

Compounding in Spanish

Franz Rainer & Soledad Varela

This paper contains a brief description of compounding in present-day Spanish. In addition to the traditional criteria of conceptual unity, syntactic islandhood and the status of free forms of all the immediate constituents we have had recourse, in our delimitation of the subject matter (1.1.), to the criterion of system adequacy. This has led us, on the one hand, to include some constructions with bound morphemes, and, on the other, to exclude some consisting only of free immediate constituents. So-called "syntagmatic compounds" have been excluded as well, since we think it is preferable to treat them as lexicalized phrases in order to avoid the duplication of syntactic rules in the lexicon. Section 2 describes the seven productive compound types of Spanish, as well as seven more or less unproductive ones, paying attention to their stress pattern, allomorphy, syntactic category, gender, pluralization, restrictions, meaning and argument structure. For reasons justified in 1.2., we have opted for specific rules of compounding as a descriptive framework. In 3.1., the ordering of compounding with respect to derivation and inflection is investigated, leading us to the conclusion that extrinsic transitive ordering might not be the right conceptual tool for capturing the existing asymmetries, some of which, at least, can be accounted for adequately by considering the meaning and restrictions of the rules involved as well as the role of lexicalization, and by allowing rules to operate on the head of their bases. In the final subsection (3.2.), it is suggested that the sharp contrast with respect to recursiveness between Germanic and Spanish compounds might also be attributable to independent factors, rather than to specific constraints.

1. Introduction.

1.1. Problems of definition and delimitation.

1.1.1. *The definition of compound.* As many other venerable concepts of linguistics, the notion of a compound seems well-defined for the prototypical instances but less so for a great many borderline cases, both on the phrasal and on the derivational side. Among the criteria traditionally considered distinctive of compoundhood the following three are the most prominent:

- (1) a. Compounds denote a unitary concept
- b. Compounds are "syntactic islands"
- c. Compounds are made up of free forms

Of these three criteria, the first two, which extend to derivatives and partly to idioms, are intended to set compounds apart from phrases, which also obey criterion (1c). By "syntactic islandhood" one generally refers to the inseparability of compounds and the inaccessibility of their constituents for syntactic rules. The third criterion, instead, is meant to set compounds apart from derivatives and, in our understanding, also from "neo-classical compounds".

As is well known, these criteria are not sufficient for a successful classification of non-prototypical cases. In order to come to grips with these borderline cases, we will have recourse to the additional criterion of system-adequacy (cf. Wurzel 1984) which will allow us both to include types which do not conform to all these criteria and exclude others which do present properties (1a) to (1c), if this buys us a classification more in line with the general architecture of Spanish word-formation. So, e.g., formations of the type *franco-español* ('French Spanish') will be considered compounds because of their perfect parallelism with a-a compounds¹ of the type *palestino-israelí* ('Palestinian Israeli') despite the boundness of the first constituent *franco*, while on the other hand formations like *sobrevolar* ('overfly') will be classified as derivatives due to their resemblance with indisputable derivatives like *sobreactuar* ('overact') and the absence of a well-established type of preposition-verb compounds in Spanish, although they respect all three criteria.

1.1.2. Problems of delimitation.

1.1.2.1. *Compounds vs. phrases.* The border between compounding and syntax poses quite intriguing problems in Spanish, problems which are far from being solved. The following remarks should therefore be considered of a rather tentative nature.

The most difficult of all is the distinction between left-headed n-n compounds and restrictive appositions. As shown in table 1, Spanish has a great variety of sequences n + n, some of which are surely phrasal, while others with an equal degree of certainty represent compounds. What remains unclear is where exactly to draw the dividing line between both types.

Table 1

	1	2	3	4	5	6	7	8	9
a	-	-	?	-	-	+	+	+	-
b	-	-	+	-	-	+	-	-	-
c	+	+	+	+	+	+	- ²	- ²	+
d	+	+	+	-	+	+	-	-	+
e	-	-	-	+	+	-	-	-	+

1. rey Juan Carlos ('king J.C.')

2. clérigo autor (de novelas) ('priest author (of novels)')

3. sector educación (lit. 'sector education', 'education sector')

4. hombre hombre (lit. 'man man', 'real man')

5. vestido violeta (lit. 'dress violet', 'violet dress')

6. tren mercancías (lit. 'train goods', 'freight train')

7. cantante autor (lit. 'singer author', 'songwriter-singer')

8. ciudad dormitorio (lit. 'town dormitory', 'bedroom-community')

9. problema clave (lit. 'problem key', 'key problem')

a. the two constituents form a conceptual unity

b. there is a feeling of a dropped preposition

c. the second constituent is modifiable by an adjective, adverb or prepositional phrase

d. the second constituent is coordinable

e. the second constituent semantically is very adjective-like

As the table shows, the criterion that most clearly circumscribes left-headed n-n compounds (6-8) is semantic unity, setting them apart both from the more phrasal sequences to its left and the *problema clave* type to its right, which comes very close to a sequence of n + a. In our understanding, in a sequence like *un clérigo poeta autor de miles de versos deleznables* (lit. 'a priest poet author of thousands of futile verses'), the first two nouns, *clérigo* and *poeta*, form a true compound, since both professions are presented as equally constitutive of the person referred to, while *autor de...* has the status of an apposition, as shown e.g. by the possibility of modifying it

² Modification here is possible only if the resulting second constituent is a lexical unit: *un restaurante-centro social* (lit. 'a restaurant social center'), *un comedor-sala de estar* (lit. 'a dining room - living room'), etc. This restriction does not hold for other n-n sequences; *bombre muy bombre* (lit. 'man very man'), *vestido violeta claro* (lit. 'dress violet light'), *moto todo terreno* (lit. 'motorcycle all terrain'), *problema absolutamente clave* (lit. 'problem absolutely key'), etc.

¹ The following abbreviations will be used throughout this paper: a = adjective, n = noun, v = verb, adv = adverb.

by a prepositional phrase. The type *fren mercancías* is very compound-like semantically, but differs by retaining some phrasal properties due, probably, to the fact that many of its members seem to have arisen by deletion of a preposition (cf. *gel ducha y baño* lit. 'gel shower and bath', *moto todo terreno* lit. 'motorcycle all terrain', etc.). The hard core of n-n compounds is thus constituted by the coordinative type *cantante autor* and the subordinative type *ciudad dormitorio*, while the borders both towards apposition and adjective phrases seem to be somewhat fuzzy.

Noun-phrases containing a prepositional phrase like *telón de acero* ('iron curtain', lit. 'curtain of steel') or an adjective (phrase) like *media luna* ('half moon') or *luna nueva* ('new moon', lit. 'moon new') are traditionally classified as "compuestos improprios" (improper compounds) in Spanish grammar if they denote a unitary concept. Such a classification is defensible if the definition of *compound* solely relies on (1a) and (1c). If, however, we take into consideration also (1b), we are forced to exclude such constructions from compounding, since their internal structure strictly obeys the rules of the Spanish noun-phrase. So, e.g., the adjectival type shows regular concord and follows the syntactic rule responsible for the position of the adjective. Considering such formations as compounds would lead to an inadmissible duplication of rules in the syntactic and lexical component. The naming function of such phrases can be attributed to the pragmatic component in a modular conception of grammar. This pragmatic rule, by the way, also applies to other linguistic units, from affixes (*los ismos*, etc.) to whole sentences (*el hágallo-usted-mismo* 'do-it-yourself', *el sálvese-quien-pueda* 'every man for himself!', etc.). Global pluralization, accentual unity and opacity to syntactic rules like adjectival/adverbial modification, properties which some of these formations — among them *media luna* — share with simple words, should be viewed as what they are, namely consequences of the naming function or lexicalization which one also finds in idioms.

If they denote useful concepts, phrases of this kind are stored in the mental lexicon as other complex signs are, thus becoming available for local analogical creations, such as *telón de bambú* ('bamboo curtain'), obviously coined after the model of *telón de acero* from which it inherits most of its idiosyncratic meaning elements. There is scarce evidence, however, that this availability for analogical coining might have led to morphological reanalysis by now. The best evidence we can find are some neologisms which do not follow the syntactic rule of adjective placement but rather some concrete model, such as *libre mercado* ('free market'), probably relying on *libre cambio* ('free exchange'), where the syntactic rule would place *libre* after the noun, or formations of the type *liberalcapitalismo*, *nacionalcatolicismo*, *socialrealismo*, *socialcatolicismo*, etc., which show the same anomaly.

In formations of the type *simergüenza* ('scountrel', lit. 'without shame'), consisting of a preposition and a noun, the change of category from PP to N might be adduced as an additional argument for including them into compounding. Since this change, however, is a natural consequence of ellipsis or exocentricity, nothing prevents us from treating them like the noun-phrases just analyzed. This is also supported by conceivable neologisms of a more phrase-like nature like *los sin una perra* (lit. 'the without a penny'). Cases which do not follow the regular syntax and semantics of the prepositional phrase, such as *entrechocar* ('to clash'), *sóbrehumano* ('superhuman'), *contraalmirante* ('rear admiral'), etc., should be treated as prefixations.

The same, finally, is true of adjective phrases with *bien* ('well'), *mal* ('badly, ill') and *siempre* ('always'), which are created productively only in syntax but can end up as lexicalized units which in some cases are close to prefixations (cf. *malcontento* = *descontento*, both meaning 'discontent', etc.).

1.1.2.2. *Compounds vs. derivatives*. In drawing the dividing line between compounding and derivation we will rely heavily on the criterion of system adequacy in the way indicated above.

Spanish has a number of formations of the following kind, which are often analyzed as right-headed compounds according to criterion (1c): *autoescuela* (lit. 'car school'), *cineclub* (lit. 'cinema club'), *ecosonda* (lit. 'echo sounder'), *farmacovigilancia* (lit. 'medicine vigilance'), *fotorreportero* (lit. 'photograph reporter'), *radioaficionado* (lit. 'radio fan'), *telenovela* (lit. 'television novel'), *videoarte* (lit. 'video art'), etc. Another reason to consider them compounds might be that the interpretation is pragmatically guided as in the Germanic compounds from which they originate. It is also possible, however, to make a case for considering them as derivatives. First, such right-headed n-n compounds would have to be limited to a fixed number of first elements (roughly those enumerated above), which is quite atypical, since compound types generally can be defined at the categorial level. This property would instead follow if we consider these first elements to be prefixes. Another property which they share with prefixes is their form: they are typically bisyllabic and end in a vowel. Substitution of words which do not share this canonical shape would normally lead to odd results: **camionescuela* (lit. 'truck school'), **baloncestoclub* (lit. 'basketball club'), **ciudadavigilancia* (lit. 'city vigilance'), etc. These first elements behave exactly like *aero-*, *agro-*, *bio-*, *cardio-*, *ciclo-*, *cosmo-*, *demo-*, *dígito-*, *drogo-*, *electro-*, *euro-*, *ferro-*, *filo-*, *fono-*, *geo-*, *gloso-*, *belio-*, *hemo-*, *hepato-*, *hidro-*, *histo-*, *moto-*, *narco-*, *neuro-*, *petro-*, *porno-*, *psico-*, *químico-*, *tecnico-*, *termo-*, *xero-*, *zoo-*, etc., which are clearly prefixal according to criterion (1c). It is this parallelism that, in our minds, warrants the inclusion of *auto-*, etc. among prefixes. The alternative would consist in setting up an independent

subtype of word-formation characterized by righthheadedness, a purely pragmatically driven rule of interpretation and severe restrictions on the number and form of elements that can serve as left constituents.

Some Spanish complex words include a first constituent coinciding with a freely occurring preposition formally and semantically, though not in terms of distribution. Such formations would qualify as compounds according to criteria (1a-c), but the additional criterion of system-adequacy leads us to include them into derivation. Concretely, the following considerations are relevant: (a) some of these first constituents have clearly prefixal synonyms (e.g. *sobrebundar* = *superabundar* 'to superabound'); (b) for many of them there exist formally identical prefixes as shown in table 2; (c) Spanish has very few right-headed compounds.

Table 2

"Prepositions"	homonymous "prefixes"
antecámara 'antechamber'	anteayer 'yesterday'
convivir 'to live together'	conmemorar 'to commemorate'
contraquerilla 'counter guerrilla'	contravenir 'to contravene'
empaquetar 'to pack'	enriquecer 'to enrich'
entrepelanta 'mezzanine'	entrebarrir 'to half-open'
sobrevolar 'to overfly'	sobrealimentar 'to overfeed'
trascoro 'retrochoir'	trasplantar 'to transplant'

Somewhat special problems of delimitation are presented by formations containing the adverbs *casicuasi* and *medio*. *Medio* ('half') has the peculiarity of showing up between the verb and the reflexive pronoun, a position otherwise reserved to prefixes: *se medio excusa* (lit. 'himself half he-excuses'), etc. *Cuasi* ('almost/quasi') in modern Spanish is confined to prefixal usage (*cuasi-delito* 'quasi-delict', *cuasi-sistemático*, etc.), while *casí* also freely occurs in syntax. Moreno (1984) claims that all uses of *casí* can be accounted for by one category-insensitive semantically based generalization. On the other hand, the parallelism with *cuasi* might also be taken as evidence that it has prefixal status when occurring inside a noun-phrase.

1.2. *Other basic concepts.* In this section we will introduce some additional concepts necessary for a description of Spanish compounding.

The head of a compound will be defined as that constituent which is a hyperonym of the whole complex word. Complex words where the head is one (or more) of the constituents, will be called *endocentric*, while complex words with no hyperonymic relationship of the kind specified will be called *exocentric*. This definition of endo and exocentricity is standard, while our definition of *head* is just one among a whole series proposed in the literature (cf. Zwicky 1985 and Bauer 1990 for a good discussion). The point of giving a privileged position to the hyperonymic constituent is that several generalizations about other properties of complex words can be best stated with the aid of the head-concept. So the head generally has the same

syntactic category and gender as the whole compound and may be the locus of morphological operations, i.e., operations referring specifically to the head of a word and not globally to the word as a whole, which will be seen to play an important role in Spanish morphology.

The *gender* of a Spanish complex word, however, does not necessarily coincide with the gender of its head. Exocentric compounds may take the gender of their hyperonym, just like abbreviations and loan-words. Since the head here is defined as the hyperonymic constituent of a complex word, head gender assignment could be viewed as a special case of hyperonymic gender assignment. Complex words which are not within the reach of the gender assignment strategies just discussed are assigned to the masculine, which is the default gender in Spanish.

The head-concept also plays a role in the syntactic rule of *adjectival agreement* which states that an adjective must agree with its head-noun in gender and number. Here, however, we are dealing with the syntactic notion of head, which is not co-extensive with the morphological notion. The other syntactic agreement rule, *verbal agreement*, states that the verbal head of a finite clause must agree with the subject in person and number.

Argument-structure is currently conceived of (cf. Pinker 1989, Grimshaw 1990) as a projection of the semantic representation, while the mapping between argument structure and syntactic structure is attributed to a special set of linking rules. In Grimshaw's system, the external argument of a verb, syntactically realized as the subject, is defined as the most "prominent" participant in the action or event denoted by that verb, e.g. the agent if there is an agent and a patient, the "theme" of intransitive movement verbs, etc. The other participants are projected onto internal arguments. If there is only one internal argument, this is generally realized syntactically as the direct object, e.g. the patient of a verb like *pegar* ('to hit'), etc.

According to the "Theta Criterion" of Chomsky (1981), all arguments have to be satisfied by exactly one word or phrase, if the resulting construction is to be well-formed. Satisfaction may be accomplished word-internally or externally, i.e. on the phrasal level. In *Juan es drogadicto*, e.g., the internal argument (*droga* 'drug') of *adicto* ('addict') is satisfied word-internally, while the external argument is satisfied syntactically by the subject *Juan*. Internal satisfaction, according to Grimshaw (1990), has precedence over external satisfaction, and if argument x is satisfied internally, no argument lower than x on the "thematic hierarchy" (cf. Randall 1984) may be satisfied externally. These somewhat rudimentary remarks about argument structure will be seen to be sufficient for a description of the behavior of Spanish compounds.

Compound rules are subject to *restrictions* just like any other word-formation rule. This fact, which may be somewhat less clear for essentially only pragmatically constrained types such as the Germanic n-n compounds, seems to us to be good reason to assume that compounding is done by *specific*

compound rules and cannot be handled by a combination of some general structure-building rule(s) in the manner of Lieber (1981) or Selkirk (1982) and lexical insertion, since such restrictions can be attributed neither to one of the constituents (as in affixation) nor to the structure-building rule(s), but only to specific compound rules. A case in point would be the compound type *pelirrojo* ('red-haired') (cf. 2.2.5.), whose quite peculiar restrictions are neither part of the lexical entry of *pelo* ('hair') nor of *rojo* ('red'), nor of a skeletal structure-building rule, but rather of this particular compound rule as a whole.

The sum of restrictions imposed on a particular rule delimits the domain of that rule. Quite frequently, the domain of a rule must be further divided into subdomains and *productivity*, in the sense of the readiness with which neologisms may be created, can vary across such subdomains. In the present survey, the description of the various restrictions will inevitably be somewhat sketchy. We would also like to mention that we have concentrated on productive types and ignored many obsolete compounds which only survive in treatises of morphology (e.g., Bustos 1986), and the Academic word cemetery.

Three aspects remain to be introduced with respect to the formal side of Spanish compounds.

There is no such thing as a unitary compound stress rule in Spanish. Rather, different compound types show different *stress patterns*: two main stresses, a secondary word-stress on the left constituent and a main stress on the right constituent or only one main stress and regular trochaic secondary stress. Furthermore, lexicalization or high frequency are often accompanied by *de stressing*, i.e., lexicalized/highly frequent words are simply treated as if they were monomorphemic. This accentual heterogeneity, by the way, is a further argument for working with individual compound rules, since stress seems to be rule-specific. We would like to stress, however, that the intricacies of Spanish compound stress are urgently in need of further investigation (despite Bustos' 1986 valid contributions), which may well somewhat alter the picture layed out in this article.

Another characteristic of many Spanish compounds is the presence of a *linking vowel -i-*. Again the presence of the linking vowel seems to be rule specific. It is convenient to distinguish a coordinative variant, occurring in coordinative compounds, clearly related to the homophonous conjunction *y* and a subordinative one occurring in right-headed compounds. The vowel to the left of the linking vowel inevitably drops.

Elision of a vowel before an identical vowel at the compound juncture is a typical instance of a "crazy" morphological rule. Although it can be limited to some compound rules, it is operative only sporadically within these rules and thus cannot be linked to the rule as a whole. Some other curious cases of deletion will be described in 3.1.2.

2. Types of compounds.

Spanish has a relatively poorly developed compounding system if compared with, say, the Germanic languages. Table 3 classifies the existing types according to the syntactic categories of the constituents.³ The letters in bold-faced type signal the head — exocentric types consequently show no bold types — and brackets, very low productivity.

Table 3

	n	v	a	adv
n-n:	<i>poeta-pintor</i>	(n-i-v): <i>maniatar</i>	(n-a): <i>drogadicto</i>	
n-n:	<i>año-luz</i>		n-i-a: <i>pelirrojo</i>	
n	(n-n): <i>castroestalinismo</i>			
n-n:	<i>nomoroeste</i>			
(n-i-n):	<i>sopicaldo</i>			
v	v-n: <i>sacacorchos</i>	(v-v): <i>bullebulle</i>		
a	(a-i-n): <i>altiplanicie</i>		a-a: <i>ético-moral</i>	
adv		(adv-v): <i>malgastar</i>	a-i-a: <i>rojiblanco</i>	

2.1. Nominal compounds.

2.1.1. *Coordinative n-n compounds.* The creation of coordinative n-n compounds is a highly productive process in Spanish, at least at the written level. The constituents mostly refer to social roles (*poeta-pintor* lit. 'poet painter'), actions (*compraventa* lit. 'buying selling'), ideologies (*marxismo-leninismo*) or places (*panadería-pastería* lit. 'bakery pastry'), but there seems to be no intrinsic limitation to these semantic fields. The reason why we have set apart this type from the subordinative n-n compounds is that they seem to contain two heads, as evidenced by the regular pluralization of both constituents: *poetas-pintores*, etc. Lexicalized compounds such as *compraventa* may show pluralization at the end only: *compraventas*. Gender, nevertheless, is determined by the left constituent in cases of conflict: *un diccionario_m-enciclopedia_f* (lit. 'a dictionary encyclopaedia'). Both constituents get main stress, though destressing can be observed in lexicalized cases (cf. *compraventa*).

³ Cardinal numbers, some of which are also compounded (e.g., *ciento dos* '102', from *ciento* '100' and *dos* '2'), have not been taken into consideration.

If both constituents coincide in their argument structure, the internal argument may be regularly realized by a phrase introduced by *de*: *escritor-traductor de novelas* (lit. 'writer translator of novels'). Second constituents without an argument structure, however, may block the realization of the internal argument of the first constituent: **escritor-maestro de novelas* (lit. 'writer teacher of novels'), ? *escritor-estrella de novelas* (lit. 'writer star of novels'), etc.

2.1.2. *Left-headed subordinative n-n compounds*. Left-headed subordinative n-n compounds show a great many semantic relationships, though only a few of them are really productive. The subtype paraphrasable by "an n_1 which is like an n_2 " is closest to the coordinative type: *ciudad dormitorio* (lit. 'town dormitory'), *empresa fantasma* (lit. 'company ghost'), etc. Other relationships are exemplified by *coche bomba* (lit. 'car bomb'), *fútbol-sala* (lit. 'football gymnasium'), *año luz* (lit. 'year light'), etc. Many of the rare relationships are due to borrowing, as in the last case. Contrary to what happens in other languages, as the Germanic ones for instance, Spanish compounds in productive use, however, only follow a restricted set of semantic types. Plural with this class of compounds is mostly realized on the head only, though some cases of double pluralization may also be found. Some second constituents which are very adjective-like in their meaning are evolving towards real adjective status, showing number-agreement with the head and sometimes also predicative use: *problema clave* (lit. 'problem key'), *problemas clave(s)*, *estos problemas son clave(s)*, etc. Even *clave*, however, which has gone further in this process of adjectivization, cannot yet be used in prenominal position: *un importante/clave problema*. Other frequent second constituents of this sort are *modelo* ('model'), *estrella* ('star'), *límite* ('limit'), *fantasma* ('ghost'), etc.

2.1.3. *Right-headed n-n compounds*. The number of right-headed n-n compounds one is willing to concede for Spanish very much depends on how one draws the dividing line between compounding and derivation. If one accepts the particular use we have made of the additional criterion of system-adequacy in 1.1.2.2., very little remains to be included in this section; in essence, marginal journalistic creations with a proper name as a first constituent, such as *castroestalinismo*, *carrillonovela*, *gonzalezmania*, etc.

2.1.4. *Exocentric n-n compounds*. Exocentricity as a semantic phenomenon can, of course, potentially affect all complex words on an individual basis. In this section we will only mention one type of n-n compound which is intrinsically exocentric, namely the formation of intermediary cardinal points. This curious compound type has some very peculiar restrictions: (a) the constituents must denote a cardinal point; (b) the meaning is the direction half-way between n_1 and n_2 ; (c) the constituents must be

adjacent cardinal points (**norsudoeste*): (d) there are precedence relationships of the following kind: *norte(sur)* ('north/south') before *este/ oeste* ('east/west') (hence **estesur*), main direction precedes intermediate direction (hence **sudestesur*).

2.1.5. *n-i-n compounds*. There are very few compounds of this type in current usage: *sopicaldo*, *carricoche*, *gallipavo*, *calicanto*, *pavipollo*, etc. The type is unproductive. Some may be considered to have the head to the right, while others (e.g. *calicanto*) are clearly exocentric. The *-i-* is a linking vowel.

2.1.6. *Righthheaded a-i-n compounds*. This type too is restricted to a small set of words: *alipianicie*, *plenipotencia*, etc. The first constituent is an adjective modifying the noun to which it is linked by *-i-*: *alto* ('high') + *planicie* ('plateau'), etc.

2.1.7. *Nominal v-v compounds*. Nominal v-v compounds are just as rare in the standard language as the two types just mentioned. The two constituents are verbs in third person singular indicative (with the exception of *tejemanaje* ('fuss, to-do'), whose second constituent formally coincides with the subjunctive), but the result exocentrically refers to a nominal concept. In one subtype, the two verbs are identical: *bullebulle* ('busybody', lit. 'boils boils'), *comecorre* ('headlong rush', lit. 'runs runs'), (*polvos*) *picapica* ('itching powder', lit. 'itches itches'), etc.; in another, antonyms: *duermevela* ('snooze', lit. 'sleeps stays awake'), *ganapierte* (a type of game, lit. 'wins looses'), etc. The plural, due to the exocentricity, is formed word-finally (though *tejemanajes*, with two plurals, is also attested). The gender is invariably masculine, probably by default.

2.1.8. (*So-called*) *v-n compounds*. So-called v-n compounds have always constituted a favorite topic in Romance word-formation. While earlier work concentrated on the origin of this compound type, recent publications are mainly concerned with the typological anomalies it presents. By far the most controversial subtopic is the nature of the first constituent, considered by some to be a verb and by others to be a noun. The best manner to proceed will therefore be to first sketch the various analyses defended in the literature and then evaluate them in the light of the various properties of this type of compound.

Among those who consider the first element to be a verb no less than six positions may be distinguished. First, the left constituent is considered by some to be an imperative, which undoubtedly is the right diachronic analysis, but cannot be maintained synchronically (and, indeed, we know of no serious defense of the synchronic relevance of the imperative thesis). Second, others also consider the first constituent to be an inflected verbform, namely a third person singular indicative. The third position, defended e.g.

by Alemany-Bolufer (1920), is very similar to the second but should not be lumped together with it. It says that the first constituent takes the *form* of a third person singular indicative, but functions as a verbal theme, i.e. an uninflected verb form. The remaining three "verbalist" positions have been proposed in the framework of generative grammar. In the late sixties and early seventies, some transformationalist accounts have been proposed, which need no longer be taken into account here since the theoretical framework as a whole is obsolete by now. More recently, Contreras (1985) has proposed to assign our compounds the structure [[DET N']_{NP} [V NP]_{VP}]_{NP}, while DiSciullo and Williams (1987: 78-88) have proposed the "marked" phrase-structure rule N → VP for the corresponding French type.

The "nominalist" positions all view our compounds as endocentric, considering the first constituent to be an agentive nominal. The differences concern the exact nature of this agentive nominal. Coseriu (1977) proposes to view it as the result of deleting the agentive suffix, Zuffi (1981) derives it by means of a zero-morpheme, while Varela (1990) advances the hypothesis of a reanalysis of the — originally inflectional — final *-a/-e* of the first constituent as derivational agentive suffixes.

Let us now review the main properties of Spanish v-n compounds and see in how far they are compatible with these various positions.

As to the formal side of the first constituent, it completely coincides with the third person singular present indicative. Since this inflectional morpheme does not seem to be present in the semantic structure of the compound, the principle of compositionality seems to force us to reanalyse it. As we have already pointed out, there have been various proposals in the literature about the exact nature of the result of this reanalysis. While Alemany assumes that it has turned into an a-functional inflection morpheme that has to be stipulated in a synchronic description, Varela holds it to have been reanalysed as a derivational affix. We see no means, presently, to choose one of these solutions over the other on principled grounds. The argument sometimes advanced in the literature that a verbal analysis would predict verb agreement, like **los sacancorchos*, where the *-n* is the number morpheme, has no bearing on an Alemany style analysis and probably not even on the one that views the final morpheme of the first constituent as a true inflectional ending, since the Spanish verbal agreement rule does not operate inside a noun phrase but between subject and verb.

With the exception of less productive types denoting places, birds, plants, games or activities, v-n compounds denote agents or instruments, which all include a feature of habituality. The instrumentals are currently quite productive and neutral as to register and connotation: *lanzacobetes* ('rocket launcher', lit. 'launch rockets'), etc. The agentives, on the other hand, are mainly productive for the creation of colloquial, more or less derogatory designations, such as *matasanos* ('quack doctor', lit. 'kill sanes'), *comeceras* ('anticlerical', lit. 'eat priests'), etc. If they refer to professions, these are

generally of little prestige: *limpiabotas* ('bootblack', lit. 'clean boots'), etc. The various analyses differ in the place they assign to the agentive/instrumental meaning element: while verbalists locate it outside the two constituents, nominalists tie it to a supposed suffixal position in the first constituent. This latter position thus implies that the originally exocentric type has meanwhile been reanalysed as endocentric.

With respect to argument-structure, the most important thing to note is that the noun has to satisfy the internal argument position of the verb. Since by the very nature of the meaning of this compound type, the verbs are invariably agentives, the internal argument is in the overwhelming majority a theme, given that agent-verb-theme is the standard case of a transitive verb in Spanish. However, one can also find some locatives in the noun position (*saltamontes* 'grasshopper', lit. 'hop woods/mountains'), which is not surprising since locatives often permit both a direct and a prepositional realization in Spanish (*saltar el muro/por encima del muro* 'to jump over the wall', etc.). Prepositional complements other than locatives, however, are excluded from appearing in v-n compounds. We should also note that indirect internal arguments, if they exist, as with other agentives cannot be satisfied externally: *quitaesmalte* ('polish remover', lit. 'remove polish'), but not **un quitaesmalte de las uñas* (lit. 'a polish remover from the fingernails'). There is one difference between verbalists and nominalists with respect to the external argument: while in the nominalist account this position is satisfied by the putative agentive suffix, satisfaction must be done in semantic structure in the exocentric account. It is unclear to us at the moment whether this difference entails any empirical consequences.

Spanish v-n compounds are nouns, but they can also be found in adjectival function: *el monstruo tragastrellas* ('the star-swallowing monster'), *yo soy más rompeteteras* ('I am more teapot-breaking'), etc. It is possible that such adjectival uses may simply be attributed to independent conversion rules. Alternatively, the noun or pronoun such adjectives are predicated of might be conceived of as satisfying the open external argument position under a verbalist analysis. The only analysis which does not view v-n compounds as nouns is Contreras'.

The gender of v-n compounds is masculine (probably by default), except for those which refer to humans and some bird and plant names, which can get hyperonymic feminine gender. Note that gender is quite independent of the gender of the non-head second constituent.

Spanish v-n compounds normally have a secondary stress on the first constituent and the main stress on the second. In very frequent words, the secondary stress may be lost by destressing. The accentual pattern thus clearly differentiates v-n compounds from left-headed n-n compounds. That means that nominalists cannot maintain that the reanalysis has yielded a normal left-headed n-n compound but rather a type of left-headed n-n compound *sui generis*. Other properties which distinguish v-n compounds

from orthodox left-headed n-n compounds are pluralization and the possibility of interpreting the second noun as an argument, a reading not available e.g. for the hypothetical left-headed n-n compound *dibujante estrella* (lit. 'drawer star').

While non-lexicalized left-headed n-n compounds get the plural on the head, v-n compounds get it at the end of the word: *girasoles*, from *girasol* ('sunflower', lit. 'turn sun'). This behavior is not surprising under an exocentric analysis, while it has to be stipulated in a left-headed n-n analysis. The internal plural of the second-constituent also deserves mention. Several researchers have noted that there is a growing tendency — at least in Peninsular Spanish — to extend this internal plural morpheme even to cases where it is not semantically plausible, such as *quitapelos* ('barber', lit. 'take-away hair'). If this tendency continues, final -s may end up as an empty morpheme constitutive of the type as such, very much as linking morphemes.

The same a-functional status has already been reached by the third person singular indicative inflection according to Alemany's analysis, who grants this inflectional morpheme a purely formal role comparable to the feminine marker on the basis of adverbs in *-mente*. Stipulative inflectional forms of this kind can be expected to occur in cases of morphological reanalysis such as v-n compounds or adverbs in *-mente*, which, as is well known, have evolved from Latin noun phrases but are now best considered as derivatives (let's note, however, that some researchers favor a phrasal or compound analysis even on synchronic grounds).

The last property of v-n compounds that we would like to note is the occasional occurrence of vowel elision, especially in lexicalized formations such as *paraguas*, from *paraguas* ('umbrella', lit. 'stop water').

To sum up, we have seen that the enigma of Romance v-n compounds is far from having been solved to everybody's content. The verbalist and nominalist positions essentially differ in how far they think that the reanalysis of the original imperative sentence has progressed. Verbalists think that the first element is still verbal in nature and that the original exocentricity is intact, while nominalists think that the agentive/instrumental feature has by now been tied to a newly created suffixal position internal to the compound. It seems to us that a definitive decision is unlikely to be reached until our understanding of general principles of morphological theory and, more specifically, morphological reanalysis will have entered in a more mature stage.

2.2. Adjectival compounds.

2.2.1. *Coordinative a-a compounds.* The formation of coordinative a-a compounds of the type *ético-moral* is among the most productive processes

of Spanish word-formation, even though it is essentially confined to the written language.

The meaning of a-a compounds can best be glossed as 'both a_1 and a_2 ': *ético-moral* ('both ethic and moral'), *árabe-israelí* ('both Arab and Israeli'), etc., though with relational head-nouns other paraphrases may sound more natural: *guerra árabe-israelí* ('war between Arabs and Israelis'). Formations of the type *francocanadiense* ('Franco-Canadian') seem to be subordinate since the first constituent picks out a subgroup of the one defined by the second one, but this may well be considered as a pragmatic effect dependent on our knowledge of the world. Thus *un acuerdo franco-canadiense* ('an agreement between France and Canada') and *de origen francocanadiense* ('of Franco-Canadian origin') would have the same abstract coordinative semantic structure, attributing the concrete reading to our knowledge of the world.

Some properties of a-a compounds are a straightforward consequence of their coordinative meaning. Both constituents generally belong to the same semantic field: *higiénico-sanitario* (lit. 'hygienic sanitary'), *antropológico-lingüístico* (lit. 'anthropological linguistic'), *sordomuto* (lit. 'deaf-mute'), *rojodorado* (lit. 'red golden'), etc. The constituents are in principle interchangeable: *matemático-estadístico/estadístico-matemático*, etc. In reality, however, the order of the constituents is subject to some additional preference rules, of which the following seem to be the most important: (a) bound elements by definition show up in non-final position: *hispano-franco-cubano* / **hispano-cubano-franco*. Bound elements of the type *franco-* (for *francés* 'French'), by the way, are quite common and can be created productively to a certain degree: *semio-lingüístico* (from *semiótico*), *participio-gerundial* (from *participial*), *culturo-lingüístico* (from *cultural*), etc.; (b) the longer constituent tends to be realized in second position; (c) if one constituent ends in -o and the other in a consonant, the vowel-final one tends to be realized in first position. This preference-rule may supersede rule (b), e.g. in *ético-moral* / **moral-ético*, etc. Systematic exceptions to this latter preference-rule are political formations of the type *nacional-católico*, etc.

There is a subset of a-a compounds which strongly resembles the coordinative type but which has to be kept apart for semantic reasons. We are referring to adjectives like *semántico-generativo*, *latino-vulgar*, *germano-occidental*, *alto-extremo*, *bajomedieval*, *grancanario*, *germano-occidental*, etc., which are paradigmatically related to noun + adjective or adjective + noun phrases: *semántica generativa*, *latín vulgar*, *Alemania occidental*, *Alta Extremadura*, *Baja Edad Media*, *Gran Canaria*, etc. These are best viewed as an adjectival counterpart of the nominal type *lingüística textual* (from *lingüística textual*), etc. Equally non-coordinative is the meaning of ordinal numbers like *decimotercero* ('thirteenth', from *décimo* 'tenth' and *tercero* 'third'), etc.

Adjectival agreement is realized on the second constituent: *unas cuestiones matemático-estadísticas*, etc. Cases with agreement-morphemes on both constituents, like *la clase media alta* ('the higher middle class'), *los prefijos negativos-privativos* ('the negative-privative prefixes'), *la lengua literaria-poética* ('the literary-poetic language'), etc., may probably be viewed as cases of asyndetic syntactic coordination. Constructions where the meaning of the two adjectives is clearly integrated never allow double agreement: **guerra palestina-israelí*, etc.

Compounds of this type have two main stresses in the regular cases, but in more frequent combinations destressing of the first constituent may be observed, especially if it is a short adjective.

2.2.2. *Coordinative a-i-a compounds*. Coordinative a-a compounds with a linking vowel *-i-* only occur with a limited set of adjectives. The only really productive subtype contains colour adjectives: *rojiblanco* (from *rojo* 'red' and *blanco* 'white'), etc. Not any colour adjective, however, may enter in this compound type, rather the first constituent seems to be reserved for the following bisyllabic adjectives in *-o* and *-e*: *blanco*, *negro* ('black'), *pardo* ('dun'), *rojo* and *verde* ('green'). The second constituent is less choosy: *verdigrís* ('green' and 'grey') or the following neologisms by C. J. Cela: *verdibermexo* ('green' and 'bright red'), *verdírubio* ('green' and 'blond'), *blanquimarrón* ('white' and 'brown'), etc.

Outside colour adjectives, we only find some isolated combinations like *agridulce* ('bittersweet') or *tontiloco* (from *tonito* 'silly' and *loco* 'crazy'). Sporadic neologisms on a local analogical basis with *agri-* and *tonit-* are attested.

Adjectives of this type do not have a main stress on the first constituent.

2.2.3. *Right-headed ado-i-a compounds*. There are only a few compounds of this type, essentially *altisonante* (from *alto* 'high' and *sonante* 'sounding') and *clarividente* (from *claro* 'clear' and *vidente* 'seeing'). The second constituent of *altilocuente* and *grandilocuente* does not occur in isolation.

2.2.4. *Right-headed n-a compounds*. Right-headed n-a compounds do not form a homogeneous set in Spanish. This type is not productive on the categorical level, but rather restricted to a fixed number of elements that can occupy the first or second position.

One subtype has a noun denoting a cardinal point in first position: *nortearagones* ('North Aragonese'), *surcoreano* ('South Korean'), etc. In journalistic usage, *oeste* ('west') can also be observed: *oeste-germano*, etc.

While in this subtype the second position can be filled with any semantically suited adjective, in the following subtype with a second constituent *hablante* ('speaking') or its bound variant *parlante*, the first

position is potentially open to any language name: *euskerabablante*, *guaraní-hablante*, *inglés-hablante*, etc. *Angloparlante*, *catalanoparlante* (from *catalan* 'Catalan'), etc. show that speakers tend to assimilate this right-headed type to the type *ético-moral* by introducing bound first elements ending in *-o-*.

In the other formations of this type both the first and the second position are fixed. One subset belongs to legal terminology: *cuentaante* ('accountant'), *derechohabiente* ('right-of-way'), *fehaciente* ('authentic'), etc. Some are only used as nouns: *causabiente* ('right-of-way'), *terratiente* ('landowner'), etc. Recent creations are mostly *calques*: *drogadicto* ('drug addict'), *capital-intensivo* ('capital-intensive'), etc. A form like *drogodependiente* ('drug dependent') shows again that speakers are uneasy with this right-headed compound-type since *droga* has been changed to *drogo*, thus assimilating the first constituent to a kind of prefix. Sporadic attempts to revive this right-headed compound type have resulted in completely artificial creations like *topoparidora* (Unamuno), *churumbelipontante* (Martín Santos), etc.

The semantic structure of these adjectives resembles that of the corresponding Germanic type. If the second constituent is deverbal, the first one may satisfy the internal argument position, while the external argument is satisfied externally by the modified noun. If the second constituent, on the other hand, is non-argument taking, as in *capital-intensivo*, the meaning is to be determined pragmatically.

Adjectives of this type invariably have a secondary stress on the first constituent and a main stress on the second one.

2.2.5. *Exocentric n-i-a compounds*. The *pelirrojo* type (from *pelo* 'hair' and *rojo* 'red') presents some quite interesting properties.

Its meaning may be paraphrased by '(intrinsicly) characterized by a (djective) n', i.e., the type is to be considered exocentric. This can be seen more clearly in cases where the second constituent cannot be considered as a hyperonym of the whole adjective: so, e.g., a person who is *cabizbajo* ('crestfallen', from *cabeza* 'head', idiosyncratically changed to *cabiz-*, and *bajo* 'low') is not necessarily *bajo* ('low, small'). Further evidence for this type's exocentricity comes from its behavior with respect to derivability: while right-headed compounds are available for head-operational derivation (*drogadicción*, *clarividencia*, etc.), exocentric adjectives refuse this type of derivation: **lengüilargura*, **cuellicortedad*, **carianchura*, etc.

The argument-structural properties of these adjectives are a straightforward consequence of their meaning. The n satisfies the external argument of the a, while the external argument position of the resulting is a projection of the meaning element '(intrinsicly) characterized by'. It is important to note that the external argument of adjectives of the type *pelirrojo* is newly created in the process of derivation, while the one of right-

headed adjectives of the type *drogadicto* is simply taken over from the head-adjective.

Another interesting feature about the *pelirrojo* type are its restrictions. As García-Lozano (1978) has pointed out, the first constituent has to be bisyllabic (note the absence of body-parts like *cabello* 'hair', *pestaña* 'eyelash', *mejilla* 'cheek', *tez* 'complexion', *pie* 'foot', *piel* 'skin', etc.) and refer to a body part. Furthermore, there is an output-restriction limiting compounds of this type essentially to concepts relevant for cattle-breeders.

The type is only scarcely productive in everyday language, due probably to the output restriction just alluded to, but it seems to be productive in cattle breeders' jargon and also in literary language. Cela's work e.g. contains many neologisms like *ojitierno*, *patirrecio*, *aliescurrido*, *aliflojón*, *pelimaro*, etc.

Compounds of this type invariably have a secondary stress on the first constituent and a main stress on the second one.

2.3. Verbal compounds.

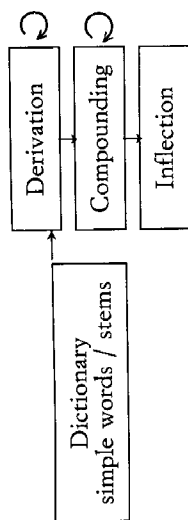
It is dubious whether Spanish has any verbal compounds, since all potential candidates diachronically are quite clearly cases of reanalysis and are not synchronically productive. This is true both of the handful of examples of the type *n-i-v* like *alicortar* (lit. 'wings to cut'), *aliquebrar* (lit. 'wings to break'), *patiquebrar* (lit. 'feet to break') and *perniquebrar* (lit. 'legs to break'), clearly backformed from the corresponding *pelirrojo* type adjectives (cf. the recent neologism *boquiabrir(se)* lit. 'mouth to-open' from *boquiabierto* 'astonished', lit. 'mouth open') and of adv-v formations of the type *bendecir* ('to bless', lit. 'well to say'), *malvender* ('to sell off cheap', lit. 'badly to sell'), *maltratar* ('to ill-treat, maltreat'), *malgostar* ('to waste', lit. 'badly to spend'), etc. which also seem to constitute a closed class. Furthermore, the only adv recurring with some frequency is *mal*. The only criterion for classifying these last formations as compounds is the clear relationship with the freely occurring adv *mal*.

3. Problems of ordering.

3.1. The relative order of compounding, derivation and inflection.

3.1.1. Level ordering, head operations and lexicalization. On the basis of previous work in level ordered morphology and his own analysis of Italian, Scalise (1984: 134) arrives at the conclusion that the lexical component may be organized as follows:

(2)



(2) is meant to capture asymmetries in the relative order of derivation, compounding and inflection across languages: while simple words may feed freely all three subcomponents, the first two of which are also internally recursive, an extrinsic ordering relationship is claimed to exist between derivation, compounding and inflection. This extrinsic order is supposed to account for the following asymmetries: (a) while compounds from derivatives are quite common, derivatives on the basis of compounds are much rarer; (b) while inflected compounds are extremely common, compounds made up of inflected words are much rarer; (c) while inflected derivatives are extremely common, derivatives on the basis of inflected words do not occur.

In the present paper, we will only be concerned with generalizations (a) and (b). Before assessing the empirical validity of (a) and (b) for Spanish, it is necessary, however, to address some general conceptual issues relevant to an evaluation of (2) which are linked to a difference in the kind of exceptionality involved in the counterexamples to (a) and (b) mentioned by Scalise.

Cases of derivation on the basis of compounds seem to be limited essentially to "strict (= lexicalized) compounds" (Scalise 1984: 121-122). If this were the case, one could try to regularize such counterexamples by assuming that lexicalized compounds are treated by speakers as simple words, which would mean that they are integrated into the dictionary and thus available for further derivation.

Although such a solution looks quite appealing at first sight, it does not seem to be viable, however. As Booij (1987) has pointed out, theories of level ordering depend on defining the "dictionary" as a list of non-complex words or stems, since the inclusion of complex words would make them available for all subcomponents, thus undermining the very idea of level-ordering. The problem for the above solution lies in the fact that there seems to be no non-*ad hoc* way of including only those compounds which permit further derivation. From a psycholinguistic perspective, there can be no doubt that even regular formations are stored in the mental lexicon, provided that they are sufficiently frequent, and there is also ample evidence that this notion of a mental lexicon plays a role in characterizing morphological operations (cf. Booij 1987, Spencer 1988). But if the criterion for the inclusion into the dictionary cannot be storage in the psycholinguistic

sense of the term, it's up to level ordered morphologists to provide a viable alternative. Some counterexamples to generalization (b) seem to follow the same pattern as the ones just described (e.g. Kiparsky's examples of the type *lice-infested*), but others are of a different kind. So Scalise proposes to account for the pluralization of left-headed Romance n-n compounds by permitting the plural morpheme to be attached by what Hoeksema (1985) has called a *head-operation*, i.e. a rule capable of looking into morphological structure in order to identify its head and perform some morphological operation on it. Thus the internal position of the plural -s in, say, *años luz* ('light years'), where *año* ('year') is the head, would cease to be a problem for (b) since it is introduced by a head-operation *after* the formation of the compound. Other counterexamples, where the inflectional morpheme appears on the non-head, as in *horno microondas* ('microwave oven'), lit. 'oven microwaves', are not amenable to such an explanation and would have to be accounted for by an exceptional "loop" from inflection back into compounding. These two additional mechanisms, i.e. head-operations and loops, however, also both seriously undermine the idea of transitive ordering which is the conceptual cornerstone of level ordered morphology.

It would thus seem fair to conclude that (2) is faced with serious problems at the conceptual level. These problems notwithstanding, let's see now whether it stands the empirical challenge constituted by the Spanish data.

3.1.2. *The ordering of subcomponents in Spanish.* The facts concerning the relative order of compounding and inflection seem to be entirely parallel in Italian and Spanish.

Left-headed n-n compounds have their plural on the one head to the left if they are subordinative (cf. 2.1.2.) and on both heads if they are coordinative (cf. 2.1.1.), while some heavily lexicalized compounds are treated as single words, showing only a plural morpheme at the end, e.g. *bocacalle* ('entry/end of a street', lit. 'mouth street'), *telaraña* (lit. 'web spider'), etc. A plural on the non-head appears in most v-n compounds (e.g. *cuentagotas* 'dropper', lit. 'count drops') and in some left-headed n-n compounds such as *horno microondas*, mentioned above, or *tren mercancías* ('freight train'), *tienda comestibles* ('food store'), etc. Formations of this type, however, might also be considered to be appositional in nature and would thus cease to disturb generalization (b).

As one can see, the description of the facts presented here requires the whole machinery introduced in 3.1.1., i.e. head-operations, loops and lexicalization, just as in Italian.

With respect to compounding and derivation, the situation in part is again parallel to the one described for Italian by Scalise, but there are some additional facts relevant for an evaluation of (2). In order to account for the data correctly, we think that it is necessary to distinguish between prefixation and suffixation on the one hand and head operational and global

affixation on the other. Syntagmatic compounds and other phrases are also taken into consideration here since their behavior does not seem to be systematically different from that of compounds.

First, there is a clear prefix/suffix asymmetry not noted for Italian by Scalise but possibly relevant for that language too: prefixes attach more freely to complex bases than suffixes and do not declench any adaptive strategies of the kind to be described below for suffixes. Thus a phrase like *guerra del Golfo Pérsico* ('Gulf war', lit. 'war of the Persian Gulf') may be prefixed with *pre-* ('before'), yielding exocentric *preguerra del Golfo Pérsico* ('pre-gulf-war-period'), leaving intact even the phrasal accentuation of the base, while a suffixal derivative from the same basis is absolutely excluded. A bracketing [*preguerra*] *del Golfo Pérsico*, which would be compatible with generalization (a), is excluded on semantic grounds, since the meaning would not be 'pre-gulf-war-period' but 'pre-war period of the Persian Gulf'. A defensible way of accounting for examples of this kind would be to allow prefixes to act also in the head-operational mode. This would explain why *pre-* has scope over the whole base though being attached to the head *guerra*. The following contrast might be taken to support such an analysis:

- (3) a. *la segunda posguerra mundial*
- b. **la possegunda guerra mundial*

(3a) is fine because *pos-* ('after') is attached to the head of the base *guerra mundial* ('world war'), while (3b) is out since *segunda* ('second') is not the head of the phrase *segunda guerra mundial*.

Formations of the type just described are quite frequent in present-day written Spanish even though this interesting set of data has not been mentioned in the literature up to now. Further examples: *contragolpe de Estado*, *ex-ministro de Cultura*, *miniparque de atracciones*, *pro-alfabetización en lenguas indígenas*, *pseudosecretos de Estado*, etc. Besides the head-operational mode, global attachment to complex bases is of course also available for prefixes: *ex-niño estrella*, *mini-paracaídas*, *pseudo-cortometraje*, etc.

In suffixation too it is indicated to distinguish a head-operational from a global mode of application.

The head-operational mode, of course, can only apply when the head of the base is to the right, which, as we have seen, is relatively rare in Spanish. We will thus not be surprised to find fewer examples of this type. A case in point is *extremoriental* (lit. 'Far Eastern'), derived from *Extremo Oriente* ('Far East') by the addition of *-al*, the suffix required by *Oriente*. Other examples: *medioriential*, *primerdivisionario*, *nuevotestamentario* (besides *neotestamentario*), *granducal*, *altoaragonés*, *drogadicción*, *clarividencia*, *cientimillonésimo*, etc. The adjective *medioambiental* ('environmental') is interesting since it presupposes that *medio ambiente* ('environment'), originally composed of a noun followed by an adjective, has been reanalyzed as consisting of a sequence a + n, a reanalysis facilitated by the fact that

ambiente as a noun means 'environment'. What forces us to set these examples apart from the ones formed according to global affixation is the choice of the affix, which is clearly determined by the head. "Choosy" suffixes like *-al*, *-ario*, *-(c)ión*, *-(c)ia*, by the way, only occur in such cases of head-operation.

A property which sets suffixed forms of both modes apart from prefixed ones is the existence of adaptive strategies with the effect of making the phrasal or compound base more word-like. This can be achieved by dropping grammatical words like prepositions (4a), determiners (4b) or conjunctions (4c), or by dropping the concordance morpheme of the first constituent (4d) or substituting it with a vowel *-o-* (4e):

- (4) a. Costa de Marfil 'Ivory Coast' → *costamarfiléño*, por cuenta propia 'on one's own account' → *cuentapropismo*
 b. La Coruña → *coruñés*
 c. noventa y ocho '98' → *noventaochista* 'of the '98 generation'
 d. primera división 'first division' → *primerdivisionario*
 e. Estados Unidos 'U.S.' → *estadounidense*
 Baja Edad Media 'Low Middle Age' → *bajomedieval*

Some formations may even show the simultaneous application of more than one of these strategies. *Río de la Plata* → *rioplatense* ('native of Río de la Plata'), *a media pensión* (lit. 'on half board') → *mediopensionista* (lit. 'demiboarder'), etc. Another strategy for avoiding complex bases consists in simply dropping the head or non-head constituent: *derecho penal* ('penal law') → *penalista* ('penalogist'), *Sendero Luminoso* ('Shining Path') → *senderista* ('S.P. follower'), etc. Elision (*Extremo Oriente* → *extremoriental*), hapology (*castañas asadas* 'baked chestnuts' → *castañasadomano* 'addict to... (Cela)') or recourse to a (pseudo-) latinism (*Buenos Aires* → *bonaerense* 'native of B.A.') may serve the same purpose.

If we have referred to such adaptations as strategies, it is because one can also find examples where they do not apply: *pormenorizar* ('to detail'), *noventaijochista*, *quintacolumnista* ('fifth columnist'), etc.

Contrary to what we have seen in our discussion of the head-operational mode of suffixation, suffixation according to the global mode is restricted to suffixes like *-ista*, *-ismo*, *-ero* *-esco*, verbal *-ar*, *-eria*, *-ez* and gentilic suffixes which are quite free in their application, looking more to the meaning of their bases than to their form or identity (lexical government). Examples: *mediocampista*, *librecambismo*, *quinceaño*, *barriolatnesco*, *cortocircuítar*, *papanatería*, *bijoputez*, *montenegrino*, etc.

Some of these cases are clearly examples of intentional creation, but on the whole it would seem that the process is not marginal as Scalise would have it for Italian. Just as in that language, it is true, however, that some degree of "lexicalization" of the basis is necessary for the process to apply, with the possible exception of *sincorbatismo* (from *sin* 'without' and *corbata* 'tie') and similar formations.

As we have already pointed out, the degree of lexicalization required by suffixes is even higher than that required by prefixes. Interestingly, the degree of lexicalization required by evaluative suffixes is again higher than that of suffixes and very close to the one observable in pluralization.

If a left-headed compound is completely transparent, diminutivization is accomplished by a head-operation as with pluralization: *añitos luz* (lit. 'years DIM light'), *niñito prodigio* (lit. 'child DIM miracle'), etc. Heavily lexicalized complex bases, on the other hand, receive their plural at the end: *telarañita* (lit. 'web spider DIM'), *bijoputita* (lit. 'son bitch DIM'). In the case of v-n compounds, things are slightly more complex. Since diminutivization here might be interpreted as referring only to the non-head n, it seems to be normally avoided: *sacacorcbitos* (lit. 'screw cork DIM'), e.g., could be interpreted as 'screw for small corks', not as 'small cork-screw'. In the more lexicalized compounds of this type, however, diminutives with scope over the whole compound are possible according to the regular infixational diminutive rule operative e.g. in *Luquitas* (from *Lucas*): *paraguítas* ('small umbrella', lit. 'stop water DIM'), *papanatitas*, *pelagatitos*, etc.

To sum up, our analysis of Spanish has yielded results that cast further doubt on extrinsic-ordering-models of the kind presented in (2). First, the order compounding before derivation has been shown to be less marginal — at least in Spanish — than Scalise would have it. Second, it has been shown that head-operations seem to be just as common in derivation as in inflection. And third, the lexicalization condition has been seen to exist in varying degrees, being strictest in inflection and diminutivization, less strict in other types of suffixation and loosest in prefixation. Although many aspects of these generalizations are still in need of clarification, they seem to warrant the conclusion that transitive ordering of subcomponents may not be the right conceptual tool to achieve this goal.

3.2. Recursiveness.

One of the more striking features of Spanish compounding is its low degree of recursiveness, especially if confronted with the situation in the Germanic languages. The question arises whether this has to be attributed to some language-specific constraint(s) or follows from other properties of the system.

In coordinative compounds, recursiveness is quite common in those of the n-n and a-a type: *el marxismo-leninismo-fascismo-castismo*, *el entramado político-económico-defensivo que supone la OTAN*, etc. The lack of recursiveness of the a-i-a type (**rojinegriblanco*, etc.) is intriguing since it certainly cannot be attributed to semantic reasons (cf. German *schwarz-rot-gold*, etc.). Probably this "lacuna" may be attributed to the independently

motivated restriction on the number of syllables allowed in this type of compounds and so would not have to be stipulated as a separate rule-specific restriction.

In subordinate compounds, cases of recursiveness are quite rare. There are a few attested examples of the structure [V [V + N]_N]_N, such as *limpiaparabrisas* ('windshield wiper', lit. 'clean stop wind'), and compounds of the form *limpiabuque-cisterna* ('tanker cleaner', lit. 'clean ship tank'), are certainly conceivable though we know of no attested examples. We also know of only one neologism of the structure [[N + N]_N], namely *decreto-ley puente* ('interim decree law', lit. 'decree law bridge').

Much of the abysmal difference in the number of recursive formations in Spanish and Germanic subordinate compounding can certainly be attributed to the differences in meaning. While Germanic n-n compounds can be either synthetic or else express any pragmatically plausible relation between two nouns, Spanish n-n compounds are restricted to a few semantic relationships in productive use. And the few productive semantic types simply do not seem to yield many useful concepts when combined. It should also be noted that Italian is closer in that respect to the Germanic languages than Spanish, allowing "monstruosities" such as *Associazione protezione uccelli* (lit. 'Association protection birds') which however do not yet form part of the core of Italian word-formation.

In Rainer (1986) it is argued that recursiveness in Spanish derivation is constrained only by independent factors. It would seem that the same can be said of Spanish compounding, but further investigation is needed in order to reach certainty about this problem. *A priori* it does not seem to be impossible that there may exist rule-specific restrictions on recursiveness, especially if it is true that the learning of the morphology is guided by some principle of conservatism on the type level, as Pinker (1989) has argued. From this perspective, the learner would not automatically infer the possibility of [[N + N]_N]_N or [[N + N]_N]_N from the existence of [N + N]_N, but rather stick to the bimembral subordinative pattern instantiated by all the examples he has heard.

Address of the Authors:

Franz Rainer
Inst. für Romanistik
Universität Salzburg
Akademiestr. 24
A-5020 Salzburg, Austria

Soledad Varela
Dept. de Filología Española
Fac. Filosofía
Universidad Autónoma de Madrid
E-9049 Madrid, España

References

- Alemaný-Bolufér, J. (1920), *Tratado de la Formación de Palabras en la Lengua Castellana*, Madrid, Librería Gral. de Victoriano Suárez.
- Bauer, L. (1990), "Be-heading the Word", *Journal of Linguistics*, 26: 1-31.
- Booij, G. (1987), "Lexical Phonology and the Organization of the Morphological Component", in E. Gussmann (ed.), *Rules and the Lexicon*, Lublin, Catholic University of Lublin: 43-65.
- Bustos-Gisbert, E. (1986), *La Composición Nominal en Español*, Salamanca, Ediciones Universidad de Salamanca.
- Chomsky, N. (1981), *Lectures on Government and Binding*, Dordrecht, Foris.
- Contreras, H. (1985), "Spanish Exocentric Compounds", in F. Nuessel (ed.), *Current Issues in Hispanic Phonology and Morphology*, Bloomington, IULC: 14-26.
- Coseriu, E. (1977), "Inhaltliche Wortbildungslehre (am Beispiel des Typs 'coupe-papier')", in H. Brekle & D. Kastovsky (eds.), *Perspektiven der Wortbildungsforschung*, Bonn, Bouvier: 48-61.
- Di Sciullo, A. M. & Williams, E. (1987), *On the Definition of Word*, Cambridge (Mass), MIT Press.
- García-Lozano, F. (1978), "Los Compuestos de Sustantivo + Adjetivo del tipo *pelirrojo*", *Ibero-Romania*, 8: 82-9.
- Grimshaw, J. (1990), *Argument Structure*, Cambridge (Mass), MIT Press.
- Hoeksema, J. (1985), *Categorial Morphology*, New York, Garland.
- Lieber, R. (1981), *On the Organization of the Lexicon*, Bloomington, IULC.
- Moreno, J.-C. (1984), "Observaciones sobre la Sintaxis de *casí*", *Dicenda*, 3: 239-45.
- Pinker, S. (1989), *Learnability and Cognition: The Acquisition of Argument Structure*, Cambridge (Mass), MIT Press.
- Rainer, F. (1986), "Recursiveness in Word-Formation, with Special Regard to Spanish", *Acta Linguistica*, 36: 197-209.
- Randall, J. (1984), "Thematic Structure and Inheritance", *Quaderni di Semantica*, 5/1: 92-110.
- Scalise, S. (1984), *Generative Morphology*, Dordrecht, Foris.
- Selkirk, E. (1982), *The Syntax of Words*, Cambridge (Mass), MIT Press.

- Spencer, A. (1988), "Bracketing Paradoxes and the English Lexicon", *Language*, 64: 663-82.
- Varela, S. (1990), "Composición Nominal y Estructura Temática", *Revista Española de Lingüística*, 20/1: 55-81.
- Wurzel, W. (1984), *Flexionsmorphologie und Natürlichkeit*, Berlin, Akademie-Verlag.
- Zuffi, S. (1981), "The Nominal Composition in Italian: Topics in Generative Morphology", *Journal of Italian Linguistics*, 2: 1-54.
- Zwicky, A. (1985), "Heads", *Journal of Linguistics*, 21: 1-29.