

## 生物言語学からみた音韻論研究の展望

三 宅 正 隆

### 1. はじめに

生成文法理論が母語話者の生得的言語能力 (faculty of language) の解明を目標として掲げ、特にアメリカ構造主義の行き詰まりを打破するとともに、新しい息吹を言語学会に吹き込んで以来ほぼ半世紀になろうとしているが、この目標自体は一貫して変わることがない。しかしこの間理論的な枠組みにはずいぶんと変化があった。この理論の象徴でもあった「変形」という冠が「生成文法理論」という呼び名から取り外されてしまったのが何よりもこの間の理論的変遷の激しさを物語っている。しかしながら、極小主義などに代表される最近の言語研究の潮流は、さまざまな言語現象分析に基づく理論構築という伝統的な枠組みを越え、遺伝子学や脳神経学など一般生物学的視点からの生得的言語機能の解明に焦点が移行し、これによって生成文法研究自体の風景が一変したように映る。

「生物言語学」という名が示すように、先端的な理論の主要な関心はコトバの「発生」「進化」の観点から言語能力を再検討することに向けられているが、この傾向は結果的には次第に「言語学」プロパーな特性の縮小と「インターフェース条件」などへの重点のシフトとなって現れている。これはある意味で、従来の「理想的な母語話者の言語能力の解明」を目指す中で重視されていた言語現象の「理想化」を置き去りにする傾向を生み出している。以下極小主義の中でこのような従来型の研究がどのように組み込めるのか、またこれからの研究方向の可能性について、特に音韻論研究に焦点を絞って考えてみたい。

### 2. 極小主義の登場

当初生成 (変形) 文法理論で強調されたのは、「説明的妥当性」を満たす普遍文法を明示的にすることであったが、これは方法論的にはあくまで個別言語の分析を通して経験的に達成するしかない目標でもあった。普遍文法と説明的妥当性を満たす個別言語の記述はとりもなおさず母語話者の持つ言語直感の説明にあたると考えられたが、この言語直感としては、運用面で

の言語現象や歴史的な変化などは直接的な考察や議論をサポートする根拠から省かれ、あくまで言語内の直感、資料が分析対象とされた。そこでは、言語のみにあてはまる特有の体系や、表示、規約、条件を見いだすことに力点が置かれた。それがとりもなおさず「言語能力」の定義であり、この定義は言語習得過程を考慮に入れれば、できるだけ可能な文法を狭く規定できる理論であるのが望ましく、これは厳密な枠組みや形式の重視という態度に表れた。しかしながら研究がすすむにつれ、個別言語の分析記述と普遍文法の構築を平行して進める方法論は、結局「記述的妥当性」と「説明的妥当性」の衝突を生み、原理的にその解消が容易でないことが次第に明らかになった。1980年代になりGB理論などの原理とパラメータ理論が提案され、慢性的な衝突問題が解消し、新たな視点から普遍性と多様性を説明しうる原理として歓迎されたが、同時にこの新たな展開によって従来からの研究の焦点もかなり変貌をとげることになる。それは「母語話者の持つ言語能力」は他の領域とは独立した、種固有のモジュールを形成するという仮定に基づく言語固有の特性、原理研究からの離脱を意味した。この傾向は極小主義 (Minimalist Program) が提案されるにいたっていっそう加速することになる。

極小主義以前の生成理論では、特に統語論の領域固有性が重視され、句構造規則や変形規則に始まり、Xバー理論に至るまで、一貫して言語能力を定義するものとして範疇や規則、原理、制約などの普遍的特性の追求にもっぱら精力が費やされた。極小主義になるとこれらの装置はほとんど捨てられ、統語論の核でもあった構造構築装置としては、二つの要素を結合する「併合」(merge)のみが残ることになる<sup>1)</sup>。結局言語表現の統語的な構造は、厳密に定義された規則群を適用することで生成されるのではなく、さらに抽象的かつ単純な併合の繰り返し適用で順次より大きな列が生成される原理に集約される。この生成を推進する装置としては、語彙項目を構成する素性間の一致関係のみが実質的に重要となるが、二項間の関係だけで順次により大きな単位に変化していくという構図は、何か生物の発生や成長を連想させ、新分野としての生物言語学領域呼称として象徴的である。

このように極小主義では言語固有の生得的な言語能力から不必要な贅肉がぎりぎりまでそぎ落とされる一方で、出来上がった語列を解釈する運用系の制約が全体のキーになるという、ある意味で逆転現象が生じ、この意味でも伝統的な生成文法研究とは袂を分つことになる。しかしながら60年代から生成理論によって繰り広げられた言語分析や議論は、統語論のみならず音韻論や意味論などの領域内部で独自の理論を発展させ、おびただしい成果をもたらした。原理やパラメータは先に述べた衝突を解消する原理ではあるが、現実には具体的な現象の衝突をほとんど解決していない。個別現象の研究が進むにつれて、それぞれの現象の「個性」が目立つ一方で、それをまとめる普遍的な規則性が見つかりにくくなり、結局パラメータの値の取り方としてこの個性を処理することで決着がつけられたが、これは原理としては機能しても現実にパラメータの値が設定された後どのように個別の現象が説明されるのかは今のところ不明

で、結局そこから先は従来の研究成果がものを言うことになるはずである。現在の極小主義をめぐる議論では、このような今まで蓄積された個別言語をめぐる経験的分析知見はいわば「化石」と化してしまった感がある。しかし「化石」は、研究者の分析手法次第でさまざまな新事実を物語ってくれるはずである。

このように「原理」の問題に焦点が集中する中、極小主義は言語能力を実現するいわば周辺領域の機能を生かす条件を言語能力の満たすべき要件としてあげる。従ってこの「周辺問題」、つまりインターフェース条件が一躍クローズアップされている。しかしながらこのインターフェースも最近ではヒト固有の能力を直接掘り下げるといふより、進化の問題として、生物の中のヒト問題として位置づけられる傾向にある。このような視点で焦点となるのは、脳と形態の進化問題である。言語の運用面については、特に調音/知覚問題として音韻論、音声学での長い研究歴史があり、形態上の特徴と音響、知覚の研究が言語能力と一体化すれば、脳からみた言語探求の視点も可能になる。

### 3. 生物言語学

極小主義は研究の対象にも大きな変化をもたらした。最近のホットな話題は言語分析ではなく言語進化の問題である。生成文法が提案された時からチョムスキーは「言語学は生物学の一部である」と明言していたが、このことは単なる「目標」としての「生得的言語能力」の研究から最近の言語学が「生物言語学」(biolinguistics)と呼ばれるようにヒト型言語の起源と進化の研究の焦点が移りつつあることを示す。長年ヒトの言語能力の固有性を強調していたチョムスキーが数年前に二人の心理学者と発表した論文では(Hauser, Chomsky, and Fitch, 2002), 今や言語学は進化生物学や人類学, 心理学, 脳神経学などの関連領域との連携なしには「進化」はあり得ないとまで言い切り、この変化はその論文の内容とともに学会の内外に波紋を投げかけている<sup>2)</sup>。この論文では著者は、「言語能力」に狭義の言語能力 (FLN, the faculty of language in the narrow sense) と広義の言語能力 (FLB, the faculty of language in the broad sense) の区別を設けている。後者にはいわゆるインターフェースに関わる能力と「再起性」の演算機構が含まれ、有限の要素から無限の表現が生成される言語の基本原理を保証するとする。これに対して、前者の狭義の言語能力に含まれるのは「再起性」という特性のみで、しかもこの部分だけが人間の言語能力の固有性であると、非常に極端で衝撃的とも言える主張をおこなっている。さらに、このFLNのみが人間固有の言語能力そのものであるにもかかわらず、進化的には言語のための適応として進化したのではなく、言語以外の理由で人間に備わった能力であると、生成理論初期の「人間特有の言語能力は習得現象を考慮すると、非常に豊かなものであるはずだ」という見解からはほど遠くなっている。「再起性」のみが人

間固有の言語能力であるという場合、可能性としては、その他の特性は人間にしかないが、言語固有の特性ではない場合と、この能力が他の種にも共通してみられる特性であるという二つの可能性が少なくとも考えられる。実際このことを論証するために、論文の大半は他の生物のコミュニケーションシステムや認知に関する能力、器官の仕組みなどとの比較、対象に割かれ、従来人間固有の言語能力であると考えられていた大部分について、「その他の特性」である可能性を例証するのに費やしている。言語の起源問題は19世紀にはあまりにも根拠のない議論がすぎたため、1865年のパリ言語学会で言葉の起源に関する論文は受け付けない決定が会則に盛り込まれたことはよく知られているが、現在と当時とは言語学内部のみならず関連分野での状況も比べ物にはならないことはもちろんであるが、経験科学としての言語学からすれば、状況はずいぶん似かよってきた印象がある。

極小主義は「プログラム」という名前が示すように、特定の言語理論を指すというよりも研究方法のありかた、言語理論の性格に関する主張でもある。その主張は同じ説明的妥当性でも従来のような「形式上の簡潔性や自然性」といったところから一歩踏み出して、生物学的な側面や言語使用といったいわば外的な要請を満たす条件が全面に押し出されることになる。生物学的な側面とは、人間の生得的な言語能力、つまり遺伝子によって発現する普遍的言語獲得装置の初期状態は他の生体器官と同じで、ある意味「完璧で、単純で、無駄のない」設計図にもとづいてできているはずという仮定であり、言語使用にかかわる要請とは、実際に意味が物理的な音声になり、逆に実際の音声の意味に解釈されることを保証する要件を満たすものでなければならないということである。運用面は運用理論に任せるという態度ではなく、もし種として言語を持つということは運用を保証するような仕組みも言語機能は満たしていなければならないということである。これがインターフェイス条件と呼ばれるもので、いわば運用に関わる言語外の条件が言語能力の妥当性条件として課せられる点、初期の生成理論から180度の転換である。

インターフェイス条件として具体的にあげられているものに「可読性の条件」(Legibility Condition)というものがある。演算で形成された統語列がインターフェイスのレベルでそれぞれ概念/意図体系と感覚/調音運動体系が読み取れるような情報であることを要求する条件である。インターフェイスでうまくつながらないと、たとえ言語機能を備え、また調音器官があっても言語として発現しない。

結局極小主義での主張にそって具体的に言語獲得装置の初期状態を研究するための方向は2つある。一つは精神/脳との関わりを明らかにすること。もう一つは言語使用の側面の解明である。前者の場合には、細胞レベルの説明ではなく脳機能としての満たすべき特性、基本原理の解明が求められる。さらに、遺伝子の具体的な設計図がどのようなものであるのかの研究も重要である。受精卵の細胞はある時期まではすべて同じで、どの部分がどのような器官に発達

するのかあらかじめ決まっている訳ではない。どのような原理に基づいて生物が全体として成長するのも依然として解明されていない。この意味で、生体器官としての言語能力の基本的な形成について依然はっきりしないのも不思議ではない。後者については、言語と現実世界との関わりである。一つは「音声」、もう一方は「意味」が絡む。

精神／脳の機能から課せられる条件は、言語の設計の適正さどう関係してくるか？ 一つの問題は、言語にはインターフェース制約の他にも必要なレベルがあるのかどうかの問題である。極小主義は言語に内在するレベルはなくて、インターフェースにかかわる可読性の条件だけで説明する方が適切であることを証明しようとする。演算操作で必要なのは例えば最低限「語」のような単位は必要であろうという最小限のものである。語は外的装置が「解釈」しなければならない単位である。そして、もう一つは語が結合してより大きな表現を構成することである。最後には語は音と意味を持つという点。つまり、「素性」とよばれる音と意味の特性、その集合としての「辞書項目」(lexical items)、それと辞書項目が連結してできた「表現」の三要素からなる。したがって、表現形式を生み出す演算体系は二つの基本的な演算をする。一つは素性をまとめて「語彙項目」を作ること。もう一つは語彙項目をより大きな統語表現にまとめあげること、である。語彙項目を構成する素性は可読性の条件の要請で次の3つの面をもつ。意味のインターフェースで解釈される「意味素性」、音声インターフェースで解釈される「音声素性」、そしてインターフェースには属さない「素性」である。素性については音韻論ではプラグ学派以来の研究伝統を持つ。したがって、次に、この分野がすでに新しい要請に答えられる用意があるのか、それとも全く新しい視点から再出発すべきなのか点検してみる。

#### 4 . Halle と Hawkins の仮説：音声素性理論と脳神経科学

Morris Halle は文字通り 20 世紀後半の音韻論研究の牽引的な役割を果たしてきた学者であるが、2002 年に著者自身の編集によるアンソロジーが出版された (Halle, 2002)。音韻論のあらゆる分野で膨大な業績を残している学者であるが、一貫して音韻現象分析を通じた母語話者の持つ言語能力の解明を目標に置いてきた。特に生成理論では音韻部門は「解釈」にかかわる部門で、統語部門からの出力に音形を付与する役割を担う部門である。このモデルが母語話者の言語能力の記述、説明であり、運用モデルでないことはあるにせよ、生成能力を統語部門に置き、音韻部門がその後に来るという順序関係からして、「聞き手」より「話者」中心のモデルである印象が強いことは否めない。そういう中で、Halle は音韻研究を常に話者と聞き手を結ぶモデルとして据え、この両者を結ぶのが「記憶」に蓄えられた「音韻表示」や規則などからなる「言語能力」であると考え、いわゆる音声学と音韻論という枠を互いに超える必要性を

強調してきた<sup>3)</sup>。このアンソロジーは特にこのような視点から取捨選択された論文が収められているが、このような研究視点はまさに音声に関わるインターフェース問題そのものであり、従ってこれからの研究にとって示唆するところは非常に大きい。

記憶の中で語や形態素に関する音韻情報がどのように表現されるのかを明らかにすることは音韻論の重要な課題の一つである。語を構成する要となる意味と発音の関係は恣意的なものである故、学習によって記憶されなければならないが、記憶される部分で発音に関する情報は基底形と一連の規則体系である (Halle, 1962, Chomsky and Halle, 1968)。基底形は分節素の列であるが分節素自体は素性の束として定義される。いわゆる余剰的な情報をすべて排除した情報が記憶には蓄えられると考えられる。それぞれのレベルで余剰な情報を特定し、これを規則や制約で復帰させ、完全な形に復活させるという考えはそれほど新しくはなく、構造言語学での音声と音素の関係も基本的にも同じ考えに基づく。特にこの点に焦点をあてた研究は“underspecification theory”として独自の研究領域をあたえられたほどである。Halleの新しい企てでは、余剰性を除いた基底表示は脳の中で記憶される場合、最小の負担で済むという解釈になる。

記憶にはメンタル辞書 (mental lexicon) と規則や適用条件などが蓄えられている。メンタル辞書には個別言語の単語または形態素について、統語、意味上の情報に加えて音声に関する情報が蓄えられている。音声に関して、それぞれの語または形態素がどのような形でどの程度の情報が必要かについては、ここ半世紀ほどの研究でかなり深い理解が得られた。素性は個別言語の分節素を定義する機能を持つとともに、調音器官への司令を司り、また分節素の音響的な特徴に対応することで、音声認知の機能も果たす。記憶上は、各語彙または形態素は互いに独立した分節素の列として表されることが考えられるが、実際の音声として発話された場合は連続した一連のスペクトルとして実現される。一方調音は、基底では分節素単位に記憶されているが、具体的な発話運動では特定の発音器官が一音まるまる調音するのではなくオーケストラのようにいくつかの独立した調音器官が独自の動きをする中で全体としてスペクトル状に時間的な流れとして実現される。調音器官が独自の動きをする際にも一つの動きから別の動きに瞬時に移れる訳ではなく、徐々に滑らかに移行するわけで、この動きは前後の音が何であるかによってかなり影響を受け、実際には同じ分節素でも前後の音によって音声学的にはかなり違った音として実現されることはよく知られた事実である。従って、基底表示 (記憶されている語句の音韻情報) の分節素 (音素) は発話で明確な境界をもった分節素としては実現されていないので、一対一の対応関係にはなく、従って、発話そのものを入力として一義的な音素列を自動的に特定することは無理である。

そこで、現実的な分析装置として考えられているのがいわゆる「合成による分析モデル」 (analysis by synthesis model) と呼ばれる手順である。これはインプットから可能性のある音素列に分析した結果を長期記憶に蓄えられている語の音韻表示と比較対照することによって

マッチしたものを特定する方法である。このような分析と演繹の両方向の分析を備えることで、対象となる言語に記憶された語彙以外のものは原則として認知されないことになるので、インプットのスペクトルがどのような分節素列なのかを比較的容易に認知する作業が可能になる。問題はどのレベルでまたどのようにして比較対照をおこなうかである。一つはインプットの分析と比較する対象をいわゆる生成音韻論という音声表示にするやり方である。つまり、基底表示からいくつかの可能な音韻規則を適用した結果を比較対照の対象とする方法である。どのような規則が適用されるのかは、主として前後の形態論的な情報が重要になる。従って、インプットの分析には、単に音声情報だけでなく、意味上の情報や統語上の情報や語用論的信息を利用して、分析対象となる発話を徐々に分節化する方法である。語境界自体を示す「ポーズ」がすべての語境界に置かれる訳ではないので、超分節素的な情報によっていくつかの可能性を内的な分析と照合しながら決定する必要がある。

このような手順を仮定すると、脚や音節、頭子音結合や韻など、音韻現象を分析した結果理論的に必要と考えられている単位が認識上も非常に重要な役目を担うと考えられる。もちろんこのような単位はあくまで理論上の概念であって実際の発話にそれぞれの単位と対応するような音響的な特徴を見つけることは無理であろうが、およそどのような構造をしているのかについては、対照回数を重ねれば可能性はしぼられるはずである。個別言語での制約は個々に特徴的でかつ限られているからである。このような単位は理論上階層を成すと考えられているが、この特性は後で触れるように重要である。いずれにせよ分節素列が特定されるには、このような「予備的」な総括的な分析、対照によって可能性を狭める手順が必要であろう。

このような操作は、結果的に最終的な分節素列を特定するわけであるが、それに至る過程で、音韻情報としては記憶されない生の発話データに含まれる個人的な特徴や余剰的な特徴を取り除く機能をはたす。このような個人的、一時的な情報は予備段階での分析では重要な役割を果たすと考えられ、必要があれば一時的な記憶からかなり長期の記憶まで、場合場合によって蓄えられると考えられる。つまり、ラングとしては余剰的な情報も個人の特定化などでは逆に主要な要因となるからである。

さて具体的な「照合」過程であるが、このレベルでは何が必要とされるか？インプットとなる発話は音響上の特徴を備えたスペクトルである。照合される基底形は脳に記憶された各調音器官への司令となる素性の束の列である。先にも触れたように記憶されている情報は分節的であるが、具体的な筋肉運動となった場合分節境界は曖昧になる。従って両者の照合は言い換えれば調音上の特徴と音響上の特徴の照合に他ならないので、直接両者を比較することはできない。一つの可能性は、辞書で調音上の特徴だけでなくそれに対応するスペクトル上の特徴もあわせて記憶されていると仮定することである。しかし音素列に対する音響上の特徴には様々な変異形があり、辞書にすべてをあらかじめ記載しておくことは記憶の負担が増えすぎる恐れがある<sup>4)</sup>

Halle and Stevens (1962) で提案されている「合成による分析」モデルではこのような記憶の負担を軽減するために辞書には基底形の情報からスペクトルに対応する特徴を生成する「規則」が記憶されていると仮定されている<sup>5)</sup>。つまりこのモデルでは、照合はインプットとなるスペクトルと、まず規則によって基底形の調音的に基づく素性の束から独自に対応するスペクトルを作りだし、その後でスペクトル同士を照合し、一致したと判断されればそのスペクトルに対応する基底形が改めて出力され、語の認識が完了するという手続きを考えている。このモデルの長所は、記憶の負担を軽減できる点に加えて知覚と調音が同じモデルで説明できる点にある。つまり知覚に際しても基底形を形成する調音系の素性からいったんそれが音声、つまりスペクトルに実現された特徴を生成する仕組みを持つからである。

しかしながら Halle (1983) で指摘されているように、スペクトル上の情報と調音上の特徴には必ずしも一対一の対応がない。例えば、よく知られているように閉鎖子音間のスペクトル上の特徴は前後の音のスペクトルに反映される。もし前後の音が母音であれば両唇閉鎖音は母音に“negative transition”を生じ、一方非円唇軟口蓋閉鎖音であれば第2フォルマントと第3フォルマントが一つのフォルマントとして現れる。しかしながら前後に子音が来る環境では、閉鎖音の種類についての音響上の決め手は閉鎖解放時の破裂に現れる違いで読み取らなければならない (p.92)。このように、同じ閉鎖子音でも環境によって調音上の特徴と音響上の特徴には厳密な対応関係が見られない。したがって、Halle (ibid.) では調音上の素性と音響上の素性を結びさらに抽象的な素性を設けて、話者/聴者に中立な母語話者の言語能力の一部と考えている。しかし、上で見たようなモデルを考えた場合は、さまざまなレベルでの余剰的情報があるので、厳密な対応がなくても予備的な分析で範囲を限定することで認識は可能であろう。この問題については後で再度触れることにする。

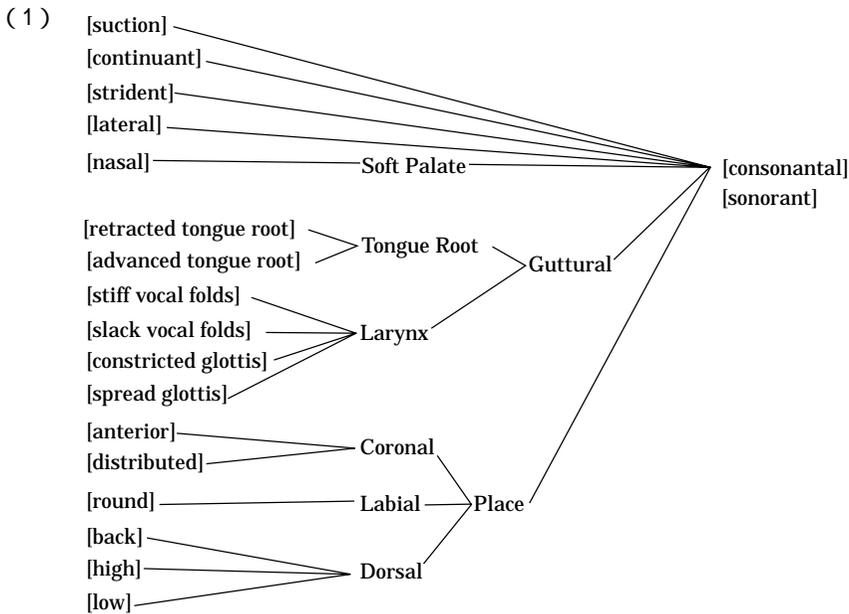
もうひとつ注意すべき点は、極小主義で要求されるインターフェース条件は、具体的な調音器官の筋肉運動や聴覚器官の仕組み、機能との連携を保証する条件であるが、たとえば聴覚上の条件といっても、先にはスペクトルの分析に基づく音列の特徴をあげたが、実際には音声は空気の振動だけではなく、自分の声を聞くときのように骨の振動などでもよいわけで、このような点を考慮に入れると、結局脳の中での刺激の反応にすべて帰さなければならないようになる。しかし、くりかえし強調しているように、このレベルでの研究は言語学の対象ではないし、言語学者としては興味もない。すなわち、一つ手前の段階で留まるべきである。となると、インターフェース条件は極小主義でのプログラムとしては受け入れられても、現段階では言語内部からの言語能力解明に重点を置くのが自然である。このような状況の中で、インターフェース条件に関わる理論として、先の合成による分析モデルの妥当性を素性理論をブリッジとして脳の仕組みに関する理論との合体の可能性を以下で示してみたい。

5 . 素性構造の階層性

従来から提案されてきた言語理論で明らかになったことの一つに、各種のレベルでの表示は階層をなすという特徴があげられる。これは構造主義の構成要素分析の特徴でもあったし、句構造をはじめとして生成理論でも重要な構造表示方法として採用されていた。音韻論も例外ではなく音節、モーラ、フットなども階層構造をなす。Halleによる上掲書の後半におさめた論文のほとんどは音韻素性の構造に関するもので、これも階層をなすのが特徴である。

先にみたように音韻素性は分節素を定義する機能を持つが、調音上の素性で表示された場合、調音運動の「司令」の機能を併せ持つと考えられるようになったことは、インターフェース条件にとっては大きな進展である。さまざまな言語での音韻現象を調べると、音声は単独ではなくグループごとに変化などを受けることが知られている。生成音韻論の初期の頃は自然音類という概念と音韻規則は自然音類を表す素性を対象にすることでこの現象を説明していた。しかしClement（1985）が素性間の関係を階層表示で行うべきであることを提案して以来、基本的には同じことであるが、どの階層に規則が適用されるかで音類の役割が特に調音運動と関連させる形でより鮮明になった。ここで、脳の新皮質の階層性と調音 / 知覚インターフェース問題の接点が浮かび上がる。

素性間の関係は「素性幾可」と呼ばれる次のような素性樹 (feature tree) によって階層的に表示できる (Halle 1995)。



(1)で 終端要素は素性、非終端記号は素性をその機能と調音上の特徴によってまとめる節点を表す。これは機能上同じ群にまとめられた素性は互いに同化現象が起こることが可能であるのに対して、異なった節点同士の要素間では同化は起こらないなど、音韻規則でまとまった単位として機能するという経験的事実を反映している。またそれぞれの節点は音声を産生する場合に具体的に活動する調音器官(唇, 軟口蓋, 喉頭(声門), 舌根, 舌葉, 舌先)に対応している。素性は特定の調音器によって生み出される効果か否かによって2種類に大別される。例えば, [nasal], [back], [anterior]などの素性はつねに特定の調音器によって実現されるが[continuant]や[lateral]などは異なる調音器官で作出すことができるので, 両者は異なる性格を持つ。特定の調音器官を持たない素性は逆にいえば, 必ずその効果が得られる調音器官を指定する必要があることを意味する。現実の発話では, 特にぞんざいな発音や早口の場合などでは, 個別の調音が完全に行われないが, このような場合も知覚上は問題が起こることはまずない。これは後者の素性に特定の調音器が特定される場合にあたる。言い換えれば, 特殊と一般といった分類に相当し, 階層状の樹形図を用いればうまく説明が可能になる。この点は後で触れるニューロンの構造とも並行的で, この意味でも非常に興味深い仮説である。

ここで音声素性に関してもう一点指摘すれば, Halleは, ヒトは特定の音響的特徴を生み出すためにどの調音器をどのように動かすのか経験的に学ぶのではなく, 生得的な能力として持っていると考えている<sup>6)</sup>。従って, Sapirが指摘したように<sup>7)</sup>, ローソクを吹き消すときの動作は経験的に学ぶかもしれないが, 同じ動作が「言語音」として使われるときには, 生得的な言語能力の一部として実現された音である。一方でオウムや九官鳥のように, 音響上はヒトのコトバに近い音を発することができるが, 言語能力からの出力がない故ヒトのコトバとは異なることになり, またヒトの言うことがある程度理解できるボノボやチンパンジーなどの類人猿は, 言語能力については不明であるが, ヒトの音声にあたる音響上の音が出せないという点で, 少なくともインターフェース条件を満たしていないといえる。

## 6. 「記憶」の形式と機能

通常脳に関わる研究にはいわゆるトップダウン方式とボトムアップ方式2つのアプローチがある。言語活動など高次の機能の解析によって, 脳神経系の働いている分子メカニズムを解明するか, または分子の解析の積み上げから, 全体としての脳神経系を解明するかのやり方があることになる。ただしボトムアップ方式の場合には, どの程度まで「ボトム」に降りるのが問題になる。脳の場合分子レベルの研究やニューロンの解析だけで脳の高次機能を解析できるかには疑問がある。なぜならニューロン自体はどこの部位をとっても基本的な形や生物機能にはほとんど違いが見られないからである。従って, 多様な機能を発現するのは個々のニューロ

ンというよりニューロンの組み合わせ、回路網にあると考えられる。ニューロンがどのような回路網を形成しているのか、またどのような機能をはたしているのかに関する問題は、言語能力やインターフェース条件を考える上で重要である。この回路網の特徴の一つに「階層」を形成していることがある。ニューロンの階層構造についてはこれまでもいくつかの説が提案されているが、その中でも素性の階層構造との平衡性を説明できるという意味で注目に値するのがHawkins and Blakeslee (以下H&B) (2004) の仮説である。彼はこの著書でニューロンの形成する階層構造に関していくつか非常に興味深い仮説を提案している。彼の仮説は言語、特に音韻理論で考えられてきた構造といくつかの点で共通性がみられ、音声脳でどのように処理されるのかなどに関して非常に有益な考察が可能になる。音声に関するインターフェース問題を考える上でも興味深い。ここではあまり詳しい議論には立ち入れないが、極小主義の主張や素性理論と関連すると思われる部分について考察する。

「統合による分析」モデルのところでも見たように、このモデルの特徴は、音声認識の対象は心理上は分節的なので、音声学的な情報のみでは語の境界やできないという困難さを解決しようとするところにあった。H&Bが提案している「記憶」についての仮説は、この問題点の解決にヒントとなる。彼等は脳の中の記憶に関して、次の4点にその特異性を見る<sup>8)</sup>: The neocortex 1) stores sequences of patterns, (パターンのシーケンスを記憶する) 2) recalls patterns auto-associatively, (パターンを自己連想的に呼び起こす) 3) stores patterns in an invariant form, (パターンを普遍形として記憶する) 4) stores patterns in a hierarchy way (パターンを階層的に記憶する)。4つの特徴に共通なのは記憶されているのは「パターン」であるという点である。この点の前提として彼はVernon Mountcastle<sup>9)</sup>の説をとりあげる。それは「新皮質そのものは機能が異なるとされる領域でも見た目は同じであるという事実から、すべての領域で起こることはまったく同じ演算手段が使われているのではないか」という仮説である。つまり、視覚や聴覚、触覚、さらには運動上の処理を行う際、演算処理に違いはなく、同じ方法で処理されていると考えている。さらにこのことから言えることは、個別のニューロンには違いがなく特定の機能もないので、ニューロン間の接続の仕方、ネットワークの組み方でさまざまな自立的な高次の機能や特異性が生成されるのではないかということである。すると、言語能力もこのようなニューロン間の回路網の結果生じていると考えられるので、生物言語学の観点からいえば、どのネットワークが生得的で、どのネットワークが経験によって形成されるのか研究の焦点となるとも言えよう。生得的とはネットワークの構築に遺伝子がかかわるということである。よく知られているように言語についてはブローカ野やウェルニッケ野など言語野と呼ばれる領域があり、この他にも新皮質にはいくつかの特定の機能に関わる部位がある。ニューロン間の回路網の構築には経験が大きな要因になるが、どのような経験によってどの程度の回路網が完成されるかについては遺伝子が決定する事項である。

Mountcastleの仮説からもう一つ言えることは、脳の柔軟性である。例えば、言語障害などに関しては、その補完能力もかなり柔軟であることなどは、回路が構築される際の入力となる刺激がどのようなものであるのかによってニューロン間の回路構築が変わりうることを示す（通常期待される刺激が入ってこない場合は、その領域は他の刺激によって新たな回路が決定されたりする）。つまり、もし新皮質のある部分に障害があれば、他の部分はその機能を引き受けることも起こる。例えば聴覚に障害があった場合でも手話など音声言語に代わる言語手段が獲得されるし、書き言葉も、身振り言語も、ことばに似た音楽も、習得できる。脳の学習能力は脅威的で、かつ柔軟性に富み、日常言語の「学習」ということばにまどわされて、その本質を誤解している恐れがある<sup>10)</sup>。つまり、経験的にどのような情報が入力されるかによってニューロンは特定の機能を遂行するおびただしいネットワークを形成するのである。このように新皮質の機能分化は経験的に入力される情報の種類に大きく依存するが、脳には特定の機能領域があることから、この柔軟さもある枠内に限られ、あらかじめ遺伝子には新皮質全体の構成が書き込まれていて、領域間のネットワークも特定されていると考えられる<sup>11)</sup>。

新皮質が均質で、脳の領域の果たす機能はニューロンの結合の違いとして説明され、演算様式も同じであるとする、もう一つ重要なことが言える。それは新皮質への入力があるような感覚器を通るにせよ、情報が一旦脳に入れば入力形式はすべて等質であると考えられることである。経験的には聴覚、視覚、触覚は全く異なったもの感じられる。聴覚は音色、高低、強弱、リズムなどからなるし、視覚は形の他にも色彩、奥行き、広がりなどであり、触覚は肌触りや温度の高低、形状などで、それぞれでかなり異なる。音は空気疎密波であり、文字は視覚で光、触覚は皮膚が受ける圧力と、それぞれ異質な特性を持つ。しかし処理のアルゴリズムが同じであるということは、感覚的には異なる入力でもすべて等価であると考えざるを得ないことになる。等価であるというのは、ニューロン自体はどこの部位をとっても基本的な形や生物機能にはほとんど違いが見られないという性質が逆に、異なる感覚を同時に、また相互に処理したり、同じ情報を別々の感覚器や運動系の器官に送ったりすることが可能になることを意味し、実際にこのような例は数多く報告されている。

異なった感覚や出来事は同じアルゴリズムで処理されるとすると、次に問題となるのはどのような形で異なった感覚、出来事として記憶されるのかという点である。同じくMountcastleはこの点に関して「パターン」という概念を使う。これは先にあげた記憶の4つの特徴にみられたキーワードでもある。我々に認識できる情報は、感覚器が何であれ「空間と時間に関わるパターン」(special and temporal patterns)に限られている<sup>12)</sup>。「空間的なパターン」とは「同時に現れている情報」(coincident patterns in time)で、同じ感覚器官の複数の受容器が同時に興奮する時生成される。視覚の場合には、光は網膜のある範囲に同時にあたるので「空間的なパターン」が生成される。これは現実の「像」とは次元の違う「電気信号からなるパター

ン」にすぎない。一方、聴覚には空間的な要素がないように見えるが、実際には空間的な側面も持つ。音波を電位に変えるのは蝸牛殻と呼ばれる渦巻き状の器官であるが、耳に入る音は成分に応じて<sup>13)</sup> 蝸牛殻の異なった部分を振動させる。つまり多数の周波数からなる音声は同時に蝸牛殻の異なった部分を振動させ、結果的に「空間的なパターン」が発生している。逆に、視覚も時間的なパターンも同時に持つ。つまり映画の場合のように、空間的なコマが時間的に流れるのが現実だからである。静止した物を見る場合には、逆に目の方が一秒に約三回視線が急激に移ることで<sup>14)</sup> 時間的な情報を形成する。

音声知覚の場合、閉鎖子音などは前後の音にどのような変化を与えるかで分節素の種類が認識されることは先の認識モデルのところでも見たが、理論的には各分節素は素性の集合として「定義」できるのにも関わらず、実際のスペクトルの分析では隣接する分節素の情報が必要なことは、音声認識も一定の時間的な幅（時間的なパターン）と分節素を構成する同時生起的素性情報（空間的なパターン）が欠かせないことを意味し、先の指摘は非常に示唆的である。そもそも音素は対立の体系をなす点が基本で、特に強勢や音調などに代表されるように単独の要素では意味がないのが本質であるからである。

「パターン」についてはこのくらいにして、次に情報がパターンとしてどのように記憶されるのかの問題に移る。まず注意すべき点は、すべての記憶はニューロン同士をつなぐシナプスに蓄えられているが、その新皮質に記憶されている事柄の数はきわめて多いのにもかかわらず、どの瞬間でも思い出せるのは記憶のほんの断片でしかないことから、記憶からからの呼び戻しに重要な役割を果たしているのはごく限られた数のニューロンとシナプスであろうと考えざるを得ないことである。これが(1)の「新皮質はパターンのシーケンスを記憶する」の意味するところである。多量のことを思い出せても、順に少しずつであって、一度にすべてを思い出すことはできないということである<sup>15)</sup>。これはコトバによる記憶の呼び戻しでも、「映像的な記憶」であっても、「感覚的な記憶」でも同じである。コトバによる場合は一度に複数の文を作れない（考えられない）ということである。文を「発話」する場合は運動であるので同時に複数の音を発音できないので、結果的に複数の単語、文を発話できず、音を順に発音するしかない、と説明される。従って、言語はいくつかの単位がどのレベルでも順に並べられることが一つの特徴とされる。しかし、これは実際に「音声」として実現されない場合にもあてはまるわけで、その意味で脳の機能の制約であると見なすことができる。「制約」でもあるが、記憶の呼び戻し方の実際であるともいえる。

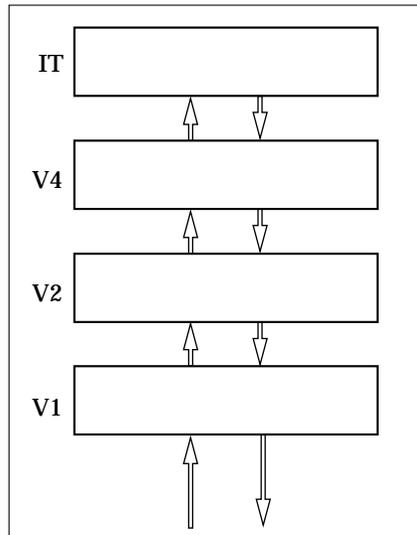
次に、二つ目にあげられている「パターンを自己連想的（self-associative）に呼び起こす」という特徴であるが、「自己連想的」とは「入力されたパターンと同じパターンが記憶から呼び戻される」という特徴で、これが保証されると、たとえ入力の一部だけであったり「変形されていても」記憶から完全なパターンが出力（知覚したり思い出したり）できるというもので、

一種の「復元能力」をいう。部分的にしか見えなくとも全体がわかったり、特別な匂いをかいだだけで昔の思い出が次から次へと思い出したりできるのも、記憶は自己連想的(かつ連鎖的)であることの証拠である。ただし、その時々「現実」の入力された情報より自己連想的な記憶の呼び戻しが勝ってしまうと、勝手な「置き換え」が起こり、「誤解」が生じることになる。「聞き違い」や「幻聴」、「空耳」なども自己連想が勝る場合である。新皮質は、自己連想によって現在の入力を補い、自己連想によって次の予測を行う。日常的にわれわれが行っている会話や一般的な認識活動では、それほど意識的に細部まで注意を払っているわけではない。誰かと話をしていても相手の話(音声自体を対象とする場合も含めて)に100%注意が向けられているわけではない。つまり、断片的な入力でも事足りていることが多い。これは以下で触れる「予測」とも関連するが、我々の認識や運動(行動)など日常的な活動にとって本質的な関わりをもつ。

記憶の三つ目と四つ目の特徴には密接な関係がある。(3)のパターンを不変形として記憶するというのは、例えば動いている対象を見る場合、視覚的な入力情報は絶えず変わるにもかかわらず同じ「モノ」として知覚できることを説明する原理である。(4)の「パターンを階層的に記憶する」とはこのような絶えず進行する変化と不変性の折り合いはパターンの記憶が「階層的」であることに起因するというものである。この点は音声認識を説明する際にも非常に大きな意味を持つ仮説といえる。ヒトがモノを覚える際も、また思い出すときも細部については曖昧、またはまったく無視していることがすぐに納得できる。この点がコンピュータなどの機械と違うところである。コンピュータなどは同じかどうかの判断は厳密な対応関係に基づくので、人間から見れば「融通がつかない」ようにうつる。この「融通が利く」とは言い換えれば「厳密さ、精密さ」に欠けるということである。細部にこだわらないということは、逆に言えば「本質」がわかっているということである。この点は先の(3)の「部分で全体がわかる」という特性とも関連する。見方を変えれば、部分は「具体性」で全体は「本質」とも言えるからである。両者は「復元可能性」で結ばれている。

この「復元可能性」を保證する具体的な新皮質の構造特徴が(4)の記憶の階層性である。この点に関するH&Bの仮説はある意味で斬新的である。新皮質の組織はコラムと呼ばれる6層からなる円柱状の構造からなり、全体としては蜂の巣状であることはよく知られているが、その機能についてはさまざまな見解がある。しかし注目すべき点は、各層がシナプスによってつながれていて、柱状構造で縦に並んだ細胞が同じ刺激に対して同時に興奮する傾向があることである。H&Bは次のような新皮質構造を提案している(P.109)

(2)



The first four visual regions in the recognition of objects.

この図は例として視覚野を示したものであるが、感覚野の階層構造はすべて同じであると考えられる。ポイントはニューロンが反応するパターンは階層ごとに性格が異なる点である。つまり最初に入力が処理される層では個々の情報が検索されるが、この信号が上層部に行くほどニューロンは「不変」部分にのみ反応するようになっていくという点である。例えば、視神経から軸索を経て運ばれる視覚信号はまず第一次視覚野V1に入力される。この信号の特徴は絶えず変化する「パターン」に反応するニューロンからなる点である。この信号はV1からV4へと順次伝達されるが、階層を上る間にニューロンの性格が変わる。つまり、V1層のニューロンは特定の入力パターンだけに反応し、絶えず変化する細かな特徴を認識するが、階層を上るに従って興奮状態はあまり変化せず、全体を認識するようになる。例えば、顔を見ているときに興奮する「顔細胞」などの存在はよく知られているが、階層の上のニューロンは、顔がどのような状態でも、またそれが絶えず動いても、その一部さえ見えている限り興奮を続けるという特徴を持つので、これは上位の細胞の機能であると考えられる。つまり、動いている物体を見る場合、動きに合わせて刻々と変わる面を認識する「下層」と、いくら変化しても、全体として「不変」の部分に反応する「上層」細胞層があり、層構造全体として動いている対象の認識が可能になるという訳である。

再度音声産生 / 知覚モデルに話を移すと、このようなモデルは「不変（普遍）性」と「多様性」という言語学の重要な課題や、言語は統語的にも音韻的にも階層理論がさげられないとい

う知見が、一般的な脳の仕組みに還元できるという点でこの考え方は非常に示唆的である。例えば、語は通常脚、音節、モーラ、分節素（コア）のように階層をなし、また（1）に示したように素性構造をとっても階層を成す。通常、階層の最上層は全体をまとめる機能を持ち、下に行くほど個別の細部の情報に関係する。例えば、言いたい語を思い出せない場合など、全体の「感じ」や音節または拍数あるいは、語の最初の音（または音列）などはぼんやりとでもわかっていても、全部が思い出せない場合がある。逆に語全部の発音を正確に聞かなくともどの語のことかわかる場合も多い。つまり、運用モデルなどでは余剰的な情報や、統計的な情報が知覚上かなり大きな働きをすることは従来から指摘されてきた。しかしこのような情報は、逆に恣意的、または可変的な要因で、言語能力の対象からは外されてきた。しかし、例えば音韻論でいう語の基底形から音声表示を派生したり、音素から異音や自由変異形を導く規則などは<sup>16)</sup>、ある意味「不変」から「具体形」への変換である（基底形では素性は二項対立なのに対して音声表示では数字による「度合い」の表示になることもこの現れであると見なせる）。逆に、個々の発音はその都度異なるのに「同じ」と感じるのは言語能力として抽象的な「不変の表現」を持っているからであると考えられる生成理論の考え方は、まさにこの新皮質の特性と見事に一致する。

個別の情報がどの程度全体の産生 / 知覚に必要であるかについては、もう一点H&Bの仮説を考慮する必要がある。これは「合成による分析モデル」が目指した、音声の産生と知覚を同じモデルの妥当性を証明するだけでなく、そのようなモデルが必要である点を明らかにもしてくれる。それは「記憶による予測」問題に関する仮説である。認識における「予測」の重要さは古くから指摘されていることで、とりたてて新しい考え方ではないし、音声認識に関しても予測を前提としたトップダウン処理の重要性がしばしば指摘されてきている。音声を聞く場合等、必ずしも音の特徴を細部まで正確に聞き取って理解するわけではなく、前後の文脈や状況に応じてあらかじめこう聞こえるはずだという予測が働いているとしか説明できないことがたくさんある。H&Bの主張は、我々の通常の行動は感覚器を通して入力された情報を新皮質がコンピュータのように逐次計算して次の行動を決定するのではなく、すでに蓄えられた記憶を使ってまず予測し、予測が行動を起こす、というものである。ここで当然起こる疑問は、我々の行動には初めての体験することも多いのに、なぜまだ起こっていないことが予測できるのかという点である。これが従来から予測の重要性はわかっても、具体的な説明や理論化が難しかった理由でもある。H&Bは、脳にはそれまでの経験から「感覚上の記憶」と「運動上の記憶」の両方が蓄えられ、その時々を外からの刺激に対して常にこの記憶からの入力加わり、記憶された感覚や運動を絶えず検証しながら認識や行動が行われると考えることによってこの疑問に答える。「感覚上の記憶」と「運動上の記憶」には対称的な特性がいくつかある。例えば投げられたボールを受けようとした場合、脳では次の三つのことが起こる。まず投げられたボールを見ることによって、自動的に過去の同じような光景の記憶が呼び起こされる。次に以前の

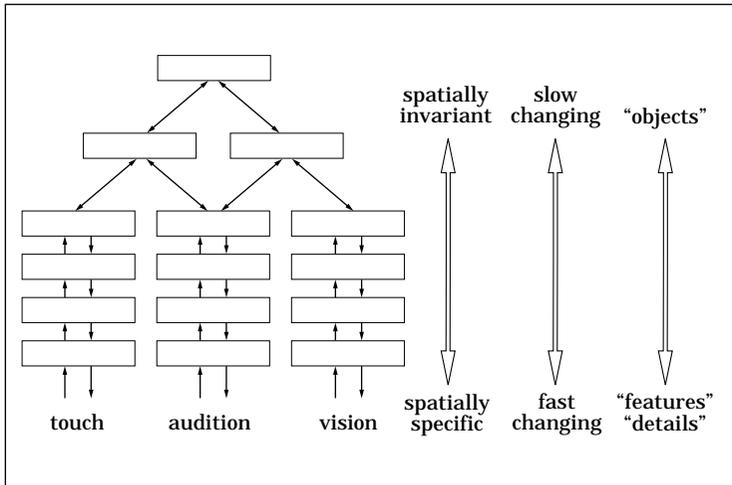
記憶から筋肉への司令の記憶が引き出され、実際に筋肉へ司令が出される。三つ目は、その瞬間の身体状況やまわりの物理的な状態が感覚器官でチェックされ、記憶によって実行される命令が絶えず微調整される。ボールを初めて受けたり、経験したことがないほど早いボールが来たりする場合には、どう行動するか司令の記憶がなかったり、調整機能が十分働かないので受け損なったり、びっくりしたりするのはこのためである。我々が日常的に感じる「違和感」や「驚き」は「予測に反したときの反応」である。重要なのは何が見えるか、聞こえるかといった感覚上の入力をチェックするだけでなく、次に自分が何をするのか、つまり「運動としての出力」も脳は「予測」していることである。そして予測に反すると「変だ」とか「驚き」の反応となって返ってくる。何も「考える必要がないときは」すべてが「予測通り」進行している場合で、「自然」である。つまりこのことは、「実際の入力」より先に「記憶からの」入力が処理されていることを示している。

現実の音声を知覚できたと言えるのは、絶えず変化する入力から、例えば特定の「語」を認知できる不変性を見つけることができた時である（同時に記憶の語との一致が確認された時）。この過程で、先にも述べたように「予測」が働くが、脳がこの予測をたてるには個別と不変の規則性をその時々的事实と比較対照する必要がある。その時々の入力情報がいかに些細でも、新皮質はさまざまな可能性を予測する。その時々の声の大きさや、誰の発話か、どのような状況での発話か、話し手の気分、意図、様々な予測が瞬時に行われる。これに記憶に基づく予測が関わっていることは、例えば初対面の人や、いわゆる異文化間でのコミュニケーションの場合、対応がスムーズに進まず、発話自体を含め周囲の状況に意識的に敏感にならざるを得ないことがあるが、これは記憶からの予測がうまく機能できない場合であると考えれば、納得がいく。

このように記憶された行動の情報が認知に働くという考え方は、ニューロンの情報が流れる方向は感覚器から内へ一方向的で、脳はこれを計算処理すると考えるのではなく、同時、もしくは外からの情報より先に記憶からも入力が流れ込むという仮説をたてることで説明がつく。はじめに指摘したように、脳は均質的で、そこでの演算アルゴリズムは機能領域に関わらず同じであることを合わせて考えると、記憶からの「不変（普遍）」情報はさまざまな感覚器官に同時に現れうる。つまり、言語で言えば言いたいことを話したり、書いたり、手話にしたりと多様に具体化が可能になることと合致する。また、音読や黙読ができることも説明がつく。

ニューロンの階層性、不変と個別の認識、処理の違い、外からの入力と記憶からの予測という「双方向」の入力の流れ、運動と感覚の普遍性、などの特性を組み込んだモデルを図示すれば次の（3）のようになる（Hawkins and Blakeslee 2004 P.114）。

(3)



Forming invariant representations in hearing, vision, and touch.

このモデルの特徴は、各感覚の階層の上位にはそれぞれの層を統合する層があることである。これは、最終的には新皮質の「連合野」に相当する領域である。われわれの経験は視覚、聴覚、触覚にまたがって起こることが多い。鳴き声を聞いただけでどのような鳥かを認識し、実際に見みることによってその具体的な存在を知覚するなどである。聴覚情報が入力されると階層の一番下から順に上に上がっていき、視覚と聴覚の共通の連合野に届く。もう一つの特徴は、情報は上層部から下降部へも流れる点である。つまり連合野から逆に情報が末梢的な視覚野と聴覚野へも下る。つまりこれこそ先に述べた「予測」に他ならない。連合野は視覚野と聴覚野へそして行動を起こす筋肉にも司令を送る回路を持つわけである。情報の流れは一方向ではなく双方向である。

## 7. 音韻論研究の展望

ヒトの生得的言語能力についての考え方の変遷を見ると、規則、演算体系や様々な制約についてはこの間かなり急激な変化をとげてきた。しかし言語はあくまでコミュニケーションをはかろうとするグループ内の構成員で共有しなければならない部分が多い。特に音声と意味は恣意的な関係にあるが故に経験的に記憶された情報に頼ることが多くなる。コトバに関する記憶の中心はそれぞれの言語の単語、または形態素に関する情報である。極小主義の台頭でこの傾向がいっそう強まっていることは先に指摘した通りである。

語や形態素の音韻的情報について長期記憶に蓄えら得る情報は基底表示と呼ばれるが、これ

は(2)で示したように示差的に機能する素性の集合が列になった形で表示されると仮定される。HBモデルの記憶の特徴を思い起こすと、この基底表示はどのように評価できるであろうか。基底表示が「不変形」であることは先に指摘したが基底形のみが「不変形」の正体ではなく、品詞や文法、意味情報すべてを含めた「不変(普遍)」が記憶にはあるはずである。頭の中だけの思考はこのような不変形の列からなる文のシクエンス的かつ自己連想的記憶の発現と言えるであろう。運動記憶とは表層の音生産/知覚の結果に関する形の記憶である。記憶では、すべての感覚に共通する形式が「不変形」でもある。従って「音」産生は調音器官だけでなく手話などとしての出力でもある。また聴覚は文字を対象とする場合視覚的情報とも互換性を持たなければならない。このレベルの普遍性の研究成果ができればインターフェース条件の研究としてはかなり実質的な成果をあげたといえよう。脳は出力としてはシクエンスに従うという制約はあるものの、一方では非常に「柔軟」である。この柔軟性の説明には学際的な視点が欠かせない。各分野の「専門用語」に固執してはこの柔軟性に対応する「不変形」の説明は不可能である。

(2)のような構造表示が「記憶」されていたとして、運用との連関、または「構造表示」の仮説に関わる制約等があるのか、簡単に見ておく。

子供が言語習得する際、特に「安定状態」にある大人との発音の違いをどう説明するかの問題は、近年OT理論などを中心によく取り上げられるトピックになっている。この中でも説明の対象となるのが「理解」と「発話」能力との差である。子供は大人の発話が理解できるのに発音は大人のものとはかなり異なる時期があるが、この時期の子供の文法、特に音韻に関する「基底表示」をどのように考えるかである。OT理論では通常基底表示は大人と同じで、規則や制約のランク付けの違いとして説明される。ここで詳しい議論に立ち入ることはできないが、自分の文法と他の人の文法が明らかに異なる例は、子供以外にも普通に見られる現象である。例えば、自分と異なる「方言」などは使えないけれども理解できるのが普通であるし、バイリンガルでも、使えないけれども「理解はできる」ケースもある。外国語の場合も然りである。このような「運動」としては「記憶」されていないが、「理解」はできるケースは、複数文法を内包しているとは考えにくい(つまり、「変異形」としての入力形は最終的には個人的な違いの段階まで、無数にある)。このような理解は、純粋なボトムズアップでは説明がつかない。語の恣意性から「知らない語」の意味は合成語とか複合語でない限り「理解できない」場合が多い。どちらの単語か判断できない場合も当然ある。トップダウンについては、統語的な情報や語用論的な意味で、「聞くはずの語」がかなり特定される。記憶されているのは何も「語」が単位ではないからである。シクエンスと自己連想故にある。しかし「語」単位で記憶されているなら「変異形」に出会ったときに「全体としてどれほど類似しているか」の判断がどのようにしてできるのかわからない。しかしもし(1)(3)に対応するような階層構造の形で

記憶されていると仮定すれば（この仮説には十分な根拠もあるが）階層が上がるごとに抽象性は増すので、現実の音声学的な音声列からは「遠くなる」が故に、可能な語の集合は大きくなるので、何らかの「外的ヒント」が与えられると、ターゲットとなる語の特定は容易になる。

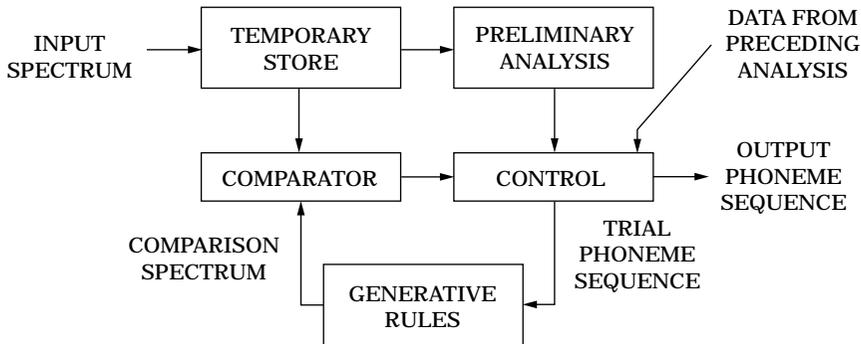
## 8 . 結び

分節素を階層的に構造づけられた素性で構成されるという素性幾何は、従来からの課題の一つであった調音と音響に関わる素性との相関をうまく説明するのに適しているばかりでなく、中和現象や同化、異化現象等を説明するのに非常に強力な記述手段であることが示されている。さらにはこのような音韻現象だけでなく、さまざまなスタイルや変異形などの言語運用に関わる現象の説明にも有効である。一方、H&Bによる脳神経の階層性に関わる仮説は、個別の感覚の認識と全ての感覚に共通した認識の仕組みを説明するだけでなく、個別の感覚認識でも特に個々の細部の変化の認識と不変（普遍）部分の認識までもみごとに説明を可能にするモデルである。この二つの「階層性」に関わる仮説は言語知識としての音韻現象や個人、社会レベルのバリエーションを説明できるだけでなく、調音器官の動きや聴覚上の機能についても同じ表示モデルで説明が可能になり、さらには従来の生成文法（特に生成音韻論）の研究成果を十分に反映できる点で「インターフェース条件」の研究の進展にとって重要な視点となりうる。

### 注

- 1) 後で触れるように、人間固有の言語能力として最後に残ると考えられる「再起性」(recursiveness)は結局は「併合」の再起性によるもので、ことばの本質ともいわれる「離散無限性」(discrete infinity)もこの仕組みから生じる。
- 2) 関連分野との連携については、チョムスキーの見解が変わってきたことに対する驚きだけで、主張自体に問題がある訳ではない。しかし、この論文での主張はかなり急進的で、すぐに多くの学者から反論が寄せられた。特にこの論文の全体にわたって反論をくり広げたものに、Steven Pinker, Ray Jackendff (2005)がある。
- 3) このアンソロジーには含まれていないが、初期にJakobson, Fantとの共著(Halle, 1952/1963)は特に分節素を定義する音声素性理論についての研究であるが、調音上の素性と音響上の素性の相関関係を考慮に入れ、音韻(音声)表示が調音/知覚の両面を備える点を強調した研究で、Halleの今述べた態度が顕著に現れている。このような研究はインターフェース条件の理解を深める上でも重要であるが、音韻論研究と音声学研究の境界を取払い、総合的な研究の必要性を広く訴えた点でも評価できる。
- 4) もちろん脳の記憶容量を実際に推測することはできないので、このような可能性を必ずしも完全に否定できる訳ではないが。

5)



6) これはもともと A. M. Liberman and I. G. Mattingly の “The motor theory of speech perception revised,” (*Cognition*, 21, 1-31) や, “A specialization for speech perception,” (*Science*, 243, 489-494), などにある説で, ここではそれに従う旨の記述がある。Halle and Stevens (1991) pp. 7-8.

7) Sapir (1925)

8) p.70. 彼はコンピューターは, インプットの演算が中心で, 記憶の果たす役割において脳と根本的に異なり, この点で人工知能には限界があると指摘している。

9) John Hopkins 大学の神経科学者で, この部分については “An Organizaing Principle for Cerebral Function: The Unit Model and the Distributed System,” in Gerald M. Edelman and Vernon. B. Mountcastel, eds., *The Mindful Brain*, MIT Press, 1978, の論文があげられている。

10) ヒトが言語を持つことと人間の脳の特徴との関連では, ヒトの脳が他の動物に比べて大きい (特に身体全体との比率で) ことが指摘される。大きいといっても全体が大きいわけではなく, 新たに新皮質が加わった分が他の動物より大きさが目立つ。新皮質でも特徴的なのは, 前部が特に発達したことである。この進化の実質的な意味は感覚野と運動野が新皮質で結合されたことにある。つまり, 新皮質の後部は第一次感覚野など感覚の大部分が認識される部分であるのに対して, 前部は高次の立案や思考に関わる領域と運動野とがあり, 人間の理性の出所である。人間以外のほ乳類では, ほとんどの行動がいわゆる「旧脳」に支配されているが, 人間の場合は新皮質が旧脳から行動を制御する機能を奪ってしまっ, 実質的な支配権を獲得している。結局新皮質が旧脳の運動システムと連動し合うようになり, 行動を支配するようになったうえ, 感覚器官と運動器官が新皮質で連結されることになり, 他の動物と異なる高次の複雑な行動, 例えば言語などが操れる能力を獲得したと考えられる。これらは, ある意味でニューロンの接続網の変化が大きく貢献したと考えられる。

11) この点は先に述べた Halle などの音声上の能力にも生得的な面があるという議論と関連する。

12) 感覚によって異なった体験をするのは, 入力領域が異なるからで, 抽象的に見れば, 基本的には同じ形式で, 同じ処理が起こる。たとえばもし犬を見ると一連のパターンの集合が視神経の繊維から視覚野に流れ込む。犬がほえるのを聞く場合は異なったパターンの集合が聴神経の繊維から聴覚野に流れ込む。犬を触れば体性感覚野に異なったパターンの集合が流れ込む。このように感覚の体験は異なる。しかし, 後で触れるように, 犬の鳴き声を聞いてその姿や体の触覚を思い出すのは簡単である。いずれにせよ, 頭の中には犬の姿も鳴き声もない。

- 13) 周波数の高い音では、蝸牛殻の硬い中心部が振え、低い周波数に対しては弾性がある外側の部分が振動する。
- 14) “Saccade”現象と呼ばれている。
- 15) 歌の中のあるフレーズを思い出そうと思ってもすぐに浮かばないとき、結局最初から順番に歌詞を歌わなければならない場合や、具体的な状況を思い浮かべないと思いつけない出来事も多いことなどは、この記憶のあり方を裏付ける。
- 16) もちろん最適性理論では規則ではなく制約の順序付けによるが、基底形と音声表示の考え方は同じである。

### 参考文献

- Clements, G. N. (1985) “The Geometry of Phonological Features,” *Phonology Yearbook* 2:225-252.
- Chomsky, Noam (1965) *Aspects of the Theory of Syntax*, The MIT Press.
- Chomsky, Noam (2000) *New Horizons in the Study of Language and Mind*, Cambridge University Press.
- Chomsky, Noam (2004) “Language and Mind: Current Thoughts on Ancient Problems,” in Lyle Jenkins, (ed.). *Variation and Universals in Bilingualistics*. Elsevier (North-Holland Linguistic Series: vol. 62): 379-405.
- Chomsky, Noam (2005) “Three Factors in Language Design,” *Linguistic Inquiry* 36:1:1-22.
- Chomsky, Noam and Morris Halle (1968) *The Sound Pattern of English*, Harper and Row.
- 藤田耕司 (2005) 「言語を言語たらしめるもの」『月刊言語』大修館書店 2005-34:30-37.
- Halle, Morris (1983) “On Distinctive Features and their Articulatory Implementation” *Natural Language and Linguistic Theory* 1: 91-105. (Halle 2002 に再録)
- Halle, Morris (2002) *From Memory to Speech and Back: Papers on Phonetics and Phonology 1954-2002*. Mouton de Gruyter.
- Halle, Morris and Kenneth N. Stevens (1964) “Speech Recognition: A Model and a Program for Research,” in J. A. Fodor and J. J. Katz (eds.), *The Structure of Language: Readings in the Philosophy of Language*. Prentice-Hall, 604-12. (Halle, 2002 に再録) .
- Halle, Morris and Kenneth N. Stevens (1991) “Knowledge of Language and the Sounds of Speech,” in Halle (2002).
- Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch (2002) “The Faculty of Language: What Is it, Who Has it, and How Did it Evolve?” *Science* 298: 1569-1578.
- Hawkins, Jeff and S. Blakeslee (2004) *On Intelligence*. Times Books.
- 北川善久・上山あゆみ (2004) 『生成文法の考え方』研究社
- Minsky, Marvin (1985) *The Society of Mind*, Simon & Schuster, Inc.
- 中島平三・池内正幸 (2005) 『明日に架ける生成文法』開拓社
- Pinker, Steven and Ray Jackendoff (2005) “The Faculty of Language: What’s Special about it?” *Cognition* 95: 201-236. (www.elsevier.com/COGNIT).
- Sapir, E. (1925) “Sound patterns in Language.” *Language*, 1, 37-51.
- Yang, Charles D. (2002) *Knowledge and Learning in Natural Language*. Oxford University press.
- 養老猛司 (1989) 『唯脳論』 青土社

( 三宅正隆, 立命館大学国際関係学部教授 )

## Phonological Studies from Bilingualistic Point of View

Recently, more and more attention is being directed toward the inquiry into the question of where the human language faculty is situated in the natural world. Research in this area considers what kind of biological system our language faculty is and how it relates to our other biological systems, leaving behind the concerns about language-internal empirical studies. The major objective of language-internal empirical studies has been to characterize the fairly narrow class of grammars that are available to language learners so that linguists can specify the notion of “human language” in a narrow and restrictive fashion. The recent new trend of theoretical linguistics, however, has abandoned most of the formal restrictions on language faculty proposed in the past. Instead, studies on language external biological systems such as sensory motor systems have become a central interest in the field.

In this essay, I have focused on the scope of studies regarding one of the two interface levels of the mind/brain systems, i.e., the level related to sound, and have pointed out how the rich results of traditional investigations regarding phonological/phonetic phenomena can be incorporated into new trends of linguistics concerns. In so doing, I have reviewed the analysis-by-synthesis model proposed by Halle and Stevens (1962) and have shown how new perspectives from neuroscience such as that by Hawkins and Blakeslee (2004) can offer new insight and contribute to forwarding our investigation on human language faculty.

( MIYAKE, Masataka, Professor, College of International Relations, Ritsumeikan University )

