Catalan Verbal Compounds and the Syntax-Morphology Competition

Susanna Padrosa Trias
Universitat Autònoma de Barcelona

1. Introduction

The Catalan compounds [NV]$_V$ have been a topic of discussion among linguists on several occasions in the literature (cf. Gràcia & Fullana (GF), 1999, 2000; Adelman, 2002; and Brunelli, 2003, just to cite a few references). The main interest in the study of this compound type lies in the fact that the internal word order of its constituents is OV, whereas Catalan is a VO language. That is, the right-headedness of the compound contradicts the left-headedness of the language. Despite being an appealing and interesting question of study by itself, the opposite word order found in the Catalan syntax and the morphological complexes presented here will only be touched upon briefly throughout the paper, mainly when discussing the possible derivations of the Catalan synthetic compound. One of the goals of the paper is then to explain the derivation of the [NV]$_V$ type of compounds in which the N indicates inalienable possession[1] and is typically interpreted as an argument of the verbal head (cf. e.g. Cabré & Rigau, 1986; Mascaró, 1986; Gavarró, 1990),[2] some examples of which are given in (1).

---

[1] The nouns that form part of [NV]$_V$ compounds are typically inalienable possession nouns (IPNs). Some examples are *cor* ('heart'), *pell* ('skin'), *cama* ('leg') and *coll* ('neck'), the meaning of which can be used in transparent compounds (like those given in (1)) and can also be subject to sense extensions to incorporate a more figurative semantics, as illustrated in *corferir* heart+hurt ('to break somebody's heart'). Note, though, that this type of compounds can also involve nouns other than strictly IPNs like *aigua* ('water') and *terra* ('earth') in compounds like *aiguabarrejar-se* (water+mix+CL 'to have waters mix') and *terratrémer* (earth+shake 'to have the earth quake'). See Gavarró (1990: 78) for a proposal according to which strict IPNs and inanimate nouns like *aigua* and *terra* are unified under a Non-Distinctness Constraint. In her terms, 'The Non-Distinctness Constraint allows for lexical chains to be formed by nominals which have identical referents', which she illustrates with the following example in which the river and its water are clearly non-distinct (Gavarró 1990: 81).

> (i) Aquest   riu   i   el Danubi s'aiguabarregen a   Alemanya.
>   This   river and the Danube CL-water-mix   in   Germany
>   'This river and the Danube join in Germany'

[2] See GF (1999: 246, 2000: 79) for a different view. They propose that the IPN is not an argument but a modifier of the complex predicate formed by the verb together with the possessor NP external to the complex verb. To illustrate the point, in *alatrencar un ocell* (wing+break a bird 'to break a bird's wing(s)'), *trencar un ocell* ('break a bird') would be the complex predicate that the IPN *ala* ('wing') modifies. A possible paraphrase could be 'to break the bird by the wing(s)'.

(1) a.   alatrencar
          wing+break
          'to break the wing(s) (of an animal)'
     b.   camatrencar
          leg+break
          'to break the leg(s)'
     c.   collportar
          neck+carry
          'to carry on one's shoulder'
     d.   pellforadar
          skin+pierce
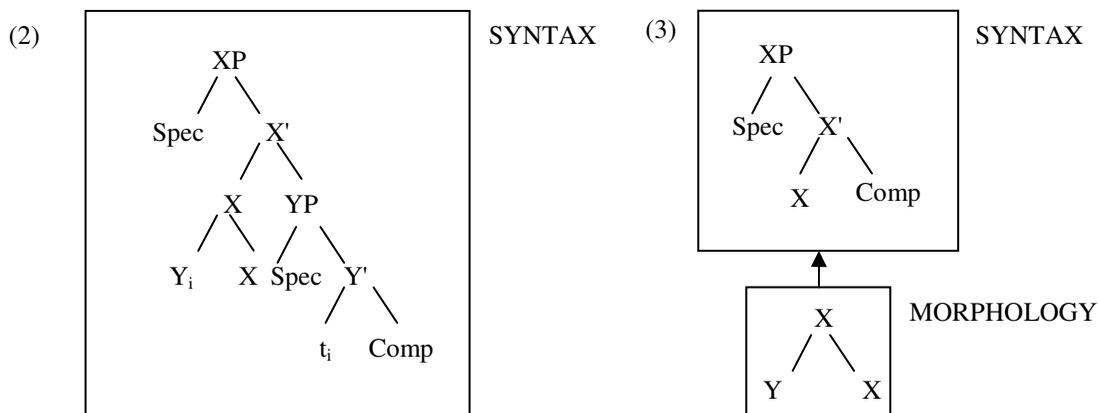   'to make a hole in something/somebody's skin'

Catalan [NV]$_V$ compounds will be argued to be generated by an independent morphological system, and not by syntactic movement. It will be shown that the syntactic account cannot easily explain some phenomena while a morphological account can do so more satisfactorily.

The second main goal of this paper is to validate the competition analysis between syntax and morphology, as put forward by Ackema & Neeleman (AN) (2004, 2007), by examining the Catalan [NV]$_V$ compounds together with their syntactic counterparts. I follow AN (2004, 2007) in the sense that I assume a model of grammar in which morphology is distributed across three independent modules, i.e. syntax, semantics and phonology. Each of them contains two submodules, which generate phrasal and word-level representations (cf. Jackendoff, 1997).

The paper is structured as follows. Section 2 contains two different views of word formation as well as the opposite predictions they make regarding some phenomena like stranding (Section 2.1), referentiality (Section 2.2), possible functions of nonheads (Section 2.3), prototypicality in incorporating languages (Section 2.4), headedness (Section 2.5) and derivational economy (Section 2.6). In Section 3 AN's (2004, 2007) theory of morphosyntactic competition is presented. The following section is divided into two subsections: Section 4.1 shows the interaction of some Catalan data with the morphosyntactic competition analysis and Section 4.2 illustrates the real use of the compound under study. Finally, Section 5 concludes the paper with a summary of the main results and poses some questions for further research.

2. Two different views of word formation

By assuming a model of grammar in which syntax has two subcomponents: one for phrasal syntax and one for word syntax (which will be referred to as S for 'syntax' and M for 'morphology' for ease of exposition henceforth), the question of where complex words are formed arises naturally. S and M can be seen as two potential components in which complex words can be generated. On the one hand, there are theories like those of Roeper & Siegel (RS) (1978), Baker (1988), Hale & Keyser (HK) (1993) and Brunelli (2003) in which complex words are formed by syntactic movement and which do not postulate an independent morphological system. On the other hand, there are theories like those of Selkirk (1982), Di Sciullo & Williams (1987), and AN (2004, 2007) which have independent components for S and M and claim that complex words can be generated by an independent morphological system. The two different theories of word formation can be illustrated with the following trees: (2) illustrates that the complex word is derived by head-to-head movement of Y, which adjoins to the higher head X; (3) shows how the complex word is derived in an independent morphological component and is then inserted into S (cf. AN 2004: 18).

(2) ... SYNTAX   (3) ... SYNTAX

(diagram 2: XP tree with Spec, X', X, YP, Y_i, X, Spec, Y', t_i, Comp)

(diagram 3: XP tree with Spec, X', X, Comp — SYNTAX; and below MORPHOLOGY box with X, Y, X)

At first sight, it looks as if the $[NV]_V$ type of compounds should be dealt with in the syntactic component. If $[NV]_V$ compounds were formed in S, a quick and plausible explanation for their rather low productivity in the language comes to mind. The word order of the internal elements to be compounded conflicts with the VO pattern of the language, which could explain why this type of compounds is not part of the Catalan speaker's everyday vocabulary. Violating the syntactic word order in Catalan makes this compounding process be used less and less, with the consequence that native speakers can understand most of the $[NV]_V$ compounds but rarely produce them (cf. Adelman, 2002; contra GF, 1999, 2000; Brunelli, 2003; cf. Section 4.2).

By contrast, if these compounds were produced in M, it would be difficult to account for their low productivity because Catalan already has right-headed morphological processes. For example, consider inflectional suffixes like the plural marker $-s$ (gat_{SG} 'cat' > $gats_{PL}$ 'cats') or the feminine marker $-a$ (nen_{MASC} 'boy' > $nena_{FEM}$ 'girl'), derivational suffixes like $-or$ which can derive nouns from adjectives (e.g. $groc_A$ 'yellow' > $grogor_N$ 'yellowness') and $-itzar$ which can derive verbs from adjectives (e.g. $industrial_A$ 'industrial' > $industrialitzar_V$ 'to industrialize'), and even other right-headed compounds like *camacurt* (leg+short 'short-legged') and *malgastar* (badly+spend 'to waste money/time'), which are $[NA]_A$ and $[AdvV]_V$ respectively. As a consequence, the right-headed compound $[NV]_V$ should not constitute a problem regarding its production, but apparently it does.

Despite appearances, closer inspection of $[NV]_V$ compounds shows that a morphological analysis is preferable to a syntactic one. In the following section some tests available in the literature will be applied to the Catalan data, which will illustrate that the morphological account of complex word formation makes the correct predictions regarding $[NV]_V$ compounds and fares better with the data than a syntactic account.

2.1. Stranding

The possibility of having stranded material modifying a nonadjacent element has been taken in the literature as an indicator of syntactic movement (cf. Baker, 1988; Kiparsky, 1997, a.o.).

If Catalan $[NV]_V$ compounds were formed by syntactic movement à la Baker (1988), there would be a derivation in which the V first selects the N as its complement and then the N undergoes head-to-head movement to incorporate into the higher verbal head. As a result, the N leaves a trace once it has incorporated into the V (cf. (2)). According to the syntactic theory of word formation, the trace should be able to license the material stranded, i.e. the material that originally modified the N prior movement. However, as the following examples illustrate, Catalan verbal compounding does not seem to be the result of N incorporation.

(4) a. En Joan trencà les ales   llargues dels   ocells
        The John broke   the wings long   of+the   birds
        'John broke the birds' long wings'

b. *En Joan alatrencà[3]   les t llargues dels ocells
   The John wing+break-PAST the long of+the birds
   'John broke the birds' long wings'

(5) a. La treballadora va foradar la pell de lleopard que tenia més a prop
    The worker went pierce the skin of leopard that had more to close
    'The worker made a hole in the leopard skin that was closer to her'
  b. *La treballadora va pellforadar la t de lleopard que tenia més a prop
    The worker went skin+pierce the of leopard that had more to close
    'The worker made a hole in the leopard skin that was closer to her'

If the incorporated N root in incorporating languages can be modified by non-adjacent material (e.g. determiners, quantifiers, relative clauses) which remains outside the complex V (cf. Baker, 1988), the examples in (4)-(5) suggest that there is no syntactic movement of the N in these Catalan verbal compounds and hence the absence of a trace explains the ungrammaticality of the material stranded, which cannot be licensed.

On the other hand, under the morphological analysis, the N and the V of the compound would merge directly in a separate module, the morphological one, and there would be no need to mention the obligatory omission of syntactic material, this being absent to start with.

AN (2004: 19) illustrate the same phenomenon with some data from English.

(6) a. *[hand$_i$made by the strong t_i of a blacksmith]
  b. *[central$_i$ize more t_i to our arguments than we thought]

The ungrammaticality of the examples in (6) shows that adjectival compounding and derivational affixes in English do not allow stranding of material accompanying the purported incorporated element either, and are also better analysed in a non-syntactic movement theory of word formation. Parallel examples to those given in (6) can be easily constructed, some of which are given below:

(7) a. *[home$_i$made in my friend's beautiful t_i in Scotland]
  b. *[sing$_i$er loudly t_i in the park]

In short, all the examples in this section support the idea that the nonhead of complex words does not move from a lower position in a syntactic tree. Rather, the nonhead and the head of complex words seem to be merged directly in an independent morphological system and inserted in the syntactic component later.

2.2. Referentiality

The possibility of the nonhead having referentiality and specificity when being compounded with another element is taken as evidence for syntactic movement. The nonhead leaves a trace in its base-generated position, the presence of which allows the object to be referentially transparent. Baker (1988) illustrates the referential transparency of incorporated Ns in languages like Mohawk and contrasts them with languages like English, in which the first element of compounds, such as *money-loser*, *tobacco-buying*, *babysitter* and *grocery-shopping*, is non-referential. They do not refer to specific items in the world.

Catalan is similar to English in the sense that the first element of [NV]$_V$ compounds is not referential either. For example, in a sentence like (8), one does not know whether the hunter broke one leg or the two of them. It is left indeterminate.

---

[3] Notice that the alleged incorporated N loses the plural marker. Brunelli (2003), who proposes a syntactic movement theory to generate the Catalan [NV]$_V$ compounds, explains the absence of plurality on the incorporated N by suggesting that incorporation implies the loss of inflection, plurality included.

(8) El caçador va   camatrencar l'ocell
    The hunter   went   leg+break   the-bird
    'The hunter broke the bird's leg(s)'

In other cases, it will be one's knowledge of the world, and not the referentiality of the N, that will tell us that the N can only refer to a single entity. Consider compounds like *colltrencar* (neck+break 'to break somebody's neck') and *corbategar* (heart+beat 'to have the heart beat').

To summarize, the non-referentiality of the first element of Catalan [NV]$_V$ compounds favours the view according to which the compounds are formed by an independent morphological system, and are not the result of movement. In other words, the anaphoric islandhood shown by these complex words supports the Lexical Integrity Principle (cf. e.g. Di Sciullo & Williams, 1987), according to which the internal structure of words is invisible to S.

2.3. Possible functions of nonheads

The two different views of word formation make opposite predictions with respect to the elements which can surface in the first position of the compound. According to the syntactic account, incorporation of the head of subjects, adjuncts and objects of prepositions must be banned, because they are islands for extraction. The trace left by movement of the N would violate the Empty Category Principle (ECP), according to which traces must be properly governed (cf. Baker, 1988). Under the morphological account, there are no restrictions on the interpretation of the first element of the compound, because it merges directly with the V. Consider the following sentences.

(9) a. En Joan porta   la Maria a   coll
         The John carries the Mary on neck
         'John carries Mary on his shoulders'
     b. En Joan collporta   la Maria
         The John neck+carries the Mary
         'John carries Mary on his shoulders'

(10) a. En Joan va   calcigar la mà   de la Maria amb el seu peu
          The John went   tread   the hand of the Mary with the his foot
          'John trod Mary's hand (with his foot)'
      b. En Joan va   peucalcigar la mà   de la Maria
          The John went   foot+tread the hand of the Mary
          'John trod Mary's hand (with his foot)'
          (Example from GF, 1999: 255)

The syntactic account has no explanation for the incorporation of *coll* and *peu* from an adjunct position to derive the well-formed compounds *collportar* and *peucalcigar*. The morphological approach to word formation imposes no restrictions on the interpretation of the nonhead in compounding, morphological complexes being merged directly in a module separate from S. The morphological analysis has then no problems in generating [NV]$_V$ compounds in which the N is interpreted as an adjunct.

Note that although the IPN of the compounds in (9) and (10) refer to the possessor in subject position, those compounds in which the IPN refers to the possessor in object position are also potential problems for syntactic movement theories of word formation. In addition to the interpretation given in (10), *peucalcigar* can also be interpreted as 'to step on somebody's foot'. Similar examples are given in (11).[4] Some authors, such as GF (1999, 2000), have provided semantic paraphrases like those presented in (11a") and (11b") for the compounds under study.

---

[4] See GF (1999: 246) for further examples.

(11) a. corferir   en Joan
          heart+hurt the John
          'to break John's heart'
      a'. ferir   el   cor   d'en   Joan
          hurt   the heart of+the   John
      a". ferir   en Joan   del   cor
          hurt   the John of+the   heart
      b. alatrencar   un   ocell
          wing+break   a   bird
          'to break a bird's wing(s)'
      b'. trencar   l'ala   de   l'ocell
          break   the+wing of   the+bird
      b". trencar   l'ocell   per l'ala
          break   the+bird by   the+wing

   If the paraphrases in (11a") and (11b") are taken seriously, the syntactic theory of movement predicts that the corresponding complex words should be ungrammatical, contrary to facts. Notice that the IPNs are introduced by a P and, as a result, they should not be able to incorporate into a higher head in S. They are objects of prepositions and the trace left will not be properly governed. To account for the well-formedness of these compounds, Brunelli (2003) suggests that the presence of these prepositions is the consequence of signalling a generic relation between the possessor and the IPN, and that no true locative, dative or genitive complement is introduced by the prepositions. Regardless of this view being true, the presence of the P is still a barrier for an object of the P to move.
   Other examples that will cause problems for the syntactic analysis of complex words are Catalan [AdvV]$_V$ compounds, in which the adverb clearly functions as an adjunct. Consider the examples in (12), which are all made up of the adverb *mal* ('badly') and a V.

(12) a. malgastar   *vs*   gastar malament[5]
          badly+spend   spend badly
          'to waste money'
      b. malacostumar   *vs*   acostumar malament
          badly+get used to   get used to badly
          'to spoil (somebody)'
      c. malvendre   *vs*   vendre malament
          badly+sell   sell badly
          'to sell (something) cheap'
      d. maltractar   *vs*   tractar malament
          badly+treat   treat badly
          'to ill-treat'

   Problematic examples for syntactic theories of movement are not restricted to Catalan. It is a fact that the N of most English compounds formed by a N plus a deverbal element can be considered the internal argument of the original V. These compounds pose no problem for a syntactic analysis. The N, being selected by the verbal head, will head the complement NP from which it will move to adjoin the V higher up in the tree. The movement of the N will leave a trace, which will be properly governed and hence no violations of the ECP will be incurred. The examples in (13) illustrate the point.

(13) a. heartbreaker
      b. typesetter
      c. money-changer
      d. habit-forming
      e. child-bearing
      f. gum-chewing
          (Examples from RS, 1978: 219, 220)

---

[5] *Mal* and *malament* 'badly' are two positional variants of the same form.

Further support for this analysis comes from the contrast in grammaticality judgements between verbal compounds in which the first nominal element satisfies the internal argument of the V (14a, 15a, 16a) and synthetic compounds in which the N bears no thematic relation to the same V (14b, 15b, 16b).

(14) a. coffee-maker
     b. *fast-making

(15) a. child-molester
     b. *early-molester

(16) a. story-teller
     b. *home-teller
       (Examples (15) and (16) from RS, 1978: 223, 243)

To summarize, the syntactic theory of movement predicts the existence of synthetic compounds whose first element is subcategorized by the verbal base, i.e. it satisfies the internal argument of the V. Otherwise, the compounds are predicted to be ungrammatical. Although the examples so far all conform to this pattern, there are other examples that question it. Consider, for example, *truckdriver* and *homemaking*. They can have two potential different readings. Regarding *truckdriver*, *truck* can be understood as the internal argument of *drive* and then the compound expresses that somebody drives trucks, but it can also be understood as a modifier of the V and then the compound can refer to a driver of a car who has a truck drawn on his T-shirt (cf. Lieber, 2003: 250). The compound *homemaking* can also have one reading in which the N is taken as the internal argument of the V *make* to express that somebody constructs houses and another one in which *home* is not subcategorized by the verbal base to refer, for example, to somebody who creates a special atmosphere in a house (cf. RS, 1978: 216). Syntactic theories of movement à la Baker cannot derive the interpretation in which the nominal element acts as a modifier/adjunct of the underlying V. In the corresponding syntactic structure, the N will be located within an adjunct phrase from which it will not be able to move without violating the ECP. At first sight, it does not look obvious why the same compound has two different underlying representations to derive the two possible readings. However, that is a consequence of the adherence to the UTAH[6] by most syntactic theories of movement (cf. Baker, 1988; HK, 1993, 1998, 2002, a. o.). There is a direct mapping between thematic roles and syntactic structure.

Other examples that pose a problem for the theories under discussion are given in (17) and (18).

(17) a. concert singer
     b. party drinker
     c. spring-cleaning
       (Examples from Selkirk, 1982: 24)

(18) a. expert-tested
     b. pan-fried
     c. snow-covered
       (Examples from RS, 1978: 242)

The underlying verbs of the compounds in (17) are transitive and require an internal argument. Instead of satisfying the subcategorization requirement imposed by the V, the nonheads add a locative

---

[6] UTAH stands for Uniformity of Theta Assignment Hypothesis and is defined by Baker (1988) in the following terms: "Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure".

(17a, b) or temporal (17c) specification to the head.[7] In other words, they are adjuncts and present the problem just discussed for *truckdriver* and *homemaker*.

A similar explanation can be applied to synthetic compounds with adjectival passives like those in (18). Again, the nominal elements (i.e. *expert*, *pan* and *snow*) are not the internal argument of the underlying verbs (i.e. *test*, *fry* and *cover*). They could be argued to be adjuncts introduced by the prepositions *by*, *in* and *with* respectively. To solve the problem of moving from a PP, RS (1978: 242-243) propose a rule that deletes the prepositions, which they take as relational elements, with the consequence that their deletion causes no problem of recoverability. I share their view that the prepositions involved in (18) are relational elements, as I believe all prepositions are, but I part company with them on the role of the semantics associated with these prepositions. Despite being relational elements, these prepositions have specific semantics which cannot be downplayed. They are semantic contentful prepositions. The semantics of agentivity, location and instrument introduced by the prepositions *by*, *in* and *with* in (18) is necessary to correctly interpret the syntactic counterparts of the compounds. RS's belief in the no-recoverability problem is just a stipulation to make their theory work.

To recap, synthetic compounds in which the nominal element is interpreted as an adjunct are a problem for syntactic movement theories of word formation. Notice, though, that they are also a problem for syntactic non-movement theories. For example, Harley (2002, 2003) adopts HK's (2002) idea that conflation is associated with Merge. HK understand conflation as copying the phonological material, the 'p-sig' in their terms, of the sister head into the higher phonological empty head, i.e. a head with a defective p-sig. Harley (2002, 2003) represents the defective p-sig by the feature [±affix],[8] which can be generated on any head. Harley (2003: 4) defines her mechanism as follows:

(19)   Only the p-sig of the label of its sister may be copied during merge of a [+affix] head. The copied p-sig is a copy of the p-sig of the head of the larger constituent. Any p-sigs within that constituent that are not in its head will not make it into the label.

This approach can easily derive compounds in which the conflated N is interpreted as the object of the V, but not as a modifier/adjunct of it. Consider how conflation derives *truckdriver*. First, *drive* is selected from the numeration. This lexical item happens to have a [+affix] feature. After selecting *truck* from the numeration and merging it with *drive*, there is copying of the p-sig of *truck* into *drive*. The result of projecting the head will be *truckdrive*, which is subsequently merged with *–er*, which also happens to have a [+affix]. Due to the affixal nature of the suffix, the label of its sister, *truckdrive*, will be copied into the defective p-sig of *–er*. After that, the head will project as *truckdriver*. The derivation is depicted in (20).

(20)

---
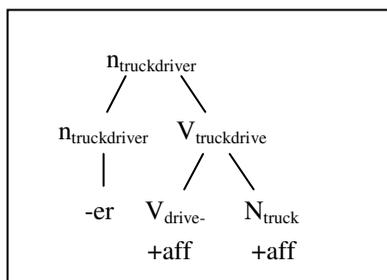
[7] One might argue that the internal argument of the verbs in (17) is implicit and consequently the subcategorization frame of the V is satisfied. Another possibility is the one proposed by HK (2002), in which the lexical semantics of the V is rich enough to identify and license its nonovert complement. Contrast *She danced*, in which the V can identify its object, with **He made* (cf. *He made trouble*), in which the light V does not contain enough information to identify its object. I do not go deeper into the different possibilities and the plausibility of each of them, since they are orthogonal to the present discussion.

[8] Following Harley (2002, 2003), the label [±affix] does not necessarily imply that the heads are necessarily morphophonologically affixal in nature, though they are most of the time.

On Harley's account, *truckdriver* will have a unique structure, the one depicted in (20), which cannot distinguish between the two possible readings of the compound. Without adding any modifications to the system, it can only generate the interpretation in which *truck* is the internal argument of *drive*.

The examples in (17) and (18) are also a problem for Harley's approach. Like RS (1978), for the theory to work, Harley has to stipulate that the prepositions intervening between the nominal elements being incorporated and the underlying verbs are null or have defective p-sigs. (18b) will have the derivation in (21).

(21)
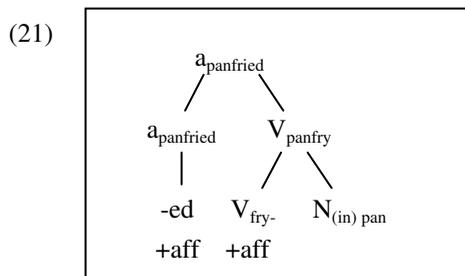
In short, synthetic compounds whose first element is a modifier cannot be easily explained by syntactic theories of word formation, be with or without movement. By contrast, in morphological approaches to word formation, as the one discussed here, the two elements of the compound are merged directly, with no need for amendments to the theory to make it work.

2.4. Incorporating languages

The discussion above has given ample evidence for the fact that Catalan is not a prototypical incorporating language in Baker's (1988) sense. The alleged incorporated element in Catalan cannot be modified by material outside the verbal complex (Section 2.1) and cannot be referential either (Section 2.2), features that define prototypical N incorporation. In addition, the nouns in Catalan $[NV]_V$ compounds can function as adjuncts, objects of prepositions in the underlying syntactic structure, which is an unexpected fact under an incorporating analysis of complex words (Section 2.3).

If one persisted in treating $[NV]_V$ compounds as incorporation, Catalan would be unique among Romance languages and within Catalan itself the process would be limited to a specific construction. In other words, if $[NV]_V$ compounding were really incorporation, one would expect to find the same process with other right-headed compounds of the language. For example, why wouldn't compounds with a structure like $[NA]_A$ (e.g. *camacurt*, leg+short, 'short-legged')[9] and $[AdvV]_V$ (e.g. *maltractar*, badly+treat, 'to ill-treat') be the result of incorporation? And, crucially, why wouldn't other categories like prepositions and verbs incorporate on a par with nouns, as is commonly found in incorporating languages?

2.5. Headedness

Recall that Catalan syntax is left-headed whereas the compounds under discussion are right-headed. If $[NV]_V$ compounds were the result of syntactic movement, one would expect these compounds to follow the syntactic laws of headedness and be left-headed. On the other hand, if morphological complexes are merged directly in a separate module, the opposite word order between S and M can be more easily accounted for. M has left-headed and right-headed structures, the type of compounds under study belonging to the latter group.

In short, a derivation without syntactic movement can deal with the word order shown by $[NV]_V$ compounds more easily than a derivation with movement.

---

[9] See Brunelli (2003) for a specific proposal of incorporation for Catalan $[NA]_A$ compounds. Recall, though, the problems an incorporation analysis has with the incorporation of objects of prepositions.

2.6. Derivational economy

It is a well-known assumption that movement should only be used as a last resort. To put it differently, movement should be avoided if it does not have some interface effect (cf. Chomsky, 1995; Rizzi, 2004). As will be seen in Section 4.1, although the thematic relations between the elements in morphologically and syntactically derived structures are the same, the two structures do not have identical semantics. It will be argued that the difference in meaning is not due to syntactic movement but to the merger of the elements in a different component.

In conclusion, all the arguments and data just discussed all point to the fact that Catalan verbal compounds are not generated by syntactic movement (cf. Gavarró, 1990; GF, 1999, 2000), but are the result of direct merger in a morphological component. It has been argued that S and M are two different components with their own principles, although some of these may be shared by the two modules (cf. Jackendoff, 1997; AN 2004, 2007).[10]

In the following section the competition analysis between S and M, as put forward by AN (2004, 2007) will be presented.

3. Competition analysis: S vs. M (AN, 2004)

AN endorse a view according to which S and M are two competitive generative systems, since they argue that in principle two lexical items can be combined in either component. AN propose that all else being equal in languages like Catalan S wins over M, although morphological merger is also possible under certain conditions, i.e. when there is no syntactic competitor.[11] There is competition between S and M when both the categories merged and the semantic relation obtained are the same in the syntactic and morphological structure. AN (2004: 51) provide the constraint in (22), which summarizes the formal and semantic conditions just mentioned.

(22)   Let α_1 and α_2 be syntactic representations beaded by α. α_1 blocks α_2 iff
   (i)   in α_1 (a projection of) α is merged with (a projection of) β in syntax, while in α_2 (a projection of) α is merged with (a projection of) β in morphology, and
   (ii) the semantic relation between α and β is identical in α_1 and α_2.

The constraint in (22) is illustrated in (23) abstractly and (24) provides a concrete example.

(23)   √a. αP   b. α   (24)   √a. They drive trucks *vs* *b. They truckdrive
         / \   / \
        α βP   β α   √a.   VP   b.   V
         | / \   / \
         β   V   NP   N   V
                                              drive | truck drive
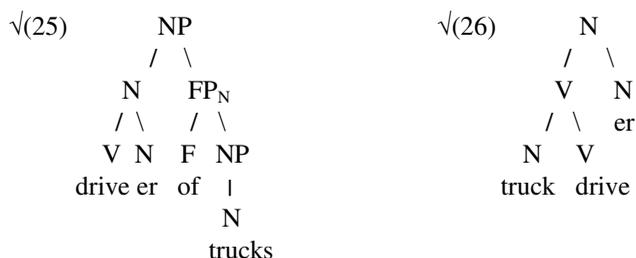   N
                                                   trucks

In both (23) and (24) there is competition between the two structures. As for (23), the same categories, α and β, are merged and the semantic relation between them is the same in the two generative

---

[10] See GF (1999: 257-259) and AN (2004: 23-45) for other arguments which allow the two accounts of word formation to be distinguished. Although GF use the tests to support a lexical analysis, they can also be taken as indicators of a morphological treatment.

[11] Whether there is a syntactic or morphological preference to combine lexical items depends on the type of language. In languages like English and Catalan syntactic merger will be the unmarked option, whereas in polysynthetic languages morphological merger will be the preferred option.
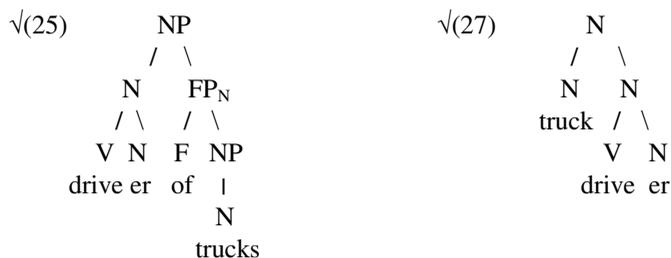
systems. Similarly, in (24) the same categories merge, i.e. a N and a V, and in both structures the N is interpreted as the object of the V. Competition is at work and gets the syntactic structure to be the winner.

As already said, morphological merger can block syntactic merger in certain circumstances, i.e. when different categories merge or the semantic relation between them is different in the two structures. AN (2004: 52) express the difference in semantics in the following terms: "Morphological merger of α and β may result in a semantics that cannot be expressed by the result of syntactic merger of the two". To illustrate the point, let us consider the syntactic derivation in (25) and contrast it with its morphological counterpart in (26).

```
√(25)        NP                  √(26)        N
           /  \                             /  \
         N    FPN                         V    N
        / \   / \                        / \   er
       V  N  F  NP                      N   V
    drive er of  |                   truck  drive
                 N
               trucks
```

Although they both involve the same semantics (i.e. *truck* is understood as the internal theta-role of *drive* in the two cases), the merger of different categories in the two structures makes the morphological merger viable. In (25) the merger of *drive* and −*er* results in a N, which in turn merges with the N *trucks* (functional projections do not count). In contrast, in (26), the merger of *truck* and *drive* crucially results in a V, which subsequently merges with the nominalizing suffix -*er*. To put it differently, only in (26) are *truck* and *drive* merged directly, which is what makes the morphological structure possible.

Let us consider (25) again and now contrast it with (27), another possible morphological derivation.

```
√(25)        NP                  √(27)        N
           /  \                             /  \
         N    FPN                         N    N
        / \   / \                       truck / \
       V  N  F  NP                            V   N
    drive er of  |                          drive er
                 N
               trucks
```

In this case, the two structures have the same categories merged but they differ in their semantics, which allows the existence of the morphological derivation. (27) is only allowed iff *truck* is not the internal argument of *drive*, but a modifier. The compound *truck driver* could refer to a driver of a car who has a picture of a truck on his T-shirt (recall the discussion in Section 2.3, cf. Lieber, 2003: 250).

AN adopt the general assumption that lexical storage should be the minimum, with the consequence that only unpredictable information will be stored. Given that syntactic merger blocks morphological merger where both can apply, morphological merger must be triggered. The trigger may be related to unpredictable or idiomatic readings of the morphological derivation. AN specify the morphological locus of merger with the diacritic M, as in <$_M$ αβ>. This will suspend the M-S competition and the morphological merger will be possible. That is the case with the root compound *colour code* in English. Contrast (28) with (29).

```
(28)     VP                      √(29)        V
       /  \                                  /  \
      V   N                                 N    V
   code  colours                         colour  code
```

The structure in (29) is possible because *colour code* is stored in the lexicon as being morphologically realized due to its unpredictable semantics. The semantics involved in (29) can only be derived in S via the P *with*. The expression *code with colours* is not in competition with *colour code*, due to the fact that different categories merge in the two derivations (see AN, 2004: 48-88 for the details of their morphosyntactic competition analysis).

4. Potential problems and extensions

4.1. Catalan data and the morphosyntactic competition

Given the morphosyntactic competition, if two lexical items can be combined both syntactically and morphologically, they should have different semantics, or the two derivations should involve merger of different categories. Since in most cases the same categories merge, this factor will not be considered to validate the competition analysis most of the time. When it becomes relevant, it will be mentioned.

At first sight, one might argue that Catalan [NV]$_V$ compounds and their syntactic counterparts share the same semantics, a stand taken by authors such as Mascaró (1986), Cabré & Rigau (1986), GF (1999, 2000), Brunelli (2003), who assume that, for example, *portar a coll* (carry on neck) is the same as *collportar* (neck+carry), and *trencar la cama* (break the leg) is the same as *camatrencar* (leg+break). Although at a superficial level, this generalization seems to be correct (with the consequence that AN's morphosyntactic competition is put into question), a deeper level of analysis shows that the semantics of the two structures is not exactly identical. Let us consider some examples.

(30) a. El caçador va alatrencar els ocells.
          The huner   went wing+break the birds
      b. El caçador va trencar les ales als   ocells.
          The hunter went break the wings to+the birds
      c. El caçador va trencar els ocells per les ales
          The hunter went break the birds by the wings
      d. El caçador va trencar les ales dels   ocells.
          The hunter went break the wings of+the birds

Because more than one paraphrase for the V *alatrencar* is possible, one might think that the M-S competition should predict the non-existence of the compound. However, it will be seen that in fact there is no competition between (a) and the rest of the structures in (30). The compound has an obligatory inalienable possession reading, which is also present in the syntactic paraphrases of (30b, c). In the case of (30b, c), the lexical prepositions (*als*, *per*) will prevent competition from taking place. Different categories will merge in the morphologically and syntactically derived structures. Regarding (30d), it cannot have an inalienable possession reading. The difference in meaning between (30a) and (30d) will suspend the competition and hence the morphological structure is allowed. Notice that the difference in semantics does not have to do with argument structure, understood broadly, but with a very fine-grained notion of semantics, that of inalienability.[12]

The following examples also show that a very detailed analysis of the semantics involved in each structure is necessary for the morphosyntactic theory to work.

(31) a. La Maria es va   trencar la cama
          The Mary CL went break   the leg
          'Mary broke her leg' (Mary can be understood as an Agent or as an Experiencer)
      b. La Maria es va   camatrencar
          The Mary CL went leg+break
          'Mary broke her leg' (Mary can only be understood as the Experiencer)

---

[12] See Brunelli (2003) for parallel examples and constrasts in Italian.

(32) a. El doctor va glaçar la sang de la Maria
     The doctor went freeze the blood of the Mary
     'The doctor froze Mary's blood', 'Mary was scared stiff'

   b. Aquella notícia terrible va sangglaçar la Maria
     That news terrible went blood+freeze the Mary
     'Mary was scared stiff by that terrible piece of news'

(33) a. En Joan porta a coll el seu fill
     The John carries on neck the his son
     'John carries his son on his shoulders', 'John carries his son (the manner not being specified)'

   b. En Joan collporta el seu fill
     The John neck+carries the his son
     'John carries his son on his shoulders'

In the case of (31), (a) can have two possible readings: one in which *Mary* is an agent and another one in which she is an experiencer. In other words, in the first reading she performs the action on purpose: maybe she broke her leg because she knew that with her leg broken, her mum would not make her go to school, and she did not want to go to school because she had an exam. As for the second reading, Mary's leg broke by accident. It was not her intention to break her leg: maybe she had a class trip to which she really wanted to go the following day but unfortunately she fell over a ball and she broke her leg. Of the two possible readings, (31b) has only the latter. This difference in meaning between the two structures would suspend the competition, which would explain the existence of the compound.

Similarly, in the case of (32) and (33) the syntactic derivation allows a wider range of interpretations than the morphological derivation. Again, this difference in semantics would explain why the compound is a possible derivation. Concerning (32), (a) can have a literal and a metaphorical reading, whereas (b) can only be understood metaphorically. Regarding (33), *portar a coll* can be understood literally as carrying somebody on one's shoulders and also as simply carrying somebody without specifying the manner; *collportar* can only have the former reading.

In short, the examples from (30) to (33) could be accommodated under the S-M competition if very fine-grained differences in meaning between syntactically and morphologically derived structures were taken into account. In other words, the constraint given in (22) is not sufficient to suspend the competition between S and M. Leaving aside the formal differences between the two derivations, the second part of the constraint takes into consideration the argument structure of the predicate. If that is the same in the two structures, competition will establish the syntactic derivation as the winner. AN are aware that by the constraint alone, the existence of some morphological alternants cannot be explained. They are then forced to assume that, despite having the same merger of categories and the same argument structure, some syntactic and morphological counterparts can coexist because, for example, the syntactic structure is interpreted literally while the morphological one figuratively[13] or because, despite hav-

---

[13] AN (2004: 84) exemplify the coexistence of syntactic and morphological derivations by giving some verb-particle constructions in Swedish (from Holmes & Hinchliffe, 1994: 321). The syntactic derivation is interpreted literally while the morphological structure is interpreted figuratively.

(i) Jag bryter av kvisten.
   I break off branch+DET
   'I break off the branch'
(ii) Jag avbryter samtalet.
   I off+break conversation+DET
   'I interrupt the conversation'

ing the same exact meaning, the morphological structure is used for official documents while its syntactic counterpart is used for more informal situations.

The fact of having unpredictable semantics (i.e. a figurative interpretation) and being used in formal situations is relevant enough for AN to allow morphological derivations to be listed in the lexicon, thus explaining their existence. The association of a figurative, and even lexicalised meaning, to the morphological structure and a transparent meaning to its syntactic counterpart can explain most of the cases in Catalan (e.g. *corferir*, 'to break somebody's heart', cf. (11a) vs. *ferir el cor*, 'hurt the heart', 'to physically hurt somebody's heart', cf. (11a)), but not all of them. For example, the fine distinctions in meaning just discussed can only be partially applied to (32). *Sangglaçar* (32b) can only have a figurative meaning, which would allow morphological specification in the lexicon. However, in addition to the literal reading, *glaçar la sang* (32a) can also be interpreted figuratively. Furthermore, notice that, in order to be able to fit the other Catalan examples into the M-S competition, even finer semantic distinctions between the two derivations are necessary. In other words, the semantics of, for example, inalienability (cf. (30)) and agent vs. experiencer readings (cf. (31)) should count as relevant enough to allow listing of the morphological structures, which would crucially allow their existence. That is, features like inalienability should be taken as idiosyncratic, thus specifying morphological realization.

In conclusion, the notion of semantics needs to be refined in AN's theory. One should know to which extent the difference in semantics the competition analysis refers to is. Otherwise, the theory cannot properly predict which morphological structures are allowed in the language.

4.2. Speaker's use of [NV]$_V$ compounds

Despite having concluded that a better understanding and definition of "semantics" should be provided for the M-S competition theory to properly make predictions, one cannot entirely discard one theory by just a few potential counterexamples. In fact, I consider that [NV]$_V$ compounding is not a productive phenomenon in the language, although some authors, such as Gavarró (1990), GF (1999, 2000), have argued against such a view. Notice that they base their conclusion on written corpora, and not on the actual use of the compound. Given the definition of productivity I adopt, their conclusion is not valid. I take the general assumption about productivity and apply it to compounding. Consequently, I understand that a compound type is productive when native speakers can understand novel combinations of the compound and can create new combinations themselves. By checking lists of [NV]$_V$ compounds with native speakers of Catalan, I found out that transparent N-V compounds can be understood but are very rarely produced (e.g. *corbategar* heart+beat) and that more opaque N-V compounds are often correctly understood but are rarely produced (e.g. *coferir* heart+hurt 'to break somebody's heart'). As already mentioned, opaque compounds like *sangglaçar* in (32) or *capgirar* in (34) do not constitute a problem for the morphosyntactic competition. Recall that the metaphorical reading associated with the morphological structure explains its existence, given that the unpredictable semantics allows listing in the lexicon.

(34) a. Avui   he   girat   el matalàs de cap
          Today   have turned   the mattress of head
          'Today I have turned the mattress from top/head to bottom/foot'
      b. Avui   he   capgirat   l'habitació   per trobar el llibre que havia perdut
          Today   have head+turned the+bedroom   for find   the book   that had lost
          'I have messed up my bedroom to find the book I had lost'

These findings confirm the results in Adelman's (2002) study of verbal compounds in Catalan. She concludes that there is a higher likelihood of using the compounds in written form than in everyday conversation, and that younger speakers would use fewer verbal compounds than older speakers both in written and spoken form. She hypothesises that younger speakers associate these forms with older speakers or an academic lexicon. Although Adelman's study is about the likelihood of using one form and not about its actual use, the results are still revealing. Nine informants aged 15-19 are 3% likely to use transparent compounds in conversation. Due to the low results regarding the usage of Catalan N-V

compounds, this compounding process cannot be considered active or productive in the language. It is possibly a potentially unstable construction in the language.[14]

If these forms were once productive, now they have fallen out of use, and possibly the ones which are still currently being used by some people are mostly being reanalysed as lexical items containing only one root. The two morphemes are no longer transparent in most of the cases. In fact, it is unclear whether all these words once really existed and were produced productively, or have simply been 'created/derived' (and stored in dictionaries) from their corresponding participial forms. Some examples favour this hypothesis. There are some cases in which the participial form (35a', b') is more common than the V (35a, b) and cases in which the participial form is the only existing form (36a vs. a', 36b vs. b') (The symbol *!* in front of a word indicates that the word is a non-existing form in the language) (cf. Badia, 1962; Mascaró, 1986 for Catalan; Booij, 2006 for Dutch).[15]

(35) a. caragirar
          head+turn
          'to change one's opinion'
      a'. caragirat
          head+turned
          'a traitor'
      b. corsecar
          heart+dry
          'to dry out somebody's life'
      b'. corsecat
          heart+dried
          'an embittered person'

(36) a. !esmaperdre
          mood+lose
          'to lose heart'
      a'. esmaperdut
          mood+lost
          'disconcerted'
      b. !llampferir
          lightning+strike
          'to strike by lightning'
      b'. llampferit
          lightning+struck
          'struck by lightning'

In addition, novel participial As can be created more easily than their corresponding infinitival verbs:

---

[14] There is a lot of discussion around the source of the internal word order of these compounds. Some authors (cf. Moll, 1952; Gavarró, 1990; Cabré, 1994) have attributed the NV order to Latin syntax (see e.g. Pinkster, 1990 for a discussion about word order in Latin), while others (cf. GF, 1999, 2000) have argued against this position on the basis that it is not plausible that Latin word order has remained constant only in this morphological construction and been lost in S entirely. I leave this question open for further research.

[15] The contrast found between the participial and infinitival forms suggests that an analysis like that of GF (1999, 2000; cf. footnote 2) in which the N is interpreted as a modifier, and not as an argument, of the verbal base is plausible. Further research may confirm the speculation according to which infinitival forms are derived from their participial forms.

(37) a. ?ullinflar
             eye+swell
    a'. ullinflat
             eye+swollen
    b. ?ulltòrcer
             eye+twist
    b'. ulltort
             eye+twisted
    c. ?culbonyegar
             bottom+lump
    c'. culbonyegut
             bottom+lumped

In short, the S-M competition cannot be really tested due to the low productivity of the [NV]$_V$ type of compound. Other types of compounds should be taken into account.

5. Conclusions

In this paper it has been argued that Catalan synthetic compounds of the type [NV]$_V$ are not the result of syntactic head-to-head movement (contra Brunelli, 2003), but are rather generated by an independent morphological system.

Regarding the competition analysis between S and M (cf. AN, 2004), it has not really been tested. Although the Catalan examples in Section 4.1 illustrate that the morphosyntactic competition needs to explain what it is understood by morphologically and syntactically derived structures having different semantics, I conclude that, because of the low presence of [NV]$_V$ compounds in Catalan, other types of compounds should be taken into account to investigate whether the S-M competition theory can be maintained with slight amendments or has to be rejected.

Finally, I would like to mention some questions which have left unresolved in the present paper and which I hope to take up in future work. One of them is to examine the reason for the current low productivity of Catalan [NV]$_V$ compounds and investigate whether they were really once productive in the language. A related question would be to make a study of the compound type under investigation across languages and investigate whether they illustrate the same phenomenon. The [NV]$_V$ has been attested in other Romance languages like Spanish (e.g. *maniatar*, hand+tie, cf. Rainer & Varela, 1992) and Aude (e.g. *gorjobadà*, throat+open, cf. Klingebiel, 1988) and in Germanic languages like English (e.g. *fingerprint*, cf. Rice & Prideaux, 1991).

References

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

Ackema, Peter and Ad Neeleman. 2007. Morphology does not equal syntax. To appear in G. Ramchand and C. Reiss, eds. *The Handbook of Linguistic Interfaces*. Oxford: Oxford University Press.

Adelman, Allison. 2002. *On the Status of N-V Verbal Compounds in Catalan*. Senior Honor Thesis, Swarthmore College.

Badia, Antoni M. 1962. *Gramática catalana*. Madrid: Gredos.

Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Change*, Chicago / London: University of Chicago Press.

Booij, Geert. 2006. Construction morphology and the lexicon. Paper presented at Forum de Morphologie – 5e Décembrettes, University of Toulouse, 7-8 December.

Brunelli, Michele. 2003. *Un grup especial de verbs compostos en català*. Tesi di Laurea. Venezia: Università Ca' Foscari.

Cabré, Maria Teresa. 1994. *A l'entorn de la paraula*. València: Servei de Publicacions de la Universitat de València.

Cabré, Maria Teresa and Gemma Rigau. 1986. *Lexicologia i Semàntica.* Barcelona: Enciclopèdia Catalana.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

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

Gavarró, Anna. 1990. *Syntactic Theory and the Grammar of Catalan Compounds*. Doctoral Dissertation, University of Edinburgh.

Gràcia, Lluïsa and Olga Fullana. 1999. On Catalan verbal compounds. *Probus* 11:239-261.

Gràcia, Lluïsa and Olga Fullana. 2000. Catalan verbal compounds: Internal order and argument interpretation. In W. Dressler, O.E. Pfeiffer, M. Pöchtrager and J.R. Rennison, eds. *Morphological Analysis in Comparison*. Amsterdam: John Benjamins. 75-87.

Hale, Kenneth and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In K. Hale, S. J. Keyser, eds. *A View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press. 53-109.

Hale, Kenneth and Samuel Jay Keyser. 1998. The basic elements of argument structure. In H. Harley, ed. *MIT Working Papers in Linguistics 32, Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*. Cambridge, Mass.: MIT Press. 73-118.

Hale, Kenneth and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge, Mass.: MIT Press.

Harley, Heidi. 2002. Why one head is better than two: Head-movement and compounding as consequences of Merge in Bare Phrase Structure. Paper presented in the Arizona Linguistics Colloquium Series, University of Arizona.

Harley, Heidi. 2003. Merge, conflation, and head movement: The First Sister Principle revisited. Paper presented at NELS 34, SUNY Stony Brook.

Holmes, Philip and Ian Hinchliffe. 1994. *Swedish: A Comprehensive Grammar*. London: Routledge.

Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Cambridge, Mass.: MIT Press.

Kiparsky, Paul. 1997. Remarks on denominal verbs. In A. Alsina, J. Bresnan and P. Sells, eds. *Complex Predicates*. Stanford, CA: CSLI Publications. 473-499.

Klingebiel, Kathryn. 1988. New compounds from the old: an unexpected source of verb+noun compounds in Romance. *Berkeley Linguistic Society* 14: 88-99.

Lieber, Rochelle. 2003. Compound Interpretation: Lexical Semantics, not syntax. In G. Booij, J. DeCesaris, A. Ralli and S. Scalise, eds. *Topics in Morphology: Selected Papers from the Third Mediterranean Morphology Meeting*. Barcelona: IULA. 241-253.

Mascaró, Joan. 1986. *Morfologia Catalana*. Barcelona: Enciclopèdia Catalana.

Moll, Francesc de B. 1952. *Gramática histórica catalana*. Madrid: Gredos.

Pinkster, Harm. 1990. *Latin Syntax and Semantics*. London: Routledge.

Rainer, Franz and Soledad Varela. 1992. Compounding in Spanish. *Rivista di linguistica* 4:117-142.

Rice, Sally and Gary Prideaux. 1991. Event-packing: the case of object incorporation in English. *Berkeley Linguistic Society* 17:283-298.

Rizzi, Luigi. 2004. On the status of the language faculty: results, developments, and perspectives. *The Linguistic Review* 21:323-344.

Roeper, Thomas and Muffy Siegel. 1978. A lexical transformation for verbal compounds. *Linguistic Inquiry* 9:99-260.

Selkirk, Elisabeth. 1982. *The Syntax of Words*. Cambridge, Mass.: MIT Press.

Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse

edited by Fabio Montermini, Gilles Boyé, and Nabil Hathout

Cascadilla Proceedings Project   Somerville, MA   2007

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Padrosa Trias, Susanna. 2007. Catalan Verbal Compounds and the Syntax-Morphology Competition. In *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*, ed. Fabio Montermini, Gilles Boyé, and Nabil Hathout, 91-107. Somerville, MA: Cascadilla Proceedings Project.

or:

Padrosa Trias, Susanna. 2007. Catalan Verbal Compounds and the Syntax-Morphology Competition. In *Selected Proceedings of the 5th Décembrettes: Morphology in Toulouse*, ed. Fabio Montermini, Gilles Boyé, and Nabil Hathout, 91-107. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1618.