

# 複層意味フレーム分析 (MSFA) に 意味役割の典型的実現値の情報を付加してシソーラス化する試み

黒田 航 井佐原 均

独立行政法人 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<sup>1)</sup>

黒田・井佐原 [14] は意味役割の一般理論 [15] によって可能となる意味役割 (semantic roles) の意味型 (semantic types) からの区別によって、意味型を基準に構築されている従来のシソーラス (e.g., 日本語語彙大系 [7]) の不備を補うことが可能であると論じたが、そのための技術的な詳細に具体的に触れていない。この発表はその詳細を補う。具体的には、黒田・井佐原 [13] の提唱する複層意味フレーム解析 (Multilayered Semantic Frame Analysis: MSFA) に現われる意味役割 (e.g., 〈逮捕. 行逮捕者〉) に、分析対象となった文中で実際に実現値になっている要素 (e.g., “検察当局” IS-A 〈逮捕. 行逮捕者〉, “ディエゴ・マラドーナ” IS-A 〈被逮捕者〉) 以外の可能な値を補充することで、意味解析の結果をシソーラスに結びつけることが可能となるという提案を行ない、実装案を示す。ただし、現在のところ、結果はあくまでも試験的なものである。

## 2 意味役割の実現値

### 2.1 事の始まり

階層意味フレーム分析 [12] の作業中、第一著者は次のような選択制限の関係に特別の注意を寄せた:<sup>2)</sup>

- (1) a. ライオンに襲われた { 獲物; ?被害者; ガゼル (?\*の群れ); \*イワシの群れ } は深刻な被害を受けた。  
b. バラクラーダーに襲われた { 獲物; ?被害者; \*ガゼル (の群れ); イワシの群れ } は深刻な被害を受けた。
- (2) a. 三人の強盗に襲われた { ?\*獲物; 被害者; アサヒ銀行; ?アサヒビル; ??アサヒ組 } は深刻な被害を受けた。  
b. 三人の鉄砲玉に襲われた { ?\*獲物; 被害者; ???アサヒ銀行; ???アサヒビル; アサヒ組 } は深刻な被害を受けた。
- (3) a. 大型台風襲われた { \*\*獲物; ?\*被害者; 被災地; 九州 (地方); ?加山町 } は深刻な被害を受けた。  
b. 火山灰に襲われた { \*\*獲物; ?\*被害者; 被災地; ??九州 (地方); 加山町 } は深刻な被害を受けた。
- (4) a. 急激な発作に襲われた { \*\*獲物; ?被害者; \*被災者; (心臓病) 患者; 佐藤さん } は深刻な被害を受けた。  
b. 急激な不安に襲われた { \*\*獲物; ??被害者; \*被災者; ?(心臓病) 患者; 佐藤さん } は深刻な被害を受けた。

このような選択制限が示唆するのは、次の関係はおのこの意味クラスとそのインスタンスとの間の実現関係であることである:

<sup>1)</sup> この研究の発表のためのデータ準備の段階で、中本敬子 (京都大学) と野澤元 (NICT) から助力があった。二人に感謝したい。

<sup>2)</sup> この結果の解釈は中本・黒田 [10] に提示されている。

- (5) a. ライオンの“獲物”と“ガゼル (の群れ)”との関係  
b. バラクラーダーの“獲物”と“イワシ (の群れ)”との関係  
c. 三人の強盗の“襲撃”の“被害者”と“アサヒ銀行”との関係  
d. 三人の鉄砲玉の“襲撃”の“被害者”と“アサヒ組”との関係  
e. 大型台風の“被災地”と“九州 (地方)”との関係  
f. 火山灰の“被災地”と“加山町”との関係  
g. 発作の“(心臓病) 患者”と“患者”との関係
- (6) ただし、“佐藤さん”がインスタンスであるような“不安”の受け手の意味クラスには“被害者”、“患者”のような名称を与えることはできない。

一部の意味クラスは“獲物”、“被害者”という形で語彙化されているが、語彙化されたもの、されていないものを区別せずに、一般に意味クラスがどんなクラスであるかを明示化することが階層化意味フレーム分析 (Hierarchical Semantic Frame Analysis: HFNA) である [9, 12, 11]。

“ $x$  が  $y$  を襲う”、“ $y$  が  $x$  に襲われる”の解釈に関して言えば、次のような 15 個の  $x, y$  の意味クラスの対によって意味フレームが特定された (三つの例のみ提示):

- (7) a. 〈捕食者:  $x$ 〉が、〈獲物:  $y$ 〉を、〈捕食:  $z$ 〉のために、{ 襲う; 襲撃 } する  
b. 〈侵略者:  $x$ 〉が、〈犠牲者:  $y$ 〉を、〈侵略:  $z$ 〉のために、{ 襲う; 襲撃 } する  
c. 〈受難者:  $y$ 〉が、〈災難:  $x$ 〉に、〈NULL:  $z$ 〉のために、{ 襲われ; 見舞われ } する

$f$  がフレーム名、 $r_i$  が  $f$  を構成する意味役割名であることを表すのに、〈 $f$ 〉、〈 $f.r_i$ 〉という表記を用いる。ただし、〈 $f.r_i$ 〉の表記で“ $f$ ”の部分は多くの場合に省略可能である。

この際、特に重要な点は次である:

- (8) a. “捕食者”、“獲物”、“災害”、“被災者”、“加害者”、“被害者”のような語彙は、状況に個別化された意味クラス〈捕食者〉、〈獲物〉、〈災害〉、〈被災者〉、〈加害者〉、〈被害者〉を表わすための語彙である。  
b. その一方、概念化可能なすべての意味クラスが語彙化されているわけではない (実際、メタファーの基盤は、語彙化されていない概念 (化) を既成の語彙を用いて近似することだと解釈できる)。

### 2.1.1 意味役割名と意味型名の関係

(語  $*W$  で表わされる) 意味クラス  $C(*W)$  が (語  $W$  で表わされる) 実例  $I(W)$  で実現されることを “ $*W \Rightarrow W$ ” で表わすことにする。“捕食者”  $\Rightarrow$  “インパラ” や “自然災害”  $\Rightarrow$  “台風” などのクラスとインスタンスとの実現関係は “意味役割名  $\Rightarrow$  意味型名” という形で区別できる。

意味型 (名) と意味役割 (名) の区別は、孤立したものではない。例えば、Gentner ら [4, 5, 2] も意味役割と意味役割名の区

図1 (9)のMSFA

別にほぼ対応するものとして、**関係のカテゴリ** (relational categories) と**関係の名(詞)** (relational names/nouns) の研究を始めている。Gentner らの分析で私たちの言う意味型と意味型名に対応するのは**対象カテゴリ** (object categories) と**対象名(詞)** (object names/nouns) の区別である(彼らの研究文脈では対象名(詞)は実体名(詞) (entity names/nouns)とも呼ばれる)。

## 2.2

以上の一般化を下敷きにして、黒田・井佐原 [14] が主張したのは、意味役割名と意味型名とを注意深く区別すると、錯綜しがちなシソーラスを再構築できるということである。簡単な実例として彼らは“犬”、“番犬”、“柴犬”の例を挙げ、その方向性を示したが、再構築の十分な詳細を与えたわけではない。特に問題になるのは、{“番犬”、“番犬”}のような意味クラスはどこに、どれだけ見つかるかという問題が明らかでないことである。

だが、幸い、問題の意味クラスの探索法は複層意味フレーム分析 (MSFA) [13] という形で定式化され、徐々に研究成果が上がっている。本研究では、これを意味型名と意味役割名の区別を反映したシソーラス構築に有効利用する。

具体的には次の通りである。MSFA は文章の個々の文で使われている個々の語句に、その文脈で割り当てられる意味役割を網羅的に特定したものだと見せる。この意味役割を意味クラスとみなし、それを実際に文章中で用いられていないが実現値として可能な語句を追加することによって、疑似シソーラス化が可能である。以下にその実装案を示す。

## 3 MSFA の疑似シソーラス化

MSFA を疑似シソーラス化する手法の概要について、一例を取り上げながら説明する。

### 3.1 MSFA の実例

次の文は京大コーパスの一文 (S-ID: 950107210-002) で、MSFA によるタグづけが行われ、結果が公開されている。<sup>3)</sup>

(9) アルゼンチンの元サッカー選手、ディエゴ・マラドーナ

<sup>3)</sup> この例を含めて、幾つかの MSFA が <http://www.kotonoba.net/~mutiyama/cgi-bin/hiki2/hiki.cgi?MSFAList> で公開されている。ただし閲覧には第一著者による認証が必要。

ナ氏が六日、同国の検察当局に一次身柄を拘束された。

図1に示したMSFAは、(9)の理解内容が、(少なくとも)次の42タイプのフレーム群によって構成されることを示す:

- (10) (〜過去性の指定〜), (〜特徴づけ〜[1],[2],[3]), (〜役割の値の指定〜[1],[2]), (〜受身〜), (〜参照〜), (〈事件報道〉), (〈報道〉), (〈報告〉), (〈知らせ(る)〉), (〈知り〉), (〈語り〉), (〈事件〉), (〈有名人の紹介〉), (〈敬意の表明〉), (〈引退〉), (〈役職からの離脱〉), (〈サッカー選手としての活躍〉), (〈スポーツ選手としての活躍〉), (〈有名人としての活躍〉), (〈活躍〉), (〈団体競技〉), (〈賃金のための労働〉), (〈チームへの所属〉), (〈所属〉), (〈権力の行使〉), (〈力の行使 [+metaphoric]〉), (〈拘束 [+metaphoric]〉), (〈逮捕〉), (〈処罰〉), (〈行動の制限〉), (〈役割の代行〉), (〈政府機関への所属〉), (〈帰属〉), (〈犯罪性の判断〉), (〈法律の適用〉), (〈犯罪の摘発〉), (〈捜査〉), (〈調査〉), (〈容疑〉), (〈疑う〉), (〈犯罪〉), (〈?職務の担当?〉)

(*〜F〜*) は *F* が ((*〜参照〜*) などのように) 文法レベルのフレームであり概念的な内容をもつとは限らないことを、(*F* [+metaphoric]) は *F* がメタファーとして機能しており、“(*F*) targets *G*” の関係で指定される転移先フレーム *G* があることを、おのおの表わす。また、(*?F?*) は *F* の認定が不確定であることを表わす。

文の含意の形成には、フレーム間関係 (frame-to-frame relations) が重要な役割を果たす。わかっている限りでフレーム間関係には次のものがある:

- (11) a. *F* constitutes *G* (構成の関係: e.g., 〈判決〉 constitutes 〈裁判〉; 〈代金の支払い〉 constitutes 〈商品の購入〉; “*G* HAS-A *F*” とほぼ同じ)  
 b. *F* elaborates *G* (詳細化の関係: e.g., 〈判決〉 elaborates 〈決定〉; “*F* IS-A *G*” とほぼ同じ)  
 c. *F* presupposes *G* (前提の関係: e.g., 〈購入〉 presupposes 〈販売〉) presupposes 〈制作〉...)  
 d. *F* presumes *G* (見こみの関係: e.g., 〈販売〉 presupposes 〈購入〉, 〈制作〉 presumes [〈購入〉 presumes 〈購入〉])  
 e. *F* targets *G* (意味の転移の関係: 見かけは *F* だが、実質の意味は *G* によって与えられる: e.g., 〈束縛〉 targets 〈逮捕〉)

### 3.1.1 意味役割の実現値

フレームはおのおの、意味役割の集合である。例えば〈捜査〉フレームは次のような意味役割からなる:

- (12) 〈Agent: 捜査者: x〉が, 〈Patient: 犯罪者: y〉を, ..., 〈GOV: 捜査〉する)

ここでは, 〈f.r〉の表記で“:” 演算子が省略され, “=” で現われる同一視(あるいは異名化)が許されると考えている。詳細は次のように書くことができる:

- (13) a. 〈捜査. 実行者〉 = 〈捜査者〉 ⇒ { 〈警察 (官)〉, 〈検察 (官)〉, 〈検察局〉 }  
b. 〈捜査. 対象〉 = 〈捜査対象〉 ⇒ { 〈犯罪〉, 〈犯罪者〉 }

ここで, “=” で表わされる同一視の関係の意味論はほぼ自明だが, “⇒” で表わされる意味クラスの実現の関係の意味論は決して自明ではない。実際, “〈捜査. 実行者〉 ⇒ 〈警察 (官)〉” や “〈捜査. 対象〉 ⇒ 〈犯罪 (者)〉” を説明するのが状況レベルの意味フレームという概念である。

### 3.2 役割実現の記述言語の仕様

表 2 に挙げるのが (9) の MSFA に用いられた意味フレームの, 実際に文章に現われた語句の他にも可能な実現形を記述したものである。

意味役割の実現を効果的に記述するためには, 記述言語の使用されている。そのための記述言語の仕様を, 表 2 に挙げた具体例を元に説明する。

#### 3.2.1 役割名のスター指定

まず, “\*捜査. 捜査者” という表記に注意が必要である。MSFA に現われる表記は “捜査者” であるが, このシートでは “\*捜査. 捜査者” としている。これを称して「意味役割名にスター指定がある」と言う。ただし, 〈捜査者〉 = 〈捜査. 捜査者〉である。

一般に “\*X” は, 〈X〉が意味クラスを指定する概念ポインターで, 語彙ではないことを表わすと取り決めておく。この規約が必要なのは以下のような理由による:

- (14) 第一に, “捜査者” という語は 〈捜査. 実行者〉を表わすことができるが, この語は (“犯罪者”, “獲物”, “犯人” などとは異なり) 語彙化された要素ではない。これは “獲物” という語が “捕獲. 対象” を表わし, かつ, その意味で語彙化されているのと同じではない。両者は区別した方がよい。
- (15) 第二に, “逮捕者” が (〈捕獲. 実行者〉 ⇒ “捕獲者” などの場合と異なり) 〈逮捕. 実行者〉の意味を表わない, 〈逮捕. 対象〉 = “被逮捕者” の語彙化であるという例外的な場合も扱える必要がある。〈逮捕. 実行者〉 = “逮捕者” は誤りで, 〈逮捕. 対象〉 ⇒ “逮捕者” が正しい。

#### 3.2.2 名称の生成

表 2 では 〈刑事捜査. 捜査者〉の可能な実現値として, “[人名]”, “警察”, “~ 署”, “(〜) 捜査官” などが挙げられている。これを網羅的に与えるのは困難だが, このような指定をなるべく詳しく行なうことが MSFA の疑似シソーラス化と呼ばれる記述方法である。

効果的な記述のためには, メタ変数, 関数の使用が必要である。この点に関して以下で簡単に説明する。

“[人名]”, “警察”, “~ 署”, “(〜) 捜査官” は実現の際の働き方が次に示すように異なる:

- (16) a. “[人名]” は人名生成関数  
b. “警察” は字面 (literal) の文字列  
c. “~ X” は X に付加要素が義務的に存在する文字列

- d. “(〜)X” は X に付加要素が随意的に存在する文字列

ただし, “[人名]” が機能するための前提となる “佐藤 (春彦)” が 〈人名〉であるという知識は, 別個に与えられているという仮定に基づく。また, “[人名]” の指定は本来なら “[捜査官名]” であるべきだが, “[捜査官名]” という関数を一般に定義することは困難, かつおそらく可能でも無意味なので, より一般的な “[人名]” で代替するのが適切だと考える。なお, “\*\*\*\*\*” は典型的な実現値がないことを表わす。

具体的には, これらは次の (18) を 〈刑事捜査〉フレームに結びつけ, 〈刑事捜査. 捜査者〉の可能な実現値であることを記述する:

(17) X は, 犯人は関係者と見て取り調べている。

(18) X ⇒ { a. 角砂糖 ([人名] あるいは [人の渾名] によって認可されるなら); b. 警察; c. 七曲署; d. (佐藤 (春彦)) 捜査官 }

特に典型的な役割名を表わす形態素 (“~ 署”, “~ 社”) を特定することは重要である。実際, 一部の派生形態素は意味役割を指定する演算子と見なしてよい側面もある [6]。実際, “~ 者” (“犯罪者”, “解答者”) や “~ 手” (“歌い手”, “持ち手”) のその機能が, 行為者タイプの意味役割の定義を可能にしている (ただし, 前述の “逮捕者” や “取っ手” は例外となる)。

#### 3.2.3 役割名の横取り

一部では普通名詞が役割名を横取りし, 占有している。〈喚問. (行) 喚問者〉 = “裁判所”, 〈裁判. (行) 裁判者〉 = “裁判官”, 〈刑事捜査. 捜査者〉 = “警察”, 〈起訴. 被起訴者〉 = (〈告訴. 被告告訴者〉) = “被告 (人)” などがその例である。実際, “(行) 喚問者”, “(行) 裁判者” のような表現は単なる “裁判所”, “裁判官” の迂言となる。

#### 3.3 意味記述の粒度と典型値の有無との相関

多くのフレームに可能な実現値を特定しようとするのとわかることだが, 典型値をもちやすい意味役割と典型値をもちにくい意味役割が存在する。(10) に挙げた 42 タイプのフレームのうち, (〜過去性の指定〜), (〜役割の値の指定〜), (〜受身〜), (〈知らせ (る)〉), (〈知り〉), (〈疑い〉), (〈敬意の表明〉), (〈語り〉), (〈報告〉) の 9 タイプに, 程度の差こそあれ, 典型値を認めがたかった。これはこれらのフレームの内容が一般的すぎて, 特定の状況を絞りこみにくいからである。この傾向は特に文法フレームには顕著であった (例外は (〜参照〜) である。これは指示詞と先行詞の関係という形でハッキリしているからである)。これは意味役割の実現値が概念階層を下がるにつれて, 状況の特定が容易になると解釈できる。

この典型値の有無という特徴は, 従って, フレームの記述粒度の評価に有効であろう。粒度がある程度細かいフレームでないと, 典型的な実現値というのは想定しがたいからである。実際, 下位レベルになればなるほど, 典型値は代表値になり, 最後には唯一可能な値になる。実際, 参与項の特定化 = 該当範囲の絞り込みの程度によってフレームの粒度を測定することも可能なはずである。これは元々, HFNA [9, 12, 11] が仮定する概念構造の特徴でもある。

## 4 終りに代えて: されど, これは正真正銘のシソーラス構築にあらず?

本稿では, MSFA の記述に現われる意味役割に, 文章に現われている実現値の他に可能な実現を作例によって補うというやり方で, シソーラスに結びつける手法を提案した。だが, この手法で正真正銘のシソーラスが構築可能かどうかは, 今のところ明らかではない。

実際, 提案手法には長所と短所がある。短所としては, 提

