

The semantics of words derived from proper names

Rafael Antonio Gutiérrez Martínez

Morphology Workshop
Université Paris Cité
Paris, November 8, 2024

The semantics of words derived from proper names

A probabilistic semantic account of word formation

A case study of Spanish words derived from proper names

Rafael Antonio Gutiérrez Martínez

Morphology Workshop
Université Paris Cité
Paris, November 8, 2024

Roadmap of the presentation

1. Introduction to the research topic
 - a. The problems of the derivation from proper names
2. Previous accounts
 - a. Semantics derivation from proper names
 - b. Semantics of derivation
3. A probabilistic approach
 - a. Toy model
4. Research objectives and what stage am I in?

1. Introduction to the research topic

The research project explores the semantics of words derived from proper names (WDPs) in Spanish.

These are words like *kantiano* (*Kant-iano*, ‘Kantian’) and *aristotélico* (*Aristóteles-ico*, ‘Aristotelian’) that are formed by adding the suffixes *-(i)ano* and *-ico* to the proper names *Kant* and *Aristóteles*, respectively.

The goal is to understand and to explain the process of meaning composition that takes place when the information carried by a proper name is linked to the information carried by an affix to create a new word that carries a new kind of information.

This process of meaning composition raises a number of questions and problems that this project aims to address.

1. Introduction to the research topic

Problem 1 (P1). What is the information that the proper name contributes to the derived word?

- It is agreed that the meaning of a complex word (e.g., *driver*) is in some way derived via an interplay of the meaning of the base (*drive*), the meaning of the morphology (*-er*) and contextual information.
- There is no consensus on what constitutes the content of a proper name.
- The most salient alternatives the *referentialist view* and the *descriptivist view* – do not provide the resources to form the meanings that the derived words can have.

Example: *kafkiano*: “said of a situation: that [it] is absurd or distressing”.

- Therefore, for words like *kafkiano* it is not clear what the semantic contribution of the base *Kafka* is to the word.

1. Introduction to the research topic

Problem 2 (P2). How to account for the tendencies in the semantics of derivatives?

- Depending on the general class of the referent of the proper name, some derivatives tend to be formed and some meanings tend to be derived.
- In Gutiérrez (2023), after studying 277 derived words of 40 proper names of different referent types (place, brand, artist, etc.), I observed that names of artists (e.g., *Franz Kafka*, *Julio Verne*) tend to give rise to quality-denoting expressions (e.g., *kafkiano*, *kafkaina*, *vernesiano*, *juliovernesco*), and avoid the creation of words denoting intellectual movements (words ending with *-ismo*).
- But this is a matter of tendency, and not something categorical. In the same study, I found one word derived from the name of an artist that referred to an intellectual movement: *gongorismo*.

1. Introduction to the research topic

Problem 3 (P3). How to develop a theory that allows us to make predictions about the interpretation of WDPs?

- Take *catalanismo*.
- What is the meaning of this form?
- The morphology provides some hints about its possible interpretations (at least 13).
 - (i) a doctrine, theory, system (*idealismo*, ‘idealism’),
 - (ii) a word or linguistic expression (*italianismo*, ‘Italianism’),
 - (iii) a disease, symptom, illness or some of its effects (*reumatismo*, ‘Rheumatism), or
 - (vi) a practice, attitude or habit (*machismo*, ‘sexism’)
- But which of these interpretations is (or are) likely or unlikely is not clear *a priori*.

1. Introduction to the research topic

Problem 4 (P4). How to account for the composition of two different sorts of meaning?

- Words derived from proper names constitute an instantiation of the situation in which two different sorts of information (the relatively discrete information from the morphology and the non-discrete information from the base) are combined and the challenge is to account for the composition of these different sorts of meaning.
 - ◆ The **discrete part** corresponds to the information more commonly associated with grammar: information about the participants in an event; about temporal features of a situation, and if it involves a change, a movement, locations, causal relations, etc.; and about the ontological types of entities.
 - ◆ The **non-discrete part** corresponds to encyclopaedic world knowledge.

2. Previous accounts 1

1. Semantics of derivatives from proper names

a. Descriptive contributions (in Spanish)

Díaz Rojo (2001), Lisyová (2005), Martín Camacho (2021)

b. Semantic contributions

Hèois (2020), Huguin (2018), Mauri and Masini (2024), Matushansky (2015), Lee (2020), Arsenijević, et al. (2014)

2. Semantics of derivation

a. Formal accounts

Dowty (1979)

b. Frame Semantics

Fillmore (2006), Plag et al (2018), Schneider (2023)

c. Relational Morphology and Construction Morphology

Jackendoff and Audring (2019, 2020)

d. Analogical modelling

Plag et al., (2023)

e. Lexical Semantic Framework

Lieber (2016)

2. Previous accounts 2

Lee (2020) and Arsenijević et al. (2014)

They propose to deal with the semantics of some types of WDPs by appealing to a relation variable R .

Lee proposes the translation for the word *Kripkean* in (1), which is the result of combining the translations of *-ian* and *Kripke*.

(1) *Kripke* [R *-ian*]: $\lambda y.Rc(y, \mathbf{Kripke})$

(2) *-ian*: $\lambda R\lambda x\lambda y.Rc(y,x)$

(3) *Kripke*: **Kripke**

-ian takes a relation R and an individual x , and produces the property of being in a contextually valued relation R_c to x (e.g. *Kripke*).

This type of account predicts that **any** relation is a good candidate for the meaning of the derivative. However,

1. Not all possible relations are good candidates. Consider the noun *rafaelista* (*Rafael-ist*, ‘Raphaelist’) from Raphael and assign it the translation $\lambda y.Rc(y, \mathbf{Raphael})$, where y is human. Being Raphael’s neighbor could in principle be a possible value for Rc . This is neither attested, nor does it seem like a valid value.
2. This does not accounts for partial productive patterns found in word formation.

2. Previous accounts 3

Matushansky (2015)

Matushansky proposes that the meaning of some WDPs is derived from the referential reading of proper names.

- The creation uses of proper names (e.g., *A Joan Miró is more durable than a Botero*) are the result of the morphological modification by a null suffix that targets entities of type *e*.

This proposal is motivated by the idea that the semantics of these words involves a reference to a definite, existing and identifiable name bearer. Is it possible to apply this sort of analysis of the content of proper names to other types of WDPs? I don't think it is.

1. The affixes that attach to proper names also attach to common nouns and the typical semantics given for the affixes are formulated to combine with common noun denotations, rather than proper noun denotations. How do we make that sort of analysis applicable to a different sort of denotation?
2. If proper names just contribute their referents, how are we going to systematically solve for *R*? It is the specific information about the referent what explains the patterns found in word formation. It is that information that has to be involved in the derivation.

2. Previous accounts 4

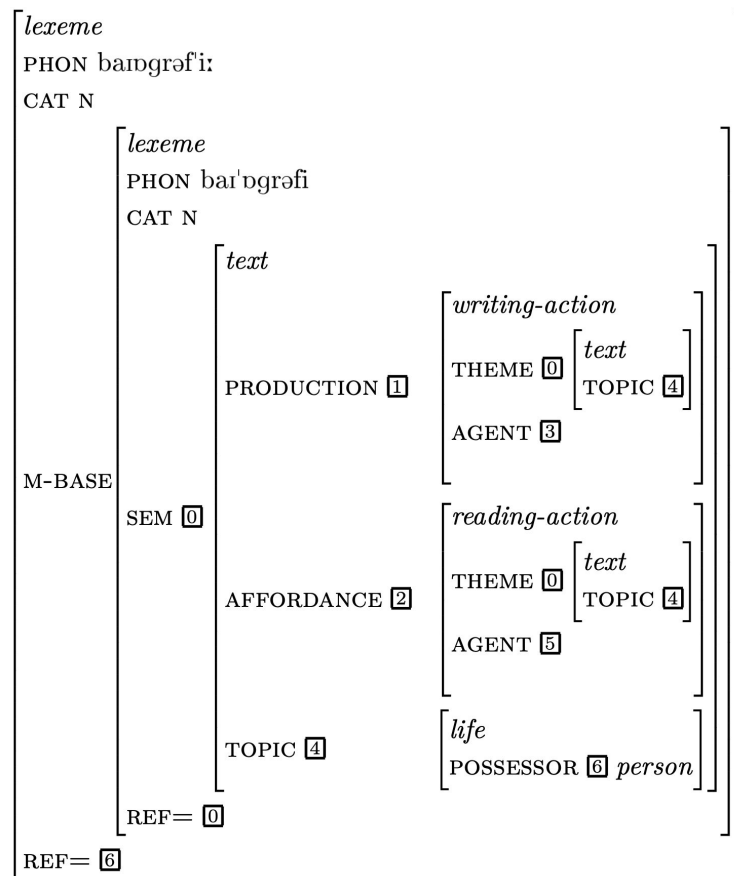
Schneider (2023)

In frame semantics, the semantic content of linguistic expressions is modeled through the use frames.

Frames are formalized as typed recursive attribute-value structures where attributes describe a concept and each attribute takes a specific value.

Schneider analyses the semantic representation of a base word and its derivative using these sorts of structures to account for why it is possible for some nominalizations in English (*biographee*) to have an event-related meaning, even though the word from which they derive (*biography*) has no event-related meaning.

However, this sort of approach do not do well with P1.



3. A probabilistic approach

Probabilistic semantics

Problems 2 and 3 suggest that the interpretation WDPs depends (at least in part) on the interplay between two sources of information: world knowledge and linguistic knowledge.

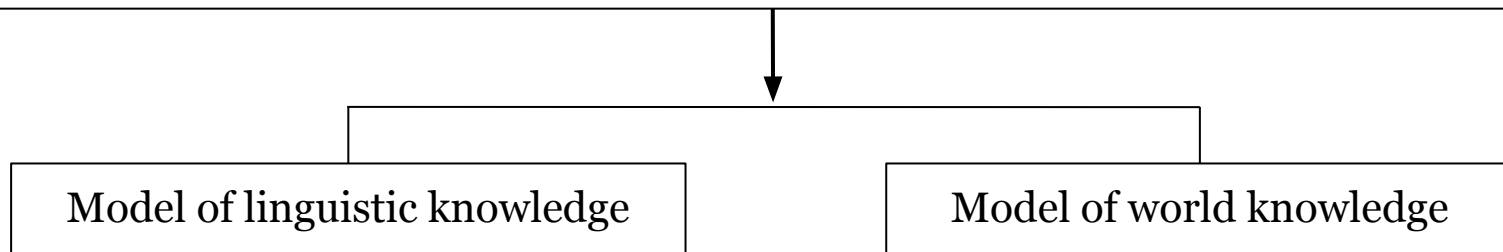
Probabilistic semantics are semantics that are design to deal with the interpretation of words (and other expressions) under different sources of uncertainty (Sutton (2015) and Erk and Herbelot (2024)).

Therefore, they seem ideal to treat the semantics of WDPs.

3. A probabilistic approach

The proposal

The interpretation of WDPs can be understood as a distribution over hypothesis concerning their meaning that depends on the interaction of the knowledge of information about the referent of the proper name and the knowledge about the semantic contribution of morphology.



Example: What are the possible readings of the words *gaudisismo*?

3. A probabilistic approach (a toy model)

Linguistic knowledge

I (provisionally) understand the **knowledge about the semantic contribution of morphology** in terms of the information provided by a descriptive grammar. According to the *Nueva gramática de la lengua española*, the set of possible interpretations of words derived from *-ismo* is reasonably well understood. It includes 13 different readings.

- (i) a doctrine, theory, system,
- (ii) attachment to something,
- (iii) a period in which a movement of thought is in effect,
- (iv) a quality or condition,
- (v) an artistic style or genre,
- (vi) a disease, symptom, illness or some of its effects,
- (vii) sport activities,
- (viii) a practice, attitude or habit,
- (ix) a word or linguistic expression,
- (x) a propensity or inclination to something,
- (xi) fans sports teams or clubs,
- (xii) a group of persons who are members of a movement, and
- (xiii) a fact or act

3. A probabilistic approach (a toy model)

Concept type	Morpheme	Examples
Intellectual movements	-ismo	<i>kirchnerismo, confucianismo</i>
Referent of the proper name	-ista, extra-	<i>el proyecto kirchnerista, extra-OTAN</i>
Quality	-mente, -esco, -dad, sub-	<i>ortegueanamente, imaginación juliovernesca, un momento de italianidad,</i>
Actions	-ada, -izar, des-	<i>kafkaneada, kirchnerizar, gongorizar</i>
Individuals	-ito, -ista, -ero	<i>kirchnerito, hitlerito, gongorista,, bolivianero</i>
Period of time	pre-, pos(t)-, proto-	<i>pre-kirschner, postkafkiano, protobizantino</i>
Historical events	-azo, -ista	<i>febreroazo, decembrista</i>
Appreciation	-ete, -ito	<i>Kirchnerete, Moisesito</i>
Word or expressions	-ismo	<i>bolivianismo, italianismo</i>
Opposition	anti-	<i>anti-kirchner</i>
Support	pro-	<i>pro-psoe</i>

World knowledge

I (provisionally) understand the **knowledge of information about the referent of the proper name** as the set of concepts that appear in the derivation from proper names.

In Gutiérrez (2023), I observed that there are 11 types of concepts that are introduced in the derivation from proper names by means of different affixes.

Not all of these concepts distribute equally in the meanings of the derivatives of different types of proper names. For example, **names of artists** introduce only doctrines, qualities, periods of time, actions, the referent, individuals, and oppositions.

3. A probabilistic approach (a toy model)

Example: *gaudisismo*

The interpretation could be understood as the result of the interaction of **(i)** the probability of a given type of concept of being associated with the name that forms the base and **(ii)** the probability of the candidate readings being associated with the derivative.

Since, in principle, all of the readings that the grammar mentions should be possible for *gaudisismo*, I assign equal probabilities to all the readings.

Using the information provided in Gutiérrez (2023), I assign different probabilities to each of the concepts that constitute our knowledge of types of referents. For instance, to artists, I assign probabilities of 0.2, 0.4, and 0.1 of being associated with doctrines, qualities, and periods of time, respectively.

What are the possible readings of the words *gaudisismo*?

Probability distribution of the possible readings of words ending with *-ismo*

Doctrine: 0.076
Quality: 0.076
Period of time: 0.076
Linguistic expression: 0.076
Action: 0.076
Attachment: 0.076
Artistic Style: 0.076
Disease: 0.076
Sport activities: 0.076
Habit: 0.076
Inclination: 0.076
Fans: 0.076
Members: 0.076

×

Probability distribution of the association of concepts with artists

Quality: 0.4
Doctrine: 0.2
Period of time: 0.1
Action: 0.1
Referent: 0.1
Individuals: 0.05
Opposition: 0.05
Linguistic expressions: 0.0
Historical events: 0.0
Appreciations: 0.0
Support: 0.0

3. A probabilistic approach: The result

1. The two most probable interpretations of *gaudisismo* are **quality** and **doctrine**.

2. *Gaudisismo* will never have interpretations of **disease**, **inclination**, and **members** because, although these interpretations are part of the list of readings of *-ismo*, they are not part of the list of concepts associated with Gaudí.

3. *Gaudisismo* will never have as interpretations **Referent**, **Individuals** and **Opposition**, since, even though these concepts are part of the concepts associated with proper names, they are not part of the interpretations of *-ismo*.

Probability distribution of the possible readings of words ending with *-ismo*

Doctrine: 0.076
 Quality: 0.076
 Period of time: 0.076
 Linguistic expression: 0.076
 Action: 0.076
 Attachment: 0.076
 Artistic Style: 0.076
Disease: 0.076
 Sport activities: 0.076
 Habit: 0.076
Inclination: 0.076
 Fans: 0.076
Members: 0.076

×

Probability distribution of the association of concepts with artists

Quality: 0.4
Doctrine: 0.2
Period of time: 0.1
 Action: 0.1
Referent: 0.1
Individuals: 0.05
Opposition: 0.05
 Linguistic expressions: 0.0
 Historical events: 0.0
 Appreciations: 0.0
 Support: 0.0

=

Hypothesis about the meaning *gaudisismo*

Quality: 0.034
Doctrine: 0.0152
Period of time: 0.0076
 Action: 0.0076
Referent: 0.0076
Individuals: 0.0038
Opposition: 0.0038
 Linguistic expressions: 0.0
 Historical events: 0.0
 Appreciations: 0.0
 Support: 0.0

Table 1. Assignment of probabilities to *-ismo* and artist names, and hypothesis about the meaning of *gaudisismo*

4. Research objectives

- **Objective 1:** Describe the morphological and semantic features of words derived from proper names.
- **Objective 2:** Develop a probabilistic semantic account that explains and captures the semantic and morphological properties and patterns identified.
- **Objective 3:** Test and evaluate the proposed probabilistic semantic account in terms of its predictive power, its relation to alternative theories and its consequences.

4. Research objectives

- ➔ **Objective 1:** Describe the morphological and semantic features of words derived from proper names.
 - ◆ I have created a dataset of 1,271 WDPs from 140 proper names of 14 different categories using two different corpora: CORPES XXI (RAE, 2024) and Corpus del Español (Davies, 2016).

References

Arsenijević, Boban; Gemma Boleda; Berit Gehrke; and Louise McNally (2014) “Ethnic adjectives are proper adjectives”. R. Baglini, et al. (eds.) *Proceedings from the 46th Annual Meeting of the Chicago Linguistic Society*, 17-30.

Davies, Mark (2016). Corpus del Español: Web/Dialects. Available online at <http://www.corpusdelespanol.org/web-dial/>.

Díaz Rojo, José Antonio (2001) "Nociones de neología. La formación de derivados y compuestos a partir de nombres propios de personas". *Panace@* 2 (5). 25-30.

Dowty, David (1979) *Word Meaning and Montague Grammar. The semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. London: D. Reidel Publishing Company.

Erk, Katrin and Herbelot, Aurélie (2024) “How to Marry a Star: Probabilistic Constraints for Meaning in Context”. *Journal of Semantics*, 1-35.

Gutiérrez Martínez, Rafael Antonio (2023) Derivational morphology in proper names: scenario knowledge and the paradigmatic lexicon. <http://hdl.handle.net/10230/58002>

Héois, Aurélie (2020) “When Proper Names Become Verbs: A Semantic Perspective”. *Lexis* [Online], 16. 1-35

References

- Huguin, Mathilde (2018) “Anthroponyms and Paradigmatic Derivation in French”. *Lingue e Linguaggio* 17: 217 – 232.
- Jackendoff, Ray and Jenny Audring (2019) “Relational Morphology and the Parallel Architecture”. *The Oxford Handbook of Morphological Theory*. 390 - 408.
- Jackendoff, Ray and Jenny Audring (2020) “Relational Morphology: A Cousin of Construction Grammar”. *Frontiers in Psychology*. Vol. 11. Art. 2241
- Lee, J. (2020). “On the asymmetry between names and count nouns: syntactic arguments against predicativism”. *Linguistics and Philosophy* (2020) 43:277–301
- Lisyová, Olga (2005) “La derivación adjetival a partir de los nombres propios”. *Moenia* 11. 249-268.
- Martín Camacho, José Carlos (2021) “El nombre propio en la creación de palabras. En torno a la eponimia”. *Archivum*, LXXI, 71. 245-277
- Mauri, Caterina and Francesca Masini (2024) “Multi-layered indexicality: When proper names become categories”. *Indexicality The Role of Indexing in Language Structure and Language Change*. Berlin: De Gruyter.

References

- Matushansky, O. (2015). The Other Francis Bacon: On Non-BARE Proper Names. *Erkenntnis*, 80(S2), 335–362
- Schneider, Viktoria (2023) “Eventualities in the semantics of denominal nominalizations”. In: *The Semantics of Derivational Morphology: Theory, Methods and Evidence*. Berlin: De Gruyter. 37-62.