

## コーパス調査による形容詞の連体形と連用形の頻度

小川典子(京都大学大学院), 李在鎬(情報通信研究機構)  
横森大輔(京都大学大学院), 土屋智行(京都大学大学院)

キーワード：形容詞、活用、構文文法、出現頻度、差異係数、相関分析

### 1. はじめに（背景と目的）

形容詞の活用は、動詞の活用とともに古くから議論されており、数多くの研究が存在する(例えば、橋本(1934)、三上(1955)、川端(1978)など)。しかし、先行研究の多くは活用の体系を記述的に捉えることを中心としており、実際の用法から網羅的に形容詞の分布を明らかにするというものではなかった。加えて、活用形による意味の違いに関してはほとんど議論されていない。

こうした現状を踏まえ、本研究では、コーパス調査の観点から活用における形容詞の活用に注目し、その分布の実態を明らかにする。調査の焦点は連体形と連用形である。これは終止形に比べて、連体形と連用形はいずれも何らかの後続要素を伴う点で複雑な構文構造をなしており、日本語学習者にとっても習得が困難な項目として挙げられることが多いためである(cf. 齋藤2002)。以上を踏まえ、本研究では実際の使用例に対する大規模な調査を行うことで、日本語教育における基礎資料を提供することを目的とする。

調査においては、新聞記事データとして「読売新聞 11 年間分」の記事を、小説データとして「新潮文庫 100 冊」を、新聞と小説の中間に位置するデータとして「新潮新書」を使用した。これらのコーパスを「茶釜」で形態素解析し、延べ 28,700 例のサンプルを収集した。さらにこれらを連体形と連用形で集計し、語彙別差異係数を求め、カイ二乗検定で統計的有意を確認した。結果として、連体形でしか使用されない形容詞(好ましい、苦い、望ましい)の存在が確認される一方、連用形でしか使用されない形容詞(いち早い、可笑しい、手早い)が存在するという二極化現象が見られた。本研究では、この結果を認知言語学における構文研究の観点から考察する。

### 2. 先行研究と問題の所在

本研究で考察対象とする形容詞連体形と連用形には、意味が大きく異なるものが見られる。例えば以下に挙げる(1)(2)の形容詞「痛い」や「恐ろしい」に関して、b の連用形の場合は程度の高さを表しており、形容詞の表す意味が a の連体形とは大きく異なっている。

- (1) a. 痛い注射  
b. 痛く感動する
- (2) a. 恐ろしい体験  
b. 恐ろしくまずい

このような相違が見られることは一部の研究で指摘されてきたが、本格的に考察されているとは言い難い。例えば、三上(1955)はこのような現象を指摘することにより、伝統文法において前提とされている形容詞の連体形と連用形を共通の語彙素からの活用形とみなすことに疑問を呈してい

る。しかし、問題提起がなされているに留まり、それ以上の考察はなされていない。また、橋本・青山(1992)は実例に基づき、形容詞の終止用法、連体用法、連用用法の三つの分布関係を調査している。ただし、(1)(2)に挙げたような連用形を文全体を修飾する副詞であると考え、考察の対象外としている。以上から、連体形と連用形の意味の違いについての調査・考察は十分ではなく、さらなる検討の余地があると思われる。

本研究では、(1)(2)に示したような相違は、連体形と連用形が形態統語的に異なることを示唆するものであると考え、定量的調査分析を通して形容詞の連体形と連用形は構文として別々の語を要請していることを明らかにする。このことから、両者はそれ自体として独立した構文であることを主張する。

### 3. データと調査方法

本研究では、実際の使用例から形容詞の活用形を網羅的に抽出し、分布実態を明らかにする。調査対象として以下の三つのテキストコーパスを使用した。

1. 新聞記事として読売新聞 11 年間分の記事 (以下、読売新聞)
2. 小説として「新潮文庫 100 冊」(以下、新潮文庫)
3. 新聞と小説の中間に位置する書籍データとして「新潮新書」(以下、新潮新書)

調査手順は以下の通りである。

1. 対象のテキスト全体を「茶筌」で形態素解析する。
2. コーパスに含まれる全形容詞の連用形と連体形を機械的に抽出する。
3. 活用形別にデータを整理・集計し、差異係数を求める。
4. カイ二乗検定で出現頻度の統計的有意を確認する。

表.1 コーパス規模

コーパス	文字数	延べ語数	異なり語数
読売新聞	7,497,353 字	4,606,346 語	52,557 語
新潮文庫	12,059,478 字	4,621,261 語	61,462 語
新潮新書	4,518,753 字	1,847,806 語	48,908 語
合計	24,075,584 字	11,075,413 語	

表 1 は、上述のコーパスを形態素解析した結果である。本研究はこれらのコーパスに含まれる全形容詞の連体形と連用形を収集した。その結果、異なり語数「548」、延べ語数「28,700」の形容詞サンプルが収集された。

## 4. 結果と考察

### 4.1. 全体の分布

まず、異なり語数で見た場合、読売新聞で 223 語、新潮文庫で 509 語、新潮新書で 293 語の形容詞が使用されていた。次に延べ語数を見た場合、表.2 の結果になった。

表.2 収集データ

	連体	連用	総計
読売新聞	4133(45.7%)	4909(54.3%)	9042
新潮文庫	5305(36.1%)	9376(63.9%)	14681
新潮新書	2128(42.8%)	2849(57.2%)	4977
総計	11566(40.3%)	17134(59.7%)	28700

以上を総合した結果、以下の分布が観察された。

1. 読売新聞の場合、規模こそ大きいものの、形容詞の使用頻度は相対的に低い。
2. 新潮新書の場合、延べ語数では読売新聞の半分程度であるが、異なり語数では読売新聞より多くの形容詞が用いられている。
3. 新潮文庫の場合、コーパスの規模は読売新聞とほぼ同サイズであるが、異なり語数と延べ語数のいずれにおいても圧倒的に多く形容詞が用いられている。
4. コーパスの相違に関係なく、連体形より連用形のほうが多く用いられている。

次にコーパス間の出現頻度に対する Pearson 相関係数を求め、相関を調べた。

表.3 相関分析の結果

	新潮新書連体	新潮新書連用	新潮新書全体	新潮文庫連体	新潮文庫連用	新潮文庫全体	読売新聞連体	読売新聞連用	読売新聞全体
新潮新書連体	1	0.339	0.832	0.684	0.305	0.499	0.86	0.235	0.605
新潮新書連用	0.339	1	0.804	0.459	0.742	0.717	0.425	0.864	0.771
新潮新書全体	0.832	0.804	1	0.703	0.63	0.738	0.794	0.657	0.837
新潮文庫連体	0.684	0.459	0.703	1	0.553	0.806	0.548	0.288	0.473
新潮文庫連用	0.305	0.742	0.63	0.553	1	0.939	0.356	0.672	0.614
新潮文庫全体	0.499	0.717	0.738	0.806	0.939	1	0.479	0.597	0.631
読売新聞連体	0.86	0.425	0.794	0.548	0.356	0.479	1	0.476	0.832
読売新聞連用	0.235	0.864	0.657	0.288	0.672	0.597	0.476	1	0.884
読売新聞全体	0.605	0.771	0.837	0.473	0.614	0.631	0.832	0.884	1

相関係数0.7以上に注目した場合、読売新聞と新潮新書は文体が類似していることもあり、相関が高く、その次に新潮新書と新潮文庫で相関が高いことが分かった。全体として1%水準で有意を示しており、あるコーパスで顕著に使用されている形容詞およびその用法は、別のコーパスでも顕著に使用されていることが明らかになった。このことから表.2の結果は調査に使用したコーパスに依存することなく、日本語の書きことばの傾向を概ね正しく反映していると考えられる。

#### 4.2. 個別の分布

調査の結果、連用形でしか使用されない形容詞の存在、連体形でしか使用されない形容詞の存在が明らかになった。以下に有意確率0.1%のものを報告する。まずは連体形である。

表.4では、ほとんど連体形でしか用いられない形容詞の一例が示されている。例えば、「数少ない」の場合、いずれのコーパスにおいても連体形での使用しか確認されない。一方、表.5に示されているように連用形でしか使用されない形容詞も存在する。

表.4 連体形での使用が顕著なもの(差異係数上位 10 位)

基本形	新潮新書		新潮文庫		読売新聞		合計		相対頻度		差異係数
	連体	連用	連体	連用	連体	連用	連体	連用	連体	連用	
数少ない	3	0	3	0	8	0	14	0	100%	0.0%	1
好ましい	2	0	5	0	25	2	32	2	94.1%	5.9%	0.882
苦い	7	0	11	2	13	0	31	2	93.9%	6.1%	0.878
望ましい	3	0	0	1	19	1	22	2	91.7%	8.3%	0.833
若い	129	8	300	46	121	5	550	59	90.3%	9.7%	0.806
新しい	465	43	265	93	701	25	1431	161	89.9%	10.1%	0.797
貧しい	13	6	40	9	21	3	74	18	80.4%	19.6%	0.608
幼い	11		26	9	8	2	45	11	80.4%	19.6%	0.607
古い	78	61	142	26	67	28	287	115	71.4%	28.6%	0.427
狭い	27	11	62	29	43	15	132	55	70.6%	29.4%	0.411

表.5 連用形での使用が顕著なもの(差異係数上位 10 位)

基本形	新潮新書		新潮文庫		読売新聞		合計		相対頻度		差異係数
	連体	連用	連体	連用	連体	連用	連体	連用	連体	連用	
いち早い	0	12	0	13	0	19	0	44	0.0%	100%	-1
遠慮ない	0	2	0	35	0	0	0	37	0.0%	100%	-1
可笑しい	0	0	0	25	0	0	0	25	0.0%	100%	-1
手早い	0	6	0	31	0	5	0	42	0.0%	100%	-1
色濃い	0	11	0	2	0	18	0	31	0.0%	100%	-1
大人しい	0	0	0	11	0	0	0	11	0.0%	100%	-1
注意深い	0	12	0	20	0	15	0	47	0.0%	100%	-1
眠い	0	5	0	14	0	0	0	19	0.0%	100%	-1
余儀ない	1	28	1	7	0	77	2	112	1.8%	98.2%	-0.964
嬉しい	0	8	1	35	0	0	1	43	2.3%	97.7%	-0.954

## 5. まとめ

以上の調査結果から、形容詞における連体形と連用形は別の形容詞を要請することがデータから示された。この結果を認知言語学における構文研究の観点から考察すると、両者は別の構文とみなすことが妥当であると言えよう。

\*謝辞: 本研究の第二著者は科学研究費補助金 若手研究(B)(課題番号: 19720111)および特定領域研究 日本語コーパス(課題番号: 19011003)の援助を受け、研究を行った。感謝申し上げます。

### 【参考文献】

- 川端善明 (1978, 1979). 『活用の研究 I・II』, 東京: 大修館書店.  
 齋藤(大関) 浩美 (2002). 「連体修飾節の習得に関する研究の動向」, 『言語文化と日本語教育』(増刊特集号), pp.45-69, お茶の水女子大学.  
 丹保健一 (1997). 「形容詞の連体、連用、終止用法の出現頻度と意味の関連性をめぐって: 「高い」「広い」「寂しい」を例として」, 『三重大学教育学部紀要 人文・社会科学』, Vol.48, pp.9-18, 三重大学.  
 橋本進吉 (1934). 『國語法要説』, 東京: 明示書院.  
 橋本三奈子・青山文啓 (1992). 「形容詞の三つの用法: 終止、連体、連用」, 『計量国語学』, Vol.18, No.5.  
 三上 章 (1955). 『現代語法新説』, 東京: 刀江書院.  
 宮島達夫 (1993). 「形容詞の語法と用法」, 『計量国語学』, Vol.19, No.2.