

音声の物理特性を通して考える失読症・自閉症の音声認知

○峯松 信明¹(みねまつ のぶあき), 櫻庭 京子², 西村 多寿子³

¹ 東京大学新領域創成科学研究科, ² 清瀬市障害福祉センタ, ³ 東京大学医学系研究科

(要旨) 性別, 年齢, 体格, マイクなど, 様々な非言語的要因によって音声の物理特性は多様に変化する。しかし人間は, この多様な音声から苦もなく言語情報を抽出する。最近の脳(聴覚皮質)研究は, 音声の言語的特徴と非言語的特徴が皮質において分離されているとのモデルを呈している。さて, 幼児が語を獲得する際, 親の声そのものを真似ようとはしない。彼らは音韻意識が希薄なため, 「音声の中の個々のモーラを認識・再生する」との説明は不適切であり, 彼らの語の獲得(即ち, 話者性を無視した親の声の模倣)の様子を発達心理学は「語全体の音形の獲得」と説明する。近年著者らにより, 非言語的特徴を消失させて音声を全体的・構造的に表象する方法が提案されている。本報告は, この新しい音声の全体的表象を通して, 失読症, 自閉症の音声認知の様子を考察する。

Key words: 非言語的要因, 音響の普遍構造, 音韻論, 失読症, 自閉症, 音声認知

1. はじめに

音声を IPA シンボル列として書き起こす。音声学を習得せんとする者が, その獲得をまず要求される能力の一つである。しかし, ここで使用される音シンボルは, 話者の違いまでを表現しない。音声をその音響的側面から考えれば, IPA シンボルは現実を抽象化した表現に過ぎない。音声工学の世界では, [a]の物理実体モデルを構築するために, 数千人の[a]の音を集め, 統計的にモデル化する。しかし, それでも電話音声の認識精度は急落する。集めた音と物理的に異なるからである。

近年の脳科学は, 聴覚皮質において音声の言語的特徴と非言語的特徴が分離されるとのモデルを呈している¹⁾。この主張が正しければ, 音声工学の方法論は, その価値を失う。例えば上記研究では, 音声の言語的特徴は, その「動き/差異」によって伝搬される, としている。音響音声学では, 母音は第一, 二フォルマントによって規定されることが多いが, これは「動き」でも「差異」でもない(上記に従えば, 言語的特徴ではない), 音の「実体」に過ぎない。これでは, 話者性は消えない。

本稿では, 音声の「動き/差異」のみを捉えることで, 話者の違いが消えた音声の全体的・構造的表象が得られることを示し, この全体表象のみを用いる認知戦略として失読症を, 全体表象が利用できない障害として自閉症を考える。

2. 音声の音響的普遍構造

話者, 年齢, 性別によって音声の音響的特徴はどのように変化するのか? その数学モデルを考え, そのモデルの上で音声を如何に変形しようと不変

な音響量が存在するならば, それが話者, 年齢, 性別不変の音響量となる。

音声工学では対数パワースペクトルを再度フーリエ変換して得られるケプストラム係数の低次項を用いてスペクトル包絡を表現する。このケプストラムベクトル c を用いると, 種々の非言語的要因による音声変形は下記のようにモデル化できる。

$$c' = Ac + b$$

ここで, 行列 A をかける演算は, スペクトルの周波数軸方向の変化(声道長差異など), b を足す演算は対数パワー軸方向の変化(マイク差異など)を表現する。これらに対する不変量として図 1 に示す, 音の動き/差異のみをバタチャリヤ距離として全て抽出することで(フォルマント等の絶対的な特性は全て捨象する)音声を構造的に表現する方法が提案されている(音響の普遍構造²⁾)。

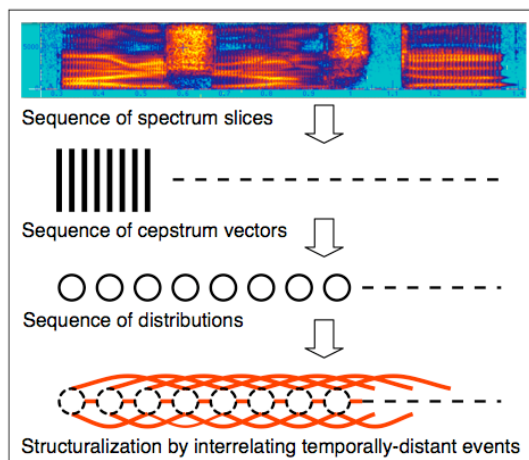


図 1 音声の全体的・構造的表象

なお、本表象を用いた音声認識も実験的に検討している。たとえば日本語孤立母音を5つ並べて語を形成した場合、語は120種類生成されるが、本表象を用いた場合、話者一人の音声参照パターンとして、不特定話者音声認識が100%の精度で行なえることを示している²⁾。

3. 音響的普遍構造の言語学的解釈

三辺の長さを規定すれば三角形は唯一に規定される。n角形の場合、nC2個だけ存在する2点間距離を規定すれば(鏡像という曖昧性は残るが)その幾何学構造を規定することになる。即ち、図1で示した全ての差異を求めるということは、音響空間において全音声事象群が張る幾何学構造を規定することになる。この構造が不変量となる。

これは、ヤコブソンが構造的音韻論にて示した音韻群の幾何学構造と非常に類似した音声表象である³⁾。更に言語学を遡れば、音響的普遍構造による音声認識は、近代言語学の祖として知られるソシュールによる主張、"The important thing in the word is not the sound alone but the phonic differences that make it possible to distinguish that word from the others"⁴⁾を、そのまま実験的に実証したに過ぎない、とも解釈できる。

図1より明らかなように、提案表象を用いても、個々の音を同定することは出来ない。音と音の差異を定量的に求め、その差異群によって語の全体的様態だけを表現している。そして、その様態は話者不変となる。その意味において本表象は、音声ゲシュタルトとして解釈されるが、語が音素の線状結合ではなく、ゲシュタルトであるとの主張は、トルベツコイに見ることができる⁵⁾。

4. 音響的普遍構造を通して考える音声認知

幼児は明確な音韻意識を有するようになる前から、親の声そのものではない「音響的何か」を模倣して、語を獲得する。発達心理学は、この現象を「語全体の音形」を獲得すると説明するが、音響的普遍構造は、一つの物理的解釈を与える。

上述したが、この音声表象を用いても、個々の音を同定することは出来ない。そもそも個々の音の絶対的な物理特性は全て捨象している。にも拘らず、2節で述べた認識タスクならば、単語同定が、入力話者に依らず100%可能である。このような頑健な枠組みを通して幼児が言語音体系を獲得し、かつ、その枠組みのみを持ち合わせた成人がいたとすれば、例えば、個々の音韻意識が非常に希薄なため、(表音)文字の読み書きに困難を示すが、音声コミュニケーションに支障は無い方々がいても何ら不思議ではない。また、仮に

このような症状を示す方が多数いたとすれば、それは、「音声⇄音シンボル列」変換を前提とする音声学・工学に対して、「その前提と言語能力とは本来無関係である」ことを実証することにもなる。

(少なくとも第一著者)は、上記の予測の下、失読症を知るに至った。「森が見えて木が見えない」症状とも解釈されているようだが⁶⁾、音響的普遍構造は、まさに森のみが見える音声表象である。

一方、音声ストリームに潜む不変項としてのゲシュタルトを抽出できない場合、即ち、音声の局所的・要素的特性ばかりを、絶対的に記憶する認知特性を持ち合わせていた場合、音声コミュニケーションは甚だ困難になるはずである。例えば異なる話者による「おはよう」という声の同一性が認知できなくても不思議ではない。これは音楽の絶対音感者が、移調された曲と原曲との同一性認知が遅れることと類似している。極端な音声の絶対音感者は、例えば、ある話者の声しか言語メッセージに変換できない場合もあるだろう。即ち、ある話者の[a]が「あ」という図形に対応するならば、他話者の[a]は「あ」ではなくなる訳である。何故なら、二つの[a]は音的に異なるからである。

(少なくとも第一著者)は、上記の予測の下、自閉症を知るに至った。「木が見えて森が見えない」症状とも記されている⁷⁾。母親の声しか言語メッセージにならない自閉症者もいるようである⁸⁾。

5. おわりに

音声に不可避免的に含まれる非言語的要因に不変な音響特徴量を模索する中で、音声工学の分野で提案した音響的普遍構造と、そこから予想される二種類の認知特性について説明し、失読症、自閉症との「奇妙な」一致について考察した。本考察の医学的妥当性については今後の検討が必要である。関係各位の意見を待ちたい。

<文献>

- 1) Belin et al., What, where and how in auditory cortex, Nature neurosciences, vol.3, no.10 (2000)
- 2) 峯松他, 音声の相対音感ー音声と音楽の同質性に関する一考察ー, 信学技法 SP2005-131 (2005)
- 3) R. Jakobson et al., Notes on the French phonemic pattern, Hunter, N.Y. (1949)
- 4) F. D. Saussure, Course in General Linguistics, McGraw-Hill, (1965)
- 5) N. S. Trubetzkoy, Principle of Phonology, Univ. of California Press (1969)
- 6) R. D. Davis, The Gift of Dyslexia, Perigee (1997)
- 7) U. Frith, Autism, Blackwall Pub. (1992)
- 8) 東田他, この地球にすんでいる僕の仲間たちへ, エスコー出版部 (2005)