

言語獲得のモデル



櫻井 彰人・酒井 邦嘉

1. まえがき

幼児は素晴らしい速度で言語を獲得する。1歳頃の片言のお喋りの可愛らしさに喜んでいるのもつかの間、4歳頃の一人前のお喋りに辟易することになる。この言語獲得の秘密を探りたくなるのは当然のことである。これまで、言語学や発達心理学、自然言語処理などの分野で非常に多くの研究がなされてきたが、まだ手つかずの問題が山積している。

言語獲得の問題の核心に、言語知識 (linguistic knowledge) がある。言語知識の3要素は、統語論 (syntax)・意味論 (semantics)・音韻論 (phonology) である。幼児はこれらを並行して獲得していく。並行して獲得するのが獲得しやすいからであろうが、それは3要素が不可分であることに起因している。しかし、言語獲得の機構を解明するには、これらを同時に考えるのは至難であり、いずれかに特化した研究が進められているのが現状である。本稿では、統語論、すなわち文法の獲得に力点をおくことにする。もちろん、文法の獲得といえども、機能語や活用規則を獲得するためには、音韻や意味の獲得から完全に切り離すことはできない。

脳における言語機能は、複雑かつ精緻なシステムによって支えられている。従って、言語機能に関わる個々の神経活動だけを追っていてもシステム全体の仕組みは分からないし、外部からシステム全体の活動のみを追ってはいそれを支える機構が分からない。言語の脳研究では、さまざまな方向からのアプローチの融合が特に必要である。本稿では、言語獲得機能に関するモデルを概観して、特に脳内のモジュール構造に対する示唆が得られるようなモデルに注意を向けることにする。現在われわれが取り組んでいる目標は、言語獲得機構のモデル化を行い、言語獲得のために自己組織化しう

る脳のモジュール、すなわち「言語獲得装置 (language acquisition device, LAD)」のメカニズムを明らかにすることである。

2. 幼児の言語獲得の特徴

ここでは、対象となる言語を母語に限定して、言語獲得とは6歳頃の臨界期以前に行われる母語の獲得を指すものとする。第2言語の習得にも6歳頃に臨界期があるが、獲得の過程や方法だけでなく、言語の脳機能の点でも、臨界期を境にして大きく変化することが報告されている¹⁾。実際、中学校の学齢以上で行われる第2言語の獲得では、語義は母語の翻訳として教示され、文法は明示的な規則として (例えば3人称・単数・現在形で一般動詞の語尾に「s」がつくというように) 教わることになる。母語であれば、このような明示的な規則を使わなくとも、直感に頼って文の文法性を判定できるので、両者の違いは明白である。幼児期に単語と事物の対応を比較的丁寧な教わることを除けば、文法や語義は、自分が生活している文脈の中で自然と獲得されていく。

言語獲得の素晴らしさは、その速度や精妙さだけでなく、文法的に不完全な文のデータを入力として、自然言語の特徴を完璧に備えた言語を獲得できることにある。Bickertonが行った詳細な研究によれば、文法規則のない単語の羅列である「ピジン」語が話されている環境で育った子供たちは、複雑な文法を備えた自然言語 (これを「クレオール」という) を生み出すことができる²⁾。言語獲得がほとんど自動的に行われることを考えれば、それを「本能」と呼んで差し支えないのであるが³⁾、そもそも入力データなしには言語獲得が成立しないのも真実である。不幸にして適切な養

育を放棄されて、全く言語的環境（音声や手話）がないところで育った幼児は、言語を獲得（生成）できなかった、という事例がいくつか報告されている⁴⁾。健聴者の両親の間に耳の聞こえない子どもが生まれた場合や子どもが言語獲得以前に失聴した場合には、両親も手話を習い、また出来るだけ早期に手話環境に接する機会を与え、子供に手話を母語として獲得させなくてはならない^{5,6)}。もちろん手話は自然言語であって、複雑な文法を駆使することができる。しかしながら、日本では口話法（読唇術）を奨励して手話を軽視する立場があり、手話が正しく認知されているとは言い難い⁶⁾。なおここでいう手話は、日本手話（JSL）やアメリカ手話（ASL）と呼ばれる言語を指し、日本の手話教室などで教わる機会の多い日本語対応手話ではない。後者は、基本的に日本語文中の単語を手話単語に置き換えただけのものであり、ピジンに相当する。

幼児に対して、文法的に誤った文が明示的に与えられること（これを否定証拠（negative evidence）という）はないとされてきたが、次節で説明するように、先験的知識や否定証拠なしに、自然言語を獲得するのは原理的に不可能である。そこで、文法に関する生得的な知識が、乳幼児の脳に存在すると仮定しなくてはならない。ただし、否定証拠が親から暗黙的な手がかりとして与えられる可能性もあるので、言語獲得に対する制約条件の中で生得的な要因が占める割合に関しては、依然として議論の余地がある⁷⁾。そこで、ある言語の文法規則を規定するためには、どのような例文をどの位の数だけ与えればよいか、ということが現実的な問題となってくる。一方、直接的な経験がほとんどないにもかかわらず、いかにして文法を獲得できるのかが、生得的な言語知識の構造を解き明かすための、重要な論点である⁸⁾。また、乳幼児が生得的な能力によって母語の音韻を獲得するにしても、言語獲得には母親語（motherese）の影響が大きいと考えられる。母親語は、発音が明瞭で韻律の変化が豊かであり、文法的かつ短文が多いことが確認されている。それがどのように幼児の言語獲得に影響しているかは、まだよく分かっていない⁹⁾。特に、母親語が文法獲得に与える影響について、今後の研究で明らかにしていく必要がある。

言語獲得をモデル化するにあたって、言語学では、そもそも言語は「獲得」されるものであって、刺激と反応の連合のような「学習」とは厳密に区別していることを強調しておきたい¹⁾。情報科学では、「学習」を

非常に広い意味に用いている。あるシステムの不完全な挙動データを元にして、その動作原理を推定したり、それと同一動作をするシステムを構成したりといった、対象システムのモデル化を陰に陽に含むプロセスを、先験的知識の多寡やアルゴリズムによらずに、「学習」または「機械学習」と呼んでいる。また、「獲得」は「知識獲得」としてエキスパート・システムの文脈で用いられるが、専門家の知識を明示的にして計算機に移植することを指している。以下では、人間は言語獲得をし、モデルは機械学習をするという用い方をする。

3. 言語獲得モデルへの最初の取り組み

Chomsky は、生成文法理論の構築当初より、個別言語を習得するための生得的な能力が人間には備わっていると¹⁰⁾。行動主義に反対しての立論であったが、後の Gold らによる計算論的な研究から明らかになるように、それは正しいものであった。

ある個別文法に従う正しい文や従わない誤った文が次々に与えられる時、その文法を推定することを文法推論（grammatical inference）という。情報科学における文法推論の問題は、きっかけは言語獲得の形式化であったが、初期は計算論的興味から、そして機械学習の枠組みとしての興味とパターン認識への応用の観点から、最近は大量のデータ内の規則性を抽出する応用（生物情報処理を含む）への観点から、着実に研究成果が上がっている¹¹⁾。

Gold は、文法推論の課題を次のように定式化した¹²⁾。生成規則で記述されるようなある個別文法 G 、 G から生成しうる有限長文字列（正例と呼ぶ）全ての集合 P 、 G から決して生成されない有限長文字列（負例と呼ぶ）全ての集合 N 、 P と N の和集合の部分集合 E 、 E のどの要素もいくつかは現れるような無限列 s_1, s_2, \dots を考える（ s_i は E の要素であって、それが正例なのか負例なのか付記されているものとする）。文法推論とは、 s_1, s_2, \dots, s_i を見た時、 G を目標とする推論を行い、それが G_i であると推定する過程であり、ある n 以上の i に対して常に $G_i \cong G$ となれば（文法表現が異なっても生成される文集合が同一であればよく、これを極限同定という）、この文法推論は成功したとする。この入出力動作をする装置が文法推論する装置 M である。

Gold は、 G がどんな文法であれ（それがあろうプログラムで等価に記述できるものであれば）、 $E = P \cup N$

(これを完全提示という)であれば文法 G を正しく極限同定できる M が存在することを示した¹²⁾。しかし、 $E = P$ の時には、正規言語(正規文法で記述される言語)でさえ極限同定できないという結論を導いた。自然言語は正規文法では記述できないゆえ、この結果は「自然言語は正例のみからでは獲得できない」ということを意味する。なお、 E を P より大きくして N の一部を含むようにしても事態はあまり変わらない。

この困難を解消する方法はいくつか考えられる。第1は、生得的な機構により、人間が生成しうる文法クラスが限定されるため、言語獲得が可能になる、というものである。第2は、言語獲得時に用いる情報として、言語データ以外にも母親や周囲の人との相互作用を通じて、自分の言ったことが文法的に正しいか否かに関する手がかりを得る、というものである。多くの言語学者は第1の立場をとっているが、前節で議論したように、どこまでが生得的な機構で説明できて、どこまで言語特有なのか、といった点が未解決である。また、第2の点は実際の言語獲得時には有効に働いていると予想されるが、必須であるとは断定できない。周囲との相互作用が制限された環境で子供を育ててみて、言語獲得がどのように影響されるかを調べるなどできない相談だから、これは乳幼児の行動実験だけでは決着がつかない問題である。

4. 記号的情報の獲得

乳児は、6ヶ月までに母語の音声の特徴を認識できるようになることが、Kuhl らによる母音の識別の実験で明らかにされた¹³⁾。さらに12ヶ月までには、外国語で必要となるような識別能力を失って、母語に特化した識別だけが可能になる¹⁴⁾。このような乳児の驚くべき言語能力は、音声の識別に限られるのではなく、7ヶ月の乳児は音声入力に潜在する規則性までも認識できることが、最近明らかにされた¹⁵⁾。ABA という音声パターン (A, B には一音節を割当てる) を2分間聞かせた乳児は、 A, B に割当てられた音節とは独立に、 ABA と ABB という二つのパターンを区別できた。この実験は、「選好注視法」といって、乳児が新しいパターンが提示される方向を好んで見るという性質を利用している。神経回路網によるモデルが、このような「人工文法」を習得できるかどうかに関しては、議論の分かれるところである¹⁶⁾。人工文法を用いた成人での行動実験では、有限状態オートマトンから生成

された音や文字の列(正規文法に従う)を予め提示しておく、刺激の種類が異なっても、同一のオートマトンから出力されたものか否かを判定できたという報告がある¹⁷⁾。

言語知識は、上の例のような人工文法よりもはるかに複雑であるが、記号で記述されうることには変わりはない。文法の獲得には、大別して2種類の記号化過程があると考えられる。第1に単語の統語範疇(名詞・動詞など)を獲得する過程であり、第2にそれらを用いた統語規則を獲得する過程である。言語獲得においては、2語文・3語文の段階で、すでに名詞・動詞の構文的な区別が明確になっている。統語範疇の獲得に生得的要素が深く関与していることを示す証拠として、次のような報告がある。健聴者を両親に持ち、手話に接したことがない先天聾の幼児(3歳から4歳)のジェスチャーには、母親のジェスチャーには見られない、自然言語の特徴が観察された¹⁸⁾。この特徴は、「能格(ergative)」と呼ばれ、他動詞目的語と自動詞主語を同じ「格(case)」として扱い、他動詞主語と区別するものである。能格言語として、エスキモー語などが知られている。この実験はアメリカと台湾で行われて同様の結果が得られたが、英語や中国語は、他動詞主語と自動詞主語を同様に主格として扱い、他動詞目的語と区別する言語(日本語も同様)である。従って、能格を持つジェスチャーは、母国語や文化の環境には関係なく、自発的かつ生得的なものだと結論できる。

次に問題となるのは、統語範疇を決める際にほとんど手がかりが存在しないという点である。文における位置やアクセントなどは、名詞や動詞を規定する手がかりとして絶対的なものではない。統語範疇は統語規則なくして定まるものではないことを示唆している。ところが統語範疇を知らずに統語規則を獲得することはできない。このジレンマは、bootstrapping 問題と呼ばれている。生得説の立場に立ってこの問題を回避するため、意味的關係を手がかりとすれば統語範疇の獲得が可能になるという仮説が、Pinker らによって提唱されている¹⁹⁾。名詞・動詞の範疇が獲得されれば、それが統語規則の獲得を助け、得られた統語的知識が統語範疇の獲得と精度の向上を助けることになる。しかし、実際に動作する計算機モデルは、まだ構築されていない。統語範疇が獲得されているという仮定のもとで、統語規則を獲得する計算機モデルの研究が行われているのが現状である。統語範疇を表す記号の列が順次与えられた時、それらの背後にある文法を推定す

ることが課題となる。しかし、理論的困難さは前述の文法推論と変わらない。

Chomsky は、自然言語の文法クラスが有限個のパラメータ（各パラメータがとる値の種類は有限個）の組み合わせで決定されるとした。例えば、主語（S）・述語（V）・目的語（O）の語順をパラメータと考えれば、日本語は SOV であり、英語は SVO である。文法の獲得は、言語データの中からこのパラメータを決定する証拠（これを trigger という）を見出し、パラメータを設定していく過程である。主要なパラメータを用いることで、統語範疇の列を入力として、実際に trigger によって文法が獲得されるか否かが検討されている²⁰⁾。特に trigger を想定しなくとも、パラメータの変更を確率的に行えば、有限マルコフ鎖のモデルを作ることができる。有限マルコフ鎖の性質から、入力に時々誤りがあっても確率 1 で正しい文法に収束することになる²¹⁾。

文法推論の研究においては、入力文字列にその構文木 (derivation tree) をヒントとして与えた場合に、文脈自由文法が獲得できるかが検討されている。構文木には統語範疇は与えられていないので、自然言語文法の獲得に引き写して言えば、文に節や句のまとまりを（まとまっているという事実のみを）ヒントとして与えたら、統語法の獲得は促進されるか、という研究である。この場合には文法の表現が限定されるが、文脈自由言語が機械学習できることが知られている²²⁾。

5. 統計的情報の獲得

人間は統計的性質を敏感に獲得する能力を生得的に持っていて、連続した音声入力を単語に分割する際にこの能力を活用している可能性が、8ヶ月の乳児を対象とした行動実験で示されている²³⁾。文法推論の研究においても、統計的情報の利用が考えられており、各生成規則に、それが適用される確率を定めるモデルがある²⁴⁾。生成規則やその確率を言語データから最尤推定すればよいので、文法推論時に入力データに誤りを許すことができるようになる。しかし、この確率分布の推定は、対象文法クラスが文脈自由文法の場合でさえ、非常に困難な問題となる²⁵⁾。確率の推定を行うときに反復計算を行うのであるが、最適値とは大きく異なる局所最適値への収束が頻繁に発生するからである。なお、対象を正規文法とすると、確率分布の推定が現実的となる。音韻規則は確率的正規文法でうまくモデル化でき

るため、音声認識には隠れマルコフ・モデル (hidden Markov model, HMM) が使われている^{25), 26)}。

最適な統計モデルを選択する際の基準の1つとして、最小記述長 (minimum description length, MDL) 原理がある。これは、観測された言語データを規則によって記述する時、規則自身を何らかの方法によって記述した時の記述の長さ（記述長）と、その規則による各言語データの記述長の総和が、最小になるように規則を定めよ、という原理である²⁷⁾。Marcken は最小記述長原理を援用し、単語分割されていない実際の英語や中国語の文字列から、教師なし学習により、語境界を好成绩で学習するのに成功している²⁸⁾。

6. 人工神経回路網によるモデル

以上の議論では、モデルといえば、それは記号操作を得意とするシステム (Turing 機械、現在の計算機) の上に構築されたものを指していた。では、より基礎的なモデルである人工神経回路網の上に言語獲得モデルを構築することは可能であろうか。現状では残念ながら極めて限定したものしかできていない。原因の一つは、まだ計算機に比べれば遥かに小規模なものしか構築できないことである*1)。そして、その簡単な人工神経回路網では（明示的にせよ暗黙的にせよ）記号処理が苦手であるということにある。

Rumelhart と McClelland は、入力された英語の動詞をその過去形に変換する、簡単だが精妙なモデルを作った³¹⁾。語数の増加に沿って、まず規則動詞と不規則動詞の両方が学習され、次に、ほとんどの動詞に“ed”をつけてしまう過一般化 (over-generalization) が起こり、そして過一般化が解消する過程が発生することを示した。このモデルの欠点は、規則変化に沿った過去形を得る規則を未知の動詞に対して適用することができない、という点である。実はこの現象（仮に「変数化の困難さ」と呼ぼう）は、人工神経回路網ではごく普通に起こることであるので別項を設けて議論する。なお、言語学からの批判に答えて、引続き人工神経回路網の研究がなされている³²⁾。古期英語における動詞の過去形については、Hare らが規則動詞間には音韻的類似性が非常に少ないことを利用して、未知の

*1) しかし、もし無限精度の計算を許すなら、単純な再帰型回路網は Turing 機械の計算能力を有する²⁹⁾。しかし、計算誤差・雑音がある場合には有限状態オートマトンまたはそれ以下の計算能力となる³⁰⁾。

動詞に適用される規則変化規則が学習可能であることを示している³³⁾。

Elman は、単純再帰型回路網 (simple recurrent network, SRN) を用いて、関係節生成を含む簡単な文脈自由文法で生成される単語列を対象とし、単語列中の単語が一つずつ入力される時、次に来るべき単語を予測するような機械学習を行った³⁴⁾。この結果、埋め込み文を超えた主語述語の一致が学習された。この場合にも、単語列中の個別の単語間の関連性が学習できただけで、単語範疇間の関係が獲得されたわけではなく、ましてや (埋め込み文を生成し得る) 再帰的な規則が学習できたわけではない。なお、少し観点を改めて、再帰型回路網を用いて記号を取り扱う研究が行われている (Adamson と Damper が現状をまとめている³⁵⁾)。

7. 変数化の困難さ

機械学習は、記号処理を用いるにせよ人工神経回路網を用いるにせよ、与えられた有限個のサンプルの裏に潜む規則性を推定すること、すなわち具体的な事例の表現からより一般的な表現を得ることを目的とする。

記号処理であれば、 $\text{past}(\text{"walk"}) = \text{"walked"}$, $\text{past}(\text{"show"}) = \text{"showed"}$, ... といった事例集合から $\forall x [\text{past}(x) = x + \text{"ed"}]$ とする。このとき新語 zork に対し $\text{past}(\text{"zork"}) = \text{"zorked"}$ と推論する。これは、定数 (例えば "walk") を変数 (x) に置き換え、それをさらに全称限量化 (\forall の導入) することで実現できる。これは、記号処理ならば簡単にできる、機械学習ではごく基本的な一般化の一手順である (これを「変数化による一般化」と呼ぶことにする)。

人工神経回路網では、これとは異なる一般化を行う。「類似による一般化」である。例えば、新語 wulk に対しては、walk の類似性から $\text{past}(\text{"wulk"}) = \text{"wulked"}$ とする。しかし、類似事例のない zork に対しては答えようがない。

変数化によれば、一般性のある規則を構成することができる。しかし、過一般化に陥る可能性は高い^{*2)}。それを避けるためには、全称限量化する変数の変域を適切に制限する必要がある。例えば上例では、 $\forall x [x \text{ is a regular verb} \rightarrow \text{past}(x) = x + \text{"ed"}]$ とすべきである。制限するためには、負例や前提知識を必要とするだ

*2) 正例のみから文法推論することの困難さは、一般化を適当なところでとどめることの困難さに由来する。

けでなく、新たな記述・概念 (上例では regular verb) を発見する必要がある。これは、帰納論理プログラミング (inductive logic programming) の補助述語発見に相当するものであり、大きな研究課題である³⁶⁾。

類似によれば、入力データに統計性があるとの前提のもと、正例のみを用いて正しい一般化をしよう。しかし、分布に対する前提知識を必要とし、それに基づく、類似性の適切な計量が必要である。また、MDL 原理を用いたとしても、文法推論のように抽象度の高い構造を推定することはできない。

これら2つの一般化手法に優劣があるわけではない。適用すべき場所が異なるだけである。しかし、文法推論に要求される一般化は類似によるものではなく、これが人工神経回路網による言語獲得のモデル化を困難にしている。解決の試みは数多くなされているが、未だ決定打はない³⁷⁾。

8. モデル研究からの示唆

音韻レベルの規則性は、音声そのものまたは人間の発声機構 (手話であれば動作・位置・表情などまたは人間の身体構造) に起因する制約に依存するため、入出力表現と直接関わっている。従って、その記述には統計的モデルが適し、その獲得には入出力表現に依存したボトムアップの推定機構と、類似による一般化が重要な役割を果たす。これに対して、統語レベルの規則性は、入出力表現とは独立のものが多く。従って、その獲得には、入出力表現に依存しないトップダウンの推定機構と、変数化による一般化が重要な役割を果たす。以上のことから、音韻の規則性を獲得する機構と、統語の規則性を獲得する機構との間に、本質的な違いが存在することが示唆される。言い換えると、音韻処理・獲得のモジュールと統語処理・獲得のモジュールが、独立して存在すると考えられる。

9. 言語機能のモジュール性

「言語機能は、他の認知機能と独立している」という意味のモジュール性は、広く受け入れられている。視覚や聴覚などが独立した機能モジュールであるのと同様に、言語もまたモジュールである。Chomsky の理論を中心とした言語学の立場は、言語知識自体が異なるモジュールに分けられることを主張してきた。冒頭で述べたように、言語知識のモジュールとは、統語論・

意味論・音韻論である。これらのモジュールが司る言語機能は、言語の情報処理に特化した極めてユニークな性質のものであり、一般の認知機構によって置き換えることはできない。言語獲得もまた、一般的な刺激と反応の連合メカニズムでは説明できない。

ブローカ失語の患者が示す失文法の症状は、文法機能がモジュール性を示す証拠として議論されてきた。前頭葉のブローカ野と統語処理との対応関係は、長く決着の付いていない重要な問題であったが、最近の脳機能イメージングで、統語モジュールとしてのブローカ野の働きが明らかになった³⁸⁾。その一方で、ウェルニッケ失語では言語理解に障害が現れることが知られている。この事実から考えると、側頭葉のウェルニッケ野が少なくとも意味処理や音韻処理に関わっていることは確かであろうが、意味論と音韻論の独立性を示すまでには至っていない。この辺りに、認知脳科学が解決していかななくてはならない課題が存在する³⁹⁾。本稿が、新しい言語研究の呼び水になることを念願している。

参考文献

- 1) 酒井邦嘉. 言語獲得の脳科学. 生体の科学, Vol.49, pp.40-53, 1998.
- 2) D. Bickerton. Pidgin and creole languages. *Scientific American*, Vol.249 (1), pp.108-115, 1983.
- 3) S. ピンカー. 言語を生み出す本能 (上, 下). 日本放送出版協会, 1995.
- 4) S. シャラー. 言葉のない世界に生きた男. 晶文社, 1993.
- 5) O. サックス. 手話の世界へ. 晶文社, 1996.
- 6) 特集 手話の世界. 言語, Vol.27, No.4, pp.20-101, 1998.
- 7) J.L. Sokolov and C.E. Snow. The changing role of negative evidence in theories of language development. In C. Gallaway and B.J. Richards, editors, *Input and Interaction in Language Acquisition*, pp.38-55. Cambridge University Press, 1994.
- 8) S. Crain. Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, Vol.14, pp.597-650, 1991.
- 9) E.L. Newport, H. Gleitman, and L.R. Gleitman. Mother, I'd rather do it myself: Some effects and noneffects of maternal speech style. In C.E. Snow and C.A. Ferguson, editors, *Talking to Children: Language Input and Acquisition*. Cambridge University Press, 1977.
- 10) N. Chomsky. *Aspects of the Theory of Syntax*. MIT Press, 1965.
- 11) Y. Sakakibara. Recent advances of grammatical inference. *Theoretical Computer Science*, Vol.185, pp.15-45, 1997.
- 12) E.M. Gold. Language identification in the limit. *Information and Control*, Vol.10, pp.447-474, 1967.
- 13) P.K. Kuhl. Learning and representation in speech and language. *Current Opinion in Neurobiology*, Vol.4, pp.812-822, 1994.
- 14) J.F. Werker and R.C. Tees. The organization and reorganization of human speech perception. *Annual Review of Neuroscience*, Vol.15, pp.377-402, 1992.
- 15) G.F. Marcus, et al. Rule learning by seven-month-old infants. *Science*, Vol.283, pp.77-80, 1999.
- 16) G.T.M. Altmann and Z. Dienes. Rule learning by seven-month-old infants and neural networks. *Science*, Vol.284, p.875a, 1999.
- 17) G.T.M. Altmann, Z. Dienes, and A. Goode. Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol.21, pp.899-912, 1995.
- 18) S. Goldin-Meadow and C. Mylander. Spontaneous sign systems created by deaf children in two cultures. *Nature*, Vol.391, pp.279-281, 1998.
- 19) S. Pinker. *Language Learnability and Language Development*. Harvard University Press, 1970.
- 20) E. Gibson and K. Wexler. Triggers. *Linguistic Inquiry*, Vol.25, pp.407-454, 1994.
- 21) P. Niyogi and R.C. Berwick. A language learning model for finite parameter spaces. In M.R. Brent, editor, *Computational Approaches to Language Acquisition*. MIT Press, 1996.
- 22) Y. Sakakibara. Efficient learning of context-free grammars from positive structural examples. *Information and Computation*, Vol.97, pp.23-60, 1992.
- 23) J.R. Saffran, R.N. Aslin, and E.L. Newport. Statistical learning by 8-month-old infants. *Science*, Vol.274, pp.1926-1928, 1996.
- 24) R.J. Solomonoff. A formal theory of inductive inference. Part II. *Information and Control*, Vol.7, pp.224-254, 1964.
- 25) C.D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.
- 26) 中川聖一. 確率モデルによる音声認識. 電子情報通信学会, 1988.
- 27) J. Rissanen. Modeling by shortest data description. *Automatica*, Vol.14, pp.465-471, 1978.
- 28) C. de Marcken. *Unsupervised Language Learning*. PhD thesis, Department of Electrical Engineering and Computer Science, MIT, 1996.
- 29) H.T. Siegelmann. *Neural Networks and Analog Computation: Beyond the Turing Limit*. Birkhäuser, 1999.
- 30) W. Maass and P. Orponen. On the effect of analog noise in discrete-time analog computations. *Neural Computation*, Vol.10, pp.1071-1095, 1998.
- 31) D.E. Rumelhart and J.L. McClelland. On learn-

ing the past tense of English verbs. In D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, editors, *Parallel Distributed Processing*, Vol.2. MIT Press, 1986.

- 32) F. Pulvermüller. On the matter of rules: Past-tense formation and its significance for cognitive neuroscience. *Network: Computation in Neural Systems*, Vol.9, pp.R1-R52, 1998.
- 33) M. Hare, J.L. Elman, and K.G. Daugherty. Default generalisation in connectionist networks. *Language and Cognitive Processes*, Vol.10, pp.601-630, 1995.
- 34) J.L. Elman. Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, Vol.7, pp.195-225, 1991.
- 35) M.J. Adamson and R.I. Damper. B-RAAM: A connectionist model which develops holistic internal representations of symbolic structures. *Connection Science*, Vol.11, pp.41-71, 1999.
- 36) 山本章博. 帰納論理プログラミングの基礎理論. 人工知能学会誌, Vol.12, pp.665-674, 1997.
- 37) A. Jogota *et al.*, editors. Connectionist symbol processing: Dead or alive? *Neural Computing Surveys*, Vol.2, pp.1-40, 1999.
- 38) D. Embick, A. Marantz, Y. Miyashita, and K.L. Sakai. A Syntactic specialization for Broca's area. *Proceedings of the National Academy of Sciences USA*, in press.
- 39) 酒井邦嘉. 言語の認知脳科学. 生体の科学, Vol.49, pp.10-22, 1998.

(さくらい・あきと, 北陸先端科学技術大学院大学 知識科学研究科・科学技術振興事業団 戦略的基礎研究推進事業)

(さかい・くによし, 東京大学大学院 総合文化研究科
・科学技術振興事業団 戦略的基礎研究推進事業)