Glossa a journal of general linguistics

Toquero-Pérez, Luis Miguel. 2020. The semantics of Spanish compounding: An analysis of NN compounds in the Parallel Architecture. *Glossa: a journal of general linguistics* 5(1): 41. 1–31. DOI: https://doi.org/10.5334/gjgl.901

RESEARCH

The semantics of Spanish compounding: An analysis of NN compounds in the Parallel Architecture

Luis Miguel Toquero-Pérez
University of Southern California, Los Angeles, CA, US
toquerop@usc.edu

Noun-Noun compounds (NN) are a concatenation of two nouns that function as a single unit both morphosyntactically and semantically. Two main challenges that a study of NN compounding faces are (i) identifying the implied semantic relations that hold between the nouns and (ii) explaining why NN compounds are or not productive (i.e. not very frequent) and creative (i.e. impromptu creation) in certain languages. Jackendoff (2009, 2016) proposed a model which considers the semantic relation between the head and modifier as established by an external function F(X, Y). Since Spanish NN compounds are not creative, the aim is to demonstrate whether the model can be applied to Spanish, and if so, identify the most productive basic functions. This can help us determine whether the lack of creativity of NN compounds in Spanish is due to a restricted set of productive functions. Our hypothesis is that only those relations that are productive are creative and are able to satisfy the Principle of Generalized Modification (Snyder 2012, 2016) operating at the syntax-semantics interface. The study also provides a comparison with the semantic relations attested in Spanish N-de-N constructions (i.e. *buque de guerra* 'ship of war' = *war ship*). Since the latter are very productive and creative, I wonder how well our hypothesis would accommodate to the data. The results indicate that few semantic relations are productive which indicates that only those should be used to interpret novel NN compounds satisfying Generalized Modification. N-de-Ns show a larger amount of functions and although the data falls short to accommodate the hypothesis, impressionistic observations tell us otherwise. Given this asymmetry, our results provide important evidence for blocking effects at the interface: the availability of the NN form blocks the surfacing of the prepositional alternative and viceversa.

Keywords: NN compounds; Spanish; Parallel Architecture; syntax-semantics interface; blocking effects

1 Introduction

This paper focuses on Noun-Noun compounds (henceforth, NN compounds) in Spanish. More specifically, it is concerned with analyzing the semantic relations that hold between the head (which denotes a subtype of the whole compound) and the modifier (which provides specific information about that head) in Spanish NN compounds based on Jackendoff's (2009, 2016) basic functions model.

Accounting for the meaning of compounds poses an interesting problem: on the one hand, the individual meanings of the elements forming the compound contribute to the meaning of the compound; on the other, there is a problem of identifying the implied semantic relations that hold between the two elements. For example, in *pepperoni hat* the individual meaning of *pepperoni* and *hat* is clear, but when they are combined in a compound, it is difficult to pick out a singular meaning since there are several possibilities (e.g. 'hat that has pepperoni', 'hat that is similar to pepperoni', 'hat that is made from pepperoni' etc.).

Identifying these implied semantic relations that exist between the two nouns is the aspect that this paper focuses on. One of the most recent frameworks that has been proposed to account for the semantic relations between the modifier and head in NN compounds is Jackendoff's (2009, 2016) basic functions model within the Parallel Architecture framework. Jackendoff (2009) proposed that the semantic relation between the head and modifier can be established by using an external function F(X, Y) which is a representation of the meaning of the constituents making up the compound. For instance, taking the example *pepperoni hat* provided above, the meaning of 'hat that has peperoni on it' could be represented in the following way: $[\mathrm{HAT}_2{}^{\alpha}; [\mathrm{HAVE}\ (\alpha, \mathrm{PEPPERONI}_1\ \mathrm{ON}\ \alpha)]]$. Jackendoff identified thirteen basic functions to account for the meaning of NN compounds in English. At the same time, he posed questions concerning the cross-linguistic applicability of the model: he wondered whether this model could be applied to other languages and to what extent the basic functions that he proposed were particular to English. Therefore, the aim is to identify the basic functions that hold in Spanish NNs, and answer the following questions: are there any functions that are present in English, but not in Spanish, and vice versa? Also, which functions are more productive and which are less productive? Finding an answer to these questions will not only shed light on Jackendoff's questions regarding the cross-linguistic applicability of his system, but will also allow us to determine whether the lack of creativity of NN compounds in Spanish correlates with their semantic inflexibility (i.e. low productivity of semantic relations). Our hypothesis is that only creative compounds can satisfy the interface condition imposed by the Principle of Generalized Modification (Snyder 2012, 2016) on the interpretation of well-formed compounds. In other words, this predicts that only those functional elements (i.e. the basic functions in our case) providing the semantic relation at the time of semantic composition that have proved creative will satisfy Generalized Modification.

In addition, given that the alternative creative structure for nominal compounding in Spanish is N-de-Ns (i.e. *casa de campo* 'house of country' > *country house*), it will be interesting to apply Jackendoff's model to the semantic interpretation of N-de-N constructions. Since these constructions are very frequent and productive, if the thirteen basic function model can be applied to them, N-de-Ns are expected to show a wider range of semantic relations as opposed to NNs. Moreover, the fact that N-de-Ns constructions can be successfully interpreted by the basic function model would provide evidence that they should be treated as NN compounds semantically, because their meaning can be accounted for in the same way. And, if so, then we would have evidence for the claim by Nicoladis (2002) that in Romance languages prepositions in 'N Prep N' constructions are becoming linking elements and are not true prepositions.

In order to test these hypotheses and answer the questions raised, I have carried out an analysis of Spanish NN compounds and N-de-Ns that are already attested. The reason why this is important is because we first need to determine what relations bundle as creative and which ones do not. And based on these results we should be able to generalize or predict that when coining new compounds, only those relations that have been marked as creative would be instantiated in the meaning of the novel items.

The items under study have been obtained from the *Corpus del Español* (Davies 2002–), the appendix in *Compound Words in Spanish: Theory and History* (Moyna 2011: 303–433), and various written sources such as articles and textbooks. I have created a database of 203 NNs and 203 N-de-Ns and analyzed them according to Jackendoff's functions. In the following sections, I will provide the analysis, present the results, and discuss their implications for the predictions and questions outlined above.

This paper is organized into 7 sections including the introduction. Section 2 provides the theoretical background and an overview of the features of NN compounds in Spanish.

Section 3 is concerned with the semantics of NN compounds and more specifically with Jackendoff's model. Section 4 presents the research questions and hypotheses. Section 5 deals with the analysis of the data collected. Section 6 discusses the main findings and how they relate to the hypotheses and research questions formulated in section 4. Section 7 concludes the paper.

2 NN compounds in English and Spanish: Features and differences

This section discusses the theoretical background necessary for understanding the semantics of NN compounds. It first focuses on outlining the characteristics of Spanish NN compounds by comparing them to NN compounds in English. Then, it provides a thorough overview of different approaches dealing with the meaning of NN compounds, which will allow to explain the hypotheses that the study is based on.

Despite the fact that NN compounds can be found in both English and Spanish, English NN compounds differ from Spanish NN compounds in two main areas: their creativity (Snyder 2012, 2016) and directionality. These syntactic properties have been widely discussed by Bauer (1983), Piera (1995), Snyder (1995, 2001, 2012, 2016) and Liceras and Diaz (2001), among others. A process is creative "if it is available for automatic, impromptu use whenever a new word is needed to fit the occasion" (Snyder 2016: 91). Creativity is often confounded with productivity which is concerned with the number of instances that an element is found in the language or listed in a dictionary. In this paper I will take productivity to mean the frequency with which a particular semantic relation is instantiated or expressed. Directionality covers the syntactic organization in terms of headedness: that is, which of the two nouns making up the compound is the semantic and categorial head and in which position the head is placed.

Contrary to Germanic languages like English, where the formation of endocentric NN compounds is utterly creative,[1] this word formation strategy is not very common in Romance languages like Spanish. This difference led Snyder (1995, 2001) to formulate The Compounding Parameter (i.e. TCP), which Snyder (2012) later redefined. This new formulation is given in (1):

(1) The Compounding Parameter (Snyder 2012)
The language (does/does not) permit Generalized Modification where Generalized Modification refers to a special type of semantic composition, operating at the syntax-semantics interface.

Generalized Modification is a principle that needs to be satisfied at the syntax-semantics interface. In other words, when the compound created in the syntax is transferred to the LF interface, the compound has to be able to semantically compose to render a suitable interpretation (William Snyder p.c.). There are two requirements: one of them is a hyponymy relation[2] and the other is an available semantic relation established between the modifier and the head (by means of a functional head, e.g. Fábregas 2008, Di Sciullo 2009). The prediction this makes is twofold: if an [XY] is neither a type of X nor a type of Y the semantic composition will crash at LF; besides, if the semantic relation involved in the semantic composition of the compound is not active (i.e. "switched off" if we assume the availability of semantic relations is parametric), the composition will also crash.

[1] Downing (1977) highlights the repeated coining of compounds *ad hoc* such as her famous *apple juice seat* which is supposed to mean 'a seat in front of which the apple juice has been placed'.

[2] Generalized Modification is defined by Snyder (2012: 10) as the following: "If α and β are syntactic sisters under the node γ, where α is the head of γ, and if α denotes a kind, then interpret γ semantically as a subtype of α's kind that stands in a pragmatically suitable relation to the denotation of β".

Thus, Snyder (2012) divides languages as [+TCP] or [–TCP] regarding whether compounding is creative or not respectively. Moreover, Snyder (2012) acknowledges that the semantic content expressed in his definition of generalized modification as "stands in a pragmatically suitable relation to" (2012: 11) is vague because the relation that holds between the modifier and the head of an NN compound in [+TCP] languages is enormously flexible. Snyder (2012), however, does not propose a way to clarify this vagueness when determining the implied semantic relation. A possible way to account for the semantic relations that hold in NN compounds is applying Jackendoff's (2009, 2016) basic function model that proposes thirteen basic functions to interpret the meaning of NN compounds. This model has been successfully applied to languages like English and Swedish which are [+TCP] and French which is [–TCP]. The former exhibit all or almost all of the functions Jackendoff proposes where the latter displays very few functions. Therefore, a plausible hypothesis is that [+TCP] languages allow a wide amount of semantic relations and therefore cause compounding to be a very creative process while [–TCP] do not. A question that is raised in this paper is whether this hypothesis is true for Spanish.

In addition to Snyder's TCP, Piera's (1995) Word Marker (i.e. WM) should be considered when analyzing compounds in languages such as English and Spanish. Piera acknowledged that English and Spanish NN compounds differ in two essential ways: English compounds are right-headed, whereas Spanish ones are left-headed (2) (head in bold); and English compounds are recursive, while Spanish ones are not (3):

(2) a. spider **man**
  b. **hombre** araña
   man spider

(3) a. movie spider man
  b. *hombre araña película
   man spider movie

Examples in (2) demonstrate that while the head of the English compound in (2a) is the right-hand element (e.g. man), the head of the Spanish compound is placed on the left (i.e. hombre). This means that directionality in Spanish is reversed compared to English. In terms of headedness, we can adopt Arnaud and Renner (2014: 2) terminology. In both English and Spanish the head corresponds to the semantic, categorial and morphological head: the semantic head is the entity that denotes the hyperonym (e.g. a "spider man" is type of man; a "hombre araña" is a type of hombre); the categorial head is the constituent whose word class determines the whole unit's word category (in this case since both constituents are nouns, there is no controversy here); and the morphological head is the component undergoing inflection such as gender or number (i.e. spider men/woman; hombre-s/mujer araña). The modifier element specifies the meaning of the head by means of an implied semantic relation: so a *spider man* is 'a man that has spider features' and bears the basic function HAVE according to Jackendoff (2016).

Piera addresses the reason why recursive NN compounds in Spanish are ungrammatical: Spanish nouns have a Word Marker (i.e. WM) at the end, which may or may not be phonetically realized, and which bears grammatical information such as number or gender. English nouns, on the other hand, do not have this WM. This difference is illustrated in (4):

(4) a. [[niñ-] **o**]
   kid-MASC.SG
   'kid'
  b. [kid]

The ‘**o**’ in (4a) is the WM that distinguishes Spanish nouns from English ones. The presence of this WM forces a Spanish noun to have a double bracket to its left, which prevents the adjunction of another noun to it, given Piera’s (1995) Double Bracket Restriction, according to which “a double bracket at the edge of a word blocks adjunction of a word” (Piera 1995: 6) as shown in example (5):

(5) a. [[niñ-] WM] [[espí-] WM]]
kid- spy
‘spy kid’
b. *[[niñ-] WM] [[espí-] WM]] [[policí-] WM]]
kid spy pólice
‘spy police kid’
c. *[[hombre] WM] [[arañ-] WM]] [[películ-] WM]]
man spider movie
‘movie spider man’

Although Piera’s Double Bracket Restriction seems to account for the directionality of Spanish compounds as well as for the non-existence of recursive NN compounds in Spanish, it does not explain why compounds such as **mesa ventana* (lit. “table window” > *window table*) should be ungrammatical. In fact, this compound shares the same internal structure as (5a). This can be seen in (6):

(6) a. [[niñ-] WM] [[espí-] WM]]
kid-WM spy-WM
‘spy kid’
b. *[[mes-] WM] [[ventan-] WM]]
table-WM window-WM
‘window table’

One reasonable hypothesis that can be proposed to account for the question why Spanish does not have NN compounds like *mesa ventana* is that at the syntax-semantics interface Generalized Modification is not satisfied: Spanish may not allow the realization of the semantic relation that should hold between the head (e.g. mesa) and the modifier (e.g. ventana) in such compounds leading to a crash at LF. That is, the ungrammaticality of such compounds may be due to semantic constraints rather than syntactic ones. More precisely, the number of semantic relations that underlies the meaning of NN compounds may be limited in Spanish, as opposed to languages like English. Thus, one reason for the restricted creativity of Spanish NN compounds may be their “semantic inflexibility”: the amount of productive semantic relations available at the syntax-semantics interface may be limited. One of the goals of this study is, precisely, to investigate whether NN compounds in Spanish do or do not allow the realization of all semantic relations that are attested in NN compounds in languages like English, and the frequency with which those relations are found, which may be two important factors causing the lack of creativity of such compounds.

3 The semantics of NN compounds

As it was mentioned at the beginning of section 2, the central question that concerns the semantics of compounds is how the meaning of the compounds relates to the meaning of the parts. If we take a compound such as *mushroom soup*, we can break down the semantic content of the nouns into two parts. On the one hand, there is a problem of how the individual meaning of the elements forming the compound (mushroom and soup) contributes to the meaning of the compound; on the other, there is a problem of identifying the implied semantic relation that holds between the two elements: is it a resemblance relation (soup *similar to* mushroom)? Or is it a possessive one (soup *has* mushrooms)? or is it a compositional one (soup *made from* mushrooms)? Determining these implied semantic relations is one of the challenges that studies concerned with the semantics of compounding try to account for.

There have been many attempts to capture the implied semantic relation between the two nouns making up the compound. For instance, Lees' (1960, 1970) attempt was based on proposing a set of underlying verbs which were deleted in the course of the derivation from a sentence to an NN. Downing (1977), however, proposed a finite set of semantic relations that can be used to account for the implied meaning of the compound and attempted to identify the semantic relations available between the head and modifier as well as any possible semantic constraints in the formation of NN compounds. Levi (1978) proposed a reduced number of predicates that can be used to establish the meaning of compounds as a whole, but as it has been argued her theory has some disadvantages. Allen (1978) introduced the importance of the lexicon: the meaning of an NN compound depends on the association that can be made between the semantic content of modifier and head which is determined by the information of the lexical entries encoded in the lexicon. However, due to the inability of these models to answer questions such as how the meaning of compounds is related to the meaning of the whole and what the status of compounds in the lexicon is, new theories and proposals have been proposed in recent years from various perspectives. In this paper, I focus on one particular framework in which the semantics of compounds is the central issue: Jackendoff's (2009, 2016) basic functions model within the Parallel Architecture framework.

3.1 A new framework for compounding: Jackendoff's (2009, 2016) basic functions

Jackendoff's model for the interpretation of compounds is base on the framework of Parallel Architecture which assumes a tripartite system based on the premise that phonology, syntax and semantics are independent generative modules. The semantic structures are not derived from the combination of syntactic units; instead, they are made of semantic units that have their own characteristics and do not have a one-to-one correspondence with syntactic categories. Therefore, semantics is a generative module of its own that is connected to syntax, and to phonology, by means of interface rules.[3] Contrary to previous accounts, instead of using paraphrases to recover the meaning of a compound, he assumes that its meaning "is a function of the meanings of its constituents": $F(X_1\ Y_2)$ (2009: 115). The question is how this function can be applied to render the meaning link between N_1 and N_2. The realization of the function (F) follows his premise that an appropriate semantic analysis should consider the individual meaning of the units. Moreover, he remarks a problem with paraphrases proposed in the 1960s and 70s: the controversy of the paraphrase lies in the fact that various perspectives can be adopted when choosing the implied predicate that links two nouns, as in example (7):

(7) a. ticket booth: booth where tickets are bought.
b. ticket booth: booth where tickets are sold.

Here in (7) we see two paraphrases of *ticket booth* that have the same event schema but involve different points of view: "buy" places the focus on the recipient of the tickets whereas "sell" places it in the owner. In order to avoid this problem, Jacekndoff considers the function as a representation of an event schema that lacks perspective or focus.

[3] The reader is referred to Jackendoff (2009) and (2016) for a more detailed overview of the Parallel Architecture model. Only the relevant details to compounding will be addressed here.

Jackendoff's model is not concerned with assigning particular and definite meaning to a compound, but to account for the possible meanings a compound might have. That is, when we encounter a context-free novel compound, we use this function $F(X_1\ Y_2)$ to connect the semantic characteristics of N_1 to those of N_2. Crucially, all possible combinations that we may think of are stored in our memory as part of the meaning of a compound, though only one becomes more salient and suitable depending on the situation. That is why he claims that NN compounds are promiscuous: a compound denoting the same entity in the real world can have several possible interpretations. This differs from Levi's (1978) notion of ambiguity because ambiguity according to Jackendoff refers to a compound that can denote different entities and therefore has a different meaning associated with each entity. This difference can be illustrated in (8):

(8) a. chocolate box: 'box that has chocolate/that is composed of chocolate/that is made from chocolate'.
b. baseball: 'a game played between two teams of nine players each in a diamond-shape field' or 'a ball used in the game of baseball'.

The compound *chocolate box* in (8a) can be interpreted in various ways depending on the context, but the denotation picked out by the compound is the same in all of them. Therefore, it is promiscuous. On the other hand, *baseball* in (8b) has two different interpretations and each of these interpretations picks out a particular denotation: the game or the ball. Thus, it is ambiguous.

In addition to this, Jackendoff adds to his function three components that are also associated with the semantics of nominals: profiling, action modality and cocomposition. Profiling means picking out an individual that is involved in an event and designating it as the one being referred to. How profiling works is illustrated in (9) for the noun *writer*, which profiles the agent of the action of writing.

(9) a. WRITE (A, B) = "A writes B".
b. λx[WRITE (x, INDEF)] = "individual who writes something"
c. [PERSON$^{\alpha}$; [WRITE (α, INDEF)]] = "a person α such that α writes something"

(9a) implies that the action of writing involves an agent A and a theme B. The agent is profiled by the noun *writer*. This is seen in (9b) where the *x* stands for the individual to be denoted. The meaning of the expression is seen in (9c) where the profiled individual is designated PERSON and is indexed to the modifier which is separated from the profiled argument with a semicolon. What makes this a well-formed expression is the fact that the modifier contains a variable α which is bound by the superscript on PERSON. Thus, as Jackendoff (2016: 21) points out, "profiling an argument of a function involves binding it to something outside the function". Profiling creates the semantic counterpart of a relative clause in syntax.

Action modalities are variant interpretations seen in agentive nominals. According to Busa (1997) and as quoted in Jackendoff (2009: 119), agentive nominals are "characters individuated by their actions". Busa (1997) points out that an agentive nominal like violinist can be ambiguous between an occupation, a habit, or an ability, as illustrated by the following examples in (10) taken from Jackendoff (2010: 17):

(10) a. She is a **violinist** in the Pittsburgh symphony but hasn't played since they went on strike.
b. She is an occasional **violinist**
c. She's a good **violinist**, but hasn't played since she sold her violin ten years ago.

In (10a) "violinist" indicates an occupation because that is how she makes a living even though she is now on strike. (10b) illustrates the habitual action of playing the violin. In (10c) she has lost the habit of playing since the sale of her violin, but still indicates that she has the ability to play.

Furthermore, Jackendoff adopts Millikan's (1984) proper function (i.e. PF) as an important action modality. PF means that something has been invented to, or is intended to perform a particular function. Only artifacts (e.g. a key), parts (e.g. the handle of a knife) and objects predetermined to become something else (e.g. a seed will eventually become a plant) have a PF. So, as an illustration we can look at (11):

(11) a. key = [KEY $^{\alpha}$; [PF (OPEN (α, INDEF)]] 'a key has a proper function which is to open something'
b. mail$_1$ man$_2$ = [MAN$_1$ $^{\alpha}$; [OCC (DELIVER (α, MAIL$_2$))]] 'man has an occupation which is to deliver the mail'

I have chosen these two examples to show the difference in the action modalities of the nouns *key* and *mailman*. The former in (11a) is an artifact and thus has a PF which is to open something; this explains why the label INDEF (i.e. indefinite) is used. The latter in (11b), on the contrary, has an occupation reading (OCC) which is to deliver the mail.

The third component is cocomposition. Sometimes nouns involve interpretations that require an activity. For instance, in 'Peter finished the movie', *the movie* specifies an action of watching which is not semantically expressed in the sentence. Thus, cocomposition is concerned with filling that internal structure of the meaning with the particular PF specified by the noun. In other words, the idea is to add a function which is not overtly expressed to build well-formed semantic links. This is portrayed more clearly in (12), where the meaning of 'Peter finished the movie' is presented:

(12) a. Peter$_1$ finished$_2$ the movie$_3$ = FINISH$_2$ (PETER $^{\alpha}$ $_1$, F (α, MOVIE$_3$)
b. Peter$_1$ finished$_2$ the movie$_3$ = FINISH$_2$ (PETER $^{\alpha}$ $_1$, WATCH (α, [MOVIE$^{\beta}$; PF (WATCH (PERSON, β))]$_3$))

The movie implies an activity which is marked by the F in (12a). This F coerces the semantic interpretation: the first argument of F is bound to *Peter* to indicate that it is him preforming the action specified by the F and the second is bound to the movie which is what undergoes the action. In (12b) the F has been replaced by *watch* because this is what the PF of the noun *movie* specifies for F.

That said, Jackendoff (2009) acknowledges that there are two ways in which the semantic relation linking N$_1$ and N$_2$ can be established: argument schema and modifier schema. The argument schema does not require an external function because the function is already expressed by the head of the compound. This head implies an explicit predicate such as "drive" in *bus driver*. Therefore, the head already specifies a function (i.e. drive) that is applied to the modifier (bus) acting as its argument (e.g. '-er who drives a bus'). However, in the case of the modifier schema, which applies to non-deverbal compounds, the function plays a fundamental role. The formulation of the modifier schema has been taken from Jackendoff (2009: 122) and it is shown in (13) below:

(13) Modifier schema:
a. [N$_1$ N$_2$] = [Y $_2$ $^{\alpha}$; [F (…, X$_1$…, α,…)]]
b. an N2 such that F is true of N1 and N2

Although this is the schema underlying every subordinate ground NN compound, the number of possibilities for the function (i.e. F) makes us return to the problem encountered by previous scholars regarding the finite or infinite set of relations that are attested. Though Jackendoff assumes that the system generates an infinite number of possibilities to realize the function, he proposes that the system is made of thirteen basic functions that can undergo profiling, exhibit action modalities and cocomposition to create more or less complex realizations of the function. The list of Jackendoff's basic functions with the schemata he provides for them and an example for each is presented in Table 1.

I have included the (+) symbol to indicate that the function is reversible and enables more than one possible reading. For instance, in the case of MAKE, we can have 'an N_2 that is made by N_1' (*spider poison*) and 'an N_2 that makes N_1' (*silkworm*). One of the weaknesses noted by Jackendoff himself is the impossibility to tell the difference between BE (X, AT/IN/ON/ Y) and PART OF (X,Y) because *garlic bread* can suit either of these (bread that has garlic on it or bread that is in part composed of garlic). This also occurs with other relations: primarily, COMP (X,Y) and MADE (X, FROM Y) though he argues that MADE (X, FROM Y) is the function that applies when the entity Y can be no longer perceived. Furthermore, in the case of the schemata for HAVE, Jackendoff leaves open the possibility that cross-linguistic studies may uncover new ones. The only function which is not as basic as the others is PROTECT because it involves an implied external argument (Z). It is the first time that such an implied relation is considered in the literature on compounding.

In addition to his thirteen basic functions, Jackendoff claims that in order to render a more accurate and faithful interpretation of those compounds whose meaning is more complex, we need to use information from the internal semantic structure of the two nouns creating the compound. He argues that we can use the PF as if it were another basic function because the PF refers to the intended purpose that an element has in a particular

Table 1: Jackendoff's thirteen basic funtions taken from Jackendoff (2016: 27–30).

Function	Schema	Example
CLASSIFY (X,Y)	$[Y_2{}^{\alpha}; [\text{CLASSIFY } (X_1(\alpha))]]$	Beta cell
BE (Y,X)	$[Y_2{}^{\alpha}; [\text{BE } (\alpha, X_1)]]$	Singer-songwriter
SIMILAR (X,Y)	$[Y_2{}^{\alpha}; [\text{SAME/SIMILAR } (\alpha, X_1)]]$	Crocodile pin
KIND (X,Y) (+)	$[Y_2{}^{\alpha}; [\text{KIND } (\alpha, X_1)]]$ $[Y_2{}^{\alpha}; [\text{KIND } (X_1,\alpha)]]$	Seal pup Pine tree
BE (X, AT/IN/ON/ Y) (+)	$[Y_2{}^{\alpha}; [\text{BE } (\alpha, \text{AT/IN/ON } X_1)]]$ $[Y_2{}^{\alpha}; [\text{BE } (X_1, \text{AT/IN/ON } \alpha,)]]$ $[Y_2{}^{\alpha}; [\text{BE}_{TEM} (\alpha, \text{AT } X_1)]]$	Lake house Inkpad November rain
COMP (X,Y) (+)	$[Y_2{}^{\alpha}; [\text{COMP } (\alpha, X_1)]]$ $[Y_2{}^{\alpha}; [\text{COMP } (X_1, \alpha,)]]$	Meat ball Sheet metal
MADE (X, FROM Y) (+)	$[Y_2{}^{\alpha}; [\text{MADE } (\alpha, \text{FROM } X_1)]]$ $[Y_2{}^{\alpha}; [\text{MAKE } (X_1, \text{FROM } \alpha)]]$	Coconut oil Rubber tree
PART OF (X,Y) (+)	$[Y_2{}^{\alpha}; [\text{PART } (\alpha, X_1)]]$ $[Y_2{}^{\alpha}; [\text{PART } (X_1,\alpha)]]$ $[Y_2{}^{\alpha}; [\text{PART } (X_1,\alpha)]]$ (only partially)	Suit pants Wheelchair Cinnamon roll
CAUSE (X,Y)	$[Y_2{}^{\alpha}; [\text{CAUSE } (X_1,\alpha)]]$	Diaper rash
MAKE (X,Y) (+)	$[Y_2{}^{\alpha}; [\text{MAKE } (X_1,\alpha)]]$ $[Y_2{}^{\alpha}; [\text{MAKE } (\alpha, X_1)]]$	Spider poison Silkworm
X serves as Y	$[Y_2{}^{\alpha}; [\text{BE } (\text{PF } (\alpha), \text{PF } (X_1))]]$	Guard dog
HAVE (X,Y) (+)	$[Y_2{}^{\alpha}; [\text{HAVE } (X_1,\alpha)]]$ $[Y_2{}^{\alpha}; [\text{HAVE } (\alpha, X_1)]]$	Gangster money Glamour girl
PROTECT (X,Y FROM Z)	$[Y_2{}^{\alpha}; [\text{PROTECT } (\alpha, X_1, \text{FROM } Z)]]$ $[Y_2{}^{\alpha}; [\text{PROTECT } (\alpha, Z, \text{FROM } X_1)]]$	Chastity belt Flea collar

situation. Thus, examples such as *eye doctor* or *bird egg* that cannot be analyzed under any of the basic functions proposed could be interpreted as it is shown in (14):

(14) a. eye_1 doctor_2 = [$\text{DOCTOR}_2{}^{\alpha}$; [OCC (EXAMINE ($\alpha$, EYES_1))]] 'doctor has an occupation which is to examine eyes'
b. bird_1 egg_2 = [$\text{EGG}_2{}^{\alpha}$; [PF (BECOME (α, BIRD_2))]] 'egg whose proper function is to become a bird'

In (14a), *doctor* has a PF of occupation because it is his job to examine the eyes of patients. In the case of (14b), although it is true that an alternative reading could be possible (such that 'egg that lays the bird'), the ultimate state of an egg is to be transformed into a bird. Therefore, it bears the PF of BECOME. Using the PF and the variants it entails (i.e. occupation, habit, ability), Jackendoff attempts to cover any possible gaps left by his thirteen basic function's framework and he himself acknowledges that "if [basic functions] were all there were to filling out the interpretation of compounds, the number of the possible relations in compounds would be thirteen, [...] clearly not enough" (Jackendoff 2016: 31).

Jackendoff's (2009, 2016) fine-grained analysis provides us with a consistent framework for establishing the meaning relations between the two members of a compound. His basic functions attempt to provide the skeleton for the interpretation which is then filled with semantic material of each individual noun. That is, with this model, framed in his theory of semantics, Jackendoff develops an elaborate conceptual system that allows to zoom in the most covert features of nouns. Therefore, even though he proposes a finite set of basic functions, their reversibility and interaction with the semantic components of nominals (e.g. profiling, action modality and cocomposition) cause an increase in the complexity and repertory of possible semantic relations between N_1 and N_2. A question he proposes is whether this system of representation and analysis can be applied to other languages or whether it is exclusive to English. This is the main research question I will return to in the hypothesis section.

4 Research questions and hypotheses

As mentioned above, the main goal of this paper is to identify the semantic relations that hold between the head and the modifier in Spanish NN compounds, based on Jackendoff's (2009) basic functions model. By comparing the results obtained from Spanish with Jackendoff's results from English, I hope to shed some light on the questions Jackendoff raised regarding the full set of basic functions and the universality of the model.[4]

A major question that this paper is concerned with is whether the lack of creativity can be understood in terms of few productive semantic relations (i.e. what was refered previously as "semantic inflexibility"). For instance, it has been argued by Marqueta-García (2017) that all compounds in Spanish can be reduced to a single semantic relation which she calls *Identifying*. Her main argument is that they can be paraphrased as *como* ('like'): *pez globo* ('globe fish') 'pez que es redondo como un globo' ('fish that is round like a globe'). If this is the case, then the *Identifying* relation would not be creative but it would be productive (i.e. all NN compounds in Spanish show this relation). So, in principle, one could then assume that [–productivity] and [–creativity] are not related in [–TCP] languages like Spanish and hypothesize the following:

(15) **Hypothesis A:** NN compounds allow few productive semantic relations, but this does not guarantee creativity.

[4] These questions have already been investigated by Rosenberg (2013) in her comparison of Swedish NN compounds and their French counterparts who found that only two basic functions are available for French NN compounds: CLASSIFY and BE (Y,X). Moreover, neither PROTECT nor KIND are attested in the data at all. These results are interesting, because they can be taken to indicate that this extreme semantic inflexibility of French NN compounds is responsible for their peripheral status.

Hypothesis A predicts that (i) there will be a small (if any) amount of semantic relations that will be instantiated with a certain frequency; and (ii) they will allow the existence of NN compounds in Spanish. However, the fact that some of these relations are productive should not entail creativity. In other words, this hypothesis predicts that few or (maybe) no semantic relations will be frequent and regularly attested as opposed to English, where it is argued that they are extremely productive and also creative. Nevertheless, this hypothesis by itself does not explain why compounding is not a creative process, potentially leading to the ungrammaticality of NN compounds, which is one of the ultimate goals of this paper. In fact, it would only predict that certain relations are repeatedly found in the data. That is why, I propose a competing hypothesis in (16):

(16) **Hypothesis B:** Only semantic relations that are productive will be creative.

(16) is along the lines of the TCP: one would expect few semantic relations to be productive. If this prediction is borne out, one could then blame the non-creativity of Spanish compounds on this low productivity of semantic relations. As a result, those relations that are productive, and thus creative, will satisfy Generalized Modification at the syntax-semantics interface; while unproductive semantic relations violate Generalized Modification in NN compounds. To put it differently, one needs to think about Generalized Modification as a kind of filter that can only see productive relations. This could provide a potential explanation for ungrammaticality of novel cases like *mesa ventana* ('window table').[5]

I need to clarify that, even though the semantic relations that will be underlying the two nouns may be limited, they may not be as restricted as suggested by Marqueta-García (2017). In fact, reducing all semantic relations to one may be too extreme given the existence of examples like *gas ciudad* ('city gas'), or *bandera pirata* ('pirate flag') which are not suitable to fall under the underlying *Identifying* relation. What is more, the fact that a compound cannot be paraphrased by a particular predicate is one of the problems early transformational approaches like Levy (1978) or Lees (1960, 1970) fell into, especially taking into account that NN compounds can be promiscuous.

This question is very much related to another aim of the study, which is the establishment of a hierarchy of semantic relations expressed in Spanish NN compounds according to their productivity. That is, the idea is to find out whether there are some relations that are more productive than others, and whether there are utterly unproductive ones. Taking into account the unproductivity of NN compounds in Spanish, I believe that few semantic relations will be favored. If this is the case, it could be assumed that the more productive a relation is, the more creative it will be. Also, given what Rosenberg (2013) found in her study regarding PROTECT and KIND (they were not attested either in French structures or Swedish NNs, see footnote 3 for details), I expect that these two semantic relations will be extremely peripheral in the data. That said, this predicts that those semantic relations that belong to the unproductive block of the hierarchy will not satisfy the interpretive conditions to be met at the syntax-semantics interface and will lead to a crash.

In addition to this, I consider it important to apply Jackendoff's model to Spanish N-de-N constructions as well, because they are very creative and productive in Spanish and are regarded as alternative constructions to NNs. This is why the semantic relations present in Spanish NN compounds (i.e. *buque escuela* 'ship school' = *training ship*) will be compared to those that could be attested in Spanish N-de-N constructions (i.e. *buque de guerra* 'ship of war' = *war ship*). Since the latter constructions are very creative in the language, I believe that they will exhibit a wider range of semantic relations. Thus, they

[5] This hypothesis does not make any predictions with respect to lexicalized compounds that bear an unproductive semantic relation and still satisfy Generalized Modification.

are expected to behave like the mirror image of compounds. That is, I suppose that most of the basic functions proposed by Jackendoff will be applicable to them. If this turns out to be the case, the results will then strongly support the claim by Nicoladis (2002) about prepositions in the equivalent French structures (i.e. prepositions in these structures are becoming linking elements: they do not bear any meaning, but act as glue sticking the two nouns together), as it may show a pattern taking place in Romance languages. Such results would also confirm that N-de-Ns should be considered as “actual” compounds with the same status as NN compounds given their conceptual unity (Kornfeld 2009: 442). As a result, if these are compounds, they should also be subject to Generalized Modification at the syntax-semantics interface. A question that arises now is whether Hypothesis B is also applicable to them.

Overall, the results obtained may shed some light on the reasons why NN compounds are so peripheral in Spanish. In other words, are NN compounds uncommon in Spanish because they are semantically inflexible due to the limited set of productive semantic relations that hold between the head and modifier?

In order to provide an answer to the questions and the hypotheses formulated here, I have carried out an analysis of Spanish NN compounds that have been obtained from the *Corpus del Español* (Davies 2002–), the appendix in *Compound Words in Spanish: Theory and History* (Moyna 2011: 303–433), and various written sources such as journal articles and textbooks. The analysis is presented in the next section.

5 Method and analysis

This section describes the methodology followed to collect NN compounds and N-de-Ns, focusing on the description of the sources and tools used to collect the data, as well as the procedure to classify the data. Then I turn to the analysis of the data in an attempt to answer the research questions formulated in section 4.

5.1 The database: Sources and classification

In order to carry out this study, I have collected the items to be analyzed from different sources: previous literature (included in the references section), the appendix in *Compound Words in Spanish: Theory and History* (Moyna 2011: 303–433) and the *Corpus del Español* (Davies 2002–). The most productive sources were the appendix in *Compound Words in Spanish: Theory and History* (Moyna 2011: 303–433) and the *Corpus del Español* (Davies 2002–). The former contained an exhaustive list of NN compounds through history, but I focused primarily on the most recent centuries, starting in the 18^{th} and moving into the 20^{th} and extracting only those which are still in use nowadays.[6] The latter is a web-based Spanish corpus of two billion words. In order to obtain the target constructions, the corpus was searched by using search terms such as “Noun Noun” or a specific word in place of a “Noun” term (e.g., “Noun *coche*” or “*coche* Noun”). It is also important to mention that N-de-Ns that had a determiner following the preposition were not selected because the appearance of the determiner indicates that the expression is a Noun Phrase rather than a compound: *casa de campo* (‘country house’) is any house in the countryside whereas *casa del campo* (‘house in the country’) refers to a known house in a particular location of the country. I also checked the context for all NNs and N-de-Ns collected to avoid concatenations of two N + N or N + PP sequences belonging to separate constituents.

Once the data were collected, a database in Excel with two spreadsheets was created, one to record NN compounds and the other to record N-de-Ns constructions. Each spreadsheet

[6] Several informants were consulted about the current usage status of these compounds before they were included in the database.

contained a total of 203 target constructions that were classified according to Jackendoff's thirteen basic functions (2009, 2016). To these thirteen functions, PF and ARGUMENT SCHEMA were added because they were necessary to account for all the meanings of the NN compounds and N-de-Ns constructions collected. In addition to classifying the NN and N-de-Ns according to their functions, a paraphrase of their meaning matching the function selected was provided: for example, *coche plátano*: 'coche que es similar a un plátano' ('banana car: car that is similar to a banana'). If an item had more than one possible reading, each meaning was assigned to the appropriate function and paraphrased.[7] Furthermore, the saliency of reading was marked by color coding.

Once the data were correctly coded in the database, I proceeded with the analysis of the data. The results of the analysis are presented in in the next section.

5.2 *Data analysis and discussion*

5.2.1 Jackendoff's (2009: 128) research questions

As mentioned in the previous section, Jackendoff (2009) raised the question of a cross-linguistic applicability of the model and the question of what a full set of basic functions is cross-linguistically. In relation to these questions, we can see from Table 2 that Spanish does not realize all of the basic functions that Jackendoff found in English.

Table 2: Functions within Spanish NN compounds in a hierarchical order.

Semantic relation	Spanish NN	English translation	Type
SIMILAR 'N_1 similar to N_2'	Cabeza buque Coche oruga	'big head' 'caterpillar car'	76
PF 'N_1 whose PF is to function as N_2'	Hombre orquesta Lápiz láser	'one-man band' 'laser pencil'	38
X serves as Y 'N_1 whose function is to function in the proper function of N_2'	Canción protesta Perro guardián	'protest song' 'guard dog'	33
BE (AT/IN/ON/) 'N_1 that is located at/in/on N_2'; 'N_1 with N_2 at/in/on it'; N_1 that takes place at time N_2'; 'N_1 whose PF is to be at/in/on N_2'; 'N_1 whose PF is to have N_2 in/on it'	Balón reloj Gas ciudad	'clock ball' 'town gas'	28
HAVE 'N_1 that has (an) N_2'; 'N_1 that N_2 has'	Bandera pirata Carril bus	'pirate flag' 'bus lane'	25
COMP 'N_1 composed of N_2'	Avión cohete Barco tanque	'rocket airplane' 'tank ship'	15
PART 'N_1 is a part of N_2'	Aguanieve Mochila arnés	'sleet' 'harness backpack	15
CAUSE 'N_1 that is caused by N_2'	Efecto placebo Guerra narco	'placebo effect' 'narc war'	14
BE 'N_1 is an N_2'	Abeja reina Perro vampiro	'queen bee' 'vampire dog'	7
ARGUMENT SCHEMA 'an N_1 of/by N_2'	Abordaje pirata Ataque pirata	'pirate boarding' 'pirate attack'	5
MAKE 'N_1 made by N_2'; 'N_1 makes N_2'	Fiesta pirata telaraña	'pirate party' 'cobweb'	5
Types of Spanish NNs			261

[7] The compounds collected were classified into the different functions according to the meaning provided by the context in the corpus search as well as judgments obtained from informants to avoid any biases.

Table 2 shows all the basic functions available for Spanish NN compounds in the collected data, together with some examples of compounds illustrating these functions and their translations in English. The column *Type* presents the number of compounds under each function. The functions are ordered according to this number. Although there was a total of 203 NN compounds in the database, due to promiscuity some compounds allowed more than one reading. Thus, the column *Type* is, more specifically, making a reference to the total number of occurrences of basic functions considering all the possible meanings for the compounds collected. This is where the number 261 comes from.

As it can be seen from Table 2, there are four functions that are not present at all in the Spanish NN compounds collected: CLASSIFY, KIND, MADE FROM and PROTECT. If we compare the availability of these functions cross-linguistically, we will see the following: KIND and PROTECT are not found in French and Swedish either, as discussed by Rosenberg (2013);[8] while CLASSIFY is one of the two basic functions attested in French NN compounds and MADE FROM is attested in Swedish NN compounds. Thus, from the data it can be deduced that the basic functions KIND and PROTECT are particular to English, as there is no other language considered in this study and Rosenberg's (i.e. Spanish, French and Swedish) that has realizations of these functions. This means that the full set of basic relations as developed by Jackendoff (2009, 2016) is particular to English. Jackendoff's model can, however, be successfully applied to other [+] and [–] TCP languages, but with certain parametric variants. That is, Spanish, French and Swedish have a reduced version of this set: the proposed version for English (and those languages that behave alike) can be taken as the superset option, whereas its variants can be regarded as the subset option.

A reviewer wonders how the concept of 'kind' is treated: on the one hand, the basic function KIND, as proposed by Jackendoff is absent; on the other, the notion of 'kind', taken from Snyder (2012) to establish a hyponymy relation, is clearly attested which the reviewer believes could be captured by SIMILAR.[9] The reviewer's main concern is how these notions receive different treatments in the classification and analysis of the data. I believe these have to be kept as separate notions. Snyder's formulation of the term Generalized Modification (footnote 2) refers to the widespread and consensual intuition that the meaning of an endocentric compound is a hyponym of the head. In other words, the meaning of [XY], where X is the head, is a kind/type of X not Y. For instance, a *computer screen* is a kind of *screen* not a type of *computer*. But this cannot be treated like a primitive semantic function because it applies to every endocentric NN compound cross-linguistically. In fact, this description is simply a formal way of stating the semantic criterion to identify what the head of the compound is: the head is the entity that provides a superset-subset relation. As to SIMILAR, I cannot think of a reason to establish a parallelism between this function and KIND since SIMILAR establishes a metaphorical or

[8] Rosenberg (2013) found comparatively few NN compounds in French. Moreover, these NN compounds expressed only two functions: BE and CALSSIFY. However, in my Spanish data, these functions were either peripheral (i.e. BE) or not even attested at all (i.e. CLASSIFY). Besides, most basic functions proposed by Jackendoff are available in Spanish. Thus, it may be the case that Spanish and French behave differently in terms of productivity with respect to TCP. Interestingly, Rosenberg also accounts for the meaning of French N-Prep-N compounds using this model and argues that more semantic relations are found in these constructions.

[9] The same reviewer is also concerned with the overlap between locative BE and PF; and wonders whether *balón reloj* ('clock ball') should be classified better under PF and whether BE should have more type realizations than PF given that the former is more basic than the latter. I agree with this reviewer that locative BE is more basic than PF and that they overlap. But the example mentioned should not be suitable to be classified under PF alone: 'a ball whose PF(purpose/goal/intention) is to have a clock'. The paraphrase sounds odd to me and to my informants. The most natural meaning for this compound is to indicate a locative relation between the head *balón* and the modifier *reloj* to render the meaning 'a ball that has a clock in it (inside)'. Moreover, this meaning matched the context found in the corpus and the real world images found online. The same was corroborated for the other examples that exhibited this subtle line between the two. In addition to this, their overlap may be related to the fact that PF can embed locative BE.

resemblance relation between the two nouns. Now, Jackendoff's concept of KIND cannot refer to this same notion Snyder is formulating. In fact, it would not make sense that the two notions of 'kind' are identical because otherwise we would expect every single NN compound to fall under this relation. What the KIND function denotes is a hyponymy relation internal to the compound: an [XY], where X is the head, means that 'X is a kind of Y' (or viceversa since according to Jackendoff this relation is reversible). So in the *pine tree* example, the meaning would be 'a pine that is a kind of tree'. Crucially, this semantic relation is expressed in Spanish through simple words. To note a few examples: *pine tree* = 'pino', *ferryboat* = 'ferry'; *puppy dog* = 'cachorro'.

To sum up, the full set of basic functions as proposed by Jackendoff can be taken as the complete model (i.e. the superset), given the languages studied so far. There seems to be cross-linguistic variation with respect to the set of functions that a given language employs, in that only a subset of the functions may be operative. Moreover, once we analyze N-de-N data below, we will see that the same type of variation may also be internal to a single language itself, given that different formal, but semantically identical structures may exhibit a distinct set of functions. Having identified the basic functions in Spanish NN compounds, I next turn to discuss their productivity.

5.2.2 Hierarchy of semantic relations based on the basic functions proposed by Jackendoff

In order to find out an answer to the question of which basic functions are more productive in Spanish NN compounds, they have been ordered in table 2 according to the frequency of their occurrences in the data. I anticipated that few semantic relations would be favored, and this is precisely what is found by using a log-linear regression model explained below: SIMILAR, PF, X serves as Y, BE (AT/IN/ON) and HAVE are the most frequent. By fitting log-linear regression models to the frequency count of different functions, I could compare every model that splits the functions into 2 groups. It was found that, amongst these models, the best fitting model was the one that makes the split just below HAVE. Goodness of fit was measured using the Akaike Information Criterion (AIC). This is shown in Table 3.

Table 3 provides evidence that, if you decide to divide the data into two blocks, this is the best way to do it. It does not provide any evidence that dividing them in two is any better than dividing them into more strata. I identify these blocks as "productive" and "unproductive". The numeral differences in the AIC column indicate whether the difference between the functions is insignificant (0–2), positive (2–6), strong (6–10) or very strong (more than 10). Therefore, the split into the two blocks is located below

Table 3: productive and unproductive basic functions according to AIC.

	AIC
SIMILAR	127
PF	121
X serves as Y	115
BE (AT/IN/ON)	111
HAVE	**106**
COMP	121
PART	131
CAUSE	138
BE	161
ARGUMENT SCHEMA	168

HAVE because there is a very strong difference between HAVE (106) and COMP (121) below, and between HAVE (106) and BE (AT/IN/ON) (111) above. The productive block is composed of the functions SIMILAR, PF, X serves as Y, BE (AT/IN/ON) and HAVE. SIMILAR proves to be the most productive function in the data: i.e. there is a 6 point difference between SIMILAR and PF which makes this evidence strong. The unproductive block, on the other hand, consists of functions COMP, PART, CAUSE, BE and ARGUMENT SCHEMA. The functions within the unproductive block generally have a 10 point or more than a 10 point difference with respect to the one below: for instance, COMP (121) vs. PART (131) or CAUSE (138) and BE (161). Since this difference is equal or larger than 10, the productivity contrasts between functions are either strong or very strong.

As seen in Table 6, the function SIMILAR is the most productive one. In other words, there is a tendency for Spanish NN compounds to express a resemblance or metaphorical relation. The function SIMILAR is illustrated in (17):

(17) SIMILAR

a. gas_1 mostaza_2: — [GAS_1^{α}; [SIMILAR (α, MOSTAZA_2)]],
gas mustard — 'gas similar to mostaza'
'mustard gas'

b. novela_1 río_2: — [NOVELA_1^{α}; [SIMILAR (α, RÍO_2)]],
novel river — 'novela similar to a río'
'river novel'

In addition to the resemblance/metaphorical relation, there are four other functions that are productive: PF, X serves as Y, BE (AT/IN/ON) and HAVE. PF indicates purpose, or occupations and abilities if talking about a person, and X serves as Y denotes an entity (i.e. the head of the compound) whose purpose is to function as another entity (i.e. the modifier). Examples of their generative schemata for F are shown in (18):

(18) a. PF

i. hombre_1 anuncio_2: — [HOMBRE_1^{α}; [OCC (WORK (α, ANUNCIO_2))]]
man commercial — 'hombre whose occupation is to work as a anuncio'
'sandwich-board man'

ii. pistola_1 láser_2: — [$\text{PISTOLA}_1^{\alpha}$; [PF (SHOOT (INDEF, LÁSER_2, FROM α))]]
gun laser — 'pistola whose PF is for people to shoot a láser from it'
'laser gun'

b. X serves as Y

i. canción_1 protesta_2: — [$\text{CANCIÓN}_1^{\alpha}$; [BE (PF ($\alpha$), PF ($\text{PROTESTA}_2$))]]
song protest — 'canción whose PF is to function as the PF of protesta'
'protest song'

ii. perro_1 guardián_2: — [PERRO_1^{α}; [BE (PF (α), PF (GUARDIÁN_2))]]
dog guard — 'perro whose PF is to function as the PF of guardián'
'guard dog'

It is interesting to note that the meaning of many NN compounds in Spanish expresses PF, even though it is not a basic function: nouns indicating purpose are all included under the function PF. Furthermore, the aspect of purpose is significant in Spanish as it is not only present here, but also in the locative function BE (AT/IN/ON). In other words, all the generative schemata suggested for PF in English are available for Spanish NN compounds. Besides, it is worth noting that the different schemata proposed to express the many senses of the locative function such as those presented in (19) are attested in Spanish NN compounds, except for temporal location (e.g. the counterpart of *summer rain* is not found in Spanish: **lluvia verano*). A sample of these schemata is provided in (19):

(19) BE (AT/IN/ON)

a.	reloj$_1$ pulsera$_2$: watch bracelet 'bracelet watch'	[[RELOJ$_1^\alpha$; [BE (α, ON PULSERA$_2$)]]	'reloj that is located on a pulsera'
b.	zona$_1$ euro$_2$: zone euro 'euro zone'	[ZONA$_1^\alpha$; ([BE (EURO$_2$ IN α)])]	'zona with euro in it'
c.	fútbol$_1$ sala$_2$: soccer room 'indoor soccer'	[FÚTBOL$_1^\alpha$; PF ([BE (α IN SALA$_2$)])]	'fútbol whose PF is to be (played) in the sala
d.	zapato$_1$ teléfono$_2$: shoe phone 'phone shoe'	[ZAPATO$_1^\alpha$; PF ([BE (TELÉFONO$_2$ IN α)])]	'*zapato* whose PF is to have a *teléfono* in it'

From the examples in (19) we can infer that the locative function also involves some sort of possession as seen in (19b) and (19d). Nevertheless, this possession is never on its own as it always appears with some location flavor. The function denoting possession and ownership by definition is HAVE. Though Jackendoff provides only two schemata for this function hoping that others will come up with a full semantic analysis of the function, no further schemata can be found in Spanish: this function seems to be restricted to the two schemata suggested. A sample of each is exemplified in (20):

(20) HAVE

a.	gorra$_1$ visera$_2$: cap hat 'peaked cap'	[GORRA$_1^\alpha$; [HAVE (α, VISERA$_2$)]]	'gorra that has a visera'
b.	perro$_1$ pastor$_2$: dog shepherd 'shepherd's dog'	[PERRO$_1^\alpha$; [HAVE (PASTOR$_2$, α)]],	'perro that the pastor has'

The high productivity of SIMILAR might imply that this function could be used to account for the "default" meaning of novel NN compounds in Spanish, when no context is available.[10] For instance, the novel compound *autobús calabaza* ('bus pumpkin' > *pumpkin bus*) is most naturally interpreted with the SIMILAR function just as *tren bala* ('train bullet' > *bullet train*). This is shown in (21) below:

(21)	a.	autobús$_1$ calabaza$_2$: bus pumpkin 'pumpkin bus'	[AUTOBÚS$_1^\alpha$; [SAME/SIMILAR (α, CALABAZA$_2$)]],	'autobús similar to a calabaza'
	b.	tren$_1$ bala$_2$: train bullet 'bullet train'	[TREN$_1^\alpha$; [SIMILAR (α, BALA$_2$)]],	'tren similar to a bala'

Another piece of evidence that can be used to support the claim that SIMILAR can be taken as the default function for novel NNs in Spanish comes from the promiscuity of

[10] By default, I mean "the most creative". In other words, The high producitiviy of similar can be taken to indicate that this function is the most producitve one when creating (novel) compounds. Impresionistically, this prediction seems consistent.

compounds. That is, whenever a compound involves the co-existence of more than one relation, and one of those relations is SIMILAR, the tendency is to use SIMILAR as the first choice. Some compounds exhibiting this property are illustrated in (22):

(22)	a.	hombre$_1$ araña$_2$: man spider 'spider man'	i. [HOMBRE$_1^{\alpha}$; [SIMILAR (α, ARAÑA$_2$)]] 'hombre similar to araña' ii. [HOMBRE$_1^{\alpha}$; [PF (BECOME (α, ARAÑA$_2$))]], 'hombre whose PF is to become araña' iii. [HOMBRE$_1^{\alpha}$; [PART (ARAÑA$_2$, α)]], 'hombre who is partly araña'
	b.	teléfono$_1$ zapato$_2$: phone shoe 'phone shoe'	i. [TELÉFONOα; [SIMILAR (α, ZAPATO)]], 'teléfono similar to zapato' ii. [TELÉFONOα; [BE (PF (α), PF (ZAPATO))]], 'teléfono whose PF is to function as the PF of zapato' iii. [TELÉFONOα; PF ([BE (α, IN ZAPATO)])], 'teléfono whose PF is to be in a zapato'

In NN compounds such as (22a) and (22b), where various interpretations including SIMILAR are possible, SIMILAR is preferred over the other interpretations to account for the most salient meaning given a context. Nevertheless, this is only a tendency and would need to be tested experimentally.[11]

Up until this point, we have examined the most productive functions found in the data. Tables 2 and 3 also show that there are six other functions that are more peripheral. COMP and PART are very similar, yet they express separate concepts: the former denotes composition, whereas the latter exhibits either a part-whole or a whole-part relation. This subtle contrast can be evidenced in (23):

(23)	a.	COMP	
		i. flota$_1$ pirata$_2$: fleet pirate 'pirate fleet'	[FLOTA$_1^{\alpha}$; [COMP (α, PIRATA$_2$)]] 'flota composed of pirata(s)
		ii. buque$_1$ cisterna$_2$: ship tank 'tank ship'	[BUQUE$_1^{\alpha}$; [COMP (α, CISTERNA$_2$)]], 'buque composed of a cisterna'
	b.	PART	
		i. perro$_1$ lobo$_2$: dog wolf 'wolf dog'	[PERRO$_1^{\alpha}$; [PART (α, LOBO$_2$)]] 'perro that has part of lobo'

[11] There are a few cases where SIMILAR seems to not be the most salient function:

a. perro$_1$ vampiro$_2$ ('vampire dog')
i. [PERRO$_1^{\alpha}$; [BE (α, VAMPIRO$_2$)]]
'perro that is a vampiro'
ii. [PERRO$_1^{\alpha}$; [SIMILAR (α, VAMPIRO$_2$)]]
'perro similar to a vampiro'
iii. [PERRO$_1^{\alpha}$; [PF (BECOME (α, VAMPIRO$_2$))]]
'perro whose PF is to become a vampiro'
iv. [PERRO$_1^{\alpha}$; [PART (VAMPIRO$_2$, α)]]
'perro that is partly a vampiro'

b. coche$_1$ hippie$_2$ ('hippie car')
i. [COCHE$_1^{\alpha}$; [HAVE (HIPPIE$_2$, α)]]
'coche that hippies have'
ii. [COCHE$_1^{\alpha}$; [SIMILAR (α); [HAVE (HIPPIE$_2$, α)]]]
'coche similar to the coche hippies have'

ii. agua$_1$ nieve$_2$: [AGUA$_1^{\alpha}$; [PART (NIEVE$_2$, α)]]
water snow 'agua that is composed in part of nieve'
'sleet'
iii. traje$_1$ pantalón$_2$: [TRAJE$_1^{\alpha}$; [PART (PANTALÓN$_2$, α)]],
suit pants 'traje that has a pantalón as a part'
'pantsuit'

COMP as suggested by Jackendoff (2009, 2016) is reversible, but we have only been able to find examples of the 'X composed of Y' schema which is the one illustrated in (23a), while the 'X that Y is composed of' schema is not attested in the data. Thus, this could be used as evidence to argue that the reversibility of this function may be language specific. On the other hand, the different schemata proposed for PART are found in (23b). We have provided an example of the variant paraphrase for mass nouns in (23b.ii). If dealing with a countable counterpart as in (23b.iii), the formal structure stays the same but the paraphrase is expressed as follows: 'X that has Y as a part'.

The function CAUSE is very restricted to examples including the head nouns *guerra* ('war'), *efecto* ('effect') and *virus* ('virus'). In addition, CAUSE can also appear in complex NN compounds that require cocomposition. This is seen in (24):

(24) CAUSE
a. efecto$_1$ placebo$_2$: [EFECTO$_1^{\alpha}$; [CAUSE (PLACEBO$_2$, α)]]
effect plcebo 'efecto caused by a placebo'
'placebo effect'
b. mochila$_1$ cohete$_2$: [MOCHILA$_1^{\alpha}$; [PF (MOVE$^{\beta}$ (α); [CAUSE (COHETE$_2$, β)])]]
backpack rocket
'rocket backpack' 'mochila that moves by a cohete causing its movement'

The next function on the productivity scale is BE. According to Fernández-Domínguez (2016), BE implies a more coordinate semantic relation, meaning that BE denotes an entity that is both N$_1$ and N$_2$ at the same time. Nevertheless, we consider that the NN compounds collected are not coordinate but subordinate ground (as defined by Saclise and Bisetto 2009) because they are mainly a hyponym of the head and only one of nouns is inflected for plural: the left hand noun is inflected as in *perro vampiro* ('vampire dog') > *perros vampiro* as opposed to *niño atleta* ('athlete kid') > *niños atletas* where both nouns are inflected for plural. A sample is given in (25):

(25) BE
a. perro$_1$ vampiro$_2$: [PERRO$_1^{\alpha}$; [BE (α, VAMPIRO$_2$)]]
dog vampire 'perro that is a vampiro'
'vampire dog'
b. príncipe$_1$ rana$_2$: [PRÍNCIPE$_1^{\alpha}$; [BE (α, RANA$_2$)]]
prince frog 'príncipe that is a rana'
'frog prince'

The most peripheral function, though, is MAKE which implies a semantic relation of source or creation. Thus, there are two possible readings for this function which are shown in (26):

(26) MAKE
a. tela$_1$ araña$_2$: [TELA$_1^{\alpha}$; [MAKE (ARAÑA$_2$, α)]]
net spider 'tela made by the araña'.
'spiderweb'

b. horno$_1$ mircroondas$_2$: [HORNO$_1^\alpha$; [MAKE (α, MICROONDAS$_2$)]]
oven microwave 'horno that makes microondas'
'microwave (oven)'

Scholars agree in the difficulty of differentiating MAKE from CAUSE. For example, Jackendoff (2016: 30) and Mellenius and Rosenberg (2016) agree that MAKE could be decomposed into 'cause something to come into existence'. While Jackendoff bases this claim on English, Mellenius and Rosenberg also note this for Swedish. The availability of readings such as (26b) for Spanish compounds adds Spanish to this list of languages.

The remaining function to be considered is ARGUMENT SCHEMA. This function is as peripheral as MAKE in our data, but it may be more productive than it appears in the data. That is because we tried to avoid collecting NN compounds whose head or modifier either had non-inflectional morphology or constituted borderline cases of deverbal derivation. In such cases, there is no need to use an external function to satisfy F since the function is already expressed by one of the elements of the compound: either the head or the modifier. Such NN compounds include examples like *ataque* ('attack'), *abordaje* ('boarding') or *salvavidas* (literally 'save + lives') and are illustrated in (27):

(27) ARGUMENT SCHEMA
ataque$_1$ pirata$_2$: [ATAQUE$_1$ (PIRATA$_2$, INDEF))]
attack pirate 'ataque on something by piratas'
'pirate attack'

Apart from this, I also believed that functions PROTECT and KIND would be extremely peripheral since they are not attested in either French or Swedish NNs, as discussed by Rosenberg (2013). As mentioned above, this is also true of the Spanish data collected. These are not the only unattested functions, though; CLASSIFY and MADE FROM are not present either. We will have to examine the Spanish N-de-N data to determine whether these functions are structure specific, or language specific. All in all, based on the frequency and productivity of the basic functions the most relevant semantic relations that hold in Spanish NN compounds are the following in (28):

(28) resemblance/metaphorical > purpose > location > possession > composition > part-whole/whole-part > cause > both (be) > source/creation

Before I move on to comparing these results with N-de-Ns, I would like to comment on one issue that I have briefly referred to while examining some of the functions: promiscuity. Jackendoff claims that many NN compounds in English can be promiscuous (i.e. they can encode more than one possible reading). Promiscuity is present in Spanish NN compounds as well, but not to the same degree as in English NNs. This is because few NN compounds in the data allow more than one reading, and because not all possible semantic relations can coexist in a Spanish NN compound, as they can in an English NN compound.[12] Despite this difference, promiscuity does exist for certain NN compounds, as shown in Table 4.

As it can be drawn from Table 4, the tendency is for Spanish NN compounds to bear only one semantic relation. This implies that NN compounds are not usually promiscuous; but if they are, it is more likely for them to have two than three possible relations. Furthermore, it is very rare to find NN compounds with more than three relations. One

[12] See Jackednoff's (2016: 20) example *boxcar*.

Table 4: Number of NN compounds with 1, 2, 3 or more possible relations.

Number of relations	Number of compounds	Examples
1	171	Sofa cama ('sleeper sofa'): BE Serpiente toro ('bull snake'): SIMILAR
2	21	Papel aluminio ('aluminum foil'): SIMILAR, COMP Coche radar ('radar car'): PF, BE (AT/IN/ON)
3	11	Tren hospital ('hospital train'): PF, X serves as Y, HAVE Barco fantasma ('ghost ship'): SIIMLAR, BE (AT/IN/ON), HAVE
More than 3	1	Perro vampiro ('vampire dog'): BE, SIMILAR, PF, PART

possibility that might explain this is related to lexicalization and productivity issues. That is, there are few attested NN compounds in Spanish compared to languages like English where NN compounding is very productive, and the few that exist are lexicalized, which entails that their meaning is fossilized. Therefore, it is very difficult to assign them a new meaning. However, since promiscuity is context bound, lexicalization may be disregarded if a compound with a not-yet-fossilized meaning appears in a novel context. Unfortunately, there are not examples in the data that can support this claim.

I will, now, apply the basic functions model to analyze the meaning of N-de-Ns. The results concerning this structure may also be able to shed some light on the status of the preposition *de* in this structure: whether it is a "dummy" element, rather than a true preposition, as proposed by Nicoladis (2002). In addition, this may clarify some of the unsolved issued posed above: are KIND and PROTECT language or structure specific?

5.2.3 Semantic relations based on the basic functions found in N-d-N constructions

In section 5.2.2. I listed the basic functions attested in the Spanish NN compounds in a hierarchical order. In order to establish a comparison of the basic functions found in NNs and N-de-Ns, Table 5 has been created. It provides a list of the functions exhibited by N-de-Ns following also a hierarchical order regarding their productivity. Similarly to what was shown in Table 2, there is a total of 203 N-de-Ns in the database, but due to promiscuity some N-de-Ns allowed more than one reading. Therefore, the column *Type* is not making reference to the number of N-de-Ns, but to the total number of occurrences of attested basic functions (242), considering all the possible meanings for the N-de-Ns collected.

Table 5 indicates that most of the basic functions present in NN compounds are also attested for N-de-Ns, though with some differences: CLASSIFY, KIND and BE are not attested in N-de-Ns, whereas MADE FROM and PROTECT which were not found in NN compounds are available for N-de-Ns.

Again here, by fitting log-linear regression models to the frequency count of different functions, we could compare every model that splits the functions into 2 groups. It was found that, amongst these models, the best fitting model is the one that makes the split just below MADE FROM. Goodness of fit was measured using the Akaike Information Criterion (AIC). This information is presented in Table 6.

Table 6 provides evidence that, if you decide to divide the data into two blocks, this is the best way to do it. It does not provide any evidence that dividing them in two is any better than dividing them into more strata. The numeral differences in the AIC column indicate whether the difference between the functions is insignificant (0–2), positive (2–6), strong (6–10), or very strong (more than 10). Therefore, the split into the two blocks is located below MADE FROM because there is a very strong difference between

Table 5: Functions within Spanish N de N constructions.

Semantic relation	Spanish N de N	English translation	Type
PF	Alcohol de heridas Casa de moneda	'alcohol' 'coin house (mint)'	44
BE (AT/IN/ON)	Estanque de patos Tormenta de verano	'duck pond' 'summer storm'	41
COMP	Bandeja de plata Circo de pulgas	'silver tray' 'flea circus'	35
PART	Corazón de manzana Reja de arado	'apple core' 'plowshare'	28
HAVE	Derecho de autor Lámpara de flecos	'copyright' 'fringe lamp'	22
MADE FROM	Aceite de girasol Caña de azúcar	'sunflower oil' 'sugarcane'	19
CAUSE	Dolores de parto Herida de navaja	'birth pains' 'knife wound'	14
PROTECT	Botas de lluvia Cinturón de castidad	'rain boots' 'chastity belt'	13
SIMILAR	Arco de herradura Sargento de hierro	'horseshoe arch' 'iron sergeant'	11
MAKE	Agujero de ratón Veneno de serpiente	'mouse hole' 'snake poison'	10
ARGUMENT SCHEMA	Calidad de vida Guía de teléfono	'life quality' 'telephone book'	4
X serves as Y	Señal de alarma	'alarm signal'	1
Types of Spanish N-de-N			242

Table 6: productive and unproductive basic functions according to AIC.

	AIC
PF	151
BE (AT/IN/ON)	127
COMP	108
PART	97
HAVE	95
MADE FROM	**94**
CAUSE	102
PROTECT	108
SIMILAR	116
MAKE	120
ARGUMENT SCHEMA	142

MADE FROM (94) and CAUSE (102) below. The productive block is composed of the functions PF, BE (AT/IN/ON), COMP, PART, HAVE and MADE FROM. PF proves to be significantly the most productive function in the data: there is a 24 point difference between PF and BE (AT/IN/ON) which makes this evidence strong. The unproductive block, on the other hand, consists of functions CAUSE, PROTECT, SIMILAR, MAKE, ARGUMENT SCHEMA and X serves as Y.

If compared to the results for NN compounds, the basic functions attested for N-de-Ns show a different hierarchical order. Functions like PF, BE (AT/IN/ON) and HAVE are still on top of the scale. Moreover, in the case of the locative relation, there are realizations of all the schemata proposed by Jackendoff (2009, 2016), including temporal location, which was not attested in NN compounds. An example is provided in (29):

(29) tormenta$_1$ de verano$_2$: [TORMENTA$_1^\alpha$; [BE$_{temp}$ (α, AT VERANO$_2$)]]
storm of summer 'tormenta that takes place at time verano'
'summer storm'

It is striking, however, that SIMILAR and X serves as Y are much less productive with N-de-Ns than with NNs, comparing Tables 3 and 6. Especially SIMILAR, which was significantly the most productive function in NN compounds, is rather peripheral here. Likewise, X serves as Y seems restricted in N-de-N compounds. On the other hand, COMP and PART are more frequent in N-de-Ns than in NN compounds. Therefore, it seems that N-de-N compounds are preferred when expressing relations of 'composition' and 'part-whole/whole-part'. Furthermore, while in NN compounds the reversibility of COMP was not available, it is possible in N-de-N, as illustrated in (30):

(30) molécula$_1$ de agua$_2$: [MOLÉCULA$_1^\alpha$; [COMP (AGUA$_2$, α)]]
molecule of water 'molécula that agua is composed of'
'water molecule'

Although according to Tables 2 and 5 it appears that there is no noticeable difference between NNs and N-de-Ns, with respect to the amount of occurrences of the functions CAUSE and MAKE, some differences can be noted with respect to the range of nouns that can encode these functions in the two types of structures. While in NN compounds CAUSE was limited to items whose head was *war*, *effect*, or *virus* or to cases that involve cocomposition (e.g. (24b)),[13] this function has a wider scope in N-de-Ns: it covers the meaning of every N-de-N that involves an originator. Some examples include the following: *herida de bala* ('wound of bullet' > *bullet wound*), *mancha de sol* ('stain of sun' > *sunspot*) or *vapor de agua* ('steam of water' > *water steam*). As for MAKE, even though it is unproductive with both NNs and N-de-Ns, it has to be noted that with N-de-Ns the function MAKE can cover any product or substance created by an agent or an instrument. On the contrary, with NNs, MAKE is limited to very few examples involving a source. An example of each schema is provided below in (31):

(31) a. veneno$_1$ de serpiente$_2$: [VENENO$_1^\alpha$; [MAKE (SERPIENTE$_2$, α)]]
poison of snake 'veneno made by the serpiente'
'snake poison'
b. gusano$_1$ de seda$_2$: [GUSANO$_1^\alpha$; [MAKE (α, SEDA$_2$)]]
worm of silk 'gusano that makes seda'
'silkworm'

In addition to this, two of the functions that were not found with NN compounds are attested in N-de-Ns: MADE FROM and PROTECT. The former denotes composition where the source is no longer visible. This function is extremely productive with N-de-N compounds; it designates edible and drinkable products such as the ones in (32):

(32) a. aceite$_1$ de girasol$_2$: [ACEITE$_1^\alpha$; [MADE (α, FROM GIRASOL$_2$)]]
oil of sunflower 'aceite made from girasol'
'sunflower oil'

[13] *Guerra narco* (*narc war*), *efecto placebo* (*placebo effect*), or *virus vampiro* (*vampire virus*) among others.

	b.	caña$_1$ de azúcar$_2$: cane of sugar ‘sugarcane’	[CAÑA$_1^{\alpha}$; [MADE (AZÚCAR$_2$, FROM α)]] ‘caña that azúcar is made from’

The examples in (32) demonstrate that MADE FROM is reversible in Spanish just as in English, as the two possible schemata for this function are active. The schema in (51b) is less frequent, though. The high productivity of this function with N-de-N compounds emphasizes the core importance of the semantic relation ‘composition’ in N-de-N constructions, as opposed to its role in NNs where it was unattested. It seems, therefore, that there is a significant tendency to express this semantic relation with *N de Ns* rather than with NNs.

As for PROTECT, this function had only been seen in English NN compounds, which led me to believe that it was language specific. Nevertheless, the fact that this function occurs in Spanish N-de-N compounds shows this not to be the case. Thus, the availability of certain basic functions in Spanish may be restricted to a particular formal structure. Future research may show whether this is true of all languages that have alternative nominal compounding constructions. The two schemata provided to indicate the reversibility of PROTECT are manifested in Spanish. This is illustrated in (33):

(33)	a.	cinturón$_1$ de castidad$_2$: belt of chastity ‘chastity belt’	[CINTURÓN$_1^{\alpha}$;[PROTECT(α,CASTIDAD$_2$ FROM Z)]] ‘cinturón protects castidad from something’
	b.	gafas$_1$ de sol$_2$: glasses of sun ‘sunglasses’	[GAFAS$_1^{\alpha}$;[PROTECT (α, EYES, FROM SUN$_2$)]] ‘gafas protect eyes from sol’

Examples such as (33a) have “what is protected” specified in the compound (i.e. chastity) and require the reader to interpret “the threat” (i.e. Z). Others like (33b) work the opposite way: the thing or person protected is omitted and has to be inferred (e.g. eyes), whereas the threat is explicitly stated (e.g. sun).

In addition to the functions that are available, the notion of promiscuity is also applicable to the interpretation of N-de-Ns. Identical to what was found in NN compounds, not every N-de-N is subjected to promiscuity, as the most frequent possibility is for these constructions to have only one interpretation. The information regarding promiscuity is given in Table 7.

The most common promiscuous N-de-Ns are those where two interpretations are possible. Interestingly, most promiscuous examples arise due to the fact that they constitute borderline cases which make the distinction between certain functions slippery: for instance, it is challenging to differentiate between PART or COMP in *tarta de chocolate*

Table 7: Number of N de N constructions with 1, 2, 3 or more possible relations.

Number of relations	Number of compounds	Examples
1	175	Cáncer de colon (‘colon cancer’): BE (AT/IN/ON) Muñeco de trapo (‘rag dummy’): COMP
2	22	Tarta de chocolate (‘chocolate cake’): PART, COMP Sargento de hierro (‘iron sergeant’): SIMILAR, COMP
3	6	Pan de ajo (‘garlic bread’): PART, COMP, BE (AT/IN/ON) Pecho de paloma (‘pigeon chest’): PART, SIMILAR, HAVE
More than 3	0	--------------

(‘chocolate cake’) even if a context is provided. What seems evident is that having three or more than three functions is not common. While for NNs I hinted that the lack of promiscuity may be due to lexicalization, claiming that lexicalization is the cause of little promiscuity in N-de-Ns is challenging: some of the N-de-Ns are clearly lexicalized (i.e. *aceite de oliva* ‘olive oil’ or *reloj de bolsillo* ‘pocket watch’), but at the same time, N-de-Ns are much more creative.[14] A possibility is that there is no correlation between N-de-Ns being creative and being promiscuous: even though an N-de-N is created *ad hoc* it may not have several meaning possibilities depending on the context; rather it may allow only one possible semantic relation just like we saw for *apple juice seat* in footnote 1. For instance, *sofá de jardín* (sofa of garden > ‘garden sofa’) is a ‘sofa whose function is to be located in the garden. However, this is just a prediction that should be further explored in future research.

Before assessing other issues, we would like to comment in a particularly interesting case by comparing (34) and (35), both involving cocomposition:

(34) barco$_1$ de vapor$_2$: [BARCO$_1^{\alpha}$; [PF (MOVE$^{\beta}$ (α)); [CAUSE (VAPOR$_2$, β)]]]
boat of steam ‘barco that moves by vapor causing its movement’
‘steam boat’

(34) is a cocomposed example of a means of transport whose formal structure is *N de N*. However, there are some recent NN compounds where *de* seems to be deleted. This is seen in (35):

(35) coche$_1$ diesel$_2$: [COCHE$_1^{\alpha}$; [PF (MOVE$^{\beta}$ (α)); [CAUSE (DIESEL$_2$, β)]]]
car diesel ‘coche that moves by diesel causing its movement’.
‘diesel car’

(35) is identical to (34) regarding the semantic structure except for the fact that it is an NN compound. It is the first time we encounter this term without the preposition as it is traditionally expressed as *coche de diesel* (‘car of diesel’ > *diesel car*). Disregarding cocomposition, another interesting pair of semantically identical but structurally different words is illustrated in (36):

(36) a. billete de tren
ticket of train
‘train ticket’
b. bono-tren
pass-train
‘train pass’

The same process can be applied here: originally, *bono* and *tren* were linked by the preposition *de* as in the original form *bono de tren* (‘pass of train’), just as in (36a); due to changes, the preposition lost its semantic import and was removed resulting in a fused NN compound as in (36b). On the other hand, this process has not affected (36a) yet, as it retains the preposition. A similar case that can support this claim is found in *telaraña* (‘web spider’ > *spiderweb*). Although it now appears as an NN compound, it was previously expressed as *tela de araña* (‘web of a spider’). The loss of the preposition resulted in *tela araña,* which due to phonology has been amalgamated.

[14] The fact that these results arise is because only a sample was collected as opposed to compounds whose search was exhausted.

A possible explanation for why some N-de-Ns are losing *de* first may be related to the lexical frequency of the items: that is, lexical items such as *coche diesel* or *bonotren* may be more common in speech than others and thus are more likely to lose the preposition. However, this would not explain why *billete de tren*, which is very frequent, has not yet lost it. The fact that there are examples where the preposition *de* seems to have been deleted provides evidence to support Nicoladis's (2002) claim that prepositions like *de*, in this context, are losing their meaning to become a linking element and are, therefore, dispensable. Thus, it seems that there is a process according to which this linking element has developed from a lexical item, causing N-de-Ns to become a conceptual unit, which can be briefly summarized in the scale in (37):

(37) preposition with full meaning > linking element > zero

Overall, it seems that the process is still on the second step in Spanish N de Ns, but there are instances of it such as (36) predicting a potential loss of the preposition. In order to prove this claim and the scale in (36), additional evidence should be gathered.

Heretofore, I have noted some of the differences and similarities between NNs and N-de-Ns with respect to the functions that are present, but in the case of the latter constructions there are also functions that do not seem to be active. One of these is BE whose non-existence seems to indicate that the semantic relation 'both' cannot be expressed formally with N-de-Ns, perhaps due to the import of the preposition *de* which is similar to "of". It is one of these functions, just like MADE FROM or PROTECT, that is structure specific. In other words, in Spanish, just like in French, some basic functions seem to be operative depending on the formal structure of the word (i.e. NN and N-de-N): a basic function may be attested for N-de-N (i.e. MADE FROM and PROTECT), but not for NN compounds and vice versa (e.g. BE). Apart from this, there is another interesting conclusion that the data lead us to: functions CLASSIFY and KIND are not present in N-de-N compounds either, just as in the case of NN compounds. In other words, these two functions are the only ones for which there is no realization at all in both NN and N-de-N compounds in Spanish. This indicates that they are both inoperative in Spanish regardless of the formal structure, which means that they should not be considered part of the full set of basic functions for Spanish. In addition, KIND is not found in Spanish, French and Swedish, and is restricted to English.

In short, considering the attested and non-attested basic functions in N-de-N compound data, a validated hierarchy of semantic relations can be established, as shown in (38):

(38) a. **NN compounds:** resemblance/metaphorical > purpose (X functions as Y) > location > possession > composition > part-whole/whole-part > cause > both > source/creation
b. **N-de-N compounds:** purpose > location > composition > part-whole/whole-part > possession > sourceless composition > cause > protection > source/creation > resemblance/metaphorical > function as Y

Some semantic relations seem to differ with respect to the position they occupy on the scale regarding their formal structure (NN or N-de-N): primarily the resemblance/metaphorical and 'X functions as Y', which are on top of the scale for NN compounds, are at the bottom for N-de-Ns; the opposite happens with relations 'composition' and 'part-whole/whole-part'. One interesting question that can be raised at this point is whether the distribution of functions differs by construction. In order to find an answer to this question a log-linear regression model of the frequency of different functions was run using the predictors: 'functional category', 'construction' (PP vs. NN), and then testing the interaction between the two. The significance of the interaction term was tested using a likelihood ratio test. The likelihood-ratio test suggests that the interaction term explains significant amounts of variance: chi-square (9) = 106, $p < 0.0001$. Thus, it could be argued that the prevalence of different functions differs by construction: the position that some semantic relations (i.e. composition) occupy on the NN compound scale is significantly different from the position that the same relations occupy on the N-de-N compound scale.

Now that I have presented the main results, I will move on to discuss these and how they fit the hypotheses formulated in section 4.

6 Discussion

In section 4, I outlined two competing hypotheses that I labeled A and B and that will be repeated here in (39) and (40) respectively:

(39) **Hypothesis A:** NN compounds allow few productive semantic relations, but this does not guarantee creativity.

(40) **Hypothesis B:** Only semantic relations that are productive will be creative.

Hypothesis A was formulated on the basis of Marqueta-García's (2017) work that there is only one semantic relation, i.e. *Identifying*, which is very productive because according to her all NN compounds have this relation underlyingly. But at the same time, this relation was not creative given that it neglected the creation of new impromptu items. Hypothesis B, on the contrary, predicted that productivity and creativity go hand in hand entailing that if a semantic relation is productive it should also be creative. The data and the analysis provided suggest that reducing all semantic relations to the one proposed by Marqueta-García (2017) is far from capturing the empirical facts. In fact, there seem to be a few of them that are possible semantic relations.

Thus, Hypothesis B seems more adequate. Considering the results obtained for NN compounds, only the following semantic relations are productive: resemblance/metaphorical, purpose (X functions as Y), location, and possession, functionally represented with SIMILAR, PF, X serves as Y, BE(AT/IN/ON), and HAVE. Thus, by hypothesis, these are the exclusive basic functions that should be creative when coining or interpreting novel NN compounds. In addition, this predicts that the other functions should not surface because this will lead to a crash at LF: Generalized Modification will not be able to provide an interpretation for these given that they are recognized by the Spanish grammar as not creative. Although at this time, no experimental results are available that can further support our hypothesis, I have data from a very small pilot study with my colleague Omar Barahona where 10 native Spanish speakers were asked to interpret the meaning of novel NN compounds.[15] The results of this pilot indicate that the most used or preferred functions to render the meaning of novel NN compounds are SIMILAR, BE(AT/IN/ON), and X serves as Y.[16] Crucially these results very much align with the functions that fall into the productive block, and thus can be taken to emphasize the optimality of the hypothesis that if a function is productive it can be creative. Therefore, this is important evidence that these functions are creative and satisfy Generalized Modification at LF during the

[15] The results of this pilot cannot be included in this work because they were conducted for a small research project in the spring of 2016 as part of coursework that did not require for us to have IRB approval. The purpose of the experiment was twofold: to find what the most common semantic relations were in novel NN compounds and to compare the formal structures used in Spanish to express what in English is expressed with an NN. Focusing only on the first part, we provided them with a Forced Choice Task and an open interpretation task. These are the results we will briefly mention. I am grateful to Omar Barahona for allowing me to mention the pilot and the results we collected together.

[16] PF was not included in the forced choice task because we were only concerned with primary basic functions.

semantic composition. However, and as a reviewer recommends, an experimental study should shed more light on this hypothesis which, though seems on the right track, as it stands now needs further empirical support.[17]

In addition to this, the purpose of establishing a hierarchy of semantic relations was to find out whether there are some relations that are more productive than others and whether there are utterly unproductive ones. The hierarchy for NNs in (38a) indicates what I have been claiming all along, and what also seems to be supported by the small pilot: the resemblance relation is the most creative one when compared to the availability of others. This is also in line with what was hinted in section 5 with respect to SIMILAR being the "default" relation. It is interesting, nevertheless, to note that N-de-Ns display a different hierarchy of creativity. A question that arises is why this is so. When answering this question one needs to consider that N-de-Ns differ from compounds in that all the possible schemata proposed for each of the functions are realized in Spanish N-de-Ns; and that some basic functions unattested in NNs are found in N-de-Ns, which already suggests divergence. The difference between the hierarchies may be due to a blocking effect (Aronoff 1976, Embick and Marantz 2008): this basically refers to the fact that languages will block the creation of one an item X if there is another item Y that means (or is) the same. If we think of an example, some compounds are unable to be expressed with a preposition *de*: *balon reloj* but **balón de reloj*. William Snyder (p.c.) proposes that a possible hypothesis that could be formulated is the following: the semantic relation required for the target construction to satisfy the principle of Generalized Modification is only available in the NN form, which would block the surfacing of the prepositional alternative. This would go hand in hand with the hypothesis that the language user may be forced to coin or produce an NN compound when there is no other choice. And the same would apply in the other direction.

All in all, the results presented in this paper provide some hints on how to address the following research question: are NN compounds uncommon in Spanish because they are 'semantically inflexible'? In other words, does the fact that there is a limited set of relations that are productive influence the grammatical creativity of these constructions in the language? "Yes" seems to be the answer towards which we have been driven. Considering that only a few semantic relations are available, it is obvious that not every desired concept that, in languages like English, can be expressed with an NN compound is possible to be captured with an NN compound in Spanish. In fact, the gaps created by the NN compound paradigm, which are quite significant, are to be filled by N-de-N constructions. Nevertheless, we should not forget that there seems to be a current tendency for the two formal structures to be merged into a single structure (i.e, NN compounds) due to the ongoing process of N-de-Ns losing *de*. The existence of cases such as *telaraña* ('cobweb') or *coche diesel* ('diesel car') demonstrates that the preposition *de* is dispensable. These were formerly N-de-Ns (*tela de araña* and *coche de diesel* respectively), but have now become NNs because the preposition is no longer used with them. The existence of these examples as well as the possibility to analyze N-de-Ns as nominal compounds in Spanish provides strong evidence for Nicoladis' (2002) claim that prepositions are linking elements in these structures in Romance languages. Furthermore, it is very likely that *de* will disappear from Spanish N-de-Ns in the future, which would constitute an ultimate confirmation of

[17] Whether hypothesis B applies for N-de-Ns is not so clear, though. Even though the results and the data from the previous section suggest that not every semantic relation is creative our intuition tells us different. For instance, MADE FROM seems very fruitful to talk about products such as *queso de cabra/oveja/vaca* (cheese of goat/sheep/cow, i.e. 'goat/sheep/cow cheese). However, not all instances of these examples were included, only a sample; and this could have influenced the statistical results. This does not mean that all functions labeled as unproductive, and thus by hypothesis not creative, behave this way.

the claim that N-de-Ns are an alternative type of nominal compounds in languages like Spanish or French.

7 Conclusions

This paper has been concerned with analyzing the semantic relations that hold in NN compounds in Spanish based on Jackendoff's (2009, 2016) basic functions model. But more importantly, it has provided an analysis of why NN compounds in Spanish are not creative. The analysis of the data has led to the following conclusions.

The model proposed by Jackendoff can be successfully applied to Spanish NN compounds as almost all basic functions are (at least) attested. In this paper, I have shown that Spanish NN compounds do not realize all of the functions proposed for English either. More specifically, functions KIND, CLASSIFY, MADE FROM and PROTECT are not present in Spanish NNs. Therefore, the full set of functions as developed in Jackendoff's model seems particular to English, while only a subset of functions is present in the four languages investigated thus far. And this fact, i.e. that languages differ with respect to the basic functions they allow, has shed some light on Jackendoff's concern about the universality of his model. However, given that there seems to be variation with respect to the functions that the languages allow, we still need to identify what the most repeated or common set of functions is possible cross-linguistically. Therefore, in order to provide evidence for this question, data from various languages should be gathered.

The most important finding, nevertheless, is the fact that only those functions that are productive are also creative. Being creative, therefore entails being able to successfully create and interpret novel compounds. This explains why some NN compounds are nonlicitly formed in Spanish: they bear non-productive (and thus non-creative) basic functions underlyingly, leading to a violation of the LF filter in charge of interpreting the structure. Although the results seem promising, new and more data should be gathered by means of experimental tasks. Only this way my hypothesis can be strengthened. But this is a good start given that now we have a better, or at least clearer, idea of why the coining or interpretation of NN compounds can be constrained, which to my knowledge had never been sketched before.

Abbreviations

MASC = masculine, SG = singular.

Acknowledgements

I would like to thank some of the faculty at West Virginia University who made this paper possible: Sandra Stjepanović, for all the times she has provided me with constructive feedback and for every time she has read and revised my NN compound mess; Jonah Katz for always encouraging me to pursue this project further, his interesting discussions and valuable comments during his out-of-office office hours, and what he calls *spreadsheeting* lessons; and María Amores-Aguera for her magnificent observations and style suggestions. I am also very indebted to my advisor, Stefan Keine, for his suggestions and comments on how to approach some of the reviewers questions. Special thanks go to William Snyder for his insight and for devoting some of his busy time in Creteling-2019 to discuss some of these issues with me; and to Omar Barahona for letting me mention the pilot we did together. Finally, I would like to thank the three anonymous reviewers of Glossa. Should there be any error or inconsistency, it is my own.

Competing Interests

The authors have no competing interests to declare.

References

Allen, Margaret. 1978. *Morphological investigations*. University of Connecticut dissertation.

Arnaud, Pierre & Vincent Renner. 2014. English and French $[NN]_N$ lexical units: a categorial, morphological and semantic comparison. *Word Structure* 7(1). 1–28. DOI: https://doi.org/10.3366/word.2014.0054

Aronoff, Mark. 1976. *Word formation in Generative Grammar*. Cambridge, MA: MIT Press.

Bauer, Laurie. 1983. *English word-formation*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139165846

Busa, Federica. 1997. Compositionality and the Semantics of Nominals. Brandeis University dissertation.

Davies, Mark. 2002–. *Corpus del español: 100 million words, 1200s–1900s*. (http://www.corpusdelespanol.org) (Accessed 2017-03-03).

Di Sciullo, Ana María. 2009. A View from Asymmetry Theory. In Rochelle Lieber & Pavol Stekauer (eds.), *The Oxford Handbook of Compounding*, 145–177. New York: Oxford University Press.

Downing, Pamela. 1977. On the creation and use of English compounds. *Language* 53. 810–43. DOI: https://doi.org/10.2307/412913

Embick, David & Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 1. 1–53. DOI: https://doi.org/10.1162/ling.2008.39.1.1

Fábregas, Antonio. 2008. Compuestos en las interfaces [Compounds at the interfaces]. In *Seminario de Lingüística Teórica LyCC*. 1–16.

Fernández Domínguez, Jesús. 2016. The semantics of primary NN compounds: From form to meaning, and from meaning to form. In Pius ten Hacken (ed.), *The Semantics of Compounding,* 129–149. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781316163122.008

Jackendoff, Ray. 2009. Compounding in the Parallel Architecture and conceptual semantics. In Rochell Lieber & Pavol Stekauer (eds.), *The Oxford Handbook of Compounding*, 105–129. New York: Oxford University Press.

Jackendoff, Ray. 2010. *Meaning and the lexicon: The Parallel Architecture 1975–2010*. New York: Oxford University Press.

Jackendoff, Ray. 2016. English Noun-Noun compounds in conceptual semantics. In ten Hacken (ed.), *The Semantics of Compounding*, 15–37. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781316163122.002

Kornfeld, Laura Malena. 2009. IE, Romance: Spanish. In Lieber & Stekauer (eds.), *The Oxford Handbook of* Compounding, 145–177. New York: Oxford University Press.

Lees, Robert. 1960. *The grammar of English nominalizations*. Bloomington: Indiana University Press and The Hague: Mouton.

Lees, Robert. 1970. Problems in the analysis of English nominal compounds. In Manfred Bierswisch & Karl Erich Heidolph (eds.), *Progress in Linguistics*, 174–186. The Hague: Mouton.

Levi, Judith. 1978. *The syntax and semantics of complex nominals.* New York: Academic Press.

Liceras, Juana & Lourdes Díaz. 2001. Triggers in L2 acquisition: The case of Spanish N-N compounds. *Studia Linguistica* 54. 197–211. DOI: https://doi.org/10.1111/1467-9582.00060

Lieber, Rochelle and Stekauer, Pavol. 2009. *The Oxford handbook of compounding.* New York: Oxford University Press.

Marqueta-García, Bárbara. 2017. Restrictions in the semantic interpretations of English and Spanish compounds. *Iberia: International Journal of Theoretical Linguistics* 9. 1–35.

Mellenius, Ingmarie & Maria Rosenberg. 2016. The Semantics of compounds in Swedish child language. In Pius ten Hacken (ed.), *The Semantics of Compounding*, 110–128. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781316163122.007

Millikan, Ruth. 1984. *Language, thought, and other biological categories*. Cambridge: MIT Press.

Moyna, María Irene. 2011. *Compound words in Spanish: Theory and history*. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.316

Nicoladis, Elena. 2002. What's the difference between 'toilet paper' and 'paper toilet'? French-English bilingual children's crosslinguistic transfer in compound nouns. *Journal of Child Language* 29. 843–863. DOI: https://doi.org/10.1017/S0305000902005366

Piera, Carlos. 1995. On Compounding in English and Spanish. In Hector Campos & Paula Kempchinsky (eds.), *Evolution and revolution in linguistic theory. Studies in honor of Carlos P. Otero*, 1–13. Washington, DC: Georgetown University Press.

Rosenberg, Maria. 2013. Semantic and formal structure: A corpus-based study of Swedish NN compounds and their French counterparts. In Pius ten Hacken & Claire Thomas (eds.), *The semantics of word formation and lexicalization*, 102–120. Edinburgh: Edinburgh University Press.

Scalise, Sergio & Antonietta Bisetto. 2009. The Classification of Compounds. In Rochelle Lieber & Pavol Stekauer (eds.), *The Oxford Handbook of Compounding*, 34–53. New York: Oxford University Press.

Snyder, William. 1995. *Language acquisition and language variation: The role of morphology*. Cambridge, MA: MIT dissertation.

Snyder, William. 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. *Language* 77. 324–342. DOI: https://doi.org/10.1353/lan.2001.0108

Snyder, William, 2012. Parameter Theory and motion predicates. In Violeta Demonte & Louise McNally (eds.), *Telicity, change and state: A cross-categorial view of event structure*, n.p. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199693498.003.0011

Snyder, William. 2016. How to set the compounding parameter. In Laurel Perkins, Rachel Dudley, Juliana Gerard & Kasia Hitczenko (eds.), *Proceedings of the 6th conference on Generative Approaches to Language Acquisition North America (GALANA 2015)*, 122–130.

ten Hacken, Pius. 2016. Three Analysis of Compounding: a Comparison. In Pius ten Hacken (ed.), *The Semantics of compounding*, 211–232. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781316163122.012

How to cite this article: Toquero-Pérez, Luis Miguel. 2020. The semantics of Spanish compounding: An analysis of NN compounds in the Parallel Architecture. *Glossa: a journal of general linguistics* 5(1): 41. 1–31. DOI: https://doi.org/10.5334/gjgl.901

Submitted: 22 January 2019 **Accepted:** 07 February 2020 **Published:** 30 April 2020

]u[*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press. OPEN ACCESS