

論文 / 著書情報
Article / Book Information

論題(和文)	神経意味論における分散表象と意味中枢—fMRIメタ分析のために
Title	Distributed Representation and Semantic Center in Neuro-semantics: For fMRI Meta-analysis
著者(和文)	赤間啓之
Authors	Hiroyuki Akama
出典 / Citation	日本認知科学会第35回大会発表論文集, , , pp. 61-67
Citation(English)	Japanese Cognitive Science Society, , , pp. 61-67
発行日 / Pub. date	2018, 9

神経意味論における分散表象と意味中枢-fMRI メタ分析のために Distributed Representation and Semantic Center in Neuro-semantic: For fMRI Meta-analysis

赤間 啓之[†]
Hiroyuki Akama

[†]東京工業大学リベラルアーツ研究教育院/生命理工学院
Tokyo Institute of Technology, Institute of Liberal Arts/ School of Life Science and Technology
akama.h.aa@m.titech.ac.jp

Abstract

This article argues that the distributed semantic processing system in the human brain can be explored through individual typological differences, based on the dataset of the 2008 Science article by Mitchell et al. on computational neurolinguistics. The crucial finding is that both modality-specific and supramodal semantic areas, which have raised critical issues in neuro cognitive semantics and its meta-analysis, could be extracted from a subgroup of subjects exhibiting mediocre precision by using different feature selection methods (ANOVA and Stability). This result might create new possibilities for human neuroscience by interlocking a single-subject analysis with methods for evaluation of individual variability and fMRI meta-analysis with collective data sources associated with various types of external knowledge system.

Keywords — fMRI, MVPA, feature selection, individual variability, semantics.

1. はじめに

人間の言語活動をめぐり、脳機能イメージングの領域に機械学習を適用し、言葉の意味処理の神経基盤を求める研究は、脳の中に辞書を地図として描く段階にまで進んでいる[1]。しかし、意味処理の神経基盤を求める研究は現在、二つの限界に直面していると言って過言でない。それは1) voxels レベルの情報が個人的に異なる分散表象を為すという個人差の捉え方の問題と2) モダリティ(感覚種や運動、感情などの情報)に依存しない意味処理の神経中枢をどこに同定するかという言語活動の本質論的な問題である。本研究では、Mitchell et al の Science 論文[2]に始まる計算神経言語学的アプローチを新たな視点で再分析することで、この限界を同時に扱おうとする具体的なメタ分析の例を提示したい。

2. 計算神経言語学とその限界

Mitchell ら[2]は、特定の単語刺激に対する神経反応(脳機能的画像データ)と、それらの単語の言語ドキュメントコーパスにおける特徴情報(意味素性と呼ばれ

る基本的語彙との共起データ)を取得し、前者を従属変数、後者を独立変数とする回帰モデルを計算すれば、神経反応情報のない未知の単語による脳の賦活も予測できることを示した。この手法は計算言語学(自然言語処理)と脳神経科学を統合する画期的なものでありながら、そのアルゴリズムはきわめてシンプルである。Mitchell らの提示した下の回帰式で、 v は voxel であり、左辺 y_v は v の(実験刺激に対応する)賦活値ベクトル、 n は意味素性の個数(ここでは基本動詞 25 個)、 $f_i(w)$ は、fMRI 実験で刺激として使われた名詞 w (60 個の名詞からなる)と i 番目の意味素性のコーパス(ここでは Google Web 1T 5-gram Version 1)における共起確率(関連性の強さの指標)、 c_{vi} は意味素性 i が voxel v の賦活に貢献するスカラーパラメーターの推定値で、 c_{vi} 以外は観察値である。

$$y_v = \sum_{i=1}^n c_{vi} f_i(w)$$

Science 論文での実験参加者は計 9 人(P1, P2, ..., P9)であり、彼らは 60 個の名詞の指示対象がそれぞれ描かれたキャプション付きのデッサンを見ながら、fMRI スキャナーの中で対象の典型的属性などを想起する。この意味処理タスクは、刺激の提示順をランダムに変えて計 6 セット行われる。Mitchell らは、上記の回帰式を解くために、この 6 セットを通じて安定した賦活を示す voxels を、刺激に関する情報を有していると考えられる素性として、交差評価のステップごとに 500 個選択した。この素性選択法は安定性(Stability)指標と呼ばれ、それぞれの voxel ごとに、セット*名詞という形の賦活値行列を用意し、各セット間のコサイン類似度による一対比較をすべて加算して、500 個の最大 Stability 値を記録した voxels を素性として選択する。

交差評価は、60 個の名詞から 2 個外して、そのつど選択された 500 個の voxels に対し、上記のモデルで機械学習を行い、賦活推定値と賦活実測値の間で相関の組み合わせを 4 通り計算した上で、同一の名詞間の相

関係数の和が異なる名詞間の相関係数の和より大きい場合に、正解とする(これを leave-two-out cross-validation という)。このステップは 1770 (=60C2) 回反復され、正解の割合をもってこのモデルの参加者ごとの精度とする。ただし、Mitchell らの論文に明確には書かれていないが、この回帰モデルには正則化 (regularization) 項が加えられ、その結果、参加者ごとの精度は、それぞれ 0.83, 0.76, 0.78, 0.72, 0.78, 0.85, 0.73, 0.68, 0.82 (平均 = 0.77) で、すべて有意であると判明した。ただし Akama [3] が示したように、正則化項を加えず、単純に古典的最小二乗法で計算すると、0.82, 0.78, 0.78, 0.78, 0.75, 0.64, 0.71, 0.6, 0.65 というように、P6, P8, P9 の精度がかなり落ちる。

このように重回帰モデルを用いて、既知の単語共起データと一部の脳反応データをもとに未知の賦活予測をする機械学習モデルは、自然言語処理と fMRI の多変量パターン分析 (Multivariate pattern analysis, 略して MVPA) を組み合わせて計算される。Mitchell らに始まるこの手法は、Murphy らにより、計算神経言語学 (Computational Neurolinguistics) の名で呼ばれ [4]、Huth らの Nature 論文 [1] につながって、現在も発展を続けている ([5], [6], [7], [8], [9])。しかし、そもそも、計算神経言語学的アプローチは、本質的に単一主体分析 (single subject analysis) が基本であり、そこから一般線形モデル (General Linear Model, GLM) のような機能的解剖学の結論を引き出すことはきわめて難しい。この分野では、実験参加者「間」のモデリングに限界があるため、実験参加者ごとの計算結果を並べ、モデル化の有効性や情報発現の類似傾向を示すという形で示すのが通例である。Mitchell ら以降も、正則化により参加者全員に対して統計的に有意なモデルを構築できたことから、個人差の問題はまったく議論されていない。そして方法論的な制約上、fMRI メタ分析もそこから個人間の偏差に関する知見を引き出すことは難しい。

個人差の問題と密接に関連するのは、それぞれの単語の意味処理が行われる voxels と領域が脳の全域にわたって分散しており、特定の少数領域に集中していないという、意味記憶 (semantic memory) の神経基盤に見られる特徴である。これは分散表象 (distributed representation) と呼ばれており、データの平滑化 (smoothing) を行わない MVPA の場合は、機械学習のために選択された素性 (voxels) が胡麻塩状に散らばる状況が見られる (Raizada ら [10] はこれを salt and

pepper effect と呼んでいる)。意味処理の分散表象を局所的にせよ解釈する場合、有効なひとつの方法は、embodiment 認知理論を適用することである。Embodiment とは、単語や文の意味理解には、その内容と関連する脳の知覚・運動系が関与・作動することから、あたかも言語の理解者が仮想的にその状況を再現・シミュレートするようなメカニズムが必要となるという考え方である。Mitchell et al. も意味素性のひとつである”eat” という動詞の賦活推定画像において、味覚領 (gustatory cortex) に大きな重みを計算出力していることで、fMRI メタ分析の知見に照らし、この数理モデルの妥当性を、Embodiment という側面から裏付けている。

Embodiment におけるモダリティの多様性は、意味処理の分散表象の内部の多様性と表裏一体になり、神経意味論の混迷を演出する要因となっている。さらにまた、知覚、運動、感情などの様々なモードを超えて意味処理一般を担う、純粋な言語活動の領域、いわゆる意味中枢に関しても、議論の対立が見られる。左側頭極の役割を強調する Hub and Spoke モデル (Patterson ら [11][12])、左中側頭回、左上側頭溝から左角回までをクロスアップする High level convergence モデル (Binder ら [13][14]) など、意味中枢に関しては、左中側頭回まわりを除き諸説があり一致しない。また古典的な問題として、言語刺激のモダリティを担う領域で既に意味判別の処理は完了しているのではないかという見方もあって否定できない。意味中枢に関する議論は、fMRI 研究におけるメタ分析が不可欠であるが、コンセンサスの欠如は、まさしく、メタ分析の対象自体の存在意義に対する懐疑の念から発している。Patterson らが主張するように、左側頭極など側頭葉前部は、磁化率の異なる組成が入り組み、fMRI では信号を検出しにくい領域である、というのがその根拠である [11]。いずれにせよ側頭葉を中心とした意味処理領域において、中枢が吻側と尾側の間で動揺しながら、Embodiment の様々なモダリティ (感覚種や運動、感情などの情報) を担う領域との間でネットワークを形成し、分散処理を行っているとするのが、神経意味論の概略的な見取り図であろう。その場合、MVPA の機械学習で素性として選択された voxels が、脳内のどの領域に位置しているかを、まず実験参加者ごとに精査することが重要な意味をもつ。

3. 計算神経言語学の再分析とメタ分析

本研究の著者は、前節で示した計算神経言語学の弱点である個人差と分散表象という観点から、Mitchellらの公開データを対象に、素性選択にまで立ち戻って再分析・メタ分析を行っている[3]。すなわち、脳の賦活から意味のある素性(voxels)を抜き出すという機械学習前のステップで、実験参加者の脳がどのように個性的に働いているかを明らかにした。そこでの発見は、正則化を入れる前の古典的回帰のみによるモデル精度の差が、まさしく言語の意味処理における個人の脳活動の質的な差に対応しているということであり、それを明らかにした再分析の手法を以下にまとめる。

まず、[3]では、素性(voxels)選択の方法として、Mitchellらを始めとして多くの計算神経言語学が採用する安定性(Stability)指標に加え、Pereiraら[15]が提案するANOVA-f値も同時に計算し、voxelごとに、二つの指標が持つ値の順位を比較・分析した。ここで言うANOVA-f値とは、刺激となる名詞ごとに6回の提示による賦活値の群を形成し、その群間差異(平方和)を群内差異(平方和)に対して評価した値である。

結果として、選択された500個のvoxelsを対象とした実験参加者別の箱ひげ図(上:Stability値、下:ANOVA-f値)が示すとおり(図1)、正則化項無しでも良好なモデル精度は、高いStability値、ANOVA-f値と関係することがわかった。特筆すべきは、P1がこれら素性値の面で、それに続くP2-P4のグループに比しても抜きん出て高いということである。P1のみが、正則化項なしでも精度にほぼ変化が見られないことと、この特異な素性値分布との間には、何らかの関連性があることが予想される。さらに素性値の中位グループ(P5, P6, P7を候補とする)を詳しく見ていくため、Stability値とANOVA-f値により選択された上位500個のvoxelsについて、スピアマン順位相関を計算したところ、P5とP7は有意な相関を記録したが、P6に対しては有意でなく、かつ、P6はP5, P7と異なり、OLSのみによる精度が70%以下、そして両指標で選択されたvoxels間の食い違いが100以上見られた。このことから、あらためてP6はP8, P9とともに最下位グループを形成すると見做した。

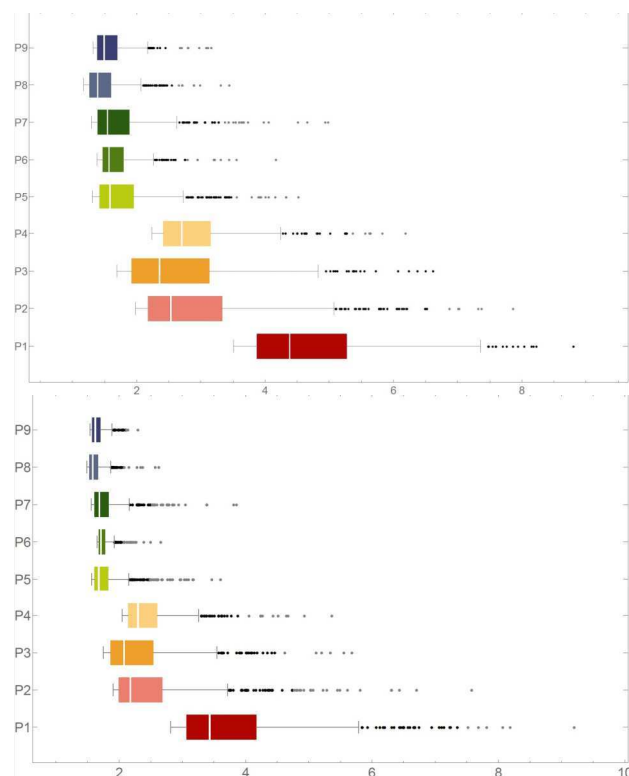


図1 選択された素性 voxels の Stability, ANOVA-f 値

このようにモデル精度の低い実験参加者では、素性選択において安定性(Stability)指標とANOVA-f値により素性選択されたvoxelsの食い違いが際立つ。しかし、後述するように、その解剖学的位置には、素性voxelsの選択法の違いに敏感なタイプと見做せるような、まさしく共通した特徴が認められ、しかもその系統的特徴は神経意味論の諸説の対立に通底するものであることが判明した。

興味深いことに、OLSのみによるモデリング精度が70%以下の参加者P6, P8, P9(下位グループ)では、安定性(Stability)指標とANOVA-f値で選択した上位500個のvoxelsのうち、100以上のものが異なっており(それぞれ183, 146, 124)、食い違い数は全体の平均より1標準偏差だけ異なっていた。以下の図2は、P1, P6を例に、両指標で選択されたvoxelsを、それらが持つ各指標での順位を座標として二次元平面にプロットしたものである(全実験参加者についての図は[3]を参照のこと)。各実験参加者に対応するグラフで、赤の水平線で区切られたものは、Stabilityで選択されたvoxelsのANOVA-f値における順位、赤の垂直線で区切られたものは、ANOVA-f値で選択されたvoxelsのStabilityにおける順位を表している。各指標で選択されたvoxelsの食い違いは、P1ではわずか13個と最小であるのに対し、P6では183個と最大値を記録している。これらの

食い違う選択結果は、図では赤線の外側(上もしくは右)に分布するドットによって示される。

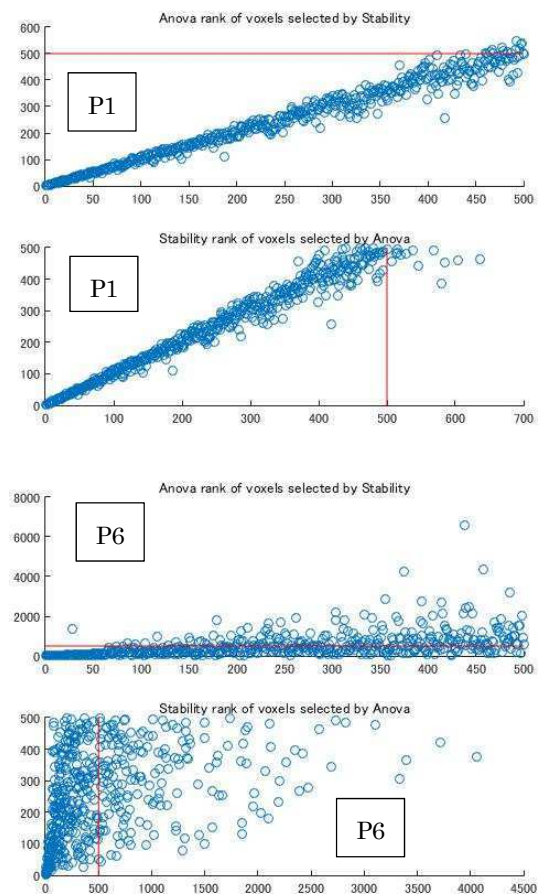


図2 素性 voxels の Stability, ANOVA-f 値の順位

P6, P8, P9(下位グループ)では、指標の排他的特有性を示す素性 voxels が、参加者ごとに 100 以上見られるので、それらに絞って空間的な分布の特徴を探っていく。そのため AAL(Automated Anatomical Labeling) [16]の脳地図を利用し、この3者において排他的特有性を示す voxels を網羅的に取り上げ、それらの解剖学的なラベルを検索し、これをもって排他的特有素性領域とした。さらに下位グループに共通する排他的特有素性領域を抽出し、そこに帰属する排他的特有素性 voxels の出現個数を指標ごとに計算した。

その結果、下位グループでは、安定性(Stability)指標で選ばれるが ANOVA-f 値では選ばれない素性 voxels(これをタイプ 2 と呼ぶものとする)の多くは、中後頭回や紡錘状回など視覚モダリティに関連している領域から抽出されていることが判明した。この指標は Mitchell らが採用したものだが、いわゆる言語の意味処理関連の領域に属する素性 voxels はほとんど見出せなかった。反対に ANOVA-f 値で選ばれるが安定性(Stability)指標では選ばれない素性 voxels(これをタ

イプ 1 と呼ぶ)は、左中側頭回を筆頭に左弁蓋部など言語の意味処理領域からも多くが抽出されることがわかった。タイプ 1 とタイプ 2 の存在領域の差は際立っており、重なる部位がほとんど存在しない。このような現象は、モデル精度の良好な実験参加者では、ふたつの素性選択法で結果がほぼ変わらないため見出せなかった。しかし、モデリングの最高精度(82%)を記録する P1 を見ると、両指標に共通する 487 個の素性 voxels は、そのほとんどが後頭葉・側頭葉の視覚領域に集中し、左下前頭回・三角部のような古典的言語野からは選ばれていないことがわかった。

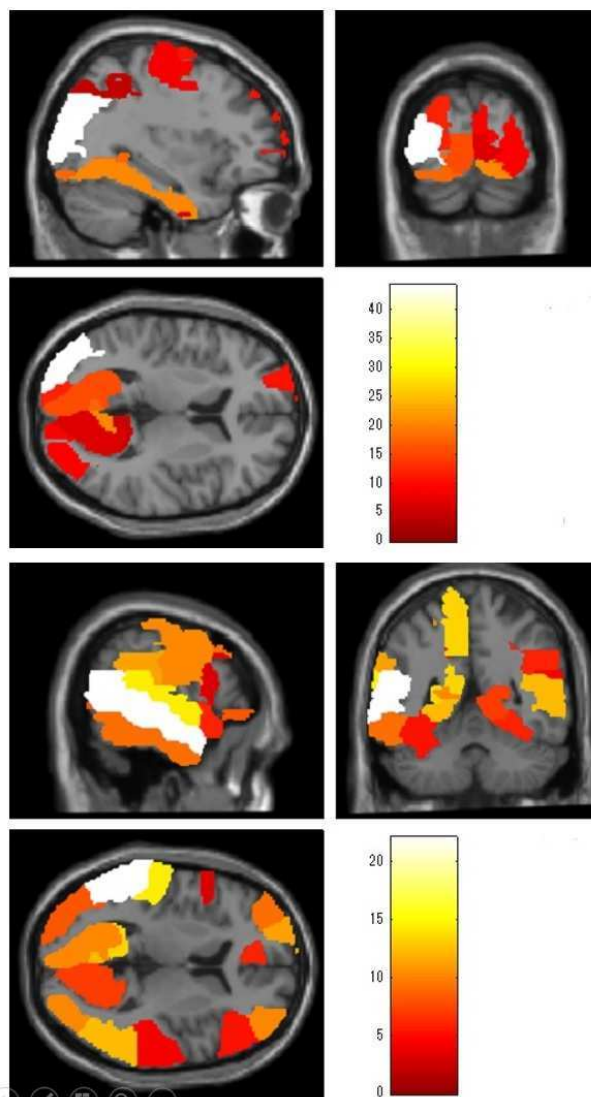


図3 素性 voxels の存在領域(下位グループ)

図3は、下位グループ内で共通な素性 voxels の存在領域に限定し、タイプ 2(上)、タイプ 1(下)の素性 voxels が何個 AAL の領域に出現するかをカラーマップで表したものである。それぞれの右下カラーバーで示したとおり、最大帰属数を記録する領域は白で塗られており、Stability でのみ選ばれるタイプ 2 の素性

voxels が中後頭回(Occipital_Mid_L)まわりに集中するのに対し、ANOVA-f 値のみ選ばれるタイプ1の素性voxels が左中側頭回(Temporal_Mid_L)に多く見られることが理解される。

こうした結果はメタ分析の視点からも重要な意味をもつ。言語の意味処理領域に関する Binder らのレビュー論文[14]が示すとおり、embodiment 認知におけるモダリティ(感覚種)の中で、視覚は知覚シミュレーションなどを通じ、超感覚種(supramodal)な意味処理への相互作用を行うという位置づけが可能である。むしろ Mitchell ら[2]が抽出した視覚システムに偏る素性選択は、言語の意味理解の深いレベルにまで降りてゆかず、単に表面的な視覚刺激の弁別だけに関与しているという見方もありうる。ただし、先述したように、Mitchell らが“eat”という動詞の推定神経反応に味覚領を見出した embodiment 解釈は、その延長で下位グループに見られるような、Stability 指標と感覚モダリティとの親和性(タイプ2)においても有効である可能性が残されている。その場合、安定性(Stability)とは、その計算定義からして概念の「同一性」を指向するものなので、刺激の参照する指示対象のインスタンスと関わりあう具体的な感覚知覚体験を反映しやすいという見方も考えられる。一方、下位グループにおいては、ANOVA-f 値と結びつく素性 voxels(タイプ1)は、Binder ら[14]の言う超感覚種(supramodal)な意味中枢、すなわちモダリティ固有の情報が高次元で収斂する脳領域に多く見出される。その代表が、左中側頭回であることは、メタ分析間でもさほどの異論はなく、ここに論理的、抽象的、超経験的な意味処理の座を求められることは、ANOVA-f 値の計算定義からしても十分に整合的である。なぜなら安定性(Stability)指標と異なり、ANOVA-f 値はここで言語象徴間の「差異」(分散)の大きさを評価しているからである。現代言語学の父、ソシュールの発想を継承する形で、言語の記号体系を差異の系統的関係によってとらえ、超感覚種(supramodal)な意味計算を行ううえで、fMRI のメタ分析が明らかにする高次元収斂領域の果たす役割には大きな意味があるだろう。

むしろ実験参加者の個人差から素性選択間の矛盾を介し脳科学の認識論、ひいては哲学まで根源的な議論を拓けていくには、まだ多くの情報が欠如していることは言うを待たない。ただし、こうした議論の端緒になるのが、fMRI の機械学習モデリングにおいて相対的に精度が「低い」実験参加者のデータであるということ

は、注目して良いことだと思われる。すなわち、弱い賦活コントラストしか示さない実験参加者の特徴に、意味論の様々な関心領域が、異なる素性選択法を通じて体系的に出現すると言う、興味深い現象が起きている。しかもそれらは、データのスムージングを行う以前の段階で、意味ある現象として記録されている。このことは、脳機能イメージングにおける、所謂「優秀でない」、「ノイズの多い」、「悪い」データのうちに、「良い」データが隠してしまう重要な情報が潜在している可能性を示唆している。この可能性は、ひいては脳機能イメージング研究のあり方そのものの問い直しにつながるのではないかとさえ思われる。

5. 考察と将来の展望

このように、脳の賦活予想を行う機械学習のモデリングは、個別の実験参加者レベルで完結するばかりでなく、正規化などの手法適用以前のデータに、ヒトの脳が細部でいかに個性的に働いているかをめぐり、かなり本質的な情報が潜んでいることが明らかになった。こうした細微な情報を活用し、個人差のモデル化を指向することは、バイオマーカーの設定と言う医療健康科学の文脈のみならず、言語機能の解明においても大いに資する点があるように思われる。脳反応の推定モデルを計算する多変量パターン分析(MVPA)では、情報の大きい領域やタスクへの応答性が個人で異なるため、横断的な実験参加者「間」脳反応解読モデル(一人の参加者で計算した推定モデルを他の参加者に適用するもの)の精度は、参加者「内」の精度に較べては十分ではない。たしかに参加者間分析の精度低下を克服するため、hyperalignment[17]や Joint Rank Feature selection[18]を始めとする様々な手法が考案されている。参加者「間」モデルの不具合は脳の構造面、機能面での個人差のどちらから、どの程度起因するのか、二つの要因間の重み付けがいかにか可能かなど、様々な問いにつながるものではある。しかし、そもそもそうした観点では、個人差をあくまでペナルティーの面から解釈するだけで、積極的な情報源としてとらえる指向性が見られない。むしろ、本研究のように、脳反応をめぐり個人差が、そこに類型を仮設定でき、個人差のモデリングにつながりうるものが予見される以上、貴重な情報源として精査し活用するのが生産的ではないだろうか。すなわち、個人差の分析を通じ、ヒトの脳をめぐる機能と構造の間のミッシングリンクを埋める可能性が開かれるという視点である。

良く知られているように、Marr は脳をコンピュータとの対比によって定義し、ハードウェアにおける実装の階層、システムの機能を扱う計算論の階層、その内部処理のメカニズムを扱うアルゴリズムの階層に分けた上で、それらを統合することによる脳研究の必要性を説いた。しかし、ヒトの認知的活動を神経システムに根差したものとして統合的に考えるに当たり、Barsalou[19]の示唆するように、脳の機能と構造を集団レベルで対応付けるマッピングだけでは、アルゴリズムの階層、特に一人ひとりの脳がどのように個性的に働いているかを明らかにすることはできない。Embodiment 認知理論の泰斗である Barsalou のこの議論は、Mitchell の回帰モデルを応用した Huth ら[1]の意味的タイリング(semantic tiling)に対する痛烈な批判の形で展開されている。Huth らの目標は各個人の脳の中に、単語や同系列単語群とその意味処理の位置を結合させ、三次元で視覚化された脳内辞書(mental lexicon)を描くというものである。この MVPA 研究において、Huth らは、意味素性—Barsalou はトピック語(topic words)と呼んでいる—と fMRI の刺激語の共起ベクトル—Barsalou は基底関数(basis function)と呼んでいる—から、集団レベルで4つの意味次元を有意な分散の大きさを示すものとして抽出している。しかし、それは全体の情報量の20%しかなく、しかも個人レベルで、基底関数に対応する神経活動を示すものとして抽出された場合、それらの主成分はたかだか35%にしか相当しない。そのため、意味処理の脳反応はきわめて強力な個人差を示すと、Barsalou は喝破している。

脳反応における個人差に体系的にアプローチするには、むしろ背景となるサンプルに十分なサイズが必要である。しかし、神経意味論に絞っても MVPA 研究での実験参加者数は、機械学習データの voxel レベルでの単変量化を除いて、せいぜい一桁を超える程度の場合が多い。だが、本研究で明らかになったのは、Mitchell らのデータ再分析が示すように、わずか9名の参加者数でもそこに類型化の兆しが見られるということであった。そこでまず考えられるのは、脳情報における個人差を、脳科学外の遺伝学的・生理学的・心理学的情報によって説明し、個性化プロファイルを行って、そこから機械学習モデリングの素性を選択するという方向性である。さらに、ここで想起すべきは、計算神経言語学が、そもそも、少ない個人の神経反応の記録を、集団的の反応の記録である言語コーパスからの共起・関連性情報と接合させ、データが存在しない条件での脳

の賦活を予測するという設計概念である。もし、fMRI データの相対的希少性にもかかわらず、神経反応の一次レベル(参加者レベル)で既に構造化の見通しがつくのであれば、そこに言語コーパス情報と同じく、外部の集団的知識体系を、しかも直接的に導入することの有効性が想定できよう。そのような外部、あるいは間接化された集団的知識体系とは、fMRI のメタ分析の結果にほかならない。Mitchell らのデータの再分析で明らかになった個人差には、言語の意味処理に関する脳神経科学のメタ分析上の矛盾として解釈できるものが垣間見られた。そもそも fMRI のメタ分析をさらに高次のメタレベルで精査すれば、脳の機能と構造のそれぞれのレイヤーを繋ぐ、様々な情報のネットワークを計算することが可能である。もし意味素性を中間の入出力層、すなわち Barsalou の言うアルゴリズムの階層において、メタ分析と言語コーパスという二つの集団的的外部知識ネットワークを接合すれば、われわれは Mitchell や Huth らの計算した個人別の差のある神経言語反応に関し、一統的な集団的学習ネットワークを踏まえ—そのパラメータ値の多様な分布、あるいは部分的な設計構造の変異としてシミュレートすることが可能になるのではないか。ただし、これは将来の課題であり、開かれた問いとして提起するにとどめておくものとする。

参考文献

- [1] Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL, (2016) Natural speech reveals the semantic maps that tile human cerebral cortex, *Nature*, Vol. 532, pp. 453–458
- [2] Mitchell TM, Shinkareva SV, Carlson A, Chang KM, Malave VL, Mason RA, Just MA, (2008) Predicting Human Brain Activity Associated with the Meanings of Nouns, *Science*; DOI: 10.1126/science.1152876.
- [3] Akama H, (2018) Individual typological differences in a neurally distributed semantic processing system: Revisiting the Science article by Mitchell et al. on computational neurolinguistics, *F1000Research*, 2018, doi: 10.12688/f1000research.14584.1
- [4] Workshop on Computational Neurolinguistics, NAACL HLT 2010 Workshops <https://sites.google.com/site/compneurowsnaacl10/>
- [5] Bullinaria JA, Levy JP, (2013) Limiting Factors for Mapping Corpus-Based Semantic Representations to Brain Activity, *PLoS ONE*; doi:10.1371/journal.pone.0057191
- [6] Anderson AJ, Binder JR, Fernandez L, Humphries CL, Conant LL, Aguilar M, Wang X, Doko D, Raizada RDS, (2016) Predicting Neural Activity Patterns Associated with Sentences Using a Neurobiologically Motivated Model of Semantic Representation, *Cerebral Cortex*; doi: 10.1093/cercor/bhw240

- [7] Pereira F, Detre G, Botvinick M, (2011) Generating Text from Functional Brain Images, *Frontiers in Human Neuroscience*; DOI: 10.3389/fnhum.2011.00072
- [8] Akama H, Miyake M, Jung J, Murphy B, (2015) Using Graph Components Derived from an Associative Concept Dictionary to Predict fMRI Neural Activation Patterns that Represent the Meaning of Nouns, *PLoS ONE*; DOI: 10.1371/journal.pone.0125725
- [9] Devereux, B., Kelly, C., Korhonen, A, (2010) Using fMRI activation to conceptual stimuli to evaluate methods for extracting conceptual representations from corpora. *Proceedings of First Workshop on Computational Neurolinguistics, NAACL HLT 2010*, pp.70-78.
- [10] Raizada, R.D.S. and Lee, Y.S. (2013) Smoothness without smoothing: why Gaussian Naive Bayes is not naive for multi-subject searchlight studies. *PLoS ONE*, *PLoS ONE* 8(7): e69566. doi:10.1371/journal.pone.0069566 PDF.
- [11] Patterson K, Nestor PJ, Rogers TT, (2007) Where Do You Know What You Know? The Representation of Semantic Knowledge in The Human Brain, *Nature Reviews Neuroscience*; DOI: 10.1038/nrn227
- [12] Rice GE, Lambon Ralph MA, Hoffman P, (2015) The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies, *Cerebral Cortex*; DOI: 10.1093/cercor/bhv024
- [13] Binder JR, Desai RH, Graves WW, Conant LL, (2009) Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies, *Cerebral Cortex*; DOI:10.1093/cercor/bhp055
- [14] Binder JR, Desai RH, (2011) The neurobiology of semantic memory, *Trends in Cognitive Science*; DOI: 10.1016/j.tics.2011.10.001
- [15] Pereira F, Mitchell T, Botvinick M, (2009) Machine Learning Classifiers and fMRI: A Tutorial Overview, *NeuroImage*; DOI: 10.1016/j.neuroimage.2008.11.007
- [16] Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M, (2002) Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain, *Neuroimage*; DOI: 10.1006/nimg.2001.0978
- [17] Haxby JV, Guntupalli JS, Connolly AC, Halchenko YO, Conroy BR, Gobbini MI, Hanke M, Ramadge PJ, (2012) A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*; doi:10.1016/j.neuron.2011.08.026
- [18] Akama H, Murphy B, Lei MM, Poesio M, (2014) Cross-participant Modelling Based on Joint or Disjoint Feature Selection: An fMRI Conceptual Decoding Study, *Applied Informatics*, Springer, doi:10.1186/2196-0089-1-1
- [19] Barsalou, LW, (2017) What does semantic tiling of the cortex tell us about semantics?, *Neuropsychologia*; doi:10.1016/j.neuropsychologia.2017.04.011.