なぜ自然カテゴリは“自然”か：曲面上に詰め込まれたカテゴリ

日高　昇平 (hidaka@cog.ist.i.kyoto-u.ac.jp)
京都大学, 日本学術振興会

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

Abstract

あるカテゴリは暗示的であるが自然で，またあるカテゴリは厳密に定義されるが役に立たないようにみえる．自然カテゴリの特定の集合はそれに固有の特徴集合をもち，自然カテゴリの大域的な構造は，特異的なカテゴリの集合群により形成される考えられる．本研究では，このような自然カテゴリの凝集的な構造の発生をカテゴリ効率性の観点から説明するモデルを提案する．提案モデルでは，固有の特徴重みを持つカテゴリ群が，特徴空間を隙間無く効率的に利用する事で，位相のそろった凝集的構造が発生する．カテゴリ凝集性の発生に必要な計算論的な条件を追究するために，提案モデルと先行研究のモデルをシミュレーションにより比較した．その結果，提案モデルのみでカテゴリ凝集性が発生する事が示された．また，モデル特性の比較から，局所的な特徴選択の連続的な接続から構成される,「曲面的な」特徴空間における効率化が自然カテゴリを構造化する事が示唆される．

Keywords: カテゴリ凝集性，カテゴリの効率性

自然カテゴリの大域的な構造

なぜ自然カテゴリは我々にとって“自然”なのだろうか．我々が最も頻繁に効率よく利用する概念はどのように構造化されているのだろうか．自然カテゴリの特定の集合は固有の特徴集合と強く関連し，別のカテゴリ集合は別な特徴集合と関連する．例えば,「羽がある」,「巣を作る」,「飛ぶ」などの特徴は,「トリ」と分類される一群のカテゴリ,「スズメ」,「タカ」などに当てはまる．また,「葉をもつ」,「幹をもつ」「実をつくる」などの特徴は「木」と分類される一群のカテゴリ,「ウメ」,「モモ」などに当てはまる．しかし，多くの場合，一つの特徴では厳密にカテゴリの境界を定義できず，多くの特徴を共有する典型的なトリ (e.g.,「スズメ」,「ハト」) から，共有する特徴の少ない非典型的なトリ (e.g.,「ペンギン」,「ダチョウ」) まで典型性にはなだらかな勾配がある (Rosch & Mervis, 1975)．このカテゴリ群と特徴群間の集団による類似構造を家族的類似性 (Family resemblance) と呼び，この性質は多くの自然カテゴリで観察される．一方,「2010 年まで青色で，それ以降は緑であるモノ」(この概念を Goodman (1955) は “grue” と呼んだ) は厳密に定義されたカテゴリである．しかし，このような人工カテゴリは不自然で，役に立たないように見える．

Muphy & Medin(1985) は，互いに関連する一群の特徴とカテゴリ (i.e., 特異的なクラスタ) が，自然カテゴリの全体像を特徴づけると主張し，これをカテゴリ凝集性 (Category coherence) と呼んだ．言い換えれば，カテゴリ群と特徴群が家族的類似性によって凝集し，異なる家族的類似性による多数のクラスタが自然カテゴリ全体の構造を形成すると考えられる．カテゴリの凝集性は，自然カテゴリの特徴選択が極めて困難な問題である事を顕在化する．多くのカテゴリ学習のモデルは，類似性を定義する特徴空間を前提として要するが，自然カテゴリではカテゴリ群固有に重要な特徴があるため，特徴選択は何を分類するかに依存する．しかし，ある事例のカテゴリを知るには，事例の類似性を定義する特徴が選択されなければならない，という循環論に陥ってしまう (詳しい議論は Gelman (1997) を参照)．このような特徴選択の循環性を根本的に解決するには，まずカテゴリ凝集性の発生するメカニズムを説明する必要がある．

本研究では，カテゴリ凝集性の発生を，カテゴリ化効率性のモデルによって説明する．提案モデルでは，カテゴリが高い識別性を保ち，より多くの事例に一般化でき，同時にカテゴリ全体の表現が経済的であるような均衡状態でカテゴリの凝集性を生み出す事を示す．この均衡状態では，あたかも荷物をスーツケースに効率よく詰め込むように，カテゴリを特徴空間に詰め込む効率性が最適化される事から，カテゴリの詰め込みモデルと呼ぶ．カテゴリの詰め込みモデルでは，個々のカテゴリに固有な特徴の重要性を仮定する．この仮定は，言うなれば，直線的・平面的な特徴空間ではなく，曲面的な特徴空間上でカテゴリの効率性を定義する事である．後述のシミュレーションでは，自然カテゴリの特徴空間の特性に関する提案モデルの仮定を検討するために，詰め込みモデルと先行研究のモデルを比較した．以下では，まずカテゴリ凝集性に関して行われてきた研究について述べ，次にカテゴリ凝集性の一面を捉える定量的な尺度と，カテゴリの詰め込みモデルについて説明する．最後に，比較対象とした先行研究について述べ，本研究の理論的な位置づけを行う．

カテゴリの特異性とその連続性

前述のように，自然カテゴリは一つの特徴 (あるいは尺度) によって記述できない特性を持つ．しかし，逆説的であるが，本研究は「単一特徴での記述しにくさ」を記述する尺度 (メタ特徴) を導入し，カテゴリの凝集性を捉える事を提案する．

この提案する尺度を動機付ける，有力な知見が語彙発達の研究から報告されている．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)．新奇語の汎用に関する多くの研究から，自然カテゴリの特徴選択がカテゴリ群に特異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．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). このような知見は，カテゴリ凝集性の一面を表す典型的な例と言える．しかし，それだけではなく，複数の研究が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している (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．

幼児の名詞カテゴリ尤度の大域的なパタンを表す模式図を図 1 に示す．図は形と肌理による 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 (事例の頻度＝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．図の最も左側にある縦長の楕円は，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，図の最も右側にある横長の楕円は，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．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．Hidaka, Smith & Saiki (2006) は，滑らかな性質が，ある特定の特徴 (e.g., 形や材質) だけではなく，広い範囲の特徴に関して見られる事を示唆した．もし，カテゴリが滑らかであれば，ある 2 カテゴリの類似性とその 2 カテゴリの事例のばらつきパタン (カテゴリ尤度) の類似性には相関があるはずである (後述の式 8)．彼らはこの予測に基づき，カテゴリ評定データの滑らかさを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) の相関と定義して検討した．その結果，16 次元特徴をもつ 48 名詞カテゴリの滑らかな性質を持つことが示唆された (図 2)．従って，特徴空間の「滑らかさ」は自然カテゴリの特徴重要性が一定ではなく，局所的に類似である (凝集性) が，大域的には連続して変化する事を捉えるメタ特徴である．次節では，滑らかなカテゴリの発生を説明する計算論的モデルについて述べる．

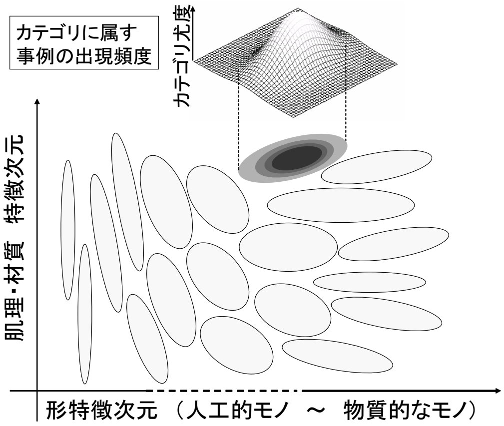

図 1: 滑らかな名詞カテゴリ分布の仮想図

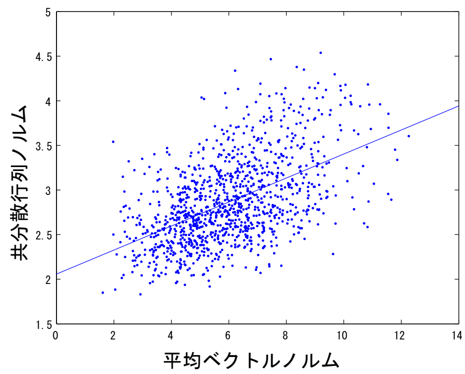

図 2: 全カテゴリ対の平均ベクトルのノルム (x 軸) と共分散行列のノルム (y 軸) の散布図 ($R = .466$).

カテゴリの詰め込み効率

本節では，モデルの理念・概要について述べ，次節で形式的な定義・理論について述べる．本研究では，カテゴリを効率化することで，滑らかなカテゴリが発生するという仮説を提案する．ここで，カテゴリの効率性は,「識別性」,「包含性」,「集約性」という 3 つの “圧力”によって構成されると仮定する．

カテゴリの識別性は，カテゴリの識別誤りで定義され，ある異なる二つのカテゴリに含まれる事例が重複しない場合に高く，含まれる事例が重複する場合に低い．カテゴリの包含性とは，あるカテゴリの参照する “範囲”であり，カテゴリに含まれる事例が多く，また多様な特徴を持つ事例を含むほど高い．カテゴリの集約性とは，カテゴリを表現する特徴空間の “サイズ”であり，特徴次元が低く，特徴空間の領域が小さいほど高い．

これらの三つはトレードオフの関係にあり，全てを独立に最大化する事はできない．この関係の概念的な図式を図 3(a) に示す．図中の楕円は，図 1 と同様に，カテゴリ尤度のあるレベルの等高線なので，楕円が大きいほど個々のカテゴリの包含性が高い．また，3 つのカテゴリ (楕円)

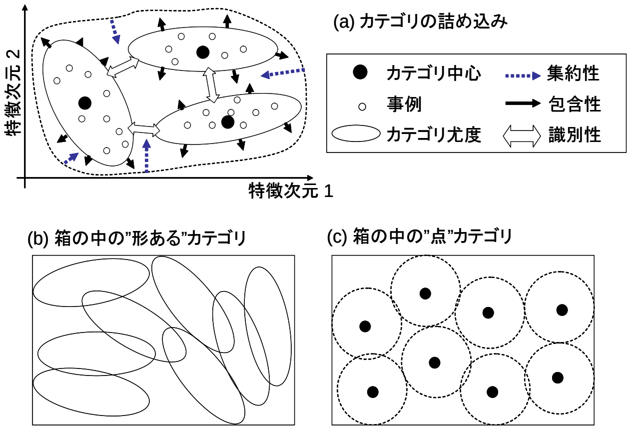

図 3: カテゴリの詰め込み

を囲む点線が，3 カテゴリの集約性を表し，これが小さいほど集約性が高い．図中の白の両側矢印は，二つのカテゴリ間の識別性を意味し，矢印が長いほど，カテゴリ間の事例の重複が小さく，識別性が高い．図 3 の関係のなかで，識別性，包含性，集約性の全てを自由に高める事はできない．例えば，集約性を高める (点線の領域を小さくする) ほど，識別性と包含性は低く (矢印が短く，楕円サイズが小さく) なり，また識別性を高める (白矢印の長さを大きく) ためには集約性と包含性が低くなる (広い領域か，楕円サイズが小さくなる) 必要がある．従って，互いに矛盾する 3 つを任意に最大化する事はできない．しかし，互いを制約として均衡を保ちながら，制約付の最大化は可能である．この均衡的な最適化は，あたかも，できるだけ大きい荷物 (重ならないカテゴリ) をできるだけ小さいスーツケース (特徴空間) に効率よく "詰め込む" 操作に相似している事から，これを本研究ではカテゴリの詰め込み効率と呼ぶ (理論的な定式化は，後の節を参照)．本研究では，詰め込み効率の高い特徴空間ほど，より滑らかな性質を有することをシミュレーションによって示す．

カテゴリ凝集性と効率性

これまでにも自然カテゴリの性質を心理的な「効率性」によって捉えようとする試みは行われてきた．この立場では，カテゴリや事例の類似性や識別性などの特性を組み合わせて一次元的な尺度を構成する事で，カテゴリの「自然さ」を記述する．本研究の詰め込み効率性もその一種である．以下ではこのような定量的な効率性に基づくモデルをメトリックモデルと呼ぶ．過去の研究では，特に基本カテゴリと呼ばれる，特別なレベルの自然カテゴリの記述に関して，多くのメトリックモデルが提案されてきた．基本カテゴリとは，最も人が好んで用いる "自然な" 抽象化レベルと考えられ，最も効率的に処理され，人のカテゴリ処理に最も適すると考えられる (Rosch et al., 1976)．例えば，人は基本カテゴリによって，写真で提示された対象の名前を速く呼ぶことができ，幼児が最も早く獲得し高頻度で発話する名詞でもある．Mervis & Rosch (1981) は，基本カテゴリは最も一般性を持ち，かつ識別性を持つ効率的なレベルであるため，これらの優位性を持つと主張した．また，その効率性のモデルとして，カテゴリ間類似性とカテゴリ内類似性の比を挙げている．このモデルでは，カテゴリ間類似性がカテゴリの識別性，カテゴリ内類似性がカテゴリの一般性を表し，両者の対比から効率性が定義される．

カテゴリ間・内類似性は，詰め込み効率性と類似して，識別性や一般性を組み合わせたメトリックである．しかし，両者が決定的に異なる点は，カテゴリを等質的に扱うか，または個々に特有の "形"(または方向性) を持つと仮定するか，という点である．二つのメトリックの違いを端的に表したのが図 3(b),(c) である．詰め込み効率では，個々のカテゴリは形，厚み，方向を持った "物体" のように扱われる (図 b) が，一方，カテゴリ間・内類似性では，個々のカテゴリは，等質な点のように扱われる (図 c)．詰め込み効率での，カテゴリ尤度 (i.e., 等高線の形) とは，事例のばらつきを表すだけでなく，その分布自体が特徴次元への重み付けを行う．言い換えれば，詰め込み効率は一つのカテゴリの周りで特徴次元の選択が局所的に行われると仮定し，カテゴリ間・内類似性はカテゴリごとではなく，特徴空間全体のみで大域的に特徴選択が行われると仮定するモデルである．本研究では，滑らかな特徴空間を形成するための十分性について，特徴表現の異なるこれらのモデルを検討する．

カテゴリ詰め込みの理論的定式化

本節では，カテゴリの詰め込みモデルを理論的に定式化し，カテゴリの滑らかさとの解析的な関係を要点のみに絞って示す．この定式化では，各カテゴリに含まれる事例を確率分布によって表現し，詰め込み効率を最適化したときの確率分布のパラメタの関係を調べる事で，最適状態においてカテゴリ群が滑らかな構造を持つ事を示す．まず，カテゴリ尤度を正規分布によって定義し，次にその確率分布パラメータによって滑らかさの数学的な定義を示す．

d 次元の特徴空間 $\Omega \supset \theta$ 上で，カテゴリ c_i $(i = 1, 2, \ldots, n)$ が特徴 θ を持つ尤度 $P(\theta|c_i)$ を以下の平均 μ_i，共分散行列 σ_i の正規分布で定義する．

$$P(\theta|c_i) = ((2\pi)^d|\sigma_i|)^{-\frac{1}{2}} \exp(-\frac{1}{2}(\theta-\mu_i)^t\sigma_i^{-1}(\theta-\mu_i)) \quad (1)$$

ただし上付きの ^{t} は転置を表す．平均ベクトル μ_i はカテゴリの中心的な特徴，共分散行列 σ_i はカテゴリのばらつきパタンまたは特徴次元に対する重み付け (Mahalanobis 距離) を表す．従って，確率分布パラメータ μ_i と σ_i によって，カテゴリの滑らかさ (類似のカテゴリが類似のカテゴリ尤度を持ち，非類似のカテゴリが非類似のカテゴリを持つ) は以下のように定義できる．

$$||\mu_i - \mu_j|| \approx h(||\sigma_i - \sigma_j||) \quad (2)$$

ただし，$h(X)$ は X の単調な関数で，$||\mu_i - \mu_j||$ と $||\mu_i - \mu_j||$ は二つの平均ベクトル，共分散行列 i, j のユークリッドノルムを表し，$\approx$ は近似を表す．

次に，識別性，集約性，包含性の三つの制約を定式化し，詰め込み効率を定義する．n カテゴリの識別性は以下のように各カテゴリ尤度に基づいて，カテゴリの誤識別率として与えられる．具体的には，最適な判断基準に

おける最大の誤り率 (Bhattarcheryya 上界) を近似的に誤識別率とする (Duda, Hart, & Stork, 2000)．n カテゴリに拡張した Bhattarcheryya 上界 F_n を以下のように定義する：$F_n = \int_\Omega \prod_i P(\theta|c_i)^{\frac{1}{n}} d\theta$．集約性 E_1 は，カテゴリの表現に用いる特徴空間の範囲に単調であるので，以下のように定義する：$E_1 = \sum_i^n \mu_i^t \mu_i$．包含性 E_2 は，一つのカテゴリの含む事例の範囲であるので，以下のように定義する：$E_2 = \sum_i^n tr(\sigma_i^{-2})$，ただし $tr(X)$ は行列 X のトレースである．詰め込み効率は F_n, E_1, E_2 の組み合わせからなり，集約性，包含性に対する相対的な識別性である．これは制約つきの最適化であるので，λ_1, λ_2 を定数とする Laggrange 乗数 $L_n = 2n\log(F_n) + \lambda_1 E_1 + \lambda_2 E_2$ を最適化する．具体的には以下の式である．

$$L_n = B^t A^{-1} B - C - n\log|n^{-1}A| - \sum_i^n \log|\sigma_i| + \lambda_1 E_1 + \lambda_2 E_2 \tag{3}$$

ただし，$A = \sum_i^n \sigma_i^{-1}, B = \sum_i^n \sigma_i^{-1}\mu_i, C = \sum_i^n \mu_i^t \sigma_i^{-1} \mu_i$ である．

次に，カテゴリ中心 μ_i または尤度パタン σ_i ($i = 1, 2, \ldots, n$) をパラメータとしてメトリック L_n の最小化した場合に，滑らかなカテゴリ (式 2) が発生する事を示す．この最適化の必要条件 $\frac{\partial L_n}{\partial \mu_i} = 0$ と $\frac{\partial L_n}{\partial \sigma_i} = 0$ ($i = 1, 2, \ldots, n$) から，以下の式が導出できる．

$$\frac{\partial L_n}{\partial \mu_i} = \frac{\partial}{\partial \mu_i}(B^t A^{-1} B - C) + \lambda_1 \sum_j^n (\frac{\partial}{\partial \mu_i}\mu_j) \cdot \mu_j \tag{4}$$

$$\sigma_i \frac{\partial L_n}{\partial \sigma_i} \sigma_i = (\overline{\mu} - \mu_i)(\overline{\mu} - \mu_i)^t + nA^{-1} - \sigma_i - \lambda_2 \sigma_i^{-1} \tag{5}$$

ただし，$\overline{\mu} = A^{-1}B$ は μ_i の重みつき平均である．さらに，式 (4), (5) をカテゴリ i, j について展開する事で，以下の近似的に滑らかな性質 (式 2) を導出できる．

$$\Delta\mu_{ij}^t \Delta\mu_{ij} = \lambda_1^2 tr\{\overline{\mu\mu^t}(\lambda_1\sigma_i - I)^{-2}(\lambda_1\sigma_j - I)^{-2}\Delta\sigma_{ij}^2\} \tag{6}$$

$$tr(\lambda_2 \Delta\sigma_{ij}) = tr(\Delta\mu_{ij}(\overline{\mu}_{ij} - \overline{\mu})^t + (\overline{\mu}_{ij} - \overline{\mu})\Delta\mu_{ij}^t) \tag{7}$$

ただし，$\Delta\mu_{ij} = \mu_i - \mu_j$ である．従って，識別性，包含性，集約性の 3 制約による詰め込み効率 (式 3) の最適化は，滑らかなカテゴリ (式 2) を導く事が解析的に示された．

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．このシミュレーションでは，カテゴリ凝集性の発生のために必要な計算論的な条件を絞るために，詰め込みモデルと，それに類似した先行研究のモデルを比較・検討する．前述のように，これらのモデルの大きな違いは，カテゴリを表現する特徴空間の形式にある．もし，先行研究のモデルのように，カテゴリ固有の方向性を無視して (図 3c)，カテゴリの識別性・一般性を最適化するだけで滑らかな構造が発生すれば，詰め込みモデルのように局所的な特徴選択を仮定する必要は無い事が示される．逆に，カテゴリ固有の尤度を考慮する，詰め込みモデル (図 3b) のみで滑らかなカテゴリが発生するのであれば，局所的な特徴選択の必要性が支持される．

具体的には，詰め込みモデルと多クラス判別分析を用いて，これらのモデルが固有に持つメトリックと特徴空間の滑らかさについて分析した．多クラス判別分析 (Multiclass Discriminat Analysis: MDA) は，与えられたカテゴリに対して，カテゴリ間・カテゴリ内の分散比を最大化する特徴次元を選択するモデルである．このモデルの指標は，Mervis & Rosch (1981) の提案する，カテゴリ内類似性に対するカテゴリ間類似性の最大化の一形式であると考えられる．また，主成分分析 (Principal Component Analysis: PCA) も同様に，比較対象に加えた．主成分分析は，データのカテゴリは考慮せず，データ分散の最大化を行う特徴選択の一手法である．本シミュレーションの場合，MDA は名詞カテゴリの識別性を分散比を基準として最大化し，PCA は，カテゴリの別を考慮せず，全ての事例の識別性を分散を基準として最大化する事になる．

一般に，特徴選択のモデルでは，与えられたメトリックを最適化 (最大・最小化) する特徴次元を計算するが，その計算過程では複数の “局所解”が現われる．例えば，MDA では 16 次元の特徴空間に対して，たかだか 16 の局所解が与えられる．局所解それぞれが最適化された特徴空間の 1 次元を表し，特徴次元ごとに MDA の基準を基にしてメトリックの最適性が与えられる．MDA では，特徴次元ごとにカテゴリ間・カテゴリ内分散比の大きさが計算され，この大きさが特徴空間の “良さ”を表すので，これを基準にして元の特徴空間より小さい部分空間を選択する事で，少数の良い特徴次元が選択できる．さて，このように特徴次元にモデルの基準固有に重み付けが与えられるので，この次元を適当に複数選択し構成した特徴空間に，メトリックと滑らかさをそれぞれ計算する事ができる．今，網羅的に，または無作為に特徴次元を選択して空間を構成すれば，良い特徴空間から悪い特徴空間まで様々な特徴空間に関して，それぞれ滑らかさを計算する事ができる．もし与えられたメトリックに関して，良い特徴空間ほど滑らかである，もしくは滑らかな特徴空間ほどメトリックが良いならば，そのメトリックの最適化と滑らかさに関連性があると考えられる．逆に，滑らかさがメトリックに影響されないならば，そのメトリックの最適化が滑らかさに関連するとは考えにくい．従って，多数の特徴空間に関して得られたメトリックと滑らかさの関係を検討した．特徴選択の基データには，48 の名詞カテゴリの 16 次元特徴についての評定データ (Hidaka, Saiki, & Smith, 2006) を用いた．

方法

滑らかさ指標　滑らかさの指標 (式 2 の一形式) 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 8) を定義する．

$$S = \frac{\sum_{i,j<i}\{(|\Delta\mu_{ij}| - \overline{|\Delta\mu|})(|\Delta\sigma_{i,j}| - \overline{|\Delta\sigma|})\}}{\sqrt{\sum_{i,j<i}(|\Delta\mu_{ij}| - \overline{|\Delta\mu_{ij}|})^2 \sum_{i,j<i}(|\Delta\sigma_{ij}| - \overline{|\Delta\sigma|})^2}} \tag{8}$$

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．

詰め込みモデルによる特徴選択　評定データからカテゴリの共分散行列を与え，カテゴリの平均ベクトルを自由パラメタとして詰め込みモデルのメトリック (式 3) を最適化する事で特徴空間を計算した．具体的には，式 4 を解析的に解く事で，768(= 48 カテゴリ ×16 特徴次元) の局所解が得られる．得られた 768 の特徴空間のそれぞれに，詰め込み効率性と滑らかさをを計算し，二変数の散布構造を分析した．

MDA・PCA による特徴選択　評定データの 16 次元特徴空間の全ての評定値と，そのカテゴリ割り当て $(1,2,\ldots,48)$ をモデルに与え，MDA・PCA それぞれの基準を最適化する 16 次元特徴空間の線形変換が計算された．MDA ではカテゴリ間・カテゴリ内分散比，PCA では残差分散が最大化の基準として用いられた (Duda, Hart, and Stork, 2000). MDA，PCA からは，16 次元の評定データに対して，16 次元の局所解が得られる．得られた 16 次元特徴のうち 3,4,5,...,13 次元を無作為にそれぞれ 200 回復元抽出する事で，合計 2200 の異なる特徴空間を作成した．例えば, 3,7,12 番目の特徴次元が選択された場合，この 3 次元のメトリックの平均が選択された特徴空間のメトリックとされた．得られた 2200 の特徴空間のメトリックと滑らかさの散布構造を分析した．

結果・考察

各特徴空間のメトリックと滑らかさの散布図を図 4,5,6 に示す．この散布図は，各モデルの選択基準と特徴空間の滑らかさとの関係を表す．低い詰め込み誤り率 (図 4)，高い分散比 (図 5)，高い残差分散 (図 6) は，各モデルで効率性の高い特徴空間を意味する．点線およびバーは，98%信頼区間を等間隔に区切った 10 区間の詰め込み誤り率の平均および標準偏差を示す．どちらのモデルでも，全ての点に関する滑らかさの平均は正の値であったが，元の評定データは正の滑らかさを持つので図 2)，これは自明な結果である．ここでは平均ではなく，メトリックに対する滑らかさのばらつき・相関の構造を考察する．

結論から述べると，詰め込みモデルだけが滑らかさと明確な関連性を示した．図 4 は 768 の特徴空間の滑らかさ指標 (x 軸) と詰め込み誤り率を表す．図中の低辺に点在する特異的な解を除いて，最も詰め込み効率のよい特徴空間は，最も滑らかな特徴空間と一致する．詰め込み指標の上位 1%(N=8) の特徴空間の平均滑らかさ指標 (0.2890) は全体の平均滑らかさ指標 (-0.0058) よりも有意に高かった $(t(7) = 5.9119, p < 0.001)$．従って，カテゴリの詰め込みによって滑らかな性質が発生する事が示唆される．ただし，散布図全体の構造は非対称な関係性を表している．つまり，高い滑らかさ指標を持つ特徴空間は効率よく詰め込まれているが，逆は成り立たず，効率よく詰め込まれた特徴空間は必ずしも滑らかとは限らない．この結果は，高い詰め込み効率性は滑らかな特徴空間の必要条件であるが，十分条件ではない，という可能性を示唆する．しかし，少なくとも，この結果は詰め込み効率性と滑らかさに強い関連があると解釈できる．

次に，MDA と PCA による結果を考察する (図 5, 図 6). 図 5 と図 6 は，それぞれ MDA の分散比の平均，PCA の残差分散 (寄与率) の平均と滑らかさ指標の散布図を表す. 高い分散比・残差分散を持つ点ほど，MDA・PCA の観点で特徴空間の効率がよい事を意味する．基本的に二つのモデルは類似した傾向を示したので，以下の考察は二つをまとめて行う．この散布図では高い分散比を持つ特徴空間は，高い滑らかさ指標を持つ特徴空間と一致せず，0.2 から 0.3 の中程度の滑らかさ指標を示した．MDA で分散比の上位 1%の特徴空間 (N=22) の平均滑らかさ指標 (0.2295) は全体の平均滑らかさ指標 (0.2518) よりも有意に低かった $(t(21) = -5.6676, p < 0.001)$．PCA で残差分散の上位 1%の特徴空間 (N=22) の平均滑らかさ指標 (0.2414) は全体の平均滑らかさ指標 (0.3254) よりも有意に低かった $(t(21) = -9.9180, p < 0.001)$．一方，分散比の低い特徴空間の滑らかさ指標は幅広い範囲にばらつきがみられた．この構造から，MDA の分散比や PCA の残差分散を基準として，より滑らかな特徴空間を選択する事は困難だと考えられる．

結論

本研究では，カテゴリの詰め込み効率性と他の効率性基準を比べて検討する事で，滑らかなカテゴリの発生メカニズムについて調べた．これらの全てのモデルは，誤り確率やカテゴリまたは事例の分散の大きさを基にして，識別性や一般性をカテゴリ効率として定義する．しかし，カテゴリの詰め込みモデルでは，個々のカテゴリがそれぞれに固有の特徴次元への重み付けを持つが，他のモデルでは全てのカテゴリが等質に扱われている点が大きく異なる．各モデルによる特徴空間の選択を分析した結果，詰め込みモデルのみが滑らかな特徴空間を明確に選択する事が示された．従って，カテゴリ凝集性が発生するためには，カテゴリ固有に局所的な特徴選択を行う必要がある事が示唆される．この結果は，次の重要な予測を導く．すなわち，自然カテゴリを記述するのに適した特徴空間は，等方向的・等質なユークリッド空間ではなく,「曲がった」空間である．自然カテゴリは「滑らか」であるために，類似したカテゴリは，類似したばらつきの構造を持ち，一方で非類似のカテゴリ同士のばらつきは大きく異なる．言い換えれば，近隣の特徴空間同士は，同じユークリッド的な特徴空間に描くことができるが，離れた特徴空間同士は別なユークリッド的な特徴空間上にある．この構造は多様体と呼ばれる曲面上の幾何学を扱う数学的概念に似ている．多様体では，局所的なユークリッド空間の「パッチ」が連続的に張り合わされて全体として歪んだ構造をなす．今後の研究では，自然カテゴリを多様体によって捉える予測が，心理実験的な手法によって検討する事が期待される．

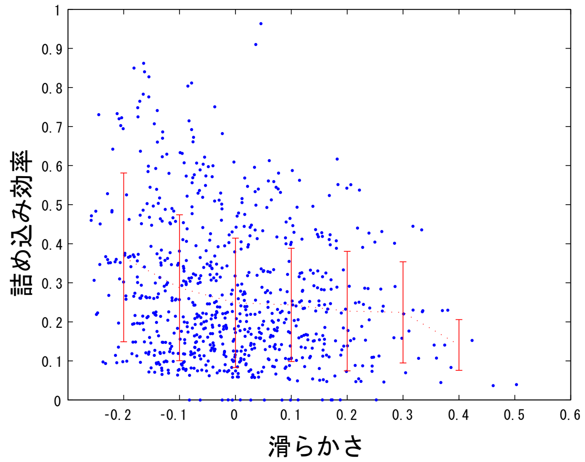

図 4: 滑らかさ (x 軸) と詰め込み誤り率 (y 軸) の散布図.

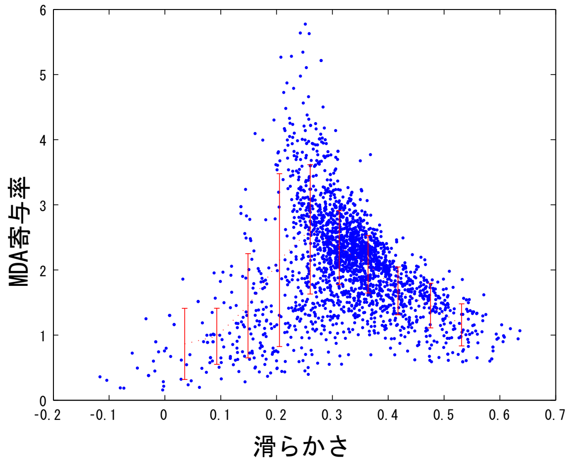

図 5: 滑らかさ (x 軸) と分散比の平均 (y 軸) の散布図.

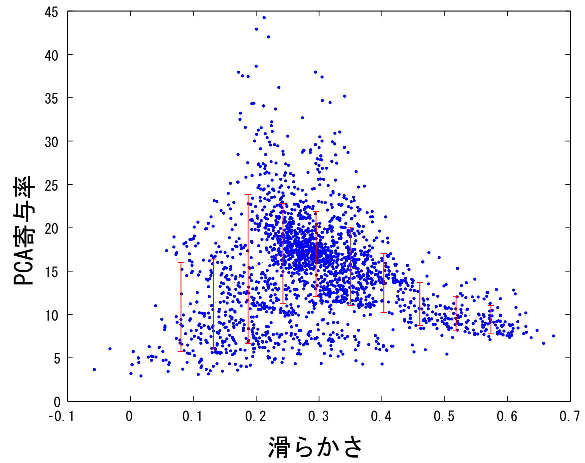

図 6: 滑らかさ (x 軸) と残差分散の平均 (y 軸) の散布図.

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.

Gelman, R., & Williams, E. M. (1997). Enabling constraints for cognitive development and learning: A domain-specific epigenetic theory. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology, volume ii: Cognition, perception and development (vol. 2 5th ed.)* (p. 575-630). 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 packing 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.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Pshycology*, *32*, 89-115.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289-315.

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. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573-605.

Samuelson, L., & Smith, L. (1999). Early noun vocabularies: do ontology, category structure and syntax correspond? *Cognition*, *73*, 1–33.

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.