

(これは校正前の原稿です。) 難波えみ・玉岡賀津雄 (2016). 様態と結果の副詞的表現と動詞の共起パターンに関するエントロピーと冗長度を指標にした検討. 計量国語学, 30(4), 195-209.

様態と結果の副詞的表現と動詞の共起パターンに関する エントロピーと冗長度を指標にした検討

難波 えみ (名古屋大学大学院生)

玉岡 賀津雄 (名古屋大学大学院)

要旨

難波・玉岡 (2014) は『毎日新聞』のコーパスを用いて、動詞句内に生起する様態と結果の副詞的表現について、基本生起位置の特定を試みた。その結果、様態の副詞的表現 ($N=23$) は(S)AdvOV と(S)OAdvV の2つを、結果の副詞的表現 ($N=17$) は(S)OAdvV のみを基本生起語順と認めた。生起位置の違いは共起する動詞と副詞的表現との選択制限の緩急が関係していると思われる。そこで、本研究では、難波・玉岡 (2014) で対象となった副詞的表現がどのような動詞と共起しているか、動詞の延べ頻度と異なり頻度を数え、エントロピーと冗長度を使って共起のパターンを指標化した。個々の副詞的表現について、両指標を使ってクラスタ分析を行った。その結果、3つのクラスタに分けることができた。様態の副詞的表現は、多様な動詞と共起するクラスタに分類され、限られた動詞と共起するクラスタは結果の副詞的表現が多く含まれた。個々の副詞的表現の共起パターンからみても、様態の副詞的表現は選択制限が緩く、結果の副詞的表現は厳しいことが示された。

キーワード : 様態の副詞的表現, 結果の副詞的表現, 共起パターン, 選択制限, エントロピー, 冗長度, 基本語順

1. 副詞的表現の生起位置について

副詞的表現¹には文中で生起可能な位置が複数あり、いずれの位置に生起しても非文となることはない。例えば、副詞的表現の「急いで」であれば、「(▲)社長は(▲)本社に(▲)応援を(▲)依頼した」の文で、▲のどの位置でも生起可能である(三原, 2008)。このように、どの▲に生起しても非文にはならず、主格名詞句の前後、与格名詞句の後、対格名詞句の後(つまり動詞の直前)に生起できることがわかる。しかしながら、文中に複数生起可能な位置を持つとはいえ、文中で最も基本となる生起位置がないというわけではない。

統語上における付加詞の分類を行った研究がある。Koizumi (1993) は統語範疇の違い

¹本研究では、様態と結果の副詞を研究対象としている。ただし、小泉・玉岡 (2006) および仁田 (2002) で、純粋な副詞だけでなく、形容(動)詞連用形も含めて「副詞類」「副詞的表現」と呼んでいる。本研究でも検索語には「熱心に」や「かたく」などの形容(動)詞の連用形も含んでいる。本研究では、副詞も形容(動)詞の連用形も一括して「副詞的表現」と表現する。

により付加詞を3分類した。MP付加詞、IP付加詞、VP付加詞である。第1に、MP付加詞は、モーダル句 (Modal Phrase, MP) 内に生起する付加詞で、「～から (because)」 「～が (but)」 などがある。文を否定した時に、否定の作用域に入らない付加詞である。第2に、IP副詞は、屈折辞句 (Inflection Phrase, IP) 内に生起する付加詞で、「～かぎり (as long as)」 や「と (if)」 などがある。IP副詞を含む文に「～わけではない」をつけることで、否定の作用域に入る付加詞である。第3は、VP副詞で、動詞句 (Verb Phrase, VP) 内に生起する付加詞で、「～ながら (while)」 や「～まえに (before)」 などがある。動詞を否定形にしたときに、否定の作用域に入る付加詞である。以上を並べると [MP MP付加詞 [IP IP付加詞 [VP VP付加詞+動詞]]]と表すことができ、階層構造を成すことが分かる。たとえば、[MP 雨が降っているので [IP 止まない限り [VP ラジオを聴きながら 掃除するだろう]]という文は正しいが、「雨が降っているので」「止まない限り」「ラジオを聴きながら」の3つの付加詞の順番を入れ替えると「(雨が) 止まない限りラジオを聴きながら雨が降っているので掃除をするだろう」と言ったように、妙な文になる。このことより、付加詞にも階層性があること、つまり、統語範疇 MP, IP, VP 内に生起する特定の付加詞があることが分かる。

表1 小泉・玉岡 (2006) の文処理実験による正順語順の判定

副詞類の種類	正順語順	例文
陳述の副詞 (MP副詞)	<u>Adv</u> SOV	たぶん次郎がボールを蹴るだろう。
時の副詞 (IP副詞)	<u>Adv</u> SOV	昔次郎がオペラを歌った。
	S <u>Adv</u> OV	次郎が昔オペラを歌った。
様態の副詞 (VP副詞)	S <u>Adv</u> OV	検事がこっそり玄関を開けた。
	SO <u>Adv</u> V	検事が玄関をこっそり開けた。
結果の副詞 (VP副詞)	S <u>Adv</u> OV	順子が細かく氷を砕いた。
	SO <u>Adv</u> V	順子が氷を細かく砕いた。

注: 下線の *Adv* が副詞, S は主語の句, O は目的語の句, V は動詞を示す。

副詞的表現も付加詞と同様に、統語範疇に対応する。小泉・玉岡 (2006) は日本語の副詞類 (陳述の副詞, 時の副詞, 様態の副詞, 結果の副詞) を MP 内に生起する MP 副詞 (陳述の副詞), IP 内に生起する IP 副詞 (時の副詞), VP 内に生起する VP 副詞 (様態の副詞, 結果の副詞) に分け、副詞類を含む文の処理に要する負担を実証した。心理言語学研究においては、処理負担が小さいことは、文理解がより容易であることを示す。すなわち、処理の負担は反応時間により検証可能である。そこで、できるだけ迅速かつ正確に文処理を行う条件で、これらの副詞類の生起位置を変えた文を被験者 (N=24) に視覚的に提示し、文の正誤判断に要する反応時間を測定した。文中における処理負担の小さい生起位置は、表1に示したとおりである。陳述の副詞は *AdvSOV* (*Adv* は副詞類, S は主語, O は目的語, V は動詞を示す) が、時の副詞は *AdvSOV* と *SAdvOV* が、様態の副詞は *SAdvOV* と *SOAdvV* の位置が、それぞれ有意に反応時間が速く、それぞれの副詞類の基本生起位置で

あると結論付けた。なお、時の副詞、様態の副詞および結果の副詞はそれぞれ2つの基本生起語順となった。様態の副詞と結果の副詞は同じ動詞句内に生起することが示された。

表2 難波・玉岡（2014）による副詞の出現位置と頻度および一様性の検定の結果

副詞の種類	(S)AdvOV		(S)OAdvV		一様性の検定	
	頻度(件)	割合	頻度(件)	割合	χ^2 値	p
様態 (N=23)	3288	48.4%	3398	50.0%	1.81	<i>ns</i>
結果 (N=17)	202	18.0%	908	80.7%	449.04	$p < .001$

注: 表のSは主語, Advは副詞, Oは目的語, Vは動詞, *ns*は有意でないことを示す。Nは検索した副詞の数を示す。

小泉・玉岡（2006）の実験結果からは、様態の副詞類も結果の副詞類もSAdvOVおよびSOAdvVの2か所が基本であるという結果であった。この結果より、さらに、難波・玉岡（2014）は様態の副詞的表現と結果の副詞的表現について、大規模新聞コーパスを用いて、より厳密な生起位置の検討した。小泉・玉岡（2006）で用いられた様態の副詞類、結果の副詞類を参考に、様態の副詞的表現23語と結果の副詞的表現17語を『毎日新聞』9年間分（1991～1999）のデータベースより検索し、目的語をともなう他動詞と共起した場合の動詞句内での生起位置を明らかにした。その結果、表2に示したように、様態の副詞的表現は、(S)AdvOVが48.4%、(S)OAdvVが50.0%で、両方の語順がほぼ同じ頻度となった。一方、結果の副詞的表現は、(S)OAdvVの語順が80.7%で、(S)AdvOVの語順は18.0%で、動詞句内の動詞の前の位置に頻繁に出現し、(S)OAdvVの語順を作ることが示された。このコーパス研究の結果にしたがうと、様態の副詞的表現は、動詞との結びつきが多様であると推測される。すなわち、動詞との選択制限が緩やかであり、目的語Oの前にも後にも柔軟に生起しやすいのではないだろうか。一方、結果の副詞的表現は(S)OAdvVの語順での生起が多数を占めた。これは、結果の副詞的表現が動詞と厳しい選択制限を持つため、動詞との結びつきが強く、動詞の前に生起する傾向があるのではなからうか。そこで、本研究では、様態と結果の副詞的表現についてどのような動詞とどのくらいの頻度で共起するか調査することにした。

2. 副詞的表現の生起位置と動詞の共起関係

2.1 目的

難波・玉岡（2014）では副詞的表現の生起位置には動詞との結びつきが関連していることが示唆された。しかし、個々の副詞的表現と動詞群との共起関係については詳細の記述はされていない。個々の副詞的表現と動詞の共起パターンを観察すること、様態の副詞的表現と結果の副詞的表現と共起する動詞の多様性について詳細に検証することが必要である。そこで、本研究では、副詞的表現と動詞の共起関係、すなわち、副詞的表現と共起する動詞の選択制限について計量的に実証する。

2.2 共起パターンの指標化

大規模コーパスを使って多様な共起表現を総括的に考察するにあたり、絶対頻度に依存しない指標が必要である。エントロピーと冗長度は、比較的絶対頻度に依存しない指標であり、個々の副詞的表現の共起関係を客観的に直接に比較することができる（この尺度の詳細は、有本, 1982; 堀, 1979; 海保, 1989; 木山・玉岡, 2011; Miyaoka and Tamaoka, 2005; 玉岡, 2011, 2016; 玉岡・木山・宮岡, 2011; Tamaoka, Lim and Sakai, 2004 を参照）。そこで、エントロピーと冗長度の両指標で、様態と結果の副詞的表現と動詞の「共起パターン」を検討する。

エントロピーは、Claude E. Shannon が『通信の数学的理論』（1948）で発表した概念であり、以下の式で求められる。

$$H = - \sum_{j=1}^j p_j \log_2 p_j$$

エントロピーは、情報の曖昧さを表す指標である。情報の乱雑さとも言え、数値が大きいほど、共起できる動詞が多様であることを示す。本研究でいえば、エントロピーは、ある副詞的表現が共起する動詞の多様性の指標となり、エントロピーの値が大きいほど、副詞的表現が多様な動詞と共起することを示す。

冗長度は、表現の種類と使用頻度に基づいて無駄の程度を表す指標である。情報の偏重性の指標とも解釈できる。冗長度 (R) は、エントロピーとエントロピー最大値（最も不規則な状態）を用い以下の式で求められる。

$$R = \left(1 - \frac{H}{H_{\max}} \right) \times 100$$

本研究では、冗長度は、ある副詞が動詞群と共起するパターンの偏重性を示す指標と考える。冗長度が高いほど、共起パターンが不規則的であり、特定の動詞との結びつきが強く、頻繁に共起していることを示す。逆に、冗長度が小さいと、共起関係において特定の動詞への偏重があまりないことになる。

エントロピーと冗長度の両指標を組み合わせることで、表現の多様性と偏重性の両側面からの考察ができる（木山・玉岡, 2011; Miyaoka and Tamaoka, 2005; 中俣, 2015; 玉岡, 2011, 2016; 玉岡・木山・宮岡, 2011; Tamaoka, Lim and Sakai, 2004）。また、これら 2 つの指標を 2 次元上にプロットリングすることで、共起パターンを 4 つに区分できる。具体的には、図 1 に示したように、エントロピーの高低、および冗長度の高低を組み合わせることで、副詞的表現を 4 つの特性に区分（以下、特性区分）できると予想する。

第 1 の特性区分に、高エントロピー・低冗長度の副詞である。ここに属する副詞的表現は多様な動詞と規則的に共起する副詞的表現である。この特性区分の副詞的表現は、共起

する動詞は多いものの、特定の動詞と共起する傾向はなく、多様な動詞と満遍なく共起する傾向がある。第2の特性区分は、高エントロピー・高冗長度の副詞的表現である。この種の副詞的表現は、多様な動詞と共起することができるものの、特定の動詞に偏重した共起関係がある。つまり、多くの動詞と共起し、ある特定の動詞との共起頻度が多くみられる副詞的表現である。第3の特性区分は、低エントロピー・低冗長度の副詞である。ここに属する副詞的表現は共起する動詞が少なく、それらの動詞と規則的に共起する特性を持つ。第4の特性区分は、低エントロピー・高冗長度の副詞である。ここに属する副詞は共起する動詞が少なく、さらに、特定の動詞との共起する頻度が高い副詞である。以上のように、動詞との共起関係について、エントロピーと冗長度を組み合わせることで4つの特性区分を想定することができる。

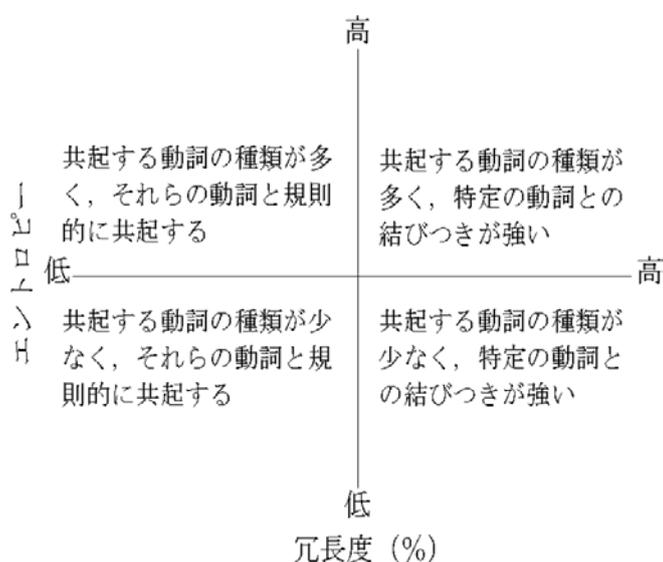


図1 副詞的表現と共起する動詞群についてのエントロピーと冗長度による4つの特性区分

本研究では、『毎日新聞』9年間分の大規模コーパスを使って、様態と結果の副詞的表現と共起する動詞の種類と頻度を計算する。そして、副詞の種類による動詞群との共起パターンを個々の副詞的表現について検討する。

2.3 仮説

本研究では、副詞的表現がどういう動詞を取りうるかという選択制限の面から、特定の副詞と動詞群との共起パターンについて仮説を立てた。

様態の副詞的表現は、動作の展開過程に内属する諸側面のあり様について言及する（仁田 2002）。ここでいう諸側面とは、たとえば、勢いや強さ、時間的な側面、動きの質などである。そうであれば、ある動詞が持ちうる側面がほかの動詞の持ちうる側面と重なれば重なるほど、その側面に言及する様態の副詞的表現は、より多くの動詞と共起する可能性がある。それならば、選択制限は緩やかになると推測される。動詞と多様な共起が可能であるため、特定の動詞との結びつきは希薄になるように思われる。

一方、結果の副詞的表現は動詞が表す変化の結果を意味する(影山, 2009). 影山(2013)によると、結果の副詞的表現は結果述語とも呼ばれ、結果述語を含む文は結果構文として成立する。結果構文では、ある動作が対象物に及んだ結果の状態が結果の副詞(二次述語)で言及される。すなわち、結果の副詞的表現と動詞は意味的に密接な関係がなければならない。さらに、結果の副詞的表現は、動詞に含意される状態をより具体的に描写するものであることより、動詞との組み合わせは、「粉々に壊す/壊れる」「ペシャンコにつぶす/つぶれる」などといったように、かなり慣例化している。結果の副詞的表現は動詞と意味的な関係が強いこと、動詞との組み合わせがかなり慣例化していることより、結果の副詞的表現より選択制限が厳しいと予想される。つまり、結果の副詞的表現については、特定の動詞群との共起に限定されると推測される。

以上の仁田(2002)および影山(2009, 2013)の議論から以下の仮説を立てた。まず、様態の副詞的表現は多様な動詞と共起し動詞との結びつきは結果の副詞的表現よりも希薄である。一方、結果の副詞的表現は限られた動詞と共起し、様態の副詞的表現と比べて特定の動詞との結びつきが強い。個々の副詞的表現の動詞群との共起パターンを、エントロピーと冗長さの指標で示すことで以上の仮説を検証する。

3. 副詞と動詞の共起表現の検索と指標化

3.1 エントロピーと冗長さの計算

難波・玉岡(2014)から、様態の副詞的表現 23 語、結果の副詞的表現 17 語を調査対象とした。1991 年から 1999 年までの『毎日新聞』9 年間分(総語数 273,541,662 形態素数)を検索エンジン『茶漉』(Purdue University, Atsushi Fukada 作成)を用いた。検索語とした副詞的表現には形容詞由来のものだけでなく、オノマトペが多く含まれており、動詞との共起例を検索するためにも、規模の大きいコーパスを使う必要がある。『毎日新聞』のコーパスを用いると時代がやや古くなるものの、本研究の検索対象の語は 20 年前でも遜色なく使われていた副詞的表現だと思われる。また、新聞コーパスであれば、広く一般大衆に向けて書かれた文章であるため、小説のように著者の個性が強く影響することもあまりないと考えられる。さらに、新聞という特定のジャンルのコーパスを使用すると、解釈が容易であるという利点もある。そのため、ジャンルを特定した大規模の新聞コーパスを使用した。なお、国立国語研究所からも『現代日本語書き言葉均衡コーパス』(BCCWJ: Balanced Corpus of Contemporary Written Japanese)が公開されており、総語数は 1 億 430 万語である(前川, 2007; Maekawa, Yamazaki, Ogiso, Maruyama, Ogura, Kashino, Koiso, Yamaguchi, Tanaka and Den, 2014)。しかし、『毎日新聞』の約半分くらいの語数であるため、本研究では使用しなかった。

様態の副詞的表現 ($N=23$): ゆっくり、ちびちび、こっそり、そっと、もりもり、さっさと、テキパキ、ペラペラ、せっせと、ころころ、すばやく、ボキッと、きっぱり、こわごわ、ぼんやり、じっと、のんびり、さらりと、どンドン、難なく、うまく、のろのろ、熱心に

結果の副詞的表現 ($N=17$): こなごなに、かちかちに、ペシャンコに、細かく、細く、

星形に、ばらばらに、人肌に、柔らかく、かたく、パリパリに、びしょびしょに、どろどろに、カリカリに、熱く、まるく、ピカピカに

上記の 40 語の副詞的表現が動詞と共に出現する 31,270 文を抽出し、個々の副詞的表現と共に出現する動詞の異なり頻度と延べ頻度を数えた。そして、個々の副詞的表現について、エントロピーと冗長度を算出した。本研究におけるエントロピーと冗長度はすでに挙げた公式に以下のように当てはめて算出した。

$$H = - \sum_{j=1}^J \frac{\text{ある動詞の頻度}}{\text{延べ頻度}} \times \log_2 \frac{\text{ある動詞の頻度}}{\text{延べ頻度}}$$

$$R = \left(1 - \frac{\text{エントロピー}}{\log_2 \text{動詞の異なり頻度}} \right) \times 100$$

3.2 結果

「ゆっくりする」「こなごなになる」のように、共起する動詞が「する」「なる」のものは「副詞的表現+する/なる」で 1 つの動詞であるとみなし、除外した。その結果、副詞的表現の副詞である「かちかちに」「びしょにしょに」は「する」「なる」としか共起していなかったため削除した。したがって、これらの 2 語は分析の対象には含めないで、全副詞的表現数は 38 語となった。また、これ以降は、様態の副詞的表現 23 語と結果の副詞的表現 15 語との比較となる。また、個々の副詞的表現と動詞の共起頻度は、特定の動詞との共起が多く、次第に少ない頻度で複数の動詞と共に出現する傾向を取る。つまり、大きな値の方向に向かって曲線が長くなだらかに裾野を伸ばしていく「べき分布」になる。そこで、共起する動詞の種類を示す共起動詞数 (type frequency, 異なり頻度, 種類頻度) と共起の総計である共起頻度 (token frequency, 延べ頻度) については、自然対数 (natural logarithm) に変換して分析した。なお、自然対数の底は、アルファベットの e で表され、ネイピア数とも言われ、 $e=2.718\cdots$ の超越数である。表 3 に、個々の副詞的表現と動詞の共起頻度、エントロピー、冗長度を示した。なお、エントロピーだけでなく、標本数に依存しない共起パターンの指標として Simpson の D (Simpson, 1949) が有効であるという報告がある (影浦, 2000; 鈴木・影浦, 2011; Yoshikane, Kageura and Tsuji, 2003)。Simpson の D は、以下の式 (影浦, 2000; 鈴木・影浦, 2011) で得られる。

$$D = \sum_{m=1}^{V(N)} V(m, N) \frac{m}{N} \frac{m-1}{N-1}$$

表 3 個々の副詞的表現の動詞との共起パターンのエントロピーと冗長度による指標化

種類	副詞	共起動詞数	共起頻度	エントロピー	冗長度(%)
様態 (N=23)	どンドン	1,044	3,639	8.64	13.86
	ゆっくり	942	3,696	8.30	15.94
	そっと	362	854	7.79	8.35
	こっそり	269	488	7.62	5.56
	さっさと	184	359	6.71	10.87
	せっせと	160	308	6.65	9.24
	熱心に	301	1,349	6.64	19.33
	すばやく	105	143	6.53	2.67
	難なく	124	226	6.39	8.16
	うまく	1,062	8,867	6.05	39.77
	のんびり	165	536	6.00	18.51
	じっと	234	1,704	5.65	28.20
	さらりと	87	218	5.40	16.13
	ぼんやり	78	280	5.16	17.98
	こわごわ	40	55	5.09	4.44
	きっぱり	96	471	4.68	28.96
	テキパキ	29	44	4.63	4.63
	のろのろ	16	25	3.64	8.92
	ころころ	17	66	2.72	33.43
	ペラペラ	9	31	2.66	15.96
	もりもり	7	18	2.14	23.75
	ちびちび	5	10	1.96	15.55
	ボキッと	3	6	1.46	7.94
	平均	232	1,017	5.33	15.57
結果 (N=15)	細かく	48	388	3.95	29.30
	ばらばらに	17	28	3.77	7.74
	柔らかく	10	11	3.28	1.33
	熱く	10	16	3.13	5.93
	細く	18	61	2.68	35.72
	人肌に	5	5	2.32	0.00
	バリバリに	5	5	2.32	0.00
	こなごなに	6	12	2.28	11.64
	ピカピカに	10	28	2.28	31.35
	かたく	5	9	2.06	11.33
	星形に	4	4	2.00	0.00
	ペシャンコに	4	11	1.49	25.44
	まるく	2	2	1.00	0.00
	カリカリに	2	8	0.95	4.56
	どろどろに	2	4	0.81	18.87
	平均	10	39	2.29	12.21

この式の N は延べ頻度であり, m は出現語数の異なり頻度である. D は相対的集中度の指

標であり、繰り返し率指標とも呼ばれる。本研究で言えば、 D の値が高い場合は、副詞的表現が特定の動詞と頻繁に共起することを意味し、 D の値が低い場合は、副詞的表現が多様な動詞と散漫に共起することを意味すると考えられる。表3の数値を利用して、Simpsonの D を計算した。さらに、1から D の値を引いて、エントロピーと同様に、値が高い方が多様であることを示すようにして（正の相関として）、エントロピーとSimpsonの D の相関係数を算出した。その結果、両者の相関係数は非常に高く（ $r=.865, p<.001$ ）、有意であった。これら2つの指標は共起する動詞の多様性を表す指標としてほぼ同じような値を示すと考えられる。エントロピーとSimpsonの D のいずれの指標を採用しても良いと思われるが、本研究では、エントロピーを採用した。エントロピーとSimpsonの D の体系的な比較については今後の課題としたい。

また、エントロピーと冗長度が同じ指標ではないかという議論もある。そこで、全副詞的表現（ $N=38$ ）について、両指標のピアソンの積率相関係数（以後、単に相関係数と呼ぶ）を計算した。その結果、 $N=38, r=.032 (ns)$ と低く、有意ではなく、両指標の関係はまったく無いことが分かった。したがって、少なくとも本研究の副詞的表現と動詞の共起パターンについては、エントロピーと冗長度は異なる側面の特性を示す指標であると考えられる。また、述べ共起頻度を自然対数に変換した値について、エントロピーとの相関係数を計算したところ、 $N=38, r=.892 (p<.001)$ と非常に高かった。全副詞のエントロピーと述べ共起頻度のプロットは図2に示したとおりである。また、冗長度との相関は、 $N=38, r=.436 (p<.01)$ で、有意ではあるが、中程度の高さであった。

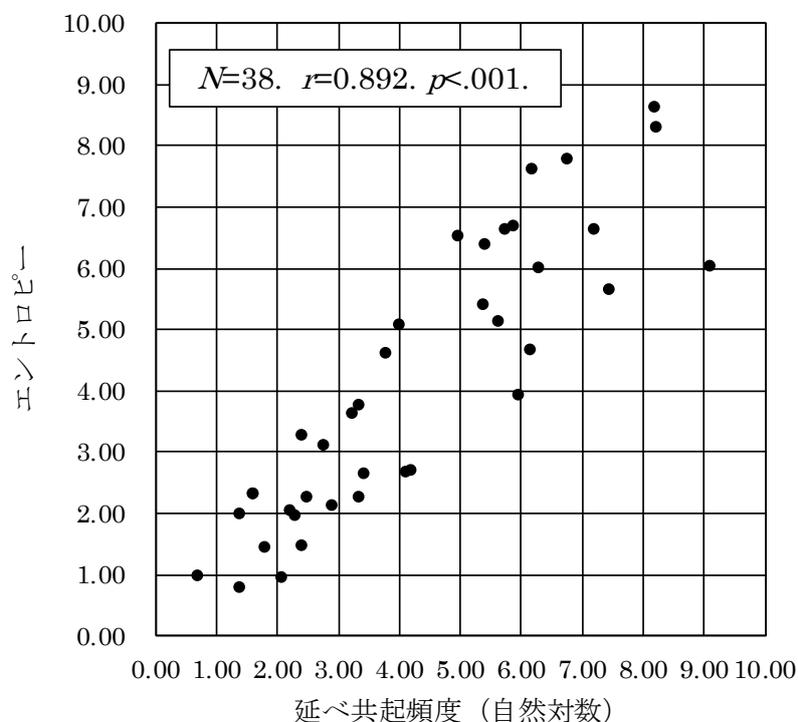


図2 様態と結果の副詞と動詞との述べ共起頻度とエントロピーのプロット

4. 様態と結果の副詞的表現の比較

様態と結果の副詞的表現を表 3 の指標を使って比較する。まず、共起動詞数については、様態の副詞的表現が平均で 232 語、結果の副詞的表現が 10 語であった。自然対数に変換した値を使って、独立したサンプルの t 検定(以下、同じ分析)で比較した結果、有意な違いがみられた [$t(36)=5.288, p<.001$]。これは、様態の副詞的表現のほうが結果の副詞的表現よりも、多様な種類の動詞と共起していることを示している。副詞的表現は、文を構成する上で、必須の要素ではないものの、コーパスにおける異なり頻度でみると、やはり結果の副詞的表現よりも、様態の副詞的表現のほうが動詞の選択制限は厳しくないようである。さらに、共起頻度についても同様に自然対数に変換した値で分析した。やはり、様態の副詞的表現のほうが結果の副詞的表現よりも延べ共起頻度も有意に多かった [$t(36)=5.004, p<.001$]。以上の分析により、共起する動詞の重なり頻度でも延べ頻度でも、様態の副詞的表現のほうが結果の副詞的表現よりも、より多く種類の動詞とより頻繁に共起していることを、コーパス頻度で示した。

共起する動詞の重なり頻度も延べ頻度の両方を考慮して、副詞的表現と動詞の共起パターンを指標化したのが、エントロピーと冗長度である。まず、両指標が相互に直接に関係を持つかどうかを検討する。両指標のピアソンの積率相関は $r=0.03$ で、低く、有意な値ではなく、エントロピーと冗長度はそれぞれ独立した指標であると言えよう。そこで、エントロピーと冗長度の値を使い、様態と結果の副詞的表現を t 検定で比較した。その結果、様態の副詞的表現のほうが結果の副詞的表現よりも有意にエントロピーが大きかった [$t(36)=5.297, p<.001$]。しかし、冗長度には有意な違いが見られなかった [$t(36)=.916, p=0.366, ns$]。エントロピーの値からみると、様態の副詞的表現のほうが結果の副詞的表現より多様な動詞と共起していることが分かった。すなわち、様態の副詞的表現の方が緩やかな選択制限を持ち、結果の副詞的表現の方が厳しい選択制限を持っていると言えよう。

5. 動詞との共起からみた様態と結果の副詞的表現の特性

ここまで、副詞の種類である様態の副詞的表現と結果の副詞的表現という分類に基づき、共起する動詞の頻度とパターンについて分析を行った。続いて、個々の副詞的表現の共起パターンの特徴を、エントロピーと冗長度の 2 つの指標から記述的に検討する。

エントロピーと冗長度はそれぞれ異なる指標である。そこで、個々の副詞的表現のエントロピーと冗長度を平均が 0、標準偏差が 1 の z 値で標準化して、様態の副詞的表現 23 語と結果の副詞的表現 15 語の計 38 語を階層的クラスタ分析 (SPSS 18.0 を使用) で分類した。クラスタ間の距離はグループ内平均連結法を用い、個々の副詞的表現の距離は平方ユークリッド距離を使用した²。その結果、25 ポイントのスケールの 19 ポイントで区切り、3 つのクラスタを得た³。なお、交差妥当化済の判別分析で、クラスタ分析で得られた

² 階層的クラスタ分析については、よく使われることの多い「ウォード法・平方ユークリッド」の組み合わせ、「グループ間平均連結法・平方ユークリッド」の組み合わせなど、いくつかの組み合わせを変えて分析した。その結果、判別分析において「グループ内平均連結法・平方ユークリッド」の組み合わせが、交差妥当化済の判別分析で、判別率の中率が最も高く、他の組み合わせに比べて、正確に分類されていたことが分かった。したがって、この組み合わせを採用する。

³ 変数がエントロピーと冗長度の 2 つのみであるためプロットを視覚的に見て判断

3つのクラスタの分類の正確さを判定したところ、判別率中率が97.4%で、非常に正確に分類されていることが示された。3つのクラスタが適切に副詞的表現を分類していることを裏付けた。図3は、エントロピーと冗長度より個々の副詞的表現を平面上にプロットリングしたものである。

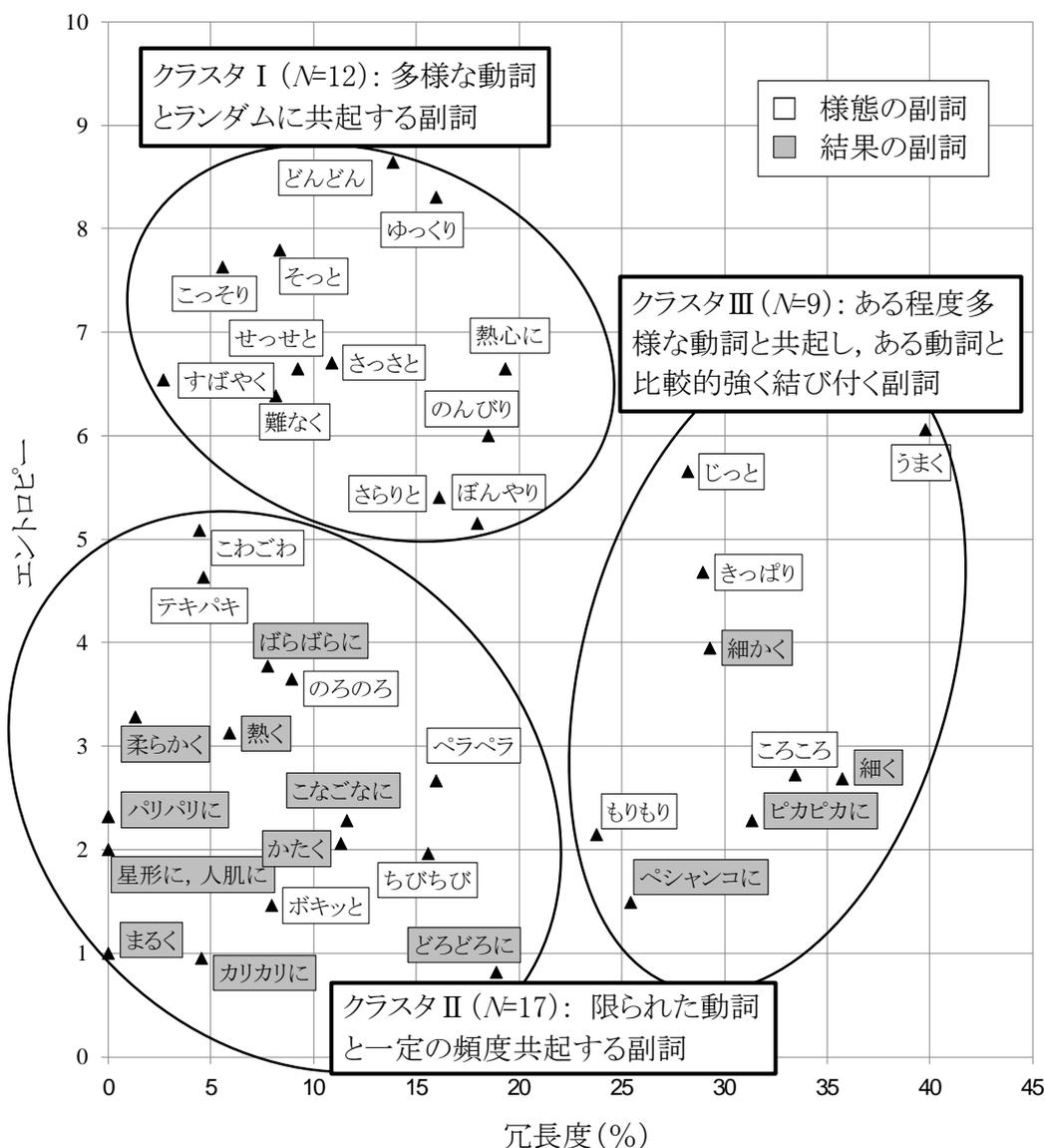


図3 副詞ごとのエントロピーと冗長度に基づくプロットリングと階層的クラスタ分析の結果

クラスタ I (N=12) は、高エントロピー、低冗長度のクラスタである。つまり、とりわけ多様な動詞と規則的に共起する副詞的表現群であるといえる。このクラスタには 12

することもできるが、より数学的かつ客観的な手法で判断するためにクラスタ分析を用いた。また、クラスタ分析を採用する利点として、判別分析によりクラスタの分類の正確さを数値で示すことができるので、より説得力のあるグループ化を示すことができる。

語の副詞が含まれ、結果の副詞は1語も含まれておらず、すべて様態の副詞的表現であった。図3に示したように、「どンドン」「ゆっくり」「そっと」「こっそり」「せっせと」「すばやく」「難なく」「さっさと」「熱心に」「のんびり」「さらりと」「ぼんやり」などで、やはり、様態の副詞的表現の共起パターンが多様であることを示している。

クラスタⅡ ($N=17$) は、低エントロピー、低冗長度の副詞的表現群である。すなわち、限られた動詞と規則的に共起する副詞的表現群であると言えよう。様態の副詞的表現は「ボキッと」「ちびちび」「テキパキ」「ペラペラ」「こわごわ」「のろのろ」の6語で、結果の副詞は「人肌に」「星形に」「パリパリに」「カリカリに」「こなごなに」「かたく」「まるく」「ばらばらに」「どろどろに」「柔らかく」「熱く」の11語であった。クラスタⅡには、結果の副詞的表現が19語中12語(63.15%)含まれ、やはり結果の副詞的表現は、動詞に対する選択制限が厳しいことを示唆している。しかし、様態の副詞的表現も17語中6語(35.29%)が含まれており、様態の副詞的表現の中にも、選択制限の厳しい副詞的表現があることが分かる。

クラスタⅢ ($N=9$) は、クラスタⅠほどエントロピーが高くなく、またクラスタⅡほどエントロピーが低くもない、中くらいのエントロピーであった。ただし、冗長度は高い副詞的表現群である。クラスタⅢに属する副詞的表現は、ある程度、多様な動詞と共起し、ある動詞と比較強く結びつく副詞的表現であると言えよう。様態の副詞的表現5語(「きっぱり」「もりもり」「じっと」「うまく」「ころころ」)、結果の副詞的表現4語(「細かく」「ペシャンコに」「ピカピカに」「細く」「熱く」)が含まれていた。

図3を概観すると、様態の副詞的表現は3つのクラスタに散らばっているのに対し、結果の副詞的表現はクラスタⅡに集中していることがうかがえる。クラスタⅡは特定の動詞と一定の頻度で共起する副詞的表現群である。結果の副詞的表現と共起できる動詞が、様態の副詞的表現ほど多様ではないこと、すなわち選択制限があることが読み取れる。しかし、クラスタⅡに含まれなかった結果の副詞的表現もあり、本研究で扱った結果の副詞的表現15語が全て同じ共起パターンを持っていないことが確認できた。

一方、様態の副詞的表現は、図3の全体に広く分布していた。影山(2009)は、結果の副詞的表現は動詞が含意する結果を補足することに対し、様態の副詞的表現は結果の副詞のように共起する動詞の含意とは関係なく、多様な動詞との共起が可能であるとしている。影山(2009)では「賢三が悲しげに話す」という例を挙げている。「話す」という動詞には「主語が悲しくなる」という意味は含意されておらず、「悲しげに」は意味的に、動詞の「話す」から独立して、「話す」の様態を表している。また、仁田(2002)は、様態の副詞的表現は動きの展開過程に関わる諸側面に言及できるという意味で雑多で多様な存在であるとしている。影山(2009)および仁田(2002)の見解より、様態の副詞的表現は多様な動詞と共起できることが考えられる。そして、様態の副詞的表現が図3の全体に分布していることは、影山(2009)および仁田(2002)で述べられていることを裏付けていると言えよう。ただし、クラスタⅡには、様態の副詞的表現が含まれているので、すべての様態の副詞的表現が多様な動詞と共起するわけではないとも示している。クラスタⅡに含まれる様態の副詞的表現については、各副詞的表現の個別の意味と意味拡張の関係を考察することでクラスタⅠおよびⅢとの様態の副詞との違いを明らかにできると思われる。しかし、本稿では個別の副詞について考察することは、本研究の範囲を大きく広げることにな

るので、稿を改めたい。

6. 結論

副詞的表現にも文中での基本的な生起位置が存在すると予想される（三原，2008；Koizumi，1993；小泉・玉岡，2006）。難波・玉岡（2014）は、大規模コーパスで、目的語を伴って出現する他動詞と共起する副詞的表現を検索し、様態の副詞的表現は目的語の前後に等しく生起しているのに対し、結果の副詞的表現は目的語の後に有意に多く生起していることを示した。文中の生起頻度の違いは副詞的表現の共起する動詞の多様性によるものであると考察した。本研究では、難波・玉岡（2014）の考察の妥当性を検証すべく、結果と様態の副詞的表現の動詞との共起パターンをエントロピーと冗長度の2つの指標で比較した。その結果、様態の副詞的表現は、結果の副詞的表現よりもエントロピーが有意に大きく、選択制限が緩やかであることが分かった。さらに、エントロピーと冗長度の2つの指標を使って、様態の副詞的表現と結果の副詞的表現 38 語の動詞との共起パターンをプロットして図 3 に示した。クラスタ分析により 3 つのクラスタに分けられた。各クラスタを見ていくと、多様な動詞と共起するクラスタは様態の副詞的表現で占められ、限られた動詞と共起するクラスタは結果の副詞的表現が多く含まれていた。個々の副詞的表現の共起パターンで見ても、様態の副詞的表現は選択制限が緩く、結果の副詞的表現は厳しいことが示された。ただし、様態の副詞的表現は、選択制限が緩やかなものから厳しいものまで広範囲であることも分かった。本研究では、様態の副詞的表現と結果の副詞的表現の共起パターンをエントロピーと冗長度の指標で計量的に分析し、同じ動詞句内に生起する副詞的表現であっても、様態と結果の副詞的表現は異なった特性を持つこと実証した。

文献

- Koizumi, M. (1993) Modal phrase and adjuncts. *Japanese/Korean Linguistics*, 2: 409-428.
- Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., Koiso, H., Yamaguchi, M., Tanaka, M., & Den, Y. (2014) Balanced corpus of contemporary written Japanese. *Language Resources and Evaluation*, 48: 345-371.
- Miyaoka, Y., & Tamaoka, K. (2005) An investigation of the right-hand head rule applied to Japanese affixes, *Glottometrics*, 10: 45-54.
- Shannon, C. E. (1948) A mathematical theory of communication. *Bell System Technical Journal*, 27: 379-423 (Part I) and 623-656 (Part II).
- Simpson, E. H. (1949) Measurement of Diversity. *Nature*, 163: 688.
- Tamaoka, K., Hyunjung L., & Sakai, H. (2004) Entropy and redundancy of Japanese lexical and syntactic compound verbs. *Journal of Quantitative Linguistics*, 11(3): 233-250.
- Yoshikane, F., Kageura, K., & Tsuji, K. (2003) A method for the comparative analysis of concentration of author productivity, giving consideration to the effect of sample size dependency of statistical measures. *Journal of the American Society for Informative Science and Technology*, 54(6): 521-528.
- 有本卓 (1982) 『確率・情報・エントロピー』 森北出版.

- 海保博之 (1989)「第 1 講：情報をはかる—エントロピー・伝達情報量・冗長度」海保博之 (編)『心理・教育データの解析法 10 講—応用編』(pp. 14-26), 福村出版.
- 影山太郎 (2009)『日英対照—形容詞・副詞の意味と構文』大修館書店.
- 影山太郎 (2013)『日英対照—動詞の意味と構文』大修館書店.
- 影浦峽 (2000)『計量情報学：図書館/言語研究への応用』丸善.
- 前川喜久雄 (2007)「コーパス日本語学の可能性—大規模均衡コーパスがもたらすもの—」『日本語科学』22: 13-28.
- 中俣尚己 (2015)「初級文法項目の生産性の可視化—動詞に接続する文法項目の場合」『計量国語学』29(8): 275-295.
- 木山幸子・玉岡賀津雄 (2011)「自他両用の「一化する」における自動詞用法と他動詞用法の比較：新聞コーパスの用例に基づく多変量解析」『言語研究』139: 29-56.
- 小泉政利, 玉岡賀津雄 (2006)「文解析による日本語副詞類の基本語順の判定」『認知科学』13: 392-403.
- 難波えみ・玉岡賀津雄 (2014)「コーパス検索による副詞の文中における基本生起位置の検討」『第 6 回日本語学コーパスワークショップ予稿集』国立国語研究所.
- 仁田義雄 (2002)『新日本語文法選書 3：副詞的表現の諸相』くろしお出版.
- 堀淳一 (1979)『エントロピーとは何か』講談社ブルーバックス.
- 三原健一 (2008)『構造から見る日本語文法』開拓社.
- 鈴木崇史・影浦峽 (2011)「名詞の分布特徴量を用いた政治テキスト分析」『行動計量学』38(1): 83-92.
- 玉岡賀津雄, 木山幸子, 宮岡弥生 (2011)「新聞と小説のコーパスにおけるオノマトペと動詞の共起パターン」『言語研究』139: 57-84.
- 玉岡賀津雄 (2011)「コーパス分析の研究例 2: 複合動詞の計量的解析」中本敬子・李在鎬 (編)『認知言語学研究の方法』(pp. 181-195), ひつじ書房.
- 玉岡賀津雄 (2016)「共起表現研究のためのコーパス検索入門」『レキシコンフォーラム』7: 239-264.