

メタファ分析に基づく置換可能な動詞カテゴリの作成

一ノ瀬 充[†] 竹内 孔一[†]

[†] 岡山大学大学院自然科学研究科
[†] {ganso, koichi}@cl.cs.okayama-u.ac.jp

動詞と名詞による表現において動詞の意味のずれにより生じるメタファが機械翻訳や要約などの自然言語処理や文章の理解において問題となる場合がある。そのメタファに内在する事実を扱うためには、前もって一般的な表現へと変換処理をする必要があると考えられる。例えば「風が遊ぶ」というメタファを「風が吹く」などの一般的な表現に言い換える処理が必要であろう。本稿ではこの処理に必要な動詞カテゴリの作成を提案する。実際に50の動詞について類似性に基づきカテゴリを作成し考察をすることで、置換可能な動詞間の類似性や共起する名詞の扱い方が動詞カテゴリの網羅性と信頼性の向上にどのように関わるかを明らかにする。

Toward Construction of Verb Thesaurus for Paraphrasing Metaphor Expressions on the Basis of Metaphor Analysis.

Mitsuru Ichinose[†] Koichi Takeuchi[†]

[†] Graduate School of Natural Science and Technology, Okayama University.
[†] {ganso, koichi}@cl.cs.okayama-u.ac.jp

Metaphor expression, especially consisting of a noun and verb, is irregular expression thus it is not easy to be dealt with in natural language applications such as machine translation, summarization, and document understanding. We believe it is important that metaphor expressions should be paraphrased into normalized basic expressions in order to extract a fact embedded in the expression: for example a metaphor expression “the wind plays” can be paraphrased into a basic expression “the wind blows”. In order to realize this task, in this paper we propose a verb thesaurus organized on the basis of semantic coherency for verbs. From the work of manual construction of a thesaurus on 50 verbs we discuss the possibility of our approach.

1 はじめに

最近の自然言語処理技術をもってしてもメタファを扱うことは容易では無い。例えば翻訳システムにおいて「風が吹く」と言う意味での「風邪が遊ぶ」や、「パソコン」が動くという意味での「パソコンが走る」という表現があった場合、もとの基本的意味である「風が吹く」や「パソコンが動く」という表現に直してから目的言語に変換していく必要があるであろう。こうした処理は自然言語処理の応用処理を行うための中間処理として必要ではなく、メタファを現実の言語処理システムで扱うための必須処理と考えられる。

本稿ではメタファのうちで特に名詞と動詞の組合せの中で動詞の意味のずれによるメタファについてもっとも基本的な表現に変換するシステムの構築法について議論する。このシステムを実現する処理の流れはメタファが否かを判断すること、適切な動詞に置き換えることの二つから成る。後者

のためには適切な動詞の選定が必要になり、既存の言語資源を用いるのみでは処理が困難なので新たな動詞の言語資源が必要となる。そこで本稿では動詞の置換によるメタファ50個から50の動詞を選出し、類語を収集することによる置換可能な動詞カテゴリの作成を提案する。

よって本稿は以下の構成をとる。我々の研究ではメタファを基本表現に置換することから言語処理における言い替え技術と密接な関連がある。そこでまず第2節で言い替え研究におけるメタファの扱いに付いて述べ、メタファ処理に必要な言語資源が足りないことを明らかにする。次に第3節で本研究であつかうメタファを明らかにしつつ、理論的アプローチ、ならびに既存の言語資源を分析し、本稿で目標とするメタファの言い替えには不十分であることを示す。こうした背景を踏まえて第4節で提案するカテゴリ作成の手順を示し、作成したカテゴリの一例を示す。第5節では第4節

で得られたカテゴリの質と量の向上に関わる知見について述べ、本研究の展望としてメタファ処理における動詞の置換処理について考察する。

2 関連研究

2.1 言い換え処理の必要性に関わる研究

本研究ではメタファ処理、すなわち言い換えを目標とした動詞カテゴリの作成を提案する。そこで本節では言い換えに関する研究を取り上げ、本研究との関連性を述べる。

まず置換や言い換えがいかなる分野において必要となるかを考えねばならない。乾ら [1] は言い換えが以下のような場面で論ずる必要性を帯びることを指摘しており、著者もその考えに準ずる立場をとる。

- 人間のための言い換え：平易な読み易い文の提供、ニュースの字幕などにおいてコンパクトな表現への言い換え
- 機械処理のための言い換え：機械翻訳の妥当性獲得、情報検索における質問拡張
- 言い換えそのものを研究の道具として：翻訳の正解例を複数作成したい場合に言い換えの技術が必要

本研究は「本来使われるべき意味の特定」ということで、上記において平易な文の提供、機械翻訳、情報検索の質問拡張に寄与するものと位置づけられる。

2.2 語義文ベースと類似度ベースの言い換え研究

実際に言い換えを試みる研究としては、語義文ベースのものと類似度ベースのものがある。言い換えの可否をいかなる観点で行っているかという点で本研究の参考とする。

語義文ベース

語義文ベースの言い換え研究は具体的には国語辞書の語義文を言い換えとみなすものであり、例えば鍛冶ら [2] は語義文と言い換え対象の文との格フレームを対応付け、不必要な格を選定することで同等句のみを抜き出している。つまり格フレームに着眼し類似度を計っている。乾ら [3] の研究では語義文の表現の重なりから単語間の意味の差を評価している。その差から制約を設けることによる良質な言い換えの生成を確認している。この手法では辞書の語釈文中の語に着目しているという点で言い換え対象の語から連想される語に着目

していると言える。

類似度ベース

類似度ベースによる言い換え研究としては Hindle[4] の単語の共起傾向に基づき類似性を測った研究がある。これは名詞に対しての研究で主語-動詞-目的語の組における共起傾向から名詞間の類似度を測る研究をしている。崔ら [5] は EDR 日本語概念体系から概念間の距離を算出し、類似度を測定に用いている。

3 メタファ

本研究ではメタファを扱うので、本節で本稿におけるメタファの形式やメタファと非メタファの境界について述べる。

3.1 本研究で扱うメタファ

メタファとは比喩の一種である隠喩のことであり、Lakoff & Johnson[6] はメタファを「概念の写像」と説明している。比喩には他にも提喩²や換喩³と呼ばれるものが存在するが、明確な区別が難しいのでここでは一括りにメタファとして扱う。

以下にメタファの代表例として「擬人化」と「有形化」を取り上げ、その性質と本稿における立場を合わせて説明する。

3.1.1 擬人化

「さるかに合戦」での「臼」のように物があたかも人間かのように意志を持って振舞っているかのような表現を擬人化メタファと呼ぶ。「風が遊ぶ」や「パソコンが走る」などがあり、それぞれ「風が吹く」「パソコンが動作する」などに情緒等の性質を付加した表現である。擬人化メタファは風やパソコンの例のようにどの語に置換して良いか分かりやすい例が多数存在し類語を予測しやすいメタファであると言える。

3.1.2 有形化

「感情」など実際には形の無いものを有形物とみなす表現を本稿では有形化メタファと呼ぶことにする。例えば「嬉しさが爆発する」や「構想を練る」などがある。大石 [7] は感情のメタファを分類しており、「感情は水である」などの有形化のメ

² 表現する事柄をそれと関係の深い付属物などで代用して表現する比喩。メトニミとも呼ばれる e.g. 「パッハが好きだ」⇒「パッハの曲が好きだ」

³ 上位概念で下位概念を表したり、逆に下位概念で上位概念に置き換えたりする比喩。換喩との違いは包含する関係にあるか否か。シネクドキとも呼ばれる e.g. 「警察に捕まる」⇒「パトカーにつかまる」

タファを扱っている。「水である」とみなすことで「感情」に「流れるもの、注がれるものである」という水の性質が付与され、「愛情を注ぐ」などの表現が可能になると説明している。本研究ではその「感情は水では無いはずである」という背景から、先述の「愛情を注ぐ」をメタファであると判断し、「水では無いものと共起する動詞」を見つけて置換処理を行う。

但し有形化のメタファでは「どの動詞に置換して良いか分からない」ものが多数存在する。「愛情を注ぐ」であれば類語として「与える」や「作る」が考えられるが、例えば「構想を練る」という表現から適した類語を想起することは難しい。一例として「構想を作る」が挙げられるが「構想を練る」に比べてむしろ不自然に感じられてしまう可能性がある。それは「練る」に対して「作る」が多様な共起名詞を許す抽象的な名詞であるからだと考えられる。本稿では不自然と感じられてしまうとしても「抽象的動詞」がどのような動詞と置換可能なかを確認することもあると有益であると考え、有形化のメタファも対象としカテゴリを作成する。

3.2 メタファと非メタファの境界

本節ではメタファか否かの判断をする際の境界について述べる。本稿においてメタファを意味のずれのある表現と定義した。その定義から境界を判断するためにまず「ずれの無い状態の意味＝基本的な意味」を定める必要がある。例として「風が遊ぶ」と「風が吹く」の二つの表現を考えると「風が吹く」の方が自然であり基本的な意味であると感じられる。また、「風」の代わりに「突風」や「木枯らし」を用いても基本的な意味と感ずることは可能であろう。

この例から動詞には理想とする共起名詞群が存在すると考えられ、この理想とする共起名詞群(以後名詞類語群と呼ぶこととする)が動詞の基本的な意味を定義すると考えられる。また動詞が多義である場合には、名詞類語群が複数存在すると言える。つまりある表現がメタファか非メタファかは表現内の名詞がある語義の動詞が求める名詞類語群に含まれるかによって判断出来るということである。

ここで動詞の語義全てが決定的な境界線を持つわけでは無いという点に留意せねばならない。国語辞書における「走る」の語義を例にとれば「犬

が走る」ことと「乗り物が走る」ことを分けているが、両者の差異は違う語義と捉えるべきか否かが非常に曖昧な差異であろう。すなわち「犬」と「乗り物」を「走る」の観点から同じ名詞類語群に存在させるべきかが非常に曖昧であると考えられるであろう。黒田[8]はメタファと非メタファの区別は連続的であるが、その区別は確かに存在すると述べている。著者も同様にメタファの区別は連続的であると考え、動詞の語義もまた連続的であると考え。それはメタファは新たな語義の一つと考えられるからである。この動詞の語義が連続的であることと、動詞の語義が名詞類語群で決まることから、名詞類語群の境界は不明瞭であり厳密で客観的な規定を行うことが困難であることを示唆していると言える。そこで本稿では著者がメタファと感じられるものはメタファとして扱うこととする。すなわち名詞類語群を著者が仮定することとする。但し主観性の過多が問題となり得るので、著者の名詞類語群感覚と EDR 電子化辞書の概念体系との差分を検分する実験を行った。

3.2.1 EDR における名詞体系を利用した名詞の類語感覚の検分

EDR 単語辞書と EDR 概念体系辞書(以下合わせて EDR)を組み合わせて、ある二つの単語間の概念距離を測ることが出来る。本節では名詞の概念距離を類似度とし、著者の考える名詞類語群の感覚と比較し評価する。具体的には以下に着目する。

- 同じ名詞類語群内の名詞間の距離が近い
e.g. 「バスが走る」「電車が走る」
- 異なる名詞類語群内の名詞間の距離が遠い
e.g. 「犬が走る」「稲妻が走る」

上記二点を念頭に、表 1 に著者の語義に対する感覚と EDR 上での概念距離との差分を記載する。ここで距離は双方の複数語義を組み合わせた中で最短のものである。

この結果から、正例として「犬」と「稲妻」は「走る」の異なる語義を生起させ距離も 10 と「電車」「バス」の「走る」の同じ語義を生起させる語間の距離の 6 よりも遠い事例が挙げられる。しかし負例として異なる語義を生じさせる「電源」「スイッチ」の距離 2 よりも同じ語義を商事させる「大根」「布地」の距離 6 の方が近い事例も挙げられる。負例が存在するということから、著者の語義決定

表 1: 語義を決定する名詞間の距離検分の表

比較する語	動詞	著者判断の語義の同異	EDR の概念体系上での距離
犬, 稲妻	走る	異	10
大根, 布地	切る	同	9
バス, 電車	走る	同	6
電源, スイッチ	切る	異	2

と EDR との間にはズレが生じている場合があることを確認出来る。

3.3 他資源のメタファへの対応

参考として動詞を扱っている他の言語資源がメタファ処理にどれほど対応出来るかを示す。各言語資源における置換の可否は各言語資源の動詞の記述体裁に合わせて以下のように考えて判断をした。

- **EDR 電子化辞書**：単語辞書と概念体系辞書を用いて単語間の概念距離を測定しそれを判断の指標とする。明確な閾値は設けていないが、距離が 0 から 5 の間であれば置換可能とする
- **分類語彙表**：“瞬く：活動, 生活, 口・鼻・目の動作”のように 3 段階に動詞を階層化して分類しており、3 階層全てが一致していれば置換可能とする。仮に 2 段階目（「瞬く」では「生活」）までの一致で置換の可否を決めると置換可能な動詞が多くなり過ぎてしまう
- **VerbNet**：動詞の語義毎に類語が与えられている。その類語の中に属していれば置換可能とする
- **日本語語彙大系**：36 の意味属性で動詞を分類しているが「走る：物理的移動」のようにかなり抽象的に分けている。その抽象的分類で一致を見ても置換可能動詞が過多となるので、語義が完全に一致した場合のみを置換可能とする
- **WordNet**：4.2.2 の体系を成しており、その中のいずれかに対象語が属していれば置換可能とする

上記の判断基準から三つのメタファ例から置換の可否と、参考として言語資源の規模を合わせて表 2 に示す。

表 2 の結果より、WordNet が言い換えの観点においてかなり網羅性の高い体系を為していることが確認出来た。但し英語資源なので日本語メタファの処理に直接の適用は出来ず、もし用いるならば

日英の対応付けという難しい中間処理が必要となるので、本研究と直接関係するということは示唆しない。

4 カテゴリの作成

本節ではカテゴリ作成の手順を説明し、実際に作成したカテゴリを明示する。

4.1 手順

本研究では動詞における意味のずれによるメタファを 50 例収集、その中の動詞 50 個についてそれぞれ類語を収集することで動詞のカテゴリを作成した。ここではそのカテゴリ生成の手順の詳細を述べる。なお、動詞の多義性を本稿では考慮するので一つの語が複数のカテゴリに属する場合もある。

4.2 カテゴリの形式

本論では一つの動詞が一つのカテゴリを形成するという定義のもとに類語収集を行う。つまり「ある語の類語の類語」が「ある語の類語」であることを保証しないということである。具体的には（財産を）「築く」の類語として（財産を）「得る」があったとしても、（意見を）「得る」の類語である（意見を）「伺う」と（財産を）「築く」が類語で無いと見なすということである。この例では括弧内に共起名詞を示しており、「築くと得る」では「財産」が共起するのに対し「得ると築く」では「意見」であり、各々の共起名詞が異なるので動詞が類語で無いのも当然であると考えられる。しかし 3.2.1 での記載のように、ここでの表現において「財産」と「意見」がそれぞれ異なる名詞類語群に属する可否かを客観的に判別することは困難である。こういった名詞類語群を判別する困難性を考慮して、本論では一つの語に対して一つのカテゴリを作成するという立場をとる。

4.2.1 類語とは

動詞の類語を集める前にまず類語とは何かを定義する必要がある。本稿では類語を「ある名詞と

表 2: 他の言語資源のメタファ対応

言語資源	規模	(パソコンが)「走る (run)」と「動く (move)」	(星が)「瞬く (twinkle)」と「輝く (brighten)」	(球を)「運ぶ (carry)」と「打つ (hit)」
EDR	27 万語	○	×	○
分類語彙表	95811 語	×	×	×
VerbNet	237 語	×	×	×
日本語語彙大系	30 万語	×	×	×
WordNet	約 15 万語	○	○	○

の共起において、置換してもある観点で意味を保持させられる動詞」と定義する。ここで置換の可否を考える上で乾ら [3] の「二つの言語表現の意味の差分を同定すること」「差分が無視出来るかを判断すること」の考えを踏襲する。

まずメタファによる意味の差分について考える。メタファは 3.1 のように「概念間の写像」により成立しているの、で「概念間の差分」こそがメタファによる差分である。それは例えば擬人化における「意志性の有無」であったり有形化における「実体としての形の有無」のことを指す。

続けてその「概念間の差分は無視出来るのかどうか」について考える。本研究では「厳密に等価である置き換え」を目指すのではなく**共通する性質が存在するか**、を判定することで分類を行う立場をとる。ゆえに「概念間の差分が無視出来るか」は「概念間の差分で共通する性質が完全に無くなってしまうことがないか」という問題に言い直すことが出来る。以上のことから本研究で置換可能、すなわち類語として収集される語に必要なのは「メタファにおいて生じるずれを考慮してもなお共通する性質が存在すること」であると言える。

共通する性質という枠で考えると膨大な数の類語が収集されてしまいそうに思われるが、例えば「ボールを右中間に運んだ」という表現を「何かが起きた」という極めて抽象的な表現に置換してもある観点では正しいであろうし、具体的なままに言い換えたとして「ボールを右中間に打った」といってもやはり違和感は無いらあろう。これらの例を踏まえて本稿では**伝わる意味に少しでも共通性が見出せる動詞**を類語と見なすことにする。

上記を踏まえて類語収集を行うにあたり、まず類語の種類、つまり差分を差し引いて残る共通の性質について考える際に英語で動詞を分類した資源である WordNet[9]が参考になると考えられるの

で説明する。

4.2.2 WordNet における動詞体系

英語の概念辞書として WordNet が存在する。動詞も分類されており、特筆すべきは様々な観点を定義した上で分類を行っていることである。具体的には以下の観点から分類を行っている。

- **上位語 (下位語)**: X が Y の行動の一種なら Y は X の上位語 e.g. X: 旅行, Y: 移動
- **troponym**: Y が X を行う際の様態なら Y は X の troponym e.g. X: 片言で話す (lisp), Y: 話す (talk)
- **entailment**: X しているとき必然的に Y しているなら Y は X に entail されている e.g. X: いびきをかく, Y: 眠る

動詞間の関係を厳密に定義することは極めて困難である。例えば上記定義においても「移動」が「旅行」の troponym であると考えられなくもない。本稿では「言い換えられるか否か」のレベルで語の関係を考えるので、Wordnet の定義の通りに分類するではなく、三つのいずれかに当てはまれば類語として収集することにする。ここで、言い換えを行う際に上記では例えば「走る」と「疲れる」を類語とみなすことが出来ないと考えられるので、これを本稿では「因果関係」として扱い類語と考える。以下に詳しく説明する。

4.2.3 因果関係

因果とはある事象が過去に起こった事象に起因することである。いわゆる時系列における関係であって、例えば「走る」と「疲れる」の関係が該当する。果たして「彼が走る」を「彼が疲れる」と言い換えることを許し「走る」と「疲れる」を類語と見なすべきかについては疑問が生じるが、現実にはそういった例が存在することを踏まえ、本研究ではこの因果関係による類語も対象とする。因果による語の関係を類語として認可することがためら

われることは「走れば必ず疲れる」わけではなく、時としてそうなる、あるいは「疲れるための方法として走るという手段がある」という偶発性に起因すると考えられる。4.2.2における「entailment」が非常に近い関係であるがそちらは「必然的である関係」であるのに対し、この因果関係は「偶発的である関係」であってその点が相違する。「風が吹けば桶屋が儲かる」のように関係の希薄なものまで因果関係として扱ってしまうと類語の数が甚大になってしまうことが最大の問題点となってしまうが、その点は次節「類語収集の流れ」で述べる「参考資源の有効活用」などにより可能な限り一般的な因果のみを扱いたいと考える。資源を参考にすることで例えば日常会話で使用される程度の汎用性を保てることが期待出来る。

4.3 類語収集手段と工程

4.2.2と4.2.3を踏まえて著者が類語か否かの判定して収集する。ここで存在する動詞の数と作業コストを考えれば自動化すべきであるが、自動化は困難である。自動的な手段の一例として格毎の共起情報を利用する手段が考えられるので、その難しさについて述べる。例えば「パソコンが走る」から類語収集を行う際に「パソコン」を共起語として考えれば「パソコンが動く」「パソコンが壊れる」などの表現が得られ「動く」「壊れる」がそれぞれ類語候補として得られるが「壊れる」は類語として不適格であろう。そこで「パソコンが高速に動く」など共起情報を増やすことで、「動く」のみを取り出せる可能性があるが、その共起情報を増やすということは共起頻度が落ちることになり、多くの類語を網羅することが困難であると考えられる。

この困難性も考慮し本稿では類語の判断は著者が行う。その際に次節の言語資源を参考にする。

4.3.1 参考とする言語資源

本研究では黒橋らの格フレーム検索⁴、そしてWeb上のサービスであるgoo英訳・和訳・国語辞書⁵を参考として類語を収集する。格フレーム検索は動詞を指定すれば、格関係毎の共起情報を調べることが出来るので、類推の助けとして用いる。今回は特にとる名詞の種類が豊富と考えられる「が格」と「を格」について考えることとする。goo辞

⁴ <http://reed.kuee.kyoto-u.ac.jp/cf-search/>

⁵ <http://dictionary.goo.ne.jp/>

書では、調べたい日本語動詞を一度英訳し、そこで得られた英語動詞を再度和訳することにより多様な日本語動詞を得られるので、その中から類語らしいものを選び出すという目的で利用する。

4.3.2 収集工程

収集について、格フレーム検索を用いる手法と辞書の往復による手法の二つについて、それぞれの工程を説明する。

● 格フレーム検索を用いる手法

1. ある一つの動詞に関して格フレーム検索より“が格対象名詞”、“を格対象名詞”を収集
2. 収集した名詞毎に格フレーム検索から共起動詞を獲得し、類語かどうかの判断を行い収集

● 辞書の往復による手法

1. ある一つの動詞を英訳し、複数英語動詞を獲得
2. 各英語動詞を和訳し、複数の日本語動詞を獲得。その中から類語を収集

4.4 対象語

青空文庫と毎日新聞コーパスより収集した動詞における意味のずれによるメタファ50例から動詞50個を選出し、それぞれについて類語を収集した。

4.5 作成したカテゴリの例

以下に作成したカテゴリを表3に示す。5つの動詞について、カテゴリ内の動詞(類語)を名詞・格関係と共に記述したものである。

5 作成カテゴリからの知見と展望

本節では動詞のカテゴリを作成して得られた知見と、目標であるメタファ処理にカテゴリをどのように用いるかについて展望を述べる。

5.1 類語の性質による大別

本研究では4.2.2や4.2.3を念頭に類語の収集を行った。表3のようなカテゴリを作成した結果、類語における性質を3つに大別出来ると考えられ、以下に記述する。

1. 同じ名詞類語群との共起を許容するが意味・ニュアンスに若干の違いのあるもの
「亀裂が走る」「亀裂が伸びる」「亀裂が貫く」
収集での着眼点：一つの名詞について共起する動詞を観察すること

表 3: 動詞カテゴリ作成例

観察したメタファと動詞	分類例
パソコンが走る	通る(道路を), 続く(線路が), 抜ける(電気が), 貫く(電気が), 巡る(電気が), 伝わる(指令が), 偏る(食生活が美食に), 伸びる(ストレートが), 動く(車が), 流れる(川が), 急ぐ(道を), 通う(登校路を), 行く(人間が), 運ぶ(運送トラックが), 風が, 香りを, 行う(計画を), 進む(道を), 去る((来た)道を), 感動する(感動が走る), 操縦する(ヨットを), 飛ぶ(鳥が), 歩く(道を), 尽力する(企業が), 滑る(スキーが)
布団を直す	戻す(設定を), 修正する・修繕する(アドレスを), 正す(姿勢を), 止める(癖を), 調整する(裾を), に対策する(ヘコミを), 畳む(布団を), 努力する(音痴を), 修理する(屋根を), 改心する(根性を直す), 回復させる(症状を), 変える(言葉遣いを), 改める(姿勢を),
時代を読む	見る(ホームページを), 予想する(行間を), 予測する(時代を), 読み込む(ファイルを), 理解する(本を), 解釈する(ファイルを), 知る(本を), 勉強する(本を読む), 話す(本を), 暗誦する(本を)
基礎を築く	作る(関係を), 固める(基礎を), 確保する(位置を), 得る(財産を), 建てる(城を), 組む(建物を), 頼られる(信頼を)
システムを破る	倒す(軍を), 勝つ(軍を), 超える(壁を), に背く(法を), 覆す(常識を), 抜ける(右中間を), 開ける(ドアを), 崩す(均衡を), 壊す(ドアを), 割る(ガラスを), に沿わない(常識を), 解く(封印を), 突破する(囲みを), 解明する(システムを), 乱す(平和を), 止める(禁煙を), 裂く(地図を), 分裂させる(地図を), 挫折させる(夢を), 犯す(掟を)

2. 同じ名詞類語群との共起を許容しないが、意味・ニュアンスがほぼ同等のもの

「人を育てる、建物を作り上げる」「人が走る、風が吹く」

収集での着眼点：異なる名詞について共起する動詞を観察すること。例えば「育てる」の類語を選ぶとしたらここでの「人」とは異なる性質を持つ名詞を選び共起動詞を観察する。「人」の「生物」という特徴がここで作用しているとするならば、「生物」で無い性質の例えば「無生物」を持つ名詞を考えれば一例として「建物」が考えられる。但し「育てる」との共起での「人」において「生物」という性質が作用していることを検出することは容易では無い。

3. 因果性による類語

「ページを捲る、ページを読む」「彼が走る、彼が疲れる」

収集での着眼点：同じ性質の名詞でも違う性質の名詞でも出現する可能性があるので、想起出来る限りのものを収集するしか無い、最も網羅の難しい類語である

上記性質に着目することで、例えば「走る」に関して類語を収集しようとする際に、1から「伸びる、貫く」、2から「吹く」、3から「疲れる」など

の動詞をそれぞれ収集出来る可能性が高まることが期待できるように、類語収集全体としての網羅性も向上を期待出来る。

5.2 メタファ処理への展望

メタファ処理として言い換えを実現するために名詞類語群の規定を自動化する展望と、本研究の手法で作成したカテゴリをいかにして利用するか展望について述べる。

5.2.1 名詞類語群の規定を自動化する展望

本論では、EDR 電子化辞書など既存の言語資源をそのまま用いて名詞類語群を規定することは困難なので、名詞類語群を著者の感覚で仮定し動詞のカテゴリ作成を行った。しかし作業コストなどを考慮すると将来的には既存の言語資源にメタファの作用における特徴や類語間の特徴(5.1)を活用するなどして名詞類語群を自動で規定したい。前者の一例を挙げると、有形化のメタファは「有形か無形か」、擬人化のメタファは「意志性」という特徴が作用しており、これら特徴は既存資源で名詞の体系として明確に記述されてある場合が多いので高い精度で判別出来ることを期待出来る。これら二つのメタファとは異なる性質のメタファも「意志性」のような特徴を明らかに出来れば、それを既存資源に適用させることで処理を行える可能

性がある。

5.2.2 カテゴリ内からの適切な語の選択

3.2で述べたようにメタファ、非メタファの判断指標は共起する名詞が基本的な意味を規定する名詞類語群に含まれるか否かである。すなわち置き換えるべき動詞は同じカテゴリ内に存在し、メタファ内の名詞を含む名詞類語群の対象である動詞となる。名詞類語群を考える上で本論での作成カテゴリでは、格情報毎に名詞類語群を考える必要があるので動詞の類語に格情報と名詞も合わせて記述している。この動詞の類語と合わせて記述した名詞から名詞類語群を仮定し、動詞を選択する方法として考慮している5.2.1で展望とした名詞類語群規定の自動化手法と合わせて考えていく必要がある。

6 まとめ

本研究では50のメタファ事例を分析し、50個の動詞について置換可能な動詞カテゴリを作成した。既存の言語資源のみによる動詞のメタファ処理は困難であり、そのためには動詞の性質毎に名詞類語群を考え動詞を分類する必要があることを明らかにした。

本研究での類語収集は手動であり人的コストを考慮すると、今後は可能な限り自動化する手段を考える必要がある。5.2.1のように名詞類語群を自動で規定することが出来れば類語か否かの判断にかかる人的コストを多に削減出来る。また、竹内⁶は同時クラスタリングにより名詞と動詞の互いの共起情報を組み合わせることによる分類の研究を行っており、その結果を本稿の動詞分類の参考に出来る可能性がある。これらを参考にしながら類語収集の質を向上させると同時に5.2で述べたメタファ処理における置換について考えていきたい。

参考文献

- [1] 乾健太郎, 藤田篤. 言い換え技術に関する研究動向. 自然言語処理, Vol. 11, No. 5, pp. 151-198, 2004.
- [2] Nobuhiro Kaji, Daisuke Kawahara, Sadao Kurohashi, and Satoshi Sato. Verb paraphrase based on case frame alignment. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (ACL2002), pp. 215-222, 2002.
- [3] 藤田篤, 乾健太郎. 語釈文を利用した普通名詞の同概念語への言い換え. 言語処理学会第7回年次大会, pp. 331-334, 2001.
- [4] Hindle and D. Noun classification from predicate argument structures. In *Proceedings of the 28th Annual Meeting of ACL*, pp. 1268-1275, 1990.
- [5] 崔進, 小松英二, 安原宏. EDR 電子化辞書を用いた単語類似度計算法. 自然言語処理研究会 93-1, pp. 1-6, 1993.
- [6] George Lakoff and Mark Johnson. Metaphors we live by. In *University of Chicago Press*, 1980.
- [7] 大石亨. 日英対照研究に向けた感情のメタファーの分類. 言語処理学会第13回年次大会, pp. 1034-1037, 2007.
- [8] 黒田航. 比喩理解におけるフレーム的知識の重要性: 比喩表現の程度の差を明示できる比喩の記述モデルの提案. メタファー研究の最前線. ひつじ書房, pp. 217-238, 2007.
- [9] George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J. Miller. Introduction towardnet. In *An online lexical database, International Journal of Lexicography*, 3(4), p. 235, 1990.

⁶ 「グラフ構造に基づく同時クラスタリングを利用した動詞の属性クラスの抽出」, 本冊子に記載