

概念体系に基づく格関係

奥村 登貴子

井原 浩子*

(株)日本電子化辞書研究所

*東京造形大学

概要

headに意味が依存する係受け(必須格)を少数の類型(headの分類である概念分類項目と、その項目ごとに定義した必須深層格のセット)に帰着させることにより、従来明確な指針のなかった必須格の定義にかえる。これにより例文解析などのデータの安定化がはかれる。また、概念分類項目がheadを類語と見る際の観点を提供するので、類語間の格解釈に整合性がとりやすくなる。この結果、機械翻訳をはじめとする「文の言い替え」に際してデータ変換の労力が軽減する。また類型が少ないことは辞書の保守、開発の見通しをよくする。

状態性概念の分類約30項目を定め、各項目に対して6個以下の必須格を定義した。基本的な関係子はこの必須格と始状態・終状態・変化分の組合せとする。各語義は概念分類項目の単なる下位ではなく、合成となる。この概念分類項目と関係子が文章の個々の係受けをよく網羅することを検証した。

A CASE SYSTEM BASED ON A THESAURUS

Tokiko Okumura

Hiroko Ihara*

Japan Electronic Dictionary Research Institute, LTD. (EDR)

Mita-kokusai-building 4-28, Mita 1-chome, Minato-ku Tokyo 108, Japan

*Tokyo University of Art and Design

2707 3-chome, Motohatsiouji-cho Hachiouji-shi Tokyo 193

Abstract

In the usual case system, the interpretation of obligatory cases is not clear. This has been an obstacle to the data analysis, in stabilizing the interpretation of a case for each head-word, and also, between the head-words that are related(e.g. similar, combined).

We present a small thesaurus, with a set of cases defined on each item. The thesaurus is a combination of two axis: One, with 30 items, defines the obligatory cases for stative concepts, and the other, with 10 items, defines which of those obligatory cases changes with the lapse of time. In other words, our cases are the combinations of obligatory cases for stative concepts, and source/goal type cases. Fairly large amount of corpus has been mapped, directly or indirectly (as a composition) to this thesaurus-case system.

*本研究は筆者が日本電子化辞書在職中に行われた。

This work has been done while the author worked for EDR.

1. はじめに

必須格の類型の必要性

従来、関係子（格）の定義は概念全体にわたるものだった。つまり、headに意味が依存する関係子（必須格）に関しては明確な定義がなく、機械翻訳システムごとにバラバラに、対訳レベルの膨大な数の類型（＝headとその必須格フレームの類型）を保持しているのが現状である。

この結果、例文解析などから新たに辞書を作ろうとする場合に必須格の関係子解釈のユレが基だしく、安定したデータがとれないという問題がある。また類語間の格解釈の整合性もとりにくい（下図）。

	例	解釈例	付与例
異なる解釈の縮退	髪を刈る 頭を刈る	髪を離す 頭を美容	obj obj
表層格のイメージの差による類語間でのユレ	be angry at me 私への怒り		obj goal
	意見に同意 意見を支持		goal obj
合成概念	reconcile SBJ reconcile OBJ	させる 和解	obj agt

これらの問題は、headにあたる語義を少数の分類項目（概念分類項目）に帰着させ、個々の係受けをそれら概念分類項目ごとに定義された必須格に帰着させることができれば解決する。

なぜなら、headの解釈を概念分類項目として決めると必須格の類型が提供されるので表層格のイメージの差が捨象しやすい。

例：「意見に同意」も「意見を支持」も、

概念分類項目「情報の離合」の関係子「情報内容」

またheadの解釈が概念分類項目として保存できるので、従来のように異なる解釈が同じ格に縮退して見えてしまうことはなくなる。

例：頭を刈る＝概念分類項目「評価」の格「評価物」

髪を刈る＝概念分類項目「付着」の格「付着物」

このような概念分類項目はheadを類語と見る際の観点を提供するので、類語間の格解釈の整合性がとりやすくなる。この結果、機械翻訳をはじめとする「文の言い替え」に際してデータ変換の労力が軽減する。また類型が少ないことは辞書の保守、開発の見通しをよくする。

この種の類型化に関してはSchankらのCD¹に始まる流れがあるが、これらは語彙や文を網羅する類型を提示したわけではない。また語彙に関して網羅性のある類型は類語辞典²などの形で存在するが、多言語にわたる網羅性の保証はなく、関係子についての検証もない。

本論文で提案する類型

以下の手順で類型（概念分類項目とその上の関係子）を設定する。基本的な類型は②までと考え、以下はその構造化で対処する。

①状態性の概念に関する分類として約30の主体系（＝表1の縦軸）を設定し、その各項目に対して6個以下の必須格（主体系格＝表2）を定義する。概念は、格1と格2、方向基準の間の関係であり、関係は素性の束である。空間概念の例について、下図と図1を参照されたい。

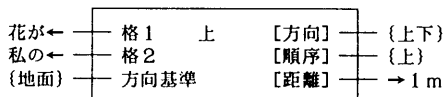
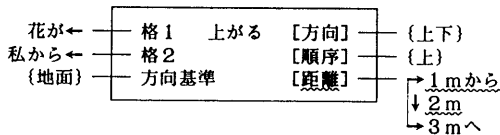


表1 縦軸に、素性を限定された概念を表示する。網がけ

したものは、さらに格の量が限定された概念である。

②主体系格のそれぞれの変化（変化性）を①に組み合わせるmatrix項目（表1の縦軸と横軸の交点）を設定する。空間概念の例について、下図と図3を参照されたい。



各matrix項目に関して、主体系格に始状態・終状態・変化（変化性格）を組み合わせた関係子を設定する（上図、表3-1、3-2）。

表1の上横軸以外に、以下の観点がある。

・非相対の場合、始状態／終状態をとるのは何か。

例：「入る」と「かぶせる」の違い

・変化分をとるか始状態／終状態のみか。

例：「向く」と「傾く」の違い

matrix項目に属するheadに「ている」がついて結果状態になった場合、変化性を状態性に戻した項目となる。

③各語義（TP=terminal concept）は、①、②の項目の下位または以下に示す合成として表現される。

・主体系として補助的概念（下表）をもつ概念が、本動詞的概念に任意個後接する。時間については表2の格1 2、その他の補助的概念については方向関係子で後接する。つまり、本動詞／助動詞の関係に準じる。

概念	機能語での表現例	語彙化された合成の例
可能	-うとする、-する、-業、 -せる、-難い、-概、-法、 -上手	指導=学ば/す、使う=動か/す 競う=勝と/うとする
評価	-始める、再-、-性、先-、 -続ける、-立て、 -がる、-ぶる、-たい、 -好き、-主観、-感、-心	けち=けちる/恒常性 休む=進行度/継続 値切る=値下げ/要求
情報	同-、別-、各-、多-、共-、 -過る、-合う、間接	探す=位置を/調査 偽る=伝達において/変える 伴う=移動において/含む

補助的概念としたのは、関係子のフィラとして再びheadになりえる概念をとりえるもの全部である。

変化と補助的概念の合成で状態性になっている「けち」などの概念は①の段階では考えず、④で合成として考える。このため、主体系の数がぐやみに増えない。

・因果など付帯状況的な合成

例：お辞儀する＝挨拶のために体を曲げる

表4と5に、headの意味に依存せずに定義できる深層格を示す。以下ではこの深層格にある係りを自由格と呼ぶ。特に表層自由格という場合は、自由格にあたって、統語上広い範囲の概念に係り得る表層格を意味する。また以下でhead/dependentとは特に断わらない限り統語上のものではなく、表2、4、5の深層の関係子のhead/dependentと解釈されたい。必須格とは、自由格及び補助的概念のdepend entとしての係り以外の係りと解釈されたい。

検証すべきこと

例文解析を中心として、以下の点について検証した。

検証範囲と方法についてはappendix 2を参照されたい。

①設定した少数の概念分類項目（の合成）が個々の語義をどこまで網羅するか。

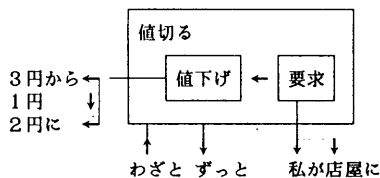
②設定した関係子が個々の係受けをどこまで網羅するか。

概念分類項目レベルでは網羅できない係受けがあるか。

分類項目の合成操作が関係子におよぼす影響（下図）。

表 1: 概念分類項目 = head にあたる概念の分類

[illegible]



- 必須格が部分概念の性質から説明できるか。
例：露店商に値切る＝値下げを要求
20円 値切る＝値下げを要求
- 自由格のheadである部分や、他の概念のdependentである部分の指定が逐一必要か。
例：わざと値切る＝「わざと値下げ」を要求
ずっと値切る＝「ずっと値下げ」を要求
- ③ 類型化のために捨象された意味はどのていどか。
- ④ 関係子と表層格との関係。

以下の節で、これらについて順に検討していく。

2. 概念分類項目の、語義に関する網羅性

1. で示した設計方法からわかるように、語義に関する網羅性は高い。特に言語間で差の大きい状態性／変化性の差や、「させる」など補助動詞的概念の付加に対する吸収力が高い。

国語辞典の守備範囲の語義は網羅できた。ただし複合的事象に関しては、関係子に寄与する部分のみをとりだした。

例：スポーツの類は「競う」「遊ぶ」と同じに解釈
なおモノについては、今回の分類のフィラとしての性質で死物、生物、組織、情報に粗く分類した。

3. 関係子の、個々の係りに関する網羅性

関係子の合成を許すと検証範囲の文はほぼ網羅できた。解釈しにくいもの（主に体言どうしの接続）は2%以下だった。なおモノとモノの直接の係りは対象外とした。

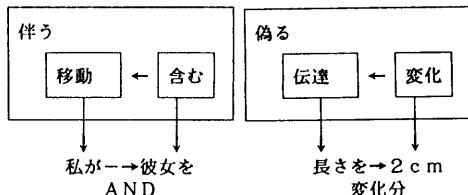
3. - 1 Headが概念分類項目の単純な下位である場合

概念分類で定義した必須格（表2、3）と、付帯状況的な自由格（表4）以外に、以下の係りがある。

概念を含む機能語

表5に格属性としてあげたものは、表層のheadとの直接の関係ではなく「{」で示したそのフィラとの関係と考え、表層のheadと直接つなぐ場合は合成関係子（表2）として記述の方が情報が多い。このうち、始状態・終状態、数量は表1にあげた関係子と同じものである。

これらの関係子の接続は、実際には補助的概念の後接と同じことである（下図）。



このように、実際には基本的な関係子ではなく、補助的概念を含む構造であるために、本来はこれにかかる関係子をもつ場合もある（表5の格属性「差」）。

例：彼女より2m伸びる＝伸びが、彼女を2m越える
補助的概念だけでなく、本動詞的概念の構造をとる機能

語もあるが、言語依存性が大いなので個別に対処することになる。その場合も、表1と表2の類型が指針を与える。

構文上の係り

主語化による係り、代動詞への関係子の転嫁、本動詞の省略、代用、隠喩については本来の係り相手への関係に戻す。文中に真の係り相手がない場合も多いので表層格との関係子合成（図2）で扱う。

代動詞に準じる問題として、補助的概念の表層必須格として本動詞的概念への係りが転嫁される場合も同様に扱う。

例：彼女への責任＝彼女へ？するという責任

必須格のフィラへの係りが転嫁する場合も同様に扱う。

例：音色で鳴る＝音色の音が広がる

関係子とはみなさない係り

慣用表現における、語源由来の係り（例：誇りを地面に捨てる）や、単なる名前（例：グリーン車）は関係子とは見なさず、ダミーなラベル idiom、name で対処する。

慣用表現か比喩かの違いは転義した語義が辞書にあるかどうかで決める。

明示的でない合成による係り

ありがたい関係相手と結びついてあたかも合成概念と同じ振舞いをする場合がある。

① 明示的でない因果合成

表層の構成語とのリンクから説明できるものが多い。この場合省略として扱う。（図2）

例：穴からいぶす＝いぶし出す

他には、前提・方法や結果としての空間的な動きが、特に、空間概念や評価概念がheadの場合に多い。

例：皿から（移動して）食べる

箱に（入れて）整理する

また、方法や目的としての情報移動、空間的概念の結果としてのモノの生成、目的としての闘争・受益などがある。これらは、自由格として表4に仮にいておく。

② 明示的でない「ている」

「ている」がついていなくても、結果状態の持続時間を示す場合がある。例：2分止まる

3. - 2 Headが合成概念である場合

関係子の網羅性

合成概念のもつ必須格は、部分概念が3-1の意味でもつ関係子に帰着できた。

全体と部分の違いは、因果関係子などの自由格のフィラの違いや、dependentとしての性質の変化にある。

例：パレットで描く（空間&情報）

? パレットで教える、パレットで動かす

○ 食事で歓談する、? 大食い歓談する

関係子に寄与する部分の指定

語彙化された合成（表層で構成語として分解できない場合）の場合のみ考える。

① Headである部分の指定

合成概念がもつ必須格に関しては、headである部分概念を逐一指定する必要がある。

自由格のheadが合成概念中のどの部分かは指定する必要はない（表8）。格属性によって持ち込まれる補助的概念は、語彙化された合成概念として後接する補助的概念の間に介在することはない。

例：卵だけ隠す＝「教えない」のを卵に限定

×「教える」のを卵に限定しない

このため、文脈を考慮しないで合成概念内で正規化を行える（表6）。

②Dependentである部分の指定

Headになりえる概念をdependentにとりえるのは補助的概念だけなのでそれだけについて考えればよい。

必ずしも全体にかかる場合だけではなく。

例：優雅に寝る＝優美な形にする／優美に形にする

4. 類型化で捨象される意味

これまでの議論では捨象してきた意味のうち興味あるものとそれへの対処についてまとめる。この他捨象した意味についてはappendix 1にまとめた。

関係子の意味の粗さ

①関係子間関係^{5, 6}

Headとなりえる概念をフィラとしえる関係子（付帯的關係子や、補助的概念のもつ関係子）については、真の意味記述を行うにはheadとdependentの両概念のもつ関係子の間の関係も記述する必要がある。関係子合成（図2）との重複付与でこれは行える。（下例では下線部がhead）

例：硬くしぼる＝[AND 原因 格1-格1]

i.e. xをしぼると、そのxが硬くなる

うきうきと歩く＝[AND 方向 格1-AGT]

概念を含む構造も同様にして記述できる。

例：私の曲＝私が作った曲＝方向{可能}+格1

②項目の粗さに伴うもの

appendix 2-2を参照されたい。

③意志性

関係子AGTとの重複付与で記述する。

例：彼が与える＝[AND 格2+始 AGT]

④複数の文の圧縮に伴う、フィラの対応関係の喪失

例：両者の商売の境界＝両者の商売場所の境界

項目の粗さ

どれほどの語義が同一主体系をもつかについて、表8を参照されたい。語義レベルの意味は、他の項目や語義との関係子（と関係子間関係）の分布で記述してゆく。

例：「叩く」は「壊れる、傷む」と因果関係が強く、「なでる」の場合はそれが弱い。

地の文との関係

語義のもつ視点情報や話法の意味も、「外側の文＝情報概念」を復元した関係子間関係で意味記述ができる⁸。

例：xへ来る＝方向{情報}+格1・始+場所2
i.e. yが「xへ来る」と言う。yはxに居る。

計算過程を含む意味

比喩や発話の意図などの解釈において使う知識も、今回の類型で表現される因果知識などと別のものと考える必要はなく、連続的なものだと考えている⁸。

5. 表層格との関係

表層格としての性質は、head、dependentともに補助概念的なほう（または全体）にある。これは3-2のスコープの問題と対応する。

例：2時に炊ける＝2時に終わる＝2時まで炊く
×2時に炊く

Dependentとしての性質

①関係子まで含む概念の問題。

例：○アメリカで買う、×アメリカンで買う

品詞で対処するか、「関係」概念にいれるべきである。

②限定済みの概念の問題。

例：「高値で買う」と「円高で買う」の違い。

この場合「品物を値段で買う→品物の値段」と「円の値段」の抵触による整合性計算から導けるが、常にこれが可

能かどうかは今後の検討を要する。

Headとしての性質

①必須格については、項目または語義ごとに表層格とのマッピングを行う。

②深層で自由格とした表層格が統語上完全に自由ではない場合がある。概念の性質から説明できる例を示す。

例：○（人が熊を）空から調査する＝「人」の場所2
×（人が）空から働く

i.e.「働く」には「人」以外のモノが必須格として関与しないので「から」のような方向性のある表層格を使わない。

6. おわりに

作業化

この類型は基本設計を与えるものである。運用の障害となるのは概念や関係子の合成化の繁雑さであろう。運用上は、概念合成による仮分類や関係子合成による仮関係子を設定するなどの調整を行う^{3, 4}。

今後の課題

類型自体の課題としては以下のものがある。

①英語に関する、この関係子の網羅性の検証。

②関係子の過剰

必須格の過剰に関しては、対応表層格がない、ということでもマスクできる。現在、フィラ限定や合成に由来する関係子の過剰に関しては表7のような現象がわかっている。

自由格についてはaspect素性やAGT関係子の有無で説明できる場合が多いことがわかっているが、表層格との対応を通して検討してゆく。

今後、表層格との対応を通じて構文解析とのリンクを進め、意味解析の枠組みとしての検証をしてゆく。その中で、今回捨象した意味（第4節）への対処や獲得方法を検討してゆきたい。

謝辞

研究の機会を与えてくださったEDRの横井所長、千葉室長、牧野前室長（現東邦大学）、例文解析を手伝って下さった中原さん始め第2研究室の皆さんに感謝します。

reference:

- 1) Schank and Abelson "Scripts, Goals, Plans, and Understanding"
 - 2) 大野、浜西 角川類語新辞典
 - 3) 井原、奥村「格関係に基づく概念体系」第38回情処全大
 - 4) 奥村、井原「格関係に基づく概念体系の検証」第37回情処全大
 - 5) 奥村「概念辞書における知識表現」第38回情処全大
 - 6) 奥村、牧野「格関係値間関係による概念規定」
- 知識工学と人工知能 研究会 55-3

合成や限定の性質	限定前	限定後（共起しなくなる）
フィラ＝ANY	彼に・目移り 内容を・忘れる	彼に？浮気な（女） 内容を？忘れっぽい
フィラ＝NONE	靴を・脱ぐ メロディで・鳴る	靴？裸足 メロディで？静まる
素性限定（ゼロ）	2m間隔に・並ぶ	（一と）間隔に？付着
矛盾する素性限定	下手くそに・話す 詳しく・知っている	下手くそに？雄弁 詳しく？無知
	詳しく・伝える	詳しく？隠す
	さしむかいで・会う	さしむかいで？行き違う
→補助概念との合成 の標準化規則	2分・炊く 2キロ・移動	2分？炊ける＝炊き終わる 2キロ？道路を断つ 2キロ？到着

↑表7：合成によって、headとしてとらなくなる関係子

	語義例	標準形	解釈例	例	語義例	scope	根拠
素性変化の否定 素性変化の完了 素性変化の開始 格値変化の否定	野放し＝教育しない 到着＝移動の完了 出発＝移動の開始 隠す＝伝えない	素性逆限定の継続 素性限定の実現 素性逆限定の実現 始状態の継続と 終状態の否定の継続	無知に留まる 距離無になる 距離有になる 始点が知に留まる 終点が無知に留まる	貸しきる 安易 まぶしい 隠す	一以外に使わせない 努力が不必要 見てもらえない 伝えない	閉 閉 閉？ 閉？	ずっとは×貸しきる ずっとは×安易 ずっとは？まぶしい ずっとは？隠す 詳しくは？隠す

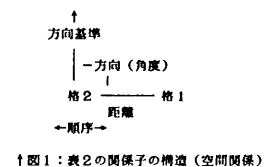
↑表8：補助的概念との合成におけるスコープ

変化性	素性変化						素性限定				
	方向変化			方向限定可			格要素変化				
	値変化		量変化	順序（配列）変化		距離変化		値変化		量変化	
	相対	非相対		相対	非相対	相対	非相対	格 1	格 2	格 1	格 2
変化性 由来の 関係子	方向・始 方向・変化 方向・終	方向・始 方向・変化 方向・終	次元・始 次元・変化 次元・終	順序・終 距離・終	順序・終 距離・終	距離・始 距離・変化 距離・終	距離・始 距離・変化 距離・終	格 1・始 格 1・終	格 2・始 格 2・終	格 1 量・始 格 1 量・変化 格 1 量・終	格 2 量・始 格 2 量・変化 格 2 量・終
	方向変化の 方向	方向変化の 方向 格 1・始 格 1・終			格 1・始 格 1・終 格 1・変化		格 1・始 格 1・終 格 1・変化			格 1 の部分 どうしの関係	格 2 の部分 どうしの関係

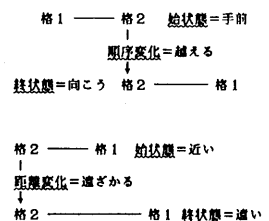
↑表3-1：変化性に伴って生じる必須格（空間関係の場合）

↓表3-2：表3-1の例（空間関係の場合）

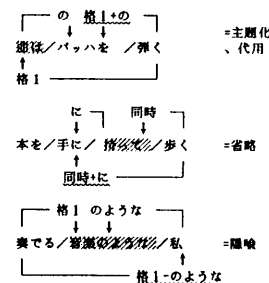
変化性格 主体系格	始状態	終状態	変化分	原形
格 1	ここから・越える	そこまで・越える	まっすぐ・近付く	私が・越える
格 2	ここから・近付く	そこまで・近付く	まっすぐ・越える	家が・近付く
	ここから・旋回	そこまで・旋回	軌道を・旋回	1 者が・挟む
距離=格12の大きさ、 格1と2の距離	5mから・近付く	3mに・近付く 2m・越える 2mに・並ぶ 長さに・つなぐ	2m・近付く 2ミリ・へこむ	川を・越える 駅に・近付く 城を・旋回 ここで・交差 ここで・折れる 2 者が・挟む
方向=格12と 方向基準の角度	18度から・傾く	斜めに・並ぶ 20度で・交差 20度で・傾く まん中に・挟む	2度・傾く	東西に・近付く
方向変化の方向	左から・傾く		右に・傾く	
方向基準				机に・並ぶ 隣りを・囲む 障に・交差 道を・近付く walk by the river ここで・つなぐ
順序=格12の配列	東から・近付く 東から・越える	西に・越える 星形に・並ぶ L形に・つなぐ	西に・近付く	
格 1 量	10人から・すく	2人に・すく	8人・すく	
格 2 量	1 か所から・散る	5 か所へ・散る		
格1, 格2+全体				棒が・縮む 棒が・交差 棒が・傾く 机が・へこむ 2 者が・近付く
格1, 格2+全体、 方向基準+全体	日英から・近付く	中韓に・近付く		棒が・折れる
方向基準+の 距離+単位		間隔で・並ぶ	歩幅で・近付く	彼女を・囲む



↑図1：表2の関係子の構造（空間関係）



↑図3：素性変化



↑図2：関係子合成 網かけ部分は省略
+は合成の方向

概念分類	語例 =以降は解釈。
関係	(札を)くずす=替える 差別・公平=違う・同じ/に見る (バカの)見本=一様、一風
時間	凝る・おかげさ=過剰 猶予=遅れ/を許可、入れ違い=時間/が異
情報的	先輩=新旧、機敏=速度=所要時間 教師=教える/役割、有名=大勢が知る 浅ましい=欲しい/恒常性 敏感=知る/のが早い スパイ=知らせる、失明=見えない ペラペラ=喋る、
可能	木化、木製=木で作る、母=生む 寛大=自由にさせる、(法)を守る 自ずから、優柔不断=決められない 使う=機能させる、盲人用=盲人が使う
機能	疲れる=機能しにくくなる、一方式 育つ=進歩する、一病、回復 ていねい=良く扱う
評価	害する=悪くする、危険=害が可能 誇る=優越すると思う、
社会的	応対、交際、交尾、威厳=上位/らしい なれなれしい=したし/け
空間的 距離	まひ=動けない、出身=居た 長さ、距離、歩幅、大きさ、ぶくぶく 歩く、バク足、血流、見る=光の移動 タックル、蹴る、さする、植える、入植 拭う、洗う、切る、ふとんむし、替る 丸、模様、あくび=変形 弱通し、開く=出入り可能化、呼吸=出入り

↑表9：主体系を共有する概念の例
/以降は補助的概念の後援

appendix 1. 今回考慮しなかった「意味」

- ①自由格の詳細な分類
順接・逆接、仮定・事実、目的・結果などの差
- ②語彙よりは構文に属する意味とみなしたもの
「が、を、は」、連体修飾、語順、表層格フレームなどが担う主題、焦点、強調、対比や限量子的な意味
- ③関係子を増やさないで今回考慮しなかった意味
 - ・評価的な意味
項目としては、「けちる=節約する」
関係子としては、「彼女なんか嫌い=彼女を嫌い」
 - ・付帯状況のうち関係子に寄与しないものは無視
行く=戻る、停止=到着、びっこ=歩く
- ④発話状況や文脈から計算される意味
 - ・視点の違い(行く=go=移動、貴方=私=人)
 - ・使用分野や用法、語法的な意味(敬語など)
 - ・限定や照応に関する意味(その=物事)
 - ・発話の意図、比喩など計算を要するもの
- ⑤単にフィラの違いに帰着する違い(在る=居る)
- ⑥事実性の違い
 - ・不可能と否定、「する」と「うとする」
 - ・構文から生じる事実性の違いも無視した。
例：「上手に歩く」と「歩くのが上手」
- ⑦言語の媒体に関するもの
 - ・だじゃれ、韻など発音に関する意味(ex.試験のシ)
 - ・形としての文字(ex.姿の中の女)
- ⑧基本語辞書がもつ情報
 - ・品詞、表層格フレーム、aspect形式素の接続性
 - ・慣用句における必須要素(油を/煮る)

appendix 2. データ

2. - 1 データの内訳

A. 分類項目の網羅性

- ①角川類語辞典掲載の全語彙を概念分類項目で解釈
- ②対訳が語彙化されていない英語語義の検討

B. 重要な語義に関する詳細な分析

Aの①の結果から選んだ以下の語義のまわりの例文解析

コーパスは朝日新聞12000文

- ①各分類項目の代表語 580語義 各30文程度
- ②境界語、合成概念 250語義 コーパス中の全部

C. 特に合成概念を中心とした分析

- ①補助用言相当語句や因果を意味する機能語を含む語義文をもつ語義の分析

②Bの②に、否定との合成とも見えるものを多く選んだ。

例：隠す、ボイコット、そっちのけ、停止、見逃す

D. 表層格との関係に関する分析

特に表層で自由格と見なせるものに関して、機械的に生成したデータの視査。

E. 機能語中心の分析

国語辞書、英和辞典中の機能語の用例分析

これから、概念を含む機能語を洗いだした。

F. 広い範囲の語義に関する(浅い)分析

- ①類語辞典掲載の用例。

- ②選ばれたコーパスの全文解析

ref3,4の仕様(合成概念や合成関係子の仮設)による。

「関係子付与不能」は2%以下「困難だが可能」は10%

朝日新聞3500文

平凡百科(自動車関係の項目)500文

2. - 2 データの作成及びチェックルーチン

2. - 1のデータB, C, Fに共通

①語彙への項目付与

- ・項目の付与方法=語義を項目の代表語(の合成)で言い替える
- ・項目の優先順位=matrix項目(合成でない)→matrix項目の合成→項目外ただし、恒常化しただけで評価、有意化しただけで機能、人間が2人関係する場合なんでも交際などとする傾向があったので、これらの項目にはなるべくとらないよう指示する必要がある。
補助的概念との合成の際に生じる多義性は、表6の標準化を行う。
- ・チェック事項=項目のユレや不足のチェック→項目追加、マージ

②係り受けへの関係子付与

- ・関係子の付与方法=headを項目の代表語に替えて、文章を言い替える。
必要ならば語順も替える。
- ・関係子の優先順位=matrix関係子→格属性→表層格シフト→付帯的→概念をもつ機能語→暗黙の合成
但し、「より」「あと一」の場合は、格属性をmatrix関係子に優先させる。
どちらもheadにとりえる場合は、表層のheadを優先。
分類項目が粗いために格1と2の違いが定義できない場合(ex. xがyの前)や、再帰的?用法(ex. カーテンがすそをたらず)でどちらが格属性かわからない場合は「がをにで」の順で決める。
- ・チェック事項=関係子のユレや過不足。
→①で語義に付与した項目の変更や合成化。
項目や関係子の細分や追加。
関係子とみない係りの設定。
但し関係子間関係の付加で区別できる違いである場合は関係子を増やさず、合成による仮関係子とする。
例：うきうきと歩く/うれいしいことに歩く

③①②を項目(と関係子)ごとにソートして見直す

同一の項目-関係子をもつ係りの意味の幅

共通性は、項目-関係子の定義を反映しているか。

落ちた意味(語義や個々の係りの違い)は何か