Root and semi-phrasal compounds: A syntactic approach

Dimitrios Ntelitheos & Katya Pertsova[*]

Abstract. We present a syntactic account of the derivation of two types of attributive nominal compounds in Spanish, Russian and Greek. These include right-headed "root" compounds, which exhibit more "word"-like properties and single stress domains, and left-headed "semi-phrasal" compounds with more phrasal properties and independent stress domains for the two compound members. We propose that both compound structures are formed on a small clause predicate phrase, with their different properties derived from the merger of the predicate member of the small clause as a root or as a larger nominal unit with additional functional projections. The proposed structures provide an explanation of observed lexical integrity effects, as well as specific predictions of patterns of compound formation crosslinguistically.

Keywords. morpho-syntax; compounds; Distributed Morphology; lexical integrity; predicate inversion; small clauses

1. Introduction. A hotly debated issue is to what extent morphological principles are independent from the syntactic ones. There is no agreement on how and why differences between "words," "phrases," and other units proposed to exist in-between these two categories arise. Syntactic approaches to word-formation such as Distributed Morphology (Halle & Marantz 1993, Marantz 1997 and related work), antisymmetry (Kayne 1994, Koopman and Szabolcsi 2000) and more recently nanosyntax (Starke 2009) derive morphosyntactic properties of different units from the derivational path of their formation. For instance, Marantz (2001) proposes a difference between units created via combinations of functional heads with roots vs. those created at higher levels. In this work, we show additional evidence for this distinction from compound structure. This evidence comes from two distinct types of attributive compounds in Spanish, Russian, and Greek that show several common asymmetries. We will predict these asymmetries in a syntactic analysis in which a distinction is made between compounds that involve merging of roots vs. those that involve merging of larger structures.

We adopt a modified Distributed Morphology/antisymmetric framework and propose that the specific compounds discussed in this paper are relative clauses with an internal small clause structure realized as a Relator Phrase (DenDikken 2006). The semantic and formal head of a compound is the subject of the small clause, and the predicate member can merge either as a root or a larger structure (e.g., nP, numP). These two types of options create two types of compounds with distinct properties summarized in Section 2. Thus, compound-distinctive properties, as well as their differences from full syntactic phrases are derived from basic assumptions about syntactic structure and operations which have been developed to account for "purely" syntactic phenomena (phrasal movement, predicate inversion, licensing, quantization, and so on). We also discuss how this account predicts the existence of lexical integrity effects which are usually taken to support a distinction between words and phrases. On our account existence of such effects is

---

straightforwardly derived from the sizes and the functional content of the syntactic structures involved, with the implementation of independently motivated syntactic rules.

The paper is organized as follows. In Section 2 we present the data, discussing the properties of the two types of compounds under investigation. Section 3 spells-out the details of the small-clause predicate structure which, we propose, forms the base for the derivation of both types of compounds. We also discuss how the proposed structures provide explanations for the different morphosyntactic and phonological properties of the two compounds, while Section 4 derives the differences between compounds and full syntactic phrases from the proposed derivational paths. Finally, Section 5, presents our conclusive remarks.

2. Properties of root- vs. semi-phrasal compounds in Spanish, Russian, and Greek. We delimit the scope of our inquiry to compounds that according to a classification introduced in Bisetto & Scalise (2005) show an attributive relation between the head and the non-head. Bisetto & Scalise classify all compounds into subordinate, attributive and coordinative (regardless of whether they are endo- or exocentric). Attributive compounds typically have a nominal head with a nominal or adjectival non-head that ascribes a property to the head. Some examples from English include *blue cheese, fruit salad, poster man.*

The notion "head" in compounding is typically defined based on both semantic and morpho-syntactic criteria. The latter include the definition of head as an element that determines the morpho-syntactic category of the compound:

(1) a. [X r Y]$_Y$ - right-headed
     b. [X r Y]$_X$ - left-headed

The above structures (from Guevara & Scalise 2008) assume that X and Y are lexical categories and 'r' is a type of grammatical relation between the two constituents. Later work within the DM framework assumes that X and Y in the structure above may not be categories but acategorial roots, with the category fixed at a higher functional level, with the addition of category-defining heads (n, a, v, and so on). For attributive NN compounds the head can be distinguished from the non-head as the element that carries the syntactically determined inflectional features (e.g., case) and/or that determines the inherent inflectional features of the compound (e.g., gender, class).

Many languages have compounds ranging on the spectrum from more phrase-like to more word-like. Here we focus on two such types of attributive NN compounds in three languages, Spanish, Russian, and Greek. We call them *root compounds* (RC) and *semi-phrasal* compounds (SC). A root compound is a compound in which at least one of the combining elements is a bare root that cannot host inflectional or derivational morphology and is often connected to the other element through a meaningless phonological element called "linking element" (abbreviated as LE henceforth[1]). Table 1 below summarizes the main differences between root- and semi-phrasal compounds in the three languages under investigation.

Root-compounds are typically head-final, form a single prosodic unit for stress assignment, and have inflectional morphology at the right-edge. In addition, they often have linking elements and tend to have idiosyncratic semantics, sometimes forming exocentric compounds. Some examples from Spanish, Russian and Greek are provided in (2)-(4). In all of these examples the head can be determined from the gender of the compound given that the two constituent elements have a gender mismatch.

---

[1] A linking element is also sometimes called an interfix or intermorph in the literature (see Scalise & Bisetto, 2009).

Root N+N compounds (RC)	Semi-phrasal N+N compounds (SC)
Head: on the right	Head: on the left
Inflectional elements: on the right edge	Inflectional elements: possible on both members
Single phonological word	Two phonological words
Semantics: tends to be idiosyncratic	Semantics: tends to be transparent
Linking element present	Linking elements absent

Table 1. Properties of root- vs. semi-phrasal attributive compounds in Spanish, Russian, and Greek

(2) a. carr-i-**cuba**   SPANISH (from Moyna 2011)
         car(MASC)-LE-barrel(FEM)
         "water cart" FEM
     b. man-i-**obra**
         hand(MASC/FEM)-LE-work(FEM)
         "maneuver" FEM
(3) a. obezjan-o-**chelovek**   RUSSIAN
         monkey(FEM)-man(MASC)
         "ape man" MASC
     b. neb-o-**sklon**
         sky(NEUT)-LE-slope(MASC)
         "skies" MASC
(4) a. arahn-o-**anthropos**   GREEK
         spider(FEM)-LE-man(MASC)
         "spiderman" MASC
     b. hart-o-**sakoula**
         paper(NEUT)-LE-bag(FEM)
         "paper-bag"   FEM

Semi-phrasal compounds on the other hand tend to be left-headed, with both head and non-head carrying gender and number morphology, and each compound member defining its own prosodic domain, carrying a main stress. Semi-phrasal compounds can also carry idiomatic readings, but to a lesser degree than root compounds. Finally, there is no linking element connecting the compound head to the non-head. Examples are provided in (5)-(8):

(5) a. **coche** cama   SPANISH (from Moyna 2011)
         car(MASC) bed(FEM)
         "sleeper car" MASC
     b. **piedra** mármol
         stone(FEM) marble(MASC)
         "marble" FEM
(6) a. **jubka** karandash   RUSSIAN
         skirt(FEM) pencil(MASC)
         "pencil skirt" FEM
     b. **jazyk** osnova
         language(MASC) base(FEM)
         "protolanguage" MASC

(7) a. **anthropos** arahni   GREEK
        man(MASC) spider(FEM)
        "spiderman" MASC
(8) b. **taxidi** astrapi
        trip(NEUT) lightning(FEM)
        "fast/sudden trip" NEUT

2.1. THE LOCUS OF INFLECTIONAL MORPHOLOGY. In Spanish semi-phrasal compounds the plural marking can show up just on the head (most common), or on both the head and the right-edge, and more marginally, just on the right edge (Guevara, 2012).

(9) hombre-**s** lobo / hombre-**s** lobo-**s** / ?hombre lobo-**s** "werewolves"
       man-PL   wolf man-PL   wolf-PL   man   wolf-PL

In root-compounds inflection is exclusively on the right edge:

(10) man-i-obra-**s**   *manosobra**s** *manosobra "maneuvers"
        hand-LE-work-PL

In Russian semi-phrasal compounds number, case, and noun-class features required by the morphosyntactic context are typically realized just on the head (while the non-head takes the NOM.SG.form), or less commonly on both the head and non-head.

(11) a. devushk-a   v [jubk-**e**   karandash-ø].   (No case agreement)
          girl-NOM.SG.II in skirt-LOC.SG.II pencil-NOM.SG.I
          "a girl in a pencil skirt"
      b. devushk-a   v [jubk-**e**   karandash-**e**].   (Case agreement)
          girl-NOM.SG.II in skirt-LOC.SG.II pencil-LOC.SG.I
          "a girl in a pencil skirt"

The same behavior is exhibited by Greek semi-phrasal compounds:

(12) a. tis   lex-**is**   klid-i   (No case agreement)
          the-GEN.SG word-GEN.SG   key- NOM.SG
          "the keyword's"
      b. tis   lex-**is**   klid-**iou**   (Case agreement)
          the- GEN.SG word-GEN.SG   key- GEN.SG
          "the keyword's"

In root compounds all inflection is realized on the right-edge and determined by the morphological properties of the right-most element, as in in (13) from Russian and (13) from Greek:

(13) a.   nos-o-rog-**ov**   *nos-**ov**-rog   *nos-**ov**-rog-**ov** "of rhinoceroses"
            nose-LE-horn-GEN.PL.   nose-GEN.PL.-horn   nose-GEN.PL.-horn-GEN.PL
      b.   arahn-o-anthrop-on   *arahn-**on**-anthrop-i   *arahn-**on**-anthrop-**on**
            spider-LE-man-GEN.PL   spider-GEN.PL-men-NOM.PL spider-GEN.PL-men-GEN.PL

In conclusion, RCs cannot host any inflectional elements on the non-head bare-root element, while SCs typically have inflection on the head which is on the left, with the non-head optionally agreeing with the head.

2.2. PHONOLOGICAL COHESION. Root compounds behave as a single phonological unit for stress assignment and some other phonological processes, while semi-phrasal compounds do not.

4

In Spanish RCs stress on the right-most element is preserved and stress on the left-most element is deleted or becomes secondary (Hualde 2007). In addition, RCs can show nasal assimilation in the context of nasal+stop (14), while SCs maintain two main stresses and have no nasal assimilation (14):

(14)　a.　balón + pié → balo**mp**ié
　　　　　ball　　foot
　　　　　"football"
　　　b.　balón+prisonéro → balón prisonéro
　　　　　ball　　prisoner
　　　　　"prison ball"

Similarly, in Russian RCs, only stress on the right-most element is preserved as main (Roon 2006)[2]. SCs carry two main stresses. This is illustrated below in (15.a) and (15.b).

(15)　a.　nós+róg → nos-o-róg
　　　　　nose horn
　　　　　"rhinoceros"
　　　b.　kov'ór+samol'ót → kov'ór samol'ót
　　　　　carpet　airplane
　　　　　"flying carpet"

Finally, the same situation regarding stress holds for Greek compounds (Nespor & Ralli 1994, 1996, Athanasopoulou & Vogel 2014, Ralli 2013[3]) (16)-(16):

(16)　a.　aráhni+ánthropos → arahn-o-ánthropos "spiderman"
　　　　　spider　　man
　　　　　"spiderman"
　　　b.　ánthropos + aráhni → aráhni ánthropos
　　　　　man　　　　spider
　　　　　"spiderman"

2.3. SEMANTIC OPACITY/IDIOSYNCRASY. We lack good quantitative corpus data about semantic idiosyncrasy, but it seems that root compounds are typically less compositional than semi-phrasal compounds (they have more unpredictable idiosyncratic meanings, contextual allosemy, or exhibit exocentricity). Thus, in Spanish, NN root compounds are rather rare, archaic, and many are exocentric (*agua* "water" + *pié* "foot" = *aquapié* "wine made by adding water to the residue of grapes, pressed by feet"; *ajo* "garlic" + *comino* "cumin" = aj-i-comino "a sauce made with garlic and cumin"). Based on a small corpus of Spanish compounds from Moyna's (2010) we estimate that only 4% of SCs are exocentric while roughly half of RCs are exocentric.

In Russian, compounds in general tend to have transparent semantics, but root-compounds are more likely to be exocentric (*nos* "nose" + *rog* "horn" = *nos-o-rog* "rhinoceros"). On the other hand, semi-phrasal left-headed compounds almost always have a predicative, including comparative predicative, or coordinate meaning (e.g., *divan-krovat'* "sofa bed").

---

[2] Long compounds in Russian can also acquire secondary stress, not necessarily on the same syllable as the stress in the source root. Interestingly, secondary stresses are typically absent in non-compound words (Gouskova 2010).

[3] In a certain type of Greek root compounds, the resulting compound has a new agreement ending, different from the agreement of either compound member, and in those cases the stress falls on the antepenultimate syllable, following standard stress assignment rules in Greek (see Ralli 2013 for discussion). However, even in those cases there is only one main stress, and thus the root compound forms a single prosodic unit.

In Greek both SC and RC appear with transparent or idiosyncratic semantics – e.g., root compound *limn-* "lake" + *thalassa* "sea" = *limnothalassa* "lagoon" and semi-phrasal compound *nomos plesio* "framework law" are both transparent and compositional. Both types of attributive compounds are not as productive as predicate-argument synthetic N(NOM)-N(GEN) semi-phrasal compounds.

In summary, RCs behave more like "words" compared to SCs: they follow the right-hand head rule of Williams (1981), inflect on the right edge, form a single phonological word, and can have unpredictable meaning.

3. A syntactic account for RC and SC formation. Our account is couched on general assumptions within a syntactic approach to word formation in which morphemes are the lexical atoms that are inserted into syntactic representations. We also assume that syntactic structure is computed and spelled-out in chunks and that the later operations may not have access to the structures that have already been spelled out. Following Marantz (2001, 2012), Embick & Marantz (2008), we assume that the difference between derivations from bare roots (atomic elements devoid of functional material) vs. higher-level derivations involving functional projections translates to a difference between idiosyncratic and regular morphology.

Ultimately, we propose that both root and semi-phrasal attributive compounds are derived syntactically from a common internal small clause structure, but they differ from each other in the number of phases (which influences their phonological and semantic properties). Additionally, since bare roots cannot be licensed on their own, the predicate root in root compounds undergoes Predicate Inversion (with a linker showing up as a result). This explains the headedness asymmetry between RCs and SCs. Whether a particular combination of concepts will be realized as an RC or a SC in a language is governed by language-specific preferences. For example, Spanish mostly utilizes semi-phrasal compounds, while Russian prefers root compounds. In all three languages, however, there are some attributive compounds that can be realized either as semi-phrasal or root compounds (cf. examples (4) and (7) in Greek) supporting our assumption that the two types share the same underlying structure.

In our analysis of attributive compounds instantiating a predication structure, we draw on the work by Den Dikken and Singapreecha (2004), Den Dikken (2006), who propose that complex noun phrases cross-linguistically may contain functional heads/linkers, which connect subject parts to predicate parts. Den Dikken (2006) proposes that predication involves an asymmetric small clause structure, in which the relationship between the subject and the predicate is mediated by a functional head (see also Adger and Ramchand 2003). In cases of predicate inversion an additional functional head, a linker, provides the projection serving as a landing site for the inverted predicate. Thus, the attributive compound relation can be viewed as subject-predicate relation, where a predicate ascribes a property to the subject. The compound-formation process is a "naming" process, in the sense that it refers to an entity which has been ascribed a specific property by the nominal predicate (i.e., not in the predication itself). In this sense, the structure resembles a relative clause. These compounds can usually be paraphrased as relative clauses (although this is less straightforward with exocentric compounds):

(17) o   anthropos arahni/arahnoanthropos   → o anthrops (pou) ine arahni
      the man   spider/spiderman   the man   (that) is spider

Den Dikken's (2006) predicative clauses (including small clauses) are asymmetric RPs, headed by a Relator Functional head. The Relator head can materialize as any functional element that

connects a subject to the predicate (e.g., a copula, a preposition, an Infl, etc.) While RP is asymmetrical and non-directional with both "subject-predicate" and "predicate-subject" orders available, the canonical order is the familiar "subject-predicate" order seen in (18).

(18)   a. Canonical order: [RP <u>Subject</u> [R' Relator <u>Predicate</u>]]
$$\text{IP: I consider [John}_{\text{Subj}}\text{ [(to be) my best friend}_{\text{Pred}}\text{]].}$$
$$\text{DP: [an island}_{\text{Subj}}\text{ [like a jewel}_{\text{Pred}}\text{]]}$$
$$\text{French: un type}_{\text{Subj}}\text{ drôle}_{\text{Pred}}$$
a   guy   funny
   b. Inverted order (<u>linker is required</u>): Predicate LE Subject
$$\text{IP: I consider [my best friend}_{\text{Pred}}\text{ *(to be) John}_{\text{Subj}}\text{].}$$
$$\text{DP: a jewel}_{\text{Pred}}\text{ *(of) an island}_{\text{Subj}}$$
$$\text{French: un drôle}_{\text{Pred}}\text{ *(de) type}_{\text{Subj}}$$
a   funny LE guy

The inverted order in (18) is derived via the process of Predicate Inversion (Moro 1990, Den Dikken 2006). Inverted structures feature an obligatory functional element, which unlike the relator cannot be omitted. We extend this insight to compounds: the linking element in root-compounds is the result of the inversion.

3.1. COMPOUND-INTERNAL SMALL CLAUSE. The head of an attributive compound acts as a subject in a small-clause structure with the non-head acting as a predicate. We adopt Den Dikken's RP label to label the structure (similar to PredP of Bowers 1993, 2008, Koster 1994). The functional head 'Relator' realizes the implicit relationship between the two parts of the compound predicate and can be phonologically null. This structure is similar to the asymmetric FP (with F a functional head) structure proposed in DiSciullo (2005) for compounds that are formed in her morphological level. We depart from DiSciullo's account in that she realizes the compound-internal linking element as the head of FP, while we assume that it is a higher Lnk(er) projection resulting from predicate-inversion. For semi-phrasal compounds (e.g. aráhn-i + ánthrop-os "man spider"), we assume the base compound-internal predicate structure in (19).

(19)
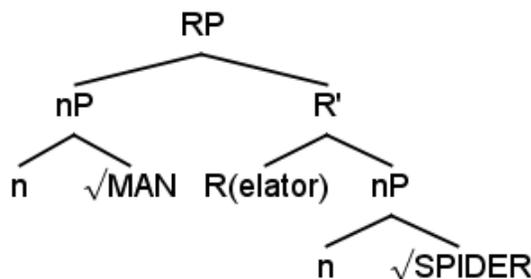
The predicate (complement of Relator) is a root, which has been assigned a category (and gender) by a nominalizing nP. In two out of three languages under consideration, nominal strings cannot surface without morphological case. There are two possibilities with respect to how the predicate nP's Case is assigned. If the size of the predicate string is just an nP and no quantization with number or D-elements has been added to the projection, then the nP appears with default case (nominative in the case of Russian and Greek). Alternatively, there is the possibility of an AGREE operation, matching features of the subject nP to the predicate nP. As we have

seen, both of these possibilities are attested in the data for all three languages (for number in Spanish, and for number and case in Russian and Greek). The Spec of RP is formed when the root √MAN merges with a nominalizing head. This nP will become the formal head of the compound and determine gender of the whole structure.

In root compounds, we assume a similar structure, albeit with the predicate being an acategorial root and not an nP:

(20)
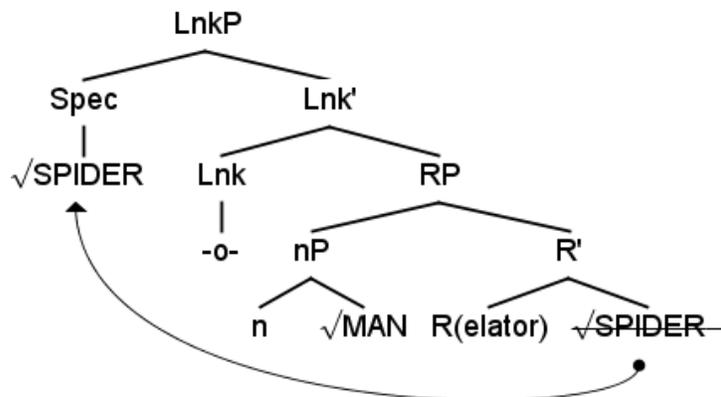
As a root, the predicate does not have any nominal category-defining morphology and no gender/number/case features. In order to be licensed it needs to invert over the subject. This generates a linker projection. The Lnk projection is just a landing site for licensing the root predicate and the whole string still behaves as a nominal string (i.e. there is no categorial, functional, or interpretive properties that Lnk contributes to the structure); it is just the landing site. Why would the root need to invert? Here, we believe that the relevant notion is that of "licensing." Nominal strings need to be licensed somehow and this is usually done by case assignment, via AGREE with a probe. However, in cases of RCs, the root is not quantized, and therefore it is ineligible for case-assignment considerations. Thus, the only way the root can be licensed is by movement to a licensing position. We propose that the linking position is such a position, for licensing a certain syntactic structure (similar to L(anding)Ps in Koopman and Szabolcsi (2000) where it would have been labelled LP$_{nP}$) (see also Alexiadou 2017). Note that the whole string is within the same phase domain (only one categorial head *n* which nominalizes the string).

3.2. DERIVING ASYMMETRIES BETWEEN RC AND SC. The differences in word order between the two types of compounds are explained by the presence/absence of Predicate Inversion. Canonical [Subj Pred] order creates left-headed structures we see in semi-phrasal compounds (where the compound "head" is the subject of the small clause) with an empty Relator. Right-headed order in root-compounds results from predicate inversion of the lower root -- this could be seen as an alternative to noun-incorporation as an account of root compounds (Baker 1995, Harley 2009).

The prosodic and interpretive properties of the derived compounds follow from the fact that RCs involve a single phase while SCs have two phases. The initial interpretation of the domain of "phase" in Chomsky 2008 includes reference to the phase being a prosodic domain. This may (and does) include the domain for stress assignment. Newell (2008) (see also Newell & Piggott 2014) proposes that stress assignment is sensitive to phase boundaries, including "weak" lower phase boundaries such as those defined by category changing morphemes. When a category-defining affix merges with a root, the resulting string is a prosodic domain for stress-assignment

purposes. Additional category-changing affixes may define domains for secondary stress assignment (see discussion in Newell 2008). In our account this predicts that RCs can only have a single main stress: the two roots are under a single category-defining head and, thus, in a single prosodic domain for stress assignment purposes. In contrast, in SCs each root is dominated by its own category-defining head, and so projects a separate prosodic domain for main stress assignment, resulting in two main stresses.

There are many investigations into the relationship between phases and the domain of semantic idiosyncrasy/idiomatic interpretation (Arad 2003, Borer 2009, Marantz 2012, Anagnostopoulou & Samioti 2013). This domain seems to be larger than the first phase domain (where a root combines with a category-defining head), but the size, in number of projections, of the syntactic structure seems to significantly control the possibility for a structure to be interpreted idiosyncratically (see for example Bruening 2014 for adjectival passives in English). This is compatible with our account and predicts that, while both root compounds and semi-phrasal compounds can have idiosyncratic meaning, root compounds will have idiomatic interpretations more frequently and in greater numbers than semi-phrasal compounds.

4. Compounds vs. Phrases. Both root and semi-phrasal compounds are distinct from attributive syntactic phrases in a number of respects, most of these differences grouped under the umbrella term "lexical integrity." Lexical integrity is defined as the inability of syntactic operations to "look inside" a word (see Anderson 1992, Lieber and Scalise 2006 for a detailed discussion). This includes ellipsis, movement (e.g. focus or topicalization operations), referentiality, and so on. In the case of compounds, it is true that most of them, including semi-phrasal ones, are not accessible to syntactic operations:

- they are not extendable (attributive member cannot be modified)
    a. a [dark blue] mug
    b. * [dark black] box; *a [black spider] man
- they are not separable (nothing can be inserted in-between two members)
    c. *a spider amazing man
- neither element can be replaced by an anaphor or elided
    d. *a superman and a spider one
    e. *a spider trapper and one man
    f. *a spider- and superman
- the non-head in compounds is typically non-referential

Our approach assumes that there is no separate morphological component in structure-building operations and that "words" are not the actual building blocks in syntactic operations. Thus, any presumed differences between words and phrases have to be explained by independently motivated syntactic operations. Therefore, the same model of syntax has to allow for the derivation of strings that may or may not participate in certain syntactic operations. The set of assumptions that has been generally pursued in the relevant literature, and which we adopt here, explains the different morphosyntactic distribution of compounds and syntactic phrases as a difference at the level of the derivation where these strings are formed. In particular, phase theory (Chomsky 2001, 2008 and subsequent work), assumes that a head defining a phase domain delimits this domain with respect to the PF and LF interfaces – an update of the earlier concept of the "cycle." If spellout takes place at the level of the phasal boundary, this means that the elements inside the phase (except its boundary) become "trapped" and are not available for further syntactic operations. Since category-changing heads delimit a phase boundary, this means that the nP formed in

root compounds is a single phase (including the landing projection LnkP). On the other hand, semi-phrasal compounds and syntactic phrases contain additional, separate phases. Thus, the compound-internal predicate phrase and its constituent parts cannot be available for further syntactic operations, such as phrasal movement and/or phonological deletion under ellipsis.

A second distinctive property of compounds containing units smaller than a DP is the issue of "quantization" (Sportiche 1999, 2005, Longobardi 2008). A bare nP/NP is unquantized, i.e. it only assumes referential properties when it is combined with a (possibly silent) determiner or other quantizing functional material above the phasal nP domain. In addition, since the root or nP predicate inside a compound does not refer to individuals, and thus is smaller than a DP, there are no case requirements and it cannot be referred to (cf. Longobardi 2008). This explains straightforwardly why there can be no reference to the compound internal nominals and the lack of case-licensing. No determiner-like elements can modify the predicate nP.

Derivation of nominal or adjectival predicates in syntax involves a Relator Phrase accommodating the subject and its predicate. Thus, for "this man is a spider", we would expect a structure of the type below:
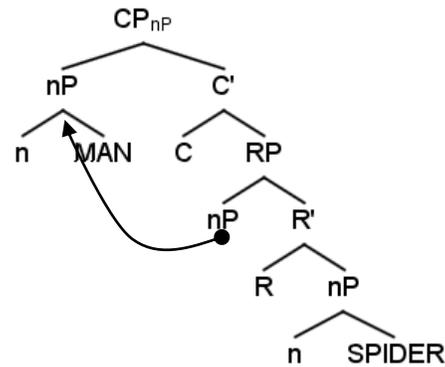
(21)

TP
　T'
T　RP
　DP　R'
　this man　R　DP
　　be　a spider
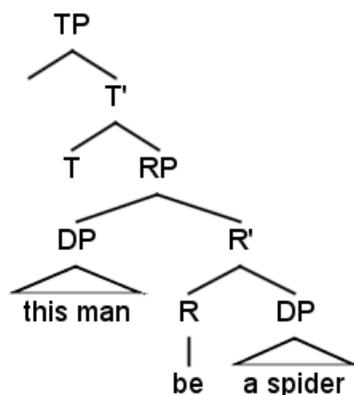
Both members of the predicate RP here are fully quantized DPs which require and receive structural case, are referential, and can be extracted via A-movement (e.g. movement of the subject to SPEC-TP for case reasons) or A'-movement (e.g. topicalization of the post-copular DP).
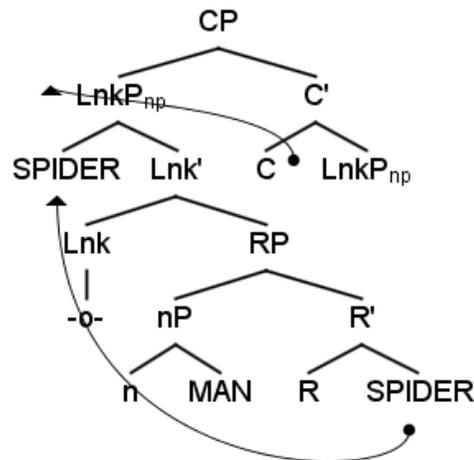
In contrast, in semi-phrasal compounds, the size of the two elements is smaller. We assume that the lower predicate nP has no additional functional projections, while the subject nP may have adjectival modification, and can be quantized by being selected by a determiner. This takes place though, after extraction to SPEC-CP$_{nP}$ which forms a reduced relative clause providing the resulting semi-phrasal compound with the interpretation of a referential element: "man (who is) spider." The CP$_{nP}$ in turn, can be selected by a D-element creating the strong-DP phase. Anything below D is not available anymore for subsequent syntactic operations. Furthermore, since D selects the full CP$_{nP}$, only the latter can be quantized and thus available for coreference. The smaller predicate nP *spider* is not quantized, and thus not available for case or reference considerations.

(22)

CP$_{nP}$
  nP   C'
 n MAN   C   RP
           nP   R'
          R   nP
         n   SPIDER

The full derivation of a root compound is shown below:

(23)

CP
  LnkP$_{np}$ C'
SPIDER Lnk'   C   LnkP$_{np}$
      Lnk   RP
      -o-   nP   R'
          n MAN   R   SPIDER

The root SPIDER cannot be licensed in base position, and thus it undergoes Predicate Inversion over a linking element. Subsequently, LnkP pied-pipes the whole predicate string to SPEC-CP to create a head-final reduced relative clause with a similar interpretation of "man who is spider." In this case, movement freezes the LnkP and everything inside it, making any extraction out of it impossible. In addition, the whole string is a single phase – single prosodic domain.

5. Conclusion. We have shown that the properties of RCs and SCs can be derived from the sizes of the structures involved, and from the differences between derivations from bare-roots, vs. categorized phrases. Unlike lexicalist approaches, our account does not need to posit two different computational components of grammar and assume a distinction between morphological vs. syntactic compounding. The same rules of syntax derive units that range from phrase-like to word-like, and their different properties are explained based on their size and the derivational path of their formation. In the future, we plan to expand our approach to cover other languages and accommodate other semantic types of compounding (e.g., subordinate, coordinate, and exocentric).

References

Adger, David & Gillian Ramchand. 2003. Predication and equation. *Linguistic Inquiry* 34. 325–359. https://doi.org/10.1162/002438903322247515.

Alexiadou, Artemis. 2017. Proper name compounds: A comparative perspective. Paper presented at the 7th Biennial Conference on the Linguistics of Contemporary English, University of Vigo. 28–30 September 2017.

Anagnostopoulou, Eleni. & Yota Samioti. 2013. Allosemy, idioms, and their domains. In Raffaella R. Folli, Christina Sevdali & Robert Truswell (eds.), *Syntax and its limits*. 218–250. Oxford: Oxford University Press.

Anderson, R. Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Arad, Maya. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language & Linguistic Theory* 21. 571–596. https://doi.org/10.1023/A:1025533719.

Athanasopoulou, Anna & Irene Vogel. 2014. Compounds in Modern Greek. 168th Meeting of the Acoustical Society of America (ASA), Indianapolis, IN.

Baker, Mark. 1995. Lexical and nonlexical noun incorporation. In Urs Egli, Peter Pause, Christoph Schwarze, Arnim von Stechow & Götz Wienold (eds.) *Lexical knowledge in the organization of language*. 3–34. Amsterdam: John Benjamins.

Bisetto, Antonietta, & Sergio Scalise. 2005. The classification of compounds. *Lingue e linguaggio*, 4(2). 319–332.

Borer, Hagit. 2009. Roots and categories. Paper presented at the 19th Colloquium on Generative Grammar, University of Basque Country.

Bowers, John. 1993. The Syntax of predication. *Linguistic Inquiry* 24. 591–656.

Bowers, John. 2008. Predication. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*. https://do.org/10.1002/9780470756416.ch10

Bruening, Benjamin. 2014. Word formation is syntactic: Adjectival passives in English. *Natural Language & Linguistic Theory* 32. 363–422. https://doi.org/10.1007/s11049-014-9227-y.

Chomsky, Noam. 2001. Derivation by phase. In: Michael Kenstowicz (ed.), *Ken Hale: A life in language*. 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory. Essays in honor of Jean-Roger Vergnaud*. 133–166. Cambridge, MA: MIT Press.

Den Dikken, Marcel, and Pornsiri Singhapreecha. 2004. Complex noun phrases and linkers. *Syntax* 7. 1–54.

Den Dikken, Marcel. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas*. Cambridge, MA: MIT Press.

Di Sciullo, Anna Maria. 2005. Decomposing compounds, *SKASE Journal of Theoretical Linguistics* 2. 14–33.

Embick, David. 2010. *Localism versus globalism in morphology and phonology* (Linguistic Inquiry Monograph 60.) Cambridge, MA: MIT Press.

Embick, David. & Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39. 1–53.

Gouskova, Maria. 2010. The phonology of boundaries and secondary stress in Russian compounds. *The Linguistic Review* 27(4). 387-448. https://doi.org/10.1515/tlir.2010.015.

Guevara, Emiliano R. 2012. Spanish compounds. *Probus*. 24(1). 175–195.

Guevara, Emiliano & Sergio Scalise. 2008. Searching for universals in compounding. In Sergio, Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*. Berlin: Springer. 101–128.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel S. Keyser (eds.), *The view from Building 20*. 111–176. Cambridge, MA: MIT Press.

Harley, Heidi. 2009. Compounding in Distributed Morphology. In Rochelle Lieber & Pavol Stekauer (eds.), *The Oxford handbook of compounding*. 129–144. Oxford: Oxford University Press.

Hualde, José Ignacio. 2007. Stress removal and stress addition in Spanish. *Journal of Portuguese Linguistics* 6(1).

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Koopman, Hilda & Anna Szabolcsi. 2000. *Verbal complexes* (Current Studies in Linguistics 34). Cambridge, MA: MIT Press

Koster, Jan. 1994. Predicate incorporation and the word order of Dutch. In Raffaella Zanuttini (ed.), *Paths towards universal grammar. Studies in honor of Richard S. Kayne*, Georgetown University Press, Washington D.C. 255–77.

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

Longobardi, Giuseppe. 2008. Reference to the individuals, person and the variety of mapping parameters. In Høeg Müller, H. & Alex Klinge (eds.), *Essays on nominal determination*. 189–212. Amsterdam & New York: John Benjamins. https://doi.org/10.1075/slcs.99.11lon.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds.), *University of Pennsylvania working papers* 4.2. 201–225.

Marantz, Alec. 2001. Words and things. Ms. MIT.

Marantz, Alec. 2012. Locality domains for contextual allosemy in words. Handout, New York University.

Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.

Moyna, María Irene. 2011. *Compound words in Spanish: Theory and history*. Amsterdam & New York: John Benjamins.

Nespor, Marina & Angela Ralli. 1994. Stress domains in Greek compounds: A case of morphology-phonology interaction. In Irene Philippaki-Warburton, Katerina Nicolaidis & Maria Sifianou (eds.). *Themes of Greek linguistics I*. 201–208. Amsterdam & New York: John Benjamins.

Nespor, Marina & Angela Ralli. 1996. Morphology-phonology interface: Phonological domains in Greek compounds. *The Linguistic Review* 13. 357-382. https://doi.org/10.1515/tlir.1996.13.3-4.357.

Newell, Heather. 2008. Aspects of the morphology and phonology of phases. McGill University.

Newell, Heather & Glyne Piggott. 2014. Interactions at the syntax–phonology interface: Evidence from Ojibwe. *Lingua* 150. 332–362. https://doi.org/10.1016/j.lingua.2014.07.020.

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

Roon, Kevin. 2006. Stress in Russian compound nouns: Head dominance or anti-faithfulness. *Proceedings of FASL* 14. 319–330.

Sportiche, Dominique. 1999. Reconstruction, constituency and morphology. Talk delivered at the 1999 GLOW conference, Berlin.

Sportiche, Dominique. 2005. Division of labor between Merge and Move: Strict locality of selection and apparent reconstruction paradoxes. *Proceedings of Division of Linguistic Labor*, the La Bretesche Workshop.

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

Williams, Edwin. 1981. On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12. 245–274. https://www.jstor.org/stable/4178218.