

## 計算機用日本語辞書の開発

桑畑和佳子 橋本三奈子 村田賢一  
情報処理振興事業協会 (IPA)

IPA技術センターでは、これまでに『計算機用日本語基本動詞辞書IPAL(Basic Verbs)』、『計算機用日本語基本形容詞辞書IPAL(Basic Adjectives)』を開発し、報告書及びフロッピーディスクで提供している。現在は『計算機用日本語基本名詞辞書IPAL(Basic Nouns)』の試作に取り組んでいる。

IPALの特徴は文法的事項を重点的に記述しているという点にある。本稿では動詞「かたまる」「かためる」、形容詞「かたい」、名詞「かたさ」「かため」を例にして、各辞書の記述内容について述べる。

### On the Lexicon of Japanese Basic Verbs, Adjectives, and Nouns.

Wakako Kuwahata Minako Hasimoto Kenichi Murata  
Information-technology Promotion Agency, Japan.

IPA Software Technology Center has developed IPAL(Basic Verbs), IPAL(Basic Adjectives). These have been supplied on documents and floppy disks. Now, We are working on IPAL(Basic Nouns).

The grammatical information contained in IPAL is detail in nature, which is one of the most salient characteristics of IPAL. In this paper, I describe the outline of IPAL using the verbs "katamaru", "katameru", the adjective "katai", the nouns "katasa", "katame" as examples.

## 1. はじめに

情報処理振興事業協会（IPA）技術センターでは、これまでに動詞と形容詞の辞書を作成し、『計算機用日本語基本動詞辞書 I P A L (Basic Verbs)』、『計算機用日本語基本形容詞辞書 I P A L (Basic Adjectives)』として、報告書形式、及びフロッピィ形式でその成果を公開してきた。現在は引き続き『計算機用日本語基本名詞辞書 I P A L (Basic Nouns)』を試作中である。

I P A L は、計算機用辞書の有り方を示すプロトタイプであると考えている。したがって、そのまま製品に活かすことができるような語数は満たせなくても、汎用的な参照用の辞書になるように配慮して開発してきた。現在までに機械翻訳、日本語ワードプロセッサ、日本語インタフェース、文書推敲支援システム、音声入出力などの様々な応用アプリケーションにおいて利用されている。さらに、日本語教育や日本語研究にも活用されている。

また、I P A L は利用するマシンを特定していない。利用者が自分のマシンで必要な箇所を好きなように編集して用いることができるように考慮してある。記述内容のコード化をできるだけ避け、理解しやすい分類名をそのまま記載し、テキスト形式で提供している。動詞辞書はフロッピィ 3 枚に、形容詞辞書はフロッピィ 1 枚に収録されている。

## 2. 4又 金録 言語

日本語の解析において比重の大きいところは、述部である。述部にくる主要品詞ということで、先に動詞、形容詞に取り組んだ。次に、述部及び項にくる主要品詞として名詞に取り組んでいる。

動詞辞書は基本和語動詞を 8 6 1 語記述した。形容詞辞書は基本形容詞を 1 3 6 語記述した。名詞辞書では、文法的に特徴のあるものを優先して約 3 1 5 0 語を記述している。

I P A L は一つの言語を辞書と文法とで相補的

に記述する立場にたっているが、特定の文法理論に依存していない。例えば学校文法では「きれい」を「形容動詞語幹」としているが、I P A L は品詞論にとらわれず、「ナ」を介して他の体言を修飾する用法を名詞の一用法として捉え、名詞辞書の枠組の中でこの用法を持つ名詞約 2 9 0 語を記述している。（一方では「形容動詞」を「ナ形容詞」と考える説も考慮し、形容詞辞書の枠組みを使った 3 0 語の試作例を既に提供している。）

尚、動詞辞書では漢語サ変動詞を扱わず、その後動詞辞書の枠組みを使った 5 0 語の試作例を提供しているだけであったが、これも名詞の一用法として捉え、サ変動詞用法のある名詞約 6 5 0 語を名詞辞書の中で記述している。

## 3. I P A L の記述

動詞、形容詞、名詞には、共通の語根を持ち、意味的にも関連している語がある。例えば次のものがそうである。

- (a1) 袖が寒さで固まっている。 《動詞》
- (b1) 子供が雪を足で固めている。 《動詞》
- (c1) この豆腐は固い。 《形容詞》
- (d1) 鉄の固さを測定する。 《名詞》
- (e1) やや固めにゆでたほうがよい。 《名詞》

下線の語の意味は、それぞれ次のようになる。

- (a2) 液体状のものが固体状に変化する。
- (b2) 液体状のものを固体状に変化させる。
- (c2) 変形したり、こわれたりしにくい。
- (d2) 変形したり、こわれたりしにくい度合。
- (e2) 変形したり、こわれたりしにくい様。

一般の国語辞典では、表記と意味と用例が書いてあるほかは、動詞の場合に自動詞、他動詞の区別と活用型が表示されているに過ぎない。計算機上での日本語の処理を人間の理解に近づけていくためには、これだけでは不十分である。人間が知識として持っていることを記述する必要がある。

また、一般の国語辞典には各品詞間の記述仕様に差はないが、動詞、形容詞、名詞には、それぞれに品詞特有の情報というものがある。よって、IPALでは品詞ごとに仕様を柔軟に変え、《統語情報》《形態情報》《意味情報》《慣用表現》について詳細な記述を行っている。

以下、実際に(a)～(e)の語を例にして、各辞書においてどのように詳細な記述がなされているかを見ていく。本稿では、主に《統語情報》を中心に述べる。動詞、形容詞の記述は既刊の辞書より抜粋している。また、名詞については現在試作中の記述から抜粋している。(尚、本稿では語源的な問題には触れない。)

#### 4. 下位区分

個々の語は意味的及び構文的な違いにより下位区分される。一つの見出し語はまず下位区分を行ってから、それぞれの下位区分ごとに記述を行った。先の(1)～(5)で挙げた例は、それぞれ以下の数だけ下位区分したうちの語義番号「01」にあたるものである。

- (a) かたまる 《動詞》 語義数：03
- (b) かためる 《動詞》 語義数：04
- (c) かたい 《形容詞》 語義数：09
- (d) かたさ 《名詞》 語義数：08
- (e) かため 《名詞》 語義数：07

各語義数が全て異なることから、品詞間には必ずしも一対一に語義の対応があるわけではないことがわかる。

- (c2) 表情がかたい。
- (d3) 表情のかたさがとれた。
- (e3) 表情がかためだ。

上の名詞「かたさ」「かため」は、形容詞「かたい(06)：自由な、のびのびとしたところがない」と関連している(カッコ内の数字は語義番号)。しかし、動詞「かたまる」「かためる」には関連する語義はない。

(a3) ×表情がかたまる。

(b3) ×表情をかためる。

また、(a)～(e)の中には他のどの語にも関連を持たない語義もある。

(b4) 彼は会社を一族で固めた。

(d4) ×一族の固さ

(e4) ×一族が固めだ。

(c5) 彼の入賞は固い。

(d5) ×入賞の固さ

(e5) ×入賞が固めだ。

(b4)は、動詞「かためる(04)：ある集団や組織を限定されたメンバーばかりにする」だけにある語義の文例であり、(c5)は、形容詞「かたい(09)：実現の可能性が高い」だけにある語義の文例である。

#### 5. 系充語彙情報

IPALは統語情報として文法的な事柄を重点的に記述している点が大きな特徴になっている。

i 文型、名詞の意味素性、名詞句の例、文例

例えば、(a)の「かたまる(01)」の記述は次のようになる。

文型	N1ガ (N2)デ	
格形式	意味素性	名詞句
ガ	CON	土、セメント、油、牛乳、のり、ゼリー
デ*	PHE/PRO	寒さ、熱/凝固剤
文例	油が 寒さで 固まっている。	

(CON:具体物, PHE:現象名詞(自然/生理), PRO:生産物・道具)

これらは動詞、形容詞について記述したほか、名詞については“述語になる用法”と“サ変動詞用法”の項目で同様の記述を行っている。

ii 述語素

動詞辞書では、格関係を表すものとして「述語素」を採用した。これは、動詞と名詞句の意味的關係を従来からある「深層格」という枠組みで捉えていたものからさらに一步すすめる試みである。

次の二例を比べてみる。

【かたまる】 語義番号：03

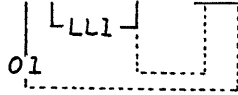
文型	N1ガ N2ニ
格形式	意味素性 名詞句
ガ	CON 荷物、店
ニ	LOC 一ヶ所、部屋の隅
文例	荷物が 一ヶ所に 固まっている。

【かためる】 語義番号：02

文型	N1ガ N2ヲ N3ニ
格形式	意味素性 名詞句
ガ	HUM 彼
ヲ	CON 荷物、店
ニ	LOC 一ヶ所、部屋の隅
文例	彼は 荷物を 一ヶ所に 固めた。

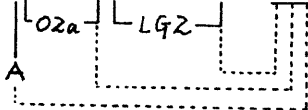
〈具体物〉がガ格にたつかヲ格にたつかという違いに、自動詞と他動詞の差が現れている。しかし同じように二格にたっている〈場所〉が、一方は〈具体物〉の存在する空間を示し、他方は〈具体物〉が移動する着点を示しているという意味の違いまでは、上の表だけではわからない。そこで述語素を用いる。

(a6) 荷物が 一ヶ所に 固まっている。



O 1 : object 対象  
L L 1 : locational locative 空間的場所

(b6) 彼は 荷物を 一ヶ所に 固めた。



A : agent 動作主  
O 2 a : object affective 対象 [受影]  
L G 2 : locational goal 空間的着点

〈LG2〉というのはヲ格の名詞句にとって二格の名詞句が着点であることを示す。例えば「先生は生徒を京都に引率した」のようにガ格の動作主にとっても〈場所〉が着点である場合には、それを示す〈LG1〉という述語素と、〈LG2〉が併記される。このように、どの名詞句にとっての着点であるかを示す点は「深層格」にはないものの一つである。

述語素には、ほかに、原因・目的の関係を表すもの、変化・作用の関係を表すもの、全体-部分の関係を表すものなどがある。

形容詞辞書でも動詞辞書でのこの考えを受け、さらに述語素を補充して記述した。名詞辞書でどのように取り入れるかは、現在検討中である。

5.1. 動詞辞書特有の統語情報

動詞の下位区分ごとに、ヴォイス、テンス・アスペクト、ムードの記述を行った。

i ヴォイス

ヴォイスとは、動詞が表す動作・作用が他に及ぶのか、他からその動作・作用をうけるのかなどを表す動詞の形態的变化をとらえる文法的カテゴリである。

①その動詞に(サ)セル形が有るか。有るならガ格の名詞句はヲ格に交替するか二格に交替するか。

(b7) 彼は 荷物を 一ヶ所に 固め送らせる。

②その動詞に(ラ)レル形が有るか。有るなら、「直接受動」「間接受動」「可能」「自発」「尊敬」のどれを表わし得るか。「直接受動」の場合は格形式の交替はどうか。

(b8) 荷物は 彼に送られて 一ヶ所に 固め送られた。  
(直接受動)  
荷物は 一ヶ所に 固め送られる。(可能)  
先生は 荷物を 一ヶ所に 固め送られた。  
(尊敬)

③ ヴォイスは次のどのタイプか。

能動：その動詞の表す行為が主語から発し、かつその行為が主語に及ばない。

例：固める、助ける、教える

相互：その動詞の表す行為が主語から発し、かつその行為が主語に及んでいる。

例：結婚する、争う

受動：その動詞の表す行為が主語から発していないが、行為が主語に及んでいる。

例：教わる、見つかる

中動：その動詞の表す行為が主語から発して、行為も主語に及んでいない。

例：固まる、休む

【かたまる】語義番号：03

ヴォイス	サセ	ヲ使役	タイプ=中動
	ラレ	間受	

【かためる】語義番号：02

ヴォイス	サセ	ニ使役	タイプ=能動
	ラレ	直受 間受、ヲ：ニヨッテ 可能、尊敬	

ii テンス・アスペクト

日本語のテンス（時制）の形は「ル形」と「タ形」である。アスペクトとは、動詞の表す動作と時間の相との関連を示す文法的カテゴリーである。

① 「ル形」（動詞の終止形）が現在を表すか、未来を表すか、通常「ル形」で用いないか。

(a9) ×荷物が 固ま~~る~~。 (用いない)

(b9) 荷物を 固め~~る~~。 (未来)

cf. 本が 固~~ま~~。 (現在)

② 「テイル形」が進行、結果残存、経験・完了、単純状態のどれを表すか。

(a10) 荷物が 固ま~~っ~~。 (単純状態)

(b10) 荷物を 固め~~ら~~。 (進行)

cf. ガラスが 割れ~~た~~。  
彼は 事件を 目撃し~~た~~。  
(経験・完了)

③ 「テイル形」の他に接続可能なアスペクト形式は有るか。

(a11) 荷物が 一ヶ所に 固ま~~っ~~。

(b11) 荷物を 一ヶ所に 固め~~ら~~。

【かたまる】語義番号：03

テア ンス スペ クト	ル	×	テイル	単純状態
	テシマウ、カケル、ハジメル、 ダス			

【かためる】語義番号：02

テア ンス スペ クト	ル	未来	テイル	進行
	テアル、テシマウ、カケル、 ハジメル、ダス			

iii ムード

ムードは、聞き手や述べる内容に対する話し手の心的態度を表す動詞の語形変化を扱う文法的カテゴリーである。

① 命令形が有るか、有るなら命令を表すか、願望を表すか。

(a12) × (荷物が) 固ま~~れ~~。 (命令形なし)

(b12) (君が) 荷物を 固め~~ろ~~。 (命令)

cf. (雨が) 降~~れ~~。 (願望)

② 意志形が有るか、有るなら意志・勧誘を表すか、推量を表すか。

(a13) ×荷物が 固ま~~ろ~~う。 (意志形なし)

(b13) 荷物を 固め~~よ~~う。 (意志・勧誘)

cf. 雨が 降~~ろ~~う。 (推量)

【かたまる】語義番号：03

ム ド	タイプ=1 (命令なし)			
	×			

【かためる】語義番号：02

ム ド	タイプ=3b (命令、意志のみ)			
	タイ・タガル、ナサイ、ナ (禁止)			

## 5.2. 形容詞辞書特有の統語情報

形容詞には、次のように述語、連体修飾、連用修飾として現れる用法をもつものがある。

- (c14) ゆで卵が固い。 (#1)  
固いゆで卵ができた。 (#2)  
 卵を固くゆでる。 (#3)

#1には終止形が、#2には連体形が、#3には連用形が現れることに着目して、形容詞の終止用法、連体用法、連用用法とそれぞれ呼ぶ。

形容詞の中には、連体用法、連用用法をもたないものもある。

- (c15) 彼の入賞は固い。  
 ×固い入賞 (連体用法なし)  
 ×固く入賞する。 (連用用法なし)

よって、下位区分ごとに各用法の有無と、有る場合にはその例を示した。

### i 連体用法

連体用法の制限用法1では、終止用法におけるガ格の名詞句が連体修飾の被修飾語にもなる名詞句を例示している。

また、同じ語義の中でも他の名詞では終止用法が可能であるのに、ある名詞では終止用法が不可能になることがある。この名詞を連体用法の制限用法2に記載した。

- ①連体修飾の際に被修飾語にくるのが「NP1」か、「NP2」か。またその名詞句の例は何か。  
 (「NP1」なら「N1」と書き、「NP2」なら「N2」と書く。)

- (c16) あのチームは 団結が 堅い。  
 (NP2) (NP1)  
 → 堅い団結  
 → N1

- ②終止用法に用いられない名詞が有るか。有るならそれは何か。

- (c17) 堅い約束。  
 ×彼の約束は堅い。 (終止用法なし)

【かたい】語義番号：07

連体用法	制限用法	1	N1	決意、団結
		2	約束	

### ii 連用用法

- ①連用用法として、連用形が「ナル」「スル」に接続可能か。

- (c18) 彼の決意はますます固くな~~な~~つた。

- ②連用用法として修飾される動詞及びその他の例は何か。

- (c19) 彼は堅く決心した。 (動詞を修飾)  
 cf. 今日はひどく暑い。 (形容詞を修飾)

【かたい】語義番号：07

連用用法	-くなる -くする2	
	動詞	決心する、約束する、誓う
	他	

## 5.3. 名詞辞書特有の統語情報

名詞辞書では、述語の項になる用法、サ変動詞用法、述語になる用法、連体修飾を受ける用法を記述する。

### i 項になる用法

名詞が文の中で「ガ・ラ・ニ」などの格助詞を伴って動詞や形容詞などの述語と共に起る用法。

- (d20) (演技の) 固さが 気になる。  
 (e20) (表現は) 硬めが よい。

ii サ変動詞用法

名詞に「スル」が後接し、動詞として働く用法が有るか。

- (d21) ×鉛筆が 固さ だ。 (サ変用法なし)  
 (e21) ×鉛筆が 固め だ。 (サ変用法なし)  
 cf. 人類が 誕生 だ。

iii 述語になる用法

文の述語になる用法が有るか。  
 有る場合は、動詞に準じて次のように書く。

【かため】 語義番号：01

文型	NP1ガ NP0ダ
格形式	意味素性 名詞句
ガ	CON このゆで卵、このガラス
文例	このゆで卵は <u>固め</u> だ。

iv 連体修飾を受ける用法

連体修飾を受ける用法が有るか。

①見出し語とその前に現れる連体修飾節(S)が全体として項あるいは述語の役割を担う用法が有るか。有るならその節は肯定か、否定か。また、「ル形」か「タ形」か。

- (d22) 坑道を掘るのに使える 硬さ  
 (肯定「ル形」)  
 (e22) ×坑道を掘るのに使える 硬め  
 (連体修飾なし)  
 cf. 猫が歩いた 跡 (肯定「タ形」)

【かたさ】 語義番号：01

S	肯定	<u>(る)</u> - た
	疑問	- るか / - たか
用例	坑道を掘るのに使える <u>硬さ</u>	

②見出し語が「という」を介して連体修飾節を取り、全体として項あるいは述語の役割を担う用法が有るか。有るならその節は肯定か、否定か。また、「ル形」か「タ形」か。

- (d23) 100倍の圧力にも耐えられるという 硬さ (肯定「ル形」)  
 cf. 映画に行かないかという 誘い (疑問「ル形」)

【かたさ】 語義番号：01

Sという	肯定	<u>(る)</u> / - た
	疑問	- るか / - たか
用例	100倍の圧力にも耐えられるという <u>硬さ</u>	

③「NPの」の形で見出し語を修飾する例は何か。

- (d24) ダイヤモンドの 硬さ

【かたさ】 語義番号：01

NPの	用例	ダイヤモンド, 石, 金属, 氷
-----	----	------------------

④その他特徴的な修飾語が有るか。

- (d25) ×このゆで卵は 固さだ。  
 (d25) このゆで卵は ちょうどよい 固さだ。

【かたさ】 語義番号：01

その他	用例	ちょうどよい, ほどよい
-----	----	--------------

⑤修飾節がついて副詞句になる用法があるか。

- cf. 調査した 結果、以下のことが判明した。

6. そのほかの情 幸段

IPALには統語情報の他にも、形態情報、意味情報、慣用表現などの記載がある。本稿では以下に記載項目を列挙するに留める。

《形態情報》

見出し語自体の形や、形の上で関連のある語に関する情報である。

〈共通〉

- ・表記
- ・異音同語（同じ意味で音韻形態に差異があるもの）
- ・合成語（複合語・派生語）

〈動詞〉

- ・活用型
- ・語幹
- ・自動詞、他動詞の区別

〈形容詞〉

- ・語幹
- ・語尾（「-い、-な、-だ」のうち接続可能なもの）

《意味情報》

語には、語彙的意味と文法的意味とがあるが、これは語彙的意味である。

〈共通〉

- ・意味記述
- ・関連語（上位語、同義語、類義語、反義語など）
- ・意味分類

〈動詞〉

- ・シソーラス（既存のシソーラスの分類名称とそのコード）

〈形容詞〉

- ・属性・感情・感覚の3分類
- ・段階性の有無
- ・評価、快・不快の意味を含むかどうか

〈名詞〉

- ・不定語（見出し語が答えになるような疑問文を作る時の不定語 ex. 誰, 何, どこ）
- ・形式名詞（「ひと, もの, ところ, とき, こと, ようす」のうち当てはまるもの）
- ・助数詞（見出し語を数えるときの数え方、また見出し語が他の名詞を数えるときの数え方）

《慣用表現》

見出し語を含む慣用句、ことわざ、よく使う言い回しなど。（各辞書とも記述している）

## 7. おわりに

限られた枚数の中では十分に辞書の記述内容を述べることはできなかったが、文法的な事柄を重点的に記述した新しいタイプの辞書というのが、I PALの特徴である。現在試作中の名詞辞書は平成6年3月に完成する予定である。

さて、今回の論文執筆にあたり、改めて動詞、

形容詞、名詞の各辞書を相互に照らしあわせながら省みたが、必ずしも記述に統一がとれているとは言い難い。動詞、形容詞、名詞辞書という順に時間をかけて開発しているので、その間に記述方法が変更されたり、記述者が変わったりしているため、記述に揺れが生じていると考えられる。今後I PALを拡張する際には、統合の試みも推し進める必要がある。

## 謝辞

共同研究者であるワーキング委員、臨時ワーキング委員の方々に深く感謝の意を表します。

## 参考文献

- (1)村木新次郎・青山文啓・六条範俊・村田賢一  
（1984）「辞書における格情報の記述」『情報処理学会自然言語処理研究会論文集』情報処理学会
- (2)情報処理振興事業協会(1987)『計算機用日本語基本動詞辞書 I PAL (Basic Verbs) 解説編』情報処理振興事業協会技術センター
- (3)情報処理振興事業協会(1990)『計算機用日本語基本形容詞辞書 I PAL (Basic Adjectives) 解説編』情報処理振興事業協会技術センター
- (4)橋本三奈子(1991)「I PALにおける統語情報——動詞・形容詞・名詞——」『I PALシンポジウム論文集』情報処理振興事業協会技術センター
- (5)橋本三奈子・桑畑和佳子・村田賢一(1991)「計算機用日本語基本名詞辞書における記述項目の構成」『情報処理学会第45回全国大会論文集』情報処理学会
- (6)桑畑和佳子・橋本三奈子・村田賢一(1991)「計算機用日本語基本名詞辞書の概要」『情報処理学会第45回全国大会論文集』情報処理学会
- (7)橋本三奈子・桑畑和佳子・村田賢一(1992)「I PAL名詞辞書の概要」『I PALシンポジウム'93論文集』情報処理振興事業協会技術センター