

複層意味フレーム分析 (MSFA)
による文脈に置かれた語の意味の
多次元的表現

黒田 航 井佐原 均

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

第六回認知言語学会(2005)

09/17/2005

お茶の水女子大学

発表資料

✻ スライドの PDF

- ✻ <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/msfa-jcla05-slides.pdf>

✻ 予稿の PDF

- ✻ <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/msfa-jcla05-handout.pdf>

✻ 関連資料

- ✻ <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/presentations.html>

発表のあらまし

✻ MSFA 開発の背景

- ✻ FOCAL (Frame-Oriented Concept Analysis) との関係
- ✻ 意味役割タグづけの仕事との関係

✻ MSFA の実例紹介

✻ MSFA の設計思想と特徴の解説

- ✻ 言及できない詳細は『認知言語学論考 No.4』収録の論文 (黒田・中本・野澤 2005) を参照願います

MSFA 開発の背景

出発点

- ✿ “語の意味(する概念)は常に何らかの理想認知モデル (ICM) = 世界知識を背景にして理解される” というのは認知言語学の常識
- ✿ 一方、従来の認知言語学はその“常識”に甘んじ、その内実の明示化や妥当性の検証を怠ってきた
 - ✿ 研究プログラムを正当化する前提の検証の意識の不足
 - ✿ 妥当な理論に基づく有用な分析の提供の意識の不足
- ✿ これに対する挑戦が FOCAL (Frame-Oriented Concept Analysis) の枠組みの提唱の動機

MSFA 開発の動機

- ✻ 自然言語文 s をヒト x が聞いたり読んだりする時に x が“理解する内容”を“なるべく自然”に記述する必要がある
- ✻ と同時に
- ✻ 特定の理論に盲従しない“良心的”言語学者が行う自然言語文 s の意味分析 $d(s)$ の蓄積を、そのまま外部者 (認知心理学者や工学者) に利用可能なデータベースにする必要がある

問題

- ✿ 認知言語学の主張が正しいなら
 - ✿ 任意の文の，任意の語の意味について，それがどんな ICM の一部なのかを分析によって示せるはず
 - ✿ その分析がうまく行っているかどうかを評価できるはず
 - ✿ その分析が一部の有能な人にできる特殊技能ではなく，(訓練次第で)誰にでも実践可能であることが示せるはず
- ✿ ICM の理論仮構性を排除するには，これらは不可避

(認知) 言語学への期待

- ✿ (認知)言語学者が本当にすぐれた意味分析をしているなら,
 - ✿ その成果はデータベース化され, 多くの人に利用可能な形で一般公開されるべき
 - ✿ 利用者からのフィードバックを受け, 不断に記述内容が改良されてゆくべき
- ✿ そうでないなら (認知)言語学の“成果”は単なる同業者内での“内輪ウケ”

宣伝

- ✻ MSFA を使った意味役割タグづけの評価版を公開しています
 - ✻ 「京大コーパス」(毎日新聞)の部分(極一部)に試験的なタグづけを行ったデータ
 - ✻ フレームの認定は Berkeley FrameNet とは独立
 - ✻ 利用に毎日新聞コーパスの購入は不要だけれど、情報通信研究機構との利用同意書を交わす必要はある
 - ✻ 連絡は私か、他の FOCAL のメンバーへ

MSFA の実例

実例

1. 大型の台風が九州を襲った.
2. ライオンがインパラの群れを襲った
3. 覆面の男が銀行を襲った.

实例

1. [“大型”, “の”, “台風”, “が”, “九州”, “を”, “襲っ”, “た”, “。”]
2. [“ライオン”, “が”, “インパラ”, “の”, “群れ”, “を”, “襲っ”, “た”, “。”]
3. [“覆面”, “の”, “男”, “が”, “銀行”, “を”, “襲っ”, “た”, “。”]

SAMPLE MSFA OF (1)

Frame ID	F1	F2	F3	F4	G1	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
Frame-to-Frame Relation		elaborates F3,F4	elaborates F4; elaborates G1	constitutes F3		elaborates F9; targets F6	constitutes F7; elaborates F9	presupposes F6; elaborates F8	elaborates F9	constitutes F8	elaborates F11			elaborates F14		realizes F14; presupposes F14
Frame Name	一時的発生	経路移動	移動	経路	非生命体の活動	攻撃 [+metaphoric]	加害	被災	経験	影響	ヒトの生活	生物の棲息	収容 [+metaphoric]	*構成*	*国土*	区別
*	GOVERNOR				活動内容1											
大型の	規模					大きさ?	規模?	規模?	規模?	規模?						
	MARKER															
台風	発生体: EVOKER	移動体: EVOKER	移動体		活動体	攻撃体: EVOKER	加害体: EVOKER	原因	内容	影響体						
*		余波														
が						MARKER	MARKER									
		GOVERNOR			活動内容2											
*	発生地	起点	起点	起点											国外	対象1
*		着点	着点	着点												
*		通過点	経過点	経過点		獲物1	被害体1	被災体1						全体[=日]	国内	対象2
九州						獲物2	被害体2	被災体2	場所		生活地域: EVOKER	棲息地域	収容体	部分: EVOKER		
(の											MARKER		収容物			
人々)						獲物3	被害体3	被災体3	経験者		生活者	生物				
*						獲物4	被害体4	被災体4			住居: EVOKER	住みか				
を						MARKER	MARKER									
襲つ						EVOKER & GOVERNOR	EVOKER & GOVERNOR	様態?								
た						EXTENDER	EXTENDER									
.																

名詞の担う意味役割

- ✿ 結びついている“状況” = “フレーム”ごとに「台風」は例えば次の意味役割を実現する
 - ✿ <(自然)発生>フレームの<(自然)発生体>; <(経路)移動体>フレームの<(経路)移動体>; <攻撃>[+metaphoric]フレームの<攻撃者>[+metaphoric]; <加害フレーム>の<加害体>; <被害>フレームの<被害の原因>
- ✿ これらは動詞 (e.g., 「襲う」) によって“付与”されるものではない
 - ✿ 動詞は単に潜在的な意味役割の集合から幾つかを選択するだけ [cf. Qualia Structure Pustejovsky (1991, 1995)]

意味役割の複合性と相対化

1. 「大型台風が九州を襲い，数千世帯に被害」

- ✿ “... を与えた” と続くか “... が発生(した)” と続くかによって台風の意味格が変わる→未指定 (underspecified) な格??
- ✿ “襲う”に対して“台風”は譬喩的に<動作主>格だが
- ✿ “被害”に対して“台風”は<原因>格
 - ✿ ゼロ照応は無用だと前提する

2. 「大型台風が九州に襲来し，数千世帯に被害」

1. “襲(う)”に対して“台風”は<加害体>で<動作主>格
2. “来(る)”に対しては“台風”は<訪問体>で<動作主>格

フレーム喚起の実例

- ※ 「アルゼンチンの元サッカー選手，ディエゴ・マラドーナ氏が，六日，同国の検察当局に一時身柄を拘束された (京大コーパス S-ID: 950107210-002)
- ※ “元~”は<<引退者 x>が，{<活動: y>から; <役職: z>を}，<理由: w>で，<GOV: 引退>する>フレームを喚起し
- ※ <引退>は<<活動者: x>が，<分野: y>で，<役職: z>として，<GOV: 活動>する>フレームを(前提の関係により)喚起
 - ※ ディエゴ・マラドーナが<引退>と<活動>フレームの x
 - ※ サッカー選手が<引退>フレームの y, z と<活動>フレームの z を実現

フレーム要素の特定

- ✿ 意味役割名の一部には状況 (= フレーム) に固有の意味役割 (= フレーム要素) の名称になっているものが存在する
 - ✿ 被害者, 犠牲者 → <被害>フレームの構成要素
 - ✿ 犯人 → <犯罪>フレームの構成要素
 - ✿ 裁判官, 被告(人), 検察(側), 弁護人 → <裁判>フレームの構成要素
- ✿ これらの名詞のフレーム喚起力は文意構築にとって重要

対象名と役割名の区別

- ✻ 対象名 n は, ある意味フレーム f を構成する意味役割 $f.r$ の**典型値**であることによって f を喚起
 - ✻ 男, 女, 子供, サメ, イルカ, イノシシ, 石
- ✻ 役割名 n^* は, ある意味フレーム f を構成する意味役割 $f.r$ の**名称**であることによって f を喚起
 - ✻ 役職名: 大統領, 首相, 社長, 奴隷, 僕 (しもべ),
 - ✻ 職業名: 運転手, 医者, 妻, 夫
 - ✻ 状況特有の(意味)役割名: 運転者, 犯人, 加害者, 被害者, 犠牲者, 被災者

MSFA の特徴

文意の並列分散性の根拠

- ✿ 文意に対する動詞以外の要素 (特に名詞) の独自の意味貢献
 - ✿ 名詞が同時多発的に, 動詞から独立に状況 = 意味フレームを喚起する
 - ✿ cf. 特質構造 (Qualia Structure)
 - ✿ 語による状況の喚起は “有か無か” の二値 $\{1, 0\}$ でなく $[0, 1]$ の連続値の重み付きの結合 (適当な判別関数)
- ✿ この喚起の効果が “意味論” の問題なのか “語用論” の問題なのかを問うことは無意味

従来の記述モデルとの比較

- ※ 語義の曖昧性の解消の従来のモデル化

- ※ 幾つかの多義の集合 $S(w) = \{s_1(w), s_2(w), \dots, s_n(w)\}$ から最適な意義 $s_i(w)$ をただ一つ選択すること

- ※ 問題点

- ※ $s_i(w), s_j(w)$ のあいだの非排他性
- ※ $s_i(w), s_j(w)$ のあいだの非分離可能性

- ※ MSFA が体現する解決策

- ※ 脱曖昧化は喚起されるフレームごとに起こる多元的な現象だと考え、語の意味記述の多次元化を実現

MSFA から SFNA へ

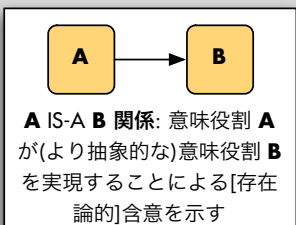
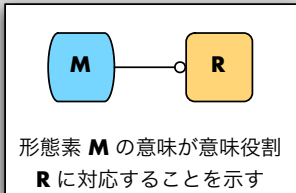
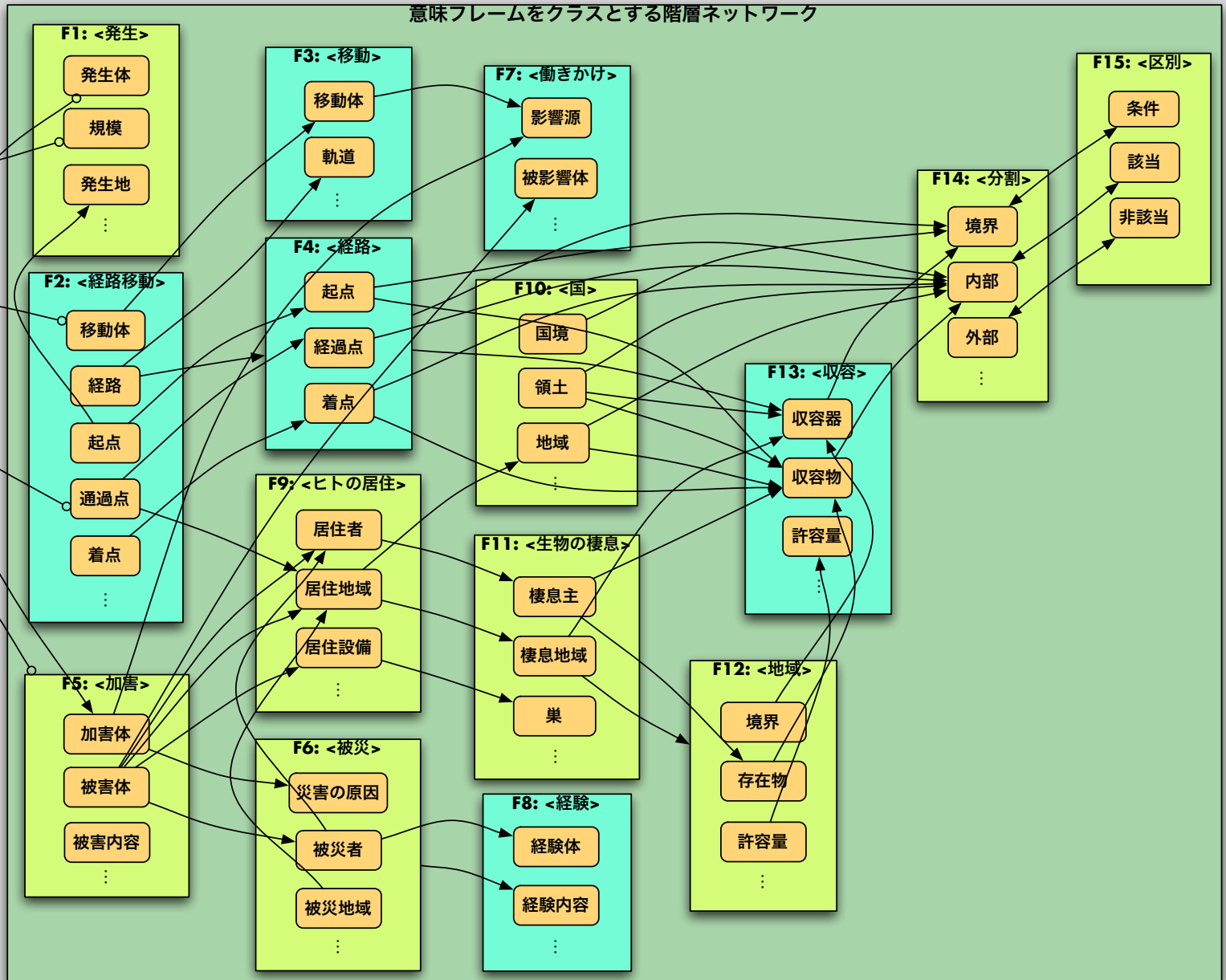
- ✻ MSFA は Semantic Frame Network Analysis (SFNA) への橋渡しの存在
- ✻ SFNA は
 - ✻ 語による状況の喚起のネットワークの形で表わしたものの
 - ✻ 事態のオントロジー階層を表わしたものの
 - ✻ 名詞のオントロジー階層を表わしたものの

SAMPLE SFNA OF (1)

意味フレームをクラスとする階層ネットワーク

自然言語文

大型
の
台風
が
九州
を
襲っ
た



注意

- ✿ MSFA の記述は「細かすぎる」とか「一般性を欠く」とか批判する人への回答
- ✿ 本当に理解が達成されるのは、具体的、個別的なレベル
 - ✿ 具体例から一般例に遡ることは可能だが、その逆はムリ
 - ✿ 情報の廃棄は容易、喪失情報の復元は至難
- ✿ 理解はスキーマ的だが
 - ✿ 理解内容を構成するスキーマは具体的で、特定のであればあるほどよい
 - ✿ いわゆる(イメージ)スキーマ類は理解内容の特定に不十分
 - ✿ 状況とは理解の単位となる下位レベルのスキーマのこと

MSFA の設計思想

理論編

MSFA が強く意識すること

- ✿ 文 $s = w_1 w_2 \dots w_n$ の意味 (文意) の記述と (s の核である) 動詞の意味記述を混同しない
 - ✿ 文意のどの部分が動詞に帰着されるかは、文意を十分に詳しく記述してから“帰納的に”決定されるべきこと
- ✿ 文意は語の意味から合成されるとは限らない
 - ✿ 文意の構成性を記述の前提にするのは、論点先取
 - ✿ 文意は理解可能な“状況”との直接結合で“選択”される
 - ✿ 全体の意味(文の意味)が部分の意味(語の意味)に優先する

構成性を仮定しない利点

❁ 謎

- ❁ 文の意味が語の意味によって構成されるのだとしたら、子供が文の意味をどうして理解できるのか、子供が語の意味がどうやって学べるかは、まったくの謎

❁ 対案

- ❁ 子供が文の意味がわかるのは、彼らが言語を使わないで状況を理解でき、状況と発話の内容とを対応づける能力が(おそらく生得的に)備わっているから
- ❁ 子供が語の意味が理解し学べるのは、語の意味がわからなくても“文”レベルの意味がわかるから

語の意味はどうしてわかるか

- ✧ 語の意味 (= 語義) は (使用) 文脈ごとにおのこの異なる
 - ✧ 語義の異なりは生起環境の異なりに等しい
 - ✧ Latent Semantic Analysis (LSA: Landauer, *et al.* 1998)
 - ✧ 同じ語義でも, 文脈的変異によって意味は異なる
 - ✧ 同義性とは類義性に基づくクラスター
 - ✧ クラスターの離散性/分離可能性を保証する何かが存在
- ✧ これが意味すること
 - ✧ 語の意味記述は“文脈”の意味記述から切り離せない

文脈の相対的定義

- ✿ 十分に詳細な語の意味記述に必要なことは
 - ✿ 語 w の意味の記述と w に対し相対的に“文脈”を形成している語群 (= 生起環境) の意味記述を行い
 - ✿ 文脈を相対的に定義すること
 - ✿ 語の意味の明示的記述と文脈の明示的記述は分離不能
- ✿ これは“文脈とは何か?”に対する答えの一つ
 - ✿ 文脈は“所与のもの”ではない
 - ✿ 文脈はそれ自体, 複数の語(の意味)によって形成される

相対的定義の帰結

- ✿ 語は共起環境の中で相互に影響しあい，それらの意味は相互に“適応”しあう
- ✿ cf. Langacker の意味調節 (semantic accommodation)
 - ✿ ただし，適応は双方向的で (Langacker の図が示唆しているように) 一方的ではない
- ✿ cf. Pustejovsky の共合成 (co-composition)
 - ✿ 語の意味の文脈的“変異”は，生起環境への適応の産物
- ✿ “主要部/項”の非対称性を弱める効果アリ
 - ✿ 名詞は単なる述語類の項として働くだけではない

語の意味は相互適応する

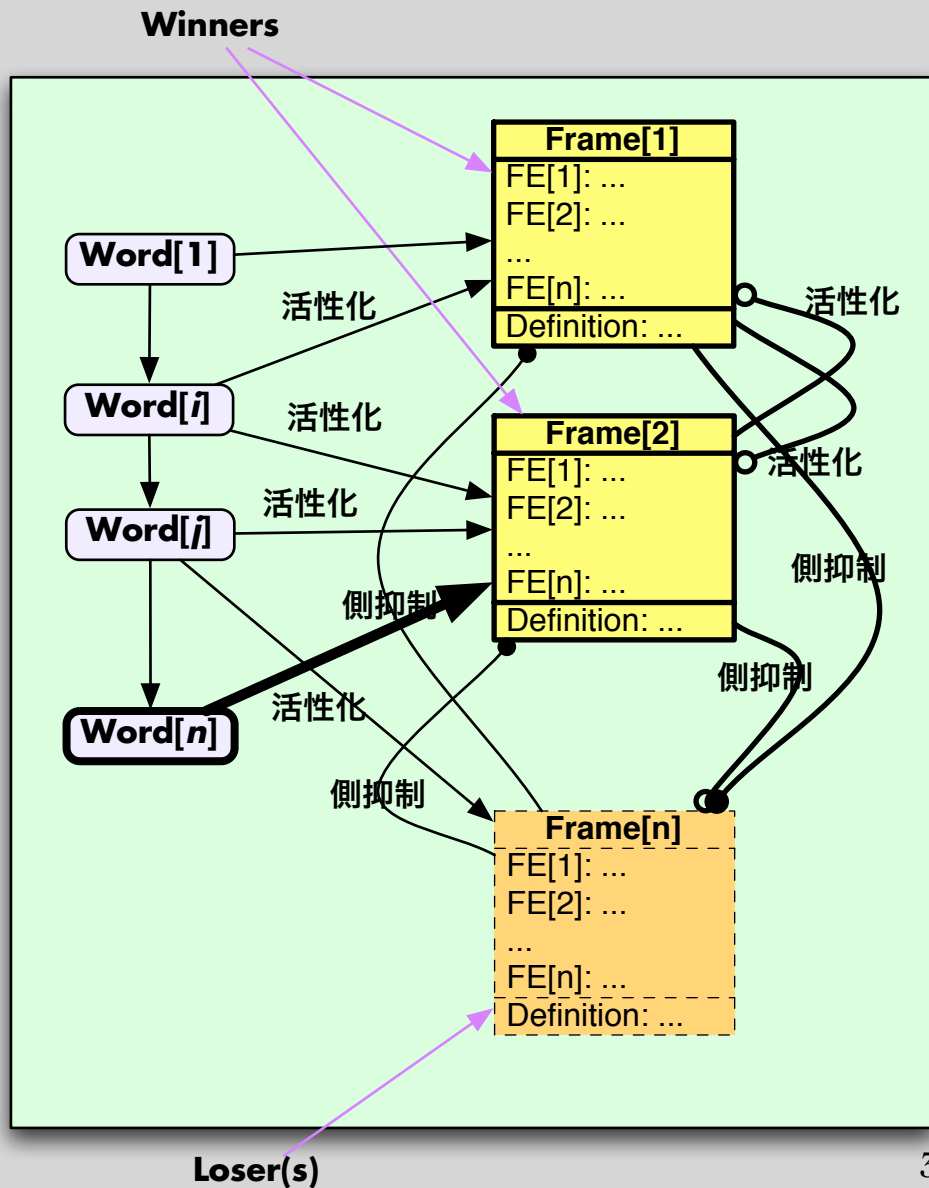
✿ 事実

- ✿ 語の意味は文(脈)という環境内で相互に調節される
- ✿ この事実を“意味の生態系”というアナロジーを使ってモデル化すると
 - ✿ 語は“文(脈)”という(複雑)生態系を構成する個体
 - ✿ それらは相互に“適応”する
 - ✿ 基本単位は“意味” = “用法”であって“語”ではない
- ✿ これが“語の意味の相互適応モデル”の意味

相互適応モデルの難点

- ✿ 語の意味の相互適応モデルでは
 - ✿ 語の意味の特定/曖昧性の解消の問題は多体問題 (n-body problem) で、トリックなしには解けない
- ✿ 提案するトリック
 - ✿ 生起環境を共有する語の意味の相互適応を並列分散処理 (Parallel Distributed Processing: PDP) としてモデル化

フレーム単位の語義の選択



- ✪ 語 w の意味の選択の単位はフレーム = 状況
- ✪ フレームの選択は自然選択に似たプロセス
- ✪ 選択の単位がフレームだということは“全体の意味が部分の意味に優先する”ということ
- ✪ 強い意味での構成性は前提にできない

まとめ

MSFA の特徴 1/2

- ✻ 語の意味の“文脈への適応” = 文脈効果をうまく記述する
 - ✻ 意味調節 (Langacker 1987, 1991) の明示化
 - ✻ 共合成 (Pustejovsky 1995) と記述的に互換
 - ✻ ただし MSFA は文脈効果の“説明”はしていない. その説明は MSFA と別に“フレームへの引きこみ効果”によって与えられる
- ✻ (並列)分散意味論 (Parallel) Distributed Semantics: PDS の実装

MSFA の特徴 2/2

- ✻ 意味論と語用論を特に区別しない
- ✻ 世界知識と言語的知識を特に区別しない
- ✻ “語(彙)の知識” (knowledge of words) と “言語の知識” (knowledge of language) = “文法” を特に区別しない
- ✻ MSFA は “これらの区別が存在しない” と教条的な主張をなすわけではないが、それは “これらの区別は言語学者が恣意的に導入するもので、言語分析の大前提とはならない” ことを示唆する

謝辞

以下の方々にお世話になりました

内山 将夫 (NICT)

金丸 敏幸 (京都大学 大学院)

李 在鎬 (NICT)

中本 敬子 (京都大学)

野澤 元 (NICT)