

2E-1

計算機用日本語形容詞辞書の試作

村田賢一 廣瀬 茂 橋本三奈子  
(情報処理振興事業協会)

1. はじめに

I P A技術センターでは、昭和56年から計算機用日本語辞書に関する研究を行ってきた。その成果の一部は計算機用日本語基本動詞辞書 I P A L (Basic Verbs) として公開されている。

現在は形容詞辞書と名詞辞書の研究に着手しているが、本論文ではこのうち形容詞について報告する。

2. 対象とする形容詞

工数に制限があるため基本的な形容詞を記述の対象とした。基本的な形容詞を選択する便法として、とりあえず「新明解国語辞典第三版」で最重要語（\*\*印）または重要語（\*印）と指定されている合計131語について実際の文書における出現頻度を調べてみた。そのために新聞約29万、理科系教科書約2.4万、計算機マニュアル約2.1万の文に使用されている形容詞を計算機で抽出したところ、表1に示すように、全出現形容詞のうち93.5%がこの131語の形容詞であった。また、文種別にみても最低92%をカバーしている。このカバー率は実用には充分ではないかもしれないが、試作する上では基本的な用語を網羅しているといえる。この観点から、上記131語を試作辞書の対象とすることにした。

表1 記述対象の形容詞のカバー率

文種	全出現回数	記述対象の形容詞	カバー率
新聞	29943	27716	92.6 %
理科系教科書	5985	5861	97.9
計算機マニュアル	503	496	98.6
計	36431	34073	93.5

3. 記述項目

記述項目として以下の24項目を設定した。このうちで最も特徴的なのは<6><7>の尺度関連情報であり、これらについては、次節で若干詳しく論ずることとする。

(1)見出し情報

- <1>エントリ 形容詞終止形の平仮名による表記
- <2>同音異義番号 同音異義語を区別するための番号
- <3>サブエントリ番号 何番目の下位区分かを示す番号

(2)意味情報

- <4>意味記述 語彙の意味
- <5>関連語 類義語、反義語、尺度共有語
- <6>尺度 気温、空間の高さ、長さなど
- <7>尺度のタイプ 段階的、非段階的の区別
- <8>評価 プラスイメージ、マイナスイメージの区別

(3)形態情報

- <9>語幹 語幹のローマ字表記
- <10>表記 漢字と送り仮名による表記
- <11>異音同語 サブエントリの異形態
- <12>活用 ナ形、ダ形の有無
- <13>派生語 助動詞・接辞が接続した形
- <14>複合語 自立語が接続した形

(4)統語情報

- <15>文型 形容詞がとる格形式のパターン
- <16>文例 その文型の例文
- <17>述語素 形容詞の統語的特徴
- <18>意味素性 名詞句の意味分類
- <19>名詞句 名詞句の例
- <20>連体用法 修飾される名詞の例
- <21>連用用法 クスル形、クナル形の有無、修飾される動詞の例

(5)その他

- <22>慣用 慣用表現の例
- <23>コロケーション コロケーションの例
- <24>備考 (1)~(4)に関して特記すべきこと

On a Experimental Lexicon of Adjectives.

Kenichi Murata, Shigeru Hirose, Minako Hashimoto.

Information-technology Promotion Agency, Japan.

#### 4. 形容詞の尺度

「Xは背が高い」のような表現は、ある尺度の上でXの取る値が大きいという意味を表すものとして、また、「XはYより背が高い」のような表現は、同一の尺度上でのXとYの値を比較して、Xの方がYより大きいという意味を表すものと考えられる。このように、「尺度」という概念を用いて、形容詞の意味の一面を明らかにすることができる。

##### 4.1 次元

ある対象物について述べる際に、そのものの持つ色々な属性・状態等の中のどれが取り上げられているのか、という点を考えてみる。例えば、ここに一つのテーブルがあるとす。このテーブルについて、「このテーブルは長い」と言えば、テーブルの持つ「長さ」という点に注目していることになるし、また、「このテーブルは重い」と言えば、「重さ」という点から、テーブルについて述べているわけである。このような、「長さ」「重さ」「大きさ」「温度」等を、ここでは「次元(dimension)」と呼ぶことにする。次元が異なれば、尺度も異なると考えられる。

##### 4.2 尺度の range

次に、「反意語」・「反義語」等と呼ばれる語の対が、単一の尺度を構成するのか、それとも、各々が異なる尺度を構成するのか、という問題がある。

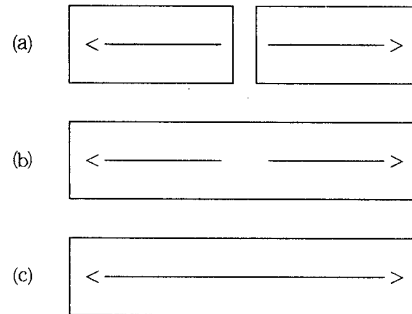
Cruse は、英語の分析において、long-shortのような反義語対は単一の尺度を構成するのに対し、stupid-intelligentのような反義語対は、それぞれが独立の尺度を構成するといっている。その根拠として、long-shortタイプでは、比較級になったときに中立的な意味で使うことができる(対象物がその次元に関して取る値の大小に係わらず、long(er), short(er) という語が使える)ということなどをあげている。これに対して、stupid-intelligentタイプでは、比較級になっても、中立的な意味にはならない。つまりlong-shortは同次元上での取り得る値(これをここでは、range と呼ぶことにする)及び方向性において対立するが、比較級になった場合には、range の違いはなくなり、対立は、方向性のみになるのに対し、stupid-intelligentでは、おそらく次元そのものが対立している関係にあるため、比較級になった場合でも、range に関して変化が起きることはない、と考える。

英語で見られたことと同じようなことが、日本語についても観察される。「高い」-「低い」の反義語対において、対象物がその次元に関して取る値の大小に係わらず、「その高さは何cmか」ということができ、「高さ」という語が中立的な意味で使われている。

図1は、(a)が「次元が異なり尺度も二つある」こと、

(b)が「次元は一つだが、range の違いから尺度が二つある」こと、そして、(c)が「次元も尺度も一つである」ことを表している。比較級の long-shortタイプの形容詞は(c)のタイプであり、stupid-intelligentタイプは(a)(又は(b))であると考えられる。

図1. 次元と尺度



##### 4.3 尺度のタイプ

尺度という点から形容詞を分類する場合、もう一つの分類方法として「段階的(gradable)」「非段階的(non-gradable)」の二つに区別することができる。「段階的」とは、連続的な値を取りうるものであり、「高い」等の形容詞は「段階的」形容詞である。段階的形容詞は、程度副詞による修飾や、二者の間の比較が可能である。

(Aはとても高い/AはBより高い。)

これに対し、「非段階的」とは、連続的な値を取らないものであり、典型的には二つの値をとる。「非段階的」形容詞は、程度副詞による修飾や、二者の間の比較ができない。「非段階的」形容詞の例としては「ない」「等しい」等があげられる。

##### 5. まとめ

形容詞の記述を終了した後、名詞辞書の試作を行う予定である。形容動詞は名詞の一用法として扱う予定であるが、この形容詞の記述枠組みで、形容動詞も充分に記述できるものと考えている。

##### 6. 謝辞

本テーマのワーキング委員の先生方、研究員の方々の御指導と御助言を頂いたことに深く感謝します。特に、尺度については臨時ワーキング委員の大名力氏の御助言を頂きました。

##### 参考文献

- ① 村田賢一他：ことばの論理を求めて、情報処理振興事業協会技術センター第6回発表論文集，pp75-99, 1987.
- ② 西尾真弥：形容詞の意味・用法の記述的研究，秀英出版，1972.
- ③ D. A. Cruse : Lexical Semantics, Cambridge University Press, 1986.
- ④ J. Rusiecki : Adjectives and Comparison in English, Longman, 1985.