



Graduate Theses, Dissertations, and Problem Reports

2018

The Semantics of Spanish Compounding: An Analysis of NN Compounds in the Parallel Architecture

Luis Miguel Toquero Perez

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Toquero Perez, Luis Miguel, "The Semantics of Spanish Compounding: An Analysis of NN Compounds in the Parallel Architecture" (2018). *Graduate Theses, Dissertations, and Problem Reports*. 7138.
<https://researchrepository.wvu.edu/etd/7138>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

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

Luis Miguel Toquero Pérez

**Thesis submitted
to the Eberly College
at West Virginia University
in partial fulfillment of the requirements for the degree of**

**Master of Arts in World Languages Literatures and Linguistics
in
Linguistics & Spanish**

Sandra Stjepanovic, Ph.D., Chair

Jonah I. Katz, Ph.D.

María J. Amores-Aguera, Ph.D.

Department of World Languages Literatures and Linguistics

Morgantown, West Virginia

2018

**Keywords: Spanish, semantics, NN compounds,
Copyright 2018 Luis Miguel Toquero Pérez**

ABSTRACT

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

Luis Miguel Toquero Pérez

Noun-Noun compounds (NN) are a concatenation of two nouns that function as a single unit both morphosyntactically and semantically. One of the main challenges that a study of NN compounding faces is their semantics, and more precisely, identifying the implied semantic relations that hold between the two elements. 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. In this model, a semantic relation between the head and modifier is established by using an external function $F(X, Y)$, which given two nouns N1 and N2 meaning $X1$ and $Y2$ respectively, yields the meaning of the compound.

Given that Spanish NN compounds are unproductive as opposed to languages like English, the aim is to demonstrate whether the model can be applied to Spanish, and if so, to identify the basic functions that hold in Spanish NNs organizing them in a hierarchy based on their frequency. Exploring these questions does not only have a potential to shed light on questions regarding the cross-linguistic applicability of Jackendoff's system, but can also allow us to determine whether the unproductivity of NN compounds in Spanish correlates with their semantic inflexibility. The study also provides a comparison of the semantic relations present in Spanish NN compounds to those attested in Spanish *N de N* constructions (i.e. *buque de guerra* 'ship of war' = *war ship*). Since the latter constructions are very productive in the language, if semantic flexibility correlates with compounding productivity, they are expected to exhibit a wider range of semantic relations as opposed to NNs. Moreover, if it turns out that *N de N* constructions can be successfully interpreted by the basic function model, we will have evidence that they should be treated as NN compounds semantically.

The results show that the model proposed by Jackendoff can be successfully applied to Spanish. Despite the fact that almost all basic functions are attested in Spanish NN compounds, these functions are not equally frequent or common. In fact, there are significant differences with respect to the productivity of the functions in the language, as it becomes evident once they are arranged on a hierarchy according to their frequency. As for the *N de Ns*, most of the functions present in NNs are also attested in *N de Ns* but in a different hierarchical order. Last but not least, taking all the facts discussed in the thesis, it could be argued that semantics is a factor responsible for the peripheral nature of NN compounds in Spanish. Since NN compounds have fewer and less productive basic functions and express fewer semantic relations than alternative *N de N* construction, it seems that NN compounds are semantically restricted, which could be one reason for their unproductivity.

TABLE OF CONTENTS

| | |
|--|-----------|
| Acknowledgments | iv |
| Chapter 1: Introduction | 1 |
| Chapter 2: Theoretical Background | 3 |
| 2.1. What is an NN compound: definition and classification | 4 |
| 2.2. NN compounds in English and Spanish: features and differences | 9 |
| 2.3. The semantics of NN compounds | 14 |
| 2.3.1. The early approaches | 14 |
| 2.3.2 New frameworks for compounding from a semantic point of view | 22 |
| 2.3.2.1 Lieber's (2009) lexical semantic approach to compounding | 22 |
| 2.3.2.2 Jackendoff's (2009, 2016) basic functions | 25 |
| 2.4. Semantic and formal structure: a corpus based study of Swedish compounds and their French counterparts, by Rosenberg (2013) | 32 |
| 3. Research questions and hypotheses. | 33 |
| Chapter 3: Method and Analysis | 35 |
| 3.1. The database: sources and classification. | 35 |
| 3.2. Data analysis and discussion | 36 |
| 3.2.1. Jackendoff's (2009: 128) research questions | 37 |
| 3.2.2 Hierarchy of semantic relations based on the basic functions proposed by Jackendoff | 39 |
| 3.2.3. A comparison with Rosenberg (2013) | 48 |
| 3.2.4. Semantic relations based on the basic functions found in N de N constructions | 49 |
| Chapter 4: Conclusions | 58 |
| References | 61 |

Acknowledgments

First of all I would like to thank my thesis committee for taking the time to read and discuss my research with me, especially my advisor, Sandra Stjepanovic who from the very first moment showed her enthusiasm to guide me and support me in this academic project. I am indebted to her for all the time she has provided me with constructive feedback and for every time she has had to read and revise my NN compound mess. Also thanks to Jonah Katz who always encouraged me to pursue this project since his Bibliography class. I much appreciate his interesting discussions and valuable comments during his out-of-office office hours as well as what he calls *spreadsheets* lessons. I also want to express my gratitude to María Amores whose personal and affectionate advice have helped me dissipate my troubles when I could not see forest for the trees. Besides, if it had not been for her visit to La Pedraja I would probably never have ended up studying at WVU and thus writing this thesis. This thesis is a product of us all.

I am also indebted to many people in the World Languages department while working at WVU. Thanks to Ángel Tuninetti, Sonia Zarco-Real and Rafael Osuna for those conversations at their kitchen tables which feel so rewarding and make you be part of their homes. And of course, thank you to Sergio Robles who, although not part of the thesis committee, has always offered me his help and personal and academic advice even when we were playing soccer together.

Thanks to those Spanish GTA colleague-friends and soccer teammate friends who I have met during my stay in Morgantown and who have always stood by my side. Special thanks to Igor and Sergio for the year we spent together and the unforgettable and miscellaneous experiences we have lived. Also thanks to Tamara, my roommate, for her help and friendship. I also want to thank Clara for being the true and *life guru* friend she is and always will be and to Lisa for all the favors, laughs and stories shared over this period of time in Morgantown. And of course, very special thanks to my girlfriend Carmen who for the last 7 months has shown an incredible amount of support and love which have helped me get through difficult times.

And finally, I must say thank you to my family and friends from La Pedraja for supporting and always encouraging me in the decisions I take. Although they are far, I always feel them near.

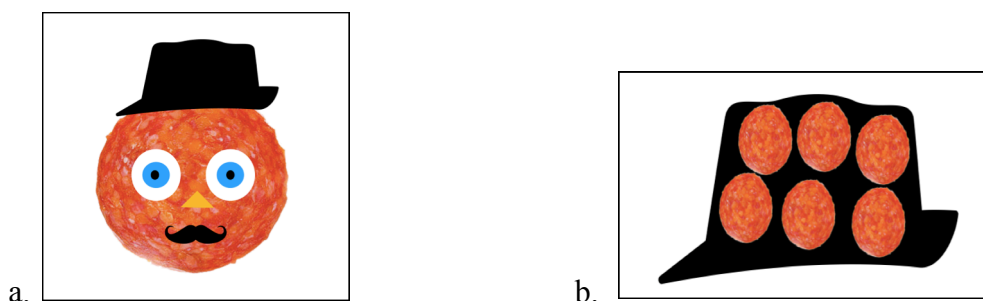
Chapter 1

Introduction

This thesis 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.). This example can be illustrated by—the pictures provided in (0) below:

(0) *pepperoni hat picture*



The pictures illustrate that the same NN compound *pepperoni hat* can be used to designate at least two different entities. The picture in (0a) stands for a 'hat for pepperoni'—since we have a pepperoni slice that is wearing a hat, while the picture in (0b) stands for a 'hat that has pepperoni on it', given that the hat is decorated with pepperoni slices.

Identifying these implied semantic relations that exist between the two nouns is the aspect that this thesis 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: this function F connects the semantic features of X (i.e. the modifier) and Y (i.e. the head). 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: $[\text{HAT}_2^a; [\text{HAVE}(\alpha, \text{PEPPERONI}_1 \text{ ON } \alpha)]]$. Another possible meaning for *pepperoni hat* is ‘hat for a piece of pepperoni’ which could be represented as the following: $[\text{HAT}_2^a; [\text{HAVE}(\text{PEPPERONI}_1, \alpha)]]$. 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. Since then the model has been successfully applied to French and Swedish by Rosenberg (2013). Swedish is very similar to English with respect to NN compounding because the process is very productive in both languages. However, NN compounding in French is not common, which is what we find in Spanish as well. That said, one goal of the thesis is to see whether Jackendoff’s basic function model can also be applied to Spanish. 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 unproductivity of NN compounds in Spanish correlates with their semantic inflexibility. In this respect, it will be interesting to compare the results obtained for Spanish NN compounds with the results obtained by Rosenberg (2013) for French NN compounds, given that NN compounding is not productive in either language. Rosenberg (2013) showed that only two of Jackendoff’s thirteen basic functions are present in French NN compounds. If it turns out that Spanish NN compounds realize only few of Jackendoff’s functions as well, then we may have evidence that their semantic inflexibility may be a factor that limits the production of new NNs in these languages.

In addition, given that the alternative productive 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 N* 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 (2001) 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 I have~~ obtained from the *Corpus del Español* (Davis 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 chapters, I will provide the analysis, present the results, and discuss their implications for the predictions and questions outlined above.

The thesis is organized into four chapters including this introductory chapter. Chapter 2 provides the theoretical background and hypotheses to be tested in the thesis. Chapter 3 describes the data collection method, presents the results of analysis of the data and provides discussion of the results. Finally, Chapter 4 is the conclusion of the thesis that summarizes the main findings. Bibliographical references follow.

Chapter 2

Theoretical Background

This chapter discusses the theoretical background necessary for understanding the semantics of NN compounds. It first focuses on the classification of compounds, in order to introduce some terminology used in the study. Then, it turns to outlining the characteristics of Spanish NN compounds by comparing them to NN compounds in English. Finally, 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.

2.1. What is an NN compound: definition and classification

This section reviews different approaches to the classification of compounds. After pointing out some problems with the traditional approaches first, it will then focus on Scalise and Bisetto's (2009) classification, which I will use when referring to different types of compounds.

The term NN compound refers to a concatenation of two nouns that function as a single unit morphosyntactically and semantically (Jackendoff 2016), or as Bauer (2003: 88) puts it: "it is the formation of a new lexeme by adjoining two or more lexemes". One interesting issue regarding compounds is their semantic interpretation, or, in other words, the question of how the meaning of the whole compound is related to the meaning of its constituents. This question will be the focus of my study. Given that, as we will see below, part of the answer to this question depends on the type of compound, I will first provide a brief overview of different types of compounds.

The classification of compounds has been the subject of much debate in the literature, but some of the classifications proposed are not very satisfactory due to terminological problems and inconsistent classification criteria. Traditionally, analyses have focused on root and synthetic compounds. Root compounds are compounds whose head is typically not deverbal, or whose non-head constituent does not function as an argument of the verb from which the head is derived; that is, the constituents are roots (Scalise and Bisetto 2009: 36-37). On the other hand, synthetic compounds are compounds whose head is a deverbal element derived by affixation (for example, an *-er* noun, an *-ing* adjective or noun, and a passive participle in English), with the non-head constituent functioning as its argument. This difference can be illustrated in example (1) below.

- (1) a. *police dog*
b. *truck driver*

In (1a), the root compound consists of two word roots, *police* and *dog*, which lack any overt verbal relation. On the other hand, (1b) involves a synthetic compound, where the noun head, *driver*, is derived from the verb *drive*, with the non-head *truck* functioning as its complement '[someone (-er)_{AGENT}] drives [a truck_{THEME}]'. In short, while the term *root* only refers to the uninflected and non-derivational status of the corresponding lexemes such as (1a), the term *synthetic* indicates a derivational nature of the head.

However, this distinction between root and synthetic compounds cannot be extended to other languages, such as Romance languages, as illustrated in (2) from Spanish:

- (2) a. *police dog*
 b. *perro policía* (dog police = ‘police dog’)
 c. *truck driver*
 d. **conductor camión* (driver truck = ‘truck driver’)

While English NN compounds such as (2a) involve the combination of two bare roots *police* and *dog* (e.g. they lack inflection and derivation), the equivalent lexemes of Spanish in (2b) have an end vowel that is not part of the root but bears grammatical information (i.e. this is called Word Marker which I will deal with in the next section): the *-o* in *perro* indicates masculine gender, for instance. Therefore, Spanish NN compounds are not constituted by bare roots only since they have an affix attached to the root adding morphological information. Besides, as pointed by Scalise and Bisetto (2005), Spanish does not seem to have synthetic compounds as defined above: while (2c) is grammatical in English, the equivalent structure in Spanish (2d) is not. Given this, we cannot use the terms root and synthetic to classify compounds, because neither term seems to be able to characterize compounds in languages like Spanish. Thus, we need to adopt a more satisfactory classification.

In addition to the terms root and synthetic, the notions of endocentricity, exocentricity and coordination have also been considered in the literature. The first two notions (endocentric and exocentric) are concerned with the presence or absence of a semantic head in the compound respectively. For example, *police dog* is an endocentric NN compound because its semantic head is *dog* which is overtly expressed in the compound (that is, a *police dog* is a type of *dog*); whereas *bird brain* is an exocentric NN compound, because the semantic head is not overtly present in the compound. Thus, a *bird brain* is not a type of *brain* but a type of *person*. The third notion (coordinate) refers to compounds whose constituents bear an implied relationship like ‘and’ or ‘or’, as in *king emperor*: the compound denotes someone that is both a king and an emperor at the same time. In some classifications such as (Haspelmath 2002: 88), these three notions are considered to be at the same level of categorization. That is, endocentric, exocentric and coordinate are separate categories at the same level. However, there is a problem with this classification, since the three categories are established based on different criteria: the distinction between endocentric and exocentric compounds is

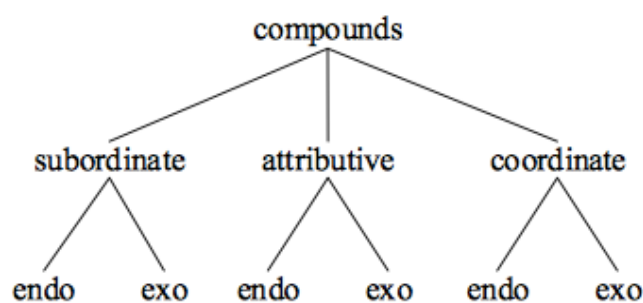
based on the presence/absence of head, whereas the notion of coordinate compounds deals with the semantic relation between the two members (i.e. coordination: an [XY] is both a kind of X and Y). Furthermore, coordinates are generally a subset of endocentric compounds as it can be seen in (3):

(3) a. *singer songwriter*

(3a) is a coordinate endocentric compound characterized by the presence of a head *songwriter* which is linked to the modifier by the implied conjunction ‘and’ that underlies its meaning (‘songwriter that is a singer’).

These are only a few of the problems that have arisen from looking at previous classifications of compounds. Scalise and Bisetto (2005) propose a classification that attempts to solve issues presented above. Their classification is based on the grammatical/semantic relations which they believe hold between the constituents of compounds. Based on these relations, they divide compounds into three types: subordinate, attributive and coordinate. This can be seen in Figure 1:

Figure 1 taken from Scalise and Bisetto (2009: 45)

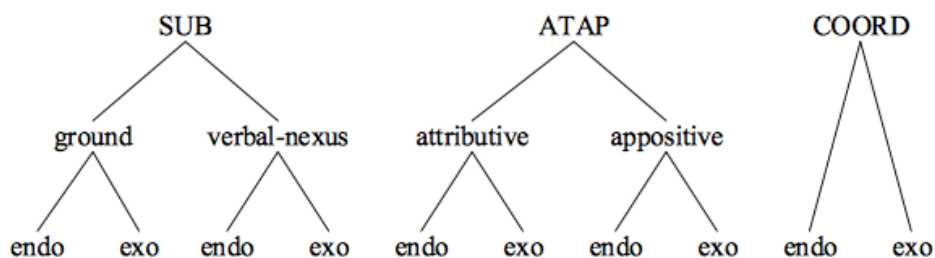


Subordinate compounds are those that can be defined as having a complement-head relation, such as *truck driver*, where *truck* is the complement of the verbal predicate *drive*, from which the head noun *driver* is derived. According to Scalise and Bisetto (2005), these kinds of compounds also include modifier-head relations such as *country house* in which *country* is a modifier and specifies the location of *house*. *Killjoy* is an example of subordinate exocentric compound: it does not refer to ‘joy’ but rather to a ‘person’ that literally *kills* the joy of everyone in a celebration’. Thus, there is an argument-predicate relation between the two constituents of the compound (*joy* is an argument of the predicate *kill*), but the semantic head is not explicitly stated. The second branch includes attributive compounds: those where the non-head (also a modifier as in *country house*) denotes, often metaphorically, a property of the head: e.g.

castle cloud meaning ‘a cloud that looks like a castle’¹. An exocentric counterpart would be *loudmouth* whose semantic head (i.e. *person*) is absent and refers to ‘tactless person that tends to talk too much in an offensive manner’. In the case of coordinate compounds, the same definition and examples as above can be used: e.g. compounds whose constituents are linked by an implicit conjunction (e.g. ‘and’, ‘or’), and they can also be endocentric and exocentric.

This could be considered the matrix level of analysis, but Scalise and Bisetto (2009) expand their analysis and distinguish an additional level to consider the different semantic relations that can hold between the non-head and head in a compound. This new proposal can be seen in Figure 2:

Figure 2 taken from Scalise and Bisetto (2009: 50)



Regarding subordinate compounds, a distinction is made between ground and verbal-nexus compounds. Ground compounds are those that can be interpreted by an implied semantic relation (for example, a Recoverably Deletable Predicate or *RDP* in terms of Levi (1978), or basic function in terms of Jackendoff (2009, 2016) discussed below) determined by the encyclopedic knowledge connected to the two constituents. The term verbal-nexus refers to the compounds traditionally known as synthetic whose interpretation depends on the base verb that is present in the deverbal head noun. This distinction is seen in example (4):

- (4) a. *olive oil*
 b. *coffee grinder*

¹ Other scholars such as Arnaud and Renner (2014) classify these two types of compounds as subordinate, where *country house* is relational because its meaning can be explained in terms of a predication with two arguments, and *castle cloud* is attributive because the semantic modification consists of attributing features of the non-head to the head. However, this classification is problematic because I believe there is no need to treat these two compounds differently: in both cases the non-head should be treated as a modifier rather than a complement.

(4a) is a ground compound because according to Jackendoff's basic functions (2009, 2016), it can be paraphrased as 'oil that is made from olives'. There is nothing explicitly determining the semantic relation between the two nouns, rather the meaning of the compound follows from our encyclopedic knowledge that *oil* can be made from *olives*. On the other hand, (4b) is a verb-nexus compound, because the verb *grind* in the base of the head noun *grinder* establishes an interpretation consistent with its argument structure: '[something (-er)_{INSTRUMENT}] grinds [the coffee_{PATIENT}]'.

The second group labeled as ATAP stands for attributive-appositive. The main difference between attributive and appositive compounds is that while in the former the 'property' denoted by the non-head is expressed with an adjective (i.e. *blue cheese*), in the latter that 'property' is expressed by a noun (i.e. *mushroom hat*). As to coordinate, no further stratum is identified as there are no available "distinctions that call into action the subtype of grammatical relation between the two constituents" (2009: 52).

The new proposal by Scalise and Bisetto offers a fine-grained and consistent classification of compounds that can be applied to various groups of languages, such as, for instance, English, French, Spanish and Swedish among others. Besides, it overcomes the terminological problem introduced by former classifications. This is why I will use it in my work, which, as we will see below, will focus on subordinate compounds, and more specifically on ground endocentric ones. Nevertheless, one of the weaknesses I find in Scalise and Bisetto's taxonomy is the unclear distinction between appositives and subordinate ground compounds. Scalise and Bisetto (2009: 52) claim that the metaphor is what allows them to tell the difference between compounds such as *castle door* (subordinate ground), where *castle* instantiates a 'part-of' relation, and *castle cloud* (appositive), where *castle* is not interpreted literally but as having some kind of shape resemblance to the entity 'castle'. However, given that Jackendoff (2016) and Krott et al. (2009) also analyze ground compounds in terms of metaphorical relations such as SIMILAR (X,Y) and LIKE respectively, this casts doubt on Scalise and Bisetto's (2009) reason to distinguish the two groups. Given that they can be analyzed identically, I will consider appositives a subtype of subordinate ground compounds. And as ten Hacken (2016: 212) states "[a]s long as we use labels of compound classes as pretheoretical descriptions, without any theoretical terms attached to them, there is no reason to bother too much about such issues". That said, as we will see below, my study will concentrate on subordinate ground compounds, including

ground compounds *per se* as well as appositives from Scalise and Bisetto's (2009) classification.

Having discussed the different types of compounds and the classification that I will follow in the thesis, I now turn to review the properties of Spanish and English compounds in order to identify the differences between NN compounds in the two languages. This will allow me to show that one of these differences, primarily the unproductivity of NN compounds in Spanish, is yet to be accounted for in terms of the semantics of compounds.

2.2. NN compounds in English and Spanish: features and differences

Despite the fact that subordinate ground compounds can be found in both English and Spanish, English NN compounds differ from Spanish NN compounds in two main areas: their productivity and directionality. These syntactic properties have been widely discussed by scholars such as Bauer (1983), Piera (1995), Snyder (1995, 2001) and Liceras and Diaz (2001), among others. Productivity deals with the frequency of such constructions in the language. Directionality covers the syntactic organization in terms of headedness: that is, which of the two nouns making up the compound is the semantic and categorical head and in which position the head is placed.

Contrary to Germanic languages like English, where the formation of endocentric NN compounds is utterly productive², 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) has later redefined. This new formulation is given in (5):

- (5) 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.

The term 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 ". This formulation implies that the

² 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'.

compound must be interpreted a hyponym of the head whose meaning is influenced by the semantic content of the modifier.

Thus, Snyder (2012) divides languages as [+TCP] or [-TCP] regarding whether compounding is productive 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, 2013) 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 while [-TCP] do not. A question that is raised in this thesis is whether this hypothesis is true for Spanish.

Furthermore, Snyder (1995, 2001, 2012) argues that [+TCP] languages like English with productive NN compounding also allow complex predicates such as (6):

- (6) a. *He hammered the metal flat.*
 b. *He picked the letter up.*

(6a) is a resultative structure which consists of an adjective *flat* that denotes the state of *the metal* as a result of being *hammered*: i.e. *he* causes *metal* to become *flat*. (6b) involves a phrasal verb composed of the lexical verb *pick* and the particle *up* which function as a single unit. On the other hand, Spanish, being [-TCP], does not allow adjectives or particles to be linked with verbs to construct complex predicates. The Spanish equivalents of the examples in (6) are illustrated in (7):

- (7) a. *Golpeó el metal (*liso)*
 b. *Cogió la carta (*arriba)*

According to Snyder, regarding NN compounds, languages³ that have the [+TCP] feature constitute the superset option of the parameter as they allow both productive NN compounding and alternative constructions where, for example, the

³ A longer list of those languages illustrating the [+/-TCP] feature of the parameter can be found in Snyder (2001) cross-linguistic survey.

modifier of the head noun is not another noun, but a prepositional phrase (e.g. *mountain top* and *the top of the mountain*). The latter are the most productive structures in [-TCP] languages like Spanish, with NN compounds being more peripheral. Prepositional phrases such as *casa de madera* ('house of wood' > *wooden house*) are more frequent than NN compounds such as *perro policía* ('dog police' > *police dog*). The [-TCP] languages, therefore, act as the subset option.

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 (8); and English compounds are recursive, while Spanish ones are not (9):

- (8) a. *spider man*
 b. *hombre araña*
 (9) a. *movie spider man*
 b. **hombre araña película*

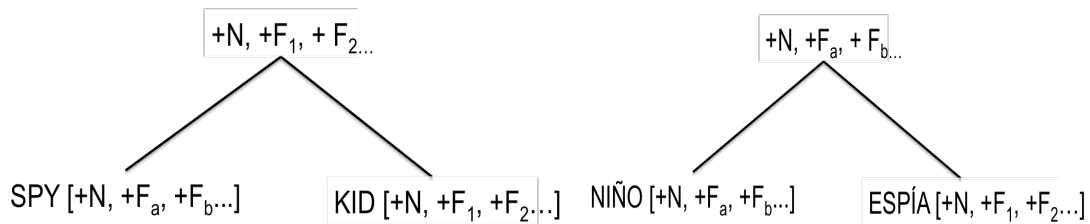
Examples in (8) demonstrate that while the head of the English compound in (8a) is the right-hand element (e.g. *man*), the head of the Spanish compound is placed on the left (i.e. *hombre*). This means that directionality in Spanish is reversed compared to English. In terms of headedness, we can adopt Arnaud and Renner (2014: 2) terminology. In both English and Spanish the head corresponds to the semantic, categorical 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 categorical 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*; *hombres/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).

Taking into account that the source of a compound is a noun plus its grammatical and semantic features, Piera (1995) uses a simplified tree in figure 3 to represent the structure of compounds in terms of headedness:

Figure 3: simplified tree following Piera's analysis (1995: 3)

(a) English

(b) Spanish



As shown in these trees, compounding is the result of the projection of the head under branching. In English, it is the features of the right-hand element that project. In Spanish compounds, though, it is the features of the left-hand element that have to project in order to be available for the whole NN compound.

Furthermore, Piera (1995) argues that the contrast in grammaticality between English (9a) and Spanish (9b) can be explained in terms of the internal structure of nouns in the two languages. He proposes that 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 WM⁴. This difference is illustrated in (10):

- (10) a. [[niñ] **o**]
 b. [kid]

The ‘**o**’ in (10a) 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” (p. 6). Adjunction to the left of noun in Spanish is, therefore, not possible. In other words, adjunction must be done to the right for a compound to be grammatical as in (11a):

- (11) a. [[niñ-] WM] [[espí-] WM]
 b. *[[niñ-] WM] [[espí-] WM] [[policí-] WM]
 c. *[[hombre] WM] [[arañ-] WM] [[películ-] WM]

The head *niño* has a double bracket to its left and thus it prevents the adjunction of another noun on its left. On the other hand, since there is only one bracket to its right, a noun (*espía*) can be adjoined to the right. In addition, the Double Bracket Restriction can be applied to account for the ungrammaticality of recursive NN compounds such as

⁴ Note that according to Piera, “English synthetic compounds pose no additional problems within the standard analysis in which their deverbal head is in turn headed by an affix: the leftmost edge of the resulting word is still single bracketted” as in [+N...{write, [+V,...]} {er. {+N,...}}].

(11b) and (9b) repeated here as (11c). The compounds in (11b) and (11c) have a double bracket to the right and disallow the insertion of another noun.

While leftward adjunction to a noun is not possible in Spanish, it is possible in English NN compounds, because nouns in English only have one bracket to their left, as it has been shown in (10b), due to the lack of a WM. Therefore, *spy* can be placed to the left of *kid* and result in a grammatical NN compound in English. What is more, this compound can be compounded further. Examples are shown in (12):

- (12) a. [*spy* [*kid*]]
 b. [*police* [*spy* [*kid*]]]

In (12a), the head of the compound is *kid* which only has one bracket to its left and thus enables the insertion of another noun on the left. (12b) is grammatical because the whole compound *spy kid*, which is acting as a unit, has only one bracket to the left and allows the adjunction of another noun on the left⁵.

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* ('table window' > *window table*) are ungrammatical. In fact, this compound shares the same internal structure as (11a). This can be seen in (13):

- (13) a. [[*niñ-*] WM] [[*espí-*] WM]]
 b. *[[*mes-*] WM] [[*ventan-*] WM]]

The representations in (13a) and (13b) demonstrate that both nouns have the same syntactic structure according to Piera's analysis. Therefore, there should be no reason to consider (13b) ungrammatical. In spite of this, **mesa ventana* is not an attested compound in Spanish. This is consistent with Snyder's (2012) claim that [-TCP] languages have no productive novel compounding and that these language only allow lexicalized compounds (2012: 5).

One reasonable hypothesis that can be proposed to account for the question why Spanish does not have NN compounds like *mesa ventana*, which would imply a locative

⁵ Piera also acknowledges that the DBR for those NN compounds involving right branching such as [[[*finance*][*committee*]][*chairman*]] holds "as long as the leftmost edge of its would-be head is single bracketed. The addition of any number of simple items in English will only increase the depth of embedding of the compound head".

relation, is that 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. 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 productivity of Spanish NN compounds may be their semantic inflexibility. One of the goals of this study is, precisely, to investigate whether NN compounds in Spanish do not allow the realization of all semantic relations that are attested in NN compounds in languages like English, which may be a factor that causes their unproductivity.

In order to learn more about the semantics of NN compounds, the next section will be concerned with a thorough overview of the different approaches that have been proposed in the literature.

2.3. 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.

This section offers an overview of the main approaches and models that have been proposed to explain the semantics of NN compounds, starting with the early generative approaches and moving towards more recent analyses.

2.3.1. The early approaches

Generativism influenced our thinking about compounding. Lees (1960) was one of the first works that applied the Chomskyan framework to the analysis of the compounding process. This work would set the bases for research on compounding for

the following two decades. According to Lees, compounds “are not sentences themselves, but rather they are noun-like versions of sentences” (1960: 54). He proposes that compounds are a result of transformations. An example of the derivation of a compound in his system is provided in (14):

- (14) a. *the soup has mushrooms*
 b. *soup with mushrooms*
 c. *soup mushrooms*
 d. *mushroom soup*

According to Lees, the compound in (14d) is derived from the sentence in (14a), which is its D-structure (i.e. D-S), by a series of transformational rules. First of all, the sentence in (14a) must be nominalized into a Noun + Preposition + Noun structure which is what we see in (14b). Next, the deletion of the preposition *with* takes place, as it can be seen in (14c). The crucial rule then is to reverse the order of the components to create a complex Noun whose structure is similar to that of a Noun Phrase (i.e. modifier + head): that is, (14d). In other words, the examples in (14) share the same D-S, but some elements have been deleted and reversed in the process of deriving (d) from (a), (b) and (c). Taking this into account, Lees classifies compounds according to the syntactic roles that the nouns play in the underlying sentence ~~undergoing deletion~~. Some of them are illustrated in table 1 below:

Table 1: label sample of Lees' compound classification

| Label sample | Examples |
|------------------------------|---|
| Subject-Predicate | <i>Boyfriend</i> (e.g. <i>The friend_{Subj} is a boy_{Pred}</i>) |
| Subject-Object | <i>Mushroom soup</i> (e.g. <i>The soup_{Subj} has mushrooms_{Obj}</i>) |
| Subject-Prepositional Object | <i>Kitchen table</i> (e.g. <i>the table_{Subj} is in the kitchen_{PObj}</i>) |

However, one of the main criticisms of Lees' work focuses on the content deleted by the transformations. Scholars like Katz and Postal (1964) note that the deletion of lexical items present at D-S would make transformations too powerful. The example they use to illustrate this is given in (15) (p. 81):

- (15) a. *John is reading an inscription on a tombstone.*
 b. *John is reading a book on morphology.*
 c. *John is reading.*

They state that (15c) cannot be derived from (15a) or (15b) by deletion of elements present in (15a) and (15b) but not in (15c), given that the meaning of sentences in (15a), (15b), and (15c) is different. Therefore, (15c) cannot share the same D-S with (15a) or

(15b). As a result, they propose a “constraint on transformational derivations” (1964: 81) according to which whatever is deleted must be explicitly stated and must be able to be recovered. Chomsky (1965: 144-5) also adopts this constraint and limits the deletion mainly to dummy elements. Given the constraint, the sentence in (15b) cannot be derived from (15a) via a deletion transformation, because the deleted material in (15b) cannot be recovered. The same logic can be applied to compounds.

Lees (1970) revised his theory of compounding and proposed a new direction in an attempt to improve it. The solution he puts forward is twofold. On the one hand, he substitutes the labels based on grammatical relations illustrated in table 1 for thematic roles (i.e. agent, instrument, location etc.). On the other hand, he suggests that NN compounds contain an underlying “generalized verb” as he puts it (1970: 181). Taking into account this implied verb, he sets a transformational sequence which is given in (16):

(16)

$$\begin{array}{ccc} V & O & I \\ | & | & | \\ N_1 & & N_2 \end{array} \longrightarrow N_2 \text{ V-S } N_1 \longrightarrow N_2 + N_1$$

The sequence in (16) shows three different stages in the derivation of NN compounds. At the D-S, O stands for complement and I for instrument. The intermediate structure would involve a sort of paraphrase such as the one in (14a), and the S-Structure presents the final result: the compound. In addition, he claims that there are two subsets of these “generalized verbs” depending on “the minimal set of semantic features which characterize all variants in the sets [of verbs]” (1970: 182). The first set is composed of the verbs *impel*, *propel*, *energize*, *activate*, *power* whereas the second is comprised of *cause*, *engender*, *produce*, *yield*. An example following Lees’ (1970) model is seen in (17):

(17) *steam engine*: energize engine_O steam_I → steam energizes the engine → steam engine

The generalized verb introduced by Lees (1970) recovers part of the meaning that is implied, which seems to be a welcome result in light of the problem of recoverability of deletion mentioned above. However, given that these generalized verbs only account for agentive predicates where the theta structure is more straightforward, Lees does not fully solve the problem of recoverability since his set of

generalized verbs leaves out non-agentive predicates. Therefore, this proposal had to be improved.

Considering Lees' work, other scholars such as Levi (1978) and Downing (1977) started to look at the semantics of compounds from a generative point of view and tried to come up with possible underlying predicates that could satisfy transformational constraints and render an accurate meaning of the compounds. Lees (1970) only offered some suggestions to further explore his previous ideas on compounding, but Levi (1978) "offers an account of the semantics of compounding that is still referred to as a benchmark today" (ten Hacken 2016: 3). The relevance of her work lies in the development of several predicates that try to solve the problem of the irrecoverably deleted material that Lees (1960, 1970) could not overcome. In other words, Levi's (1978) account does not involve violations of the constraint on transformational derivations proposed by Katz and Postal (1964). Like Lees (1960), she assumes that NN compounds are derived from sentences by the deletion of the predicate, but she suggests a finite set of Recoverably Deletable Predicates or *RDPs* to capture the underlying relation between the two members of the compound. She limits the number to nine, but three of them are reversible (e.g. that is they can be used in the active or passive form). The RDPs she proposed are listed below in table 2:

Table 2 RDPs taken from Levi (1978: 76-77)

| RDPs | Examples |
|------------------------------|-------------------------|
| CAUSE ₁ (active) | <i>Accident weather</i> |
| CAUSE ₂ (passive) | <i>Peer pressure</i> |
| HAVE ₁ (active) | <i>Picture book</i> |
| HAVE ₂ (passive) | <i>Student power</i> |
| MAKE ₁ (active) | <i>Honey bee</i> |
| MAKE ₂ (passive) | <i>Stone wall</i> |
| USE | <i>Watermill</i> |
| BE | <i>Queen bee</i> |
| IN | <i>Mountain top</i> |
| FOR | <i>Doghouse</i> |
| FROM | <i>Coconut oil</i> |
| ABOUT | <i>Crime story</i> |

At D-S these twelve RDPs are attached to the head noun as a modifier in the form of an embedded clause. Levi dedicates many pages to explain how this complex structure is derived into an NN compound. A reduced and simplified version of Levi's (1978) explanation is given in (18) below for the compound *peer pressure* using the predicate CAUSE:

- (18) *peer pressure*
- a. pressure which peers cause
 - b. pressure which is caused by peers
 - c. pressure which is peer-caused.
 - d. peer-caused pressure.
 - f. peer pressure.

According to Levi, (18) is derived from an active relative clause which is then turned into a passive one in (18b). The next step is the creation of a compound adjective in (18c). After that, the WH- relative is deleted and predicate is preposed as in (18d). Finally, the former predicate is deleted because the semantic relation between the modifier *peer* and the head *pressure* is assumed to be CAUSE.

As it can be observed in table 2, the first three predicates can be reversed. In other words, they can be used either actively or passively. One of the consequences of Levi's analysis is the ambiguity of compounds: that is, a novel NN compound that is yet to be lexicalized should have the twelve possible readings since there are twelve RDPs. Some of these readings are more likely than others, and sometimes ambiguity diminishes due to the semantic content of lexical items or the extralinguistic knowledge, as she notes (1978: 158). Another consequence is the generality of RDPs. Levi (1978) assumes that these general predicates have broad meanings and include more specific relations that otherwise should have been individually listed. For instance, Levi (1978) proposes that interpretations involving *on* and *at* are surface realizations of the underlying predicate IN which also includes more specific relations such as INHABIT or DURING. However, some problems arise from this account. Instead of providing the particular meaning of a compound, the predicates seem to render the matrix set of interpretations that native speakers assign to NN compounds and complex nominals, which results in a very poor meaning. If we take the example *doghouse* in table 2, analyzed under the RDP FOR, the possible paraphrase according to Levi's (1978) system is 'a house for the dog' where *dog* simply is the recipient. However, the RDP FOR does not specify the purpose of the house: a house for the dog *to live in*. That is, RDPs assign the basic meaning to the compound, but leave aside more complex aspects of meaning. As we will see below, this is what Jackendoff (2009, 2016) attempts to improve with his model. In addition, she claims that each compound should have has twelve possible interpretations, although some of them may be more far-fetched than others; for instance, *apple cake* in (19):

- (19) a. the cake HAS apples.

- b. the cake IS MADE OF apples.
- c. % the cake CAUSES the apples.
- d. % the cake is FOR apples.

Apple cake in (19) can be paraphrased by HAVE or MADE OF; it seems obvious that the meanings in (19c) and (19d) are extremely unlikely in light of our world knowledge, but it is equally obvious that they can be made felicitous by just creating a discourse context where, for example, a woodland fairy goes around giving cakes to various pieces of vegetation, and this one is the cake for the apples. That is why the % is used in (19c-d): they are semantically odd given our world-knowledge, but there are not semantic or grammatical restrictions to them and thus they could be possible.

To this we have to add the notion of the vagueness of the meaning assigned by the predicates she proposes because rather different semantic relations can be involved under the same predicate. A classic example of this controversy is the one introduced by ten Hacken (1994) presented here as (20):

- (20) a. *fertility pills*
 b. *headache pills*

Both (20a) and (20b) could be analyzed as having the same relation FOR. Nevertheless, the purpose of each of the *pills* is different according to the information provided by the semantic non-head: (20a) exemplifies the purpose of *inducing* fertility while (20b) refers to *curing* headaches. Therefore, another drawback is the extreme polysemy of the predicates. The extreme polysemy of the predicates often prevents an NN compound from being assigned a unique reading.

These two early approaches are based on the premise that compounds can be derived from sentences. They assume that NN compounds are reduced versions of sentences at S-Structure. However, a relevant study questioning this is the analysis of compounding in English offered by Downing (1977).

As opposed to Lees (1960, 1970), she argues that even though compounds can be paraphrased by a sentence, they do not actually come from sentences because sentences and compounds have distinct types of meanings. Overall, sentences have truth conditions, whereas compounds refer and do not assert. Furthermore, establishing the meaning of a compound with a paraphrase does not mean that the compound has to be derived from an underlying structure.

In her study, she tries to discover the conditions on compounding by looking at novel compounds and evaluating “the nature and relative frequency of the semantic

relationships underlying attested but non-lexicalized compounds” (Downing 1977: 817). She collects the compounds from different written sources and classifies them according to their underlying semantic relations. The semantic relations that she proposes can be seen in table 3:

Table 3 Semantic relations proposed by Downing (1977)

| Semantic relations | Examples |
|--------------------|---|
| Whole-part | <i>Duck foot</i> ‘a foot that is part of a duck’ |
| Half-half | <i>Giraffe-cow</i> ‘animal that is both a cow and a giraffe’ |
| Part-whole | <i>Pendulum clock</i> ‘a clock of which a pendulum is a part’ |
| Composition | <i>Stone wall</i> ‘a wall composed of stone’ |
| Comparison | <i>Crocodile pin</i> ‘a pin that is similar to a crocodile’ |
| Time | <i>November rain</i> ‘rain that occurs in November’ |
| Place | <i>Apple juice seat</i> ‘a seat where you can drink your apple juice’ |
| Source | <i>Olive oil</i> ‘oil made from olives’ |
| Product | <i>Honey bee</i> ‘bee that produces honey’ |
| User | <i>Flea trampoline</i> ‘trampoline used by fleas’ |
| Purpose | <i>Hedge hatchet</i> ‘a hatchet that is used for the hedges’ |
| Occupation | <i>Piano man</i> ‘man whose occupation is the piano’ |

Just like Levi, Downing proposes 12 semantic relations, but some of these relations clearly overlap (i.e. part-whole and whole-part, and user and purpose). Besides, her results prove that there are indeed interpretative constraints on compounding, as there must be a logical relation between the modifier and head, but the meaning of these two elements must not overlap. This means that for speakers to provide an interpretation of a compound, they must be able to establish some common ground between the meaning of the modifier and the meaning of the head, but at the same time modifier and head cannot overlap in meaning. This can be seen in (21):

- (21) a. # / % *night democrat*
 b. # / % *cow pony*

According to Downing, (21a) shows a lack of association between *night* and *democrat*. One cannot simply be a *democrat* only at *night*. Thus, it is semantically odd. As to (21b), its semantic anomaly resides in the fact that the two elements denote related entities; in fact, *cow* and *pony* are cohyponyms of ‘mammal’ that would not interbreed well. Therefore, there is co-occurrence. Nevertheless, the restrictions proposed by Downing (1977) should be taken with a pinch of salt because for some speakers both compounds can be considered correct: the former could be ‘someone who is hiding their political affiliation during the day’, and the latter ‘a mixture of a pony and a cow’ which

would perfectly fit her category of *half-half*. Thus, the % symbol has been used to indicate that the meaning may be possible for some speakers. So, a good question to ask is to what extent the constraints proposed by Downing are valid.

Downing suggests that “a small set of relationships is generally favored; and the appropriateness of a given relationship is also dependent on its permanence, its predictability in context, and on the semantic class of the head noun” (1977: 836). In other words, she proposes that although compounds should not have a completely predictable relation between their parts, interpretability has to be guaranteed taking the meaning from the elements in the compound or from the context.

Although Downing (1977) starts to question the claims made by previous studies such as Lees (1960, 1970), she is unable to come up with a set of semantic relations that can render all possible readings for NN compounds. Furthermore, her account faces the problem of generality of predicates as well, because the relations she proposes are wide and do not allow more specific interpretations. That is why the introduction of lexicalist ideas into our thinking about the meaning of compounds was important, because it provided a way to resolve some of the problems that early accounts such as Downing (1977) and Levi (1978) faced. Up until this point, scholars had tried to establish the meaning of a compound focusing on the compound as a whole instead of examining the individual parts of which it is made (though Downing (1977) starts exploring this field). Allen (1978) addressed compounds from a lexicalist point of view and proposed two rules to analyze the meaning of NN compounds which are illustrated in (22) (taken from Hacken (2009: 72))

- (22) a. IS-A Condition:
If $[X\ Y]_Z$ is a compound, Z is a Y .
b. Variable R condition:
If $[X\ Y]_Z$ is a primary compound, the relationship between X and Y is variable. The meaning of Y makes a number of slots available and X may fill any of these slots that is compatible with its own meaning.

The proposition in (22a) is based on the claim that the head of the compound is the hyperonym of the whole entity: e.g. a *pineapple pen* is a type of *pen* and not a type of *pineapple*. The second proposition, presented here in (22b), puts an emphasis on the meaning of the individual component. According to her model, the Variable R condition in (22b) provides a range of possible meanings for a particular compound. This range of meanings is specified in terms of sets of semantic features of each compound constituent: that is, the semantic content of the first constituent can “fill in any one of

the available feature slots in the feature hierarchy of the second constituent” (Allen 2000: 203). However, in order to fill the slots, there must be some compatibility of the semantic features. This model is very similar to the one proposed by Lieber (2009), which is illustrated in the next section.

If we compare the approaches described above that came after Lees (1960, 1970), their proponents all agree that Lees’ (1960, 1970) account involves impossible transformations due to the deleted content which cannot be recovered. Downing (1977), however, tries to propose a finite set of semantic relations that can be used to account for the implied meaning of the compound and attempts 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) proposes 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. The conditions proposed by Allen (1978) introduce 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 thesis, I review two frameworks in which the semantics of compounds is the central issue: Lieber’s (2009) lexical semantics model and Jackendoff’s (2009, 2016) basic functions model within the Parallel Architecture framework.

2.3.2 New frameworks for compounding from a semantic point of view

2.3.2.1 Lieber’s (2009) lexical semantic approach to compounding

Lieber develops a decompositional system in which lexical items are dissected into minimal semantic units made of features. She breaks down the meaning of words into two basic parts: the semantic/grammatical skeleton and the semantic/pragmatic body. The skeleton consists of seven features that are relevant to syntax and are hierarchically organized: [+/- material], [+/- dynamic], [+/-IEPS], [+/-Loc], [+/- B], [+/-

| | |
|--------------------------|-------------------------|
| interested in sports...} | |
| b. <i>day bed</i> | |
| [+material ([i])] | [+material ([i])] |
| <i>day</i> | <i>bed</i> |
| <+temporal> | <+artifact> |
| {24 hours, | <3-dimensional> |
| opposed to night} | <function> |
| | {for sleeping, flat...} |

(23a) is a coordinate compound, which can be seen in its structure: not only the features of the skeleton are coindexed because they are identical (i.e. [i]), but there is also a perfect match between the features of the bodies in angle brackets. It is only when we look at the encyclopedic features of the body that they differ. Given this parallelism, we can conclude that both nouns are given the same status and have no modification pattern. Conversely, in (23b) there is nothing that matches (apart from the skeletons bearing the [+material] feature indicating we are dealing with a noun). Even though the encyclopedic information of the body is not identical, a logical link can be established between them: beds are used for sleeping usually at night; but sleeping can also take place during the day. Thus, the meaning link is utterly determined pragmatically in this case. The point here is that there is nothing relevant to morphology or syntax that can guide interpretation. In this sense, Lieber's proposal resembles Jackendoff's approach, given that it highlights the inability of syntactic and productivity rules to establish a meaning link between the components of this type of compounds. In addition, Lieber's (2009) theory resembles prototype theory (Rosch et al 1976), considering that the implied semantic relation is based on the matching of prototypical and peripheral encyclopedic features of both nouns. In other words, for a meaning A to be a hyponym of B, it is not necessary for A to have all the attributes of a typical B: for instance, in *snail mail*, *snail* has the property of {gastropod} but also of {being slow} and *mail* the property of {it takes time}; thus, we could take the latter property of *snail* and *mail* and characterize their meaning.

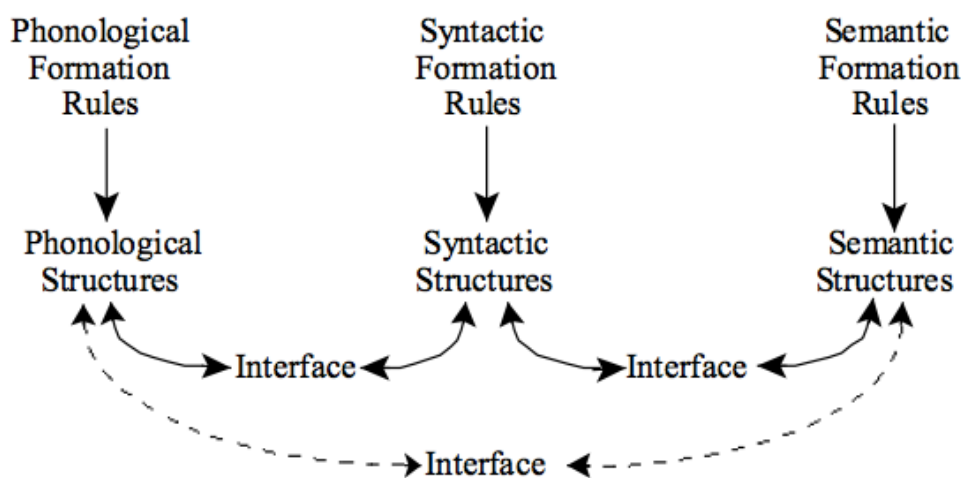
This framework of analysis of compounds seems to provide a more detailed account for the semantics of coordinate and verbal-nexus compounds than for the semantics of ground subordinate compounds because in the former compounds, there is clear relation between the features of the skeleton, on the one hand, and the features of the body, on the other: it is by matching the features of skeletons and bodies of both nouns that the interpretation is made available. However, the only way to establish a link between nouns in ground subordinate compounds is by logical matching of

properties in the unsystematic part of the semantic body (i.e., the extralinguistic knowledge of speakers). I will turn now to discuss Jackendoff's (2009, 2016) model which provides a detailed account on the semantic interpretation of ground subordinate NN compounds.

2.3.2.2 Jackendoff's (2009, 2016) basic functions

Jackendoff's model of conceptual semantics is 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. The result of this hypothesis is a Parallel Architecture of the grammar which is illustrated in figure 4:

Figure 4 Parallel Architecture Model taken from Jackendoff (2009: 106)



This model contrasts with mainstream generative approaches where the phonological and semantic properties are subordinated to syntax, and also with traditional cognitive grammar according to which syntactic structure is triggered by semantics. The model seems to stand in the middle of these two traditional approaches. This also means that there is no strict division between the lexicon and the grammar. Rather, as Jackendoff proposes there is “a continuum between idiosyncrasy and regularity” (2009: 108): words are found on one side and universal rules on the other, respectively. He adopts such a position because a precise division between the lexicon (i.e. words) and the grammar (i.e. set of rules generating the language) implies compounds are problematic:

what is the position compounds occupy in this spectrum? Individuals must store a great amount of lexicalized compounds with very specific meanings (including extralinguistic information), but not all compounds can be stored in the lexicon because of their extreme productivity. Besides, speakers may differ with respect to which compounds they have in their lexicon. Thus, language users must come up with a set of principles to interpret novel compounds, taking into account the mechanisms on which they have relied to learn the meaning of lexicalized ones. However, the meaning of the compounds is more than just the sum of the meaning of the two nouns making up the compound, which entails that pragmatics plays an important role. Thus, syntactic rules cannot be the main factor that is responsible for productivity in compounding. Taking into account the relative insignificance of syntactic rules in compounding, Jackendoff recognizes compounding as a relic of protolanguage which is characterized by having vocabulary and pragmatics, but lacking syntax or morphology (2009: 111-113).

Since his analysis is based on conceptual semantics, he is interested in the totality of meaning: this includes lexical semantics and its interaction with word formation, as well as pragmatics. 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 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 (24):

- (24) a. *ticket booth*: booth where tickets are bought.
 b. *ticket booth*: booth where tickets are sold.

Here 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, Jackendoff considers the function as a representation of an event schema that lacks perspective or focus.

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 (25):

- (25) 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 (25a) 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 (25b) 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. According to Jackendoff, 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 (26) for the noun 'writer', which profiles the agent of the action of writing.

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

(26a) 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 (26b) where the x stands for the individual to be denoted. The meaning of the expression is seen in (26c) 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 (27) taken from Jackendoff (2010: 17):

- (27) 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 (27a) *violinist* indicates an occupation because that is how she makes a living even though she is now on strike. (27b) illustrates the habitual action of playing the violin. In (27c) she has lost the habit of playing since the sale of her violin, but still indicates that she has the ability to play. Jackendoff (2009) refers to these variant interpretations as “the *action modalities* under which a nominal can be understood (119)”.

Furthermore, Jackendoff adopts Ruth Millikan’s (1984) proper function (PF) as an important action modality. PF means that something has been invented to, or is intended to perform a particular function. Only artefacts (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 (28):

- (28) a. *key* = [KEY^α; [PF (OPEN (α, INDEF))]] ‘a key has a proper function which is to open something’
 b. *mail₁ man₂* = [MAN₁^α; [OCC (DELIVER (α, MAIL₂))]] ‘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 (28a) 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 (28b), 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 (29), where the meaning of *Peter finished the movie* is presented:

- (29) a. $Peter_1 finished_2 the\ movie_3 = FINISH_2 (PETER^{\alpha}_1, F(\alpha, MOVIE_3))$
 b. $Peter_1 finished_2 the\ movie_3 = FINISH_2 (PETER^{\alpha}_1, WATCH(\alpha, [MOVIE^{\beta}; PF(WATCH(PERSON, \beta))])_3)$

The movie implies an activity which is marked by the F in (29a). This F coerces the semantic interpretation: the first argument of F is bound to *Peter* to indicate that it is him performing the action specified by the F and the second is bound to the movie which is what undergoes the action. In (29b) 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* (these are what Scalise and Bisetto (2009) call verbal-nexus compounds). 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 subordinate ground compounds, the function plays a fundamental role. For this reason, subordinate ground compounds will be the focus of my investigation. The formulation of the modifier schema has been taken from Jackendoff (2009: 122) and it is shown in (30) below:

- (30) Modifier schema:
 a. $[N_1 N_2] = [Y_2^{\alpha}; [F(\dots, X_1, \dots, \alpha, \dots)]]$
 b. an N_2 such that F is true of N_1 and N_2

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 4:

Table 4 Jackendoff’s thirteen basic functions taken from (Jackendoff 2016: 27-30)

| Function | Schema | Example |
|----------------|---|------------------|
| CLASSIFY (X,Y) | $[Y_2^{\alpha}; [CLASSIFY(X_1(\alpha))]]$ | <i>Beta cell</i> |

| | | |
|----------------------|---|--------------------------|
| BE (Y,X) | [Y ₂ ^a ; [BE (α, X ₁)]] | <i>Singer-songwriter</i> |
| SIMILAR (X,Y) | [Y ₂ ^a ; [SAME/SIMILAR (α, X ₁)]] | <i>Crocodile pin</i> |
| KIND (X,Y) | [Y ₂ ^a ; [KIND (α, X ₁)]] | <i>Seal pup</i> |
| (+) | [Y ₂ ^a ; [KIND (X ₁ ,α)]] | <i>Pine tree</i> |
| BE (X, AT/IN/ON/ Y) | [Y ₂ ^a ; [BE (α, AT/IN/ON X ₁)]] | <i>Lake house</i> |
| (+) | [Y ₂ ^a ; [BE (X ₁ , AT/IN/ON α,)]] | <i>Inkpad</i> |
| | [Y ₂ ^a ; [BE _{TEM} (α, AT X ₁)]] | <i>November rain</i> |
| COMP (X,Y) (+) | [Y ₂ ^a ; [COMP (α, X ₁)]] | <i>Meat ball</i> |
| | [Y ₂ ^a ; [COMP (X ₁ , α,)]] | <i>Sheet metal</i> |
| MADE (X, FROM Y) | [Y ₂ ^a ; [MADE (α, FROM X ₁)]] | <i>Coconut oil</i> |
| (+) | [Y ₂ ^a ; [MAKE (X ₁ , FROM α)]] | <i>Rubber tree</i> |
| PART OF (X,Y) | [Y ₂ ^a ; [PART (α, X ₁)]] | <i>Suit pants</i> |
| (+) | [Y ₂ ^a ; [PART (X ₁ ,α)]] | <i>Wheelchair</i> |
| | [Y ₂ ^a ; [PART (X ₁ ,α)]] (only partially) | <i>Cinnamon roll</i> |
| CAUSE (X,Y) | [Y ₂ ^a ; [CAUSE (X ₁ ,α)]] | <i>Diaper rash</i> |
| MAKE (X,Y) | [Y ₂ ^a ; [MAKE (X ₁ ,α)]] | <i>Spider poison</i> |
| (+) | [Y ₂ ^a ; [MAKE (α, X ₁)]] | <i>Silkworm</i> |
| X serves as Y | [Y ₂ ^a ; [BE (PF (α), PF (X ₁))]] | <i>Guard dog</i> |
| HAVE (X,Y) | [Y ₂ ^a ; [HAVE (X ₁ ,α)]] | <i>Gangster money</i> |
| (+) | [Y ₂ ^a ; [HAVE (α, X ₁)]] | <i>Glamour girl</i> |
| PROTECT (X,Y FROM Z) | [Y ₂ ^a ; [PROTECT (α, X ₁ , FROM Z)]] | <i>Chastity belt</i> |
| | [Y ₂ ^a ; [PROTECT (α, Z, FROM X ₁)]] | <i>Flea collar</i> |

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 N₂ that is made by N₁’ (*spider poison*) and ‘an N₂ that makes N₁’ (*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 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 (31):

- (31) a. *eye*₁ *doctor*₂ = [DOCTOR₂^α; [OCC (EXAMINE (α, EYES₁))]] ‘doctor has an occupation which is to examine eyes’
 b. *bird*₁ *egg*₂ = [EGG₂^α; [PF (BECOME (α, BIRD₂))]] ‘egg whose proper function is to become a bird’

In (31a), the *doctor* has a PF of occupation (i.e.) because it is his job to examine the eyes of patients. In the case of (31b), 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” (2016: 31).

Having reviewed Lieber’s and Jackendoff’s models, we can see that one of the main differences is, as ten Hacken (2016: 229) also points out, “the degree of analysis”. Jackendoff’s conceptual structure is more complex and elaborate than Lieber’s skeleton. Regarding what I call ground compounds, the skeletons of both components of the compound are only made of the [+ material] feature. It is only in the semantic body that more features can be applied. Furthermore, the meaning of this kind of compounds is established by identifying the merging of encyclopedic features as it has been illustrated in (23b). However, as we have seen above, Jackendoff’s basic functions model provides a complex and detailed conceptual structure for ground compounds, which makes it more accurate.

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 my hypothesis section.

But before getting into the hypothesis section, I consider it appropriate to mention a study of the semantic and formal structure of Swedish compounds and their French counterparts by Rosenberg (2013), where Jackendoff's model is put into practice.

2.4. Semantic and formal structure: a corpus based study of Swedish compounds and their French counterparts, by Rosenberg (2013)

Swedish and French exhibit the same type of differences with respect to NN compounds as English and Spanish do. NN compounds in Swedish are very productive, just like in English, and it is relatively easy to form novel compounds. On the other hand, just like Spanish, French also prefers alternative constructions to NN compounds. Given these differences, the main goal of Rosenberg's (2013) study is to identify the formal structure of French counterparts of Swedish NN compounds as well as the semantic relations expressed by the Swedish NN compounds and the corresponding French constructions. Furthermore, by adopting Jackendoff's (2009) model, Rosenberg attempts to shed light on his questions: "What is the full set of basic functions, and how uniform are they cross-linguistically? And to what extent are they special to language - or English?" (2009: 128)."

In order to investigate these issues, Rosenberg (2013) gathered the data from a parallel corpus that contains only written texts. The texts that conform the corpus are drawn from debates taking place in the European Parliament. Rosenberg gathered 983 NN compounds that corresponded to 940 French counterparts. Although the results point to some tendencies that require further investigation, they indicate that French counterparts of Swedish NNs are mainly *N de N* and NA (i.e. Noun Adjective), whereas few NN compounds have been collected. It is *N de N* and NA that can render almost any semantic relation which is expressed by Swedish NN compounds. As for French NN compounds, few of them are attested in the data, and only two basic functions are available: CLASSIFY and BOTH (BE (Y,X) in Jackendoff's terminology). Moreover, according to Rosenberg's (2013) data, there are four main dominant relations and two that are unattested in Swedish or French: PURPOSE, COMP, LOC and CLASSIFY

(the function of PURPOSE has been taken from Gagné and Shoben (1997) and is aimed to replace Jackendoff's PF) are dominant, while 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. However, the author acknowledges that the results only illustrate a strong tendency and may be influenced by the type of corpus chosen for the study: i.e. a parallel corpus made of written texts collected from debates at the European Parliament. Apart from this, another interesting outcome of the study is the frequent equivalence of semantic relations between Swedish NN compounds and French counterparts. That is, a Swedish NN compound *X* usually has the same semantic relation as its French counterpart *Y*.

In sum, one of the conclusions of the study is that there are several structures that are used in French to express concepts that are realized with NN compounds in Swedish. Furthermore, a hierarchy with respect to the most productive French formal structures expressing the same semantic relations found in Swedish NN compounds can be established: NA > N *de* N > Simple Nouns > NNs. The first two structures express almost every semantic relation, which is in the line of Nicoladis' (2002) claim about the preposition *de* in French: *de* in French may be becoming a linking element. As a result, it can be assumed that "N *de* N constructions have the same status as compounds in French" (Rosenberg 2013: 118). Besides, given that French NN compounds can express only few semantic relations, this semantic inflexibility may be a factor that causes their unproductivity.

Rosenberg's (2013) study is relevant to the purpose of this thesis as it compares a language with productive NN compounding (i.e. Swedish) to a language with unproductive NN compounding (i.e. French). Besides, its goal is to account for the number of semantic relations that hold in NN compounds in the two languages using Jackendoff's (2009) basic functions model. Thus, it attempts to test the universality of the model and shed some light on the differences between the languages. In addition to that, it emphasizes the possibility that NN compounds could be unproductive in some languages due to the semantic inflexibility of the constructions. These are some of the same issues I tackle in my study with respect to Spanish, and they are carefully explained in the next section of the thesis.

3. Research questions and hypotheses.

As mentioned above, the main goal of this thesis 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. Furthermore, as it has been shown above, these questions have already been investigated by Rosenberg (2013) in her comparison of Swedish NN compounds and their French counterparts. Given her results, it will be interesting to find out whether there are any differences between French and Spanish in terms of the basic functions and semantic relations expressed in their NN compounds, considering that NN compounding is not productive in either language. My hypothesis is that, in spite of being typologically similar, the types of semantic relations attested in Spanish NN compounds will be different from those available in French NN compounds. One of the reasons why this may be true is that the preliminary analysis of my data reveals the availability of the metaphorical relation between modifier and head as in *lápiz cohete* ('pencil rocket' = rocket pencil). Such compounds provide instances of the SIMILAR function, which is not attested in the French data.

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, and that there will not be a significant difference between those relations that are used the most and those that are less productive. That is, even though a relation like HAVE may be more productive than MADE FROM, the difference with respect to their frequency may not be significant. Also, given what Rosenberg (2013) found in her study regarding PROTECT and KIND (they were not attested either in French structures or Swedish NNs), I expect that these two semantic relations will be extremely peripheral in my data.

In addition to this, I consider it important to try to apply Jackendoff's system to Spanish *N de N* constructions as well, because they are very frequent 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 with 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 productive in the language, I hypothesize that they will exhibit a wider range of semantic relations. 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 French counterparts of these constructions mentioned above (i.e. that prepositions in these structures are becoming linking elements), as it may show a pattern taking place in Romance languages. Such results would also confirm that *N de N* should be considered as ‘actual’ compounds with the same status as NN compounds given their conceptual unity (Kornfield 2009: 442).

Finally, 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 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 I have been obtained from the *Corpus del Español* (Davis 2002-), the appendix in *Compound Words in Spanish: Theory and History* (Moyna 2011: 303-433), and various written sources such as articles and textbooks. The analysis is presented in the next chapter.

Chapter 3

Method and Analysis

This chapter 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 it turns to the analysis of the data in an attempt to answer the research questions formulated in Ch. 2 section 3.

3.1. 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 (cited in the works cited section), the appendix in *Compound Words in Spanish: Theory and History* (Moyna 2011: 303-433) and the *Corpus del Español* (Davis 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 (Davis 2002-). The former contained an exhaustive list of NN compounds through history, but I focused primarily on the most recent centuries, starting in the 18th and moving into the 20th and extracting only those which are still in use nowadays. The latter is a web-based Spanish corpus of two billion words. In order to obtain the target constructions, I searched the corpus 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: e.g. *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, I created a database in Excel with two spreadsheets, one to record NN compounds and the other to record *N de N* constructions. Each spreadsheet 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 N* constructions collected. In addition to classifying the NN and *N de N* 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' (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. 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 the next section.

3.2. Data analysis and discussion

The previous chapter posed some of Jackendoff's (2009) questions that I would attempt to answer. More specifically, these questions were concerned with what the full set of basic functions is and how uniform it is cross-linguistically, and to what extent these functions are specific to English or other languages. In order to find an answer to these questions I will analyze the Spanish NN compounds I collected and compare the

results to those that Jackendoff obtained for English. As a result of this analysis, we will be able to see whether the functions that Jackendoff (2009, 2016) proposes are applicable to Spanish NN compounds, and, more precisely, whether the full set of them applies, or not. If the latter turns out to be the case, the goal would be to come up with a possible explanation of this cross-linguistic variation.

In Chapter 2, I also outlined a related goal, which is to organize the basic functions hierarchically according to the frequency of their occurrence in the data. This order will allow us to see which relations are most and least frequent, which can shed some light on why NN compounds are unproductive in Spanish.

Once the analysis of Spanish NN compounds is completed, I will make a more detailed comparison of these results with those found by Rosenberg (2013) in her analysis of French and Swedish NN compounds. I hypothesized that despite the typological similarities between French and Spanish NNs, the types of basic functions attested in Spanish would be different from those available in Rosenberg's data.

Apart from NN compounds, I will also analyze the *N de N* data. The purpose is to find out whether the model proposed by Jackendoff can also account for the meaning of *N de N*s. If it turns out that *N de N*s can be analyzed using the same model as that for NN compounds, this fact will strongly support the Nicoladis' (2002) claim about prepositions in these constructions losing their true prepositionhood and becoming linking elements. Another goal of the analysis is to try to confirm the hypothesis that a wide amount of basic functions and semantic relations will be available in *N de N*, considering that they are very productive.

Last but not least, I will compare the results for both NNs and *N de N*s in Spanish, in an attempt to provide an answer to whether semantic inflexibility is a factor responsible for unproductivity of NN compounds in Spanish.

3.2.1. 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 5 below that Spanish does not realize all of the basic functions that Jackendoff found in English:

Table 5 Functions within Spanish NN compounds in a hierarchical order

| Semantic relation | Spanish NN | English translation | Type |
|-------------------|------------|---------------------|------|
|-------------------|------------|---------------------|------|

| | | | |
|--|--|---|-----|
| | | | |
| SIMILAR 'N ₁ similar to N ₂ ' | <i>Cabeza buque</i> <i>Coche oruga</i> <i>Tren bala</i> | 'big head' 'caterpillar car' 'bullet train' | 76 |
| PF 'N ₁ whose PF is to function as N ₂ ' | <i>Hombre orquesta</i> <i>Lápiz láser</i> <i>Mecánico dentista</i> | 'one-man band' 'laser pencil' 'dental technician' | 38 |
| X serves as Y 'N ₁ whose function is to function in the proper function of N ₂ ' | <i>Canción protesta</i> <i>Perro guardián</i> <i>Sombrero</i> <i>paraguas</i> | 'protest song' 'guard dog' 'umbrella hat' | 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' | <i>Balón reloj</i> <i>Gas ciudad</i> <i>Zona euro</i> | 'clock ball' 'town gas' 'euro zone' | 28 |
| HAVE 'N ₁ that has (an) N ₂ '; 'N ₁ that N ₂ has' | <i>Bandera pirata</i> <i>Carril bus</i> <i>Coche patrulla</i> | 'pirate flag' 'bus lane' 'police car' | 25 |
| COMP 'N ₁ composed of N ₂ ' | <i>Avión cohete</i> <i>Barco tanque</i> <i>Flota pirata</i> | 'rocket airplane' 'tank ship' 'pirate fleet' | 15 |
| PART 'N ₁ is a part of N ₂ ' | <i>Aguanieve</i> <i>Mochila arnés</i> <i>Papel carbón</i> | 'sleet' 'harness backpack' 'carbon paper' | 15 |
| CAUSE 'N ₁ that is caused by N ₂ ' | <i>Efecto placebo</i> <i>Guerra narco</i> <i>Virus zombie</i> | 'placebo effect' 'narc war' 'zombie virus' | 14 |
| BE 'N ₁ is an N ₂ ' | <i>Abeja reina</i> <i>Perro vampiro</i> <i>Príncipe rana</i> | 'queen bee' 'vampire dog' 'frog prince' | 7 |
| ARGUMENT SCHEMA 'an N ₁ of/by N ₂ ' | <i>Abordaje pirata</i> <i>Ataque pirata</i> <i>Bote salvavidas</i> | 'pirate boarding' 'pirate attack' 'life boat' | 5 |
| MAKE 'N ₁ made by N ₂ '; 'N ₁ makes N ₂ ' | <i>Fiesta pirata</i> <i>Horno</i> <i>microondas</i> <i>telaraña</i> | 'pirate party' 'microwave oven' 'cobweb' | 5 |
| Types of Spanish NNs | | | 261 |

Table 5 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 5, there are four functions that are not present at all in the Spanish NN compounds I 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), 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 mine and Rosenberg's study (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.

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 discussing their productivity.

3.2.2 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, I have ordered them in table 5 according to the

frequency of their occurrences in the data. I predicted 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, we 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 in table 6:

Table 6 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 |

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. These blocks are identified 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 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 (32):

- (32) SIMILAR
- a. *gas*₁ *mostaza*₂ (*mustard gas*): [GAS₁^α; [SIMILAR (α, MOSTAZA₂)]], ‘*gas* similar to *mostaza*’
 - b. *novela*₁ *río*₂ (*river novel* or *saga*): [NOVELA₁^α; [SIMILAR (α, RÍO₂)]], ‘*novela* similar to *río*’

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 (33):

- (33) a. PF
1. *hombre*₁ *anuncio*₂ (*sandwich-board man*): [HOMBRE₁^α; [OCC (WORK (α, ANUNCIO₂))]], ‘*hombre* whose occupation is to work as a *anuncio*’
 2. *pistola*₁ *láser*₂ (*laser gun*): [PISTOLA₁^α; [PF (SHOOT (INDEF, LÁSER₂, FROM α))]], ‘*pistola* whose PF is for people to shoot a *láser* from it’
- b. X serves as Y
1. *canción*₁ *protesta*₂ (*protest song*): [CANCIÓN₁^α; [BE (PF (α), PF (PROTESTA₂))]], ‘*canción* whose PF is to function as the PF of *protesta*’
 2. *perro*₁ *guardián*₂ (*guard dog*): [PERRO₁^α; [BE (PF (α), PF (GUARDIÁN₂))]], ‘*perro* whose PF is to function as the PF of *guardián*’

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 (34) 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 (34):

(34)

BE (AT/IN/ON)

- a. *reloj*₁ *pulsera*₂ (*bracelet watch*): [RELOJ₁^α; [BE (α, ON PULSERA₂)]], ‘*reloj* that is located on a *pulsera*’
 b. *zona*₁ *euro*₂ (*euro zone*): [ZONA₁^α; ([BE (EURO₂ IN α))], ‘*zona* with *euro* in it’
 c. *fútbol*₁ *sala*₂ (*indoor soccer*): [FÚTBOL₁^α; PF ([BE (α IN SALA₂)])], ‘*fútbol* whose PF is to be (played) in the *sala*’
 d. *zapato*₁ *teléfono*₂ (*phone shoe*): [ZAPATO₁^α; PF ([BE (TELÉFONO₂ IN α))], ‘*zapato* whose PF is to have a *teléfono* in it’

From the examples in (34) we can infer that the locative function also involves some sort of possession as seen in (34b) and (34d). Nevertheless, this possession is never on its own as it always appears with some locality flavor. The function denoting possession and ownership by definition is HAVE. Though Jackendoff provides only two schemata for this function hoping that scholars 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 he suggests. A sample of each is exemplified in (35):

(35)

d. HAVE

1. *gorra*₁ *visera*₂ (*peaked cap*): [GORRA₁^α; [HAVE (α, VISERA₂)]], ‘*gorra* that has a *visera*’
 2. *perro*₁ *pastor*₂ (*shepherd’s dog*): [PERRO₁^α; [HAVE (PASTOR₂, 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: 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 (36) below:

- (36) a. *autobús*₁ *calabaza*₂: [AUTOBÚS₁^α; [SAME/SIMILAR (α, CALABAZA₂)]], ‘*autobús* similar to a *calabaza*’
 b. *tren*₁ *bala*₂: [TREN₁^α; [SIMILAR (α, BALAZA₂)]], ‘*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 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 the SIMILAR relation as the first choice. Some compounds exhibiting this property are illustrated in (37):

- (37) a. *hombre*₁ *araña*₂ (*spider man*)
 1.[HOMBRE₁^α; [SIMILAR (α, ARAÑA₂)]]

‘*hombre* similar to *araña*’

2. [HOMBRE₁^α; [PF (BECOME (α, ARAÑA₂))]], ‘*hombre* whose PF is to become *araña*’

3. [HOMBRE₁^α; [PART (ARAÑA₂, α)]], ‘*hombre* who is partly *araña*’

b. *teléfono*₁ *zapato*₂ (*shoe phone*)

1. [TELÉFONO_α; [SIMILAR (α, ZAPATO)]], ‘*teléfono* similar to *zapato*’

2. [TELÉFONO_α; [BE (PF (α), PF (ZAPATO))]], ‘*teléfono* whose PF is to function as the PF of *zapato*’

3. [TELÉFONO_α; PF ([BE (α, IN ZAPATO)])], ‘*teléfono* whose PF is to be in a *zapato*’

In NN compounds such as (37a) and (37b) 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 since there are a few cases where SIMILAR is not the most salient function, as illustrated in (38):

(38) a. *perro*₁ *vampiro*₂ (*vampire dog*)

1. [PERRO₁^α; [BE (α, VAMPIRO₂)]], ‘*perro* that is a *vampiro*’

2. [PERRO₁^α; [SIMILAR (α, VAMPIRO₂)]], ‘*perro* similar to a *vampiro*’

3. [PERRO₁^α; [PF (BECOME (α, VAMPIRO₂))]], ‘*perro* whose PF is to become a *vampiro*’

4. [PERRO₁^α; [PART (VAMPIRO₂, α)]], ‘*perro* that is partly a *vampiro*’

b. *coche*₁ *hippie*₂ (*hippie car*)

1. [COCHE₁^α; [HAVE (HIPPIE₂, α)]], ‘*coche* that *hippies* have’

2. [COCHE₁^α; [SIMILAR (α); [HAVE (HIPPIE₂, α)]]], ‘*coche* similar to the *coche* *hippies* have’

Contrary to what has been presented in (37), (38) shows that SIMILAR need not always be the main function. It can also be a secondary one.

Up until this point, I have examined the most productive functions found in the data. Tables 5 and 6 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 (39):

(39) a. COMP

1. *flota*₁ *pirata*₂ (*pirate fleet*): [FLOTA₁^α; [COMP (α, PIRATA₂)]], ‘*flota* composed of *pirata*(s)’

2. *buque*₁ *cisterna*₂ (*tank ship*): [BUQUE₁^α; [COMP (α, CISTERNA₂)]], ‘*buque* composed of a *cisterna*’

b. PART

1. *perro*₁ *lobo*₂ (*wolf dog*): [PERRO₁^α; [PART (α, LOBO₂)]], ‘*perro* that has part of *lobo*’

2. *agua*₁ *nieve*₂ (*sleet*): [AGUA₁^α; [PART (NIEVE₂, α)]], ‘*agua* that is composed in part of *nieve*’
3. *traje*₁ *pantalón*₂ (*pantsuit*): [TRAJE₁^α; [PART (PANTALÓN₂, α)]], ‘*traje* that has a *pantalón* as a part’

COMP as suggested by Jackendoff (2009, 2016) is reversible, but I have only been able to find examples of the ‘X composed of Y’ schema which is the one illustrated in (39a), while the ‘X that Y is composed of’ schema is not attested in my 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 (39b). I have provided an example of the variant paraphrase for mass nouns in (39b.2). If dealing with a countable counterpart as in (39b.3), the formal structure stays the same but the paraphrase is expressed as it 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 (40):

- (40) CAUSE
1. *efecto*₁ *placebo*₂ (*placebo effect*): [EFECTO₁^α; [CAUSE (PLACEBO₂, α)]], ‘*efecto* caused by a *placebo*’
 2. *mochila*₁ *cohete*₂ (*rocket backpack*): [MOCHILA₁^α; [PF (MOVE^β (α); [CAUSE (COHETE₂, β)])]], ‘*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 N₁ and N₂ at the same time. Nevertheless, I consider that the NN compounds collected are not coordinate but subordinate ground (as defined in Ch. 2, section 2.1) because they are mainly a hyponym of the head and only one of nouns is inflected for plural: e.g. the left hand noun is inflected as in *perro vampiro* (‘dog vampire’ > *vampire dog*) > *perros vampiro* as opposed to *niño atleta* (‘kid athlete’ > *athlete kid*) > *niños atletas* where both nouns are inflected for plural. A sample is given in (41):

- (41) BE
1. *perro*₁ *vampiro*₂ (*vampire dog*): [PERRO₁^α; [BE (α, VAMPIRO₂)]], ‘*perro* that is a *vampiro*’
 2. *príncipe*₁ *rana*₂ (*frog prince*): [PRÍNCIPE₁^α; [BE (α, RANA₂)]], ‘*príncipe* that is a *rana*’

The most peripheral function, though, is MAKE which implies a semantic relation of source or creation. The reversibility of function proposed by Jackendoff is

also available in Spanish. Thus, there are two possible readings for this function which are shown in (42):

- (42) MAKE
- a. *tela*₁*araña*₂ (*cobweb*): [TELA₁^α; [MAKE (ARAÑA₂, α)]], ‘*tela* made by the *araña*’.
 - b. *horno*₁ *microondas*₂ (*microwave oven*): [HORNO₁^α; [MAKE (α, MICROONDAS₂)]], ‘*horno* that makes *microondas*’

Scholars agree in the difficulty of differentiating MAKE from CAUSE. For example, Jackendoff (2016: 30) as well as 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 (42b) 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 my data, but it is more productive than it appears in my data. That is because I 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 (43):

- (43) ARGUMENT SCHEMA
1. *ataque*₁ *pirata*₂ (*pirate attack*): [ATAQUE₁ (PIRATA₂, INDEF)], ‘*ataque* on something by *piratas*’
 2. *bote*₁ *salva*₃*vidas*₂ (*life boat*): [BOTE₁^α; [SALVAR₃ (INDEF, VIDAS₂, WITH α)]], ‘*bote* that someone *salva vidas* with’

These 6 functions, though documented, constitute a more peripheral block in Spanish NNs. It is important to mention that there is a noticeable difference in productivity between peripheral functions BE or MAKE among others and the more frequent functions such as SIMILAR or PF; and on top of that, there are significant differences within the peripheral block of functions itself. The fact that there is a noticeable division between (more) frequent and (more) peripheral functions in Spanish rejects my hypothesis that there would not be a significant difference between the functions attested. Given that Spanish NN compounds are not productive, I expected that the few functions attested would not be significantly different in terms of frequency. In other words, they would all be equally peripheral. However, the fact that

some functions are significantly more productive than others (i.e. SIMILAR vs. any other function) refutes this hypothesis.

Apart from this, I also thought 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 I collected. These are not the only unattested functions, though; CLASSIFY and MADE FROM are not present either. Thus, Spanish does not allow these functions to satisfy *F* in NN compounds. I 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 (44):

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

Before I move on to comparing these results with Rosenberg's (2013), I would like to comment on two issues that I have briefly referred to while examining some of the functions: i.e. promiscuity and cocomposition. Regarding the former, 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 my 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 (recall Jackendoff's (2016: 20) example *boxcar*) Despite this difference, promiscuity does exist for certain NN compounds, as shown in table 7:

Table 7 Number of NN compounds with 1, 2, 3 or more possible relations

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

As it can be drawn from table 7, 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 possibility that might explain this is related to lexicalization and productivity issues. That is, there are few 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 my data that can support this claim.

As for cocomposition which I spoke briefly of when talking about CAUSE (in section 3.2.1), Spanish allows NN compounds with complex meanings, which a single basic function fails to interpret. Therefore, the solution is to use a cocomposed function to render the accurate meaning of the compound. The most frequent combination is CAUSE and PF to denote vehicles or machines used for transport as (40b) and (45):

- (45) *coche₁ diesel₂* (*diesel car*): [COCHE₁^α; [PF (MOVE^β (α)); [CAUSE (DIESEL₂, β)]]], ‘*coche* that moves by *diesel* causing its movement’.
 also *coche gasoline* (*gas car*), *proyector cohete* (*rocket missile*), *puño cohete* (*rocket fist*), *trineo cohete* (*rocket sled*)

The PF of *coche* (‘car’) is to move, but it moves because *diesel* is the material that makes a car run. As a result, both meanings are cocomposed to create the structure in (45).

Animals are also particular cases where combinatorial options need to come into use to generate a more complex relation. For instance, *pez espada* (‘fish sword’ > *sword fish*) and *tiburón martillo* (‘shark hammer’ > *hammerhead*) are a fish that has a part that is similar to a sword and a shark that has a part that is similar to a hammer, respectively. This already implies that there are two functions involved: PART and SIMILAR. Their schema is represented in (46):

- (46) *pez₁ espada₂* (*sword fish*): [PEZ₁^α; [PART ([Z^β; SIMILAR (β, ESPADA₂)], α)]]

Another similar case is found in *hormiga bala* (‘ant bullet’ > *bullet ant*). It is not an ant that makes bullets, but an ant that makes wounds similar to those caused by

bullets. In this case *F* is satisfied by MAKE which takes the implied noun *wounds* as argument. Besides, it is the wounds that are SIMILAR to bullet wounds which at the same time should be decomposed into ‘CAUSE wounds’. This is illustrated in (47):

- (47) *hormiga*₁ *bala*₂ (*bullet ant*): [HORMIGA₁^α; [MAKE ([WOUNDS^β; SIMILAR (β, BALA₂; CAUSE (β))]) α]]

The possibility to coerce extra functions in order to obtain the underlying semantic relations and meaning of the compound provides further evidence for the fact that the generative system proposed by Jackendoff can be properly applied to account for the semantics of NN compounds in Spanish. It would be a challenge for the system if NN compounds such as the ones from (45) to (47) existed but the system was unable to derive their meaning. Since this is not the case, and taking into account that most of the basic functions are attested, these constitute strong support for the claim that Jackendoff’s conceptual semantics model is suitable for Spanish as well as for English, as well as other languages, as demonstrated by Rosenberg (2013).

In the next section, I will establish a comparison between the results of my study and those obtained by Rosenberg (2013) with respect to French. I will try to confirm the hypothesis that, despite being typologically similar, Spanish NN compounds will exhibit distinct basic functions and, thus, distinct semantic relations.

3.2.3. A comparison with Rosenberg (2013)

Even though a brief comparison with Rosenberg has already been made in section 3.2.1., I would like to elaborate a bit more on the differences and similarities between Spanish NN compounds and Rosenberg’s (2013) results. The aim is to determine whether there is a difference between Spanish and French NNs with respect to their basic functions. As I predicted in Ch. 2, section 3 from a quick look at my data, the types of basic functions that are expressed in NN compounds in these two languages differ.

Rosenberg (2013) found comparatively few NN compounds in French. Moreover, these NN compounds expressed only two functions: BE and CLASSIFY. However, as we have seen above, in my Spanish data, these functions were either peripheral (i.e. BE) or not even attested at all (i.e. CLASSIFY). Conversely, as I have explained in the previous sections (e.g. 3.2.1. and 3.2.2.), most basic functions proposed by Jackendoff are available in Spanish. Therefore, the first conclusion is that the basic functions and, therefore, semantic relations implied in NN compounds are not the same

in the two languages, which confirms my prediction. Since fewer basic functions and semantic relations are present in French, this might indicate that one could expect to find fewer NN compounds in French than in Spanish despite the fact that both languages are [-TCP] languages in the sense of Snyder (2012).

Despite this difference, there is also a similarity which can as well be extended to Swedish, regarding the non-attested functions mentioned in section 3.2.1.: neither KIND nor PROTECT are found. Since these two functions are not attested in any language from Rosenberg's, Jackendoff's and my study other than English, a possibility is that they are particular to English. Another possibility at least for Spanish (this was not the case for French) is that these two functions are realized by a different formal structure: *N de Ns* (which will be dealt with in the next section).

To sum up, given the results of my and Rosenberg's study, we can make two conclusions: on the one hand, Spanish displays a larger amount of basic functions than French where they seem more restricted; on the other hand, functions such as KIND or PROTECT seem unproductive, which may entail that their appearance in the semantics of compounds is restricted to certain languages or formal structures.

Similar to what Rosenberg does with French N-Prep-N compounds, in the next section I will analyze *N de N* constructions in Spanish to see if Jackendoff's model is also suitable for accounting for their meaning. I will, therefore, 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 question of whether the preposition *de* in this structure is a mere linking element, rather than a true preposition, as proposed by Nicoladis (2002), as well as on some of the unsolved questions posed above: i.e. are KIND and PROTECT language or structure specific?

3.2.4. Semantic relations based on the basic functions found in N de N constructions

In section 3.2.2. 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*, I have created table 8. 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 5, 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 8 is given below:

Table 8 Functions within Spanish N de N constructions

| Semantic relation | Spanish <i>N de N</i> | English translation | Type |
|--------------------------------|---|--|------|
| PF | <i>Alcohol de heridas</i> <i>Casa de moneda</i> <i>Huevo de dinosaurio</i> | 'alcohol' 'coin house (mint)' 'dinosaur egg' | 44 |
| BE (AT/IN/ON) | <i>Estanque de patos</i> <i>Mesa de jardín</i> <i>Tormenta de verano</i> | 'duck pond' 'garden table' 'summer storm' | 41 |
| COMP | <i>Bandeja de plata</i> <i>Circo de pulgas</i> <i>Molécula de agua</i> | 'silver tray' 'flea circus' 'water molecule' | 35 |
| PART | <i>Corazón de manzana</i> <i>Manecilla de reloj</i> <i>Reja de arado</i> | 'apple core' 'clock hand' 'plowshare' | 28 |
| HAVE | <i>Derecho de autor</i> <i>Lámpara de flecos</i> <i>Sombrero de policía</i> | 'copyright' 'fringe lamp' 'police hat' | 22 |
| MADE FROM | <i>Aceite de girasol</i> <i>Caña de azúcar</i> <i>Estofado de carne</i> | 'sunflower oil' 'sugarcane' 'beef stew' | 19 |
| CAUSE | <i>Dolores de parto</i> <i>Herida de navaja</i> <i>Turismo de sol</i> | 'birth pains' 'knife wound' 'sun tourism' | 14 |
| PROTECT | <i>Botas de lluvia</i> <i>Cinturón de castidad</i> <i>Gafas de sol</i> | 'rain boots' 'chastity belt' 'sunglasses' | 13 |
| SIMILAR | <i>Arco de herradura</i> <i>Cara de vinagre</i> <i>Sargento de hierro</i> | 'horseshoe arch' 'sour face' 'iron sergeant' | 11 |
| MAKE | <i>Agujero de ratón</i> <i>Cera de abeja</i> <i>Veneno de serpiente</i> | 'mouse hole' 'bee wax' 'snake poison' | 10 |
| ARGUMENT SCHEMA | <i>Calidad de vida</i> <i>Guía de teléfono</i> <i>Número de teléfono</i> | 'life quality' 'telephone book' 'telephone number' | 4 |
| X serves as Y | <i>Señal de alarma</i> | 'alarm signal' | 1 |
| Types of Spanish <i>N de N</i> | | | 242 |

Table 8 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 9:

Table 9 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 |

Table 9 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 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: i.e. 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 (48):

- (48) a. *tormenta₁ de verano₂* (*summer storm*): [TORMENTA₁^α; [BE_{temp} (α, AT VERANO₂)]], ‘*tormenta* that takes place at time *verano*’
also *chaqueta de invierno* (*winter jacket*), *deporte de invierno* (*winter sport*),
lluvia de primavera (*spring rain*)

It is striking, however, that SIMILAR and X serves as Y are much less productive with N de Ns than with NNs, comparing tables 6 and 9. 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 (49):

- (49) a. *molécula₁ de agua₂* (*water molecule*): [MOLECULA₁^α; [COMP (AGUA₂, α)]], ‘*molécula* that *agua* is composed of’

Although according to tables 5 and 8 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*⁷ or to cases that involve cocomposition (e.g. (40b)), 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* (‘spot 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. Moreover, just as with NN compounds, this function is also reversible. An example of each schema is provided below in (50):

- (50) a. *veneno₁ de serpiente₂*: [VENENO₁^α; [MAKE (SERPIENTE₂, α)]], ‘*veneno* made by the *serpiente*’
also *rayo de sol* (*sunbeam*), *agujero de ratón* (*mouse hole*), *cueva de grillo*
(*cricket hole*)

⁷ *Guerra narco* (*narc war*), *efecto placebo* (*placebo effect*), or *virus vampiro* (*vampire virus*) among others.

b. *gusano*₁ *de seda*₂ (*silkworm*): [GUSANO₁^α; [MAKE (α, SEDA₂)]], ‘*gusano* that makes *seda*’

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, designating edible and drinkable products such as the ones in (51):

- (51) a. *aceite*₁ *de girasol*₂ (*sunflower oil*): [ACEITE₁^α; [MADE (α, FROM GIRASOL₂)]], ‘*aceite* made from *girasol*’
 also *cerveza de trigo* (*wheat beer*), *crema de cacahuete* (*peanut butter*), *estofado de carne* (*beef stew*), *harina de maíz* (*cornstarch*), *harina de trigo* (*wheat flour*) etc.
 b. *caña*₁ *de azúcar*₂ (*sugarcane*): [CAÑA₁^α; [MADE (AZÚCAR₂, FROM α)]], ‘*caña* that *azúcar* is made from’

The examples in (51) 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 (52):

- (52) a. *cinturón*₁ *de castidad*₂ (*chastity belt*): [CINTURÓN₁^α; [PROTECT (α, CASTIDAD₂ FROM Z)]], ‘*cinturón* protects *castidad* from something’
 also *funda de teléfono* (*cell phone case*), *casco de ciclista* (*cyclist helmet*)
 b. *gafas*₁ *de sol*₂ (*sunglasses*): [GAFAS₁^α; [PROTECT (α, EYES, FROM SOL₂)]], ‘*gafas* protect eyes from *sol*’
 also *collar de pulgas* (*flea collar*), *crema de sol* (*suncream*)

Examples such as (52a) have what is protected specified in the compound (i.e. chastity) and require the reader to interpret the threat (i.e. Z). Others like (52b) 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 notions of promiscuity and cocomposition are 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 10:

Table 10 Number of *N de N* constructions with 1, 2, 3 or more possible relations

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

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* (*chocolate cake*) even if a context is provided. What seems evident is that having three or more than three functions is not common. It is interesting that while for NNs I claimed 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 productive. A possibility is that there is no correlation between *N de Ns* being productive 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. For instance, *sofá de jardín* (sofa of garden = ‘garden sofa’) is a ‘*sofá* whose function is to be located in the *jardín*’. However, this is just a prediction that should be further explored in future research.

As for cocomposition, there are also cases of *N de Ns* with complex meanings that require more than one function to account for their meaning. Thus, cocomposition must come into play to coerce several functions together. A particular example of cocomposition in *N de N* compounds is given in (53):

- (53) *barco₁ de vapor₂ (steam boat)*: [BARCO₁^α; [PF (MOVE^β (α)); [CAUSE (VAPOR₂, β)]]], ‘*barco* that moves by *vapor* causing its movement’

(53) is a cocomposed example of a means of transport whose formal structure is *N de N*. However, it is very interesting that there are some recent NN compounds with cocomposition where *de* seems to be deleted. This is seen in the NN compound in (45) repeated in (54):

- (54) *coche₁ diesel₂ (diesel car)*: [COCHE₁^α; [PF (MOVE^β (α)); [CAUSE (DIESEL₂, β)]]], ‘*coche* that moves by *diesel* causing its movement’.

(54) is identical to (53) regarding the semantic structure except for the fact that it is an NN compound. It is the first time I 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 (55):

- (55) a. *billete de tren* (‘ticket of train’ > *train ticket*)
 b. *bonotren* (‘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 (55a); due to changes, the preposition lost its semantic import and was removed resulting in a fused NN compound as in (55b). On the other hand, this process has not affected (55a) yet, as it retains the preposition. A similar case that can support this claim is found in *telaraña* (‘web spider’ > *cobweb*). 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. 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 (56):

(56) 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 (54) predicting a potential loss of the preposition. In order to prove this claim and the scale in (56), 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*): i.e. 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 (57):

- (57) a. NN compounds: resemblance/metaphorical > purpose (+ function 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 ‘function as Y’ which are on top of the scale for NN compounds are at the bottom for *N de N*s; 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.

The hierarchies in (57) and what has been explained in this section confirm the hypothesis that more semantic relations are found in *N de N* compounds than in NNs. Besides, all the possible schemata proposed for each of the functions are realized in Spanish *N de N*s, as opposed to NNs, which highlights the semantic flexibility of the *N de N* structures. Furthermore, some schemata are available for *N de N*s but unavailable for NN compounds (primarily, temporal and spatial locations expressed with BE, and reversible COMP). All this provides extra evidence for the proposal that in Romance languages *N de N* and NN structures should be considered the same type of compound semantically: even though they differ morphosyntactically, semantically they are twins.

The results presented in this section together with those in the previous sections also shed light on the following research question: are NN compounds uncommon in Spanish because they are semantically inflexible due to the limited set of semantic relations that hold between the head and modifier? Though Spanish NN compounds express almost all relations established for English NN compounds, which was unexpected, the relations in Spanish are more restricted and unproductive when compared to English because there are fewer of them found and they may not allow all possible schemata proposed for the functions. This, thus, indicates that NNs are not very flexible semantically. It seems, though, that in the future NN compounds may become more frequent by virtue of *N de N*s losing their Ps (i.e. *de*). On the other hand, *N de N*s

are semantically identical to NN compounds but are more common because the relations expressed exhibit a higher productivity and allow all the combinatorial schemata for the functions, which demonstrates they are extremely flexible.

Now that the data has been analyzed and discussed, I will turn to the last chapter of this thesis: the conclusion. In this chapter I will summarize the main findings and highlight important contributions of the study, as well as outline potential research in some aspects that may require further evidence or clarification.

Chapter 4

Conclusions

This thesis has been concerned with analyzing the semantic relations that hold in NN compounds in Spanish based on Jackendoff's (2009, 2016) basic functions model within the PA framework. 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 attested, despite the fact that Spanish is a language where compounding is said to be unproductive (Snyder 2001, 2012). However, some of the functions that Jackendoff proposed for English seem not to be found in NN compounds in other languages: for instance, Rosenberg (2013) already showed this to be true for French NN compounds, where the only functions attested are CLASSIFY and BE (X,Y). In this thesis, we have seen 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. Furthermore, languages differ with respect to the functions they allow even if they are typologically similar, such as French and Spanish. The fact 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: it seems that not all the functions that Jackendoff proposes are present in all the languages to which this model has been applied. 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.

In addition, despite the fact that almost all basic functions are attested in Spanish NN compounds, these functions are not equally frequent or common. In fact, there are significant differences with respect to the productivity of the functions in the language, as it became evident once they were arranged on a hierarchy according to their frequency. In Spanish the most dominant basic function is SIMILAR, and an NN compound is very likely to bear a metaphorical or resemblance semantic relation, rather than, for example, a whole-part/part-whole relation yielded by the basic function PART, which is not very productive. The fact that the metaphorical relation is the most productive one could mean that NN compounds in Spanish can encode this relation by default. This, however, should be explored further. On the other hand, semantic relations like protection, classification, sourceless composition and kind cannot be expressed formally with NN compounds in Spanish. This is due to the absence of basic functions PROTECT, CLASSIFY, MADE FROM and KIND in Spanish NNs. A possibility for future research is to test the validity of the hierarchy I propose in (57) by gathering more data in Spanish. In addition, data from other languages can be collected to establish a comparison of the productivity of the semantic relations cross-linguistically.

Furthermore, I found that NN compounds in Spanish also allow promiscuity and cocomposition. While NN compounds in Spanish typically allow only a single interpretation, there are cases where various interpretations are possible. This means that some NN compounds are promiscuous; though the most frequent pattern seems to permit one reading only. On the other hand, cocomposition is available for Spanish NN compounds and seems not to be restricted by any particular reason; in other words, if the meaning underlying a compound is complex, Spanish can resort to cocomposition to describe its meaning (e.g., *hormiga bala* ‘bullet ant’). Some frequent cases involving cocomposition, though, include means of transport or machines (*coche diesel* ‘diesel car’ COMP + PF) and animals (*tiburón martillo* ‘hammerhead’ PART + SIMILAR).

As for the *N de Ns*, most of the functions present in NNs are also attested in *N de Ns* but in a different hierarchical order. A striking difference is SIMILAR which is very unlikely in *N de Ns*. Thus, it could be argued that if a function is very productive with a particular formal structure such as NNs, that function will tend not to be productive elsewhere: i.e. *N de Ns*. This conclusion is related to another finding: the availability of some functions seems to be restricted to a particular formal structure. That is, some functions are available for *N de Ns* but not for NNs and vice-versa. An example is

PROTECT: the semantic relation of protection is not operative in NNs but it is possible in *N de Ns*. This tendency is not only found in Spanish, but is also found in French, where PF, COMP and BE (IN/ON/AT) are found in *N de Ns* but not in NN compounds. A possible conclusion is that some basic functions and their respective semantic relations are structure specific in Romance languages.

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

Last but not least, taking all the facts discussed in the thesis, it could be argued that the peripheral nature of NN compounds in Spanish can be explained semantically. Since NN compounds have fewer and less productive basic functions and express fewer semantic relations than alternative constructions such as *N de Ns*, NN compounds are semantically restricted which limits their productivity. On the other hand, *N de Ns* are extremely flexible semantically which makes it easier to come up with new examples. Nevertheless, given the 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*, it seems that perhaps in the future NNs may become more productive.

The analysis of the factors affecting the productivity of compounds in Spanish using Jackendoff's (2009, 2016) model provides interesting insight in the semantic properties of NN compounds. In spite of the conclusions reached, it is important to mention that this study was carried out with a significant but restricted number of items. Thus, further research should consider creating a larger database to explore the tendencies indicated and the claims made in this thesis.

References

- Allen, Margaret R. 1978. *Morphological Investigations*. PhD thesis. University of Connecticut.
- Allen, Margaret R. 2000. Chapter 13. In Stekauer (ed.), *English Word Formation: a History of Research 1960-1995*: 201-218. Tübingen: Müller+Bass.
- Arnaud, Pierre J.L. and Renner, Vincent. 2014. 'English and French [NN]_N lexical units: a categorial, morphological and semantic comparison. *Word Structure* 7.1: 1-28.
- Bauer, Laurie. 1983. *English Word-Formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 2003. *Introducing Linguistic Morphology* (2nd ed.). Washington DC: Georgetown University Press.
- Busa, Ferica. 1997. *Compositionality and the Semantics of Nominals*. PhD thesis. Branderis University.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Davies, Mark. 2002-. *Corpus del Español: 100 million words, 1200s-1900s*. March, 3rd 2017. <http://www.corpusdelespanol.org>.
- Di Sciullo, Ana María. 2009. A View from Asymmetry Theory. In Lieber and 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.
- Fernández Domínguez, Jesús. 2016. The Semantics of Primary NN compounds: from form to meaning, and from meaning to form. In ten Hacken (ed.), *The Semantics of Compounding*: 129-149. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 2002. *Understanding Morphology*. London: Arnold.
- Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jackendoff, Ray. 2009. Compounding in the Parallel Architecture and Conceptual Semantics. In Lieber and 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.
- Katz, Jerrold J. and Postal, Paul M. 1964. *An Integrated Theory of Linguistic Descriptions*. Cambridge: MIT Press.
- Kornfield, Laura Malena. 2009. IE, Romance: Spanish. In Lieber and Stekauer (eds.), *The Oxford Handbook of Compounding*: 145-177. New York: Oxford University Press.
- Krott, Andrea, Gagné, Christina L. and Nicoladis, Elena. 2009. How the parts relate to the whole: Frequency Effects on Children's Interpretations of Novel Compounds. *Journal of Child Language* 36: 1-28.
- Lees, Robert B. 1960. *The Grammar of English Nominalizations*. Bloomington: Indiana University Press and The Hague: Mouton.
- Lees, Robert B. 1970. 'Problems in the analysis of English nominal compounds'. In Bierswisch and 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 and Díaz, Lourdes. 2001. Triggers in L2 Acquisition: the Case of Spanish N-N Compounds. *Studia Linguistica* 54: 197-211.
- Lieber, Rochelle and Stekauer, Pavol. 2009. *The Oxford Handbook of Compounding*. New York: Oxford University Press.
- Lieber, Rochelle. 2009. A Lexical Semantic Approach to Compounding. In Lieber and Stekauer (eds.), *The Oxford Handbook of Compounding*: 79-104. New York: Oxford University Press.
- Lieber, Rochelle. 2016. Compounding in the Lexical Semantic Framework. In ten Hacken (ed.), *The Semantics of Compounding*: 38-53. Cambridge: Cambridge University Press.
- Mellenius, Ingmarie and Rosenberg, Maria. 2016. The Semantics of Compounds in Swedish Child Language. In ten Hacken (ed.), *The Semantics of Compounding*: 110-128. Cambridge: Cambridge University Press.
- 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.
<http://site.ebrary.com.www.libproxy.wvu.edu/lib/nwvu/detail.action?docID=10481793>.
- Piera, Carlos. 1995. 'On Compounding in English and Spanish'. In Campos and Kempchinsky (eds.), *Evolution and Revolution in Linguistic Theory. Studies in Honor of Carlos P. Otero*: 1-13. Washington DC: Georgetown University Press.
- Rosch, Eleanor, Mervis, Carolyn, Gray, Wayne, Johnson, David and Boyes-Braem, Penny. 1976. 'Basic objects in natural categories'. *Cognitive Psychology* 8: 382-439.
- Rosenberg, Maria. 2013. Semantic and Formal Structure: a Corpus-Based Study of Swedish NN Compounds and their French Counterparts. In ten Hacken and Thomas (eds.), *The Semantics of Word Formation and Lexicalization*: 102-120. Edinburgh: Edinburgh University Press.
- Scalise, Sergio, Bisetto, Antonietta and Guevara, Emiliano. 2005. Selection in Compounding and Derivation. In Dressler, Kastovsky, Pfeiffer and Rainer (eds.), *Morphology and its Demarcations*: 133-150. Amsterdam: John Benjamins.
- Scalise, Sergio and Bisetto, Antonietta. 2009. The Classification of Compounds. In Lieber and 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. PhD thesis. Cambridge: MIT Press.
- Snyder, William. 2001. 'On the Nature of Syntactic Variation: Evidence from Complex Predicates and Complex Word-Formation'. *Language* 77: 324-342.
- Snyder, William, 2012. Parameter Theory And Motion Predicates. In Demonte and McNally (eds.), *Telicity, Change and State: a Cross-Categorial View of Event Structure*: n.p. Oxford University Press.
- ten Hacken, Pius. 1994. *Defining Morphology: A Principled Approach to Determining the Boundaries of Compounding, Derivation, and Inflection*. Hildesheim: Olms.
- ten Hacken, Pius. 2009. Early Generative Approaches. In Lieber and Stekauer (eds.), *The Oxford Handbook of Compounding*: 54-77. New York: Oxford University Press.

- ten Hacken, Pius. 2016. Introduction: Compounds and their Meaning. In ten Hacken (ed.), *The Semantics of Compounding*: 1-11. Cambridge: Cambridge University Press.
- ten Hacken, Pius. 2016. Three Analysis of Compounding: a Comparison. In ten Hacken (ed.), *The Semantics of Compounding*: 211-232. Cambridge: Cambridge University Press.