

状況の喚起と文意の決定 (≠構築)

生成辞書理論を超えて

★★★★

黒田 航/NICT

★★★★

第25英語学会ワークショップ

「語彙情報と文法現象」

2007/11/10 名古屋大学

辞書と文法の関係

- * 理論言語学者の多くは“辞書”と“文法”を峻別する
 - * 関心の中心は文法情報で語彙情報はせいぜいオマケ
- * だが文法の定義から考えて辞書なしの文法は無意味
 - * $G = (S, R, C, V)$, S = symbols, R = rules, C = constraints, V = vocabulary.
 - * C には S, R 側に属するものと V 側に属するものがある

方法論的基準

* “はじめに文法情報ありき”は論点先取

* 説明責任から考えて、語彙的情報では説明できないもの(のみ)を文法情報とすべきであり、その逆ではない

* だが、この基準が意味をもつには次の条件が必要

* 語彙情報とは何か、特にどんな意味情報が語彙情報でないかが判明している。

二種類の区別

* 区別 1:

* 文法情報と語彙情報の区別 (内側の境界条件)

* 文法 (grammar) と辞書 (lexicon) の区別に対応

* 区別 2:

* 語彙情報と非語彙情報 = 一般知識との区別 (外側の境界条件)

* 意味論 (semantics) と語用論 (pragmatics) の区別に対応

私が模索する方向

- * 情報を辞書に書くか文法に書くかを二者択一し、記述の冗長性を排除する方向性は好ましいか?
- * 従来のように記述の経済性の指標を記述量に限定せず、それに実行/処理時間も含めた場合、本当に経済的な記述とはどんなものか?
- * 文法と辞書、意味論と語用論の原理的な区別は、その場合でも必要なものか?

私の理論上のバイアス

* Implementational Occam's Razor

- * パフォーマンスの別に反映しない区別はしないに限る
- * “エレガント”なモノより，ゴチャゴチャしていても“チャンと動く”モノの方が魅力的

* 従って

- * 実装上の必要性が不明なら，文法と辞書，意味論と語用論を区別しない

発表の流れ

- * 状況の喚起の分類と例の提示
- * “(yの)犯人”の意味の分析
 - * “(yの)犯人”の特質構造 (Pustejovsky 1995, 2001)
 - * “犯人”の格フレーム辞書 (河原・黒橋 2003) データ
 - * “犯人”のコーパス分析 (断片)
- * 考察

状況の喚起の分類



状況の喚起とは？

* 定義1

- * 表現 e (典型的には文 s) の解釈で、特定の状況 s が (その理由は無問にして) 読み取られること
- * 喚起はフレーム意味論 (Fillmore 1982, 1985; Fillmore et al. 2003) の *evocation, evoke* の訳
- * 喚起は想起 (reminding) (Schank 1999, 2001) の一種だろう
- * 語彙的 (lexical) 喚起と超語彙的 (superlexical) 喚起を区別

喚起が自明ではない理由

- * 定義1が意味をもつための条件
 - * 状況のリストは e の意味記述から独立して与えられている
 - * さもなければ論点先取であり，仮定は常に成就する
- * 独立性の条件は従来の言語学では考慮されていない
 - * FrameNet (Baker et al. 1998; Fillmore et al. 2003) などが状況のデータベースを構築中

状況の語彙的喚起

- * 定義 2:

- * 語句 w_i が単独に使われた時に状況 s の喚起が起こること

語彙的喚起の例

(1) a. ... 警察 ...; b. 犯人 ...; c. ... 逮捕 ...

(1)i. “警察”とは、〈犯罪者〉の〈逮捕〉などで〈治安〉の〈維持〉
に貢献する機関

(1)ii. “犯人”とは、いつか、どこかで〈犯罪〉(e.g., 〈強盗〉, 〈誘
拐〉)を犯した人物

(1)iii. “逮捕”とは〈〈警察〉による, 〈犯罪者〉の, 〈逮捕〉〉のこと

複合的=超語彙的喚起とは

* 定義 3:

- * 語句 w_i と w_j とが単独に使われた時には状況 s の喚起は起こらないが、 w_i, w_j の両者が同時に現われた時に(は) s の喚起が起こること

超語彙的喚起の例

(2) 彼の逮捕は [身から出た錆] だ。

(2)i. [錆(x)] の意味と [身(y)] の意味と [x が y から出た] の意味を合成して t : [当然の報い] という意味を得るのは (事後承諾以外の形では) 無理

(2)ii. “身から出た錆” の t という意味は超語彙的 (superlexical) に喚起されていると言う (しかない)

- ただし慣用表現の意味は超語彙的喚起の特殊な場合でしかない

分散的喚起とは

* 定義 4:

- * 語句 w_i と w_j の超語彙的喚起が ($w_i + w_j$ の一語句ではなく) 連続しない位置に “分散して” 現われた時にも s の喚起が起こること

* 注意

- * 分散的喚起は超語彙的=複合的喚起の特殊な場合

分散的(超語彙的)喚起の例

(3) 彼らは ... 犯人を ... 捕まえた.

(3)i. “犯人”の意味は [犯罪(y)の犯罪者(x)] であるとして,

(3)ii. “捕まえた”の意味は (“犯人”の意味との共合成を通じて) [逮捕した] に詳細化され

(3)iii. “彼ら”の意味は (同じ共合成の結果として) [警察 or 検察] に詳細化される

注意

* 次の文で“捕まえた”は“逮捕した”と詳細化されない

(4) 彼らはトラを捕まえた。

* “ z が x を捕まえた”の意味が $\langle z$ が x を逮捕した \rangle になるのは
(当然と言えば当然だが) x が[犯人]で z が[警察]か[検察]
の場合に限る

* だが、なぜ?

* この問題を「あたり前」で済ませない

分散的喚起の概略 1/6

(1) 警察が犯人を逮捕した.

A. [S が犯人を逮捕した] evokes $F(-S) = \{f^s_1, f^s_2, \dots, f^s_b\}$

B. [警察が O を逮捕した] evokes $F(-O) = \{f^o_1, f^o_2, \dots, f^o_m\}$

C. [警察が犯人を V (した)] evokes $F(-V) = \{f^v_1, f^v_2, \dots, f^v_n\}$

* $F = \{F(-S), F(-O), F(-V)\}$ からの Winner-Take-All 型の競合

* ただし勝者は一人ではない

分散的喚起の概略 2/6

- * 次の二種類の作用の連鎖反応のサイクルの結果 (=安定状態) が解
 - * 関連のあるフレーム f_i と f_j は強めあう
 - * 関連のないフレーム f_i と f_j は弱めあう
- * 並列計算なら時間=サイクル数は意外とかからない

分散的喚起の概略 3/6

* (1)A, B, C は (1)D, E, G の語彙的喚起の単位と別のレベルに存在する超語彙的喚起の単位

(1) 警察が犯人を逮捕した。

D. [警察が O を V (した)] evokes $F(-O, -V)$

E. [S が犯人を V (した)] evokes $F(-S, -V)$

F. [S が O を逮捕した] evokes $F(-S, -O)$

分散的喚起の概略 4/6

- * (1)F は動詞中心の語彙意味論 (e.g., 影山1993, 影山 (ed.) 1996; Jackendoff 1991) の記述対象
- * (1)D, E は生成辞書 (Pustejovsky 1995, 2001) の特質構造 (qualia structure) などの語彙意味論の記述対象
- * 注意
 - * (1)A, B, C は語彙意味論の記述対象ではないが、これがないと文の意味記述と (1)D, E, F の語の意味記述を繋げない

分散的喚起の概略 5/6

* Pattern Matching Analysis (PMA) (Kuroda & Nakamoto 2007) の用語で
言う

* (1)A, B, C 超語彙的パターン (superlexical subpatterns: SLPs)

* (1)D, E, F は語彙的パターン (lexical subpatterns: LPs)

* Kuroda (2000, 2001) 版の PMA ではこの区別はなかった

* SLPs は Construction Grammar (Fillmore 1988; Goldberg 1995) で言う
“構文” (constructions) の形式とほぼ同一視可能

分散的喚起の概略 6/6

- * 次の二つの可能性のいずれが、どの場合に妥当するかは一般には不明
 - A. 個々の LPs によって独立に喚起された状況が一つの SLP に対応する状況に絞り込まれる
 - B. SLPs は LPs とは独立に状況を喚起する
 - * B は慣用表現に顕著で、分散性を許さない傾向があるように思える

補足

- * 状況の超語彙的喚起は池原ら (2005, 2006) の言う表現の
非線形性 (nonlinearity) の源泉

語順の影響

- * オンライン処理と考えた場合,

- * 後から現われる要素 (e.g., V) の意味は, すでに現われている要素 (e.g., S, O) が要求している情報を (語義の一部として) 使うことができる

- * ただし

- * (相対的に) 初めに現われる要素の意味は(相対的に) 後に現われる要素の意味で変更される可能性もある

“*Y*の犯人”の意味記述

推論 1

(3) 彼らは犯人を捕まえた。

* 推論

(3)i. “犯人”とは<犯罪>を犯した<犯罪者>のこと

(3)ii. “捕まえた”というのは<逮捕した>のこと

(3)iii. “彼ら”というのはおそらく[警察(官)]か[検察(官)]のこと

推論 2

(3) 彼らは犯人を捕まえた。

* 推論

(3)i. “犯人”は<犯罪者>と言うほど悪いヤツではなく単なる<つまみ食い手>で

(3)ii. “捕まえた”というのも<逮捕>という意味ではなく、<誰が該当者かを確認めた>といった程度

推論 3

(3) 彼らは犯人を捕まえた。

* 推論

* “彼ら”が[マフィア]で

* “犯人”は<裏切り者>で

* “捕まえた”というのも<逮捕>という意味ではなく、<非合法的に捕捉/拘束>という意味

推論 4

(3) 彼らは犯人を捕まえた。

* 推論

* “彼ら”が[モグラ]で

* “犯人”は<畑を荒らす者>で

* “捕まえた”というのも<逮捕>という意味ではなく、<(悪さをする)動物の捕獲>という意味

許されない推論 1

(3) 彼らは犯人を捕まえた。

* 推論

* “彼ら”が〈猟師〉で

* “犯人”は〈猟の獲物〉で

* “捕まえた”というのも〈獲物の捕獲〉という意味

問題の定式化

- * 許される推論は色々あるが“何でもアリ”ではない
 - * 問題は、ある解釈が不可能だということではなく、ある解釈 I は他の解釈 I' より (処理負荷がかかり) 難しい
- * (一定労力内で) 許される推論の範囲の特定は言語学が答えるべき正統な問題の一つ
 - * でないと意味の認知科学に貢献できず、(工学, 教育目的に) 有用な辞書も書けない

“(yの)犯人”の意味論

(5) 彼らは y の犯人を捕まえた。

- a. 彼らは強盗の犯人を捕まえた。
- b. 彼らはつまみ食いの犯人を捕まえた。
- c. 彼らはボス暗殺の犯人を捕まえた。
- d. 彼らは畑荒らしの犯人を捕まえた。

注意

- * “犯人”という名詞が関係名詞 (西山 1991, 2003; Gentner & Kurtz 2005) と言う (= 文法情報の提供?) だけでは不十分
- * [(y の) 犯人(x)] の y は<犯罪>に限定されないが
 - * メタファー的拡張 (Lakoff & Johnson 1980, 1999) を許す?
- * [(y の) 犯人(x)] の解釈は一定の範囲に収まっている
 - * y を特定するにはキメ細かい状況の区別が必要

語義を一般化すれば？

* “(yの) 犯人(x)” の意味として, (5) の代わりに (6) を考えるのは被覆率を上げる一方で精度を下げる

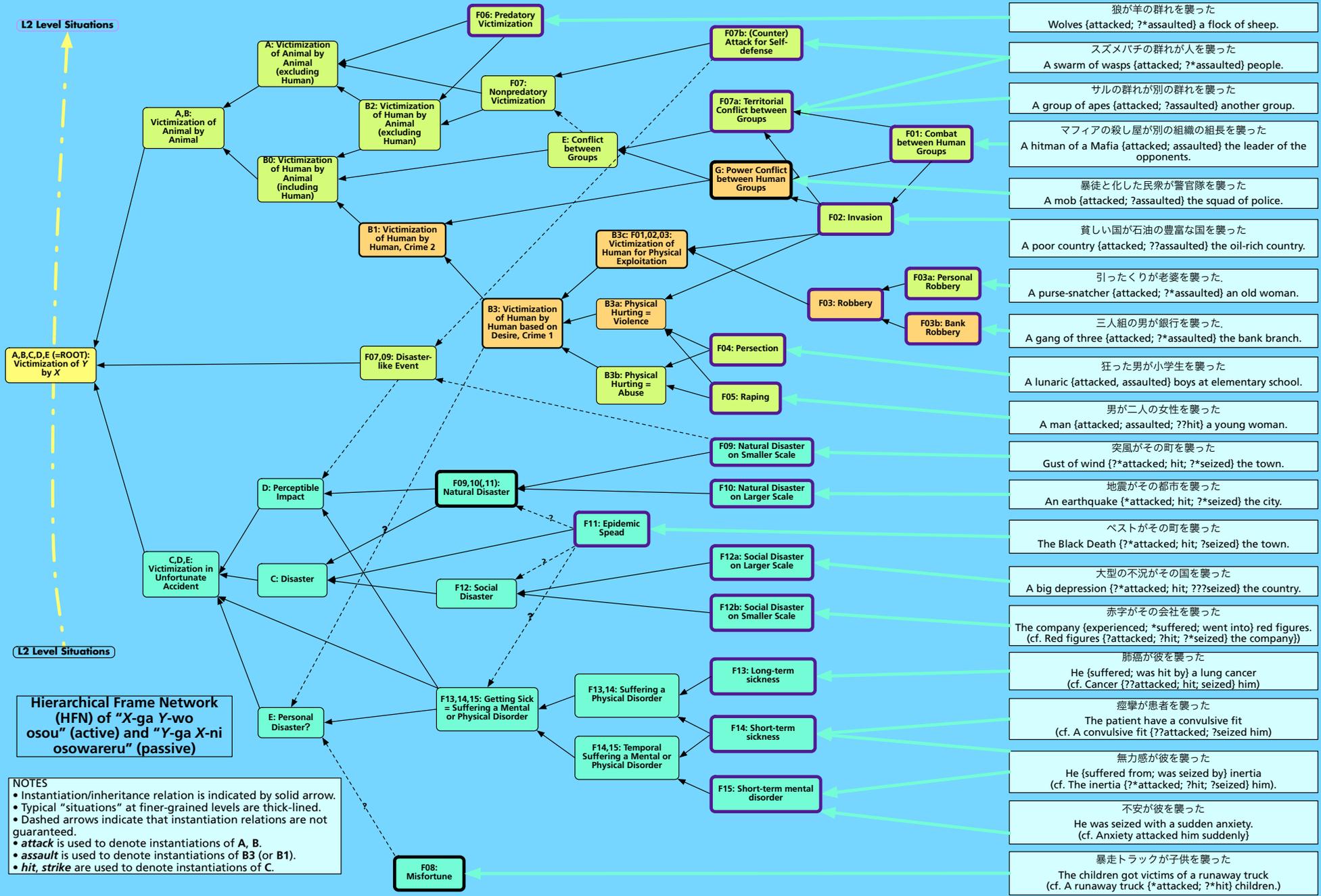
(6) <<犯人(x)>が, <時間(t)>に, <場所(l) で>, <被害者(v)> を生む<犯罪(y)> を犯す>

t は “犯行時間”, *l* は “犯行現場”

(7) <<加害体(x)>が, <時間(t)>に, <場所(l)>で, <被害者(v)>に害(y) を与える>

More Abstract, Coarse-grained

More Concrete, Finer-grained



Frame ID	文 (n(評定者) = 11)	襲った	攻撃した	襲撃した	見舞った	襲いかかった	文 (n(評定者) = 1)	襲い	(の)襲撃	(の)攻撃	(の)殺害	(の)虐待	(の)侵略	(の)強盗	(の)強姦	(の)破壊	(の)停滞
F01	敵対する二人の組員が組長を__	1.0	0.8	1.0	0.1	1.0	敵対する二人の組員が組長__の犯人だった	0.0	1.0	1.0	1.0	0.5	0.0	0.0	0.0	0.0	0.0
F01	暴徒と化した民衆が警官隊を__	1.0	0.9	0.9	0.0	1.0	暴徒と化した民衆が警官隊__の犯人だった	0.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
F01	森の西側の部族が北側の部族を__	1.0	1.0	1.0	0.1	0.9	森の西側の部族が北側の部族__の犯人だった	0.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
F02	資源の乏しい国が隣国を__	1.0	1.0	0.8	0.1	0.7	資源の乏しい国が隣国__の犯人だった	0.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	0.0	0.0
F02	ドイツの戦車部隊がブタベシュを__	1.0	1.0	0.9	0.0	0.8	ドイツの戦車部隊がブタベシュ__の犯人だった	0.0	1.0	1.0	0.0	0.5	1.0	0.0	0.0	1.0	0.0
F02	テロリストの集団がアメリカ軍基地を__	1.0	1.0	1.0	0.0	0.7	テロリストの集団がアメリカ軍基地__の犯人だった	0.0	1.0	1.0	0.0	0.0	0.0	0.5	0.0	1.0	0.0
F03	二人組の強盗が都内の銀行を__	1.0	0.4	1.0	0.0	0.4	二人組の強盗が都内の銀行__の犯人だった	0.0	1.0	0.5	0.0	0.0	0.0	1.0	0.0	1.0	0.0
F03	外国人のグループが現金輸送車を__	1.0	0.4	1.0	0.0	0.7	外国人のグループが現金輸送車__の犯人だった	0.0	1.0	0.5	0.0	0.0	0.0	1.0	0.0	1.0	0.0
F03	覆面をかぶった男が銀座の宝石店を__	1.0	0.3	0.9	0.0	0.4	覆面をかぶった男が銀座の宝石店__の犯人だった	0.0	1.0	0.5	0.0	0.0	0.0	1.0	0.0	1.0	0.0
F04	通り魔がその小学生を__	1.0	0.4	0.5	0.0	1.0	通り魔がその小学生__の犯人だった	0.0	0.5	0.5	1.0	1.0	0.0	0.0	0.5	0.0	0.0
F04	5,6人の少年たちが公園にいた浮浪者を__	1.0	0.5	0.7	0.1	1.0	5,6人の少年たちが公園にいた浮浪者__の犯人だった	0.0	1.0	0.5	1.0	1.0	0.0	1.0	0.5	0.0	0.0
F04	23歳無職の男性が何人かの通行人を__	1.0	0.4	0.7	0.0	1.0	23歳無職の男性が何人かの通行人__の犯人だった	0.0	0.5	0.5	1.0	0.0	0.0	1.0	0.5	0.0	0.0
F05	ストーカーが近所の女性を__	1.0	0.1	0.4	0.1	1.0	ストーカーが近所の女性__の犯人だった	0.0	1.0	0.5	1.0	1.0	0.0	1.0	1.0	0.0	0.0
F05	無職の男が一人暮らしの若いOLを__	1.0	0.1	0.3	0.1	1.0	無職の男が一人暮らしの若いOL__の犯人だった	0.0	1.0	0.5	1.0	1.0	0.0	0.5	1.0	0.0	0.0
F05	店長がアルバイトの女子店員を__	1.0	0.3	0.2	0.1	1.0	店長がアルバイトの女子店員__の犯人だった	0.0	0.5	0.5	1.0	1.0	0.0	0.5	1.0	0.0	0.0
F06	サメが傷ついたイルカを__	1.0	0.5	0.4	0.0	0.9	サメが傷ついたイルカ__の犯人だった	0.0	1.0	0.5	0.0	0.5	0.0	0.0	0.0	0.0	0.0
F06	ライオンがインパラの群れを__	1.0	0.5	0.5	0.0	1.0	ライオンがインパラの群れ__の犯人だった	0.0	1.0	0.5	0.0	0.5	0.0	0.0	0.0	0.0	0.0
F06	ハイエナの群れが国立公園の監視員を__	1.0	0.4	0.5	0.0	1.0	ハイエナの群れが国立公園の監視員__の犯人だった	0.0	1.0	0.5	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F07	スズメバチの大群が子供達を__	1.0	0.6	0.6	0.0	1.0	スズメバチの大群が子供達__の犯人だった	0.0	1.0	0.5	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F07	イノシシがキノコ採りに来ている男性を__	1.0	0.4	0.4	0.0	0.9	イノシシがキノコ採りに来ている男性__の犯人だった	0.0	1.0	0.5	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F07	毒蛇が41歳の登山客を__	1.0	0.4	0.3	0.0	1.0	毒蛇が41歳の登山客__の犯人だった	0.0	0.5	0.5	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F08	交通事故が買い物婦の主婦を__	1.0	0.0	0.0	0.2	0.6	交通事故が買い物婦の主婦__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
F08	暴走したトラックが大通り沿いの保育園を__	0.7	0.0	0.0	0.1	0.6	暴走したトラックが大通り沿いの保育園__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0
F08	大きな災害が原発付近の住民を__	1.0	0.0	0.0	0.3	0.8	大きな災害が原発付近の住民__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F09	突風がテレビのリポーターを__	1.0	0.0	0.1	0.4	0.7	突風がテレビのリポーター__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F09	土砂崩れが民家を__	1.0	0.0	0.1	0.4	0.6	土砂崩れが民家__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	1.0	0.0
F09	鉄砲水が避難する住民を__	1.0	0.0	0.1	0.5	0.5	鉄砲水が避難する住民__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F10	大洪水が東海地方を__	0.9	0.0	0.0	0.2	0.5	大洪水が東海地方__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
F10	大型台風が日本列島を__	1.0	0.0	0.0	0.4	0.8	大型台風が日本列島__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0
F10	直下型地震が神戸を__	1.0	0.0	0.0	0.3	1.0	直下型地震が神戸__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
F11	悪性のインフルエンザがアジア諸国を__	1.0	0.0	0.0	0.4	0.6	悪性のインフルエンザがアジア諸国__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.5
F11	ベストが川沿いの町を__	1.0	0.0	0.1	0.4	0.6	ベストが川沿いの町__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.5	0.0
F11	エイズがアフリカの国々を__	0.9	0.0	0.0	0.4	0.5	エイズがアフリカの国々__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.5
F12	大型の不況が南米の国々を__	1.0	0.0	0.0	0.4	0.6	大型の不況が南米の国々__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5
F12	株価の暴落が株式市場を__	1.0	0.0	0.0	0.4	0.6	株価の暴落が株式市場__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0
F12	狂牛病問題が畜産業界を__	0.8	0.0	0.0	0.4	0.7	狂牛病問題が畜産業界__の犯人だった	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0
F13	悪性のガンが働き盛りの男性を__	0.9	0.0	0.0	0.3	0.5	悪性のガンが働き盛りの男性__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F13	脳卒中が隣のおじさんを__	0.7	0.0	0.0	0.3	0.5	脳卒中が隣のおじさん__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F13	拒食症が知人の妹を__	0.7	0.0	0.0	0.3	0.4	拒食症が知人の妹__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F14	痙攣が難病の少女を__	0.8	0.0	0.0	0.2	0.7	痙攣が難病の少女__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F14	眠気が雪山遭難者を__	1.0	0.0	0.0	0.2	0.6	眠気が雪山遭難者__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F14	脱力感が仕事中の男性を__	1.0	0.0	0.0	0.3	0.4	脱力感が仕事中の男性__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F15	言いようのない不安が敏腕の社長を__	1.0	0.0	0.0	0.2	0.7	言いようのない不安が敏腕の社長__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F15	激しい嫉妬がいつも温和な若者を__	0.8	0.0	0.0	0.2	0.5	激しい嫉妬がいつも温和な若者__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
F15	罪悪感が万引きした少年を__	0.9	0.0	0.0	0.2	0.5	罪悪感が万引きした少年__の犯人だった	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0

分散喚起を認める必要性

- * 分散的喚起を仕組みとして認めないと、高い精度と高い被覆率の両立はおそらくムリ
- * もちろん状況の分散的喚起の計算機上の実装は困難
 - * かなり大規模なコネクショニストモデルなどが必要だろう

意味論か?語用論か?

- * これは意味論の問題ではなく語用論の問題である (ので言語学者は係わらなくて良い) という論者は多い
- * だが, これは明らかに
 - * 言語の個体獲得 (≠ 文法の個体獲得) の部分問題の一つ
 - * 言語 (の意味) の認知科学の部分問題の一つ
 - * “使える”辞書の編纂の部分作業の一つ

注意 1

- * 自分は言語学の外部に対して何の貢献もしないと言うのでない限り，ここで語用論と意味論の区別をもち出しても**免罪符**にはならない
- * ただし，単に意味論と語用論の区別をなし崩しにしても**問題の解決**にはならない
- * それは必要なだけであって，十分ではない

注意 2

- * 百科事典的知識を使った記述にも制約が必要
 - * 百科全書的知識は原理的には何でも記述可能 (=被覆率は最大) だが、精度は最低
- * 百科事典的知識の利用と言えども、効率化のための制約を被っていると考えるのは妥当
 - * が、これを理由に「規則がある」とは言うのは論点先取

特質構造は救いになるか？



“犯人”の特質構造

* 直観的に次は明らかに思える

* Agentive: [犯人(x)が, 犯罪(y) を, 犯す] \Rightarrow [NP y の犯人(x)]

* Telic: [警察(z) or 検察(z) が, 犯罪 (y) の犯人(x) を, 逮捕する]

* だが, このような特質構造を設定するための独立した根拠はあるのか?

* 説明のアドホック性を回避するにはそれが不可欠

“犯人”の格フレームデータ

* [y の犯人(x)] の Telic 役割を示唆

* z_1 が (y の) x を搜索

* z_2 が (y の) x を逮捕

<http://reed.kuee.kyoto-u.ac.jp/cf-search/>

<http://reed.kuee.kyoto-u.ac.jp/~ryohei/ncf-search/index.cgi>

捕まえる:動4	ヲ格	626
追う:動11	ヲ格	493
探す:動42	ヲ格	415
いる:動1	ガ格	401
捕まる:動3	ガ格	341
逮捕:動3	ヲ格	303
特定:動9	ヲ格	302
捜す:動10	ヲ格	295
見つける:動46	ヲ格	280
追い詰める:動2	ヲ格	257
突き止める:動6	ヲ格	239

“犯人”の格フレームデータ

* 低頻度の小節構造内のト格との共起が Agentive 役割を示唆

* z_1 が x {を, が} y の犯人と疑う

* z_2 が x {を, が} y の犯人と断定

* z_3 が x {を, が} y の犯人と主張

断定:動3	ト格	70
言う:動3	ト格	56
疑う:動59	ト格	55
確信:動30	ト格	47
思う:動348	ト格	46
考える:動979	ト格	38
思い込む:動18	ト格	36
決めつける:動6	ト格	36
主張:動234	ト格	20
対峙:動36	ト格	16
する:動2	ト格	16
戦う:動369	ト格	15
対決:動29	ト格	14
睨む:動151	ト格	11
みなす:動213	ト格	11
決まる:動1	ト格	11

“犯人”の格フレームデータ

- * ただ、次のようなデータは解釈が難しい
 - * z_1 が (犯罪(y)の) 犯人(x) を (人(p)と) 特定
- * これは Telic かつ Agentive?

特質構造理論の限界1/2

- * (Agentiveと) Telic の連鎖構造=シナリオをどうする?
 - * [犯人(x) が 犯罪(y) を 犯す] prepares
 - * [捜索者(z_1) 犯罪(y) の 犯人(x) を 捜索する] presumes
 - * [逮捕者(z_2) が 人物(w) を 犯罪(y) の 犯人(x) の 容疑者(c) として 逮捕する] prepares
 - * [犯罪(y) の (犯人(x)) を 決める) 裁判で, 裁定者(z_3) が 容疑者(c) が 犯人(x) か どうか 裁定する]

特質構造理論の限界2/2

- * 今のままではメタファーがうまく扱えない
- * エイズの「犯人」(3例)
- * 円高不況の犯人(1例)
- * 株価暴落の犯人(1例)
- * 交渉つぶしの“犯人”(1例)
- * 一連の事故の「犯人」(1例)
- * 日英新聞記事対応づけデータ (JENAAD: 内山・井佐原 2003) では約4% (= 7/190) がメタファー例

Corpus	S-ID	L	T	R	E	犯人の 訳	yの値	犯人の値
p11-1989-20 01.eucj.txt	19980619J1TYMAA 1400060/1998061 9E1TDY01D000010 /10	<J>一連の事故の「	犯人	」は、上流の紙パ ルプ工場の未処理 廃水だった。</J>	<E>The pollution was traced to an upstream pulp mill that was releasing untreated wastewater into the river. </ E>	NONE	一連の事故 [+metaphoric]	上流の紙パ ルプ工場の 未処理廃水
p11-1989-20 01.eucj.txt	19960301J1TYMAC 1400120/1996030 1E1TDY06C000060 /13	<J>▽八四年九月、 仙台市で開かれた第 六回国際ウイルス学 会議で、エイズの	犯人	はウイルスである ことが確定した。 </J>	<E>-At the 6th International Virology Conference in Sendai in September 1984, experts acknowledged the existence of a virus that led to AIDS. </E>	NONE	エイズ [+metaphoric]	ウィルス
pnm-1989-2 001.eucj.txt	19960301J1TYMAC 1400120/1996030 1E1TDY06C000060 /11	<J>▽八三年の早い 時期までに、厚生省 や研究班の中では、 ウイルスがエイズの	犯人	であるとの疑いを 持ち、血液や血液 製剤を介して感染 する危険を認識し ていた。</J>	<E>-By early 1983, the ministry and its AIDS study group was aware that HIV might be responsible for destroying the human immunological system. They also recognized the danger that HIV might spread via blood and blood products. </ E>	X [+inferred] responsible for Y	エイズ [+metaphoric]	ウィルス
p11-1989-20 01.eucj.txt	19930905J1TYMAG 1400010/1993090 5E1TDY06C000010 /25	<J>いずれにせよ、 ここで円高不況の	犯人	探しをしても意味 がない。</J>	<E>It is meaningless to try to identify the party responsible for allowing the current recession, which was caused by the strong yen, to take on such grave proportions. </E>	party	円高不況 [+metaphoric]	UNSPEC
p11-1989-20 01.eucj.txt	19990318J1TYMAC 1400060/1999031 8E1TDY06D000060 /18	<J>急騰の立役者 が、何かのきっかけ で株価暴落の	犯人	に転じる心配は常 につきまとう。</ J>	<E>The sort of money that has been instrumental in causing share prices to soar could also make them plummet. </ E>	X [+inferred] instrumenta l in causing Y	株価暴落 [+metaphoric]	急騰の立役 者[+potential]

注意

- * 全体的に“犯人”の英語訳は語彙的変動が大きい
- * 理由はおそらく
 - * 英語には“犯人”ほど generic な語はなく
 - * *culprit, criminal, suspect, perpetrator* は“犯人”ほど用法が広くない
 - * 特定の犯罪者 (e.g, *hijacker, kidnapper, killer*) で記述されるか
 - * 一般指示 (e.g., *they, people, those (who ...)*) で済まされる

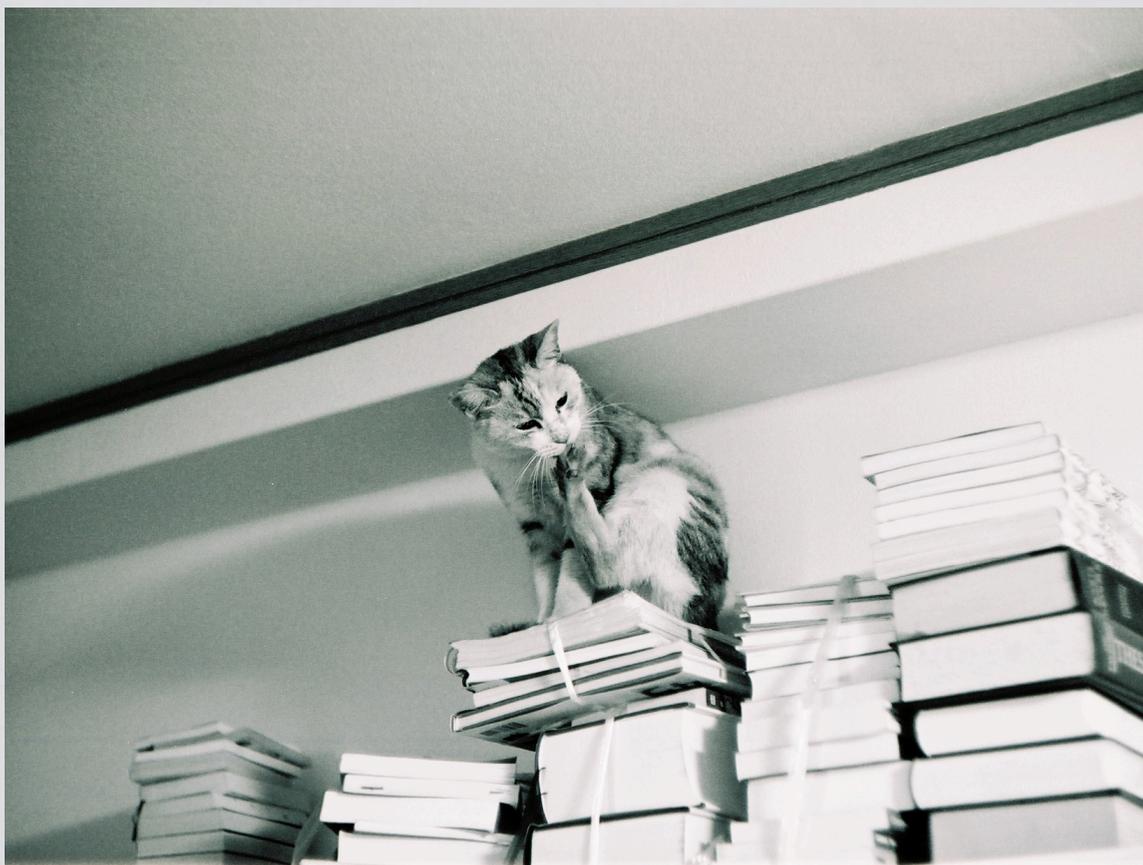
注意

- * “ y の犯人(x)”の一部は日本語特有のメタファー表現?
 - * y が<害>である<加害体 (harm-causer)>が先概念
- * ただし [[加害体は犯人]] のような概念メタファーは過剰般化 (= 高被覆率 AND 低精度)
 - (8) ?* 絶望が彼の自殺の犯人だった.
 - (9) ?* 地震が家屋倒壊の犯人だった.

(健全な) 懐疑

- * 特質構造は実在するのか? もっとごちゃごちゃした知識の近似ではないのか?
- * (例えば精度の良い格フレーム辞書という形で) 十分な共起情報があれば, それを近似できる
- * 特に Agentive 役割と Telic 役割の必然性は疑わしい
- * 実態は, 状況のネットワーク構造 N と個体の N への参与のインデックス集合では?

考察



文意の決定に関する一般化

- * 脱曖昧化は動詞の意味に対してのみ起こるのではなく、名詞(句)の意味に対しても起こる。
- * 意味の相互調節 (mutual accommodation) (Langacker 1987, 1991)
- * 共合成 (co-composition) (Pustejovsky 1995, 2001)
- * 語義の脱曖昧化=曖昧性の解消 (sense disambiguation) はどれも(状況の)分散的な(超語彙的)喚起である
- * ただし SLPs の意味指定が LPs の意味指定に優先される

注意

- * 次のような理論的”見通し”には根拠がない:
 - * 述語 p の意味が決まってから項 a_1, \dots, a_n の意味が決まる
 - * 項の意味が決まってから述語の意味が決まる
- * 実態は次以外のものではない:
 - * 述語の意味と項の意味が同時に決まる

“選択”としての文意の決定

- * 次の選択的処理としか考えられない
 - * 述語 p の意味 m_i と項群の意味(列) $m(N | m_i) = \langle m(a_1 | m_i), \dots, m(a_n | m_i) \rangle$ は対 $\langle m_i, \langle m(a_1 | m_i), \dots, m(a_n | m_i) \rangle \rangle$ になっている
 - * 述語とその項群の意味の曖昧性の解消は、あらかじめ決まっている k 個の候補 $\{m_1, m_2, \dots, m_k\}$ の中からの選択
- * 厳密に言うと文意は選択されるが、構築されない
 - * 構築されるのは“文意の候補”

語義の数え上げ?

- * 文意の (非構築的) 選択理論の前提
 - * 述語 p の k 個の語義があらかじめ与えられている
- * だが, 待て
 - * 生成辞書理論 (Pustejovsky 1995) の “語義の数え上げ” (sense enumeration) への批判を思い出せ
 - * 生成辞書 (Generative Lexicon) とは語義を生成する (sense-generative) 辞書

私の見解

- * “正しく” 数えれば語義は数え上げ可能

- * 従来の辞書の難点は方法論的/原理的なものではなく，技術的なもの

- * “正しい” 数え上げの条件

- * 語義の体系性/階層性を表現し，粒度を一定にして数える

- * 従来の理論，辞書の語義の記述の粒度は粗すぎる

何を数えるのか?

- * 数え上げの対象は正確には
 - * 述語の p の (曖昧な) 語義 m_1, \dots, m_n ではなく
 - * 語義で参照されている具体的状況 t_1, \dots, t_n
- * これらは
 - * Berkeley FrameNet (BFN) (Fillmore et al. 2003) の記述の対象で,
 - * オントロジー研究 (溝口 2005) の研究対象の一部

語義は幾つあるのか？

- * 正直なところ，必要な語義の数は膨大
 - * ヒトはコトバを使って自然界と社会の多様性，複雑性をそれなりに記述できるのだから，数が多くて当然
- * ここで再度，問い直す
 - * 語義の数え上げは本当に言語学者の仕事か？
- * 反語的な答え
 - * 言語(科)学者がやらないなら，いったい他の誰がやるの？

THANK YOU
