

動詞語義の階層的分類に関する一考察

加藤 恒昭

東京大学大学院総合文化研究科

kato@boz.c.u-tokyo.ac.jp

1 はじめに

語義知識の構成において、語義（概念）をノードとし様々な意味関係をアークで関連づけたネットワークとすることが一般的である。意味関係の中で特に包摂関係（上位下位関係）は、ネットワークの背骨としての役割を果たし、いわゆる階層的シソーラスを構成する。包摂関係は「 x は y の一種である」のような説明がなされるが、名詞に対する概念の場合に比べ、動詞のそれは必ずしも明確ではない。

英語の語義知識である WordNet は動詞に対する包摂関係を「 x することはある仕方である y することである」と定義し、様態に着目して troponymy と呼んでいる [1] が、この関係によるネットワークは動詞語義をひとつの DAG にまとめない。一方で、日本語語義知識の代表である電子化辞書 (EDR) [8, 9] と述語項構造シソーラス [11, 12] は、動詞語義全体をひとつの DAG (述語項構造シソーラスでは木) にまとめあげて、階層的な分類を行っているが、後述のようにその構造は大きく異なる。また両者の構造におけるノードは概念ではあるが、特定の語と結びついた語義ではない。

動詞語義にはいくつかの分類の観点がありうるので、「正しい」階層的分類は一つではないだろうが、それを特徴付けたり、それが一貫しているかを検証するような分析は可能であると思われる。本稿では、このような分析に別の意味関係との関係を用いることを提案する。動詞語義においては様々な意味関係が存在するので、それらとの関係をみることの有効性が期待できる。本稿では、自動詞・他動詞の対立という意味関係に注目し、それと包摂関係との関係を観察し、動詞語義の階層的分類について考察した結果を報告する。分析対象としては、EDR と述語項構造シソーラスをとりあげた。分析のために作成した言語データについても説明し、このような分析の方法論の得失について検討する。

2 着眼

語義知識に限らず、オントロジ全般において包摂関係の設計は困難な課題である。その設計を検証するための方法論として OntoClean が提案されている [2]。これは概念を定義する属性の特徴を明らかにし、その特徴を用いて包摂関係となりうる概念を制約するものである。こ

れを通じて、部分全体関係やクラスインスタンス関係を包摂関係と誤ることを防げるとしている。包摂関係への制約というこの発想を語義知識に適用し、語義をつなぐ (包摂関係以外の) 意味関係をその制約として用いることを考えた。つまり、同じ意味関係にある語義の対は、同じ包摂関係のトポロジを持つべきであるという制約である。あわせて、ある意味関係がどのような包摂関係のトポロジと対応づけられるかでその語彙知識を特徴付けられると考える。動詞語義の場合、そのような意味関係として、語彙的「態」である自動詞・他動詞の対立 (自他対立) や可能態、自発態などの文法的態 (が語のレベルに現れたもの) [13]、反義関係があげられる。本稿では、この中で自動詞・他動詞の対立について検討している。

語彙概念構造 (LCS) の考え方に基づくると、自他対立、より厳密には非対格動詞と使役他動詞との関係は、以下のように表すことができる [3]。LCS では他動詞「開ける」の語義を

(x ACT-ON y) CAUSE (BECOME (y BE AT-open)) と表現する。これは「 x が y に働きかけることが、 y が AT-open という状態に変化することを引き起こす」ことを表している。ここで (x ACT-ON y) を上位事象、(BECOME (y BE AT-open)) を下位事象と呼ぶ。この他動詞語義の下位事象が、「開ける」と対立する自動詞「開く」の語義となる。対立する自動詞と他動詞のこのような意味関係に着目する。

このような意味の構造を考えた場合、動詞語義の分類のひとつの方針として、語義の構造全体に着目して、(x ACT-ON y) CAUSE ...) と (BECOME ...) との違いを最上位の分類とし、その後、それぞれを細分類していくことが考えられるが、前者 (他動詞語義) の細分類をその下位事象に従って行くとすれば、それは自動詞語義の表現であるので、その後の分類は両者で並行するはずである。もちろん、前者を上位事象の x , y に着目して細分類することもできるが、 y は下位事象にも含まれるし、 x の特徴はそれほど深い階層をもたらしなないと考えられるので、それらの場合も並行性が現れると予想される。一方で、このような対立は同じ事象においてそれを引き起こした動作主が表現されているか否かというたかだかひとつの素性の値の違いと考えることもできる。この場合、他動詞語義は下位事象に従って分類され、自動詞と他動詞の対立は階層的分類の葉に近いところで表現されると予想される。いずれの場合にせよ、対立する自動詞

と他動詞の分類の間には並行性が現れるはずであるし、べきであるということが着眼である。

3 言語データ

3.1 分析対象

電子化辞書 (EDR) と述語項構造シソーラスについて分析を行なった。

EDR は複数の辞書の集合体で、日本語英語各 20 万語に約 40 万の概念が単語辞書によって関係づけられ、概念辞書がそれらの概念を包摂関係に基づいて体系化し、階層的分類 (概念体系) を構成している。分析に用いたものは、この概念体系において、〈事象〉を根とした日本語動詞に関わる部分 [4] である。〈事象〉は最上位では、〈現象〉〈移動〉〈行為〉〈変化〉〈状態〉¹ に分類されている。

述語項構造シソーラスは動詞、形容詞、形容動詞という述語を網羅して概念体系を設計し、電子化辞書 Lexeed [7] の語義に対して、意味役割を含む項構造や意味分類を付与したものである。ここではその中の動詞を対象とする。その概念体系の設計では前節と同様の議論がなされ、上位事象と下位事象を持つ語義も下位事象だけからなる語義も下位事象に着目して分析するという方針が示されている [11]。最上位の分類は〈状態変化あり〉〈状態変化なし (活動)〉〈状態変化なし (状態)〉となっており、最大 7 階層の分類が行われている。

3.2 データ準備

形態素解析用辞書 IPADIC [6] に含まれる動詞について、自他対立を含む形態的意味的な関係を注釈した言語データを作成している [4]²。これを用いて、以下のふたつの言語データを作成した。

EDR 語義自他対立データ EDR 概念辞書に含まれる動詞のうち、上記の言語データに含まれ意味的な自他対立を持つとされたものを対象に語義レベルでの対応 (自他対立) を人手で判断した。対象となったのは動詞 450 対の語義、自動詞語義延べ 1803 件、他動詞語義延べ 1753 件、計 3,555 件である³。判断は日本語概念見出しと日本語概念説明に基づいて行い、相手側のひとつの語義と対応する (1 対 1 対応)、相手側の複数の語義をまとめたものと対応するような 1 対 n 等、1 対 1 でない対応がある (その他)、対応する語義がないがそのような語義が想定できる (想定可)、対応する語義が想定できない (想定不可)、語義が不明確などで判断ができない (無印)、に分類し、対応する語義がある場合は対応先を記述した。結果として 626 対の 1 対 1 対応の語義の対を得た。作業例を表 1 に示す。

¹概念名や概念見出しを (...) で示す。

²自他の対立についてはいわゆる自動詞他動詞だけではなく、「読む」「読ませる」のような他動詞とその使役態も含んでいる。

³EDR において概念は概念識別子で一意化される。異なる語が同じ概念を語義とする場合がある。

EDR-Lexeed 語義対応データ EDR 概念辞書と Lexeed (述語項構造シソーラス) とに共通して含まれる動詞のうち、上記の言語データに含まれ自他対立など形態的意味的な関係を持つとされたものについて、EDR と Lexeed で付与されている語義の同義性を人手で判断した。対象となった語は 1,291 語、EDR の語義は延べ 4,711 件、Lexeed の語義は 4,258 件である。判断は EDR の日本語概念見出しおよび日本語概念説明と Lexeed の語釈文とを照らし合わせることで行なった。相手側のひとつの語義と同義である (1 対 1 対応)、相手側の複数の語義をまとめたものと同義である等の対応がある (その他)、同義となる語義がない、に分類し、同義である語義がある場合はその語義を記述した。結果として 2,825 対の同義の語義の対を得た。作業例を表 2 に示す。

3.3 分析用データ

作成した言語データから、EDR において 1 対 1 の自他対立となっている語義の対 (EDR 語義自他対立データから得られる) で、その対をなす語義の両方が Lexeed の語義との間で 1 対 1 対応しているもの (EDR-Lexeed 語義対応データから得られる) を抜き出すことで、165 対の語義の対応を得た。これらの対がそれぞれの概念体系 (包摂関係に基づく階層的分類) でどのような位置関係にあるかを分析した。

4 分析と考察

4.1 EDR の分析

EDR の概念体系は多重継承を許し、これが分析を複雑にしている。分析の対象となった概念を上位に辿った際に、多重継承を経るものが、延べ数で 330 件中 100 件あり、そのうち、69 件が複数の最上位分類を持つ。

多くみられるのは、上位事象による分類と下位事象による分類の両方を包摂関係とするものである。例えば、〈具体物を移動する〉は〈空間移動〉と〈対物活動〉の 2 つを上位概念として持ち、それらは〈移動〉と〈行為〉に繋がっている。前者は下位事象、後者は上位事象に基づく分類であると考えられる。同様に変化や現象を引き起こす行為を〈変化〉や〈現象〉であり〈行為〉であるとしたり、状態の変化を〈状態〉であり〈変化〉であるとしている場合がある。分類を上位事象と下位事象のいずれに基づいて行うかは概念体系設計の判断であろうが、それが混在している場合、一般に包摂関係で成り立つと考えられている推移律が成り立たなくなることが問題となる。これに加えて EDR の設計方針と関連する多重継承が 2 種類ある。ひとつは多面性の扱いで、もうひとつは複合動詞の扱いである [8]。いずれも推移律を前提とすると適切とはいえない帰結を生む場合がある [5]。

自他対立する概念のトポロジについては、165 対中 26 対がそれらと同じ概念もしくは 1 段か 2 段上位の概念を共有するようになっている。その割合が比較的小さいことから、EDR は原則として上位事象による分類を行

表 1: 自他对立の対応づけ (「浮かぶ/浮かべる」の例)

対立語	語義	概念識別子	概念見出し-概念説明	対応	先
浮かぶ ウカブ (自)	I1	0e585a	浮かぶ [ウカ・ブ]- (出来事や現象が) 表面に現れる	想定不可	
	I2	1faf1a	浮かぶ [ウカ・ブ]- 面目が立つ		
	I3	3c2fe5	浮かぶ [ウカ・ブ]- 考えが心の中に湧く	その他	T1
	I4	3c2fe6	浮かぶ [ウカ・ブ]- 死者が成仏する	想定不可	
	I5	3d05ae	浮かぶ [ウカ・ブ]- 水面や空中で沈んだり落ちずにいる	1 対 1	T3
浮かべる ウカベル (他)	T1	3cf1f8	思い出す [オモイダ・ス]- 過去の出来事を思い出す	その他	I3
	T2	3d0435	表明する [ヒョウメイ・スル]- 心の中のものを出す	想定可	
	T3	3d05ad	浮かせる [ウカセ・ル]- 水面や空中で沈んだり落ちないようにする	1 対 1	I5

表 2: EDR/Lexeed 語義の対応づけ (「浮かべる」の例)

語	辞書	語義	概念識別子 語義番号	概念見出し-概念説明 語釈	対応	先
浮かべる ウカベル	EDR	E1	3cf1f8	思い出す [オモイダ・ス]- 過去の出来事を思い出す	その他	L3
		E2	3d0435	表明する [ヒョウメイ・スル]- 心の中のものを出す	1 対 1	L2
		E3	3d05ad	浮かせる [ウカセ・ル]- 水面や空中で沈んだり落ちないようにする	1 対 1	L1
	Lexeed	L1	2002900-1	水面や水中や空中などに浮くようにする	1 対 1	E3
		L2	2002900-2	表面に表す	1 対 1	E2
		L3	2002900-3	考えや幻想などを意識の中に上らせる。思い出す	その他	E1

なっていると推察される。そのような中で近くに配置される概念の対を見ると、「くじなどで金品などを得る」という [当たる/当てる]⁴, [腫れる/腫らす] など, 上位事象と下位事象という構造を想定すべきか迷うものも多いが, 「発行(される/する)」という [出る/出す] 等, その配置に疑問のあるものもある。

それ以外の 139 対について, 前述の多重継承を外部事象による分類と考えられるものを選択して分析した。他動詞語義の 133 件が〈行為〉に属する。これらについて, 対立する自動詞語義の最上位分類毎にどのような下位分類が対応しているかをまとめた。結果を表 3 に示す。〈行為〉はその下の 1 段または 2 段が上位事象の働きかけ先 (y) に基づく分類になっている。その下位の分類を目安に頻度の多いものをあげている。括弧内の数字が頻度である。一定の傾向はあるが, 自動詞語義が属する概念と他動詞語義が属する概念との関係はそれほど明確ではない。

4.2 述語項構造シソーラスの分析

下位事象に基づいて分類を行うという設計方針を反映して, 自他对立は一般に非常に近く, かつ系統だった分類になっていることが多い。165 件中 88 件が, 自他对立の語義を同じ概念とするか同じ概念を直接の上位として共有している。系統的にそれとは異なる設計をしているものとして, 〈心理的变化〉をその下位を〈感情変化〉と〈感情変化の引き起こし〉に分け, その下で怒りや驚きを分類している (6 件)。[怒る/怒らす], [驚く/驚かす] はこの構造で対比される。「行動を(起こす/起こさせる)」という [動く/動かす] と [振れる/振る] とい

⁴カギ括弧内の語釈文は Lexeed のものを要約して作成した。自他对立部分を (.../...) で示している。対立する語義の対はそれ付与された語を並べて [.../...] で示す。

う〈活動を引き起こす活動〉という他動詞概念は, 〈状態変化なし(活動)〉, 〈使役〉につながる〈ACT の引き起こし〉に属している。[傷つく/傷つける] や [温まる/温める] は, 〈状態変化有り〉の直接の下位で〈主体の変化〉と〈対象の変化〉に別れている (6 件)。

その他については, 系統的な分類は難しいが, 次で述べる方法論の問題に起因するものが多い。ただし, 他動詞概念を〈状態変化あり〉, 自動詞概念を〈状態変化なし(状態)〉としているもの, 「交差(する/させる)」という [交わる/交える] 等, 4 件は, 上位事象の分類が混入していることを疑わせる。

4.3 方法論の問題点

述語項構造シソーラスの分析において, 自他对立の関係を系統的に説明できない部分については, 自他对立という関係の問題点, それを前提とする言語データの問題があると考えられる。大きく 3 種類に分類できる。

ある事象は作用の連鎖という構造を持っており, 言語化はその一部を切り取ったものである [10]。自動詞はその事象の結果に注目し, 他動詞はその過程に注目する。この過程が結果を得ることと直接結びつかなかったり, それ以上の含意を持つことがある。例えば, 「隙間なく(満ちる/満たす)」という [詰まる/詰める] は, 〈充足〉という〈関係の変化〉と〈吸入〉という〈位置変化〉に位置づけられる。この〈関係の変化〉と〈位置変化〉の対比は他にも見られる。また, 「必要なもの(が/を)全部(備わる/備える)」という [揃う/揃える] は, 〈充足〉と〈準備〉に位置付けられている。このような相違は上位事象の分類の混入というより, 語義における注目点の違いの影響であろう。面白い例に「漬物に(なる/する)」という [漬かる/漬ける] がある。前者は〈科学的变化・反

表 3: 行為における自動詞語義と他動詞語義の対応

自動詞語義分類	他動詞語義分類 (〈行為〉の下位概念)
現象 (34)	〈変える〉→〈ものを対象とする行為〉→〈対象行為〉 (7)
	〈物の性質を変化させる行為〉→〈対物活動〉→〈対象行為〉 (5)
	〈物に力を加える行為〉→〈対物活動〉→〈対象行為〉 (3)
	〈情報の移動〉→〈ものを対象とする行為〉→〈対象行為〉 (3)
変化 (33)	〈変える〉→〈ものを対象とする行為〉→〈対象行為〉 (11)
	〈行為や物事の進行にかかわる活動〉 (7)
	〈覆う〉→〈対物活動〉→〈対象行為〉 (4)
	〈ものともとの関係にかかわる行為〉→〈対象行為〉 (3)
移動 (28)	〈具体物を移動する〉→〈対物活動〉→〈対象行為〉 (14)
	〈物に力を加える行為〉→〈対物活動〉→〈対象行為〉 (4)
	〈情報の移動〉→〈ものを対象とする行為〉→〈対象行為〉 (3)
状態 (21)	〈具体物を移動する〉→〈対物活動〉→〈対象行為〉 (3)
	〈情報の移動〉→〈ものを対象とする行為〉→〈対象行為〉 (3)
	〈行為や物事の進行にかかわる活動〉 (3)
	〈変える〉→〈ものを対象とする行為〉→〈対象行為〉 (3)
行為 (17)	〈害を与えること〉→〈対人行為〉→〈対象行為〉 (4)
	〈させる〉→〈対人行為〉→〈対象行為〉 (3)

応)であり、後者は〈付着〉という〈位置変化〉である。

動詞がとりうる項が意味分類に影響する。「横に(なる/する), 横に(寝る/寝かす)」という[横たわる/横たえる]は、この語釈文からは自他对立が判断されるが、典型的な用法では、ヒトが横たわり、モノを横たえる。この区別が語義の位置付けに影響する場合がある。また項がヒトである場合とモノ場合とを区別しないでひとつの語義となっている場合もあるが、それを区別しないと概念体系での配置ができない場合もある。

目的語の意味役割が対象でないため、あるいは、慣用的な用法から自他对立が典型的でない場合がある。「ものを通して向こう側(が見える/を見る)」という[透ける/透かす]、「時計の針(が/を)先を指すように(なる/する)」という[進む/進める]、「(温度が高く湿気がこもる/蒸されるようにする)」という[蒸れる/蒸らす]などが例となる。これらは、自他对立の定義の問題で、作成した自他对立データが不適切であり、そのための雑音であるとも考えることもできるが、形態的な自他对立があることは確かであるし、トートロジーに陥らず、「適切な」ものを選ぶのは困難である。

5 おわりに

自他对立の意味関係と照らし合わせることで、語彙知識の概念体系(包摂関係による階層的分類)を分析することを試みた。このような分析は、一方で自他对立の意味関係の広がりを感じさせてくれて、それが概念体系を一意に制約するものではないことが示されたが、それを通じて、包摂関係設計における判断をあらためて明らかにし、ある事象の捉え方(上位事象と下位事象への分解)、語義の分割基準等への示唆を与えるものであった。自他对立の意味関係が関連するのは概念体系の一部であるので、その他の意味関係との関係がその他の部分について同様の示唆を与えるかを検討していきたい。

謝辞 本研究は JSPS 科研費#17H01831 および #15K12873 の助成を受けた。ご支援に深く感謝する。

参考文献

- [1] Fellbaum, C. eds. WordNet: An Electronic Lexical Database. A Bradford Book, 1998.
- [2] Guano, N. and Welty, C. Evaluating Ontological Decisions with OntoClean, CACM, 45(2) pp. 61–65, 2002.
- [3] 影山太郎. 動詞意味論—言語と認知の接点—. くろしお出版, 1996.
- [4] 加藤恒昭, 林良彦, 伊藤たかね. 形態素解析用辞書における動詞の形態・意味関係. 情報処理学会 自然言語処理, 2012-NL-207-11, 2012.
- [5] 加藤恒昭, 林良彦. 日本語動詞に関する EDR 概念辞書の分析. 情報処理学会 自然言語処理 2015-NL-224-12, 2015.
- [6] 桑畑和佳子, 橋本三奈子, 村田賢一. 計算機用日本語辞書の開発. 情報処理学会 人文科学とコンピュータ, 18-4, pp. 27–34, 1993.
- [7] NTT コミュニケーション科学基礎研究所監修. 基本語 データベース. 学習研究社, 2008.
- [8] 荻野考野, 中尾由雄, 小笠原あゆみ, 長澤陽子. 日本電子化辞書研究所における概念体系. 情報処理 情報学基礎, 1993-FI-032-4, pp. 27–34, 1993.
- [9] 荻野考野, 中尾由雄, 小笠原あゆみ, 長澤陽子. EDR 電子化辞書における概念体系. 言語処理学会 第 1 回年次大会, pp. 197–200, 1995.
- [10] 大堀壽夫. 認知言語学. 東京大学出版会, 2002.
- [11] 竹内孔一, 乾健太郎, 藤田篤, 竹内奈央. 意味の包含関係に基づく動詞項構造の細分類. 言語処理学会 第 14 回年次大会, pp. 1037–1040, 2008.
- [12] 竹内孔一, 石原靖弘, 竹内奈央. 述語項構造のシソーラス分類と意味役割の設計について. 人工知能学会 第 27 回全国大会 2D4-OS-03a-1, 2013.
- [13] 寺村秀夫. 日本語のシンタクスと意味 I. くろしお出版, 1982.