論　文

因子分析を用いた程度副詞と述語等の共起関係分析の試み

―― 新聞コーパスのデータから ――

服　部　　匡
学芸学部・情報メディア学科

Abstract

The cooccurrence restrictions between degree adverbs and predicates have been studied using intuition-based approaches. In this paper, the statistical method of factor analysis is applied to data from a newspaper corpus to extract properties of degree adverbs that characterize the possibility of their coocurrence with predicates. Four factors were extracted that can be interpreted to correspond to semantic features of predicates.

1　は　じ　め　に

近年、電子化された大量データを用いた日本語研究の可能性が開けてきた。現在のところ最も有効に利用しうるのは新聞のデータである。新聞は、言語資料として偏りがあるもののように思われがちであるが実際には（著作権の関係で電子化されないものを除いても）投書欄・座談会・インタビュー・エッセー・小説などを含んでおり、多様な文体特徴・内容の文章から構成されている（田野村（2000）参照）。

しかしながら、こうした大量データを用いた、発見的・探索的な言語研究の可能性は未だ十分に追求されているとは言えない。

本稿は、電子化データを用いた発見的研究の新たな可能性を探るため、程度副詞と述語の間の共起関係を、大量データ（新聞記事）に多変量解析（因子分析）の手法を適用することによって分析し、それに基づいてそれぞれの程度副詞の特徴づけを行うことを試みるものである。

先入観なく実際のデータからそれぞれの程度副詞の、述語との共起における特性を抽出するため、新聞記事の大量データから程度副詞－述語等の組み合わせを切り出し、その組み合わせに対して因子分析の方法を適用することにより、各々の程度副詞の共起特性を明らかにする。その上で従来の・内省（あるいは用例視察）に基づいた分類・体系化の提案と対照することで本稿の方法の有効性を検証し、また逆に従来の方法による分析の妥当性を検討する。

今回は、一般に程度副詞とされる語類のうち、次の15語を当初の対象とした。

きわめて、はなはだ、とても、大変、非常に、かなり、ずいぶん、相当、少し、多少、少々、だいぶ、やや、結構、なかなか

2　データについて

本研究で使用したのは、毎日新聞の1993年から2005年までの13年分のデータであり、およそ8億字（見出しと本文）からなる。この

A Factor-analysis Approach to Cooccurrence Restrictions between Degree Adverbs and Predicates in a Newspaper Corpus

データから、次のような手順によって程度副詞から文末までの部分を切り出した。

手順 1：「**程度副詞**――。」という文字列をすべて抜き出す。ただし、――は 10 字以内の文字列であり、この部分は以下で言う「**述語等**」に当たるべきものである。

つまり、ここで問題にするのは、程度副詞の後に 10 字以内の文字列が来て文が終止する場合だけである。文末から離れた位置で用いられている程度副詞は問題としない。その理由は、文中のすべての程度副詞を扱うとするとそれとの共起関係を問題にすべき述語等の範囲を確定するのが難しいからである。もちろん、このような限定が分析結果に影響を及ぼす可能性はある。この点については 5 で再び取り上げる。

ここで、述語等については、多義語の区別や異表記のまとめは行っていない（ただし「いる（要る／居る）」のような仮名表記の明らかな異義語については区別した）。程度副詞については異表記・異形をまとめている。例えば「大分・だいぶ・だいぶん」をまとめて一項目とみなしている。

手順 2：「程度副詞と、意味的にそれがかかわる述語等」という関係を含んでいない文字列を排除する。

例えば「相当」の場合を例にとれば、“相当する”、“相当）を盗んだ疑い”、“相当なもの”などといった文字列が排除される。[1] “相当に”“ずいぶんと”“多少は”なども排除する。

「かかわる」ということについて補足する。例えば「とてもいいことだと思う」では「とても」は前部の「いい」にだけかかわると分析可能であるが、「いいことだと思う」全体を述語等として取る。逆に、「ずいぶん様子が違う」では「ずいぶん」は後部の「違う」のみにかかわるとも分析可能であるが、「様子が違う」全体を当該述語等として取る。「かかわる」範囲の判定が困難な例もあり恣意的な判定を避けるため分析せず必ず文末までの範囲を取ることにした。

手順 3：10 回以上出現（横の合計が 10 以上）の述語等のみを残す。

この結果表 1 のような共起度数表を得る（合計しての共起度数の多い述語等から順に 20 位

表 1　共起度数

	きわめて	非常に	大変	とても	相当	ずいぶん	はなはだ	かなり	大分	けっこう	なかなか	多少	やや	少々	少し	合計
難しい	203	192	54	68	11	0	1	61	0	26	332	0	8	2	2	960
異例	845	0	0	0	0	0	0	2	0	0	0	0	0	0	0	847
高い	239	125	8	20	31	0	1	156	1	13	1	0	20	3	8	626
大きい	243	134	17	20	11	0	1	39	0	4	0	0	16	0	8	493
違う	0	5	1	0	22	77	0	110	73	1	0	2	16	7	127	441
少ない	266	95	7	10	1	0	4	30	1	2	1	0	17	0	0	434
うれしい	1	52	103	244	0	0	0	1	0	1	0	0	0	0	0	402
多い	39	104	17	35	2	9	1	51	0	113	0	0	22	0	2	395
珍しい	298	41	7	4	1	0	0	1	0	0	0	0	0	0	0	352
低い	195	25	5	2	5	1	0	58	0	1	0	0	30	0	2	324
残念	43	141	81	31	0	0	2	1	0	0	0	0	2	1	9	311
重要だ	146	115	28	11	0	0	0	2	0	0	0	0	0	0	0	302
ある	0	8	0	3	29	13	0	134	0	83	0	0	1	0	8	284
異例だ	250	0	0	0	0	0	0	0	0	0	0	5	0	0	0	250
厳しい	89	63	19	11	11	2	0	52	0	3	24	0	1	1	0	276
残念だ	41	103	82	16	0	0	2	1	0	0	0	0	1	5	5	256
強い	95	77	3	20	12	0	0	29	0	8	0	0	2	0	0	246
いい	7	45	4	74	1	0	0	12	7	4	47	0	0	0	0	201
遺憾	142	15	35	0	0	0	7	0	0	0	0	0	0	0	0	199
遺憾だ	107	11	49	0	0	0	7	0	0	0	0	0	0	0	0	174

まで並べ、以下を略したものである)。

手順4：共起度数表で、1以上の値をすべて1に変換する。

この結果、表2のような行列を得る（同様に20位まで示し後を略す)。つまり「共起例があるかないか」の二値データになる。共起の量的側面（頻度）を捨象するのは乱暴なようであるが、用例というのは「あるかないか」が最も重要と思われる。もちろん、量的な側面に注目することが必要な問題もあるに違いないが、量を考慮することでかえって現象の本質が見えにくくなることもありうる。発見的・探索的な研究においては、どの方法が適切は結果として分かるという面もあるのであり、多様な方法が試みられることが望ましい。

他に、一定の閾値（例えば5回）を決めて、それを超えて共起している組み合わせだけを1にするといった方法も考えられる。

表2において、述語等は555個、副詞は15個あり、掛けると8325の欄が存在する。そのうち1（共起例あり）の欄は2526個、0（共起例なし）の欄は5799個である。なお述語等のうち丁寧体のものは65である。

参考までに示すが、これらの欄について、対象とする新聞記事データの範囲を「91年のみ」から始めて1年分づつ増やしていくと、1の欄の数は表3のように推移する。比率の伸び方が徐々に鈍っていくことが分かる（年によって記事の量が異なるため厳密な統計にはならない)。もっとも、さらにデータを増やしていけば一定の比率に収束するのかどうかは何とも言えない。

これは、このような共起関係の調査にはどの程度の量のデータが必要かを考える上で興味深い結果である。もちろん、上記は15年間で10度以上程度副詞と共起する述語等に限定した場合のことであって、より低頻度の述語もすべて含めて調査する場合は、当然、その種類自体がデータの範囲の拡大とともに増えていく。

表2　共起の有無

	きわめて	非常に	大変	とても	相当	ずいぶん	はなはだ	かなり	大分	けっこう	なかなか	多少	やや	少々	少し
難しい	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1
異例	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
高い	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1
大きい	1	1	1	1	1	0	1	1	0	1	0	0	1	0	1
違う	0	1	1	0	1	1	0	1	1	1	0	1	1	1	1
少ない	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0
うれしい	1	1	1	1	0	0	0	1	0	1	0	0	0	0	0
多い	1	1	1	1	1	1	1	1	0	1	0	0	1	0	1
いる	0	0	0	0	1	1	0	1	1	1	0	0	0	0	0
珍しい	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0
低い	1	1	1	1	1	1	0	1	0	1	0	0	1	0	1
残念	1	1	1	1	0	0	1	1	0	0	0	0	1	1	1
重要だ	1	1	1	1	0	0	0	1	0	0	0	0	0	0	0
ある	0	1	0	1	1	1	0	1	0	1	0	1	1	0	1
異例だ	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
厳しい	1	1	1	1	1	1	0	1	0	1	1	0	1	1	0
残念だ	1	1	1	1	0	0	1	1	0	0	0	0	1	1	1
強い	1	1	1	1	1	0	0	1	0	1	0	0	1	0	0
いい	1	1	1	1	1	0	0	1	1	1	1	0	0	0	0
遺憾	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0

表3　1（共起あり）の欄の比率の推移

95年のみ	96年まで	97年まで	98年まで	99年まで	00年まで	01年まで	02年まで	03年まで	04年まで	05年まで
18.05%	20.41%	22.41%	23.96%	25.48%	26.68%	27.66%	28.49%	29.15%	29.83%	30.34%

なお、最初に述べたように新聞は様々なスタイル・内容の文章を含んでいるので本稿のように日本語としてあり得る表現の網羅的調査を行うには適していると言えるが、他の種類のデータとの比較も当然行わなければならない。[2] 他の種類の、計算機処理可能な現代日本語のデータの蓄積を待ち、取り組んで行きたいと考えている。

3　因　子　分　析

3. 1　分析の概要

Gorsuch（1983）によれば、因子分析の通常の目的は、「概念化の補助手段として、諸変数間の相互関係を簡潔ながらも正確な仕方で要約すること」である。

それぞれの程度副詞（があるものと共起するかしないか）を変数と見、述語等をケースと見て、変数の背後にある共通因子を抽出する。ただし、予備的な分析において低い共通性を示した「はなはだ」と「少々」の2語を当初の15語から除き、13語を分析対象とする。データは表2のものである。

主因子法・プロマックス回転（斜交回転）によって4因子を抽出した（初期の固有値はそれぞれ3.07、1.87、1.34、1.19である）。因子間相関行列は表4のようになり、第1・第2因子の間に負の相関（－0.41）があり、第3・第4因子の間に正の相関（0.32）がある。因子パターンは表5のようになった。

「ずいぶん・大分・少し・かなり・相当・多少・やや」との共起は第1因子に規定されるところが大きいことになる。一方、「非常に・とても・大変」の3語との共起は第2因子に強く規定されている。また、「結構・なかなか」との共起は第3因子に規定され、「きわめて・非常に」との共起は第4因子に規定される。以上それぞれのグループの副詞には意味的に共通点があるのは明らかだが、その解釈はしばらく保留し、因子得点を参照して検討することとする。

なお、朝日新聞の88－96年の記事データに対して同様の手順による分析を行った場合にもほぼ同様の結果を得ている（2006年9月24日、「日本語コーパス」日本語学班研究会において報告）。よって、今回の結果は新聞における当該表現の一般的な使用傾向を反映したものと言える。

表4　因子間相関行列

	1	2	3	4
1	1.00	**－0.41**	0.22	0.16
2	**－0.41**	1.00	0.12	－0.06
3	0.22	0.12	1.00	**0.32**
4	0.16	－0.06	**0.32**	1.00

表5　因子パターン行列

	因　子			
	1	2	3	4
ずいぶん	**0.757**	0.120	－0.139	－0.238
大分	**0.620**	－0.005	－0.117	－0.196
少し	**0.528**	－0.004	－0.013	－0.053
かなり	**0.497**	－0.187	0.229	0.270
相当	**0.487**	0.045	0.252	0.031
多少	**0.441**	0.035	－0.077	－0.140
やや	**0.355**	－0.169	－0.078	0.210
非常に	0.138	**0.723**	－0.064	***0.391***
大変	－0.037	**0.612**	－0.030	0.016
とても	－0.022	**0.611**	0.236	－0.121
けっこう	0.011	0.030	**0.687**	－0.251
なかなか	－0.162	0.077	**0.561**	0.039
きわめて	－0.230	0.075	－0.142	**0.647**

3. 2　因子の解釈

それぞれの因子に対する各ケース（述語等）の因子得点を表にして示し、因子の解釈を行っていく。因子得点とは、抽出された各因子の持つ傾向を、各々のケースがどの程度強く有しているかを示したものであり、平均0、分散1の標準化得点として示される。

まず、第1因子の因子得点の高い述語等・低い述語等をそれぞれ30位まで示すと表6・表7のようになる。

高い因子得点を示す述語等は、基準との差異、基準からの変化、物事の存在（零との差異とみなせないことはない）を表すものが多い。

表６　第１因子の因子得点（高得点順）

	ずいぶん	大分	少し	かなり	相当	多少	やや	非常に	大変	とても	けっこう	なかなか	きわめて	因子得点
変わる	1	1	1	1	1	1	1	1	0	0	0	0	0	2.71
違う	1	1	1	1	1	1	1	1	1	0	1	0	0	2.69
減った	1	1	1	1	1	1	1	0	0	0	0	0	0	2.63
異なっている	1	1	1	1	1	1	1	0	0	0	0	0	0	2.63
異なる	1	1	1	1	1	1	1	0	1	0	0	0	0	2.55
あった	1	1	1	1	1	1	0	1	0	0	1	0	0	2.50
増えている	1	1	0	1	1	1	1	1	0	0	0	0	0	2.40
変わった	1	1	1	1	1	1	0	0	0	0	0	0	0	2.35
違ってくる	1	1	1	1	1	1	0	0	0	0	0	0	0	2.35
変わってきた	1	1	1	1	1	0	1	1	0	0	0	0	0	2.25
違っていた	1	1	1	1	0	1	1	0	0	0	0	0	0	2.22
違った	1	1	1	1	1	0	1	0	0	0	0	0	0	2.17
増えた	1	1	1	1	1	0	1	1	0	1	0	0	0	2.16
ある	1	0	1	1	1	1	1	1	0	1	1	0	0	2.12
狭い	1	1	1	1	1	0	1	1	1	0	0	0	1	2.08
違いがある	1	1	1	1	0	1	0	0	0	0	1	0	0	2.00
楽になった	1	1	1	1	1	0	1	0	1	1	0	0	0	2.00
安い	1	1	1	1	1	0	1	1	1	1	0	0	1	1.99
増えてきた	1	1	1	1	1	0	0	1	0	0	0	0	0	1.97
長い	1	0	1	1	1	1	1	1	1	1	1	0	1	1.95
違うようだ	1	1	1	1	0	1	0	0	0	0	0	0	0	1.94
違います	1	1	1	1	0	1	0	1	0	1	0	0	0	1.94
進んでいる	1	1	0	1	1	0	1	1	0	1	1	0	0	1.91
回復した	1	1	0	1	0	1	1	0	0	0	0	0	0	1.91
悩んだ	1	1	1	1	1	0	0	1	0	1	0	0	0	1.88
進んだ	1	1	0	1	1	0	1	0	0	0	0	0	0	1.85
差がある	1	1	0	1	1	0	1	0	0	0	0	0	0	1.85
残っている	1	0	1	1	1	1	0	0	0	0	1	0	0	1.84
改善された	1	1	1	1	0	0	1	1	0	0	0	0	0	1.84

表７　第１因子の因子得点（低得点順）

	ずいぶん	大分	少し	かなり	相当	多少	やや	非常に	大変	とても	けっこう	なかなか	きわめて	因子得点
読みやすい	0	0	0	0	0	0	0	0	1	1	0	0	1	− 1.04
印象的だった	0	0	0	0	0	0	0	0	1	1	0	0	1	− 1.04
印象的でした	0	0	0	0	0	0	0	0	1	1	0	0	1	− 1.04
興味深い	0	0	0	0	0	0	0	1	1	1	0	1	1	− 1.04
便利	0	0	0	0	0	0	0	1	1	1	0	1	1	− 1.04
難しいですね	0	0	0	0	0	0	0	1	1	1	0	1	1	− 1.04
元気	0	0	0	0	0	0	0	1	1	1	0	1	1	− 1.04
魅力的	0	0	0	0	0	0	0	1	1	1	0	1	1	− 1.04
美しい	0	0	0	0	0	0	0	0	0	1	1	1	1	− 0.97
おもしろかった	0	0	0	0	0	0	0	0	1	1	1	1	0	− 0.96
勉強になる	0	0	0	0	0	0	0	0	1	1	1	1	0	− 0.96
重要	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
重要です	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
大事だ	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
残念に思う	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
大事です	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
残念だ」と話している	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
大事	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
分かりやすい	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
残念だ」と話した	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
よく似ている	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
重要なことだ	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
興味がある	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
重要な問題だ	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
貴重だ	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
わかりやすい	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
残念だ」と述べた	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
重要だった	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96
残念なことだ	0	0	0	0	0	0	0	1	1	1	0	0	1	− 0.96

例えば「違う」、「違いがある」のように差異・相違を言うもの、「ある」のように存在を言うもの、「増えてくる」のように増減を言うもの、「回復した」「改善した」「楽になった」など元の状態からの進展的変化の結果を言うものなどが含まれる。

一方、低い因子得点を示す述語等は、全体に、差異や存在としては解釈しにくい。

例えば「困難（だ）」「便利（だ）」「元気（だ）」といった表現は、文脈的によってはある状態との異なりを表すものと解しうることもあるが（例えば「元気だ」が、病気だった時の状態に比べてのことを言う場合のように）、これらの表現固有の意味の中に差異という要素が含まれているとは言いにくい。

次に、第２因子の因子得点の高い述語等・低い述語等をそれぞれ30位まで示すと表８・表９のようになる。

因子得点の高いものは人の感情や評価に関わるものが多い。例えば、「楽しかった」「うれしかった」など感情の形容詞、「面白い」「おいしい」など、人による評価を含むものが含まれる。これとは対照的に、因子得点の低いものは、「低くなった」「減った」「異なっている」など、それ自体の意味の中には感情や評価の要素を含まない、いわば、客観的な事柄を意味するものが多いことが分かる。第１因子の高得点群と第２因子の低得点群、および逆の組み合わせにはある程度共通のものがあり、因子間の逆相関の事実に対応する。

表８　第２因子の因子得点（高得点順）

	ずいぶん	大分	少し	かなり	相当	多少	やや	**非常に**	**大変**	**とても**	けっこう	なかなか	きわめて	**因子得点**
参考になった	1	0	0	0	0	0	0	**1**	**1**	**1**	0	1	0	**1.27**
楽しかった	0	0	0	0	0	0	0	**1**	**1**	**1**	1	1	0	**1.26**
面白かった	0	0	0	0	0	0	0	**1**	**1**	**1**	1	1	0	**1.26**
楽しみです	0	0	0	0	0	0	0	**1**	**1**	**1**	1	1	0	**1.26**
おいしい	0	0	0	0	0	0	0	**1**	**1**	**1**	1	1	0	**1.26**
難しいのです	0	0	0	0	0	0	0	**1**	**1**	**1**	1	1	0	**1.26**
難しいものです	0	0	0	0	0	0	0	**1**	**1**	**1**	1	1	0	**1.26**
難しい問題です	0	0	0	0	0	0	0	**1**	**1**	**1**	1	1	0	**1.26**
良かった	1	0	0	0	0	0	0	**1**	**1**	**1**	0	1	1	**1.25**
難しかった	0	0	1	0	0	0	0	**1**	**1**	**1**	1	1	1	**1.18**
厳しかった	0	0	1	0	0	0	0	**1**	**1**	**1**	1	1	1	**1.18**
素晴らしい	0	0	0	0	0	0	0	**1**	**1**	**1**	0	1	0	**1.18**
魅力的だ	0	0	0	0	0	0	0	**1**	**1**	**1**	0	1	0	**1.18**
興味深かった	0	0	0	0	0	0	0	**1**	**1**	**1**	0	1	0	**1.18**
興味深い	0	0	0	0	0	0	0	**1**	**1**	**1**	0	1	1	**1.16**
便利	0	0	0	0	0	0	0	**1**	**1**	**1**	0	1	1	**1.16**
難しいですね	0	0	0	0	0	0	0	**1**	**1**	**1**	0	1	1	**1.16**
元気	0	0	0	0	0	0	0	**1**	**1**	**1**	0	1	1	**1.16**
魅力的	0	0	0	0	0	0	0	**1**	**1**	**1**	0	1	1	**1.16**
勉強になった	1	0	0	0	0	0	0	**1**	**1**	**1**	0	0	0	**1.15**
勉強になりました	1	0	0	0	0	0	0	**1**	**1**	**1**	0	0	0	**1.15**
参考になりました	1	0	0	0	0	0	0	**1**	**1**	**1**	0	0	0	**1.15**
うれしかった	0	0	0	0	0	0	0	**1**	**1**	**1**	1	0	0	**1.14**
危険です	0	0	0	0	0	0	0	**1**	**1**	**1**	1	0	0	**1.14**
喜んでいる	0	0	0	0	0	0	0	**1**	**1**	**1**	1	0	0	**1.14**
喜ばれた	0	0	0	0	0	0	0	**1**	**1**	**1**	1	0	0	**1.14**
重要である	0	0	0	0	0	0	0	**1**	**1**	**1**	1	0	1	**1.13**
満足している	0	0	0	0	0	0	0	**1**	**1**	**1**	1	0	1	**1.13**
多いのです	0	0	0	0	0	0	0	**1**	**1**	**1**	1	0	1	**1.13**

表9　第2因子の因子得点（低得点順）

	ずいぶん	大分	少し	かなり	相当	多少	やや	**非常に**	**大変**	**とても**	けっこう	なかなか	きわめて	**因子得点**
低くなった	0	1	1	1	0	0	1	**0**	**0**	**0**	0	0	1	**－1.68**
持ち直した	0	1	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.67**
盛り返した	0	1	0	1	1	1	1	**0**	**0**	**0**	0	0	0	**－1.63**
劣る	0	0	1	1	0	1	1	**0**	**0**	**0**	0	0	0	**－1.63**
高くなっている	0	0	1	1	1	0	1	**0**	**0**	**0**	0	0	1	**－1.62**
深刻	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	1	**－1.62**
上回っている	0	0	1	1	1	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
大きめ	0	0	1	1	1	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
上回った	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
高め	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
下回った	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
上回る	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
出遅れた	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
落ちる	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
興奮気味	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
後退した	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
下回る	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
事情が異なる	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
縮まった	0	0	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.61**
減った	1	1	1	1	1	1	1	**0**	**0**	**0**	0	0	0	**－1.60**
異なっている	1	1	1	1	1	1	1	**0**	**0**	**0**	0	0	0	**－1.60**
違っていた	1	1	1	1	0	1	1	**0**	**0**	**0**	0	0	0	**－1.60**
違った	1	1	1	1	1	0	1	**0**	**0**	**0**	0	0	0	**－1.58**
前	1	1	1	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.58**
低くなる	0	0	0	1	0	0	1	**0**	**0**	**0**	0	0	1	**－1.56**
遅れそうだ	0	0	0	1	1	0	1	**0**	**0**	**0**	0	0	0	**－1.55**
優勢	0	0	0	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.54**
改善した	0	0	0	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.54**
苦戦	0	0	0	1	0	0	1	**0**	**0**	**0**	0	0	0	**－1.54**

次に、第3因子に対する因子得点の高い述語等・低い述語等をそれぞれ30位まで示すと表10・表11のようになる。

概ね、得点の高いものは、対象や事態が容易でない感じ、一筋縄で行かぬ感じ、軽視できない感じなど、いわば「手ごたえ」の大きさを表すものであり、得点の低いものはそれに遠いもののように思われる。

表10　第3因子の因子得点（高得点順）

	ずいぶん	大分	少し	かなり	相当	多少	やや	非常に	大変	とても	**けっこう**	**なかなか**	きわめて	**因子得点**
深刻だ	0	0	1	1	1	0	0	1	1	1	**1**	**1**	1	**2.29**
強かった	0	0	1	1	1	0	0	1	1	1	**1**	**1**	1	**2.29**
難しい	0	0	1	1	1	0	1	1	1	1	**1**	**1**	1	**2.29**
重い	0	0	1	1	1	0	1	1	1	1	**1**	**1**	1	**2.29**
つらい	0	0	1	1	1	0	1	1	1	1	**1**	**1**	0	**2.29**
怖い	0	0	1	1	1	0	1	1	1	1	**1**	**1**	0	**2.29**
大変です	0	0	0	1	1	0	0	1	0	1	**1**	**1**	0	**2.28**
迫力がある	0	0	0	1	1	0	0	1	0	1	**1**	**1**	0	**2.28**
面白い	0	0	0	1	1	0	0	1	1	1	**1**	**1**	1	**2.27**
忙しい	0	0	0	1	1	0	0	1	1	1	**1**	**1**	0	**2.27**
勇気がいる	0	0	1	1	1	0	0	0	1	1	**1**	**1**	0	**2.22**
高い	0	1	1	1	1	0	1	1	1	1	**1**	**1**	1	**2.12**
いい	0	1	0	1	1	0	0	1	1	1	**1**	**1**	1	**2.10**
少ない	0	1	0	1	1	0	1	1	1	1	**1**	**1**	1	**2.10**

深い	1	0	0	1	1	0	1	1	0	1	**1**	**1**	1	**2.03**
厳しい	1	0	0	1	1	0	1	1	1	1	**1**	**1**	1	**2.03**
うるさい	0	0	1	1	1	0	0	1	0	0	**1**	**1**	0	**2.02**
厳しいものがある	0	0	0	1	1	0	0	1	1	0	**1**	**1**	1	**1.99**
むずかしい	0	0	1	1	0	0	1	1	1	1	**1**	**1**	1	**1.90**
うまい	0	0	1	1	0	0	0	1	1	1	**1**	**1**	0	**1.90**
説得力がある	0	0	0	1	0	0	0	1	0	1	**1**	**1**	1	**1.89**
大変だ	0	0	0	1	0	0	0	1	0	1	**1**	**1**	0	**1.88**
役に立つ	0	0	0	1	0	0	0	1	1	1	**1**	**1**	1	**1.88**
広い	0	0	0	1	0	0	1	1	1	1	**1**	**1**	1	**1.88**
おもしろい	0	0	0	1	0	0	0	1	1	1	**1**	**1**	0	**1.88**
しんどい	0	0	0	1	0	0	0	0	0	1	**1**	**1**	0	**1.81**
大変	0	0	1	0	1	0	0	1	0	1	**1**	**1**	0	**1.68**
しっかりしている	0	0	0	0	1	0	0	1	0	1	**1**	**1**	0	**1.66**
きつい	0	0	1	1	0	0	0	1	1	0	**1**	**1**	0	**1.61**

表 11　第 3 因子の因子得点（低得点順）

	ずいぶん	大分	少し	かなり	相当	多少	やや	非常に	大変	とても	**けっこう**	**なかなか**	きわめて	**因子得点**
強くなった	1	1	0	0	0	0	0	0	1	0	**0**	**0**	1	**− 1.32**
楽になりました	1	1	1	0	0	0	0	0	0	0	**0**	**0**	0	**− 1.30**
お世話になった	1	0	0	0	0	0	0	0	1	0	**0**	**0**	0	**− 1.15**
迷惑をかけた	1	0	0	0	0	0	0	0	1	0	**0**	**0**	0	**− 1.15**
前だった	1	0	1	0	0	0	1	0	0	0	**0**	**0**	0	**− 1.12**
、違う	0	1	0	0	0	0	0	0	0	0	**0**	**0**	0	**− 1.07**
大きくなった	1	0	1	0	0	0	1	1	0	0	**0**	**0**	1	**− 1.05**
和らいだ	0	0	1	0	0	1	1	0	0	0	**0**	**0**	0	**− 0.99**
光栄に存じます	0	0	0	0	0	0	0	0	1	0	**0**	**0**	0	**− 0.90**
ご迷惑をかけた	0	0	0	0	0	0	0	0	1	0	**0**	**0**	0	**− 0.90**
趣が違う	1	1	1	0	1	0	1	0	0	0	**0**	**0**	0	**− 0.90**
遺憾である	0	0	0	0	0	0	0	0	1	0	**0**	**0**	1	**− 0.90**
重要であります	0	0	0	0	0	0	0	0	1	0	**0**	**0**	1	**− 0.90**
遺憾です	0	0	0	0	0	0	0	0	1	0	**0**	**0**	1	**− 0.90**
不適切	0	0	0	0	0	0	0	0	1	0	**0**	**0**	1	**− 0.90**
困難な状況にある	0	0	0	0	0	0	0	0	1	0	**0**	**0**	1	**− 0.90**
有利か	0	0	0	0	0	0	1	0	0	0	**0**	**0**	0	**− 0.89**
優位か	0	0	0	0	0	0	1	0	0	0	**0**	**0**	0	**− 0.89**
欠ける	0	0	0	0	0	0	1	0	0	0	**0**	**0**	0	**− 0.89**
優位	0	0	0	0	0	0	1	0	0	0	**0**	**0**	0	**− 0.89**
低め	0	0	0	0	0	0	1	0	0	0	**0**	**0**	0	**− 0.89**
不良」	0	0	0	0	0	0	1	0	0	0	**0**	**0**	0	**− 0.89**
小ぶり	0	0	0	0	0	0	1	0	0	0	**0**	**0**	0	**− 0.89**
低調	0	0	0	0	0	0	1	0	0	0	**0**	**0**	1	**− 0.89**
消極的だ	0	0	0	0	0	0	1	0	0	0	**0**	**0**	1	**− 0.89**
異例だ	0	0	0	0	0	0	0	0	0	0	**0**	**0**	1	**− 0.89**
異例のことだ	0	0	0	0	0	0	0	0	0	0	**0**	**0**	1	**− 0.89**
不十分だ	0	0	0	0	0	0	0	0	0	0	**0**	**0**	1	**− 0.89**
困難な情勢だ	0	0	0	0	0	0	0	0	0	0	**0**	**0**	1	**− 0.89**

最後に、第 4 因子に対する因子得点の高い述語等・低い述語等をそれぞれ 25 位まで示すと表 12 ・表 13 のようになる。因子得点の高いものは文章語的文脈でよく出現しそうに思われるもの、しかも評価的にマイナスのものが多い。一方因子得点の低いものは口頭語的文脈で生じ易く評価的にプラスのものが多いように思われる。因子得点の下位 50 形式中で丁寧体のものは 16 形式あるのに対して、上位 50 形式中で丁寧体のものは 1 形式しかない。

表 12　第 4 因子の因子得点（高得点順）

	ずいぶん	大分	少し	かなり	相当	多少	やや	***非常に***	大変	とても	けっこう	なかなか	**きわめて**	**因子得点**
険しい	0	0	0	1	1	0	1	***1***	0	0	0	1	**1**	**1.81**
厳しくなる	0	0	0	1	1	0	1	***1***	0	0	0	0	**1**	**1.60**
難しくなった	0	0	1	1	0	0	1	***1***	1	0	0	1	**1**	**1.57**
大きくなる	0	0	1	1	1	0	1	***1***	1	0	0	0	**1**	**1.56**
難しそうだ	0	0	1	1	1	0	0	***1***	0	0	0	1	**1**	**1.53**
苦しい	0	0	1	1	1	0	1	***1***	1	1	0	1	**1**	**1.53**
薄い	0	0	1	1	0	0	1	***1***	0	0	0	0	**1**	**1.42**
慎重だ	0	0	1	1	0	0	1	***1***	0	0	0	0	**1**	**1.42**
遅かった	0	0	1	1	0	0	1	***1***	0	0	0	0	**1**	**1.42**
高いという	0	0	0	1	0	0	1	***1***	0	0	0	0	**1**	**1.40**
困難になった	0	0	0	1	0	0	1	***1***	0	0	0	0	**1**	**1.40**
乏しい	0	0	0	1	0	0	1	***1***	0	0	0	0	**1**	**1.40**
高いことが分かった	0	0	0	1	0	0	1	***1***	0	0	0	0	**1**	**1.40**
小さい	0	0	1	1	1	0	1	***1***	0	1	0	0	**1**	**1.39**
遅い	0	0	1	1	1	0	1	***1***	0	1	0	0	**1**	**1.39**
心もとない	0	0	0	1	0	0	1	***1***	0	1	0	1	**1**	**1.38**
ユニークだ	0	0	0	1	0	0	1	***1***	0	1	0	1	**1**	**1.38**
難しい	0	0	1	1	1	0	1	***1***	1	1	1	1	**1**	**1.36**
重い	0	0	1	1	1	0	1	***1***	1	1	1	1	**1**	**1.36**
疑問だ	0	0	1	1	0	0	1	***1***	1	0	0	0	**1**	**1.36**
不満だ	0	0	1	1	0	0	1	***1***	1	0	0	0	**1**	**1.36**
高くなる	0	0	0	1	1	1	1	***1***	1	0	0	0	**1**	**1.33**
問題がある	0	0	0	1	1	1	1	***1***	1	0	0	0	**1**	**1.33**
厳しくなった	0	0	1	1	1	0	0	***1***	0	0	0	0	**1**	**1.32**

表 13　第 4 因子の因子得点（低得点順）

	ずいぶん	大分	少し	かなり	相当	多少	やや	***非常に***	大変	とても	けっこう	なかなか	**きわめて**	**因子得点**
楽になりました	1	1	1	0	0	0	0	***0***	0	0	0	0	**0**	**− 1.79**
お世話になった	1	0	0	0	0	0	0	***0***	1	0	0	0	**0**	**− 1.57**
迷惑をかけた	1	0	0	0	0	0	0	***0***	1	0	0	0	**0**	**− 1.57**
おいしかった	0	0	0	0	0	0	0	***0***	1	1	1	0	**0**	**− 1.53**
気に入っている	0	0	0	0	0	0	0	***0***	1	1	1	0	**0**	**− 1.53**
好きだ	0	0	0	0	0	0	0	***0***	1	1	1	0	**0**	**− 1.53**
好きです	0	0	0	0	0	0	0	***0***	0	1	1	0	**0**	**− 1.47**
楽しいです	0	0	0	0	0	0	0	***0***	0	1	1	0	**0**	**− 1.47**
好き	0	0	0	0	0	0	0	***0***	0	1	1	0	**0**	**− 1.47**
、違う	0	1	0	0	0	0	0	***0***	0	0	0	0	**0**	**− 1.37**
お世話になりました	0	0	0	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.36**
うれしく思う	0	0	0	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.36**
驚いています	0	0	0	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.36**
感激しました	0	0	0	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.36**
きれい	0	0	0	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.36**
楽しみにしています	0	0	0	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.36**
光栄に思います	0	0	0	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.36**
元気です	0	0	0	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.36**
違いがある	1	1	1	1	0	1	0	***0***	0	0	1	0	**0**	**− 1.36**
温かい	0	0	1	0	0	0	0	***0***	1	1	0	0	**0**	**− 1.34**
おもしろかった	0	0	0	0	0	0	0	***0***	1	1	1	1	**0**	**− 1.32**
勉強になる	0	0	0	0	0	0	0	***0***	1	1	1	1	**0**	**− 1.32**
趣が違う	1	1	1	0	1	0	1	***0***	0	0	0	0	**0**	**− 1.30**
美しかった	0	0	0	0	0	0	0	***0***	0	1	0	0	**0**	**− 1.30**

4　内省等に基づく分析との関連

次に、当分析の結果を従来の、主に内省や用例視察に基づく分析と比較する。まず、渡辺実氏（1990）「程度副詞の体系」の分析を取りあげる。渡辺氏は、程度副詞には体系があるとされ、それぞれの副詞が「比較構文」「計量構文」に立つか立たないかに基づく分類を示された。比較構文とは「X は Y より__A だ」の構文であり、計量構文とは「X は__A だ」の構文である。渡辺氏の分類の概要を示すと次のようになる。

この副詞分類は当分析の結果とよく一致していることが分かる。つまり、「とても類」は第２因子に、「結構類」は第３因子に、「多少類」は第１因子に対応する。

唯一の重要な例外は「ずいぶん」であり、当分析から見ると、この語はむしろ「多少」の仲間のように思われる。実は「ずいぶん」については他にも渡辺氏の分類と一致しない事実がある（方言差や話者の年代による差異が関与する可能性もある）。

なお、渡辺氏は別の観点から程度副詞（を含む様々な表現）を下記のように「わがこと系」と「ひとごと系」に分類することも提案されている（1991）。「わがこと系」とは、「私は__うれしい」のように話し手の心中を表す場合などに用いられるものである。

ひとごと系　―　ずいぶん、ずっと、よほど、いっそう、はるかに、いたって、大分、相当、かなり、なかなか、わりに、けっこう、ごく

わがこと系　―　大変、非常に、極めて、はなはだ、とても、ばかに、やけに、ちょっと、少し、いささか

こちらは、当分析の結果には直接の対応を見出しにくいようである（第２因子に関して観察したように「非常に、とても、大変」に感情・感覚等の表現との共起傾向があるのは事実であり、その多くは一人称での使用と思われるが）。

森山卓郎氏（1985）は程度副詞を「量的程度副詞」と「純粋程度副詞」とに二分して共起述語類の相違を説明された。「ゆれもあろうが、いちおうの簡単なリスト」として示されたところでは前者は「かなり、随分、結構、やたら、なかなか、比較的、相当、大分、わりあい、少し、ちょっと、多少、少々、ある程度、いささか／もっと」などであり、後者は「非常に、大変、はなはだ、著しく、極めて、ごく、すこぶる、あまりに／ずっと、より、最も、一番」などである。「もっと、ずっと」のような他者との比較の語など本稿で扱っていないものを別にすれば、氏の言われる「量的程度副詞」は当分析の第１または第３因子に強く関連し、「純粋程度副詞」は第２または第４因子に強く関連するとおよそ言える。

表 14　渡辺実氏の分類

		比較構文	計量構文	判断構造	評価	表現性	語　例
発見系	とても類	×	○	発　見	±	驚　嘆	とても・はなはだ・すこぶる・大変・極めて・非常に・**ずいぶん**
発見系	結構類	×	○	望外発見	＋	脱懸念	けっこう・なかなか・わりに・ばかに・やけに
比較系	多少類		○	潜在比較	―	反期待	多少・すこし・ちょっと・やや・いささか・かなり
比較系	多少類	○		潜在比較	±	反期待	同上

服部（1996）に示したように、各々の副詞の生起度数に対する比較構文での生起度数の比は次の通りである（朝日新聞の 1988－1994 年のデータ）。つまり、「ずいぶん」は「○○よりずいぶん大きい」のような形の文[3]に生起する。

とても	0‰
非常に	0.71‰
随分／ずいぶん	6.84‰
相当	17.2‰
多少	17.4‰
だいぶ	22.8‰

なお森山氏の類別の基準は「～が〔程度副詞〕ある」のような存在量叙述の構文への出現可能性（量的程度副詞のみがこれに立つ）である。参考までに示せば、氏の純粋程度副詞に「賃金の差は非常にある」、「共感するところが非常にある」、「～という気持ちはとてもあります」などの例がある（これらを存在量と見るかどうかは別として）。

本分析の第1因子に強く関連する副詞は工藤浩氏（1983）が量副詞としての用法を持つとされるものにおよそ対応する。

5　ま　と　め

述語等との共起の有無のデータを基に、程度副詞の共起傾向に関わる因子を抽出し、因子得点と述語の意味的（文体的）特性を参照しながら意義付けを試み、従来説の妥当性を検証した。

もちろん今回の方法は、程度副詞の性質の一側面（文末にある述語等との関係）を問題とするものにすぎない。例えば、ある副詞が主文中によりもむしろ特定の形式の節（例：「～ても」）の中に生じやすいといった大局的な要因に関連した傾向は無視している（例えば「少し」と「多少」を比較する上ではそのような観点も必要であろう）。

このような単純な方法で、従来の内省（実例視察）による分析の一部とかなり一致する結果が得られたことは興味深いことであり、今後このような方法が未知の現象の解明につながる発見手順として利用される可能性を示すものである。

今後は、従来あまり研究されていない語類に同様の方法を適用することにより、共起傾向に関する事実を発見することを試みたい。仮にそれが原理的には内省等による方法で到達しうる結果であったとしても、現実の研究者の有する時間は有限である以上、こうした方法を発見手順として用いることは無意義とは言えないと思われる。[4)]

付　記

本研究は、文部科学省科学研究費補助金（特定領域研究「日本語コーパス」）による研究成果の一部である。

新聞記事のデータは著作権者である新聞社の許諾を得て研究用に使用している。

なお因子分析の適用に関して、柳井晴夫・聖路加看護大学大学院教授よりご助言を頂いた。お礼申し上げる。

注

1）「とても」、「なかなか」については、“とてもできない”、“なかなか来ない”等のように否定形式の述語と共起しているものを排除した。

厳密に言えば、例えば「なかなか来られない」のような用法（非・程度副詞的）と「なかなか来るのが難しい」のような用法（程度副詞的）とは連続的なように思われる（服部（1994）参照）がここでは敢えて否定形式かどうかで線を引いた。

もっとも「なかなか悪くない」「とても好ましくない」のような例はむしろ程度副詞的用法とみるべきであるが、問題となる実例は今回はなかった。

2）例えば、「はなはだ（甚だ）」に関して国会会議録での程度副詞との共起例を調査すると、当調査では共起例のない組み合わせのうちの少なからぬものが出現する（前川喜久雄氏のご教示により確認した）。

3）渡辺氏の言われる比較構文と範囲が全く同じではない可能性がある。なお前稿の調査範囲外（毎日新聞2005年）に、「とても」に関して「震災以前よりとても生活が苦しい」、「通常よりとても割高」のような例が見られる。

参 考 文 献

工藤浩（1983）程度副詞をめぐって　渡辺実る編『副用語の研究』明治書院

田野村忠温（2000）用例に基づく日本語研究 ―コーパス言語学―『日本語学』19－5、192－201.

服部匡（1994）副詞「なかなか」の意味用法の分析『言語学研究』13号、79－90

服部匡（1996）程度副詞と比較基準 ―「多少」、「少し」を中心に―『同志社女子大学学術研究

年報』47巻Ⅳ号、269－284
服部匡（2006）多変量解析の手法を用いた、程度副詞－述語の共起関係の分析　特定領域研究「日本語コーパス」日本語学班研究会発表資料
森山卓郎（1985）程度副詞と動詞句『京都教育大学国文学会誌』20、60－65
渡辺実（1987）程度副詞の体系『上智大学国文学科紀要』4、1－16
渡辺実（1991）「わがこと・ひとごと」の観点と文法論」『国語学』165、1－14
柳井晴夫他（1990）『因子分析、その理論と方法』朝倉書店
Gorsuch, Richard L.（1983）*Factor Analysis*, second edition, Lawrence Erlbaum Associates.