MDPI

Article

The Spell-Out of Non-Heads in Spanish Compounds: A Nanosyntactic Approach

Bárbara Marqueta Gracia

Department of Linguistics and Hispanic Literatures, University of Zaragoza, 50009 Zaragoza, Spain; marqueta@unizar.es

Abstract: This paper considers the morphophonological variants found in the non-head position of Spanish productive compound patterns. In the literature it has been noted that compounds resort to both *I*- and *O*-ending stems and words (even inflected ones) to spell-out compound non-heads. The study takes a nanosyntactic approach to compound syntax and examines the functional structure of the main classes of Spanish compounds. The goal of the analysis is to show that the functional structures of compounds determine to some extent which spell-outs are chosen in lexical innovations within the schemas, and also to address the issue of why certain allomorphs never appear in a given context while others seem to compete for lexical insertion. Research was conducted using a self-compiled sample of 1250 compounds extracted from contemporary Spanish written corpora. The results provide empirical support for syntacticist approaches to compounding, in that the proposed functional structures capture the predominantly compositional semantics of compounding, and also because traditional "linking vowels" are reconsidered to be the stable and systematic spell-outs for relational categories inside compounds.

Keywords: compound; word syntax; Spanish morphology; Nanosyntax; spell-out

1. Introduction

Compounds constitute one of the most controversial domains in morphological theory. There is scant consensus on where to set the boundary between compounds and phrases, on which kind of units are combined to create compounds (whether they are words or not), and on what kind of rules or processes generate compounds (see Scalise and Vogel 2010 for a summary of these issues).

Spanish, like other inflectional languages, is quite heterogeneous in terms of the type of units that serve to generate its compounds morphophonologically (henceforth *spell-out*).[1] There are four conditions (1–4):

Compounds containing a unit which is morphologically non-equivalent to any Spanish word, which I will refer to as a *stem* (S). These units, unlike morphosyntactic words (W), do not admit gender or number inflection, and are never seen in phrasal syntax (1):

1.   a.   [N+A]   <u>oj-i-azul</u> (S: *oji* W: *ojo*)
                     eye-I-blue
                     'blue-eyed'
     b.   [N+N]   <u>célul-o</u>-terapia (S: *célulo* W: *célula*)
                     cell-O-therapy
                     'cell therapy'
     c.   [A+A]   <u>guald-i-negro</u> (S: *gualdi* W: *gualda*)
                     yellow-I-black
                     'yellow-and-black'
     d.   [A+A]   <u>celt-o</u>-gaélico (S: *celto* W: *celta*)
                     Celt-O-Gaelic
                     'Celtic and Gaelic'

Compounds where a unit which is morphologically equivalent to a Spanish word appears in a compound context where units never inflect for number or gender (2):

2.   a.   [N+A]   <u>Google</u>-adicto
                        Google-addict
                        'google-addict'
      b.   [N+N]   <u>video</u>-llamada
                        video-call
                        'video call'
      c.   [A+A]   <u>verde</u>-azul
                        green-blue
                        green-and-blue'

Compounds where a unit which is morphologically equivalent to a Spanish word appears in a compound context where it can be (optionally) inflected for number[2], and for gender and number when it is an adjective (Rainer 1993, p. 280). These are the same contexts where, sporadically, compound units may experience phrasal expansion–in (3), *delanteros* 'front' modifies only to *cristales*, and not *levantacristales*, and *de policía* 'police' modifies only to *jefe*, and not *inspector-jefe*, so they do not behave as external modifiers of the whole compounds, but are parts of them–[3]:

3.   a.   [V+N]   quita-<u>pelo</u>-(s)   levanta-<u>cristal</u>-es   eléctric-o-s
                                                              delanter-o-s

                               remove-hair(PL)   lift-glass-PL   electric-M-S
                                                              front-M-S
                               'hair remover'   'front electric windows lift system'
      b.   [N+N]   hotel-es-<u>escuela</u>-(s)   inspector-jefe de policía
                               hotel-PL-school-(PL)   inspector-chief of police
                               'school-hotels'   'police chief inspector'
      c.   [A+A]   judí-(o/a)-   judí-(o/a)-(s)-
                               estadounidense-s   estadounidense-s
                               Jewish-(M/F)-   Jewish-(M/F)-(PL)-
                               American-PL   American-PL
                                      'Jewish-American'

Compounds where a unit is a Spanish word that appears in a compound context which is obligatorily inflected for number and gender, establishing agreement relations with phrases external to the compound (4):

4.   a.   [N+A]   adulto-<u>céntric-o-s</u>
                        adult-centric-M-PL
                        'adult-centric'
      b.   [A+A]   verd-i-<u>blanc-o-s</u>   hispan-o-<u>árabe-s</u>
                        green-I-white-M-PL   Spanish-O-Arab-PL
                        'green-and-white'   'Hispano-Arabic'
      c.   [N+N]   bici-<u>carril-es</u>   <u>avion-es</u>-espía
                        bike-lane-PL   plane-PL.spy
                        'cycle lanes'   'spy planes'

The words highlighted in (4) are the morphological heads of the compound. As the examples show, Spanish has both right-headed (4a–b) and left-headed (4c) compounds. This paper address only the spell-out of compound units in the contexts exemplified in (1–3), because the presence of a morphosyntactic word as the head of a compound is mandatory and systematic.

To sum up, *I*- and *O*-ending stems (non-words) are seen in (1), non-inflecting words in (2), and optionally-inflecting words in (3).

The aim of this study is to provide an analysis of a phenomenon that has thus far received little attention in the specialized literature. Through working with written contemporary Spanish corpora, it becomes evident that the same compounds are sometimes attested with different spell-outs, and hence it is possible to find a series of morphological variants for which there are no semantic repercussions, that is, the compound denotes the same concept in all of them. However, variation is not arbitrary. So:

The stems of (1a–c) co-appear with words in contexts like (2a–c), where inflection is not possible, as shown in (5a–c), but they never substitute words in the contexts of (3) (note that in (5d), impossible spell-outs are in parentheses).

5. 　　a.　　pech-i-frío　　　　　　　　　　　　pecho-frío
　　　　　　　chest-I-cold　　　　　　　　　　　　chest-cold
　　　　　　　'non-commited football player'
　　　b.　　danz-o-terapia　　　　　　　　　　danza-terapia
　　　　　　　dance-O-therapy　　　　　　　　　dance-therapy
　　　　　　　'dance therapy'
　　　c.　　blanqu-i-celeste　　　　　　　　　blanco-celeste
　　　　　　　white-I-light blue　　　　　　　　white-light blue
　　　　　　　'white-and-light-blue'
　　　d.　　lava-cara　　　　　　　　　　　　　lava-cara-s (S: *cari, *caro)
　　　　　　　wash-face　　　　　　　　　　　　　wash-face-s
　　　　　　　'wash basin'
　　　　　　　hotel-es escuela　　　　　　　　　hotel-es escuela-s (S: *escueli, *escuelo)
　　　　　　　hotel-PL-school　　　　　　　　　hotel-PL-school-PL
　　　　　　　'school hotels'

The adjectival *O*-stems of the type illustrated in (1d) co-occur with optionally inflected words, but not with I-stems (6a). The same is the case with noun stems of (1a–b), which are not interchangeable (6b).

6. 　　a.　　árab-o-isrraelí-es　　　árabe-(s)-isrraelí-es　　　(S: *arabi)
　　　　　　　Arab-O-Israeli　　　　　Arab-(s)-Israeli
　　　　　　　'Arab-Israeli'
　　　b.　　piern-i-corto　　　　　　　(S: *pierno)
　　　　　　　leg-I-short'
　　　　　　　'short-legged'

Paradoxically, then, some stems cannot alternate but can be replaced with words with no semantic repercussions.[4]

We can also observe that the same grammatical relation–coordination–can be spelled-out with any of the forms: *I*- and *O*-stems (7a), words that never inflect (7b), and optionally-inflecting ones (7c).[5]

7. 　　a.　　lect-o-escritura　　　　　　　　　verd-i-negro
　　　　　　　lect-O-writing　　　　　　　　　　green-I-black
　　　　　　　'reading and writing'　　　　　　'black and green'
　　　b.　　bata-manta　　　　　　　　　　　　azul-crema
　　　　　　　robe-blanket　　　　　　　　　　　blue+cream
　　　　　　　'sleeved blanket　　　　　　　　　'blue and cream'
　　　c.　　cliente-s empresa-(s)　　　　　　croata-(s)-musulmana-s
　　　　　　　client-PL-company-(PL)　　　　Croatian-PL-Muslim-PL
　　　　　　　'client companies'　　　　　　　'Croatian Muslims'

In what follows I will argue that compounds, regardless of the spell-out of their units, are syntactic structures that are at least partially compositional. As also happens with phrases, their interpretation depends on the meaningful contribution of the grammatical

relationships between their constituent parts and not only on the combination of the lexical meaning of the units themselves. It is thus necessary to overcome the customary definition of compounds as the "sum or roots or lexemes", due to the fact that, as shown in (3), the structural complexity of some units can go beyond the limits of the morphosyntactic word.[6] In line with neoconstructionist approaches, I take the view that compounds are syntactic projections that arise from the merging of concepts with functional categories (Mendívil 2019).

I develop a functional structure for compound units, which are merged through relational categories that determine their grammatical relations (in line with Di Sciullo 2005), even if these relations are not marked morphologically as in phrases. In fact, previous studies have defended the idea that the grammatical relationships in compounds receive no morphological marking (Guevara and Scalise 2009), but in this paper I will argue that relations can sometimes be revealed by the choice of certain stems. Thus, the elements in (1), traditionally considered as "linking elements" in the Spanish morphological tradition —meaningless sound material appearing between compound members—(Val 1999, p. 4813; Buenafuentes 2020, pp. 287, 298), are seen as the spell-outs of the relational structures of the compounds (Delfitto and Melloni 2009; Marqueta 2019b).

I assume a nanosyntactic approach (Starke 2009; Fábregas 2016), showing how the principles of the syntax–lexicon interface of this framework capture the systematic connections between compound patterns and specific types of allomorphs, and thus helping to predict the syntactic contexts in which these allomorphs compete for spell-out (5–6). Besides, contrary to the view of the lexicon found in Distributed Morphology, Nanosyntax allows for the storage of syntactically complex objects, such as idioms and compounds.

The remainder of the paper is organized as follows: In Section 2, I present some methodological considerations. I also specify the notion of compound as it is used in this study. In Section 3, I describe the basic assumptions of Nanosyntax in terms of the nature of word syntax and its interface with lexical entries. In Section 4, I explore the different functional structures of the five most productive Spanish compound patterns, providing an analysis of each one, with a focus on their (im)possible spell-outs. Section 5 discusses the results of this study.

2. Materials and Methods

The analysis is based on a self-collected sample of 1250 compounds drawn from the *Corpus del Español NOW* (https://bit.ly/3DcgE6W) (accessed on 23 December 2021) and the database of neologisms from the *Observatori de Neologia* (https://bit.ly/322JyZF) (accessed on 23 December 2021). Hence, only compounds that contemporary native speaker use are involved. This is crucial, in that elsewhere in the literature synchronic descriptions of compounds sometimes use examples that are not truly representative of current usage,[7] in that they are infrequently used words or are not especially familiar to normal speakers: *carricuba* lit. car+i+barrel, 'watercart', *ajicomino* lit. garlic+i+comin, 'sauce with garlic and cumin', and *piedra mármol* lit. stone+marble, 'marble stone'. As Ricca (2010) points out, it is not uncommon for very old compounds to undergo arbitrary lexicalization processes and to specialize their meanings, which is the reason why a sample of low-token frequency words and *hapaxes* within a productive compound pattern will, I believe, afford us a more realistic picture of the functioning of compound rules.

As noted in the previous section, this study is concerned with analyzing the morphological variants linked to contemporary compound patterns, that is, those that give rise to a considerable number of new compounds today. According to my corpora-based sample, these are the following:

A. Left-headed N+N structures (559 examples)

dieta milagro   *bailarín coreógrafo*

diet miracle   dancer choreographer

'miracle diet'   'choreographer-dancer'

B. Coordinated adjectives: (229 examples)

materno-filial   *catalan-o/a-aragones-a*

mother-O-child-AL   Catalan-O/M/F-Aragonese-F

'mother-child'   'Catalan-Aragonese'

C. VN compounds (**227 examples**)

riza-pestañas   *limpia-hogar*

curl-eyelashes   clean-home

'eyelash curler'   'household cloths'

D. Right-headed N+N compounds (175 examples)[8]

gastro-turismo   *Cristiano-manía*

gastro-tourism   Cristiano-mania

'food tourism'   'Cristianomania'

E. N+Adj inalienable compounds (62 examples):

pel-i-teñid-a   *oj-i-rroj-o*

hair-I-dyed-F   eye-I-red-M

'hair-dyed'   'red-eyed'

There is no agreement in the morphological literature as to which constructions should be classified as compounds (Lieber and Štekauer 2009). I assume a structural-based delimitation of compounding: a compound in Spanish is formed whenever one of its members (the non-head) is merged[9] with the other before its full syntactic projection is concluded.[10] In the examples for A, C and D, the nouns are combined with the other members of the compound as bare singular/plural nouns, but never as Determiner Phrases (DPs) (8a), Prepositional Phrases (PPs) (8b) or Coordination Phrases or &Ps (Progrovac 1998) (8c), even if the structural meaning of these phrases would be roughly the same as the one of the compounds.

8. a. *rizapestazañas* 'eyelash curler' Phrase: *riza las pestañas* 'it curls the eyelashes'
    b. *gastroturismo* 'food tourism' Phrase: *turismo de gastronomía* 'tourism of gastronomy'
    c. *bailarin coreógrafo* 'choreographer-dancer' Phrase: *bailarín y coreógrafo* 'both a dancer and a choreographer'

In the examples in B, adjectives show optional agreement, contrary to phrasal coordinated adjectives, where agreement is mandatory (9).

9. *la corona catalano/a-aragonesa*
    the-F crown Catalan-O/(F)-Aragonese-F
    'Catalan-Aragonese crown'
    *la corona catalan-o y aragonesa   *la corona catalan-a y aragonesa*
    the-F crown Catalan-O-aragonese-F   the-F crown Catalan-F and aragonese-F
    'the Catalan and Aragonese crown'

According to this delimitation of the compound, N+P+N constructions such as *diente de león* 'tooth of lion' 'dandelion', and N+Adj ones like *perrito caliente* 'little dog+hot' 'hot dog' are not compounds, but rather phrasal lexical units, even if they do, like compounds, name entities for unitary concepts. They are often referred to as *phrasal compounds* in the Spanish morphological tradition (Bustos-Gisbert 1986), but also as syntactic compounds (Rio-Torto and Ribeiro 2009), syntactic words (DiSciullo and Williams 1987), or, in analyses of other Romance languages, lexicalized phrases (Fradin 2009). The latter concept is the most suitable to reflect the fact that they are phrases that must be stored in the lexicon

because of their non-compositional meaning, as many VP phrases are (without being considered compounds for that reason) (10):

10.	Lexicalized NPs:	canto de cisne	marcha-atrás
		song of swan	walk-back
		'last work of a dying person'	'withdrawal'
	Lexicalized VPs:	cantar las cuarenta	irse por los cerros de Úbeda
		to sing the forty	to go by the hills of Úbeda
		'to tell sb off'	'to beat around the bush'

According to our approach, compounds are neither necessarily idiomatic (in fact, the majority of the examples in our sample are fairly transparent for native Spanish speakers), nor necessarily one-stressed. They differ from phrases in that the relations between constituents are not expressed through agreement, prepositions or conjunctions, and in that they may present special allomorphs. In Spanish, special allomorphy is restricted to the non-head position of the compound, where words and stems alternate, as shown in (3).

In the external position, Greek and Latin stems are not allomorphic variants of Spanish words, but productive bound stems (*parquímetro* 'parking meter'; *seriéfilo* 'series fan') because the Spanish composition system requires full (and stressed) morphosyntactic words in that position. However, some neoclassical patterns give rise to new words quite often, such as the stems above, and others like *-plastia* (*gluteoplastia* 'gluteoplasty') and *-fóbo* (*lesbófobo* 'lesbophobic'), and we do occasionally find pairs of compounds ending either in a bound stem or a Spanish word, such as *gastroplastia/gastrocirugía* 'gastroplasty; stomach surgery' and *gordófobo/gordofóbico* 'fatphobic'.

3. A Nanosyntactic Approach to Spanish Compounds

It was proposed in the previous section that compounds are distinguishable from phrases because of the restricted syntactic projection of one of their members and their special allomorphy. This implies the existence of structural differences between compounds and phrases to some extent. As we know, there are, on the one hand, lexicalist approaches to word formation, which assume that compounds and phrases are formed through different rules.[11] On the other hand, there are non-lexicalist approaches (prior and subsequent to Lexicalism) which account for compound formation with syntactic rules (Lees 1960; Lieber 1992; Kornfeld 2004; Harley 2009; Marqueta 2019b)[12].

Neoconstructionist approaches, such as Distributed Morphology and Nanosyntax, deviate from lexicalist approaches in that they assume that there is one single structure-building module, the syntax, which is responsible for sentence structure as well as for word structure. There is no "morphology" or a "lexicon", understood as the loci where words are formed before they are used to build syntactic structures.

A key assumption of these models is that lexical items are not the minimal building blocks of structures, but rather syntactic features. So, it is equally possible that a set of these features are "spelled-out" as a word in one language, but as a phrase in another.

Syntax is assumed to be a vehicle for expressing grammatical semantics, arranging syntactico-semantic features from a (perhaps universal) pre-syntactic lexicon. So, there is a clear-cut distinction between the (grammatically relevant) meaning of such features, and their extralinguistic or conceptual meaning, which is not relevant for syntax and is thus linked to the structures post-syntactically through the lexicon (*Late-Insertion hypothesis*), or in the conceptual–intentional component of human grammars.

Nanosyntax is a generative approach linked to cartographic assumptions. It assumes that syntactico-semantic features are arranged in functional projections where each feature is assumed to be an independent head that projects[13]. A syntactic structure is constructed by merging the features one by one, starting from the bottom of a functional projection and stopping at the points where *lexical insertion* happens.

The functional structures codifying grammatical meaning in compounds are comprised by categorizing heads (*N*, *Adj*), spelled-out in Spanish through theme vowels

(Bermúdez-Otero 2013), interpretable and non-interpretable projections with features of *Number* and *Gender*, these merged above the categorizing heads, and a series of relational categories conditioning the grammatical relationships between members: *Relation*, *Possession*, *Identity*, and *Conj(unction)*, whose properties will be addressed in the next section.

The lexical insertion system of Nanosyntax differs from that of Distributed Morphology in that spell-out does not necessarily target terminal nodes or heads, but whole chunks of structure or syntactic trees (*Phrasal Spell-out*). This is possible because Nanosyntax assumes a more conventional view of lexical entries: while DM divides the lexicon into roots and functional categories, Nanosyntax contemplates full lexical entries with slots for syntactico-semantic structures, phonological forms, and conceptual information.[14] This does not imply that syntax "projects" from the lexicon, as in mainstream generative grammar, because the lexicon in Nanosyntax is only an interpretative component: it "fills" the syntactic structures, matching them with the appropriate material from the lexicon.

The lexicon in Nanosyntax is an interface phenomenon. It is constructed through the language acquisition process by the storage of specific structural configurations, since syntax is continuously producing syntactico-semantic trees, some of which merit storage in the lexicon linked to a phonological as well as–though only in the case of contentful units–conceptual information (either in the format of words, compounds or phrasal idioms).

This implies that a novel compound can be formed by syntax and not necessarily stored. In other words, it is predicted that not all the possible, well-formed compounds in a language are stored in the speaker's lexicon,[15] just as not all possible phrases are, even if one cannot deny that syntactic structures of compound size seem to be more prone or "privileged" to storage than fully-fledged sentences, for instance.

Because lexical entries can target larger pieces of syntactic structure, this model is able, without a post-syntactic component of morphological readjustment rules, to account for phenomena such as *syncretism* (where the same lexical entry maps syntactic trees with more or fewer features). Furthermore, in Nanosyntax, lexical entries compete when there is more than one lexical entry able to spell-out the same syntactic tree (allomorphy), and there is a series of spell-out principles that regulate lexical insertion. For the purposes of our present study, there are three of these principles that can help us to account for the variation within Spanish compounding: the *Exhaustive Lexicalization Principle*, the *Superset Principle*, and *Panini Condition*.

The *Exhaustive Lexicalization Principle* (*ELP*) (Fábregas 2007) posits that every syntactic feature must be lexicalized. Let us recall the compounds with non-word constituents in (1): *ojiazul*, *céluloterapia*, *gualdinegro*, *celto-gaélico*. Units such as *oji* and *célulo* never appear in phrases. The traditional view of such units in Lexicalist models is that morphology constructs words with different units than syntax does (Ralli 2013). However, the *ELP* principle above would prevent these units from entering phrases if the syntactico-semantic tree associated with them lacks gender and number features. The nouns and adjectives in a phrase always enter into an agreement process, so the lexicon must feed phrases with units able to identify these features. So, the *ELP* rules out the possibility that noun and adjective allomorphs unable to inflect for gender and number can spell-out NPs and AdjPs in phrases.

The *Superset Principle* (Caha 2009) assumes that a lexical entry is inserted into a syntactic tree if it matches at least one subconstituent. Let us recall compounds with Spanish words in the context of (2): *Googleadicto*, *videollamada*, *verdeazul*. Here, nouns never inflect for number or establish agreement relationships (contrary to the heads of these compounds). So, Spanish words can spell-out compound members even if they contain more features (gender, number) than required in a given context.

In the context of (3), number and/or gender features also appear that are not required: *quitapelos (quitapelo)*, *hoteles escuelas (hoteles escuela)*, *judías-estadounidenses (judeo-estadounidenses)*. Each of these patterns must be examined individually, taking into account the *Panini Condition*, which states that, when two lexical units are available for the same

syntactic node, the speaker will select the more specific lexical item (the more similar form in number and the nature of features) (Fábregas 2016)[16].

In the next section, I will show the structural contexts and the lexical entries available for spelling these out.

4. Results

4.1. Right-Headed N+N compounds

The type of compounds exemplified by words such as *videoartista* 'video artist' is currently becoming more and more significant as a word-formation pattern in Spanish (Moyna 2020, p. 314). Its growth is clearly linked to the effects of globalization—products, techniques and concepts spreading throughout the world which have a name recognizable in all countries—and to the enormous influence of English worldwide, especially notable among users of the internet and social networks, which are a breeding ground for creativity.

Spanish shows two types of spell-out for the noun non-head in these compounds.[17] It uses common Spanish words, as in the examples in (11), and *I-* and *O*-ending stems inspired by those appearing in neoclassical compounds, this a shared feature of diverse Indo-European languages (12):

11. Café-manía   lider-dependencia   Marca-adicto
    Coffee-obsession   leader-dependency   Marca-addict
    'Coffee lover'   'leader dependency'   'Marca addict'

12. Vertebr-o-terapia   fungu-i-turismo   islam-o-fascismo
    Vertebra-O-therapy   fungus-I-tourism   Islam-O-fascism
    'Vertebra therapy'   'fungus tourism'   'Islamofascism'

Both Spanish words and stems are legitimate spell-outs, as shown by the fact that the same compound is attested in both ways in some cases (13):

13. *danza-terapia* and *danzo-terapia* 'dance therapy'
    can-i-cross and *perro-cross* 'canicross'
    blog-o-novela and *blog-novela* 'blog-novel'
    ciclo-turismo and *bici-turismo* 'bycicle tourism'
    drogo-dependencia and *droga-dependencia* 'drug addiction'

One of the variants (the one on the left) is the more frequent, but asymmetries in frequency are to be expected, considering that most new terms end up spreading in only one form. Prior to lexicalization and "institutionalization" processes, the Spanish speaker can use both Ws and Ss to create new compounds. The choice would depend on unsystematizable factors, such as the so-called "family size" of the compound member (Van Jaarsveld et al. 1994).[18]

However, it is possible to predict (and formalize) the impossible spell-outs of the noun modifier. Neither Ws nor Ss can be inflected for number (**drogos/drogas-dependencia* meaning 'dependency to more than one kind of drug'). Gender, an arbitrary classifier in Spanish nouns which determines agreement within the DP, is also irrelevant in this context, because the non-heads do not enter into agreement processes in such a context (*el biciturismo* and not *la biciturismo).

In Figure 1, a model of lexical entry for right-headed Spanish compounds is provided. To the left, the acceptable and unacceptable spell-outs are provided; in the center we see the functional structure assigned to these compounds, focusing on the context for lexical insertion of non-head allomorphs; to the right are both the structural semantics and the conceptual meaning of the compound.
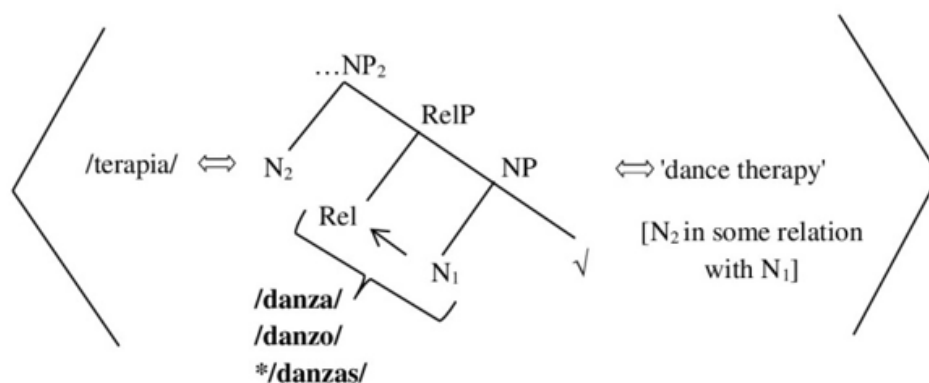
Figure 1. Lexical entry for a right-headed N+N compound.

Figure 1 shows that the noun-head of the compound is the projection of a categorial head N_1 (NP), with no further development of its functional structure (The asterisk in */danzas/ indicates that the example would be ill-formed). The merge with the head of the compound is mediated by the most basic form of the relational category, the so-called R function (Downing 1977; Guevara and Scalise 2009). I assume, in line with Delfitto and Melloni (2009), that the merge of two identical objects (N_1 and N_2) causes a point of symmetry that is broken by a conflation movement (indicated with the arrow) of one of the nouns (the non-head) to a functional head (*Relation* in this case). The mediation of *Relation* and no other functional head guarantees interpretative freedom. Both encyclopedic and pragmatic information combine to determine the relationship that is understood between the two compound members. The conflated structure of Rel > N can be spelled-out with a Spanish word like *danza* (following the *Superset Principle*), and with a more precise spell out (*danzo*) without associated gender and number features (following the *Panini Condition*).

4.2. N+Adj Compounds Codifying Inalienable Possession

The type of compounds represented by words such as *pelirrojo* lit. hair+I+red, red-haired' has been relatively productive as a word formation pattern in Spanish since c. 1400 (Moyna 2020, p. 311). The main feature of these compounds is that they codify a specific kind of grammatical relationship: *inalienable possession*. The non-head (*peli* above) is understood as an inseparable, constitutive part of an inferable or explicit possessor (typically a person or an animal).

Spanish exhibits two types of spell-out for the noun non-head in these compounds: one by default is an *I*-ending stem (14) and the other is—very occasionally, and with only nouns denoting body parts—a common Spanish word (15). Some alternations between them are found in (16).

14.	oj-i-azul	franj-i-verde	corn-i-cort-o	pel-i-teñid-a
	eye-I-blue	stripe-I-green	horn-I-short-M	hair-I-dyed-F
	'blue-eyed'	'green-stripped'	'short-horned'	'with her hair dyed'

15.	picha-brav-a	pecho-lobo
	dick-boastful-F	chest-(of)-wolf
	'boastful'	'man with a hairy chest'

16.	*pecho-amarill-o* and *pech-i-amarillo*	
	chest-yellow-M	chest-I-yellow-M
	'great kiskadee'	
	piqu-i-negr-o and *pico-negr-o*	
	bill-I-black-M	bill-black-M
	'black-billed thrush'	

However, inalienable compounds such as (15) can be N+N, like *pecholobo* above, and show variable internal agreement (*picha-s-brava-s*, *pecho-s-lobo-s*), contrary to compounds with *I* stems in (14). This leads us to think that inalienable possession with bodyparts is codified through different structures: there are exocentric (*bahuvrīhi*) compounds such as (15), formed when a noun (*picha*, *pecho* above) merges with its modifier (*brava*, *de lobo*) and refers metonymically to a possessor (a male human), and there are endocentric compounds, such as (14), for which the part–whole relationship is codified in the compound structure, as explained below.

More importantly, we have reasons to suspect that the *I*-ending stems in (14) do not spell-out the same structure as the *I*-stems in the previous section (Rel > N). From a morphological point of view, they never alternate with *O*-ending stems.[19] Semantically, the compounds in (14) are not interpretatively free like the compounds described in Section 4.1: the non-heads are obligatorily understood as inalienable possessees of some external possessor.

In Figure 2, a model of lexical entry for right-headed Spanish inalienable compounds is provided. To the left, the acceptable and unacceptable spell-outs are provided; in the center, the functional structure assigned to these compounds is shown, focusing on the context for lexical insertion of non-head *I*-ending allomorphs; to the right, both the structural semantics and the conceptual meaning of the compound are shown.
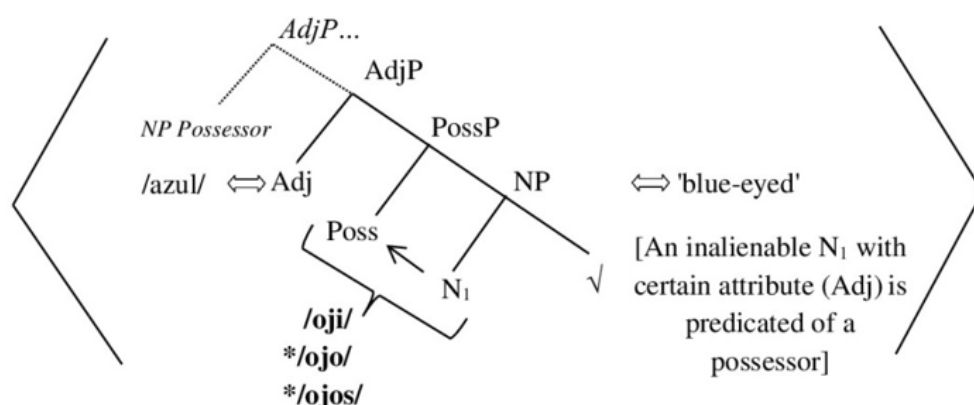

Figure 2. Lexical entry for a right-headed inalienable possession compound.

Figure 2 shows that the noun-head of the compound is the projection of a categorial head N_1 (NP), with no further development of its functional structure. The merge with the head of the compound is mediated by a semantically specific relational category, the so-called *Possession* (Gil Laforga 2014; Marqueta 2019a). This head causes its NP complements to be obligatorily understood as inalienably possessed for an entity external to the compound, although possessor and possessee does not form a syntactic constituent (the possessee is adjoined to the adjectival head, not to the NP possessor). But this is not problematic, because, as Guéron (1992) suggests, inalienable possession relationships can be treated as a form of anaphoric binding. The conflated structure of Poss>N can be spelled-out in Spanish only with a specific I-allomorph, because any other spell-out would not lexicalize Poss (according to the *Exhaustive Lexicalization Principle*).[20]

4.3. Left-Headed N+N Compounds

The pattern exemplified by *cláusula suelo* lit. clause+floor 'floor clause' is most frequently used as a neological strategy of nominal compounding (considering the number of hapaxes and low token frequency examples in the corpora). Its current productivity is higher than the one predicted by data extracted from lexicographical sources (Moyna 2020, p. 310), perhaps because many of the innovations within this structure never become institutionalized enough to be included in dictionaries, and maybe because of its "phrasal-like" features.

These compounds are left-headed structures made up of two phonological words, which show regular head inflection and compositional semantics. In fact, at first glance they are barely distinguishable from Spanish syntactic appositions. There are some proposals on how to distinguish between compounds, understood as morphological objects, and syntactic appositions (Bartoš 1999; Garcia-Page 2011), as well as proposals that account for their differences, assuming that compounds are also syntactic structures (Fábregas 2005; Marqueta 2020).

In addition to this, N+N seems to be the shared spell-out for at least three different grammatical relationships: compounds made up of a head and a modifier where the latter subclassifies the former (17a); compounds made up of a head and a noun predicate (17b), and coordinating compounds (17c):

17. a. película río   capital golondrina   avión espía
         film river   capital swallow   plane spy
         'novel sequence'   'swallow capital'   'spy plane'
     b. hora bruja   bono basura   fiesta protesta
         hour witch   bond junk   party protest
         'period when a baby cries' 'junk bond'   'protest party'
     c. casa museo   bailarín coreógrafo
         house museum   dancer choreographer
         'house museum'   'choreographer dancer'

The possibility of finding a plural non-head is expected to reflect these different grammatical relations to some extent. For subordinating compounds (17a), only head inflection is expected; for attributive compounds (17b), pluralization of the non-head could reflect their "adjective-like" status, or the fact that examples in (17b) are phrases (Buenafuentes 2020, pp. 296–97); for coordinating compounds, number inflection of the non-head is much more frequent, maybe reflecting the fact that they possess two semantic heads.

Left-headed NN constructions, then, show two types of spell-out for the noun non-head in these compounds, with or without plural inflection. Therefore, we find plural compounds that are attested only with a singular non-head (18). The most frequent situation, however, is to find plural compounds attested with a non-head either in singular or plural, being much more frequent with the singular (19). The rarest situation is to find plural compounds attested in both forms, these being more frequent with a plural non-head (20). Finally, there are a considerable number of compounds where the non-head is only attested in the plural (21):

18. *hombres anuncio* (*hombres anuncios) 'sandwich men'
     capitales riesgo (*capitales riesgos) 'risk capitals'
     claúsulas túnel (*cláusulas túneles) 'tunnel clauses'

19. *hombres bomba(s)* 'bomb men'
     partidos bisagra(s) 'hinge parties' (they align with different ideologies)
     aeropuertos fantasma(s) 'ghost airports'
     naciones estado(s) 'nation states'
     villas miseria(s) 'shanty towns'
     casas cueva(s) 'earth shelters'

20. *aviones espía(s)* 'spy planes'
     hombres dios(es) 'God men'
     categorías reina(s) 'the queen categories/premier classes'
     delanteros tanque(s) 'tank/strong forwards'
     hoteles restaurante(s) 'restaurant hotels'

21. *cementerios museos* 'museums graveyards'
    pintores poetas 'painters poets'
    escritores periodistas 'journalists writers'
    padres abuelos 'parents grandparents'

The alternating categories of (19) and (20) host subordinate, attributive and coordinate relationships, so number variation does not seem to be determined by those. However, we do find some tendencies, such as the fact that animate predicate non-heads favor plurality (Fábregas 2005), and that plural attestations are far more prolific with noun non-heads which have a high token frequency (as opposed to *hapaxes* and low token frequency nouns), such as *bomba, basura* and *fantasma*.

In Figure 3, a model of lexical entry for left-headed Spanish NN compounds is provided. To the left, the acceptable and unacceptable spell-outs are provided; in the center, the functional structure assigned to these compounds is shown, focusing on the context for lexical insertion of non-head allomorphs; to the right, both the structural semantics and the conceptual meaning of the compound are shown.
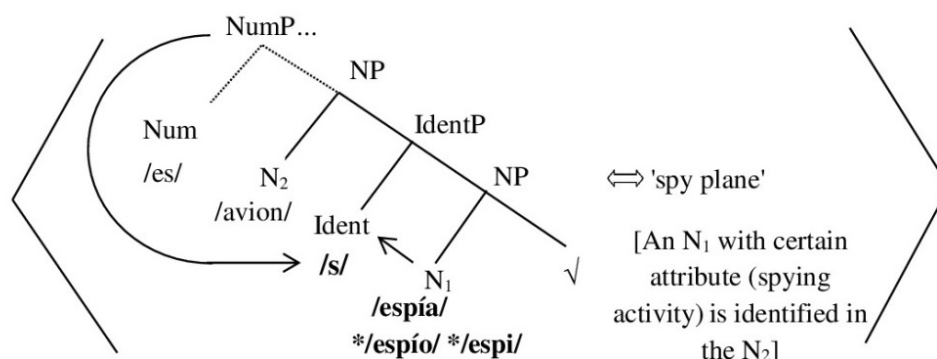
Figure 3. Lexical entry for a left-headed N+N compound.

Figure 3 shows that the non-head of the compound is the projection of a categorial head N_1 (NP), with *a priori* acceptable development of its functional structure (Num, Gen, D), because its spell-out is a morphosyntactic and phonological (stressed) word (*espía*) (Delfitto and Melloni 2009). A number-inflecting N head such as *espía(s)* cannot be merged with the *Relation* head, as described in Section 4.1: note that the spell-out of that head can be a stem. Number-inflecting Ns combine in syntax with a well-known relational head: the preposition. However, in a compound such as *avión espía* a preposition does not mediate the merge between the members (*avión de espía*)—in such a case we would have a phrase, not a compound—so the functional category that relates nouns in left-headed compounds is "prepositional-like" without being spelled-out as a preposition. I assume that this happens because the structural semantics associated with the functional head in the compound is not comparable to Spanish prepositions. The relation between the noun members of NN left-headed compounds is mediated by a functional category called *Identification* (Marqueta 2017).

While right-headed compounds and N+P+N phrases establish different types of semantic relations between their members, such as locative (*Eurotúnel* 'Eurotunnel') or causative (*herida de bala* 'bullet wound')—or any other of those stipulated for English N+N compounding in Jackendoff (2009)—left-headed Spanish NN compounds only allow the non-head to be understood as contributing to the meaning of the compound because some of its semantic features are "identified" in the head itself.[21] As such, for subordinating compounds like *hombre bomba*, the head is identified by some features of the non-head *bomba* (the fact that the man can explode) and not by another (e.g., the compound does not mean that the man has a round belly, like a prototypical bomb). For attributive compounds such a *categoría reina*, the head *categoría* is identified with a feature of the non-head (important like a queen), and never interpreted as, for instance, 'a category where only queens can

participate'. In coordinating compounds, as Lieber (2009) points out, there is always some kind of coincidence or "identification" between the compound members; the ones in *pintor poeta* share the feature 'human occupation', *casa cueva* are both locations where people can live, etcetera. Even if the grammatical relationships are different, I propose that they are constructed by a common, basic structure (Figure 3).

With regard to spell-out, *Identification* lacks a specific functional word to spell it out (a preposition or a conjunction), as noted above. Due to the fact that number inflection for the non-heads is not semantically relevant, being neither referential nor quantifying (the non-head does not refer to a "group" of espías in *aviones espías*, but only to more than one plane), we cannot propose a Number Phrase for these non-heads.

Neither can the number inflection of *espías* be understood as syntactic agreement at first, because non-heads such as *espía* are not adjectives. However, I suggest that the alternations of (18–21) are not so rare, but, on the contrary, quite predictable if we look at the behavior of noun predicates and modifiers in identificative contexts in Spanish phrasal syntax.

Nouns also take plural inflection when predicating a plural subject in copular identificative sentences (22a). NPN phrases with identificative modifiers (introduced by preposition *con* 'with') can alternate between singular and plural spell-out (22b). Even coordinate nouns in identificative contexts can alternate (22c):

22. a. *Sus hermanos son ángeles/*angel* 'They brothers are angels' (good people)
  b. *Son hombres con bomba/bombas* 'They are men with a bomb/bombs'
  c. *Algunos son escritor(es) y periodista(s)* 'Some are writer(s) and journalist(s)'

So, I suggest that plural inflection can spell-out the Identification Phrase (after the conflation between the NP and this category, signaled in Figure 3 above with arrows). Co-plurality, like agreement, would only serve to mark morphologically the fact that there exists an identification between these nouns, and thus it can in a sense be predicted that the more semantic features the nouns share, the more likelihood there is for co-plurality to appear, as happens with coordinating compounds.

4.4. V+N Compounds

The type of compounds represented by words such as *cazanazis* lit. 'hunt+nazis' 'nazi-hunter' is one of the most productive compounding patterns in Spanish since the origins of the language, being one of the hallmarks of the evolution from Latin to Romance (Moyna 2020, p. 311). One of the main features of such compounds in Spanish is that the noun non-head of these compounds is spelled-out either in a singular form or in a plural one, the latter being far more frequent. There has been some discussion in the literature as to an explanation for why the nouns appear in plural, since plurality is not required for agreement purposes (Varela 1990; Moyna 2011, p. 63; Val 1999, p. 4797; Buenafuentes 2014). Other studies have examined the contexts in which the nouns appear in singular, this being the marked option (Alvar 1984, p. 84; Moyna 2011, p. 209; Von Heusinger and Schwarze 2013).

As for N+N compounds in Section 4.4, we need to account for four possibilities. First, there are compounds where the non-head only appears in the plural (23), which is clearly the most frequent scenario when V+N compounds refer to animate entities in light of their flaws or some sort of problem. Second, we find compounds whose non-heads are attested both in singular and plural, the latter being far more frequent (24). Third, we find alternating nouns that are more frequently attested in the singular (25). These present some of the characteristics that, according to the literature, favor the singular (for example, nouns with unique referents, as well as mass nouns). Finally, in my own data I found just two compounds attested with the noun only in singular, *guardabotellón* lit. 'keep+big bottle', 'street-drinking keeper' and *manchapecho* lit. 'stain+chest' 'carapulcra'[22].

23. abraza-farola-s   destroza-matrimonio-s   junta-letra-s
  hug-lamp post-PL   destroy-marriage-PL   joint-letter-PL
  'a drunk person'   'a mistress'   'a mediocre writter'
  corta-rollo-s   limpia-playa-s   porta-servilleta-s
  cut-roll-PL   clean-beach-PL   carry-serviette-PL
  'Roll cutter'   'windland machine'   'serviette-carrier'

24. asaltabanco(s) 'rob+bank(s)' 'burglar'
  marcalibro(s) 'mark+book(s)'bookmark'
  pagafanta(s) 'pay+fanta(s)' 'friend-zoned'
  vendepatria(s) 'sell+fatherland(s)'traitor'
  matapasion(es) 'kill+passion(s)' 'not sexy underwear'
  batefuego(s) 'beat+fire(s)' 'extinguising tool'
  guardatiempo(s) 'keep+time(s)' 'timepiece'
  quemagrasa(s) 'burn+fat(s)' 'fat-burning'

25. mataleón(es) 'kill+lion(s)' 'rear naked choke'
  cubretecho(s) 'cover+ceiling(s)' 'tent'
  quitalluvia(s) 'remove+rain(s)' 'small plane'
  limpiahogar(es) 'clean+home(s)' 'house cleaner'

In Figure 4, a model of lexical entry for Spanish VN compounds is provided. To the left, the acceptable and unacceptable spell-outs are provided; in the center, the functional structure assigned to these compounds is shown, focusing on the context for lexical insertion of non-head allomorphs; to the right, both the structural semantics and the conceptual meaning of the compound are shown.
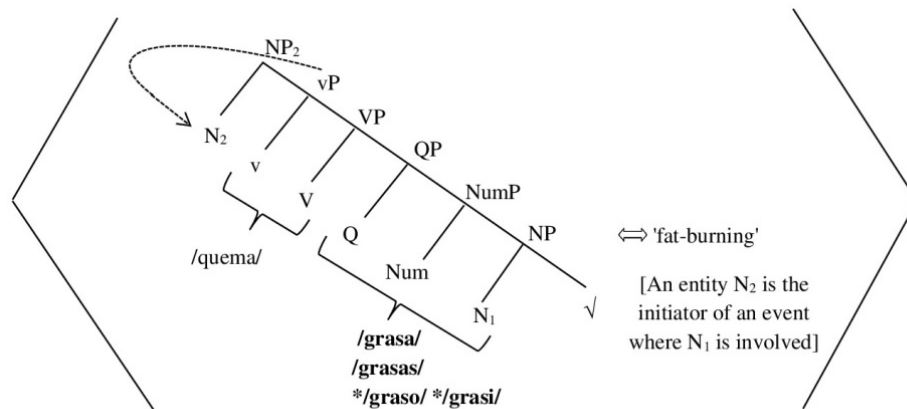

Figure 4. Lexical entry for a V+N compound.

Figure 4 shows that the noun-head of the compound is the projection of a categorial head N_1 (NP), with acceptable development of its functional structure (Num, Q), because its spell-out is a morphosyntactic and phonological (stressed) word (*grasa*). A number-inflecting N head such as *grasas* can be merged directly as the internal argument of a *V* head, spelled-out with a verbal lexical entry (provided with a thematic vowel) (*quema*). Following Marqueta (2018), it is assumed that /*quema*/also spells-out the functional category (little) *v*, which explains why the resulting denotation of the compounds is restricted to initiators of the event denoted by the verbal predication (Ramchand 2008). Through a remerge movement to the specifier position of *v* (Gärtner 2002), the compound can be recategorized as a noun preserving the denotation of the event (thus, a *quemagrasas* would denote a specific fact-burning product). However, this recategorization process is not necessary when the specifier position of *v* is occupied by an NP external to the compound, as in *actividad quemagrasas* (fat-burning activity), where *quemagrasas* remains as a predicate. Franco (2015) has also developed a Nanosyntactic approach to VN compounds which

focuses on explaining why they are not always nouns. He treats them as reduced relative clauses that, via phrasal spell-out, are reintroduced/rebooted in the syntax from the lexicon. My only reservation with this proposal is that it assumes the loss of all the functional material in the lexicalization process, so it is difficult to figure out why prepositions should remain in examples such as *saltimboca* lit. 'jump in mouth' 'saltimbocca' and *cantambanco* lit. 'sing in table' 'busker' (Franco 2015, p. 84) if relative heads must disappear.

In contrast to the N+N left-headed compounds of the previous section, it is possible to postulate the presence of a Number Phrase, because, although the non-head does not (apparently) enter into an agreement operation with a D head (although, see Scalise et al. 2009), we cannot clearly affirm that plural inflection is not semantically relevant in V+N compounds. Drawing on observations in the literature—that number is semantically relevant because of the generic aspect and habituality of the predication—I assume that the non-head projects a Quantifier Phrase, able to check the Num features. Plural non-heads appear by default with count nouns (*parachoques* 'stop+crashes') to refer to an unspecific quantity of Ns, but plurality can be extended over mass nouns like *paraguas* by virtue of the plurality of "water-stopping" events in which an umbrella is implied. Despite that, mass nouns tend to appear in the singular since, being unbounded, homogeneous entities, imply a plurality of points. Finally, count nouns appearing in singular (*mataleón*) are understood as unique referents, and thus they express cardinality. It seems to me that the choice between singular and plural depends on the specific way in which a speaker perceives the event; for instance, for *mataleón*, the speaker may have perceived that the rear naked choke is more easily performed only with one lion (person). So, he or she favored a singular noun/unique referent.

4.5. Adj+Adj Compounds

The type of compounds represented by examples such as *judeo-estadounidense* 'Jewish+American' is one of the most productive compounding patterns in contemporary Spanish written texts (Val 1999, p. 4807). Despite this, there are barely any studies devoted specifically to these compounds (Grossmann and Rainer 2009), although there is some discussion on the status of coordinate structures as compounded or phrasal ones (Padrosa 2011; Felíu 2016), and on whether they are one-stressed or double-stressed constructions (Val 1999, p. 4812; Hualde 2007).

We can identify two different types of adjectival compounds. On the one hand, we find compounds composed of color and qualifying adjectives, whose first member is typically an *I*-ending stem (26a). On the other hand, we have two-word compounds made up of relational adjectives—those referring to nationalities are quite common—whose first member is typically an *O*-stem (26b).

26.   a.   blanqu-i-celeste   gord-i-buen-a
               white-I-sky   blue fat-I-good-F
               'blue-white'   'curvy woman'
     b.   (frontera) grec-o-turc-a   (periodo) hispán-ic-o-colonial
               (border) Greek-O-Turkish-F   (period) Hispan-IC-M-colonial
               'Greek-Turkish border'   'Colonial-Spanish period'

Let us start with the former class. The more frequent alternations are found in color terms. The default for paroxytone disyllabic color terms such as *rojo* 'red' and *verde* 'green' is to be spelled-out with *I*-stems (27a), although they can alternate with a spell-out as non-inflecting words (27b). They cannot alternate with a spell-out as *O*- stems (27c):[23]

27.   a.   *guald-i-negro*   *verd-i-blanco*
                  gold-I-black   green-I-white
                  'gold and black   'Green and white'
     b.   *verd-i-amarilla/verde-amarilla* 'green (and) yellow'
                  blanqu-i-celeste/blanco-celeste 'with (and) sky blue
     c.   *gualdonegro,   *verdoamarilla
                  *gold+o+black,   *Green+o+yellow

Spell-out with *I*-stems is ruled out for oxytone color terms like *azul* 'blue' or *marrón* 'brown', and for polysyllabic ones (*amarillo* 'yellow'), including morphologically-derived color terms (28):

28.   *azulcrema* (*azul-i-crema) 'blue and cream'
        pardogrisáceo (*pard-i-grisáceo) 'dun and greyish'
        verd-os-o-blanqu-ecin-o (*verdos-i-blanquecino) 'greenish and milky'
        roj-iz-o-anaranjado (*rojic-i-anaranjado) 'reddish orange'

The appearance of *I*-ending stems is then restricted from a prosodic point of view; they favor the construction of unmarked Spanish syllables or dysyllabic feet (CV.CV) (Fábregas 2004). However, there are other differences between the classes of adjectival compounds of (26a) and (26b).

The main difference is that compounds of (26b) are not necessarily interpreted as coordinated structures. In the compounds of (29), one of the members of the compound can be interpreted as modifying the other:[24]

29.   *intervenciones quirúrgico-estéticas* ([aesthethics surgery] operations)
        placa dento-bacteriana (bacterian plaque on the teeth)
        identidad celto-gaélica (Celtic identity of Welsh people)

The second relevant difference is that *O*-stems, contrary to *I*-stems, can replace morphologically-derived adjectives (30), and thus they do not seem to be prosodically conditioned. In fact, there is some discussion as to whether the *O*-stems of these compounds receive a primary or a secondary stress (Val 1999, p. 4812), while *I*-stems are clearly unstressed.

30.   democrat-ic-o-liberal-es
        democracy-IC-M-liberal-PL
        'liberal-democratic'
        obsess-iv-o-compuls-iv-a
        obsession-IV-M-compulsive-F
        'obsessive-compulsive'

The third difference concerns the morphological variants of *O*-ending stems. While I-ending stems alternate with non-inflectable Spanish words (27b), *O*-stems alternate not only with these (31), but also with adjectives inflecting for number and gender (32):

31.   (fedederación) croat-o-musulman-a/croata-musulman-a
        (federation) croatian-O-muslim-F/croatian-muslim-F
        guerras árab-o-isrraeli-es/(conflicto) árabe-israelí
        (wars) Arab-O-Israeli-PL/(conflict) Arab-Israeli
        (literatura) jude-o-árabe (problema) judí-o-marroquí
        (literatura) jude-O-Arabic/(problem) Jewish-M-Moroccan
        dem-o-cristian-o-s/demócrata-cristian-o-s
        dem-O-Christian-M-PL/democratic-Christian-M-PL

32.   (frontera) grec-o-turc-a/grieg-**a**-turc-a
      (Frontier) Greek-O-Turkish-F/Greek-F-Turkish-F
      (cultura) hispan-o-visigod-a/hispan-**a**-visigod-a
      (culture) hispanic-O-Visigothic-F/Hispanic-F-Visigothic.F
      comunidad jude-o-estadounidense/representante-s judí-**o-s**-estadounidense-s
      (community) jude-O-American/(representatives-PL) Jewish-M-PL-American.PL
      trastorno-s obsesiv-**o-s** compulsiv-o-s
      (disorders-PL) obsessive-PL-compulsive-M-PL

O-ending stems and morphosyntactic words exhibit clearly opposing properties. The former never project number and gender (*judeos vs. judíos). According to our analysis, number and gender allomorphs can never appear in right-headed nominal compounds, but (genderless and numberless) stems can (33):

33.   *afr-O-descendencia* (**afric-an-o descendencia*)
     afr-O-offspring (*African-M-offspring)
     'African offspring'
     hispan-o-parlante (**hispanic-o-parlante*)
     hispanic-O-speaking (Hispanic-M-speaking)
     'Spanish-speaking'
     angl-o-fobia (**ingles-es-fobia*)
     angl-O-phobia (*English-PL-phobia)
     'Anglophobia'

In Figures 5 and 6, a model of lexical entry for each of the adjectival compound structures is provided. To the left, the acceptable and unacceptable spell-outs are provided; in the center, the functional structure assigned to these compounds is shown, focusing on the context for lexical insertion of non-head allomorphs; to the right, both the structural semantics and the conceptual meaning of the compound paroxytone disyllabic color are shown.

Figure 5 shows that the internal adjective of the compound is the projection of a categorial head Adj$_1$ (AdjP), with neither further development of its functional structure (e.g., *Degree*) nor agreement potential. The merge with the head of the compound is mediated by a functional head, *Coordination*, which restricts the interpretation of the grammatical relationship between the members of the compound to a coordinating one. I assume that coordination conjunctions are projecting heads that take one of the coordinated members as their complement and the other as their specifier (Progrovac 1998). Normally, the Coordinated Phrase will inherit the categorial features of its arguments, but exocentricity is common in coordinate compounds (Scalise et al. 2009).

With regard to the spell-out of the compound constituent to the left, the evidence that Spanish compounds have specific *I*-allomorphs linked to coordinating conjunctions surpass this class of compounds, in that these allomorphs also appear in compound numbers (*veint-i-dos* 'twenty+AND+two'). After the conflation process of Rel and Adj$_1$, the structure can be spelled-out with that more specific allomorph (following the *Panini Condition*), provided that the aforementioned prosodic conditions are fulfilled. The external adjective (Adj$_2$) will act as the morphological head of the compound (*camisetas verdiblanc-a-s* 't-shirts-F.PL. green-I-white-F-PL' 'green and white t-shirts').

Figure 6 shows a rather different structure. As noted above, the compound is not necessarily understood as coordinated, so it is not possible to postulate a CoordP. Hence, I once more assume the basic form of Relation (Rel), so the interpretation of the relationship would depend on pragmatic and encyclopedic knowledge. By postulating Rel, we also capture the essence of why the default spell-out of the conflated Rel > Adj constituent is a *O*-ending stem (*greco*), as was the case for right-headed N+N compounds (*grecoparlante* 'Greek-speaking'), which contain Rel and are spelled-out with the same stem class.
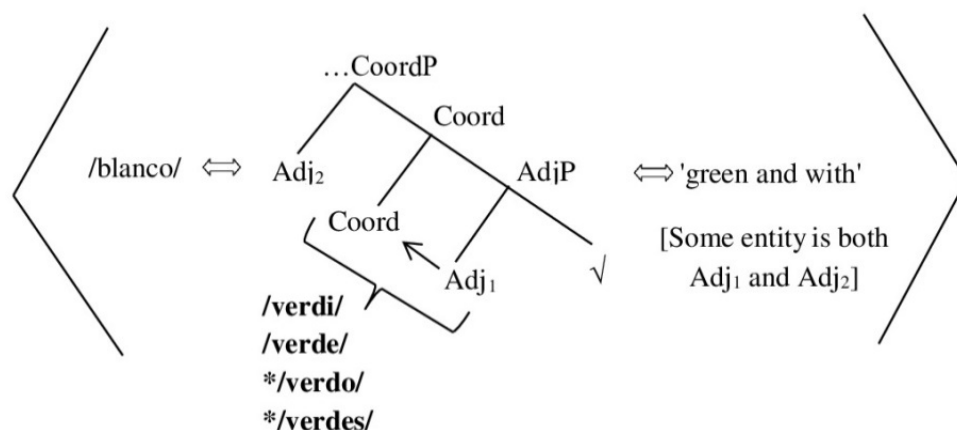
Figure 5. Lexical entry for a coordinated Adj+Adj compound.
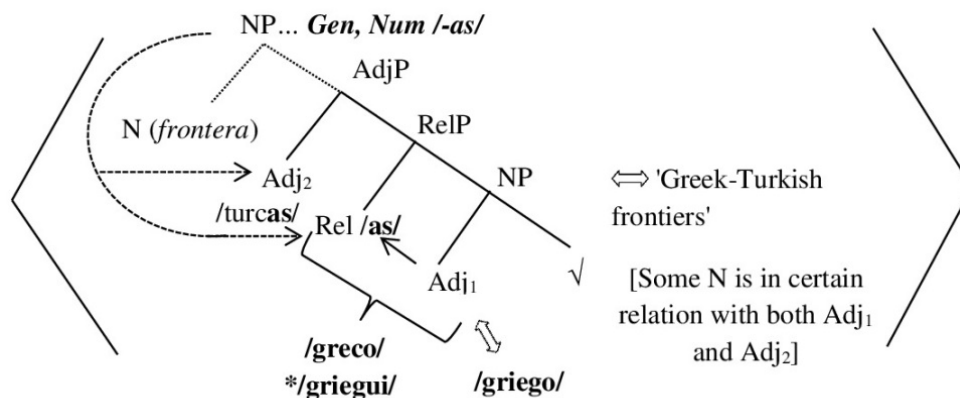

Figure 6. Lexical entry for juxtaposed Adj+Adj compounds.

However, when Adj_1 is spelled-out with a morphosyntactic adjective (*griego*), I propose that the *Relation* head is spelled-out by the agreement morphemes determined by the NP to which the compound modifies (*fronteras*). While agreement is obligatory for the external adjective (*turcas*), it is optional for the internal one, even when there is a morphosyntactic word like *griego* (not *greco*) spelling-out the Adj_1 node (*fronteras griego-turcas* is also fine). This never happens with *greco* because, as my approach predicts, *O*-ending stems spell-out *Rel*, so *Rel* is no longer available for spell out. A strong argument against the treatment of *i* and *o* as linking vowels therefore emerges: *O*- and *I*-ending stems are in complementary distribution with agreement because they are, just like agreement, spell-outs of grammatical relations.

Thus, I believe that we can take constructions such as *griegas-turcas* to be phrasal structures, considering that *griegas* seems to enter in an agreement relationship with either *fronteras* or *turcas*. However, this agreement is "rare" in the sense that it is not obligatory, and I therefore suggest that what we find is similar to the (also optional) spell-out of *Identity* in NN left-headed structures such as *pintores poetas*: the morphological marking of the non-head serves as a means of stressing the fact that *griego* and *turco* share their main features (in this case, they both refer to countries). I predict, then, that it is more unlikely that the internal adjective would be inflected when its features are distinct from those of the external one, and, intuitively, **placas dentales-bacterianas* (dental+bacterian plaques) is ill-formed in Spanish for that reason.

5. Conclusions

In this paper, I have explored the connections between the syntactico-semantic structures of different classes of productive Spanish compound patterns and their morphological spell-out. I have argued for some unprecedented correspondences between structures and

morphological forms: the appearance of O-stems where the relations between the members are less structurally determined, as opposed to I-stems; the fact that only the former serve as a replacement for adjectival agreement; the formalization of both interpretable and non-interpretable number inflection in N+N and V+N compounds. The analysis with corpora has revealed that the morphological "instability" of compounds, that is, the existence of more than one form of the same compound, is neither an accidental phenomenon nor is it restricted to specific classes of compounds. In future research, corpus-based quantitative analyses might usefully provide more detailed insight into the variants addressed in the present study. Additionally, the examination of the functional structures and their available spell-outs in other Romance languages could reveal stronger evidence (or counterevidence) for my claims, which has been applied here only to Spanish.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

Notes

1   See Montermini (2010) for an examination of the variety of morphological units within compounds in a larger sample of languages, including German, Russian, Welsh and Greek.

2   The presence of optional number inflection within compounds has been observed in other languages (Booij 2005; Villoing 2012).

3   Note that plural in the non-head position of VN compounds is the default, contrary to the case in English, where it is typically not permitted (*flies-swatter*). On the other hand, phrasal non-heads, which are rare in Spanish, appear frequently in English. In any case, compounding with inflected nouns and with phrasal constituents are both problematic for those (weak) lexicalist approaches that assume that morphological rules operate with non-inflected lexical units (Punske 2016).

4   The lack of semantic repercussions in this choice constitutes empirical evidence against proposals that link Spanish "root-compounding" (1) with semantic arbitrariness as opposed to "word-compounding" (2) with compositionality (Ntelitheos and Pertsova 2019).

5   It is possible to consider the examples in (7c) as phrases, due to the presence of fully-inflected words in the non-head position. In any case, the possibility of constructing the same grammatical relationship with either clear morphological words (7a-b) and phrasal-like compounds (7c) with no difference of meaning constitutes evidence against models of competition between word and phrasal syntax (Ackema and Neeleman 2004).

6   Contrary to Botha (1981, p. 18) and Carstairs-McCarthy (2002), phrasal modifiers within compounds in (2) are neither "quotes" nor relexicalized idioms. See Ackema and Neeleman (2004, p. 126), and Sato (2010), for compositional accounts of phrasal compounding.

7   This is also the case with Adj+N compounds such as *altavoz* lit. 'high+voice' 'speaker', and *medianoche* 'midnight', which are very rare, and occasional univerbations from NPs such as *guardiacivil* lit. guard+civil, 'civilguard', and *telaraña* lit. web+spider' 'spiderweb', resulting from the dropping of a preposition (Rainer and Varela 1992).

8   Many of these right-headed N+N constructions alternate with derived adjectives: *Oasismanía > Oasismaniaco* (Oasismania > maniac), *Juvedependiente > Juvedependencia* (Juvedependent > dependency), *adultocentrismo > adultocéntrico* (adultcentrism > centric).

9   I assume a form of the so-called set-merge, and hence there is an inherent assimetry in the combination (one of the merged members would project as the head). Merge can combine two possible complex syntactic objects.

10 Syntactic projection is a common term in all generative frameworks to refer to the result of combining two lexical items/features into one object that inherits the properties of the head. Given that nanosyntax postulates functional sequences of features that are a priori always present, it is necessary to assume that the projection of structures can be *truncated* at a certain layer, preventing the higher functional categories for projecting (Haegeman 2003).

11 Different versions of the Lexicalist Hypothesis exist: some place limits on compound recursivity, and others prohibit the modifiability of individual compound members (see Lieber and Scalise 2006 for a review). The validity of these predictions depends crucially on whether or not the more phrasal-like combinations of nouns are considered to be compounds (Villoing 2012), yet typically multiword expressions comply with lexicalist predictions despite being clearly phrasal: if their structure is altered, they lose their idiomatic meaning. The predictions of the Lexical Integrity Hypothesis also depend on the language

examined: for compounds in Classical Sanskrit, for example, most of the lexicalist predictions do not hold (Lowe 2015); see also counterevidence in Georgian (Harris 2006). I take the position that, while lexicalist predictions prove to be empirically adequate most of the time in a language like Spanish, they fail to establish a clear-cut distinction between complex words and phrases, and thus I consider that they should be seen as the *result* of the characteristics of word-like syntactic configurations (e.g., structures lacking certain functional features as a consequence of an early spell-out).

[12] One can see compounds as the products of transformations on syntactic structure, as in the earlier models, or as the consequence of incorporation (a case-driven phenomena in more recent approaches, such as Harley 2009, or Moyna 2011 for Spanish compounds). I opt for the (theoretically simpler) assumption that compounds can be base-generated. See Punske (2016) for arguments against incorporation approaches to compound formation.

[13] There are many proposals as to the exact features which make up the functional projections. See Abney (1987) and Ritter (1991) for nouns (*D, Q* and *Number* are indisputable; see Picallo 2008 for *Gender* in Romance Languages), Cinque (1999) for the areas of Tense, Aspect, and Mood, and Cinque (2010) for Adverbs and Adjetives, Caha (2009) and Svenonius (2010) for Case and Prepositions, Ramchand (2008) for the event structure of Verbs, and Rizzi (1997) for the left-periphery or informational structure.

[14] Even if the MD lexicon is more economical, and the same category-less root can be merged in many syntactic contexts, it has the drawback of learnability: it is far from obvious that speakers (at least in a language like Spanish) learn/store lexical entries linked to specific categories, and not category-less units. Roots are only learnt as an abstraction from a conjunct of related words, and it is difficult to see when words share the same root or not, putting diachronic considerations aside (since etymological links are not a part of the natural knowledge of the language for a normal speaker).

[15] Following Von Heusinger and Schwarze (2013), novel compounds would be compositional structures that, once stored, become lexicalized, restricting and changing the meaning predicted by the structure, and even losing its internal structure (in which case they will be accessed as simple words). This is easily observed for VN compounds: at first, *limpiabotas* lit. clean-boots, 'shoe shiner' can refer to any individual or artifact able to perform the depicted action, but it becomes lexicalized so as to refer to only one of the possible referents, and even to acquire a derogatory meaning.

[16] Other principles are *Cyclic override*, which predicts that previous lexicalizations can be overriden by later, more precise ones, and the *ABA theorem, which restricts phrasal spell-out to adjacent features (Bobaljik 2007)

[17] Some of these modifiers, such as *cripto-* (*criptomoneda* 'cryptocurrency'), *Euro-* (*Eurobanco* 'Eurobank'), *video-* (*videotexto* 'videotext'), *narco-* (*narcoedificio* 'drug flat') and, recently, *corona* (*coronadivorcio* 'Corona divorce'), appear in so many words that they may even be experiencing a grammaticalization process towards prefixation.

[18] If one stem appears in many words or in high-frequency ones, it would have a greater potential for attracting analogical innovations, so it would have a higher token frequency. But family size can clearly vary among speakers, in that they do not necessarily have identical lexicons.

[19] I- and O-ending stems in neoclassical-inspired compounds do alternate (*agricomercio/agrocomercio* 'agrotrade'). Notice that one must not confuse Spanish words ending in an O theme vowel, such as *pelo* 'hair', with the O-ending stems for nouns, which substitute theme vowels (*pierna → piern-i/*piern-o* 'leg').

[20] In English, for productive inalienable possession compounds, one can identify the (word-like) spell-out of a (modified) N/*(blue)-eye*/and the spell-out of *Poss* with a suffix/*-ed*/.

[21] In Delfitto and Melloni (2009, p. 93-ss.) the exact semantic features are retrieved from the Quale structure of the nouns.

[22] Although lexicographical sources document more "only-singular" compounds, my results are based on neological and contemporary low token frequency compounds.

[23] We should not confuse Spanish words ending in a *o* vowel theme such as *rojo* with the o-ending stems for nouns, which can substitute other theme vowels (*croata → croato*).

[24] I accept that compounds with color terms are coordinating, even if it is true that one of the compounded colors can be the prevailing one, while the other can shade the former (Padrosa 2011).

References

Abney, Stephen. 1987. The English Noun Phrase in its Sentential Aspect. Ph.D. dissertation, MIT, Cambridge, MA, USA.

Ackema, Peter, and Ad Neeleman. 2004. *Beyond Morphology. Interface Conditions on Word Formation*. Oxford: Oxford University Press. [CrossRef]

Alvar, Manuel. 1984. De nuevo sobre los compuestos de verbo más sustantivo. In *Actas del II Simposio Internacional de Lengua Española*. Edited by Manuel Alvar Exmo. Gran Canaria: Cabildo Insular de Gran Canaria, pp. 83–97.

Bartoš, Ludovic. 1999. En torno a las formaciones binominales yuxtapuestas en español. *Études Romanes de Brno* 29: 49–59.

Bermúdez-Otero, Ricardo. 2013. The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features. *Probus* 25: 3–103. [CrossRef]

Bobaljik, Jonathan. 2007. Auspicious compounds. *McGill Working Papers in Linguistics* 17: 65–71.

Booij, Geert. 2005. Compounding and derivation: Evidence for Construction Morphology. In *Morphology and Its Demarcations*. Edited by Wolfgang Dressler, Dieter Kastovsky, Oskar Pfeiffer and Franz Rainer. Philadelphia: John Benjamins, pp. 109–32. [CrossRef]

Botha, Rudolph. 1981. *Morphological Mechanisms, Lexicalist Analyses of Synthetic Compounding*. New York: Pergamon Press.

Buenafuentes, Cristina. 2014. The analysis of inflection in Spanish compounds. *Borealis: An International Journal of Hispanic Linguistics* 3: 1–21. [CrossRef]

Buenafuentes, Cristina. 2020. Main compounding types in Spanish: Synchronic issues. In *The Routledge Handbook of Spanish Morphology*. Edited by Antonio Fábregas, Víctor Acedo-Matellán, Grant Armstrong, María Cristina Cuervo and Isabel Pujol Payet. Oxford: Routledge, pp. 285–302.

Bustos-Gisbert, Eugenio. 1986. *La composición Nominal en Español*. Salamanca: Universidad de Salamanca.

Caha, Pavol. 2009. The Nanosyntax of Case. Ph.D. dissertation, University of Tromsø, Tromsø, Norway.

Carstairs-McCarthy, Andrew. 2002. *Current Morphology*. Oxford: Routledge. [CrossRef]

Cinque, Gugielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Cinque, Gugielmo. 2010. *The Syntax of Adjectives: A Comparative Study*. Cambridge: MIT Press. [CrossRef]

Delfitto, Denis, and Chiara Melloni. 2009. Compounds don't come easy. *Lingue e Linguaggio* 8: 75–104.

Di Sciullo, Anna Maria. 2005. *Asymmetry in Morphology*. Cambridge: Cambridge University Press. [CrossRef]

DiSciullo, Anna Maria, and Edwin Williams. 1987. *On the Definition of Word*. Cambridge: MIT Press.

Downing, Pamela. 1977. On the creation and use of English compound nouns. *Language* 53: 810–42. [CrossRef]

Fábregas, Antonio. 2004. Prosodic constraints and the difference between root and word compounding. *Lingue e Linguaggio* 2: 303–39.

Fábregas, Antonio. 2005. La Definición de la Categoría Gramatical en una Morfología Orientada Sintácticamente. Nombres y Adjetivos. Ph.D. dissertation, Universidad Autónoma de Madrid, Madrid, Spain.

Fábregas, Antonio. 2007. An Exhaustive Lexicalisation Account of Directional Complements. *Nordlyd* 34: 165–99. [CrossRef]

Fábregas, Antonio. 2016. *Las Nominalizaciones*. Madrid: Visor.

Felíu, Elena. 2016. En torno a los compuestos coordinativos en español: El caso de "relación madre-hija". *Lingüística Española Actual* 38: 291–313.

Fradin, Bernard. 2009. Compounding in French. In *Oxford Handbook of Compounding*. Edited by Rochelle Lieber and Pavol Štekauer. Oxford: Oxford University Press, pp. 417–435.

Franco, Ludovico. 2015. VN compounds in Italian and some other Romance languages, phrasal spell-out and rebooting. *Rivista di Grammatica Generativa* 37: 81–98.

Garcia-Page, Mariano. 2011. Hombre clave, hombre rana, ¿Un mismo fenómeno? *Verba* 38: 127–70.

Gärtner, Hans M. 2002. *Generalized Transformations and Beyond. Reflections on Minimalist Syntax*. Berlin: Akademie-Verlag. [CrossRef]

Gil Laforga, Irene. 2014. La interacción de los Componentes Gramaticales en la Formación de Palabras: Adjetivos Posesivos Derivados y Compuestos. Ph.D. dissertation, Universidad Autónoma de Madrid, Madrid, Spain.

Grossmann, Maria, and Franz Rainer. 2009. Italian adjective-adjective compounds: Between morphology and syntax. *Rivista di Linguistica* 21: 71–96.

Guéron, Jacqueline. 1992. Inalienable Possession and Locative Aspect. *Anuario del Seminario de Filología Vasca Julio Urquijo* 27: 135–46.

Guevara, Emiliano, and Sergio Scalise. 2009. Searching for universals in compounding. In *Universals in Language Today*. Edited by Sergio Scalise, Elisabetta Magni and Antonietta Bisetto. Berlin: Springer, pp. 101–28. [CrossRef]

Haegeman, Liliane. 2003. Conditional clauses: External and internal syntax. *Mind & Language* 18: 317–39.

Harley, Heidi. 2009. Compounding in Distributed Morphology. In *The Oxford Handbook of Compounding*. Edited by Rochelle Lieber and Pavol Štekauer. Oxford: Oxford University Press, pp. 204–83.

Harris, Alice. 2006. Revisiting Anaphoric Islands. *Language* 82: 114–30. [CrossRef]

Hualde, Juan Ignacio. 2007. Stress removal and stress addition in Spanish. *Journal of Portuguese Linguistics* 6: 59–89. [CrossRef]

Jackendoff, Ray. 2009. Compounding in the Parallel Architecture and Conceptual Semantics. In *The Oxford Handbook of Compounding*. Edited by Rochelle Lieber and Pavol Štekauer. Oxford: Oxford University Press, pp. 162–203.

Kornfeld, Malena. 2004. La Formación de Palabras en la Sintaxis desde la Perspectiva de la Morfología Distribuida. Ph.D. dissertation, UBA, Buenos Aires, Argentina.

Lees, Robert. 1960. *The Grammar of English Nominalizations*. The Hague: Mouton.

Lieber, Rochelle. 1992. *Deconstructing Morphology. Word formation in Syntactic Theory*. Chicago: Chicago University Press.

Lieber, Rochelle. 2009. A Lexical Semantic approach to Compounding. In *The Oxford Handbook of Compounding*. Edited by Rochelle Lieber and Pavol Štekauer. Oxford: Oxford University Press, pp. 78–104.

Lieber, Rochelle, and Pavol Štekauer. 2009. Introduction: Status and definition of compounding. In *The Oxford Handbook of Compounding*. Edited by Rochelle Lieber and Pavol Štekauer. Oxford: Oxford University Press, pp. 1–24.

Lieber, Rochelle, and Sergio Scalise. 2006. The Lexical Integrity Hypothesis in a new theoretical universe. *Lingue e Linguaggio* 6: 7–32.

Lowe, John. 2015. The syntax of Sanskrit compounds. *Language* 91: 71–115. [CrossRef]

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

Marqueta, Bárbara. 2018. Una propuesta de categorización sin morfemas para el compuesto lavaplatos. *Estudios de Lingüística de la Universidad de Alicante* 32: 219–36. [CrossRef]

Marqueta, Bárbara. 2019a. The syntactic structure of pelirrojo compounds. *Glossa: A Journal of General Linguistics* 4: 1–23. [CrossRef]

Marqueta, Bárbara. 2019b. La Composición, la Arquitectura del Léxico, y la Sintaxis de las Palabras con Estructura Interna en Español. Ph.D. dissertation, Universidad de Zaragoza, Zaragoza, Spain.

Marqueta, Bárbara. 2020. ¿De qué están hechos los compuestos?: La interfaz entre constituyentes sintácticos y exponentes léxicos del español. *RILCE* 35: 575–602. [CrossRef]

Mendívil, José Luis. 2019. If everything is syntax, why are words so important? An a-morphous but non-lexicalist approach. *Linguistics* 57: 1161–215. [CrossRef]

Montermini, Fabio. 2010. Units in Compounding. In *Cross-Disciplinary Issues in Compounding*. Edited by Sergio Scalise and Irene Vogel. Amsterdam: John Benjamins, pp. 77–92. [CrossRef]

Moyna, Maria I. 2011. *Compound Words in Spanish. Theory and History*. Amsterdam: John Benjamins. [CrossRef]

Moyna, Maria I. 2020. The diachrony of Spanish compounding. In *The Routledge Handbook of Spanish Morphology*. Edited by Antonio Fábregas, Víctor Acedo-Matellán, Grant Armstrong, María Cristina Cuervo and Isabel Pujol Payet. London: Routledge, pp. 303–16. [CrossRef]

Ntelitheos, Dimitrios, and Katya Pertsova. 2019. Root and semi-phrasal compounds: A syntactic approach. *Proceedings Lingistic Society America* 30: 1–14. [CrossRef]

Padrosa, Susana. 2011. Complex Word-Formation and the Morphology-Syntax Interface. Ph.D. dissertation, Universitat Autònoma de Barcelona, Barcelona, Spain.

Picallo, Carmen. 2008. Gender and number in Romance. *Lingue e Linguaggio* 7: 47–66.

Progrovac, Ljiljana. 1998. Structure for Coordination. *Glot International* 3: 3–9.

Punske, Jeffrey. 2016. Compounding. *Language and Linguistics Compass* 10: 382–93. [CrossRef]

Rainer, Franz. 1993. *Spanische Wortbildungslehre*. Tübingen: Niemeyer. [CrossRef]

Rainer, Franz, and Soledad Varela. 1992. Compounding in Spanish. *Rivista di Linguistica* 4: 117–42.

Ralli, Angela. 2013. *Compounding in Modern Greek*. New York: Springer. [CrossRef]

Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon. A First Phase Syntax*. Cambridge: Cambridge University Press. [CrossRef]

Ricca, Davide. 2010. Corpus data and theoretical implications: With special reference to Italian VN compounds. In *Cross-Disciplinary Issues in Compounding*. Edited by Sergio Scalise and Irene Vogel. Amsterdam: John Benjamins, pp. 237–54. [CrossRef]

Rio-Torto, Graça, and Silvia Ribeiro. 2009. Compounding in contemporary Portuguese. *Probus* 24: 119–45. [CrossRef]

Ritter, Elizabeth. 1991. Cross-linguistic evidence for number phrase. *Canadian Journal of Linguistics* 37: 197–218. [CrossRef]

Rizzi, Luiggi. 1997. The fine structure of the left periphery. In *Elements of Grammar*. Edited by Lilianee Haegeman. Dordrecht: Kluwer, pp. 281–337. [CrossRef]

Sato, Yosuke. 2010. Complex phrase structures within morphological words: Evidence from English and Indonesian. *Lingua* 120: 379–407. [CrossRef]

Scalise, Sergio, and Irene Vogel. 2010. Why Compounding? In *Cross-Disciplinary Issues in Compounding*. Edited by Sergio Scalise and Irene Vogel. Amsterdam: John Benjamins, pp. 1–18. [CrossRef]

Scalise, Sergio, Antonio Fábregas, and Francesca Forza. 2009. Exocentricity in Compounding. *Gengo Kenkyū* 135: 49–84.

Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36: 1–6.

Svenonius, Peter. 2010. Spatial p in English. In *Mapping Spatial PPs: The Cartography of Syntactic Structures*. Edited by Guglielmo Cinque and Luigi Rizzi. Oxford: Oxford Scholarship Online, pp. 127–60. [CrossRef]

Val, José Francisco. 1999. La composición. In *Gramática Descriptiva de la Lengua Española*. Edited by Ignacio Bosque and Violeta Demonte. Madrid: Espasa-Calpe, pp. 4757–841.

Van Jaarsveld, Henk J., Riet Coolen, and Robert Schreuder. 1994. The role of analogy in the interpretation of novel compounds. *Journal of Psycholinguistic Research* 23: 111–37. [CrossRef]

Varela, Soledad. 1990. Composición nominal y estructura temática. *Revista Española de Lingüística* 20: 55–81.

Villoing, Florence. 2012. French compounds. *Probus* 24: 29–60. [CrossRef]

Von Heusinger, Klaus, and Christoph Schwarze. 2013. Italian V+N compounds, inflectional features and conceptual structure. *Morphology* 23: 325–50. [CrossRef]