

国立国語研究所学術情報リポジトリ

An automatic processing of “conjugation”

メタデータ	言語: jpn 出版者: 公開日: 2017-03-31 キーワード (Ja): キーワード (En): 作成者: 江川, 清, EGAWA, Kiyosi メールアドレス: 所属:
URL	https://doi.org/10.15084/00000997

「活用形処理」の自動化における一方式

江 川 清

現在、国立国語研究所において、Computer を利用した「新聞の語い調査」が進行中である。そこでは、「入力データの自動分ち書き」および「漢字の読みがなづけ」などを始めとして多くの面で作業の自動化が進められているが、まだ未解決のまま残されている課題も多い。

その中の1つに、活用形の代表形(基本形)集約の機械化という課題がある。

例えば、「行く」という動詞について調べてみると、この動詞は文中のある箇所では「行く」という形で、また、別の箇所では「行っ」あるいは「行け」などに活用して、使用されていることがわかる。これらの「行く」「行っ」および「行け」は語い的意味という面では、同一であり、代表形を「行く」という形で統合されうるものである。

語い調査の中では、これらの語形の異なる活用形を代表形に集約させて、処理したり、カウントしたりする必要がある。

ところが、Computer にこれらの語形の異なる活用形を処理させようとする場合、通常、これらの語は別語とみなされてしまう。換言すれば、Computer にとっては、これらの各語形が「行く」という形で代表される語の活用形であるという認識が困難であるといえる。

ここに、「活用形の処理のアルゴリズム」の研究の必要性が存在する。

活用形処理システム

活用形を代表形に変換するためのアルゴリズムとしては、種々のものが考えられるが、本論文では、入力データである活用形が漢字かな混じりであり(必ずしも重要な要件ではないが)、かつ、語の認識、即ち、単位切りがなされて

いるという前提に立って、「活用形変換ルーチン」を考える。

この場合、大雑端に考えてみても、以下の4つの方式が考えられる：

1. 入力データを一旦、ローマ字に変換して、活用語尾の処理を行ない、しかる後に、再び原表記に変換するといった方式。
2. 「自立語—付属語間の接続様式」に着目する方式
3. 活用語テーブル（一種の辞書）を利用して変換を行なう方式
4. あらかじめ、人手によって付加されている何らかの文法情報（特に、活用情報）を変換の手がかりにする方式

本論文では、さいわい、新聞の語い調査システムの中で、Pre-edit の1つとして、「語種、品詞、活用」などに関する文法情報（我々の用語で言えば、「付加情報」）が与えられているので、それを利用する第4の方式を用いる。

この方式を用いた別の理由として、現時点では辞書の作成が困難であること、および、語い調査のシステムの中で、既に、自立語と付属語とが分断されてしまっているため、第2、第3の方式を用いることが不可能であるという消極的な理由もある。

A. 付加情報

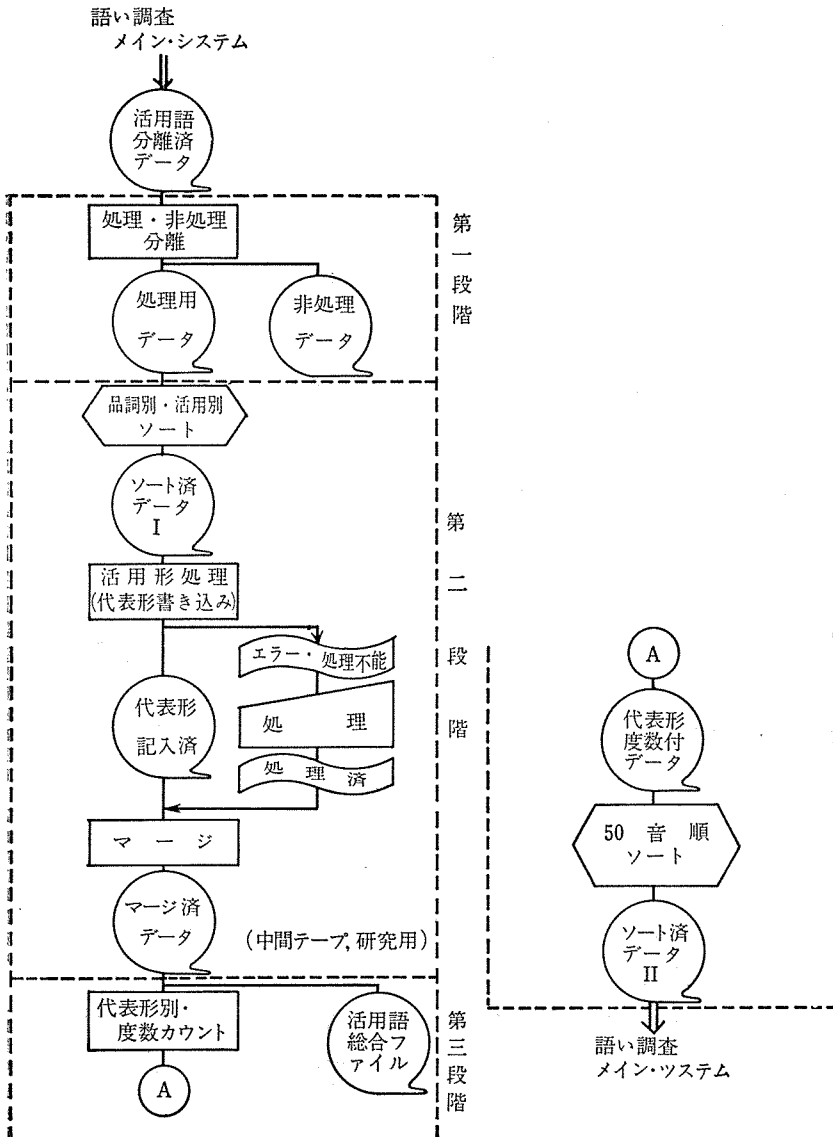
付加情報の内容については、本報告書の中で中野（P38—54）が紹介しているので、詳しい説明は省いて、ここでは、活用形処理に関連するものに限って記述する。

活用形処理における relevant な情報は、品詞情報と活用情報との2つである。

1) 品詞情報

品詞情報に関しては、全部で23種の Code が用意されているが、ここで必要なものは、動詞、形容詞、助動詞（形容動詞の語尾を含む）などを示す情報の他に、動詞性接辞（例えば、「食べ／すぎる」あるいは「走り／疲れる」における「すぎる」「疲れる」など）および形容詞性接辞（「書き／にくい」「感じ／やすい」の「にくい」「やすい」など）を含む5つの Code である。

図1 活用形処理の BLOCK-CHART



2) 活用情報

活用情報には、2種類の情報（それぞれを活用情報「A」「B」とする）がある。動詞においては、活用情報「A」は五段・下一段などの活用の型を示すものであり、活用情報「B」は活用の「行」を示すものである。例えば、「居る」という活用形に対しては、「上一段（活用）ーア行」という形で情報づけがなされており、「書け」という活用形には、「四、五段ーカ行」という情報が付加されている。

この情報は、動詞および動詞性接辞について考慮されたものであるが、形容詞に対しては、文語、口語の別が、また、助動詞については、助動詞そのものと形容動詞の語尾との区別がなされている。

B. システムの概要

図1は、活用形処理の Block Chart である。

語い調査の main systemある段階で、活用語情報を含むデータが全データ中より分離されて、このシステムに送り込まれる（図1の活用語分離済データ）。

1) 第1段階：代表形交換への準備

まず、活用形分離済データの中から、今回の処理の対象となるもの（処理用データ）と処理対象から除外される「非処理データ」とに分割される。

ここでいう非処理データとは、下記の基準の1項以上に該当するものの意味である。

[非処理データの基準]

1. 助動詞および形容動詞（の語尾）

現在の語い調査システムの中では、助動詞の代表形別カウントまでは要求されていないこと、および助動詞には同形異語が多いため、代表形変換が困難であることなどから助動詞の処理を除外した。従って、形容動詞の語尾もこれと同列に扱った。

2. 付加情報が同時に2組み以上与えられているもの

付加情報は本来ならば、1語について1組みの情報しか付加されえない

はずのものであるが、作業手順の関係から、中には複数の付加情報を有するデータも混入している。従って、このような語を代表形に変換しても、あまり重要な意味がないと考えられるからである。

3. 活用の形が明確でないため、代表形を決めたいもの

これは小説における会話文などのように話ことばをそのまま記述しようとした際に生じるものであり、「いらしっ／た」「みりゃ／いいんだ」「すりゃ／いいんだ」の「いらしっ」「みりゃ」「すりゃ」などがその例である。

従って、ここで処理の対象となるデータは、品詞について言えば、動詞、形容詞、動詞性接辞および形容詞性接辞に限られており、かつ、代表形が一義的に決定されうるものである。

2) 第2段階：代表形書込みルーチン

Input data である活用形を代表形（本論文では、終止形）に変換する際に、品詞および活用の段（五段とか上一段などの）あるいは活用の行の差異によって、処理の様式が異なってくる。

例えば、四・五段活用の動詞においては、特殊な活用形^(注1)を除いては、一律に最後の1文字のみ（即ち活用語尾）を代表形語尾に置き換えるだけで、目的が果たせるのに対して、上一・下一段の動詞においては、未然・連用両活用形の場合（活用語尾は最後の1文字のみ）とそれ以外（活用語尾は最後の2文字）との場合では処理の手続が異なる。また、同じ下一段の動詞の場合であっても、ラ行とそれ以外の行の活用形では別途の処理が要求される。

このような理由から、データが品詞情報別・活用情報別にグルーピングされるように配列（ソート）されてあれば、処理の能率を高めることができることがわかる。（図1の品詞別、活用別ソート）。

ソートされた結果、四・五段活用動詞群を先頭に、以下、上一・下一……変格（更に、カ変・サ変……ラ変の順）と並び、最後に形容詞（口語・文語の順）が並んでいる磁気テープが作成される（ソート済データⅠ）。

（注1）新聞の紙面では、広告欄に多く見られる形で、「社保寮有」あるいは「乞来社」などの「有」「乞」などのように、活用語尾が省略されているもの。

以下、このテープが入力データとなる。

ソート済データ I M/T FORMAT

7 桁	5	40	40	3	21	1	10+V	V	V	E/I
度数	配列情報 (B0)	仮名文字 (B1)	漢テレ文字 (B2)	インデックス	付加情報	位置情報	仮名文字 (A1)	漢テレ文字 (A2)	ルビ付漢テレ文字 (B3)	

(注) 付加情報は

個数 1桁

語種コード1 }
品詞コード1 } ×5の構成になっている。
活用コード2 }

付加情報は最大5組みまで認められている。

またVは VARIABLE の意味である。

〔代表形書き込みの手順〕

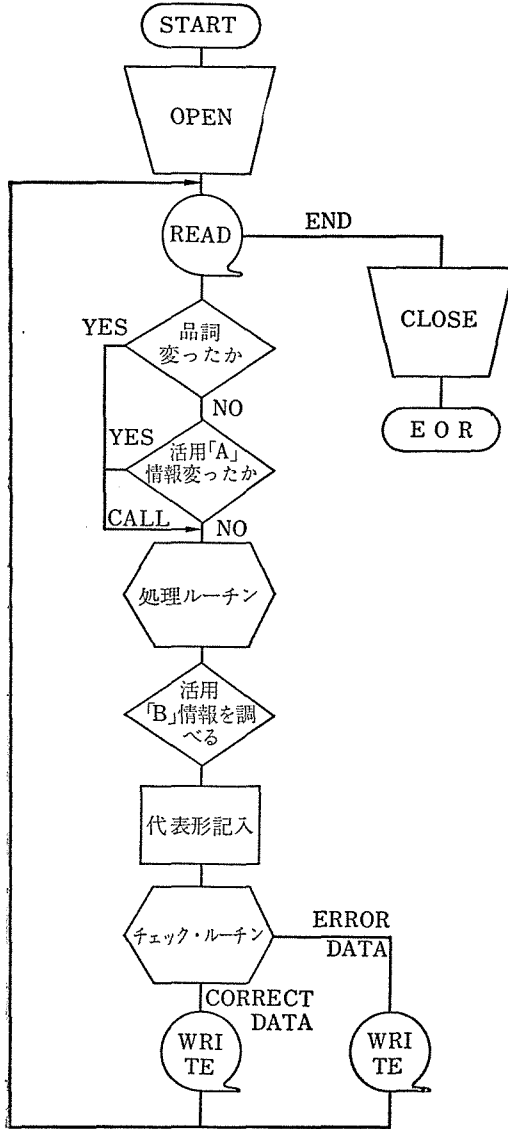
代表形の記入は、原則として、以下の順序で進められる(図2参照)。

- a. データ(ソート済データI)を読み込む。
- b. 活用「A」情報を判定し、必要な処理ルーチン呼び出す(注2)。
- c. 活用「B」情報を判定し、代表形(終止形)語尾をストアする。
- d. 入力データの活用語尾を発見し、それを代表形語尾に変換する。
- e. 次にエラー・データのチェックをする(error 検出の方法については、後述)。
- f. (i) もし、データが正しければ、6つの活用形(未然形あるいは終止形など)(注3)の情報を付加し、所定の磁気テープ(代表形記入済データ)に記入し、次のデータを読む。(ii) エラー・データの場合は、別途に処理し、次のデータを読む。

(注2) 実際の処理においては、1データ毎に処理ルーチン呼び出すのではなく、同じ活用「A」情報を有する活用形が一箇所に集まっているので、情報が異なる段階までは同じ処理ルーチンを通ることになる。

(注3) 同じ、活用形の場合であっても語尾が異なるものには、異なった情報が付加されている。例えば、四五段活用の連用形の音便形とそれ以外では別の情報が付加されている。

図2 代表形書き込みルーチン



代表形記入済データ M/T FORMAT

7	5	40	40	40	1	40	3	21	1	10+V	V	V	E/I
度数	配列情報 (B0)	仮名文字 (B1)	漢テレ文字 (B2)	仮名代表形 (R1)	活用形情報	漢テレ代表形 (R2)	インデックス	付加情報	位置情報	仮名文字 (A1)	漢テレ文字 (A2)	ルビ付漢テ (B3)	

以上の処理が終了した段階で、エラー・データおよび機械処理が不能であったデータ（後述）を人手によって、修正し、それを「代表形記入済データ」に書き加える（マージ）。

3) 第3段階：メイン・システムへの引き渡し

以降の処理は、語い調査の main system へ引き渡すための作業である。

各活用語を代表形別にまとめ、その使用度数を合計する（代表形度数付データ）。

代表形度数付データ M/T FORMAT

7	5	40	40	3	21	E/I
度数	配列情報 (B0)	仮名代表形 (R1)	漢テレ代表形 (R2)	インデックス	付加情報	

最後に、代表形を50音順に配列し（ソート済データⅡ）、main system に引き渡す。

[ERROR CHECK ROUTINE の構成]

機械処理において、最も難解なもの1つが、データ・チェックであるといえる(注4)。勿論、ここで問題にしているエラー・チェックはパリティ・エラー

(注4) 語い調査システムにおける一般的なエラーチェックについては既に木村によって述べられている。(木村繁 1968「漢テレ・入力データのチェック」国立国語研究所報告31「電子計算機による国語研究」pp 133-150)

などの主としてメカ的な原因によるものではなく、むしろ、付加情報をつける際の人間による誤りである。

そのためには、人間の誤りのパターンを心理学的に研究しておく必要があるが、ここでは、人間の誤りは正反応に近いところで生ずることが多いという単純な前提に立って、エラーの検出にせまりたいと思う。言い換えれば、もし「書く」という活用形には付加情報を付加する場合、「動・五・ガ行」という情報を付加するという誤りは生じても、「動・上一・タ行」というような誤りを犯すことはめったにないと考えられる。

その前提にたつて、活用形の語尾についてみてみると、そのパターンはごく限られたものであることがわかる。そうすると、活用語尾と付加情報との関係から、エラーの検出をはかることも可能であると思われる。

具体的には、ある付加情報を有する語、例えば、「書く（動・五・カ行）」についていえば活用語尾「く」の適否を判定するため「活用語尾テーブル」^(注5)を作成しておけばよい。

このテーブルによって、以下の種類のエラー・チェックが可能となる。

例

活用形	誤った付加情報
書く	(動・五・マ行) / (動・上一・ア行)
似	(動・五・ナ行) / (動・下一・ナ行)
対する	(動・変・ザ行) / (動・変・カ行)
⋮	

しかし、このようなチェックでも、次のような3つの型の誤りは見のがされてしまうことになる。

例1

活用形	誤った付加情報とそれに基づいて変換される形
書い	(動・五・ガ行)→書く/(動・上一・テ)→書いる/(形・口語)→書いて
早く	(動・五・カ行)→早く/(動・上二・カ)→早く
⋮	

この種のエラーが検出されえない原因は、入力データにおける活用語

(注5) 「活用語尾テーブル」の原型を付表1, 2に示しておいた。

尾が誤った付加情報が要求する活用語尾の範囲にも入っているということから生ずる。例えば例1の「書い」の付加情報の誤りがチェックされない理由は「書い」の活用語尾「い」は五段動詞カ行連用形「イ音便」の活用語尾であると同時に五段動詞のガ行連用形「イ音便」あるいは上一段動詞ア行（未然・連用形）ないしは口語形容詞（終止・連体形）などの活用語尾でもありうるからである。従って、そのような誤った付加情報に基づいて表形変換を行なった結果、表に示したような奇妙な代表形「書く」「書いる」「書い」などが存在することになる。

例2 本来は、ザ変動詞であるべきはずの「信じ（代表形は‘信ずる’）」に対して、上一段（代表形は‘信じる’）の付加情報が付加されてしまったような類のもの。即ち、同形異語のチェックは不可能となっている。

本システムのように、語単位で入力される方式の中では、これらのエラーを検出することは、ほとんど不可能に近い。

例3

活用形	誤った付加情報とそれに基づいて変換される形
切る	動・上一・カ→切る
着る	動・五・カ →着る
⋮	

この型のエラーは、代表形への変換という点では何ら支障はないが、本来の活用語とは別語としてカウントされることになる。(注6)

〔処理不能なデータ〕

数は少ないが処理不能なデータ（または処理を保留したデータ）も存在する。これは主として、形容詞に関係するものである。その典型的な例として、形容詞ウ音便の処理をあげることができる。

例えば、「寒い」の音便形は「寒う」であるのに対し、「赤い」「正しい」はそれぞれ「赤こう」、「正しゅう」となっている。この音便の現象には、一

(注6) 我々の「語い調査」では、語形、付加情報の総てが一致したものを同語とみなす約束になっているからである。

往の法則が存在するが、この法則は「代表形→音便形」変換にとっては有力であるが、その逆の「音便形→代表形」への方向の変換に対しは決め手になりえない。(註7)そのため、今回の処理はあきらめざるをえなかった。

また、語形の不完全な語(例えば、「良/すぎる」の「良」)の処理も、今回は見送った。

活用形処理によるOut-put

以上の処理の結果から、現在でも得ることのできる Out-put の一例をあげておく。

1. 代表形別・活用形別度数表

代表形	活用形		漢テレコード	度数	度数計
カク (動・五カ行)	未然形		書く		pt
					P ₁
	連用形	カカ	書か	f ₁	P ₂
		カコ	書こ	f ₂	
	終止・連体形	カキ	書き	f ₃	P ₃
		カイ	書い	f ₄	
假定・命令形	カク	書く	f ₅	P ₄	
	カケ	書け	f ₆		
マルイ (形・口語)	未然形		まるい		pt
					P ₁
	連用形	マルカロ	まるかろ	f ₁	P ₂
		マルク	まるく	f ₂	
		マルウ	まるう	f ₃	
	終止・連体形	マルカッ	まるかっ	f ₄	P ₃
假定形	マルイ	まるい	f ₅	P ₄	
	マルケレ	まるけれ	f ₆		

2. 品詞別・活用形別度数表

品詞	活 用		度数	度数計	異なり語数	異なり語数計
動詞	四・五段	未然形	f_1	pt p ₁	n ₁ n ₂ n ₃ n ₆	mt m ₁
		連用形	f_2			
		終止・連体形	f_3			
		假定・命令形	f_4			
形容詞	文語	未然・連用形	f_1	P ₁ p ₁	n ₁ n ₂ n ₃ n ₄ n ₅	mt m ₁
終止形	f_2					
連体形	f_3					
已然形	f_4					
命令形	f_5					

3. 同音・異表記語一覧表

音	品詞	活用形		漢テレコード	度数	度数計
トレ	動詞	四・五	假定命令	とれ	f_1	P ₁ P _m P ₁
				トレ	f_2	
				取れ	f_3	
				採れ	f_4	
				捕れ	f_5	
				執れ	f_6	
				撮れ	f_7	
		下一	未然連用	とれ	f_1	P ₂
			トレ	f_2		
			取れ	f_3		
			捕れ	f_4		

この表では、活用語のみが扱われているが、語彙調査の main system に残されている、非活用語ファイルを利用することによって、完全な同音異表記語

の一覧表を作成することができる。

4. 以上の他にも、種々の Out-put が可能である。

また、これらの表の場合でも、度数順あるいは50音順など、種々の形に変形して、Out-put することも可能である。更に、必要ならば、使用順位、使用比率その他の情報をつけ加えることもできる。

今後の課題

1. 本研究で開発された「代表形変換ルーチン」は、活用語の語形と付加情報の利用を特徴としているために、変換が容易になっている反面、それ故に内在する問題点も多い。

ここでは、付加情報に関する問題点の指摘にとどめる。

第1に、先に述べたような付加情報のつけ誤りによって、本来、変換さるべき代表形とは異なった代表形が産みだされることになる。

第2に、「短単位」語（いわゆる単語単位よりも細かい言語単位）に分割された段階で、付加情報が付加されるという語い調査の作業手順から生じる問題点もある。例えば、「……そう／なすっ／た／のに……」という文があった場合、短単位では／の部分で分割されることになっている。そこで「なすっ」という活用語だけを見て付加情報を付加することになるので、代表形「なする」が想定され、それに対する付加情報が与えられてしまう。しかし、これは本来は、非処理データの基準で述べた第3項——活用の形が明確でないもの——の1例である。勿論、Computer は付加情報に従うので「なすっ」を「なする」に変換して、処理を終える。その結果、1で生じたと同様の問題が起こる。また、先に述べた「信じ」あるいは「き／ない」の「き」も同様の例である。

同じ理由から、「かく」という語に対しては、動詞の「書く」「欠く」など

(注7) 形容詞が「ウ音便」化する場合は、(1)語幹が母音「a」で終るものは、それが「o」に変化した上で「u」が付与される(2)語幹末が「キ」か「シ」のものは「ウ」が「ェウ」に変化する(3)1、2以外のものは、単に「ウ」が付与されるという法則がある。しかし、この法則を「音便形→代表形」の変換に適用しようとした場合、「アオ(青)ウ」が「アアイ」に変換されることになる。

の他に名詞の「核」「画」「格」などが想定される結果、複数個の付加情報をもつ語がつくりだされることになる。

本システムのように、1語1語が切り離され、語単位で入力されている現状では、これらの問題点を解決することは不可能に近い。(自立語と付属語とを同時に処理することで部分的に解決される場合があるとしても)。この点から現状では非常に困難ではあるが、本格的な言語の自動処理へ進むためには、センテンス単位での言語処理が要求されてくるといえる。

2. 現在の語い調査システムの一作業としての「活用形処理」という観点からみれば、各活用形を代表形に変換するためのルーチン、即ち、「代表形変換ルーチン」を開発したことで一往の役割を果たしていることになる。

しかし、将来の発展を考えるならば、この問題を単なる「代表形変換 Routine」の開発のみに限定してしまうだけではなく、「言語情報処理 (LDP)」あるいは Computational Linguistics (CL) という広い視点から考えておく必要がある。

第1には、この問題は、終局的には、現在、LDP あるいは CL 研究の中で課題の1つとなっている「同語異語の判別」につながる問題であり、直接的には、その中の「同形異語の判別」への1ステップであるとも考えられる。

第2に、「代表形変換」からの産物——磁気テープ (M/T) ファイル——は、これだけでも今後、活用語テーブルとして利用されうるが、これを更に拡大、発展させることによって、活用語に関する豊富な情報を含有しうるファイルとなり、ひいては、自動構文解析あるいは機械翻訳などの一部に寄与しうる可能性を有するものとなると考えられる。——たとえば、「活用語総合ファイル」とでも称せられる形において。今後、これらの面へのアプローチを続けていきたい。

付表

Computer で活用語尾を見いだそうとする場合、最後の文字から探索していく方が、能率的である。そこで、附表1、附表2の表が作成された。

付表1は、活用語尾を50音順配列し、その語尾をとる品詞と対応させた図表である。付表2は、付表1に、一層詳しい情報をつけ加え、逆50音順に配列し直したものである。

付表1 50音順活用語尾一覧表

活用語尾	品 詞	活用語尾	品 詞
イ	形。上一。四。	ケル	下一。
イヨ	上一。	ゲル	下一。
イル	上一。	ケレ	形。下一。
イレ	上一。	ゲレ	下一。
イロ	上一。	ケロ	下一。
ウ	助動。形。四。	ゲロ	下一。
エ	四。下一。	コ	四。カ変
エヨ	下一。	ゴ	四。
エル	下一。	コイ	カ変
エレ	下一。		
エロ	下一。	サ	四。サ変。
オ	四。	サセ	助動。
カ	四。	サセヨ	助動。
カッ	形。	サセル	助動。
カロ	形。	サセレ	助動。
ガ	四。	サセロ	助動。
キ	上一。四。カ変	シ	四。サ変。
ギ	上一。四	ジ	上一。サ変。
キヨ	上一。	シメ	助動。
ギヨ	上一。	シメヨ	助動。
キル	上一。	シメル	助動。
ギル	上一。	シメレ	助動。
キレ	上一。	シメロ	助動。
ギレ	上一。	ジヨ	上一。
キロ	上一。	ジル	上一。
ギロ	上一。	ジレ	上一。
ク	四。形。	シロ	サ変
グ	四。	ジロ	上一。サ変。
クル	カ変。	ス	四。文サ変。
クレ	カ変	ズ	助動。
ケ	下一。四。	スル	サ変。
ゲ	下一。四。	ズル	サ変。
ケヨ	下一。	スレ	サ変。
ゲヨ	下一。	ズレ	サ変。

セ	助動。四。サ変。下一。	ナ	助動。形動。四。
ゼ	サ変。下一。	ナイ	助動。
セヨ	助動。サ変。下一。	ナカッ	助動。
ゼヨ	サ変。下一。	ナカロ	助動。
セル	助動。下一。	ナク	助動。
ゼル	下一。	ナケレ	助動。
セレ	助動。下一。	ナラ	助動。形動。
ゼレ	下一。	ニ	助動。形動。四。上一。
セロ	助動。下一。	ニヨ	上一。
ゼロ	下一。	ニル	上一。
ソ	四。	ニレ	上一。
タ	助動。四。	ニロ	上一。
ダ	助動。形動。	ヌ	助動。四。
タイ	助動。	ネ	助動。四。下一。
タカッ	助動。	ネヨ	下一。
タカロ	助動。	ネル	下一。
タク	助動。	ネレ	下一。
タケレ	助動。	ネロ	下一。
ダッ	助動。形動。	ノ	四。
タラ	助動。		
タロ	助動。	バ	四。
ダロ	助動。形動。	ビ	上一。
チ	上一。四。	ビ	上一。四。
チヨ	上一。	ビヨ	上一。
チル	上一。	ビヨ	上一。
チレ	上一。	ヒル	上一。
チロ	上一。	ビル	上一。
ソ	四。	ヒレ	上一。
テ	四。下一。	ビレ	上一。
デ	助動。形動。下一。	ヒロ	上一。
デシ	助動。	ビロ	上一。
デシヨ	助動。	ブ	四。
デス	助動。	へ	下一。
テヨ	下一。	べ	四。下一。
デヨ	下一。	へヨ	下一。
テル	下一。	べヨ	下一。
デル	下一。	ヘル	下一。
テレ	下一。	ベル	下一。
デレ	下一。	へレ	下一。
テロ	下一。	ベレ	下一。
デロ	下一。	へロ	下一。
ト	四。	ベロ	下一。
トウ	助動。	ボ	四。

マ	四。	ラシカツ	助動。
マイ	助動。	ラシク	助動。
マシ	助動。	ラシケレ	助動。
マシヨ	助動。	ラシユウ	助動。
マス	助動。	ラレ	助動。
マスレ	助動。	ラレヨ	助動。
マセ	助動。	ラレル	助動。
ミ	四。上一。	ラレレ	助動。
ミヨ	上一。	ラレロ	助動。
ミル	上一。	リ	上一。四。
ミルレ	上一。	リヨ	上一。
ミロ	上一。	リル	上一。
ム	四。	リレ	上一。
メ	四。下一。	リロ	上一。
メヨ	下一。	ル	四。
メル	下一。	レ	助動。四。下一。
メレ	下一。	レヨ	助動。下一。
メロ	下一。	レル	助動。下一。
モ	四。	レレ	助動。下一。
ユウ	形	レロ	助動。下一。
ヨウ	助動。	ロ	四。
ラ	四	ワ	四。
ラシイ	助動。	ン	助動。四。

付表2 逆引50音順活用語尾一覧表

活用語尾	代表形語尾	品詞・活用形	操作機能
イ	イ	形A・B・C・終・体	0
	イル	ア上一未1・2・4・5用1・2	＋ル
	ウ	ワア四用1	→ウ
	ク	カ四A用2	→ク
	グ	ガ四用2	→グ
	ル	ラ四C命1・2	→ル
コイ	クル	カ変命1	→クル
ラシイ	ラシイ	助動ラシイ終・体	0
タイ	タイ	助動タイ終・体	0
ナイ	ナイ	助動ナイ終・体	0
マイ	マイ	助動マイ終・体	0
ウ	イ	形A用2	→イ
	ウ	ワア四終・体	0
ウ	ウ	助動ウ終・体	0
コウ	カイ	形B用3	→カイ

トウ	タイ	助動タイ用3	→タイ
ラシユウ	ラシイ	助動ラシイ用3	⇒イ
ユウ	イ	形C用4	→イ
ヨウ	ヨウ	助動ヨウ終・体	0
エ	ウ	ワア四仮・命1・2	→ウ
	エル	ア下一未1・2・4・5用1・2	＋ル
オ	ウ	ワア四未5	→ウ
カ	ク	カ四A・B未1・2・3	→ク
ガ	グ	ガ四未1・2・3	→グ
キ	キル	カ上一未1・2・4・5用1・2	＋ル
	ク	カ四A・B用1	→ク
	クル	カ変用1・2	→クル
ギ	ギル	ガ上一未1・2・4・5用1・2	＋ル
	グ	ガ四用1	→グ
ク	ク	カ四A・B終・体	0
	イ	形A・B・C用1	→イ
ラシク	ラシイ	助動ラシイ用1・2	⇒イ
タク	タイ	助動タイ用1・2	⇒イ
ナク	ナイ	助動ナイ用1	⇒イ
グ	グ	ガ四終・体	0
ケ	ク	カ四A・B仮・命1・2	→ク
	ケル	カ下一未1・2・4・5用1・2	＋ル
ゲ	グ	ガ四仮・命1・2	→グ
	ゲル	ガ下一未1・2・4・5用1・2	＋ル
コ	ク	カ四未A・B5	→ク
	クル	カ変未1・2・4・5	→クル
ゴ	グ	ガ四未5	→グ
サ	ス	サ四未1・2・3	→ス
	スル	サ変A未3	→スル
シ	ス	サ四用1・2	→ス
	スル	サ変A・B未2・5用1・2	→スル
デシ	デス	助動デス用4	⇒ス
マシ	マス	助動マス用4命2	⇒ス
ジ	ジル	ザ上一未1・2・4・5用1・2	＋ル
	ズル	ザ変未2・5用1・2	→ズル
ス	ス	サ四終・体・文サ変終	0
デス	デス	助動デス終	0
マス	マス	助動マス終・体	0
ズ	ヌ	助動ヌ用1・2	→ヌ
セ	ス	サ四仮・命1・2	→ス
	スル	サ変A・B未1・4	→スル
	セル	サ下一未1・2・4・5用1・2	＋ル

セ サセ マセ ゼ ソ	セル サセル マス ズル ゼル ス	助動セル未1・2用2・3・4 助動サセル未1・2用2・3・4 助動マス未1命1 ザ変未1・4 ザ下一未1・2・4・5用1・2 サ四未5	十ル 十ル →ス →ズル 十ル →ス
タ タ ダ ダ チ ツ ツ カッ ラシカッ タカッ ナカッ ダッ ダッ テ デ デ ト	ツ タ ダ ダ チル ツ ツ ウ ク ツ ル イ ラシイ タイ ナイ ダ ダ ツ テル ダ デル ダ ツ	タ四未1・2・3 助動タ終・体 形動A・B終 助動ダ終 タ上一未1・2・4・5用1・2 タ四用1 タ四終・体 ワア四用2 カ四B用2 タ四用2 ラ四A・B・C用2 形A・B・C用5 助動ラシイ用4 助動タイ用4 助動ナイ用4 形動A・B用1 助動ダ用4 タ四仮・命1・2 タ下一未1・2・4・5用1・2 形動A・B用2 ダ下一未1・2・4・5用1・2 助動ダ用2・3 タ四未5	→ツ 0 0 0 十ル →ツ 0 →ウ →ク →ツ →ル →イ →イ →イ →イ →ダ →ダ →ツ 十ル →ダ 十ル →ダ →ツ
ナ ナ ニ ニ ヌ ヌ ネ ネ ノ	ダ ヌ ダ ダ ヌ ニル ダ ヌ ヌ ヌ ネル ヌ ヌ	形動A体 ナ四未1・2・3 助動ダ体 形動A・B用3 ナ四用1 ナ上一未1・2・4・5用1・2 助動ダ用1 ナ四終・体 助動ヌ終1・体1 ナ四仮・命1・2 ナ下一未1・2・4・5用1・2 助動ヌ仮 ナ四未5	→ダ →ヌ →ダ →ダ →ヌ 十ル →ダ 0 0 →ヌ 十ル →ヌ →ヌ

バ	ブ	バ四未1・2・3	→ブ
ヒ	ヒル	ハ上一未1・2・4・5用1・2	＋ル
ビ	ビル	バ上一未1・2・4・5用1・2	＋ル
	ブ	バ四用1	→ブ
ブ	ブ	バ四終・体	0
へ	ヘル	ハ下一未1・2・4・5用1・2	＋ル
べ	ベル	バ下一未1・2・4・5用1・2	＋ル
	ブ	バ四仮・命1・2	→ブ
ボ	ブ	バ四未5	→ブ
マ	ム	マ四未1・2・3	→ム
ミ	ム	マ四用1	→ム
	ミル	マ上一未1・2・4・5用1・2	＋ル
ム	ム	マ四終・体	0
ヌ	ム	マ四仮・命1・2	→ム
	メル	マ下一未1・2・4・5用1・2	＋ル
シメ	シメル	助動シメル未1・2用2・3・4	＋ル
モ	ム	マ四未5	→ム
イヨ	イル	ア上一命2	→ル
エヨ	エル	ア下一命2	→ル
キヨ	キル	カ上一命2	→ル
ギヨ	ギル	ガ上一命2	→ル
ケヨ	ケル	カ下一命2	→ル
ゲヨ	ゲル	ガ下一命2	→ル
ジヨ	ジル	ザ上一命2	→ル
セヨ	スル	サ変A・B命2	→スル
	セル	サ下一命2	→ル
セヨ	セル	助動セル命2	→ル
サセヨ	サセル	助動サセル命2	→ル
ゼヨ	ズル	ザ変命2	→ズル
	ゼル	ザ下一命2	→ル
チヨ	チル	タ上一命2	→ル
テヨ	テル	タ下一命2	→ル
デヨ	デル	ダ下一命2	→ル
ニヨ	ニル	ナ上一命2	→ル
ネヨ	ネル	ナ下一命2	→ル
ヒヨ	ヒル	ハ上一命2	→ル
ビヨ	ビル	バ上一命2	→ル
ヘヨ	ヘル	ハ下一命2	→ル
ベヨ	ベル	バ下一命2	→ル
ミヨ	ミル	マ上一命2	→ル
メヨ	メル	マ下一命2	→ル
シメヨ	シメル	助動シメル命2	→ル
リヨ	リル	ラ上一命2	→ル

レヨ	レル	ラ下一命2	⇒ル
レヨ	レル	助動レル命2	⇒ル
ラレヨ	ラレル	助動ラレル命2	⇒ル
デシヨ	デス	助動デス未2	⇒ス
マシヨ	マス	助動マス未2	⇒ス
ラ	ル	ラ四A未1・2・3 B未1・3 C未1・2	→ル
タラ	タ	助動タ仮	→タ
ナラ	ダ	形動A・B仮	→ダ
ナラ	ダ	助動ダ仮	→ダ
リ	ル	ラ四A・B・C用1	→ル
	リル	ラ上一未1・2・4・5用1・2	→ル
ル	ル	ラ四A・B・C終・体	0
イル	イル	ア上一終・体	0
エル	エル	ア下一終・体	0
キル	キル	カ上一終・体	0
ギル	ギル	ガ上一終・体	0
クル	クル	カ変終・体	0
ケル	ケル	カ下一終・体	0
ゲル	ゲル	ガ下一終・体	0
ジル	ジル	ザ上一終・体	0
スル	スル	サ変A・B終・体	0
ズル	ズル	ザ変終・体	0
セル	セル	サ下一終・体	0
セル	セル	助動セル終・体	0
サセル	サセル	助動サセル終・体	0
ゼル	ゼル	ザ下一終・体	0
チル	チル	タ上一終・体	0
テル	テル	タ下一終・体	0
デル	デル	ダ下一終・体	0
ニル	ニル	ナ上一終・体	0
ネル	ネル	ナ下一終・体	0
ヒル	ヒル	ハ上一終・体	0
ビル	ビル	バ上一終・体	0
ヘル	ヘル	ハ下一終・体	0
ベル	ベル	バ下一終・体	0
ミル	ミル	マ上一終・体	0
メル	メル	マ下一終・体	0
シメル	シメル	助動シメル終・体	0
リル	リル	ラ上一終・体	0
レル	レル	ラ下一終・体	0
レル	レル	助動レル終・体	0
ラレル	ラレル	助動ラレル終・体	0
レ	ル	ラ四A・B・C仮A・B命1・2	→ル

	レル	ラ下一未1・2・4・5用1・2	＋ル
レ	レル	助動レル未1・2用2・3・4	＋ル
イレ	イル	ア上一仮	⇒ル
エレ	エル	ア下一仮	⇒ル
キレ	キル	カ上一仮	⇒ル
ギレ	ギル	ガ上一仮	⇒リ
クレ	クル	カ変仮	⇒ル
ケレ	イ	形A・B・C・D仮	→イ
	ケル	カ下一仮	⇒ル
タケレ	タイ	助動タイ仮	⇒イ
ナケレ	ナイ	助動ナイ仮	⇒イ
ラシケレ	ラシイ	助動ラシイ仮	⇒イ
ゲレ	ゲル	ガ下一仮	⇒ル
ジレ	ジル	ザ上一仮	⇒ル
スレ	スル	サ変A・B仮	⇒ル
マスレ	マス	助動マス仮	→マス
ズレ	ズル	ザ変仮	⇒ル
セレ	セル	サ下一仮	⇒ル
サセレ	サセル	助動サセル仮	⇒ル
セレ	セル	助動セル仮	⇒ル
ゼレ	ゼル	ザ下一仮	⇒ル
チレ	チル	タ上一仮	⇒ル
テレ	テル	タ下一仮	⇒ル
デレ	デル	ダ下一仮	⇒ル
ニレ	ニル	ナ上一仮	⇒ル
ネレ	ネル	ナ下一仮	⇒ル
ヒレ	ヒル	ハ上一仮	⇒ル
ビレ	ビル	バ上一仮	⇒ル
ヘレ	ヘル	ハ下一仮	⇒ル
ベレ	ベル	バ下一仮	⇒ル
ミレ	ミル	マ上一仮	⇒ル
メレ	メル	マ下一仮	⇒ル
シメレ	シメル	助動シメル仮	⇒ル
ラレ	ラレル	助動ラレル未1・2用2・3・4	＋ル
リレ	リル	ラ上一仮	⇒ル
レレ	レル	ラ下一仮	⇒ル
レレ	レル	助動レル仮	⇒ル
ラレレ	ラレル	助動ラレル仮	⇒ル
ロ	ル	ラ四A・B・C未5	→ル
イロ	イル	ア上一命1	⇒ル
エロ	エル	ア下一命1	⇒ル
カロ	イ	形A・B・C未	→イ
タカロ	タイ	助動タイ未2	⇒イ

ナカロ	ナイ	助動ナイ未2	→イ
キロ	キル	カ上一命1	→ル
ギロ	ギル	ガ上一命1	→ル
ケロ	ケル	カ下一命1	→ル
ゲロ	ゲル	ガ下一命1	→ル
シロ	スル	サ変A・B命1	→スル
ジロ	ジル	ザ上一命1	→ル
	ズル	ザ変命1	→ズル
セロ	セル	サ下一命1	→ル
サセロ	サセル	助動サセル命1	→ル
セロ	セル	助動セル命1	→ル
ゼロ	ゼル	ザ下一命1	→ル
タロ	タ	助動タ未2	→タ
ダロ	ダ	形動A・B未	→ダ
ダロ	ダ	助動ダ未2	→ダ
チロ	チル	タ上一命1	→ル
テロ	テル	タ下一命1	→ル
デロ	デル	ダ下一命1	→ル
ニロ	ニル	ナ上一命1	→ル
ネロ	ネル	ナ下一命1	→ル
ヒロ	ヒル	ハ上一命1	→ル
ビロ	ビル	バ上一命1	→ル
ヘロ	ヘル	ハ下一命1	→ル
ベロ	ベル	バ下一命1	→ル
ミロ	ミル	マ上一命1	→ル
メロ	メル	マ上一命1	→ル
シメロ	シメル	助動シメル命1	→ル
リロ	リル	ラ上一命1	→ル
レロ	レル	ラ下一命1	→ル
レロ	レル	助動レル命1	→ル
ラレロ	ラレル	助動ラレル命1	→ル
ワ	ウ	ワア四未1・2・3	→ウ
ン	ム	マ四用2	→ム
	ブ	バ四用2	→ブ
	ヌ	ナ四用2	→ヌ
ン	ヌ	助動ヌ終・体	→ヌ

(注) 助動詞「ヨウダ、ソウダ、フウダ、ミタイダ」は助動詞「ダ」の処理をもって、その処理にかえる。従って、助動詞「ダ」の連用形に「ニ」を加える。

また、助動詞「ヨウデス……ミイデス」の処理は助動詞「デス」の処理でかえる。

表の左のコラムは、活用語尾の種類を示している。第3のコラムは、その語尾をとる

品詞および活用形を指している。第2のコラムは、第3のコラムに記入された活用形の代表形語尾である。最後のコラムは第1コラムの活用語尾を Computer によって、代表語尾へ変換するための公式である。(附表A参照)

なお、コラム3における、「未1, 2」「命1」などの記述は、原則として、「岩波国語辞典」(西尾・岩淵編)による活用形の細分類に従っている。それらを表の形で示すと、附表Bのようになる。

附表一A

機能	意味	例			
		記号	活用語尾	操作	代表形語尾
0	そのまま	0	イ	+ \emptyset	イ
+	加える	+ル	イ	+「ル」	イル
→	置き換え	→ウ	イ	-「イ」+「ウ」	ウ
⇒	一部置き換え	⇒イ	ラシュウ	-「ユウ」+「イ」	ラシイ

附表一B

〔動詞〕

活用形	接続の様式
未然形 1	ぬ* (打消)
2	ない
3	せる・させる, れる・られる
4	う・よう
連用形 1	たい ←または言いさしの時
2	た(だ), て(で)

命令形の1, 2は単に語形上の差異を意味しており、

1は「〇ろ」, 2は「〇よ」の形をとるものを示す。

* 打消しの「ぬ」に続く形の意。以下同様。

〔形容詞口語〕

連用 1	}ウ音便*	言いさしや副詞的な連用修飾語になる時
2		
3		
4		
5		助動詞「た」に続く時

*岩波国語辞典では細分類されていないが、ここでは、次の基準で分類した。

3は「赤い」→「赤こう」, 「早い」→「早よう」などの形で変化するもの

2は単に語幹に「う」がつくもの

4は「大きい」→「大きゅう」、「正しい」→「正しゅう」などの形で変化するもの

〔形容詞文語〕

活用形	接続の様式
未然形 1	仮定の助詞「ば」に続く形
2	助動詞「む」などにつく形
連用形 1	口語の1と同じ
2	助動詞「き」に続く形
連体形 1	連体修飾語になる時などの形
2	助動詞「べし」に続く形

〔付記〕

ここで示した付表1, 2は, 国立国語研究所研究所補助員益子芳江・堀江久美子両嬢の協力を得て作成したものである。

また, プログラミングの段階で, 木村繁氏から多大の助言を得ている。

未筆ながら, これらの各氏に謝意をささぐ。