3B-10

共起関係による形容詞の分析

横田　英司、　荻野　孝野、　中原　賞子、　前葉　玉緒
泉　佑二、　小宮山　由香
㈱日本電子化辞書研究所

1．はじめに

筆者らは、概念体系辞書を開発している。概念体系は品詞に依存しないが、品詞と体系が対応する傾向にあるのも事実である[文献1]。よって、形容詞と他の語との共起関係に着目して、その共起関係のパターンから、形容詞的概念の体系化をボトム・アップ的に試みている。（他の品詞的概念については、別途、試作中である。）

先ず、実際の文中での形容詞の係り先の単語を抽出し、次にこのデータをもとにして係り先を上位概念に置き換えた。最後に、計算機でソート処理して、形容詞的概念をパターン分類して体系化する。

今回は、係り先を上位概念で置き換えたところまでを報告する。

2．分析方法

2．1　分析対象データ

（1）資料名　　日本語品詞列集成[文献2]
（2）分析対象語数　　異なり語数：　762
　　延べ語数：　3416

2．2　分析手順

①用例による係り先の抽出

資料[文献2]より、分析対象語とその係り先を抽出し、以下の形式で計算機入力した。

入力形式：　／”読みかな”／”分析対象語”／”格表示”／”係り先の単語”
入力例：　／きいろい／黄色い／ガ／海

②意味情報および関係子による記述

次に、①のデータをもとに、係り先語例をまとめて意味情報で記述した。かつ、分析対象語と係り先との関係を関係子で表示した。また、意味情報での記述については、深層的観点から係るものと、表層的観点から係るものを区別して、データを２種類作成した。

なお、意味情報は、文献３、文献４を使用し、関係子は当社作成のものを使用した。

入力形式：　識別番号／関係子／意味情報，意味情報，……／格表示
　　／関係子／意味情報，意味情報，……／格表示
　　……
　　／形容詞
　　／補足説明／＊

③ソート、及びパターンの検討

②のデータを、関係子、意味情報、格表示をキーとしてソートし、パターンを検討した。表層的観点からみた形容詞の係り先の例を図１に示し、深層的観点からみた形容詞の係り先の例を図２に示す。

An Analysis of Adjective from the point of Cooccurrence
Eiji YOKOTA, Takano OGINO, Shoko NAKAHARA, Tamao MAEBA, Yuji IZUMI, Yuka KOMIYAMA
Japan Electronic Dictionary Research Institute ltd.

図１　表層的観点からみた形容詞の係り先の例

```
12005/object/1.6:場所,1.222L:照明具,1.252?:明るさを持つ物理現象による物,
1.254:視覚認識現象物,1.233:天体物/ガ
/明るい
/光が十分である/*
12006/object/1.22342:絵画・図画,1.22343:映像記録物/ガ
/明るい
/色を持つ具象物の色が明るい/*
12007/object/1.5?:未来を含んだ時間,1.4132?:希望/ガ
/明るい
/希望が持てる/*
```

図２　深層的観点からみた形容詞の係り先の例

```
12005/object/1.431221?:明るさ/ガ
/明るい
/光が十分である/*
12006/object/1.431221?:色/ガ
/明るい
/色を持つ具象物の色が明るい/*
12007/object/1.4321:経過/ガ
/明るい
/希望が持てる/*
```

３．分析の問題点

３．１　ハ格の取扱いについて。

「ハ格」は、文脈から判断して適切な格助詞に置き換えた。但し、提題の係助詞や、置き換えると同じ助詞が単文中に２つ以上になってしまうものは、「ハ格」とした。理由は、次の３つである。

（１）無理に「ハ」を置き換えて単文中に同じ格がでてくれば、実際には存在しない格パターンを記述したことになる。

（２）係り受けパターンと格パターンが厳密に対応するなら、置き換える必要はない。対応しないのなら、なおさら、置き換えるべきではない。格が表層に依存するのなら、係り受けの意味から、置き換えるべきでない。

（３）意味に依存した係り受けパターンは関係子で記述する。

３．２　意味情報置き換えの基準

係り先の意味情報を「広くとる」か、「狭くとる」かでのゆれが生じた。「広くとる」とは表層的に係り得るものすべてを係り先と認める立場で、「狭くとる」とは、深層的に係るものだけを係り先と認める立場である。

例えば、「大きい」という形容詞の係り先は、「広く」とると「具象物」であり、「狭く」とると「大きさ」という属性と考えられる。これは、次のことを仮定している。

形容詞　＝　属性値
狭い係り先　＝　属性
広い係り先　＝　属性　＋　属性の所有概念

よって、この２つの視点から意味情報に置き換え、別データとした。

４．おわりに

本稿では、形容詞と他概念との共起パターンを見出だすために行った、形容詞と係り先の抽出データとその抽出、分析方法を紹介した。今後は、係り先認定の見直しを行って、データを整理した上で、パターン化していく予定である。

参考文献

［１］井原浩子、奥村登貴子：「概念の体系化に関する考察」，情報処理学会第３６回大会
［２］田中穂積、荻野孝野、荻野綱男：「新編　日本語品詞列集成」，電子技術総合研究所
［３］荻野孝野：「日本語の意味分類体系」，計量国語学１６巻３号
［４］荻野孝野：「日本語の意味分類試案」，計量国語学会第３１回大会発表資料