

音声の構造的表象に基づく発音ポートフォリオの提案と学習者分類

峯松 信明[†] 牧野 武彦^{††} 朝川 智[†] 鎌田 圭[†] 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: {mine, asakawa, k-kamata, hirose}@gavo.t.u-tokyo.ac.jp, mackinaw@tamacc.chuo-u.ac.jp

あらまし 音声の音響的実体には、不可避免的に、話者/性別/年齢などの非言語的要因が混入する。これらの要因を表現する次元そのものを消失させて音声を物理的及び構造的に表象する方法論を提案している [1], [2]。本構造的表象は、構造的音韻論の物理的実装として位置づけられている。本表象を用いて学習者の発音を記述すると、純粹に、母国語に起因する発音の訛のみに着眼することが可能であり、それは、構造的歪みとして観測されることになる [3]。本報告では、擬似的に作成された「米語単母音を学習する日本人学習者」の音声データを用いて、1) 学習者の発音変遷の記述、2) 話者性に依存しない学習者分類、及び、3) 優先的に矯正すべき母音順序の推定、について検討したので報告する。

キーワード 音声の構造的表象、発音ポートフォリオ、発音変遷、学習者分類、母音矯正順序

Pronunciation portfolio and classification of learners based on the structural representation of speech

Nobuaki MINEMATSU[†], Takehiko MAKINO^{††}, Satoshi ASAKAWA[†], Kei KAMATA[†], and
Keikichi 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: {mine, asakawa, k-kamata, hirose}@gavo.t.u-tokyo.ac.jp, mackinaw@tamacc.chuo-u.ac.jp

Abstract Acoustic substances of speech inevitably contain non-linguistic features, such as speaker identity, gender, age, etc. A novel structural representation of speech was recently proposed, where the dimensions indicating these factors are effectively removed. This representation is considered as physical implementation of structural phonology. With the representation, it is possible to describe learners' pronunciations by purely focusing on foreign accents caused by the mother tongue and effectively ignoring the non-linguistic factors of speech. In this report, using simulated speech data of Japanese students learning American English monophthongs, some results are shown about experiments of 1) recording history of learning monophthongs, 2) classification of learners only based on foreign accents not on the non-linguistic factors, and 3) automatic estimation of the order of vowel correction.

Key words Structural representation of speech, pronunciation portfolio, pronunciation transition, classification of learners, vowel correction order

1. はじめに

“How are you?” “Fine, thank you. And you?” 英語学習の場で、繰り返されてきた会話である。筆者らはこの会話に強い違和感を感じている。“How are you?”を「あなたはどのような状況にありますか?」と解釈した場合、問いかけている教師は、各学生の「状態/状況」をどの程度正しく把握しているのだろうか。筆者らが考える正しい応答は、“I don't know how I am. You don't know either, do you!”である。如何だろうか。

教育現場では、教師対学生は1対50、時には1対数百という状況もある。この場合、個々の学生の状態を把握し、それに基づいて指導法を考える、というのはそもそも不可能である。またこの状況は、語学のみならず、全ての教育現場に当てはまる状況である。結局、自分がどのような状況にあり、次に何をすべきか、という意識は学生自身の姿勢に任されるべきであり、行なうべきは、各学生の「状態」について学生自身が知ることができる機会を提供すること、更には、かつてその「状態」にいた学生が、何をし、何を得、最終的にどのような状態へと至っ

たのか、という情報を提供すること、であるとする。

非母国語との接触が多い欧州では、各個人の（欧州の各言語に対する）言語運用能力を記述した「言語パスポート」を定義する動きがある（ELP, European Language Portfolio [4]）。当事者の社会活動を規定することになる情報であり、「何をもってその言語が操れると考えるのか」について、種々の議論が行なわれている。本研究は、日本人学習者が英語（例えば米語）を学習する際に、その発音状態についてのパスポート（発音ポートフォリオ）についての一提案をするものである。外国語の学習目的は、対象言語を用いたバーバル・コミュニケーションであるため、「どの程度通じるのか」についての情報を学習者に与えるべきである。しかし、国際語としての英語 [5] を考えた場合、日本人英語が日本人には最も分かり易い英語であるように^(注1)、「どの程度通じるのか」は聞き手を規定して初めて議論可能となる。当然、この「真の了解度」の定義は困難であるため、本研究では、次のように発音ポートフォリオを捉える。

どのような学習環境であれ、学習者が目標としている発音は必ず存在する。例えば北米英語（より具体的には、ある地域で話される北米英語）であったり、あるいは、ある特定の教師の英語を直接的かつ具体的な目標として、日々励んでいる学習者も多いだろう。明確な目標発音（目標話者）があるのであれば、それと現在の学習者の発音がどう違うのか、どこから矯正するのが効率的か、そして、かつて同様の状態にあった学習者がどう変遷していったか、を示すことができれば、これは学習者にとっては非常に有益な情報となると考える。

このような情報提示をするためには、学習者の音声と教師の音声を直接的に比較する必要がある。しかし、音響量を直接比較した場合、例えばある女性教師の発音を獲得したいと考える男性学生には、「まずは女声を出しなさい」との教示を与えることになる。学習者の音声と教師の音声の一体「何が」等しい時に、それは「目的の発音を獲得できた」と判断すべきであろうか。本研究では、これを音韻構造に求める。個々の音素の絶対的な音響量ではなく、ある音素が他の音素群と正しい関係（目的発音に見られる関係と等価な関係）を結んでいる時に、正しい発音が獲得された、と判断するものである。音素の音響的実体そのものを考えれば、必ず話者性が観測される。では、音素と音素の関係に着目した場合、そこには話者不変の物理量が存在するのだろうか？以下、話者/年齢/性別/マイクなどの違いが音声の音響量にどのような影響を及ぼすのか、に関する数学モデルの上に、音素関係の抽出が非言語的要因に不変な物理量の抽出となることを示し、その後、米語単母音群の学習を念頭に置いた、発音ポートフォリオについて検討する。

2. 話者不変の音声の構造的表象

話者/年齢/性別/マイクが変わると、音声の音響量は如何なる変形を被るのか？本節ではまず、この問いに対する数学モデルを考える。音声認識では、非言語的要因によって音声が被る

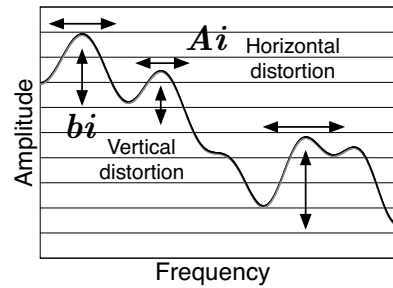


図1 スpekトルに対する水平/垂直方向の音響歪み

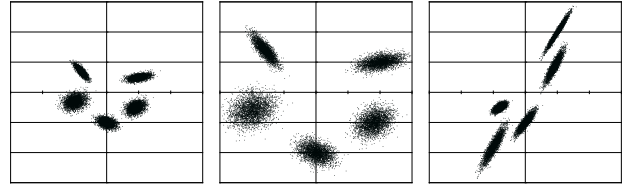


図2 音声に内在する音響の普遍構造

音響歪みは三種類に分類される。背景雑音による重畳性雑音（歪み）、声道長の違いによるフォルマント周波数の高低変動、声道長以外の話者性の違いやマイク特性の違いによるスペクトル変化、である。ここで、重畳雑音は不可避ではないため、本研究の対象としない。図1に第二/三の音響歪みの様子を示す。前者はスペクトルの水平方向の歪みを生じさせ、後者はスペクトルの垂直方向の歪みを生じさせる。スペクトル包絡をケプストラムベクトル c で表現した場合、水平方向歪みは $c' = Ac$ として、垂直方向歪みは $c' = c + b$ として表現されるため、これらの混合歪みは $c' = Ac + b$ として表現される。これが最も簡素な非言語的特徴による音声変動の数学モデルである。

結局、この数学モデルにおいて不変なる音響特徴量を求めればよいことになるが、その一つの解答が、音声の構造的表象である（音響の普遍構造 [1], [2]）。ケプストラム空間に分布として存在する音響事象を考える。分布群が N 個ある場合、 ${}_NC_2$ だけ存在する二分布間距離を、バタチャリヤ距離尺度を用いて測定する。距離行列を規定すれば、（鏡像という曖昧性を残すが）幾何学的構造を規定することになる^(注2)。その様子を図2に示す。3つの5分布群はどれも、距離行列が等しく、構造は等しくなる。これらが歪んで観測されるのは、ユークリッド空間（この場合は平面）の上で構造を観測するからである。一次変換（アフィン変換）は構造を歪める変換としてCGの世界では頻繁に使用されるが、バタチャリヤ距離尺度は、空間を歪める形で距離を規定しており、言い換えれば「構造が歪まないように空間を歪ませながら構造を抽出する」ということになる。

話者不変の構造的表象は、ヤコブソンによって提唱された構造的音韻論の物理的解釈として位置づけられる（図3）[6]。この構造的表象を用いて、日本人による米語単母音群を距離行列として構造的に表象した場合、観測される構造歪みは、純粋に、母語からの影響（訛）によるものと解釈される。以降、発音変遷の記述、非言語的特徴に影響を受けない学習者分類、矯正すべき母音の順序推定について実験的に検討する [7], [8]。

(注1)：例えば、米語/英語母語話者として、日本人にとって分かり易い英語を話したければ、日本人英語を習得すべきである、という議論すら可能であろう。

(注2)：三辺の長さを全て測定すれば三角形が規定されるように、 N 角形の場合は全ての2点間距離を定めれば、図形の幾何学的構造は規定される。

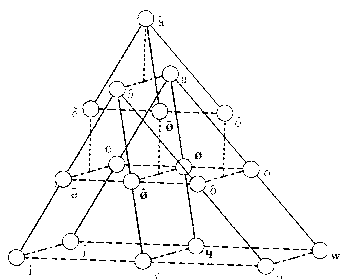


図3 仏語の母音及び準母音に対する音韻構造

表1 日本語母音・米語母音の置換表	
Japanese vowels ↔ English vowels	
a	ɑ, ʌ, æ, ɶ, ɒ
i	i, ɪ
u	u, ʊ
e	ɛ
o	ɔ

3. 学習者の発音変遷の様子の記述

3.1 日本人学習者の模擬音声の収録

英語劇及び発音指導経験を持つ日本人男性に米語 11 単母音を /b V t/ の形で発声させた。ここで有意味語が存在しない場合は、/t/を/d/とするなどの処置を行ない、11 単母音の収録を 1 回ずつ行なった。次に、/b V t o/ という形で日本語母音を 5 回ずつ収録した。これらの音声資料を用いて、一部の米語母音を日本語母音と置換する形で、種々の日本人英語発音を模擬した。表 1 に用いた母音置換表を示す。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。これらの置換パターンを用いることで、1) 日本語母音に全て置換した発音、2)/a/に近い母音を矯正、3)/i/に近い母音を矯正、4)/u/に近い母音を矯正、5)/e/に近い母音を矯正、6) 全ての母音を矯正、と 6 種類の発音状態を規定した。

3.2 発音変遷の記述

米語及び日本語の母音図を図 4 に、及び、6 種類の発音状態の様子を図 5 に示す。11×11 の距離行列で表現される構造を、Ward 法により樹形図化している^(注3)。本分析対象のデータは飽くまでも模擬音声であるが、完全なる日本語母音の使用による英語発音から米語発音に至るまでの一つの遷移の様子が、各母音 1 回ずつの発声を用いるだけで、明確に示されている。

4. 母音構造に基づいた学習者群の分類

要素群の距離行列があれば、ボトムアップクラスタリングが可能となる。前節では 11 母音の距離行列を求めることで、6 種類の発音状態（11 角形）を樹形図として視覚化した。さて、11 角形と 11 角形の構造間距離を規定することができれば、それは学習者と学習者の距離を、純粋に発音訛に着眼した形で（非言語的特徴を無視した形で）定義することとなり、学習者が N 人いた場合、 $N \times N$ 行列の計算、及び学習者分類が可能となる。

(注3)： N 次元空間の幾何学構造を 2 次元に写像しているため、当然失われる情報もあることに注意する必要がある。

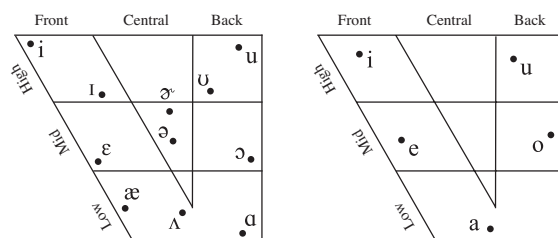


図4 米語及び日本語の母音図

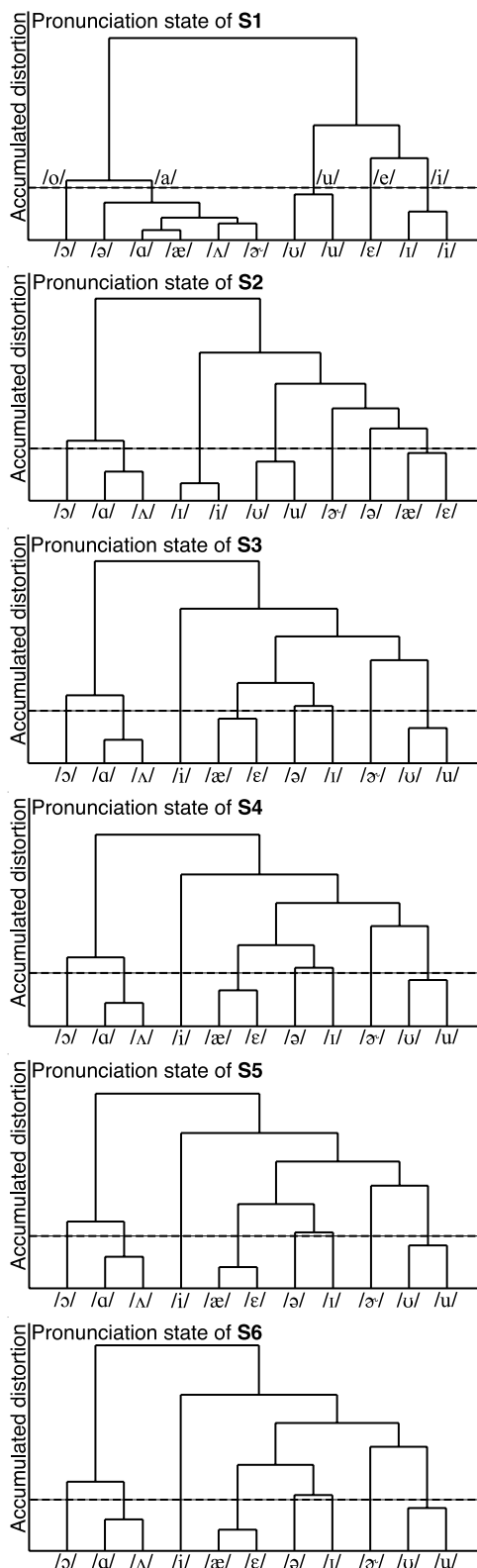


図5 日本人英語発音から米語発音へ（模擬実験）

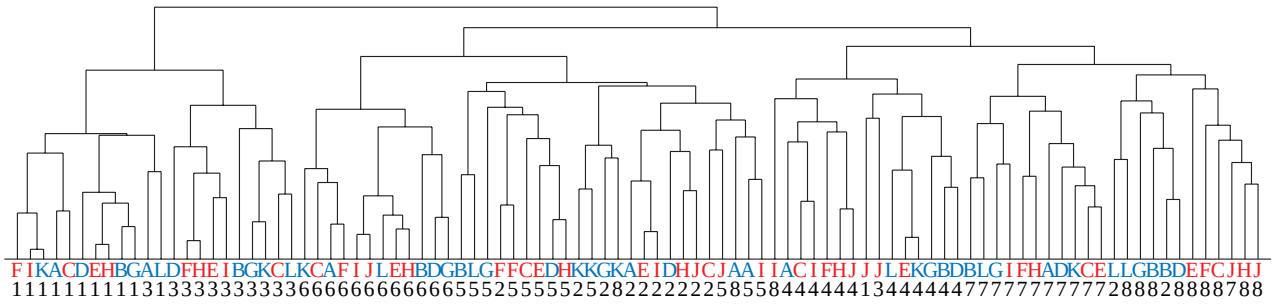


図 6 母音間差異の比較に基づく 96 発音構造 ([12 話者 A~L]×[8 発音状態 1~8]) の分類 (D_1)

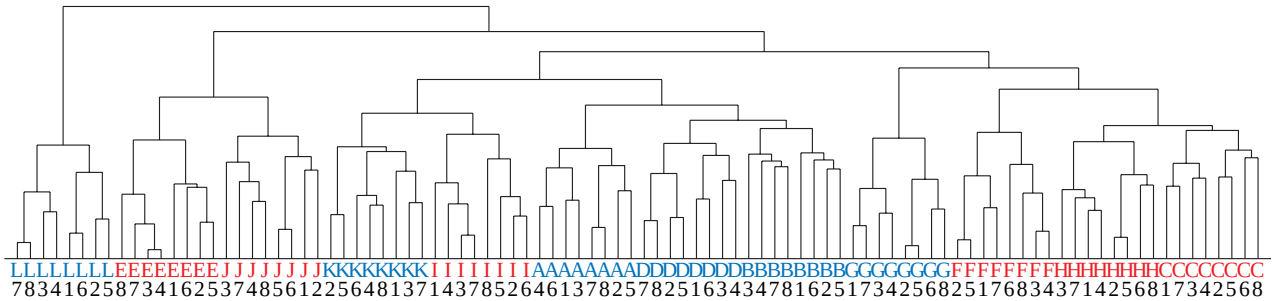


図 7 母音実体の比較に基づく 96 発音構造 ([12 話者 A~L]×[8 発音状態 1~8]) の分類 (D_2)

表 2 母音置換によって模擬された 8 種類の発音状態

	α	æ	Λ	ə	ɜ	ɪ	i	u	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

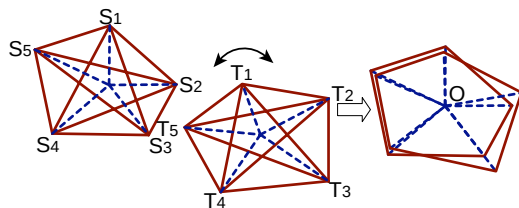


図 8 回転とシフトによる構造群の接近と構造間距離

4.1 実験に使用した音声資料

帰国子女、或いは英語劇経験者の日本人 12 名 (男性 6 名、女性 6 名、A~L) より、11 種類の米語母音と 5 種類の日本語母音を第 3.1 節と同様に、各々 1 回、5 回ずつ収録した。これらの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、様々な発音状態を規定する。本実験では、表 2 に示す置換パターンを用い、話者ごとに 8 種類の発音状態を用意した。なお、米語・日本語間の母音置換の対応づけは表 1 の通りである。このようにして模擬された 96 種類 (12 話者 × 8 発音) の発音状態に対して、下記で定義される二つの距離尺度に基づく距離行列を計算し、各々を樹形図化することで分類を試みる。

4.2 二種類の構造間距離尺度

構造間距離は図 8 に示すように、一方の構造を回転及びシフ

トして他方に近づけた際に観測される、対応する 2 点間距離の総和の最小値として規定できる^(注4)。この最小値は、二つの距離行列間のユークリッド距離として近似できることが示されており、この距離尺度を使用する。これは、母音間距離の差異、即ち、差異の差異 (二次差分) を求めることに相当する。

$$D_1(S, T) = \sqrt{\frac{1}{M} \sum_{i < j} (S_{ij} - T_{ij})^2} \quad (1)$$

ここで、 S_{ij} は構造 (学習者、距離行列) S の (i, j) 成分であり、 M は母音数、11 である。

これとは別に、母音の音響的実体を構造間 (話者間) で直接観測することで規定される構造間距離についても考える。これは、従来の DP や HMM で行なわれる実体間の差異を求める操作 (一次差分) に基づく距離尺度である。

$$D_2(S, T) = \sqrt{\frac{1}{M} \sum_i BD(v_i^S, v_i^T)}, \quad (2)$$

ここで、 v_i^S は構造 (話者) S の i 番目の母音であり、 $BD(x, y)$ は分布 x, y 間のバタチャリヤ距離である。

4.3 96 種類の発音構造の分類

2 種類の距離尺度 (D_1 と D_2) による分類結果を図 6、図 7 に示す。樹形図のリーフノードにおける数字 (1~8) が発音状態を意味し、A~L が話者を意味する。二つの樹形図は、全く同一のデータに対して、構造間距離尺度のみを変えて求めたものであるが、母音の音響的実体の直接比較に基づく分類 (図 7) は完全なる話者分類となっているのに対し、母音間差異に基づく比較 (図 6) では、凡そ発音状態の分類となっている。発音状態の定義や種類についてはより詳細な検討が必要であると考

(注4) : なお、一次変換によって距離行列が不変であるため、 A の乗算は回転として、 b の加算はシフトとして解釈される。

えられるが、構造的表象を用いることで、話者や性別などの非言語的特徴に影響されることなく、発音の分類、即ち、言語的な分類が可能となることが実験的に示された。これは例えば、英語を学ぶ数億人の学習者を（性別や年齢といった要因を切り離して）話者を単位として分類できることを意味する。このような分類は方言研究に応用することも可能である。

図 6 において、発音状態が混合したクラスが一部見受けられる。これは、各話者で米語を身につけた過程が異なり（地方、及び、滞在期間など）、各々発音の癖や訛があることをが原因の一つであると考えている。英語の母音構造は方言に強く影響を受けるため、厳密には、同じ発音状態番号であっても話者ごとに発音状態は異なってくる。例えば、全て米語発音である P8 が、P2 や P5 のクラスに紛れ込んでいる話者がいる。表 2 から考えるに、/ɪ, i, e, ɔ/ の発音が日本語的な発音になっている、と推測されるが（特に /ɪ/ ）、詳細に関しては該当話者音声に対する分析が必要である。

なお、図 6 における 8 クラス分類は、凡そ発音状態に対応しているが、8 クラスに至る分類の様子についての妥当性について、調音音声学の観点から検討を行なっている。聴取に基づいた 96 種類の発音構造の母音図を英語音声学の専門家（第二著者）が作成し、これらを用いて 96 発音構造を分類した結果、図 6 と非常に類似した樹形図が得られている [9]。

5. 優先的に矯正すべき母音の自動順序推定

本節では、学習者の発音構造と目標とする発音構造を比較することによって、発音構造のどの部分（母音）から矯正するのが最も効率的か、矯正順序を推定することを試みる。前節で定義した 8 種類の発音状態に対し、P8 を教師発音構造とした場合、P1～P7 の発音状態にある学生に対して矯正すべき母音順序を、表 2 を参照することなく、学生の母音構造（距離行列）と教師のそれとを比較することのみで推定する。

5.1 構造間の要素歪みの定義

第 1 節で定義した D_1 は、二つの構造間差異の総和に対応する量であるが、この定義から、特定の母音 I に起因する要素的な構造歪みを以下のように定義する。

$$d(I) = \sum_{j=1}^M |S_{Ij} - T_{Ij}| \quad (3)$$

2 種類の構造が与えられた場合、各々の母音に起因する構造歪みを上記式より算出し、「より大きな構造歪みを呈する母音」を「より優先的に矯正すべき母音」として示すことを考える。

5.2 矯正母音順序の自動推定

前節で求めた 96 種類の発音構造を、各々の発音状態毎に 8 人の平均をとり、各発音状態の平均距離行列を求めた。平均行列に対して式 (3) を用いることで、各状態の学習者が状態 8 に遷移するために矯正すべき母音の順序を推定した。なお、自動推定の際に使用する情報は、2 つの距離行列のみであり、表 2 に示している置換パターンの情報は一切用いていない。

結果を図 9 に示す。縦軸が各母音に起因する構造歪みであり、横軸が母音の矯正順序である。灰色で色付けされているものは

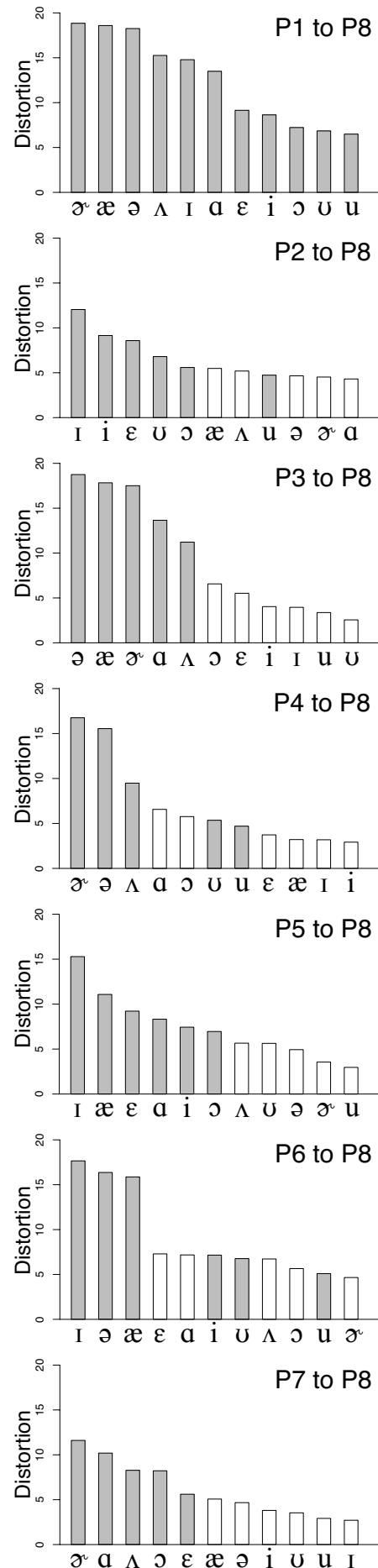


図 9 各発音状態に対する矯正母音順序推定（模擬実験）

日本語母音に置換された母音である。まず P1 について考察する。11 種類の母音は構造歪みによって大きく 3 つのグループに分かれる。/ɔ, æ, ə/ の母音は構造歪みが大きく、矯正の優先度が高いと判定されている。構造歪みが中程度の母音として /ʌ, ɪ, ɑ/ が、歪みが小さい母音として /ε, i, ɔ, u, ʊ/ が示されている。歪み大と判定された /ɔ, æ, ə/ は、日本人がその発音を苦手とする母音である。また、歪み小のグループに属する母音群は、日本語の /e, i:, ɔ, u, ʊ:/ と音色が似ており、初心者を対象とした発音教則本などでは、日本語の母音と置き換えても支障の無い母音として位置づけられている [10]。これらのことを考えると、P1 に対する推定結果は、日本語・英語音声学的に考えて、非常に妥当性の高い結果であると言える。

P2~P7 については、日本語母音で置換された母音が歪み大として推定されている傾向を、明確に確認できる。P2, P4, P6 では日本語で置換された母音が優先度小として推定されている結果もあるが、これらは /i, ʊ, ʊ/ のいずれかである。上記した様に、これらは日本語の /i:, ʊ, ʊ:/ と音色が非常に類似した母音であり、置換後も歪み小と推定されるべき母音である。

6. 考察 ～母音の絶対的物理量は必要か？～

話者、年齢、性別などの要因は、発音教育とは一切無関係でありながら、音声は「ある話者が発声したものである」以上、音声に不可避免地に混入する要因である。従来検討されてきた（音素の）音響モデリングは、この不必要な要因を消失させるために、数千、数万という話者の音声を集めてきた。その上で、各母音の絶対的物理量を、分布としてモデル化してきた。

本研究では、各母音の絶対的物理量をモデル化し、絶対量に基づいて学習者と教師を比較することよりも、非言語的特徴というノイズを表現する次元を消失させることを最優先し、母音群の関係だけを捉える方法論を提案した。そして、個々の母音の絶対的特性は一切参照していないにも拘らず、どの母音を矯正すべきか、という問題に対しても適切に解答することが可能であることを実験的に示した。非言語的要因による音声変動に対して最も簡素な数学モデルを仮定した議論ではあるが、特定の教師発音と特定の学生発音を、直接比較することが可能となった、と考えている。従来の方法論では、学習者の音声は数万の母語話者の平均分布と比較され、評価されてきた。提案手法では、例えば、特定の映画俳優の発音と学習者の発音とを比較し、どこから矯正するのがその俳優に最も効率よく近づくのか、というアプリケーション開発も可能となる^(注5)。発音学習は多くの場合、退屈な学習となる。学習の動機付けも、継続学習に対する大きな要因であるが、提案手法は、その問題を容易にクリアすることができる技術であると考えている。

幼児の言語獲得の場合、動機付けが問題となることはなく、両親とのコミュニケーションそのものが動機付けとなる。この場合、親の声を模倣する形で語を獲得する訳だが [11]、彼らは親の声など模倣していない [12]。親の声の「物真似」などしな

い。言い換えれば、彼らは音響的模倣することなく、言語的模倣を通して言葉を獲得する。両者の違いは、どこにあるのだろうか？ [13] 少なくとも、音の絶対的物理特性を捉えてきた従来法は、数万の母語話者の平均的物理特性を「物真似」することを学習者に要求してきた。音響的模倣である。この指導は正しいのだろうか？或いは、根本的に誤っているのだろうか？提案手法は、音そのものではなく、音と音との差異／コントラストを学習することを主張する方法論である。そして「差異群の学習」と「音群の体系の学習」間の、数学・幾何学的等価性については **第 2 節** で示した通りである。言うまでも無いが、「体系の無い要素は無価値」と主張するのが言語学である [14]。

図 1 に、非言語的要因による音響歪みの様子を示した。例え話になるが、鬼多郎の目玉親父やウルトラマンのフォルマント周波数は人類の領域を遥かに超えたところに位置する。即ち、彼らが「英語を学習したい」と主張しても、人類の提供するデータでは網羅できず、音響的模倣に基づく方法論では対処できないが、本手法なら、彼らを学習者の一人として受け入れることに、何ら問題を生じ得ない。50 センチ強の個体が、2 メートル弱の個体を出す音響信号を通して言語を獲得するというのは、音響的模倣の枠組みでは「不条理」である。しかし、提案手法が主張する枠組みなら、何ら問題を生じ得ない。

7. ま と め

筆者らが提案する音声の構造的表象に基づく発音学習支援に関して総括すると共に、(恐らく全ての先行研究の枠組みである) 音響的模倣に対して懐疑的な考察を行なった。今後実際の学習者音声に基づく検討と、提案手法に基づいた教材作成など、より現場に近い環境での実験的検証を行なう予定である。

文 献

- [1] 峯松信明, “音声に内在する音響的普遍構造とそれに基づく語学習者モデリング”, 信技技報, SP2003-179, pp.25-30 (2004-1)
- [2] 峯松信明他, “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 (2005-5)
- [3] 峯松信明, “構造音韻論の物理実装に基づく学習者表象と発音学習支援”, 信技技報, TL2004-47, pp.47-52 (2004-12)
- [4] 吉島茂他訳・編, 外国語教育 II - 外国語の学習, 教授, 評価のためのヨーロッパ共通参照枠, 朝日出版社 (2004)
- [5] C. Burt, “What is international English?” Working Papers in TESOL & Applied Linguistics, vol. 5, no. 1 (2005)
- [6] R. Jakobson *et al.*, Notes on the French phonemic pattern, Hunter, N.Y. (1949)
- [7] 朝川智他, “音声の構造的表象に基づく非母語話者の英語発音分析”, 信学技報, SP2005-24, pp.25-30 (2005-6)
- [8] 朝川智他, “発音の構造的表象に基づく学習者分類と発音教生生成”, 信学技報, SP2005-156, pp.37-42 (2006-1)
- [9] 鎌田圭他, “音声の構造的表象を用いた英語学習話者の分類に関する実験的検討”, 信学技報 SP2006, (2006-11)
- [10] 英語音声学研究会, 大人の英語発音講座, NHK 出版, 東京 (2003)
- [11] 早川勝廣, 言語獲得と育児語, 月刊言語, 大修館書店 (2006-9)
- [12] N. Minematsu *et al.*, “Universal and invariant representation of speech,” Proc. Int. Conf. Infant Study (2006-6)
- [13] 峯松信明他, “音声の構造的表象に基づく音声認知と外国語発音学習に対する一考察”, 信学技報 TL2006 (2006-11)
- [14] F. d. Saussure, Course in General Linguistics, McGraw-Hill (1965)

(注5): ブラット・ピットの「口」と貴女の「口」を近付けるアプリケーション
言えば、本提案技術の“恐ろしいほど魅力的な”意味が分かるだろう。