

日本語の Pattern Matching Analysis (PMA) の簡単な実例*

— “ x が y を襲う” の分析を一例にして並列分散意味論 (PDS) の基礎を解説する—

黒田 航

情報通信研究機構 けいはんな情報通信融合研究センター

1 はじめに

Pattern Matching Analysis (PMA) [6, 7, 8] は私が生成文法で主流となっている統語構造の分析の対案として私が提案した枠組みである。

PMA は安易な樹状構造による記述を受けつけない日本語のような言語の統語構造を記述するために開発されたものである。だが、日本語の分析に PMA を適用したものは—諸々の理由から—これまで存在しなかった。日本語への適用例を示す必要性は常々自覚されていたが、私の研究全体では比較的優先順位が低かったので、長らく後回しになっていた、というのが実情である。

だが、運の好いことに、ちょうど先日、FOCAL 関係者に PMA を使って比喩の効果を説明する機会があり、その結果を示すことで、日本語への比較的詳細な PMA の適用例を示すことが可能となった。

1.1 表記法の変更と記述概念の部分的改訂に関する注意

なお、表記はこれまでに公開された PMA の基本文献 [6, 7, 8] のものと部分的に異なっている。この論文で新たに採用された $[x^*]$, $[x^{**}]$ のような記法は基本文献では用いられていなかったことは明示しておく。これまでは、(i) $[x^*]$, $[x^{**}]$ はいずれも太字で示されていて、(ii) $[x^*]$ と $[x^{**}]$ の区別はなかった。 $[x^*]$ と $[x^{**}]$ の区別を導入したという点は、基本文献の内容を部分的に改訂していることになる。 $[x]$ は形式と意味との両方をあらわし、 $[x^*]$ は文脈の影響のない $[x]$ の意味を表わし、 $[x^{**}]$ は文脈におかれた $[x^*]$ の意味を表わすと考えてもらえたらよい。

* この論文の準備、改訂にあたって、中本敬子 (京都大学教育学研究科 PD)、横森大輔 (京都大学大学院) のお二人との議論が参考になった。この場を借りて、お礼を申し上げたい。

2 日本語の PMA

この論文で取り上げる例は (1) である:

(1) ?その日、直下型地震が島課長を襲った。

ただし、簡略化のため、以下の説明では (1) での “その日、” の効果は無視しよう。

まず、文 (1) は自然な文というより、些か奇妙な文である。地震は個人規模でヒトを襲うタイプの自然災害ではなく、島課長一人が被害にあったとは考えにくい (詳細は [17, 14] などを参照されたい)。従って、(1) は次の (2) のような文に較べると、相対的に奇異である:

(2) その日、直下型地震が東海地方を襲った。

最終的に、PMA の知見を生かして、(1) の奇妙さがどうやって “解消” されるかを §3 で説明することになるが、その前に、まず準備が必要である。

2.1 PMA の基本的な仮定

(1) の PMA は、図 1 に指定した情報を基にして、次のような仮定から始まる:

(3) $s = \text{“直下型地震が島課長を襲った”}$ [= (1)] の統語構造と意味構造は、次の条件を満足するような $s = w_1^{**} \cdot w_2^{**} \cdot w_3^{**} \cdot w_4^{**} \cdot w_5^{**}$ として必要十分に記述しうる:

(4) $w_1^{**} = \text{[直下型地震**]} = f(S_1, \dots, S_5)$ (ただし $S_1 = \text{[直下型の地震*]}$),

$w_2^{**} = \text{[が**]} = f(P_1, \dots, P_5)$ (ただし $P_2 = \text{[が*]}$),

$w_3^{**} = \text{[島課長**]} = f(O_1, \dots, O_5)$ (ただし $O_3 = \text{[島課長*]}$),

$w_4^{**} = \text{[を**]} = f(Q_1, \dots, Q_5)$ (ただし $Q_4 = \text{[を*]}$),

s		直下型の地震**	が**	島課長**	を**	襲った**
w1	直下型の地震	直下型の地震*	P1	O1	Q1	V1
w2	が	S2	が*	O2	Q2	V2
w3	島課長	S3	P3	島課長*	Q3	V3
w4	を	S4	P4	O4	を*	V4
w5	襲った	S5	P5	O5	Q5	襲った*

図1 (1)の PMA

$w_5^{**} = [\text{襲った}^{**}] = f(V_1, \dots, V_5)$ (ただし $V_5 = [\text{襲った}^*]$)
 S, O はおのおの主語名詞 (句) S(SUBJECT) NP, 目的語名詞 (句) O(OBJECT) NP という要素を — 正確には, それらの素性の塊を — エンコードする. P, Q は後置詞 (Postposition: P) のクラスに属する要素で, おのおの S, O をマークするものだとする.

ここで, $x \cdot y$ は x と y が隣接していることを表わし (演算子は “ \cdot ”)¹⁾, $f(x_1, \dots, x_n)$ は $\{x_1, \dots, x_n\}$ の素性の統合 (feature-integration) だとする (演算子は “ f ” だが, これは統語的には不可視である).

2.1.1 PMA の要点を極く大雑把に

PMA は元々, 意味構造を (S, O のような記号化によって) 見えにくいようにしている. 意味構造を記述するとは, $X = \{S, P, O, Q, X^*\}$ のような要素の「中身」を考えるとということであり, それは X の意味素性をしっかり特定するということである.

これに対し, PMA はかなり明確に統語構造を規定する. $X = \{S, P, O, Q, X^*\}$ がどういう順番に現われ, 他の層にあるどの要素と一致 (match) するかによって統語構造が必要十分に規定されると考えられるからである. この点に関する実例は基本文献には腐るほど例が出てくるが, この論文ではその詳細には触れず, 簡単な例を示すに留める. 例えば, 句構造 (5) は表 1 の C/D 表 (Composition/Decomposition Table) で記述される:

$$(5) [s [NP_1 x] [VP [v y] [NP_2 z]]]$$

¹⁾ 細かいことを言うと, \cdot は隣接性を表わす必要はなく, もっと弱い “ x が y に (時間的に) 先行する” という出現順序を表わす演算子であればよい. この点はこの論文の記述範囲では瑣末過ぎる点なので, 便宜のために問題を簡略化した. 詳細は [7] を参照されたい.

表1 (5)の PMA 表記

	x^{**}	y^{**}	z^{**}
$x.$	x^*	V	(O)
$y.$	S	y^*	O
$z.$	S'	R	z^*

ただし, $R = \{V, P\}$ ($V = \text{Verb}, P = \text{Preposition}$) で, 更に $[+S(x)] \text{ OR } [+O(x)] \rightarrow [+N(x)]$ という素性の含意の関係 (“ x が S か O ならば, x は N である”) を仮定する²⁾. 詳細に興味のある方は基本文献を参照されたい.

2.1.2 統合のイメージは「重ねあわせ」

統合は複雑な処理を含むが, その直観的理解は難しくない. 統合のイメージを簡単に言うと, $[x^{**}]$ は $[x_1], \dots, [x_5]$ を “縦に重ねあわせたもの” である. 例えば, $[w_1^{**}] = [\text{直下型地震}^{**}]$ は

$$(6) \{ [\text{直下型地震}^*], [S_2], [S_3], [S_4], [S_5] \}$$

を “縦に重ねあわせたもの” である (ただし, $[\text{直下型地震}^*] = [S_1]$).

[7] では強調しなかったが, 次のことは生成系の派生 (derivation) を用いた統語分析との比較のために強調しておいてもいいだろう:

- (7) $\{S, S', O, (O), V, P, R, \dots\}$ は, それらを含むパターンの内部ではゼロ要素としてふるまう. 例えば,
- a. $S, S', O, (O)$ はゼロ代名詞として,
 - b. V はゼロ代動詞として,
 - c. P はゼロ前置詞として
- ふるまう.

²⁾ これは $N: \{S, O\}$ という型階層 (type hierarchy) を考えることに等しい.

これは牽強付会ではない。実際、動詞 y の主語句部分である S は PRO と等価であるし、目的語句 z の主語句部分である S' も同様に PRO 相当の存在体としてふるまうはずなのだが、今のところ生成言語学では目的語に PRO が随伴しているという段階まで過激な分析は進んでいないようだ³⁾

これには次のような重要な含意がある:

- (8) PMA では PRO/pro のような大部分のゼロ代名詞のふるまいは、代名詞の存在を仮定しないゼロ照応 (zero anaphora) の関係として記述される。

これは次の理由による:

- (9) パターンの列ごとの重ね合わせは、Chomsky 派生成言語学では十八番となっている「統語派生」で生じる横方向の「移動」の効果と等価である。

2.1.3 線状化

このような重ねあわせは、[直下型地震**], [が**], [島課長**], [を**], [襲った**] の列ごとに起こる。重ねあわせの結果を線状に並べたもの (linearization) が、[直下型地震**, が**, 島課長**, を**, 襲った**] である。これの異表記が $s = w_1^{**} \cdot w_2^{**} \cdot w_3^{**} \cdot w_4^{**} \cdot w_5^{**}$ である。

2.1.4 統合と単一化との関係

素性の統合は素性の単一化 (feature-unification) を基本とするが、それに限られるわけではない。一定の条件下で素性値の不一致を許す。詳細は以下の議論を参照されたい。

以下では統合の際に何が起こるかを記述しよう。これが §3 で説明するような (1) の多様な読みを説明する。

2.2 意味素性指定の特徴

以下では素性統合に利用される意味特徴がどのように語彙的に表示されているかを説明するが、そのために、まず表記法を確定しておこう。

2.2.1 表記法

図 1 に指定した情報は、次のようにも表わせる:

- (10) $s = f(w_1, \dots, w_5)$, ただし
 $w_1 = [\text{直下型地震}^*, P_1, O_1, Q_1, V_1]$
 $w_2 = [S_2, \text{が}^*, O_2, Q_2, V_2]$
 $w_3 = [S_3, P_3, \text{島課長}^*, Q_3, V_3]$
 $w_4 = [S_4, P_4, O_4, \text{を}^*, V_4]$
 $w_5 = [S_5, P_5, O_5, Q_5, \text{襲った}^*]$

ただし、 $[x_1, x_2, \dots, x_n]$ は $[[x_1], [x_2], \dots, [x_n]]$ の略記で、その意味は $[x_1] \cdot [x_2] \cdot \dots \cdot [x_n]$ と等しいとする。

2.2.2 意味表示がいかにエンコードされるか

この際に重要なのは、

- (11) w_1, \dots, w_n がおのおの自律的に意味フレーム (semantic frames) という形で記述できる意味表現と結びつけられ、
 (12) i 番目の語 w_i の意味と統語が $[x_1, \dots, x_i^*, \dots, x_n]$ という統語パターン (syntactic pattern) にエンコードされている

と考えることである。

これは、意味フレームは語ごと—正確には形態素ごと—にコードされているということである。これは非常に冗長な表示であるが、並列分散処理 (Parallel Distributed Processing: PDP) [2, 13, 11] の洞察に基づいて、ヒトの意味処理、統語処理は本質的にこの種の冗長性を利用していると考えるのが PMA の最大の特徴である。

2.2.3 意味の分散表現

重要な点を繰り返そう。

s を構成する語 w_1, \dots, w_5 はいずれも、おのおの独自に意味素性を指定し、表示している。 w_i ($i = \{1, 2, \dots, 5\}$) の指定は、それぞれの S_i, O_i のような“項” (arguments) の形で実現されている。喩えて言うと、 w_i の意味情報は、その項 S_i, O_i に“乗って”いる。これ故、項は意味フレームの観点からすると、意味役割 (semantic role) の“乗り物”である。

重要なのは、

- (13) 意味素性の指定は、同一の統語的単位に関するもの (e.g., S, O) であっても、語ごとに異なっていない

という点である。より正確に言うと、例えば、

³⁾ このような分析をすると格マーキングや Agr-O がうまく説明できるはずなのに、この可能性は Minimalism の枠組み [1] でも気づかれていないようだ。

s		直下型の地震**	が**	島課長**	を**	襲った**	
L1	直下型地震…島課長…	直下型の地震*	P1	島課長*	Q1	V1	名詞列の意味
L2	…が…を…	S2	が*	O2	を*	V2	格パターンの意味
L3	…襲った	S3	P3	O3	Q3	襲った*	動詞の語意的意味

図2 日本語の三層意味コーディング [(1) の PMA の場合]

- (14) w_3 の S_3 の素性値指定の内容と, w_5 の S_5 の素性値指定の内容は —両者が同じく S についての指定であっても— 統合可能であれば同一でなくてもよい

ということである。これは,

- (15) 語が自律的に状況をコードしていること,
 (16) 意味情報と統語情報が共に語彙情報に分散的表現されていること

を意味する。

ただし, 動詞や名詞のような実詞類の意味素性指定と助詞類の意味素性指定の仕方は違う。まず, このことを簡単に説明しておこう。

2.2.4 動詞や名詞のような実詞類の意味素性指定

例えば, $w_1 = \text{“直下型地震”}$, $w_3 = \text{“島課長”}$, $w_5 = \text{“襲った”}$ では, 意味素性の指定は具体的で, 個別的な状況として特定できるような意味フレームに結びつけられている。

2.2.5 抽象的な議論へ, ちょっと脱線

[x] を記述するパターンの性質について, 一つだけ抽象的な議論をつけ加えておく。小難しい話が嫌いな人は, 無視してよい。

n 個の語からなる文 s の i 番目の語 w_i は,

$$(17) w_i = [x_1, \dots, x_{i-1}, x_i^*, x_{i+1}, \dots, x_n]$$

で表現され, その意味は, おのこの要素の意味素性の指定という形で実現されている。この指定を大別すると, 次の

- (18) a. $L(x_i^*) = [x_1, \dots, x_{i-1}]$
 b. $R(x_i^*) = [x_{i+1}, \dots, x_n]$

になる。 $L(x_i^*)$ の指定と $R(x_i^*)$ の指定とは興味深い非対称を示すので, これらの区別は理論的には重要なのであるが, この区別はこの論文の範囲では気にしなくていい。

2.2.6 例1

例えば, $w_1 = \text{“直下型地震”}$ の意味特性は

$$(19) [\text{直下型地震}^*, P_1, O_1, Q_1, V_1]$$

の P_1, O_1, Q_1, V_1 への意味素性の指定という形で実現されている。この際, O_1, V_1 への意味指定が重要で, P_1, Q_1 への意味指定は特に重要でないことには注意が必要である。つまり, O_1, V_1 への意味指定は語彙的であるが, P_1, Q_1 への意味指定はそうではない。

具体的には, V_1 は例えば, [襲う*] のような動詞と一致するが, [笑う*] のような動詞とは一致しない素性指定をもつ。同様に, O_1 は例えば, [東海地方*] のような名詞と一致するが, [島課長*] のような名詞とは一致しにくい素性指定をもつ。

2.2.7 曖昧性の解消

語の意味がどれくらい特定の状況に結びつけられているかは, 語ごとに大きく異なる。例えば, w_1, w_5 が結びつけられている状況はかなり特定されたものだが, w_3 が結びつけられている状況は w_1, w_5 に較べると, 圧倒的に漠然としている。

おのこの語に結びつけられている意味フレームの数は一つではない。通常, 一つの語には複数の意味フレームに結びつけられており, 脱曖昧化は素性の統合の際に実現される。

2.2.8 助詞類の意味素性指定

これに対し, $w_2 = \text{“が”}$ や $w_4 = \text{“を”}$ の項である S_2, O_2 や S_4, O_4 は個別的な状況として特定できるような意味フレームと結びついておらず, それらは強い意味指定をしない。それらがするのは関係的な指定であり, これが “が” や “を” のような要素の意味が 「文法的」 だと言われる理由である。

補足的に言うておくと, 例えば “ x が” や “ y を” の “が” や “を” の意味を x, y の意味から分離して記述することは事実上不可能である。従って, 文献に 「ガの意味の分析」とか 「ヲの意味の記述」とか

あった場合、それは多くの場合、「後置詞句 “x が” の意味の分類」や「後置詞句「後置詞句 “x が” の意味分類」の (言い) 誤りである。

2.2.9 意味フレームとの結びつきに関する注意

具体的な意味分析に移る前に、複層意味フレーム分析 (MSFA) [15, 16] の知識のある人への注意を一つ: $[x^{**}]$ は意味フレームには対応しない。意味フレーム対応するのは、 $W = \{w_1, \dots, w_n\}$ から適当な要素の取り合わせたもの (e.g., $\{w_1, w_3\}$) の方である。つまり、PMA と MSFA は意味フレームの表現法が反対—正確には裏返しの—である。この点に関する詳細は、次の節で明らかになるだろう。

2.3 日本語の意味エンコーディングの特徴

この節では日本語のような形態論が豊かな言語の意味エンコーディングの典型的特徴を述べる。

2.3.1 三層エンコーディング

英語のような言語と異なり、日本語の意味構造のエンコーディングは基本的に三層からなる。その三層とは以下の L1, L2, L3 のことである:

- (20) L1: 直下型地震 ... 島課長 ... = [直下型地震*, P_1 , 島課長*, Q_1 , V_1]
 L2: ... が ... を ... = [S_2 , が*, O_2 , を*, V_2]
 L3: ... 襲った = [S_3 , P_3 , O_2 , Q_3 , 襲った*]

同じことを図 2 にも示した。

L1 の情報は図 1 の w_1, w_3 の情報が部分的に統合されたものであり、L2 の情報は w_2, w_4 の情報が部分的に統合されたものである。この融合の際、図 1 の w_i の添え字 i と図 2 の L_i の添え字 i とは一致させなかったため、添え字の食い違いには注意された。 L_i の添え字 i はマクロ指標である。

一見して明らかのように、L1 は名詞群の意味に、L2 は格パターンの意味に、L3 は動詞の語彙の意味に対応している。L2 がエンコードするのは、事態の比較的抽象的な側面 (e.g., 移動の有無とか) である。英語では、L2, L3 がほとんど区別できない⁴⁾。

2.3.2 L1 は意味フレームを喚起する

次の点には注意が必要であろう:

- (21) PMA は (L2, L3 と並んで) L1 にも状況のエンコーディングの効果認め、意味フレームの喚起の効果認める。

これには以下に述べるような些細とは言い難い含意がある。

2.3.3 日本語の「構文」的意味

(21) が正しいとすると、日本語の「構文」的意味 (constructional meaning) をエンコードするのが L1 なのか、L2 なのか、あるいは L1, L2 の複合なのかはハッキリしないことになる。これが日本語の構文文法 (Construction Grammar) に対して意味することは軽微ではないだろう。特に記述モデルを Goldberg [3] に求めた場合には、そうである。

私の見る限り、日本語の構文文法研究者は、L1, L2 の効果をうまく分離し損なっているのではないかと思う。L1, L2 を一緒に扱うことは、それはそれで誤りではないが、私としては、L1 の L2 から分離された独自の意味論的貢献を重視する。

この立場に立つ場合、一つ心配なのは、もし日本語構文文法の研究者が L2 のみを構文と見なし、L1 を構文と見なさないとしたら、それは本末転倒ではないかということである。実際、Goldberg の研究で構成体 (constructions) と言われている効果の源泉は、明らかに L1 に帰属する情報である。

もちろん、問題は、日本語が「語順が相対的に自由な言語」だと考えられているところにある。従って、L1 は単独では機能せず、強い構文効果は {L1, L2} の複合のレベルで現われる。ところが実際には、同一の格助詞が現われている場合でも、L1 の内容が異なる場合には、理解される状況が異なる。例えば、次の例は、いずれも動詞がつけるで同一の [... は ... に ... を ...] という格パターンを取るが、L1 に現われる名詞句の内容によって、容認性が異なる:

- (22) a. ?彼はいつも変わった色を髪につける。
 ((着色) フレーム)
 b. 彼はいつも髪に変わった色をつける。
 ((着色) フレーム)
 (23) a. 彼はいつも変わった色を身につける。
 ((着衣) フレーム)
 b. ???彼はいつも身に変わった色をつける。
 ((着衣) フレーム)

⁴⁾ これは実際にはそれほど正確な言明ではない。英語でも *to, from, with* のような前置詞でエンコードされている情報は、L1, L2 に相当する。

〈Z が X を身につける〉は、[身につける] が複合述語化、単一動詞化しており、「つける」の基本パターンである〈Z が Y に X をつける〉をから特殊化、個別化している。

なお、(22a), (22b) の [変わった色**] は [変わった色の染料**] に、(23a), (23b) の [変わった色**] は [変わった色の服**] に、おのおの補正される必要があることは、正確さのために指摘しておく。このような意味の調節 (semantic accommodation) [9, 10]、あるいは“(意味) 型の強要” (type coercion) [12] がなぜ生じるのかを説明するという目標を考えただけでも、日本語の構文効果の記述、説明がいかに困難であるかわかる。

このような理由もあって、私は L2 が日本語の構文だと主張することは—特に無意味だと言うわけではないが—あまり意味のないことだと考えている。L2 の状況喚起力は、明らかに L1 よりも弱いし、L2 の喚起も抽象的な効果に限られる。この事実は、L1 が L2 には還元できない、独自の意味論的貢献をなしており、かつ、それが文の意味解釈の中核的な部分であることを意味している。

2.3.4 L2 は冗長か?

語彙概念構造論 (Lexical Conceptual Structure Theory: LCS Theory) [4] を日本語で実践している研究者 [5, 21] の立場からすると、おそらく L2 のエンコードする情報は冗長である。その理由は「L2 は L3 から予測できるから」というものであろう。

だが、この主張は実際には正しくない。少なくとも、謙虚に現象を観察した場合にはそうである。李 [18, 20] が多くの実例を挙げて論じているように⁵⁾、L2, L3 の情報には、しばしばズレがある。それらのズレは確かに、「解消」可能だが、ズレが存在すること事態は、L3 から予測できることではない。従って、L2 は L3 の情報の一部であるとする概念構造意味論の主張は、経験的に妥当なものだとは見なしがたい。実際、L2 が L3 から予測され、冗長であると強く主張する研究者は、事実の観察に基づいてそう言っているより、肩入れしている表示理論によるバイアスからそう言っている可能性が高い。文法が複雑系であることを知らず、冗長性を嫌い、単純な文法を求める時代は、とっくの昔に終わっている。

⁵⁾ 里は、[19] で [NP に] の挙動の定量的分析も試みている。

既存理論の批判はこれぐらいにして、日本語の PMA の具体的な側面をもう少し解説しよう。

2.4 三項述語の場合

(1) の [襲う] は二項述語だった。ここで、(22b) の PMA を考え、三項述語の場合に何が起こるか見てみよう。

2.4.1 〈生物〉が〈体の一部〉に〈何か〉をつける

(22b) の文の PMA 案を図 3 に与えた。ただし、表記は以下のように変更した。

- (24) $M(\alpha)$ は役割 α のマーカーである (P, Q はおのおの, $M(S), M(O)$ である)。
- (25) $M(G)$ は移動の到着点 (Goal) をコードする。
- (26) $M()$ は $M(\alpha)$ の役割 α が“未指定” (underspecified) であることを表わす。 $M()$ は $M(S)$ にも $M(O)$ にも $M()$ にも、あるいは他のマーカーにも一致する (が、マーカー以外のものとは一致しない)。
- (27) ADV は副詞 (素性) をコードする。
- (28) ? は文法役割素性のエンコードがあるか不明な箇所である。

図 3 の PMA は実験的なもので、未解決な問題も多いが、この分析が主張するは、次のことである:

- (29) $P_1: [\dots \text{が} \dots \text{に} \dots \text{を} \dots]$, $P_2: [\dots \text{が} \dots \text{を} \dots \text{に} \dots]$ という格パターンは、おのおの
 - a. $P_a: [\dots \text{が} \dots \text{を} \dots]$,
 - b. $P_b: [\dots \text{が} \dots \text{に} \dots]$,
 - c. $P_c: [\dots \text{に} \dots (\text{を} \dots)]$
 - d. $P_d: [(\dots \text{が}) \dots \text{を} \dots \text{に}]$
 という三つの小規模な二項的格パターンの複合体であるが、 P_a, P_b, P_c, P_d はおのおの独自 (かつ冗長的) に、(関係的) 意味をエンコードする。
- (30) Y は Z の直接目的語である。別の言い方をすれば、X は「を」格でマークされているが、この文の (「二番目に目立った参与者」という意味での) 直接目的語ではない。
- (31) 〈Z が Y に X をつける〉は Z が引き起こす X の Y への (抽象) 移動が読み取れるばかりでなく、Z が引き起こす Y の X の (抽象) 所有も読み取れる。

s		彼**	は**	いつも**	髪**	に**	変わった色**	を**	つける**
w1	彼	彼*	M(S)1		O1	M(O)1			V1
w2	は	S2	は*	ADV					V2
w3	いつも	S3	M(S)3	いつも*					V3
w4	髪	S4	M(S)4		髪*	M()4			V4
w5	に	?			S5	に*	O5		V5
w6	変わった色	?			S6	M()6	変わった色*	M()6	V6
w7	を	S7	M(S)7				O7	を*	V7
w8	つける	S8	M(S)8	ADV	G8	M(G)8	O8	M(O)8	つける*

図3 日本語の三層意味コーディング [(22b) の PMA の場合]

s		彼**	は**	いつも**	変わった色**	を**	髪**	に**	つける**
w1	彼	彼*	M(S)1		O1	M(O)1			V1
w2	は	S2	は*	ADV					V2
w3	いつも	S3	M(S)3	いつも*					V3
w4	変わった色	S4	M(S)4		変わった色*	M(O)4			V4
w5	を	S5	M(S)5		O7	を*			V5
w6	身	?			S8	M(S)6	髪*	M(G)6	V6
w7	に	?			S9	M(S)7	G7	に*	V7
w8	つける	S8	M(S)8	ADV	O8	M(O)8	G8	M(G)8	つける*

図4 日本語の三層意味コーディング [(23a) の PMA の場合]

(32) [Y に X] の部分 ($\{w_4, w_5, w_6\}$) の意味エンコーディングは動詞の意味エンコーディングから独立しており、ここで述べた (抽象) 所有の読み取りは自律的に起こる。これが、[変わった色*] の意味役割は [髪*] から見ると O であるという事態を許す⁶⁾。

これらの主張の根拠は、以下の二つの観察である:

(33) [Z が Y に X を V] という格パターンをもつ文で [Y に X] の部分に Y による X の (抽象) 所有が読み込まれるという傾向は非常に顕著であるので、これは日本語の顕著な構文効果だと考えてよいと思う。ただし、これは Y, X の実現値に依存する。

(34) これに対し、[Z が X を Y に V] という格パターンをもつ文で [X を Y に] の部分に X の Y (抽象) 移動が読み込まれるという傾向

が非常に顕著である。これも、当然 Y, X の実現値に依存する。

以上の二つは重要な一般化だが、残念ながらこの論文はそれを詳細に立証するための試みとしては不正確であろう。

2.4.2 〈生物〉が〈何か〉を〈体 (の一部)〉につける因みに、(23b) が (23a) に較べて逸脱していることも、ここで取りあげた一般化から説明可能である。(23a) の PMA の概略は、図4に示した。

ただ、[変わった色**] は正確には [変わった色の染料**] であること、また、それは移動体 (theme = mover) とは言えないことにも注意が必要である。従って、[髪**] を G と見なすことは、可能であるとしても、無条件に許されるわけではない。

2.5 まとめ

この節では、PMA を日本語に当てはめるために必要な修正、拡張を、幾つか議論した。ただ、この論文で紹介した PMA —特に図3は— は試験的なものであり、今後、緻密化、前提の正当化が必要であろう。いずれ機会があれば、それも行ないたい。

以上の注意の下で (1) の具体的な読みを記述しよう。

⁶⁾ Y による X の (抽象) 所有を Y による X の (抽象) 経験と見なし、これに Y の X への事態変化と読み取ることも可能であるようだ。だが、この複数の読み取りの基盤はよく分かっていない。それはメタファー的なものかも知れないし、そうでないかも知れない。参照点構造が関係しているかも知れないし、そうでないかも知れない。詳しい調査が待たれる。

3 統合の際に何が起こるか

もっとも単純な素性統合は単一化として記述できるが、自然言語の現実(は残念ながら)それほど単純ではない。素性統合の際に、非常に頻りに素性値の不一致が起こり、素性統合は、その不一致を解消するプロセスを含む。

不一致の場所が一ヶ所ならば解消は比較的簡単で、解消法も自明であるが、不一致の箇所が複数ある場合にはそうはいかない。

不一致が複数箇所あった場合にそれを解消するための基本原理は、複数ある不一致を評価し、そのうちもっとも「軽微なもの」のみを解消するということである。ただし、どれが取るに足らないと見なせるかは、もちろん自明ではない。それは観点によるもので、どんな観点がどの特徴を尊重するものであるかを特定することは、残念ながら PMA の目標ではない。その説明には別の説明原理が必要である。

重要なのは、可能な観点の数は有限であり、観点の一つ一つについて、可能な不一致の解消法を特定することが可能であるという点である。以下の記述では、何らかの理由で特定の観点が選ばれたと仮定し、選ばれた観点の下での不一致の解消法の記述を提供する。

以下の説明では、添え字は図 2 ではなく図 1 で使用したものである。

3.1 (1) の読み 1

w_1, w_3 の意味素性指定が優先されると、 w_5 : [襲う*] の意味素性指定 $[S_5, O_5]$ はそれを許容するように修正される。これは比喩の基本的効果である。

3.1.1 動詞の意味の(弱い)不変性

ここで (1) の [襲った**] の意味を (35) の [襲った**] の意味と比べて見よう。

(35) ライオンがガゼルを襲った

(36) $s' = [\text{ライオン}^{**}, \text{が}^{**}, \text{ガゼル}^{**}, \text{を}^{**}, \text{襲った}^{**}]$

の構成部分である動詞 [襲う] の構造記述

$$w'_5 = [S'_5, P'_5, O'_5, Q'_5, \text{襲う}^*]$$

では、 S'_5, O'_5 に関して (1) の w_5 の S_5, O_5 と語彙的に異なる意味素性の指定をもつ必要がないことに

注意されたい。

(1) が (主に) 〈自然災害の発生〉 フレームに、(35) が (主に) 〈動物の襲撃〉 フレームに結びつけられるのは、動詞の意味によって決まることではなく、[直下型地震, ..., 島課長, ...], [ライオン, ..., ガゼル, ...] によって、つまり名詞(群)の動詞から独立した意味論的貢献によって決まることである。

この仮定の下では、[襲う] の意味、統語記述は、次のようになり抽象化できる。

(38) [襲う] =

$$\begin{aligned} & [S[?animate, ?human, ?visible, \dots, \\ & \quad -predictable, +risk-indicative, \dots], \\ & P[+S-marking, -O-marking, \dots], \\ & O[+animate, ?human, ?visible, \dots], \\ & Q[-S-marking, +O-marking, \dots], \\ & \text{襲う}^*] \end{aligned}$$

この定義は、[襲う] の出現位置に依存しない。

ただ、これは抽象的なレベルの [襲う] の中核的意味の表示を考えた場合であって、もっと個別的で具体的な [襲う] の語意表象が存在していけないわけではない。冗長性は矛盾に導かない限り、問題ではない。

3.1.2 助詞の意味の(強い)不変性

同様の理由から、 $s = (1), s' = (35)$ について、 P_i, Q_i, P'_i, Q'_i に関しては、事実上、 $P_i = P'_i, Q_i = Q'_i$ だと考えてよい。ただし、動詞と助詞を較べた場合、意味の不変性は前者の方が低いと考えられる。

3.2 (1) の読み 2

w_1, w_5 の意味素性指定が優先されると、 w_3 : [島課長*] は $[O_1], [O_5]$ の指定に適応し、〈自然災害〉の被害者集団の代表値となるように補正される。これはメタファーを保存しつつ、 $[O_3]$ に関して微弱的なメトニミー的補正が生じている読みである。

3.3 (1) の読み 3

w_3, w_5 の意味素性指定が優先されると、 w_1 : [直下型地震*] と $[S_3], [S_5]$ との値の食い違いは w_1 : [直下型地震*] が〈個人を襲う突発的な事件〉の値 (e.g., [ライバルの陰謀によって準備された左遷(の知らせ)]) となるように修正される。これは [直下型地震*] に局所的に生じるメタファーである。

4 終わりに

この論文では、簡単ながら日本語の PMA の実例を示すことを目的とした。この短い論文が、Chomsky 学派の UG の存在仮説への肩入れによって現実を無視した空虚な統語論、認知言語学派のイメージ先行で空虚な意味論の脱却のために一石を投じることができれば、幸いである。

参考文献

- [1] N. Chomsky. *The Minimalist Program*. MIT Press, Cambridge, MA., 1995.
- [2] J. L. Elman, E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. *Rethinking Innateness: A Connectionist Perspective on Development*. MIT Press, Cambridge, MA, 1996. [邦訳: 『認知発達と生得性: 心はどこから来るのか』(乾敏郎・山下博士・今井つみ訳). 共立出版.]
- [3] A. D. Goldberg. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press, 1995.
- [4] Ray S. Jackendoff. *Semantic Structures*. MIT Press, 1990.
- [5] 影山太郎. 動詞意味論: 言語と認知の接点. くろしお出版, 1996.
- [6] Kow Kuroda. Outlining the *Pattern Matching Analysis*: A theoretical framework proposed for a realistic description of natural language syntax. 言語科学論集, Vol. 5, pp. 1–31, 1999. 京都大学基礎科学科, 京都大学.
- [7] Kow Kuroda. *Foundations of PATTERN MATCHING ANALYSIS: A New Method Proposed for the Cognitively Realistic Description of Natural Language Syntax*. PhD thesis, 人間・環境学研究科. 京都大学, 2000. [<http://cls1.hi.h.kyoto-u.ac.jp/~kkuroda/kuroda2000/> から章ごとに入手可能].
- [8] Kow Kuroda. Presenting the *Pattern Matching Analysis*, a framework proposed for the realistic description of natural language syntax. *Journal of English Linguistic Society*, Vol. 17, pp. 71–80, 2001.
- [9] R. W. Langacker. *Foundations of Cognitive Grammar, Vol. 1: Theoretical Prerequisites*. Stanford University Press, 1987.
- [10] R. W. Langacker. *Foundations of Cognitive Grammar, Vol. 2: Descriptive Applications*. Stanford University Press, 1991.
- [11] J. McClelland, D. Rumelhart, and The PDP Research Group. *Parallel Distributed Processing, Vol. 2*. MIT Press, 1986.
- [12] J. Pustejovsky. *The Generative Lexicon*. MIT Press, 1995.
- [13] D. Rumelhart, J. McClelland, and The PDP Research Group. *Parallel Distributed Processing, Vol. 1*. MIT Press, 1986.
- [14] 中本敬子, 野澤元, 黒田航. 動詞「襲う」の多義性: カード分類課題と意味素性評定課題による検討. 認知心理学会第二回大会口頭発表, p. 39, 2004. [<http://cls1.hi.h.kyoto-u.ac.jp/~kkuroda/papers/Nakamoto-et-al-CogPsy2004-Original.pdf>].
- [15] 中本敬子, 黒田航, 野澤元, 金丸敏幸, 龍岡昌弘. FOCAL/PDS 入門: フレーム指向概念分析/並列分散意味論の具体的紹介. [未発表論文: <http://cls1.hi.h.kyoto-u.ac.jp/~kkuroda/papers/introduction-to-focal.pdf>], 2004.
- [16] 黒田航, 井佐原均. 意味フレームを用いた知識構造の言語への効果的な結びつけ. 電子情報通信学会技術研究報告, 第 104 (416) 巻, pp. 65–70. 電子情報通信学会, 2004. [増補改訂版: <http://cls1.hi.h.kyoto-u.ac.jp/~kkuroda/papers/linking-1-to-k-v3.pdf>].
- [17] 黒田航, 中本敬子, 野澤元. 状況理解の単位としての意味フレームの実在性に関する研究. 日本認知科学会 第 21 回大会 発表論文集, pp. 190–191, 2004.
- [18] 李在鍋. 他動詞のゆらぎ現象に対する「構文」的アプローチ. 言語科学論集, Vol. 7, pp. 1–20, 2001.
- [19] 李在鍋. 助詞「に」の定量的分析の試み: 語法研究の新たな手法を求めて. 日本認知言語学会論文集第 4 巻, pp. 55–65. 日本認知言語学会 (JCLA), 2004.
- [20] 李在鍋. 認知事象の複合的制約に基づく結果構文再考: 構文現象の体系的記述を目指して. 山梨 正明ほか (編), 認知言語学論考 No. 3, pp. 183–262. ひつじ書房, 2004.
- [21] 影山太郎 (編). 日英対照: 動詞の意味と構文. 大修館, 2001.