

スペイン語における前置詞後続名詞句の数・定性

—名詞句の現働化による 7 前置詞のクラスタリング—

喜多田 敏嵩

Toshitaka KITADA

1. はじめに

名詞句の数・定性は、スペイン語学において活発な議論が行われてきた分野であり、スペイン語と日本語の両方で多くの先行研究が存在する。しかし、その大半は、文における統語機能を果たす名詞に焦点を当てたものであり、前置詞句内をはじめとする、統語的要請を受けにくい位置にある名詞句の数・定性を体系的に記述した研究は、管見の及ぶ限り存在しない。本稿は、*de* および *de* と意味の重なりを有する 6 つの前置詞 *con, desde, en, para, por, sobre* の計 7 つについて、後続する普通名詞の数・定性を変数とした階層的クラスター分析を実施し、Fernández Ramírez (1986) と Bosque (1996) による前置詞句の数・定性に関する分類的記述の不透明性を指摘するものである。

2. 先行研究

2.1. Fernández Ramírez (1986)

管見の及ぶ限りにおいて、前置詞句における冠詞の生起に関する最も詳細な記述を行っているのが Fernández Ramírez (1986: 166-169) である¹。Fernández Ramírez は、明確な空間指示機能を有する前置詞は無冠詞単数形を後続させにくいとし、*tras, detrás de, junto a, encima, desde, debajo, hasta, sobre, dentro* の 9 つの前置詞（句）を挙げている。

2.2. Bosque (1996)

スペイン語の無冠詞名詞句に関する詳細な記述を行っている Bosque (1996: 53-55) は、無冠詞名詞句単数形が前置詞に後続する構造の容認度の低さに触れながら、実際に無冠詞名詞単数形が項となって形成される前置詞句の特性に関する記述を行っている。その中で、前述の Fernández Ramírez (1986) の分類基準を参照しながら、*con, de, en, por* の 4 つを、無冠詞名詞単数形を後続させやすい前置詞であるとしている。

2.3. 問題の所在

ここまで、本稿の分析に有用な 2 つの先行研究を挙げたが、これらの論考が、*en* をうまく分類しえない記述となっている点を問題として指摘したい。Fernández Ramírez (1986) は、明

確な空間指示性の有無により前置詞の分類を試みていたが、この性質が確固とした指標であるとは言いがたい。また、Bosque (1996) の挙げる 4 つの前置詞 *con, de, en, por*において、*en* は空間的意味を有しており、空間指示性の強弱を尺度にした分類において、*en* が弱者に分類されるというのは、スペイン語学習者としての経験的印象に反するものである。したがって、前置詞に被制語として後続する名詞句の数・定性を分類する尺度としての空間指示性の強弱は、その適性に疑問が残る。

本稿では、コーパス調査を通じて、Bosque (1996) の挙げる 4 つの前置詞が、後続名詞句の数・定性について同様の分布を見せるかを検証し、先行研究の挙げる分類尺度の不透明性を考察する。

3. 分析手法

3.1. 理論的枠組み

本稿では「現働化」という概念を援用する。「現働化」はフランス語学における術語 *actualisation* の訳語であり、横井 (1994) の簡潔な紹介によれば「*LANGUE* のレヴェルにある ‘lexique’ に貯蔵されている名詞を、発話行為によって *DISCOURS* に出現せしめる (127)」、言い換えれば、「言語記号としての名詞が現実世界の事物に対応し指示するようになる」現象を指す語である。名詞に付随することで現働化を促す形態素は、現働化詞と呼ばれ、スペイン語においては限定詞と複数形態素がそれにあたる。このように「現働化」は数と定性を「任意の名詞に付随することで、言語記号である名詞を現実の事物に対応させる働きをもつ形態素」として同範疇に属するものと解釈する概念であり、名詞句の形態に関する 2 変数である数と定性を、名詞に付隨する同一範疇の形態素として捉えることができる点で「現働化」を分析の枠組みに据えるのは有益であると考える。

3.2. 分析対象の限定

本稿で分析対象とするのは、普通名詞が現働化して前置詞に後続する語列であるが、まず各品詞の定義を決めておかねばならない。普通名詞については、前述のとおり、固有名詞ではない名詞の類とし、限定詞については、寺崎 (1998) の定義にしたがうことにする。寺崎 (1998: 75) によれば、スペイン語の限定詞には、冠詞、指示形容詞、所有形容詞、関係形容詞、疑問形容詞、数詞、不定形容詞の 7 つが該当するが、本稿では、4.1. で後述する理由から、関係形容詞と数詞を除いた 5 つを限定詞として分析する。

また、前章で挙げた先行研究では、意味の重なりが考慮されていない多数の前置詞が記述されている。そこで本稿では、López (1972) の記述に依拠しながら、スペイン語において最

頻出の前置詞 *de* と、*de* に関連のある前置詞のみを分析対象とすることにする。López(1972)は、スペイン語における前置詞の意味の対立・中和について図式などを用いた詳細な分析を行っており、その中で *de* との意味の中和を有し、可換な場合がある前置詞として、*con, desde, en, para, por, sobre* の 6 つを挙げている。なお、これら 6 つは、以下のような場合に *de* と置換可能であると述べられている²。

表 1. *de* と置換可能な例

日本語	<i>de</i> を使う場合	可換な例
雪に覆われた	cubierto <u>de</u> nieve	cubierto <u>por</u> la nieve
マドリードから来る	viene <u>de</u> Madrid	viene <u>desde</u> Madrid
1つのティーセット	un juego <u>de</u> té	un juego <u>para</u> té
赤く塗られている	pintado <u>de</u> rojo	pintado <u>en</u> rojo
人生について話す	hablar <u>de</u> la vida	hablar <u>sobre</u> la vida
グラス 1 杯の水	un vaso <u>de</u> agua	un vaso <u>con</u> agua

前述の Bosque (1996) が挙げる現動化を求めていく 4 つの前置詞 *con, de, en, por* は、これら 7 つの前置詞 *con, de, desde, en, para, por, sobre* (以下、「7 前置詞」と呼ぶ) に包含されている。Bosque (1996) の挙げていない 3 つの前置詞のうち、*desde, sobre* は Fernández Ramírez (1986) が空間指示性を持つものとして挙げており、記述のない *para* についても、その空間的意味に加えて、*pro/per + ad*、すなわち *pro* あるいは *per* に *ad* が付随して形成されたものであるとする通時の見解³に留意すれば、後続名詞句に現動化を求めるやすい *a* と類似した分布を見せることが予想される⁴。したがって、López(1972) の挙げる 7 前置詞は、後続名詞に現動化を求めていく *con, de, por* と、求めやすい *desde, para, sobre*、そして両者に含まれる可能性がある *en* により構成されていることから、本研究で行う分析に適した対象であると言える。

3.3. 分析の手順

本分析の手順は、以下のとおりである。まず、既存のスペイン語コーパスから、7 前置詞に後続する普通名詞の現動化に関する頻度、すなわち数・定性の頻度を収集する。その後、データの標準化を経て、各前置詞に関する変数を導出する。そして、求めた変数を用いて 7 前置詞に関する階層的クラスター分析を実施し、形成されたクラスターを考察する。なお、変数間の非類似度の算出にはユークリッド距離を、クラスター間の距離算出にはウォード法を使用する。また、分析には統計ソフト R を使用する。

3.4. 使用コーパスの詳細

本分析では、esTenTen 11 というコーパスを使用する。esTenTen 11 は、コーパス検索ツール Sketch Engine にて公開されているコーパスであり、2011 年に公開されていたスペイン語圏 19か国のウェブページをもとに構築されたスペイン語均衡ウェブコーパスである。総収録語数は約 95 億語となっており、管見の及ぶ限り、約 450 億語を収録する通時コーパス Google Books に次ぐ、最大規模のスペイン語コーパスである。スペイン語には大別して、スペインで使用されるヨーロッパスペイン語と、中南米で使用されるアメリカスペイン語があるが、このコーパスを分析に用いる理由は、スペイン・中南米両地域のデータを豊富に含んでいる点と、CQL (Corpus Query Language) を用いた柔軟な検索が可能である点による⁵。

表 2. esTenTen 11 の構成国内訳

順位	国名	比率
(1)	アルゼンチン	30.39 %
(2)	スペイン	21.35 %
(3)	メキシコ	17.94 %
(4)	チリ	10.63 %
(5)	コロンビア	4.52 %
(6)	ペルー	3.08 %
(7)	ベネズエラ	2.68 %
(8)	キューバ	2.55 %
(9)	ウルグアイ	1.92 %
(10)	その他	4.96 %
合計		100.02 % ⁶

4. 分析

4.1. データ収集

コーパスからデータとして収集したいのは、名詞句が現動化せずに 7 前置詞に後続する頻度であり、con NP, de NP, desde NP, en NP, para NP, por NP, sobre NP という語列を検索しなければならない。しかし、前述のとおり、スペイン語には後置修飾句・節が頻繁に生起するため、後続する語列に何らかの制限を設けないと、NP の現動化を、前置詞か NP のいずれかに求めることができなくなる可能性がある。したがって、抽出する語列を「前置詞+(限定詞)+普通名詞+ピリオド」という構造に限定して検索式を作成した。以下がその詳細である。

品詞タグ

普通名詞	普通名詞単数形	限定詞 ⁷	ピリオド
NC	NC.S	D	Fp

検索式 1 : [word="前置詞"] [tag="NC.*"] [tag="Fp"]

抽出例: con discapacidad. (障害を抱えて), para niños. (子供のために)

検索式 2 : [word="前置詞"] [tag="D.*"]? [tag="NC.S*"] [tag="Fp"]

抽出例: desde el exterior. (外から), sobre la salud. (健康に関して)

検索式 1 は、普通名詞が限定詞を伴わず 7 前置詞に後続し、直後にピリオドが来る語列を抽出する式であり、これをもとに高頻度上位 100 の普通名詞の総用例数に占める、単数形用例数の相対頻度 α を導出した。また、検索式 2 は、普通名詞単数形が限定詞を任意で伴って前置詞に後続したのちにピリオドが生起する語列を抽出する式であり、これにより、高頻度上位 100 の普通名詞単数形の総用例数に占める、限定詞を伴わない用例数の相対頻度 β を導出した。なお、普通名詞のタグ NC/ NC.S で抽出された語が以下に該当した場合は、ノイズとして除外した。

- A) 固有名詞: internet, iPhone など
- B) 代名詞: ti (前置詞格代名詞 2 人称単数形) など
- C) 副詞: atrás (後ろへ), mañana (明日) など
- D) 月を表す名詞: enero (1 月), febrero (2 月) など
- D) 規範的でない、あるいは正書法に即していない表現: tod@s (みんな), dia⁸ (日) など
- E) 複合名詞のため、形態上单複の区別がつかないもの: portaaviones (空母) など
- F) 名詞の語彙素性が数性を帶びているもの: tijeras (ハサミ), pantalones (ズボン) など
- G) 使用が特定の地域に集中している語: camote (サツマイモ), ómnibus (バス) など
- H) その他: URL など

以上の操作から導出された 2 変数 α , β と、 α , β を標準化した Z_α , Z_β は次のとおりである。

表 3. α , β および Z_α , Z_β の値⁹

	α	β	Z_α	Z_β
con	0.91	0.44	0.66	0.78
de	0.78	0.55	0.11	1.27
desde	0.77	0.06	0.03	-0.92
en	0.97	0.22	0.98	-0.20
para	0.28	0.02	-2.18	-1.09
por	0.94	0.55	0.81	1.24
sobre	0.67	0.02	-0.41	-1.08

4.2. 数性・定性の無相関検定

本稿では、7 前置詞をケース、 Z_α, Z_β を変数としたケースクラスター分析を実施するわけであるが、その前に 2 変数 Z_α, Z_β の母集団、すなわち前置詞に被制語として後続する名詞句の数性と定性における相関の有無を確認しておく。帰無仮説 H_0 を「母相関係数は 0 である」、対立仮説 H_1 を「母相関係数は 0 ではない」として t 検定を行うと、2 変数 Z_α, Z_β の相関係数 r は 0.61 となり、検定統計量 T は 1.72 となる。これは自由度 5 の t 分布にしたがうので、検定量 T は有意水準 5 % の臨界値 2.57 を下回る。よって、帰無仮説 H_0 は棄却されず、数・定性の相関は有意なものであるとは言えない。したがって、本分析では、これらの標本である Z_α, Z_β を独立した 2 変数であるとして、クラスタリングに使用することにする。

4.3. クラスター分析

以上をふまえ、統計ソフト R を使用して、7 前置詞をケース、 Z_α, Z_β を変数とする階層的ケースクラスター分析を実行した。ユークリッド距離にて算出したケース間の距離を示した非類似度表、2 変数を軸とした平面上に占める各ケースの座標を示した散布図と、ウォード法を用いたクラスタリングの樹形図は以下のとおりである。分析の結果、7 前置詞から {con, de, por} {desde, en, para, sobre} の 2 つのクラスターが形成された。

表 4. 非類似度表

	con	de	desde	en	para	por	sobre
con							
de	0.73						
desde	1.82	2.19					
en	1.03	1.70	1.20				
para	3.41	3.29	2.21	3.28			
por	0.48	0.70	2.30	1.45	3.79		
sobre	2.15	2.41	0.47	1.65	1.77	2.62	

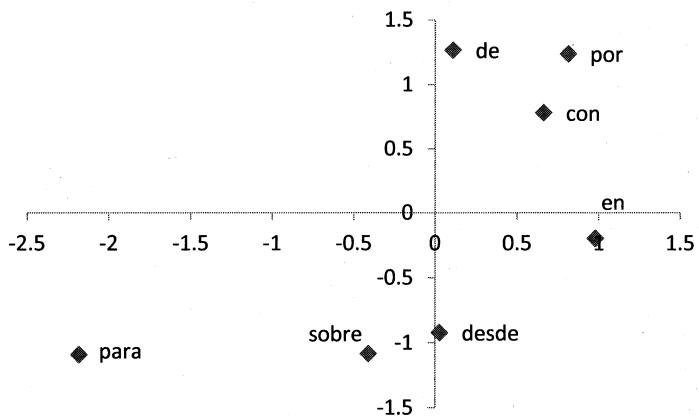


図 1. 散布図

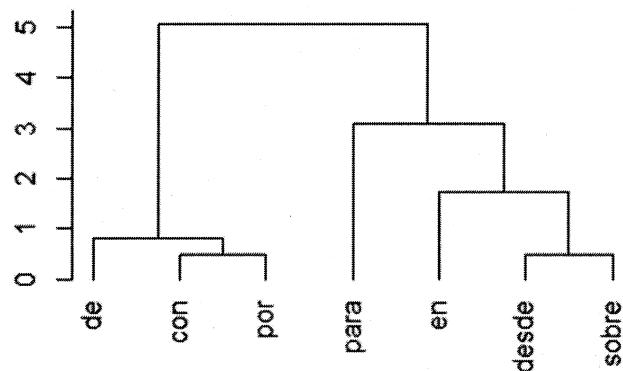


図 2. デンドログラム

4.4. 分析結果の考察

クラスタリングの結果、7 前置詞は $\{con, de, por\}$ $\{desde, en, para, sobre\}$ の 2 群に分類された。Bosque (1996) の挙げる 4 つの前置詞 con, de, en, por は、後続名詞句の数・定性について同様の分布を見せず、2.3. で指摘した問題点を裏付ける結果となった。

しかし、散布図における 7 前置詞の座標を確認してみると、 en が $\{con, de, por\}$ と $\{desde, para, sobre\}$ の中間に位置していることが分かる。デンドログラムでは、 en が早い段階で後者のクラスターに組み入れられているが、これは $para$ の孤立が関与しているものであり、 en と $\{con, de, por\}$ の分布上の類似性は否定できるものではない。2 つの検索式を用いたコーパス調査によれば、前置詞 en は、以下の例 (1) のように、現動化していない名詞句を後続させる頻度が高く、これは con, de, por と共にした特徴である。

(1) en curso, en equipo, en español, en televisión¹⁰

一方で *en* は (2) のように場所を表す普通名詞を、定冠詞を伴って後続させることも多い。これは *desde, para, sobre* と共にした特徴であり、*en* の空間記述性は無視できるものではない。

(2) en el país, en el mundo, en la ciudad, en la región

したがって、このような *en* 前置詞句の特徴を考慮すれば、7 前置詞に後続する名詞句の数・定性に関する分類指標として空間記述性は不十分であると言える。そこで、新たな指標を加えた複合的な記述が必要となるが、前置詞に関する意味的指標であった空間記述性のほかに、具体的に以下のような統語・形態的性質を分類基準として想定することができる。

統語的性質に関して *con, de, en, por* は、関係代名詞と共に起する際に定冠詞の付与が任意となる性質を共有している¹¹。たとえば、以下の例文 (3) では、下線部が *a la que* ではなく *a que* となっており、定冠詞単数女性形 *la* が省略されている。

(3) Esta es la persona a que me refería antes.

また、形態レベルでは、単音節性が共通の性質として挙げられる。3.1.において *para* の成り立ちを考察したが、*para* をはじめとする複音節の前置詞の中には、複数の語が組み合わさって現在の形となり、結果として、より具体的な意味を有しているものもある。たとえば、*desde* はラテン語の前置詞句 *de ex de* の縮約形であり (Corominas & Pascual, 1980-91)、複音節性と意味の具体性の相関も検討すべき尺度であると言える。

5. むすび

本稿では、これまでに体系的な分析が行われていない、スペイン語における前置詞句の数・定性について、現働化という概念に基づきながら、Fernández Ramírez (1986: 167) および Bosque (1996) が用いた空間指示機能の強弱が、分類の指標として適当であるかを、クラスター分析を通して考察した。その結果、7 前置詞 *con, de, desde, en, para, por, sobre* は *{con, de, por}* *{desde, en, para, sobre}* の 2 群に分類された。これにより、Bosque (1996) が後続名詞句に現働化を求めるにくい前置詞として挙げる 4 つの前置詞 *con, de, en, por* の中で、前置詞 *en* は、後続普通名詞に求める数・定性の頻度が他の 3 つよりも高く、その分布上の特異性が明らかになった。これにより、分類基準としての空間記述性が不十分であることが分かったほか、

本稿では、先行研究がとった意味的分類を補いする形態・統語的分類基準の提案も試みた。今後の研究では、前置詞句の数・定性を、形態・統語・意味の3つの視座から分析し、複合的な詳述を行っていくほか、とりわけ前置詞 *en* に後続する普通名詞の数・定性に関して、過去の論考とは異なった位置づけを提案していくことが肝要である。

最後に、本稿で明らかになった2点の研究課題を挙げる。1点目は、コロケーションとの親和性に着目した調査である。本稿のコーパス調査で得られた限定詞を伴わない前置詞句の中には、イディオムとして学習書などで見られるものが相当数存在した。ゆえに、コロケーションの生産性と限定詞を伴わない名詞句の後続頻度の関連性を考察していくことが求められる。2点目は、地域的有意差の検定である。本稿では、語彙以外に地域的差異の影響が表れることはないという立場をとったが、前置詞句の数・定性に関する体系的記述には、地域的有意差の有無を分析していくことも必要である。

註

- 1) Solís García (2011:111) 参照。
- 2) 表1はLópez (1972) の挙げる例を引用したものであるが、一部修正して掲載している。
- 3) *para*の形成に関しては、*pro + ad > pora > para*とする見解(RAE, 2014)と、*per + ad > (par + ad) pora > para*(Corominas & Pascual, 1980-91)とする見解があるが、いずれの形もaの要素(ad)を有している。本稿では、*para*の形成に関する議論に関しては態度を保留し、aを含んだ前置詞であるという見解の一致のみを援用する。
- 4) 前置詞aに後続する名詞句の数・定性はBosque (1996: 78-79)などで言及されている。
- 5) 詳細はKilgarriff & Renau (2013) や Kilgarriff *et al.* (2014) 参照。
- 6) 構成国内訳の比率総和は100.02%になるが、esTenTen 11のホームページに掲載されている情報に従って記載した。(https://www.sketchengine.eu/esteten-spanish-corpus/, 2019年3月15日確認)
- 7) コーパスの限定詞タグDには数詞と関係形容詞が含まれていない。しかし、数詞は他の限定詞との共起が唯一可能である点から、純粋な限定詞ではないと言えないため、分析から除外した。関係形容詞についても、直後にピリオドが来る語列には表れにくい語であるため、分析対象外とした。
- 8) *tod@s*は*todos*(全員)の男女両形を総称する非規範的な表現である。*dia*には、正書法上必要なアクセントが付されていない。
- 9) 数値は小数点第3位を四捨五入して表記している。

- 10) 現動化することなく前置詞 *en* に高頻度で後続する普通名詞は、*en general* (一般的に), *en particular* (とりわけ), *en desarrollo* (発展中の), *en efectivo* (現金で) のように、*en* と結びつくことでイディオムとして機能するものが多く、これらの連語表現を何らかの形で分析から排除した場合、複数形語尾による現動化の頻度も高まることが予想される。
- 11) RAE & ASALE (2009: 3302, §44.2e) 参照。本稿では扱っていないが、*a* も含まれる。

参考文献

欧文

- Bosque, I. (ed.) (1996): *El sustantivo sin determinación*, Visor Libros.
- Corominas, J. & Pascual, J. (1980-91): *Diccionario crítico etimológico castellano e hispánico*, Gredos.
- Fernández Ramírez, S. (1986): *Gramática española* (2.^ª ed.), vol.3.2: *El pronombre*, Arco/Libros.
- Kilgarriff, A. & Renau, I. (2013): “esTenTen, a vast web corpus of Peninsular and American Spanish”, *Procedia-Social and Behavioral Sciences*, 95: 12-19.
- Kilgarriff, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, & V. Suchomel (2014): “The Sketch Engine: ten years on”, *Lexicography*, 1: 7-36.
- López, M. (1972): *Problemas y métodos en el análisis de preposiciones*, Gredos.
- RAE (Real Academia Española) (2014): *Diccionario de la lengua española* (23.^ª ed.), <http://www.rae.es/rae.html> (2019年3月15日確認)
- RAE & ASALE (2009): *Nueva gramática de la lengua española*, Espasa.
- Solís García, I. (2011): *El concepto de referencia y su utilidad en la didáctica del español como lengua extranjera*, Tesis doctoral, Universidad de Oviedo.

和文

- 寺崎英樹 (1998)『スペイン語文法の構造』, 大学書林.
- 藤田健 (2011)「フランス語とスペイン語における不定冠詞の分布について」『北海道言語文化研究』, 9, 北海道言語研究会, 1-22.
- 横井雅明 (1994)「フランス語の冠詞と現動化レヴェル」, *Artes Liberales*, 55, 岩手大学人文社会科学部, 127-134.

使用コーパス

Sketch engine. esTenTen 11 (9 497 213 009 palabras):

<https://www.sketchengine.eu/>よりアクセス可能 (2019年3月15日確認)