

日本語慣用句機械辞書

首藤公昭[†] 高橋雅仁^{††} 田辺利文[†]

意味,あるいは構文上の構成性が成立しないと考えられる複単語表現,すなわち慣用句の取り扱いが機械処理において古くからの問題点である。筆者らは3,669個の典型的と思われる日本語慣用句を収録した機械処理用辞書を開発したのでその概要を報告する。本辞書はより広範な複単語表現辞書 JDMWE(首藤ほか 2010, Shudo et al. 2011) において特に非構成性が強いと判断される表現を抜粋したものである。

A Dictionary of Idioms for Machine Processing of Japanese

Kosho Shudo,[†] Masahito Takahashi^{††}
and Toshifumi Tanabe[†]

There still exist various problems around the idiom processing in NLP. This paper introduces a new dictionary of 3,669 typical idioms for Japanese language processing, which is characterized by a detailed description of syntactic function, structure and an indication of the possibility of internal modification of constituent words for each idiom.

1. はじめに

Ivan. A. Sagらの論文“Multiword Expressions: A Pain in the Neck for NLP”が2002年に発表されて以来,慣用句を代表とする,言語的に特異な複単語表現(MWE)の取り扱いの問題がNLPにおいてクローズアップされている。しかしながら,これらの表現は,まさに「特異」であり「個別的」であるが故に統一的取り扱いが難しく,各国言語においてもいまだ十分な研究成果は得られていないようである。筆者らは年数をかけて以下の表現を収録したMWEの大規模機械辞書JDMWEを編纂し,その概要を(首藤ほか 2010, Shudo et al. 2011)で報告した[a]。

1. 慣用句(idioms)
2. 準慣用句(quasi-idioms)
3. 諺, 格言, 故事成句(proverbs, sayings)
4. 決まり文句(cliches, set expressions)
5. 準決まり文句(quasi-cliches)
6. 四字熟語(four-character-idioms)
7. オノマトペやメトニムを含む表現(onomatopoeic or mimetic multiword expressions)
8. 非構成性あるいは弱構成性の複合語(non- or weak-compositional compounds)
9. クランベリー型表現(cranberry-type expressions)
10. 慣用される比喩, 直喩, 張喩, 換喩表現(generally-used metaphors, generally-used similes, generally-used hyperboles, generally-used metonymies)
11. 複合連結詞(multiword discourse connectives)
12. 支援動詞構文(support verb constructions: SVC), 軽動詞構文(light verb constructions:LVC)
13. 支援形容詞構文(support adjective constructions: SAC)
14. 慣用される挨拶, 応答, 呼びかけ, 独言表現(generally-used expressions for greeting, responding, calling, soliloquys)

本論文では,上記のなかで際立った特異性を持つ1,すなわち,意味に明確な非構成性が認められる,いわゆる日本語慣用句3,669個を収録した機械辞書について,その概要を述べる。本辞書の収録表現は,新聞記事,雑誌,テレビ音声・映像などの生データから人手によって収集したものを中心に,既存の事典類を利用して確認・補強を施したものであり,3,669行,8欄(A~H欄)からなるMicrosoft Excel形式に纏められている。各欄には,見出し,分か

[†] 福岡大学
Fukuoka University

^{††} 久留米工業大学
Kurume Institute of Technology

a) ただし,以下は必ずしも排他的な分類ではない。

ち書き情報, 異表記情報, 機能・種別情報, 構造情報, 前方文脈条件情報, 後方文脈条件情報, 連体・連用・動詞化情報がこの順に記載されている. 以下, 各欄の情報について解説する.

2. 見出し(A欄)

この欄には慣用句の平仮名ベタ書き見出し(header)を与える. 読みベースであり, 漢字が複数の読みを持っている場合は可能な読みごとに見出しを与える. 例えば, 「九-死-に-一-生-を-得-る」の「得る」は「える」, 「うる」の両方の読みが可能なので, 「きゅうしにいっしょうをえる」, 「きゅうしにいっしょうをうる」を別見出しとして立てる, などである.

3. 分かち書き情報(B欄)

この欄には, 見出し表現を単語, 接辞(接頭語, 接尾語, 接頭造語要素, 接尾造語要素)を単位として分かち書きした表現(segmentation)を与える. ここで, 造語要素とは造語能力が比較的強く, 単独で用いられることのない形態素である[b]. 接頭語, 接尾語, 接頭造語要素, 接尾造語要素は以下の様に分類している.

(1) 接頭語(P)

- (1.1)名詞, 動詞連用形, 形容動詞語幹に前接する接頭語
- (1.2)用言に前接する接頭語

(2) 接頭造語要素(Q)

- (2.1)名詞, 動詞連用形, 形容動詞語幹に前接する接頭造語要素
 - (2.1.1)名詞を与えるもの
 - (2.1.2)形容動詞語幹を与えるもの
- (2.2)用言に前接する接頭造語要素

(3) 接尾語(S)

- (3.1)名詞, 動詞連用形, 形容動詞語幹に後接する接尾語
 - (3.1.1)サ変名詞以外の名詞を与えるもの
 - (3.1.2)サ変名詞を与えるもの
 - (3.1.3)形容動詞語幹を与えるもの
 - (3.1.4)形容詞を与えるもの
 - (3.1.5)動詞を与えるもの
- (3.2)用言に後接する接尾語

b) 音読みの一漢字が多い.

- (3.2.1)名詞を与えるもの
 - (3.2.2)形容動詞語幹を与えるもの
 - (3.2.3)形容詞を与えるもの
 - (3.2.4)動詞を与えるもの
- (4)接尾造語要素(R)
- (4.1)名詞, 動詞連用形, 形容動詞語幹に後接する接尾造語要素
 - (4.1.1)サ変名詞以外の名詞を与えるもの
 - (4.1.2)サ変名詞を与えるもの
 - (4.1.3)形容動詞語幹を与えるもの
 - (4.2)用言に後接する接尾造語要素
 - (4.2.1)名詞を与えるもの
 - (4.2.2)形容動詞語幹を与えるもの

本辞書の収録表現で使われている接頭語, 接尾語, 接頭造語要素, 接尾造語要素の具体例を付録 1 に示す. 分かち書きはドット「.」あるいはハイフン「-」で示している. ドットは要素間の区切りを与えると同時にその直後の語句がさらなる修飾を受け得ること(内部修飾可能性)を表わす. 例えば, 「あぶら-を.うる」の「うる」の前のドットは「油-を-いつも-売る」の様に「売る」が「いつも」などの様な内部修飾を受ける可能性が有ること, 同様に, 「.わな-に.かかる」のドットは「巧妙-な-罠-に-再び-掛かる」などの可能性があることを表わす. 単語であっても, その一部が異なった字種で表記可能な場合は字種の変り目に別種の区切り記号, アンダースコア「_」を入れている. 例えば, 「りゅう_いん-が.さがる」の単語「溜飲」は「りゅう飲」と表記されることがあると考え, 「りゅう_いん」と区切っている. これと C 欄の漢字情報「溜飲」とから「溜飲」, 「溜いん」, 「りゅう飲」, 「りゅういん」という4つの表記形が組み合わせ的に導出できる. ただし, 「善悪」, 「白黒」, 「幽明」など, 対立的意味の漢字の組み合わせは二語の並列句と見なして「善-悪」, 「白-黒」, 「幽-明」としている. 原則として活用語尾は語幹から切り離していない. ただし, 形容動詞の活用語尾とも見なされる, 「な」, 「に」, 「なる」, 「たる」などは助動詞「だ」, 「なり」, 「たり」の活用形と見なして語幹と切り離す.

4. 異表記情報(C欄)

片仮名表記, 漢字表記, 送り仮名の有無など, 日本語特有の表記の多様さをコンパクトに記載した欄である. 表記, n_1, n_2, \dots, n_m が交換可能であるとき, $(n_1/n_2/\dots/n_m)$ と括弧内に記号「/」で区分して記載し, 表記 n が無くてもよいとき(n)と記載している. 例えば, 「(思/想/惟/懐)い-も-及ば-ない」における「(思/想/惟/懐)い」の部分は「思い」, 「想い」, 「惟い」, 「懐い」の

4つの可能性が有ることを表わす。また、「真(っ)-赤-な-嘘」の「真(っ)」の部分は「真」と「真っ」の2つの場合が有ること、さらに、「右-も-左-も-(分(か)/解/判)ら-ない」の「(分(か)/解/判)ら」は、「分から」、「分ら」、「解ら」、「判ら」の4つの場合があることを意味する。これらとB欄の仮名書きを加えれば、「わからない」、「分らない」、「分らない」、「解らない」、「判らない」などの異表記がカバーされる。

B, C 欄の情報によれば、本辞書は実質的に 15,000 表記程度をカバーすると推定される。

5. 機能・種別情報(D 欄)

この欄は機能と種別表示欄である。機能・種別は、概略、以下に英字で示すようなコードで記載されている。

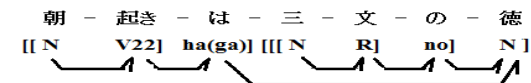
1. 形容詞性表現の類
Adj 例: 「愛想-が-(良/善/好)い」
2. 形容動詞性表現の類
AdjVerb 例: 「(顎/アゴ)-が-落ちる-様」
3. 連体修飾表現の類
Adnom 例: 「有り-と-有らゆる」
4. 連用修飾表現の類
Adv 例: 「明け-ても-暮れ-ても」
5. 連結詞(文接続詞)性表現の類
DiscCon 例: 「事-に-(因/依/拠)っ-たら」
6. 名詞性表現の類
Noun 例: 「赤-の-他人」
7. 動詞性表現の類
Verb 例: 「ああ-(言/云/謂)え-ば-斯う-(言/云/謂)う」
8. 格言、諺、故事成句、決まり文句の類
_ProvClich 例: 「朝-起き-は-三-文-の-徳」
9. 応答表現の類
_Res 例: 「冗談-も-休み-休み-(言/云/謂)え」
10. 独言の類
_Self 例: 「何-と-(言/云/謂)う-事-だ」

活用など、慣用句の後続語に係る機能は原則として末尾の単語の同機能に準じるが、かなり

の例外が見られる。例えば、「杓子-定規」の末尾の単語[定規]は単純な名詞であるのに「杓子定規な意見」と「な」をおくって連体修飾が行なわれる。この種の細かな情報は後述の H 欄で与える。機能・種別コードの詳細は付録 2 に記す。この情報により通常の形態・構文処理への組み込みが容易になる。

6. 構造情報(E 欄)

ここでは表現の内部構造を括弧表現で与える。構造の基本的枠組みとして、いわゆる係り受け構造を採用し、修飾子、被修飾子の対を括弧 [] で括った句表示 (phrase marker) で構造記述を与える [c]。日本語では修飾子は文頭側、被修飾子は文末側と決っているので、係りの方向指示は行なわない。句構造文法の非終端記号に相当するものは用いない。構造記述は、自立語は品詞記号、機能語(および相当語)は英小文字による綴りで与える。すなわち本辞書は品詞付き、係り受け構造付きの慣用句辞書である。例えば、「朝-起き-は-三-文-の-徳」の E 欄の構造記述 [[NV22]ha(ga) [[[N]no]N]] は下図の係り受け関係を示す。



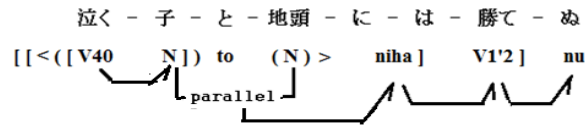
ここで、N は名詞、R は接尾造語要素、V22 は動詞の連用形であり、ha(ga) は、深層で「が」格として使われた副(係)助詞「は」を意味する。品詞とその記号体系は付録 3 および付録 4 に与える [d]。慣用句には古語的表現がよく現れる。本辞書で採用した現代語、古語を併せた活用と記号体系の概略を付録 5 に示す。多くの場合、句表示の最外殻の括弧は省略している [e]。また、機能性名詞「よう(様)」、「ほど(程)」、機能動詞(あるいは軽動詞)の「する」、「なす」、「す」、「ある」、「なる」および機能性形容詞「ない」は活用形を含めて英小文字で表記した。例えば、「足枷-に-なる」の構造記述は[*Nni]naru とした。表記中のアスタリスク[*] は後述するように、それに続く「足枷」が「重い足枷になる」のように内部修飾句を取る可能性があることを示す。

並列構造は括弧表現 < > または << >> で、並列される要素は括弧 () で表わしている。例えば、「泣く-子-と-地頭-に-は-勝て-ぬ」の句表示 [[<<[V40N]to(N)>niha]V1'2]nu は次の構造を意味する。

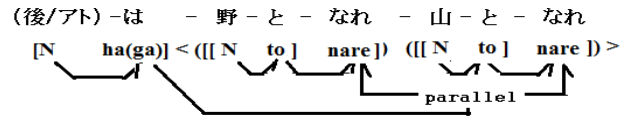
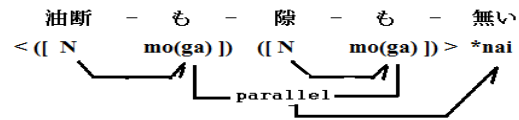
c) 係り受けとは言いにくい文節内部の語の接続も本辞書では便宜上係り受けと同じ括弧、[] で句構造表示を行なっている。

d) D 欄に与えた表現全体の機能・種別とは異なる記号系である。

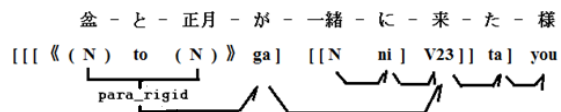
e) それ以外で括弧が省略されているときは、左分岐、 $[w_1 w_2 w_3 \dots w_n] = [[[[w_1 w_2]w_3] \dots]w_n]$ と約束する。



「泣く-子-と-地頭」の並列される要素は「泣く-子」と「地頭」であることがそれぞれ () で括って示され、「泣く-子-に-は-勝て-ぬ-and-地頭-に-は-勝て-ぬ」とほぼパラフレーズできる分配型並列句であることが括弧 < > で示されている。V40は動詞連体形、V1'2は古語動詞の又接続未然形を意味する。以下に同様の分配型並列句の二例を示す。



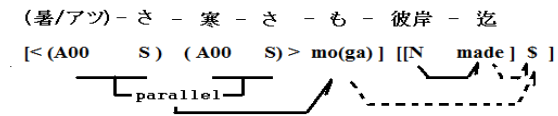
また、「盆-と-正月-が-一緒-に-来-た-様」の句表示，[[[<((N)to(N))ga][[Nni]V23]]ta]you は次の構造を意味する。



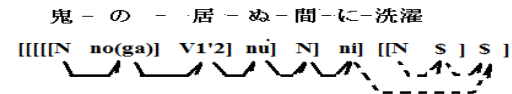
V23は動詞のタ、ダ接続の連用形を表わす。この例に含まれる並列句「盆-と-正月」は「一緒に」という語句との共起で強い結合が与えられており、「盆-が-一緒-に-来-た-and-正月-が-一緒-に-来-た」と分配的にパラフレーズすることはできない。これは水谷(1972)が東ねの並

列句と呼んだ、分配則の利かない並列構造である。このように < > は分配的な並列構造 (distributive coordination), << >> は東ねを代表とする硬い並列構造 (rigid coordination) を意味している。

慣用句や決まり文句にはいわゆる句構造をなしていない表現が含まれる。本辞書ではこの種の表現の構造記述に空要素記号「\$」を用いている。例えば、「(暑/アツ)-さ-寒-さ-も-彼岸-迄」には句表示 [[<(A00S)(A00S)>mo(ga)][[Nmade]\$]] によって次の構造を与える。

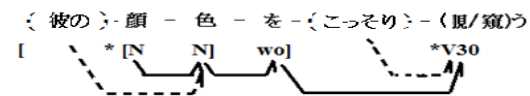


「\$」は句を成立させるためのダミー記号である。この例では\$によって「続く」などの述語が想定されている。同様に「鬼-の-居-ぬ-間-に-洗濯」の場合は次のとおりである。



この例では末尾に「を-する」などの構成要素が想定されている。V1'2は古語動詞の又接続の未然形を意味する。

構造記述内のアスタリスク「*」はB欄のドット「.」と同じ内部修飾可能性を句構造表示上に示す。アスタリスクの直後に続く句は内部修飾を受ける可能性があることを意味している。例えば、「顔-色-を-(視/窺)う」の場合、B欄に「.かお-いろ-を.うかがう」、E欄に「*[NN]wo]*V30」と記載されているが、修飾を受け得るのは「顔」ではなく「顔-色」であり、この様にE欄の方がより正確に内部修飾可能性を記述できる。「*」の与える情報は、例えば、下図のとおりである。V30は動詞終止形を意味する。



内部修飾句の入り得る場所は係り受けの非交差条件によって自動的に決まる。この情報によれば、慣用句を単語 (words with spaces) として扱う弊害を避け、慣用句の構文的柔軟性に対処する事が出来る。

7. 前方文脈条件(F 欄)

例えば、「(羽目/ハメ)-に-なる」, 「(目/眼)-に-(会/遭/逢/遇)う」は、それぞれ、「苦しむ-(羽目/ハメ)-に-なる」, 「つらい-(目/眼)-に-(会/遭/逢/遇)う」のように修飾句(「苦しむ」, 「つらい」など)を必須的に要求する。F 欄にはこの条件が〈adnom. modifier〉* と記載されている。これらの条件には制約(必須)条件と選好条件とがあり、制約(必須)条件の場合はアスタリスク[*]を付す。この種の条件には、例えば、以下の種類がある。

- ・ 〈adnom. modifier〉 : 連体修飾句をとる
- ・ 〈beginning of sentence〉 : 文頭に来る
- ・ 〈wo-adv. modifier〉 : 「を」格による連用修飾句をとる
- ・ 〈ga-adv. modifier〉 : 「が」格による連用修飾句をとる
- ・ 〈ni-adv. modifier〉 : 「に」格による連用修飾句をとる

8. 後方文脈条件(G 欄)

例えば、副詞性表現「(一/1)つ-と-し-て」は後方の否定句に係る事が必須であり、G 欄にはこの様な後方の文脈条件を記載する。この種の条件には、例えば、以下のものがある。

- ・ 〈~なんて〉, 〈~とは〉, 〈~などと〉 : 「なんて」, 「とは」, 「などと」などの文末側語句と共起する
- ・ 〈end of sentence〉 : 文末にある
- ・ 〈negation〉 : 否定句をとる

制約(必須)条件の場合はアスタリスク[*]を付す。

9. 連体・連用・動詞化用法(H 欄)

形容動詞性表現 AdjVerb, 状態記述名詞 Noun/statdescr, 連用修飾(副詞性)表現 Adv, 格言・諺 AdvClich は広い意味で物事の様態を表わし、連体修飾句, 連用修飾句として用いられ、動詞化して用いられる場合がある。H 欄はその可能性と変化の仕方を与える。この情報は三つ組, $\alpha-\beta-\gamma$ の形式で与える。ここで, α , β , γ は、それぞれ、連体修飾句化, 連用修飾句化, 動詞化させる際に必要な後接表現の集合である。例えば、形容動詞性表現「意気-揚(揚/々)」に対しては{taru, tosita, no}-{tosite, to, ε}-{tosuru}と記述している。即ち、「意気揚々たる」, 「意気揚々とした」, 「意気揚々の」で連体修飾, 「意気揚々として」,

「意気揚々と」, 「意気揚々」で連用修飾, 「意気揚々とする」と動詞化することを示している[f]。 $\alpha-\beta-\gamma$ のパターンは様態複合表現の形態・構文的機能の種類すなわち一種の精密化された品詞である。例えば、{na}-{ni}-{ninaru}というパターンは典型的な形容動詞を意味する[g]。 α , β , γ には空集合 ϕ の場合もある。 $\alpha-\beta-\phi$, $\alpha-\phi-\phi$ は、それぞれ、 $\alpha-\beta$, α と略記している。これらの変化形も慣用句としてカウントすれば本辞書は C 欄の異表記とは別に約 7,000 慣用句を収録している事になる。

10. 関連研究

NLP の立場から日本語慣用句に焦点をあてた初期の研究としては、「名詞+格助詞+用言」の形の慣用表現に限定してその構文的な硬さを解析に利用する手法を考察した(奥 1987)や慣用句の形態をより総括的に調べ、許容・禁止される変化形の種類を網羅的に整理・考察した(首藤ほか 1987, 1988, 今枝ほか 1987)がある。これらの研究は慣用句辞書の編纂や慣用句解析に向けた予備的な考察を行ったものであるが、首藤(1989)は日本語慣用句候補、約 21,000 表現を収録した辞書を編纂し、公開した。本稿で示した辞書は、この辞書の見出しを典型的慣用句に限定するいっぽう、辞書項目を機械処理向きに大幅に拡充して電子化したものとなっている。

佐藤(2007)は(金田一(監修) 2005), (宮地裕(編) 1982), (米川ほか(編) 2005)など、5 種類の市販の慣用句辞典にどのような表現が収録されているかを調べ、3,629 表現を採集した。これらの表現の多くは本論文で述べた辞書にも収録されている。

テキストの意味解析を行う際、慣用句に見える表現が本当に慣用句として使われているのか、文字通りの意味で使われているのかを判定する必要がある場合がある。これには慣用句としては許されない変化形で使われているという判定容易な場合も有るが、それ以外の一般的な判定基準が必要になる。Hashimoto ほか(2009)は素性セットを巧く設定すれば対象とした 146 個の慣用句については機械学習によってかなりの精度でこのような多義解消が可能である事を示した。

11. おわりに

本稿で述べた慣用句辞書の特徴を要約すれば以下の通りである。

1. 漢字, カタカナの用法や送り仮名の有無など、表記の多様性に配慮している。
2. 慣用句の構文(一部, 意味)機能を細かく規定し、通常の構文処理とのリンクを容易にして

f) この欄で現れる「ε」は空表現を意味する。

g) これらのパターンは全部で 300 種程度が存在する。

- いる。
- 慣用句の柔軟性の指標である内部修飾可能性(internal modifiability)を個別的に細かく記載している。
 - 慣用句内部の係り受け構造および並列構造を精密に記載している。
 - 構造的に不完全な表現(ill-formed expressions)にも対応している。
 - 様態を表す慣用句に対しては文中での用法を体系的に記載している。

参考文献

- Hashimoto, C. and Kawahara D.: Compilation of an Idiom Example Database for Supervised Idiom Identification. Language Resource and Evaluation, Vol.43, No.4 (2009).
- 今枝紀代司, 吉村賢治, 首藤公昭: 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 (1987)
- 金田一秀徳(監修): 小学生のまんが慣用句辞典, 学研 (2005).
- 宮地裕(編): 慣用句の意味と用法, 明治書院 (1982).
- 水谷静夫, 田中幸子: 語の並列結合子, 計量国語学, No.63, pp.19-36 (1972).
- 奥雅博: 日本語慣用表現の分析と日英翻訳への適用, 情報処理学会研究報告, 1987-NL-062 (1987).
- 尾上兼英(監修): 成語林-故事ことわざ慣用句, 旺文社 (1993).
- Sag, I. A., Baldwin, T., Bond, F., Copestake, A., and Flickinger, D.: Multiword Expressions: A Pain in the Neck for NLP, Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics, CICLING2002, pp.1-15 (2002).
- 佐藤理史: 基本慣用句五種対照表の作成, 情報処理学会研究報告, 2007-NL-178 (2007).
- 首藤公昭, 吉村賢治, 武内美津乃, 津田健蔵: 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 5 (1987).
- 首藤公昭, 吉村賢治, 武内美津乃, 津田健蔵: 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-, 情報処理学会研究報告, 1988-NL-066 (1988).
- 首藤公昭: 日本語の固定的複合表現, 昭和63年度文部省科学研究費(1)「情報ドキュメンテーションのための言語の研究」63101005 報告書 (1989).
- 首藤公昭, 田辺利文: 日本語の複単語表現辞書: JDMWE, 自然言語処理, Vol.17, No. 5, pp.51-74 (2010).
- Shudo, K. Kurahone, A., and Tanabe, T.: A Comprehensive Dictionary of Multiword Expressions, Proceedings of the 49th Annual Meeting of the ACL, pp.169-177 (2011).
- 竹田晃: 四字熟語・成句辞典, 講談社 (1990).
- 田島諸介: ことわざ故事・成語慣用句辞典, 梧桐書院 (2002).
- 米川明彦, 大谷伊都子(編): 日本語慣用句辞典, 東京堂出版 (2005).

付録1 本辞書に現れる接頭語, 接頭造語要素, 接尾語, 接尾造語要素

1 接頭語

1.1 名詞, 動詞連用形, 形容動詞語幹に前接する接頭語 (P)

表現	出現例
「御」, 「お」, 「ご」, 「おん」, 「み」	「御-株-を-奪う」, 「御-灸-を-据える」, 「御-里-が-知れる」, 「御-茶-を-濁す」, 「御-機嫌-斜(め)」, 「錦-の-御-旗」
「どん」	「貧乏-の-ドン-底」
「空」, 「そら」	「他人-の-(空/ソラ)-似」

1.2 用言に前接する接頭語 P

表現	出現例
「打(ち)」, 「うち」	「尾羽-打(ち)-枯らす」

2 接頭造語要素

2.1 名詞, 動詞連用形, 形容動詞語幹に前接する接頭造語要素 (Q)

2.1.1 名詞を与えるもの

表現	出現例
「逆」, 「さか」, 「ぎゃく」, 「さかさ」	「(逆/サカ)-手-に-取る」
「真」, 「真」, 「ま」, 「まつ」	「真-綿-で-(首/クビ)-を-(絞/締)める」
「大」, 「おお」, 「だい」	「(瘦/瘠)せ-の-大-(食/喰)い」, 「見る-と-(聞/聴)く-と-は大-違い」, 「大-風-が-吹け-ば-桶-屋-が-儲(か)る」
「不」, 「ふ」, 「ぶ」	「出来-不-出来」, 「医者-の-不-養生」
「無」, 「む」, 「ぶ」	「無-味-乾燥」
「両」, 「りょう」, 「もろ」	「(諸/両)-刃-の-剣」, 「両-手-に-花」, 「(両/諸/モロ)-(肌/膚)-を-脱ぐ」, 「文-武-両-道」, 「一-挙-両-得」, 「(クルマ/車)-の-両-輪」, 「(ケンカ/喧嘩)-両-成敗」
「小」, 「しょう」, 「こ」	「小-鼻-を-(膨/脹)らます」, 「小-腹-が-減る」, 「小-回り-が-利く」, 「小-脇-に-抱える」, 「小-脇-に-挟む」
「難」, 「なん」	「難-(癖/クセ)-を-付ける」
「多」, 「た」	「多-事-多-難」, 「多-種-多-様」
「丸」, 「まる」	「面目-(丸/マル)-潰れ」

「生」, 「なま」	「蛇-の-生-殺し」, 「生-唾-を-(飲/呑)み-込む」, 「生-木-を-裂く」, 「生-兵法-は-怪我-の-(本/基)」
「半」, 「はん」	「半-死-半-生」, 「一-言-半-句」
「悪」, 「あく」	「悪-名-が-高い」, 「悪-夢-から-覚める」
「前」, 「ぜん」	「前-門-の-(虎/トラ)-後-門-の-狼」
「後」, 「こう」, 「あと」	「前-門-の-(虎/トラ)-後-門-の-狼」
「片」, 「かた」	「片-足-を-突(つ)-込む」
「角」, 「かく」	「(目/眼)-を-三-角-に-する」

2.1.2 形容動詞語幹を与えるもの

2.2 用言に前接する接頭造語要素 Q

表現	出現例
「逆」, 「さか」, 「ぎやく」	「柳眉-を-(逆/サカ)-立てる」
「相」, 「あい」	「幽-明-相-隔てる」, 「肝胆-相-照らす」, 「同病-相-(哀/憐)れむ」

3. 接尾語 (S)

3.1 名詞, 動詞連用形, 形容動詞語幹に後接する接尾語

3.1.1 サ変名詞以外の名詞を与えるもの (なし)

3.1.2 サ変名詞を与えるもの S

表現	出現例
「ごっこ」	「(鼬/イタチ)-ゴッコ」

3.1.3 形容動詞語幹を与えるもの (なし)

3.1.4 形容詞を与えるもの (なし)

3.1.5 動詞を与えるもの S30

表現	出現例
「めく」	「時-めく」

3.2 用言に後接する接尾語

3.2.1 名詞を与えるもの S

表現	出現例
「味」, 「み」	「高-み-の-見物」 (注1)
「さ」	「(暑/アツ)-さ-寒-さ-も-彼岸-迄」, 「可愛-さ-余(つ)-て-憎-さ-百-倍」 (注2)

注1 形容詞の語幹に接続

注2 形容詞, 形容動詞の語幹に接続

3.2.2 形容動詞語幹を与えるもの (なし)

3.2.3 形容詞を与えるもの S30

表現	出現例
「難しい」, 「がたい」	「有(り)-難-迷-惑」 (注3)
「易い」, 「やすい」	「熱し-易く-冷め-易い」 (注4)

注3, 4 動詞連用形に接続

3.2.4 動詞を与えるもの (なし)

4 接尾造語要素 (R)

4.1 名詞, 動詞連用形, 形容動詞語幹に後接する接尾造語要素

4.1.1 サ変名詞以外の名詞を与えるもの R

表現	出現例
「枚」, 「まい」	「(首/クビ)-の-皮-一-枚」, 「(首/クビ)-の-皮-一-枚-で-繋がる」, 「役者-が-一-枚-上」, 「一-枚-噛む」
「貫」, 「かん」	「裸-一-貫」
「中」, 「ちゆう」, 「じゆう」, 「ぢゆう」	「不幸-中-の-幸い」, 「年-がら-年-中」, 「十-中-八-九」
「物」, 「ぶつ」, 「もつ」	「(腹/ハラ)-に-一-物」
「者」, 「しゃ」	「二-者-択一」
「人」, 「にん」, 「びと」	「二-人-三-脚」, 「十-人-十-色」, 「三-人-寄れば-文殊-の-知恵」
「面」, 「めん」, 「づら」	「(吠え/ホエ)-の-(面/ヅラ)-を-搔く」
「度」, 「ど」	「(仏/ホトケ)-の-顔-も-三-度」, 「二-度-(有/在)る-事-は-三-度-(有/在)る」, 「御-百-度-を-踏む」
「本」, 「ほん」, 「ぼん」	「一-本-調子」, 「一-本-取る」
「事」, 「じ」	「多-事-多-難」
「点」, 「てん」	「紅-一-点」
「種」, 「しゆ」	「多-種-多-様」
「氣」, 「げ」, 「け」	「(虞/恐/怖/畏)れ-氣-も-無く」, 「毒-氣-を-(抜/脱)か-れる」, 「何-氣-無し-に」
「部」, 「ぶ」	「一-部-始-終」
「様」, 「よう」	「多-種-多-様」
「前」, 「まえ」	「朝-(飯/メシ)-前」
「代」, 「だい」, 「しろ」	「一-世-一-代」

「手」, 「しゅ」, 「て」	「(押/圧)し-の-手」
「足」, 「そく」	「二-足-の-草鞋-を-履く」
「度目」, 「どめ」	「三-度_目-の-正直」
「歩」, 「ぼ」, 「ほ」	「五十-歩-百-歩」
「言」, 「こと」, 「ごと」, 「ごん」, 「げん」	「一-言-居士」, 「一-言-一-句」
「番」, 「ばん」	「い-の-一-番」
「頭」, 「とう」	「一-頭-地-を-抜く」
「生」, 「せい」, 「しょう」	「半-死-半-生」
「倍」, 「ばい」	「可愛-さ-余-つ-て-憎-さ-百-倍」
「文」, 「もん」	「朝-起き-は-三-文-の-徳」, 「二-束-三-文」
「杯」, 「はい」, 「ばい」	「一-(杯/パイ)-(食/喰)う」, 「胸-が-一-杯」
「里」, 「り」	「千-里-の-道-も-一-歩-より-始まる」, 「悪事-千-里-を-走る」, 「一-望-千-里」, 「五-里-霧中」
「日」, 「にち」	「十-年-一-日-の-如く-に」, 「日-が-な-一-日」, 「人-の-(噂/ウワサ)-も-七十五-日」
「品」, 「ひん」, 「びん」	「天下-一-品」
「歳」, 「さい」, 「さい」	「青-二-(オ/歳)」
「山」, 「さん」, 「ざん」	「天王-山」
「匹」, 「引き」, 「びき」	「(猫/ネコ)-の-子-一-匹-居-ない」, 「(大/太)山-鳴動-して-(鼠/ネズミ)-一-匹」
「寸」, 「すん」, 「ずん」	「一-寸-の-虫-に-も-五-分-の-魂」, 「舌-先-三-寸」, 「一-寸-先-は-(闇/暗)」
「銭」, 「せん」	「(泥棒/ドロボウ)-に-追(い)-銭」
「分」, 「ぶん」, 「ぶん」	「四-分-五-裂」, 「五-分-五-分」, 「申し-分-が-無い」, 「一-寸-の-虫-に-も-五-分-の-魂」, 「九-分-九-厘」
「厘」, 「りん」	「九-分-九-厘」
「屋」, 「や」	「餅-は-餅-屋」, 「大-風-が-吹け-ば-桶-屋-が-儲(か)る」
「先」, 「さき」	「(目/眼)-先-が-(替/変/代)(わる)」, 「(鉾/矛/ホコ)-先-を-向ける」, 「鼻-先-で-あしらう」, 「舌-先-三-寸」
「丁」, 「ちょう」	「口-八-丁-手-八-丁」
「敗」, 「はい」, 「ばい」	「一-敗-地-に-塗れる」
「階」, 「かい」	「二-階-から-(目/眼)-薬」
「世」, 「せ」, 「せい」	「一-世-一-代」

「夢」, 「む」	「悪-夢-から-覚める」
「時」, 「じ」, 「とき」	「茶-腹-も-一-時」, 「一-時-の-(恥/辱/ハジ)」
「発」, 「はつ」	「百-発-百-中」
「策」, 「さく」	「窮余-の-一-策」, 「万-策-尽きる」
「年」, 「ねん」	「一-年-の-計-は-元旦-に-(有/在)り」, 「十-年-一-日-の-如く-に」, 「桃-栗-三-年-柿-八-年」, 「(災い/禍)-も-三-年-経て-ば-用-に-立つ」, 「石-の-上-に-も-三-年」
「車」, 「しゃ」	「前-車-の-轍-を-踏む」

4.1.2 サ変名詞を与えるもの R

表現	出現例
「化」, 「か」	「千-変-万-化」
「視」, 「し」	「白眼-視」

4.1.3 形容動詞語幹を与えるもの R00

表現	出現例
「式」, 「しき」	「(芋/イモ)-(蔓/ヅル)-式」

4.2 用言に後接する造語要素

4.2.1 名詞を与えるもの R

表現	出現例
「様」, 「よう」, 「ざま」	「例え-様-が-無い」, 「手-の-打ち-様-が-(有/在)る」, 「(物/モノ)-は-(言/云/謂)い-様」

(注 5)
 注 5 動詞連用形に接続

4-2-2 形容動詞語幹を与えるもの R00

表現	出現例
「放題」, 「ほうだい」	「遣り-たい-放題」 (注 6)
「げ」	「事-も-無-げ」 (注 7)

注 6 動詞連用形, 助動詞「たい」の連体形に接続
 注 7 形容詞語幹, 一部の動詞, 形容動詞語幹に接続

付録2 収録表現の機能・種別

1. 格言、諺、故事成句、諷刺文句の類

コード	説明	出現例
_ProvClich_Adj	末尾が形容詞	「(蟻/アリ)-の-這い-出る-隙間-も-無い」
_ProvClich_Noun	末尾が名詞	「朝-起き-は-三-文-の-徳」
_ProvClich_Noun_gotosi	末尾が名詞+「の如し」	「帰心-矢-の-如し」
_ProvClich_Noun_kara	末尾が名詞+「から」	「病(い)-は-気-から」
_ProvClich_Noun_made	末尾が名詞+「まで」	「(暑/アツ)-さ-寒-さ-も-彼岸-迄」
_ProvClich_Noun_nari	末尾が名詞+「なり」	「事実-は-小説-より-も-奇-なり」
_ProvClich_Noun_wo	末尾が名詞+「を」	「(目/眼)-に-は-(目/眼)-を-齒-に-は-齒-を」
_ProvClich_Verb	末尾が動詞(終止形)	「悪事-千-里-を-走る」
_ProvClich_Verb(imp)	末尾が動詞(命令形)	「当(た)っ-て-砕ける」
_ProvClich_Verb_besi	末尾が動詞+「べし」	「後生-畏る-べし」
_ProvClich_Verb_ka	末尾が動詞+「か」	「鬼-が-出る-か-蛇-が-出る-か」
_ProvClich_Verb_gotosi	末尾が動詞+「如し」	「過ぎ-たる-は-及ば-ざる-が-如し」
_ProvClich_Verb_mai	末尾が動詞+「まい」	「雉(子)-も-鳴か-ず-ば-(撃/射/打)た-れ-まい」
_ProvClich_Verb_nai	末尾が動詞+「ない」	「転ん-でも-(只/唯/タダ)-で-は-起き-ない」
_ProvClich_Verb_nu	末尾が動詞+「ぬ」	「泣く-子-と-地頭-に-は-勝-て-ぬ」
_ProvClich_Verb_reru	末尾が動詞+「れる」	「飼(い)-(犬/イヌ)-に-手-を-噛ま-れる」
_ProvClich_Verb_tari	末尾が動詞+「たり」	「(下手/ヘタ)-の-考え-休む-に-似-たり」
_ProvClich_Verb_zu	末尾が動詞+「ず」	「悪銭-身-に-付か-ず」
_ProvClich_Verb_na	末尾が動詞+「な」	「日光-を-見ない-内-は-結構-と-(言/云/謂)う-な」
_ProvClich_Verb_ta	末尾が動詞+「た」	「泣い-た-(鳥/カラス)-が-もう-笑-った」
_ProvClich_Verb_zaraq	末尾が動詞+「ざらん」	「精神-一到-何事-か-成ら-ざらん」

2. 応答表現の類

コード	説明	出現例
_Res	応答表現	「冗談-も-休み-休み-(言/云/謂)え」

3. 独言の類

コード	説明	出現例
_Self	独言	「(目/眼)-に-(物/モノ)-見せ-て-くれる」

4. 形容詞性表現の類

コード	説明	出現例
Adj	末尾が形容詞(終止形)	「愛想-が-(良/善/好)い」
Adjnai	末尾が形容詞+「ない」	「(目/眼)-に-入れ-ても-(痛/イタ)く-ない」

5. 形容動詞性表現の類

コード	説明	出現例
AdjVerb		「(目/眼)-の-前-が-真っ-暗」
AdjVerb_hodo	末尾が「程」	「(穴/アナ)-の-(空/開)く-程」
AdjVerb_you	末尾が「様」	「(顎/アゴ)-が-落ちる-様」

6. 連体修飾(連体詞性)表現の類

コード	説明	出現例
Adnom	連体修飾表現	「有り-と-有らゆる」
Adnom_Adj	末尾が形容詞(連体形)	「当(た)り-障り-の-無い」
Adnom_Noun_no	末尾が名詞+「の」	「(腕/ウデ)-に-覚え-の」
Adnom_Verb	末尾が動詞(連体形)	「足-(下/元/許)-に-火-の-付く」
Adnom_Verb_nai	末尾が動詞+「ない」	「得体-の-知れ-ない」
Adnom_Verb_nu	末尾が動詞+「ぬ」	「(言/云/謂)い-知れ-ぬ」
Adnom_Verb_ta	末尾が動詞+「た」	「打っ-て-変(わ)っ-た」
Adnom_youna	末尾が「様な」	「気-の-遠く-なる-様-な」

7. 連用修飾(副詞性)表現の類

コード	説明	出現例
Adv	連用修飾表現	「明け-でも-暮れ-でも」
Adv_Adj_made	末尾が形容詞+「まで」	「完膚-無き-迄」
Adv_Verb_zu	末尾が動詞+「ず」 副詞的に用いられる場合が多いが、古語的に文末述語として用いられる事もある	「相-も-変(わ)ら-ず」

8. 連結詞(文接続詞)性表現の類

コード	説明	出現例
DiscCon	連結詞(文接続詞)性表現	「早い-話-が」

9. 名詞性表現の類

コード	説明	出現例
Noun	名詞性表現	「赤-の-他人」
Noun/dynamic1	「する」が後接してサ変動詞化するもの	「一-念-発起」
Noun/dynamic2	「を」が後接してサ変動詞化するもの	「一-生-の-御-願い」
Noun/dynamic3	「する」あるいは「を」が後接してサ変動詞化するもの	「暗中-摸索」
Noun/statdescr	物事の様態を表わすもの	「阿吽-の-呼吸」

10. 動詞性表現の類

コード	説明	出現例
Verb	動詞性表現	「ああ-(言/云/謂)え-ば-斯う-(言/云/謂)う」
Verb_nai	末尾が動詞(あるいは助動詞)+「ない」,(注、形容詞の活用をする。)	「開い-た-ロ-が-塞がら-ない」
Verb_nu	末尾が動詞(あるいは助動詞)+「ぬ」(注、連体詞的に用いられる場合が多いが、古語的に文末述語として用いられる事もある。)	「得-も-(言/云/謂)わ-れ-ぬ」
Verb_zu	末尾が動詞(あるいは助動詞)+「ず」,古語的に文末述語として用いられる。	「イザ-知ら-ず」

付録3 本辞書に現れる自立語---英大文字の品詞表記

品詞	コード	品詞	コード	品詞	コード
名詞	N	動詞	V	形容詞	A
形容動詞	K	副詞	D	連体詞	T
接続詞	C	感動詞	E	オノマトペ	O

注 接頭語, 接頭造語要素, 接尾語, 接尾造語要素は, それぞれ, P, Q, S, R と表記

付録4 本辞書に現れる機能語性表現---英小文字の綴り表記

格助詞

綴り	表現	綴り	表現	綴り	表現	綴り	表現
wo	を	ga	が	ni	に	de	で
kara	から	made	まで	ori	より	he	へ
to	と	no	の	no(ga)	の(が格)	made(wo)	まで(を)

副(係)助詞

綴り	表現	綴り	表現	綴り	表現	綴り	表現
ha(wo)	は(を格)	ha(ga)	は(が格)	ha(ni)	は(に格)	mo(wo)	も(を格)
mo(ga)	も(が格)	demo	でも	demo(wo)	でも(を格)	ka	か
bakari	ばかり	dake	だけ	mo	も	ha	は
koso(ga)	こそ(が格)	sae	さえ				

助詞-助詞の組み合わせ

綴り	表現	綴り	表現	綴り	表現	綴り	表現
nimo	に-も	niha	に-は	hemo	へ-も	heto	へ-と
deha	で-は	desae	で-さえ	mademo	まで-も	toha	と-は
tomo	と-も	yorimo	より-も	womo	を-も	ja	じゃ=で-は

接続助詞

綴り	表現	綴り	表現	綴り	表現	綴り	表現
ba	ば	to	と	tara	たら	temo	ても
te	て	nagara	ながら	tari	たり	dari	だり
tomo	とも	domo	ども	teha	て-は	tutu	つつ
tu	つ	ya	や=ばの音便	do	ど	nara	なら

並立助詞

綴り	表現	綴り	表現	綴り	表現	綴り	表現
to	と	ka	か	ya	や	no	の

終助詞

綴り	表現	綴り	表現	綴り	表現	綴り	表現
na	な	yo	よ	yara	やら	ya	や

強意の助詞

綴り	表現
simo	しも=し-も

例: 「無き-に-し-も-((有/在)ら/非)-ず」

特殊

綴り	表現	綴り	表現
gana	がな	kare	かれ=く-あれ

例: 「日-がな-一日」, 「良-かれ-悪し-かれ」

助動詞

綴り	表現	綴り	表現	綴り	表現	綴り	表現
ta	た	da	だ	nai	ない	reru	れる
rareru	られる	seru	せる	rare	られ=「られる」 の連用形	re	れ=「れる」の 連用形
rero	れる=「れる」の命令形	se	せ=「せる」の未然形	u	う	mai	まい

tai	たい	nu	ぬ	tari	たり	taru	たる=「たり」の連体形
nari	なり	nara	なら=「なる」の未然形	nare	なれ=「なる」の仮定/命令形	zu	ず
ne	ね=「ず」の仮定形	zara	ざら=「ざり」の未然形	zaru	ざる=「ざり」の連体形	gotosi	ごとし
gotoku	ごとく=「ごとし」の連用形	na	な=「だ」の連体形	besi	べし	beki	べき=「べし」の連体形
bekara	べから=「べかり(推量)」の未然形	q	ん=「ず、ぬ(打消)」の連体形「ぬ」、または「む(推量)」の終止/連体形「む」	ri	り(完了)		

支援動詞

綴り	表現	綴り	表現	綴り	表現	綴り	表現
suru	する	si	し=「する」の未然/連用形	se	せ=「する」の未然形	sa	さ=「する」の未然形
sure	すれ=「する」の仮定形	seyo	せよ=「する」の命令形	aru	ある	at	あつ=「ある」の連用形(音便)
aro	あろ=「ある」の未然形	are	あれ=「ある」の仮定/命令形	nasu	なす	naru	なる
nat	なつ=「なる」の連用形(音便)	nari	なり=「なる」の連用形	su	す	ari	あり=「ある」の連用形

支援形容詞

綴り	表現	綴り	表現	綴り	表現	綴り	表現
nai	無い	naku	無く=「無い」の連用形	nasi	無し	naki	無き=「無し」の連体形

形容動詞性表現を与える機能性名詞

綴り	表現	綴り	表現
you	様	hodo	程

付録5 活用の記号系

動詞

	現代語			古語		
	語尾	コード	備考	語尾	コード	備考
語幹	---	V00		略	V0'0	
未然形	略	V11	ナイ接続			
	略	V12	ヌ、ズ接続	略	V1'2	ヌ、ズ、ジ、バ、シム接続
	略	V14	ウ、ヨウ接続			
連用形	略	V15	レル、ラレル、セル、サセル接続	略	V1'5	ラル、ル接続
	略	V22	中止用法、名詞化用法	略	V2'2	ツ、タリ接続
	略	V23	タ、ダ、テ、デ接続			
音便形	V23'	タ、ダ、テ、デ接続				
終止形	略	V30		略	V3'0	
連体形	略	V40		略	V4'0	
仮定形	略	V50	バ接続	略	V5'0	バ接続、ド接続
命令形	なし	V60		略	V6'0	リ(完了)接続
	ろ	V61				
	よ	V62		略	V6'2	

形容詞

	現代語			古語		
	語尾	コード	備考	語尾	コード	備考
語幹	---	A00			A0'0	
未然形				「から」	A1'2	ヌ、ズ接続
	「かろ」	A13	ウ接続			
連用形	「く」	A22	中止用法、ナイ接続	「く」、「かり」、「しく」、「しかり」	A2'2	
	「う」音便形	A22'	ナイ接続	「う」音便形、「しゅう」音便形	A2'2'	ナイ接続
	「かつ」	A23	タ接続			

終止形	「い」	A30		「し」	A3'0	
連体形	「い」	A40		「き」、「かる」、「しき」、「しかる」	A4'0	
仮定形	「けれ」	A50	バ接続	「けれ」、「しけれ」	A5'0	バ、ド接続
命令形	「かれ」	A60		「かれ」、「しかれ」	A6'0	

形容動詞

	現代語			古語		
	語尾	コード	備考	語尾	コード	備考
語幹	---	K00			K0'0	
未然形	「だろ」		ウ接続、助動詞「だ」の未然形	「なら」		ヌ、ズ接続、助動詞「なり」の未然形
				「たら」		ヌ、ズ接続、助動詞「たり」の未然形
連用形	「で」		中止用法、ナイ接続、助動詞「だ」の連用形	「なり」		助動詞「なり」の連用形
	「だつ」		タ接続、助動詞「だ」の連用形	「たり」		助動詞「たり」の連用形
	「に」		助動詞「だ」の終止形	「と」		助動詞「たり」の連用形
終止形	「だ」		助動詞「だ」の終止形	「なり」		助動詞「なり」の終止形
連体形	「な」		助動詞「だ」の連体形	「たり」		助動詞「たり」の終止形
	「たる」		助動詞「たり」の連体形	「たる」		助動詞「たり」の連体形
	「なる」		助動詞「なり」の連体形	「なる」		助動詞「なり」の連体形
仮定形	「なら」		バ接続、助動詞「だ」の仮定形	「なれ」		バ接続、ド接続、助動詞「なり」の仮定形
				「たれ」		バ接続、ド接続、助動詞「たり」の仮定形
命令形				「なれ」		助動詞「なり」の命令形
				「たれ」		助動詞「たり」の命令形