

# 助詞「に」の定量的分析の試み 語法研究の新たな手法を求めて

京都大学大学院 人間・環境学研究所 李在鎬

## 1. はじめに

本稿の目的は、二点ある。一点目に、従来多様な用法を持つことから、(記述・説明レベルで)一般化が難しいとされてきた助詞「に」に対して、生起文脈の制約に基づく記述・分析を提案する。二点目として、助詞「に」をケーススタディとし、文法の記述に定量的手法を提案することである。とりわけ、当面の目的として、以下の(1)に見られる助詞「に」の機能的・文法的意味の多義性について考察していくと同時に、その動機づけを明らかにする。

- (1) a. 車が壁にぶつかった。
- b. 武が二階に荷物を運んだ。
- c. コックさんが肉を平らに叩いた。
- d. 体が大きい割に口は小さい。

日本語の助詞「に」は、(1) a や b のように動作の着点、ないしは移動の方向を表す用法、c のように変化後の結果状態を表す用法、さらには、d のように副詞句を表す用法など、多様な使用文脈を持つ。こうした問題に対して、本稿では、まず、「に」格の問題に対する従来のアプローチの問題点を明らかにする。とりわけ深層格で代表される動詞中心の分析モデルを批判的観点から検討する。そして従来の分析が抱えた記述の循環論に対する具体的な問題提起を行った上、その代替案となるモデルを示す。そのモデルは二本の柱で構成される。第一の柱として認知言語学および構文文法の理論的・記述的立場に準拠し、使用文脈の本質的重要性を認め、文脈ベースのアプローチを採用する (cf. Langacker 1987, 1990, 2000, Goldberg 1995)。第二の柱として定量的言語分析の重要性を認め、大量言語データに対する統計的評価法を採用する。

最終的には次の論点を明らかにする。語法研究の観点から「に」の意味的分布は(構文や名詞といった)使用文脈の制約に支えられていることを示す。具体的には、「に」の使用の背後には、5つのクラスタの存在が見られる点を指摘する。同時に、「に」の動機づけをめぐる問題として、3つのクラスタ化の傾向が見られる点を明らかにする。それには、1) 共起する名詞に強く制約されるタイプ、2) 統語的環境に強く制約されるタイプ、3) 名詞と統語的環境の両方に制約を受けるタイプが存在することを示す。そして、各クラスタ間の相互関係を明らかにする。最後に、分析手法の観点から本研究のアプローチが意味拡張を定量的に捉えるものであると同時に、文脈といった複雑な言語現象に対しても、システムティックな記述・分析を与えることが可能であるという見通しを示しておく。

## 2. 問題提起と代替案

国立国語研究所(1997)に代表される従来の意味格をめぐる分析においては、動詞こそが文の中心であり、格現象はそこから生じる付随的現象と位置づけられる。しかし、動詞の意味に対する過剰な信念は場合によっては一貫性の欠けた分析となることがあり、とりわけ、次の3点の問題点が指摘できる。

- A. 無理な一般化
- B. 記述の歯止め
- C. 観察記述の対象が不明確

動詞を中心とした分析では A で示すように動詞の意味から予測困難な場合、無理な一般化をせざるを得ない (cf. Goldberg 1995, 李 2003)。また B で示すように動詞の意味の数だけ深層格を増やさなければならず、記述の歯止めがきかない (cf. Langacker 1991)。最後に A、B) の問題から派生的生じる分析上の問題として C) が

ある。この点に関連し、具体例として村木（1994）による（2）を示す。

- |        |             |      |
|--------|-------------|------|
| (2) a. | 子供が親に甘える。   | 「態度」 |
| b.     | 彼をゲストに迎える。  | 「資格」 |
| c.     | 父親が娘をひざにだく。 | 「部分」 |

村木（1994:150-165）

「に」の意味的役割についての村木（1994）の分析には、次の問題点が指摘できる。それは従来の研究によく見られる傾向として分析者自身は「に」格を記述しているつもりであっても、その分析が（2.a）のように動詞の意味を記述しているとしたか考えられないもの、または（2.b）のように名詞句全体を記述しているもの、さらには（2.c）等のように名詞句と名詞句の関係を記述しているとしたか考えられないものが含まれており、観察対象の一貫性の問題に対して必ずしも明示的でない。

こうした問題点を踏まえ、本発表では新たな方法論を導入する。それは、次のような理論的含意を持つ。まず、動詞といった単一の構成要素の制約をプリミティブなものとして仮定する手法を放棄する。そして、本質的制約を前後の使用文脈、すなわち共起する語や統語的文脈から生じる構文の制約で捉えなおす。この種の方法論は、一言で言うならば、「意味 = 生起文脈・生起環境」と定義する手法である<sup>4</sup>。

一見、極論とも思われるが、こうした手法を用いることで、上記の3点の問題点をうまく回避できる。とりわけA)の無理な一般化に関する問題と、C)の観察対象の一貫性に関する問題は明らかにクリアできると考えられる。なぜなら、このアプローチは、動詞の制約とは、別のレベルで意味格を捉えることを目的としており、「に」が生じる使用の文脈を観察の基本単位と定めているからである。

しかし、本研究が提案する方法論においても、いくつかクリアしなければならない問題点がある。それは、この手法においても前節で示した従来の研究の問題点、B)の記述の歯止めに関する問題が生じ、同様の循環論を引き起こす可能性が排除できない。なぜなら、本研究の手法をとった場合、文脈の制約を基本とする以上、(動詞の数だけ深層格を認定すること同様に)文脈が異なればそれは異なる深層格と考えなければならず、「に」の具体例の数だけ深層格を認定することになりかねないからである。となると、このアプローチでは、「に」のすべての具体例を観察記述する必要があり、事例が増えるたび「に」の意味を書き換える、ないしは書き加えていかなければならず、先行研究以上に記述の歯止めが利かなくなる。

こうした問題を克服するため、本研究では次の手法を導入する。それは、頻度効果の観点から問題を捉えなおす。具体的には、以下の二点の分析手法を導入する。

- 1) コーパスベースアプローチ: 大量言語データからサンプルを採集する。
- 2) 統計解析ベースアプローチ: 採集されたサンプルに対して一貫性をもった評価を与える。

以下の本論では、1)に関連しては、複数のコーパスからサンプルを採集し、コンコーダンスを作成する。2)に関連しては、17の変数による1/0の分散表示を行った上でクラスタ分析を試みる。

### 3. 方法（使用データと分析の手順）

本研究では、テキストが持つバイアスを極力排除しつつ、分析の便宜上の理由から次のデータを使用した。まず、一人の書き手によるバイアスを排除するため、8つの異なる書き手によるテキストデータを利用した<sup>5</sup>。次に、分析の便宜上、対象が助詞であることから、トークン頻度が高いことを考慮し、コーパスサイズを限定した。具体的には、前述の8つの小説の冒頭から、(内容に関係なく)それぞれ1000文字のみを取り出した。その結果、8人の書き手による総計8000字のミニコーパスができた。そこから、テキスト検索を行い、コンコーダンスを作成した。そして、語中の「に」(例えば、「こんにちは」の「に」)といったノイズを除去した結果、最終的に「に」を含む203の文がサンプルとして採集できた。

次に、サンプルの分析に際しては、多変量分析の一手法で、数理アルゴリズムを利用した統計学的分類法である「クラスタ分析」を用いた。以下にその概略の手順を示す。

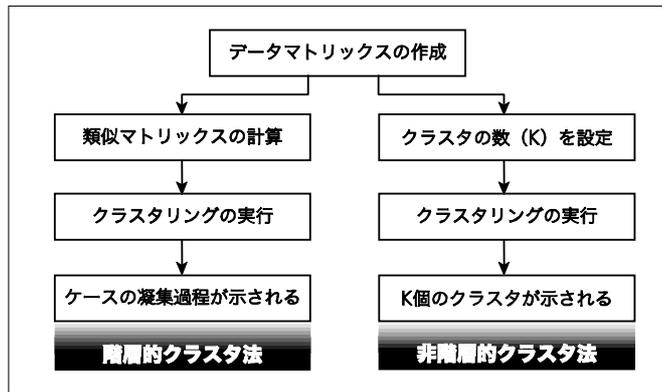


図.1

まず、データマトリックスの作成に関して簡単に述べる。この作業は、一言で言うならば各々のサンプルを計算機が認識できる（数値としての）表示に変換する作業である。実際の作業では、二つの変数セットを用いた。一つ目として「に」と共起語する語の意味（FS1）をめぐって、6変数を用いた<sup>v</sup>。二つ目として「に」が生じる統語関係を中心とし、項構造の基本的関係を構成する前後の文脈環境（FS2）として、11変数を用いた。

区分	変数	具体例
FS 1	主体	{ ガードマン/ 彼女/ 社員/ 客/ 田中/ 妻} に
	場所	{ 仙台/ 玄関/ 会社/ 部屋/ ポケット/ 壁} に
	具体物	{ 容器/ エレベーター/ 電車/ 花瓶/ カメラ} に
	事	{ 具体的/ 頻繁/ 一緒/ 本当/ 妙/ 氷づけ} に
	抽象物	{ 時間/ 気持ち /習慣/ 立場/ 問題/ 不安} に
	抽象的關係	{ 最後/ こと/ 逆/ 時/ 以外/ 第一/ 日} に
FS 2	文頭	X に... (東京に....)
	後続_が	... ~が X に... (...みわこが東京に....)
	後続_を	... ~を X に... (...みわこを東京に....)
	後続_に	... ~に X に... (...一緒に東京に....)
	後続_から	... ~から X に... (...京都から東京に....)
	先行_動詞	...X に動詞... (...東京に送る....)
	受動態	...X に受動態動詞... (...東京に送られる....)
	先行_が	...X に ~が... (...東京に春が....)
	先行_を	...X に ~を... (...東京に生徒を....)
	先行_に	...X に ~に... (...東京に行くことに....)
先行_形容詞	...X に形容詞 (名詞述語) ... (...東京に多い....)	

表.1

表1では、データマトリックス作成において、利用した変数を示した。FS1に関しては、具体的な共起語の例を示し、FS2に関しては構文スキーマを示した。なお、FS2の変数設定に関連し、次の点を補足しておく。「に」の生起環境を形成する文脈は、「で」や「の」といった格関係、あるいは「は」や「も」といった取立詞など、表1のもの以外にもたくさんのディテールが存在する。さらには、埋め込み節や複文といった統語的制約によって実現される構造なども存在する。これらのディテールが存在するにも関わらず、本研究がたったの11変数のみで解析を行った背景として、次の点に注目してほしい。本研究の基本的目的は構文文法の観点から、項構造の問題を捉えるところにある。これは、文文法における項としての「に」を捉えるものであり、この点においてのみ、FS2の変数は十分なものと位置づけることができる。

次に、各ケースに個々の変数が示す特徴が存在する場合は「1」、存在しない場合は「0」で表した（cf.

Romesburg 1989, Kohonen 1995) 具体例として、以下の(3)を例に考えてみよう。

- (3) a. 彼は訪問者をエレベーターに**に**乗せた。  
 b. みわちゃんに**に**は3歳の息子がいる。

まず、(3) a を例に考えてみたい。下線部に注目した場合、助詞「に」と直接的な共起関係にある語「エレベーター」は、場所であり、具体物であるという特徴を示す。同時に、助詞「に」が生じる文脈環境を考えてみた場合、「を」格の後ろに生じ、能動態の動詞の前に生じている。このことを0と1で表現した場合、以下のようなになる。

区分	主体	場所	具体物	事	抽象物	抽象的関係	文頭	後続が	後続を	後続に	後続から	先行動詞	受動態	先行が	先行を	先行に	先行形容詞
(3) a	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0
b	1	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0

表2

表2で示すコード化作業の結果、2233の1/0の値で構成されたデータマトリックスが得られた。次に、図1の手順に従って、データマトリックスを計算機に与え、解析を実行してみた(統計解析アプリケーション SPSS Ver10.1 Win版を使用)。なお、解析の際には、非階層的手法(K-means)と階層的手法(ワード法、平方ユークリッド距離で測定)をとともに用いて検証を行った。

#### 4. 結果

##### 4.1. 5つのクラスタ

本節では、前節の203のサンプルを17の変数を用いて評価した結果を報告する。考察を始めるに当たって、まず階層的クラスタ分析の結果を図2に示す。なお、クラスタの結合法は、ワード法で平方ユークリッド距離を用いている。

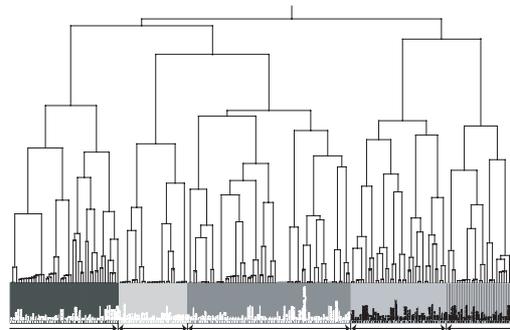


図2

図2では、個々のケースが相互に結合していくことで、それぞれのクラスタが形成されていく過程を示した。この延長で非階層的クラスタ法を用いて、解析を試みた結果、5つのクラスタ(K=5)で最適な分離が得られた。表3は、K=5においてその最終クラスタ中心とケースの数を示すものである(K=5/反復回数は12/初期の中心間の最小距離は2.236)。

区分	クラスタ1	クラスタ2	クラスタ3	クラスタ4	クラスタ5
主体	0.021	0	0.131	0.118	0.172
場所	0.936	0.020	0	0.176	0.724
具体物	0.596	0	0.016	0	0.414
事	0	1	0.574	0.647	0
抽象物	0.021	0.122	0.443	0.353	0.069
抽象的関係	0.277	0.429	0.443	0.471	0.276
文頭	0	0.492	0.148	0	0.483
後続が	0.660	0.286	0.672	0	0.310
後続を	0.106	0	0	1.000	0
後続に	0.021	0.041	0.016	0	0
後続から	0.021	0	0	0	0

先行_動詞	1.000	0.102	1	0.824	0
受動態	0.170	0	0.115	0.118	0
先行_が	0	0.306	0	0	0.414
先行_を	0.021	0.224	0	0	0.517
先行_に	0	0.143	0	0.176	0
先行_形容詞	0	0.204	0	0	0.069
ケース数	47	49	61	17	29

表3

## 4.2. 各クラスタの詳細

本節では、4.1 で示された5つのクラスタの詳細を具体例とともに示す。結論的には次の論点を明らかにする。まず、クラスタ1の「に」は、自動詞移動表現における到達点として特徴づけられる。次に、クラスタ2は副詞、クラスタ3は自動詞結果構文における様態、クラスタ4は他動詞結果構文における変化後の状態、クラスタ5は使役移動表現における移動先および受け手として特徴づけられることを示す。

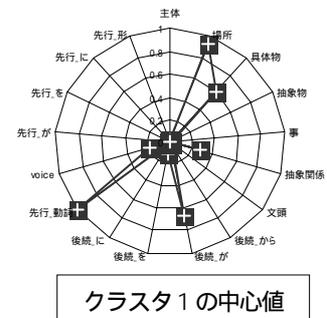
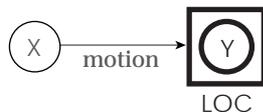
### 4.2.1 クラスタ1 (自動移動表現の到達点)

#### ■ 具体例

- (4) a. 私がエレベータに乗ったとたんに、・・・。  
 b. けっこう夜おそくに誰かが社に戻ってきたりするので・・・。  
 c. 彼女の出勤ローテーションに合わせて、その店に行くことにした。

#### ■ 形式: 「XがYにVする」

#### ■ 意味

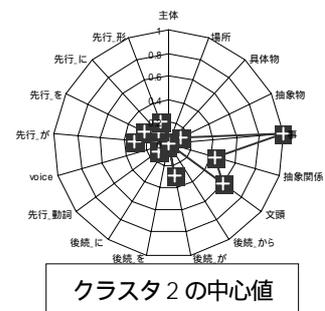


### 4.2.2 クラスタ2 (副詞)

#### ■ 具体例

- (5) a. チップ代わりにマッチ棒が配られ、・・・。  
 b. 体のわりに小さなおちょぼ口をしているのがユーモラスで、・・・。  
 c. 壁に投げつけたときのような妙に扁平な音が聞こえた。

#### ■ 形式: 「XにYがVする」



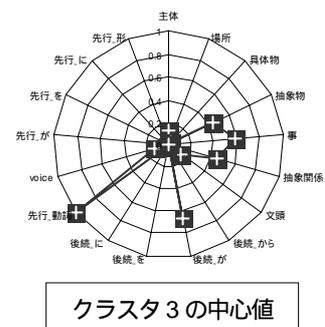
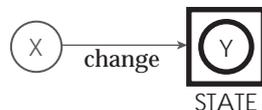
### 4.2.3 クラスタ3 (自動詞結果構文の結果様態)

#### ■ 具体例

- (6) a. そのため私は無防備な気持ちになった。  
 b. 左側の脳で全く別の計算をし、真っ二つに割れた西瓜を合わせるみたいにその二つを合体させるわけである。  
 c. 試合はたちまち夜更けの大勝負にエスカレートしていった。

#### ■ 形式: 「XがYにVする」

#### ■ 意味

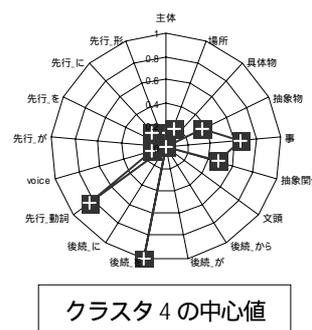
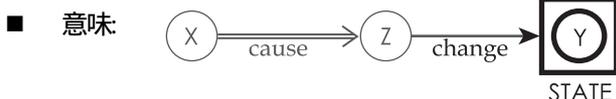


#### 4.2.4 クラスタ 4 (他動詞結果構文の変化後の様態)

##### ■ 具体例

- (7) a. 普段は駐車場になっている部分を道路に作り変え、・・・  
 b. 水島は竖琴をかきたてました。曲は「春らんまん」をマーチ風に変えたものです。  
 c. 女性は長い髪の毛を馬の尾のように無造作に垂らしたり、神社の巫女のように・・・。  
 d. 私はいつもポケットに小銭をためておき、暇があれば・・・。

##### ■ 形式: 「XがZをYにVする」

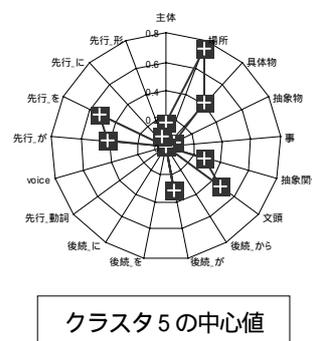
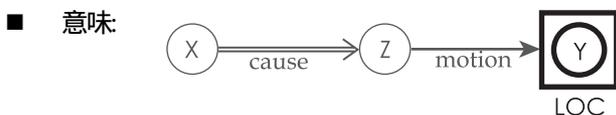


#### 4.2.5 クラスタ 5 (使役移動表現の移動先)

##### ■ 具体例

- (8) a. 私は右側のポケットに百円玉と五百円玉を入れ、左側に五十円玉と十円玉を入れる。  
 b. ステンレス・スチールの壁にぴたりと耳をつけてみたが、それでもやはり音は聞こえなかった。  
 c. 大きなトランクの中に詰められ、その上に重い鎖をぐるぐると巻きつけられ・・・。  
 d. 会社に行き、床やテーブルを掃除し、全員にお茶を淹れる、というきまりがあった。

##### ■ 形式: 「(Xが)YにZをVする」



#### 4.2.6 形式と意味の相関

これまで示したクラスタ分析の結果は、一見したところ、従来の見解を検証したに過ぎず、これといった新しい事実の発掘までには及んでいないと思われるであろう。実際のところ、本研究の分析手法は二つの側面から評価できる。第一の側面として、従来の見解を科学的に検証し、事実のレベルで言語現象を分析・記述する側面を持つ。第二の側面として、新たな見通しを示唆する側面を持つ。ここでは、以上の分析のまとめを兼ね、これらの側面に関連して、二点の事実を指摘する。一点目として、「に」の振舞いを記述する上で、有意義な制約とは何かを示唆する。FS1 に関して言えば、場所や具体物性は貢献度が高いのに対して、主体性といった変数は貢献度が低いことが分かる。次に、二点目として、FS1 とFS2 に規則的相関が見られる。このことは、動詞の制約とは独立に存在する、形式と意味の対からなる説明項としての構文 ( construction ) を実証的に示すものである。

#### 4.3. クラスタ間の非類似度から見た相互の関係

本節では、クラスタ分析を用いる最大の意義でもあるクラスタ間に関する分析結果を報告する。とりわけクラスタ間の類似・非類似の観点から、その親疎関係を明らかにする。

本研究の手法を用いることは、単純な数理モデルに基づく分析を意味するものではない。クラスタ分析を用いる目的の一つには、分析者の直感を科学的に検証し、それを踏まえた分析精度の向上につなげていくところにある。具体的に言えば、「に」の用法と関連し、(多くの分析者に共有されている知見として) 場所を表す「に」と移動先や相手を表す「に」が意味的に類似している、あるいは連続線上にあることが知られて

いる。反対に、副詞を表す「に」と場所や方向の「に」では、その意味的振舞いに大きな差が感じられる。こうした分析者の直感を検証する際、本稿のアプローチは大いに役立つ。特に親疎関係をめぐる記述的要請に関連し、本研究のアプローチは以下のようなクラスタ間の距離から明らかに示すことができる。

クラスタ	1	2	3	4	5
1	-	1.940	1.640	1.676	1.386
2	1.940	-	1.294	1.440	1.572
3	1.640	1.294	-	1.258	1.697
4	1.676	1.440	1.258	-	1.873
5	1.386	1.572	1.697	1.873	-

表4

表4では、各クラスタ間の類似・非類似度が距離として表されている。例えば、クラスタ1の場合、クラスタ2と「1.940」の距離を示すと同時に、クラスタ3とは「1.640」の距離を持つ。こうした数値化は、上記の分析者の直感をうまく捉えている。というのは、クラスタ1とクラスタ5の距離（1.386）から両者が相対的に近い関係におかれていることが分かる。また、クラスタ2とクラスタ1の距離（1.940）から両者が遠い関係におかれていることが分かる。

ところが、表4の非対称性は何に基づいているのであろうか。この問題は、4.1節の表3および図3に照らし合わせて考えてみた場合、自然と明らかになる。この問題に関連し、説明を要する事実は、次の二点である。第一に、クラスタ5とクラスタ1が近い関係を形成できるのは、なぜかという問題、第二にクラスタ2とクラスタ3とクラスタ4が近い関係を形成できるのは、なぜかという問題である。この二点に問題点に対して、(表3が示すように)クラスタ1とクラスタ5はともに場所性や具体物性に高い反応を示すのに対して、クラスタ2、クラスタ3、クラスタ4はともに、事性や抽象物性に高い反応を示していることが分かる。このことを踏まえて、相互の関係を図示した場合、図5のように示すことができる(それぞれの数値は表4のクラスタ間の距離を示す)。

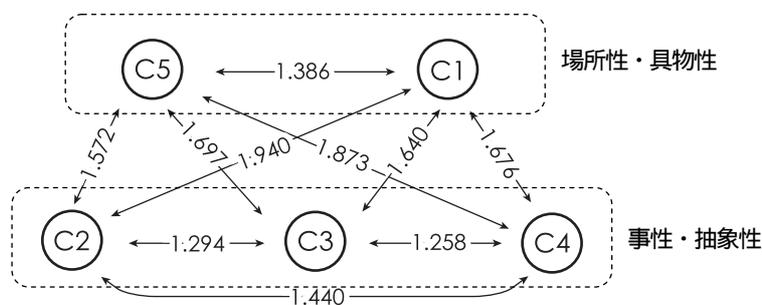


図5

図5が示すように、本研究のアプローチでは、表3の数値データからクラスタ1とクラスタ5が場所性、具物性に高い反応を示していることから、二者の親疎関係を理解することができる。同様の理由からクラスタ2、クラスタ3、クラスタ4の親疎関係に対する記述的要請をうまく満たすことができる。

## 5. 考察(「に」の多義とその動機付け)

本節では、4節で示したクラスタ分析の結果が従来、助詞「に」をめぐって行われてきた分析をどのように評価し、かつ統合しうるかを検討していく。主として「に」をめぐり構造的な多義の動機づけに関する問題を考察し、結論的には三つの非対称的分布関係が見られる点を指摘する。まずは、(8)から(10)の具体例から考えてみよう。

- (9) a. ビスケット生地を円形タルト型に焼いたものに、果物やソースを...  
 b. もえは食の細い先生に色々工夫をこらしたビスケットを焼いたところだった。
- (10) a. 一文銭で銀を見つけたのは金花を嫁に決めたあとと分かり.....  
 b. 子どもにこの学校を決めた理由
- (11) a. 進歩党は幣原を空席の総裁に受け入れると発表した。  
 b. ボランティア活動の機会の提供として施設にボランティアを受け入れることである。

(9) から (11) の例は、a と b の対は、仮に語順による意味的制約を捨象した場合、構造的には同じ構成要素で成り立っていると言える。(9) で言えば、動詞「焼く」に対して、「が格」「を格」「に格」といった要素によって構成されている。これらの例については、次の点に注目してほしい。(9) a で、「に」格は、ビスケット生地を焼いた結果として生じる変化後の状態・様態を表すのに対して、(8) b では、単に焼いたビスケットの移動先・受け手を表している。前節のクラスタで言えば、(10) a はクラスタ 4 の具体例であるのに対して、(10) b はクラスタ 5 の具体例である。同様の交替が (10) や (11) のペアからも観察できる。一方自動詞表現に関する問題として、(12) を考えてみよう。

- (12) a. 旧日本鉄道株式会社（現 J R 東北本線）宇都宮～矢板間旧線は 19 世紀に消えた。  
 b. 船が霧の中に消えた。  
 c. 二日目になって匂いが完全に消えた。

(12) は、動詞「消える」と「に」の共起をめぐる構文現象の具体例である。この例に関しては、まず、そのいずれも同じ文型で生じている点を確認しておきたい。しかし、「に」の意味の問題を考えてみた場合、いずれも異なっている点は、注目に値する。というのも、(12a) では単なる時空間を示すだけであるのに対して、(12b) では船の移動による到達点を示している。さらに、(12c) では副詞を表している。

さて、次の問題として、(9) から (12) でみた様々な交替現象を動機付けるものは何であろうか。果たして、動詞といった単一の要素のみの制約として、これらのデータを説明することは可能であろうか。本研究のこれまでの問題提起に従うのであれば、以上の交替現象は、むしろ動詞の制約とは別の次元で起きているものと捉えるべきであろう。このことは決して、動詞の制約を否定するものではない。というのは、以上の考察は構文現象が決して一枚岩ではないことを示すものであり、その動機付けに関しても単一の制約を仮定することは正しくないことを示唆するからである。とりわけ名詞や構文の制約に注目し、表 3 から問題を捉えてみた場合、次のような非対称的(動機づけの)システムが見えてくる。クラスタ 1 の場合、場所性(0.936) や具体物性(0.595) といった共起語に制約されつつ、動詞に先行(1) し、主語に後続する(0.659) という統語環境に制約されていることが分かる。一方、クラスタ 2 の場合、事性(1) や抽象的關係(0.428) といった共起語に制約されるものの、文頭(0.491) であること以外には統語環境による制約はそれほど顕著ではない。さらに、クラスタ 3 の場合、事性(0.573) といった共起語の制約があるものの、主語に後続(0.672) し、動詞に先行する(1) という統語関係に強く制約されていることが分かる。クラスタ 4 やクラスタ 5 についても同様の視点から観察した場合、以下のようにまとめることができる。

区分	クラスタ 1	クラスタ 2	クラスタ 3	クラスタ 4	クラスタ 5
共起語 (FS1) の制約			×	×	
統語関係 (FS2) の制約		×			

表.5 動機づけの非対称性

表 5 は、クラスタ 1 からクラスタ 5 が、どの変数セットにより強い制約を受けるかを示している。それによると、三つの動機づけのシステムが立ち表れてくる。まず、一つ目として共起語に強い制約を受けるクラ

スとしてクラスタ2、二つ目として、統語関係に強い制約を受けるクラスとしてクラスタ3とクラスタ4、最後に、3つ目として共起語と統語関係の両方に制約を受けるクラスタ1とクラスタ5がある。

#### 6. 最後に（文法記述の定量化がもたらすメリット）

本節では、最後として本研究の分析手法が示す見通しを簡単に述べておきたい。主として本研究が用いたクラスタ分析（広くは多変量解析の分析手法）がもたらす文法記述への利点について考えてみたい。

本研究では、動詞に基づく従来の枠組みを批判的關係から検討した上で、文脈の制約に基づく「に」の意味分析を試みた。また、分析に際して多変量解析の手法を用いることで「文脈」といった複雑な現象に対しても、（計算機を活用して）一貫した基準から定量的に分析することが可能であることを示した。だが、これまでの考察においては、「なぜ定量的分析の必要か」という疑問、さらには、「なぜ文系学問に数値的アルゴリズムを導入する必要があるのか」といった疑問に、必ずしも明確な答えは出していない。こうした問題点に関連して、本研究の分析手法そのものがもたらす利点としては、以下のものをあげることができる。

##### 分析の精緻化

- 言語現象の隠れた性質を比較的容易に見つけ出すことができる。
- 主観の関与をコントロールできる。
- 複雑で抽象度の低い大量の言語データを一貫した基準で分析・記述することができる。

##### 再分析可能性

- 主観の混入を後から特定できるので、議論の検証と再構築に役立つ。

以上の利点をよりグローバルな視点から評価してみた場合、定量的・統計学的アプローチは、言語分析の精度に関わる質的向上に役立つものと位置づけることができる。最終的には一解釈としての言語分析のレベルから事実としての言語分析へのレベルアップを指向するものであり、今後、その踏み台としての役割が多いに期待されると言えよう。

#### 参考文献

- Croft, William (1991) *Syntactic Categories and Grammatical Relations*. Chicago: The University of Chicago Press.
- Goldberg, Adele. E.(1995) *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- 国立国語研究所 (1997) 『日本語における表層格と深層格の対応関係』, 三省堂.
- Kohonen, T (1995) *Self-Organizing Maps*, Berlin: Springer Series in Information Sciences, Vol. 30.
- Langacker, Ronald W.(1987) *Foundations of Cognitive Grammar, Vol. 1, Theoretical Prerequisites*, Stanford, Calif.: Stanford University Press.
- Langacker, Ronald W. (1990) *Concept, Image, and Symbol. The Cognitive Basis of Grammar*, Cognitive Linguistics Research, 1, Berlin-New York: Mouton de Gruyter.
- Langacker, Ronald W.(2000) "A Dynamic Usage-Based Model." Michael Barlow and Suzanne Kemmer (eds.) *Usage Based Models of Language*. Stanford, Calif. : CSLI Publications, pp.1-63.
- Michaelis, Laura A.(2003) "Word Meaning, Sentence Meaning and Constructional Meaning." In H. Cuyckens, R. Dirven and J. Taylor, (eds.), *Cognitive Perspectives on Lexical Semantics*. Amsterdam: Mouton de Gruyter. pp.163-210.
- 村木新次郎 (1991) 『日本語動詞の諸相』, ひつじ書房.
- 仁田義雄 (1993) 「日本語の格を求めて」、仁田義雄編著 『日本語の格をめぐる』 くろしお出版. PP.1-37
- Romesburg, C. (1989) *Cluster Analysis for Researchers*. Florida: Robert E. Krieger Publishing Company.

Taylor, R. John. (2003) *Cognitive Grammar*. Oxford: Oxford University Press.

寺村秀雄 (1982) 『日本語のシンタクスと意味 1』, くろしお出版.

山梨正明 (1983) 「格文法理論」、安井稔編 『意味論英語学大系』, 大修館書店, PP.467-547.

山梨正明 (2000) 『認知言語学原理』, くろしお出版.

李在鎬 (2001) 「他動詞文のゆらぎ現象に関する「構文」的アプローチ」, 『言語科学論集』(京都大学) No.8, pp. 1-20.

李在鎬 (2003) 「複合的認知事象の制約に基づく結果構文再考: 構文現象の体系的記述に向けて」, 『認知言語学論考 No.3』, ひつじ書房.

#### データベース

- ・ 新潮社、「CD-ROM 版 新潮文庫の 100 冊」
- ・ NTT コミュニケーション科学基礎研究所、「日本語語彙大系 CD-ROM 版」, 岩波書店.

---

\* 本稿は 2003 年 第 4 回 日本認知言語学会における口頭発表に基づくものである。発表時において貴重なコメントをくださったフロアの方々および京都大学 山梨正明研究室の院生の方々にこの場を借り、感謝申し上げる。

<sup>i</sup> 紙幅の都合上、詳細は述べないが、日本語研究の観点からこの問題に関連する考察として李 (2004) を参照されたい。

<sup>ii</sup> この種の手法は、計算言語学において用いられるものであり、連合や共起を語彙の意味と定義する手法、すなわち潜在的意味解析法 (LSA; Latent Semantic Analysis) を基本に置いたものである。LSA に関する詳細は、Landauer & Dumais (1997) および <http://lsa.colorado.edu/> を参照されたい。

<sup>iii</sup> いずれのテキストも「CD-ROM 版 新潮文庫の 100 冊」のもので、詳細には、赤川次郎『女社長に乾杯!』/大江健三郎『死者の奢り・飼育』/村上春樹『世界の終わり』とハードボイルド・ワンダーランド』/椎名誠『新橋烏森口青春篇』/宮本輝『錦織』/筒井康隆『エディプスの恋人』/五木寛之『風に吹かれて』/司馬遼太郎『国盗り物語』となっている。また、データの公正性の問題に関連し、次の点を述べていく。本研究がテキストデータとして (会話や新聞といった他のコーパスを使わず) 小説を利用した背景には、今回の調査で用いた手法に関する問題が深く関与している。それは、今回用いた分析手法は文脈情報をデータとして利用したものであり、省略に (分析精度の面で) 弱い。このことから、助詞の省略が激しい会話コーパス、あるいは (見出しなどで) 漢語表現による独特の省略が行われる新聞コーパスは利用しなかった。なお、こうしたバイアスによる影響は別の手法を用いることで十分に克服できると考えられるものであり、本研究の本質的限界を示すものではないことに注意を喚起しておきたい。

<sup>iv</sup> FS1 の変数は、『CD-ROM 日本語語彙大系』に準拠している。

<sup>v</sup> 紙幅の都合上、詳細な解説は避けるが、このことは、認知言語学が提唱する記号としての言語観の経験的妥当性を実証的に示すものでもあることに注目してほしい。同時にこの見方は、日本語の文法を根底から見直す重要な視点でもあることを強調しておく。この点に関連した詳細な記述分析、さらに具体的研究成果は、Langacker (1987, 1990) および、山梨 (2000)、Michaelis (2003) 李 (2001, 2004) を参照されたい。

## **abstract**

### **Quantitative analysis of Japanese Particle “*NI*” - A New Approach to Descriptive Grammar -**

LEE, Che-ho (Graduate School of Kyoto University)

The purpose of this paper is to examine the usage of Japanese particle *NI* by restrictions of the context. In previous studies, *NI* is generally analyzed as a postpositional particle and acquires various meanings such as the goal of a motion, the result of the change, and the marker of an adverbial phrase. However, the relation between them is not so clear.

Two questions I try to answer in the paper are as the following: (1) how many usage of *NI* we find based on form-meaning pairs, (2) what kind of relation exists between each example of *NI* usage. In order to provide a solution, I would like to introduce a new research methodology.

Firstly, in order to observe the usage of *NI* I collect *NI* sentences as samples and create a corpus. This is the very first level to put this methodology into practice. Secondly, I analyze my data by conducting a quantification of the usage, based on a multi-variable analysis and also based on a theoretical/descriptive position of Construction Grammar. To give a more specific account, the following procedure is performed, a) all of the samples are collected from a corpus, b) all samples are identified as 1 and 0 by Distributed Representations, c) the samples are statistically analyzed using cluster analysis.

As a concluding remark, the analysis result suggests the following facts. 1) Three kinds of asymmetric restrictions are observed while clustering. They are the meaning of the coinciding noun, the syntactic relation and both coinciding noun and syntactic relation. All these three restrictions affect the clustering. 2) The usage of the postpositional particle *NI* is understood by restrictions coming from the context.