

# **On the edge between Nouns and Verbs. The heterogeneous behavior of Spanish deverbal nominalizations empirically verified.**

## **Abstract**

This study investigates the intricate semantic and morphosyntactic behavior of morphological nominalizations. Besides offering an in-depth quantitative and qualitative description of a set of Spanish nominals, it empirically verifies to what extent the more “nouny” items more frequently adopt nominal features as opposed to the more “verby” items. The starting point is the tripartite semantic distinction between referential, event and state nominals, whose morphosyntactic behavior is then described on the basis of a detailed corpus analysis. The following properties are analyzed: determination, modification, number, and argument structure. The corpus study makes it possible to verify a series of theoretical assumptions on the processes of deverbalization and substantivization, and in particular the applicability of the hierarchical clines proposed within a typological perspective. The article also explores the possibility of creating an integrated transcategorization cline, and therefore combines all morphological features – nominal as well as verbal ones – into a multifactorial (classification tree) model.

## **Keywords**

morphological nominalizations; Spanish; transcategorization; usage-based; classification tree method.

## **1. Introduction**

Nominalizations, and particularly deverbal nominalizations, constitute a challenging case in point for categorization theories. Not only are they involved in a complex process of transcategorization implying a loss of verbal properties and acquisition of nominal properties, the outcome of this process may be quite diverse and entails intracategorical gradience.

First, nominalizations are said to involve both decategorization and recategorization (Croft, 1991; Malchukov, 2004, 2006). The former implies that when verbs and nouns are not used in their primary functions, they lose some of the morphosyntactic characteristics associated with this function. The latter then logically means that when a linguistic entity is used in an extended function, it acquires properties of the newly associated category. Nominalization thus equals deverbalization and substantivization. From a conceptual point of view, these transcategorial operations are not unexpected given that both word classes, verbs and nouns, are composed by the same set of features, albeit in an inverse order of dominance: V (predicative > referential function), N (referential > predicative function) (Anderson, 1997; Croft, 2001). Moreover, due to a clash between their entity reading (i.e. their referential function) and event reading (i.e. their predicative function), nominalizations entail a complex conceptual process (Dam-Jensen, 2008; Stekauer, 1998).

In Spanish, these complex mapping relations between the entity and the situation meaning can be illustrated by the morphosyntactic and functional differences that appear within and between the classes of derived nominalizations (*la adquisición de la tecnología* – *the acquisition of the technology*), infinitival nominalizations (*(el) comer es una fiesta* – *to eat is a party*) and complementizer phrases (*me extrañaba (el) que no hubiera habido especulaciones*

– *it seemed strange to me (the fact) that there had been no speculations*) (Dam-Jensen 2008). This complexity is even increased in English, where a fourth class of gerundive nominalizations must be added, that can be subdivided into nominal (*There is no general monopoly on the giving of legal advice*) and verbal (*I remembered him using the word malfunction*) gerunds (Heyvaert 2008; Fonteyn, Heyvaert & Maekelberghe 2015).

From a distributional perspective, the process of shifting a verb into a noun through the addition of a suffix comprises both the gradual loss of verbal features and the acquisition of nominal ones. To wit, a deverbal noun loses the ability to assign syntactic functions to other constituents in the clause, to mark voice morphologically and to express tense, aspect and mood distinctions, but gains the aptitude to pluralize and to combine with adjectival modifiers, possessive markers and other determiners (Comrie & Thompson, 1985; Givón, 2001b; Hopper & Thompson, 1984, 1985; Koptjevskaja-Tamm, 1993; Mackenzie, 1985).

Second, the shifting processes are argued to be of a gradual nature. Both from a functional and formal point of view the outcome can be quite diverse, causing a high degree of internal variation within the nominalization category. A nominalization indeed involves “conceptual reification” (Hartmann 2014: 158) which is, however, a matter of degree: the interpretation of a nominalization can be perceived as static or dynamic, and it can highlight the symptoms of both identities (i.e. entity or situation). This idea of semantic diversity is rendered by the classical distinction between result nominals (e.g. *the translation of this book is of good quality*), event nominals (e.g. *the translation of this book took me two hours*) (Dam-Jensen, 2008; Picallo, 1999), and more recently, the category of state nominals (e.g. *his interest in this problem is very strong*) (cf. Section 2). This semantic diversity is said to be mirrored by variable distributional patterns: some members display increased nominal properties (e.g. determiners, pluralization), whereas others preserve verbal characteristics, such as an overt argument realization.

What is highly interesting in this respect, and will also be at the core of this study, is the idea that categorial features leak, but “in an orderly, hierarchical fashion” (Anderson 1997: 73). Or, to put it differently, there is an ordering of the properties that are lost (decategorization or deverbalization) and acquired (recategorization or substantivization) during nominalization. In order to account for the order of these processes, several hierarchies have been proposed. One of the first detailed attempts was the *desentailization scale* elaborated by Lehmann (1988), yet the most influential theory has been proposed by Malchukov (2004). His cline of deverbalization (Figure 1) describes the loss of verbal features according to the principle that the more external a layer to the verbal nucleus is, the more easily it will be lost.

[[[[[V]VAL]TAM]AGR]IF]
------------------------

**Figure 1.** Cline of deverbalization (Malchukov 2004: 27)

Specifically, a nominalization first loses the ability to express Illocutionary Force (IF), before Subject Agreement is rejected (AGR). Then, it loses the markers of Tense, Aspect and Mode (TAM) and, finally, also the features of Valency and Voice (VAL). The acquisition of nominal features is described through the cline of substantivization represented in Figure 2. The more external a layer to the nominal nucleus, the more easily it will be acquired in the nominalization process.

[[[[[N]CL]NB]POS]DET]Case]
----------------------------

**Figure 2.** Cline of substantivization (Malchukov 2004: 47)

In a first phase, the nominalized form thus allows Case markers (Case). Then, it can combine with both the Definite article and Demonstrative determiners (DET), and only after this does it allow Possessive determiners and other Genitive markers (POS). Finally, it seems difficult to encounter examples that present the feature of Number (NB), and thus pluralize and combine with numerals, without manifesting certain markers that Malchukov (2004) considers typical of the nominal Class (CL), such as adjectival modifiers. He therefore assumes that the Number and Class layers are acquired simultaneously.

The result of these gradual shifts is a heterogeneous category: some nominalizations acquire more prototypically nominal features, whereas others preserve more features of the verbal prototype. This comes to say that the set of nominalizations displays a kind of categorial shading from a central, more “nouny” core to a more peripheral “verby” boundary. The more nominal properties a construction acquires, the more it tends to refer to a concrete entity and thus approach the semantic prototype of a noun, and vice versa (Comrie & Thompson, 1985; Dik, 1985; Givón, 2001b; Hartmann, 2014; Hopper & Thompson, 1985; Koptjevskaja-Tamm, 1993; Mackenzie, 1985). This proposal somehow relates to the model of Croft (1991, 2001), who assigns to both nouns and verbs a prototypical semantic class, as well as a pragmatic function. Concretely, a noun prototypically refers to an object, whereas a verb tends to predicate events. However, both categories possess a number of peripheral members, that do not meet these descriptions. Moreover, a clear relationship can be established between the degree of prototypicality of a given form and its morphosyntactic behavior, as only the most representative cases display the entire range of features that can be assigned to the category<sup>1</sup>.

The clines of deverbalization and substantivization have been developed and applied mainly within the field of typological studies to demonstrate their cross-linguistic validity. However, to the extent of our knowledge, they have not yet been empirically tested on a language-specific group of deverbal nominalizations. This lack of an intensive corpus-based empirical validation (Bekaert & Enghels, 2014, 2015) is an imperfection that many studies on morphological nominalizations suffer from. Although this is understandable, as it is often a side effect of differently oriented research questions, we nevertheless set ourselves four tasks to remedy this imperfection:

- to offer an in-depth quantitative and qualitative usage-based analysis of a set of frequently used and highly polysemous Spanish nominalizations;
- to empirically verify to what extent the more semantically defined “nouny” items indeed adopt more nominal behavior and the “verby” items more verbal behavior;
- to examine the above-mentioned theoretical assumptions on the processes of de- and recategorization, and more particularly to verify the applicability of both clines presented

---

<sup>1</sup> On the contrary, from a more generative-style framework, Baker (2003) argues that a prototypical and gradient perspective on word classes does not make strong predictions on their morphosyntactic behavior. However, this article shows that a prototypical behavior can be defined for each category, and that it consists of a number of features which must not be displayed by all members of the category.

in Figures 1 and 2 to the Spanish nominalization process. Can these clines be empirically (and even statistically) confirmed, or is there any interaction with language-specific structural properties, as Malchukov (2006) himself suggests?

- to explore the possibility of an integrated cline, that combines the decategorization and the recategorization process in one model.

The structure of this paper is as follows. Section 2 further elaborates on the idea of intracategorical heterogeneity of the nominalization category, and discusses the classical distinction between event and result nominals. It is argued that from a semantic point of view, three types need to be distinguished, namely the event, state and referential nominals. Section 3 outlines the methodology applied in this paper. The results of the analysis of the morphosyntactic features of a set of eight nominalized forms are presented in Section 4. This part successively examines their nominal (determination, pluralization, and modification) and verbal features (argument structure). It also aims at offering an integrated transcategorization cline for the Spanish nominalization process by statistically verifying the internal order of the assumed clines. The conclusions are presented in Section 5.

## 2. Intracategorical heterogeneity: Event, Result and State nominals

In order to meet the idea of intracategorical heterogeneity of the nominalization class, a first distinction can be made between event nominals (EvNom) and result or referential nominals (RefNom)<sup>2</sup> (Alexiadou, 2001; Grimshaw, 1990; Picallo, 1999). An event nominal denotes an event taken as an entity and is therefore “frozen” or “abstracted” from its situational properties (i.e. time and modality). As can be verified in (1a), the nominalized form *building* displays some nominal features (e.g. the presence of a determiner), but preserves verbal properties, such as an overt argument structure (*of the bridge* as its created object, and *by the soldiers* as the agent), and the expression of aspectual content (namely the trait [+ duration], as testified by the verbal form *took*). The referential nominal (1b) takes the recategorization process one step further as it refers to a concrete entity. It allows for more nominal features to appear, such as a wider range of determiners (e.g. the indefinite article), while it rejects most verbal properties (e.g. *He lives in an enormous building* \**by the soldiers*).

- (1) a. The building of the bridge by the British soldiers took 5 weeks.  
b. He lives in an enormous building.

However, this strict dichotomy between event nominals and referential nominals has been questioned. So, more recently a third type has been described, viz. state nominals (StateNom), as illustrated in (2a) (Barque, Fábregas, & Marín, 2011; Fábregas & Marín, 2012a; Fradin, 2011). On the one hand, these nominals exhibit the same pattern as event nominals, because they denote a situation with a certain temporal extension (as attested by the verbal form *lasted*). On the other hand, they lack dynamicity compared to event nominals. For instance, it would be

---

<sup>2</sup> The traditional term Result Nominal has been shown to be inadequate, because those nominalizations that have abandoned their eventuality denotation and thus approximate the nominal prototype do not necessarily refer to the result of the corresponding verb. In fact, they can refer to all participants more or less tightly related to the eventuality described by the verb (Bisetto & Melloni, 2005; Scott, 2010; Melloni, 2011; Fábregas & Marín, 2012a). Following Melloni (2011), we will therefore adopt the more neutral term Referential Nominal. This category contains the following semantic subsets: product/result (e.g. *construction*), means (e.g. *connection*), psychological stimulus (e.g. *attraction*), path (e.g. *prolongation*), agentive-collective (e.g. *administration*), and locative (e.g. *entry*).

impossible to replace *lasted* in (2a) by the dynamic verb *took*. In this property, state nominals approximate the nominal prototype and thus behave like referential nominals. They are therefore said to occupy an intermediate position between the event and the referential nominals, which is confirmed by various facts. For instance, while they inherit the argument structure of their corresponding verb (e.g. (2a) relates to *John is preoccupied with the economy*), they reject pluralization (2b) (Fábregas & Marín, 2012a).

- (2) a. John's preoccupation with the economy lasted/\*took several years.  
b. \*John's preoccupations with the economy lasted several years.

The particular behavior of state nominals can easily be explained by the status of their corresponding verbs. It is commonly accepted that state verbs are less prototypical instances of the verbal category and thus also reject some typical verbal features, and even adopt certain nominal properties (Hopper & Thompson, 1984, 1985; Croft, 1991, 2001).

The three-part semantic classification and its mapping onto the nominal/verbal continuum serves as a starting point for this study. It is hypothesized that the more “nouny” nominalizations (i.e. referential nominals) exhibit more nominal features, whereas the nominalizations that keep on referring to a situation (i.e. event nominals and state nominals) preserve more verbal features, albeit in varying degrees. This hypothesis is tested throughout the case study of the morphosyntactic behavior of Spanish nominalizations. It will be interesting to examine not only the absence or presence of particular morphosyntactic features, but also what the specific behavior of the semantically defined types of deverbal nominals reveals about the typological clines. The corpus and methods that are applied in this analysis are presented in the section below.

### 3. Materials and methods

This study looks into the functional and structural behavior of a set of frequently used Spanish nominalizations. However, the selection of a representative sample was not a straightforward process. Three different criteria have been taken into account. First, it is based on a list of frequently cited items in the literature on Spanish nominalizations (Picallo, 1991, 1999; Fábregas & Marín, 2012a, b; Jaque Hidalgo, 2012a, b). Second, only forms that are sufficiently attested in the corpora (in concrete, the *Corpus de la Real Academia Española* (CREA)) are included in the sample. For each item, the number of tokens has been analyzed, and the items that did not give sufficient examples have been removed from the initial list. This was the case, for instance, for the form *reparación* [reparation]. Finally, the forms and their different uses are necessarily equally distributed over the three semantic types of deverbal nominals. This selection process resulted in the following set of nominalizations: *creación* [creation], *construcción* [construction], *traducción* [translation], *destrucción* [destruction], *preocupación* [preoccupation], *interés* [interest], *vinculación* [linking], and *interrupción* [interruption]. Moreover, as is shown by Table 1, these items are highly polysemous as their tokens spread over two or three possible interpretations. This shows that the intracategorical heterogeneity of the nominalization class cannot be limited to the difference between various forms, but that even one form can exhibit different semantic interpretations. The lexeme *creación*, for instance, can refer to the event of creating itself (3a), or to its result (3b); similarly, *vinculación* can

denote the act of linking up two or more things (4a), the result of the linking event (4b), but also the static relationship between two entities (4c)<sup>3</sup>.

- (3) a. La creación de una lista de reproducción es un proceso trivial. [The *creation* of a playlist is not a trivial process.] (*EvNom*)<sup>4</sup>  
 b. Y cuando el genial artista malagueño contempló aquella maravillosa creación, dijo: "No hemos inventado nada". [And when the genius artista of Málaga contemplated that wonderful *creation*...] (*RefNom*)
- (4) a. La vinculación de la territorialidad humana con la animal ha sido habitual [...]. [The *connection* of the human territoriality with the animal one has been usual...] (*EvNom*)  
 b. Trasladando estos conceptos a la pintura, no tardan en detectarse estrechas vinculaciones y no poca promiscuidad entre metáfora, metonimia, sinécdoque y perífrasis; [...]. [... it doesn't take a long time to detect tight *connections* and considerable promiscuity between metaphor, metonymy, synecdoche and periphrasis] (*RefNom*)  
 c. El portavoz del PP, Rafael Hernando, aseguró que esta operación demuestra la vinculación de ETA con la "kale borroka" [...]. [... guaranteed that this operation shows the *connection* of ETA with the "kale borroka"] (*StateNom*)

Given that the semantic recategorization of these items is not clear-cut but gradient and highly context-dependent, they all constitute interesting subjects of research.

	EvNom	StateNom	RefNom
<i>Creación</i> [creation] <i>Construcción</i> [building]	+		+
<i>Traducción</i> [translation] <i>Destrucción</i> [destruction]	+		+
<i>Preocupación</i> [preoccupation] <i>Interés</i> [interest]		+	+
<i>Vinculación</i> [linking] <i>Interrupción</i> [interruption]	+	+	+

**Table 1.** Set of Spanish nominalizations included in the corpus

For this set, a random sample was compiled of 3,000 tokens, all selected from the CREA corpus. The following parameters were taken into account: (i) *location*: Peninsular Spanish, (ii) *period*: 2000-2004, (ii) *genre*: press, fiction and non-fiction books. In a first phase all the examples that respond to these parameters were extracted from the CREA corpus, which

<sup>3</sup> Both event and state readings still refer to eventualities, and (4a) and (4c) can thus be easily paraphrased by the corresponding verbal forms: "it has been usual to *connect* human territoriality with the animal one" (4a), and "this operation shows that Eta and "kale borroka" *are connected*" (4c). The only difference between events and states is that only the former are dynamic. Referential nominals, on the other hand, lack temporality and thus acquire a more concrete meaning. As a consequence, in (4b), *vinculaciones* refers to an entity (or entities), and thus to "what links metaphor, metonymy, synecdoche and periphrasis".

<sup>4</sup> All examples cited in this article are selected from the CREA corpus. The more detailed bibliographical information of each case is not reproduced in the present text for practical reasons but can easily be retrieved in this database. The relevant parts of the Spanish examples are translated into English but note that these translations are often literal translations containing a nominalized form, which would perhaps not always be the most natural solution for a native speaker of English.

resulted in a base corpus of 8,887 cases. Afterwards, a random sample of 3,000 tokens was conducted on this base corpus. Then, the tokens were manually annotated according to the semantic interpretation of the nominalized form (i.e. event, referential or state nominal). However, a number of examples were semantically ambiguous and were therefore replaced by other unquestionable examples randomly selected from the base corpus.

The selection procedure gave the following distribution: *creación* – 639 ex.; *construcción* – 741 ex.; *traducción* – 100 ex.; *destrucción* – 125 ex.; *preocupación* – 364 ex.; *interés* – 904 ex.; *vinculación* – 102 ex.; *interrupción* – 25 ex. However, given that our primary concern was to establish this classification on an objective and non-intuitive basis, avoiding a circular reasoning<sup>5</sup>, only the cases with an overt formal diagnostic of their semantic classification – whose significance has been previously demonstrated – were selected for further analysis (Balvet et al., 2011; Barque et al., 2009; Bekaert & Enghels, 2015; Fábregas & Marín, 2012a; Fábregas, Marín & McNally, 2012; Huyghe & Marín, 2007; Marín 2011; Marín & Vilhoing, 2012; Vanhoe, 2004). In particular, formal elements pointing toward the event classification of a nominal are (a) the dynamic verbs *producirse* [to occur] or *tener lugar* [to take place], (b) other verbs that describe properties of dynamic eventualities, such as *empezar*, *comenzar* [to begin], *acabar* o *terminar* [to end], or (c) the prepositional phrase *en curso de* [during] (e.g. (5a), and more examples below). Next, although positive diagnostics of stativity are quite rare, frequent collocations of state nominals are *un estado de* [a state of], emotion verbs like *sentir* [to feel], or verbal groups like *dar muestras de* [+/- to show] (5b). Finally, referential nominals tend to appear with lexical elements that emphasize their concrete meaning, like the verbs *ofrecer* [to offer] and *leer* [to read], or concretizing modifiers as *pétreo* [stone] and *de madera* [wooden] (5c).

- (5)
  - a. Si **se produjera** la pérdida o *destrucción* accidental de la copia. [If the loss or the accidental *destruction* of the copy would occur]
  - b. Fernández Villaverde **sentía** *preocupación* por la abundancia de billetes que circulaban. [Fernández Villaverde felt *preoccupation* with the abundance of banknotes that circulated]
  - c. Poco se conoce de los arquitectos que diseñaron estas emblemáticas *construcciones pétreas*. [Little is known about the architects who drew these emblematic stone *buildings*]

Moreover, some diagnostics appear in the larger context, as is the case in the following example:

- (6) Las *construcciones* más interesantes que han llegado hasta nuestros días son de carácter religioso. **La Iglesia Parroquial de San Pedro** es un templo renacentista. [The most interesting *constructions* that made it to our days have a religious character. The Parochial Church of Saint Peter is a Renaissance temple.]

---

<sup>5</sup> This circular reasoning would involve semantically classifying the nominals intuitively and possibly under the influence of their morphosyntactic behavior (e.g. the presence of a determiner, their argument structure), and then verifying the morphosyntactic behavior of these instances.

In this example, *construcciones* refers to the church that is described in the following sentence, which clearly emphasizes its referential meaning<sup>6</sup>. Table 2 gives a complete overview of the diagnostics that have been applied:

<b>EvNom</b>	[+ Dynamic]	Verbs: <i>tener lugar, producirse, proceder a, empezar, comenzar, acabar, terminar</i> Modifiers: <i>rápido, lento, acelerado</i>
	[+ Temporal]	Verbs: <i>coincidir (con), retrasar, acelerar, estar en marcha</i> Modifiers: <i>simultáneo, instantáneo, repentino, progresivo, gradual</i> Prepositions: <i>durante, en curso de, antes de, después de, tras</i> Larger context (e.g. <i>acción, proceso</i> )
<b>StateNom</b>	[- Dynamic]	Verbs: <i>sentir, dar muestra(s) de, (de)mostrar, dejar muestras de</i> Collocation: <i>un estado de</i>
	[+ Temporal] <sup>7</sup>	Verb: <i>coincidir (con)</i> Modifiers: <i>simultáneo, instantáneo, repentino, constante, permanente</i>
<b>RefNom</b>	[+ Referential]	Verbs: <i>nacer, presentar, ofrecer, poseer, leer, publicar, editar, etc.</i> Modifiers: <i>bueno, malo, magnífico, principal</i> Concrete appositions and predications Larger context (e.g. <i>iglesia</i> )

**Table 2.** Overview of the formal diagnostics

As a result of this manual and labour-intensive selection process, the concrete working set was reduced to a corpus of 589 tokens. The tokens are distributed as follows: *creación* – 142 ex.; *construcción* – 177 ex.; *traducción* – 52 ex.; *destrucción* – 29 ex.; *preocupación* – 74 ex.; *interés* – 94 ex.; *vinculación* – 14 ex.; *interrupción* – 7 ex. These cases were annotated with the morphosyntactic categories that are pointed at by the transcategorization clines (Figures 1 and 2) and are relevant for Spanish nominalizations<sup>8</sup>: (a) presence and type of determiner, (b) singular or plural form, (c) presence and type of modifier, and (d) development of argument structure<sup>9</sup>.

In what follows, the results of the analysis are described in more detail. Section 4.1 examines the nominal features attesting the degree of recategorization or substantivization of the forms, while Section 4.2 elaborates on the extent to which they preserve the verbal feature of argument

<sup>6</sup> The fact that some elements appear in the larger context explains why some examples cited in the article do not display a formal diagnostic.

<sup>7</sup> Events and states are both defined as temporal eventualities. Therefore, some formal diagnostics overlap. Nevertheless, these tests can still be used with the forms that only combine two different interpretations, as *creación*, that combines an eventive meaning with a referential one, and *preocupación*, that manifests a stative and a referential use.

<sup>8</sup> Spanish, like English, lacks case marking, and thus does not allow us to make any predictions in this respect, as opposed to other languages.

<sup>9</sup> In fact, we did not only compare the behavior of the different classes of deverbal nominalizations, but also looked into the behavior of the items separately. However, these comparative data will not be discussed in the present article, because the items display a quite uniform behavior.



structure. With regard to the applied methodology, we first operationalize each morphosyntactic feature by defining its levels and formulating a working hypothesis. These hypotheses are then empirically verified through an exploration of bivariate statistics characterizing the relation between two variables: an independent one (i.e. the semantic type: event nominal vs. referential nominal vs. state nominal) and a dependent one (each of the morphosyntactic features listed above). The following questions guide the analysis: (1) To what extent can the semantic classification of the nominalized forms be correlated with the presence of each morphosyntactic feature, and in what order? (2) Does a more “nouny” – referential – interpretation indeed coincide with a higher degree attained in the substantivization cline, and does a more “verby” – event – reading match up with a lower degree in the substantivization and deverbalization clines? (3) And to which kind of extreme (nouny or verby) are state nominals drawn? Section 4.3 presents the results of a multifactorial analysis, viz. the classification tree or decision tree test. This test measures the relative importance of each (morphosyntactic) variable in the process of nominalization within the whole of factors, and shows to what extent the transcategorial clines can be empirically verified for Spanish.

## 4. Results and discussion

### 4.1. Acquisition of nominal features

It can be assumed that referential nominals approximate the nominal prototype, and thus will acquire more nominal features than event and state nominals. Following the substantivization cline in descending order (Figure 2), this section examines to what extent this hypothesis can be confirmed by the analysis of the determiners which the three types of deverbal nominals combine with (4.1.1), their possibility to pluralize (4.1.2) and their modification pattern (4.1.3).

#### 4.1.1. Determination

Event and state nominals are said to only admit the definite article and the possessive determiner, whereas referential nominals combine with the entire range of determiners, including indefinite articles and demonstrative determiners (Alexiadou, 2001; Barque, Fábregas & Marín, 2011; Dam-Jensen, 2008; Grimshaw, 1990; Mourelatos, 1978; Picallo, 1991, 1999). Although these rigid assumptions have been refined by several authors (Dubois & Dubois-Charlier, 1999; Rauh, 2004; Sleeman & Brito, 2010a, b)<sup>10</sup>, it can be assumed that referential nominals allow for a wider range of determiners. This theory can be examined by analyzing the data in Table 3.

	EvNom		StateNom		RefNom	
	#	%	#	%	#	%
Definite article	170	<b>80.19</b>	27	<b>31.40</b>	117	40.21
Possessive	30	14.15	29	<b>33.72</b>	39	13.40
Indefinite article	4	1.89	17	19.77	47	16.15

<sup>10</sup> Sleeman and Brito (2010a, b), for instance, show that event nominals can combine with a demonstrative determiner, which has then a contrastive effect (e.g. *The journalists were watching that construction of the bridge, when the bomb fell.*)

Demonstrative	2	0.94	4	4.65	18	6.19
Indefinite quantifier	1	0.47	3	3.49	25	8.59
Numeral	0	0.00	0	0.00	2	0.69
Interrogative	0	0.00	0	0.00	1	0.34
Zero determiner	5	2.36	6	6.98	42	14.43
	212	100	86	100	291	100

**Table 3.** Semantic class and determiner type<sup>11</sup>

The quantitative analysis shows that event nominals generally combine with the definite article (7a) (80.19%) and, albeit less frequently, with a possessive determiner (7b) (14.15%). A combination with other determiners is not excluded, but highly uncommon in the corpus. By contrast, referential and state nominals show a more balanced distribution: besides the definite article and possessive determiner, they also frequently combine with a demonstrative determiner (8a), an indefinite article (8b) or another indefinite determiner (8c)<sup>12</sup>. Finally, only referential nominals combine with a numeral (9a) or an interrogative determiner (9b).

- (7) a. Debido a que el tiempo de vida medio de las células transfundidas es corto y **la** *destrucción* eritrocitaria continua puede ser necesaria una nueva transfusión. [Given that the mean lifetime of transfused cells is short and the erythroid *destruction* continuous...]  
b. Los petroleros de doble casco empezaron a construirse generalizadamente desde comienzos de los 90 y **su** *construcción* no entraña ningún problema. [... and their *construction* did not entail any problem.]
- (8) a. Y cuando el genial artista malagueño contempló **aquella** maravillosa *creación* [...]. [And when the genius artist from Málaga contemplated that wonderful *creation*] (*RefNom*)  
b. Carl Sagan no dejó de mostrar **un** desmedido *interés* por buscar una explicación al enigma de Sirio. [Carl Sagan kept on showing an excessive *interest* in searching for an explication for the Sirian puzzle.] (*StateNom*)  
c. Prueba de ello son los kioscos y **otras** *construcciones* que se sitúan en las proximidades del camino. [Evidence of that are the kiosks and other *buildings* that are situated close to the road.] (*RefNom*)
- (9) a. [...] **doce** viejas *construcciones* se están convirtiendo en lo que se ha bautizado como Maisons folie [...]. [...twelve old buildings are being transformed in what was baptized as Maisons folie]  
b. Y **qué** *traducción* electoral puede tener todo esto. [And what electoral translation can all this have.]

In sum, the theoretical prerequisites are largely confirmed for the Spanish event and referential nominals: the former tend to combine only with a definite article or a possessive determiner, whereas the latter allow for a wide range of different determiners. Besides the basic approach to prototypicality in terms of frequency (the underlying rationale being that more prototypical items of a class occur more frequently), what is highly interesting for our approach is its

<sup>11</sup> Given that this table is a larger than 2 x 2 table, we calculated a two-tailed Fisher's Exact test. The p-value was extremely low, pointing toward a correlation between the semantic category and the determiner type.

<sup>12</sup> Illustrating each different category with an example would lead us too far and is beyond the scope of this paper. Please note, however, that we have strived to provide a well-balanced representation of the different phenomena.

definition in terms of distribution. Indeed, more prototypical elements are taken to be less formally constrained and therefore appear in a wider variety of contexts (Gries, 2006; Jansegers et al., 2015). So, taking into account the multiplicity of determiners they can combine with, referential nominals clearly confirm their higher degree of prototypical nouniness. Particularly interesting is also the behavior of state nominals. They seem to take an intermediate position between the other two semantic types: they show a more balanced distribution than event nominals, but reject some determiners that are allowed by referentials, in particular numerals and interrogative determiners.

#### 4.1.2. Pluralization

As referential nominals often refer to concrete objects, they are said to pluralize easily (Croft, 2001; Grimshaw, 1990). Events, on their part, tend to appear in singular form but do allow pluralization when referring to delimited events (Bisetto & Melloni, 2005; Marín & Villoing, 2012; Melloni, 2011; Picallo, 1999, among many others).<sup>13</sup> State nominals, in turn, lack aspectual delimitation, which means that they should by definition reject pluralization (Barque, Fábregas & Marín, 2011; Fábregas & Marín, 2012a; Marín & Villoing, 2012; Varela, 2012). Table 4 establishes a correlation between the semantic class of the nominalized forms and their ability to pluralize.

	EvNom		StateNom		RefNom	
	#	%	#	%	#	%
Singular	208	<b>98.11</b>	86	<b>100</b>	201	69.07
Plural	4	1.89	0	0	90	30.93
	212	100	86	100	291	100

**Table 4.** Semantic type and pluralization

The chi-square test confirms that the semantic categories significantly differ as to the variable pluralization ( $\chi^2 = 96.246$ ,  $df = 2$ ,  $p < 0.001$ ), and the association between the variables is shown to be strong (Cramér's  $V = 0.404$ )<sup>14</sup>. Indeed, only referentials frequently pluralize, both when they refer to a concrete object (10a), and when they acquire a more abstract meaning through some metonymical transposition (10b). Event nominals can also pluralize, although this is highly exceptional (1.89%) and causes the nominalized form to refer to a sequence of temporally delimited events, such as the achievements in (11). Finally, state nominals clearly stick with the event nominals and reject pluralization.

- (10) a. El conjunto arquitectónico está integrado por dos grandes bloques de edificios: la Basílica, que es accesible y las **construcciones** para turistas y peregrinos. [... and the *buildings* for tourists and pilgrims.]

<sup>13</sup> This follows the idea that time can be conceptualized in terms of space (e.g. Lakoff & Johnson, 2003). Thus, a bounded region in time (being an event with fixed temporal boundaries) more easily allows pluralization.

<sup>14</sup> Cramér's  $V$  is a correlation coefficient which measures the strength of a correlation between two variables. The result is a coefficient between 0 and 1 showing no relationship between both variables or a strong one, respectively. The value 0.404 points toward a strong association.

b. Se manifiesta de forma difusa, se muestra como imágenes sociales del espacio, como **construcciones** ideológicas y simbólicas. [It appears in a diffuse way, as social images of space, as ideological and symbolic *constructions*]

- (11) No es normal la frecuencia con la que se están produciendo **interrupciones** en el suministro eléctrico. [The frequency with which *interruptions* in the electricity supply occur is not normal.]

To conclude, only instances of the semantic class that approximates the nominal prototype – namely referential nominals – admit pluralization. The semantic types that refer to situations – event and state nominals – tend to reject it in varying degrees.

#### 4.1.3. Modification

Although verbs are typically modified by adverbs and nouns carry adjectival modifiers, the different types of deverbal nominalizations are said to all combine with adjectival modifiers (Alexiadou, 2001; Azpiazu, 2004; Comrie & Thompson, 1985; Dik, 1985; Givón, 2001a; Varela, 2012). However, former studies have shown that there are some crucial differences between nominalization types with regard to modification. Picallo (1999) assumes that adverbial adjuncts can be transposed to the construction including an event nominalization without adding a preposition (12). By contrast, in similar contexts referential nominals require the preposition *de* [of] (e.g. *una magnífica creación de 20 años más tarde*, ‘a magnificent *creation* of 20 years later’). Little is known about the modification of state nominals, but as they refer to an eventuality, it can be expected that they will also admit the appearance of adverbial adjuncts.

- (12) Por efímero que hubiera sido este decreto de Salamanca, no cabe regatearle la gloria de haber abierto el camino que condujo a la *creación* de la peseta, **20 años más tarde**. [... the path that led to the *creation* of the peseta, 20 years later]

In order to describe in more detail the modification pattern of the three semantic types in the Spanish corpus, Table 5 first examines to what extent each one combines with one or more modifiers, or none.

	EvNom		StateNom		RefNom	
	#	%	#	%	#	%
0	144	<b>67.92</b>	47	<b>54.65</b>	73	25.09
≥ 1	68	32.08	39	45.35	218	<b>74.91</b>
	212	100	86	100	291	100

**Table 5.** Semantic type and presence of modifier

Again, the chi-square test confirms that the three categories differ significantly with respect to the presence or absence of a modifier ( $\chi^2 = 94.941$ ,  $df = 2$ ,  $p < 0.001$ ), and that the association between both variables is strong (Cramér’s  $V = 0.401$ ). Whereas referential nominals mostly carry at least one modifier (74.91%) (13a), event (e.g. 7b *supra*) and state nominals (e.g. 3b *supra*) most frequently appear without any modifier being expressed (resp. 67.92% and 54.65%). However, modification of both types is not excluded (13b-c), especially for state nominals.

- (13) a. La versión castellana es deudora de la **magnífica traducción** de mi hermano Xúlio R. Trigo. [... the excellent *translation* of my brother Xúlio R. Trigo.] (*RefNom*)  
 b. Y entre esos hechos consumados se cuenta también la **construcción progresiva** de un aberrante y racista muro **en Cisjordania** [...] [... the progressive *building* of an aberrant and racist wall in the West Bank] (*EvNom*)  
 c. Ella no mostraba el **mismo interés**. [She did not show the same *interest*] (*StateNom*)

However, the analysis of the type of modifiers with which the three types combine reveals a more nuanced pattern. In particular, the variable “modifier type” comprises three levels: (a) adjectival modifiers (*la constante preocupación* [the constant preoccupation]), (b) non-argumental adnominal complements<sup>15</sup> (*las construcciones de mayor nivel económico* [the buildings of the highest economical level]), and (c) relative clauses (*el interés que suscitaba* [the interest that it provoked]). Table 6 details the modifier type of each semantic class.

	EvNom		StateNom		RefNom	
	#	%	#	%	#	%
Adjective	14	19.44	34	<b>79.07</b>	182	<b>65.70</b>
Adnominal adjunct	57	<b>79.17</b>	4	9.30	57	20.58
Relative clause	1	1.39	5	11.63	38	13.72
	72	100	43	100	277	100

**Table 6.** Semantic type and modifier type<sup>16</sup>

Referential and state nominals tend to combine with adjectival modifiers (cf. *supra* 13a-c), and, to a minor extent, with relative clauses (14). This co-occurrence is not unexpected given that relative clauses are said to behave like adjective clauses, prototypically modifying a nominal nucleus (Real Academia Española, 2009a, 2009b).

- (14) a. Es un modelo racional, una **construcción que parte, de hecho, de un “número limitado de hechos y observaciones”**. [It is a rational model, a *construction* that departs, in fact, from a “limited number of facts and observations”.] (*RefNom*)  
 b. Ya lo anticipó Alfredo Kraus, con el **interés que mostró en sus momentos finales por este tenor singular**. [Alfredo Kraus already anticipated it, with the *interest* that he showed in his final days for this unique tenor.] (*StateNom*)

In contrast, event nominals mostly combine with adnominal adjuncts (79.17%). These complements are prepositional phrases that correspond to adverbial adjuncts in a clause. Although they are mainly spatial (15a) or temporal (15b) complements, adjuncts expressing finality are also observed (15c). Since the accomplishment of the event expressed in the final clause always depends on the realization of another event, namely the one expressed by the

<sup>15</sup> Non-argumental adnominal complements can be defined as the equivalents within the NP domain of adjuncts at the clause level (Escandell Vidal, 1995). As opposed to the argumental complements (analyzed in detail in Section 4.2), they are not selected by the nucleus. There is thus no restriction with regard to the quantity of concatenated non-argumental adnominals, which then add different types of semantic specifications to the NP (referring to the (alienable) possessor, or adding temporal, spatial, modal or attributive information).

<sup>16</sup> Given that this table is a larger than 2 x 2 table, we calculated a two-tailed Fisher’s Exact test. The p-value was extremely low, pointing toward a correlation between the semantic category and the modifier type.

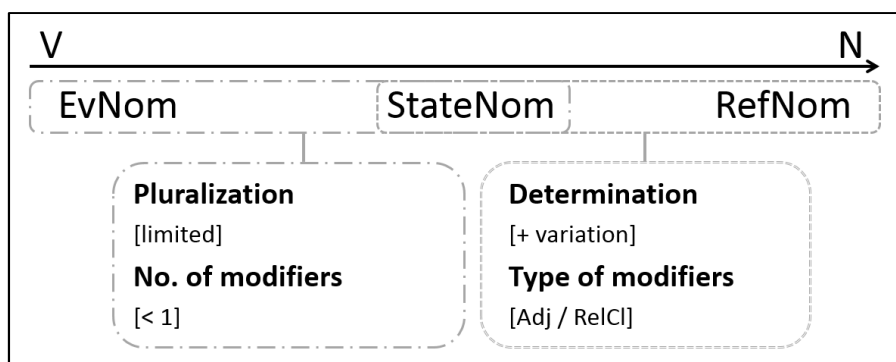
nominalized form, the presence of a final adjunct clearly emphasizes the situational character of event nominals (Croft, 1991; Picallo, 1999; Vanderschueren, 2013).

- (15) a. Uno de sus trabajos, la *construcción* de carreteras **en las selvas del Amazonas**, denunciado por los conservacionistas de la naturaleza, se había realizado manteniendo al personal en auténticos búnkers. [One of his jobs, the *building* of roads in the Amazon jungles ...]
- b. El Parlamento aprobó en junio la *creación* de 30.000 nuevas plazas de guardería **a lo largo de la legislatura**. [The Parliament approved in June the *creation* of 30.000 new childcare places throughout the term.]
- c. [...] la banda terrorista ETA se propone la *creación* de un centro de impresión “**para publicar cualquier cosa, periódicos, libros, carteles, pegatinas**”. [... the terrorist group ETA decided on the *creation* of a printing center “to publish anything, newspapers, books, posters, stickers”.]

To sum up, the semantically closer a nominalization is to the nominal prototype, the more frequently it combines with one or more modifiers, and vice versa. Moreover, each semantic class displays some preference as to what kind of modifier it combines with: the more “nouny” referentials select adjectives and relative clauses; the event nominals, which have the least affinity with the nominal prototype, mostly select adnominal adjuncts, equivalent to clausal adverbial adjuncts. The state nominals can be situated in between: they appear almost as often with or without a modifier, but when they do take a modifier, it is mostly a typically nominal one (i.e. an adjective or relative clause).

#### 4.1.4. Interim conclusions

The empirical analysis of the acquisition of nominal features by Spanish deverbal nominalizations shows that in general terms, they all recategorize and thus substantivize, albeit to varying degrees. First, the morphosyntactic behavior of Spanish referential nominals confirms that they approximate the nominal prototype: (a) they can combine with the entire range of determiners, (b) they can easily pluralize, (c) and they frequently combine with one or more modifiers, particularly with adjectival modifiers and relative clauses. Secondly, the event nominals, in contrast, acquire fewer nominal features: (a) they almost exclusively appear with a definite article or possessive determiner, (b) they rarely pluralize, and (c) they disfavor modifiers. When a modifier appears, this tends to be an adnominal adjunct that resembles a clausal adverbial complement. Finally, state nominals occupy an intermediate position between the other two semantic classes: (a) their determination pattern is more balanced than the one of event nominals, but less varied than the one of referentials, (b) they tend to reject pluralization, and (c) only half of the examples combine with a modifier, which is nevertheless typically a nominal one (i.e. adjectives and relative clauses). Figure 3 plots the acquisition of nominal features on a continuum from the “verby” extreme on the left to more “nouny” behavior on the right.



**Figure 3.** Semantic type and acquisition of nominal features

#### 4.2. Loss of verbal features: the Argument Structure

We already know that Spanish deverbal nominalizations are strongly nominalized. Consequently, they allow for a proper quantitative analysis of only one verbal feature, namely their argument structure (AS). As event and state nominals semantically refer to an eventuality or situation which automatically implies one or more participants, these semantic types are expected to preserve a more elaborate argument structure. In the subsequent analysis, the notions of First Complement (FirstC) and Second Complement (SecC) are used to refer to different constituents which can be part of the argument structure of a nominalized form:

- The First Complement refers to the nominal argument that approximates the proto-agent (Dowty, 1991; Van Valin, 1999, 2005) (e.g. *la creación de un excelente literato* [the creation of a great man of letters]; *la destrucción de los ídolos por el hermano del príncipe* [the destruction of the idols by the brother of the prince]);
- The Second Complement corresponds to the nominal argument that approaches the proto-patient (Dowty, 1991; Van Valin, 1999, 2005) (e.g. *la traducción de la Biblia* [the translation of the Bible]; *la destrucción de los ídolos por el hermano del príncipe*).

With regard to the argument structure of the three semantic classes, the following differences can be pointed out:

- Event nominals require the overt realization of the SecC, introduced by the preposition *de* [of] or formalized as a possessive determiner.<sup>17</sup> The realization of the FirstC is optional, but when it appears, it tends to carry the preposition *por* (*parte de*) [by] (Azpiazu, 2004; Giammatteo, Albano, & Ghio, 2005; Melloni, 2011; Picallo, 1999).
- State nominals are said to combine with an obligatory FirstC, which refers to the experiencer of the state and is introduced by the preposition *de* [of] or appears as a possessive determiner. The SecC, referring to the cause of the state, can be omitted, but when it is realized, it can be introduced by a variety of prepositions (Barque et al., 2009; Fábregas & Marín, 2012a; Marín & Villoing, 2012; Melloni, 2011).
- Given that referential nominals prototypically refer to a concrete object that does not imply any participant, their argument structure is less elaborate. First, they admit the realization of the FirstC. It is important to note that this complement does not refer to a participant of the eventuality and should instead be defined as a possessor in a broad sense. It is introduced by the preposition *de* [of] or expressed as a possessive determiner. Second, most of the

<sup>17</sup> Koptjevskaja-Tamm (1993) and Heyvaert (2008) have fine-tuned these rigid assumptions, showing that the Second Complement can be omitted, in a nominal ellipsis.

referential nominals reject the realization of the SecC, which would automatically trigger an event or state reading (Alexiadou, 2001, 2009; Amador Rodríguez, 2009; Azpiazu, 2004; Harley, 2009; Picallo, 1999; Varela, 2012). However, one type of referentials, namely those that combine with an affected SecC (like *traducción* [translation] and *destrucción* [destruction]), do combine with a SecC, which is introduced by the preposition *de* [of] or realized as a possessive determiner (Bisetto & Melloni, 2005; Escandell Vidal, 1995; Melloni, 2010, 2011; Sleeman & Brito, 2010a).<sup>18</sup>

Table 7 details the different argument patterns displayed by the three types of deverbal nominalizations.

	EvNom		StateNom		RefNom	
	#	%	#	%	#	%
[-FirstC][-SecC]	19	8.96	14	16.28	191	<b>65.64</b>
[+FirstC][-SecC]	1	0.47	6	6.98	79	27.15
[-FirstC][+SecC]	190	<b>89.62</b>	32	<b>37.21</b>	19	6.53
[+FirstC][+SecC]	2	0.94	34	<b>39.53</b>	2	0.69
	212	100	86	100	291	100

**Table 7.** Semantic type and argument structure

Event, state and referential nominals appear to be significantly different with regard to their argument structure ( $\chi^2 = 558.051$ ,  $df = 6$ ,  $p < 0.001$ ), the correlation between both variables being strong (Cramér's  $V = 0.688$ ).

The first difference is that referential nominals are completely deverbalized, as they tend to occur without any complement expressed (65.64%). When an argument is present, it is in general the FirstC (27.15%), introduced by the preposition *de* [of] (16a) or realized as a possessive determiner (16b). Moreover, the data confirm that only referentials that carry an affected SegC admit its syntactic realization (16c-d). The realization of both complements is extremely rare with RefNoms (0.69%) (16e).

- (16) a. Según Lampérez, aquí se pierde la uniformidad bernarda para adquirir cierto carácter de regionalidad, fenómeno inusual en las *construcciones* **de los monjes blancos**. [... an unusual phenomenon in the *buildings* of the white monks.]  
 b. La primera vez que escribe una letra, por ejemplo, exhibe triunfalmente **su** “*creación*”. [The first time he writes a letter, for instance, he triumphantly displays his “*creation*”.]  
 c. Cabe señalar que, de acuerdo con las *traducciones* más fidedignas **de la Biblia**, los gigantes de los que se habla en el Génesis serían, etimológicamente, los Nefilim. [It should be noted that, in accordance with the most reliable *translations* of the Bible ...]

<sup>18</sup> As a result, it is necessary to distinguish between referentials with a created Second Complement and referentials with an affected Second Complement. The verbs *crear* [create] and *construir* [build] combine with a theme created through the action of the verb (it was thus non-existing before the event took place). By contrast, the theme of the verbs *traducir* [translate] and *destruir* [destroy] already exists before the actions take place, so that it is only affected by the action of the verb.



d. Estas dos últimas frases no son mías, ni lo es **su** bella *traducción*. [These last two sentences aren't mine, neither is their beautiful *translation*.]

e. La *traducción de la obra de Jacques Novicow por Nicolás Salmerón*, aludida más arriba, fue publicada por Daniel Jorro, Editor, de Madrid, en 1914. [The *translation* of the work of Jacques Novicow by Nicolás Salmerón, referred to above, was published...]

As opposed to this argument-poor pole, both event and state nominals manifest a more elaborate argument structure. To begin with, event nominals mostly realize their SecC, which then carries the preposition *de* [of] (17a) or is replaced by the possessive determiner (17b) (89.62%). However, and against what was postulated by Grimshaw (1990), the absence of the SecC does not impede an event reading (8.96%) (17c). The FirstC is rarely present, but when it appears, it is introduced by the preposition *por* [by] (17d).

(17) a. De hecho, los descubrimientos de Colón inician la *destrucción de la imagen del mundo*, es decir, de las concepciones sobre las que se había asentado esa imagen a lo largo de la Edad Media. [In fact, the discoveries of Columbus initiate the *destruction* of the world view ...]

b. Durante **su** *construcción* se hundió el viaducto de tres arcos que lo sostenía. [During its *construction*, the viaduct of three arches that sustained it collapsed ...]

c. El objetivo de la sonda Dawn será caracterizar las condiciones y los procesos de la época más temprana del Sistema Solar, investigando detalladamente dos de los protoplanetas más grandes que permanecen intactos desde la *creación*. [... that remain intact ever since the *creation*.]

d. Y quizá el aldabonazo más duro lo ha dado el terrorismo, y las consecuencias de este usando los medios técnicos descubiertos hasta ahora, que pueden hundir el avance occidental, como hace temer lo ocurrido con la *destrucción, por fanáticos islamistas usando aviones suicidas, de las dos emblemáticas torres del World Trade Center en Manhattan* (Nueva York). [... as leads to fear what occurred with the *destruction*, by fanatical Islamists using suicide planes, of the two emblematic towers of the World Trade Center in Manhattan.]

In fact, these data suggest that the syntactic structure of the event nominal, at least in Spanish, must be related to the passive voice. This does not come as a surprise, given the more static nature of the passive compared to the active construction (Givón, 1982: 34). Not only does the FirstC carry the same preposition as the agent in a passive sentence, both structures foreground the event as a whole, backgrounding its agent or cause. This event only includes the internal argument or patient (syntactically realized as SecC) (Shibatani, 1985; Solstad & Lyngfelt, 2006).

Next, when the state nominal syntactically realizes only one complement, it tends to be the SecC, as opposed to the assumptions generally made in the literature. This SecC can be introduced by a wide variety of prepositions (18a-b). State nominals also frequently combine with both complements. In this case, the FirstC tends to be realized as a possessive determiner (18c), although it can also be introduced by the preposition *de* [of] (18d). State nominals also appear without any complement expressed (18e) or merely with a FirstC (18f), but these constructions are rather uncommon.

- (18) a. En este artículo se analiza cómo se plasmó en su obra un *interés por la arquitectura* que fue constante en toda la trayectoria del artista catalán. [This article analyzes how an *interest* for the architecture is captured in his work ...]
- b. Solbes agregó que sus homólogos de la eurozona mostraron “una cierta *preocupación*” **sobre si era el momento más oportuno para hacer una propuesta de este tipo** [...]. [Solbes added that his counterparts of the Eurozone showed “a certain *preoccupation*” as to whether or not it was the most appropriate moment to make a proposal of this kind ...]
- c. “Quiero felicitar al Consejo de Administración y al jugador por el esfuerzo que han hecho por llegar a un acuerdo”, apuntó Víctor Fernández, que días atrás había mostrado **su preocupación** [...] **por la posible marcha del central**. [...who days ago had shown his *preoccupation* [...] for the possible departure of the central defender]
- d. El ministro de Economía y Hacienda, Pedro Solbes, mostró este miércoles el *interés del Gobierno por buscar una solución para la financiación sanitaria*. [Wednesday, the Finance Minister, Pedro Solbes, showed the *interest* of the Government to search for a solution for the financing of Healthcare]
- e. En la última década del siglo XX persiste esa *preocupación*. [In the last decade of the 20th century that *preoccupation* persists]
- f. Con las espadas en alto continuarán las partes implicadas hasta el mes de febrero, con la consiguiente *preocupación de todos*, en especial de los empresarios [...]. [... with the resulting *preoccupation* of everyone, especially of the employers]

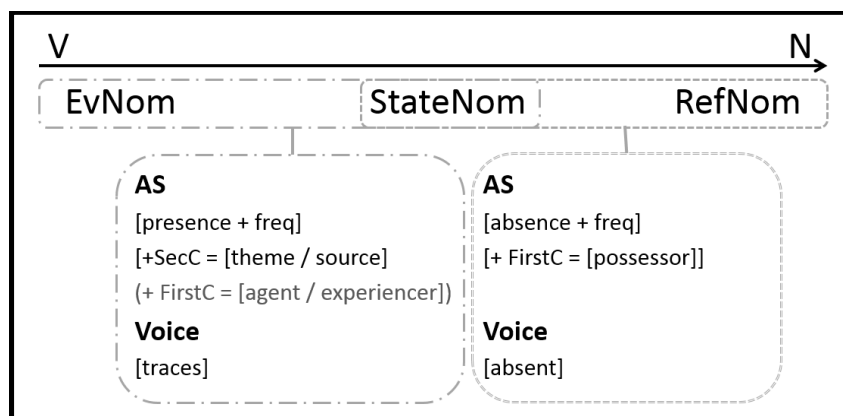
Spanish psychological verbs, from which these state nominals are derived, allow for various syntactic configurations (Marín & McNally, 2011; Marín, 2015), namely an accusative (19a), a dative (19b), a reflexive (19c) and a resultative construction (19d).

- (19) a. Marta lo preocupa. [Marta (actively) worries him.acc]
- b. El humo le preocupa. [The smoke worries him.dat]
- c. Josep se preocupa. [Josep gets.SE worried]
- d. Marta está preocupada. [Marta is worried]

However, the present analysis suggests that state nominals are fundamentally derived from the reflexive construction (of the type *preocuparse de*, *interesarse por*): this construction also combines with a series of different prepositions, and is the only one that has been analyzed as a straightforward stative one (Belletti & Rizzi, 1988; Marín, 2011; Marín & McNally, 2011; Vanhoe, 2004). Interestingly, Vanhoe (2004) analyzes this variant as a middle construction. Although we did not actively examine the presence of Voice markers within the set of Spanish deverbal nominals, we thus observe that both event and state nominals retain some explicit traces of Voice: the passive in the case of the event nominal, and the middle in the case of the state nominals. Moreover, the passive and the middle construction share an important number of semantic and syntactic features. Both have to be analyzed as intransitive constructions, with a subject that does not play an active role and an oblique complement that can easily be omitted. Moreover, in both cases, the oblique complement is frequently introduced by the preposition *por* [by] in Spanish (Vanhoe, 2004; Solstad and Lyngfelt, 2006).

In conclusion, as visualized in Figure 4, referential nominals approximate the nominal prototype both semantically and syntactically, which explains why they mostly appear without any complement expressed. When a complement appears, it will generally be the FirstC, which

should be interpreted as a possessor rather than the more “verby” agent participant. This complement can thus also be analyzed as a genitive marker, which constitutes a nominal feature. Event and state nominals display a more elaborate argument structure. However, they exhibit some important differences, which in turn can be explained by their verbal origin. Event nominals tend to combine with a mere SecC because of their connection with the passive structure. State nominals are clearly derived from the reflexive variant of Spanish psychological verbs: their SecC admits a large range of different prepositions, and they also frequently appear with both complements expressed.



**Figure 4.** Semantic type and deverbalization

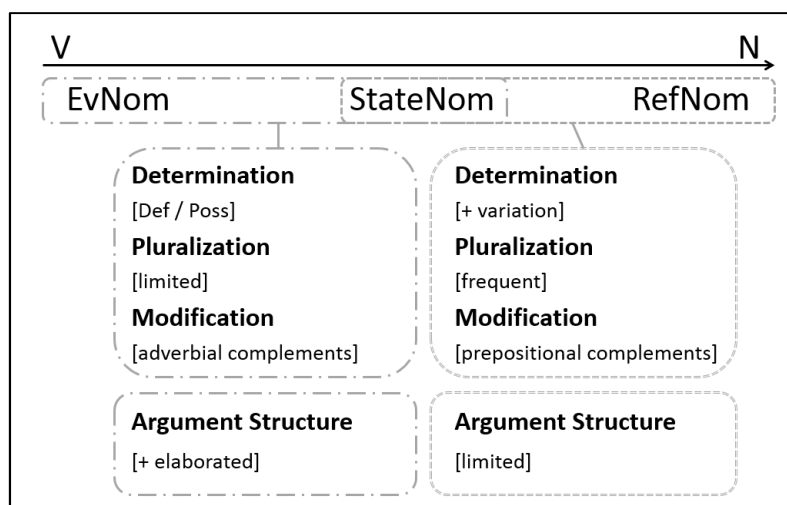
#### 4.3. Toward an integrated transcategorization cline for Spanish: a multifactorial account

As was pointed out in the introduction (Section 1), the substantivization and deverbalization clines developed by Malchukov (2004) are based on typological data. Following Koptjevskaja-Tamm (1993), he shows that deverbal nominalizations in different languages have reached different stages of the categorial shift. This methodology is highly valuable because it reveals universal patterns, but nevertheless presents some drawbacks too. Particularly, it does not allow us to observe subtler differences within a particular language where one single form can be recategorized to varying degrees (cf. the cases in Table 1), nor does it take into account the existence of other intermediate categories, viz. state nominals.<sup>19</sup> Therefore, a first task is to assess to what extent the thorough monolingual empirical analysis we have just performed confirms the predictions made by typologists.

To begin with, the monofactorial analysis of the different nominal and verbal features confirms the hypothesis formulated at the beginning of this article (Section 2), according to which “nouny” nominalizations (i.e. referential nominals) exhibit more nominal features and event nominals preserve more verbal features. The semantic heterogeneity of the state nominals also seems to reflect upon their morphosyntactic behavior. As a consequence, our thorough

<sup>19</sup> Another possible disadvantage of this methodology is that it assigns at the same layer Tense, Aspect and Mode, while various authors have shown that deverbal nominals indeed reject Tense and Mode markers, but preserve a certain degree of Aspectuality (Alexiadou, 2001; Barque et al., 2009; Cano & Jaque, 2011; Fábregas & Marín, 2012a, 2012b; Grimshaw, 1990; Varela, 2012). However, an Aspectuality analysis is beyond the scope of this study, because it requires a different approach. Particularly, Aspect can hardly be quantified and thus asks a more qualitative and descriptive study. It will therefore be the subject of a separate paper (see also Bekaert & Enghels, 2015).

quantitative and qualitative empirical analysis confirms the validity of the general model for Spanish, as shown in Figure 5.

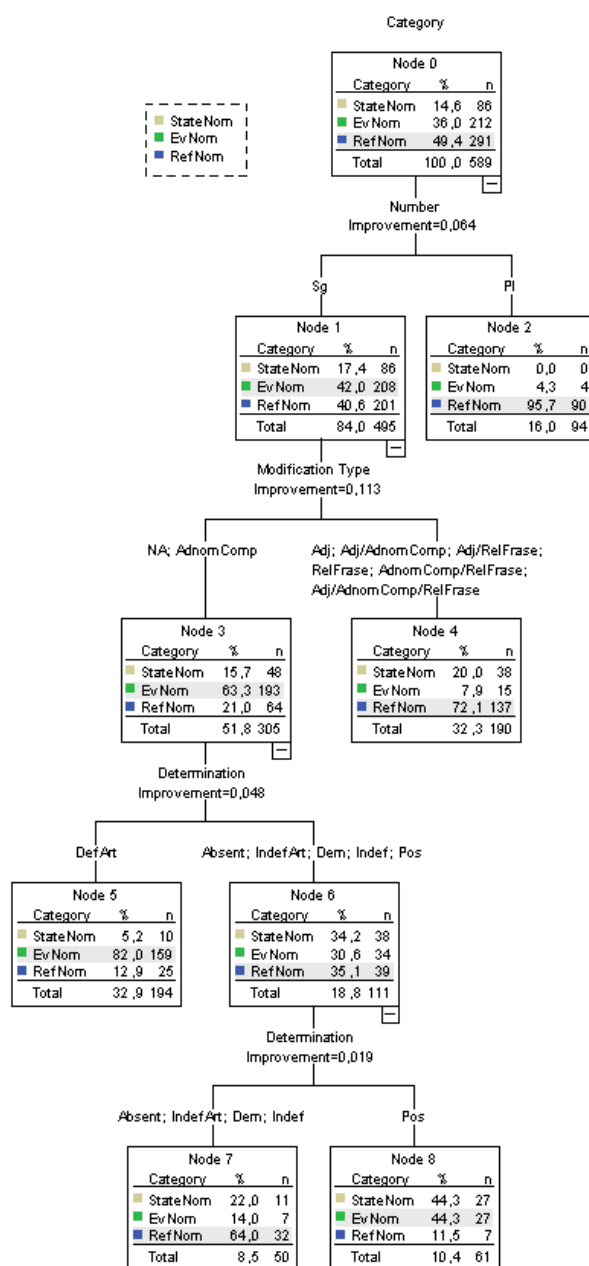


**Figure 5.** Semantic types and morphosyntactic features

However, at this point, it would be interesting to take the analysis one step further, and to examine whether the typological clines themselves – the order of the loss or acquisition of verbal and nominal features – can also be empirically verified. This would require a multifactorial analysis that examines the relative importance of each of the morphosyntactic features in the recategorization process. This goal can be achieved by means of the Classification Tree Method or Decision Tree test.<sup>20</sup> More specifically, the procedure builds a tree-based classification model by predicting values of a dependent variable based on values of independent or predictor variables (IBM Corporation, 1989, 2012). Among other things, it allows us to identify interactions between the dependent variable and the predictor variables. In this study, the dependent variable will be the semantic category with the values “EvNom”, “StateNom” and “RefNom”, while the morphosyntactic properties that have been analyzed above function as predictor variables. The variable inducing the first split should be seen as the variable that presents the strongest degree of interaction with the categorization, and is therefore identified as the variable through which the three categories are most distinct. The subsequent splits are gradually less decisive and thus, in terms of the goals of this study, reveal an empirically validated cline. The trees have been built using the CRT method (*Classification and Regression Trees*), which proposes binary splits to reinforce the dichotomy between more nominal and more verbal features.

<sup>20</sup> This test is preferred to other multifactorial methods because it is less exigent than for instance Logistic Regression, and allows for a more plain interpretation through its lucid visualization (see Piper et al., 2011: 95 for a more detailed argumentation).

The model presented in Figure 6 gives an account of the relative importance of the substantivization variables, namely “determination”, “number” and “modification type”<sup>21</sup> (risk estimate<sup>22</sup> = 0.244).



**Figure 6.** Classification tree: substantivization properties

The model that emerges suggests that the “number” variable presents the strongest relationship with the categorization of deverbal nominalizations. It shows that pluralization is strongly

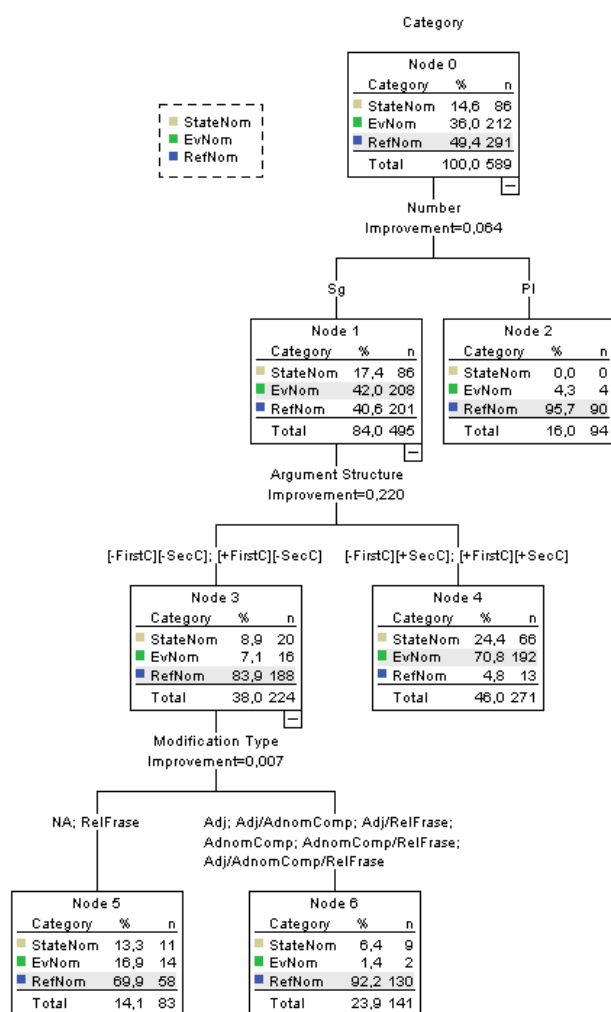
<sup>21</sup> The variable “modification type” also implies the value “absent” and thus incorporates the variable “presence/absence of modification”.

<sup>22</sup> The risk estimate evaluates the strength of the model, because it offers the proportion of incorrectly classified cases. A risk estimate of 0.224, for instance, implies that the model is incorrect in 22.4% of the cases and correct in 77.6% of the cases.

associated with the referential type, the closest to the nominal prototype, indicating that number is only acquired in a posterior, more advanced phase of the nominalization process. The second split is based on the variable “modification type” and indicates that nominal modifiers are acquired in an earlier stage of the nominalization process than pluralization, but later than determination, which is responsible for the last two splits. Indeed, determiners seem to be acquired early in the nominalization process, although a distinction must be made between definite articles and possessive determiners, on the one hand, and the other determiners on the other hand – the former being acquired earlier than the latter. In fact, if Figure 6 were turned horizontally counterclockwise, the different variables would appear on a continuum which, read from right to left, would reflect the features acquired early on and later in the nominalization process. This model can be compared with the substantivization cline proposed by Malchukov (see Figure 2 above).

Although Malchukov’s substantivization cline proposes that the feature Class, mainly expressed by adjectival modifiers, is closer to the nominal prototype than the feature Number, Malchukov (2004) suggests that both features seem to be acquired simultaneously. However, the model in Figure 6 suggest that, for Spanish, the features Class and Number need to be inverted: pluralization is restricted to the prototypical nominal (i.e. referentials), while adjectival modification extends to all three types. Secondly, the cline conjointly analyzes the definite article and the demonstrative determiner, under the [DET]-label. Yet, in Spanish, the definite article and the possessive determiner appear first in the substantivization process, while the demonstrative determiner is only acquired in a more advanced stage.

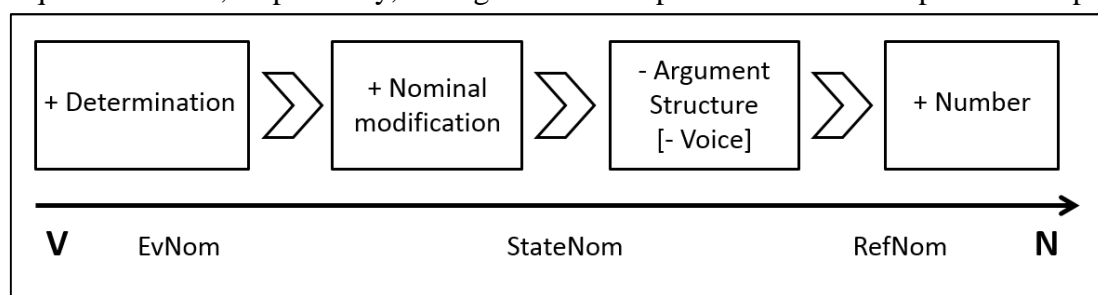
In Figure 7 the verbal feature of argument structure is integrated into the model. This second classification tree allows us to explore the possibility of arriving at an integrated cline, which would combine the decategorization and the recategorization features in one model (risk estimate = 0.202).



**Figure 7.** Classification tree: integrated model

Interestingly, the “argument structure” property is inserted between the variables “number” and “modification type”. The model thus suggests that the argument structure is lost only in a posterior, more advanced phase of the nominalization process, namely after the acquisition of determiners and nominal modifiers, but before the acquisition of number.

To conclude, the multifactorial analysis leads toward the integrated cline proposed in Figure 8 which, from left to right, shows the order in which both nominal and verbal characteristics are acquired and lost, respectively, throughout the complex nominalization process in Spanish.



**Figure 8.** Integrated transcategorial cline for Spanish nominalization

## 5. Conclusion

The micro-level descriptive study of a set of morphological nominalizations performed in this study has generated a number of significant theoretical and methodological insights .

In the first place, it has contributed to a better understanding of the transcategorial nominalization process in Spanish. By using a case study on “vagueness in grammar”, we have shown that the phenomenon of gradience is not only present in semantics and pragmatics, but also in morphosyntactic markings. Indeed, the class of deverbal nominalizations is not static but dynamic, as different members display different degrees of distance with regard to the prototypical noun. This distance is defined by the functional properties of the nominalized form, and more particularly by the degree to which its meaning overlaps with noun-like (for referentials) or verb-like (for event nominals) semantics. The state nominals are argued to be situated in between both groups. Moreover, it is interesting to observe how one particular lexeme can travel on this semantic continuum as it can be used differently according to the context. In a next phase of the analysis, the validity of the syntax-semantics interface principle was clearly demonstrated as the semantic continuum could be correlated to a morphosyntactic hierarchy: the more “nouny” forms (referentials) clearly exhibited more nominal features than the more “verby” ones (event nominals), as was expected, with the states once again displaying more indecisive behavior.

A second significant contribution of this study is that it has led toward a better comprehension of the radius of the typological de- and recategorization clines. These hierarchies were shown to be highly valuable as they offer a solid framework for studying the complex nominalization process. Also in Spanish, the categorial shift can be defined as an ordered acquisition and loss of particular features, but with regard to the proposed clines, some flexibility was required in order to account for its language-specific behavior. Among other things, it was shown that according to the behavior of the Spanish nominalizations in the corpus, the internal order between the acquisition of Number and other Class features (modification) should be inverted, and that the feature Determination needs to be refined. The syntactic complexity of the base verb was also shown to strongly influence the morphosyntactic behavior of the nominalized form.

Finally, at this point it is important to underline the crucial input of the detailed analysis of authentic corpus data, and especially the analytical statistical analysis. Not only did they allow us to verify in a very detailed manner the correlation between the semantics of the forms and their morphosyntactic behavior, the multifactorial model also led toward an integrated transcategorization cline for Spanish nominalizations, merging the nominal and verbal categories. The loss of argument structure properties could be situated in a rather late phase of the nominalization process, just before the acquisition of number. However, the proposed cline only concerns Spanish and not necessarily extends to other languages. It would thus be interesting to compare the results of this case study, and in particular the proposed integrated cline, with those of a comparable of deverbal nominals in (an)other language(s).

## References

Alexiadou, A. (2001). *Functional structure in nominals : nominalization and ergativity*. Amsterdam/Philadelphia: Benjamins.



- Alexiadou, A. (2009). On the role of syntactic locality in morphological processes: the case of (Greek) derived nominals. In A. Giannakidou & M. Rathert (Eds.), *Quantification, definiteness, and nominalization* (pp. 253-280). New York: Oxford University Press.
- Amador Rodríguez, L. A. (2009). *La derivación nominal en español: nombres de agente, instrumento, lugar y acción*. Frankfurt am Main: Peter Lang.
- Anderson, John M. (1997). *A notional theory of syntactic categories*. Cambridge: Cambridge University Press.
- Azpiazu, S. (2004). *Las estrategias de nominalización : estudio contrastivo del estilo nominal*. Frankfurt am Main: Peter Lang.
- Baker, M. (2003). *Lexical categories: verbs, nouns and adjectives*. Cambridge: Cambridge University Press.
- Balvet, A., Barque, L., Condet, M.-H., Haas, P., Huyghe, R., Marín, R., & Merlo, A. (2011). La ressource Nomage: Confronter les attentes théoriques aux observations du comportement linguistique des nominalisations en corpus. *TAL*, 52(3), 129-152.
- Barque, L., Fábregas, A., & Marín, R. (2011, April 18-20) *On the (un)countability of stative deverbal nouns*. Chronos 10, Aston University, Birmingham.
- Barque, L., Huyghe, R., Jugnet, A., & Marín, R. (2009). *Two types of deverbal activity nouns in French*. Paper presented at the 5th International Conference on Generative Approaches to the Lexicon, Pisa.
- Bekaert, E., & Enghels, R. (2014). Nominalizations of Spanish perception verbs at the syntax-semantics interface. In O. Spevak (Ed.), *Noun Valency* (pp. 61-88). Amsterdam/Philadelphia: Benjamins.
- Bekaert, E., & Enghels, R. (2015). On the aspect of deverbal nominals: a corpus study of perception nominalizations in Spanish. *Verbum*, 37(1), 41-68.
- Belletti, A., & Rizzi, L. (1988). Psych-verbs and  $\theta$ -theory. *Natural Language and Linguistic Theory*, 6(3), 291-352.
- Bisetto, A., & Melloni, C. (2005). *Result nominals: a lexical-semantic investigation*. Paper presented at the 5th Mediterranean Morphology Meeting, Fréjus.
- Cano, M. Á., & Jaque Hidalgo, M. (2011). *On the aspectual interpretation of deverbal formations in Spanish: weak and strong patterns of suffixation*. Paper presented at the The 4th Workshop on Nominalizations, University of Stuttgart, Germany.
- Comrie, B., & Thompson, S. A. (1985). Lexical nominalization. In T. Shopen (Ed.), *Language typology and syntactic description* (Vol. 3: Grammatical categories and the lexicon, pp. 349-398). Cambridge: Cambridge University Press.
- Croft, W. (1991). *Syntactic categories and grammatical relations*. Chicago/London: The University of Chicago Press.
- Croft, W. (2001). *Radical construction grammar*. New York: Oxford University Press.
- Dam-Jensen, H. (2008). Determination, nominalisation and conceptual processing. In H. Høeg Müller & A. Klinge (Eds.), *Essays on nominal determination: from morphology to discourse management* (pp. 287-308). Amsterdam/Philadelphia: Benjamins.
- Dik, S. C. (1985). Formal and semantic adjustment of derived constructions. In A. M. Bolkestein, C. de Groot, & J. L. Mackenzie (Eds.), *Predicates and terms in functional grammar* (pp. 1-28). Dordrecht/Cinnaminson: Foris Publications.
- Dubois, J., & Dubois-Charlier, F. (1999). *La dérivation suffixale en français*. Paris: Nathan.

- Escandell Vidal, M. V. (1995). *Los complementos del nombre*. Madrid: Arco/Libros.
- Fábregas, A., & Marín, R. (2012a). The role of Aktionsart in deverbal nouns: state nominalizations across languages. *Journal of Linguistics*, 48, 35-70.
- Fábregas, A., & Marín, R. (2012b). State nouns are kimian states. In F. Irene, S. Lusini, & A. Saab (Eds.), *Romance Languages and Linguistic Theory 2010: Selected papers from 'Going Romance' Leiden 2010* (pp. 41-64), Amsterdam/Philadelphia: Benjamins.
- Fábregas, A., Marín, R., & McNally, L. (2012). From Psych Verbs to Nouns. In V. Demonte & L. McNally (Eds.), *Telicity, change and state: a cross-categorical view of event structure* (pp. 162-185). Oxford: Oxford University Press.
- Fonteyn, L., Heyvaert, L., & Maekelberghe, C. (2015). How do gerunds conceptualize events? A diachronic study. *Cognitive Linguistics*, 26(4), 583-612.
- Fradin, B. (2011). Remarks on state denoting nominalizations. *Recherches linguistiques de Vincennes*, 40, 73-99.
- Giammatteo, M., Albano, H., & Ghio, A. (2005). Clases de predicados y nominalización. In J. Cuartero Otal & G. Wotjak (Eds.), *Algunos problemas específicos de la descripción sintáctico-semántica*. Berlin: Frank und Timme.
- Givón, T. (1982). Evidentiality and epistemic space. *Studies in Language*, 6(1), 23-49.
- Givón, T. (2001a). *Syntax: An introduction* (Vol. I). Amsterdam/Philadelphia: Benjamins.
- Givón, T. (2001b). *Syntax: An introduction* (Vol. II). Amsterdam/Philadelphia: Benjamins.
- Gries, Stefan Th. (2006). Corpus-based methods and cognitive semantics: The many senses of to run. In S. T. Gries & A. Stefanowitsch (Eds.), *Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis* (pp. 57-99). Berlin/New York: Mouton de Gruyter.
- Grimshaw, J. (1990). *Argument structure* (4th print. ed.). Cambridge: MIT press.
- Harley, H. (2009). The morphology of nominalizations and the syntax of vP. In A. Giannakidou & M. Rathert (Eds.), *Quantification, definiteness, and nominalization* (pp. 320-342). New York: Oxford University Press.
- Hartmann, S. (2014). *The diachronic change of German nominalization patterns: an increase in prototypicality*. Paper presented at the Selected papers from the 4th UK Cognitive Linguistics Conference.
- Heyvaert, L. (2008). The periphrastic realization of participants in nominalizations: Semantic and discourse constraints. In B. Lewandowska-Tomaszczyk (Ed.), *Assymetric events* (pp. 245-259). Amsterdam/Philadelphia: Benjamins.
- Hopper, P. J., & Thompson, S. A. (1984). The discourse basis for lexical categories in universal grammar. *Language*, 60(4), 703-752.
- Hopper, P. J., & Thompson, S. A. (1985). The iconicity of the universal categories 'noun' and 'verb'. In J. Haiman (Ed.), *Iconicity in syntax* (pp. 151-183). Amsterdam/Philadelphia: Benjamins.
- Huyghe, R., & Marín, R. (2007). L'héritage aspectuel des noms déverbaux en français et en espagnol. *Faits de Langues*, 30, 265-273.
- IBM Corporation. (1989, 2012). IBM SPSS Decision Trees 21 *IBM SPSS Statistics 21*.
- Jansegers, M., Vanderschueren, C. & Enghels, R. (2015). The polysemy of the Spanish verb *sentir*: a behavioral profile analysis. *Cognitive Linguistics*, 26(3), 381-421.
- Jaque Hidalgo, M. (2012a). Las nominalizaciones deverbales estativas en el diccionario monolingüe español. In A. Nomdedeu Rull, E. Forgas Berdet, & M. Bargalló Escrivà

- (Eds.), *Avances en lexicografía hispánica* (Vol. 2, pp. 519-530): Universitat Rovira i Virgili.
- Jaque Hidalgo, M. (2012b). *Objetos, estados y estructura verbal: El caso de las nominalizaciones no eventivas en español*. Paper presented at the X Congreso Internacional de Lingüística General, Universidad de Zaragoza.
- Koptjevskaja-Tamm, M. (1993). *Nominalizations*. London; New York: Routledge.
- Lakoff, G. & Johnson, M. (2003). *Metaphors we live by*. Updated Edition. Chicago: University of Chicago Press.
- Lehmann, C. (1988). Towards a typology of clause linkage. In J. Haiman & S. A. Thompson (Eds.), *Clause combining in grammar and discourse* (pp. 181-225). Amsterdam/Philadelphia: Benjamins.
- Mackenzie, J. L. (1985). Nominalization and valency reduction. In A. M. Bolkestein, C. de Groot, & J. L. Mackenzie (Eds.), *Predicates and terms in functional grammar* (pp. 29-47). Dordrecht/Cinnaminson: Foris Publications.
- Malchukov, A. L. (2004). *Nominalization / verbalization: constraining a typology of transcategorical operations*. Munich: Lincom.
- Malchukov, A. L. (2006). Constraining nominalization: function/form competition. *Linguistics*, 44(5), 973–1009.
- Marín, R. (2011). Casi todos los predicados psicológicos son estativos. In Á. Carrasco Gutiérrez (Ed.), *Sobre estados y estatividad* (pp. 26-44). Munich: Lincom.
- Marín, R. (2015). Los predicados psicológicos. Debate sobre el estado de la cuestión. In R. Marín (Ed.), *Los predicados psicológicos* (pp. 11-50). Madrid: Visor.
- Marín, R., & McNally, L. (2011). Inchoativity, change of state, and telicity: Evidence from Spanish reflexive psychological verbs. *Natural Language and Linguistic Theory*, 29(2), 467-502.
- Marín, R., & Vilhoing, F. (2012). Nominalisations: nouveaux aspects. In R. Marín & F. Vilhoing (Eds.), *Nominalisations: nouveaux aspects* (pp. 7-19). Villeneuve d'Ascq: Presses du Septentrion.
- Melloni, C. (2010). Action nominals inside: lexical-semantic issues. In M. Rathert & A. Alexiadou (Eds.), *The semantics of nominalizations across languages and frameworks*. Berlin: Mouton de Gruyter.
- Melloni, C. (2011). *Event and result nominals - A morpho-semantic approach*. Bern: Peter Lang.
- Mourelatos, A. P. D. (1978). Events, processes, and states. *LaPh*, 2, 415-434.
- Picallo, M. C. (1991). Nominals and nominalizations in Catalan. *Probus*, 3.3, 279-316.
- Picallo, M. C. (1999). La estructura del sintagma nominal: las nominalizaciones y otros sustantivos con complementos argumentales. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 363-393). Madrid: Espasa.
- Piper, M. E., Loh, W.-Y., Smith, S. S., Japuntich, S. J. & Baker, T. B. (2011). Using decision tree analysis to identify risk factors for relapse to smoking. *Substance use and misuse*, 46(4), 492-510.
- Rauh, G. (2004). Nomina, referentielle Argumente und syntaktische Effekte. In R. Kailuweit & M. Hummel (Eds.), *Semantische Rollen* (pp. 269-298). Tübingen: Gunter Narr Verlag.

- Real Academia Española. Database (CREA) [online]. Corpus de referencia del español actual. <<http://www.rae.es>> [23/03/2016]
- Real Academia Española. (2009a). *Nueva gramática de la lengua española. Morfología. Sintaxis I*. Madrid: Espasa Libros.
- Real Academia Española. (2009b). *Nueva gramática de la lengua española. Sintaxis II*. Madrid: Espasa Libros.
- Scott, A. K. (2010). Accounting for the semantic extension of derived action nouns. *Journal of Linguistics*, 46, 711-734.
- Shibatani, M. (1985). Passives and related constructions. *Language*, 61, 821-848.
- Sleeman, P., & Brito, A. M. (2010a). Aspect and argument structure of deverbal nominalizations: A split vP analysis. In A. Alexiadou & M. Rathert (Eds.), *The syntax of nominalizations across languages and frameworks* (pp. 199-218). Berlin: Mouton de Gruyter.
- Sleeman, P., & Brito, A. M. (2010b). Nominalization, event, aspect and argument structure: a syntactic approach. In M. Duguine, S. Huidobro, & N. Madariaga (Eds.), *Argument structure from a crosslinguistic perspective* (pp. 113-129). Amsterdam/Philadelphia: Benjamins.
- Solstad, T., & Lyngfelt, B. (2006). Perspectives on demotion: Introduction to the volume. In B. Lyngfelt & T. Solstad (Eds.), *Demoting the agent. Passive, middle and other voice phenomena* (pp. 1-20). Amsterdam/Philadelphia: Benjamins.
- Stekauer, P. (1998). *An onomasiological theory of English word-formation*. Amsterdam/Philadelphia: Benjamins.
- Vanderschueren, C. (2013). *Infinitivo y sujeto en portugués y español. Un estudio empírico de los infinitivos adverbiales con sujeto explícito*. Berlin: De Gruyter.
- Vanhoe, H. (2004). *Aspectos de la sintaxis de los verbos psicológicos en español : un análisis léxico funcional*. Frankfurt am Main: Lang.
- Van Valin, R.D.J. (1999). Generalized semantic roles and the syntax-semantics interface. In F. Corblin, C. Dobrovie-Sorin, & J.-M. Marandin (Eds.), *Empirical issues in formal syntax and semantics 2* (pp. 373-389). Den Haag: Thesus.
- Van Valin, R.D.J. (2005). *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Varela, S. (2012). La interacción de las nominalizaciones con la voz, el aspecto y la dimensión temporal. In E. Bernal, C. Sinner, & M. Emsel (Eds.), *Tiempo y espacio en la formación de palabras del español* (pp. 91-106). München: Peniope.