

Janet DeCesaris - e.mail: janet.decesaris@trad.upf.es  
Maria Rosa Bayà - e.mail: rosa.baya@trad.upf.es  
Elisenda Bernal - e.mail: elisenda.bernal@trad.upf.es  
Institut Universitari de Lingüística Aplicada - Universitat Pompeu Fabra (Spain)

## NOUN-NOUN SEQUENCES IN CATALAN AND SPANISH

**Abstract:** This paper examines noun-noun sequences in Catalan and Spanish in which the second, non-head noun appears to be modifying the first, head noun. It has been previously claimed that most of these sequences correspond to a specific semantic class, but the examples show that such a restriction cannot account for the attested data. The prototypical interpretation for these sequences involves assigning a figurative sense to the non-head noun. It is suggested these sequences are generated in the syntax, thus expanding the nature of noun phrase structure in these languages.

### 0. Introduction

Compounding is at the heart of much research in morphology because it lies at the intersection of two components, one of which is responsible for forming words and the other of which is responsible for forming phrases and sentences. In languages like Catalan and Spanish, it has been claimed (Rainer and Varela 1992) that there are not always observable differences between the morphosyntactic behavior of lexical compounds and that of syntactic strings, and therefore crucial arguments in favor of classifying a sequence as a compound as opposed to as a syntactic phrase are often based on semantics, and many linguists feel uncomfortable with classifications based on claims on what a sequence means as opposed to its formal behavior.<sup>1</sup> Exactly what constitutes a compound in any given language is thus not always straightforward. This is the case for English, if we consider that many of the strings claimed to be compounds by Lieber (1992) have been argued to not be compounds but rather be mere quotations of syntactic strings (cf. Weise 1996).<sup>2</sup> The characteristics of compounds in Romance languages, particularly noun-noun compounds, are even less well defined, to judge

---

<sup>1</sup> We are not endorsing this lack of confidence in semantically based analyses, but rather stating what to us seems to be a fact in morphological analysis: linguists prefer to base claims for lexical compounds as opposed to syntactic phrases on easily quantifiable behavior such as pluralization, type of complement allowed, relationship to other word-formation processes, etc. rather than on assertions of meaning, which are more open to differing interpretations.

<sup>2</sup> We cite Weise (1996) because he specifically discusses forms claimed to be compounds by Lieber, but take this opportunity to note that the idea of quoting syntactic strings and inserting that information into a context did not originate with Spencer (1991) cites Bauer (1988) on this, and it is unfortunate that Weise makes no mention of these two well-known, easily accessible sources.

from the discussion in the literature (Di Sciullo and Williams (1987), Corbin (1994), and several articles in the 1992 issue of *Rivista di Linguistica* edited by Scalise).

This paper examines several types of noun-noun sequences in Catalan and Spanish to see if they should be analyzed as lexical compounds or as syntactic phrases. If it were to be shown that the syntaxes of Catalan and Spanish produce noun-noun sequences in which the non-head noun acts as a modifier of the head noun, some traditionally accepted assumptions about possible noun phrase structure and agreement processes within noun phrases may need revision. Before taking that step, however, we first need to establish the range of possible noun-noun sequences in these two languages. In order to frame our analysis, we first discuss the analyses of Spanish noun-noun compounds (Rainer and Varela 1992) and French noun-noun compounds (Zwanenburg 1992). Our discussion concentrates on two types of noun-noun sequence: sequences Rainer and Varela call "subordinative compounds,"<sup>3</sup> which are exemplified by *ciudad dormitorio* 'commuter town' or *empresa fantasma* 'ghostlike company, fly-by-night firm,' and sequences in which the second noun is a proper noun such as *cuervo Danone* 'Dannon body' or *madre Nutribén* 'Nutriben mother'. These two types of sequences were chosen because specific claims about their structure have been made in the past for Spanish and French, respectively, yet it is our opinion that these claims merit further scrutiny because they do not seem to account for many of the noun-noun strings commonly found in current prose.

None of the authors cited above specifically dealt with Catalan data, so it might be argued at the outset that any comparison between Spanish and Catalan or French and Catalan is unfair and simply not worthy of academic discussion (this would not apply to comparison of Rainer and Varela's analysis of Spanish data and our discussion of Spanish). In terms of compounding, however, we believe that parallelisms between Romance languages are not uncommon, and in fact references to "compounding in Romance" as opposed to "compounding in Germanic" abound in the literature.<sup>4</sup> It is therefore not unreasonable for us to assume that the basic characteristics of compounding in Catalan could, at the very least, be similar to those of Spanish and/or French, and as such we entertain these proposals while being aware of their possible limitations.

### 1. Data under consideration

Even a cursory glance at current prose in Catalan and Spanish yields numerous examples of noun-noun sequences that are candidates for compound status:

---

<sup>3</sup> Throughout this paper we will use their terminology in order to aid the reader by not introducing our own terminology, although in our view *subordinate* (contrasting with *coordinate*, and not *coordinative*) would have been a more felicitous choice of words.

<sup>4</sup> The following passage from Zwanenburg (1992: 222) illustrates this well: "Another property which opposes Romance compounding to Germanic compounding is that many compounds resemble syntactic phrases."



In fact, one of the common threads underlying his discourse is that most of these forms are unnatural in the language in some sense and should be replaced by more Spanishlike formations. He is probably correct in assuming that many are loan translations from English, but the fact remains that modern Spanish has been able to incorporate these strings to such extent that the structure is productive.

Noun-noun sequences take the gender marking of the left noun. Plural formation of these sequences varies, not only with the string but also with the speaker. As Rainer and Varela note for Spanish (1992: 126), plural marking is usually only found on the left noun, but can occasionally also appear on the right noun. This is also the case for Catalan; e.g. conservative speakers will say *problemes clau* 'key/main problems,' but all three of us have heard, both on the radio and in conversations we ourselves have had, *problemes claus*.<sup>5</sup>

## 2. Rainer and Varela's discussion of the data

Rainer and Varela point out that "left-headed n-n compounds and restrictive appositions" are the hardest sequences in Spanish to classify as either compounds or phrases. They provide examples of nine types of noun-noun sequences, and suggest that "the hard core of n-n compounds is thus constituted by the coordinative type *cantante autor* ['singer-songwriter'] and the subordinative type *ciudad dormitorio* ['bedroom community'], while the borders both towards apposition and adjective phrases seem to be somewhat fuzzy" (1992: 120). They go on to claim that contrary to what occurs in Germanic, Spanish compounds of this sort "in productive use only follow a restricted set of semantic types" (1992: 126), although they do not identify what exactly defines these semantic types.

Let us consider their idea in more detail, particularly with reference to other word-formation processes in Spanish and Catalan. Obviously, semantics plays an important role in determining whether a given formative can actually combine with a given stem in derivation. There are surely semantic restrictions on compounding as well, so that a combination like Sp. *ciudad cantante* / Cat. *ciutat cantant* is not likely to exist, because in the real world that we refer to we know of no 'singer cities' nor are we likely to require this reference. But this does not seem to be the type of semantic restriction Rainer and Varela are alluding to. Rather, they state that the words *modelo* ('model'), *estrella* ('star'), *límite* ('limit'), *fantasma* ('ghost'), and *clave* ('key') form a semantic class of nouns which lends itself to an adjectival interpretation, and thus these forms are beginning to act like adjectives in terms of pluralization and predicative use. It is not clear to us exactly what this semantic class is or how it can be described. We note that even if this semantic class existed, it would be unique in terms of word-formation in Catalan and Spanish. Restrictions on word-formation can be based on argument structure or semantic features such as [animacy], but we know of no semantic feature grouping together the nouns in question. Moreover, the paraphrase offered by Rainer and Varela, "an  $n_1$  which is like an  $n_2$ " (1992: 126) is too general to effectively

---

<sup>5</sup> We note that Rainer and Varela have the same intuitions about *clave* in Spanish.

identify a semantic class. Rainer and Varela's comments seem to be based on the idea that these are the only (or practically only) nouns used productively in this sort of construction, but that is not the case, as witnessed by the examples given above.

### 3. The semantics of subordinative compounds

Although we are not convinced that there is a single semantic feature or set of features shared by nouns appearing in non-head position, it does seem to be the case that these nouns are typically not interpreted literally but rather figuratively. Let us consider the following examples:

- (3) a. el plan/programa/proyecto piloto 'pilot plan/program/project'  
 b. el paper moneda 'bill' (lit. paper coin)

The only possible way to interpret *piloto* is figuratively, as neither plans nor programs nor projects can be pilots because pilots must be animate. Likewise, paper is not like coins in a literal sense but rather serves the same purpose as coins do.

The behaviour of the noun *bomba* 'bomb' is particularly interesting. In both languages, *bomba* combines with other nouns to refer to the bomb's packaging:

- (4) cotxe bomba (Catalan) / coche bomba (Spanish) 'car bomb'  
 paquet bomba / paquete bomba 'letter bomb'

Although Rainer and Varela analyze *coche bomba* as a subordinative compound, we suggest that the literal interpretation given to both nouns would put it on a par with *sofà cama*: i.e., in our view it is a coordinative compound. *Bomba*, however, can also be understood figuratively:

- (5) Els preus bomba (Catalan) 'shockingly low prices'  
 Una noticia bomba (Spanish) 'shocking news item'

If semantics is the basis for comparison of coordinative vs. subordinative compounds, then *preu bomba* must be different in structure from *cotxe bomba* because *bomba* is not understood in the same fashion.

Another interesting set of examples are those listed in (6), taken from a document on pharmaceutical procedures written in Spanish.

- (6) a. la solución patrón 'standard reference solution'  
 b. la solución problema 'test solution'  
 c. la solución reactivo 'reagent solution'

The sequences in (6a) and (6b) can be analyzed as subordinative compounds in that the solution serves as the standard reference or as the test case being run through the equipment. Neither (6a) nor (6b) appears to have arisen as a result of preposition

deletion, as the corresponding phrases *\*solución de(l) patrón/problema* do not occur.<sup>6</sup>

The phrase in (6c), however, may be related to *la solución del reactivo*, with deletion of the preposition *de* 'of' and of the definite article.

Preposition deletion has not received much attention in Spanish or Catalan. Rainer and Varela mention that forms which seem to have lost a preposition, such as *tren mercancías* or *moto todo terreno* (their examples) probably do not qualify as lexical structures but rather as syntactic strings, and we would agree. Examples of noun-noun sequences that appear to be missing a preposition are commonplace in advertising, where they often coexist with the full forms:

(7) Examples taken from hardware store catalogue in Catalan, summer 1996

- |                    |                      |                    |
|--------------------|----------------------|--------------------|
| a. Both forms:     | ventilador sostre    | 'ceiling fan'      |
|                    | ventilador de sostre | 'ceiling fan'      |
| b. No preposition: | bústies jardí        | 'garden mailboxes' |
|                    | mobles jardí         | 'garden furniture' |
|                    | carret jardí         | 'garden cart'      |
|                    | dutxa jardí          | 'outdoor shower'   |
|                    | taules resina        | 'resin tables'     |
|                    | taula ordinador      | 'computer desk'    |

In an advertising context, in which a speaker's knowledge of real world references is all-important, the semantic information contributed by the preposition is easily obtained without the preposition itself. It is what we know about the world, and not some characteristic of noun-noun sequences, that allows us to interpret *taula resina* as a table made out of a certain substance yet *taula ordinador* as a table designed to hold a computer while in use. That explains why the specific preposition that can be deleted is not always *de* 'of', which is surely most often used to relate nouns, but can also be Catalan *per a* 'for', *amb* 'with', Spanish *por*, *para* 'for', *con* 'with', etc. The semantics of these noun-noun sequences is thus crucially dependent on external information, whereas the semantics of true subordinative noun-noun sequences is more dependent upon a non-literal interpretation of the non-head noun.

#### 4. Noun-noun sequences with proper names

A subset of subordinative noun-noun sequences are those strings in which the righthand noun is a proper name:

- |     |                    |                        |
|-----|--------------------|------------------------|
| (8) | una cafetera Bosch | 'Bosch coffeemaker'    |
|     | ventanas Velux     | 'Velux-brand windows'  |
|     | la garantía Nissan | 'the Nissan guarantee' |

Zwanenburg, in his discussion of the parallel forms in French, suggests that this structure is used for name-giving and then becomes more or less fixed in meaning

---

<sup>6</sup> *La solución del problema* would be interpreted as the answer to the problem.

(1992: 225). It is certainly true that this kind of structure can become lexicalized, as in the Spanish phrase *cuerpo Danone* 'body by Dannon Yogurt', which is not used to describe the bodies of yogurt-eating individuals but rather to describe a young, physically fit person, but the great majority of these strings do not exhibit this behavior. The prototypical interpretation for these strings is 'made, manufactured, produced by...', but again, the role of what a speaker knows about the real world is important in being able to interpret the phrase. We may compare two phrases:

- (9) a. madre Nutribén 'a Nutriben mother, a mother who uses Nutriben products'  
b. potito Nutribén 'a Nutriben jar of baby food'

The possible relationship between the nouns in (9a) cannot be the same as that in (9b) because as people we know that mothers are not made by the manufacturers of commercially prepared baby food. Unlike the subordinative sequences discussed above, however, the non-head noun *Nutribén* is not understood figuratively. Rather, this seems to be a syntactic structure in which a commercial company is related to a head noun. This structure probably originated from one with the preposition *de* 'of', but can no longer only be explained by preposition deletion precisely because of examples like (9a), as *\*una madre de Nutribén* is ungrammatical. Nor does the structure seem to display the semantic unity that is typically claimed for lexical compounds. We therefore suggest that these strings are the result of a syntactic structure combining two nouns.

##### 5. Lexical vs. syntactic origin of compounds

As we stated at the outset, one of the major debates in morphology addresses whether compounds are lexical or syntactic in nature. Both views have eloquent defenders: Lieber (1992) claims compounds are syntactic because all morphological processes are syntactic; some years earlier Di Sciullo and Williams (1987) argued that what had previously been claimed to be compounds in French were lexicalized syntactic phrases listed in the lexicon and not the output of word-formation in the lexicon, which is basically the view expounded in Zwanenburg (1992). Alternatively, a recurring theme in the work of linguists like Carstairs-McCarthy and Spencer is that word-formation is just that –word-formation– and not sentence formation. Choosing one approach or the other depends on one's overall view of what the organization of a grammar should be, what generalizations one wants to make in that grammar, and what tools one wants to use or has available to make those generalizations. Even for a single structure, those decisions can only be made when a substantial variety of word-formation structures have been closely examined. Since this paper has only dealt with a small number of examples of a few structures, it would be rather presumptuous to make any far-reaching claims based on such little data. We can suggest, however, whether the types of noun-noun sequences discussed should be generated by the word-formation component or by the syntax. To date, many of the arguments for a syntactic origin entail lexicalization of a set phrase, or at least a phrase naming a specific reference. We believe these two issues need to be separated from one another. The productive use of

nouns in the non-head position would seem to simply that lexicalization of a set phrase is not necessary for a noun to be used as a noun modifier. Strings like *piso muestra* (literally 'model apartment, apartment sample') or *preu bomba* are not lexicalized in our view, yet we think they should be generated in the syntax as opposed to in the word-formation component. This, in turn, means that nouns can act as modifiers of other nouns in Catalan and Spanish, which in effect expands the possible type of noun phrase structure usually assumed for these languages. The acceptability of these structures in the modern languages is probably aided by external factors such as increased influence from English, especially in journalistic prose, and an advertising style in which recoverable information is deleted to save space, as well as by the existence of other noun-noun sequences resulting from apposition. These syntactic phrases may lexicalize, as in *fecha límite / data límit* 'deadline', but it is equally as possible for the non-head noun to become increasingly more adjectivelike, as is the case with *clave / clau* 'key'.

## 6. Concluding remarks

The analysis outlined in this paper expands the roles nouns play in Catalan and Spanish. If this approach is correct, then we should expect to find nouns in other non-prototypical contexts, and we do:

- (9) Spanish nouns in adverbial contexts:
- |  |                             |
|--|-----------------------------|
| Me lo pasé <i>muy bien / bomba</i> .             | 'I had a blast.'            |
| No hagas nada <i>bestia</i> (lit. 'beast').      | 'Don't do anything stupid.' |
| Les gustó <i>una barbaridad</i> (lit. 'horror'). | 'They like it a lot.'       |
- Catalan nouns in adverbial contexts:
- |   |                                 |
|---|---------------------------------|
| No m'agrada <i>gota</i> (lit. 'drop').                | 'I don't like it at all.'       |
| Li van impressionar <i>quantitat</i> (lit. 'amount'). | 'They impressed him/her a lot.' |

The alternative, considering subordinative noun-noun sequences compounds, in our view would open the flood gates as to what constitutes a compound in these languages. That in itself is not argument against it, but lexical origin seems ill-equipped to account for proper nouns as non-head nouns, and if those noun-noun sequences are to be treated like other subordinative sequences –and we think they should– then a syntactic approach is preferable.

Noun-noun sequences in Catalan and Spanish are interesting proving ground for morphological analysis. We hope to continue gathering data on potential compound structures, and specifically on the semantics involved in subordinative noun-noun sequences, not only to achieve a better understanding of how these two languages work, but also to shed light on compounding and word formation in Romance. For now, we hope to have shown that the range of semantic relationships between the two nouns is wider than previously thought.

## References

- Bauer, L. 1988. *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.
- Corbin, D. 1994. "Locutions, composés, unités polylexématiques: lexicalisation et mode de construction". *Actes du colloque «La locution, entre lexicque, syntaxe et pragmatique*. Paris - Saint Cloud.
- Di Sciullo, A. M.; Williams, E. 1987. *On the definition of word*. Cambridge, MA: MIT Press.
- Lieber, R. 1992. *Deconstructing Morphology*. Chicago: University of Chicago Press.
- Martínez de Sousa, J. 1996. *Diccionario de usos y dudas del español actual*. Barcelona: VOX-Biblograf.
- Rainer, F.; Varela, S. 1992. "Compounding in Spanish". *Rivista di Linguistica*, 4/1, 117-142.
- Scalise, S. (ed.) 1992. *The Morphology of Compounding*. Special issue of *Rivista di Linguistica*, 3/1. Turin: Rosenberg & Sellier.
- Spencer, A. 1991. *Introduction to Morphological Theory*. Oxford: Blackwell.
- Weise, R. 1996. "Phrasal Compounds and the Theory of Word Syntax". *Linguistic Inquiry*, 27/1, 183-193.
- Zwanenbourg, W. 1992. "Compounding in French". *Rivista di Linguistica*, 4/1, 221-240.