非言語的要因に不変な音声の全体的・構造的表象とそれに基づく失読症・自閉症に対する一考察*

○峯松信明（東大・新領域），櫻庭京子（清瀬市福祉センタ），西村多寿子（東大・医学系）

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，数千・万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することで多様性問題の解決を図って来た。

幼児は両親という非常に限られた音環境で，電話越しのお婆ちゃんの音声にも対応できるようになる。限られた音環境での音声活動は一生続く。人の聞く声の半分は自分の声であり（speech chain），話者バランスのとれた音声コーパスの聴取など，不可能である。機械と人間の本質的な差異はどこにあるのだろうか？

発達心理学では「音声模倣を通して言語獲得する」と説明するが [1]，音を模倣する幼児はいない。父親の太い声を出そうと努力する幼児はいない。音韻意識が未熟な幼児は，聴取音声を個々のモーラに分割し，各モーラを自らの口で生成する，という技は不可能である。幼児と異なり，音を模倣するのが九官鳥である。優秀な九官鳥は聞けば飼い主が分かるが [2]，優秀な幼児を聞いても飼い主は同定できない。幼児は親の声の何を真似ているのだろうか？筆者らはこの問いを多くの発達心理学者に問いかけたが，その回答は得られなかった。その一方で，「語全体の音形，語ゲシュタルトを獲得する」と説明する [1]。即ち，この語ゲシュタルトの物理的定義を求めないまま，その存在を仮定した議論を重ねているのが実情である。

相対音感者がメロディーをドレミで書き起こす場合，そのメロディーを移調したとしても，書き起こされるドレミ（階名）は変わらない。相対音感者は個々の音の同定を，その音の絶対的な物理特性ではなく，他の音群（曲全体）に対する相対的特性，即ち，音の機能的価値を感覚して行なう。全体把握から要素を同定し，長調の主音を「ド」と同定する。物理的に異なる音に対して「ド，と聞こえる」と答える。

異なる母音の音色の違いは，その母音を生成する際の口腔形状の違い，即ち，定常波の共鳴周波数の違いに起因する。これは，異なる形状の楽器が異なる音色を生成するのと物理的に等価であり，日本語には五つの母音楽器が存在する。音楽の音ストリームは音高の変化によって構成されるが，音声（特に母音）の音ストリームは，音色の連続変化によって構成される。既に，音声聴取を「音色変化に対する相対音感」として解釈することで，頑健なる話者不変表象，語ゲシュタルト，が数学的に存在することを示している [3]。

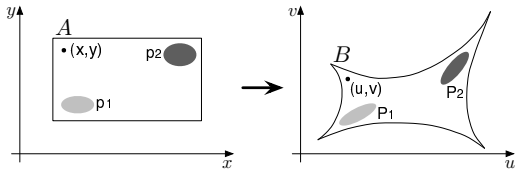

Fig. 1　一対一対応関係を有する二つの空間 A と B

本稿では，筆者らが提案する，非言語的要因に不変な音声の全体的・構造的表象を前提として人間の音声認知を考察する。すると，医学や障害学の世界ではパラドックスとまで言われる失読症の症状（頭が良いのに本が読めない）[4] が，当然の現象として，予測できることを示す。少なくとも第一著者は，物理的かつ数学的な思考に基づいて，失読症という障害を知らずに，上記症状を予言した本人である。また，認知特性として，刺激の要素的かつ絶対的な物理特性に執着した形の情報処理が強い場合，音声言語が著しく困難になると予想されるが，該当する症状は一部の自閉症者に見られる。以下，第一著者が物理的及び数学的な思考の下で，各々の症状を予言しつつ失読症や自閉症と出会った経緯について記述する。技術者（第一著者）と臨床家（第二著者）との対話である。

2　非言語的要因に不変な音声の全体的表象

図 1 に示す二つの空間を考える。両者には一対一の対応関係があり，空間 A の点 (x, y) は空間 B の点 (u, v) へ写像され，逆もまた成立する。但し，その写像関数は与えられていない。空間 A における事象 p_1 と p_2 を考える。全ての事象は点ではなく，確率密度関数として存在する。p_i に対応する空間 B の事象を P_i とする。P_i も確率密度関数となる。この時，バタチャリヤ距離は両空間で常に等しくなる [3]。

$$\begin{aligned} BD(p_1, p_2) &= -\ln \iint \sqrt{p_1(x,y)p_2(x,y)}dxdy \\ &= -\ln \iint \sqrt{P_1(u,v)P_2(u,v)}dudv \\ &= BD(P_1, P_2) \end{aligned}$$

この不変性は，非線形変換を含む，幅広い変換群に対して成立する。また，カルバックライブラ距離，ヘリンジャ距離なども満たす一般的性質である。音声認識の話者適応，音声合成の話者変換いずれも，異なる話者間において音響空間の二者間写像を定義する操作に他ならない。上記の議論は，その写像に対して，頑健なる不変量が存在することを保証する。

---

* A consideration of dyslexia and autism based on the holistic and structural representation of speech which is invariant with non-linguistic factors in speech. by N. Minematsu, K. Sakuraba, and T. Nishimura

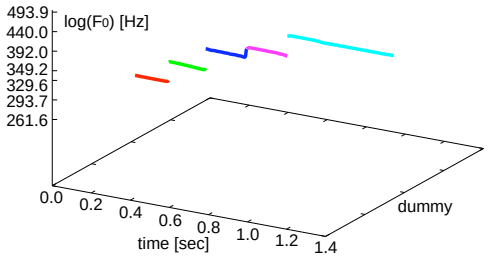

Fig. 2　CDEFG に対する基本周波数の遷移パターン

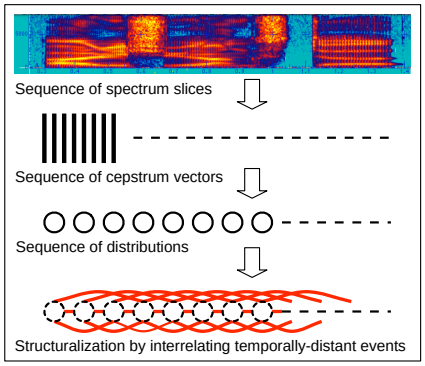

Fig. 3　事象間差のみを抽出して構成される不変構造

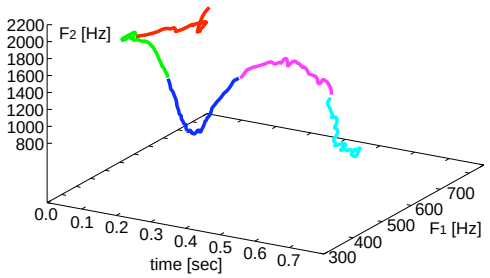

Fig. 4「あいうえお」に対する F1/F2 の遷移パターン

3　多次元音楽としての音声モデリング

相対音感者は，音高変化を伴う音ストリームに対して，その変化のみを感覚し，個々の音を階名として同定する。この能力はヒトのみが有しており，動物は基本的に絶対音感者である。個々の音要素の同定が行なえない幼児は，親の声と自らの声の何が一致した時に「同一」を感覚し，満足するのだろうか？

図 2 に CDEFG と鍵盤を叩いた場合の基本周波数の様子を示す。これを「ドレミファソ」と書き起こす相対音感者は GABCD を聞いても「ドレミファソ」と聞こえる。音高の遷移パターンが同一だからである。提唱している音声表象は，音声ストリームを分布系列へと変換し，全ての分布間距離を求め，距離行列（幾何学構造）を算出することで求まる（**図 3** 参照）。この操作を**図 2** 同様，位相空間で考えたものが，例えば**図 4** である。「あいうえお」という発声を，F1/F2 平面に対して時間軸が延びた三次元空間を用いて示している。音高という一次元物理量に対する変化が音楽であり，音色（この図では F1/F2）という二次元物理量に対する変化が音声である。即ち多次元音楽として音声を捉えることに相当する（音色の表現には，複数次元が必要）。前節の議論は，この遷移パターンを変換不変とすることが可能であることを意味する。

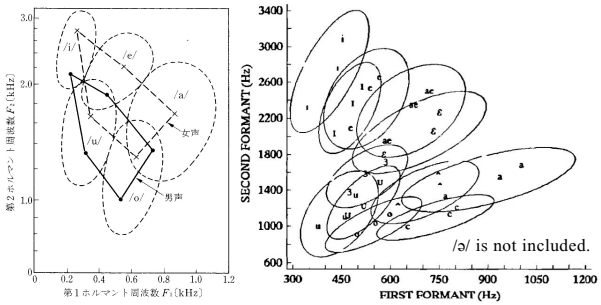

Fig. 5　日本語 5 母音と米語 12 母音の F1/F2 図

4　技術者と臨床家との対話

既に第一著者は，相対音感的な音声の全体的表象に基づく音声認知を実験的に検証した。全体的表象は，音声の話者性が時不変であることを前提としているが，話者性を時変とした合成音声（無意味モーラ列）に対する同定率が有意に低下する様子を報告した [3]。

「あ」と単独発声しても，それは「あ」と分かる。否定し難い絶対音感である。しかしこれは，各音韻が異なる音響音として実現されることを要求する。日本語のように少ない母音数であればよいが，例えば米語などは**図 5** に示すように大きな重なりを呈する。

音韻の音響実体に対して，重なりが大きくなれば，要素音同定は，相対音感（階名）的にならざるを得ない。ここで興味深いのは，階名同定が困難な相対音感者である。ハミングする形で（**図 2** の遷移パターンを追従して）聴取したメロディーを再生する。これでも音楽は楽しめる。「音色変化に対する相対音感」として音声聴取を考えた場合，**図 4** の遷移パターンを追従して音声生成する人が予想される。この場合，音韻同定が困難な者として観測されるはずである。音声言語には（知的にも）難は無いが，文字だけが極端に苦手，という一見不思議な現象が起きるはずである。

一方，絶対音感者の中には，C の音で始まらない「カエルの歌」は「カエルの歌」では無いと主張する者もいる。異なる音は異なる。音声の研ぎすまされた絶対音感者（全体的なパターンを捉えることが困難で，要素刺激の絶対的物理特性に固執する者）は，二話者の「おはよう」の同一性の認知が困難になる。この場合，音声言語は存在せず，耳に届く音声は，音であって言語では無くなるはずである。

第一著者は以上の問いを第二著者に投げかける形で失読症，自閉症を知った。第一著者にとって失読症はパラドックスではない。なお，音声言語を持たない三歳の自閉症児が，他者の声をそっくり模倣する例も知られている [5]。そこに通常の意味は無い。あるのは音である。本稿で示した物理的及び数学的な考察が，医学及び言語障害学の分野に対して，どのように貢献できるのかは未知数である。しかし，「謎の解明」に少しでも助力することができれば幸いである。

[1] 早川, 月刊言語, 35, 9, 62–67, 大修館書店, 2006.
[2] 宮本, “音を作る・音を見る,” 森北出版, 1995.
[3] 峯松他, 人工知能学会, SIG-Challenge-0624-6, 35–42, 2006.
[4] S. Shaywitz, “読み書き障害のすべて,” PHP 研究所, 2006.
[5] 深見, “ひろしくんの本 (V),” 中川書店, 2006.