

## 音声の相対音感 ～音声と音楽の同質性に関する一考察～

峯松 信明<sup>†</sup> 西村多寿子<sup>††</sup>

<sup>†</sup> 東京大学大学院新領域創成科学研究科 〒277-8562 千葉県柏市柏の葉5-1-5

<sup>††</sup> 東京大学大学院医学系研究科 〒113-0033 東京都文京区本郷7-3-1

E-mail: <sup>†</sup>mine@gavo.t.u-tokyo.ac.jp, <sup>††</sup>nt-tazuko@ams.odn.ne.jp

**あらまし** 音声コミュニケーションには、話者・環境・聴取者に起因する音響歪みが不可避免的に混入する。これら静的な非言語的歪みを数学的にモデル化し、そのモデルの上で、音響歪みを表現する次元を完全に失った音声の物理表象を提案している [1]。個々の音声事象の絶対的な物理特性は一切捨象し、音声事象間の関係のみを、全ての二事象間差異（コントラスト）の集合、即ち、ある幾何学構造として抽出する。この新しい物理表象は、構造音韻論の物理的実装として解釈されている。事象間のコントラストのみを捉える処理は、音楽の相対音感に類似した処理と考えられるが、本稿ではその提案表象を、言語学、心理学、言語障害学、神経生理学、脳科学、及び音楽学の観点から再度考察、解釈する。その中で、音素を音響空間内で定位する従来の方法論の是非について検討する。

**キーワード** 構造不変の定理、ゲシュタルト、失読症、相対音感、大脳皮質の均一性、側抑制、メタ表象

## Relative Sense of Speech

### — A Study of Equality between Speech and Music —

Nobuaki MINEMATSU<sup>†</sup> and Tazuko NISHIMURA<sup>††</sup>

<sup>†</sup> Grad. School of Frontier Sciences, Univ. of Tokyo, 5-1-5, Kashiwanoha Kashiwa, Chiba, 277-8562 Japan

<sup>††</sup> Grad. School of Medicine, Univ. of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0031 Japan

E-mail: <sup>†</sup>mine@gavo.t.u-tokyo.ac.jp, <sup>††</sup>nt-tazuko@ams.odn.ne.jp

**Abstract** In speech communication, acoustic distortions are inevitably involved by speakers, channels, and listeners. In our previous study, these distortions were mathematically modeled, and on that model, a novel speech representation was proposed where the distortions cannot be observed [1]. Absolute properties of speech events are completely discarded and only their interrelations are extracted as a full set of phonic differences or contrasts. The set is mathematically equal to a certain geometrical structure. This new representation is considered as physical implementation of structural phonology. Extraction of contrasts between two events is viewed as a process similar to hearing music, i.e. relative pitch. In this paper, the new representation is reconsidered from viewpoints of linguistics, psychology, language disabilities, neurophysiology, brain science, and musicology. Here, the conventional paradigm where a phoneme is localized absolutely at a certain point in an acoustic space is also reconsidered.

**Key words** Theorem of invariant structure, Gestalt, dyslexia, relative pitch, uniformity of the neocortex, lateral inhibition, meta-representation

### 1. はじめに

音声の生成、送受信、そして聴取。全ての過程において、不可避免な非言語的要因が音声の音響特性を変動させ、多種多様な声質・音質を生み出す。その結果、同一物理特性の音が、異なる話者の異なる言語音として同定されることもある [1]。音声工学はこの多様性問題を、膨大なる音声収集によって解く方法論を選択した。一方、乳児が体験する「個人性」は、母親（及び周囲の人間）が大部分を占めているにも拘らず、やがて、初対面の声でも容易に聴取できるようになる。更に、人の聞く声の半分は自らの声である事実を考えた場合、個人が体験する非

言語的要因というのは極めて大きな「偏り」があることは自明である。言い換えれば「人の耳が不特定話者音響モデルを有するのは困難である」と考えるのが自然である。この場合「話者、マイクの変化に対して、随時、話者・環境適応を行なっている」と音声工学者は主張するが、それは音声メディアを「最も忙しいメディア」とであると定義することとなり、彼等が「音声インタフェース」の実現を求める理由を、自己否定することになる。市販されている音声認識アプリケーションには「学習話者数が35万人」という事実を宣伝文句にしているものもあるが、そもそも話者数が35万人も存在しない言語は数多く存在している。

Poverty of stimulus. 乳児に対する言語入力 of 乏しきにも拘

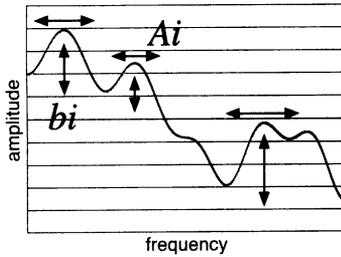


図1 水平方向及び垂直方向への不可避的な音響歪み

らず、乳児は言語を獲得する。チョムスキーはこの事実を説明するために普遍文法を提案し、言語の生得性を主張した。言語遺伝子など種々の議論があるものの、環境との相互作用、及びヒトという種が進化を通して獲得した認知基盤への着眼も必要であると筆者らは考える。さて、話者の多様性に対する適応処理も遺伝的に組み込まれているのだろうか？とすると、ヒトは電話機の発明を予測しながら進化したことになる。電話音声と自然音声等しい個体を、少なくとも筆者らは知らない。

Poverty of stimulus. 個人が体験する非言語的要因の多様性の乏しさの問題を解くために、筆者らは音響的普遍構造を提案した [1]。パラ言語情報の伝搬に寄与する基本周波数の情報を音声物理から分離するためにスペクトルスムージングが頻繁に行なわれる。であれば、非言語情報を伝搬する物理要因を音声物理から分離してしまえばよい。話者・環境・聴取者の違いが生む音響歪みを数学的にモデル化し、その上で構築された普遍的な音声の物理表象である。話者・マイクが見えない表象である。

本稿では、この音響的普遍構造について簡単に説明した後、この表象についての言語学、心理学、言語障害学、神経生理学、脳科学、音楽学各々における意義について検討する。なお、本稿は先行研究 [1] に対して更なる考察を付与する形をとっており、[1] を読まれてから本稿に目を通すことを勧める。

## 2. 音響的普遍構造

### 2.1 不可避的な非言語的要因による音響歪みのモデル化

音声認識では、音声に混入する音響歪みを加算性雑音、乗算性歪み、線形変換性歪みの三種類に分類することが多い。この中で、加算性雑音は物理的消去が可能という意味において不可避ではない。よって、ここでは考察対象としない。

乗算性歪みはマイクや伝送特性など、フィルタリングとして混入する歪みであり、GMM による話者モデルを考えると、話者性の一部も乗算性歪みである。これらは、ケプストラムベクトル  $c$  に対する加算  $c' = c + b$  として表現される。

話者間の声道長差異や、聴取者間の聴覚特性差異はいずれもスペクトルに対する周波数ウォーピングとして解釈され、これは、ケプストラムベクトル  $c$  に対して行列  $A$  を掛ける演算として精度良く近似される [2] ( $c' = Ac$ , 線形変換性歪み)。

結局不可避的な音響歪みは、スペクトルの水平方向 ( $A$ )、垂直方向 ( $b$ ) の変動に分解され、複数の変形要素による統合歪みは  $c' = Ac + b$  となる (アフィン変換)。本稿ではこれを非言語的要因による不可避な音響歪みの数学モデルとして採用する。

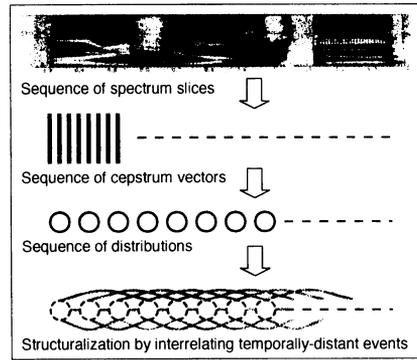


図2 発声の構造化に基づく音響的普遍構造

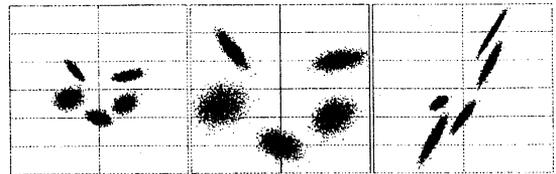


図3 ユークリッド平面への写像による構造歪み

### 2.2 音響的普遍構造 ~発声の構造化~

ケプストラムの一次変換によって不変でかつ、音声認識が可能な音声の物理表象を図2のようにして導出する。音声をケプストラム系列に変換し、更にそれを分布系列に変換する。ここで分布形状としては混合ガウス分布を仮定する。次に、全ての二分布間距離をバタチャリヤ距離、カルバック・ライブラ距離、ヘリング距離などを用いて算出する (距離行列の算出)。得られた距離行列をもって、この発声の物理表象とする。

上記の分布間距離尺度は、両分布に共通の一次変換を施しても距離は不変であるとの性質を有しており、その結果、距離行列は不変である。 $n$  点 (事象) から成る幾何学構造は、その距離行列によって構造を一意に規定することができることを考えると、上記性質は、音響空間において  $n$  事象が張る幾何学構造がアフィン変換によって不変であることを意味する (構造不変の定理、 $A =$  回転、 $b =$  シフト)。例えば図3の三つの分布群は何れも、行列  $A$  の乗算によって変換可能な分布群 (構造) であり、これらの分布群が呈する距離行列が等しくなる。即ち、音響的普遍構造はアフィン変換によって構造が歪まないように「空間を歪めて」解釈する方法である。図3はユークリッド平面に写像して観測するため、構造が歪んで「見えるだけ」である。

提案表象による音声認識が可能であることが示されている [3]~[5]。孤立母音系列 (系列長=5、即ち、PP=120) という原始的なタスクではあるが、1) 一人の話者の音声で不特定話者音声認識が可能であること、2) チャネル特性の違いによる影響を受けないこと、3) 加算性雑音下であっても非常に高い認識率を示すこと、4) 4000人以上の話者から構築した HMM よりも性能が良いこと、などが示されている。さて、提案手法は事象と事象の差異 (コントラスト) だけを捉え、事象群を構造的に表象し、事象の絶対的な物理特性は一切捨棄するという方法論をとっている。以下、この方法論と言語学、心理学、言語障害学、神経生理学、脳科学、音楽学との接点について検討する。



## 5.2 要素認知に困難を示すように見える障害

提案表象に基づく音声認識は、音声コミュニケーションにおいて個々の音声事象（単音）を絶対的に解釈することが不要であることを示唆する。これは、音の連続ストリームとして入力される音声を分割し、個々の単音やモーラの同定することはできないが、音声コミュニケーションは難なくこなせる話者、例えば「しりとり」が出来ない音声の使い手の存在を示唆する。発達性の失読症（難読症）がそれに相当する。「知的な遅れや視覚障害がなく、十分な教育歴と本人の努力にも拘らず、その知的能力から期待される読字能力が低い」症状を意味する<sup>(注1)</sup>。言語能力のうち「文字を読む」という側面のみが阻害されている状態である（日本語の場合平仮名の音読が困難となる）。当初視覚障害としての可能性が追求されたが、現在では、多くの研究者が音韻意識の欠如を原因として考えている。彼等は聴覚提示された単語をモーラや単音に分割すること、及び、個別に提示されたモーラや単音を繋げて単語を形成することが極端に困難である。即ち、失読症者は意味のある事象に名前を振り（語）、同定することはできるが、それ自体に意味の無い個々の音に名前を振って同定することが困難である。そして彼等の存在<sup>(注2)</sup>は、音声コミュニケーションにおいて、そもそも音韻意識は不要であることを示唆している。それは、音声を文字という媒体に変換する際に初めて必要となる能力でしかない。

## 5.3 失読症と相対音感

個々の（意味の無い）要素の同定が困難であるにも拘らず、要素群の全体を容易に認知する過程は音楽の相対音感と非常に類似している。絶対音感という言葉を「与えられた音に対して、他の音との相対的な比較を介さずに、有限個のラベルの一つを正確に割り当てる能力」として定義するならば、音韻意識は音声の絶対音感であり、失読症者はこれを有していない。音声の絶対音感が音声活動における必要条件で無いように、音楽の絶対音感も音楽活動における必要条件では無い。音楽を音符という媒体に変換する際に初めて必要となる能力でしかない[13],[14]。音楽の絶対音感者にとって「音楽を聞く」とは、音を音符に変換することであり、「歌う」とは、音符列を音に変換することであり、そこには常に音符の心的イメージが付きまとう[14]。健常者の音声聴取においては（例えば日本語なら）常に平仮名の心的イメージが付きまとう。筆者らにとって、このイメージを伴わない音声聴取を想像することは甚だ困難であるが、これは絶対音感者も同様である。彼等は、音符を伴わない音楽の聞き方、歌い方が理解できない。相対音感者が何故音楽を楽しむのか、歌えるのか、が理解できない。音素を物理的、絶対的に定義し、その系列でもって語を表現する方法論は、音声コミュニケーションの必要条件では無いのである。

上記は、失読症者と音楽の相対音感者間のビヘイビオレブルでの類似性であるが、以下に、脳活動における類似性を考察する。殆どの場合、言語活動の優位脳は左側脳半球に存在し、大

脳皮質の言語野は右半球の相同部位よりも拡大している事実は数々の脳科学研究が報告している。音楽聴取時の脳活動における左右半球の機能差に関しても数々の報告があるが、興味深いのは音楽従事者と非従事者における活動部位の差異である<sup>(注3)</sup>。音楽非従事者は右半球でメロディータスクを行なう傾向があり、従事者は左半球でメロディーを分析的に認知する傾向がある[16]。非従事者を対象とした場合、音を個別に提示すると左半球が、メロディーを提示すると右半球がより活性化される[17]。絶対音感者の左半球の聴覚野は右半球のそれより40%大きい[18]。専門的な音楽教育によって読譜時の賦活される脳領域が、右後頭葉外側部から左後頭葉外側部に変わるならば、それを指標として音楽教育の成果を評価できる可能性がある[19]。

失読症者の言語活動時の脳活動で特徴的なのは、左半球の優位性（言語野の非対称性）が減少していることが解剖学的にも、画像解析によっても観測されている事実である[20],[21]。更に興味深いのは、文字のエキスパートとも言える速読者が、速読時に示す脳活動が失読症者同様、左半球の優位性が減少する[22]。両者のビヘイビアは「読字が極めて困難」「読字が極めて早い」であり、一見共通項が全く存在しないように見えるが、前者が音韻意識が元来無い場合の読字活動であり、後者が故意に音韻意識を除外した（音読しない、音韻ループを介さない）読字活動であることを考えると、両者が類似した脳活動パターンを示しても、何ら不思議では無い。なお、筆者らが知る限り、失読症に関してその歴史、国際的な取り組みと教育界への提案、神経生理学的要因、更には、失読症が存在することの人類学・言語学的意義にまで言及した報告書として[23]がある。興味ある読者は、まずこれを参照することを勧める。

## 5.4 失読症と音響の普遍構造

筆者らの提案する音声の物理表象から「しりとりができない音声の使い手」の存在を予見し、その存在を確認することができた。しかし、彼等が本当に音声事象を相対的に捉えているかどうか、音的差異だけを用いた音声活動をしているかどうかについては、実験的な確認ができていない。現在、失読症研究者との議論を開始しているところである。提案表象が失読症の解明に貢献することができればと考えている。第5.1節において音素の絶対的音響定義と自閉症との類似性、第5.2節において音素の相対的音響定義と失読症との類似性を考察した。発達障害の研究において、この両者を対置して議論することは行なわれていないが、失読症（dyslexia）に相対する症状として、hyperlexia という症状がある。一般に予想される暦年齢よりも早い時期に文字を読む能力が表れ、文字、数字、スベル、単語に対する強い興味を示す。その一方で、文の理解が困難であり、言語発達は遅延する。特に聴覚からの情報の理解が苦手である。言語発達以外の側面においても、社会的スキル（人とかかわり合い）の発達が障害されている、常同行動が散見されるなど、自閉症や高機能自閉症との類似性は広く認知されていることを追記しておく。発達障害研究者との議論を重ねる予定である。

(注1)：知的水準は健常者の平均より上である、との報告もある。また文化的な著名人、政治家の中にも失読症者（だった方々）は存在する。

(注2)：欧米では（軽症も含めると）約5~15%の子供が失読症であると言われている。神経行動学的障害の中では最も多い障害と言われている。

(注3)：但し、左右半球の機能差を議論することの是非についての結論が出ている訳ではない。音楽従事者と非従事者の違いは、あくまでも部位の差であって、左右差とみるのはやや恣意的であるとの指摘もある[15]。

## 6. 神経生理学から見た音響的普遍構造

### 6.1 視覚生理学と聴覚生理学

「コントラストを抽出する」という処理系を視覚生理学の観点から考える。聴覚刺激は視覚刺激と異なり瞬時に消えるなどその扱いや視覚化が難しく、また、言語音を使った動物実験は意味を成さないなど、聴覚生理研究における制約の大きさは否めない。聴覚系と視覚系では「古い脳」の意味するところが異なるなど、差異はある[24]。また、大脳皮質の各領域の機能には聴覚・視覚・言語などの機能的差異もある。しかし、皮質の各領域の生理学的構造を考えた場合、各領域の差異よりも、その均一性にこそ着眼すべきとの議論もあり、皮質のコラム構造を提唱したマウントキャッセルは皮質処理の均一性を提唱している[25]。事実、網膜からの情報を視床で切り替えて聴覚野と接続しても「見る」ことができる例などが報告されている。

### 6.2 視覚情報処理系におけるコントラスト抽出

ニューロンの結合はシナプスを通して主に化学的に行なわれるが、この結合には興奮性と抑制性の二種類がある。興奮性ニューロンの出力が抑制性ニューロンを介して近隣のニューロンを抑制したり、興奮性ニューロンへの入力に抑制性ニューロンを介して近隣のニューロンを抑制するなどの回路が生体において発見されており、反回性・順向性側抑制回路と言われる。ニューロン間の（ネットワークとしての）近さが刺激の空間的近さを表現していれば空間微分の演算子となり、時間的近さを表現していれば時間微分の演算子となり、空間的あるいは時間的差異に反応する回路となる。視覚野での視覚刺激の表象の一つは、輪郭のみが抽出された表象であるが、これは空間微分の成せる技である。一方、皮質ではないが網膜の視細胞は光の強度の「差」に反応することが知られている（時間微分。dPCMと言った方が分かり易いだろうか）。この場合差異に反応するため、刺激を静的なものにすると世界が消失することになるが、この消失を防ぐために眼球は常に微動している（固視微動）。固視微動を踏まえた上で網膜の像を固定すれば十秒ほどで本当に世界は消失する<sup>(注4)</sup>。さて、側抑制を微分演算子と考えた場合、時間的・空間的に隣接した領域の差異のみが対象となるが、視野のある場所に表れた面の見え方が、空間的に、或は、時間的に離れた刺激によっても影響されることが知られている。所謂マスキングであるが、面が現れた30~100ms後に現れた刺激によって最初の面の見えが顕著に低下する（差が強調される）現象が知られており、メタコントラストと言われている[26]。

このように、差異に着眼する処理は、感覚器である網膜のみならず、中枢でも（例えば視床の側抑制的な受容野特性など）広く観察される特性である。さて、第一次視覚野には、固視微動による「動き」の情報を含めたイメージが到着する。このままでは例えば「地」が動いていると解釈される。固視微動は末梢系の処理であるため、脳はモニターすることができず、固視微動による「動き」情報の分離は視覚野に到着した情報にのみ

基づいて行なう必要がある。[27]では、「視野全体の動き」の情報から「地の動き」の情報を差し引くことによって（相対運動に落とす）この問題を解決するモデルを提案している。視野全体に渡る大局的な（共通の）動きを差し引き、分離する訳である。

### 6.3 コントラストの差異（差異の差異）の認知

「動き」という情報は如何にして視覚的に検出されるのだろうか。一般に、視野の各領域における輝度に変化することによって「動き」認知が行なわれると考えられるが、輝度変化に依らない「動き」が存在することが知られている。各領域の平均輝度は変化しないが「動いている」との認知がされる訳である。それは、コントラストの変化に基づく「動き」である。前者は一次運動（光の「差」による動き）、後者は二次運動（「差の差」による動き）と呼ばれている。二次運動は自然界の中では稀であるにも拘らず、視覚系は二次運動を検出できる。何故出来るのか、ついて興味深い議論がある。[29]では二つの仮説を提案しており、一次運動の認知処理が、たまたま、二次運動も認知できるようになった、とするものと、二次運動の認知が一見するよりも重要な生物学的意味を持ち、生体はそれを検知するメカニズムを発達させてきた、とするものである。

第2.2節で、提案表象による音声認識について述べた。音的差異を全て求めることで単語の表象としている（距離行列）。ここで構造間差異（音響的照合）は距離行列をベクトルと見なした際のユークリッド距離を求める処理を基盤としている。こうすることで、片方の構造を回転&シフトさせた後の対応する事象間距離の総和の最小値が近似できるからである[3]。即ち「音的差異の差異」を検出して音声認識を行なわせている。

### 6.4 普遍的なトップダウン的メタ表象に対する探求

本節では視覚に限らず、薬学、脳科学、神経心理学における研究者達の、事象の抽象化に基づく普遍的なメタ表象に対する「書籍による」興味深い主張を示す。「記憶が厳密なものであれば、変化を続ける環境の中では無用な知識となります。脳は絶対的な事象の抽出ではなく、その範囲における相対的な文脈の連辞関係を記憶しています」[30]<sup>(注5)</sup>「同様の抽象化が新皮質全体に渡って行なわれていると確信している。記憶される形式は関係の本質を捉えたものであり、ある瞬間の詳細ではない。新皮質は極めて具体的に詳細な入力を受け取って、それを普遍の表現に変換する。記憶も、それと比較される新しい入力パターンも、普遍の表現である」[31]、「感覚表象そのものにはクオリアはなく、それが経済的に符号化される過程、あるいは扱い易いチャンクとして準備される過程でクオリアを獲得し、その結果、高次のメタ表象ができていく」[32]、「二次表象は、一次表象のある面を強調し、その後に行なう新規な様式の計算—即ち、内部で連続的なシンボル操作（思考）をするための計算か、あるいは一次元の音声の流れ（言語）を通して他者とアイディアをやりとりするための計算のどちらかを円滑にするトークンを作り出すのに役立っています」[32]。以上の主張は、第1.節に示した筆者らの思考の妥当性を十分示唆するものである。事

(注4)：ソシュールは「言語は概念的差異と音的差異の体系である」と主張したが、網膜は「視覚世界は光的差異の体系である」と確信に語っているのだろうか。

(注5)：そのすぐ横で、「ソシュールは、意味や言語は相対的な差異の構造を反映しているに過ぎない、と述べていますが」と記されているのも興味深い。

実、「不変的な関係」の抽出の様子を神経生理学的に突き止めた研究もある[33]。更には、物理刺激が存在しなくても（実際に見なくても）ある対象物をメタ表式的に想像することを可能たらしめるトップダウン信号が見つまっている[34]。実際に「見る」場合は感覚野からのボトムアップ信号と前頭葉からのトップダウン信号とが相互作用する。提案表象は各事象の絶対的な物理特性を一切捨象した、事象群全体を描く抽象的な表象である。これをトップダウン表象と考えれば、これと、従来の個々の事象を描くボトムアップ表象とを融合させた形での処理系が、本来求められるべき処理系であると筆者らは考えている。

## 7. 脳科学・音楽学から見た音響の普遍構造

脳活動イメージングによって、音楽聴取時の脳活動の詳細がより明らかになってきている。特に興味深いのは、音楽における音的構造認知と言語の統語構造認知が同一の脳部位によって支えられている（左右差が存在することもある）、という実験事実である[15],[35],[36]。言語の専門領域と考えられたいブローカ野が音楽聴取に関与している報告もある[37]。音構造の物理的表象として[38]に記載のある全和音から成る距離行列や、[39]で提案された全ての音事象から成る距離行列を考えた場合、筆者らが提案する音声構造と言語の統語構造と神経生理学的な関係があることが示唆される。[1]において、言語の統語構造の起源を音声構造に求める記述をしたが、上記の論文はその考察の妥当性を示唆する。なお、聴覚刺激の構造化、体制化を言語の統語構造の起源と考えるのは筆者らだけではない[40]。

## 8. まとめ

筆者らが提案している音響の普遍構造に対して幅広い観点から再考、熟考し、音声表象としての妥当性について検討した。失読症者が音声認識装置を作ろうとした時に、彼等は音素を音響的に、絶対的に、定位する方法論を生み出すのだろうか？筆者らの問いかけの真意を理解して戴けたら幸いである。

### 文 献

- [1] 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”, 信学技法, SP2005-12, pp.1-8 (2005-5)
- [2] M. Pitz *et al.*, “Vocal tract normalization equals linear transformation in cepstral space,” IEEE Trans. Speech Audio Processing, vol.13, no.5, pp.930-944 (2005)
- [3] 峯松信明他, “音声の構造的表象とその距離尺度”, 信学技法, SP2005-13, pp.9-12 (2005-5)
- [4] 村上隆夫他, “音声の構造的表象を用いた日本語母音系列の自動認識”, 信学技法, SP2005-14, pp.13-18 (2005-5)
- [5] 村上隆夫他, “音声の構造的表象を用いた雑音環境下における日本語母音系列の自動認識”, 信学技法, SP2005 (2005-12, 予定)
- [6] H. A. Gleason, An introduction to descriptive linguistics, New York: Holt, Rinehart & Winston (1961)
- [7] F. D. ソシュール, 一般言語学講義, 岩波書店 (1940)
- [8] R. Jakobson *et al.*, Preliminaries to speech analysis, MIT Press, Cambridge, MA (1952)
- [9] R. Jakobson *et al.*, Notes on the French phonemic pattern, Hunter, N.Y. (1949)
- [10] N. S. トルベツコイ, 音韻論の原理, 岩波書店 (1958)
- [11] U. フリス, 白閉症の謎を解き明かす, 東京書籍 (1991)
- [12] 小嶋秀樹他, “ロボットに「心の理論」は教えられるか?”, 発達心理学シンポジウム資料 (1998-3)
- [13] 宮崎謙一, “絶対音感”はどこまで分かったか?, 音響学会

- 誌, vol.60, no.11, pp.682-688 (2004)
- [14] 最相葉月, 絶対音感, 小学館 (1998)
- [15] R. Zatorre, “Music, the food of neuroscience?” Nature, 434, pp.312-315 (2005)
- [16] T. G. Bever *et al.*, “Cerebral dominance in musicians and nonmusicians,” Science, 185, pp.537-539 (1974)
- [17] D. Breitling *et al.*, “Auditory perception of music measured by brain electrical activity mapping,” Neuropsychologia, 25, pp.765-774 (1987)
- [18] G. Schlaug *et al.*, “In vivo evidence of structural brain asymmetry in musicians,” Science, 267, pp.699-701 (1995)
- [19] 岩田誠, 脳と音楽, メディカルレビュー社 (2001)
- [20] K. R. Pugh *et al.*, “Functional neuroimaging studies of reading and reading disability,” Mental retardation & developmental disabilities research review, 6, pp.207-213 (2000)
- [21] M. Eckert, “Neuroanatomical markers for dyslexia: a review of dyslexia structural imaging studies,” Neuroscientist, vol.10, no.4, pp.362-371 (2004)
- [22] N. Fujimaki *et al.*, “Neural activation dependent on reading speech during covert reading of novels,” NeuroReport, vol.15, no.2, pp.239-243 (2004)
- [23] 石井加代子, “読み書きのみの学習障害への対応策”, 文部科学省科学技術政策研究所・科学技術動向研究センター, 科学技術動向 45, pp.13-24 (2004-12)
- [24] 辻野広司, “脳型情報処理から見たロボット聴覚:「脳とからだをもった耳」”, 人工知能学会研究会資料, SIG-Challenge-0522-6, pp.35-40 (2005-10)
- [25] V. Mountcastle, “An organizing principle for cerebral function: the unit module and the distributed system,” in G. M. Edelman and V. B. Mountcastle, eds., The Mindful Brain, MIT Press (1978)
- [26] H. Kondo *et al.*, “Suppression on neuronal responses by a metacontrast masking stimulus in monkey V4,” Neuroscience Research, 36, pp.27-33 (2000)
- [27] I. Murakami *et al.*, “Visual jitter: evidence for visual-motion-based compensation of retinal slip due to small eye movements,” Vision Research, 41, pp.173-186 (2001)
- [28] S. Nishida *et al.*, “Neuroimaging of direction-selective mechanisms for second-order motion,” J. Neurophysiology, 90, pp.3242-3254 (2003)
- [29] A. T. Smith *et al.*, Visual detection of motion, Academic Press (1994)
- [30] 池谷裕二, 記憶力を強くする, 講談社 (2001)
- [31] J. ホーキンス, 考える脳, 考えるコンピュータ, ランダムハウス講談社 (2005)
- [32] V. S. ラマチャンドラン, 脳のなかの幽霊, ふたたび, 角川書店 (2005)
- [33] P. Paavilainen *et al.*, “Neuronal populations in the human brain extracting invariant relationships from acoustic variance,” Neuroscience, 265, pp.179-182 (1999)
- [34] H. Tomita *et al.*, “Top-down signal originating from the prefrontal cortex for memory retrieval,” Nature, 401, pp.699-703 (1999)
- [35] A. D. Patel, “Syntactic processing in language and music: different cognitive operations, similar neural resources?” Music perception, vol.16, no.1, pp.27-42 (1998)
- [36] S. Koelsch, “Neural substrates of processing syntax and semantics in music,” J. Neurobiology, 15, pp.1-6 (2005)
- [37] B. Maess, “Musical syntax is processed in Broca’s area: an MEG study,” Nature neuroscience, vol.4, no.5, pp.540-545 (2001)
- [38] C. L. Krumhansl, “The geometry of musical structure: a brief introduction and history,” ACM Computers in Entertainment, vol.3, no.4 (2005)
- [39] 後藤真孝, “リアルタイム音楽情景記述システム”, 情報処理学会音楽情報科学研究会資料, 2002-MUS-47-6, pp.27-34 (2002)
- [40] 科研費研究プロジェクト基盤 (S)「聴覚の文法: 言語と非言語を包括する体制化の研究」(研究代表者: 中島祥好@九州大学)