確な音韻モデルの構築を支援し、聴覚障害者の自学自習が可能な発音訓練の基礎理論・訓練法を確立する。

従来、種々のセンサを用いて、母音スペクトルを表示したり、鼻振動、呼気流、声帯振動を捉えたりする発音訓練システムが主に児童向けとして実用化されている。しかし、このシステムは発声された結果である音声そのものが、どの程度、正確であるかを直接的に示すものではない。一方、音声認識技術を利用して発音の正確さを計測することも試みられてきた。

しかし、一般の音声認識は文章入力ソフトやカーナビなどのコマンド入力として実用化しているものの、単語を基本単位とする認識技術であるため、あくまでも単語全体で見たときの確からしさを評価するのみである。したがって、単語を構成する音韻、さらに細かな音素の確からしさは確認できない。そのため、システムから発音が悪いとの指摘を聴覚障害者が受けても、単語内のどの音素が適切に発音できていないのかの手がかりがない。

### 2. 研究の目的

聴覚障害者には発音障害が伴うことが多いことは良く知られており、その発音障害の程度は個人差が非常に大きい。聴覚を失った時期、聞こえの程度（たとえば聞こえる周波数帯域）、発音訓練の経験やその程度などによって様々な個人差があると考えられている。残存する聴力がある程度認められる場合もあり、また補聴器や人工内耳の補助によるフィードバックも、発話特性に大きな影響を及ぼしていると考えられている。研究期間内において、次の点を明らかにする。

- ・聴覚障害の様々なタイプによって苦手な音韻等の発話特性の違いを系統的に調査する。
- ・そのために、音声セグメントシステムを用いて、聴覚障害者の発声データを収集する。
- ・収集データから様々な聴覚障害者にとって、どのような音韻が発音し難いかを分類する。
- ・それぞれの音韻の発音を改善するための訓練法としてどのような形態が効果的か明確化する。
- ・聴覚障害者などが本システムを違和感無く使用できるように、模範発音の提示方法や発音内容をリアルタイムかつ一目で分かりやすいように提示する方法を開発する。

### 3. 研究の方法

次の方法によって研究を行った。

- ・聴覚障害者の障害の程度・内容の多様性を前提とし、様々な障害のタイプによる発話特性の違いについて調査する。調査方法としては、聴覚障害者の承諾を受けた上でのインタビューと、改良する音声セグメントシステムを用いて聴覚障害者の発声データを収集する。
- ・現在の歯科向けの発話機能評価システムでは発音が終了された後に表示されるが、発音訓練ではできるだけリアルタイムに結果を表示することが、訓練者にとって望ましい。したがって、リアルタイム性を目標として音声セグメントシステムの再構築を行う。
- ・発声データ収集を、20名を目標として行い詳細に調査する。
- ・平成22年度および平成23年度に収集した

発声データから聴覚障害者にとって、特に発音し難い音韻や発声の特性を明らかにする。事前調査では、人にもよるが例えばラ行・サ行音、母音の中では舌が最も奥へ動く「お」や語尾が難しいという結果が得られた。

・発音の困難な音韻の発音訓練法として、それぞれの音韻に対して効果的な形態について研究する。

・発音訓練ソフトの仕様を作成する。模範発音の提示方法と訓練者の発音結果を例えば仮名文字表記、調音器官の断面図などによって分かりやすく提示する方法を検討する。

・上記の成果を踏まえて発音訓練ソフト version1 を試作する。

・発音訓練ソフト version1 を聴覚障害者に試用していただき改良を加え、完成度を上げた発音訓練ソフト version2 を試作する。

・発音訓練法について、訓練を飽きずに実施し、達成感を実感できるようにするため、ゲーム性を取り入れる等の工夫を行う。

#### 4. 研究成果

##### (1) 平成 22 年度の成果

本年度はすでに実装している子音領域推定機能に加えて母音領域推定アルゴリズムを策定し、音声分析システム VoiceAnalyzer に母音領域推定機能として組み込んだ。また、母音領域の表示・非表示を選択できるようにし、従来同様の簡素な使い易さも維持するようにした。ラベル表示については従来からあるアルファベット 2 文字からなる表示に加え、なるべく平仮名で表示する方式を開発し切り替えられるようにした。これによって、音声の専門家でなくても理解しやすくなった。

また、アイコンを押すだけで音声セグメントラベル系列を CSV ファイルとして保存し、この音声セグメントラベル系列に基づいて筑波技術大学で開発したマッチングプログラムで解析するために、連携起動する機能を実装した。

さらに、静岡県立大学で開発しているモーラ判定機能および音素の妥当性を判定する機能を連携起動するアイコンも実装した。これら機能は Excel ベースのマクロを直接プログラミングすることなく、比較的簡単なユーザインタフェースで開発可能な StIL を用いてプログラミングした。モーラ判定機能および音素判定機能については系統的に分析し、訓練者に理解しやすい形式としてまとめることを次年度以降の目標としていた。

筑波技術大学では、聴覚障害者の矯正のための訓練手法について検討した。すなわち、

自分の発音と模範的な発音とを比較して、異なる個所、同一とみなせる個所を把握し、異なる個所について修正すべく練習する手法である。VoiceAnalyzer をこのような目的に用いるとき、音声訓練の専門家ならば VoiceAnalyzer の出力（音声セグメントラベル系列）を比較して異同を指摘することができる。しかし、必ずしも専門家ではない指導者（たとえば母親）あるいは聴覚障害者自身がこの異同を識別することは容易ではない。ここでは、指導者と訓練者の音声セグメント列を比較して部分ごとの一致不一致を自動検出するプログラムを作成し、訓練者の間違いを指摘する機能の可能性を示した。

##### (2) 平成 23 年度の成果

本年度は音声分析システム VoiceAnalyzer の使い勝手の向上のために収集した音声データの音量が小さすぎる場合に大きさを調整する振幅調整機能を組み込んだ。この機能によって、分析できなかったデータのうちのかなりの部分が使用できるようになった。従来は、別のソフトで対応する必要があり、手間がかかっていた。

また、研究効率の推進のために音声ラベル数を自動的にカウントする機能と多数のデータを分析し、ラベルデータ保存を自動的に実行する、いわゆるバッチ処理機能も実装した。これによって大量のデータを一括で処理できるようになり、研究の効率化が図れる体制を整えた。さらに、特に音韻長の計測精度向上のためにパワー計算機能を実装した。

一方、筑波技術大学で開発したマッチングプログラムと VoiceAnalyzer との連携については音声セグメントラベルのデータの自動保存機能を実装し、使い勝手の向上を図った。

静岡県立大学が担当している、EXCEL のマクロを用いた開発ツール StIL を用いて実装した判定機能については、「胃痛」「一通」について促音の判定、長音の判定、一部の音韻の判定を行い、聴覚障害者にとっても分かりやすいコメント機能によって結果を表示するようにした。

また、StIL のバッチ処理機能を用いて、大量のデータに対する分析を自動化することによって、判定アルゴリズムの改良—評価のサイクルを繰り返し回すことができ、開発を加速化した。

筑波技術大学では、聴覚障害者の 235 単語 × 10 人分のデータを収集した。

##### (3) 平成 24 年度の成果

VoiceAnalyzer は今まで、Windows XP、Vista、

7 への対応であったが、Windows 8 への対応を行った。また、Windows8 で新たに加わったタッチパネルのタッチ入力への対応を行った。タッチ入力については特に子音領域をポイントする範囲がかなり狭く扱いにくいところがあったが、調整することによって、使いやすくなった。また、スマートフォン(スマホ)と PC 上の VoiceAnalyzer の連携による分析・評価システムを構築した。このシステムによって、どこでも気軽に発音評価を行う環境が整った。スマホと発音評価システムとの連携は次の手順で行われる。

- (1) スマホ (Android 端末等) で音声入力 (PCM 録音: 22kHz, 16bits) を行い、WAV ファイルを作成する。
- (2) WAV ファイルをメールに添付し、パソコンへ送付する。
- (3) パソコン (解析 PC) に実装している VoiceAnalyzer で分析し音声セグメントラベル列を求め、それらを発音評価システム入力し評価した解析結果をスマホへメールで送信する。
- (4) スマホで結果を表示する。

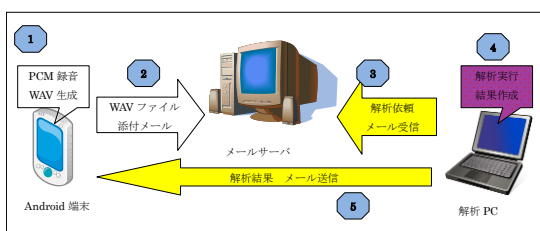


図1 スマホと発音評価システムとの連携

既存機能の「ラベル選択範囲対応」は、ラベル領域上で任意の範囲を選択後に実行することで、選択されたラベル領域に合わせて波形領域上の選択範囲を変更(同期)する、というものであった。

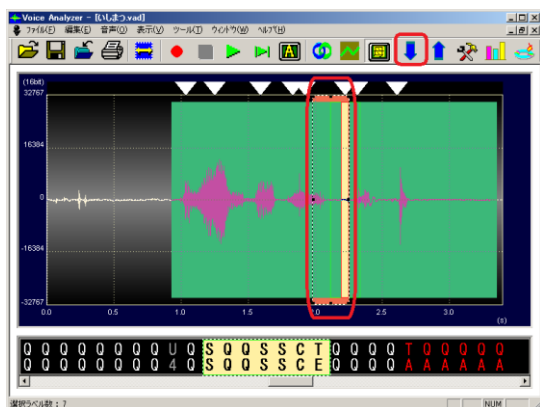


図2 ラベル選択範囲対応

これに対して、「波形選択範囲対応」を新規に追加した。これは「ラベル選択範囲対応」とは逆で、波形領域上で任意の範囲を選択後に実行することで、選択された波形領域に合わせてラベル領域上の選択範囲を同期する。

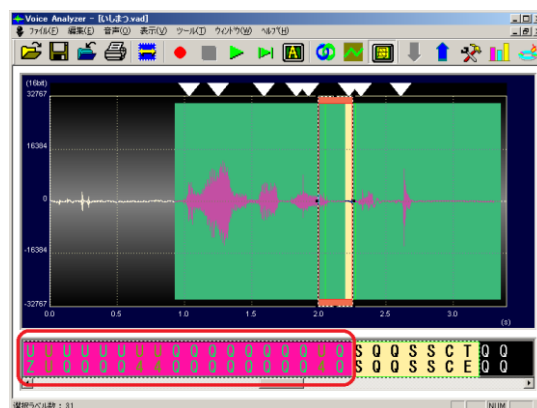


図3 波形選択範囲対応

また、データの一括閉じを作成するなどユーザーインターフェースの改良を行った。すなわち、従来は開いたファイルを一括で閉じるためには、VoiceAnalyzer 自体を閉じるしかなかった。新たにアイコンを用意して、これを実行すればいつでも複数ファイルを一括で閉じられるようにした。

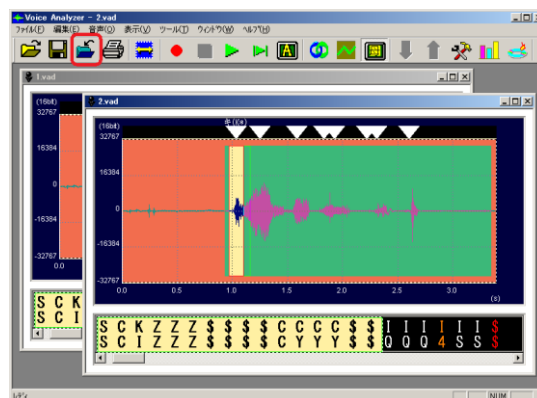


図4 データの一括閉じ

従来のVoiceAnalyzerでは、音声データ中に300msec以上の無音区間があると、自動的に音声区間を分割して別の単語として分析する。日本人の通常の発声である問題ないが、外国人などでは、「さっぽろ」のように促音が含まれる場合に、「さっ」と「ぼろ」で音声区間が分割され別の単語として解釈されてしまうことがあった。

今回、このような促音が含まれる場合でも、音声区間が安易に分割されないようにするた

めに、分割の判断基準である「単語中の無音区間の長さ」をユーザが設定できるようにした。なお、無音区間の長さは、最小300msから最大1000ms間で任意の値とした。

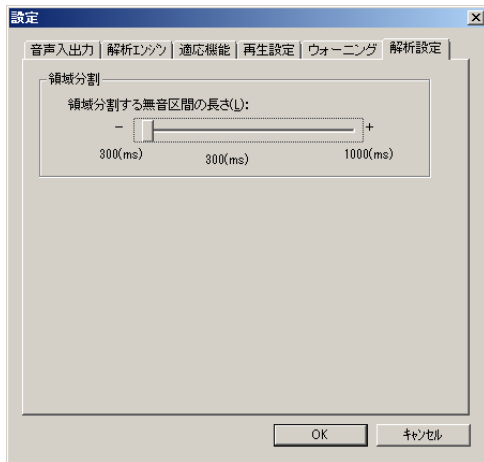


図5 単語中の無音区間長の設定

科研費の終了を控え筑波技術大学を訪問し、聴覚障害者の方にVoiceAnalyzerが使えるかについて検証した。田中哲男教授(昨年度末に退職により分担者から外れた)が23年度にデータ収集した学生のうちの1名が岡崎教授のもとで修士課程に在籍していたので、再発声してもらった。彼は「こんにちは」が「こんにちは」と言うように、省略しているほか、所どころあいまいだったが、発声指導によって改善がみられた。

しかし、「さしすせそ」は「かきくてと」と言うように、全く発声できない。サ行は歯音に分類され、上下前歯間・舌口蓋間に瞬時に狭い空隙を作るのが困難だったためと考えられる。また、一般に聴覚障害者は濁音の発声が難しいようである。聴覚障害者の中には人工内耳を装着したため、かなり聞こえて障害認定されない人もいるようである。今回の話者の中にも確かに発声が、かなり良い人もいるが、それでも弱点は残っている。

筑波技術大学の加藤伸子先生にも同席していただき、聴覚障害者の実態を尋ねたところ、彼らは子供の時に訓練しているが、自分の声が聞こえないため、フィードバックの機会を失い、次第に悪くなるそうである。親に注意されたりすることがあるそうであり、大学に在籍している間にも悪くなることに、先生方も気づかれているとのことである。社会に出ると最初は良くても、だんだん自分が言っていることが理解してもらえないことに気づき、発声に関心を持つ人が多いとのことである。

来年度の4年生で発声に興味がある学生が

おり、ぜひVoiceAnalyzerを使って卒業研究をやってみたいという申し出をいただいた。加藤先生は手話で通訳ができるぐらいの達人であり、聴覚障害の学生ときちんと意思疎通ができるため、科研費の研究期間終了後も継続して効果的に共同研究を行っていきたいと考え、VoiceAnalyzerを提供した。

## 5. 主な発表論文等

[雑誌論文] (計4件)

- ①秀島雅之, 松浦 博: 義歯と発音機能 – 音声認識システムによる発語明瞭度の客観的評価 – ; 東京都歯科医師会雑誌, Vol. 61, 1, pp. 4-13 (2013.1)
- ② Junichiro Wada, Masayuki Hideshima, Shusuke Inukai, Tomohiro Ando, Yoshimasa Igarashi, Hiroshi Matsuura: Influence of the major connector in a maxillary denture on phonetic function; Journal of Prosthodontic Research: Volume 55, Issue 4, Pages 234-242, (2011.10).
- ③Hideshima M, Wada J, Ando T, Inukai S, Igarashi Y, Matsuura H : Assessment of the Consonants Pronunciation in Subjects with Sound Dentition; Journal of Prosthodontic Research , 査読無, 580, special issue B, (2010)
- ④松浦 博, 秀島雅之, 犬飼周佑, 安藤智宏, 和田淳一郎, 五十嵐順正, 永江尚義: 日本語発話の発音誤り検出における留学生と日本人学生との対比 – 音声セグメント技術による「じ」と「ち」の弁別を中心として – ; 日本音響学会誌, 査読有, 66, 370-380 (2010)

[学会発表] (計8件)

- ①T. Kaneshima, J. Wada, S. Inukai, H. Ohmori, K. Shimazaki, M. Hideshima, M. Taira, T. Kurabayashi, Y. Igarashi, H. Matsuura, T. Ono: Effects of the orthodontic appliance on brain activity during syllable pronunciation: An fMRI study. 第77回口腔病学会学術大会, 東京医科歯科大学(2012.12.1).
- ②松浦 博, 秀島雅之, 和田淳一郎, 犬飼周佑, 五十嵐順正: 留学生/日本人学生の促音・長音の発話特徴と音声セグメントラベルに基づく発語自動評価; 日本音響学会2012年秋季研究発表会抄録集, 1-3-13 信州大学 (2012.9.19)
- ③犬飼周佑, 秀島雅之, 和田淳一郎, 安藤智宏, 五十嵐順正, 松浦 博: 下顎部分床義歯

の連結装置の形態が発音に及ぼす影響について；日本補綴科学会 第 121 回学術大会，1-4-6，横浜産業貿易センタービル (2012. 5. 26)

④和田淳一郎，秀島雅之，安藤智宏，犬飼周佑，五十嵐順正，松浦 博：上顎可撤性義歯の大連結子の幅および断面形態が発語機能に及ぼす影響について；日本補綴科学会 第 121 回学術大会，1-4-7，横浜産業貿易センタービル (2012. 5. 26)

⑤佐野友理香，松浦 博，秀島雅之，和田淳一郎，安藤智宏，犬飼周佑，五十嵐順正：音声セグメントラベルに基づく留学生/日本人学生による「胃痛，一通」発話データの自動判定；日本音響学会 2012 年春季研究発表会抄録集，3-Q-8，神奈川大学 (2012. 3. 15)

⑥松浦 博，佐野友理香，秀島雅之，和田淳一郎，安藤智宏，犬飼周佑，五十嵐順正：留学生による「画商，合唱」発話データの音声セグメントラベルに基づく自動評価；日本音響学会 2011 年秋季研究発表会抄録集，1-Q-23，島根大学 (2011. 9. 20)

⑦和田淳一郎，秀島雅之，犬飼周佑，安藤智宏，五十嵐順正，松浦 博：上顎可撤性義歯の大連結子が発語機能に及ぼす影響；第 75 回口腔病学会学術大会抄録集，25，査読無，22，東京医科歯科大学 (2010. 12. 3)

⑧上田聡太. 指導；秀島雅之，和田淳一郎，犬飼周佑，安藤智宏，関西崇史，五十嵐順正，松浦 博：下顎義歯の連結装置の形態が発音に及ぼす影響 —上顎との対比—. 2010 年度東京医科歯科大学歯学部 4 年生研究体験実習学生発表会 ポスター発表，東京医科歯科大学 (2010. 9. 21)

[その他]

ホームページ等

<http://ai.u-shizuoka-ken.ac.jp/~matsuura/>

## 6. 研究組織

### (1) 研究代表者

松浦 博 (MATSUURA HIROSHI)  
静岡県立大学・経営情報学部・教授  
研究者番号：60451085

### (2) 研究分担者

田中 哲男 (TANAKA AKIO)  
筑波技術大学・産業技術学部・教授  
研究者番号：10369308  
(研究期間：2010-2011)

### (3) 研究分担者

岡崎 彰夫 (OKAZAKI AKIO)  
筑波技術大学・産業技術学部・教授  
研究者番号：20516679

### (4) 研究分担者

秀島 雅之 (HIDESHIMA MASAYUKI)  
東京医科歯科大学・歯学部附属病院・講師  
研究者番号：50218723