

語彙概念構造による動詞辞書の作成

竹内 孔一

岡山大学工学部情報工学科

koichi@it.okayama-u.ac.jp

1 はじめに

文章を解析する上で文の中心的役割を果たす動詞とそれを取りまく名詞との意味関係を解析することは大変重要である。意味関係を解析する第一歩として表層の格を伴って現れる名詞と動詞の関係が深層での意味の結び付きを概形的にとらえることが一文の意味を解析する鍵となる。

語の意味構造の記述した辞書として EDR と IPAL の辞書がある。人手をかけて良く整備されているが個別の動詞の表層の格関係の情報と意味分類との関係の記述が重視されているため、語彙としての動詞間のタイポロジーの見通しや構文構造で関連する言い換えについてうまく処理できない。

一方、語彙意味論研究において名詞と動詞の関係を意味役割で記述する項構造 [1] が発展している。この考え方をさらにすすめて項同士の概形的意味関係を限られた意味述語で記述する語彙概念構造が発展している。語彙概念構造は項構造、アスペクト分析を取り込んで表層表現との関係を意識しつつ概形的な深層の意味構造を明確に表現できるところが利点である。

本研究では理論で展開されてきた語彙概念構造を日本語の動詞、とくにサ変名詞について人手で実際に付与を行い理論の検証と整理を行っている。本発表では 1) 理論をどのように整理して概念構造を付与してきたか、2) 約 1300 語の動詞を付与した段階で問題となった点、3) 全語彙概念構造パターンについて明らかにし作成した語彙概念構造辞書がどう応用可能であるかについて明確にする。作成している語彙概念構造は web 上で公開する。

2 語彙概念構造の理論背景

理論上で議論されてきた語彙概念構造には様々なアイデアが盛り込まれている。我々はこの中からなるべく単純で明確な理由があり言い回しによるテス

トによってその構造が推定できる要素を取り込んで構築した。

本研究の語彙概念構造 (TLCS と呼ぶ) に採用した考え方は基本的に 1) 項構造 ([1]), 2) アスペクト分析 ([3]), 3) 概念の意味記述 ([2]) の 3 つである。順に説明する。

2.1 項構造

項構造は動詞がとる名詞の意味役割を記述した物でいくつかの意味役割が提案されている [1]。名詞の意味役割は文の表層において必須格と結び付くことからどの範囲の意味役割を扱うかは必須格をどこまで認めるかという問題がある。本研究では意味役割集合は最も基本である Agent, Theme, Goal の 3 つを中心に記述することにした¹。これらは日本語では表層格としてガ格、ヲ格、ニ格を伴って現れ、態を入れ換えた場合にその差異が明らかになる。

3 その文を英語に翻訳する/その文が英語に翻訳される

4 道路を歩く/*道路が歩かれる²

この例からは上文のヲ格は Theme に対応するが下文のヲ格は異なる深層の意味関係であることがわかる。この構造が語彙概念構造にどう反映されているか例で示そう。

5 [x CONTROL [BECOME [y BE AT z]]] (翻訳)

6 [x MOVE] (歩く)

x, y, z は項 (Argument) と呼ばれ名詞をさしており CONTROL 等は項同士の関係を記述する意味述語である³。語彙概念構造では Agent, Theme, Goal とラベルを張るかわりに構造で表層と深層の対応をとる。表層との対応ルール (linking rule) は単純で

- 角括弧の外側から順にガ, ヲ, ニ格に結び付ける (ただし CONTROL と ACT がある LCS)
- それ以外は順にガ, ニ格と結び付ける

¹ 必要に応じて Source も記述する。

² *印は意味を目的とする意味をなさないことを表している。

³ 詳細は [8] を参照。また全パターンは 4.2 を参照。

となる。上記の LCS では x, y, z の順にガマ二格に対応する。また深層の意味役割との対応も同様で

- 角括弧の外側から順に Agent, Theme, Goal に結び付ける (ただし CONTROL と ACT がある LCS)
- それ以外は順に Theme, Goal 格と結び付ける

である。上記の場合 5 の例では $x=Agent, y=Theme, z=Goal$ となり、6 では $x=Theme$ である。

2.2 アスペクト分析

アスペクト分析とは動詞の動作様相を分類する方法のことで Vendler [3] による 4 分類をはじめ細分類が検討されている。アスペクト特性は LCS の意味述語に記述されており概略, ACT が継続的な動作 (活動) を表し, BECOME が動作の達成, BE が状態を表す。例えば 4.2 節の L1 は継続的な動作の動詞であり, L2 は動作が変化して y がある目標 z になる (達成) ことを表している。

2.3 概念の意味記述

語彙概念構造はその名のとおり語彙の基本的な意味の骨組みを記述するものである。上述の項構造とアスペクトによる分析を基に意味述語の組合せで記述され、その内容が動詞の持つ意味とあっている必要がある。1つの動詞が表現できる動作には基本的なフレームが存在することが提案されており「行為 (者) 道具 対象 (物) 結果状態」[2] という行為の連鎖が最大である。語彙概念構造にこの考え方は反映されており、例えば 4.2 節の TLCS では L1 は (行為), L2 は使役 (行為 (者) 対象 (物) 結果状態), L8 は結果状態に対応する。

3 概念構造の付与

語彙概念構造を動詞に付与するために上記に示したアイデアに付随する言い替えテストを整理して作業用マニュアルを整備した。本章ではどのような作業手順で概念構造辞書を作成したかを明らかにすることで辞書の信頼度を浮き彫りにする。

付与作業は上述の項構造の分析, アスペクト分析, 概念の意味構造のチェックという手順で概念構造パターンを付与する。具体的にどのような分析を

行うか以下簡単に記述する。

項構造の分析

以下に関連するテストを行う。

- ガ格とヲ格の入れ換え (能格動詞かどうか)
- 態を入れ換える (Agent, Theme かどうか)
- 二格に関する分析

最後の二格に関する分析が理論でもまだはっきりしないためなかなか難しい。二格が深層の Goal であるかを調べるため主に以下の分析を行う。

- 二格が受身をとれるかどうか所有関係の分析
- 二格と他の格との言い換えのテスト (例えば「彼に教わる/彼から教わる」の場合は「二格」の項構造は Source と分析し意味述語 FROM を対応させる。)
- オノマトペや副詞は二格と認めない

こうした分析の結果どう分類を進めていくかの詳細については辞書と共に公開する付与マニュアルに記載する予定である。

アスペクト分析

副詞との共起による動詞の意味解釈で分類を行う。語彙意味論でさまざまな調べ方が提案されており、その中で意味述語との結び付きがはっきりしている言い回しのテストについて選んで整理した。以下に一部を記述する。

- 「10分で」および「10分間」をつけて動作の終了特性を調べる
- 「てある」をつけて状態変化が起こったかどうかを調べる (あれば BECOME という述語が想定される)
- 副詞「たくさん」を付与して何が多いかを調べる
「たくさん笑う」(笑う量がたくさん ACT)
「たくさん発生する」(発生したものがたくさん BECOME)

これにより概念構造に含まれる意味述語を想定する。概念構造の構築

上記2つの分析で動詞のとり項と意味述語がわかることから以下の手順で語彙概念構造を決定する。

- 意味述語間の構築に関する制約 (影山 [8]) に基づいて語彙概念構造を構成
- 構成された語彙概念構造の意味が動詞の持つ概念的な意味として合っているかをチェック
- 外れている場合は再構築

問題となる部分は概念構造が動詞の意味と合わない

場合である。この場合、再構築のために以下の2点について検討を行う。

- 他の動詞とどう振るまいが違うかについて様々な言い回しを適用して明確にする
- 違いについて整理して振る舞いの似た動詞を集める

違いを整理した後、語彙意味論の分析を参考に新たに概念構造を仮定し更新をつづける。次章で困難な付与例に付いて取り上げどのように分析したか明らかにする。

4 語彙概念構造の付与結果

上記の付与手順をもとに約1300語の動詞に対し語彙概念構造を付与した。作業で問題となったものの中で主な事として以下の3つがあげられる。

- 必須格の範囲の設定
カラ格、ト格などどこまで認めるか
- 二格に関する分析
表層で「～に」という表現はかなり多く、項構造との対応が不明なものが見つかった
- 概念意味の不一致
項構造とアスペクト分析から普通に概念構造を設定すると動詞の持つ意味とまったく一致しない動詞がいくつか見付かった。とくに心理動詞、要求動詞について整理を行う必要があることがわかった。

これらの問題は独立しているわけではなく相互に絡んでいる。以下ではこの問題点についてどこまで対処できたかを述べ最後に最新版の語彙概念構造のパターンを示す。パターンの種類は最初は14種類で設定したが分析の結果16種類に増えた⁴。

4.1 問題と再分析結果

必須格について

必須格としてガ、ヲ、二格以外にカラ格を扱う必要性がでてきた。しかし完全な付加詞であるもの(例「東から風がふく」)とは区別したい。そこで言い換えによりガ、ヲ、二格に置き換えればそのカラ格を認めFROMという意味述語で記述することにし

⁴ FILLED という埋め込みタグの違いで実際は約2倍の概念パターンであるが下位分類であるのでLCSパターンとしては16種類と考える。

た⁵。例えば以下の例においてカラ格は必須項とみとめ概念構造に記述する。

- 風呂から上がる/風呂を上げる
- 先輩から教わる/先輩に教わる

それぞれのLCSは

[x CONTROL [x MOVE FROM y]] (上がる)
[BECOME [z BE WITH [y MOVE FROM x TO z]]] (教わる)

とした。表層のカラ格とFROMをリンクする(影山[4]を参照)。

二格に関する問題

二格で表現される動詞の中で以下の相手を表す二格[7]を語彙概念構造に反映できていない。

- 7 教師に反抗する
- 8 作戦に関与する
- 9 太郎に賛成/反対する

これらの二格は必須格であると考えられる⁶が項構造で考えると何かが到着するGoalではない。明らかに対応する深層の意味が見当たらないのである。現段階ではこの「相手の二格」は付加詞扱いとして構造に反映していない。アスペクト分析の観点から例7~9の語彙概念構造は[x ACT]である。相手の二格をどう構造で反映すべきか他の二格の現象をさらに洗い出すことで見直したい。

概念構造の意味の問題

項構造とアスペクト分析に基づいて使役構造をもとに概念構造を作成すると概形の意味が合わないものがある。主に下の2つのタイプの動詞が問題になることがわかってきた。

- 心理動詞(語彙意味論で常に問題)
- 要求動詞(今回の付与で再発見)

ここでは要求動詞について問題と再分析の結果について説明する。

例えば「要求する」は「ファンが松井にサインを要求する」のように完全な他動詞である。しかし

[x CONTROL [BECOME [y BE AT z]]]

という使役の語彙概念構造は意味をなさない。なぜならこの構造だとy=サインをz=松井に与えるようにx=ファンが動作することを意味するからである。

これも二格の分析の問題にかかわるが、再分析として文の格関係と意味関係を確認し意味的に構造を

⁵ 項構造ではSourceという意味役割に対応する。

⁶ 「教師が反抗された」と態の入れ換えに関与するため。

構築し直してみる．態を変えると

- 松井がサインを要求される
- サインが松井に要求される

のようにどちらの格も受身が可能であるので英語の間接目的語，直接目的語が並ぶ構文 (ex. I gave her a dress.) に対応する．この場合 [5] の分析から「松井」と「サイン」の間には所有関係が成り立つことが分かってるので「要求する」は「松井」にあるサインが自分のところに来るように求めることである．これを概念構造で記述すると 4.2 節の L16 のパターンとなる． x, y, z の表層への対応は同じである．FROM z は 2.1 節のリンキングルールから表層では二格となるが意味的には Source であることを示している．現段階ではこれを支持する確実な言い回しは分かっていないが他の動詞とのふるまいが違うことを示す意味でこの構造を採用する．

4.2 TLCS パターン

下記に最新の全 TLCS パターンを示す．

- L1 [x ACT ON y] 操作，運転
- L2 [x CONTROL[BECOME [y BE AT z]]]
処理，翻訳
- L3 [x CONTROL[BECOME [y NOT BE AT z]]]
遮へい，抑止
- L4 [x CONTROL [y MOVE TO z]] 伝送，伝搬
- L5 [x=y CONTROL[BECOME [y BE AT z]]]
回復，終了
- L6 [BECOME[y BE AT z]] 飽和，分布
- L7 [y MOVE TO z] 移動，遷移
- L8 [x CONTROL[y BE AT z]] 管理，維持
- L9 [x CONTROL[BECOME[x BE WITH y]]]
認識，予測
- L10 [y BE AT z] 関係，位置
- L11 [x ACT] 会議，行列
- L12 [x CONTROL[BECOME [[FILLED]y BE AT z]]] 署名
- L13 [x CONTROL[BECOME x BE WITH [[FILLED]z BE AT y]]] 尊敬，信用
- L14 [BECOME[y BE WITH [[FILLED] BE AT z]]] 感動
- L15 [BECOME [z BE WITH [y MOVE FROM x TO z]]] 受ける，授かる，教わる
- L16 [x CONTROL [y MOVE FROM z TO x]] 要

求，求める

5 おわりに

語彙の概念的な意味記述法である語彙概念構造の必要性ならびに利便性に付いてその理論背景をもとに説明した．現在構築中である語彙概念構造辞書にどのように理論が反映されてどの程度の情報がどのくらいの信頼度で記述されているかについて付与作業と現在分析中の問題から浮き彫りにした．現在この辞書を応用として言い換えシステムの構築 [6] の研究が始まっている．この辞書は web 上で公開するので広く利用されることを望む⁷．

6 謝辞

奈良先端大自然言語処理学講座の降幡，藤田，乾の 3 氏にはたくさんの有益なコメントを頂いたことを感謝する．本研究の一部は科学研究費補助金若手研究 (B) 課題番号「日本語複合語解析における語彙の概念構造の構築」の援助を受けた．

参考文献

- [1] Grimshaw, J.: *Argument Structure*, MIT Press (1990).
- [2] Langacker, R.: *Foundations of Cognitive Linguistics, Vol.2 :Descriptive Application*, Stanford University Press (1991).
- [3] Venlder, Z.: *Linguistics in Philosophy*, Cornell University Press (1967).
- [4] 影山太郎: 非対格構造の他動詞, 東京大学出版会, chapter 5, pp. 119–145 (2002).
- [5] 岸本秀樹: 非対格構造の他動詞, 大修館書店, chapter 5, pp. 127–153 (2001).
- [6] 降幡健太郎, 藤田篤, 乾健太郎, 松本裕治, 竹内孔一: 語彙概念構造を用いた機能動詞結合の言い換え, 第 10 回言語処理学会年次大会 (2004). B4-5.
- [7] 寺村秀夫: 日本語のシンタクスと意味 I, くろしお出版 (1982).
- [8] 影山太郎: 動詞意味論, くろしお出版 (1996).

⁷ <http://cl.it.okayama-u.ac.jp/> または <http://chezdico.ex.nii.ac.jp/>