

## 日本語フレームネット: 文意理解のためのコーパスアノテーション

小原 京子 (慶應義塾大学)

### 1. はじめに

コーパス、特に代表性と均衡性をもつ BCCWJ が資料としてのコーパスの利用を促進するだけでなく日本語研究のあり方にも影響を与えるようになってきている。これまではコーパスへのアノテーションと言えば自然言語処理システムで用いるためのタグ付けを指すことが多かった(『言語処理学事典』:6)。しかし日本語研究においても、コーパスを用例の収集元として用いるだけでなく、言語分析結果をコーパスへのアノテーションとして蓄積し、分析例として共有化したり、コーパスにアノテーションを行うことで言語分析の理論的枠組みの検証を行ったりということが可能になってきている。

本論文では、日本語フレームネットにおける意味アノテーションを紹介し、語彙の意味記述と文の意味記述のために必要なアノテーションの枠組みを検討する。日本語フレームネットは、フレーム意味論に則ってコーパスデータの意味分析を行い、その結果を意味アノテーションとして蓄積した言語資源である(小原 2012, Ohara 2012a, b)。フレーム意味論とは、フレームの概念を使って文やテキストの意味を理解するための枠組みである。

フレーム意味論に基づき BCCWJ に語義のアノテーションを行ったところ、そのアノテーション体系は係り受け関係にある語とその他の文要素との意味関係の記述に有効であることが分かった。しかしながら、係り受け関係以外の文の要素間の意味関係については十分に記述できないことも明らかとなった。そこで、フレーム意味論と親和性のある構文文法と呼ばれる理論に基づきアノテーション体系を拡張することを提案する。これにより、構文の意味もアノテーションできるようになり、文全体の意味が記述できる。

以下では、まず日本語フレームネットの意味アノテーションのバックボーンであるフレーム意味論の特徴を概説した後(第2節)、フレーム意味論に基づく語義アノテーションの有効性を示す(第3.1節)。次に、フレーム意味論に基づく語義アノテーションの限界について指摘(第3.2節)し、文の意味記述のためにアノテーション体系の拡張が必要であることを見る(第4節)。言語分析結果に基づきコーパスへのアノテーションを試みることにより、アノテーション体系の背後にある言語理論の妥当性を検証することがで

きる。

### 2. 文の意味記述のための枠組み: フレーム意味論

フレーム意味論とは、言語形式とその意味との関係を、背景的知識(フレーム)との関連で捉える意味論の枠組みである。1970年代に Fillmore が自らが提唱した格文法を発展させた理論的枠組みとして、言語使用や理解におけるフレームの重要性に着目して提唱したものである(Fillmore 1976, 『言語処理学事典』:272)。フレームを用いて語の意味を捉え、その語を含む文全体、文章全体の意味を理解しようとする。

私たちの知識は典型化された日常的な場面をもとに成り立っている(大堀 2002:36)。このような日常的な場面や出来事を理解する枠となる認知構造がフレームである。フレームは日常の経験がスキーマ化されて形成された、容易に想起可能な知識である。フレーム意味論では、言語形式がフレームを喚起することで、話者がそのフレームを想起すると捉える。つまり、フレームは話者が言語形式の意味を解釈する際に中心的な役割を果たす。

たとえば、「これと、これと、これ! 日本語でメニューを指差した。」や「モンゴル名物の馬乳酒も飲んでみたかったが、これは日持ちがしないとかでメニューにはなかった。」などの文においては、「メニュー」という語がレストランフレームを喚起する。その結果、話者はレストランでウェイターやウェイトレスの持ってくるメニューに従い食事を注文するフレームを想起することになる。ところが、「このOSには、よく使うメニューだけを優先的に表示し、あまり使わないメニューを一時的に非表示にする機能があります」や「表計算ソフトウェアならばこのようなめんどろな関数を入力しなくても、データ処理メニューの中に頻度分布を求めるサブメニューがあって便利だ。」などの文では、「メニュー」という語がコンピュータ画面上に表示される実行可能な機能の一覧を指している(『明鏡国語辞典』)。この場合、「メニュー」という語はレストランフレームではなく、コンピュータのソフトウェアフレームを喚起する。

個々のフレームを構成する意味的な要素のことをフレーム意味論では「フレーム要素」と呼ぶ。たとえば、レストランフレームのフレーム要素に

は、客、ウェイター、食事、コックなどが含まれる。フレーム要素は個々のフレームにおいて定義づけられる、フレームに依存したカテゴリーである。格文法における意味役割が有限個であり普遍的なものであるとみなされていたのとは対照的である。また、フレーム要素は個々のフレームによって定義づけられることから、他の理論的枠組みや言語資源で用いられる「意味役割」よりも粒度が高い。

たとえば、以下の各文の[ ]で囲んだ二つの名詞句に、有限個の普遍的な意味役割集合から適切な意味役割を割り当てるのは容易ではない。

- (1) a. 一つの王朝がたおれ、[新しい政権が]  
[それに]とってかわった。
- b. [真理が]たえず批判され、吟味され、[新しい真理に]とってかわられる。
- (2) a. [ヒトから][機械に]安易に置き換えようというは間違っている。
- b. 恋愛から抜けだすとき、人はしばしば  
[カップル用の習慣を][シングル用に]  
置き換えるだけだと勘違いしやすい。
- c. [この「神」を][「非日常」と]置き換えてみよ。

フレーム意味論では、動詞「とってかわる」と「置き換える」が、代替フレームを喚起すると分析する。代替フレームとは、以前ある役目を担っていたもののかわりにその地位に新しいものを据えることに関する背景知識である。代替フレームのフレーム要素としては、Patient, Theme, Contentなどの抽象的で普遍的な意味役割ではなく、このフレームに即した、動作主、役目、新、旧などが定義されている。その結果、(1)と(2)の各文の[ ]の名詞句のそれぞれには次のようにフレーム要素を割り当てることができる。

- (1') a. 一つの王朝がたおれ、[新しい政権が<sub>新</sub>]  
[それに<sub>旧</sub>]とってかわった。
- b. [真理が<sub>新</sub>]たえず批判され、吟味され、  
[新しい真理に<sub>旧</sub>]とってかわられる。
- (2') a. [ヒトから<sub>旧</sub>][機械に<sub>新</sub>]安易に置き換えようというは間違っている。
- b. 恋愛から抜けだすとき、人はしばしば  
[カップル用の習慣を<sub>旧</sub>]  
[シングル用に<sub>新</sub>]置き換えるだけだと勘違いしやすい。
- c. [この「神」を<sub>旧</sub>][「非日常」と<sub>新</sub>]置き換えてみよ。

また、フレーム意味論では動詞の意味分類に際し、格フレームではなく「結合価パターン」を用いる。格文法では個々の動詞の選択する深層格の集合体を動詞の格フレームと呼んだ(『言語処理学事典』:270)。日本語処理の分野でも格フレー

ムを用いて動詞を分類することが一般的である。ただし、個々の動詞の撮りうる表層格(助詞)のパターンをリストアップしたものを指すことが多い。

しかし、表層格の情報のみでは動詞の意味は記述できない。一例として格助詞「に」を伴う名詞句を取る動詞を見てみよう。通常格助詞「に」を伴う名詞句は、「九月末、チンギス・ハーンはいよいよホラズム国境の近くに着いた」のように意味的には終点を表すことが多い。以下の(3a)でも名詞句「太郎に」は動作の終点を表す。

- (3) a. 花子が[太郎に]チョコレートをあげた。

ところが、(3b)では、同じ名詞句「太郎に」は動作の終点ではなく起点を表している。

- (3) b. 花子が[太郎に]チョコレートをもらった。

このことは、表層格の情報のみでは動詞の意味を記述できないことを表している。フレーム意味論では、深層格(意味役割)のみから成る格フレームでもなく、表層格のみから成る格フレームでもなく、次のような4層から成る結合価パターンで動詞を意味分類する<sup>1</sup>。

- (4) フレーム要素  
文法機能(主語、目的語など)  
句タイプ(名詞句、副詞句など)  
助詞(表層格)

フレーム意味論の体系で(3a)と(3b)の、主動詞を中心とする文の意味を記述すると以下のようになる。まず、(3a)の「あげる」は贈与フレームを喚起する動詞である。贈与フレームとは、贈与者がある対象を贈与者から受け手に移すことに関する認知構造である。贈与フレームのフレーム要素としては、贈与者、受領者、対象などがある。(3a)の「太郎に」は受領者に相当する。

- (3') a. [花子が<sub>贈与者</sub>][太郎に<sub>受領者</sub>]  
[チョコレートを<sub>対象</sub>]あげた。

動詞「あげる」の結合価パターンには、{ [贈与者. 主語. 名詞句. が] [受領者. 間接目的語. 名詞句. に] [対象. 直接目的語. 名詞句. を] }を含めることになる。

次に、(3b)の「もらう」は受領フレームを喚起する。受領フレームは、受領者と贈与者との協同

<sup>1</sup> 英語フレームネットでは、フレーム要素、文法機能、句タイプから成る3層の結合価パターンを採用しているが、日本語フレームネットでは、さらに助詞(表層格)も含めた4層の結合価パターンを用いている。

の行為の結果、受領者がある対象を所有するに至ることに関する認知構造である。受領フレームのフレーム要素としては、贈与者、受領者、対象がある。(3a)とは異なり(3b)の「太郎に」は贈与者に相当する。

- (3') b. [花子が<sup>受領者</sup>] [太郎に<sup>贈与者</sup>]  
[チョコレート<sup>対象</sup>を] もらった。

そして、動詞「もらう」の結合価パターンとしては、{ [受領者. 主語. 名詞句. が] [贈与者. 間接目的語. 名詞句. に] [対象. 直接目的語. 名詞句. を] } を含めることになる。

このように、フレーム意味論では、動詞の意味が異なればその動詞にかかる名詞句が同じ格助詞を伴っていても異なる意味をもつことを、フレーム、フレーム要素、結合価パターンを使って記述できる

さらに、フレーム意味論では、動詞「あげる」と「もらう」の意味関係はフレーム間関係を用いて記述される。すなわち、「あげる」が喚起する贈与フレームと「もらう」が喚起する受領フレームはいずれも移行に関するフレームであるが、それぞれ移行に関して別個の視点を持つフレームである。

以上見てきたように、フレーム意味論は個々の言語形式が喚起するフレームがその言語形式を含む文全体、文章全体の理解や解釈にどう結び付くかを分析するための理論的枠組みである。従って、文の意味記述の目的に適った枠組みであるといえよう。次節では、フレーム意味論を用いた、日本語フレームネットにおけるコーパスへの語彙の意味アノテーションの実際を見ていく。

### 3. 日本語フレームネットにおける語彙の意味アノテーション

日本語フレームネット・プロジェクトでは、二つのアノテーションモードでBCCWJへのアノテーションを行っている。一つ目は、辞書編纂モードという、フレームごと、語ごとに、コーパスから例文を選びアノテーションを行っていくモードである。もう一つは、全文テキストアノテーションモードという、テキストに出現する自立語(フレームを喚起する語)すべてに対してアノテーションを行っていくモードである。

まず次節で、辞書編纂モードでのアノテーションにおいて明らかになった日本語フレームネットの意味アノテーションの有効性について述べる。次に、全文テキストアノテーションモードで発見された日本語フレームネットのアノテーション体系の限界について論じる。

#### 3.1 日本語フレームネットにおける語彙の意味アノテーション

日本語フレームネットでは、動詞のみならず、形容詞、形容動詞、副詞、名詞<sup>2</sup>のアノテーションも行っている。動詞のアノテーションでは、その動詞のいわゆる項にあたる名詞句のみならず、係り受け関係にあれば副詞句、動詞句なども対象にフレーム要素その他をアノテーションしている。次の(5)で、主動詞の「驚かす」は感情経験フレームを喚起すると分析できる。感情経験フレームとは、ある刺激が経験主に特定の感情を呼び覚ますことに関する認知構造である。この感情経験フレームのフレーム要素としては、刺激、経験主などが挙げられる。

- (5) それまで、貞子がダンスをするシーンなど想像したこともなかったため、貞子の踊りを間近で見て、遠山はかなり驚かされた。

この文で、刺激に相当するのは、「貞子の踊りを間近で見て」である<sup>3</sup>。しかしながら、「貞子の踊りを間近で見て」は、「驚かす」の項に相当する名詞句ではなく、項の外にある副詞句である。フレーム意味論に基づく日本語フレームネットの語意アノテーションの体系では、タグ付け対象は、動詞と項構造にある必要はない。従って、(5')のようにアノテーションすることができる。

- (5') それまで、貞子がダンスをするシーンなど想像したこともなかったため、[貞子の踊りを間近で見て<sup>刺激</sup>]、[遠山は<sup>経験主</sup>]かなり驚かされた。

また、日本語フレームネットの語意の意味(語意)アノテーションでは、文の構成素がフレームを喚起する語(フレーム喚起語、Frame Evoking Element, FEE; ターゲットともいう)と係り受けの関係にありさえすれば、中心的なフレーム要素(core frame elements, core FEs)に対応する構成素だけでなく、周位的なフレーム要素(core frame elements, core FEs)に対応する構成素もタグ付けしている。

以下の(5'')では、「かなり」が感情経験フレームの周位的フレーム要素である程度に相当している。

- (5'') それまで、貞子がダンスをするシーンなど想像したこともなかったため、[貞子の踊りを間近で見て<sup>刺激</sup>]、[遠山は<sup>経験主</sup>]かなり驚かされた。

<sup>2</sup> 名詞に付いては事態性名詞に限定してアノテーションを行っている。

<sup>3</sup> 感情経験フレームのフレーム要素・刺激に相当するのは「貞子の踊りを間近で見て」の一部である「貞子の踊り」である、とする分析も可能である。日本語フレームネットでは二次アノテーションとしてこのような分析もアノテーションに含めている。

[かなり<sub>程度</sub>]驚かされた。

### 3.2 日本語フレームネットにおける語彙の意味アノテーションの限界

以上見てきたように、日本語フレームネットの語彙の意味アノテーションは文の意味の記述に貢献している。しかし、日本語フレームネットの語意アノテーションで記述できるのは主にフレーム喚起語と係り受け関係（叙述関係、修飾関係、補文関係など）にある語の意味関係に限定される。全文テキストアノテーションモードでアノテーションを行うことにより、文の意味は係り受け関係以外のものも含むことが明らかになった。外の関係と内の関係の両方をもつ様々な言語形式が文の意味に貢献している。Fillmore et al. 2012 は、文の意味に貢献する言語形式として、1) 統語的イディオム、2) 時制、アスペクト、モダリティを表す言語形式、3) 意味論的、語用論的、対話的フレームを喚起する構文を挙げている (Fillmore et al. 2012:338-339)。

BCCWJ への全文テキストアノテーションにより、係り受け関係以外に文の意味に貢献する日本語構文としては以下が含まれることが分かった。(6)、(7)、(8)はそれぞれ上の 1)、2)、3)のタイプに対応している。

#### (6) 比較構文

これの方があれより長い。

#### (7) 伝聞構文

聞くところによるとルンバは素晴らしいらしい。

#### (8) 結果の残存構文

濃い霧がおりている。

### 4. 日本語フレームネットにおける構文の意味アノテーション

今現在日本語フレームネットは英語フレームネットと連動し、「構文」の意味も記述できるようにアノテーション体系を拡張中である。これについては別の機会にその詳細を論じる予定である (cf. Ohara 2012b)。

### 5. おわりに

本論文では、日本語フレームネットの意味アノ

テーションを紹介し、語彙と文の意味記述のためのアノテーションの枠組みを検討した。このように BCCWJ の整備により、日本語語処理のみならず日本語研究の立場からコーパスへのアノテーションを行ったり、言語学の枠組みに基づいたアノテーションの体系を構築したりということが可能になってきている。

### 謝辞

本発表における研究は一部科学研究費助成事業（学術研究助成金（基盤研究（C））による助成を受けた（課題番号 24520437）。日本語フレームネット構築には、国立国語研究所業務委託金による支援を受けた。

### 主要参考文献

- Fillmore, Charles J. 1976. "Frame semantics and the nature of language." In *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*. Vol. 280: 20-32.
- Fillmore, Charles J., Russell R. Lee-Goldman, and Russell Rhodes. 2012. "The FrameNet Constructicon" Boas, H.C. and Sag, I.A. (Eds.) *Sign-based Construction Grammar*.
- Fillmore, Charles J. and Collin Baker. 2010. A frames approach to semantic analysis. In Heine, Bernd and Heiko Narrog (Eds.) *The Oxford Handbook of Linguistic Analysis*. 313-339. Oxford University Press.
- 小原京子. 2012. 「日本語フレームネットにおけるコーパスデータの多義性分析」, 『言語処理学会第 18 回年次大会予稿集』.
- Ohara, Kyoko Hirose. 2012a. Semantic Annotations in Japanese FrameNet: Comparing frames in Japanese and English. LREC2012. <http://www.lrec-conf.org/proceedings/lrec2012/index.html>
- Ohara, Kyoko Hirose. 2012b. Japanese FrameNet: Toward Constructicon Building for Japanese. Plenary Lecture given at the Seventh International Conference on Construction Grammar (ICCG7). Hankuk University of Foreign Studies, Seoul, South Korea. August 11th, 2012..
- 言語処理学会編. 2009. 『言語処理学事典』 共立出版.