

日本語の慣用的表現について

— 語の非標準的用法からのアプローチ —

首藤公昭, 吉村賢治, 武内美津乃, 津田健蔵

福岡大学 工学部

広範な言語現象を対象として自然言語の文解析を考える際、種々のルール化を難しくする原因の一つは語の標準的用法と非標準的用法との境界があいまいで、どこまでをルール化の対象とし、どこからを例外的処理にゆだねるかの見通しが立てにくい点である。この問題を克服するには標準的用法の方から攻める通常の方法に加えて、非標準的と思われる用法を網羅的に収集・整理したうえで双方向から境界づけを試みる必要があるであろう。本稿ではこのような観点から筆者らがしている日本語大量データに基づく慣用的表現の調査・収集の現時点での結果を概観し、これらの表現の変化形の可否をどの様に整理するかについて論じる。

On the Idiomatic Expressions in Japanese Language

— An Approach Through the Close and Extensive
Investigation of Non-Standard Usage of Words —

Kosho SHUDO, Kenji YOSHIMURA, Mitsuno TAKEUCHI and Kenzo TSUDA

Department of Electronics, Fukuoka University

8-19-1, Nanakuma, Jonan-ku, Fukuoka-shi

814-01 JAPAN

Whenever we try to make rules such as "case frames" for verbs or "thesaurus" for the selectional restriction of nouns for the natural language understanding we immediately encounter the difficulty of fixing the boundary between the standard and non-standard (exceptional) usage of words. Authors believe one basically important way to remedy this problem is to make close and extensive investigation of the usage of words which is (or seems to be) non-standard, prior to the design of the rule systems.

In this paper, first, the results of the extensive investigation of Japanese non-standard expressions and the like, including so-called "idioms" is overviewed.

Second, the problem of the treatment of the variants of these expressions is discussed.

1. はじめに

自然言語文の解析を目差した研究で基本的に重要なことは簡単に言えば、(1) 語の共起の仕組みを明らかにしておくことと、(2) 語の共起に伴って語の意味から全体の意味がいかに統合・生成されるかのルールを明らかにしておくことである。そして、これらのルール化を行う際、語の持つ「属性」あるいは属性を共有するものを一体化して扱う「類」の概念が重要であり、具体的にどの様な属性情報あるいは分類情報を定めておくかが大きな問題となる。この問題は広範な言語現象を対象とする場合、単にルールや意味の記述および操作をどう形式化するか、どの様なソフトウェア・システムやツールで実現するかといった副次的な問題に比べ、より本質的である。従って、この問題に取り組むことは自然言語処理において不可欠であるが、これを難しくしている原因の一つは言語現象には個別的(例外的)現象が多く、どこまでを属性や類概念でルール化できるかの見通しが立てにくい事である。たとえば、述語と共起する格要素を格フレームの形で規定したり、そのため名詞のシソーラスを作成したりする場合、語のどの様な用法までを標準的なものとしてこれらの体系に乗せるかは作業上もただちに直面する切実な問題である。従来、これらのルール化は、出来得る所までをまず試みた上で、カバーしきれない現象は例外的に対処するというアプローチが一般的であったと考えられるが、筆者らは、(1)、(2)のルール化の難しいと思われる表現をできるだけ網羅的に収集・整理しておくという逆からのアプローチも上記の問題を軽減する上で必要であろうと考えている。この様な観点から筆者らは語の非標準的用法を特定の語との共起を捉える立場で収集しているが、本稿ではこれまでに収集したこの種の表現の概要を示し、関連する問題点について述べる。ここで対象とする表現は、国語学で従来論じられてきた慣用句や機能動詞結合⁽⁴⁾を包含した、より広範な固定的共起表現であり、ここではこれらを慣用的表現と呼ぶことにする。

2. 基本的表現の収集

実際の科学技術文データ約10万文、および市販の国語辞典4種中に現れる文表現すべてを対象として、慣用句、機能動詞結合などと呼ばれている現象を含めた語の固い結合と思われる表現を網羅的に収集した。収集に際しては疑わしきものはすべて収集するという方針をとった。これらの表現は語の結合が特に強く全体を単語的に捉えられると判断できるものとそうでないものに大別できる。以下、前者を一語性表現、後者を多語性表現と呼び、その概要を述べる。

2.1 一語性表現

語が隣接して強く結合しており、他語の挿入や語の交換が通常許されない表現を収集した。これらは構成語の意味から全体の意味を合成することが出来ないと思われるもので、処理上単語として一括して取り扱うことになる。これらは大きく、自立語的なものと付属語的なものに分かれる。

2.1.1 自立語的表現

名詞的	1700	根くらべ、不幸中の幸い、負け犬、負い目、骨折り損、袋の鼠、どんぐり眼、物のはずみ、目の毒、門前払い、尻拭い、…
動詞的	1000	はじき出す、青ざめる、思い当る、思いつく、煮つめる、言い張る、言い逃れる、一癖ある、一息入れる、…
形容詞的	200	何でもない、泣くに泣けない、目も当てられない、途方もない、…
形容動詞的	450	温室育ち、ごきげん斜め、藪から棒、しらみつぶし、これみよがし、…
副詞的	350	暇にあかせて、目に見えて、取るものも取りあえず、折にふれ、驚いたことに、着のみ着のまま、言うに事かいて、臆面もなく、…

連 体 詞 的	50	史上まれにみる、名にし負う、 音に聞く、持って生まれた、…
接 続 詞 的	10	さもなくば、従って、 しかしながら、…

2. 1. 2 付属語的表現

付属語的に扱ひ得る表現については、筆者らによって以前から収集・整理が行われており、拡張文節モデルの語彙レベルの資源としてまとめられている⁽⁶⁾⁽⁷⁾。これらの表現は大きく、助詞的なものと助動詞的なものに分けられ、それぞれ、関係表現、助述表現と呼んでいる。

（助 関詞 係の 表 現 ）	500	だけではなく、に基づいて、 について、によって、における、 に対して、にも拘らず、 のように、でさえも、 などといった、…
（助 動 詞 表 現 ）	2500	なければならない、かもしれない、 に違いない、とは限らない、 ている、てばかりいる、 たほうがよい、たものである、 べきである、…

2. 2 多語性表現

一語性表現に比べると結合の度合いが弱く、表現中に他語の挿入も時として許されるものをここでは多語性表現と呼ぶ。これらは、概念語同志の共起によるものと付属語が大きく関与するものに分けて示す。以下でPは述語、Nは名詞、Vは動詞、Sは文を意味する。

2. 2. 1 概念語同志の共起

NをP	2700	けんかを売る、けりをつける、 峠をこす、手を切る、骨を折る、 愛想をつかす、口火を切る、 生き恥をさらす、身を立てる、 腹を立てる、鼻をあかす、…
NがP	1550	頭が痛い、台所が苦しい、 気が重い、胸が痛む、 油が乗る、おつりがくる、 足がつく、山がはずれる、 注文がつく、…
NにP	950	逃げ腰になる、計算に入れる、 ふるいにかける、世評に高い、 ペールに包まれる、 眼鏡にかなう、槍玉にあげる、 踏みつけにする、…
NでP	30	鼻であしらう、金でしばる、 足でかせぐ、色めがねでみる、 金で釣る、きれいな事ですます、 肌で感じる、地で行く、…
NからP	15	根底からくつがえす、 ゼロから出発する、 一から出直す、 どん底からはい上がる、…
NまでP	3	骨までしゃぶる、…
NとP	4	面と向う、師と仰ぐ、…
NのN をP	60	財布の底をはたく、 背水の陣をしく、…
NのN がP	40	二の句がつけない、 化けの皮がはげる、…
NのN にP	30	元のさやに収まる、 玉のこしにのる、…
P(連体) NをP	50	寝た子を起す、痛い所をつく、 なす術を知らない、 知らないふりをする、…
P(連体) NがP	10	見る目がある、 取りつく島がない、…
P(連体) NにP	10	痛い目にあう、 高いものにつく、…
P(連体) NでP	3	白い目でみる、 長い目でみる、…
NもN もP	15	もともこもない、 身も心も捧げる 酸いも甘いもかみわかる、…

NにN をP	45	手に職をつける, 金に糸目をつけない, 顔に泥を塗る, 門前に市をなす, …
NをN にP	15	首を縦に振る, 世界を股にかける, …
NにN がP	17	足下に火がつく, 話に花がさく, 身に覚えがある, …
NがN にP	5	手が留守になる, 足が棒になる, …
NでN をP	9	肩で息をする, 陰で糸をひく, …
NがN をP	6	へそが茶をわかす, 金が物をいう, …
Nから NがP	5	喉から手が出る, 目からうろこが落ちる, …
その他 (格言, 諺など を含む)	1400	目は口ほどに物を言う, 当たって砕ける, 画竜点睛を欠く, 毒をもって毒を制す, 同じ釜の飯をくう, 天を仰いで唾する, 名を捨てて実を取る, 芸は身を助く, …

2. 2. 2 付属語が直接関与する共起

2. 2. 1が概念語同志の共起であるのに対して付属語が主役として共起に関与する場合である。これらには2項ばかりでなく3項の関係もある。

付属語が 直接関与 する共起	250	決して…ない, もし…ば, …しか…ない, いったい…か, とかく…がちだ, すでに…ている, どんなに…ても…ない, …
----------------------	-----	---

2. 3 その他の語の共起

特定の語の共起と言うより, 語を特定しない慣用的共起パターンが見られる。これらは, 未だ収集が完了していないが, たとえば次の様なものがある。

P(連体)なら P(連体)で	書くなら書くでちゃんと書け
P(連体)こと はP(終止)	面白いことは面白いが
NはNで	アメリカはアメリカで考えている
NというN	入口という入口は
NまたN	人また人だった
N(連用)にV	飲みに飲んだ

3. 変化形について

2. で(特に2. 2. 1で)示した慣用的表現における語の共起の固定化の度合いは千差万別であり, どこまで変化形が許されるかは, それぞれの表現によって異なる。また, 可能な変化形すべてを辞書式に羅列しておく事も量的に困難と思われる。従って, 各表現に対してどこまでの変化形を許すかを整理し, 明示しておく必要がある。国語学でも慣用句, 機能動詞結合としてどの様なものを認めるかという観点から, 変化形について種々の考察がなされているが⁽⁴⁾, 以下ではこれまでの断片的な知見を総合し, 機械処理の立場から考察する。

A. 付加

まず, 基本形に何らかの語が付加され得るか否か, 付加される場合はどの様な語が許されるかが明らかにされなければならない。これらは, 大きく, 修飾要素の付加, 助動詞等の付加, 副助詞等の付加に分けられる。

(a) 修飾要素の付加

① 連体修飾語句

名詞に対する修飾語句については, 表現によって, 必須的に要求される場合, 禁止される場合, その他の場合がある。また修飾句の形態に制約のある場合とない場合が考えられる。さらに, 修飾語句に対する意味上の制約も規定しておく必要がある。

例：○必須的な場合

S きれいだがある
S : <現象>, ...
P 顔をする
P : <感情>, ...
N の目に狂いはない
N : <人>, ...

○禁止される場合

N の業を煮やす
N : ϕ

○その他の場合

N の手に余る
N : <人>, ...
N の首を切る
N : <人>, ...

②連用修飾語句

述語に対する修飾語句についても必須的な場合（必須格）、禁止される場合、その他の場合が考えられる。また、動詞が本来取らなかった格が新たに取りられる場合や、その逆の場合もある。これらについても通常の動詞に対する格要素の規定と同様のことを行っておく必要がある。

例：○必須格

N につばをつける
N : <行為>, ...
N にバトンを渡す
N : <人>, ...

○禁止される場合

無い袖はぶらぶら振れない

○新たな格

N と（対象）手を切る
N : <人>, ...
N に（対象）気がある
N : <人>, ...
N を（対象）棒に振る
N : <地位>, ...

○消える格

彼から（起点）指揮をとる
（←→結晶から不純物をとる）

(b)助動詞等の付加

述語性の慣用表現の場合、特定の助動詞等が必須的に要求される場合や、禁止される場合がある。また、含まれる動詞本来のAspect素性が変わる場合もある。

例：○必須的な場合

話にならない
手がつけられない

○禁止される場合

頭が切れている
つぶしが利きははじめる
割に合ったばかりだ

(c)副助詞等の付加

副助詞等の付加も場合によって許されたり許されなかったりする。

例：○許されない場合

喉からは手が出る

○許される場合

気が気ではない

B. 削除

基本形に含まれる助詞を取りさっても慣用的表現として働くものがある。この種の削除も各表現について可能か否かを明らかにしておく必要がある。また、削除した結果の表現と複合動詞、複合形容詞等との区分も、少くとも取扱い上明確にしておく必要がある。

例：○許される場合

攻守が所をかえる
気味が悪い
汗水をたらす

○許されない場合

気がおける
身を持ちくずす

C. 置換

Aの(c)と類似の現象として格助詞を副助詞で置き換え得るかどうかも表現ごとに異なる。

例：○許されない場合

きまりしか悪くない

○許される場合

きまりも悪い

D. 語順の変更

格要素の順序を入れ換えてもよい場合と悪い場合がある。

例：○許される場合

口を横から出す

油に火をそそぐ

○許されない場合

手が喉から出る

棒に足がなる

E. 受身化, 使役化

例：○許される場合

金に物を言わせる

(←→金が物を言う)

○許されない場合

荷に勝たれる

(←→荷が勝つ)

へそに茶をわかさせる

(←→へそが茶をわかす)

F. 倒置による体言化

含まれる格要素としての名詞を末尾に回して連体被修飾語にかえる事が許される場合と許されない場合がある。

例：○許される場合

棒になった足

枯れた芸

○許されない場合

ない仕方

(←→仕方がない)

つけた気

(←→気をつける)

4. むすび

本稿では、まず、大量の言語データに基づいて行った慣用的表現の調査・収集の結果を示した。ここでの慣用的表現とは、従来、国語学で論じられてきた慣用句や機能動詞結合を含む、より広範な語の固定的結合体である。このデータは今後の日本語処理で各種ルール化を目差す際、ルール化の範囲を規定するための(いわば対極的な)基礎資料として有効と考えている。また、これらのデータは強い語の結合に関するものであるから構文解析の段階で解析のあいまいさを減らすためにもある程度有効であろう。次に、ここでの表現がどのような変化形について考慮されなければならないかについて、その基本的な事項を整理した。実際の処理ではこの種の情報は与えられた入力文字列が慣用的表現と解釈すべきか、標準的用法の表現と解釈すべきかの判定のためにも利用される事は言うまでもないが、たとえば、「金に糸目をつけない」や「たたらを踏む」などのように慣用的表現に特有の語が含まれる場合も整理しておくことも有効である。さらに広範な文脈処理等からの情報によってこの種のあいまいさを減らす問題は今後の重要な研究課題である。

本研究は一部、文部省特定研究「言語情報処理の高度化」による。

参考文献

- (1) 首藤, 吉村, 武内, 津田: 日本語文における慣用的表現について(1), 昭和62年度電気関係学会九州支部連合大会論文集, 808, 1987.
- (2) 今枝, 吉村, 首藤: 日本語文における慣用的表現について(2), 昭和62年度電気関係学会九州支部連合大会論文集, 809, 1987.
- (3) 田口, 吉村, 首藤: ユニフィケーション文法における慣用的表現の取り扱いについて, 昭和62年度電気関係学会九州支部連合大会論文集, 802, 1987.
- (4) 特集・慣用句, 日本語学, VOL. 4, 1月号, 1985.

- (5) 奥：日本語慣用表現の分析と日英翻訳への適用，情処学研報，87-NL-62, 1987.
- (6) 首藤，檜原，吉田：日本語の機械処理のための文節構造モデル，信学論，62-D, 12, 1979.
- (7) 首藤，檜原：日本語の文構造のわく組みを与える表現，福岡大学総合研究所報，63, 1983.