

言語処理を意識した語彙概念構造の構築

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

インターネットの発展により音声・画像・言語といった様々なメディアによって情報が利用できるようになった。その中でも特に言語は意見や事実関係といった重要な情報ははっきり伝える媒体としてその重要度が増しており、コンピュータによる言語処理が期待されている。文の基本的な意味は述語(特に動詞¹)とそれを取りまく名詞によって構成されるため、表層の格を伴って現れる名詞と動詞の関係がどのような意味関係にあるかを捉えることは文の理解のために重要な手がかりとなる。

動詞と名詞との関係を記述する方法として文の構造(構文解析)を基にした項構造(argument structure)、語彙概念構造(lexical conceptual structure)、生成語彙(generative lexicon)、FramNetなどが提案されている。本研究では項構造を包含する語彙概念構造(LCS)に着目し日本語に対してLCS辞書と応用モデルの構築を行う。LCSは項同士の関係を少数の意味述語で関係付けて動詞を取りまく名詞がどのような関係に至るかまで記述できるため文の理解モデルの基礎言語資源として利用できる可能性が高い。また生成語彙やFrameNetに比べて記述の具体化に必要な分析が蓄積されてきている点が大きな理由である。LCSは意味記述であるが項構造を介して表層の構文情報とリンクされておりLCSの記述を行うことができると表層から意味構造へのインターフェースの役割を担う(図1参照)ことが期待される。

LCSは現在も発展し続けているが処理という観点からすべての動詞に対するLCSの体系をまず第一に明らかにしたい。そこでLCSで

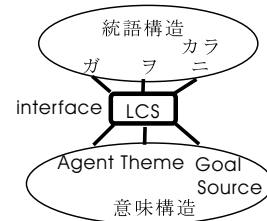


図 1: 語彙概念構造の言語処理における位置

採用されているアイデアを整理しなおしてなるべく簡易な形で多くの動詞に対してLCSを付与することを目標とする。こうして作成されたLCSをここではTLCSと呼ぶ。²日本語の漢語動詞を中心に約1300語についてTLCSの付与を行った。以下ではどのようにアイデアを整理して、何を記述したのかを記述する。付与の結果、全体としてのTLCS体系を明らかにすると同時に、TLCSからみてかなり異なる動詞のタイプが存在することが明らかになった。まず次の章ではTLCSはどのような動詞のタイプを予想して設定したかを明らかにする。次に実際動詞を付与して当てはまらなかった例を明らかにして動詞のタイプの拡張について説明する。

2 TLCSの構築

LCSは必須項とよばれる動詞が必要とする名詞に着目し、名詞同士の関係を意味述語と呼ばれる限定した関係子で構造的に結びつけることで動詞の意味を表現する方法である。意味述語をどう設定し、必須項をどう制約するかで多彩な動詞の意味を記述できる。理論上で議論されてきたLCSには様々なアイデアが盛り込まれているが、本研究で提案するTLCSには(1)

²TLCSの'T'は著者の名前から由来するが小さいという意味でtinyの意味も含んでいる。

¹サ変名詞を含む

項構造, (2) アスペクト分析, (3) 概念の意味記述の3つのアイデアを取り込んだ。後に示すようにこれら3つの見方だけでも数多くの TLCS タイプが明らかになり動詞のクラス分けを行うことができる。

TLCS を動詞に付与する場合, 対象とする動詞がどういう TLCS かを明らかにするには様々な例文を観察して名詞や副詞との共起から判断する必要がある。判断の材料として言い回しによるテストがある。上述のアイデアでは (1) 項構造に関するテストと (2) アスペクト分析に関するテストがいくつか利用できる。こうしたテストを参考に意味述語を組み合わせて (3) 概念の意味構造を表す。現段階では TLCS の全体系はわかっていないが用意した意味述語と典型的な動詞との関係である程度のパターンはわかっているのでそれを記述しておく。

以下では項構造, アスペクト分析, 概念の意味構造について具体的に説明し, どう仮定したのか, どれだけの事例を最初に設定して設計したかについて明らかにする。次に, こうした仮定を基にどう具体的な作業にもとづいて付与したかを明らかにする。これにより我々の構築した TLCS がどの程度の限界であるかを浮き彫りにする。

2.1 項構造

項構造は動詞がとる必須の名詞に対して意味役割を割り当てることでその動詞の振る舞いを記述する方法である。意味役割とは名詞の振る舞いに対する概念的な意味記述で Agent(動作主) Theme(対象) Goal(目的) Source(起点) など文の構造に即した分類が提案されている。理論的(Grimshaw[2]), 実践的(Dorr[1]) 研究からこうした意味役割は語によって定まっているため語彙の知識として記述する有効性が示されている。

項構造を付与するためには, 個々の動詞に対して (1) 意味役割の種類はいくつあるのか, (2) 動詞を取り巻く名詞においてどれが必須の名詞であるかの2点を明らかにする必要がある。

(1) に対しては意味役割のタイプをあらかじめ

め決めておく必要がある。本研究では必須項のみに注目しているのも最も基本的な Agent, Theme, Goal, Source の4タイプのみという限定した形から記述をはじめることにした。(2) は個別の動詞に対する判定を行う必要がある。動詞をとりまく名詞の格が必須格かどうか, さらにどの意味役割に対応するかを調べる完全な判定テストは存在しないが態の入れ換えによるテストを利用することで推測することができる。さらに格の入れ換えや品詞を利用して判定していくことが必要である。以下に実例を示そう。

他動詞においてヲ格が Theme に対応する場合としない場合がある。この場合能動態と受動態で比較すると差異が明らかになる。

1 その文を英語に翻訳した/その文が英語に翻訳された

2 道路を歩く/*道路が歩かれる³

1の例ではヲ格は Theme に対応するが2の例では対応していない。この場合のヲ格は TLCS では必須項ではない付加詞(adjunct)として扱う。

必須格かどうかで判断が難しいのはニ格である。ニ格は意味役割では Goal に対応する動詞が多いが, 一方で「6時に」「カチカチに」など明らかに副詞としての作用もあり区別を行いたい。現段階では「時間」と「オノマトペ」におけるニ格の場合必須項ではないと判断している。さらにニ格と意味役割との対応についてはまだすべての現象を観察していないので後の3章の分析で現段階の分析結果を報告したい。

表層格におけるカラ格は Source(意味役割)に対応するため必須格として扱うが, 一方で「6時から」「最初から」といった付加詞との区別が難しい。文の構造で生成的に「～から」が付いて意味を成す場合は語彙に対して情報を付与することは避けたい。そこでまだ検証段階であるがカラ格と他のガ, ヲ, ニ格への言い換え可能かどうかでカラ格が必須格であるかどうかを判断する [13]。以下に例を示そう。

3 風呂から上がる/風呂を上がる

4 東から風が吹く/*東に風が吹く

この3の例の場合は起点を表す意味役割を自動

³ 印は目標とする意味を成さないことをあらわす

詞「上がる」が持っているが他方 4 の「吹く」の場合はカラ格は付加詞と判定して TLCS には反映しないことにする。⁴

以上のように分析された結果、項構造と TLCS がどのように記述されるか示そう。

5 翻訳 (Agent (Theme (Goal)))

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

6 歩く (Theme)

[x MOVE]

丸括弧を使って記述されているのが項構造である。その下に TLCS を記述している。小文字の x, y, z は動詞がとる名詞を表し, CONTROL, BECOME などの大文字英単語が意味述語である。項構造の分析で用いられていた Agent, Theme.. という表記を用いる代わりに, TLCS の構造と x, y, z の位置関係で意味役割と結び付けておく。上記の 5 の例では x, y, z が Agent, Theme, Goal に対応しており, 表層格に対してガ格, ヲ格, ニ格にそれぞれ対応する。6 の場合では x は Theme に対応しており, 表層格ではガ格に対応する。意味述語のタイプや組み合わせについては 2.3 で説明する。

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ひとつであり, Vendler に代表される 4 分類 (状態動詞, 活動動詞, 到達動詞, 達成動詞) をはじめ日本語に対しても理論的 [10][12] または处理的 [6] な側面から分析されている。上記の 4 分類を簡単に説明すると以下のようになる。⁵

- 状態動詞は一切の変化や運動とは関係ない動詞 (例:維持, 位置)
- 活動動詞は運動がつづけられる動詞 (例:操作, 行列)
- 到達動詞は自らがある状態に変化する動詞 (例:移動, 飽和)

⁴TLCS の事例であげている「操作」「処理」「飽和」といったまったく性質の異なる動詞に対しても「右から/操作/処理/飽和する」という文を考えた場合うまく理解できず。つまり語の特異性ではなく文における要素(つまり付加詞)と考える。

⁵TLCS の動詞分類はこの下位分類に属する。

- 達成動詞はある動作主が対象を変化させる動詞 (例:編集, 処理)

項構造の分析により動詞のとる名詞の役割が明らかになった後, アスペクト分析を行うことによって名詞間の関係が明らかになる。これにより動詞の意味を名詞間の関係で記述することができる。⁶具体的に例を挙げて説明する。

項構造の分析で

- 編集 (Agent (Theme))

という 2 項をとることが分かったとする。次にアスペクト分析によって達成動詞であることが分かったとすると 主語 (Agent) が 目的語 (Theme) に対して変化をさせて結果 (編集したもの) を生成するという意味があることが確認できる。⁷

アスペクト分析を動詞に対して行うには様々な文を作成してどのように解釈できるかテストを行う。特に時間に対して対象が変化したのかしなかったのかが重要になるためそれを浮き立たせる表現を利用することが有効である。以下にその表現と分析例を示す。

- 「10分で」や「10分間」を付与して動作の終了特性を調べる
「10分で翻訳する」(翻訳した結果ができる→達成)
「10年間維持する/*10分で維持する」(変化しない→状態)
- 「である」を付与して状態変化した結果があるかどうか調べる
「*行列してある」(運動→活動)
「編集してある」(結果がある→達成)
- 「たくさん」を付与して何が多いかを調べる
「たくさん運転する」(運転動作がたくさん→活動)
「たくさん発生する」(発生した量がたくさん→到達)

基本的には上記のアプローチでアスペクト分析を行うが最終的には人間による解釈で判断する。

⁶こうした事実関係の記述による動詞の体系化は言語処理において大変有用である。

⁷TLCS では「編集」: [x CONTROL [BECOME [y BE AT z]]] である。

アスペクト分析にはさらに Pustejovsky[4], 影山 [12] や畠山 [9] が指摘しているように動詞のとり名詞によってアスペクトが異なるという特徴がある。

- 杖をにぎる/おにぎりをにぎる
- ピアノを弾く/ソナタを弾く

どちらの例も左側は活動動詞であるのに対して右側は達成動詞である。

- # ピアノを 10 分で弾く⁸
- ソナタを 10 分で弾く

こうしたアスペクトの違いは動詞に対する語義の違いに対応するのではないかと考えられる。つまり「にぎる」「弾く」に多義があり名詞との組み合わせでその意味が異なる。これは TLCS の項 x, y, z の定義域を定めるという方法で実現することができる。例えば弾くに対して

- 弾く (活動) [x ACT ON y]
- 弾く (達成) [x CONTROL [BECOME [y BE AT z]]]

としておき, y に対する定義域を「楽器」「曲」としてそれぞれ記述する。しかし名詞の階層的意味分類体系が必要になるので現在では定義域の記述は行っていないが記述として分けておくことは重要である。

アスペクト分析の結果から分析対象の動詞が持つ TLCS がどのような意味述語を持っているか推測することができる。状態動詞ならば BE, 活動動詞ならば ACT, 到達動詞ならば BECOME (または MOVE), 達成動詞ならば CONTROL と BECOME の両方を持つ。これらがどのような構造になるかは最後の概念の意味記述の作業で行うことになる。

2.3 概念の意味構造

概念の意味構造とは TLCS が意味述語と項の組み合わせで示す語の基本的な意味の骨組みのことである。上述の項構造, アスペクト分析を基に意味述語を組み合わせで動詞の意味と合致するもっとも簡易な構造の TLCS を構築する。

TLCS で現在用いている意味述語は影山の日

本語の動詞に対する分析結果 [11] [12] を参考にしている。意味述語をなるべく少なくすることで TLCS タイプ間の相関が浮き彫りになるように注意した。現段階では以下に示す 11 種類の意味述語で記述している。

- CONTROL, BECOME, MOVE, BE, AT,

TO, FROM, WITH, NOT, ACT, ON 各意味述語がどのように設定されているかは影山 [12] に譲るとしてこれら意味述語を組み合わせについて説明する。様々な組み合わせが考えられるが人間の物の見方に対する暗黙の決まった規則 (ピリヤードボールモデル) が認知言語学で分析されており [3] これに即した形で意味述語を組み合わせる。上述の Vendler の 4 分類もこの分析に当てはまるので異なる分析結果を一つの記述形態で表すことができるのが利点である。Vendler の分類における達成動詞の場合を取り上げて以下に示す。

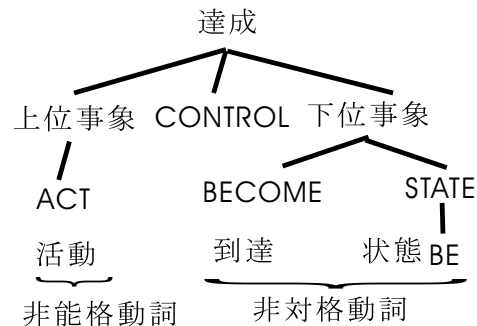


図 2: 概念構造の構築

達成動詞は上位事象の行為を通して下位事象で対象が変化してある状態になることを示している。活動, 到達, 状態も下位分類として位置づけられるので部分的な事象として意味述語を組み合わせで概念構造を構築する。

こうした外形の決定の他に項構造に関する以下の分析結果をこの構造の反映させる。

- 能格動詞の導入 (x=y)
- 定義域の設定 (FILLED)
- 非能格動詞と非対格動詞それぞれについて説明する。
 - 能格動詞とは表層のガ格とヲ格を入れ換えることができる動詞である (影山 [12] を参照)。

⁸# 印は完了の意味がはっきりしないことを示す。

- フランスが領土を拡大する / 領土が拡大する
- 領土拡大 / *フランス拡大

「拡大」という動詞の概念構造を考えた場合、態を入れ換えることなくこうした表現が可能なのは $x=Agent$ が $y=Theme$ と等しくなると考えられる (分析 [12] 参照). これを TLCS では $x=y$ という記号を用いて第 4 章の L5 のように示した.

b. 動詞がとる名詞の項 x, y, z に関する定義域の制約は行っていない. もし設定できれば詳細な意味を記述できる (2.2 節参照). しかし項が何か埋め込みで既に埋まっているかどうかを記述することは可能である. そこで FILLED という記号を利用してこれを表す. 例えば署名という動詞は「協定に署名する」のように記述する「名前」が既に動詞の中に取り込まれている. この意味構造を TLCS で表すと

- 署名: [x CONTROL [BECOME [[FILLED]y BE AT z]]]

となる. 項として名詞を取るかどうかを TLCS の構造に記述することができる.

c. 非能格動詞と非対格動詞は上述の図 2 に記述されているとおりに自然にアスペクト分析に取り込まれている概念であるが処理の観点から重要である. 非能格と非対格動詞はどちらも自動詞であるが前者は主語が Agent であり後者は Theme であることが大きく異なる. この違いを記述できることで複合名詞の語構成に対して分析が可能になる.

- (非能格) 客が行列する / *客行列
- (非対格) 状態が遷移する / 状態遷移

非対格動詞では表層のガ格が Theme にあたるため複合名詞を構成することができる. 一方, ガ格が Agent である場合は複合名詞を構成することはできない [12][14].

以上のアイデアを盛り込んで語彙概念構造を具体的に付与する. 第 4 章に約 1300 語の動詞に対して TLCS を付与した場合の全基本パターンを示す. 各 TLCS について個別の議論が存在しそれぞれを深めることも重要であるが, それよりも動詞の機能と意味に注目して TLCS という観点からどの程度のタイプが存在するのか

を明らかにすることを重視した. よって分析が未完であるタイプ (感情に関する動詞など) もあるがまず TLCS でカテゴリー分けを行なってパターンの可能性を明らかにすることを目標にした.

以下ではこれらの基本アイデアと付与法をもとに動詞を分析した結果分析が難しいいくつかの現象についてその問題点を明らかにする. さらに難しいながらも現段階での TLCS の付与案も提示する.

3 TLCS 付与からの知見

TLCS の付与は項構造の分析, アスペクト分析, 概念の意味構造の分析の順で行なわれる. 日本語の漢語動詞を中心に約 1300 語に対して付与を行なった結果さまざまな事例を観測することができた. ほとんどのものは設定した枠組み内の範囲で分析できたが, いくつかの動詞は大きく異なる分析結果を得た. 主なものを取り上げると次の 2 つである.

- 二格に関する問題
TLCS との対応で設定外の振る舞いをする動詞
- 概念的意味の特異性
項構造とアスペクト分析の結果から導き出される TLCS の意味が直感に反する動詞

以下ではこれらについて具体的に取り上げて何か問題でどう判断したかについて記述する. これを通して想定外の動詞に出会ったときの実践的な TLCS 付与法の例を示したい.

3.1 二格に関する問題

動詞を観測していくと二格の多様性から良く使われる動詞であっても上記で設定した枠組みを超える振る舞いをする動詞が見つかる. それらのうち以下の 3 つについて取り上げる.⁹

7. H が G に賛成 / 反対する
8. H が G に会う / ぶつかる
9. H が G に感動する / 悩む (感情動詞)

⁹最初の 2 例は寺村 [7] にも整理されている.

それぞれ項構造とアスペクト分析を行い TLCS を推定する。

7. 「賛成/反対」の分析

項構造の分析を行なう。態を入れ換えてみよう。

- 彼が親に反対する/親が彼に反対される
- 住民はその建築に反対する/その建築は住民に反対された

この分析からガ格、二格は必須項であることがわかる。また受身形があるので他動詞でガ格は Agent である。二格に関してはテストは無いが意味から推測すると Theme と判断できる。アスペクト分析では「である」や「10分で」との共起を考えても時間的な終了が感じられないため状態動詞か活動動詞かのどちらかである。この場合人の活動と判断すると結局以下のような TLCS を得る。

- 項構造: (Agent Theme)
- アスペクト分析: 活動動詞
- TLCS: [x ACT TO y]

この動詞が他の動詞と比べて大きく異なるのは他動詞であるにもかかわらずヲ格をとらずに二格を取る点である。上記の TLCS では [TO y] という構造を付与したが現段階ではこれを積極的にサポートする根拠を見つけていない。またこの構造からはなぜヲ格をとらないか理由は明確でない状況である。

8. 「会う/ぶつかる」の分析

項構造を分析してみよう。まず態を入れ換えや格を変えてテストしてみる。

- 彼が友達に会う/*友達が彼に会われる
- 彼が友達に会う/友達が彼に会う
- 彼が友達と会う/ぶつかる

受動態がないことから自動詞であることがわかる。ガ格と二格を入れ換えることは可能であるが意味として対等ではない。アスペクト分析を行うと「10分で」という表現と共起せず¹⁰また「である」とも共起しないので状態動詞か活動動詞である。「たくさん」との共起を考えると会った人やぶつかったものがたくさんあるという解

¹⁰「10分で会う/ぶつかる」は10分後にそれぞれの事象が起こり始めることをあらわしていて動作の完了の読みにはならないという意味。

釈になるので状態動詞と判断できる。よってこれらを総合すると TLCS は以下ようになる。

- 項構造: (Theme (Goal))
- アスペクト分析: 状態動詞
- TLCS: [y BE WITH z]

つまり「彼は友達と会う/といる」と同様の TLCS であることを示している。¹¹ただし依然この意味構造が積極的に支持されるテストなど見つかっていない。また「ぶつかる」と「会う」はその動作の起こる時間的な幅に違いが感じられるが TLCS では記述されていないなど検討の余地が残っている。

9. 「感動する/悩む」の分析

感情に関する動詞である。必須項の分析をするとガ格と二格をとり受身形がない。

- 彼が映画に感動する/*映画が感動される
- 彼が騒音に悩む/*騒音が悩まれる

またアスペクト分析を行うと「10分で」との共起や「である」との共起でははっきりしないが大石 [6] の分析から瞬間に変化する動詞 (到達動詞の細分類) と考えられる。よって非対格動詞である。項構造はガ格は Agent, 二格は Goal である。このまま TLCS を付与すると *[BECOME [y BE AT z] (y=「彼」が z=「映画」の状態に変化するという意味) となり意味が合わない。そこで英語との対応で考え直す。英語では使役形で表現する。

- 映画が彼を感動させた
- The movie moved him.

[x CONTROL [BECOME [y BE WITH [感動]z]]]

使役形で TLCS を考えると単純である。x=「映画」(Agent) が y=「彼」(Theme) を z=「感動の状態」(Goal) の状態に変化させる。この場合日本語では英語の受身に当たる形で表現されている。このような場合の語彙概念構造は影山 ([12]:118) にある BE WITH という形を使って以下のように TLCS で表現しておく。

¹¹ガ格が人であるにもかかわらず Theme であるということは問題ない。例えば「依存する」の場合「彼が会社に依存する」と人をとるがこの場合「彼」=Theme, 「会社」=Goal で TLCS は [BECOME [y BE AT z]] と非対格動詞である。

- 項構造: (Theme (Goal))
- アスペクト分析: 到達動詞
- TLCS: [BECOME [x BE WITH [[FILLED]y BE TO z]]]

TLCS の読みは x=「彼」(Theme)が WITH 以下の状況に変化した。その状況とは y=「感動する気持ち」が z=「映画」に対して (TO) ある (BE) である。これによって「感動」という意味述語のある側面を記述したがそれぞれの意味述語の存在や構造に対して直感以外のテストというものとは現在ではない。ただ他の TLCS と異なるという記述を重視して記しておく。

3.2 概念的意味の特異性

動詞の中で概念の意味構造が基本として設定した達成動詞の構造 (図 2 参照) と大きく異なるものがある。上述の心理動詞以外に、分析の結果から

10. 「教わる」「授かる」といった動詞

11. 要求動詞

という 2 タイプの動詞が存在することが明らかになった。これらについて以下に分析を行う。

10. 授かる動詞

項構造から分析してみよう。表層格を入れ換えたり態を入れ換えてみる。

- 太郎が先輩に / から 数学を教わった
- *先輩が太郎から 数学を教わられた
- 先輩が太郎に 数学を教えた

受身形がないことから「教わる」は自動詞である。アスペクト分析を行うと「10分で」との共起から動作の完了を読み取ることができる。これらの分析から動詞は非対格動詞で到達動詞であることがわかる。表層格をみるとガ格、ヲ格、ニ格がそろっており他動詞に見えるが意味から考えるとガ格は Agent に対応していない。そこで能動態「教える」の文と比較すると「先輩」から「太郎」に対して「数学」が伝わる通常他動詞となっている。よく観察すると「教わる」はこの反対 [7] になっている。

x 太郎 (Goal) <----- z 先輩 (Source)
y 数学 (Theme)

となる。つまり x=太郎はガ格であるが項構造では Goal の意味役割を果たし、z=先輩はカラ格/ニ格のどちらでも許される Source の意味役割を果たす。英語では必ず Source は from という前置詞を利用することに比べて対照的である。¹²さて、このような動詞の TLCS はどのように記述すると良いであろうか。影山 [5] の分析を参考に以下のように記述する。

- 項構造: (Goal (Theme (Source)))
- アスペクト分析: 到達動詞
- TLCS: [BECOME [z BE WITH [y MOVE FROM x TO z]]]

TLCS 内の z, y, x は順に Goal, Theme, Source を表しており y=「数学」が x=「先輩」から z=「太郎」に伝わることを示されている。¹³この構造もそれぞれの意味述語の存在を決定するテストが存在していないが意味分析は対応をとれているので現段階では以上のように分析する。

11. 要求動詞

項構造を調べるために態を入れ換えた表現と比較する。

- ファンが松井にサインを要求する
- 松井がサインを要求される
- サインが松井に要求される

ニ格、ヲ格も受身形をとることができる他動詞である。アスペクト分析を行うと「10分で」と共起した場合、要求そのものが完了するという意味にはならず「要求してある」のように要求が相手に伝わったという意味が認識できる。達成動詞の一種として扱う。この分析をもとに TLCS を構築すると

[x CONTROL [y MOVE TO z]]

となるがこの場合 x=「ファン」が y=「サイン」を z=「松井」に与えることになってしまい全く意味が合わない。ここで上述の「教わる」動詞の分析においてニ格は意味によっては Source の意味役割に対応することが分かった。この「要求する」の場合も同じくニ格は起点である (た

¹²寺村にもあるように Source であるにもかかわらずニ格という通常 Goal の意味役割を指す表層格と同じものが利用できるのは日本語と朝鮮語ぐらいで世界的にみても特異な現象である。

¹³能動態の「授ける」の TLCS は [x CONTROL [y MOVE TO z]] である。

だしカラ格とは言い換えられない). よって意味役割 Source に対応する意味述語 FROM を導入して以下のように TLCS を設定する.

- 項構造: (Agent (Theme (Source)))
- アスペクト分析: 達成動詞
- TLCS: [x CONTROL [y MOVE FROM z TO x]]

上記の意味に対応した TLCS を付与できた. しかしながらこうした要求動詞にはさらに複雑な事象を記述することができる.

- 監督が松井にトレーニングを要求した
- 監督が松井にトレーニングするように要求した

要求する内容がサ変名詞など動作を持つ場合, その動作主が二格で表されている. こうした場合の TLCS をどう記述すべきかまだ検討段階である. この場合先ほどの TLCS では不十分でありサ変名詞を取り込む機構が必要となる. さらに上例にあるように「～するように」と二格をとって要求内容をさすことができる. 必須格か付加詞かどうか記述すべきか考察の段階である.

以上で実際の付与での問題点と現段階での分析結果を明らかにした. 大きく振る舞いの異なる動詞をいくつか検出できたがまだ千数百語程度であるので未知の振る舞いをする動詞に出会うかもしれない. しかし見方は統一しているので上記に示したような同様のアプローチをとって TLCS を付与することになる.

4 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]y BE TO 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
- L17 [x CONTROL [x MOVE FROM z]] 要求, 求める
- L18 [x ACT TO y] 賛成, 反対

5 他の言語資源との対応

英語では既に Dorr が LCS を構築して Web 上で公開している.¹⁴ Dorr の LCS は主に翻訳処理を意識して構築されており 31 種類の意味役割, 時間, 場所に関する付加詞までも含めた体系になっている. また英語のほとんどすべての動詞について分析を行っている点も評価できる. 英語の LCS と日本語の LCS の体系はおそらく異なるため, 日本語の場合を構築することで理論的な部分を比較できる一方, LCS を利用した翻訳など応用も期待される.

他の言語資源として電子化されている辞書 EDR[15], IPAL[8] がある. 表層格と文の表現による分類の検討がなされており, こうした先行するデータとの対応をとることが様々な分析・処理の可能性をもたらすと考えている. 上述の Dorr の LCS では WordNet など語の意味辞書とのリンクも構築している.

¹⁴http://www.umiacs.umd.edu/~bonnie/LCS-Database_Documentation.html

6 TLCS の利用

構文解析から意味解釈を行う場合、表現の違いを頼りに意味を推測するので正確に分析された辞書が必要になる。TLCS は表層格へのリンキングルールを考慮して作成しているのでインターフェースとしてより詳細な意味記述との接点となる役割をすることが予測される。

また語の意味構造が重要になる処理として複合名詞の解析があげられる。複合名詞は格による指示がないため、より語の意味と機能が率直に意味に反映される場合が多い。こうした分析に対して TLCS を利用した応用が考えられている。

多くの動詞に対して概念構造を付与できれば動詞の種類について展望することができるので文の分析以外にも利用できるのではないかと考えている。

7 TLCS さらなる発展

本研究では動詞の必須項に限定したが付加詞に対する記述枠組みも処理の点から重要である。既に杉岡 [16] が理論的考察から提案しており将来の課題としたい。また名詞に対する定義域の設定を行なう。既に名詞に対する実践的な分類を持つ電子化辞書は多数存在する [15] [8]。よってこれらの意味分析をよく観察して記述の可能性を考えたい。

8 謝辞

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] Dorr, B.: Large-Scale Dictionary Construction for Foreign Language Tutoring and Interlingual Machine Translation, *Machine Translation*, Vol. 12, No. 4, pp.

- 271-325 (1997).
- [2] Grimshaw, J.: *Argument Structure*, MIT Press (1990).
- [3] Langacker, R.: *Foundations of Cognitive Linguistics, Vol.2 :Descriptive Application*, Stanford University Press (1991).
- [4] Pustejovsky, J.: *The Generative Lexicon*, MIT Press (1995).
- [5] 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).
- [6] 大石亨, 松本裕治: 格パターン分析に基づく動詞の語彙知識の獲得, 情報処理学会論文誌, Vol. 36, No. 11, pp. 2597-2610 (1995).
- [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] 情報処理振興事業協会技術センター情報処理振興事業協会 技術センター: 計算機用日本語動詞辞書 IPAL (1986).
- [9] 畠山真一, 坂本浩, 加藤恒昭, 伊藤たかね: 日本語動詞の LCS 推定に関して, 情報処理学会 2005-NL-165, pp. 1-8 (2005).
- [10] 工藤真由美: アスペクト・テンス体系とテキスト, ひつじ書房 (1995).
- [11] 影山太郎: 文法と語形成, ひつじ書房 (1993).
- [12] 影山太郎: 動詞意味論, くろしお出版 (1996).
- [13] 竹内孔一: 語彙概念構造による動詞辞書の作成, 第 10 回言語処理学会年次大会, pp576-579 (2004).
- [14] 竹内孔一, 内山清子, 吉岡真治, 影浦峯, 小山照夫: 語彙概念構造を利用した複合名詞内の係り関係の解析, 情報処理学会論文誌, Vol. 43, No. 5, pp. 1446-1456 (2002).
- [15] 日本電子化辞書研究所: EDR 電子化辞書使用説明書 (第 2 版) (1995).
- [16] 杉岡洋子: 動詞の意味構造と付加詞表現の投射, 研究報告 (2) 先端的言語理論の構築とその多角的な実証 (2-A), pp. 341-363 (1998). 神田外語大学.