# 自然物体の “種類” に固有な新奇語の汎用

**日高　昇平 (shhidaka@indiana.edu)**
Indiana University

**Linda B. Smith (smith4@indiana.edu)**
Indiana University

**Abstract**

近年の発達心理学的研究から，2，3 歳児はさまざまな新奇な物体に付与された新奇語に対して，その物体の “種類” に固有な汎用を行う事が報告されている．例えば，硬い木材でできた物体に対してはその形状を，液状の物体に対してはその材質を基に汎用する場合が多い．また，このような新奇語の汎用は，たった 1 事例の新奇物体に対してでも見られる事から，この効率的な名詞学習は即時マッピングと呼ばれる．本研究では，新奇語の汎用を，ある特徴空間上での確率分布を推定するカテゴリ化問題として定式化し，計算論的モデルの提案を行った．原理的に，多数の事例を観測するほど，その確率分布の推定精度は高くなる．しかし，逆に，真に新奇なカテゴリから観測された唯一の事例だけでは，その確率分布を推定する事は困難である．この意味で，即時マッピングには何らかの事前的な制約条件が必要である．本研究の提案モデルでは，観測した事例の確率分布の推定において，既に学習したカテゴリを含むカテゴリ集団全体の識別性の最大化を制約として導入した．この提案モデルは成人の名詞カテゴリと類似の “種類” に固有な構造が自己組織的に形成でき，また，新奇語の汎用シミュレーションでは，新奇な事例に対して高い精度でカテゴリ分布を推定可能であった．

**Keywords:** 新奇語の汎用，即時マッピング，特徴選択，種類固有性

## カテゴリの種類固有性

カテゴリ化の最も基礎的な機能は，特徴を選択し情報を圧縮する事である．例えば,「ボール」というカテゴリには様々な色や模様の物が含まれるが, 概してそれらは共通に球形をとる事が多い．つまり,「ボール」という名前の意味を理解するには, そのカテゴリに本質的な特徴 (球形) を選択し, 関係のない特徴 (色や模様など) を無視する必要がある．さらに, 一つの上位的な特徴の下で，カテゴリ群が集団的な共通性を持つ事もある (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002)．例えば,「ボール」には「球形」が共通するが,「カップ」には「円筒」であり,「バット」には「棒状」など, これらはそれぞれに「固有の形状」を持ち, また「形がくずれない硬さ」という特徴を共通に持つ．このような上位的な共通性を持つカテゴリ群の形容として, 上位カテゴリ (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), Over hypothesis (Goodman, 1955), The second order generalization (Smith et al., 2002) なども用いられるが, ここではこれを「同じ種類」のカテゴリと呼ぶ．

このように自然カテゴリの特定の種類のカテゴリ群は, 特定の共通的な特徴を基にカテゴリ化される事が多い．従って, 種類の構造化メカニズムを解明する事は, 自然な概念形成を理解するための基礎となると考えられる．本研究では, 特に幼児が初期に獲得する典型的な名詞群に関して, 種類固有的なカテゴリの発生を説明する計算論的モデルの提案を行う．このモデルでは種類固有的なカテゴリ群の構造を,「種類に固有」ではないメカニズムによって説明する事で, より一般的で統一的な視点を提案する事を目的とした．以下では, まず本研究の背景として種類固有的な新奇語の汎用についての研究を紹介し, 幼児の効率的な新奇名詞の学習と, そこから推定されるカテゴリ群の集団的な構造について述べる．次に, これらの新奇語名詞の学習を理論的に定式化し, 本研究のモデルを提案する．

## カテゴリの種類固有性とその連続的分布

新奇語汎用課題 (Novel word generalization task) は，ある物体に付与された新奇語の他の物体への汎用をみる課題で，幼児が初期に形成する名詞カテゴリの構造を調べるために広く用いられる (Markman, 1989)．多くの研究で，対象に与えられた新奇名詞を汎用する際に，2，3 歳児は堅固な物体 (e.g., 木製のブロック) に対してその形類似性を重視し，また堅固でなく液状である物質 (e.g., クリーム) に対してその材質類似性を重視する事が知られている (Jones, Smith, & Landau, 1991; Soja, Carey, & Spelke, 1991)．また，これ以外にも様々な状況で，新奇な名詞が付与された新奇な物体であるのにも関わらず，幼児は系統的な汎用を行う事から，この効率的な名詞学習は特に即時マッピング (Fast mapping) と呼ばれる．このような知見は，名詞カテゴリの種類固有性を表す典型的な例と言える．しかし, 新奇語の汎用に関する最近の研究から，自然カテゴリの特徴選択がカテゴリ群に固有的であるだけではなく，連続的・漸次的である事が示唆されている．その一つとして，堅固な物体と堅固ではない物質の「間」にあるカテゴリがその中間的な構造を持つ事を示唆されている (Imai & Gentner, 1997; Colunga & Smith, 2005)．Colunga & Smith (2005) は名詞の付与される新奇事例が形の典型的な人工的なモノ (複雑で，角ばった輪郭を持つ) から，物質的なモノ (単純な形を持つ) へと変わるに従って，形と材質類似性に基づく幼児の新奇名詞の汎用が次第に変化する事を示した．

**カテゴリ尤度の大域的な分布**　以降の議論では，新奇語の汎用を確率的に表現されたカテゴリの推定の問題として定式化し，先行研究から推測されるカテゴリ分布について説明する．幼児の名詞カテゴリ尤度の大域的な仮想的なパタンを図 1b, c に示す．図 1b は形と肌理による 2 次元の特徴空間を表す．また，個々の楕円は，あるカテゴリに属する事例が出現する一定頻度の等高線を表す (図 1a; 事例があるカテゴリに含まれる頻度＝カテゴリ尤度: 3 次元図の z 軸)．新奇語の汎用や自然カテゴリでは，カテゴリに含まれる事例に共通する特徴はより重視され，事例間でばらつきのある特徴は軽視される事が示唆されている (Samuelson & Smith, 1999; Rips, 1997)．従って本研究では，あるカテゴリに属する事例の特徴空間上の出現頻

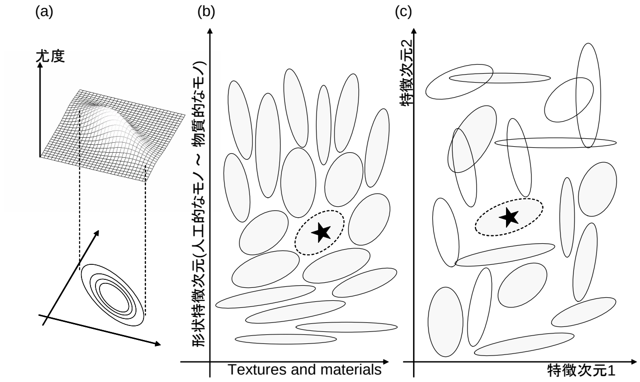


図 1: 仮想的なカテゴリの尤度分布 (等高線)

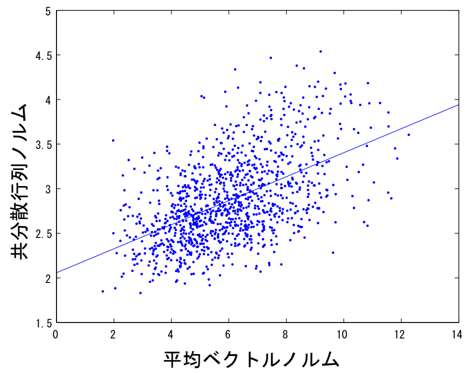


図 2: 全カテゴリ対の平均ベクトルのノルム (x 軸) と共分散行列のノルム (y 軸) の散布図 ($R = .466$).

度分布 (以下カテゴリ尤度と呼ぶ) を特徴の重要性と同一視する．つまり，楕円の形は，カテゴリごとの事例のばらつきを表し，またどの特徴がそのカテゴリの判断に重要であるかを示す．図の最も下側にある横長の楕円は，そのカテゴリに属する事例は肌理特徴ではばらつきがあるが，形特徴ではばらつきが小さく事例の類似性が高い事を示す．従って，最左側にある縦長の楕円は人工物が形類似性に基づき汎用されるパタンを表す．一方，図の最も上側にある縦長の楕円は，カテゴリに属する事例が肌理特徴で類似し，形特徴にばらつきがある事を示す．従って，この楕円は物質が材質や肌理類似性に基づき汎用されるパタンを表す．そして，形が類似した左側のカテゴリ群と，材質が類似した右側のカテゴリ群の間には，形と材質がほどほどに類似した中間的なカテゴリ群が漸次的に横たわっている．

以上の知見をまとめると，自然カテゴリでは，類似したカテゴリほど類似の特徴重要性を持ち，非類似のカテゴリほど非類似の特徴重要性を持つ．これは，つまり，特定の領域の “同じ種類” のカテゴリ群がその種類に固有的な特徴を重視し，また別の領域のカテゴリ群はその種類に固有的な特徴を重視する事を意味している．さらに，この種類固有性は特徴空間上で連続的に変化するのではないかと考えられる．本研究では，このような大域的なカテゴリ分布を持つ特徴空間を「滑らかな」特徴空間と呼ぶ．もし，カテゴリが滑らかであれば，カテゴリの中心間の距離 (i.e., 平均ベクトルノルム) とカテゴリ尤度の類似性 (i.e., 共分散行列ノルム) には正の相関があるはずである．Hidaka, Smith & Saiki (2006) はこの予測を，成人による特徴の評定実験によって検討した．その結果，幼児が初期に獲得する 48 名詞カテゴリは滑らかな性質を持つことが示された (図 2).

## 滑らかなカテゴリと即時マッピング

もし自然カテゴリが滑らかであったとすると，どのような利点があるのだろうか．滑らかさは単にカテゴリの記述的な特性であるだけではなく，幼児の種類固有的な新奇語汎用を促進する可能性がある．なぜなら，滑らかなカテゴリ群は，そうでないものに比べて，未知のカテゴリに関する予測性が高いと考えられるからである．予測性とカテゴリ分布の関係を，図 1b と c を用いて説明する．ここで，ある学習者にとって図 1b の点線で表記したカテゴリは未知のもので，それ以外の実線で表されたカテゴリを既に学習ずみであると仮定しよう．図 1b の場合，未知のカテゴリの尤度パタン (i.e., 楕円の形状) を予測する事は比較的容易であると考えられる．なぜなら，カテゴリ分布が全体的に連続的である事から，未知なるカテゴリも近隣のカテゴリと類似の尤度パタンを持つと推測できる．一方，図 1c のように滑らかではない構造を持つ場合はどうだろうか．この場合，カテゴリ分布に連続性・規則性が見られないため，新奇なカテゴリ尤度の推測は他のカテゴリを利用して行う事が困難である．従って，カテゴリ分布全体に渡る一貫性・連続性は新奇なカテゴリの予測を促進する事がわかる．

## カテゴリ識別性の最適化

それでは，滑らかなカテゴリ群 (種類固有性および連続性) はどのように発生するのだろうか．ここでは，滑らかなカテゴリ群の発生を説明するモデルを提案し，その中心的な考え方について述べ，次節で形式的な定義・理論について述べる．

一般に，カテゴリはそれに属する事例の分布をよく反映するべきであり，つまり，事例への適合性はカテゴリの形成を制約する一つの必要条件である．本研究で提案するモデルでは，これに加えて，カテゴリ間の識別誤りの最小化，という原則を取り入れる．もし与えられた事例に対して 2 つのカテゴリ A, B が十分に適合していたとしても，そのカテゴリ A と B が互いに識別できないほど類似していれば，どちらが冗長だと考えられる．このような考えから，同じ適合性を持つカテゴリ群であれば，より識別性の高いカテゴリのほうが，学習者にとって無駄がない事例の表現であるとみなす．この意味で，カテゴリ間の重複のある図 1c よりも図 1b のカテゴリ分布はより識別性が高い．従って，滑らかなカテゴリであるためには少なくとも識別性が高い必要があると考えられる．

識別性を保つためには，カテゴリ間の尤度の重複をできる限り避ける必要があり，この自己組織的な分布形成では，複数のカテゴリが競合的に事例の分布を占有する過

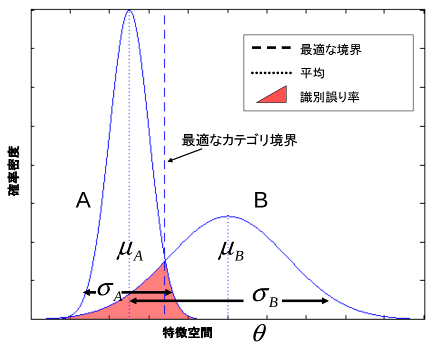


図 3: カテゴリ尤度分布と識別誤り率

程がみられる．カテゴリ間の競合によって生じる均衡 (最適な状態) は，あたかも，限られた空間 (事例の分布) をもつスーツケースに，多数の荷物 (重複のないカテゴリ) に効率よく“詰め込む”操作に想起させる事から，本研究ではこれをカテゴリの詰め込みモデルと呼ぶ．また適合性と識別性のバランスを詰め込みコストと呼ぶ．

## カテゴリ詰め込みモデルの定式化

この理論の定式化では，識別性をカテゴリを誤識別する確率，適合性を与えられた事例に対するカテゴリの尤度として定義する．識別性と適合性を結合し，この両方を結合した条件つきの尤度関数を詰め込みコスト関数と呼び，その最適化の結果得られるカテゴリ分布について述べる．まずは，2 つのカテゴリだけの単純な場合から考えて，それを一般の場合に拡張する．即時マッピングにおける確率分布の推論は，1 事例しか与えられないため尤度関数が縮退するため，一般の導出の特殊な場合として与えられる．

**カテゴリの識別性**　まず，図 3 のようにカテゴリ A と B の尤度が 1 次元特徴空間上に分布する場合に，ある特徴を持つ事例がカテゴリ A または B のどちらに属するのか判断するとしよう．ある事例に対して，それが属する最も尤度の高いカテゴリと判断するのが最適であるので，図 3 の場合，破線を境界として左側の事例をカテゴリ A と，右側の事例をカテゴリ B とみなす場合に最も誤りの少ないカテゴリ判断となる．また，この最適な判断においても不可避な誤りの確率は，図 3 の色付けされた部分の面積 (i.e., 二つのうちより小さいカテゴリ尤度の和) で表される．さらに，最小値に関する不等式を用いると，誤識別の確率 $\varepsilon_{AB}$ の上界は次のように表現できる: $\varepsilon_{AB} \le F_{AB} = \int_\Omega \{P(\theta|A)P(\theta|B)\}^{\frac{1}{2}} d\theta$, ただし，$\theta$ と $\Omega$ はそれぞれ特定の特徴値とその特徴空間である．我々の目的は誤りの最小化であるので，誤りの上界 $F_{AB}$ によって識別率を定義する．加えて，本研究ではカテゴリ尤度 $P(\theta|A), P(\theta|B)$ が正規分布 (平均 $\mu_i, \mu_j$ および共分散行列 $\sigma_i, \sigma_j$) に従うと仮定し，特にこの場合，誤識別率 (上界) の対数は，以下のように一種の重みつき距離の形で表される: $-\log(F_{ij}) = \frac{1}{4}(\mu_i-\mu_j)^t(\sigma_i+\sigma_j)^{-1}(\mu_i-\mu_j) + \log|\frac{1}{2}(\sigma_i+\sigma_j)| - \frac{1}{2}\log(|\sigma_i||\sigma_j|)$．これは特に Bhattercheryya 上界と呼ばれている (Duda, Hart, & Stork, 2000).

次に，より一般的な場合，$D$ 次元特徴空間上 $\Omega \supset \theta$ の $N$ 個のカテゴリについて考える．$c_i$ 番目のカテゴリから特徴が $\theta$ である事例が出現する条件つき確率を $P(\theta|c_i)$ とし，その正規分布の平均ベクトルを $\mu_i$，共分散行列を $\sigma_i$ とする．このとき，$P(\theta|c_i) = ((2\pi)^D|\sigma_i|)^{-\frac{1}{2}} \exp(-\frac{1}{2}(\theta-\mu_i)^t\sigma_i^{-1}(\theta-\mu_i))$ ただし 上付きの $^t$ はベクトル・行列の転置を表す．$N$ カテゴリのときの識別性 $\bar{F_N}$ をすべてのカテゴリ対の Bhattercherrya 上界の総和とすると，

$$\exp(\bar{F_N}) = \sum_i \sum_j P(c_i)P(c_j) \int_\Omega \{P(\theta|c_i)P(\theta|c_j)\}^{\frac{1}{2}} d\theta \quad (1)$$

.

**尤度関数**　次に事例に対するカテゴリの適合度を定義する．与えられた事例 $(k=1,2,\ldots,K)$ に対するカテゴリに対数尤度 $\bar{G_K}$ は，カテゴリのラベルが未知の場合 (教師なし学習)，以下のように混合正規分布モデル (Gaussian Mixture Model, GMM) として与えられる．

$$\bar{G_K} = \sum_{k=1}^{K} \log\left\{\sum_i P(c_i)P(\theta = x_k|c_i)\right\} \quad (2)$$

ただし $x_k$ は $k$ 番目の事例の特徴を表す．混合正規分布モデルではある事例 $x_k$ がカテゴリ $i$ に属する条件つき確率 $R_{ik} = P(c_i|x_k)$ も同時にパラメタ推定されるが，カテゴリラベルが与えられた場合 (教師あり学習) を $R_{ik}$ を適当な定数とした場合として含む．従って，教師あり・なしの場合を含めて混合正規分布モデルを尤度関数とする．

**詰め込みコスト関数**　詰め込みコスト関数は以上で定義した識別性 $\bar{F_N}$ がある値のときの尤度関数 $\bar{G_K}$ と定義する．これは，以下のようにある定数 $\lambda$ を用いた Laggrange 乗数法による $L_N$ 制約つき最適化として定式化できる．すなわち，$L_N = \bar{G_K} + \lambda(\bar{F_N} - \log(F_0))$，ただし，$F_0$ はある定数とする．

**詰め込み最適なカテゴリ分布の性質**　次に，最適詰め込みにおけるカテゴリ分布の解析的な性質を示す．これは，詰め込みコスト関数 $L_N$ のカテゴリのパラメタ $\mu_i$, $\sigma_i$ $(i=1,2,\ldots,N)$ による偏微分がゼロになる解として与えられ，以下に最適な共分散 $\sigma_i$ の性質を導出する．詰め込みコスト関数 $L_N$ の $\sigma_i$ $(i=1,2,\ldots,n)$ による偏微分から，$\frac{\partial L_N}{\partial \sigma_i^{-1}} = \sum_j Q_{ij}\left\{(\mu_i-\bar{\mu}_{ij})(\mu_i-\bar{\mu}_{ij})^t + \bar{\sigma}_{ij} - \sigma_i\right\} + \lambda\sum_k^{K_i} R_{ik}\{(x_{ik}-\mu_i)(x_{ik}-\mu_i)^t - \sigma_i\} = 0$ ただし，$Q_{ij} = P(c_i)P(c_j)F_{ij}$，$R_{ik} = P(c_i|x_k) = \frac{P(x_k|c_i)}{\sum_i P(x_k|c_i)}$，$\bar{\sigma}_{ij} = 2\sigma_i(\sigma_i+\sigma_j)^{-1}\sigma_j$，$\bar{\mu}_{ij} = \frac{1}{2}\bar{\sigma}_{ij}(\sigma_i^{-1}\mu_i + \sigma_j^{-1}\mu_j)$ である．この方程式を $\sigma_i$ $(i=1,2,\ldots,N)$ についてそれぞれ解くことで，

$$\sigma_i = \frac{\sum_j Q_{ij}\left\{\hat{S}_{ij} + \bar{\sigma}_{ij}^{-1}\right\} + \hat{\lambda}S_i}{\sum_j Q_{ij} + \hat{\lambda}\sum_k R_{ik}} \quad (3)$$

ただし，$\hat{S}_{ij} = (\mu_i-\bar{\mu}_{ij})(\mu_i-\bar{\mu}_{ij})^t$ , $S_i = \sum_k^{K_i} R_{ik}(x_k-\mu_i)(x_k-\mu_i)^t$, $\hat{\lambda} = \lambda\frac{\sum_i^N\sum_k^K R_{ik}}{\sum_i^N\sum_j^N Q_{ij}}$. 式 (3) は共分散行列 $\sigma_i$ が 3 つの要素の重みつきの平均 (i.e., $Q_{ij}$ $(j=1,2,\ldots,N)$ と $\lambda_i$ を確率密度とする) から構成される事を意味する．その構成要素は，

データの散布行列 $S_{ij}$，カテゴリ集団の共分散行列の調和平均 $\bar{\sigma}_{ij}^{-1}$，そしてカテゴリ間の散布行列 $\bar{S_i}$ である．ここで，確率密度 $Q_{ij}$ がカテゴリ $c_i$ と $c_j$ の距離によって指数的に減衰する事を踏まえると，カテゴリ間の共分散行列 $\hat{S}_{ij}$ はカテゴリ $c_i$ を中心として，その近傍の他のカテゴリ $c_j$ へのばらつきを表している．これを幾何学的に解釈すれば，図 1b の破線で示した楕円ように，周囲にある既存のカテゴリのパタンから "空白地" の形状を表す事がわかる．

また，共分散行列の調和平均 $\bar{\sigma}_{ij}$ は，カテゴリ $c_i$ の近傍の他のカテゴリの平均的な共分散行列を表しており，詰め込み最適解において，カテゴリ間の距離が共分散行列と相関を持つことを示唆している．以上の事をまとめると式 (3) は一般に，詰め込み最適なカテゴリ群が "滑らか"(i.e., 類似のカテゴリが類似の共分散パタンを持つ) である事を示唆している．より具体的なカテゴリ分布についてはシミュレーションで後述する．

**唯一つの新奇事例からの分布推定**　ここで，典型的な即時マッピングの場合，つまり新奇カテゴリに唯一の事例が与えられその確率分布 (ここでは共分散行列) を推定する場合，について考える．これは前述の一般的な導出おいて，データ共分散行列 $S_i \approx$ を近似的に 0 とみなせる特殊な場合 (すなわち，$K_i$=1 かつ $x_k \approx \mu_i$) に対応する．このとき，新奇カテゴリの共分散は以下のように推定できる．

$$\sigma_i = C\frac{\sum_j Q_{ij}\left\{\hat{S}_{ij}+\bar{\sigma}_{ij}^{-1}\right\}}{\left|\sum_j Q_{ij}\left\{\hat{S}_{ij}+\bar{\sigma}_{ij}^{-1}\right\}\right|} \tag{4}$$

ただし，$\lambda$ は $\frac{\partial L_N}{\partial \lambda^{-1}} = \bar{G}_K - \log(C) = 0$ ($C$ は適当な定数) を解く事で消去した．

**詰め込み最適な平均ベクトル**　停留点の満たす必要条件から，$\frac{\partial L_N}{\partial \mu_i} = 0$．これより，$\left\{\sum_k R_{ik}\sigma_i^{-1}(\mu_i - x_k) - \hat{\lambda}\sum_j Q_{ij}(\sigma_i+\sigma_j)^{-1}(\mu_i-\mu_j)\right\} = 0$ を得る．これを $\mu_i$ について解く事で，以下の最適解を得る．

$$\mu_i = Z^{-1}\left\{\sum_k R_{ik}\sigma_i^{-1}x_k + \hat{\lambda}\sum_{j\neq i} Q_{ij}(\sigma_i+\sigma_j)^{-1}\mu_j\right\} \tag{5}$$

ただし，$Z = \left\{\sum_k R_{ik}\sigma_i^{-1} + \hat{\lambda}\sum_{j\neq i} Q_{ij}(\sigma_i+\sigma_j)^{-1}\right\}^{-1}$

これまでの導出で得られた式 (3), (4) や (5) では，右辺の一部の項 ($Q_{ij}$ や $R_{ik}$ など) に求めるべき $\mu_i, \sigma_i$ が含まれるため解析的に求める事ができないが，最適な数値解は EM アルゴリズム (Dempster et al, 1977) を用いた繰り返し計算によって求める事ができる．

## シミュレーション

前節では，カテゴリの詰め込みが滑らかなカテゴリを形成する事を解析的に示した．ここでは二つのシミュレーションを報告する．第一のシミュレーションでは，詰め込みコスト関数の最適化を行い生成されたカテゴリと，成人から得られたカテゴリとの比較を行った．このカテゴリ分布の再現シミュレーションでは，モデルに与える事例の構造を最小限にする事で，提案モデルの特性である識別性が構造化するカテゴリ分布に焦点を当てる．このシミュレーションは，成人の自然カテゴリにおける種類特異的な構造を再現する事で，モデルの持つ識別性の妥当性を検討する．次に，新奇な 1 事例からその未知のカテゴリの確率分布を推定するシミュレーションを行った．これは幼児が新奇語の汎用を行う状況の再現に相当する．もしモデルが 1 事例から高い精度でカテゴリ分布を推定できれば，モデルに仮定した識別性の制約が即時マッピングに深く関連する事を示せる．モデルで用いた名詞カテゴリは，全て幼児が 3 歳までに獲得する標準的な名詞リスト (the Mac Arthur-Bates Communicative Development Inventory, MCDI; Fenson et al., 1994) から選択されたもので，その特徴は後述の実験で成人が評定を行った．

### 方法

**カテゴリ分布の再現**　シミュレーションでは，モデルには最小限の事例の構造が与えられた．具体的には，全てのカテゴリに対してゼロベクトル ($K=1$, $x_1=0$) のみが与えられた．どのような事例の分布であっても特徴空間上に少なくともある一点を含み，一般性を失わずそれを原点としてよいので，これはモデルが事例分布に関する情報を全く持たない場合と想定できる．言い換えれば，事例への適合性を排除し，純粋に識別性が形成するカテゴリ分布パタンを調べる事になる．具体的なカテゴリ生成の手続きは，まず 2 次元空間上に 50 カテゴリの共分散行列を乱数を用いて生成し[1]，これを元に各カテゴリの最適な平均ベクトルを求める事によって詰め込み最適なカテゴリ群を生成する．最適化は EM アルゴリズムによって行った．まず乱数により生成した $\mu_i^{(0)}$ の初期値とし，$t$ 回目の最大化 (M)-ステップでは無作為に選んだカテゴリ $i$ に繰り返し式 (5) を適用し，得られた $\mu_i^{(t)}$ ($i=1,2,\ldots,N$) より確率 $Q_{ij}, R_{ik}$ を計算し (期待値 (E)-ステップ)．同様に，$t+1$ 回目の反復でも最大化ステップと期待値ステップを計算した．この反復計算の後収束した解を分析に用いた．種類に固有な構造を定量化するために滑らかさ指標 (後述) を用いる事で，モデルの生成した詰め込み最適なカテゴリ分布と，成人の自然カテゴリの特徴評定の結果を比較する．

**即時マッピング**　このシミュレーションでは，モデルは 47 の自然カテゴリについてある程度の事例を観察済みとし，別の 1 つのカテゴリに関しては全く未知であると仮定する．このとき未知のカテゴリから得られた 1 事例を基にして，その未知のカテゴリのばらつきパタン (共分散行列) を予測する事がモデルの課題である．事例への適合度だけでは与えられた 1 事例から共分散行列を推定する事は困難であるので，モデルでは与えられた新奇事例を平均として，詰め込み最適な共分散行列 (式 (4)) によって推定を行った．こうして推定された共分散行列と成人の評定値とを比較する事で，即時マッピングとしての推定精度を分析した．多くのいくつかのモデル研究ではモデルの学習した事例の数が，新奇語汎用 (e.g., Colunga & Smith, 2005) このシミュレーションでの推定精度には，既に学習済みのカテゴリ (既知カテゴリ) の構造が大きく影響する

[1]共分算行列は正定値であるため，2 次元の固有値を対数正規乱数 (平均 0, 標準偏差 0.5) から，回転角 ($0 \leq \theta \leq 2\pi$) を一様乱数から生成した．

と考えられる．そこで，既知カテゴリの平均・共分散行列の推定精度を操作するために，MCDI リストに記載された各名詞の獲得割合 (月齢ごとに 16ヶ月から 30ヶ月まで) に基づいて，モデルが観察する事例の数を統制した．具体的には，標本平均・標本共分散行列がモデルが学習するべき真の共分散行列 (成人の評定値) になるような 2000 の標本を乱数によって生成した．MCDI リストのある月齢のある名詞が 100%の獲得率である場合，モデルはそのカテゴリの 2000 標本のうち 50 標本を学習し，0%の獲得率のときは 0 標本を学習したと想定した．このように，仮想的な月齢にあわせた既知カテゴリの事例数に伴って変化するモデルの新奇語の汎用パタンを分析した．

**滑らかさ指標**　滑らかさの指標として，2 カテゴリの平均ベクトル間ノルムと 2 カテゴリの共分散行列間のノルムの相関 $S$(式 6) を定義する．

$$S = \frac{\sum_{i,j<i}\{(|\Delta\mu_{ij}| - |\overline{\Delta\mu}|)(|\Delta\sigma_{i,j}| - |\overline{\Delta\sigma}|)\}}{\sqrt{\sum_{i,j<i}(|\Delta\mu_{ij}| - |\overline{\Delta\mu_{ij}}|)^2 \sum_{i,j<i}(|\Delta\sigma_{ij}| - |\overline{\Delta\sigma}|)^2}} \qquad (6)$$

ただし，$|\Delta\mu_{ij}| = \sqrt{(\mu_i - \mu_j)^t(\mu_i - \mu_j)}$，$|\Delta\sigma_{ij}| = \sqrt{tr((\sigma_i - \sigma_j)^t(\sigma_i - \sigma_j))}$ はそれぞれカテゴリ $i, j$ の平均ベクトルのユークリッド距離 ($L_2$ ノルム) と共分散行列のユークリッド距離 ($L_2$ ノルム)，$|\overline{\Delta\mu}| = \frac{2}{n(n-1)}\sum_{i,j<i}|\Delta\mu_{ij}|$，$|\overline{\Delta\sigma}| = \frac{2}{n(n-1)}\sum_{i,j<i}|\Delta\sigma_{ij}|$ は平均ベクトル距離の平均，共分散行列距離の平均である．

**名詞カテゴリの特徴データ**　日本人の大学生を協力者として得られた 48 の名詞カテゴリの 16 次元特徴による評定データ (Hidaka & Saiki, 2004; Hidaka, Smith & Saiki, 2006) が，特徴選択のシミュレーションに用いられた．48 の名詞カテゴリは 30ヶ月児の 50%が獲得する語彙の記載されたチェックリスト MacArthur Communicative Development Inventory (Fenson et al., 1993) の “animals”, “body parts”, “clothing”, “food and drink”, “furniture and rooms”, “outside things”, “small household items”, “toys”, “vehicles”の 9 カテゴリから選択した．

## 結果・考察

図 4 に詰め込み最適化によって生成されたカテゴリの分布の一例を示す．図の各点が各カテゴリの平均位置，楕円は一定の共分散を表す等高線 ($\sigma/20$) を表す．生成されたカテゴリ分布では，各カテゴリの近傍にはそれと類似の共分散パタンを持つカテゴリが配置された傾向がある．成人の特徴空間におけるカテゴリ分布の滑らかさ (図 2; 滑らかさ 0.466) と同様に，滑らかさ指標を用いてモデルの生成したカテゴリを分析したところ，同程度の滑らかさを示した (図 5; 平均 0.434, 標準偏差 0.119, $N = 10$)．この結果はモデルに仮定した識別性に起因している．従って，成人の自然カテゴリの集団における滑らかな構造が，識別性の最適化によって発生する事を示唆する．また，滑らかさとは種類固有性の連続的な変化の尺度であるので，この結果は種類固有性が発生したと言い換えられる．モデルに仮定した識別性は，領域固有的な (すなわち特定の “種類” だけに通用する) 知識・制約ではないにも拘わらず，カテゴリ集団において “種類” に相当する局所的な偏り・

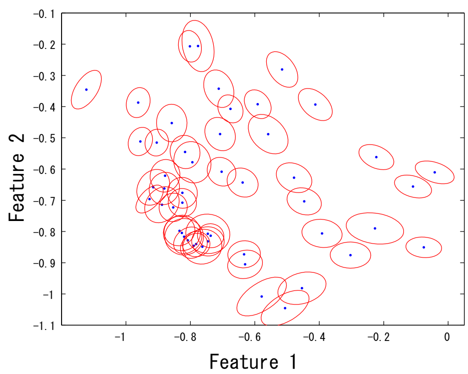


図 4: モデルの生成した詰め込み最適なカテゴリ分布例

クラスタが発生した．次に，モデルによる即時マッピングのシミュレーションの結果について述べる．図 6 に仮想月齢 (16ヶ月から 30ヶ月) における成人の評定値と，新奇カテゴリの分散・共分散の予測値 (点線) および既知カテゴリの分散・共分散の推定値 (実線) の相関係数 (48 カテゴリの平均値，またエラーバーは標準偏差) を示す．仮想 30ヶ月において，分散・共分散の予測値の相関係数は 1%水準で有意であり，それぞれ 0.578, 0.559 であった．また，図 7 に既知カテゴリの平均獲得率 (MCDI 名詞リストより) と滑らかさ指標を表す．

まず，図 6 の結果から，仮想 20ヶ月齢までは，新奇カテゴリの分散・共分散の予測値は既知のカテゴリと同程度の精度を持つ事がわかる．図 7 から，20ヶ月では約 50%の獲得率 (平均の観測事例数にして 25 相当) であり，一方，新奇カテゴリの予測には 1 事例のみが用いられる．これを考慮すれば，モデルの予測は既知カテゴリの精度 (理論的な上限) に近く，非常に高い精度を持つと言える．従って，成人の名詞カテゴリと特徴空間において，提案モデルは即時マッピングに成功したと考えられる．また，新奇カテゴリの予測性は，既知のカテゴリの精度と強い相関があった (事例数と予測相関係数との月齢を通じた相関係数は 0.911, $p < 0.01$)．実際の幼児関する縦断的な研究から，幼児の獲得語彙数と新奇語の汎用パタンとの相関が報告されており (Smith et al., 2002)，モデルの結果はこれに一致する．一方，既知カテゴリの滑らかさは (図 7) カテゴリ間の平均の距離と共分散の距離との相関を表すため，詰め込み最適化ではなく，この関係を利用する事でも新奇語の予測は可能である．この場合，新奇の予測は既知カテゴリの滑らかさ (相関係数) を超えないと考えられる．しかし，各月齢で予測相関値は既知カテゴリの滑らかさを有意に上回っているため，モデルの識別性には，単に既知カテゴリの滑らかな構造を利用する以上に積極的な予測機能がある可能性が挙げられる．

## 結論

本研究では，カテゴリと事例の適合性だけでなく，カテゴリ間の識別性を導入する事で，特徴空間上に漸次的・連続的なカテゴリ分布が発生する事を示した．これは局所的にある種の特徴によってカテゴリ化される “種類” 固有性

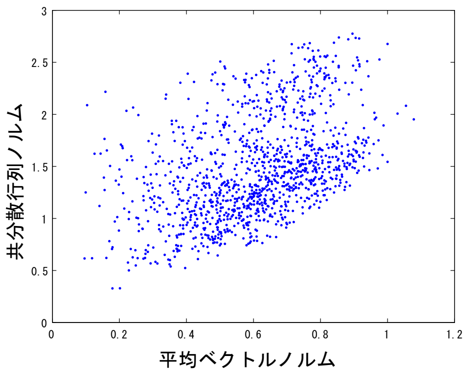


図 5: 全カテゴリ対の平均ベクトルのノルム (x 軸) と共分散行列のノルム (y 軸) の散布図 ($R = .423$).

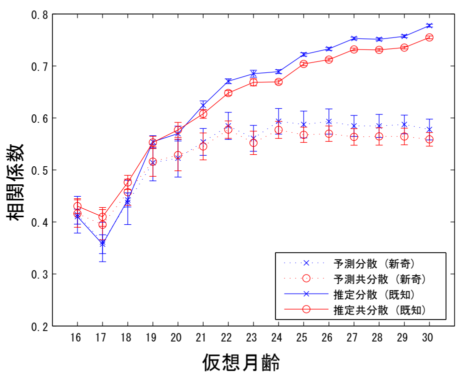


図 6: 成人の評定値と，新奇カテゴリの分散・共分散の予測値および既知カテゴリの分散・共分散の推定値の相関

の自己組織化であると考えられる．また，こうして生成される特徴空間上での大域的な構造を利用する事で，識別性の最適化による新奇なカテゴリの予測性が高まる事が示された．以上の事から，本研究のモデルは幼児の効率的な新奇語の汎用が既知カテゴリの構造的特性および学習における識別性に支えられている事を指摘する．今後の課題としてこの予測を心理実験的手法により検討する事が挙げられる．

## 参考文献

Colunga, E., & Smith, L. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, *112*, 347-382.

Duda, R. O., Hart, P. E., & Stork, D. G. (2000). *Pattern classification (2nd ed.)*. New York: John Wiley and Sons.

Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge: Harvard University Press.

Hidaka, S., Saiki, J., & Smith, L. B. (2006). Semantic pack-

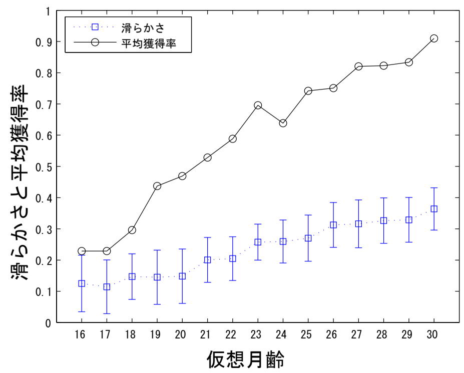


図 7: 既知カテゴリの平均名詞獲得率と滑らかさ

ing as a core mechanism of category coherence, fast mapping and basic level category. In *Paper presented at twenty eighth annual cognitive science society* (p. 1500-1505).

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: universal ontology and linguistic influence. *Cognition*, *62*, 169–200.

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child development*, *62*, 499-516.

Markman, E. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Rips, L. J. (1997). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (p. 21-59). Cambridge: Cambridge University 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., & Smith, L. (1999). Early noun vocabularies: do ontology, category structure and syntax correspond? *Cognition*, *73*, 1–33.

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 terms. *Cognition*, *38*, 179–211.