

音声の構造的表象に基づく発音矯正必要度の計算手法の検討

鎌田 圭[†] 朝川 智[†] 峯松 信明[†] 牧野 武彦^{††} 広瀬 啓吉^{†††}

[†] 東京大学大学院新領域創成科学研究科 〒277-8561 千葉県柏市柏の葉 5-1-5

^{††} 中央大学経済学部 〒192-0393 東京都八王子市東中野 742-1

^{†††} 東京大学大学院情報理工学系研究科 〒113-0033 東京都文京区本郷 7-3-1

E-mail: [†]{k-kamata, asakawa, mine, hirose}@gavo.t.u-tokyo.ac.jp, ^{††}mackinaw@tamacc.chuo-u.ac.jp

あらまし 音声コミュニケーションは、音声の生成、収録、伝送、再生、聴取の何れの過程においても非言語的特徴が不可避免的に混入するが、これを表現する次元をおよそ保有しない音響的普遍構造が提案されている。この構造的表象を用いて、発音矯正の必要度を学習者別に推定可能であることが実験的に示されている。本稿では、この推定方法が教育的に妥当であるかを、英語音声学を専門とする音声学による学習者発音の母音図表記を通して検討する。また、この推定方法について、改善可能な点を示し、より妥当な推定方法を検討する。

キーワード 音声の構造的表象, CALL, 母音矯正度, 英語音声学, 相関分析

Automatic estimation of the vowel correction priority for language learners based on the structural representation of pronunciation

K. KAMATA[†], S. ASAKAWA[†], N. MINEMATSU[†], T. MAKINO^{††}, and K. HIROSE^{†††}

[†] Grad. School of Frontier Sciences, Univ. of Tokyo, 5-1-5, Kashiwanoha Kashiwa, Chiba, 277-8561 Japan

^{††} Faculty of Economics, Chuo University, 742-1, Higashi-nakano, Hachioji, Tokyo, 192-0393 Japan

^{†††} Grad. School of Info. Sci. and Tech., Univ. of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0031 Japan

E-mail: [†]{k-kamata, asakawa, mine, hirose}@gavo.t.u-tokyo.ac.jp, ^{††}mackinaw@tamacc.chuo-u.ac.jp

Abstract Speech communication induces inevitable variations of speech acoustics due to differences in speakers, microphones, etc. Recently, a novel acoustic representation of speech was proposed where these variations were effectively suppressed. In other words, the dimensions indicating the variations were mathematically removed. Using the new structural representation for CALL, automatic evaluation of the pronunciation was investigated. In this study, some correlation analysis is done using the vowel charts of the individual learners plotted by an expert researcher of English phonetics. Further, an improved method of estimating the vowel correction priority is devised.

Key words structural representation of speech, CALL, priority in vowel correction, English phonetics, correlation analysis

1. ま え が き

従来の音声工学システム構築においては、音響音声学に基づき音声スペクトルをその物理表象として用いてきた。しかし、音声生成の際には話者の声道形状の特性、収録や伝送の際には音響機器の特性、さらには聴取の際には聴取者の聴覚特性、といった非言語的特徴が音声には不可避免的に混入する。このため、音声スペクトルに基づいて構築されたシステムは、不可避な歪みを内包することとなり、不一致問題を抱えることになる。例えば、従来構築された CALL (Computer Aided Language Learning) システムは、不安定性が指摘されており、教育導入

に対して懐疑的な報告もされている [1]。

近年、不可避免的に混入する非言語的特徴を表現する次元そのものを保有しない「音響的普遍構造」が提案された [2]。本表象は、複数の音声事象の関係のみに着眼する物理表象であり、音素の相対的定義に基づいた音響モデリングである。この新しい音声の構造的表象を発音学習に応用することにより、英語学習者の発音状態を非言語的特徴の影響を受けずに記述可能となることが示されている [3]。これにより、従来の手法では不可避であった不一致問題を解消した発音教材を作成することが可能である。

この構造的表象に基づく発音教材の一つとして、発音矯正順

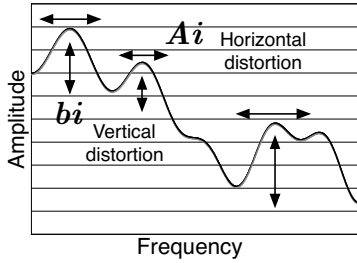


図1 スペクトルに対する水平/垂直方向の音響歪み

序の自動推定が提案されている [3]. この自動推定では、発音事象間の距離のみを用いて発音矯正の必要度を推定するという特殊な手法が用いられている. 音声の構造的表象では、物理実体は捨象されており、発音矯正の必要度を推定する際に頼るべき絶対的な空間、即ち座標軸を持ち合わせていないからである.

本稿では、絶対的な座標軸を持ち、かつ非言語的特徴を有さない母音図に着目し、発音矯正必要度の推定方法の妥当性を検証する. このため、従来より使用している音声資料に対して、英語音声学を専門とする音声学者（第四著者）による母音図作成を行った. そして、得られた母音図を用いて、絶対的な座標軸を基準として得られる発音矯正必要度と、提案されている推定手法から得られる発音矯正必要度との相関分析を行うことにより、妥当性を検証する. また、この提案されている推定手法を見直し、より高い妥当性が得られる推定手法を検討する.

2. 音声の構造的表象

2.1 音声に不可避的に混入する非言語的特徴

音声に不可避に混入する非言語的特徴は、乗算性歪みと線形変換性歪みの2種類に分類される. 乗算性歪みとしては、マイクロフォンなどの伝送特性や話者の声道形状の違いの一部が挙げられ [4]、ケプストラムベクトル c に対するベクトル b の加算 $c' = c + b$ で表現される. 線形変換性歪みとしては、声道長の差異や聴覚特性の差異が挙げられ [5] [6]、 c に対する行列 A の乗算 $c' = Ac$ に相当する. 以上より、音声に不可避的に混入する非言語的特徴は、ケプストラムベクトル c に対して $c' = Ac + b$ としてアフィン変換でモデル化することができる.

なお、音声に混入する非言語的特徴として、背景雑音に代表される加算性雑音が挙げられる. しかし、これは必ずしも不可避ではないため、本研究では対象外とする.

2.2 音声に内在する音響的普遍構造

空間内に存在する n 点に対して、 ${}_nC_2$ 個の全ての二点間距離（即ち距離行列）を求めると、その n 点で張られる構造は一意に規定される. 即ち、事象群に対して距離行列を求めると、その事象群を構造的に規定することになる. ケプストラム空間内の n 点に対して構造を考えた場合、その構造は非言語的特徴（アフィン変換）によって歪みを生じるが、非言語的特徴に不変な距離尺度を介することにより、不変な構造を得ることが可能となる. つまり、事象群が観測される空間を歪ませることで、歪んだ構造から不変な構造を抽出するのである.

音声事象を全て分布として記述し、二事象間距離としてパッチャリヤ距離を考える. 二つの分布の確率密度関数をそれ

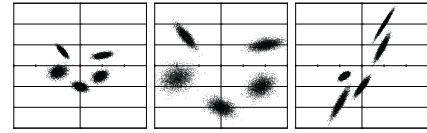


図2 音声に内在する音響的普遍構造

ぞれ $p_1(x)$, $p_2(x)$ とすると、

$$BD(p_1(x), p_2(x)) = -\ln \int_{-\infty}^{\infty} \sqrt{p_1(x)p_2(x)} dx \quad (1)$$

と表される. $0 \leq \int_{-\infty}^{\infty} \sqrt{p_1(x)p_2(x)} dx \leq 1$ を確率として解釈すれば、式 (1) は自己情報量となり、単位は [bit] となる. また、二つの分布がガウス分布で表現されているとき、

$$BD(p_1(x), p_2(x)) = \frac{1}{8} \mu_{12}^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} \mu_{12} + \frac{1}{2} \ln \frac{(|\Sigma_1 + \Sigma_2|/2)}{|\Sigma_1|^{\frac{1}{2}} |\Sigma_2|^{\frac{1}{2}}} \quad (2)$$

となる. μ_{12} は $\mu_1 - \mu_2$ である.

さて、式 2 に対して共通の如何なるアフィン変換^(注1)を施しても、距離不変であることが知られており、これは、非言語的特徴による歪みに対して距離不変となる. 例えば図 2 に示す三つの五分分布群は全て等しい距離行列を持つ. これらの分布群はユークリッド平面上で観測すれば異なる構造として観測されるが、歪んだ非ユークリッド曲面上で観測すれば、不変構造となる. これは話者/マイクに普遍に観測され、音響的普遍構造と呼ぶ.

3. 音声の構造的表象に基づく発音矯正必要度

本節では、音声の構造的表象を用いて発音矯正必要度を導出し、発音教示を自動生成する方法について、実験を通して述べる [3]. まず、実験に用いる音声資料について述べ、次に、発音矯正必要度を得るための導出手法について述べる. 最後に、音声資料にその導出手法を適用した結果を見て、この手法が適切であることを確認する.

3.1 日本人学習者の模擬音声の収録

音声資料として英語を学習する日本人学習者の音声を用意した. この際、後の考察のために、日本人学習者の音声を模擬したものを使用した.

日本語・英語双方を話せる話者（帰国子女や英語劇話者）の日本人 12 名（男性 6 名、女性 6 名）の音声を収録した. 収録した音声は、米語 11 単母音と日本語 5 母音である. 収録の際には、米語は /bVt/、日本語は /bVto/ を発声させた. なお、米語は /bVt/ という発音の有意味語が存在しない場合は、/b/ を /p/、/t/ を /d/ とするなどの処置を行った. また、後述の作業のために、日本語は各単語を 5 回収録した.

収録した米語母音と日本語母音を同一話者内で置換することにより、米語教師や日本人米語学習者の音声を模擬した. 置換パターンには表 1 を用い、1 話者の音声から 7 人の日本人米語学習者 (P1~P7) と 1 人の米語教師 (P8) 相当の発音を用意した. なお、米語と日本語の母音置換には、表 2^(注2) のように

(注1)：さらには一対一対応であれば非線形写像に関しても不変である [7].

(注2)：日本語母音は /a/, /i/, /u/, /e/, /o/ と表記するのが慣習であるが、本稿では米語母音との混同を避けるため、/あ/, /い/, /う/, /え/, /お/ と表記した.

表 1 母音置換によって模擬された 8 種類の発音状態

	ɑ	æ	Λ	ə	ɜ	ɪ	ɪ	ʊ	u	ε	ɔ
P1	J	J	J	J	J	J	J	J	J	J	J
P2	A	A	A	A	A	J	J	J	J	J	J
P3	J	J	J	J	J	A	A	A	A	A	A
P4	A	A	J	J	J	A	A	J	J	A	A
P5	J	J	A	A	A	J	J	A	A	J	J
P6	A	J	A	J	A	J	J	J	J	A	A
P7	J	A	J	A	J	A	A	A	A	J	J
P8	A	A	A	A	A	A	A	A	A	A	A

A : 米語母音を使用, J : 日本語母音で置換

表 2 日本語母音・米語母音の置換表

日本語母音	↔	米語母音
/あ/		/æ/, /Λ/, /ɑ/, /ɜ/, /ə/
/い/		/ɪ/, /ɪ/
/う/		/u/, /ʊ/
/え/		/ε/
/お/		/ɔ/

日本人学習者によく見られる置換パターンを用いた。異なる米語母音が同一の日本語母音と置換される場合は、同一母音・異発声のサンプルを用いた。

このようにして作成された 96 人（12 話者×8 状態）の米語母音発音群に構造的表象を適用することにより、96 個の米語発音構造（11×11 距離行列）を得る。

3.2 発音構造間の要素差異に基づく発音矯正必要度の導出

異なる二つの構造を比較し、構造を構成する各要素（構造の頂点）の構造歪みを求める。構造歪みは各要素が構造を歪ませることへの貢献度を表し、二つの構造を一致させるための障害の大きさを表す指標である。米語教師の発音構造と日本人米語学習者の発音構造を比較した場合、この構造歪みは当該母音に関する発音矯正必要度と捉えることができる。

二つの構造 S と T の、要素 v の構造歪みを次式で定義する。

$$d(S, T, v) = \sum_{j=1}^m |s_{vj} - t_{vj}| \quad (3)$$

m は構造を構成する要素の数（前節の米語発音構造ならば 11）、 s_{ij} と t_{ij} は構造（距離行列） S と T の要素 i と要素 j の要素間距離（バタチャリヤ距離）を表す。

3.3 構造的表象に基づく発音矯正必要度

式 (3) を前述の米語発音構造に適用した結果が、図 3 である。図中の P1~P8 は、12 話者平均構造である。各図は米語教師相当の構造 P8 と日本人学習者状態の P1~P7 を比較したときの構造歪みを表している。教師と学習者の比較であるので、この構造歪みは各母音に対する発音矯正必要度と捉えることができる。図の縦軸は構造歪み（発音矯正必要度）を表し、色づけられているものは日本語母音に置換した母音、白抜きのものは米語発音の母音である。

半分の母音が置換された発音状態である P2~P7 について見てみると、日本語に置換された母音の多くが、米語母音よりも高い発音矯正必要度を示している。また、すべての母音が置換された P1 に関しても、日本人が発音を苦手とする /ɜ/, /æ/, /ə/, /ɪ/

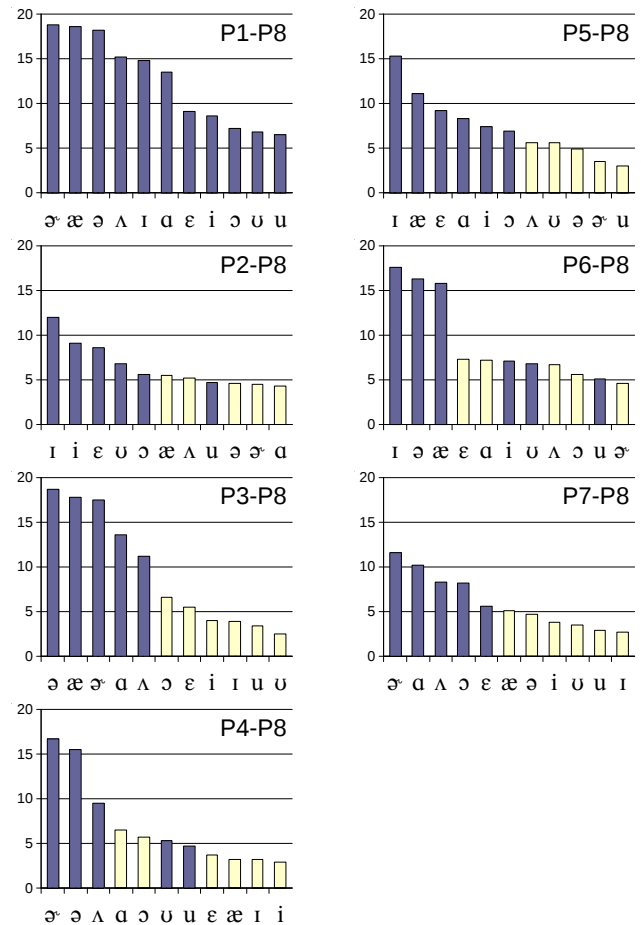


図 3 自動推定された母音矯正順序

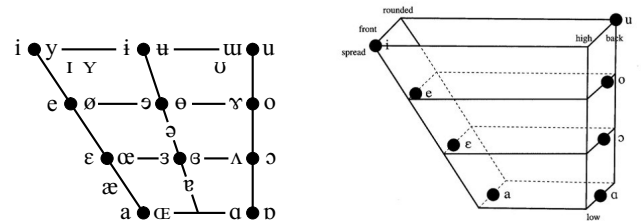


図 4 母音図

が高い結果となった。このことより、構造的表象に基づいて各母音の発音矯正必要度を導出可能であることが実験的に示された。

4. 母音図から見る発音矯正必要度の妥当性

この節では、前節で提示された発音矯正必要度の導出手法を母音図を通して見るにより、妥当性の検討を行う。まず、前章の音声資料を母音図化する。そして、2 通りの方法で各母音の発音矯正必要度を評価し、この 2 つの評価に対して相関分析を行うことで、発音矯正必要度の妥当性を検証する。

4.1 音声資料の母音図化

4.1.1 母音図

母音図（図 4）とは、調音音声学で用いられる母音の発音を視覚化する際に用いられる図である。母音図では、母音発音時の舌の最高点を記述することで、発音をプロットする。

母音図空間は 2 つの特徴を持つ。一つは、発音から話者性などの非言語的特徴が取り除かれているということである。もう

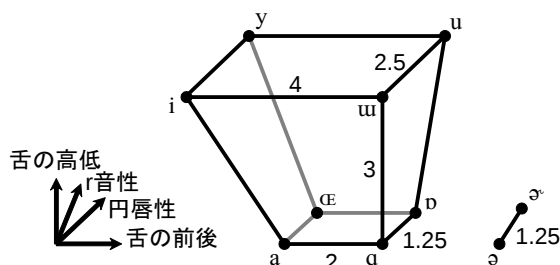


図5 母音図に基づく母音間距離を算出する際の次元の重み

一つは、母音図空間には座標軸が存在し、各発音は座標値を持つということである。この2種類の特徴により、教師の発音と学習者の発音を直接比較して発音矯正必要度を導出することが可能である。構造的表象では、座標軸を持たないために、発音矯正必要度を相対的な関係に頼らざるを得なかった。だが、母音図表象では、直接比較による絶対評価と相対的な関係による相対評価の両方が可能である。

4.1.2 音声資料の母音図化

前章の音声資料（96人×11母音）を英語音声学の専門家に母音図を用いて評価してもらった。母音図化を担当した英語音声学の専門家は、[8], [9]などの日本人向けの発音教則本や、[10]などの英語辞書の発音記載を担当する第4著者である。

発音の評価は4つの評価軸で行った。1) 舌の前後、2) 舌の高低（口の開き具合）、3) 円唇性、4) r音性の4つである。まず、母音図上にプロットすることで、発音時の舌の前後、高低について評価した。円唇性に関しては無円唇、弱い円唇、強い円唇の3段階で評価した。r音声に関してはr音声あり（+）、r音声なし（-）の2段階で評価した。

4.1.3 母音図上の距離尺度

母音図上に表現された発音間の距離には、ユークリッド距離を用いた。本来ならば、各母音間の知覚的距離などを参考すべきであるが、ここでは、母音図という従来より調音音声学において広く使われる母音群の構造的表示形式（4:3:2の台形／2次元表示）を重要視し、舌の前後／高低については台形表示におけるユークリッド距離を用いることとした。これに円唇性、r音性を付与することで4次元空間の二点間距離として母音間距離は定義できる。しかし、各次元の尺度が異なるため、追加する2次元には適当な重み付けが必要である。円唇性、r音性の重みに対しても調音音声学における構造的表示形式からの要請に基づいた重みを議論すべきだが、これについては参考となる議論が見つからなかったため、各母音間のケプストラム距離を参考にして、円唇性、r音性の距離における重みを決定した。

最終的に4:3:2の台形／2次元表示に対して、円唇性の次元は無円唇・強円唇の距離を2.5に、r音性の次元は、r音性の有無の距離を1.25として設定した（図5）。

4.2 母音図に基づく発音矯正必要度

前述の作業によって得られた96枚の母音図を用いて発音矯正必要度を導出する。発音矯正必要度の導出は、以下の2通りの方法で行う。

4.2.1 発音の絶対的評価による発音矯正必要度

ある母音図を教師の発音、異なる母音図を学習者の発音とす

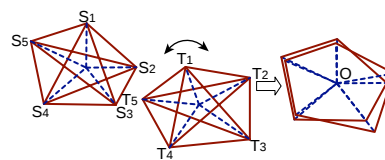


図6 回転とシフトによる構造の重ね合わせ

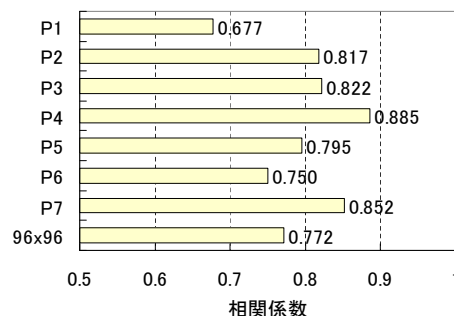


図7 絶対的評価と相対的評価の相関係数

る。この二つの母音図を用いて発音矯正必要度を考える場合、発音を表す座標値を用いて、教師の発音と学習者の発音を直接比較し、その距離をもって発音矯正必要度とする方法が考えられる。つまり、二つの母音図を母音図の軸（枠）を重ねて比較し、発音矯正必要度を導出する方法である。本稿では、この方法で導出された発音矯正必要度を、母音図の軸（枠）という絶対的な基準に基づいてることから、「絶対的評価」と称する。

4.2.2 発音の相対的評価による発音矯正必要度

次に、構造的表象で用いた式（3）により、発音矯正必要度を導出することについて考える。これは母音図の軸（枠）が消失し、母音だけがプロットされた状態で、二つの発音群を比較し発音矯正必要度を推定する問題となる。この場合、図6のように構造を回転&シフトさせ、二つの構造を重ね合わせた後に、発音矯正必要度を推定することとなる。構造的表象と同様にすべての発音間距離を求めることにより、母音図に基づく発音間距離行列を得る。そして、この距離行列に対して式（3）を適用し、発音矯正必要度を導出する。本稿では、この方法で導出された発音矯正必要度を、発音間の相対的な関係に基づいてることから、「相対的評価」と称する。

4.3 絶対的評価と相対的評価の相関分析

前述の絶対的評価と相対的評価に対して相関分析を行うことで、構造的表象における発音矯正必要度の導出式（3）の妥当性について検討する。妥当性の評価には相関係数を用いる。

まず、ある話者のP8（米語教師相当の発音）と、同一話者のP1（日本人米語学習者の発音）を比較し、発音矯正必要度を2通りの方法（絶対的評価と相対的評価）で求める。この12組×11母音に対して絶対的評価と相対的評価の相関分析を行う。同様に、P8と同一話者のP2, P3, ..., P7を比較し、相関分析を行う。最後に、教師も学習者も限定せず、96人すべてを教師・学習者として相関分析を行う（ ${}_{96}C_2 \times 11$ 母音）。

4.4 考察

各比較の相関係数についてまとめたものが図7である。横軸は相関係数である。また、“96x96”は96人すべてを教師・学習者とした場合を表す。

96人総当たりの場合の絶対的評価と相対的評価の相関係数は0.772であり、強い相関が確認された。この事実から、式(3)は一定の信頼性があると考えられる。

P8とP1,P2,...,P7を比較した場合の相関係数について見てみると、P2~P7は相関係数0.750~0.885と高い水準を示しているのに対し、P1のみが相関係数0.677と他の学習者群よりも弱い相関であることがわかる。これにはP1がP2~P7と比べてP8から乖離した発音構造であることが起因していると考えられる。P1はすべての母音を日本語発音に置換しており、置換量が半分であるP2~P7よりも歪んだ発音構造となっている。

5. 構造歪みを用いた発音矯正必要度の再導出

この節では、発音矯正必要度を導出する式(3)を改善する方策について検討する。まず、式(3)の問題点について述べ、これを改善する導出式を示す。そして、改善式によって導出された発音矯正必要度に対して、前節と同様の相関分析を行い、改善の効果を確認する。

5.1 構造歪みに基づく重み付け

構造歪みを測る式(3)は、構造のすべての要素を等価に扱っている。これは式(3)を行列の演算に書き換えた式(4)を見れば明らかである。

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{pmatrix} = \begin{pmatrix} 0 & \delta_{12} & \dots & \delta_{1m} \\ \delta_{21} & 0 & & \delta_{2m} \\ \vdots & & \ddots & \vdots \\ \delta_{m1} & \delta_{m2} & \dots & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad (4)$$

ここで、 d_v は構造 S と構造 T を比較したときの要素 v の構造歪みを表す。また、 δ_{ij} は、要素 i と要素 j の要素間距離の差($\delta_{ij} = |s_{ij} - t_{ij}|$)である。

式(4)において、すべての要素が1からなる列が存在するが、この列は各要素の重みと考えられる。すべての重みが一致している式(4)ではすべての要素は等価に扱われている。 v 以外の要素がおよそ正しい場合、この扱いは妥当であるが、正しくない要素がある場合は妥当性が低下することになる。そこで、正しいと思われる要素に重みを置き、正しくないと思われる要素の影響度を下げる導出手法を試行する。

ここでは、式(4)によって導出された構造歪み d_v を要素の正しさを測るための値として利用し、構造の各要素の重み w_v を求める。重み w_v を求める式を次式のように定める。

$$w_v = c \times \frac{1}{d_v} \quad (5)$$

c は重み列正規化係数($\sum_{i=1}^m w_i = m$)である。式(5)によって求められた重みを用いて、構造歪み d'_v を再導出するには次式を用いる。

$$\begin{pmatrix} d'_1 \\ d'_2 \\ \vdots \\ d'_m \end{pmatrix} = \begin{pmatrix} 0 & \delta_{12} & \dots & \delta_{1m} \\ \delta_{21} & 0 & & \delta_{2m} \\ \vdots & & \ddots & \vdots \\ \delta_{m1} & \delta_{m2} & \dots & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} \quad (6)$$

さらに、 d_v の代わりに d'_v を用いて、重み w'_v を得ることによ

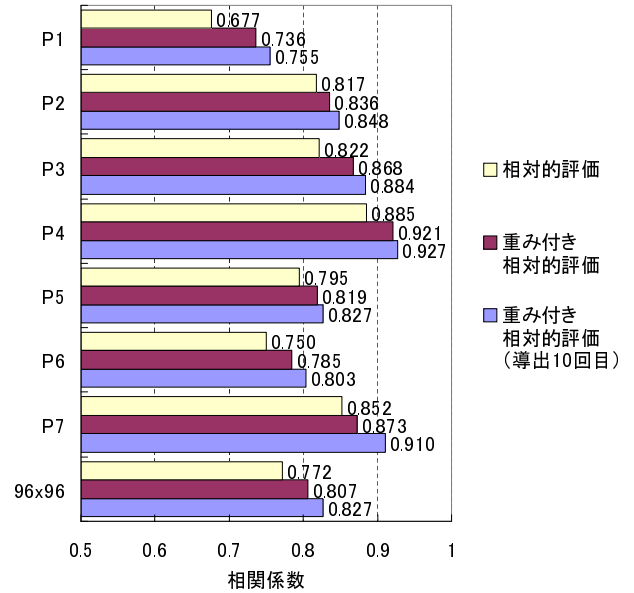


図8 絶対的評価と重み付き相対的評価の相関係数

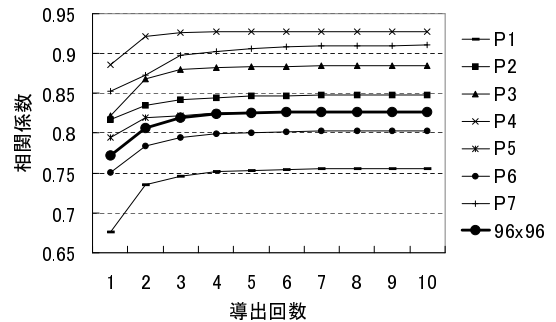


図9 導出に伴う相関係数の推移

り、 d''_v を定義することも可能である。式(4)で導出された要素 v の構造歪み d_v を相対的評価としたことに倣って、式(6)で再導出された構造歪み d'_v を「重み付き相対的評価」と本稿では称する。

5.2 絶対的評価と重み付き相対的評価の相関分析

前節と同様に、絶対的評価との相関分析を行うことで、式(6)の効果を測る。

まず、ある話者のP8(米語教師相当の発音)と、同一話者のP1(日本人米語学習者の発音)を比較し、重み付き相対的評価を求める。この12組×11母音に対して絶対的評価と重み付き相対的評価の相関分析を行う。同様に、P8と同一話者のP2,P3,...,P7を比較し、相関分析を行う。最後に、教師も学習者も限定せず、96人すべてを教師・学習者として相関分析を行う(${}_{96}C_2 \times 11$ 母音)。

求められた相関係数をまとめたものが図8である。

5.3 考察

96人総当たりのときの絶対的評価と重み付き相対的評価の相関係数は0.807となり、相対的評価の相関係数0.772と比べ、3.5ポイント改善した。また、P8とP1,P2,...,P7を比較して相関分析をした場合でも、すべての場合で改善を示した。これにより式(6)の有効性が示された。

次に、式(5)の相対的評価 d_v の代わりに重み付き相対的評価 d'_v を適用することにより、さらなる相関係数の上昇させる

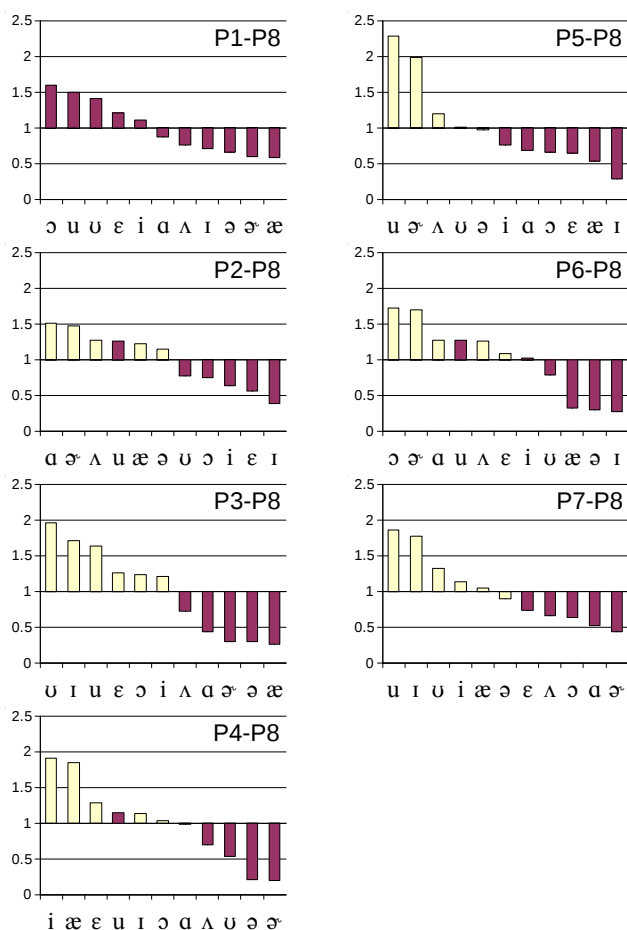


図 10 発音矯正必要度導出時（10 回目）の重み

ことを考える。この発音矯正必要度の再々導出を行った結果が図 9 である。ここでは、再々導出で得られた d_v'' を利用して d_v''' を導出する、というようにに再帰的に 10 回導出した。結果はさらなる改善を示したが、やがて相関係数は収束した。10 回導出したときの相関係数についてまとめたものが、図 8 である。96 人総当たりのときの絶対的評価と重み付き相対的評価（導出 10 回目）の相関係数は 0.827 となり、相対的評価と比較して 5.0 ポイント、重み付き相対的評価（導出 2 回目）と比較しても 2.0 ポイントの改善を示した。また、P8 と P1, P2, ..., P7 を比較して相関分析をした場合でも、すべての場合で改善を示した。

6. 構造的表象に基づく発音矯正必要度の再導出

前節では母音図表象に発音矯正必要度の再導出を適用した。本節では構造的表象に対して適用する。そして、適用した際の重みを調べることで、再導出の効果について考察する。

6.1 発音矯正必要度の再導出の構造的表象への適用

3.1 で収録した音声資料から得られた構造に対して、前節の手法で発音矯正必要度を導出する。導出時の重みについてまとめたものが図 10 である。図中の P1~P8 は、12 話者平均構造である。各図は米語教師相当の構造 P8 と日本人学習者状態の P1~P7 を比較したときの発音矯正必要度導出時（10 回目）の重みを表している。図の縦軸は重みを表し、色づけられているものは日本語母音に置換した母音、白抜きのものは米語発音の母音である。

6.2 考 察

発音矯正必要度の導出時に用いられる重みは、各母音の信頼度と捉えることができる。図 10 を見ると、米語発音母音（白抜き）の多くが重く評価されている。米語発音の母音は信頼できる母音であるから、これは妥当な結果と言える。

また、日本語母音に置換されている母音では、/i/, /o/ が比較的大きな重みを保有している。これらの母音は、日本語の /い/, /お/ に近い母音とされており [8]、この事実と一致する結果となった。しかし、/a/, /u/, /u/ といった母音も大きめの重みとなっており、重みの算出方法について検討の余地がある。

7. おわりに

本稿では、提案されていた発音矯正必要度の推定手法について、母音図を通して妥当性を検証した。母音図に基づく母音間距離を用いて、座標軸がある場合に下される発音矯正必要度と提案されていた導出手法による発音矯正必要度を求め、相関分析を行った。その結果、相関係数 0.772 と良好な相関を示した。しかし、比較する二つの構造が大きく異なると、この相関係数は低下することが確認された。

発音矯正必要度の導出手法について改善策を提案した。構造歪みを考慮した重み付けを取り入れた結果、前述の発音矯正必要度の相関はより強くなった（相関係数 0.807）。また、再導出された構造歪みを用いて構造歪みを再々導出することによって、相関が一層強くなることが確認された（相関係数 0.827）。さらに、この導出手法を構造的表象に適用し、導出時の重みを調べたところ、信頼されるべき母音の重みが高くなる、という妥当な結果を得た。

今後は、子音や連続発声への拡張について検討する予定である。

文 献

- [1] A. Neri, C. Cucchiari, and H. Strik, "Automatic speech recognition for second language learning: How and why it actually works," Proc. Int. Congress of Phonetic Sciences, pp.1157-1160, Aug 2003.
- [2] 峯松信明, 西村多寿子, 西成活裕, 櫻庭京子, "構造不変の定理とそれに基づく音声ゲシュタルトの導出," 信学技報, SP2005-12, pp.1-8, May 2005.
- [3] 朝川智, 峯松信明, 広瀬啓吉, "発音の構造的表象に基づく学習者分類と発音教示生成," 信学技報, SP2005-156, pp.37-42, Jan 2006.
- [4] D.Reynolds and L.P.Heck, "Speaker verification:from research to reality," Proc. Int. Conf. Acoustics, Speech, and Signal Processing, tutorial session, May 2001.
- [5] M. Pitz and H. Ney, "Vocal tract normalization equals linear transformation in cepstral space," IEEE Trans. Speech and Audio Processing, vol.13, pp.930-944, Sep 2005.
- [6] 江森正, 篠田浩一, "音声認識のための高速最尤推定を用いた声道長正規化," 信学論 (D-II), vol.J83-D-II, no.11, Nov 2000.
- [7] 峯松信明, 朝川智, 広瀬啓吉, "線形・非線形変換不変の構造的表象とそれに基づく音響モデリングに関する理論的考察," 日本音響学会春季講演論文集, 1-P-12, pp.147-148 (2007-3)
- [8] 牧野武彦, "日本人のための英語音声学レッスン," 大修館書店, 2005.
- [9] 英語音声研究会, "大人の英語発音講座," NHK 出版, 2003.
- [10] 浅田幸善, Paul E. Davenport, 藤田崇夫, 牧野武彦, 宮井捷二, "グランドセンチュリー英和辞典 第 2 版," 三省堂, 2005.