

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成 25 年 5 月 31 日現在

機関番号：17401
 研究種目：若手研究（B）
 研究期間：2011～2012
 課題番号：23700606
 研究課題名（和文） 調音状態をフィードバックする発話訓練システムの開発
 研究課題名（英文） Development of speech training system
 that feeds back the articulation state
 研究代表者
 坂田 聡（SAKATA TADASHI）
 熊本大学・大学院自然科学研究科・助教
 研究者番号：80336205

研究成果の概要（和文）：

音声のホルマント情報を用い、話者の調音状態を正規化母音構音空間と声道形状表示により視覚フィードバックする発話訓練システムの開発を行った。開発した訓練ツールを聴覚障害児の発話訓練において試用した結果、正規化母音構音空間は構音状態のフィードバックとその定量的な評価において有効であると判断されたが、声道形状表現は口唇部分の計算式において正規化が不十分であったため、その変位にばらつき出ることが確認された。

研究成果の概要（英文）：

We have developed an utterance training system using the formant information of audio, visual feedback to the vocal tract shape display and normalized vowel articulation space articulation state of the speaker. Result of the trial in speech training of hearing impaired children, it is determined normalized vowel articulation space is effective in the quantitative assessment and real-time feedback of articulation state, vocal tract shape representation, formula lip portion since the normalization is insufficient, it is out variation in the displacement is confirmed, it is believed that it is necessary to further improvements.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
交付決定額	1,800,000	540,000	2,340,000

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：音声学，可視化，教育工学，ユーザインターフェース，リハビリテーション

1. 研究開始当初の背景

音声言語の可視化手法の1つとして、複数の音声特徴量を変換し画像パターンとして提示する音声画像の開発が行われてきた。この音声画像は、ホルマント周波数群で定まる色彩パターンに、ニューラルネット出力によって定まる子音特徴量を重畳し、ピッチ周波数の変化（イントネーション）により画像のパターン高を変化させた統合化画像である。本システムは、聴覚的な音声イメージを視覚イメージとして伝達することを目的として設計さ

れたものであり、単語音声表現における読み取りにおいて高い効果が得られることが実験的に確認されている。開発当初、音声画像はマイク入力された音声をおフライン処理し被験者に提示していたが、昨今の汎用PCの演算性能とグラフィック処理性能の向上を背景に、リアルタイム処理を実現したことで、自己の発声を音声画像に変換するリアルタイムフィードバックが可能となった。

われわれは、これまでに、リアルタイム処理音声画像処理に基づく発話訓練システム

を構築し、聴覚障害児・者の発話訓練や、発話障害者の診断補助、外国語母語話者の日本語発話学習等へ実験的に試用を行ってきた。被験者からは、「聴覚的に比較することが困難な自己の発話を、視覚イメージとして比較することで客観的に判断できる」と好意的な意見が聞かれた。その一方、唇や舌等の調音器官をどのように変化させれば正しい発音に近づくの分かり難いとの意見も見受けられた。これは、本システムによって、発話訓練を行う際に、リアルタイムで変化する画像を参考に、被験者が主体的に口形や舌の位置を試行錯誤的に変化させて、正しい調音方法を習得することを目的としているためである。しかし、訓練に際し、発話訓練を補助する人間がいる場合は、調音方法を大まかに示唆することで訓練を円滑にすすめることが可能であるが、発話障害児や外国語母語話者など、日本語の調音に関する知識が不十分な訓練者や訓練者が単独でシステムを利用する場合は、正しい調音がわからないため、混乱する要因となっていると考えられる。

2. 研究の目的

そこで、本研究は、調音の変化に対して声道の共振周波数であるホルマント周波数がどのように変動するのかを明らかにすることを目的とする。ホルマント周波数から調音状態を逆推定することが可能になれば、訓練者に適切な指示を与えることができるため、より訓練効率の良い発話訓練システムに近づくと考えられる。

3. 研究の方法

本研究を遂行するために、以下の様な項目について検討を行った。

(1) 音声分析手法の精度評価

発話訓練が学校のパソコン室等の実環境で実施されることを考慮し、訓練システムのエンジン部分であるホルマント周波数の推定は、安定かつ高精度に行えることが望ましい。そこで、システムの有効性を左右するホルマント推定方式について検討を行う。

(2) 正規化構音空間の提案と構音状態との比較

発話時の構音状態を表現する一般的な方法は、ホルマント周波数そのものを利用するホルマントグラム (F1-F2 空間) であるが、ホルマント周波数は、話者の声道長に依存するパラメータであるため、性別や年齢層によってその分布位置が大きく異なることが知られている。このような問題を解消した手法が、我々が先に開発した音声画像化方式による色彩表現であるが、構音状態と色相の対応関係を事前に学習する必要があるため、児童(特

に就学前や小学校低学年) が利用するには難しいと考えられる。

そこで、構音状態を視覚的にフィードバックする新たな空間を提案し、通常発話や顎を拘束した状態で発声した音声を用いて、提案した空間の構音状態表現能力について検討する。

(3) リアルタイム発話訓練システムの開発と実証実験

発話訓練において、自己発話状態がタイムラグなしにリアルタイムで訓練者にフィードバックされることが望ましい。我々はこれまでに、リアルタイム音声特徴抽出技術の研究を行っており、これらをベースにしたリアルタイム発話訓練ソフトウェアの開発を行い、聴覚支援学校の協力の下、聴覚障害児の発話訓練を行う。

(4) 声道形状表示機能の検討

声道形状を発声手本として呈示する発声学習ツールは知られているが、自発声の声道形状のフィードバックは、音声特徴量からの逆推定問題の複雑さとリアルタイム処理の困難さのために実現困難であった。そのため、古くは簡単な声道図形をホルマント周波数の変化でコントロールする方法が用いられる事もあった。そこで、リアルタイム音声特徴技術に基づき、正規化ホルマントベースでの声道形状表現の検討を行った。

4. 研究成果

(1) 音声分析手法の精度評価

さきに我々が提案した逆フィルタ制御 (IFC) 法は、音声信号の零交差情報を用いてホルマントを推定するため、スペクトル情報を利用する他の推定方式とは全く異なる方式である。

そこで、IFC 法に用いられる零交差情報の有効性を、現在も用いられることが多い線形予測分析 (LPC) 法に用いられるスペクトル情報の比較において明らかにした。まず、両ホルマント推定法を用いて実音声から得た F1-F2 ダイアグラム上の真の正規分布とみなされる領域を明示する「データ縮減法」を提案する。その分布の 3σ 楕円領域の外側にあるデータをホルマント推定エラーと定義することにより、両方式のホルマント推定精度を比較した。その結果、IFC 法が LPC 法よりも優れていることが示された (図 1)。次に、その結果を合成音声を用いたモンテカルロ法の精度チェックにより確認した (図 2)。最後に、分析過程の視察に基づいて、IFC 法が高精度になるのは、励振点付近の不規則な共振が零交差周波数の荷重平均によって取り除かれることに

よることが確認された。

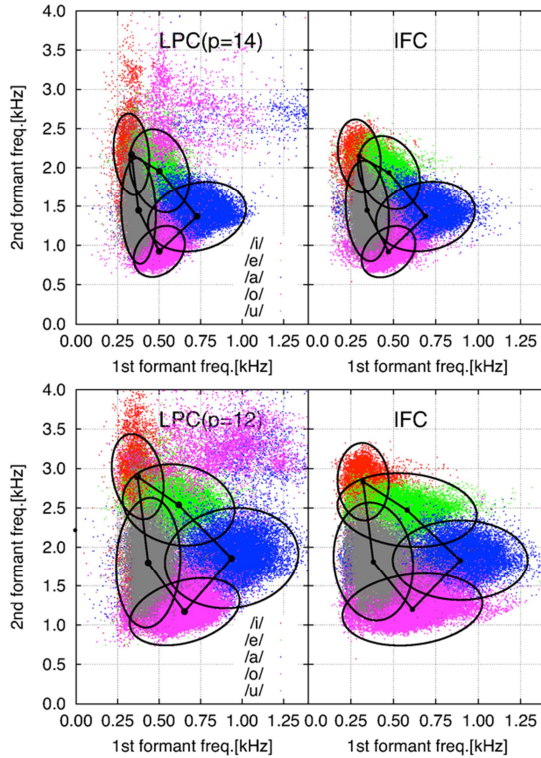


図 1. 各方式によって抽出されたホルマント分布から、データ縮減法によって推定されたデータ分布の 3σ 確率楕円

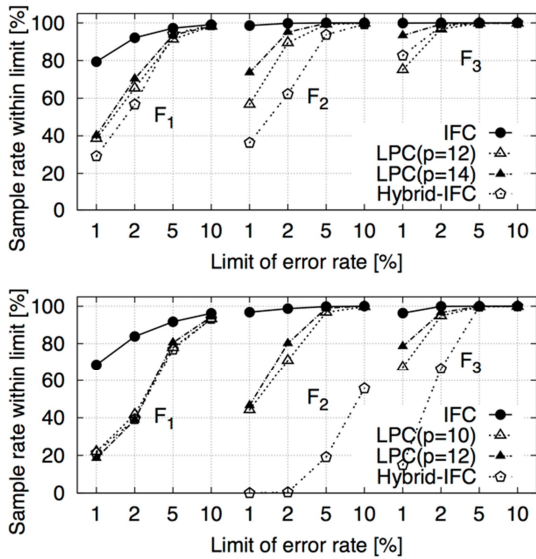


図 2. 合成音声に対して、ホルマント推定値が一定の誤差範囲に入るサンプルの割合

(2) 正規化構音空間の提案と構音状態との比較

先に提案した音声画像化手法は、話者の年齢や性別に関係なく、母音音韻性に関する正規化効果が期待できる方式であった。さらに、発話者が構音を学ぶための空間として、話者正規化と同時に、母音構音時の舌や顎などの幾何学的位置関

係の可視化を考慮した hv 平面 (正規化構

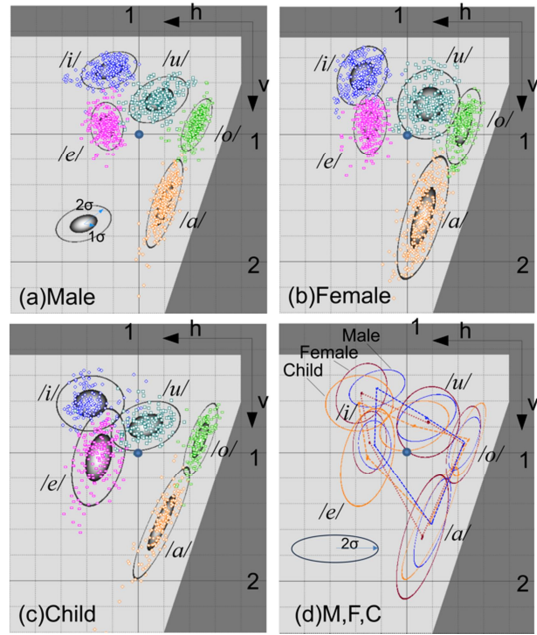


図 3. 各話者群の日本語 5 母音マッピング

音空間) を提案した。この hv 平面において、h 座標は舌の前後位置に、v 座標は舌・顎の高低位置に対応し、中性母音は話者に依存せずに中性点 $(h, v) = (1, 1)$ に位置するという特徴を持つ。また、計算式より、連比 $(F1:F2:F3)$ が等しい母音はマップ上の 1 点に写像される。このような特徴から、ホルマント空間で見られるような、性別や年齢の違いによる分布の分散が減少し、同時に母音発声時の構音位置との関連が直観的に理解できる空間となっている (図 3)。

この hv 平面の構音表現能力の検証を目的として、構音障碍の一つである顎変形症を模擬するために、顎の開きを拘束した状態での発声音声を hv 平面で分析した。その結果、開口率を変化させることで v 値が変化することが確認され、構音時の顎の開きを適切に表現できることが確認された。また、h 値は、拘束条件により大きな変化が見られなかったことから、拘束の有無による舌の前後位置の変化は少ないと考えられるが、拘束による hv 値の遷移から構音状態の推定が行えることが示された (図 4)。

(3) リアルタイム発話訓練システムの開発と実証実験

(2) において提案した、hv 構音空間を利用する発話訓練ツールの構築を行った。本ツールは、Windows-PC 上で動作し、サウンド入出力処理とグラフィック処理には Win-API を利用している。ホルマントを含む音声特徴量は、10ms フレーム周期で抽出され、hv 構音図上にリアルタイム表示される。また、児童の

発話学習を想定し、hv 座標情報と併せて画面

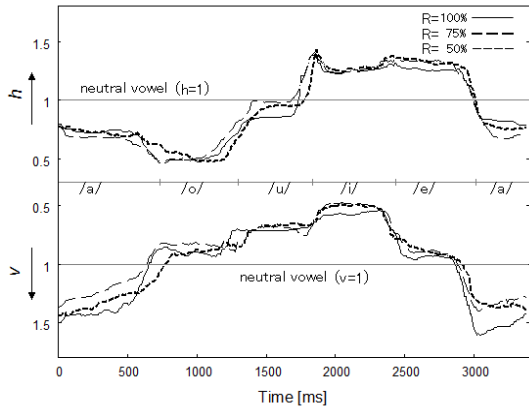


図 4. 拘束発話状態における hv 値の変化 (発話者:成人男性 M4, 発話音声:/aouiea/)

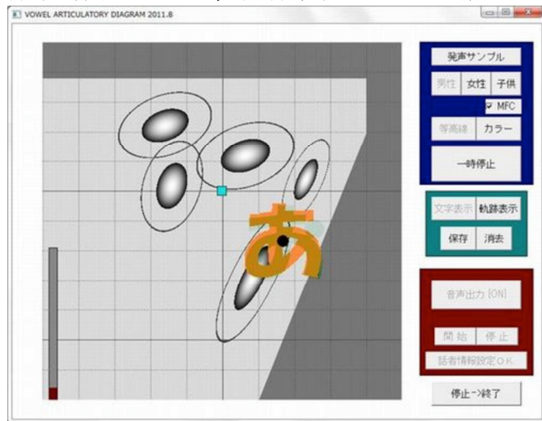


図 5. 母音構音発話ツールの表示例 (女声 /a/)

上に音韻情報として文字画像を表示するモードも付加した(図 5).

ここで開発した母音構音学習ツールを、聴覚支援学校の協力のもと、聴覚障碍児の発声学習に利用した。被験者は、小学 5 年の男女児各 1 名であり、4 分法による平均聴力レベル男児 M が右 95dBHL, 左 84dBHL, 女児 T が右 99dBHL, 左 100dBHL と両者とも高度の難聴に分類される。学習期間は約 1 年間で、通常の指導に加えてツールを使った訓練を 2 週間に 1 回(45 分)程度の割合で実施した。訓練は、指導教員と一緒にいき、画面上に表示される構音空間における自己発話の座標点を確認しながら、舌や顎の動きを修正するという方式で行った。発話訓練効果を、ホルマントの軌跡と hv 空間における母音データ群の分離度によって定量的に評価した。男児 M は訓練の前後でホルマント軌跡の分散値が減少し、hv 空間に於ける母音分離度が 2.7 倍向上した。女児 T は、ホルマントの分散は減少したが、分離度は 1.6 倍程度の向上にとどまった。主観的な聴覚印象としては、両者とも母

音の音韻性や明瞭度の向上が確認された。また、指導教員の感想として、被験者が

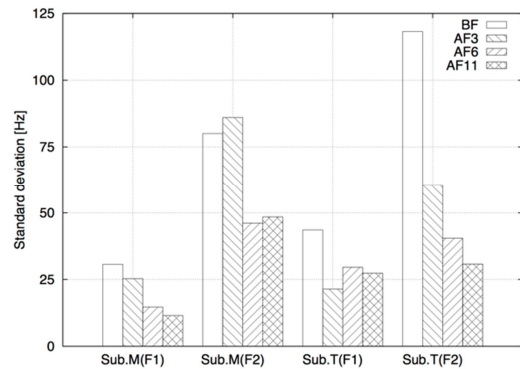


図 6. 訓練の経過に伴うホルマント軌跡における標準偏差の変化

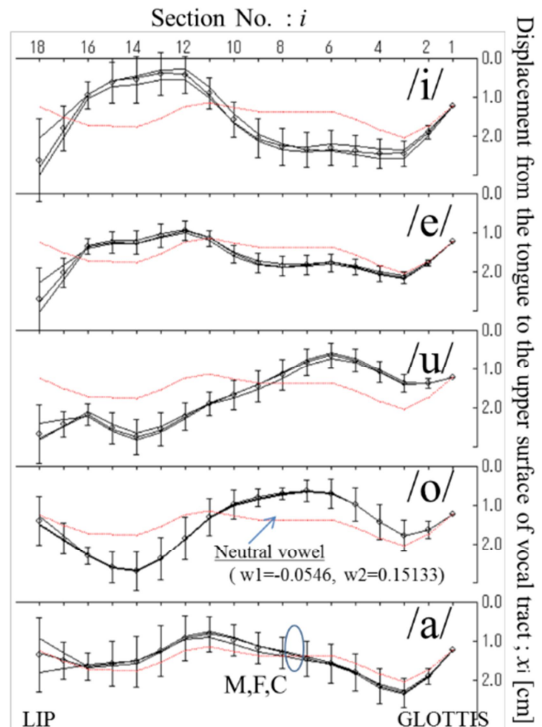


図 7. 実音声から推定した日本語 5 母音の平均声道形状

自己の発話を自分で評価できることから、訓練に取り組む姿勢に変化が見られたとの好意的な意見が寄せられた。

(4) 声道形状表示ツールの作成

構音状態の直接的な可視化手法としては正規化ホルマントベースでの声道形状表現の検討を行った。これには、P. Ladefoged (1978) による実験的なモデルを採用し、これまでの研究において作成しているリアルタイム音声特徴ベクトル推定エンジンによって出力される高精度ホルマント情報を用いて、上部構音器官各部の点から舌体表面点までの変位を関数によってモデル化する手法である。本表現方式の適切さについて、日本語母音

の平均的なホルマントデータを用いて検討した結果、各点における話者間の形状

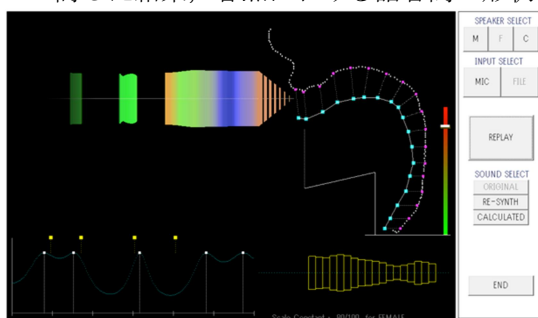


図 8. 母音色彩及び声道形状表示ツール

変動は比較的小さい範囲に収まっているが、唇と歯の部位における変動が話者群間で大きいことが示されたため、今後はこの影響をどのように補正するかを検討する必要がある(図 7)。最後に推定された声道形状を発話訓練システムに組み込み、リアルタイム動作することを確認した(図 8)。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 4 件)

- ① 上田裕市, 坂田聡, 新中須真奈, 平原成浩, 五味晁憲, “構音障碍診断支援のための音声特徴量の統合化と可視化に関する検討,” 電子情報通信学会技術研究報告, SP2011-162, pp. 39-44, 2012, 査読無
- ② 坂田聡, 上田裕市, 渡邊亮, “母音ホルマント分布に基づく男女間平均声道長比の推定,” 電子情報通信学会技術研究報告, SP2012-09, pp. 47-52, 2012, 査読無
- ③ 坂田聡, 上田裕市, 渡邊亮, “逆フィルタ制御ホルマント推定における零交差情報の有効性,” 日本音響学会誌, Vol. 67, No. 10, pp. 459-469, 2011, 査読有
- ④ 坂田聡, 佐伯勇哉, 柴田航, 上田裕市, “母音発声のリアルタイム視聴覚フィードバックのための正規化構音空間の検討とその応用,” 電子情報通信学会技術研究報告, ET2008-27, pp. 55-60, 2011, 査読無

[学会発表] (計 8 件)

- ① 坂田聡, 三代国将, 上田裕市, “正規化構音マップでの視覚フィードバックを用いた母音構音学習ツールの開発,” 日本音響学会 2013 年春季研究発表会, 2013. 3. 13-15, 東京工科大学八王子キャンパス(東京)

- ② 三代国将, 榎木涼太, 平金紋佳, 坂田聡, 上田裕市, “正規化構音図を用いた母音構音学習とその安定性の評価手法の検討,” 第 38 回感覚代行シンポジウム, 2012. 12. 3-4, 独立行政法人産業技術総合研究所臨海副都心センター(東京)
- ③ 金子卓, 山本恭磨, 坂田聡, 上田裕市, “構音診断支援ツールにおける母音の構音特徴表現に関する検討,” 平成 24 年度電気関係学会九州支部連合大会, 2012. 9. 23-24, 長崎大学文教キャンパス(長崎)
- ④ 上田裕市, 佐伯勇哉, 坂田聡, “聴覚障碍児の母音発声学習におけるホルマント情報のリアルタイム呈示様式とその系統的利用法,” HCG シンポジウム, 2011. 12. 7-9, サポートホール高松(香川)
- ⑤ 佐伯勇哉, 坂田聡, 上田裕市, “カラー構音マップ上での発声フィードバックによる母音発声学習システムの開発,” 第 37 回感覚代行シンポジウム, 2011. 12. 5-6, 独立行政法人産業技術総合研究所臨海副都心センター(東京)
- ⑥ 佐伯勇哉, 坂田聡, 上田裕市, “構音マッピングによる母音可視化法を用いた発声フィードバックに関する検討,” 日本音響学会九州支部・第 9 回学生のための研究発表会, 2011. 11. 26, 大分大学旦野原キャンパス(大分)
- ⑦ 佐伯勇哉, 坂田聡, 上田裕市, “母音構音可視化のための構音空間マッピング手法とその応用,” 平成 23 年度電気関係学会九州支部連合大会, 2011. 9. 27, 佐賀大学本庄キャンパス(佐賀)
- ⑧ 坂田聡, 新中須真奈, 平原成浩, 上田裕市, “額変形症音声の定量評価手法の検討,” 日本音響学会 2011 年秋季研究発表会, 2011. 9. 20-22, 島根大学松江キャンパス(島根)

6. 研究組織

(1) 研究代表者

坂田 聡 (SAKATA TADASHI)
熊本大学・大学院自然科学研究科・助教
研究者番号：80336205

(2) 研究分担者

(3) 連携研究者

上田 裕市 (UEDA YUICHI)
熊本大学・大学院自然科学研究科・教授
研究者番号：00141961