

# 動詞意味類型の曖昧性解消に向けた格フレーム情報との関連調査

岡田 正平 山本 和英

長岡技術科学大学 電気系

{okada, yamamoto}@jnlp.org

## 1 序論

日本語における用言は動詞と形容詞に分類される<sup>1</sup>。一般に動詞はウの韻で終わり、外的に観測可能な人や物の動きや変化を表すとされている。しかし動詞の中には「優れる」のような物事の特性を表す語や「むかつく」のような人の感情を表す語が存在し、形態的な分類と意味的な分類は一致していない。

そこで中山ら [1] は形態上の特徴に関係なく、その意味によって用言を「作用性用言」と「形状性用言」に分類することを提案した。そのうえで、分類の前段階として動詞、形容詞と作用性用言、形状性用言の中間概念と言える4種類の意味類型(動作, 変化, 感覚・感情, 形容)を定義し、人手でIPA品詞体系日本語辞書(1)(以下, IPA辞書)の動詞に付与した。この作業では動詞の持ち得る意味を落とすことの無いよう、無理に一意に決めようとせず、関係すると思われる意味類型全てを付与している。

動詞が多義性を持つ場合にも複数の意味類型が付与されることがあるが、この場合は前後の文脈から意味類型を決定できると考える。そこで本研究では多義性によって複数の意味類型が候補として付与された動詞に対して、文脈中での意味類型の推定を行うことを目的とする。本研究は、形態素解析の結果によって得られる情報のみを用いて推定を行うことを想定しており、形態素解析結果に直接情報を付加することができると考えている。形態素解析はあらゆる言語処理において基本となる処理であり、その波及効果は動詞・形容詞という品詞情報を用いる全ての処理に及ぶと考えられる。近年、自然言語処理技術の発達により、テキストマイニングツールなどの形で専門分野外の間でもその技術の恩恵に与ることができるようになってきた。そのような現状において、各種ツールの処理プロセス中に組み込むことが出来れば、専門知識を必要とせずにより意味に直結した情報を扱えるようになり、その

意義は大きい。

本稿では、予備実験として動詞の意味類型の違いによる格フレームの変化の観察を行った。

## 2 中山らの意味類型

中山らは用言の形態的な分類と意味的な分類が一致しない問題に対して、用言の新しい分類法として「作用性用言」と「形状性用言」を提案し、その分類における問題と課題を検討した。形状性用言と作用性用言はイの韻、ウの韻といった形態的な分類にはこだわらず純粋に意味によって分類される。例えば、動詞「優れる」は人や物の性質を表すため形状性用言となる。また、動詞「走る」は通常作用性用言であるが、「虫唾が走る」という1つの表現は形状性用言となる。

この分類を行う上で中山は「動詞については動作動詞や変化動詞など外的に現れる動きや、心の動きなど内的な活動で客観的に捉えることの出来ない表現が含まれている。さらには『優れる』など性質を表すものまで含まれているため、分類の前処理として各動詞が持つ意味をある程度まとめる必要がある」と述べたうえで、動作, 変化, 感覚・感情, 形容という4種類の意味類型を定義した。この意味類型は1動詞に対して1つの意味類型が必ず決定するものではなく、複数の意味類型をもつことを許容している。以下に各意味類型の定義を示す。

**動作** 客観的に観測者が捉えることのできる運動で、その運動が終了すると運動前の状態に戻り、結果状態を表さない動詞を意味類型の”動作”とする。例えば「泳ぐ」や「食べる」などが”動作”である。

**変化** 主体に現れる運動の結果状態を表す動詞を意味類型の”変化”とする。主体が意思を持たず結果のみを表し、かつ運動が終了しても運動前の状態には戻ら

<sup>1</sup>本稿では特に言及しない限り「形容詞」に形容動詞を含めて考える

ない表現である。例えば「乾く」や「死ぬ」などが”変化”である。

感覚・感情 目, 耳, 皮膚といった感覚器官の活動と, 頭脳や心の働きなどを表す知情意を意味類型の”感覚・感情”とする。定義では, 意思性のある感覚器官の活動を”動作”とし, 「見える」や「感じる」など非意志的なものを感覚とするとしている。また「愛する」など, 心の動きだけでなく客観的に観察可能な動きの意味も含んでいる動詞に対して, “感覚・感情”の意味類型だけでなく”動作”も持つと定義している。

形容 人や物の様子や性質, 形, 存在, 関係を表す表現を意味類型の”形容”とする。“形容”は一般的な形態上の分類である形容詞の集合と, 意味的な分類の上で同義になると考えられている。

”動作”または”変化”の動詞は作用性用言に, “形容”である動詞は形状性用言に漏れ無く対応する。

中山らは IPA 辞書中の各動詞 (12,648 表現) に対して人手で意味類型を付与した。この際, 動詞が持ち得る意味を落とすことの無いよう, 関係すると思われる意味類型全てを付与した。

### 3 意味類型の自動推定に向けた調査

中山らが意味類型を付与した動詞の中には, 動詞の多義性によって複数の意味類型が付与されたものがある。これらは, 前後の文脈が決定すれば意味類型も決定できると考える。例えば, 動詞「満たす」は”動作”と”形容”の意味類型が付与されているが「コップに水を満たす」という文中ではコップに水を注ぎ入れ満杯にするという運動を表すので意味類型は”動作”, 「条件を満たす」という文中では条件に適合するという状態や性質を表すので”形容”であると言える。そこで本研究では, 多義性によって複数の意味類型が候補として付与された動詞に対して, 文脈中での意味類型の推定を行うことを目的とする<sup>2</sup>。

本稿では意味類型推定のための予備調査として, 意味類型の違いによって動詞の文中での使われ方が異なるのかを調べる。そこで, 単一の意味類型を付与された動詞を対象に, 京都大学格フレーム (2) (以下, 格フレーム辞書) を用いて各々の意味類型が付与された動詞が取る表層格および項について調べた。格フレー

<sup>2</sup>ただし複数の意味類型が付与された理由のアンノテーションはされていない。そのため, あくまで多義性により複数意味類型が付与されたと思われる動詞について推定を行うものとする。

表 1: 単一の意味類型が付与された動詞数 (分母) および格フレーム辞書中に共通で出現する動詞数 (分子)

動作	変化	感覚・感情	形容
1516/6637	418/1530	360/1441	78/358

ム辞書は JUMAN 辞書に基づいており, IPA 辞書とでは単語の単位が異なるため IPA 辞書の動詞の表記と格フレーム辞書の述語 (動詞) の表記が一致したもののみを対象とした。この際, 態の違いは考慮していない。また, 格フレーム辞書中では複数形態素での表現も登録されているが, その場合は IPA 辞書の動詞が最初の形態素と一致したものを対象とした。単一の意味類型が付与された動詞数および, その中で格フレーム辞書にも出現した動詞数は表 1 の通りである<sup>3</sup>。

#### 3.1 表層格の頻度

意味類型ごとに取りやすい表層格に違いがあると仮定し, それぞれ格フレーム辞書中の 9 種の表層格の出現頻度を累計した。また, 複数の意味類型が付与されている動詞が文中においてどの意味類型になるかは一様な確率であると仮定し, 表層格ごとにどの意味類型の動詞に取られやすいかを集計した。ただし単一の意味類型が付与された動詞全てを対象に集計すると, “動作” が付与された動詞の絶対数が多いため偏りが生じてしまう。そこでランダムサンプリングにより意味類型ごとの動詞数を, 最も動詞数の少なかった”形容”と同じ 78 に揃えた上で表層格ごとの意味類型の出現頻度を累計した。

各意味類型の表層格の出現頻度を表 2 に, その内訳を図 1 に示す。サンプリングした動詞の表層格の出現頻度を表 3 に, 全動詞の場合とを比べた各意味類型の表層格出現頻度の相関係数を表 4 に示す。いずれの意味類型でも 0.9 以上の相関係数が得られたので, このサンプリングは妥当であると言える。サンプリング後の表層格ごとの動詞の意味類型の割合を図 2 に示す。

ガ格とト格を取る動詞は, ”形容” が付与されたものが 7 割超存在し大きな割合を占めている。”形容” が付与された動詞は, 主語となる名詞とセットでその名詞状態を説明する働きがあるためガ格を取りやすく, また文中で省略されにくい。一方でその他の意味類型が付与された動詞については, 主語が既に話題に挙がっ

<sup>3</sup>同じ語の漢字表記とひらがな表記に関して IPA 辞書では別のエントリであるのに対して, 格フレーム辞書中では同一のエントリにまとめられていることに注意されたい。共通の動詞の数が少ないように見えるのはこのことが大きな原因の一つである。

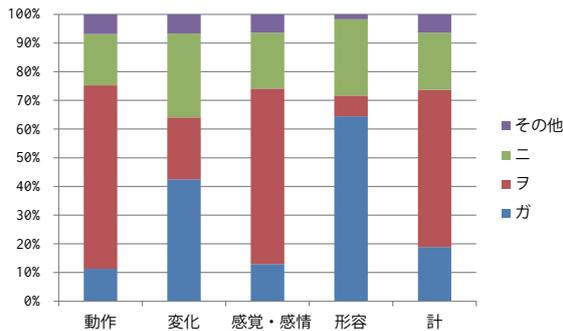


図 1: 動詞の意味類型ごとの動詞の取る表層格の割合

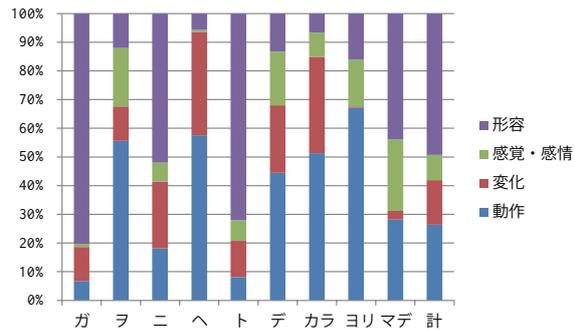


図 2: サンプル後の動詞に対する表層格ごとの取られる動詞の意味タイプの割合

表 2: 全動詞における意味類型ごとの総表層格数

動作	変化	感覚・感情	形容
2億 2664 万	3022 万	2278 万	2424 万

表 3: サンプル後の意味類型ごとの表層格数

動作	変化	感覚・感情	形容
1296 万	764 万	432 万	2424 万

ている場合に主語が省略される場合が多く、その他の情報を説明することが多いため、ガ格以外の表層格を取るのだと考える（例えば「私が食べる」というより「りんごを食べる」という用法の方が多い）。ト格については、「形容」以外の意味類型が付与された動詞については「ゆっくりと」などと言った副詞的な用法が多いが、「形容」が付与された動詞はそれに加えて「異なる」や「似る」のような動詞の比較対象を表すためによりト格を取りやすいのだと考える。

ヲ格、ヘ格、カラ格、ヨリ格については、「動作」が付与された動詞が 5 割超存在し、大きな割合を占めている。「動作」が付与された動詞は、動詞全体に対する割合が大きく多様な動詞を含んでいるため多くの表層格を取るのだと考える。これらの表層格の情報は、「形容」と「動作」の 2 つが付与された動詞の意味類型推定を行う場合に、ひとつの大きな手がかりになる。一方で、「変化」や「感覚・感情」は表層格ごとに割合の「変化」はあるが、全体で大きな割合を占める表層格は無かった。したがって、{ 形容 | 動作 } と { 変化 | 感覚・感情 } の間で判別が必要な場合には表層格の種類の情報には推定に貢献しにくいと考える。

表 4: 全動詞とサンプルされた動詞間での、意味類型ごとの表層格頻度の相関係数

動作	変化	感覚・感情	形容
.998	.942	.989	1.000

### 3.2 項の種類

表層格ごとにその意味類型を持つ動詞にしか取られない項の名詞意味属性がその意味類型を特徴付けると仮定し、IPA 辞書と格フレーム辞書に共通に出現した全動詞を対象に意味類型固有の表層格と名詞意味属性のペアを収集した。名詞意味属性には、日本語語彙大系 (3) の名詞・固有名詞意味属性（以下、まとめて名詞属性）を利用した。

IPA 辞書と格フレーム辞書に共通な動詞全てを対象にして、意味類型ごとに得られたその意味類型を持つ動詞にしか取られない項の名詞属性の数を表 5 に示す。また、得られた名詞属性中から著者 1 人の主観により、比較的抽象度の高いと思われるものの一部を以下に示す。

動作 カラ格:事象, 性質, 具体, ヘ格:精神, 具体, ト格:事象, 性質, 主体, ガ格:抽象, 国名, マデ格:事, 行為, 状態, 形状, 数量, 組織, 人間, デ格:表情, 性質, 衣料, ヨリ格:出来事, 地方名, 都市, ヲ格:事象

変化 ヘ格:感覚, 出来事, ト格:生命現象, マデ格:形, デ格:無生物, ヨリ格:食料

感覚・感情 ヨリ格:事, 行為, 道具

形容 ヘ格:宗教名, ト格:乗り物, マデ格:特徴, ヨリ格:形, 生物

表 5: 意味類型固有の表層格-名詞属性ペアの数

動作	変化	感覚・感情	形容
3243	196	198	137

これらの表層格-名詞属性のペアが文中に現れた場合は、意味類型推定のための決め手になると考える。実際には意味類型が4種類付与されていなければ、付与されている意味類型内で固有であればいいので、扱うことのできる情報はより多い。また、日本語語彙大系の名詞属性は木構造になっているため、得られた属性の親属性が意味類型に固有かどうかを調べていくことで、より多くの情報を得ることができる。

## 4 関連研究

本研究は、語義を4種類に規定した上での動詞の語義曖昧性解消タスクであるとみなすことができる。文中の動詞に語義を付与するシステムには、竹内ら [2] による意味役割付与システムが挙げられる。意味役割付与システムは入力文を述語項構造解析を行い、その後述語の語義の同定と項の意味役割付与を行う。言語資源を人手で構築することで、深い処理に向けた高級な情報を付与している。また本研究とは述語に付与する語義も異なっている。本研究では動詞自身の語義のみを扱っているのに対し、竹内らは動作対象の状態変化などを含むより広い範囲を扱っている。

また、動詞をその意味により分類することは、日本語教育の上でも意義がある。谷口 [3] によれば、日本語を正しく使いこなす上で動詞の種類についての理解が重要であるとし、その動詞の分類法の一つにアスペクトによる分類を挙げた。これは金田一 [4] が提唱した分類法であり、具体的には第二種の動詞（継続動詞）は中山の意味類型の”動作”と、第三種の動詞（瞬間動詞）は”変化”と、第一種の動詞（状態動詞）及び第四種（常に「テイル」を伴う動詞）は”形容”と相関が強いと考える。このことから、意味類型の自動推定を用いることで日本語の文に日本語学習者に向けた情報を付与できると考える。

## 5 まとめ

中山らは日本語の用言を意味的に分類する前段階として、4種類の意味類型を提案し、IPA 辞書の動詞について人手で付与を行った。本研究では、多義性によ

て複数の意味類型が付与された動詞について、文脈から意味類型の推定を行うことを目的とする。今回はその予備実験として、単一の意味類型が付与された動詞について、意味類型ごとにその動詞がとる表層格と、項の違いを調べた。表層格に関しては、ガ・ト格を取る動詞は意味類型の”形容”が付与された動詞が、ヲ・ヘ・カラ・ヨリ格を取る動詞は”動作”が付与された動詞が大きな割合を占めることがわかった。また、“変化”と”感覚・感情”については表層格以外の情報が必要であることがわかった。項については日本語語彙大系上の名詞属性を利用して、意味類型に固有の表層格-項の名詞属性の組を得ることができた。

今後の展望として、今回得られた結果から実際に自動推定器を作成し評価とフィードバックを繰り返しながら精度の向上を目指す。完成した推定器は公開する予定である。

## 使用した言語資源およびツール

- (1) IPA 品詞体系日本語辞書 . Ver.2.7.0 . <http://mecab.sourceforge.net/src/>
- (2) 河原大輔, 黒橋禎夫 . 京都大学格フレーム Ver 1.0 . 言語資源協会, 2009 . <http://www.gsk.or.jp/catalog/gsk2008-b>
- (3) 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 大山芳史, 林良彦 . 日本語語彙大系 CD-ROM 版 . 岩波書店, 1999 .

## 参考文献

- [1] 中山匠, 山本和英. 用言の新しい意味類型 — 作用性用言と形状性用言 — . 言語処理学会 第 17 回年次大会 発表論文集, pp. 560–563, 2011.
- [2] 竹内孔一, 土山傑, 守屋将人, 森安祐樹. 類似した動作や状況を検索するための意味役割及び動詞語義付与システムの構築. 電子情報通信学会技術研究報告. NLC, 言語理解とコミュニケーション, Vol. 109, No. 390, pp. 1–6, 2010.
- [3] 谷口秀治. 日本語教育における動詞の分類について. 大分大学留学生センター紀要 2 号, pp. 53–63, 2005.
- [4] 金田一春彦. 日本語動詞の一分類. 日本語動詞のアスペクト, pp. 5–26. むぎ書房, 1976 (原論文 1950) .