



## Kobe Shoin Women's University Repository

Title	PHONON : 音韻のプロトタイプカテゴリー Phonon as a Phonological Prototype Category
Author(s)	松井 理直 (Michinao Matsui)
<i>Citation</i>	Theoretical and applied linguistics at Kobe Shoin, No.3 : 47-81
Issue Date	2000
Resource Type	Bulletin Paper / 紀要論文
Resource Version	
URL	
Right	
Additional Information	

# PHONON — 音韻のプロトタイプカテゴリー\*

松井理直

---

## Phonon as a Phonological Prototype Category

Michinao Matsui

Categorization is fundamental to all higher cognitive activities. Yet two different approaches to categorization raises deep philosophical and linguistic problems: one is discrete category (e.g. binary features), the other is continuous category (e.g. prototype categories).

In this paper, I will propose a integrated concept called '*phonon*', which is a basic unit of phonological representations on Head-driven Phrase Structure Grammar (HPSG). Furthermore, Phonon has both qualitative inner structures which achieve the same effects as the underspecification theories and quantitative inner structures including the interaction systems between articulatory and perceptual cost-functions. This idea can also combine phonology with phonetics.

カテゴリー化は、高次の認知活動にとって必要不可欠なものであり、その性質について多くの議論がなされている。本稿では、音韻表象の基本単位であり、定性的性質と定量的性質を併せ持つ *phonon* という概念を提案する。定性的性質は、いわゆる不完全指定理論と同質の効果をもたらし、定量的性質は、調音コスト関数と知覚コスト関数の相互作用として表される。

---

\*本研究は、平成 11・12 年度文部省・学術振興会科学研究費 (奨励研究 (A), 課題番号: 11710291) 「日本語動詞形態素の獲得に関する理論的・実証的研究」の援助を受けている。

*Theoretical and Applied Linguistics at Kobe Shoin* 3, 47-81, 2000.

© *Kobe Shoin Institute for Linguistic Sciences.*

## 0. 序論

心的表象 (mental representation) とは、ある情報やカテゴリー化を明示する形式系であり、<sup>1</sup>同時にその明示方式も示したものである。

認知機構において、こうした情報の表現方式はきわめて重要である。表象の種類によって、どのような情報が明示され、どの情報が明示されないか、また様々な実体がいかなるカテゴリーに属するかが限定されてしまうからである。例えば 2 進法で “1001” という表現を持つ実体と、10 進法で “9” と表現された実体は、実体としては等価である。しかし、演算体系として、もし 2 のべき乗であるかどうか重要な場合には、2 進法の表現が好ましい。コンピュータの内部計算がこれに相当する。一方、10 のべき乗であるかどうか重要な場合には、10 進法の表現を採用したほうが、演算そのものも単純になる。「そろばん」の機構は 2 進法よりも 10 進法の表象に適したものであろう。音の世界でいうならば、octave 感覚は 2 進法の表現、dB 単位系は 10 進法の表現がより適切である。このように、ある実体を表現する際に適切な表象を用いなければ、情報を犠牲にし演算や復元を極めて困難なものとしてしまう。

言語認知機構においても事情は同じであり、言語表象の性質は避けることのできない問題である。中でも、音声／音韻に関わるモジュールは音声特有の問題に直面する。このモジュールが、音響としての物理的な実体と、言語としての抽象的な実体を共に処理しなければならないからである。ここで、物理的な実体である「音」は連続的なものであり、抽象的な実体である「音韻」は有限個の離散量であると考えられる。したがって、音韻表象の問題は、連続量と有限な離散量とを結びつけるカテゴリー化の問題と言い換えることもできる。本論文では、こうした点を踏まえ、音韻表象の内部構造であるカテゴリーの性質およびその計算体系の妥当なレベルを見いだすことを目的とする。

なお、議論に入る前に、2つの前提を置くことにする。第一の前提は、言語認知機構は 1つの独立した心的モジュールであり、言語に関わる表象が心的に実在するというものである。二つめは、認知機構には記号的操作を行うシステムと前記号的計算を行うシステムがあるが、言語モジュール自体は明示的な記号操作システムであるという前提である。いずれの前提も自明なものではなく、心的表象は不必要であるという議論 (Brooks 1991) や、明示的操作体系なしに、一般的学

<sup>1</sup>形式系とは、それが操作規則を持つ記号の集合であることを意味する。

習機構で言語知識を習得するシステムに関する研究 (Elman et al. 1996) が存在する。本稿では、表象の心的実在性と記号操作体系を前提とし、議論の直接の対象にはしないが、独立して議論されるべき重要な問題であることは、強調しておくべきであろう。

## 1. 言語カテゴリーに関する先行研究

表象の性質、すなわちカテゴリーの考え方については、大きく分けて2種類のアプローチが議論されてきた。1つはある属性を持つか持たないかという二項対立に基づく離散的カテゴリーの考え方であり、もう一方は属性の連続的な変化を認めるプロトタイプカテゴリーの考え方である (Tayler 1995, Lakoff 1987, 河上 1996, 山梨 1995)。

離散的カテゴリーは古典的カテゴリーとも呼ばれており、(1) に示すような特徴がある。

- (1) a. 素性はそれ以上分解できない原子要素的 (atomic and primitive) な性質を持つ。
- b. 素性の値は二項対立的である。
- c. カテゴリーは必要十分な素性の連言によって定義される。

(1) は、カテゴリーが矛盾律と排中律に従うことを意味している。1つのカテゴリーは、ある素性を有するか有さないかのいずれかであり、また、ある対象は1つのカテゴリーに属するか属さないかのいずれかでなければならない。言い換えるならば、カテゴリー間に、明確な境界が存在することになる。

言語において、こうした二項対立に基づく離散的カテゴリーを用いる最大の利点は、カテゴリーを素性によって差異の体系 (Saussure 1916) として表示することが可能になる所にある。そのため、言語理論においても、Jakobson et al. (1957) から Optimality Theory (Prince and Smolensky 1993) に至るまでの音韻論で標準的に使われている枠組みとなっている。また、Liberman (1957) らが行った子音の弁別実験において、子音間のカテゴリー境界が極めて急峻であることが示されており (子音の範疇的知覚)、離散的カテゴリーはある種の心的過程とも矛盾しない枠組みといえよう。

これに対し、プロトタイプカテゴリーの考え方には、いくつかの異なったアプローチがある。中でも重要なものは、Wittgenstein (1953) の「家族的類似性」に基づく哲学的考察、および Roshe (1977) による心理学的な研究であろう。これらの多くのアプローチに共通する「プロトタイプ」の概念は、次のようなものである (Tayler 1995, 河上 1996)。

- (2) a. カテゴリー間には重複が許される (分布性、家族的類似性)。
- b. プロトタイプとは、中心的な成員の抽象的表象である。

こうしたプロトタイプカテゴリーの性質は、自然種を表す語の意味を考える上で特に有効なものである。音韻論においても、Anderson and Durand (1987) らによる *dependency phonology* では、母音の音質 (quality) を 4 つの成分 ([i], [u], [a], [ə]) の混合と依存関係によって示すという枠組みを用いており、これは母音を相互排他的な類と見なさないという点で、プロトタイプカテゴリーのアプローチに近い。*Particle Phonology* (Shane 1983) や *Government Phonology* (Charette 1991, Harris 1995) も同様である。さらに、プロトタイプカテゴリーが、音声の心的性質として有効な枠組みであることを示す証拠もいくつか存在する。例えば、Fry (1962) は母音の知覚カテゴリーの実験で、母音カテゴリー間の境界が緩やかに連続的に変化していくこと (母音の連続的知覚) が示したが、これはプロトタイプカテゴリーの特徴である。また、Jaeger and Ohala (1984) は、音声の有声性や sonority に関する判断が範疇的に変化するのではなく、段階的に変わりうることを示している。日本語音声では、北条 (1982) が子音の類似性判断の実験から、子音の心的空間を得た研究がある。この研究によると、独立した弁別素性に対応する空間軸はなく、[strident] と [coronal] が関わっている軸、[-sonorant] かつ [+voiced] かどうかを対立させる軸などが抽出されている。これらの結果は、離散のカテゴリーよりもプロトタイプカテゴリーの心的実在性をサポートするものといえよう。

このように、心的表象のカテゴリーとしては、離散のカテゴリーとプロトタイプカテゴリーのそれぞれに長所と短所があることがわかる。事情は、言語認知機構で使われる心的表象においても同様である。さらに困難なことに、心的実在としての音韻表象は、連続量と有限な離散量とを結びつけるカテゴリーでなければならない。こうした問題を解決する 1 つのアプローチとして、本論文では離散のカテゴリーとプロトタイプカテゴリーの性質を共に持てるような音韻表象である

*phonon* という概念を提案してみたい。<sup>2</sup>

## 2. *phonon* の特性

### 2.1 音韻表象の基本単位

まず、*phonon* の基本的な性質について見てみよう。これらの性質は、Charette (1991) による *government phonology* で議論されている *phonological elements* と、ほぼ同一の概念であるが、いくつかの点で重要な拡張が行われている。

(3) a. *phonon* は音韻表象の基本単位である。

b. 各 *phonon* は、音声知覚と関連する音響的性質、および調音運動に関わる性質を内部構造として持つ。

(3a) の性質は、音素や音声 (基底形や表層形) は、*phonon* 単独で表されるか、あるいは *phonon* が複数集まった集合体として表示されることを意味している。例えば、日本語の音声は A (開口性)、I (口蓋性)、U (円唇性)、@ (軟口蓋性)、R (舌端性)、q (破裂性)、h (摩擦性)、N (鼻音性)、V (有声性)、v (無声性) という 10 個の *phonon* を基本単位とし、母音 a は A 単独、母音 e は I と A の複合体、子音 t は v, q, R の複合体などとして表されることになる。

一方、(3b) の性質は、*phonon* が弁別素性に相当するような性質を内部構造として持ちうることを意味する。例えば、鼻音性に関わる N という *phonon* は、調音運動の素性として [+声帯振動, +鼻腔共鳴] のような内部構造を持ち、同時に音響知覚特性の性質として [+有声性, +鼻音性] というような内部構造を持つ。

さらに (3) は、*phonon* を用いる音韻理論の体系 *S* に、3 つの重要な特性を与える。まず、この性質は、*phonon* という音韻の基本単位が最小単位ではないことを意味する。したがって、*S* は要素還元的ではあるが、最小要素には還元されないため、強い要素還元主義の持ついくつかの問題点から逃れることができる。次に、最もコンパクトな *S* を構築する場合においても、基本単位間の性質が完全に独立である必要はなく、互いに関連し合うことが許される。これは、例えば音声の家族的類似性や不完全指定 (*underspecification*) に関わる現象などを、ごく自然に説明することにつながる。最後に、*phonon* が内部構造として音響知覚的特徴と

<sup>2</sup>音韻表象 (*phon*) の基本要素 (*-on*) を表すものであり、物理学で使われる用語とは無関係である。

調音の特徴を持つことから、理論体系  $S$  は「調音運動と音声知覚のループ」を本質的な特性として備えることになる。後に見るように、調音運動と音声知覚の相互作用は、音声の心的処理過程における重要な特性の1つである。次節では、これら3つの特性を具体的に観察してみよう。

## 2.2 還元主義と表示のレベル

この論文では、phonon を心的表象の1つと見なしていることから分かるように、心理主義の立場を前提としている。すなわち、言語理論は言語の心的実在あるいは心的処理過程に関するなんらかの説明原理であるという立場である。

心理学的事象の説明概念の水準を微視的な最小基本単位(心理は生命体に属するものであるから、これはしばしば生理学的単位と等価である)に帰着させる立場が、心理学における還元主義(reductionism)である。しかし、ゲシュタルト心理学(Gestalt psychology)が明らかにしたように、感覚・知覚の領域では、「全体は部分の総和以上のもの(ゲシュタルト質)」であり、心理的構造は多くの場合ゲシュタルト質を持つことが知られている。例えば、あるメロディを移調して演奏しても、メロディ全体としての感じは大きくは変わらない。これは、メロディが全体としてまとまるゲシュタルト質を帯びているからである(ゲシュタルトの移調性)。さらに、このメロディが個々に組み合わせられた上位構造である「音楽」は、調性や音楽の動機(motif)という独自の情報を持つ。逆に、メロディの下位構造である個々の複合音は、いくつかの純音から構成されているが、純音単独では持ち得ないある種の「音色」を情報として持つ。このように、心理学的事象は、各レベルごとに全体としてまとまった1つの特性が現出する。したがって、説明のための概念の水準を微視的単位に還元することは困難であるといえよう。ある心理事象は、特定のレベルにおいて、創発(emergent)するものなのである。

もし、言語理論が心的現象の説明原理であるならば、同様のことが成立しなければならない。まず、可能ならば、理論の基本単位となる表示レベルも、下位構造を持つことが望ましい。こうした下位構造が存在した場合、「言語」としての心的過程と、それよりもより微細な認知過程との関係を規定することが可能になるからである。phonon が音韻の基本単位であり、かつ内部構造を持つと定義されているのは、この条件を満たすためである。

次に、ある属性(下位構造)を持った表示が、より上位の表示の下位構造になれるような枠組みを持っていることが望ましい。こうした性質は、人工知能にお

いては、タイプ継承 (type inheritance) として知られている。多くの生成文法理論は、階層構造を用いることにより、この条件を満たしている。その中でも、タイプ継承の考え方を最も明示的に用いている文法理論が、*Head-driven Phrase Structure Grammar* (HPSG; Pollard and Sag 1987, 1994) および *Japanese Phrase Structure Grammar* (JPSG; Gunji 1987, Gunji and Hasida 1998) である。

HPSG や JPSG では、言語情報を取り扱う際の形式的な装置として、素性名 (feature name) と素性値 (feature value) の対である素性構造 (feature structure) と呼ばれる次のようなデータ構造を用いる。

$$(4) \left[ \begin{array}{ll} \mathbf{Feature1} & \mathit{value1} \\ \mathbf{Feature2} & \mathit{value2} \end{array} \right]$$

素性値は、それ以上細かな構造を持たない原子シンボル (atomic symbol) で表現される原子素性構造 (atomic feature structure) の形を取ることもあれば、以下に示すように素性名と素性値から成る複合素性構造 (complex feature structure) を取ることもできる。なお、カテゴリーやタイプの考え方から見ると、(5) における **Feature3** は **Feature1**, **Feature2** における上位カテゴリーや上位タイプとして捉えられる。

$$(5) \left[ \begin{array}{ll} \mathbf{Feature3} & \left[ \begin{array}{ll} \mathbf{Feature1} & \mathit{value1} \\ \mathbf{Feature2} & \mathit{value2} \end{array} \right] \\ \mathbf{Feature4} & \mathit{value4} \end{array} \right]$$

この素性構造を使うと、日本語の分節音の構造は、phonon を原子シンボルとする次のような構造によって示される。ここで、curly brackets “{ }” は、素性値として複数の可能性、すなわち選言的關係を表すものとする。

$$(6) \left[ \begin{array}{ll} \mathbf{seg} & \left[ \begin{array}{ll} \mathbf{major} & \{V, v, N\} \\ \mathbf{manner} & \{q, h, \langle q, h \rangle\} \\ \mathbf{place} & \{A, I, U, R, \emptyset\} \end{array} \right] \end{array} \right]$$

この枠組みを使うと、例えば e 音や t 音などは以下のように表現される<sup>3</sup>。なお、これらの表示を簡略して、[A, I], [v, R, q] のように表記することもある。

<sup>3</sup>タグは素性の共起關係を示す。



$$(7) \quad e = \left[ \begin{array}{c} \text{seg} \left[ \begin{array}{c} \text{major} \\ \text{manner} \quad \{A, \textcircled{1}\} \\ \text{place} \end{array} \right] \\ \text{head} \quad \textcircled{1} \text{ I} \end{array} \right]$$

$$(8) \quad t = \left[ \begin{array}{c} \text{seg} \left[ \begin{array}{c} \text{major} \quad v \\ \text{manner} \quad \textcircled{1} \\ \text{place} \quad R \end{array} \right] \\ \text{head} \quad \textcircled{1} q \end{array} \right]$$

両者は、いずれも複数の phonon の集合体として表現されており、かつそのうちの 1 つの phonon が中心的な役割を果たしている (**head** 素性になっている) 点で共通している。ただし、(7) のような母音は、同じカテゴリー (**manner** 素性) に属する phonon が融合してできたものであるのに対し、(8) のような子音は、一般に異なったカテゴリーに属する phonon の集合体として表現されている点が異なる。

さらに、これらの要素は上位の構造に取り込まれる。JPSG に基づく音韻論では、明示的な素性として表現される最上位の音韻構造は **mora** 素性である。例えば、(8) と (7) を組み合わせた [te] という拍 (モーラ) は、(9) のような構造で表される。

$$(9) \quad [te] = \left[ \begin{array}{c} \text{mora} \left[ \begin{array}{c} \text{onset} \left[ \begin{array}{c} \text{seg} \left[ \begin{array}{c} \text{major} \quad v \\ \text{manner} \quad \textcircled{1} \\ \text{place} \quad R \end{array} \right] \\ \text{head} \quad \textcircled{1} q \end{array} \right] \\ \text{peak} \left[ \begin{array}{c} \text{seg} \left[ \begin{array}{c} \text{major} \\ \text{manner} \quad \{A, \textcircled{2}\} \\ \text{place} \end{array} \right] \\ \text{head} \quad \textcircled{2} \text{ I} \end{array} \right] \end{array} \right] \end{array} \right]$$

JPSG 音韻論では、**mora** 素性以上の音韻構造は明示されず、タイプ *mora* がどのような上位タイプに属するかという、タイプ継承によって表現される。上位の韻律構造を、明示的な構造ではなく、タイプ継承で表す理由は 2 つある。第一の理由は、日本語の音韻論においては、**mora** を支配する上位の音韻構造である **syllable** と **foot** が純粹に階層構造を保たないことである。どちらのタイプも、**mora** を直

接支配するものであり、両者間の関係は alignment などの制約によって影響され合うだけである。こうした木構造でいう枝の交差は、素性構造のような再帰的な構造では基本的に許されない。また、日本語音声の心的処理においては、後に見るように、*mora* の影響は頑強であるが、それ以上の音韻構造は、アクセントやイントネーションなどの音調に対して、その状況に依存してのみ影響する。こうした変動の大きい現象は、明示的な素性に関わる制約というよりも、あるタイプに特定の音調に関する制約が属しており、そのタイプが実現された時のみ制約が働いて、下位の構造には見られない現象が創発すると考えたほうがより適切であろう。吉本 (1993) は、こうした観点から HPSG のタイプ継承を用いて、日本語アクセントの様々な現象を説明している注目すべき研究である。

### 2.3 家族的類似性と不完全指定

Jaeger and Ohala (1984) は、音声学の訓練を受けたことのない英語話者を被験者にし、彼らが [voiced] などの弁別素性に基づいて、英語の語頭音を分類できるかどうかを検証した。被験者はいくつかの子音について [voiced] 素性による分類の訓練を受け、その後、訓練では出てこなかった子音を含むテストを受けた。その結果、子音は [+voiced] / [-voiced] のように明確に区別されるのではなく、鼻音が最も有声性の強いカテゴリーであり、無声破裂音が最も有声性の弱いカテゴリーとなるような、一つの連続体を形成することが示された。

phonon の持つ内部構造は、こうしたカテゴリー間の連続性を保証するものである。例えば、素性構造 (6) は、**major** というカテゴリーに、V, v, N という要素が属することを示したものと捉えることもできる。タイプ継承の枠組みで考えるなら、これら 3 つの phonon は、**major** というタイプに属するトークンである。これらの phonon は、有声性、無声性、鼻音性を表す基本単位であり、内部構造として、[+voiced, -nasal], [-voiced, -nasal], [+voiced, +nasal] という音響知覚的性質を持つ。この時、[+voiced] という性質は、N と V を関連づけ、[-nasal] という性質は、V と v を関連づけている内的性質と見なすことができる。これらの性質によって、**major** の要素が、N, V, v という順序で連続体をなす事が保証される。これは、Wittgenstein (1953) のいう家族的類似性 (*family resemblance*) であり、(2) にも示したように、プロトタイプカテゴリーの特徴の 1 つといえよう。また、こうした phonon の内部構造は、[+nasal] であるなら [+voiced] であるといった含意関係をも内的な性質として自然に示しうる。不完全指定理論では、指定されてい

ない値を補填するために、文脈自由・文脈依存規則が要求されるが、phonon のシステムではこうした規則が不必要になり、文法システムの経済性の点からも1つの優れたアプローチであるといえよう。

#### 2.4 音声産出と音声知覚の連鎖

音声の心的処理システムを考える場合、音声産出と音声知覚の相互依存関係は極めて重要な性質である。いくつかの実験、例えば、遅延聴覚フィードバック (Lee 1950)、ランバード効果、変換聴覚フィードバック (河原 1995)、日本人の /r-l/ 習得実験 (Yamada 1995) は、こうした相互作用の明確な証拠となるものである。また、代表的な音声知覚仮説である *articulatory reference theory of speech perception* (Liberman 1957) や *motor theory of speech perception* (Liberman et al. 1962), *Analysis-by-Synthesis model* (K. N. Stevens 1960) など、音声産出と音声知覚の相互作用をモデルの中心においている。例えば、*Analysis-by-Synthesis model* では、聞き手は聴覚末梢系で得られる音響的パターンの照合によって音声知覚を行うと考えられている。この仮説では、発話された音声を、抽象的で離散的な言語符号が生成過程において言語規則や調音を経て変換された結果と見なす。聞き手は、生成過程におけるこれらの規則や変換と同じものを使って、音響的パターンを合成し、この合成パターンを聴取した音声の音響的パターンと照合する。両者のパターンが一致すれば照合が完了するが、一致しない場合には、誤差に基づいて修正された合成パターンが作成され、再び照合が行われる。このアプローチは、川人 (1996) による音声産出過程の計算論的モデルなどにも取り込まれている。

音声知覚過程と音声産出過程は相互依存적であるという考え方には、いくつかの反論も提出されている。特に、産出と知覚において処理の違いがあることから、両者に共通した言語処理過程が使われているに疑問が出されることも多い。しかし、これらの点に関しては、たとえ共通した言語知識を仮定しても、そこにある種の経済性が関係しているならば、非対称性が生じ得ることを理論的に説明可能である (松井 1999)。また、変換聴覚フィードバックなどの極めて明確な現象を、知覚と産出の相互作用なしに、どのように説明すべきかは全く明らかでない。そこで、本稿では、産出と知覚に共通した言語知識が使われており、それ故、両者の間に相互作用が存在するという立場を採用する。

(3) で示した phonon が内部構造として調音的性質と音響知覚的性質を同時に持つという性質は、こうした産出過程と知覚過程における知識の同一性、相互依存

性の1つのモデル化である。これらの内部構造は、産出時には phonon が調音的性質に翻訳され、運動処理機構に対する調音指令になる。一方、音声知覚場面においては、一般的聴覚機構で抽出された音響的特徴が、phonon の内部構造にしたがって翻訳され、音韻表象としてに統合される。この流れは、一般的聴覚理論である *auditory scene analysis* (Bregman 1990) の「分析と統合の過程」という枠組みにも合致している。換言するならば、phonon は、言語機構における音声・音韻の基本単位であると同時に、言語機構と運動系および言語機構と知覚系とのインターフェースにもなるということである。

### 3. phonon の定性的内部構造

#### 3.1 phonon と弁別素性との関係

次に議論すべきことは、実際に phonon はどのような内部構造を持つのかという問題である。前述したように、音声の場合は定性的な性質と定量的な性質の両面を考える必要がある。本節では、まず phonon とその定性的内部構造である弁別素性との関係を見てみよう。

まず、phonon と弁別素性の関係を、以下のように定義する。

- (10) a. 分節音 (segment) はいくつかの phonon の集合であり、head となる phonon と dependent となる phonon がある。<sup>4</sup>
- b. 各 phonon に内在する弁別素性 (DF) は、hot DF と cold DF の 2 種類があり、各 DF は segment に引き継がれる。
- c. ただし、引き継がれる DF の強さは以下の通りである。
1. head 要素の DF を segment に引き継ぐと [cost -E] が掛かる。
  2. hot DF を segment に引き継ぐと [cost -D] が掛かる。

この規則から、head となる phonon の hot DF が最も強く、次に dependent である phonon の hot DF, head 要素の cold DF という順で続き、最も弱いものが dependent となる phonon の cold DF という序列になることが分かる。例として日本語の [e] 音を取り上げ、(10) の具体的な働きを見てみよう。すでに見たように、日本語の

<sup>4</sup>head となる phonon は、(7) などの素性構造において、head 素性として示されている。簡易表記では下線の引かれている phonon が head 要素である。

[e] は phonon の表示で (7) のような構造になり、head となる phonon として I 要素を、dependent 要素として A 要素を持っている。ここで、各 phonon の持つ弁別素性と (9) の計算規則により、日本語の [e] は音声知覚における一般的聴覚過程とのインターフェースにおいて、次のような弁別素性とマッチすることになる。<sup>5</sup>

(11) phonon と定性的音響特性との関係

<i>dependent</i>	<i>head</i>		<i>segment</i>
[A,	I]	↔	/e/
- <b>DIFFUSE</b>	+diffuse		- <b>DIFFUSE</b>
+grave	- <b>GRAVE</b>		- <b>GRAVE</b>
-flat	-flat		-flat
-creak	+ <b>CREAK</b>		+ <b>CREAK</b>
+cont	+cont		+cont
-strid	-strid		-strid
-murmur	-murmur		-murmur
+voiced	+voiced		+voiced

同様に、音声産出過程とのインターフェースにおいては、phonon の表示から以下のような調音的素性が展開され、運動指令の離散的表現となる。

(12) phonon と定性的調音特性との関係

<i>dependent</i>	<i>head</i>		<i>segment</i>
[A,	I]	↔	/e/
- <b>HIGH</b>	+high		- <b>HIGH</b>
+low	-low		-low
+back	- <b>BACK</b>		- <b>BACK</b>
-labial	-labial		-labial
-paratal	+ <b>palatal</b>		+ <b>palatal</b>
+sonorant	-sonorant		-sonorant
-shrill	-shrill		-shrill
-nasal	-nasal		-nasal
+slack	+slack		+slack

<sup>5</sup>hot DF を大文字ゴシック体で表示する。

### 3.2 産出と知覚の非対称性

調音的弁別素性と音響知覚的弁別素性を phonon の内部に取り込んでしまうというアプローチを採用する理由は、不完全指定の自然な表現および表示の経済性を達成するという目的ばかりではない。言語認知機構として見た場合、調音的特徴と音響知覚的特徴が個別に計算されていると考えるべき現象が存在するのである。その一例として、日本語を母語とする幼児の *r* 音獲得の過程を見てみよう。音声獲得中の幼児は、*r* 音の調音に関して困難が見られることがあり、しばしば *d* 音で発音する (風間・阿部 1997)。しかし、こうした幼児であっても、*r* 音の知覚は正確にできることが知られている。例えば、「リング」という語を [dingo] と発音する幼児が、親が「ディング」というと、その発音の誤りを指摘できる。また、この時期の幼児は、既に *d* 音の変異音である [dʒ(i)] あるいは [ʒ(i)] を正確に発音できるにも関わらず、*r*(i) 音については [d(i)] 音を使い、[dʒ(i)] あるいは [ʒ(i)] を使用することはほとんどないことが観察されている。さらに興味深いことに、*r* 音の *d* 音エラーは、発達と共に *r* 音に修正されていくが、この時「真」の *d* 音が *r* 音に過剰修正されることはほとんどない。このことから、この時期の幼児は、内的には成人と同様に *r* 音と *d* 音の区別がなされており、既に正しい表象が獲得されていると考えられる。産出のエラーは、表象自体が原因なのではなく、その表象を何らかの形で「解釈」する際のエラーなのである。

こうした音声産出と音声知覚の間に存在する非対称性は、音響的弁別素性の値が調音的素性から予測されている (あるいはその逆) というアプローチでは、説明することが困難である。しかし phonon のアプローチでは、この現象を以下のように捉えることができる。

今、正常な *r* 音と *d* 音の違いは、調音的には [continuant] 素性によって、音響的には [sonorant] 素性によって示されるとする<sup>6</sup>。すなわち、*r* 音は [+cont, +son] の素性を持ち、*d* 音は [-cont, -son] という素性を持つ。また、*r* 音の表象は [q, R], *d* 音の表象は [V, R, q] であり、R 要素を持つ [+Cont, +Son] 素性は hot DF, q 要素を持つ [-Cont, -Son] 素性も hot DF になっていることが最終的な状態であるとする。ここで、成人は内的に正しい表象を獲得していると考えられるため、phonon の集合全体は、*r* 音については (13) のように、*d* 音に関しては (14) のように解釈

<sup>6</sup>その他、[coronal], [stiff] を調音的 DF、[sonorant], [grave], [voiced] を音響知覚的 DF として用いている。

される。表から分る通り、おとなの *r* 音と *d* 音は、調音的には [cont] 素性において違いがあり、音響的には [sonorant] 素性において差異が存在する。

(13) 成人の *r* 音における定性的調音特性および音響特性

<i>dependent</i>	<i>head</i>		<i>segment</i>
[q,	<b>R]</b>	↔	<i>r</i> ([r],/r/)
-Cont	<b>+Cont</b>		<b>+Cont</b>
	<b>+Coronal</b>		<b>+Coronal</b>
-stiff	<b>+stiff</b>		<b>+stiff</b>
<b>-Sonorant</b>	<b>+Sonorant</b>		<b>+Sonorant</b>
	<b>-Grave</b>		<b>-Grave</b>
-voiced	<b>+voiced</b>		<b>+voiced</b>

(14) 成人の *d* 音における定性的調音特性および音響特性

<i>dependent</i>		<i>head</i>		<i>segment</i>
[v,	<b>R,</b>	<b>q]</b>	↔	<i>d</i> ([d],/d/)
	<b>+Cont</b>	<b>-Cont</b>		<b>-Cont</b>
	<b>+Coronal</b>			<b>+Coronal</b>
<b>+Stiff</b>	<b>+stiff</b>	<b>-stiff</b>		<b>+Stiff</b>
	<b>+Sonorant</b>	<b>-Sonorant</b>		<b>-Sonorant</b>
	<b>-Grave</b>			<b>-Grave</b>
<b>+Voiced</b>	<b>+voiced</b>	<b>-voiced</b>		<b>+Voiced</b>

一方、*r* 音の獲得最終段階にいるこどもは、*r* 音 (15) と *d* 音 (16) について以下のような DF を持っている。成人との違いは、**R** の調音的性質である [+cont] 素性が hot DF になっていないところにある。これは生理学的理由により、幼児の、**R** 要素の発音<sup>7</sup>が不安定であることを意味する。なお、音響的要素については、成人との間に違いはない。知覚面に関しては、生理的発達は無関係と考えられるためである。

<sup>7</sup>英語の *r* 音に相当する。

(15) 幼児の *r* 音における定性的調音特性および音響特性

<i>dependent</i>	<i>head</i>		<i>segment</i>
[q,	<u>R</u> ]	⇔	<i>r</i> ([d],/r/)
-Cont	+cont		-Cont
	+Coronal		+Coronal
-stiff	+stiff		+stiff
-Sonorant	+sonorant		-Sonorant
	-Grave		-Grave
-voiced	+voiced		+voiced

(16) 幼児の *d* 音における定性的調音特性および音響特性

<i>dependent</i>		<i>head</i>		<i>segment</i>
[V,	R,	<u>q</u> ]	⇔	<i>d</i> ([d],/d/)
	+cont	-Cont		-Cont
	+Coronal			+Coronal
+Stiff	+stiff	-stiff		+Stiff
	+sonorant	-Sonorant		-Sonorant
	-Grave			-Grave
+Voiced	+voiced	-voiced		+Voiced

この音声の獲得中期にいる幼児が、既に表象としてはおとなと同じ知識 — すなわち [q, R] と [V, R, q] を獲得している点に注意されたい。また、R 要素の持つ [cont] 素性の値である “+” 自体には違いがないことにも注目されたい。各音韻要素の持つ弁別素性の値は、固定的 (生得的) であり、おとなと子どもの間で差はない。両者の違いは、インターフェースにおける弁別素性の解釈にあり、R 要素における [+cont] という調音的素性が cold DF に解釈されてしまっている点だけである。しかし、この素性が hot DF になっていないことが、正しい知覚が可能であるにも関わらず、調音の狂いが生じる原因になっているのである。

こうした現象は、調音的弁別素性と音響的弁別素性はインターフェース・レベルで個別に解釈されるものであり、言語の表示レベルでは両者を統一した単独の表象があることを支持する 1 つの例といえるだろう。また、Piaget によると、子



どもは2歳頃から知覚と運動のつながりの間に、経験的な表象といえるものを扱えるようになるという。この表象は、初めは知覚に依存した静的なものだが、次第に知覚的経験とは独立に、内的に操作できるようになる。調音的性質と音響知覚的性質を内部構造に持つ phonon のアプローチは、こうした Piaget の考え方も矛盾しないものである。

### 3.3 プロトタイプとしての phonon

phonon の内部に弁別素性を持たせることで、phonon は家族的類似性 (2a) 以外に、もう一つプロトタイプカテゴリーとしての特徴である (2b) も満たすことができる。(2b) は、単独のプロトタイプは、それ自体で自立した存在であることを意味する。言い換えるならば、プロトタイプとプロトタイプの集合体とは、同じレベルに属するのである。

弁別素性は、離散的カテゴリーであるため、強い不完全指定理論の立場に立たない限り、こうした性質を持たない。例えば、[+coronal] という指定は、それ自体では音声にならないが、弁別素性の集合体である [-voiced, -cont, +coronal...] は [t] という音声になる。したがって、弁別素性と弁別素性の集合体は、異なったタイプに属するのである。<sup>8</sup> 一方、前節で示した phonon の内部構造から分かる通り、多くの弁別素性から構成されているため、各 phonon は基本的に単独でも「音声」として存在しうる。例えば、A は [a], I は [i] として自立した音価を持つ。<sup>9</sup> また、複合要素である [A, I] は [e] という音価に相当する。したがって、phonon とその集合体は、同じレベルに属する表象であり、phonon それ自体はプロトタイプカテゴリーといってよいだろう。

ただし、このことは、音韻における離散的カテゴリーを否定するものではない。phonon は、内部構造として離散的カテゴリーを持っており、これは「差異の体系」を保持するために必要不可欠である。また、前項で見たように、離散的カテゴリーは、他のモジュールとのインターフェースにおいて重要な役割を果たす。本稿での主張は、あくまで独立した言語認知機構の音韻基本単位を考える限り、

<sup>8</sup>なお、強い不完全指定理論の立場に立つと、例えば [+coronal] は他に条件がない限り、[t] などの音声として実現され得る。この点において、不完全指定理論における弁別素性の「タイプ」がどのような性質を持つのか、興味深い議論が可能である。

<sup>9</sup>例外は、声帯振動に関わる要素である V と v であるが、*Government Phonology* の立場では、これらは音調の high tone, low tone という音声的实现形を担うと考えている (Charette 1991, Harris 1993)。

それはプロトタイプカテゴリーであってもよいというものである。<sup>10</sup>

## 4. phonon の定量的内部構造

### 4.1 産出および知覚におけるコスト関数

次に phonon の定量的性質について議論を行う。この性質においても、定性的性質同様、産出に関するものと音響知覚に関するものが必要である。ここで、産出に関わる性質を、ある音響的性質 X を実現するための調音運動に必要な運動量として定義する。一方、知覚に関わる性質は、ある音響的性質 X を分節音 Y として同定できる確率として表現されていると考える。これらは、いずれも調音運動の困難度、あるいは知覚的同定の困難度を表す指標である。したがって、phonon の定量的性質とは一種の経済性であり、運動量が増加するほど産出における負荷が高く、また分節音同定の困難度が増すほど知覚における負荷が高くなるようなコスト関数のことであると言い換えてもよい。<sup>11</sup>

こうしたコスト関数により、phonon は定量的性質として、ある種の分布を持ったプロトタイプカテゴリーであることが保証される。これは、定性的性質における家族的類似性によって特徴づけられるされるプロトタイプの性質とも矛盾しない。さらに、この性質により、音韻のプロトタイプ性を巡る 1 つの問題点を回避することができる。

第 2 節で見たように、子音は範疇的に知覚され、母音は連続的知覚になることが実験的に確認されている。もし、音韻の基本単位である phonon がプロトタイプカテゴリーであるとするならば、離散カテゴリーの証拠と考えられる子音の範疇的知覚の要因をなんらかの形で説明しないとイケない。ここで、再び (7) と (8) に注目されたい。両者の違いは、(7) が同じ上位カテゴリー (**manner** 素性) に属する phonon の集合であるのに対し、(8) は異なった上位カテゴリーに属する phonon の集合体として表現されている点にある。同一の上位カテゴリーに属するものは、分布の重複が大きい (これが定性的な家族的類似性を保証する) のに対し、異なっ

<sup>10</sup>また、プロトタイプカテゴリーは認知言語学・認知意味論の分野で主に議論されているが (Lakoff 1987, Taylor 1995, 河上 1996, 山梨 1995)、生成文法の枠組みと矛盾するというわけでもない。HPSG などの生成文法の枠組みで、プロトタイプカテゴリーがどのような効果をもたらすのかという点は、本稿の隠れた目的の 1 つである。他の隠れた目的は、phonology と phonetics の自然な接続である。

<sup>11</sup>音韻要素と物理量との関係については、Harris and Lindsey (1990) の議論がある。彼らのアプローチは、音響的性質を直接扱うもので、産出・知覚の相互作用は考慮されていない。

たカテゴリーに属する要素同士は、共通した性質が少ないため、分布の重なりは狭くなる。したがって、同一カテゴリーに属する phonon の集合が作り出す分布の分散に比べ、異なったカテゴリーに属する phonon の集合によって作られる分布の分散は極めて狭くなり、結果的に全体としてのカテゴリーの独立性が強くなる。これが、子音知覚の範疇的性質の要因と考えることができる。

#### 4.2 最適解：産出／知覚の安定点

次に、phonon に内在する定量的性質を具体的に見てみよう。例として、A 要素に内在するコスト関数を取り上げる。phonon A は、(11), (12) に示される定性的性質から分かる通り、単独では低舌母音 [a] に相当する性質を持つ。舌の高さは、舌体の動きを固定した場合、顎の開閉に伴う回転運動にほぼ一致し、その運動量の変化は、口腔の運動モデル (Mermelstein 1973, McGowan 1994) から推定できる。今ここで、最も自然な母音の発音位置として、中舌母音 ə (第 1 フォルマントの値でほぼ 500Hz) を基準に定め、ə の位置からの運動量を産出におけるコストとする。この時、A に内在する産出的性質は、Figure 1 に示すようなコスト関数として表される。<sup>12</sup>

<sup>12</sup>実際にはこれらの数値は相対的に変動し得る。なぜなら、絶対的基準においては、ある話しての母音 X のフォルマント振動数は、別の話しての母音 Y のフォルマント周波数に対応することが頻繁に起こりうるからである。

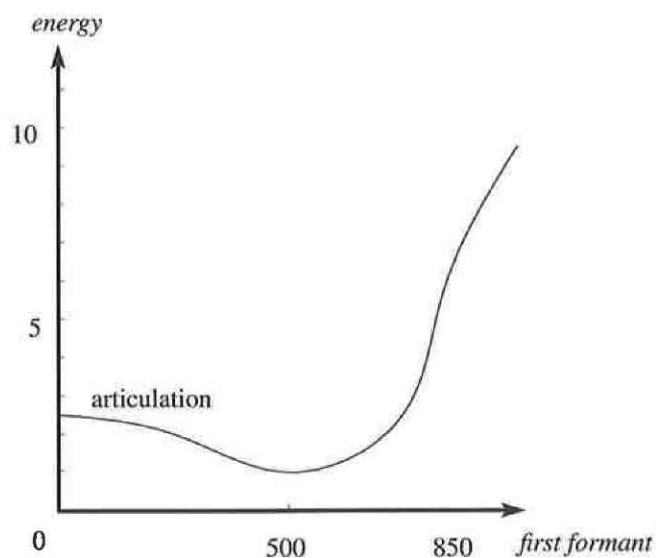


図 1: A に内在する産出的コスト

一方、母音の知覚には第 1 フォルマントおよび第 2 フォルマントが重要になるが、中でも第 1 フォルマントは舌の高さと相関を持つことが知られている。したがって、低舌母音 [a] においては、第 1 フォルマントの値が特に重要な手がかりになる。したがって、第 1 フォルマントの変化 (第 2 フォルマントは第 1 フォルマントに従属して変化) に伴う [a] の同定率の変化が、A の持つ音響知覚的性質であると考えて良い。この時、A に内在する知覚的性質は、Figure 2 に示すようなコスト関数として表される。

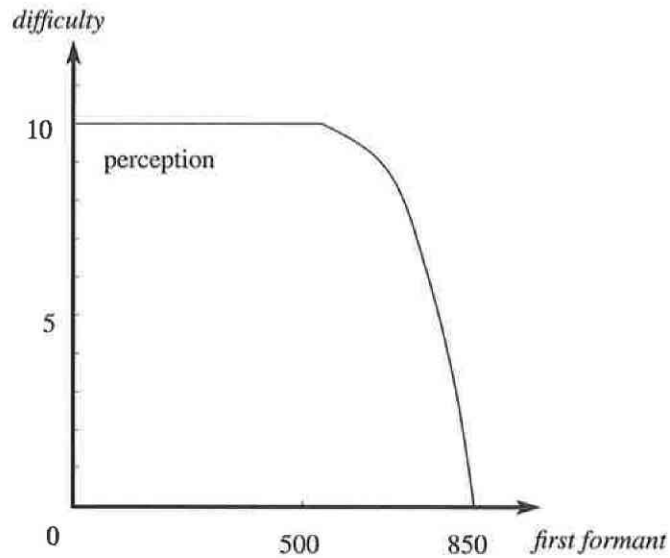


図 2: A に内在する知覚的成本

ここで、[a] の理想的な実現値を得たい場合、それは調音運動としても、音響知覚処理としても、最も安定する点になると考えられる。なぜなら、調音運動に掛かるコストが極めて低いものであっても、それによって得られる音響特性が、知覚的に困難度の高いものであるならば、実現値としては理想的なものとはいえないであろうし、また容易に知覚できる音響特性であっても、それを実現する生理運動が困難なものであるなら、やはり実現値としては理想的とはいえないからである。したがって、A の定量的最適値は、産出のコスト関数と知覚のコスト関数を組み合わせた交点になるといってよいだろう。以後、こうした産出コスト関数と知覚コスト関数の相関モデルを、運動知覚モデルと呼ぶ。A の標準状態における運動知覚モデルは、Figure 3 のようになる。

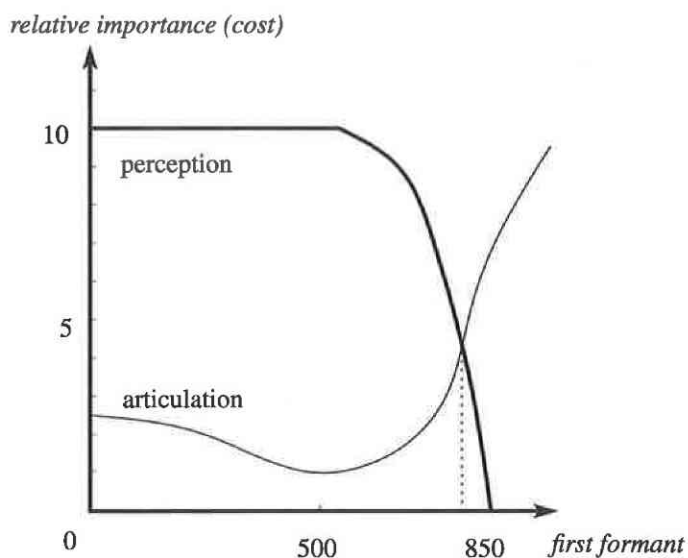


図 3: 標準状態での最適コスト

ここで重要なことは、運動知覚モデルにおける産出コスト関数と知覚コスト関数との相対的な比率が、言語環境などの状況によって比率が変わりうるという点である。例えば、前後の音環境から極めて大きな調音運動を必要とする場合や、語用論的理由によって発話スピードを速める必要がある場合には、知覚面での重要度は標準状態と変わらないが、産出面での相対的成本は増加すると考えられる。一方、英語のストレスの掛からないシラブルなどでは、調音運動のコストは標準状態と変わらないが、知覚の相対的成本は減少するといっていよう。これらの違いを運動知覚モデルで表すと、次のようになる。まず、産出コストと知覚コストのバランスが取れている標準状態の運動知覚モデル Figure 3では、第1フォルマントの値が、ほぼ 820 Hz となる。これに対し、Figure 4 に示すグラフは、早口ことばなど、極めて複雑な調音運動が必要になり、産出コストの相対的重要度が増加した場合の運動知覚モデルである。この時の第1フォルマントの値は、約 770Hz であり、標準状態に比べ、第1フォルマントの値が減少していることが分かる。これは調音点でいうならば、舌の位置が /a/ の標準位置に比べ、よ

り高くなっていることを意味しており、母音が多少訛った状態ということができる。一方、Figure 5は、知覚コストの相対的重要度が極めて低くなった場合の運動知覚モデルであり、第1フォルマントの値は、約650Hzにまで落ち込んでいる。この場合の第1フォルマントの落ち込み方は、産出の相対的重要度が増加した場合よりもはるかに劇的であり、[a]というよりは、むしろ[a]の音に近づいていることが分かる。これは英語の非ストレス母音のə化、およびそうした弱化母音のəの変異が広範囲に渡るといふ現象をうまく説明する。

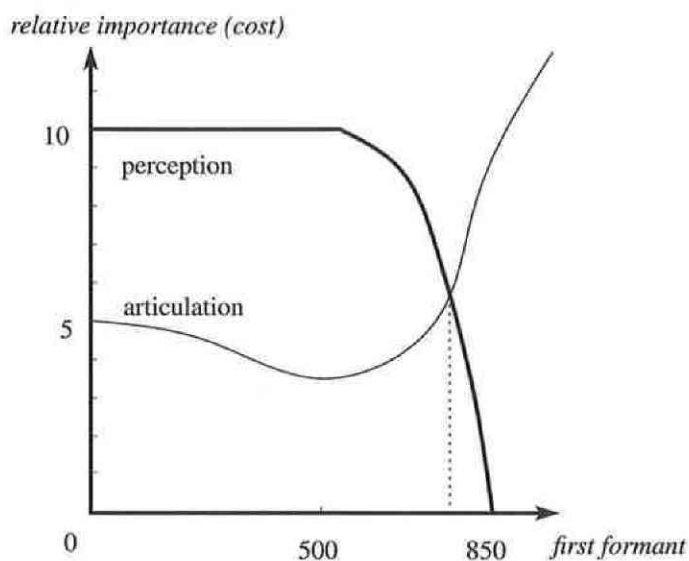


図4: 調音的理由による声のなまり

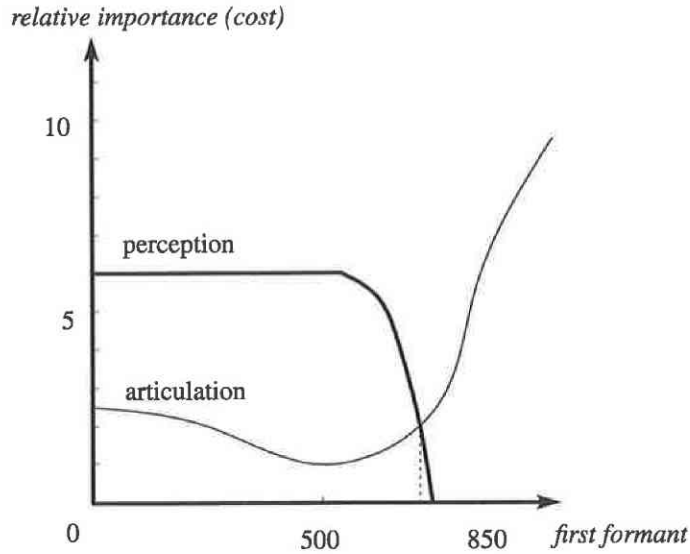


図 5: 非ストレス母音の schwa 化

同様の運動知覚モデルは、子音のような複数の phonon が組合わさった場合にも同様に構築できる。例として、前節でも取り上げた *r* 音: [q, R] と *d* 音: [V, R, q] の関係を見てみよう。ここでは特に状況を単純化して、R 要素と q 要素の複合的な性質のみを考えることにする。この時、*d* 音と *r* 音は phonon の headedness の違いによって表現されることになり、音響的には VOT とフォルマント遷移によって決定される sonorancy の違いに反映される。また、産出コスト関数の違いはほとんど無視できるものになるため、*d* 音の産出コスト関数を代表として用いる。この時、*d* 音の運動知覚モデルのコスト関数は、Figure 6 のようになり、*d* 音の運動知覚モデルは、Figure 7 のように表現される。また、横軸は voice onset time (VOT) とフォルマント遷移との複合タイミングを表し、sonorancy の尺度で考えるなら、軸の右側にいくほど sonorancy が下がることを示す。両者のグラフを比較すると、標準状態においては、*d* 音の sonorancy は、*r* 音に比べ明らかに低くなっていることがわかる。



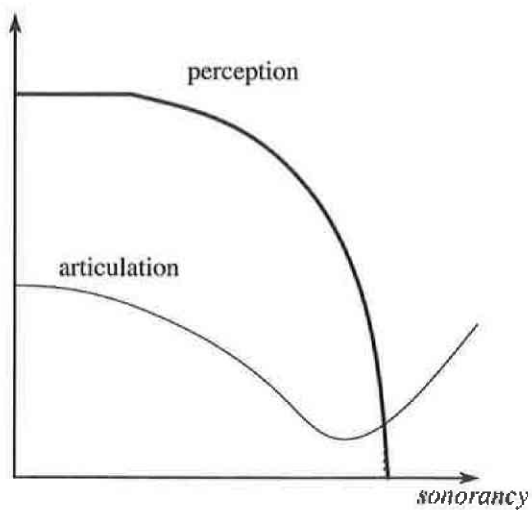


図 6: d 音の運動知覚モデル

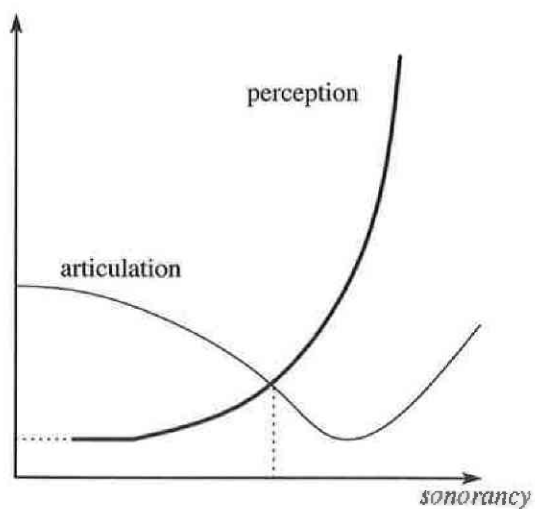


図 7: r 音の運動知覚モデル

次に、前節で見た幼児の *r* 発音における歪みについて考えてみよう。すでに述べたように、幼児は、*r* 音の表象は獲得されており、それゆえ、知覚においては大人とほぼ同様のパフォーマンスが可能である。すなわち、知覚コスト関数は標準状態と同一と見なしてよい。しかし、産出に関しては、[+ cont] の性質が hot になっていないため、調音の歪みが起こる。このようなある定性的性質が hot になっていないということは、その性質が activate されにくいということであり、経済性の観点からいえば、よりコストが掛かることを意味する (Smolensky 1993)。したがって、幼児の *d* 音発話を運動知覚モデルに当てはめるならば、前述した早口ことばでの発話と同じく、知覚コスト関数は標準状態と変わらないが、産出コスト関数の割合が相対的に高くなっている状態に相当する。この時、運動知覚モデルは Figure 8 のようになり、*r* 音の安定解が標準状態の *d* 音の解に近づいていることがわかる。また、*d* 音の安定解も標準状態の *d* 音の値よりやや小さくなっており、幼児の発音の不明瞭性を示す結果となっている。

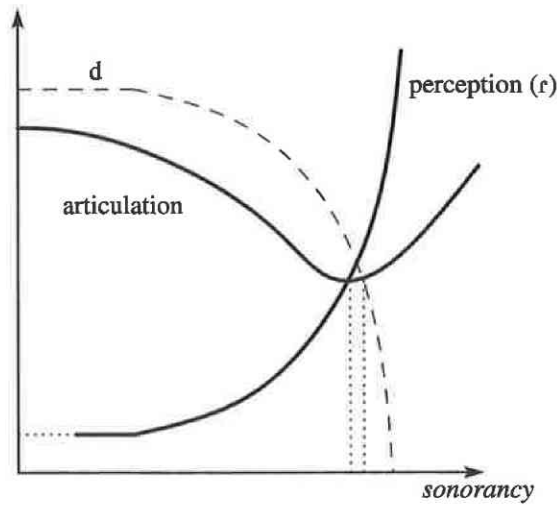


図 8: 幼児の *r* 音の歪み

最後に非ストレス母音の弱化と似たような現象を見てみよう。子音においては、例えばアメリカ英語で観察される、母音間に挟まれた *d* 音の弾音化現象がこれに相当するものと思われる。南條 (1996) は、この異音過程を、より望ましい音

節構造を作り出すために、音節末子音がより母音的になるという一種の母音化 (vocalization) であると見なしている。音節の中心となれる音は、一般に母音や接近音などの sonority の高い音であるが、これは sonority の高い音は「知覚的」に目立ちやすく、リズムの中心を担いやすいという理由による。換言するならば、sonority の高い音は、知覚しやすく、知覚に必要なコストが真子音にかかるコストよりも小さくて良い。そこで、d 音の運動知覚モデルにおいて、知覚コスト関数が相対的に減少したとすると、Figure 9 のようなグラフが得られる。この時、弱化した d 音の sonorancy が、「標準状態」における r 音の sonorancy とほぼ一致しており、d 音の弾音化が実現されることになる。

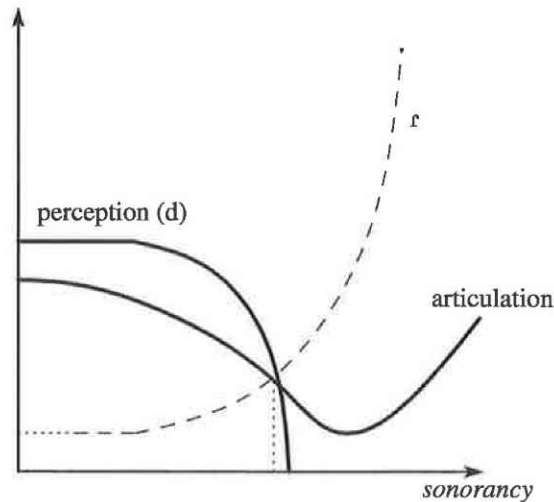


図 9: 共鳴音間の d 音の変異

このように、phonon を用いるアプローチでは、音の変異を定量的数値として導出することができ、かつ弁別素性によって表現される定性的性質と矛盾なく対応させることが可能である。しかし、音韻表象に定量的性質を与えた場合、解決しなければならない2つの問題点が存在する。1つは、運動知覚モデルの安定解は何を意味する値なのかという問題であり、残りの問題は、連続的なコストの計算の行われる処理単位の問題である。実のところ、これらの問題は自明なものではなく、むしろ理論的に定義すべき問題と思われる。

### 4.3 定量的性質を巡る問題

まず、運動知覚モデルの安定解の問題について考えてみよう。最も単純な解釈は、安定解を、「実際に実現される」音響的な値であり、また「実際に知覚されるべき」音響的数値として解釈することである。確かに、コスト関数の与え方によっては、こうした解を導出するよう定義することも可能である。しかし、我々はこのアプローチを採用するべきではないように思われる。実際に実現される数値は、非言語的な要因、例えば単なる話し手のくせや、再現性のない調音運動のゆがみなどの影響を受けた結果である。こうした偶発的な要因まで言語モジュールが計算すべきことではない。むしろ、言語モジュールで処理されるべき情報は、何らかの形で「言語としての意味機能」を担う情報であり、それ以外の情報は他のモジュールに任せてしまうほうが合理的な設計であろう。

そこで、本稿では運動知覚モデルの安定解を、外部モジュールにおける神経指令の「目標値(target value)」であると考えてみよう。この「目標値」というものを設定するには根拠がある。例えば、詳細な音響的計量により、発話音の音価は明瞭なゆっくりした発話での音価とは時に著しく異なっていることがわかっている。発話がうち解けたものであるか、あるいは発話速度が速いものであるほど、理想的な音価とはかけ離れたものとなる。これは、理想的な環境ならば目標値が実現される可能性があるが、状況により理想状態が保てない環境においては、より小さな値となったり、あるいは目標値を越えてしまう場合もあり得ることを意味する。また、こうした逸脱した値についても、人間の認知機構は正確な音韻知覚が可能であり、これは、言語認知機構が実際の値から理想値を復元していることを示唆している。

Figure 10 に示した図は、標準状態において目標値を実現した場合と、極めて早口で話した場合などに見られるように、目標値に到達せず、「音声のなまり」が生じた場合とを模擬的に示したものである。こうした歪みは、発話場面で典型的に観察されるが、音声知覚場面においても起こりうる。例えば、相川・津崎・河原(1995)による聴覚系の基本周波数追跡モデルでは、急速に変化する音に対して、聴覚系で処理される値と実際の物理的な数値の間にズレが生じること、また場合によっては、実際の数値よりも高い値を検出する場合があることを示している。したがって、我々の立場では、同一の phonon 表示であっても、実際の音声は、2つの要因で音響的変異が起こりうることになる。1つは、phonon の定量的性質で

ある運動知覚モデルによって得られる目標値の変動であり、もう1つの要因は言語外モジュールにおける歪みである。

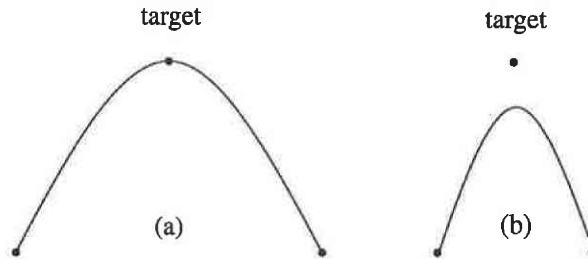


図 10: 単位処理時間と目標値とのずれ (a) 通常の処理時間 (b) 短い処理時間

残る問題は、運動知覚モデルの最適値が計算される範囲についてである。これは音韻理論がどの範囲までの言語現象を扱うのかという問題とも絡む。例えば、/ti/ が [tʃ] として現れる拘束変異から、鼻子音に後続する母音が鼻母音化するような自由変異までの全ての文脈依存性を、phonon による記号レベルで計算するのであれば、運動知覚モデルは分節音ごとに計算すればよい。一方、工学的な音声認識のように、文脈依存性を全て定量的なレベルで計算するのであれば、記号レベルでの文脈依存性はなくなり、CVC および VCV 連鎖の範囲で、運動知覚モデルを構築することになる。

この問題に対する明確な解答は今のところ存在しないように思われる。そこで、我々は音声処理の心理的単位に関する実験結果を手がかりに、暫定的な答をだしておくことにしよう。Otake et al (1993) は、日本語音声の知覚単位を実験的に検証し、その単位がほぼモーラであるという結論を得ている(一部の現象ではシラブルを示唆する結果が得られている)。また、心的辞書へのアクセス単位を調べた天野 (1997) の研究でも、やはりモーラ単位でのアクセスが行われるという結果が示されている。これらの結果は、心的な音声言語処理の単位は、分節音や分節音の連鎖のレベルではなく、モーラあるいはそれ以上の構造化された単位であることを示唆している。このことから、本稿でも運動知覚モデルの計算を行う単位を暫定的にモーラとしておく。この仮定を置いた場合、極端に定量的計算に偏ることもなく、また全てを記号的計算に頼る必要もない。モーラ単位はコスト計算の

レベルとしては極めて妥当な情報量があり、拘束変異のような頑強な言語現象は記号レベルで計算を行い、自由変異のような外的要因の強い現象に関しては、運動知覚モデルで計算を行くことが可能になる。

## 5. 総合論議

本稿では、音韻の心的表象がどのような性質を持つべきかについて議論を行った。我々の得た結論は、心的表象の基本単位である phonon は、定性的および定量的内部構造を持ったプロトタイプカテゴリーであることが望ましいというものである。この内部構造を持つという性質により、phonon は基本単位でありながら、「最小要素」還元性のもたらず欠点から逃れうる。また、この内部構造の定性的性質が二項対立的弁別素性であることから、このプロトタイプカテゴリーは一定の規則 (10) により、離散的カテゴリーに翻訳できるという特徴も持つ。同時に、この定性的性質は phonon 間の家族的類似性およびカテゴリー単独での自立性をも保証する。すなわち、phonon は離散的カテゴリーを取り込んだプロトタイプカテゴリーといえる。

phonon の内部構造におけるもう 1 つの重要な点は、調音的性質および音響的性質を同時に表現している所にある。この性質は、表示の経済性に優れるという理論的特性のみならず、音声産出と音声知覚の相互作用を自然に表現するためにも必要なものである。これにより、定性的性質である調音的弁別素性と音響的弁別素性は、互いに関連し合いながらも、phonon の集合体である segment レベルにおいて、異なった値を持ちうるということが説明される。一方、定量的性質である運動知覚モデルにおいては、調音と知覚の相互作用がさらに本質的な役割を果たす。phonon の定量的性質は、ある発話状況における神経指令の具体的な目標値を導出するが、この値は調音コスト関数と知覚コスト関数の安定解として得られるものである。また、このアプローチは、従来の音韻論が扱わなかった音声現象を説明することもできる。代表的な現象は、ある種の発話状況における音声の歪みや、自由変異の分布などで、これらは調音コスト関数と知覚コスト関数の相対的重要度の変化として捉えられる。このように、phonon のアプローチでは、産出と知覚の相互作用は、本質的な性質としてモデル化されており、より多くの現象を説明できる枠組みと思われる。

次に、phonon と上位カテゴリーの関係についてまとめておく。前述したよう

に、phonon 同士は完全に独立した基本単位ではなく、互いに家族的類似性を持つ。このことは、ある phonon 同士が、同一の上位カテゴリー、すなわち同一タイプに属するトークンであることを意味する。こうしたタイプ継承の問題は、様々な方法で形式化可能であるが、本稿では phonon を HPSG の素性構造の中に組み込むことにより、上位カテゴリーの明示化が可能であることを見た。この方法のメリットは、母音は同一カテゴリーに属する phonon の集合体として表現されるのに対し、子音は異なったカテゴリーに属する phonon の集合体として表されるようになる点にある。これにより、母音の分布よりも子音の分布の境界が極めて狭くなるという定量的性質がもたらされる。この性質は、母音の知覚が連続的知覚様式を持つのに対し、子音は範疇的知覚に従うという心理現象を説明する。範疇的知覚という離散のカテゴリーに有利な現象を、分散の小さいプロトタイプカテゴリーで説明するという方法は、定性的性質としてタイプ付きの離散のカテゴリーを持ち、かつ定量的性質を合わせ持つ phonon のようなアプローチにより、初めて可能になるものである。

しかし、残された課題も多くある。まず、現在のところ、定量的性質が明確になっていない phonon がいくつか残されている。特に、知覚コスト関数に関しては、心理実験によってその分布を推測する以外に方法がないため、今後、この方面での実験が必要不可欠である。

また、(7) のような、同一タイプに属する phonon 同士が結合して新たなカテゴリーを作った場合に、認知空間の再編化がどのように行われるのかという点も検証しなければならない。この問題に関しては、ファジー理論で提案されているメンバーシップ関数の正規化 (Kay and McDaniel 1978, 藪内 1990) や Between 関数 (吉川 1992) などが重要な示唆を与えるものと思われるが、特にカテゴリー境界の扱いなどに関して問題を残しているように思われる。

さらに、本文でも述べた通り、運動知覚モデルを当てはめる言語的単位の決定についても検討の余地が残されており、それに伴って、記号レベルにおける音韻制約に関する検討も行う必要がある。また、運動知覚モデルによって得られる目標値と、調音結合モデルなどの外部モジュールとの接続についても考えなければならない。しかし、これらの点に関しては、今後の課題としたい。

## 参考文献

- 相川清明・津崎実・河原英紀 (1995). スイープ音追跡系の動特性. 日本音響学会聴覚研資料, H-95-31.
- 天野成昭 (1995). 音韻・単語知覚における相互作用の時間的側面. 東京大学大学院文学研究科博士論文.
- Anderson, J. and Durand, J. (1987). *Explorations in dependency phonology*. Dordrecht: Foris Publications.
- Bird, S. and Klein, E. (1993). Enriching HPSG phonology. *Research Paper EUCCS/RP-56*, Edinburgh university.
- Bird, S. and Ellison, M. (1994). One-level phonology: autosegmental representations and rules as finite automata. *Computational Linguistics*, 3 (20), 455–491.
- Bregman, A. (1990). *Auditory Scene Analysis*. The MIT Press, Cambridge, Mass.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159.
- Burns, E. M. and Ward, W. D. (1978). Categorical perception—phenomenon or epiphenomenon—: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, 63, 456–468.
- Charette, M. (1991). *Conditions on Phonological Government*. Cambridge University Press, Cambridge.
- Coleman, J. (1995). Declarative Lexical Phonology. in Durand, J & Katamba, F., *Frontiers of Phonology*. Longman.
- Drescher, G. L. (1991). *Made-up minds*. The MIT Press, Cambridge, Mass.
- Elman, J. L. et al. (1996). *A connectionist perspective on development*. MIT Press.
- 藤崎博也・川崎崇子 (1971). 合成音の弁別と言語音知覚機構のモデル. 日本音響学会誌, 27, 453–462.



- Getty, D. J. (1975). Discrimination of short temporal intervals: a comparison of two models. *Perception and Psychophysics*, 18, 1–8.
- Gunji, T. (1987). *Japanese Phrase Structure Grammar*. D. Reidel, Dordrecht.
- Gunji, T. and Hasida, K. (1998). *Topics in Constraint-based grammar of Japanese*. Kluwer Academic Publishers. Dordrecht and Boston.
- Harris, J. (1994). *English sound structure*. Oxford, U.K., Cambridge, Mass. Blackwell.
- Harris, J. and Lindsey, G. (1990). Segmental decomposition and the signal. 7th International Phonology meeting.
- 北条 弘 (1982). 日本語子音素の類似性。— INDSICAL と林の数量化理論第 I 類による分析 —. *The Japanese Journal of Psychology*, vol.53, No.2, 72–79.
- Jakobson, R., Fant, C., and Halle, M. (1952). *Preliminaries to Speech Analysis*. Mouton, The Hague.
- Itô, J. and Mester, R. A. (1989). Feature predicatability and underspecification : parallel prosody in Japanese mimetics. *Language*, 65, 258–293.
- Itô, J. and Mester, R. A. (1992). Feature structure and lexical organization. Paper presented at the Xth National Conference of the English Linguistic Society of Japan.
- Kay, P. and McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, 54, 610–645.
- Kaye, J. (1989). *Phonology: a cognitive view*. LEA.
- 河上誓作 (1996). 認知言語学の基礎. 研究社出版.
- 河原英紀 (1995). 音声知覚・生成相互作用の伝達特性について. 日本音響学会聴覚研資料, H-95-35.
- 川人光男 (1996). 脳の計算理論. 産業図書.

- 窪藺晴夫 (1994). 日本語の音節量について. *国語学*, 178, 7–17.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press. (池上・河上ほか (訳) 認知意味論：言語からみた人間の心. 紀伊國屋書店)
- Lee, B. S. (1950). Effects of delayed speech feedback. *Journal of Acoustic Society of America*, 22 (6), pp. 824–826.
- Liberman, A. M. and Harris, K. S. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal Experimental Psychology*, 54, pp. 358–368.
- Liberman, A. M. (1957). Some results of reserch on speech perception. *Journal of the Acoustical Society of America*, 29, pp. 117–123.
- Liberman, A. M. and Cooper, F. S. (1962). A motor theory of speech perception. *Proceedings of Speech Communication Seminar*, paper-D3, Stockholm.
- 松井理直 (1999). 協調行為として見た音韻論の経済性. 杉藤美代子 (編), *文法と音声 II*, くろしお出版. pp.109–126.
- Mermelstein, P. (1973). Articulatory model for the study of speech production. *Journal of Acoustical Society of America*, 53, 1073–1082.
- McGowan, R. S. (1994). Recovering articulatory movement from formant frequency trajectories using task dynamics and a genetic alogrithm: Preliminary model tests. *Speech Communication*, 14, 19–48.
- 南條健助 (1996). 英語の異音過程. 音韻論研究会 (編), *音韻研究 — 理論と実践—*, 開拓社. pp.117–120.
- Otake, T., et al. (1993). Listener's representaions of within-word structure: A Cross-linguistic and corross dialectual investigation. *Proceedings of the 4th European Conference on Speech Communication and Technology*, 3, pp. 1703–1706.

- Otake, T. and Cutler, A. eds. (1996). *Phonological Structure and Japanese Processing: Cross-linguistic studies*. Mouton de Gruyter.
- Pollard, C. J. and Sag, I. A. (1987). *Information-based Syntax and Semantics, Vol. 1: Fundamentals*. CSLI Lecture Notes Series. No.13. Center for the Study of Language and Information, Stanford University, Stanford.
- Pollard, C. J. and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago.
- Prince, A. and P. Smolensky. (1993). *Optimality Theory : Constraint Interaction in Generative Grammar*. Technical report 2, Center for Cognitive Science, Rutgers University.
- Saussure, F. de. (1916). *Course in General Linguistics*. New York: McGraw.
- Smolensky, P. (1993). Harmony, Markedness, and Phonological Activity. *Rutgers Optimality Workshop 1 (ROA87)*
- Stevens, K. N. (1960). Toward a model for speech recognition. *Journal of the Acoustical Society of America*, 32, 47–55.
- Taylor, J. (1995). *Linguistic Categorization: Prototypes in Linguistic Theory*. (2nd edition). Oxford: Clarendon Press.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Basil Blackwell. (藤本・坂井 (訳). 哲学探究. 法政大学出版会)
- 藪内稔 (1990). 色彩と範疇化: ファジィ集合論的アプローチ. *日本ファジィ学会誌*, 2, 456–469.
- Yamada, R. A. et al. (1995). The effect of perception training on Japanese speakers' production of the American English /r/-/l/ contrast. *日本音響学会春期発表講演論文集*, pp.386–386.
- 山梨正明 (1995). *認知文法論*. ひつじ書房.

吉川歩 (1992). *Between* 集合の数学的性質. 日本ファジィ学会誌, 4, 150–159.

**Author's E-mail Address:** [matsui@sils.shoin.ac.jp](mailto:matsui@sils.shoin.ac.jp)

**Author's web site:** <http://sils.shoin.ac.jp/~matsui/>