

米語母音発音に対する絶対的評価と相対的評価の相関分析*

◎朝川智[†], △鎌田圭[‡], 峯松信明[†], △牧野武彦[‡], 広瀬啓吉[†]([†]東大, [‡]中央大)

1 はじめに

年齢, 性別, 個人性, 音響機器など, 音声には不可避的に非言語的情報が付随し, これが音声の音響的特徴を変形させ, 音声記号としては同一音であったとしても, 物理的には異なる特性を持った音響現象として観測されることになる。従来の音声工学では, 複数の話者の音声から同一記号音を収集して統計的にモデル化し, 発音学習応用においては, この統計モデルと学習者音声とを直接的に比較する方法論をとってきた。しかし, 学習話者と評価話者の話者性/収録環境の不一致は避けられず, これらが要因となってシステムの不安定性を呈することになる。

筆者らは, この不一致問題を, 話者や音響機器に不変な音声の音響的表象を導入することで解決を図っている^[1]。この音声表象は, 個々の音事象の絶対的な物理特性は捨象し, ある音事象を, 他の全音事象群からの関係のみで規定する方法論を採る。その結果, 個々の音韻の音響的実体は, それだけでは評価対象にはなり得ず, 他の音韻の音響的実体があって初めて評価対象となる。既に発音評定応用を目的として, この関係性のみに着眼する発音表象を用いて, 矯正すべき母音の優先順位を, 学習者別に推定することが可能であることを実験的に示した。即ち, 母音の絶対的な音響的特性は考慮せず, どの母音をまず矯正すべきなのか, に答えるシステムを構築している^[2]。

本稿ではまず, 様々な訛を有する米語母音発音を, 英語音声学/調音音声学を専門とする音声学者に母音図上にプロットさせ, 各発音の母音構造を規定する。次に, その(訛った)母音構造と, 本来あるべき母音構造を二種類の方法論で比較する。一つは訛った構造と正しい構造とを, 各母音毎に絶対的に比較する方法であり, 他方は母音構造同士を, 相対的に比較する方法である。筆者らが提案する自動評定方法は, この後者の方法に対応する。本実験結果を通して, 個々の音韻の絶対的な音響実体を完全に無視する評定方法の教育的妥当性について, 再度検討する。

2 話者不変量としての音声の構造的表象

音声に混入する不可避的な非言語的特徴は, 乗算性歪みと線形変換性歪みに大別され, 対数パワースペクトルに対して, 前者は垂直方向の変形として, 後者は水平方向の変形として現れる。これらの歪みはケプストラムベクトル c に対して $c' = Ac + b$ として線形変換でモデル化することができ, このモデルに対する不変物理量を用いることで話者不変の音声表象が構成される。筆者らは, 分布間距離として知られるバタチャリヤ距離がこの性質を満たすことに着目

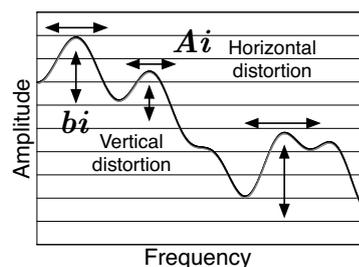


Fig. 1 スペクトルに対する水平/垂直方向の音響歪み

し, 音声事象を全て分布として推定し, 全ての分布間距離を求め, 距離行列として音声事象群を表象することを提案している。距離行列は幾何学的構造を規定するため, これは音声の構造的表象となる。なお, この構造的表象は線形のみならず, 非線形変換に対しても不変性を満たすことが示されている^[1]。

一つの発音を分布系列へと変換し, 全ての二分布間距離を求める形で, 発音を一つの幾何学構造として表象できる。この方法論は, 音声を, (音高ではなく)音色が変化する音楽として捉えることに相当する^[1]。相対音感があるメロディーを(階名としての)ドレミに書き起こす場合, そのメロディーが移調してもドレミ列は一切変わらない。この場合, 個々の音の絶対的な物理特性は完全に無視され, ドレミとしての同定は, 各音の, その曲全体(その曲を構成する個々の音群)との関係に基づいて行なわれる。筆者らは, 音声の知覚も相対音感に基づく処理が存在すると考えている^[3]。個々の音の絶対的な音響特性を無視する相対音感者であっても, 音痴の歌を聞いた場合に, 矯正すべき部位を特定できる。同様に, 提案手法では, 個々の音韻の絶対的な音響特性を無視した形で, 相対音感的に矯正すべき母音を同定する^[2]。

3 聴取による距離行列(母音図)の作成

帰国子女の日本人12名(男女6名ずつ, A~L)に米語11母音を/b V t/の形で, 及び, 日本語母音を/b V t o/の形で発声させた。米語は母音当たり1発声, 日本語は母音当たり5発声収録した。この米語母音の幾つかを日本語母音と置換する形で日本人英語音声を模擬した。日米母音の対応を表1に, この表に基づいて定義した8種類の日本人英語母音発音を表2に示す。全12話者各々に対して8種類の発音が定義されるため, 合計96種類の母音発音が得られる。

96種類の母音発音中に存在する全ての母音(1,056個)を英語音声学/調音音声学を専門とする音声学(第四著者)に, 各発音(11母音)毎に聴取させ, 母音図としてプロットさせた。通常, 2次元の母音図は舌の位置を指定するために使われるが, 本稿では

* A correlational analysis between absolute and relative evaluation of American English vowels produced by Japanese learners. by S. Asakawa, K. Kamata, N. Minematsu, T. Makino, and K. Hirose

Table 1 日本語母音・米語母音の対応表

日本語母音	↔	米語母音
/あ/		/æ/, /ʌ/, /ɑ/, /ɔ/, /ɒ/
/い/		/i/, /I/
/う/		/u/, /ʊ/
/え/		/ε/
/お/		/ɔ/

Table 2 母音置換に基づく8種類の母音発音

	ɑ	æ	ʌ	ə	ɔ	ɪ	i	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

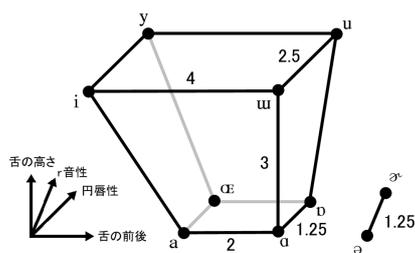


Fig. 2 4次元母音図と各軸の重み

これに加え、円唇性を3段階、r音性を2段階で評定した。得られた結果(2次元母音図+円唇性+r音性、即ち4つの数値の組)から、次のようにして各母音を4次元空間にプロットすると共に、距離行列を計算した。まず、2次元母音図については、広く用いられている4:3:2の台形表示を踏襲した。円唇性及びr音性に対する次元の重みについては、参考数値を調音音声学の文献に見つけることが困難であったため¹、母音間のケプストラム距離、及び4:3:2の台形表示を考慮して、図2に示す4次元母音図を用いた。各母音発声は4次元空間の1点として表現されるが、11母音群を11×11の距離行列として表象すると、個々の母音の絶対的な位置座標の情報は失われる。

4 絶対的評価と相対的評価間の相関分析

12話者、8種類、合計96種類の母音群発音が存在する。ある母音群発音を教師とし、異なる母音群発音を学習者とする。「どの母音から矯正すべきか」を考える場合、学習者/教師間で、4次元母音図における当該母音の距離を測り、「距離が大きい母音ほど要矯正」と考える。即ち、2つの母音図を母音図の軸(棒)を重ねて比較し、ズレが大きい母音ほど要矯正と判断する。本稿ではこれを絶対的評価²と定義する。

¹母音位置+円唇性による3次元母音図などが提案されているが、円唇性の軸(奥行き)の長さに対する定量的な議論は無い。

²各話者の母音を、聴取に基づいて母音図にプロットした時点で、話者性は正規化される。その意味において、本稿で言う絶対的評価も厳密には相対的評価である。本稿で定義する絶対的評価とは、各母音の評価に対して、4次元母音図中の座標値を使用するか否か、に着目して、絶対/相対という言葉を使い分けている。

Table 3 絶対的/相対的評価間の相関分析

P1	P2	P3	P4	P5	P6	P7
0.68	0.82	0.82	0.89	0.80	0.75	0.85

筆者らが提案する発音の構造的表象は、ある空間における絶対的な座標情報を有しない。即ち母音図における台形の軸(棒)が存在せず、母音だけがプロットされた状態で、2つの母音群(母音散布)を比較し、どの母音から矯正すべきか、を推定する問題となる。この場合、片方の構造を回転&シフトさせて他方と重ね合わせ、どの母音から矯正すべきかを議論する。この問題を下記のように解決し、相対的評価と呼ぶ。

回転&シフトさせて得られる、対応する頂点間距離の総和の最小値は、距離行列間のユークリッド距離に近似的に比例することが示されている。

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

S, T は2つの距離行列であり、 M は母音数である。次に、各母音の矯正要請度を下記式で定義する^[2]。 $D_G(S, T)$ を2つの構造間の全体的な歪みと考えた場合、 D_G に対する各母音の貢献度に相当する。

$$D_L(S, T, v) = \sum_{j=1}^M |S_{vj} - T_{vj}| \quad (2)$$

絶対的/相対的評価、いずれも、各母音に対して矯正要請度が定量的に算出される。この要請度の相関分析を絶対/相対の間で行ない、各母音の音響特性を無視した提案手法の妥当性を検討する。教師として各話者のP8を用い、その話者のP1~P7をP8と比較する形で矯正要請度を二通りの方法で求めた。全話者のP1~P7に対して各母音の矯正要請度を算出した後、 P_i 別に絶対/相対間で相関分析を行なった。結果を表3に示す。また、各話者の各 P_i とその話者のP8とを比較することで得られる矯正要請度全体の相関値は0.78であり、教師をP8に限定せず、任意の P_i を教師に、それ以外の(他話者を含む) P_j を学習者に見立てた場合の相関値も0.78であった。

絶対/相対間の高い相関値は、「各母音の絶対的な音響特性が入手できなくても、その母音と他の母音群との関係(距離)が得られた場合、学習者/教師間で母音構造を比較することで、どの母音から矯正すべきかの指針が得られる」との主張を強く支持する。しかし、P1に対する相関値は低い。これは式(2)で定義される局所的な歪み D_L を、母音 v に全て帰着させることに起因する。 v が適切な場合でも、それ以外の母音 j の多くが不適切な発音をしている場合、 D_L は大きくなり、 v は不適切な発音と評価される。母音図の軸が与えられない以上、P1のように全母音に対して不適切な発音をされると、本手法の精度は落ちる。実用的には問題無いと考えられるが、アルゴリズム的な解決について、今後、検討したい。

参考文献

- [1] 峯松他, 人工知能学会, SIG-Challenge-0624-6, 35-42, 2006.
- [2] 朝川他, 電子情報通信学会, SP2006-77, 37-42, 2006.
- [3] 峯松他, 電子情報通信学会, SP2005-12, 1-8, 2005.