

## 幼児の新奇語カテゴリ化のモデル研究

日高 昇平・齋木 潤

It is well known that children can learn words rapidly. Recently, several studies suggested that knowledge about the relationship between vocabulary and perceptual objects works as a constraint to enable children to generalize novel words quickly. Many experimental results of novel word generalization task supported this hypothesis, but its mechanism remains unclear. In this study, we examined a past proposed model explaining its mechanism and showed that they could not simulate novel word generalization task well in certain conditions. Therefore, in stead of the previous model, we proposed a model that could learn optimal feature attention for specific prototype. Our proposed model works well even with a multidimensional vocabulary set including rich perceptual information that the past model could not work. It suggested that statistical learning could be powerful enough to solve feature selection problem even in noisy information source. Furthermore, it also suggested that the basis of children's word learning was prototype-specific feature attention.

Keywords: vocabulary acquisition (語彙獲得), novel word generalization (新奇語汎化), shape bias (形バイアス), feature selection (特徴選択), prototype model (プロトタイプモデル), linguistic cue (言語的手掛かり)

### 1. はじめに

幼児は生後 18ヶ月または 50 語前後の語彙獲得後から急速な語彙獲得速度を示すことが知られている (Schafer & Plunkett, 1998). 幼児はたった一度の提示で新奇語を参照する概念領域に対応付けられることが示唆されており, この現象を即時マッピングと呼ぶ (Heibeck & Markman, 1987). 即時マッピングは一見ごく当たり前の現象だが, 新奇語は常に多くの可能な解釈を含み参照領域の特定は不良設定問題となる. 例えば, 幼児にとっては未知の赤いボールと新奇語「ボール」が与えられた場面において, 新奇語「ボール」が「球形物体」または「赤い物体」のどちら (または他の無数の可能性) を参照するのか特定する事は不可能である.

A Model Study of Infants' Novel Word Categorization, by Shohei Hidaka (Graduate School of Informatics, Kyoto University) and Jun Saiki (Graduate School of Human and Environmental Studies, Kyoto University).

即時マッピングは幼児の急速な語彙獲得に対して重要な役割を果たすと考えられるが, どのように実現されるのだろうか. この問題提起に対し, 多くの発達心理学的実験を根拠として, 参照対象や語彙のありように関する事前知識を「制約」として利用することで, 語彙獲得の促進を行うという仮説が提案されている (Markman & Hutchinson, 1984; Landau, Smith, & Jones, 1988). このような制約知識は一般に新奇語汎化課題を用いて調べられる事が多い (Colunga & Smith, 2000a; Imai & Gentner, 1997; Landau et al., 1988; Markman & Hutchinson, 1984; Samuelson, 2002; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; Yoshida & Smith, 2003). 新奇語汎化課題では協力者はまず標準刺激として未知の物体を提示, 命名され, 次に提示される操作された複数の物体から選択, または標準刺激と選択刺激の類似度を答える事で新奇語の汎化パターンを調べる. このような幼児の心理実験から「事物全体バイアス」, 「分類

学的カテゴリバイアス」(Markman & Hutchinson, 1984) や「形バイアス」(Landau et al., 1988) などのバイアスを幼児が持つことが示唆され、これらの新奇語汎化バイアスは語彙獲得に関する制約知識を反映すると考えられている。本研究では特に物体の知覚的特徴に関する知識である形バイアスに焦点を当てて研究を進めた。

### 1.1 形バイアス

本研究では概念発達過程、特に語彙獲得における物体カテゴリの形成を主題に議論する。カテゴリ形成は物体が持つ無数の特徴の中から、カテゴリを記述するのに役立つ特徴を取捨選択する問題として定式化できる。幼児はある新奇物体に新奇語が付与されると、その新奇語を形が類似した物体に汎用することが知られており、形バイアスと呼ばれる(Landau et al., 1988)。本研究では幼児のカテゴリ化に顕著に見られる形バイアスの発現メカニズムをモデルにより検討することで幼児の概念発達過程を議論する。形バイアスの計算論的議論をする理由、利点として以下の3点が挙げられる。

- (1) 幼児が発達初期段階で獲得する多くは知覚的特徴により記述される基本カテゴリ (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) とよばれるものであり形, 材質, 固形性などの物体属性は幼児にとって決定的なカテゴリ形成要因である。
- (2) 形バイアスの発現は多くの文脈において発達心理学実験で調べられている現象であるので、課題、刺激、言語的文脈と選択的注意(バイアス)の関係が得られている (Imai & Gentner, 1997; Colunga & Smith, 2000a; Jones & Smith, 2002; Landau et al., 1988; Samuelson, 2002)。
- (3) 初期獲得名詞の知覚的類似性や可算/非可算名詞の用法などに関しての定量的知見が得られている。(Samuelson & Smith, 1999; Colunga & Smith, 2000b)

(2) に関して、例えば、幼児は固形性の物体 (solid objects) について形状の類似 (形バイアス) に、非固形性の物質 (nonsolid substance) を材質の類似に基づいて新奇語の汎用をする事が知られている (Colunga & Smith, 2000a; Imai & Gentner, 1997; Samuelson, 2002; Soja, Carey, & Spelke, 1991)。

幼児が固形性を考慮して存在を分類する理由として、物体/物質の存在論的知識 (Soja et al., 1991) や英語の統語クラスである可算/非可算名詞の学習 (Quine, 1960) などが提案されている。Quine (1960) によれば可算/非可算名詞の統語クラスの学習が先行し、それに伴い物体/物質の存在論的クラスが獲得されるが、Soja et al. (1991) は逆に生得的な物体/物質の存在論的知識が統語クラスの学習を促進している。これに対して Imai & Gentner (1997) は英語話者と、それに対して可算/非可算名詞の明確な区別を持たない日本語話者の新奇語汎化課題の成績を比較し検討した。実験は標準刺激に対し形類似、材質類似の二つの刺激から強制選択する課題であり、幼児の固形、非固形刺激に対する選択刺激の選好は形-固形性の知識を反映していると考えられる。

その実験において複雑な物体、非固形な物質に対しては日本語、英語話者共に形、材質バイアスをそれぞれ示し、Quine (1960) のような強い言語相対性の主張を棄却した。しかしそれらの中間と考えられる単純な物体に対しては英語話者が形バイアスを示したのに対して日本語話者はチャンスレベル (形、材質類似刺激の二肢強制課題) の反応であったため、統語的クラスの影響も示唆した。

このような形-固形性のバイアスは新奇語獲得に文脈特異的に表れ、新奇語獲得の重要な現象であり、この現象を統計的な語彙学習により説明するモデルが提案されている (Colunga & Smith, 2000a, 2000b; Samuelson, 2002)。しかし統計的学習によるモデルは発展途上であり、どのような学習が適するののかという議論は十分ではない。そこで統計的学習の立場から語彙学習に必要な計算原理を追究するために先行研究のモデルと本研究での提案モデルを比較し議論する。

### 1.2 学習モデル

統計的学習モデルでは、形バイアスは既に知識を持つ大人から与えられるラベル付与を経験することにより発現するとされ、生得的な知識は必要とされない。それゆえにモデルの論点は、主に学習環境 (大人のカテゴリ知識) がどのような性質のもので、それをどのような統計学習で抽出可能であるかという二点である。既に形バイアス獲得のモデルとして連続値ホップフィールドネットワークを用いたモデ

ル (Colunga & Smith, 2000a, 2000b; Samuelson, 2002) が提案されているが、それに対して本研究では事例を基本とするモデルを提案し、それらの学習メカニズムの違いを検討する。提案モデルの背景として、発達やカテゴリ学習の実験、モデル研究において事例を単位とする統計学習を支持する知見を以下で紹介する。

Smith et al. (2002) は個々の事例から個別のカテゴリそしてさらに上位のカテゴリへと体系的な知識が獲得されると主張しており、縦断的訓練実験からその仮説を支持する結果を得ている。例えば、彼らの仮説によれば「個別のボール」、「球形の物体」、「特定の形をなす物体」、「特定の特徴をもつ物体」のように帰納的に汎化の範囲が拡大することになる。また大人のカテゴリ学習の実験成績を説明するモデルとして ALCOVE (Kruschke, 1992) などの事例モデルが提案されている。ALCOVE は特徴次元に対する選択的注意のメカニズムを持ち、中間層に事例を表す放射状基底関数を持つ RBF (Radial Basis Function) ネットワークモデルの一種である。カテゴリ学習のモデルは大きく事例かプロトタイプを基にするモデルに分かれるが、プロトタイプは個々のカテゴリの代表点であるので本研究では広い意味でその両方を「事例的モデル」と呼ぶ。幼児の語彙獲得を実験系より大きな規模のカテゴリ学習とみなし、事例的モデルで幼児のカテゴリ学習が説明できれば既に得られているカテゴリ学習と連続的に接続できる事になる。

一方で事例ベースではなく、形バイアスを再現するモデルとして Colunga & Smith (2000a) や Samuelson (2002) は連続値ホップフィールド型のモデルを提案している。連続値ホップフィールドモデルのユニットは双方向に全結合し、確率的に動作が決定するのが特徴である。このモデルは事例やプロトタイプ単位ではなく、特徴空間全域を特徴次元ごとに更新する学習を行なう。このようなモデルを本研究では事例的モデルに対して「連続値ホップフィールドモデル」と呼ぶことにする。

本研究では既に提案されている連続値ホップフィールドモデルと新たに提案する事例的モデルを比較検討することで、どのような統計学習が妥当であるか議論する。学習方式は学習対象の構造と常に表裏一体であり、任意の入力情報に対して有効な学習アルゴリズムは存在しない (Wolpert & Macready,

1997) ので、学習対象も同時に検討する必要がある。

### 1.3 初期獲得語彙の性質

Samuelson & Smith (1999) はマッカーサー乳幼児言語発達質問紙 (MacArthur Communicative Development Inventory; MCDI) (Fenson, Dale, Reznice, Bate, Hartung, Pethick, & Reilly, 1993)<sup>1)</sup> に記載されている幼児の典型的な獲得名詞のリストの中から 'I', 'mother' などの人物の名詞を除く 312 の名詞について成人被験者による評定実験を行っている。評定は統語的クラス { 可算名詞/非可算名詞/曖昧な名詞 } と固形性 (solidity) { solid/non-solid/曖昧 } とカテゴリ共通の知覚的特徴 { 形/色/材質/曖昧 } について行われている。この実験によれば、幼児の名詞は大きく 2 つの群にわかれ、固形 (solid) で、可算名詞であり (e.g. pillow, flower), 形の類似性で名詞の指示対象が定義される名詞群と、非固形 (non-solid) で非可算名詞であり (e.g. snow, butter), 材質の類似性で名詞の指示対象が定義される名詞群である。以後これを形-固形性の語彙性質と呼ぶ。Colunga & Smith (2000a), Samuelson (2002) はこのような初期獲得語彙の統計的性質を単純化した仮定をモデルの学習コーパスに取り入れている (詳細は実験 1 を参照)。重要な点は「固形」であれば「形」が名詞カテゴリを定義しており「非固形」であれば「材質」が名詞カテゴリを定義しているという点で、モデルではこの非明示的規則を統計的学習により獲得し、新奇な対象に対しても応用するとされている。

### 1.4 モデルの論点

本研究では形バイアス獲得の問題を以下の 3 点で検討することで、カテゴリ形成の一般的な計算原理に必要な学習環境、学習メカニズムまたその関係を議論する。

- (1) 連続値ホップフィールドモデル、事例的モデルではどのような学習を行なうか、またその学習は形バイアス発現を説明するのに適しているか。
- (2) どのような初期獲得語彙の性質が形バイアス再現に必要なか、また十分か。
- (3) 限定された知覚次元 (形, 材質, 固形性) に

1) 18-30ヶ月の幼児の発話、理解などの言語発達を調査するため開発された用紙で、この時期に子供が獲得する語彙が記載されている。

についてのモデルでなく、より現実的なモデルに拡張可能か。

(1) 先行研究で提案されている連続値ホップフィールドモデルと事例的モデルそれぞれの学習・汎化特性を検討する事で単に統計学習というだけでなく、さらに深い議論をすることが目的である。

(2) 学習対象の特性は前述のように(1)と不可分で双対した問題であり、学習対象にどの程度の条件が必要かを議論する事で(1)同様に計算論的議論を行なう。

(3) 連続値ホップフィールドモデルでは知覚属性は形、材質、固形性に限定されているが、現実の幼児は無数の特徴の中から「形バイアス」を抽出しているはずである。それゆえ、モデルはより制約の少ない現実的物体の学習で形バイアスを抽出できるか拡張可能性を検討することは重要である。

以上の(1),(2)は実験1,2で連続値ホップフィールドモデルの追実験を行い検討を行なった。また、実験1,2で明らかになった先行研究の問題点を考慮して実験3では新たな事例的モデルを提案した。さらに実験4では提案モデルにおいて(3)に焦点を当てて検討を行なった。

## 2. 実験1

Samuelson & Smith (1999) は形-固形性の語彙性質を主張しているが、これを満たす条件下でも学習コーパスにはまだ自由度が残されているので、様々な学習コーパスの条件でモデルを評価し、学習メカニズムを探ることができる。そこで実験1では先行研究で提案されている特徴モデルの詳細なメカニズムを解明するために、学習コーパスの持つ統計的性質を3つの条件で統制し、新奇語汎化課題のシミュレーションを行なった。

まずは先行研究の連続値ホップフィールドモデルが学習する学習語彙について考えてみよう。語彙調査からの知見である形-固形性の語彙性質は「初期獲得名詞のうち可算名詞は固形で、特定の形で材質は多様であり、非可算名詞は非固形で、特定の材質で多様な形であるものが多い」事を示唆している。一方、連続値ホップフィールドモデルは数値実験の単純化のため「初期獲得名詞のうち可算名詞は特定の形を持ち任意の材質はであり、非可算名詞は特定の材質であり、任意の形である」と形-固形性の語彙性質を単純化した語彙モデルを用いている。この

語彙モデルは、特定の形を持つカップの材質は「任意」つまり「どんな材質へも等頻度でなりうる」事を仮定している。例えば、材質が皮である物体Xがカップまたはボール(共に固形で特定の形を持つ)のどちらか推量する事を考えてみれば、我々は皮製のボールを多く経験しているので物体Xはボールだろうと推測できる。しかし先行研究の仮定する世界では、ボールもカップも等頻度で皮製でも陶器製でも在り得るため判断がつかない(一様性)。また先行研究ではモデルが語彙を学習後に新奇語汎化課題で行い、モデルの汎化パターンを幼児の結果と比較している。モデルを現実に置き換えると、例えば、幼児が陶器やプラスチック製の「カップ」を経験し形類似性で定義される事を学習した場合に、新奇語汎化課題では既に経験した陶器やプラスチック製の刺激はテストされるが、未経験の皮製の刺激はテストされない。つまり課題でテストされる刺激は既に経験済みの特徴の未経験の組み合わせのみが提示される(既知性)。

つまり Samuelson (2002) のシミュレーションでは、学習する語彙カテゴリを規定しない特徴次元(固形物の材質、非固形物の形)が等頻度で任意に出現する一様性、また新奇語汎化課題で学習済みの特徴の組み合わせで刺激を作成する既知性が仮定されている。一様性と既知性は数値実験上の単純化による副産物であり本来意図されていないが、形-固形性の語彙性質以上に強い制約として働いている可能性もある。そこで、Samuelson (2002) のシミュレーションと同じ標準条件、未学習の特徴が実験刺激として出現する未知条件、可算名詞でも各カテゴリが材質特徴において固有の散らばりを持つ多様条件の3条件を設定してモデルシミュレーションを行い、モデルに必要な条件や学習メカニズムを検討した。

### 2.1 先行研究

Samuelson (2002) は次のような実験を行い、15~21ヶ月の幼児が固形性を基にして形バイアスを持つことを示し、また連続値ホップフィールドモデルでシミュレーションを行なっている。

(1) 実験者が標準刺激を提示し、その刺激に対して「This is my **wug**.」と未知の単語「**wug**」のラベル付けを可算/非可算名詞のどちらでもない、中立的な文法で行った。

(2) 2つの選択刺激（形類似刺激と材質類似刺激）を提示し「Can you get your wug?」とどちらの物体が「wug」であるかを幼児に強制選択させた。

2つの選択刺激は形、材質、固形性の特徴次元で統制されており、形類似または材質類似、固形または非固形の2×2の条件で行われた。固形物はビスケットやチョーク、非固形物はジェルやローションで作られた刺激が用いられた。

以上の実験結果から幼児が固形物に対しては形類似刺激を選択するというバイアスをもつことが示されたが、同時に非固形物に対しては形類似刺激を選択するという「未熟な形バイアス」が示された。3歳ぐらゐの幼児は、固形物、非固形物に対して異なる反応をするが (Colunga & Smith, 2000a), Samuelson (2002) の実験の2歳以下の幼児は形バイアスを過剰般化したと考えられている。本実験では Samuelson (2002) の追実験を行なうため2歳児を対象としたシミュレーションを行なう。

## 2.2 シミュレーション手続き

Samuelson (2002) のモデルと同様のニューラルネットワークを構築した。各ユニットが全てのユニットと結合を持つホップフィールド型のネットワークで、入力層、隠れ層、出力層を持ち Contrastive Hebbian Learning (CHL) (Movellan, 1990) で結合荷重を最適化する (学習、出力動作の詳細は付録Aを参照)。入力層は形状 (12次元)、材質 (12次元)、その他 (12次元)、固形性 (3次元)、統語 (3次元) の計42次元で構成され物体の知覚的、可算/非可算名詞カテゴリを表す。また物体に対応する語彙カテゴリを出力層に22次元、隠れ層に20次元のユニットを持つネットワークを構成した (図1)。ネットワークは後述する物体と語彙カテゴリを学習後、新奇語汎化課題を行い幼児の反応パターンと比較し議論を行なった。

### 2.2.1 先行研究の語彙モデル

Samuelson & Smith (1999) の知見 (表1) に基づき Samuelson (2002) は形、材質、その他、固形性を物体の表現としてネットワークの入力とし、本実験でも同様の符号化を行った。形類似、材質類似、その他のカテゴリ化がなされているものを、定義される次元は特定値 (0,1 の2値) で、他の次元を事例ごとに一様乱数 (0-1 の連続値) で値を生成した。

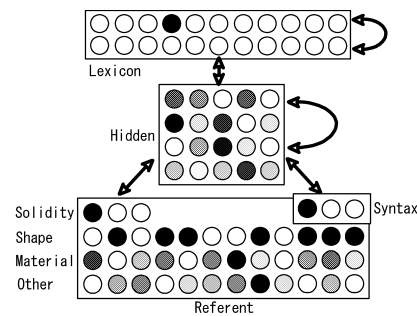


図1 Samuelson(2002) のモデル

例えば、形類似によりカテゴリが形成される語彙カテゴリの場合、形は「101101000101」のように12次元が固有の値で表現されるが、材質、その他の次元は提示事例ごとに一様乱数 (0-1 の連続値) により生成した値を割り当てた。また、物体の固形性は固形である場合「100」、非固形である場合「010」、曖昧の場合「001」で表現した。同様に可算/非可算の区別も、可算名詞なら「100」、非可算名詞なら「010」、曖昧なら「001」で表現した。22個の語彙カテゴリは出力層において局所的に表された (e.g. 「0010...00」)。

学習コーパス構成比率は (可算名詞、固形物、形類似カテゴリの割合)、先行研究同様 Samuelson & Smith (1999) のコーパス調査結果に従った (詳細は表1を参照)。Samuelson (2002) によれば、コーパス中の大部分の名詞は、形類似-固形-可算であり、この比率や相関が形バイアス獲得には重要であることが示唆されている (Samuelson, 2002)。

### 2.2.2 語彙モデルの条件

先行研究と同じ標準条件に加えて、新奇語汎化課題において未学習の特徴が提示される未知条件、名詞カテゴリを規定しない特徴次元で各カテゴリごとに固有の散らばりを持つ多様条件、の3条件を設定した。3条件の学習語彙の統計的性質を概念的に示したのが図2である。この図では形、材質、固形性がそれぞれ1次元で、その他を除いた計1次元で物体が表現され、2つの固形物体と2つの非固形物体のカテゴリの分布が白で表示されている。

#### 標準条件

標準条件 (図2a) はカテゴリを規定しない特徴次元がどのカテゴリも一様頻度で出現し (一様性)、新奇語汎化課題で提示される刺激 (図2の黒領域)

表1 Samuelson & Smith (1999) の調査による初期獲得語彙の固形性カテゴリ, 統語的カテゴリ, カテゴリ化の比率. 括弧の隣の数値はモデルでの値で括弧内は評定実験の結果である

	solidity classification			syntactic noun category			category organization		
	solid	nonsolid	ambiguous	count	mass	ambiguous	shape	material	ambiguous
Total	.64(.63)	.07(.04)	.29(.32)	.64(.74)	.14(.10)	.21(.16)	.50(.48)	.29(.16)	.21(.39)
solid				.79(.85)	.00(.02)	.22(.13)	.79(.61)	.00(.11)	.22(.31)
nonsolid				.00(.00)	1.0(.79)	.00(.00)	.00(.00)	1.0(.51)	.00(.39)
ambiguous				.50(.65)	.25(.16)	.25(.19)	.00(.30)	.75(.15)	.25(.50)
count noun	.79(.71)	.00(.00)	.22(.29)				.56(.59)	.11(.06)	.33(.38)
mass noun	.00(.13)	.50(.00)	.50(.52)				.00(.00)	1.0(.58)	.00(.42)
ambiguous	.67(.56)	.00(.06)	.33(.38)				.67(.28)	.00(.43)	.33(.39)

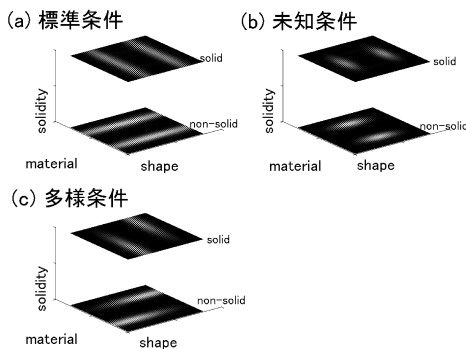


図2 名詞カテゴリの知覚特徴空間. 白領域から学習刺激, 黒領域から標準刺激が提示される.

は学習した特徴 (図2の白領域) と同じ範囲から提示され (既知性), Samuelson (2002) と同じ条件である.

#### 未知条件

一様性を持ち既知性がない学習を行なう条件で, カテゴリを規定しない特徴次元の値域を標準条件の1/5にした (図2b). 図2bのテスト刺激 (黒領域) が学習刺激 (白領域) より広いので未学習の特徴の組み合わせが課題で提示される.

#### 多様条件

既知性を持ち一様性がない学習を行なう条件で, カテゴリ特徴の分布の平均値が異なる (図2c). この条件では固形物でも材質次元でカテゴリ固有の範囲を持つ (白領域がずれている) のである程度のカテゴリの判別が可能である.

### 2.2.3 新奇語汎化課題

モデルでは学習後に新奇語汎化課題として, 未学習の標準刺激とその標準刺激と形, 材質が異なる選択刺激から強制選択する課題を行った. Samuelson (2002) と同じく固形, 非固形刺激の2条件があ

り, 未学習の標準刺激はそれぞれ固形次元が  $\text{solidity} = \text{solid}$ ,  $\text{non-solid}$  (図2(a), (b), (c)の黒で表された超平面. 実際には  $\text{solidity}$  は3次元で表される.) で固定され, その他の形, 材質などの次元は特徴は一樣乱数 (0-1の連続値) により生成された. 次に, 2つの選択刺激は標準刺激と形または材質以外は同じ特徴値を持つが, 形不一致刺激の場合は形次元のみ再び一樣乱数で特徴値を決定した (選択刺激は図2(a), (b), (c)の黒の超平面上で標準刺激から形, 材質軸に平行移動した点となる). 標準刺激と選択刺激 (形, 材質不一致) の刺激に対する出力ベクトル (語彙カテゴリ) のパターンを式(1)(2)を用いて形, 材質の選択率を算出した. ただし標準刺激, 選択刺激に対する出力ベクトル (語彙カテゴリ値) をそれぞれ  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  とし, 材質不一致刺激の選択確率  $P_{\text{shape}}$  (形類似選択) は式(2)のように標準刺激と材質, 形不一致選択刺激の類似度  $s_{\text{shape}}, s_{\text{material}}$  により決定された.  $a$  は距離から類似度に変換する際のスケージングを表すパラメータである.

$$s = \exp\left(-a\left(\sum_i (x_i - y_i)^2\right)^{\frac{1}{2}}\right) \quad (1)$$

$$P_{\text{shape}} = \frac{s_{\text{shape}}}{s_{\text{shape}} + s_{\text{material}}} \quad (2)$$

連続値ホップフィールドモデル ((Colunga & Smith, 2000a), (Samuelson, 2002)) では隠れ層ベクトルの距離を物体の心理的距離としているが, 心理的距離から類似度の算出法はそれぞれ異なっている. 本研究では類似度を心理的距離の指数関数とする Colunga の算出法を採用した. しかし先行研究と異なる点は, 本実験ではすべて隠れ層ベクトルではなく出力層ベクトル距離を評価に用いた点である. これは続く実験3, 4で行う異なるネットワークアーキテ

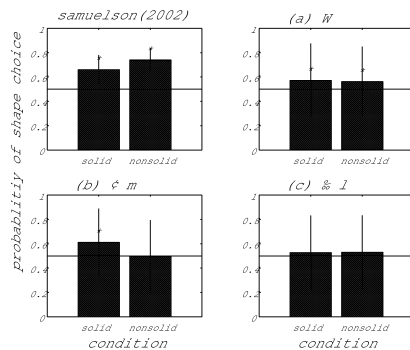


図3 形類似刺激選択率. チャンスレベル50%に対し有意に平均が大きい条件を\*で示した.

クチャのモデルと評価基準をそろえるためである.

### 2.3 結果・議論

3条件における100回(初期化の異なる4つのネットワークの各25回)の新奇語汎化課題の平均の形類似刺激選択率の結果を図3に示す. 各条件における平均値は全てチャンスレベル(50%)との $t$ 検定を行った. まず標準条件においてSamuelson(2002)の行動実験(図3左上), シミュレーションの結果と同様に固形, 非固形条件でチャンスレベルに対して有意に高かった(固形条件:  $t(99) = -2.57, p < 0.01$ , 非固形条件:  $t(99) = -1.68, p < 0.5$ ). 未知条件では固形条件では標準条件の固形条件のみ有意に高かった(固形条件:  $t(99) = -3.65, p < 0.001$ , 非固形条件:  $t(99) = -0.29, p = 0.38$ ). 多様条件では2条件とも有意差は見られなかった(固形条件:  $t(99) = -0.79, p = 0.22$ , 非固形条件:  $t(99) = -0.79, p = 0.22$ ). 以上の結果から, 標準条件に対し未知条件では固形条件の形バイアスは再現できたが, 非固形条件での形バイアスの汎化(未熟な形バイアス)は再現できず, また多様条件では固形, 非固形の2条件共にチャンスレベルとなり形バイアスを再現できなかった(図3右下).

語彙モデルの単純化を緩和した条件の形バイアスの再現性から, 連続値ホップフィールドモデルでは形-固形性の語彙性質以外の一様性, 既知性が形バイアスの再現に強く影響する事が示唆される. 続く実験2では3つの学習条件における詳細な汎化メカニズムを明らかにし実験1の考察を裏付けるために, 小規模な連続値ホップフィールドモデルによる実験を行った.

## 3. 実験2

実験1では連続値ホップフィールドモデルは語彙モデルに非常に敏感で, 形バイアス獲得メカニズムが形-固形性の語彙性質だけではなく, 一様性, 既知性などの制約に依存することが示唆された. 実験1では42次元の特徴で表された22の語彙カテゴリを学習したが, 実験2では個別カテゴリの汎化パターンを可視化して分析するため, 最低限の3次元の特徴で表された4つの語彙カテゴリを学習する. つまり実験2で用いる語彙モデルは形, 材質, 固形性の3次元で表された図2そのものである.

### 3.1 方法

ネットワークの入力は形, 材質, 固形性の各1次元計3次元からなり, 出力はカテゴリに対応する語彙が4次元で局所的( $(1, 0, 0, 0)$ など)に表現され, 隠れ層に8つのユニットを持つ. モデルの学習セットの各学習条件における語彙モデルをは表2に示した. 特徴次元数, カテゴリ数が非常に小さくなり, コーパスの構成比率(固形物: 非固形物 = 1:1)を変更した以外は各カテゴリの確率分布の性質は実験1と変わらず, この語彙モデルは実験1と同じ統計的性質を持つ. 特徴空間が3次元に限定されており, モデルの挙動が容易に可視化できるので, 特徴空間における新奇語汎化課題の標準刺激の汎化パターンを分析した. 新奇語汎化課題の際の固形, 非固形の標準刺激は固形領域( $solid = 0, 0 < shape < 1, 0 < material < 1$ ), 非固形領域( $solid = 1, 0 < shape < 1, 0 < material < 1$ )の2つの領域からそれぞれ一様乱数で生成した. それに対して形類似の選択刺激は材質次元のみ標準刺激と異なり乱数で再生成された. つまり標準刺激は固形平面上の任意の一点で表され, それに対し形, 材質次元軸に垂直に平行移動した点が選択刺激である.

### 3.2 結果と考察

モデルが新奇語汎化課題で提示される固形刺激は3次元の入力空間上の固形平面( $solid=1$ )の正方形領域に表されるので, それに対する出力パターンを見ることで直接モデルの汎化形式を分析した. 標準条件, 未知条件, 多様条件で新奇語汎化課題で生成された新奇な入力に対するネットワークの汎

表 2 実験 2, 3 で用いた語彙モデル.

各 2 つずつの固形カテゴリ (obj0,1), 非固形カテゴリ (sbc0,1) が形 (shape), 材質 (material), 固形性 (solidity) の 3 次元特徴で表現される. カテゴリは出力 4 次元の局所表現で表された. R は 0 から 1 の一様乱数で .4\*R は 0 から 0.4 の一様乱数を表す. 図 2 の入力パターンを実装したのものになる.

	出力	標準条件			未知条件			多様条件		
		shape	material	solidity	shape	material	solidity	shape	material	solidity
obj 0	1000	0	R	1	0	.4*R+.3	1	0	R-.3	1
obj 1	0100	1	R	1	1	.4*R+.3	1	1	R+.3	1
sbc 0	0010	R	0	0	.4*R+.3	0	0	R-.3	0	0
sbc 1	0001	R	1	0	.4*R+.3	1	0	R+.3	1	0

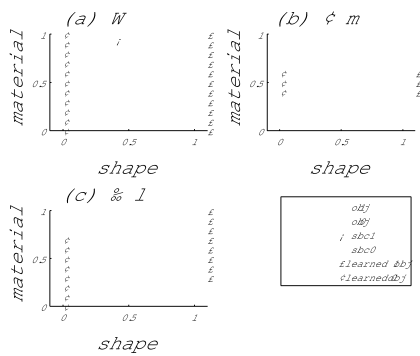


図 4 連続値ホップフィールドモデルの固形平面上的汎化パターン. shape 軸に (a) 直交, (b) 直交, (c) 斜交した汎化境界が見られる.

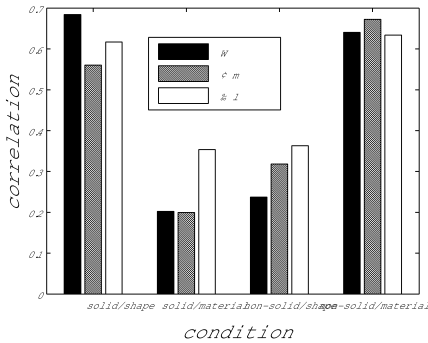


図 5 固形/非固形刺激の汎化パターンと形/材質次元との相関. 標準, 未知条件では形, 材質バイアスが見られるが, 多様条件ではバイアスが弱い.

化 (最大出力値を示したカテゴリ: obj0, obj1, sbc0, sbc1) を solid=1 の平面上の格子点 (10 × 10: .05 から .95 までの .1 間隔) にプロットした結果を表す (図 4). モデルの動作は確率的であるため, 図 4 は 1 つのネットワークの学習例における 100 個の標準

刺激に対する 30 回の最頻汎化パターンを示している. 提示された標準刺激 (図 4 の各点) に対して形類似刺激は形軸に平行移動, 材質類似刺激は材質方向に平行移動した点で表されるので, 出力パターンが形軸に平行に同じ出力パターンであれば標準刺激に対して形類似刺激の類似度が高まり, 選択確率が高くなる (形バイアス). 固形領域で obj0, obj1, 非固形領域で sbc0, sbc1 が対称的に学習され汎化パターンも同様であったので以下では固形領域のみを議論する. 標準条件・未知条件の固形領域 (solid=1) においてほぼ obj0 または obj1 の 2 カテゴリのどちらかに汎化し, 形次元に対して垂直に汎化の分離面 (obj1: ● と obj0: ○ の分離面) ができているので形バイアスが強いと考えられる. 一方, 多様条件では汎化分離面が形, 材質軸に対して斜めに発生し, 形バイアスが弱いと考えられる.

汎化分離面が形, 材質軸に対して垂直なパターンを持つか測るための定量的指標として汎化パターン (入力空間の形状, 材質座標値) と汎化カテゴリ値 (obj0, obj1, sbc0, sbc1) = (1, 2, 3, 4) の Kendall 順位相関<sup>2)</sup> の絶対値を用いて分析した (図 5). 例えば汎化カテゴリ値が汎化パターンの形次元に従い変化する場合, 形次元の順位相関の絶対値は高くなる (0 ≤ |R| ≤ 1) ので, 汎化カテゴリ値と形状との相関が高ければ形バイアスが強いことを意味する. 標準条件・未知条件では固形領域・非固形領域でそ

2) ある二つの組になった系列  $S_1, S_2$  があり, 系列  $i, j$  の対を  $O_{1ij} = \text{sign}(S_{1i} - S_{1j})$  と符号化する. ただし  $\text{sign}$  は正値なら 1, 負値なら -1, 0 なら 0 に写像する符号関数である. このとき同順位修正を行わない Kendall 順位相関係数は以下で表される

$$R_{12} = \frac{\sum_i \sum_j O_{1ij} O_{2ij}}{\sum_i \sum_j 1} \quad (3)$$

この場合  $S_1, S_2$  がそれぞれ特徴次元の座標値とカテゴリ値に対応している.



れぞれ形・材質次元に高い相関を持ち、形、材質バイアスを示すが、多様条件では固形領域でも材質次元に比較的高い相関を持ち、固形領域の形バイアス、非固形領域の材質バイアスが弱くなる事が示された。

標準条件や多様条件での汎化パターン(図4の左下)から特徴モデルは obj0,obj1 などの隣接するカテゴリから等距離の点にカテゴリ境界を作成する(ボロノイ分割)ように汎化パターンを決定するため、学習カテゴリの配置に対して敏感であると考えられる。ボロノイ分割は空間を分割する様々な問題で最適解として現れるので、モデルが学習サンプルをこの形式でカテゴリに分割するパターンを示すのは十分考えられる。しかし、その形式でカテゴリ分割し、かつ形バイアスの獲得を示すためには必然的に知覚的特徴空間上の配置に制約(つまり「一様性」)を求める必要がある。しかし前述のように一様性は現実的には必ずしも成り立たない条件であるので、特徴モデルに何らかの条件が必要である事が示唆される。そこで続く実験では新たに事例的モデルを提案し、本実験と対照して検討を行なった。

#### 4. 実験3

実験3では実験2で検討した連続値ホップフィールドモデルに対比して事例的モデルではどのような学習を行うか検討する。代表的な事例モデルのALCOVE (Kruschke, 1992)は特徴空間全体に対して明示的な特徴注意ユニットを持ち、成人協力者のカテゴリ学習過程をシミュレートできる事が知られている。しかし、特徴空間全体で距離基準が均一という点では連続値ホップフィールドモデルと同じであるので、これにSmith et al. (2002)の「事例からカテゴリに特化した特徴注意の学習をする」という仮説を取り入れてアーキテクチャを変更、拡張した。新たに提案するモデルはプロトタイプモデルで、個別のプロトタイプに対して特徴注意を学習することからPrototype-Specific Attention Learning (PSAL)モデルと呼ぶ。実験1,2で分析した連続値ホップフィールドモデルは「カテゴリ特異的に重要な特徴がある事」を仮定しないモデルであったため、特徴空間における個別カテゴリの分布よりも、隣接するカテゴリとの配置が汎化パターンに大きな役割を持っていた。一方PSALは「個別カテゴリの特徴に対し重み付ける」モデルであるので、

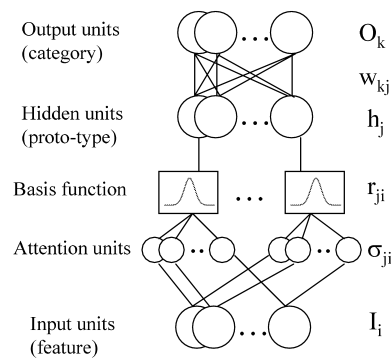


図6 PSALモデルのネットワークアーキテクチャ

隣接配置するカテゴリとの関係よりも、個別の特徴重み付けが汎化パターンに強い影響を持つと考えられる。これは、例えば、「個別のボール」からプロトタイプ(特徴空間上での基底位置)を学習し、「ボール」のプロトタイプに重要な特徴「球形」を個別に特徴注意学習で抽出することに対応する。

#### 4.1 方法

PSALモデルは入力層、隠れ層、出力層の3層からなる一般化RBFネットワークの一種で、隠れ層の各基底ユニットごとに注意ユニットを持ち、教師信号に対する誤差を最小化することで学習するモデルである。出力層ユニットの活性値 $o_k$ は隠れ層ユニットの活性値 $h_j$ の荷重線形和で、以下の式で表される。

$$o_k = \sum_j w_{kj} h_j \quad (4)$$

また、隠れ層ユニットの活性値 $h_j$ は入力層ユニットベクトル $I$ と基底ユニットベクトル $r_j$ の荷重付き( $\sigma_{ji}$ )ユークリッド距離で決定され、以下の式で表される。ただし $m$ は正の入力空間の距離基準パラメータで $c$ は定数のスケーリングパラメータである。

$$h_j = \exp\left(-c \left(\sum_i \sigma_{ji} (r_{ji} - I_i)^m\right)^{\frac{1}{m}}\right) \quad (5)$$

学習は教師信号と出力との誤差を逆伝播し、層間の荷重と注意ユニットの重み、隠れ層基底ベクトルを最適化する(詳細は付録Bを参照)。

提案モデルのベンチマークテストとして実験2と同じ入出力パターンを以て学習を行い、実験2と同様の汎化パターンの分析を行った。モデルの各パラメー

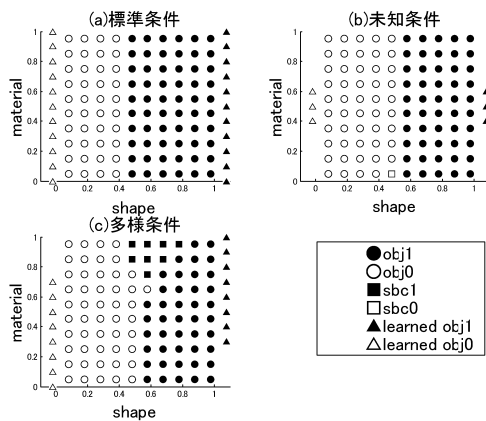


図7 固形平面上の PSAL モデルの汎化パターン

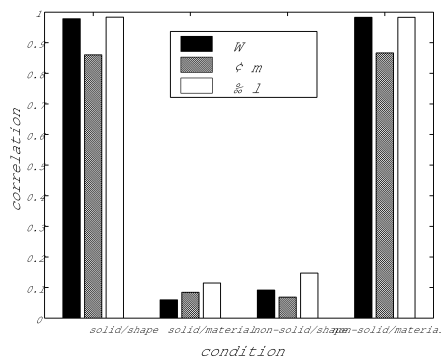


図8 PSAL モデルの相関パターン。

全ての条件で形・材質バイアスが見られた。

タは  $c = 1$ ,  $m = 1$ ,  $\lambda_w = 1$ ,  $\lambda_\sigma = 0.1$ ,  $\lambda_r = 0.1$  とし学習回数は 100 エポック 400 回だった (パラメータの詳細は付録 B を参照)。また基底ベクトルの初期位置は 4 つの入力カテゴリの平均位置とした。

#### 4.2 結果・考察

実験 2 同様に出力パターンと形・材質次元との相関の分析を行った (図 8)。その結果全ての条件で実験 2 の標準条件同様に「形バイアス」を獲得する結果となった。つまり、PSAL モデルは、連続値ホップフィールドモデルと異なり、カテゴリ配置に依存した境界面を形成せず、各カテゴリ特異的に分散の大きい特徴 (固形カテゴリの材質次元など) の重み付けを小さくするような汎化パターンを示すことが示唆された。では連続値ホップフィールドモデルと PSAL モデルは一体何が違うのだろうか。実は PSAL モデルでも特徴モデルと同じく「等距離分割」をして

カテゴリ境界を決定していると考えられるが、その「距離基準」が異なっている。PSAL モデル中の特徴空間の距離基準はプロトタイプに特異的な特徴注意によって決定されるため、カテゴリを「等距離」で分割しても異なる結果になると考えられる。

### 5. 実験 4

実験 1, 2, 3 の特徴選択問題では形, 材質, 固形性だけの限定した知覚的属性のみ物体属性として扱い, 既にある程度の「特徴選択」済みの状況であった。知覚的, 概念的, 言語的情報が混在する日常で統計的学習だけで物体のカテゴリ化に有効な特徴選択が可能だろうか。本実験では形のみではなく, 材質, 動き, 大きさなどの豊富な特徴を持つ知覚空間における特徴選択問題に対する PSAL モデルの有効性を検討することを目的としてシミュレーションを行った。さらに形訓練を受けた 18ヶ月児と通常の 33ヶ月児の新奇語汎化課題の結果と本実験の結果を比較し, モデル内で統語的カテゴリの有無による形バイアス変化の予測を行った。実験 4 で扱う「現実的な」コーパスは Hidaka & Saiki (2004) が行った初期獲得語彙に対する知覚属性の評定実験の結果を利用する。コーパスは Samuelson & Smith (1999) の調べた MCDI の名詞の一部 (312 語中の 48 語) に対して 16 対の形容詞を用いて調べられている。対象となる属性は形, 材質などの特定の属性のみに偏っていないので PSAL モデルの能力を試すのに適していると考えられる。

#### 5.1 方法

##### 5.1.1 学習する語彙モデル

Hidaka & Saiki (2004) は幼児が獲得する典型的な語彙について形容詞対に用いた Semantic Differential (SD) 法 (Osgood, Suci, & Tannenbaum, 1957) により語彙の心理的属性の定量化を行った。彼らは対象語彙を MCDI (Fenson et al., 1993) の 9 つの上位カテゴリ<sup>3)</sup>からほぼ同じ割合で 48 語選択し (詳細は付録 C を参照), それらについて主に知覚的属性を表現する 16 対 32 語の形容詞対 (図 10 を参照) を用いて日本語 (104 人), 英語 (20 人) 成人モノリンガルを協力者として調査を行った。評定は 1, 2, ..., 5 の 5 段階評価 (e.g. 非常に大きい, 大

3) 'animals', 'body parts', 'clothing', 'food and drink', 'furniture and rooms', 'outside things', 'small household items', 'toys', 'vehicles'

きい, どちらでもない, 小さい, 非常に小さい)で行われた。日本語と英語の結果は非常に類似しており, 翻訳や固有言語の問題はないと考えられ, 本実験では日本語のデータを利用した。Hidaka & Saiki (2004) は調査データの主成分分析から幼児の初期獲得語彙に対する主な物体属性は固形性, 動きなどの属性であることを示した。Hidaka & Saiki (2004) はこのような現象に対して調査結果を用いた計算論的モデルを用いて次のような現象を説明した。日本語話者は動物と非動物の中間にあたる刺激を動物とみなし分類するのに対して, 英語話者は非動物とみなし分類を行ない (Yoshida & Smith, 2003), また物体と物質の中間にあたる刺激に対しても両言語話者は異なる反応をする (Imai & Gentner, 1997)。彼らの計算論的モデルは調査結果を共通の知覚属性空間とし, 「いる」/「ある」などに見られる日本語の動物/非動物の名詞クラスと英語の可算/非可算名詞クラスが日本語話者, 英語話者それぞれの付加的な入力として働くことで言語差を説明した。Hidaka & Saiki (2004) のデータセットは物体を多次元の心理的属性で表現し, 固さ, 動きなど物体の複数の要因を含むので実験 1, 2, 3 で扱った人工コーパスに較べてより現実的なコーパスとして適していると考えられる。

### 5.1.2 学 習

PSAL モデルは入力層に物体属性 (16 次元), 隠れ層にはプロトタイプ (48 次元), 出力層には名詞カテゴリ (48 次元) を表すユニットで構成された。一回の学習には一つの名詞カテゴリが選ばれ, それに対応する無作為抽出された協力者個人の評定値 (形容詞次元) が入力として用いられ物体の属性は各次元 1-5 で表現された。調査データの 0.1% が欠損値であり, その値は 3 (評定値 1-5 の中央値) で代用された。また教師信号は選出された名詞カテゴリが  $(0, 1, 0, \dots, 0)$  のように局所的に表現された。

学習条件として, 単に名詞カテゴリを学習する条件と, 名詞カテゴリと可算/非可算名詞カテゴリを同時に学習する条件を設定し, 学習後に新奇語汎化課題の成績を評価した。統語あり条件の場合は出力層は Hidaka & Saiki (2004) に従い (付録 C を参照), 可算名詞, 非可算名詞のそれぞれを  $(1, 0)$ ,  $(0, 1)$  と符号化し出力層を  $48 + 2 = 50$  次元で構成した。モデルの各パラメータは  $c = 1.5$ ,  $m = 1$ ,

$\lambda_w = 0.15$ ,  $\lambda_\sigma = 0.15$ ,  $\lambda_r = 0.05$  (パラメータの詳細は付録 B) とし学習回数は 48 回の学習を 1 エポックとし 400 エポックだった。また基底ベクトルの初期値は 48 の各カテゴリの全協力者評定の平均値とした。

### 5.1.3 新奇語汎化課題

Colunga & Smith (2000a) や Samuelson (2002) の実験と同じく標準刺激が固形, また非固形である条件を設定し, 形類似刺激の選択率で成績を評価した。新奇語汎化課題は, 実験 1, 2, 3 と同様の方法で行ったが, 実験 4 で学習したデータセットの特徴次元の表現が異なることに注意しなければならない。実験 1, 2, 3 では特徴は「形」「材質」など抽象的な上位概念であったが, 実験 4 では「大きい (小さい)」「丸い (四角い)」などの形容語対による具体的特徴による分散表現であるので, 直接比較が難しい。しかし, Hidaka & Saiki (2004) では形容語がどの知覚次元を表現するか (例えば, 「丸い」は形をよく表現し, 材質は表現しないなど) の評定も行っている。そこで, 先行研究の言語尺度の知覚表現適性に基づいて, 形, 材質表現に適する上位 2 形容詞対を形, 材質次元として選択した。形次元は「丸い (四角い)」(形: 0.83, 材質: -0.25) と「大きい (小さい)」(形: 0.63, 材質: -0.458) の 2 次元, 材質次元は「潤った (乾いた)」(形: -0.75, 材質: 0.08) と「滑らかな (いびつな)」(形: 0.54, 材質: 0.45) の 2 次元とした。但し, 言語尺度の形と材質の表現適性値は適切 (1) ~ 不適切 (-1) で表現された協力者の平均評定値である。固形刺激条件では, 標準刺激の「硬い (軟らかい)」の特徴次元が 5 であり, 非固形刺激条件では 1 とした。標準刺激の固形性以外全ての特徴次元は一様乱数で 1-5 の離散値で生成された。それに対して形類似の選択刺激は材質次元 (上記の 2 次元) を一様乱数で再生成し, その他の次元は標準刺激と同じ値で生成された。また材質類似の選択刺激も形次元のみ同様に再生成された。

選択は実験 1 同様に式 (1), (2) で定義した出力層ベクトルによる物体間の類似度に基づき行った。シミュレーションは各条件 100 回 (初期化を変えた 4 つのネットワークの各 25 回) の平均を選択率とした。式 (1) のスケールパラメータは全ての条件で  $a = 50$  とした。

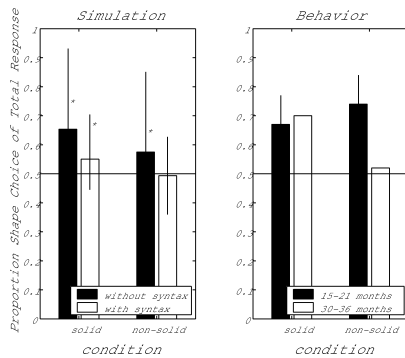


図9 新奇語汎化課題の形類似刺激選択率

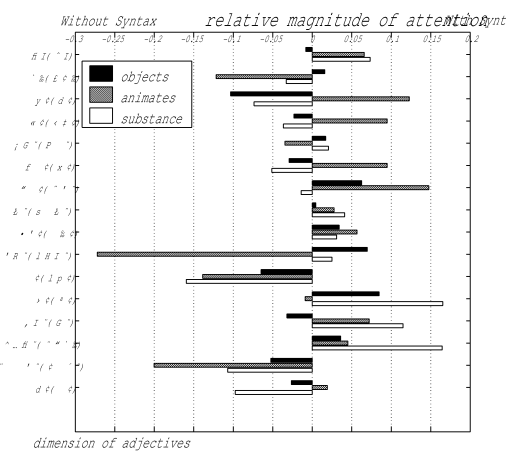


図10 注意ユニットの分析: 上位カテゴリ (動物, 物体, 物質) の平均注意の学習条件間の差

## 5.2 結果・考察

図9は可算・非可算の統語情報を含む、または含まない学習条件での新奇語汎化課題の形類似刺激の選択率を示している。統語なし固形条件 ( $t(99) = 5.52$ ,  $p < 10^{-4}$ ), 統語なし非固形条件 ( $t(99) = 1.68$ ,  $p < .05$ ), 統語あり固形条件 ( $t(99) = 5.75$ ,  $p < 10^{-4}$ ), でチャンスレベル (50%) より形類似刺激の選択率が有意に大きかったが、統語あり非固形条件 ( $t(99) = -.473$ ,  $p = .68$ ) では有意差は無かった。まず統語あり非固形条件を除く条件で、物体に対する形類似刺激の選好がチャンスレベル以上であることから、PSALモデルが複雑な情報源を学習し形バイアスを示す事が示唆される。ただし、図9のように15ヶ月齢の英語話者は固形物、非固形物を区別せず (Samuelson, 2002), 30ヶ月齢では区別し

てを適用する (Colunga & Smith, 2000a; Imai & Gentner, 1997)。一方モデルでは統語なし学習条件ではその区別が見られず15ヶ月の反応パターンに類似し、統語あり条件では区別があり30ヶ月の反応パターンに類似している。モデルの結果は30ヶ月齢のような固形性の区別のある形バイアスの獲得のために統語情報が働く可能性を示唆している。なぜなら心理実験から可算名詞/非可算名詞の統語情報の産出は2歳から2歳半の間に、理解は24ヶ月齢で可能であることが示唆されている (Soja et al., 1991) ので15ヶ月、30ヶ月齢は丁度この前後に相当し、モデルの学習条件と合致する。Imai & Gentner (1997) の実験では可算/非可算の統語クラスを学習しない日本語話者でも固形/非固形物を区別して形バイアスを適用することが示されており、本実験の結果はこれに矛盾する。この矛盾の説明として、一つには存在論的概念が知覚的属性の統計学習で十分に説明出来ず、何らかの推論システムが必要である可能性があり、もう一つは日本語、英語話者が学習するコーパスの統計的性質の違いにより幼児が異なる汎化バイアスを持つ可能性がある。日本語と英語の初期獲得語彙の構成は類似しているが、固形物、非固形物を参照する名詞の割合など異なる性質を持つ (Colunga & Smith, 2000b) ため、形バイアスは学習語彙の違いによっても影響されると考えられる。この仮定の下では提案モデルは英語コーパス学習したので、日本語話者ではなく統語クラスの学習経験が浅い英語話者と類似のパフォーマンスが得られたと解釈できる。しかし、コーパスのどのような要因が日本語・英語話者のバイアスの違いを説明できるのか現時点で不明であり、今後の詳細の研究が必要である。

### 5.2.1 注意ユニットの分析

PSALモデルの隠れ層にはプロトタイプユニットがあり、各々が持つ注意ユニットは各名詞カテゴリでの特徴バイアスの大きさを表している (e.g. 「素早い」-「遅い」の特徴は「ネコ」では重要だが「カップ」では無意味だ、など)。名詞カテゴリを動物、物体、物質という上位カテゴリに分けた場合、領域ごとの特徴注意はどのように変化するか、また統語あり/なしの学習条件でどのように変化するか。幼児のように固形、非固形物を区別している場合、上位カテゴリ領域ごとに優位な特徴注意は異なると

考えられる。

前述のように Hidaka & Saiki (2004) は英語における可算/非可算名詞、日本語における動物的名詞に対する「いる」、非動物的名詞に対する「ある」の用法を言語固有カテゴリとして扱っている。この言語的カテゴリの分類に従い、可算名詞かつ「ある」(非動物)に属するカテゴリを物体 (objects), 可算名詞かつ「いる」(動物)に属するカテゴリを動物 (animates), 非可算名詞かつ「ある」に属するカテゴリを物質 (substance) とした。特定名詞カテゴリに対し最も活性の大きい中間層のユニットを、そのカテゴリのプロトタイプユニットとし、動物、物体、物質の上位カテゴリの平均注意はそれに属する下位カテゴリのプロトタイプユニットの特徴注意の平均値で計算した。上位カテゴリの平均注意ベクトルのノルムを 1 に正規化し、学習の 2 条件における平均注意の差を表したのが図 10 である。統語あり条件では「自然な・人工的な」のように動物カテゴリの分別に有効な次元が相対的に減少し、一方、「丸い・四角い」、「真っ直ぐな・曲がった」のような主に形特徴が物体、物質カテゴリで相対的に増加している事が分かる。つまり、統語的情報なしの場合、元来含まれる、動物・非動物など上位カテゴリ識別など他の要因も固形性による物体・物質の分類と同様になされているが、統語識別を学習に課すことで、物体・物質カテゴリの分類感度がより高まっていると考えられる。これに対応して、固形性次元に近いと考えられる「硬い・軟い」特徴は、統語あり条件では物体の「硬さ」に注目するのに対して物質ではより減少していることが分かる。以上の分析結果と新奇語汎化課題の結果を総合すると、可算、非可算名詞の上位カテゴリの学習は動物-非動物の分類に役立つ特徴より物体-物質の境界に注意するバイアスを高め、結果として各下位カテゴリでの固形性や形に対する分類感度を高めていると考えられる。

## 6. 総合考察

本研究では幼児の新奇語カテゴリ化の計算原理を究明することを目的として研究を行なった。形バイアスは幼児が新奇語を物体に適用する場合に顕著な現象であり、概念形成の重要な側面を反映していると考えられる。形バイアスが何に起因しているかは論争中であるが、本研究では統計的学習に起因するとした立場からその計算論的メカニズムを議論した。

特に既に提案されているボルツマンマシンモデルとカテゴリ学習などで用いられる事例的モデルのどちらが適しているか検討を行なった。その結果、事例的モデルである PSAL モデルが小規模な仮想コーパスで先行研究より安定した形バイアスを示し、現実的なコーパスでも形バイアスの発現を示した。

### 6.1 プロトタイプ学習

提案モデルの特性を考慮すると、幼児の新奇語カテゴリ化が事例またはプロトタイプを単位としており、それを基に上位の概念に拡張、汎化を行なうという事が示唆される。ではなぜ事例単位カテゴリ化を行なうのだろうか。一つの可能性として、問題領域の分割を行なう事により計算的な利点があるということが挙げられる。提案モデルは内部にプロトタイプを表すユニットを持ち、学習時には事例に近いプロトタイプについて集中的に学習する。例えば、提案モデルでは「リンゴ」のある事例について学習を行なった場合、類似しない別のカテゴリ「カップ」についてはほとんど学習しない。一方、連続値ホップフィールドモデルでは「リンゴ」で修正されたある特徴への重みを「カップ」が共通して持つ場合、影響する可能性がある。幼児の語彙獲得は大人が実験室で学習する人工カテゴリよりも遥かに大規模で複雑な問題である。大規模な問題を解く場合に、「リンゴ」に関する知識が「カップ」や「ミルク」に「適用されてしまう」と逆に学習が遅くなるのではないだろうか。つまり事例単位の学習をするとは問題領域を分割し、「全ての領域で応用できる知識」を期待しないという事になる。ただし、プロトタイプごとに完全に独立に学習しては極端に汎化能力が落ちる事になる。しかし実験 4 の上位カテゴリごとの注意分析で示されたように (全てでも一つでもない) 中程度の概念領域では物体、物質などそれぞれに特化した選択的特徴を持つ事が示されている。これは動物、物体、物質などの上位カテゴリは知覚的特徴空間上でクラスタを形成している (Hidaka & Saiki, 2004) ので近傍のプロトタイプ (e.g. ドアとテーブル、牛乳と水など) が強い影響を及ぼす一方で遠いプロトタイプ (e.g. ネコとはさみ、自動車とピザなど) は殆ど独立に学習されるので知識の「汎化領域」が自然と定まるためと考えられる。

プロトタイプに特化した獲得は計算的な利点だけではなく、語彙以外の獲得において発達心理学的知

見があり、語彙だけでなく一般的な認知過程である可能性もある。例えば動詞の獲得は初期段階において特定の動詞に関して個別に用法を習得することが知られている (Tomasello, 1992)。また代名詞でも同様の習得法が見られることも報告されている (Pine, Lieven, & Rowland, 1998)。

## 6.2 名詞カテゴリーの影響

実験4で用いたデータセットは形、材質のみならず固さ、動きなどの要因も豊富に含む多次元知覚空間で構成されていたが、PSALモデルは可算/非可算カテゴリーを同時に学習する事で固さと形の関係を抽出することができた。一方、統語情報を学習しない条件において、非固形物に対する形バイアスの過剰汎化が起こった。これを前述の議論を考慮して考えると、統語情報が知覚空間全体で大局的な問題領域の分割を行なうことで、物体/物質の存在論的カテゴリー知識を促進する効果があると考えられる。これは純粋な言語相対性と異なり、統語情報が潜在する知覚的に重要な要因に対する重み付けを変更する役割として働く可能性を示唆する。

## 6.3 これからの展開

Smithの獲得モデルの1次汎化メカニズムの明示的な訓練は語彙獲得数を増加させることが示されている (Smith et al., 2002; Samuelson, 2002)。本研究で提案したPSALモデルは1次汎化メカニズムを取り入れたモデルであるので、その学習過程で形バイアスの強弱と語彙獲得速度などを分析することで学習後だけでなく発達過程を再現できる可能性があり、検討すべき課題である。また今回のシミュレーションでは強制選択パラダイムを対象としたが、新奇語に対する刺激を自由に選択する課題では、日本人幼児が刺激の形状または材質の類似した刺激を迷う(形:材質=55%:44%)が、どちらかしか選択せず、「同じもの」を選択するよう促した場合も迷う(形:材質=69%:73%)が、両方を選択し得る事が示されている (Imai & Mazuka, 2003)。提案モデルは選択行動について精緻化されていないので、現時点ではこの現象を説明する事は出来ないが、こうした言語発達の様々な側面に対してモデルの整合性を検討することが今後の課題として挙げられる。

## 文献

- Colunga, E. & Smith, L. (2000a). Committing to an Ontology: A Connectionist Account. *The Twenty Second Annual Meeting of the Cognitive Science Society*.
- Colunga, E. & Smith, L. (2000b). Learning to Learn Words: A Cross-Linguistic Study of the Shape and Material Biases. *The 24th Annual Boston University Conference of Language Development*.
- Fenson, L., Dale, P., Reznice, J., Bate, E., Hartung, J., Pethick, S., & Reilly, J. (1993). MacArthur Communicative Development Inventories. San Diego.
- Heibeck, T. & Markman, E. (1987). Word learning in children: an examination of fast mapping. *Child Development*, **58**, 1021-1034.
- Hidaka, S. & Saiki, J. (2004). A mechanism of ontological boundary shifting. *The Twenty Sixth Annual Meeting of the Cognitive Science Society*, 565-570.
- Imai, M. & Gentner, D. (1997). A cross-linguistic study of early word meaning: universal ontology and linguistic influence. *Cognition*, **62**, 169-200.
- Imai, M. & Mazuka, R. (2003). *Re-evaluation of linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation*, 430-464. Cambridge, MA: MIT Press.
- Jones, S. & Smith, L. (2002). How children know the relevant properties for generalizing object names. *Developmental Science*, **5**, 219-232.
- Kruschke, J. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **99**, 22-44.
- Landau, B., Smith, L., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, **3**, 299-321.
- Markman, E. & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, **16**, 1-27.
- Movellan, J. (1990). Contrastive Hebbian learning in the continuous Hopfield model. *Connectionist Models Proceedings of the 1990 Summer School*, 10-17.
- Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Univ. of Illinois Press.

- Pine, J. M., Lieven, E. V. M., & Rowland, C. F. (1998). Comparing different models of the development of the English verb category. *Linguistics*, **36**, 807–830.
- Quine, W. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382–439.
- Samuelson, L. (2002). Statistical Regularities in Vocabulary Guide Language Acquisition in Connectionist Models and 15-20 Month Olds. *Developmental Psychology*, **38**, 1016–1037.
- Samuelson, L. & Smith, L. (1999). Early noun vocabularies: do ontology, category structure and syntax correspond? *Cognition*, **73**, 1–33.
- Schafer, G. & Plunkett, K. (1998). Rapid word learning by 15-month-olds under tightly controlled conditions. *Child Development*, **69**, 309–320.
- Smith, L., Jones, S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, **13**, 13–19.
- Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meanings: object terms and substance term. *Cognition*, **38**, 179–211.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Wolpert, D. & Macready, W. (1997). No Free Lunch Theorems for optimization. *IEEE Transactions on Evolutionary Computation*, **1**, 67–82.
- Yoshida, H. & Smith, L. B. (2003). Shifting ontological boundaries: how Japanese- and English- speaking children generalize names for animals and artifacts. *Developmental Science*, **6**, 1–34.

## 付 録

### A. Contrastive Hebbian Learning

本研究では Colunga & Smith (2000a) と Samuelson (2002) がシミュレーションに用いた Contrastive Hebbian Learning (Movellan, 1990)) を追実験に用いている。以下にその計算手続きを示

す。ネットワーク全体のエネルギー関数は以下の式で表される。

$$F = E + S \quad (6)$$

ただし  $E$  と  $S$  は

$$E = -\frac{1}{2} \mathbf{a}^t \mathbf{W} \mathbf{a} = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i w_{ij} a_j \quad (7)$$

$$S = \sum_{i=1}^n \int_{rest}^{a_i} f_i^{-1}(a) da \quad (8)$$

ただし  $a_i$  はユニット  $i$  の活性値,  $w_{ij}$  はユニット  $i$  と  $j$  の結合荷重,  $rest = f(0)$  である。ネットワークのエネルギー関数は以下の式に従って更新することで単調に減少する。

$$\frac{da_i}{dt} = \lambda(-a_i + f_i(net_i)) \quad (9)$$

ただし  $f_i$  はシグモイド関数で  $T$  は温度パラメータである。

$$f(x) = \frac{1}{1 + e^{-\frac{x}{T}}} \quad (10)$$

ネットワークの初期状態から十分に上記の更新をする事で定常状態の活性値  $\hat{a}$  が得られる。CHL では隠れ層のユニットのみを自由に更新する固定フェーズと隠れ層と出力層のユニットを自由に更新する自由フェーズを持ち、それぞれのフェーズでの定常状態の活性値を  $\hat{a}_i^{(+)}, \hat{a}_i^{(-)}$  とする。このネットワークで入力に対する出力を生成するには自由フェーズを動作させ、出力層の定常状態の活性値を出力パターンとして定義する。また学習時にはまず固定フェーズで動作させ、一定回数更新し定常状態後に自由フェーズで一定回数更新し定常状態の活性値を得る。ユニット  $i, j$  間結合  $w_{ij}$  は2つのフェーズの活性値を用いて以下のように更新する。

$$\Delta w_{ij} \propto \hat{a}_i^{(+)} \hat{a}_j^{(+)} - \hat{a}_i^{(-)} \hat{a}_j^{(-)} \quad (11)$$

結合荷重は全て  $-0.5$  から  $0.5$  の範囲の連続値の擬似一様乱数で初期化された。実験 1 では固定、自由フェーズの更新回数 1000 回, 更新率  $\lambda = 1$ , 温度  $T = 0$  (ヘビサイド関数), 重みの学習率 0.72 で学習サンプルの誤差の絶対和が 0.01 以下になった

場合に更新を終了した。1回に1サンプルを学習するオンライン学習で異なる初期化を行なった4つのネットワークで平均227回の重み更新が行なわれた。実験2では固定、自由フェーズの更新回数は500回、重みの学習率は0.005、重みの更新は異なる初期化の4つネットワークの500回に固定して行なわれた。また温度  $T$  を0.5から0まで固定、自由フェーズの更新回数に比例して下げて行なわれた。他の条件は実験1と同じである。

## B. PSALの数学的定式化

PSAL (Prototype-specific Attention Learning) モデルは3層のフィードフォワード型で、隠れ層のユニットの活性化関数に指数族関数などの放射状の関数を持つ RBF (Radial Basis Function) の1種である。隠れ層の基底ユニットは入力空間におけるカテゴリのプロトタイプ的位置を表し、エラー駆動型の学習によりその位置を最適化する。

出力層ユニットの活性値  $o_k$  は隠れ層ユニットの活性値  $h_j$  の荷重線形和で、以下の式で表される。

$$o_k = \sum_j w_{kj} h_j \quad (12)$$

また、隠れ層ユニットの活性値  $h_j$  は入力層ユニットベクトル  $I$  と基底ユニットベクトル  $r_j$  の荷重付きユークリッド距離で決定され、以下の式で表される。ただし  $m$  は正の入力空間の距離基準パラメータで  $c$  は定数のスケールパラメータである。

$$h_j = \exp(\text{net}_{ji}) \quad (13)$$

ただし

$$\text{net}_{ji} = -c \left( \sum_i \sigma_{ji} (r_{ji} - I_i)^m \right)^{\frac{1}{m}} \quad (14)$$

出力層ユニットの活性値  $o_k$  とそれに対する教師信号  $t_k$  の誤差を以下の式で定義する。

$$E = \frac{1}{2} (t_k - o_k)^2 \quad (15)$$

PSAL モデルは教師信号と出力との誤差を逆伝播し、層間の荷重と注意ユニットの重みを以下の更新式で最適化する。ただし  $\Delta W_{kj}$  は隠れ層と出力層間の結合荷重の更新量、 $\lambda$  は定数の学習率、 $\Delta \sigma_{ji}$

は注意ユニットの更新量、 $\Delta r_{ji}$  は隠れ層ユニットの基底ベクトル、 $\lambda_w, \lambda_\sigma, \lambda_r$  は学習率である。

$$\Delta w_{kj} = -\lambda_w \frac{\partial E}{\partial w_{kj}} = -\lambda_w \frac{\partial E}{\partial o_k} \frac{\partial o_k}{\partial w_{kj}} \quad (16)$$

$$\Delta w_{kj} = -\lambda_w \delta_k h_j \quad (17)$$

ただし

$$\delta_k = t_k - o_k \quad (18)$$

また隠れ層の荷重  $\sigma_{ij}$  の更新式は偏微分の連鎖則より

$$\begin{aligned} \Delta \sigma_{ji} &= -\lambda_\sigma \frac{\partial E}{\partial \sigma_{ji}} \\ &= -\lambda_\sigma \sum_k \frac{\partial E}{\partial o_k} \frac{\partial o_k}{\partial h_j} \frac{\partial h_j}{\partial \text{net}_{ji}} \frac{\partial \text{net}_{ji}}{\partial \sigma_{ji}} \end{aligned} \quad (19)$$

$$\sum_k \frac{\partial E}{\partial o_k} \frac{\partial o_k}{\partial h_j} = \sum_k \delta_k w_{kj} \quad (20)$$

$$\Delta \sigma_{ji} = -\lambda_\sigma \left( \sum_k \delta_k w_{kj} \right) \frac{1}{2} c h_{ji} \text{net}_{ji}^{\frac{1}{m}-1} (h_{ji} - I_i)^m \quad (21)$$

また同様に基底ユニットの位置  $r_{ji}$  の更新式も  $E$  の偏微分より

$$\begin{aligned} \Delta r_{ji} &= -\lambda_r \frac{\partial E}{\partial r_{ji}} \\ &= -\lambda_r \sum_k \frac{\partial E}{\partial o_k} \frac{\partial o_k}{\partial h_j} \frac{\partial h_j}{\partial \text{net}_{ji}} \frac{\partial \text{net}_{ji}}{\partial r_{ji}} \end{aligned} \quad (22)$$

(20) 式より

$$\begin{aligned} \Delta r_{ji} &= -\lambda_r \left( \sum_k \delta_k w_{kj} \right) c h_{ji} \text{net}_{ji}^{\frac{1}{m}-1} (\sigma_{ji} (r_{ji} - I_i))^{m-1} \end{aligned} \quad (23)$$

## C. Hidaka & Saiki (2004) の評定語彙およびその言語的カテゴリ

日本語は各名詞に対する動詞の「いる」と「ある」の用法を基準に分類したカテゴリであり、英語は各名詞が可算名詞 (count) と非可算名詞 (mass) のどちらに属するかを基準としたカテゴリである。

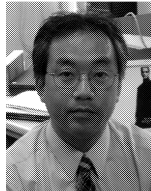


	英語	日本語		英語	日本語
蝶	count	いる	椅子	count	ある
ネコ	count	いる	ドア	count	ある
魚	count	いる	冷蔵庫	count	ある
カエル	count	いる	テーブル	count	ある
ウマ	count	いる	雨	mass	ある
サル	count	いる	雪	mass	ある
トラ	count	いる	石	count	ある
腕	count	ある	木	count	ある
目	count	ある	水	mass	ある
手	count	ある	カメラ	count	ある
膝	count	ある	カップ	count	ある
舌	count	ある	鍵	count	ある
ブーツ	count	ある	お金	mass	ある
手袋	count	ある	紙	mass	ある
ジーンズ	count	ある	はさみ	count	ある
シャツ	count	ある	植木	count	ある
バナナ	count	ある	風船	count	ある
タマゴ	count	ある	本	count	ある
アイスクリーム	count	ある	人形	count	ある
牛乳	mass	ある	糊	mass	ある
ピザ	count	ある	飛行機	count	ある
塩	mass	ある	列車	count	ある
トースト	count	ある	自動車	count	ある
ベッド	count	ある	自転車	count	ある



日高 昇平 (学生会員)

1979年生. 2002年, 九州大学理学部生物学科卒業. 2004年, 京都大学大学院情報学研究科修士課程修了. 現在, 京都大学大学院情報学研究科博士後期過程在学中. 概念形成, 語彙獲得の理論的研究に従事.



齋木 潤 (正会員)

1962年静岡県に生まれる. 1985年東京大学教育学部教育心理学科卒業. 1996年カリフォルニア大学ロスアンゼルス校大学院心理学系修了. Ph.D. (心理学)・名古屋大学大学院人間情報学研究科助手, 京都大学大学院情報学研究科助教授を経て, 2005年より京都大学大学院人間・環境学研究科助教授. 視覚認知の行動実験, 機能的脳イメージング, ニューラルネットワークモデリングの研究に従事.

(Received 16 July 2004)

(Accepted 11 May 2005)