

生成語彙理論を用いたイベント構造の自動生成

高橋 幸

東北大学大学院国際文化研究科

sachi@insc.tohoku.ac.jp

1 はじめに

日本語のような膠着言語では、使役や時制を表す接辞が付加し複雑な述部構造を作り出す過程も一種の語形成と考えられる。日本語では、複合述部 (complex predicate) が多用される。構成は極めて多様であり、全ての構成パターンを網羅的に記述することは事実上不可能である。本稿では、動的な辞書モデルを用いて基体となる動詞の意味が接辞によって意味が変化していく過程を、接辞の線形的な順序に基づき分析を進める。なお、記述に用いる枠組みは生成語彙理論 (Generative Lexicon: GL) [1] である。

2 動詞の意味

動詞がどのような項構造 (格パターン) を持ち、そして項がどのような意味役割を持つかは、その述語の語彙的な意味と密接な関連があることが指摘されている。特に、完了 (telic) / 未完了 (atelic) といった語彙的なアスペクトが述語の意味を考える上で重要である。

本稿では、まず項構造における述語の分類によって、各々の語彙的なアスペクトを大まかに以下のように設定する (文献 [2] の分類に基づき、S は State, P は Process, T は Transition を示す)。日本語の形容詞、形容名詞の語彙的なアスペクトは、S に分類される。自動詞の場合は、S (状態動詞: 「ある」「いる」)、P (非能格: 「走る」)、T (非能格: 「飛び出す」) のように、3つのパターンが考えられる。一方、他動詞は基本的に語彙的なアスペクトは T に分類される。主語の行為によって、主語自身もしくは対象である目的語に何らかの変化をもたらすものと考えられるからである (Incremental Theme)。次に、文献 [3] [4] [5] の手法を採用し、以下のように動詞の分類を行った。

STEP-1: State vs. Non-state

状態動詞の非過去形は現在の動作を表すが、非状態動詞の非過去形は未来の動作や習慣的な動作を表す。

(1) a. 山が 見える。

b. 花子が 走る。

STEP-2: Activity vs. Non-activity

「もし途中で動作が止まったとしても、その動詞が表す事象を終えたと言えるか」

「「V-テイル」が「V-シタ」を含意するか」

これは activity を表す動詞の atelicity を検査するものである。以上の質問に yes と答えることができる動詞は、終結性 (telicity) がなく、activity を示す動詞だということができる。

(2) a. 太郎が 走っている (「太郎が走った」を含意している)

b. 母がケーキを 作っている。(「母がケーキを作った」を含意していない)

STEP-3: Achievements vs. Accomplishments

「「～分間後, V スル」と「～分で, V スル」の意味が同じであるか」

これは Achievements 動詞の瞬間性 (punctuality) を計るテストである。

(3) a. 孝子が 1時間後に / 1時間で 出かける。(同じ意味を表す)

b. 私の姉が 1時間後に / 1時間で ケーキを作る (異なる意味になる)

operator のない状態で、動詞の語彙的なアスペクトは 5 つに大別できる。まず、単一のイベントからなる S と P, T の下位イベントの組み合わせとして、Vendler の Accomplishments に対応する T1 と Achievements に対応する T2, そして主辞事象が曖昧な T3 の 3 パターンを考えた。S, P のように下位イベントが存在しないもの、「死ぬ」のように予め主辞事象が規定できるものに関しては、辞書に主辞事象を記載する。しかし、曖昧なものに関しては、underspecified (未指定) にしておき、文として組み上げた際に決定する。

3 アスペクト解釈の計算

アスペクトとは、基準時間における動きの時間的展開のあり方を示す。述語が本来持つ語彙的なアスペク

トは、述語の種類や意味と密接に関係する。しかし、文全体のアスペクトの意味は、述語の語彙的アスペクトのみで規定されるものではない。近年の研究は動詞句の分類と各々の意味を記述していく段階から、副詞的成分等の関わりを含め、アスペクト解釈決定のプロセスを整理していく方向へと発展している。

本節では、述語の語彙的アスペクトから、「-タ/-テイル」のアスペクト形式、副詞、付加詞等（これらの要素は以下、operator とする）が与える意味的な操作によって、包括的な時間的解釈を導き出す。

副詞や付加詞の機能は、述語の持つイベント構造（語彙アスペクト）に変化を与えるものとして、4つに分類することができる。

- | |
|--|
| <ol style="list-style-type: none"> 1. 1つの下位イベントを修飾する。 2. 主辞事象（head event）を特定化する。 3. “Transition” 構造を形成する。 4. イベントの回数を表す。 |
|--|

図 1: operator の機能

1の機能を持つのは、「がさがさ」等のオノマトペや「だんだん」「徐々に」等の動きの早さを表す副詞、「とっさに」等の瞬間性を表す副詞、「じっと」「いっせいに」などの様相を表す副詞である。文献 [2] のイベント構造は階層性を持ち、付加詞や副詞によって修飾するイベントが異なる。State, Process, Transition のいずれの（下位）イベントを修飾するかは副詞の種類によって異なる。

2の機能に含まれるのは、「早く」のような動きの様相を表す副詞、「既に」「もう」のような時間副詞、「学校に/学校の方に/学校まで」といった Path/Cause/Goal を表す付加詞、そしてアスペクト形式（-タ、-テイル）である。Transition 構造の中で、State と Process のどちらを主辞事象として特定するかは、各々異なる。

3の機能に当てはまるのは、「3時まで」「40分」のような telic な意味を持つ時間副詞、「10km 走った」の「10km」等の距離や「90度傾く」等の量を表す副詞、そして「部屋で」「真二つに」「粉々に」「ぺちゃんこに」といった Goal/Locative/Result を表す付加詞である。同様に、「～なる」「～いく」のような述語につく助動詞¹は基となる述語のイベント構造に新たな Transition 構造を加えることにより、その複雑な事象を説明できる。

「ステーキを 500g 食べた」のような量化詞を含む文は、「ステーキを 300g 食べて、その後 200g 食べた」のように複数のイベントの解釈が可能である。ま

¹「～らしい」「～だろう」等のムード・モダリティを表す助動詞については本稿の分析対象とせず、今後の課題とする。

た「太郎はスイスの教会で結婚している」のように、perfect な意味を持つテイル形や habitual な意味を持つテイル形は下位イベントではなく、イベント全体に言及し、イベントの完結性を示唆する。

4 複合述部の形成

次に、日本語で生産的な「持ってくる」のような「V1-テ-V2」型と「飛び上がる」のような「V1+V2」型²の形成について考察する。

4.1 内部関係と形成制約

V1+V2 の関係は、(4) の 5 つのパターンに分類することができる [7]。(4b),(4d),(4e) では、V1 と V2 の時間関係がほぼ同時期であるが、(4a),(4c) では、V1 が先行して V2 という形である。

- (4) a. 手段：V1 することによって V2 . 切り倒す，踏み潰す，押し開ける，折り曲げる，切り分ける，…
- b. 様態：V1 しながら V2 . 尋ね歩く，転げ落ちる，遊び暮らす，忍び寄る，持って来る，…
- c. 原因：V1 の結果 V2 . 歩き疲れる，抜け落ちる，おぼれ死ぬ，…
- d. 並列：V1 かつ V2 . 泣き喚く，忌み嫌う，恋い慕う，慣れ親しむ，…
- e. 補文関係：V1 という行為/出来事を(が) V2 . 見逃す，死に急ぐ，聞き漏らす，晴れ渡る，使い果たす，食べてしまう，…

複合動詞では 2 つの動詞の主語が一致しなければならない [8] [9] . 動詞と動詞が融合する際の項構造における制約として、文献 [6] では「他動詞調和の原理」(Transitivity Harmony Principle) を挙げている。これにより「他動詞と非対格 (*洗い落ちる, *染め変わる)」「非能格と非対格 (*走り転ぶ, *飛び落ちる)」「非対格と他動詞 (*揺れ落とす, *売れ飛ばす)」「非対格と非能格 (*痛み泣く, *転び降りる)」の可能性が排除される。

- (5) 他動性調和の法則 (Transitivity Harmony Principle):
動詞 + 動詞の複合において、同じタイプの項構造を持つ動詞同士が複合される。

意味的な制約としては、文献 [10] [11] の「一義的意味経路の制約」(Unique Path Construction) が一般的には満たされなくてはならない。

²文献 [6] では、「語彙的複合動詞」と「統語的複合動詞」に分類しているが、本稿ではこの差異には触れない。両者の統語的な振る舞いの違いを GL でどう表すかは今後の課題とする。

(6) 一義的意味経路の制約 (Unique Path Construction):

X という物体について, 単文内で2つ以上の異なる経路 (path) について叙述することはできない. 単一の経路 (a single path) という概念は次の2つの場合を規定している.

(a) X は特定の時点において2つの別々の位置に移動するようには叙述できない.

(b) 移動は単一の情景 (landscape) の中で1つの経路を辿らなくてはならない.

一義的経路の制約を違反すると, 「彼は椅子からくすぐり落とした」のような間違った複合語ができてしまう [12]. 「くすぐる」という状態と「落とす」という物理的な経路には異なりがあるため, 許容できない. これが一義的経路制約の違反である. 以上より, 状態変化と位置変化の両方を1つの述語を用いて表現することはできない.

4.2 現象のモデル化

本節では, 時間的關係や (5) や (6) の制約を考慮し, (4) で挙げた各々の事例を GL を用いて表示する.

(7) 手段: 太郎が木を [V₁切り] [V₂落とす].

$$(8) \left[\begin{array}{l} \text{切る} \\ \text{eventstr} = \left[\begin{array}{l} e1 = \boxed{e1} \text{ process} \\ e2 = \boxed{e2} \text{ state} \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 太郎} \\ \text{arg2} = \boxed{Y} \text{ 木} \\ \text{arg3} = \boxed{Z} \text{ artifact_tool} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{formal} = \text{cut}(\boxed{e2}, \boxed{X}) \\ \text{agentive} = \text{cut_act}(\boxed{e1}, \boxed{X}, \boxed{Y}, \boxed{Z}) \end{array} \right] \end{array} \right]$$

$$(9) \left[\begin{array}{l} \text{倒す} \\ \text{eventstr} = \left[\begin{array}{l} e3 = \boxed{e3} \text{ process} \\ e4 = \boxed{e4} \text{ state} \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 太郎} \\ \text{arg2} = \boxed{Y} \text{ 木} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{formal} = \text{be_down}(\boxed{e4}, \boxed{X}) \\ \text{agentive} = \text{throw_down_act}(\boxed{e3}, \boxed{X}, \boxed{Y}) \end{array} \right] \end{array} \right]$$

$$(10) \left[\begin{array}{l} \text{切り倒す} \\ \text{eventstr} = \left[\begin{array}{l} e1 <_{\infty} e4 \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 太郎} \\ \text{arg2} = \boxed{Y} \text{ 木} \\ \text{arg3} = \boxed{Z} \text{ artifact_tool} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{formal} = \text{be_down}(\boxed{e4}, \boxed{X}) \\ \text{agentive} = \text{cut_act}(\boxed{e1}, \boxed{X}, \boxed{Y}, \boxed{Z}) \end{array} \right] \end{array} \right]$$

(11) 原因: 花子は [V₁歩き] [V₂疲れた].

$$(12) \left[\begin{array}{l} \text{歩く} \\ \text{eventstr} = \left[\begin{array}{l} e1 = \boxed{e1} \text{ process} \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 花子} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{agentive} = \text{walk_act}(\boxed{e1}, \boxed{X}) \end{array} \right] \end{array} \right]$$

$$(13) \left[\begin{array}{l} \text{疲れる} \\ \text{eventstr} = \left[\begin{array}{l} e2 = \boxed{e2} \text{ process} \\ e3 = \boxed{e3} \text{ state} \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 花子} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{formal} = \text{be_tired}(\boxed{e3}, \boxed{X}) \\ \text{agentive} = \text{R}(\boxed{e2}, \boxed{X}) \end{array} \right] \end{array} \right]$$

$$(14) \left[\begin{array}{l} \text{歩き疲れる} \\ \text{eventstr} = \left[\begin{array}{l} e1 <_{\infty} e2 \circ e3 \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 花子} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{formal} = \text{be_tired}(\boxed{e3}, \boxed{X}) \\ \text{agentive} = \text{walk_act}(\boxed{e1}, \boxed{X}) \end{array} \right] \end{array} \right]$$

手段や原因の場合には, V1 の process の結果, V2 の state が導き出されるという causal chain が成り立つ. この場合, 項構造はそのまま複合述語に継承され, unification の形をとる. 「e1 <_∞ e2」は e1 が終了して, 次に e2 が起こるといった時間関係を示し, 「e1 ∘ e2」は e1 と e2 が同時発生的に起こること, 「e1 <_∞ e2」は始めに e1 が起こることを意味する. (10) の R は疲れたという状態を引き起こす process を示し, それが V1 に一致する.

(15) 様態: 息子は母を [V₁尋ね] [V₂歩いた].

$$(16) \left[\begin{array}{l} \text{尋ねる} \\ \text{eventstr} = \left[\begin{array}{l} e1 = \boxed{e1} \text{ process} \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 息子} \\ \text{arg2} = \boxed{Y} \text{ 母} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{agentive} = \text{ask_act}(\boxed{e1}, \boxed{X}, \boxed{Y}) \end{array} \right] \end{array} \right]$$

$$(17) \left[\begin{array}{l} \text{歩く} \\ \text{eventstr} = \left[\begin{array}{l} e2 = \boxed{e2} \text{ process} \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 息子} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{agentive} = \text{walk_act}(\boxed{e2}, \boxed{X}) \end{array} \right] \end{array} \right]$$

$$(18) \left[\begin{array}{l} \text{尋ね歩く} \\ \text{eventstr} = \left[\begin{array}{l} e1 \circ e2 \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 息子} \\ \text{arg2} = \boxed{Y} \text{ 母} \end{array} \right] \\ \text{qualia} = [\text{agentive} = \text{ask_act}(\boxed{e1}, \boxed{X}, \boxed{Y}) \wedge \text{walk_act}(\boxed{e2}, \boxed{X})] \end{array} \right]$$

様態と「太郎は [V₁泣き] [V₂喚いた.」の並列は同様の構造で表すことができる.

(19) 補文関係: 花子はゲームを [V₁見] [V₂逃した].

$$(20) \left[\begin{array}{l} \text{見る} \\ \text{eventstr} = \left[\begin{array}{l} e1 = \boxed{e1} \text{ process} \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 花子} \\ \text{arg2} = \boxed{Y} \text{ ゲーム} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{agentive} = \text{watch_act}(\boxed{e1}, \boxed{X}, \boxed{Y}) \end{array} \right] \end{array} \right]$$

$$(21) \left[\begin{array}{l} \text{逃す} \\ \text{eventstr} = \left[\begin{array}{l} e2 = \boxed{e2} \text{ process} \\ e3 = \boxed{e3} \text{ state} \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 花子} \\ \text{arg3} = \boxed{Z} \text{ T} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{formal} = \text{be_missed}(\boxed{e3}, \boxed{Z}) \\ \text{agentive} = \text{miss_act}(\boxed{e2}, \boxed{X}, \boxed{Z}) \end{array} \right] \end{array} \right]$$

$$(22) \left[\begin{array}{l} \text{見逃す} \\ \text{eventstr} = \left[\begin{array}{l} e1 \circ e2 \end{array} \right] \\ \text{argstr} = \left[\begin{array}{l} \text{arg1} = \boxed{X} \text{ 花子} \\ \text{arg2} = \boxed{Y} \text{ ゲーム} \\ \text{arg3} = \boxed{Z} \text{ T} \end{array} \right] \\ \text{qualia} = \left[\begin{array}{l} \text{formal} = \text{be_missed}(\boxed{e2}, \boxed{Z}) \\ \text{agentive} = \text{miss_act}(\boxed{e1}, \boxed{X}, \boxed{Z}) \end{array} \right] \end{array} \right]$$

4.3 V2 の漂白化

「V1-テ-V2」の型の場合, V2 の意味が漂白化される場合がある.

- (23) a. 太郎は学校にお弁当を持って きた.
 b. 太郎は僕に電話して きた.
 c. 花子は震えて きた.

(23a) では「太郎が学校に行った」という物理的行動が含まれるが、(23b) では、太郎の物理的行動が含まれるか否かは曖昧である。後者の読みをするなら、V2 の意味が bleaching しているといえる。(23c) の場合には明らかに V2 が物理的行動が含まれているとは考えず、状態変化である。3 節で述べた Transition 構造の再構築として説明できる。

5 解析

以上の規則を基に、正しく述部構造が形成されるか実験を行った。本稿で用いたプログラムでは、解析結果は GL のフォーマットに基づく、句ごとの語彙構造とその要素同士を結びつける木構造として出力する。語彙構造は、語の意味情報を属性-属性値のマトリックス構造で表現するものである。

意味解析には様々な手法が採られてきたが、段階性と underspecification の原則を取り入れた Minimal Recursion Semantics (MRS) [13] のような研究が主流である。段階的に解析を行う手法によって、例外的な文法・語彙規則が解析結果に及ぼす悪影響を減らすことができる。本稿では、underspecification と default の破棄可能性の原理を取り入れており、意味が 1 つに決定できない単語に関しては、デフォルトの意味を与えておき、段階的に解析する過程で、デフォルトが文脈にふさわしくないと判断すれば、意味変化を与える規則をかけて別の意味を生成させる。デフォルトの意味は underspecified にしておき、処理上破棄できるような状態にしておく。

対象データは『天声人語』1994 年の 1 年分と CD-ROM 版『新潮文庫の 100 冊』(1995 年 新潮社) に収録の 5 作品から抽出した 500 文である。500 文とも人手により、本稿で扱っている現象を含んでいることを確認した。現在は単文レベルの処理を行っているため、単文に区切ってからプログラムで処理した。

図 2 は動的語彙モデルのシステムの構成図である。先ず、日本語入力文が形態素解析システムに渡され、得られた形態素解析結果から入力文語彙規則を生成する。形態素解析システムには『茶筌』[14] を使用している。

6 おわりに

実験の結果、72.8% の割合で正しいイベント構造を出力することができた。本稿では、GL の記述によって複雑な述語構造を形成する過程を説明した。今回のモデルでは、V1 と V2 の外項(主語)が一致する述部構造を扱ったが、今後は複数の predicate-argument

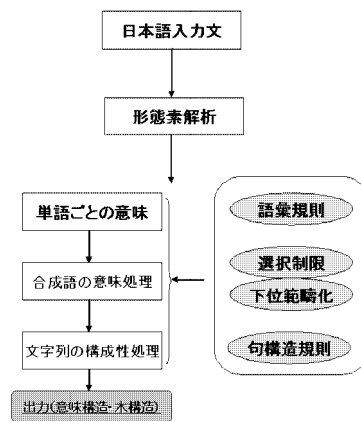


図 2: 動的語彙モデルの解析システムの構成

構造を持つと考えられる「花子は彼にぶたれた」(「花子」が「れる」, 「彼」が「ぶつ」) のような間接受動文を GL の枠組みで扱っていくよう発展させていく。

参考文献

- [1] Pustejovsky, James (1995) The Generative Lexicon. Cambridge, Mass.: MIT Press.
- [2] Pustejovsky, James (1992) The syntax of event structure. In: Beth Levin and Steven Pinker (eds.) Lexical and Conceptual Semantics, pp.47-81. Oxford: Blackwell.
- [3] Vendler, Zeno (1957) Verbs and times. Philosophical Review 66, pp.143-160.
- [4] Dowty, David R. (1979) Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ. Dordrecht: Reidel.
- [5] Shirai, Yasuhiro (2000) The semantics of the Japanese imperfective -teiru: An integrative approach. Journal of Pragmatics, 32, pp.327-361.
- [6] 影山太郎 (1993) 『文法と語形成』東京: ひつじ書房.
- [7] 影山太郎 (1999) 『形態論と意味』東京: くろしお出版.
- [8] 影山太郎・由本陽子 (1997) 『日英語比較選書 8 語形成と概念構造』東京: 研究社出版.
- [9] 松本曜 (1998) 「日本語の語彙的複合動詞における動詞の組み合わせ」『言語研究』114 号, pp.37-83.
- [10] Goldberg Adele E. (1991) It can't go down the chimney up: Paths and the English resultatives. Berkeley Linguistic Society, 17, pp.368-378.
- [11] Goldberg Adele E. (1995) Constructions: A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press.
- [12] 竹内孔一・内山清子・吉岡真治・影浦峯・小山照夫 (2000) 「文法的制約を用いた複合語解析モデルの作成」『学術情報センター紀要』12 号, pp.7-15.
- [13] Copestake, Ann, Dan Flickinger, Ivan Sag and Carl Pollard (1999) Minimal Recursion Semantics: An Introduction. Manuscript, Stanford University.
- [14] 松本裕治・北内啓・山下達雄・平野善隆・松田寛・高岡一馬・浅原正幸 (2003) 『日本語形態素解析システム『茶筌』version 2.3.3』奈良先端科学技術大学院大学.