

連体修飾節のデ格の用法について

阿辺川 武

東京工業大学大学院 総合理工学研究科

abekawa@lr.pi.titech.ac.jp

奥村 学

東京工業大学 精密工学研究所

oku@pi.titech.ac.jp

1 はじめに

日本語では名詞を修飾する形式は多岐に渡りますが、その 1 つに連体修飾節による修飾形式がある。

- (a) さんまを焼く男
- (b) さんまを焼く匂い

連体修飾節内の動詞と被修飾名詞の関係をみると、(a)のように名詞が動詞に対して格要素となり得る関係と、(b)のように格要素になり得ない関係に大別される。寺村 [8] は前者を「内の関係」、後者を「外の関係」と呼んでいる。内/外の関係は、動詞・名詞間の共起関係を考慮し決定されるが、その他にも名詞が外の関係をとれるか、被修飾名詞に他の連体要素が付属しているかなど様々な要因により決定される。共起関係から内/外の関係は被修飾名詞が連体修飾節内の動詞の必須格として挿入される場合は容易に判断できるが、任意格、特にデ格の場合、共起関係だけでなく名詞の意味を考慮しなければ内/外の判断ができない。

- (c) 肖像権を認める判決

(c) の場合「判決で肖像権を認める」ということもできるし、連体修飾節が「判決」の内容を説明しているとも考えることもできる。一般に連体修飾節において内/外の判別をする場合、動詞と名詞の共起関係に基づいて判断する手法が多い。この手法では「判決」と「認める」との間に共起関係が存在すれば、内の関係になるし、存在しなければ外の関係になる。そして新聞コーパス上では実際に共起関係は存在する。しかし共起関係が存在するからといって、これを内の関係とすることはできない。そもそもこのような事例の場合、内/外の関係のどちらであるかという明確な基準が存在しない。判断は解析の目的によって異なるし、正解を付与する人によっても異なる。このような内/外の関係の判別に迷う事例はデ格の格要素となる場合が多いようである。そこで本稿ではこのようなデ格と外関係の判断が揺れる原因について探っていく。

2 連体修飾節とデ格の用法

2.1 デ格の用法

デ格の用法といった場合、まず問題になるのは用法の分類基準である。デ格ではガ格、ヲ格など比べ多くの用法があり、どのような基準によって分類するかは研究者によっても様々である。例えば国立国語研究所 [2] では 14 種類、石綿 [1] は 11 種類、EDR [6] ではデ格に該当する用法に限定しても 13 種類ある。用法の分類がはっきりしていると思われる [場所] でも「公園」「川底」などの具体的な場所と「会議」「コンサート」など実際の空間に射影できない場所を [場面] として区別しているものもある。

本稿では後半で EDR コーパスを用いることから、EDR で採用されているデ格の用法の分類を利用することにする。表 1 は、デ格の用法別に単文の格要素を取り出して連体節化できるか、また逆に連体節から単文に言い換えられるかを記したものである。ただ EDR では用法を「事象・事実表現のための関係子」と呼んでおり、本稿の意味する用法とは若干異なるかもしれない。同様な考察が丸元ら [3] においてもなされているが、デ格の分類として田窪らの分類を用いており、本稿で必要な分類とは異なるため、特定の連体修飾節の形式を扱っていない。

2.2 内の関係とデ格の用法

連体修飾節と被修飾名詞の関係を考えたとき、内の関係の基本的な構造は動詞と格要素となるかどうかである。ガ格、ヲ格に対しては動詞の必須格として共起関係という枠組みで考えることができるが、二格の一部やデ格のように動詞の任意格の場合、共起関係だけではうまく現象を捉えることができない。例えばデ格の用例の 1 つである [場所] を考えたとき、その用法をとる名詞は、たいていの動詞に対してデ格の格要素となることができると考えられる。一方、[原因・理由] を考えたとき、その用法をとる名詞は、ある動詞では格要素となれるが、格要素になれない動詞も数多いと

表 1: デ格の用法別の特徴

用法	EDR での呼称	単文→連体節	連体節→単文	例文
動作主体	agent	○		チーム全員で様々な考えを巡らした．
場所	place	○	○	カーネギーテックで教鞭をとった．
場面	scene	○	○	連絡会議で取材状況などを伝える．
時間	time	○	○	その時点でヤマ場を迎えていた．
材料	material	○	○	大理石でビルを建てた．
手段・道具	implement	○	○	長いしっぽで体の釣り合いをとっている．
原因・理由	cause			豊作でキャベツが暴落した．
条件	condition			直進時で 4%の省エネになる．
様態	manner			起き抜けで編集局に駆けつけた．
目的	purpose			ガソリンの不足を補う目的で使用された．
単位	unit	×	×	15 年で 3 倍になった．

推測される．この違いはなんであろうか？それは名詞の持つ意味である．[場所][時間][動作主体]をとる名詞を考えたとき，それらはすべて具体的な意味を持つ．具体的な意味を持つが故にデ格で動詞の格要素となったときに，用法が一意に決定される．一方[原因・理由][様態]などの用法をとる名詞では，名詞単独では意味が足りないことが多い．それが故に名詞の意味を補充するためにある外の関係という形式が成り立つのである．外の関係を取り得る名詞であっても連体修飾節以外のなんらかの修飾により意味が補充されると，連体修飾節をとったとしても外の関係にならず，内関係になってしまう．

先程も言及したように，デ格は任意格のため動詞との結びつきは弱く，名詞の具体性によって用法が決まる．外の関係をとる名詞では連体修飾節があってもその名詞単独では意味が弱いので，デ格として内関係にはなれない．なんらかの修飾要素が付随すれば意味が補充され，内関係として動詞の格要素になれるのである．

ただし前後の文脈により，名詞の持つ意味の補充性は異なってくる．外の関係をとる名詞でも，あらかじめ前の文脈で具体的な内容が言及されていれば，十分な意味を有しており，連体修飾節で使用されても内関係となることができる．例えば，冒頭で示した「肖像権を認める判決」は内ともいえるし外ともいえるとしたが，これは文脈により異なるのである．「判決」についてなんの裁判かどこで結審したかなど事前に説明があれば，単独でデ格の格要素となり「判決で肖像権を認める」といっても意味が通る．一方判決について初出であると単独でデ格の格要素となっても一体何の判決かはわからなくなるため，外関係として判断す

べきであると思われる．内か外かで揺れる連体修飾節は，その時点での被修飾名詞の意味の補充度を考え，単独でデ格の格要素として意味が通るなら内であるし，意味を補充しないといけない場合は外の関係とした方がよい．

つまりこのような場合の内/外の判別は，動詞と名詞の共起関係はいうに及ばず，そのときの文脈に則した名詞の意味の補充度を考えながら，総合的に判断しなければならないということである．

意味がはっきりしている名詞，つまり具体名詞の多くはデ格で用いられても，デフォルトの用法があるので，単独で使用されても違和感はない．一方，抽象的な名詞の多くは，意味が不完全なために単独でデ格で用いられても，全体の意味も不完全になってしまう．「豊作」や「円高」などの抽象名詞の中でも意味がかなり限定されている場合は，単独でも使用できる．

動詞との兼ね合いで名詞の意味に現実性を帯びるということもある．

2.3 外関係とデ格の用法

[様態]は外関係と関係がある．「裸で子供が駆け回る」「猛スピードで車が走る」といった[様態]で使われる場合は「裸」や「猛スピード」は具体性のある名詞なので「?子供が駆け回る裸」「?車が走る猛スピード」のように連体修飾節にしても文意がおかしくなる．一方「猛スピード」でも具体性をなくした「スピード」では「車が走るスピード」となり外関係として文意が成立する．逆に具体性のない「スピード」のみでは「?スピードで車が走る」と単独でデ格として使用することができない．

具体性を持つ名詞はデ格として使用されても，用法

はほぼ固定される．逆に 1 つの用法では，具体性，抽象性の両方をとることはあまりないと思われる．抽象性・具体性の名詞をとる用法は，修飾する要素により意味が補充されているわけで，その意味により用法が変化すると思われる．

シソーラスを使った手法からの観点と，共起関係を使った手法からの観点共起関係ではうまく扱えない例

- 母が入院したショック
- ショックで母が入院した
- 公道を走る練習
- 練習 [で] 公道を走る

「原因」「様態」といった用法はその他に比べて結びつきが弱い．なんとかして説明できないものか？結びつきが強いものと弱いものがある．こういうときは，その他の説明する要素が多いのではないだろうか？つまり文脈が冗長にならざるを得ない．

- その法案を審議した国会
- 国会 [が/で] 審議した (動作主体)
- 与党だけが審議した国会
- 国会 [で] 与党だけが審議した (場面)

複数の候補がある場合でも既に格要素が埋まっていれば，候補は限定される．

連体修飾時は，連体修飾節の中でどの格になるかは，ほとんど意識されない．単文に戻そうとしたとき，格の間で揺れようが，用法間で揺れようが気にしていない．わざわざ戻そうとするから，問題が生じる．

飛行機が墜落した事故→外の関係足を骨折した事故→事故で足を骨折した

彼が出家する動機 (ある) 動機で彼が出家する

3 デ格の用法のクラスタリング

前節で見てきたようにデ格は，同じ名詞が前接しても動詞によって用法が異なる．しかし，すべてのデ格の用法がとれるわけではなく，いくつかの用法のみに限定されると思われる．例えば「喧嘩」で使えるのは「場面，原因，手段・道具，様態」の 4 種類ぐらいである．そしてそれらの用法の中で解釈がゆれる．

同様に他の名詞をとっても，取り得るデ格の用法はある程度限定されていると思われる．逆にデ格の用法から見れば，前接する名詞もすべての名詞ではなく，限定された名詞の集合である．

以上の観点から，デ格の用法をクラスタリングを行なった．予想では，判断がゆれる用法同士は，おたがいに類似度が高いはずである．

3.1 手法

実例に基づいたクラスタリングを行なうために，本研究は EDR 日本語コーパス [6] を利用した．このコーパスには文節ごとに意味情報が付与されており，デ格の用法を得ることができる．ここから名詞に対して格助詞デが後接する用例が収集した．ただし接続助詞や判断のダの連用形など格助詞以外の用法で使用されている場合を考慮して，助詞デの後ろに係助詞や句点が存在する文は収集の対象としない．

EDR コーパスでは，ガ格やヲ格といった他の格に対しても用法を付与しており，用法の種類はデ格に付与される数よりも多い．そして 20 万文という大量の文に対して用例を付与しているため，実際には誤っているものもある．そのため本来デ格ではとりえない用法も付与されている．したがってデ格として取り得る用法のみを独自に判断して使用している．実際に収集された用例別の数を表 3.1 に示す．ただし用法のラベリングが間違っているといっても，はっきりしたものは間違えていないものと仮定する．間違っているとすれば，用法の揺れるところで，つまり間違いやすいところで間違えていると思われる．

デ格を収集した際，格要素の名詞に対して，EDR 概念体系の概念識別子がそれぞれ付与されている．クラスタリングにはこの概念識別子を用いる．ただしこのままでは事例数に対して概念識別子の種類数が多すぎるので，概念体系の階層構造を用いて汎化を行なう．ルートから 7 階層より深い概念識別子は，祖先にあたる 7 階層の概念識別子を用いた．そして用法別にどの概念識別子が何回出現したかの頻度を求める．

各用法の概念識別子をベクトルの成分と考えれば，各用法をベクトルで表すことができる．このベクトルを用いてデ格の用法をクラスタリングする．

動作主体 abcdef:20, ..., 000adf:1

場所 abcdef:0, ..., 000adf:2

クラスタリングの手法は，各ベクトルの類似度を計算し，最も類似度の大きいクラスタ同士を結合し，ボトムアップにまとめていく手法とする．類似度の計算にはコサイン距離を用いる．2 つのベクトル $\vec{x}, \vec{y}$ に対してコサイン距離は次の等式で計算される．

$$\cos(\vec{x}, \vec{y}) = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$$

この類似度を使用して組み上げた木が図 1 である．

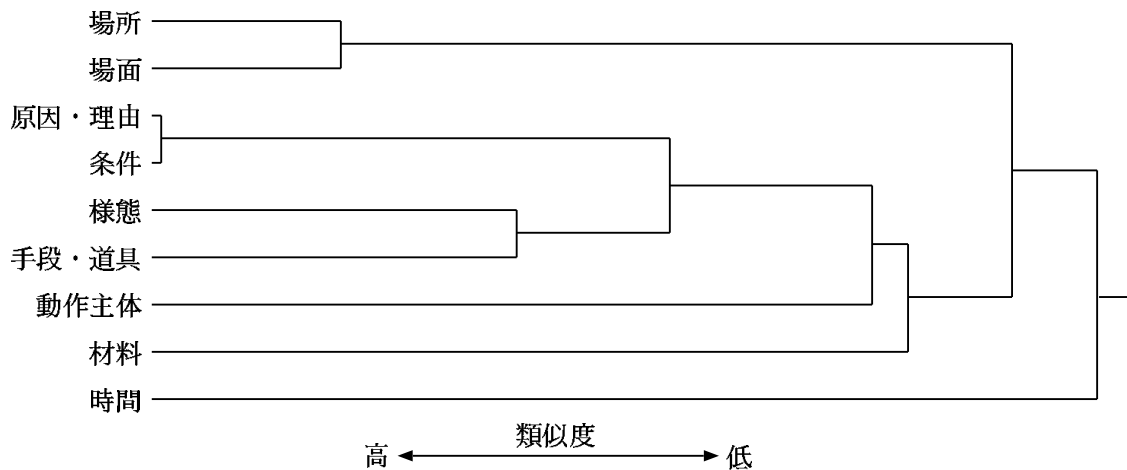


図 1: 意味素性によるデ格の用法の分類

表 2: 用法別の頻度

場所	場面	原因・理由	条件	様態	手段・道具	動作主体	材料	時間
253 日本	320 中	52 疑い	11 中	242 形	108 手	79 自分	5 材料	52 歳
196 中	135 分野	44 円高	10 影響	74 自分	81 方法	37 みんな	2 木	46 時点
136 東京	127 間	38 事故	8 問題	57 共同	78 電話	25 人	2 板	30 後
136 各地	90 なか	21 影響	8 円高	48 状態	77 形	9 署	2 土	28 年
125 米国	69 市場	20 問題	8 意味	47 年	67 車	9 社	2 大理石	22 あと
87 全国	54 面	18 事件	7 人	40 意味	65 目	8 間	2 人	21 日
85 世界	51 段階	16 病気	6 形	39 円	49 円	8 家族	2 樹脂	18 段階
83 国内	39 選挙	15 理由	6 疑い	38 歳	47 機	7 同士	2 者	17 途中
79 病院	38 世界	15 現行犯	5 理由	37 姿	44 自分	7 皆	2 紙	17 今年
78 ホテル	36 大会	14 容疑	5 登場	35 途中	40 テレビ	6 夫婦	2 金	15 現在
7156 計	3999 計	1616 計	761 計	3523 計	6661 計	385 計	170 計	452 計

木構造の葉に近い場所で結合しているほど類似度が高く、葉から遠ざかるほど類似度は低いことを表している。

3.2 考察

図 1 より [原因・理由][条件] の用法が一番高い類似度となることがわかった。研究者によりこれらの用法は特に区別されていないことから、非常に判断の揺れる用法であると思われる。先に述べたように [原因・理由] になるか [条件] になるかは、事象が既に生じたのかそれともこれから生じたのかという差でしかない。したがって名詞だけ用いて分類を行った今回の手法では、区別できなかったのであろう。

次に、[場面] と [場所] の類似度が高い。これも研究者によっては区別していない用法である。両用法共に「中」が高い頻度で出現しているが、これも用法が区別できない一因である。今回の手法では助詞デに前接している名詞のみを見てきたが、名詞「中」はそれだけでは意味をなさない場合が多い。「会議の中」であれば、[場面] であろうし、「お店の中」であれば [場所] で

あろう。用法自体が判断に揺れがある上に、前接する名詞だけを使用したことから、類似度が高くなったものと思われる。

森山 [5] は、事態の背景の客観性について述べており、[場所][時間][材料][手段] などの場合では事態成立との関わりは、かなり客観的に同定されうるが、[原因] の場合は必ずしも客観的であるとは言い切れない面があるといっている。図 1 では、[場所][時間][材料] などの用法は、類似度の低いところで結合されていることから、ラベル付与者間に揺れは存在しない、もしくは名詞に用法の揺れは存在しないことがいえ、森山の主張を裏づけている。

4 おわりに

[3],[7],[4] を参照本稿では、デ格の用例に着目し、デ格の用例の分類が研究者によって異なることを説明した。実際にラベル付された文を用いてデ格の用例のクラスタリングを行った結果、分類の揺れる用例は類似度が高いことが実証された。また、連体修飾節の解析

との関連についても述べることができた。

参考文献

- [1] 石綿敏雄. 現代言語理論と格. ひつじ書房, 1999.
- [2] 国立国語研究所. 日本語における表層格と深層格の対応関係. 三省堂, 1997.
- [3] 丸元聡子, 乾裕子. 連体修飾を受ける体言の格構造の復元・コーパスに基づく「内の関係」の分析. 言語処理学会 第 6 回年次大会 発表論文集, pp. 16–19, 2000.
- [4] 松本善子. 日本語名詞句構造の御用論的考察. 日本語学, Vol. 12, No. 11, pp. 101–114, 1993.
- [5] 森山新. 認知的観点から見た格助詞デの意味構造. 日本語教育, Vol. 115, pp. 1–10, 2002.
- [6] 日本電子化辞書研究所. EDR 電子化辞書仕様説明書第 2 版. Technical Report TR-045, 1995.
- [7] 仁田義雄. 格のゆらぎ. 言語, Vol. 24, No. 11, pp. 20–27, 1995.
- [8] 寺村秀夫. 連体修飾のシンタクスと意味—その 1～その 4—. 「日本語・日本文化」4 号～7 号, 1975–1978.