

## 弱教師学習を基礎とした文書の編集ナビゲーター構築の試論

### 一抽象絵画鑑賞を例に一

中村潤（中央大学）、永吉実武（静岡大学）

#### 1. 研究背景と目的

一般には理解しづらい抽象絵画を鑑賞し、人間がどのような感想・創造・想像を行ったのかを記述してもらおう。同時に、生成AIに対しても同一のプロンプトを与え、同一の抽象画に対する評価文を生成させる。このようにして、これまで筆者らは、人間と生成AIの文書特徴を比較し可視化を手掛けてきたが、本研究では、これら人間による感想文と生成AIによる評価文の関係を分析し、「どこをどのように修正・補強すればよいか」を示唆する“文書編集ナビゲーター”の構築を目的とする。現時点は、理論的な枠組み（アルゴリズム）の開発段階である。

#### 2. 力学的メタファーの導入

本研究では、自然科学における計算モデルである「ニュートンの法則」(1)および「クーロンの法則」(2)に着想を得た力学的メタファーを社会科学的な文書分析に導入する。これらの法則は、二点間の相互作用が、質量（または電荷）の積に比例し、二点間の距離の二乗に反比例する、という共通の構造をもつ。

$$F_{grav} = G \cdot \frac{m_1 m_2}{r^2} \quad (1)$$

$$F_{elec} = k \cdot \frac{q_1 q_2}{r^2} \quad (2)$$

文章は静的なものというより、互いに影響しあう場の中の存在、という立場にたつと、従来の限界は、コサイン類似度・KL divergenceなどは近い／遠いは言えるが、なぜ引き寄せられるのかは説明できない。また、共起ネットワークや KeyGraph では単語中心で、文章の意味的運動を扱えない、といった課題があった。本研究の新規性は、「質量×質量÷距離の二乗」という力学的構造を、文章間の単なる距離測定ではなく、編集可能性を示唆する意味的相互作用量として再定義した点にある。生成AI文を重力源とすると、感情の多様性がある人間文がどこで引き寄せられて／反動されているかの可視化をすることを目指している。

現時点でのデータは、人間文が30文と生成AI文が1文ある。30の人間文を学習することで、一定の編集ナビゲーターの基盤を構築する

### 3. 意味的影響力 F の定義

ここで人間が記述した文章の意味的質量（後述）を  $m_1$ 、生成 AI が生成した文章の意味的質量を  $m_2$  とする。距離については、文章間の意味的差異の測定値であり、初期段階はコサイン類似度や TF-IDF に基づく既存の文章間距離を用いる。いずれ意味次元別距離（後述）へと拡張する。

### 4. 意味的相互作用行列による可視化

生成 AI による文章を意味的重力源とみなし、人間の文章が各意味次元においてどの程度引き寄せられているか／拒否されているかを、意味的相互作用行列（Semantic Interaction Matrix）としてヒートマップ表示する。この可視化を通じて、抽象絵画を起点とした物語生成・創造的解釈（アイデア）のための“編集ナビゲーター”の構築を試みる。

### 5. 質量 m の解釈（理論的枠組み）

分子にある「質量  $m_i$ 」は、文章が他の文章に及ぼす意味的・解釈的影響の源として解釈される。本研究では、以下の3つの観点から構成される複合概念としてとらえる。

- ① Narrative Attractiveness（物語としての引力）：注意・共感・解釈を引き寄せる力
- ② Interpretive Coherence（解釈の筋の良さ）：説明の一貫性・説得力
- ③ Aesthetic Engagement（美的関与）：感情・比喩・具体性による関与度

### 6. SNEH 意味次元による分解

SNEH は、文章の意味を直接表現する概念分類ではなく、意味的影響力 F を計算・分解するための基底空間として導入される。次のように質量  $m_i$  を  $m_i \in [0,1]$ （大きいほど魅力度・中心性・引力源になりやすい）として設計し、複合スカラーとして考える：

$$M_i = w_1 S_i + w_2 N_i + w_3 E_i + w_4 H_i \quad (3)$$

表 1：SNEH 意味次元の意味

要素	意味（弱教師ラベル）	チェックポイント
Structure (S)	構造的完成度（起承転結）：①②	セグメント長・接続詞
Narrative (N)	物語性・展開度：②	トピック遷移量
Emotion (E)	感情エネルギー：①③	感情極性 x 強度
Human (H)	人間らしさ/主体的解釈：③	人間分 vs 生成 AI 文の識別確率

ChatGPT などの生成 AI は一般に人間らしさ/主体的解釈 (H) が低く、構造的性 (S) が高い一方で、人間文は感情的揺らぎ (E) や解釈の多様性 (N) が大きい傾向をもつことが、先行研究および当研究室での実験により確認されている。これらの特徴を「弱教師あり学習」の枠組みとして利用し、下記に示す(4)式にて相互作用の強さ（引き寄せ・反動の力）の基本モデルを表現する。本研究では、 $F$  をスカラー値としてではなく、SNEH 意味次元に分解された相互作用行列として定義す

る。この相互作用行列は、どの意味次元において人間文が生成 AI 文に引き寄せられているか、あるいは反動しているかを示す「意味的影響力の断面」として解釈される。 $m_i$ は、スカラーではなく、次元分解された SNEH 空間上のベクトルとしての質量分布である。分母の  $r$  は  $F$  を減衰させる要因であり、表現のずれ（減衰）を調整する項でもあり、影響の届きにくさを表す。ここで意味方向（感情極性）の向きを考慮し、 $F$  が + なら引き寄せ、- なら反動、0 なら無関係、という SNEH の図（意味的相互作用行列: Semantic Interaction Matrix）を成果物として設計する。

$$F_{kl} = G \frac{m_1^{(k)} \cdot m_2^{(l)}}{r_{kl}^2} \quad (k, l \in \{S, N, E, H\}) \quad (4)$$

各意味次元  $d \in \{S, N, E, H\}$ 、および意味次元  $d$  を測るために用いる個別の観測特徴すなわち弱教師ルールである  $k = 1, 2, \dots, K_d$  については、下記の通り整理した：

例：  $d = s(\text{Structure})$  の場合、

- k=1 接続語の数
- k=2 文数の適正度
- k=3 結論表現の有無
- k=4 冗長さの低さ

意味次元  $d(S/N/E/H)$  のスコアは、その次元を測るための複数の観測特徴  $k$  を重み付きで合成し、正規化すると、下記(5)式で表現される。

$$SNEH_d = \sigma\left(\sum_k w_{d,k} \cdot f_{d,k}(\text{text})\right) \quad (5)$$

$\sigma$  : 0 - 1 に収める正規化 (min-max / sigmoid)

$w_{d,k}$  : 重み (最初は等しい重みとし、学習していくことを考えている)

$f_{d,k}$  : 観測可能な特徴量 (カウント、比率、有無)

ここで、片側の  $m_i$  を計算するにあたって、SNEH に分解する処理手順を以下に示す。

A) 前処理として、句点「。」で分割し、形態素解析を行うが、接続語（が、さらに、また）を保持しておく（例：Structure のため）。

B) 観測可能な特長量（カウント・比率・有無）を Z スコア正規化ではなく、期待値基準正規化による計算をする。理由は、サンプル数が少ない（本研究で扱う学習対象の文書数は 30 本程

度) ので、学習するうえでの分布が安定しないことと、編集ナビゲータ (編集支援) のためには、SNEH の各次元を 0-1 で捉えたいため。また、平均より多い・少ないといった統計分布ではなく、編集支援のためには心理的・機能的な基準を設けて、「多すぎる」、「少なすぎる」、「丁度いい」などを算出するため。

C) SNEH の各特徴量の正規化の方針を下記とする

表 2 SNEH の特徴量の正規化方法

特徴量のタイプ	正規化方法
密度・頻度	期待値基準 (÷期待値)
有無	0 または 1
冗長度	1 - 比率
強度	Min-Max または logistic

## 6. 意味的相互作用行列の構成

上記の F は、下記のような意味的相互作用行列のような定義を考えている。

縦軸 (k) : 人間文がどの意味次元で発信しているか

横軸 (l) : 生成 AI 文がどの意味次元で受け止めているか

$m_1$  : SNEH 空間上の質量ベクトル

$m_2$  : 同上

$r^2$  : SNEH 次元毎の減衰行列

F : SNEH x SNEH の相互作用行列

分母である  $r^2$  については、段階を経て検討を進めていく。

初期段階は、下記(6)式を想定している。

$$r^2 = \|x_1 - x_2\| \quad (6)$$

そのうえで、下記(7)のように意味次元別距離を定義し、S 次元では近いが、E 次元では遠い、などを表現する。このため、F は行列となる。

$$r^2 = \begin{bmatrix} r_S^2 & 0 & 0 & 0 \\ 0 & r_N^2 & 0 & 0 \\ 0 & 0 & r_E^2 & 0 \\ 0 & 0 & 0 & r_H^2 \end{bmatrix} \quad (7)$$

上記のうち、SNEH の質量ベクトルの具体的な計算例を 8. にて述べる。

## 7. 編集ナビゲーターとしての解釈

$F$ が意味次元別に分解されているため、編集ナビゲーターによる編集示唆は下記の形に該当すると考えている。

表3 計算モデルの記号と編集ナビゲーターによる示唆案

原因	編集示唆 (案)
$m$ の値が小さい	その意味次元を補強する (H次元が低い→主体的解釈を強めに！)
分子 ( $r$ ) が大きい	表現のずれを調整する
符号が逆	解釈の方向性を明確化する
他次元に比べ突出	強みとして残す・伸ばす

## 8. 具体的な計算例

ここでは最も重要な質量の計算のみ紹介する。2025年8月25日の第77回ことば工学研究会にて発表したピカソの絵（右下参照）に対する一人の人間の創造したシナリオの書き出しは下記であった。これをもとにSNEHのSNの例を挙げて確認する。

「左右どちらの絵も非常に抽象的で対象を一目で捉えることは難しいが、左側の絵は上部の頭らしき部分に着目すると人間の上半身が浮かび上がり、人間の存在の気配を感じた。さらに、断片的に構成された胴体の要素と合わせることで、描かれているものが人間であると想像した。また、背景が暗いことで、中心部のハイライトが際立ち、背景から浮かび上がっているような視覚効果があると思う。」



### A) Structure の質量計算：

- ① 接続・展開語数（が、さらに、また等）が3個あり、うち文内逆接「が」を除くと2個ある。また、文数は3つ。接続語密度を $x$ とすると、接続語数 / (文数 - 1) とした場合、0の解釈は「文が断片的」、0.3は「弱い結束」、0.6は「ほぼ意味が成立」、1.0は「十分に構造化」、1.0より多い場合は「過剰（飽和）」と解釈する。そのうえで期待値基準正規化を定義すると、下記(8)式の通り：

$$f_{conn} = \min\left(1, \frac{x}{\tau}\right) \quad (8)$$

$\tau$ : 「これくらいあれば十分」という設計値

$$x = \frac{\text{接続語数}}{\text{文数} - 1}$$

従って、接続語密度 $x$ は、 $x=2/(3-1)=2/2=1.0$

「これくらいあれば十分」という設計値は1。理由は、全ての文間に接続語が必要ではないものの、主要な文間は1つあれば十分 (Halliday&Hasan, 1976) から。

すると、正規化後は1.0となる。

- ② 文の数 (パラグラフ毎に2~5文を最適と期待値仮定: (Stanford NLP Group などの既存の感情分析研究に基づく上限) が3文あり: 1.0
- ③ 結論・判断表現の有無 (~と考える) が1個あり: 1.0
- ④ 冗長度 (同一語の反復率: 名詞・動詞・形容詞の総語数[分母]に対し、2回出たら1、3回でたら2と加算[分子]し計算する。冗長度それ自体のスコアの定義は、下記(9)式とおく。

$$Reundancy = \frac{\sum_{w \in V} \max(0, f(w) - 1)}{N} \quad (9)$$

しかし冗長度は低いほどよいので、(10)式の通り[1 - 正規化値]で反転させる。

$$f_{S, redundancy} = 1 - Reundancy \quad (10)$$

- ・ V: 内容語語彙
- ・  $f(w)$ : 語  $w$  の出現回数
- ・ N: 内容後総数

例) 内容語数 = 40

人間: 3回→2、絵: 2回→1、浮かび上がる: 2回→1 合計=4

冗長度スコアは $(1 - 4/40) = 0.90$

よって、 $S = (1.0 + 1.0 + 1.0 + 0.9) / 4 = 0.975$

#### B) Narrative (推論の流れ: 解釈が前に進んでいるか) の質量計算:

- ① 視点移動語 (着目すると、合わせることで) が2個カウントできる。理由は、まず視点移動、すなわち、観察対象・注視点が明示的に切り替わるのは、下記4つある:
  1. 「左右どちらの絵も非常に抽象的で…」: **全体視点**
  2. 「左側の絵は上部の頭らしき部分に着目すると…」: **局所視点 (上部)**
  3. 「断片的に構成された胴体の要素と合わせることで…」: **別局所 (胴体)**
  4. 「背景が暗いことで、中心部のハイライトが際立ち…」: **背景/全体構図視点**明確な視点移動は3回あるが、初期視点と全体視点を除いて2個となる。  
次に、視点移動が多ければ良いわけではないので、次(11)式のように期待値基準正規化を設ける:

$$f_{shift} = \min\left(1, \frac{c_{shift}}{\tau_{shift}}\right) \quad (11)$$

- ・ 0 : 静的
- ・ 1 : 最低限
- ・ 2~3 : 最適 (文章の長さに依存するが、今回は 3 と設定)
- ・ 4 以上 : 散漫

すると、最適上限である 3 を分母とし、分子は 2 であるため、 $0.67 \div 0.7$

- ② 部分→全体統合表現は、「描かれているものが人間であると想像した」すなわち、「頭部の観察」と「胴体要素の観察」を合わせている Narrative 的にも重要な操作が 1 個あり、量ではなく成立条件として 1 とした。
- ③ 因果・推論語 (~ことで、~ため、~ので、その結果) など、理由と結果を結ぶ語として次の 2 箇所と判断：
  1. 「断片的に構成された胴体の要素と合わせることで」
  2. 「背景が暗いことで、中心部のハイライトが際立ち」
 そのうえで、因果語も多すぎると説明的になりすぎるので、このパラグラフの量で期待値 2 を最適値として①式により  $2/2=1.0$  (本研究では初期値として 1.0 を与えるが、今後の弱教師学習により重みが調整される可能性がある)
- ④ 仮説形成は、観察された要素から確定でない解釈が提示されたかどうかであり (人間であると想像した) がそれに該当する。ここで注意すべきは、「~と考える」は判断・評価であり、新しい解釈の生成とは限らないので仮説形成に含めないが、「~ではないかと考える」「~と仮定すると」などは仮説形成に含める。ここは有無の判断で 1.0。

### C) Emotion (感情) の質量計算：

感情語密度 (E1 : density)、強度語 (E2 : intensifier)、情緒表現 (E3 : affective imagery) に分解し、 $(E1+E2+E3) / 3$  を計算する：

- ① 感情語密度 (E1) は、日本語感情辞書 (PN Table) に登録された語の密度、例えば、喜・哀・怖・不安・美・不気味・緊張・落ち着くなどを指す。ここでは 500 字程度で「十分感情がでていいる」とみなす期待値上限を 5 とした (cf. Stanford NLP Group) 場合、6.①での式と同様にすると(下記(12)式)、感情語は「気配」(準感情語)のみであり  $E1=1/5=0.1$  となる。

$$E_1 = \min\left(1, \frac{n_{emotion}}{\tau_{emotion}}\right), \text{with } \tau_{emotion} = 5 \quad (12)$$

- ・ 0 語 : ほぼ無感情静的
- ・ 1-2 語 : 感情の芽
- ・ 3~5 語 : 十分に感情がでていいる
- ・ 6 語以上 : それ以上増えても方向性はかわらない

- ② 強度語 (E2) は、感情の強さを増幅する語であり、非常に・強く・際立って・はっきり・妙に、などがあげられる。感情表現における短文の実行上限期待値は3 (cf. Stanford NLP Group) とすると、今回の例文では「非常に」「際立ち」の2語であるが、続けての一分文なので1とカウントし、 $E2=1/3 \div 0.33 \div 0.3$  とした。
- ③ 情緒表現 (E3) は、雰囲気・感覚を喚起する表現であり、浮かび上がる・暗い・際立つ・気配・視覚効果が含まれているが、あり／なしの1/0ではなく、最大値を0.4 (情緒表現あり) と0.0 (情緒表現なし) と設定した。0.4の理由は、情緒は感情に寄与するものの、感情よりは弱いと判断したため。

以上の①②③より、 $E=(0.2+0.3+0.4)/3 \approx 0.3$  とした。

#### D) Human/Hermeneutic (主体的解釈) の質量計算：

認知動詞 (H1)、仮説語 (H2)、主体性 (H3)、不確実性の明示 (H4) に分解し、 $(H1+H2+H3+H4) / 4$  を計算する：

- ① 認知動詞 (H1) は、思考・知覚・判断を示す動詞であり、着目する・想像する・感じる・考える・見える、と5つもあるが、飽和する期待値は3 (cf. Stanford NLP Group) とした場合に、 $H1=3/3=1.0$  となる。
- ② 仮説語 (H2) は、~らしき、~のような、~かもしれない、といった確定を避け、解釈の仮説を示す語ととらえ、あり／なしの1/0ではなく、最大値を0.9 (仮説語あり) と0.0 (仮説語なし) と設定した。0.9の理由は、仮説性はHに寄与するものの、過剰な曖昧さを満点にしないと判断したため。従って、今回は「頭らしき」をカウントし、 $H2=0.9$ 。
- ③ 主体性 (H3) は、文章の中で解釈の主体が誰かが明確であるかどうかで判定した。その判定ルールは下記とした。今回の例文では一人称はないものの、認知動詞が連続しており、暗黙の主体があると判断し、 $H3=0.8$ 。

(ア) 1人称明示 (私は~)	: 1.0
(イ) 認知動詞が主語省略で継続	: 0.8
(ウ) 客観記述中心	: 0.3
(エ) 無主体	: 0.0

- ④ 不確実性の明示 (H4) は、~らしき、~と考える、~と思われる、といった解釈が確定的ではないことを示す表現ととらえ、あり／なしの1/0ではなく、最大値を0.8 (不確実性あり) と0.0 (不確実性なし) と設定した。0.8の理由は、不確実な表現は人間的ではあるが、多すぎると責任回避になると判断したため。従って、今回は「らしき」「考える」で、[あり]なため、0.8とした。

以上の①②③④より、 $H=(1.0+0.9+0.8+0.8)/4=0.875$ とした。

E) 最終的な SNEH ベクトルは下記(13)の通り。

$$m_1 = \begin{bmatrix} 0.975 \\ 1.000 \\ 0.300 \\ 0.875 \end{bmatrix} \quad (13)$$

これまでの手法により文章の質量が求められれば、(4)式の分子は計算できる。分母の距離は、既存のベクトル空間によるコサイン類似度等の方法で計算できる。

今後は、人間の30文を学習させ、意味的相互行列を作成する(A)。生成AI文の行列(B)、31人目の新規ユーザー文の行列(C)、があるとすると、編集ナビゲーターは、新規テキストを直接トレーニングコーパスと比較するのではなく、人間の解釈によって定義された意味領域内にテキストを位置付け、AI生成の参照テキストとの意味的相互作用を評価することで、方向性のある解釈可能な修正指針を次のような推敲案を提供する。

例1：(C)のEが(A)の平均よりも低く、(B)よりも低い場合、「感情を1語だけ補うと、人間的にもAI的にも安定する」。

例2：(C)のHが(A)での許容範囲であるが、(B)と強く反動する場合、「主體的解釈は強みなので残しつつ、構造的根拠を補強する」

## 9. 先行研究

- ① 本研究の「 $r^2$ 」の初期定義の理論的背景は、Landauerら(2007)が提唱した「ベクトル空間」による類似度を測るアルゴリズムが基礎にある。そのために、単語x文書行列をもとに特異値分解(SVD: Singular value decomposition)による次元削減を行う「計算機科学」と、人間の意味理解という「認知モデル」の両面から論じている。

T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch. (2007). Handbook of Latent Semantic Analysis. New Jersey: Lawrence Erlbaum Associates, Publishers.

② Narrative attractiveness の理論的裏付けとしては、人間の意味的理解は「物語構造」に依存することを提唱した Bruner(1991)。すなわち、Bruner(1991)は、人間が「現実」を客観的に写し取るのではなく、ナラティブ(物語)を通じて構成することを論じている。科学的・論理的思考とは異なり、物語は意図、規範からの逸脱、解釈可能性を含み、文化や社会的文脈の中で意味を獲得する。ナラティブは検証可能性ではなく「もっともらしさ」によって受容され、人間の経験理解や文化的交渉を可能にする認知様式であると位置づけられる。

J. Bruner. (1991). The Narrative Construction of Reality. *Critical Inquiry*, 18(1), 1-21.

③感情 (Emotion) を計算モデルに組み込む試み。感情は意思決定、知覚、学習などにおいて不可欠な役割を果たしており、つまり感情は合理的な思考のメカニズムそのものに影響を与えている。感情を認識し理解する能力、さらには感情を持ち表現する能力を計算モデルに与えるため、人間の感情に関する背景知識、感情知能の要件に触れ、計算モデルの設計と構築について論じた。

R. W. Picard. (1997). *Affective Computing*. The MIT Press.

④文書を静的特徴ではなく 意味遷移をもつ存在として扱う論文。Topic models としての LDA (latent Dirichlet allocation) は文書等の統計的分析に有用なツールとなり得るものの、トピック間の相関を表現できていない。このため、トピックの割合がロジスティック正規分布を通じて相関を示す CTM (correlated topic model) を提唱した。本研究では、Narrative (N) 次元の理論的支柱として位置づけられる。

D. M. Blei, & J. D. Lafferty. (2007). A Correlated Topic Model of Science. *The Annals of Applied Statistics*, 1(1), 17-35.

⑤AI を評価基準ではなく「参照点・触媒」とする立場。生成 AI の台頭は、人間と AI の相互作用におけるパラダイムシフトを推進し、人間中心の AI に新たな機会と課題をもたらしている。従来の入力-出力モデルを超え、新たな相互作用として、人間と AI の共創を確立する転換がそれである。共創システムの理論的基盤、すなわち、相互作用を通じて意味が動的に生成されることを示し、ユーザーの主体性と協働という HCAI 原則に沿った共創パートナーシップ設計の視点を提供した。本研究では、ChatGPT 文を意味的重力源とする発想と整合している。

N. Davis, M. Clements, Rezwana, J., & E. Browne. (2025). *Human-AI Co-Creation: A New Interaction Paradigm for Human-AI Interaction*. In: Xu, W. (eds) Handbook of Human-Centered Artificial Intelligence. Springer, Singapore. [https://doi.org/10.1007/978-981-97-8440-0\\_76-1](https://doi.org/10.1007/978-981-97-8440-0_76-1).

⑥ラベル不在環境で意味的属性を学習する枠組み

機械学習システムの展開において、大量の正解な教師データのラベリングが最大のボトルネックとなるケースが増加している。Weak Supervision は、「完全には正しくないが、そこそこ当たる」情報源を多数かき集め、そのノイズを統計的に補正して大規模な学習用データを合成することで、このボトルネックを解消しようとするシステムを提案。ユーザーは、未知の精度や相関を持つ可能性のある任意のヒューリスティックを表現するラベリング関数を記述することで、コストを大幅に下げた。本研究では、M 推定の方法論的基礎として弱教師ラベルの導入に寄与した。

Rather A., Bach, S. H., Ehrenberg, H., Fries, J., Wu, S., & C. Ré. (2017). Snorkel: Rapid Training Data Creation with Weak Supervision. In *Proceedings of the VLDB (Very Large Data Bases) Endowment*, 11(3), 269-282.

⑦この古典的研究では、談話の結束性は接続語・照応・省略、などによって形成される。特に接続語は文間関係を明示する最小単位となる。重要なのは、すべての文間に接続語が必要なわけではないが、主要な文間には1つあれば十分である。これが「1文間1接続語」基準の理論的原点である。

M. A. K. Halliday, & R. Hasan (1976). (結合性理論) Cohesion in English. Addison-Wesley Longman Ltd

⑧Stanford NLP Group の文献：スタンフォード NLP グループの Socher らによる先行研究では、感情強度、主観性マーカー、構成的構造が意味解釈に非線形に寄与し、少数のマーカーで性能が飽和することが示された。これらの知見に基づき、筆者らは頻度ベースの正規化ではなく、小さな操作上限を採用する。

- Socher et al., 2013, EMNLP  
*Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank*  
※E2 (強度語) の期待値3の実証的裏付け
- Socher et al., 2012, EMNLP  
*Semantic Compositionality through Recursive Matrix-Vector Spaces*  
※Nの統合表現—意味は単語の足し算ではなく、構成 (composition) によって構成される
- Socher et al., 2011, EMNLP  
*Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions*  
※H1 (認知動詞) の期待値3の「飽和」論理の実証的補強。一人称・認知動詞・主観マーカーの存在が人間的判断の検出に寄与する。出現回数は1~数回で頭打ち。
- Socher et al., 2010, NIPS  
*Learning Continuous Phrase Representations and Syntactic Parsing with Recursive Neural Networks*  
※SNEHのS・Nを下支え：文献では、意味は構造と合成で決まると提言している。

以上