

上代日本語のコーパスにおける 意味役割の付与

国立国語研究所、405号室

2011年1月20日、10:30-12:30

スティーブン・ライト・ホーン
PD研究員
stephen.horn@orinst.ox.ac.uk

VSARPJ: People

Principal Investigator

Bjarke Frellesvig

Co-Investigator

Peter Sells (SOAS)

Postdoctoral Researchers

Stephen Wright Horn (AHRC)

Kerri L Russell (AHRC)

Doctoral students

Zixi You

External Members

VSARPJ: 日本語史に於ける動詞の意味と項の実現

目標：

言語学の分析が出来る様なコーパスの開発
形態素と構文を表すタグ付与
項の実現を通時的に捉える

<http://vsarpj.orinst.ox.ac.uk/nihongo.html>

項の実現

多くの場合、一つの述語には複数の項のパターンがありうる。

(1) 壁にペンキを塗る。

‘smear paint on the wall’

(2) ペンキで壁を塗る。

‘smear the wall with paint’

これは場所格交替動詞の用例。

閉鎖コーパスの問題点

(1) と (2) は結合価の上での minimal pair で、語彙素を固定要因とするから、場所格交替が浮き彫りになる。同じパターンに参加する別の動詞の用例を見たら「移転物」と「場所」の意味役割を抽象することができる。

しかし、上代語をデータとする閉鎖コーパスではこの方法を使えない。

項の実現を記述するために、そしてコーパスの操作だけで項の交替が割り出せるように、どんなアノテーションが必要なのか。

コーパスの設計

- Text Encoding Initiative (TEI) 規定に基づいた xml マークアップ。
- ローマ字で音韻を表記
- 底本の表記法を区別：音・訓仮名、訓読み、無標記、など
- 品詞、活用形
- 形態素の種類と機能
- 語彙素や形態素にID番号を振る
- 構成部の階層的関係を表す統語的構造

検索可能： 主題、右方転移構文、取り立て、構成部の主要部、焦点、助詞のスコープ、関係節の係、節の種類、複合動詞、名詞句の格標示・無表示とその順位、など

依拠するテキスト

古事記歌謡	712	
日本書紀歌謡	720	
風土記歌謡	730	
仏足石歌	753	以降
万葉集	759	以降
続日本紀歌謡	797	

統計

poems: 4976

words: 89,356

morphemes: 15,795

verbs: 22,855

verbs with ID: 16,867

verbs without ID: 5,988

adjectives: 3,601

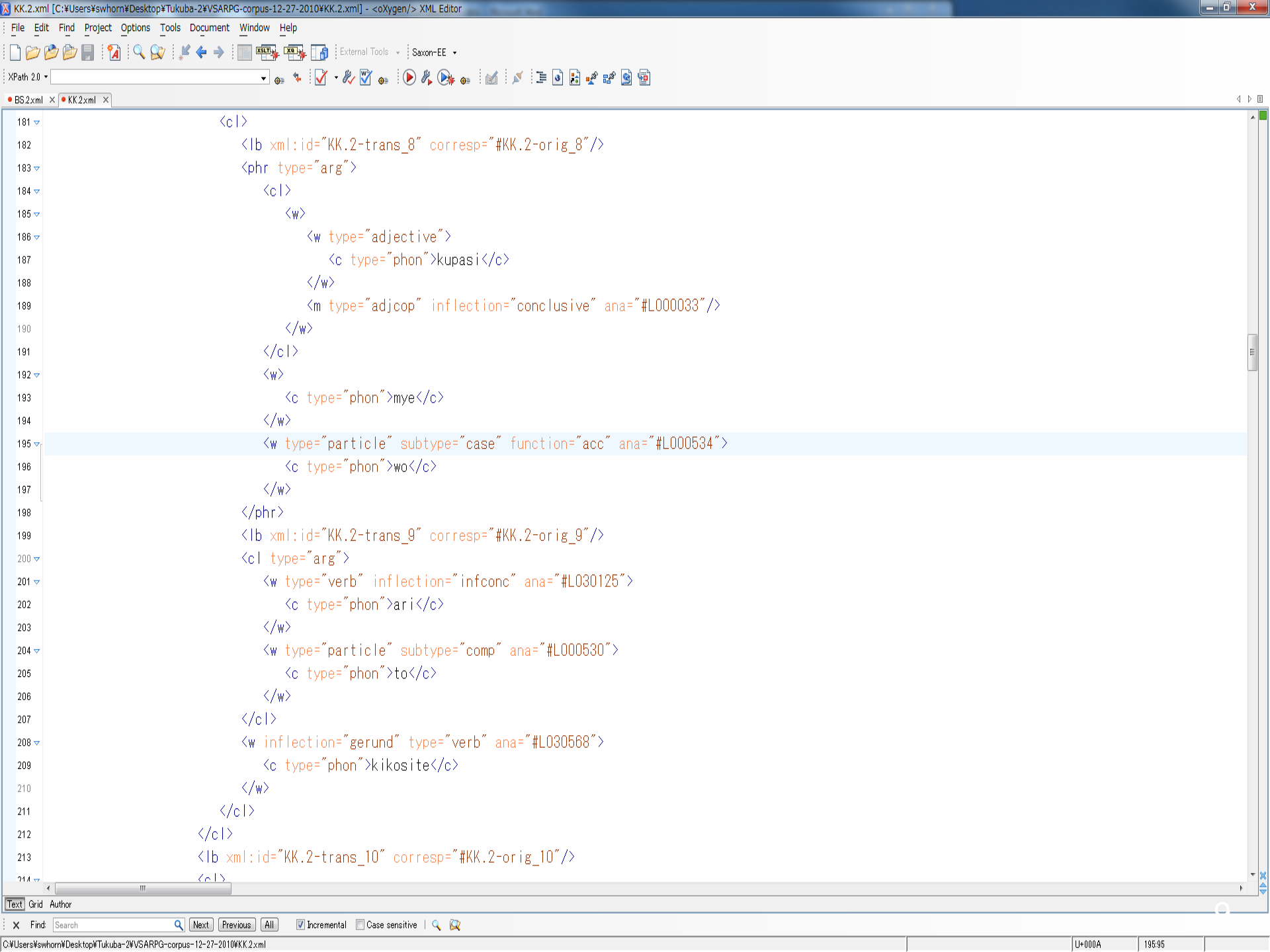
copulas: 2,942

adjcops: 2,922

phrases: 26,927

argument phrases: 11,265

argument phrases containing no case particles at all: 6012



予備実験

<http://verbs.colorado.edu/~mpalmer/projects/verbnet.html> を参考

に意味役割リストの作成 : *Agent, Patient, Theme, Causee, Source, Path, Goal, Location, Experiencer, Beneficiary, Maleficiary, Stimulus, Instrument, Manner, Name*

七つの動詞に適用 : 聞く、漕ぐ、持つ、おもほゆ、取る、渡る、遣る

付与された意味役割の延べ数 : 1045

作業の方針

- **述語の種類**：動詞のみ、複合動詞を含めて
- **意味役割付与の対象**：名詞句のみ
- **統語的範囲**：主要部動詞の節に内在する名詞句、または節に隣接した主題名詞句のみ
- **コーディングの方法**：名詞句のラベル付け
(e. g., `<phr role= “agent” />`)
- **コーパスとの関係**：コープスデータに直接付け加える

意味役割のリスト

限られた数の意味役割をもって、どんな意味的フレームでも網羅して記述できることを目指す。

そのために一般化した役割の間に相互排除の関係が成り立つように定義しなければならない。

コーディングのばらつきを抑えながら、できるだけ語彙の多義性を反映させる。

動作主 (Agent)

動作を意図的に行って、動因主となるもの。何らかの変化・影響・作用をもたらすことが多い。有生性のあるもの、或いは内在制御機能を持つ機械。使役文の動因主。

此夕 零来雨者 男星之
早滂船之 賀伊乃散鴨

このゆふべ 降りくる雨は ひこほしのAgent
はや漕ぐ船の かいのちりかも (MYS. 10. 2052)

被動者 (Patient)

動作の影響を受けて変化する物。他動性が強い動詞の目的語。非対格動詞や能動態動詞の主語。

夜光 玉跡言十方 酒飲而
情乎遣尔 豈若目八方

よるひかる たまといふとも さけのみて
こころをPatient やるに あにしかめやも (MYS. 3. 346)

対象 (Theme)

受動者で、作用を受けながら影響を受けない物。移動するもの。場所に位置をしめるもの。状態や属性の帰属先。経験者の判断の対象。所有物。などなど。

箆毛與 美箆母乳 布久思毛與
美夫君志持 此岳尔 菜採須兒

こもよ みこ Theme もち ふくしもよ
みぶくし Theme もち このをかになつますこ
(MYS. 1. 1)

被使役者 (Causee)

使役形の述語のなかで動作・作用させられるもの

大夫登 念有吾乎 如是許
令戀波 小可者在來

ますらをと おもへるわれを Causee かくばかり
恋せしむるは あしくはありけり (MYS. 11. 2584)

起点 (Source)

移動の出発点、物交換の譲渡者

瀧上乃 三船山従 秋津邊
夾鳴度者 誰喚兒鳥

たぎのうへの みふねのやまゆ Source あきづへに
きなきわたるは たれ呼ぶこどり (MYS. 9. 1713)

経路 (Path)

移動・移転する間の状況・経路・経由点

可豆思加乃 麻万能宇良未乎 許具布祢能
布奈妣等佐和久 奈美多都良思母

かづしかの ままのうらを^{Path} こぐふねの
ふなびとさわく 波たつらしも (MYS. 14. 3349)

着点 (Goal)

移動の到着点、物交換の受領者

奥去哉 赤羅小船尔 裏遣者
若人見而 解披見鴨

おきゆくや あからをぶねにGoal つとやらば
けだしひとみて ひらきみむかも (MYS. 16. 3868)

場所 (Location)

時間・空間的場所。個体の存在する場所・時間。活動が行われる場所・時間

虎尔乘 古屋乎越而 青淵尔

鮫龍取將來 劔刀毛我

虎に乗り ふるやを超えて あをふちに Location

みつちとりこむ つるぎたちもが (MYS. 16. 3833)

経験者 (Exper i encer)

刺激を受けて何らかの心理状態を経験する者。五感を使って何かを認知する者。或いは思考を発揮し判断する者。

今造 斑衣服 面影

吾尔所念 末服友

いまつくる まだらのころも おもかげに

われにExper i encer おもほゆ いまだ着ねども

(MYS. 7. 1296)

受益者 (Beneficiary)

該当者なし？理由・動機づけを表す句はある：

妹為 菅實探 行吾

いもがため すがのみとりに ゆきしわれは

(MYS. 7. 1250)

結合価を重視すると「いもがため」が状態 (manner) の付加詞に当たる。事象構造を重視すると「いも」が受益者に当たる。

被害者 (Maleficiary)

該当者なし？それとも迷惑受け身の主語？

白髪爲 子等母生名者 如是
將若異子等丹 所詈金目八

しろかみし ころも Maleficiary おひなば かくのごと
わかかけむころに のらえかねめや (MYS. 16. 3793)

刺激 (Stimulus)

経験者に心理的な影響を及ぼす。認知の対象。考慮

- ・判断の場合は、刺激ではなくて「対象」(theme)とする。

山近 家哉可居 左小牡鹿乃

音乎聞乍 宿不勝鴨

やまちかく いへやをるべき さを鹿の

こゑを Stimulus 聞きつつ いねかてぬかも

(MYS. 10. 2146)

手段・道具 (Instrument)

主語と別に動作主の存在が示唆される場合、主語も道具の役割を持ちうる（「酸が銅版に像を刻んだ」というのはやや不自然な日本語だが）。

山科 強田 馬雖在
歩吾來 汝念不得

やましなの こはたのやまを 馬はあれど
かちゆ Instrument わがこし なをおもひかね
(MYS. 11. 2425)

状態 (Manner)

動作・作用がどんな状態で、あるいはどんな程度で行われるかを表す

久佐麻久良 多妣爾之婆之婆 可久能未也
伎美乎夜利都追 安我孤悲乎良牟

くさまくら 旅にしばしば^{Manner} かくのみや^{Manner}
君を遣りつつ あがこひをらむ (MYS. 17. 3936)

名前 (Name)

コーパスでは名づける行為を表す動詞の補文を名詞句としてマークアップしている。

梅花 先開枝乎 手折而者
裏常名付而 与副手六香聞

うめのはな まづさくえだを たをりてば
つとと Name なづけて よそへてむかも (MYS. 10. 2326)

適用に当たっての問題点

項目の曖昧性、相互排除のむずかしさ：自動詞は **Agent** をとるか。
「惑ふ」の主語は **Patient** なのか、**Experiencer** なのか。

一つの述語の用例に当たり二つ以上のNPに同じ役割を付与する可能性
が残る：**Location**, **Manner**, など

意味役割と **Phrase type** をミスマッチする恐れ：「**手**に取り持ちて」の「手に」は統語上の結合価を重視すると **Goal** だが、事象構造を重視すると **Instrument** と分析しがちだ。「もちて」が主要部動詞の従属節全体が **Instrument** を示す場合が多いが、その中の項がその役割を果たしているとは間違いしやすい。**Beneficiary** の用例と同じ様に。

他動詞としての「漕ぐ」

動作主と対象が両方現れる用例：

海末通女 棚無小舟 榜出良之
客乃屋取尔 梶音所聞

あまをとめ^{Agent} たななしをぶね^{Theme} こぎづらし
たびのやどりに かぢのおと きこゆ (MYS. 6. 930)

自動詞としての「漕ぐ」

「ふね」が主語（Theme? Instrument?）の用例：

可豆思加乃 麻万能宇良未乎 許具布祢能
布奈妣等佐和久 奈美多都良思母

かづしかの ままのうらみを こぐふねの Theme
ふなびとさわく なみたつらしも (MYS. 14. 3349)

「ふなひと」は「こぐ」の主語じゃない：CNPC 違反

自・他の区別付かず

大船 真楫繁拔 榜間
極太戀 年在如何

おほぶねに まかぢしじぬき こぐほとも
ここだくこひし としにあらばいかに
(MYS. 11. 2494)

「おほぶね」は「こぐ」の Theme なのか。必ずしも Binding Condition C を違反してはいないが、上位節の null argument を束縛できない。

「漕ぐ」は使役起動交替するか？

たとえそうだととしても、仮に使っている意味役割ではその
区別できない。

天漢 八十瀬霧合 男星之
時待船 今滂良之

あまのかは やそせきらへり ひこほしの
ときまつふねは^{Theme} いましこぐらし (MYS. 10. 2053)

「は」が付くと主・対格標示が吸収される。そもそも自動
詞の主語も他動詞の目的語も格が無標示な場合が多い。

「漕ぐ」で見えてきたように、意味役割の交替や格の交替だけでは多動性の違いが割り出せない場合がある。

しかし、意味役割の後退で語彙の多義性がわかる場合がある：

動詞の多義性：「とる」

文法役割を固定要因にしたら、意味役割の交替で動詞に多義性が見いだせる。

波斯多₁能 久良波斯夜麻袁 佐賀志美登
伊波迦伎加泥₂ 和賀₃登良須母

はしたての くらはしやまを さがしみと
いはかきかねて わが手_{Theme} 取らすも (KK. 69)

動詞の多義性：「とる」

本都延波 登理韋賀良斯
志豆延波 比登登理賀良斯

ほつえは 鳥みがらし
しづえは Patient 人採りがらし (KK. 43)

比婆理波 阿米迹迦気流 多迦由玖夜
波夜夫佐和気 佐邪岐登良佐泥

ひばりは あめにかける たかゆくや
はやぶさわけ さざき Patient 獲らさね (KK. 68)

格の交替で動詞の多動性交替がわかる場合もある：

英語にない交替：「覆ふ」

The tarpaulin covered the ground.

C. f. *Bill covered a tarpaulin (over the ground). (Dowty 1993)

天下 須泥爾於保比低 布流雪乃
比可里乎見礼婆 多敷刀久母安流香

あめのした すでにおほひて ふるゆきの Transferred item
ひかりをみれば たふときもあるか (MYS. 17. 3923)

英語にない交替：「覆ふ」

不奉仕 立向之毛 露霜之 消者消倍久 去鳥乃
相競端尔 渡會乃 齊宮従 神風尔 伊吹或之
天雲乎 日之目毛不令見 常闇尔 覆賜而

まつろはず たちむかひしも つゆしもの けなば
けぬべく ゆくとりの あらそふはしに わたら
ひの いつきのみやゆ かむかぜに いふきまと
はし あまくもを *Transferred item* ひのめもみせず
とこやみに おほひたまひて (MYS. 2. 199a)

しかし、仮の意味役割リストでは、(MYS. 17. 3923)の「覆ふ」が「覆われたもの」と「移されたもの」という二つの **Theme** を取る。Bi-uniqueness に適っていない。はっきりさせるために **Transferred item** という新しい項目を設けなければならなかった。

使役起動交替する「置く」

夕凝 霜置來 朝戸出尔
甚踐而 人尔所知名

ゆふこりの しも^{Theme} おきにけり あさとでに
はなはだふみて ひとにしらゆな (MYS. 11. 2692)

多麻能宇良能 於伎都之良多麻 比利倣礼杼
麻多曾於伎都流 見流比等乎奈美

たまのうらの おきつしらたま^{Theme} ひりへれど
またそおきつる みるひとをなみ (MYS. 15. 3628)

「置く」の場合、用例に格の交替や項の数の変化が認められない限り、多動性の交替が割り出せないが、閉鎖コーパスでは用例の少ない動詞が多い。

Lexical Conceptual Structure や意味フレームなど
では割り出せるものの、そういった上位概念を取
り入れない限り、コーパス操作だけでは割り出せ
ない現象がある：

使役起動対の一本化：「開く」

久方乃 月夜乎清美 梅花
心開而 吾念有公

ひさかたの つくよをきよみ うめのはな
こころひらけて わがおもへるきみ (MYS. 8. 1661)

於久夜麻能 真木乃伊良度乎 等杼登之弓
和我比良可武爾 伊利伎弓奈左祢

おくやまの まきのいらとを とどとして
わがひらかむに いりきてなさね (MYS. 14. 3467)

意味役割と格だけでは説明ができない交替の中に以下のものもある：

格交替が何と共変動するか

「飽く」の Theme : に、を、が、の

「忘る」の Theme : を、が、の

「のむ」の Goal : を、に

「知らなく」の Theme : を、の

使役構文の Causee : の、に

格の交替：「忘る」

「を」格の Theme :

靡相之 宣君之 朝宮乎
忘賜哉

なびかひし よろしき君が あさみやを Theme
忘れたまふや (MYS. 2. 196a)

尊敬の副動詞

格の交替：「忘る」

「が」格の Theme :

従蘆邊 満来塩乃 弥益荷
念歟君之 忘金鶴

あしへより 満ち来るしほの いや増しに
おもへか君が Theme 忘れかねつる (MYS. 4. 617)

可能性を問題にする副動詞

格の交替：「忘る」

「の」格の Theme :

面形之 忘戸在者 小豆鳴
男士物屋 戀乍將居

おもかたの Theme 忘れてあらば あづきなく
をのこじものや こひつつをらむ (MYS. 11. 2580)

状态的 analytic predicate (Agentを背景化?)

文法役割付与が必要なのか？

類型論で使われる文法役割の分類化：

Agent: 他動詞の主語

Subject: 自動詞の主語

Object: 他動詞の目的語

非対格動詞の同定

上代語では **Subject** も **Object** も、位置を変えない限り、格が無標示のままが普通だから、格をあてにできない。

文法役割と意味役割とを組み合わせれば、多動性が分かる場合がある。

例えば、同じ名詞句に **Subject** と **Patient** が重なると、動詞が非対格動詞あるいは能動態動詞ということが分かる。

(**Patient** でない **Subject** をとる非対格動詞もあるが)

「惑ふ」

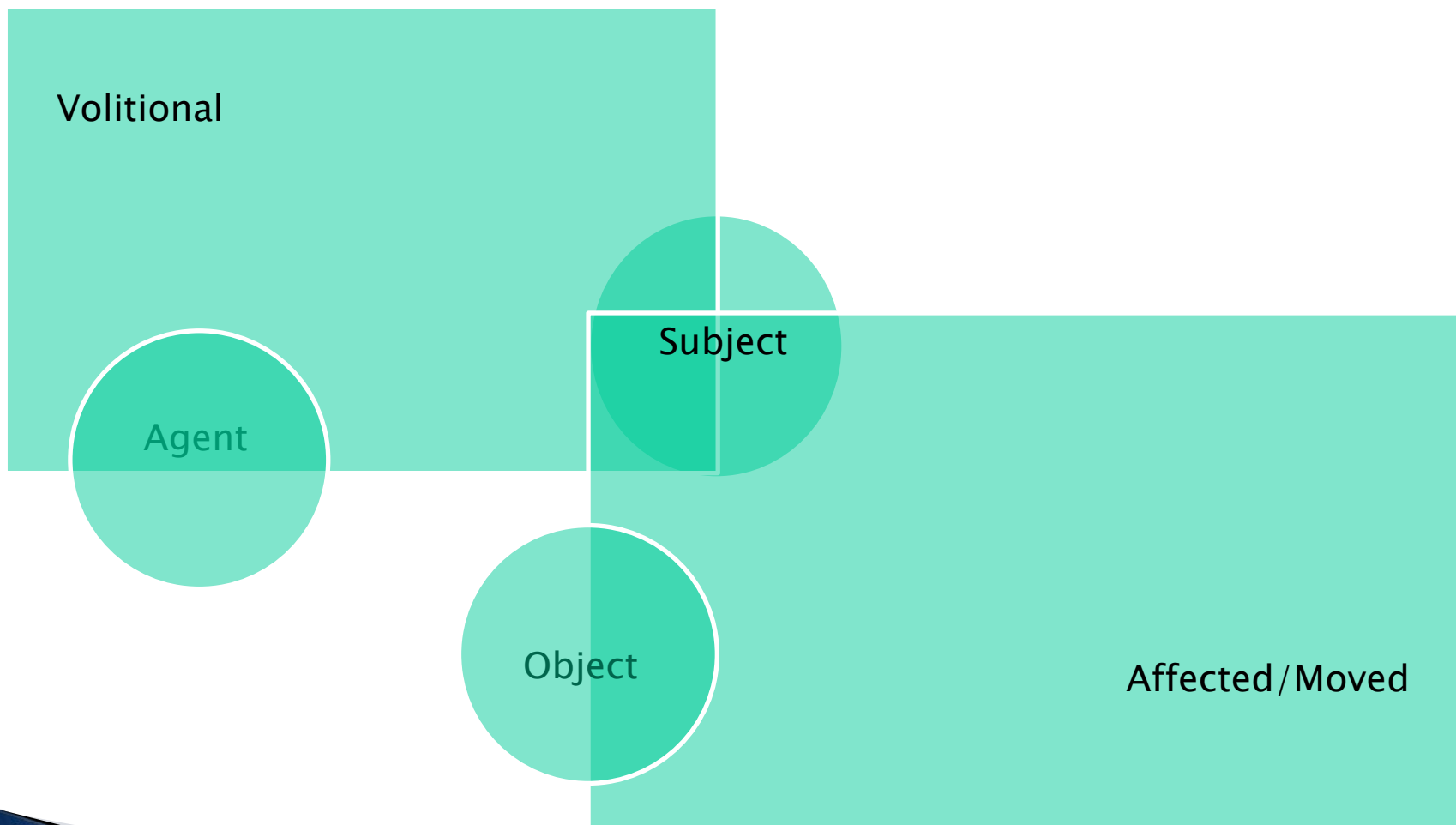
虚蝉之 常辞登 雖念
繼而之聞者 心遮焉

うつせみの つねのことばと おもへども
つぎてしきけば **こころ** Patient/Subject まとふ
(MYS. 12. 2961)

非対格性を確認するには統語的な診断の方が望ましいが、閉鎖コーパスではテストと成る構文が必ずしも表れない。

仮のリストを文法役割と組み合わせてもそれだけでは非対格動詞のすべてを割り出すことができない。**Subject** が影響を受けるか移動するかのどちらかというのがその動作・作用の意味特徴の条件。とすると、移動さえすれば、**Theme** や **Agent** も非対格動詞の **Subject** として現れる。

文法役割と動作・作用の意味特徴： 非対格性の場合



文法役割を取り入れるのなら、意味役割を再考察しても良いかも。

動作・作用に様々な意味特徴

項の実現に関わる要因には

- Volitional/Non-volitional
- Incremental/Non-incremental
- Affected/Unaffected
- Motion/Position
- Telic/Atelic
- Causative/Inchoative
- Figure/Ground
- Eventive/Property-describing, etc.
- De dicto/de re

一つの項に二つ以上の意味特徴がある場合が多い。
仮に使っている意味役割だけでは表せきれない。
。

Proto-role entailments?

個々の項に当てはまる entailment をリストアップする？

Agent Proto-role:

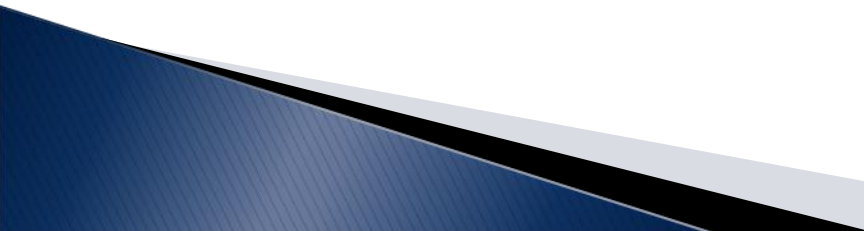
- a. volitional involvement in the event or state
- b. sentience (and/or perception)
- c. causing an event or change of state in another participant
- d. movement (relative to the position of another participant)
- (e. exists independently of the event named by the verb)

Patient Proto-role:

- a. undergoes change of state
- b. incremental theme
- c. causally affected by another participant
- d. stationary relative to movement of another participant
- (e. does not exist independently of the event, or not at all)

(Dowty, 1991)

名詞句の意味特徴をコーパスに？

- Animate/inanimate
 - Sentient/non-sentient
 - Mass/count
 - Generic/definite/non-specific
 - Abstract/concrete
 - 1st person/2nd person/3rd person
 - Singular/plural
- 

コーディングの方法

名詞句のラベル付けでは複数の役割を表せない。

天海 月船浮 桂梶
懸而滂所見 月人壮子

あめのうみに つきのふねうけ かつらかぢ

かけて こぐみゆ つきびとをとこ Agent, Subject



(MYS. 10. 2223)

複数の役割：提案 その1

先行者がある場合、コーパスに null argument を書き込むか。

春霞 流共尔 青柳之
枝喙持而 鶯鳴毛

はるかすみ ながるるなへに e_i Agent あをやぎの

えだ Object くひもちて うぐひす i Subject なくも

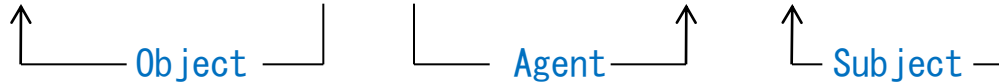
(MYS. 10. 1821)

複数の役割：提案 その2

矢印にラベルを付けるか：

はるかすみ ながるるなへに あをやぎの

えだ くひもちて うぐひす なくも



(MYS. 10. 1821)

Discourse antecedents をコーパスに？

コンテキストに先行者を求める null argument をどう扱う？ (c. f. FrameNet: Null Instantiation)

能動態の Agent, 命令文の Subject, 第一・二人称の項、などは、多くの場合は無標記。

事象構造を重視して、コーパスに(何らかの方法で)盛り込むか。データを客観的に捉えることにとどめるか。

フレームを個別に？

意味役割・文法役割を core data に書き込まないで、フレームのレキシコンを個別に作る？そして、フレームのレキシコンに基づいて動詞の分類化を行った後、個々の動詞の verb class の所属を語彙素の見出しに書き込む？

利点：同じ構文にある他の動詞を無視してもよくなるから、名詞句にラベルを付けるだけでいい。動詞群ごとにその特有な意味役割を設けることができる。それで動詞群間の一貫性がそれほど必要でなくなる。先入観を core data に持ち込まない。

欠点： core data に文法役割の付与があるといろいろ役に立つ。例えば、矢印が non-local NP を指しているると null argument の存在が分かる。せっかくの分析結果を他のところで生かしたい。

Mapping を制限するか？

Bi-uniqueness: 個々の動詞の実例に当たり、各々の意味役割を一つ以上の名詞句に付与してはいけない。

利点：意味役割を Argument-indexing に使える。

欠点：項目を増やさないと Discontinuous argument が必要になる。

Phrase type を増やすか？

- 結果状態を表す節
- 命題的内容を表す補文
- 原因、理由、方法、継起などを表す従属節

利点：より精密な記述ができる。

欠点：項目を増やさないとできない。

意味役割リストをどう変える？

リストを絞る

利点：一貫性が高まる。

欠点：項の実現に関わる意味要因を見落とす恐れがある。

リストを広げる

利点：事象構造をより忠実に記述。予測していない要因を見出す。

欠点：統合との関連付け (Linking, argument indexing) が難しくなる。項目を増やせば増やすほどその間の総合排除が難しくなる。コーディングの一貫性が失われる。

意味役割を同定する基準を どう変える？

コーディングの一貫性だけが問題ではない。

客観的で操作的な基準を設けないかぎり、試験再現性がない。

既成の分析枠組みを無理して当てはめようとする
と誤謬につながる。

項の実現を通時的に記述するには

個々の用例の広義での結合価パターン

- 形態・統語の情報（内的構造・外的位置）
- 項の文法役割
- 項の意味役割

その集まりから抽象した幾つかの基本的フレーム

同じパターンに参加する動詞を同定

他時代との比較（個別の動詞でも、動詞群でも）

参考文献

- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language*, 67:547-619.
- Kipper-Schuler, Karin. VerbNet: A broad-coverage, comprehensive verb lexicon. Ph.D. Dissertation. U. Penn., 2005.
- Kipper-Schuler, Karin. VerbNet v3.1.
<http://verbs.colorado.edu/~mpalmer/projects/verbnet/downloads.html>
- Kipper, Karin; Anna Korhonen; Neville Ryant; Martha Palmer. A Large-scale Classification of English Verbs, *Language Resources and Evaluation Journal*. 42(1), pp. 21-40, Springer Netherland, 2008.
- Kipper-Schuler, Karin; Anna Korhonen; Susan W. Brown. VerbNet overview, extensions, mappings and apps, Tutorial. NAACL-HLT 2009, Boulder, Colorado.
- Levin, Beth and Malka Rappaport Hovav. *Argument realization*. Cambridge, UK: Cambridge University Press, 2005.
- Ruppenhofer, Josef; Michael Ellsworth; Miriam R. L. Petruck; Christopher R. Johnson; Jan Scheffczyk. *FrameNet II: Extended theory and practice*. September 14, 2010.
http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=126