# *Abrelatas* And *Scarecrow* Nouns:
# Exocentric Verb-Noun Compounds
# As Illustrations of Basic Principles of Cognitive Grammar

DAVID TUGGY *
*SIL Znternational, Dallas*

**ABSTRACT**

Spanish and English have exocentric verb+object = subject/instrument compounds, such as *abrelatas (opens-cans)*'can-opener' and *scarecrow*. They share a general constructional pattern, consist of "clumps" or subfamilies of forms, and have a negative or jocular tendency. They differ in their individual compounds, subfamilies and constructional prototypes. The Spanish construction is a widely productive, major mechanism for naming instruments; the English construction names subjects, and is a minor pattern currently productive only in one subfamily. Exceptional forms in both languages approach each other's prototype. In both languages the category fits into wider families or categories of constructions, but those wider families are different.

These patterns illuminate basic tenets of Cognitive grammar, including: (1) usage-based grammar. (2) Multiple patterns. (3) Lower-level outranking higher-level patterns. (4) Functional motivation, but (5) persistence of patterns despite absence of functionality. These considerations underline (6) the insufficiency of models positing innate, absolute, few and simple rules.

KEYWORDS: Spanish, English, exocentric compounds, Cognitive grammar, functional motivation.

*Address for correspondence*: David Tuggy, C/o SIL, 16131 N Vermon Drive, Tucson AZ 85739-9395,USA. Phone: (520) 825-6000. **Fax:** (520) 825-6116. E-mail: david_tuggy@sil.org

# I. INTRODUCTION

Spanish and English share with a number of other European languages' a striking type of exocentric (headless) noun compounds. In them a transitive verb combines with a noun which is understood as its object. However, the combination of the two does not designate either the verb nor the object. Rather it designates either the subject, the one that does the verb to the object, or the instrument, a thing a subject would use to do the verb to the object. *Abrelatas* [opens-cans] 'can-opener'[3] and *scarecrow* are typical of the two constructions, and I will use them as convenient handles, referring to these nouns as *abrelatas* and *scarecrow* nouns, and speaking of the *abrelatas* and *scarecrow* constructions.

The *abrelatas* and *scarecrow* constructions resemble each other in a number of ways, including some rather surprising ones, but are dissimilar in others. This article is an attempt to describe and discuss some of these similarities and differences, from the standpoint of the Cognitive grammar framework (CG, Langacker 1987, 1991a, 2000). Some relevant features of CG are presented in section II; then successive sections discuss, from the perspective of CG, similarities and distinctions between the *abrelatas* and *scarecrow* constructions. Section *IX* closes the paper with a summary and a few general reflections.

# II. SOME BASIC CONCEPTS OF CG
## II.1 Schemas, Categorization and Prototypes

Much linguistics of the past century, especially those currents heavily influenced by the work of Chomsky (1965), consciously adopted the assumption that the central mechanisms of language ("linguistic competence") were largely innate and fixed, and quite separate from and mostly unaffected by linguistic "performance". Linguistic usage, of course, belonged squarely in the latter category, and was thus peripheral to the primary concerns of linguistics.

An increasing number of linguists are coming to accept (or return to)' a sharply-contrasting vision of language as usage-based, with all kinds of linguistic structures, including syntactic structures, growing out of the ways language-users have found it possible, useful, and eventually comfortable and habitual to communicate with each other. "On this view language and the rules for its use in a particular society are a set of social conventions which have evolved in the particular way they have in that society [...] These conventions [...] become habits of the individual spcakers of the language". (LaPolla to appear). People want to communicate. to describe, to express and shape attitudes, to entertain and tickle. A language, including everything from phonology to semantics, from individual specific morphemes to the patterns of entire discourses and conversations, is the incredibly flexible, complex, and beautiful system of conventionalized habits that results.

CG is such a usage-based, "bottom-up" model of language, in which actual attested language data form the foundation of the linguistic system. Highly specific structures such as lexical items, and low-level generalizations, coexist with the higher-level generalizations

traditionally recognized as "rules". All linguistic structures are subject to the "content requirement" (Langacker 1987:53-54, 1991:2): they must be either (parts of) actually attested, overtly occurring structures, direct generalizations from such structures, or categorizing relationships linking the two. Generalizations are called schemas, and are represented graphically by an arrow (representing the categorizing relationship) from the schema to the more fully-specified structure it subsumes: thus $A_1$ is a subcase of generalization (schema) $A$ in Figure 1. Schemas are abstracted (by language users, henceforth speakers, or secondarily by linguists) from the actually attested low-level (highly-specific) data they encounter and learn to use. Thus, suppose a language has an array of specific structures $A_1, A_2$, and so forth, all conforming to a pattern $A$. Speakers are likely to recognize pattern $A$, and use it to categorize $A_1, A_2$, etc. The resulting cognitive configuration, a typical kind of category, is represented as a part of Figure 1.
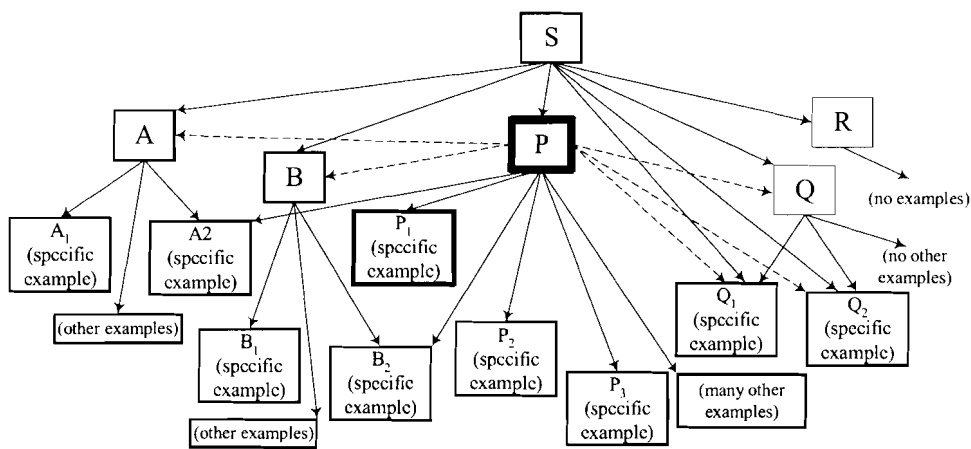


Figure *1*: Schernas and subcases

Note that the existence of $A$ as a cognitive structure in speakers' minds does not preclude or supplant the existence of the subcases $A_1, A_2$, etc.; rather both the schema and its subcases can exist simultaneously in people's minds. In fact they do coexist to the extent that both are entrenched as conventional in speakers' minds by repeated, *salient* (cognitively energetic) usage. (That is, they become habitual for each speaker, and speakers know that they are habitual for each other). But such structures will vary in their degrees of entrenchment and conventionality, depending on their usage (which itself can vary from speaker to speaker). There is no presumption that a specific example (or lower-level schema) will have all its important characteristics specified in the schema that subsumes it; rather it is expected (if not necessary by definition) that it will exhibit significant specifications that go beyond what is stipulated by the overall pattern. In other words, $A$ underspecifies $A_1$ and its other subcases, and cannot

properly be thought of as explaining them away. Making a generalization, under CG, does not entail excising the particulars.

The same language may have an array of similar specific structures $B_1$, $B_2$, etc., and speakers may recognize pattem B in them, i.e. they may extract schema B. And the similarities between A and B may prompt the extraction of schema S, which subsumes the two related pattems and expresses their commonality, uniting them into a higher-order category. It becomes clear that schematicity is not a plus-or-minus quality, but a matter of degree; structures of this sort can grow to include any number of levels. In fact what we are calling specific examples are themselves schemas or pattems, and actual usage events involving them will have details that are not specified by them (e.g. phonetic details, or semantic variations; cf. the subcases of scarecruw and *espantapájaros* represented in Figure 5 [see section III.2], which in their tum are also schemas from actual usage events).

Often, even typically, a structure of this sort will have a niid-level schenia that is significantly more prominent or salient than the others. Thus in Figure 1 P is represented, by the boldfacing of its box, as being more salient than A or B, its sister subcases under schenia S. Such enhanced salience will result naturally from greater usage; thus a pattem (schema) with a greater number of subcases, will, ceteris paribus, be more prominent than one with fewer,' and P, with many exaniples, is naturally more salient than A or B with relatively few. Such a relatively prominent subcase of a general pattem is a kind of cognitive prototype, and its subcases (such as $P_1$, $P_2$, etc.) are *prototypical* examples of category S in a derived sense: whether or not they themselves are relatively salient (and they will typically differ in their degrees of salience; thus P, is represented as more salient than the others) they are examples of the most salient pattem (P).

In contrast, Q, which only has two subcases, may never be extracted as a separate generalization at all and is likely to be tenuous if it is. If it is not, Q, and $Q_2$ remain linked to the overall pattem as subcases of S. Even more marginal, in fact almost certain to be non-existent in most speakers' minds,' would be a logically possible pattem R with no examples.

In a structure like Figure 1 the differences in status among such established pattems as A and P, tenuous generalizations like Q, and hypothetical ones like R, is directly representcd. This is an important difference between CG and other models, which have too often assumed that characterizing the topmost schema (S) is enough to account linguistically for a category. On the CG view the topmost schema(s) of a category are important in that they characterize what all their subcases have in common, but the prototype(s) of the category, and indeed the whole structure, must be taken into account for a complete analysis.

Two other features of Figure 1 need to be mentioned. Note that $A_2$ and $B_2$ are represented as subcases of A and B respectively but also of P. It is very common for specific examples thus to instantiate more than one pattem or schema siniultaneously.

Note also that dashed-line arrows extend froni P to A and B and also to Q, Q, and $Q_2$. These arrows indicate relationships of *extension*, or *partial* schematicity. It is a cognitive commonplace that lesser-known or less-entrenched concepts are typically compared with better-

known, more-strongly-entrenched ones. When such a comparison does not show any discrepancies between the standard and target of comparison a judgment of full schematicity results: all specifications of the standard hold true of the target. Such relationships are naturally more salient than relationships of partial schematicity, where there is some distortion or conflict in specifications: our cognitive systems get excited when the compared items are fully compatible. But when there is a standard which is already highly salient in its own right, because it is highly entrenched or because its salience is enhanced by context, comparisons anchored to that standard acquire greater importance. In a structure like that of Figure 1, one might say that the membership of A, B, Q,, etc., in the same category with P is as much because they are seen as slightly distorted examples of P, as because they are well-formed examples of S.

## II.2 A Headed Cornpound under CG

Figure 2.a represents some important aspects of how a typical headed compound is analyzed under *CG*. (1) both the component words or stems (tow and truck) and the cornposite *structure* in which they participate (towtruck) are bipolar, having one pole in "semantic space" and the other in "phonological space", with a conventional symbolic linkage joining them. (2) The semantic pole of one of the components (truck in this case) is schematic for the composite semantic structure (thus the arrow from TRUCK to TOWTRUCK). This is what makes *truck* the head of *towtruck*. We here follow the convention of representing by all capital letters the designatum (also called the *profile*, or *profiled entity*) of a semantic structure. The designata of *truck* and of towtruck correspond: towtruck merely adds semantic detail about the kind of truck designated. (3) The notion of towing carries within it the notion of two prominent participants, something that causes motion by pulling with a rope or chain and something else that is attached to the rope or chain and is caused to move. The former, more prominent participant is, traditionally within *CG,* called the *trajector* (abbreviated Tr in the diagrams or TR when profiled), and the latter, secondarily prominent participant is called the *landmark* (Lm or LM). (4) The designatum of the second component (the TRUCK) is construed as identical with the trajector of the first; this identification is represented graphically by the dotted arced line between the two, and the resulting integration is represented within the semantic pole of the corresponding composite structure. Such identification of a nominal entity (one designating a Thing) with the trajector of a verbal entity (one designating a process) amounts, under *CG,* to subject status for the nominal entity. (5) There is no important integrative process between the two phonological components beyond juxtaposition in a certain order. This is what makes the construction a compound.[6] It is not the case, for instance, the case that [trsk] brings with it the expectation of a phonological stem to precede it and that [to"] fulfils that expectation. If it were, [trsk] would be, to that extent, suffixal). (6) Primary stress in the composite phonological structure falls on the material corresponding to the first component, i.e. on [tó$^w$] rather than on [tr∂k].
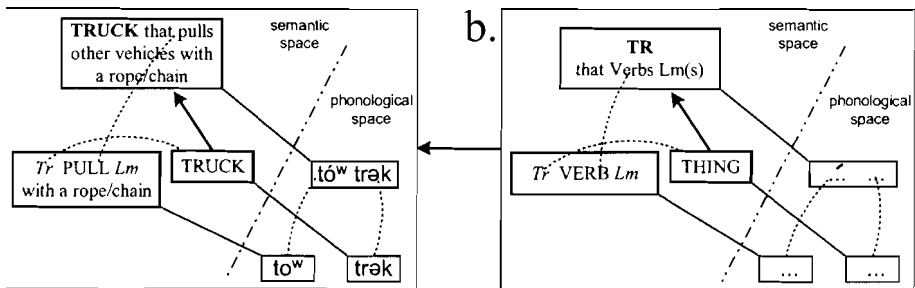
*Figure* 2: A hcadcd compound

Figure 2.b is a schema generalizing over *towtruck* and similar compounds such as *dumptruck, scrubwoman, cutgrass* or *prybar*.[7] It captures (*mutatis mutandis*) the defíning characteristics for this kind of compound as surnmarized in points 2-6 above (point 1 is also true of it and in fact of all syntagmatic structures). It is the CG equivalent of a rule or template defíning the class.

## 11.3 Independence, Unpredictability and Opacity of Composite Structures

Representing the composite structure separately from the components, as we do in Figure 2.a-b, is a recognition of the fact that composite structures, while certainly not unrelated to their components and perhaps initially not accessible apart from them, easily achieve an independent cognitive existence and acquire characstenstics not attributable to or predictable from the components. Components and their pattems of integration are better seen as scaffolding than as building blocks (Langacker 1987: 461). For instance, for most speakers a *towtruck* is used only for towing broken-down or crashed vehicles, and for some speakers it may actually be a truck with a tilting flatbed onto which such vehicles can be winched up and on which they are then carried rather than towed. It must be a truck specially modifíed for such towing: a pickup towing a car by a rope or chain (whether or not the car is broken down or crashed) is not a *towtruck,* nor is a truck which is specially built to tow a trailer, or which happens to be towing one. These characteristics pertain to the composite structure and are not attributable to the components. Such discrepancies can obtain at the phonological pole as well as at the semantic pole: the *scarecrow* noun *breakfast* is so pronounced as to violate in some degree the norms for both its cornponent words: one says [brE'kfəst] rather than [bré'kfæst].

Some degree of such specialization orpartial unpredictability, particularly at the seinantic pole, of *scarecrow* and *abrelatas* nouns, is absolutely typical, as should be clear in the many examples to follow. Such meanings are often described as "arbitrary", and there is a sense in which that is right. They are conventional. They are meanings of the words just because they are, not because they have to be. They are not arbitrary, however, if that means there is no reason for

them or reasoning behind them. Rather their idiosyncrasies are often reasonable in some degree once you come to think about them. For instance, it is not predictable that *matarratas* [kills-rats] should mean, as it does, 'rat poison' (but not 'rat trap') or 'rotgut whisky', nor is it predictable that *rotgut* should mean 'cheap whisky'. But one can see the whimsical reasoning behind the names, and it would be much more nearly arbitrary if they meant 'comb' or 'quark' or 'nostalgia'.

Such semantic or phonological idiosyncrasies are often taken as a proof of, or as a *sine qua non* for, lexical status. On the CG view this is mistaken. Lexicalization is best seen as simply the conventionalization and cognitive entrenchment of a fully-specified structure. Not only may perfectly regular formations become lexicalized in this sense, but, as Langacker (1987:452-456) points out, even novel structures often display characteristics underspecified by and unpredictable from their components and sanctioning patterns of formation. In other words, idiosyncrasies are neither necessary nor sufficient criteria for demonstrating lexicalization.

A composite structure may lose its connection to its components to the point where they are discerned only sporadically if at all. This is equivalent, in a diagram such as Figure 2, to the components gradually fading out of the picture, leaving only the composite structure on its own; in Langacker's metaphor the scaffolding has been dismantled, but the building remains. Such a loss of analyzability, or increase of opacity, as it is sometimes termed, afflicts a number of the English forms we discuss below, and relatively fewer of the Spanish forms. *Breakfast* is again a convenient example: many speakers of English have never thought of it as composed of *break* and *fast*. Most of what is said below of the *scarecrow* and *abrelatas* nouns is true only to the extent that they remain analyzable."

## 11.4 Sanction and Productivity

Productivity, under CG, is a matter of speakers utilizing existing patterns (i.e. schemas) of the language in a more or less direct fashion to structure and render intelligible novel forms whose semantic structures match the notions they want to cornmunicate. Schemas, to the extent that they are well entrenched and cognitively salient, *sanction* or legitimize their subcases. Relationships of partial schematicity afford weaker *partial sanction,* which is strengthened as fewer specifications of the sanctioning structure are contradicted by the target (i.e. as the partial schematicity more nearly approaches full schematicity). Sanction is also enhanced as the sanctioning structure is a close match (matching in many details) with the target structure. This means that, *ceterisparibus,* lower level schemas will sanction a form more strongly than high-level ones. As we have seen, a cognitive structure may be a subcase of more than one schema at the same time; it follows that a target structure may receive sanction from a number of different patterns simultaneously.

There is no threshold beyond which productive usage of a schema is completely impossible.' but as productive usage of a schema occurs repeatedly, the expectation of further such usage becomes attached to it. When we speak of productive schemas in the following

discussion we will be referring to such structures, one of whose specifications is expectation of productive use. In a convention not used elsewhere that I know of, I have marked such productive schemas with gray backgrounding in Figures 7-10."

## III. SPECIFIC EXAMPLES OF *ABRELATAS* AND *SCARECRO* WNOUNS

Scarecrow nouns include such words as breakwater, *catchfly*, *cureall*, dreadnought, *killjoy*, pickpocket, *spitfire*, and *spendthrift*. Abrelatas nouns form (currently) a numerically much larger class; they include such words as guardaespaldas [guards-backs] 'bodyguard', matamoscas [kills-flies] 'flyswatter', pasamanos [passes-hands] 'handrail', *quemacocos* [burns-coconuts] 'sunroof (of an automobile)'," andsacacorchos [takes.out-corks] 'corkscrew'. The basic structure of the *abrelatas* and *scarecrow* constructions is highly similar, which of course need not surprise us very much since they were chosen precisely for their similarity." Each consists of a transitive verb in combination with its object. And in both cases the compound as a whole designates neither the verb nor the object, but something or someone that does (or is used to do) the verb to the object.

Somewhat separate from or incidental to these definitional characteristics is the fact that in both cases the verb precedes its object.

I will use as a sort of informal shorthand to refer to these commonalities the formula V **+ O = S/I**, which means that a verb combines with a following element, its object, to designatc either its subject or instrument. (Similar formulas, hopefully self-explanatory, will be used for other, related, pattems as well).

### III.1 Specific Exarnples: *Espantapájaros* and *Scarecrow*
As is fitting in a bottom-up, usage-based model, we begin by looking at somc specific examplcs. Scarecrow and espantapájaros [frightens-birds] 'scarecrow, bird-scarer' are a closely matching pair of nouns from the two languages. Figure 3 displays them in a manner comparable to that of towtruck in Figure 2.a. As in that case, (1) the component words or stems and the composite structure are all bipolar. However, unlike the case with *towtruck* (2) neither componcnt semantic pole is schematic for the composite semantic structure. This is what makes these compounds exocentric. [13] (3) As with the notion of towing, the concept of frightening (the semantic pole of scare or of espanta) implies a trajector, something that causes fright, and a landmark, an animate thing that experiences the fright. (4) The profile of the second stem or word (the thing designated by crow or pájaros) is identified with the landmark rather than the trajector of the verb. This means direct object status for it. (5) The phonological components are again juxtaposed but not integrated as expected companions: these are compounds rather than stem-affix constructions.

As with towtruck, (6) the composite semantic structure in both cases designates the Trajector, the thing that does the frightening. It happens to be (though this would not have been predictable) an inanimate Thing. The specification of frightening birds away from a crop is also

common to both structures, though it is not predictable. (7) The semantic structure of pájaros remains essentially unchanged in the composite structure espantapájaros, but the semantic structure of crow, designating a particular kind of black bird, corresponds to the more generalized concept of a bird in scarecrow. Thus there is an arrow of schematicity from **bird** to CROW in the diagram. (8) Espanta (unlike scare) specifies a third-person trajector (subject), present tense, and indicative mood, and pájaros designates a group of birds. These specifications do not hold true of scare or crow. The morphological complexity of the Spanish forms (*espant-a-ø* [frighten-indicative.thematic.vowel-3rd.pers.sg] and pájar-o-S [bird-masculine-plural]) is not represented. (9) The Spanish compound is grammatically masculine." Since English does not have grammatical gender, the English compound naturally lacks that specification.[15]
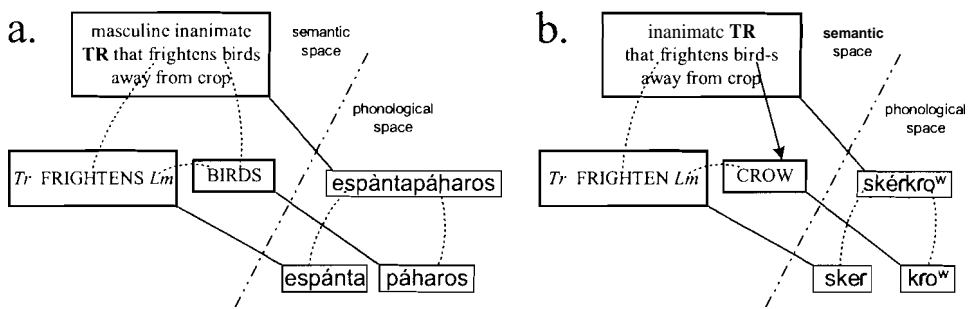


**Figure 3:** *espantapájaros* and **scarecrow**

One might question whether the designata of espantapájaros and scarecrow are subjects or instruments. The prototypical subject (trajector) in causational situations is an agent, a human being who purposefully and actively does something energetic that causes a process to occur. The prototypical instrument is an inanimate thing that is crucial to such an agent's bringing the process about and which is physically wielded by the agent to that end, but which could not or would not, by itself apart from being wielded, produce that result. The designata of scarecrow and *espantapájaros* can be assimilated to either category but are central members of neither. They are inanimate, and are actively set up by people who purpose that the process (of frightening birds) thereby occur, and to that extent they are instrument-like. However, the humans do not physically wield them and in fact are not even present when the process comes about, and the designata exhibit, apart from human intervention, attributes such as physical motion and shapes characteristic of animacy; such features give the illusion of animacy and that illusion is crucial to the birds' being frightened. To that extent they are "good" subjects.

The definitional (most highly schematic) characterization of "trajector", and thus of "subject", is "most salient participant (figure) in a relation (such as a process)" (Langacker 1987:217-220, 231-236, 1991b:9-10). Whether a particular verb allows non-prototypical

participants to attain that status is a matter of conventional construal. Both *scare* and *espantar* 'frighten' conventionally and easily allow inanimate frighteners, whether set up by humans or not, to be coded as their clausal subjects, so it is quite reasonable to consider them subjects in these cases. However, it wouldn't be unreasonable to consider them instruments, either. Such an analysis is likely to be enhanced in Spanish by the fact that the prototype for the *abrelatas* category is an instrument (cf. 9.a contrasted with 9.b). English, more than Spanish, as a rule permits instruments and other inanimates to function as subjects, especially those not immediately wielded by a human, in effect leaving any such human out of the picture. A number of *scarecrow* nouns designate such subjects *(scarecrow* itself, *dreadnought*, *breakwater*, etc. Under *CG* the two analyses need not be strictly distinguished, and even if they are, both can be entertained by speakers simultaneously.[16]

### III.2 *Prototypical Vs. Peripheral Meanings of Espantapájaros and Scarecrow*

From Figure 3 one might suppose that the meanings of *espantapájaros* and *scarecrow* are almost identical. In fact, however, they differ markedly, at least in my own speech (and I have informally confirmed this with other speakers of American English and Mexican Spanish), in that they have rather different prototypes. Practically anything used to scare birds from a crop can be called an *espantapájaros,* but the most typical *espantapájaros* is a strip of foil or other thin metal, or a rag or a piece of plastic, that is suspended where the wind can move it, and it is that motion, and sometimes an accompanying noise. that scares the birds. An inflatable plastic owl, or a human-like effigy, can also be called by the name, but they are not central to the category. For *scarecrow*, in contrast, the designatum almost has to be a human-like effigy. I can overcome a certain reluctance to call an inflatable plastic owl a *scarecrow*, but it is only with difficulty that I could call a rag on a pole, or a suspended piece of foil, by that name. I would typically use a hedged phrase like *a sort of scarecrow* if I wanted to describe one of them by the word. My wife says she might call them *scarecrows*, but it would only be because there isn't any other word for that sort of thing. The photograph in Fig. 4 can be unproblematically described as a *milpa con tres espantapájaros,* but much less easily as a *cornfield with three scarecrows*.

English glosses, whether in small caps or not, are less than fully adequate representations of semantic structures. Some aspects of meanings can sometimes be better represented by another less than fully adequate means, namely drawings. Figure 5 uses drawings to represent the facts described above. It shows only the semantic poles of the compounds: thus 5.a is an expansion or filling out in detail of the structure abbreviated as "masculine inanimate TR that frightens birds away from crops" in 3.a, and 5.b similarly expands on the corresponding abbreviated form in 3.b. Essentially, for the abbreviatory purposes of Figure 3 we used only the relatively schematic but still prominent structures 5.a.ii and 5.b.ii, but they are not the whole story. One should think of the phonological pole [espantapáharos] as conventionally linked to each of the structures in 5.a, but most strongly to the more prototypical structures, and similarly for the phonological structure [skérkro$^w$] in 5.b.

Figure 4:  Milpa con tres espantapájaros

Clearly on this view the semantic poles of *espantapájaros* and *scarecrow* are not single unitary structures, but in fact semantic hierarchies or networks of schemas and subcases in the style of Figure 1. Such polysemy is typical of lexical items. Note particularly that the high-level schema 5.a.ii is relatively prominent in the case of *espantapájaros,* but the corresponding 5.b.ii for *scarecrow* is rather less so. This records the judgment that (for me at least) this schematic meaning (anything that scares birds from a crop) is more prominent for the Spanish than for the English word. We characterized both *espantapájaros* and *scarecrow* as designating inanimates, and that is generally true. Both words have at least one kind of exception to this, however. In English a person can be conventionally called a *scarecrow* who is tall and skinny or who has raggedy clothing, i.e. who looks like a prototypical scarecrow. This metaphorical extension from the prototype of 5.b.i is represented in 5.b.iii. It is possible and likely that a schema containing the commonality of the two is extracted (it would essentially say "thing that looks human but is unusually skinny and/or ill-dressed"), but it is not necessary, and would probably be tenuous; in any case I have not represented such a structure here. Spanish also has such an extension for some speakers, though it is not represented in the diagram (the somewhat synonymous form *espantajo* has this meaning more strongly entrenched). Spanish, however, can also, peripherally but still conventionally, allow the word to be used of a person (usually a child) who is assigned to protect a crop from birds (5.a.iii.). This meaning is, I would judge, an extension of both the prototype (5.b.i) and the most prominent schema above it (5.b.ii); it is similar to the prototype in that a waving motion is likely to be used to frighten the birds, but unlike it in that the motion is not wind-induced. In any case, it is reasonable to suppose that speakers extract a schema like 5.b.iv to include it into the category. But (I would judge) such a schema would be much less prominent than 5.b.ii, since human bird-scarers are (in my experience) only rarely called *espantapájaros*.
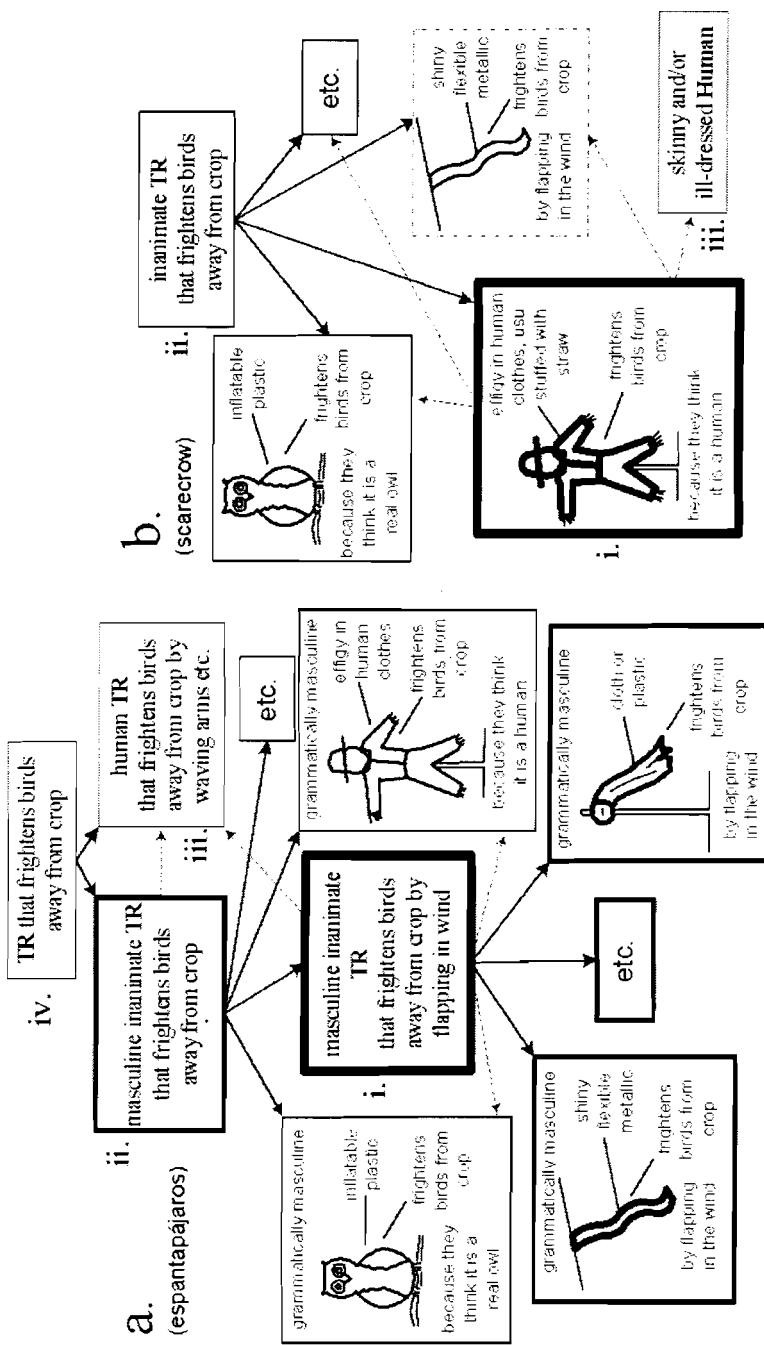
*Figure 5:* Prototypical *vs.* peripheral meanings of *espantapájaros* and scarecrow

One further point of note: where a human is designated by the Spanish form, and she happens to be female, the composite form will be feminine instead of masculine. Thus *la espantapájaros* would be 'the girl who scares birds from the crop', whereas *el espantapájaros* could be 'the boy who scares birds' but will more likely be 'the inanimate thing that scares birds'. This pattern of gender assignment for *abrelatas* nouns designating humans is general. In any case. this is why the specification 'masculine' is lacking in 5.b.iii-iv.

Figure 5.a-b gives an indication —quite inadequately, to be sure— of a considerable complexity of meaning in the specific examples *scarecrow* and *espantapajaros.* Such complexity is quite typical. We cannot (of course) take the time and space to replicate here the same sort of discussion for all the other examples we will be dealing with, but it is worth bearing in mind that the semantic structures represented by glosses or simply by the orthographic forms of the words are likely to be of equal complexity with these two, and to contain as many surprises, features that make sense but are not predictable from the component parts or the mode of their integration.

At the level of detail we have been discussing there is substantial likelihood that structures will differ significantly from one speaker or group of speakers to another. (It is not that such differences occur only at this level: they do so at more abstract levels as well. They are just less likely to show up in a given case at higher levels). For instance, it would not surprise me to find that for some English speakers a wind-moved foil bird-scarer is a better *scarecrow* than it is for me. I know Spanish speakers even in Mexico for whom a human effigy is the typical *espantapájaros,* and I am given to understand that this is likely to be so in Spain. Such sociolinguistic variation does not mean that differences such as those represented here are not important. Insofar as they are established for large and relevant groups of speakers they are important.

How do such distinctions come to exist? By speakers experiencing (and producing) different patterns of usage. And it is of course inextricably intertwined with cultural (in this case agricultural) considerations. In the cultures in which I learned Spanish people used wind-activated bird-scaring devices more frequently than effigies, and that naturally affected the meaning of *espantapájaros.* Similarly in my English-speaking culture human effigies were more commonly used to frighten birds, which helps explain why *scarecrow* has a different prototype. The prominence in the American culture of the Scarecrow in *The Wizard of Oz* probably has something to do with many Americans' strong sense that a good scarecrow ought to be stuffed with straw, and at least one Spanish speaker who thinks of an effigy as central to the *espantapájaros* category admits to likely being influenced by Spanish versions of that work.

In short, such variations are to be expected from one language to another, or to lesser degrees within a language, because different people in different cultures and situations have different experiences and make different choices with respect to language usage, and usage is the basis on which the cognitive structures constituting a language are established in speakers' minds.

## III.3 Non-Coincidence of Specific Examples

These same factors account for a striking set of differences between the *scarecrow* and *abrelatas*
categories, namely the fact that by and large specific examples do *not* coincide across the
languages. We have seen that *espantapájaros* and *scarecrow* do not coincide in all respects;
nevertheless they are exceptional in that they correspond as well as they do. Their component
stems and their meanings, their modes of integration, and their composite meanings, are quite
comparable. *Matarratas* 'rat poison, rotgut' and *rotgut* correspond in the designata of one of their
composite meanings, and in the mode of integration of the components, but the components
themselves are very different, and thus the background against which the composite designatum
(the whisky) stands out as figure is strikingly different (the notion of killing rats is not very
similar to that of causing one's intestines to decay). Perhaps the best-matched pair *ispasatiempo*
[pass-time] and *pastime*, which likely owe much of their remarkable similarity to both being
calqued on the French *passe-temps.* Three other good matches, again likely results of calquing, are
*cortabolsas* [cuts-bags/purses/pockets] and *cutpurse, rompeolas* [breaks-waves] and *breakwater,*
and *guardarropa* [guards-clothingj and *wardrobe." Lavacoches* [washes.cars] or *lava autos*
[washes autos]" are rather good matches with *car wash;* again calquing may have been involved.
But note that *car wash* is **O + V**, not V + **O;** thus it is a cousin rather than a sister of *scarecrow*
and the other forms in its category (see Section VIII).

    Most other specific examples do not match nearly as well. Typically a *scarecrow* or
*abrelatas* noun will be translated into the other language by some other kind of construction.
*Abrelatas* itself ('can-opener') or *matamoscas* 'flyswatter' have no corresponding scarecrow
nouns. A *rotnpecabezas* [breaks-heads] is a *puzzle,* not a *\*break-head* or a *\*bust-brain*, and a
*pasamontañas* [passes-mountainsj is a *ski mask.* Similarly, *apickpocket* is gcnerally called (at
least in Mexico) a *carterista* 'wallet specialist',ⁱ" and a *picklock* is a *ganzúa.* In each of these
cases and many others, someone in one language chose to coin a V + **O = S/I** compound and
others picked it up and used it until it became standard, while those speaking the other language,
for whatever reason, used forms of some other type to name the same kind of entity.

    In yet other cases, what is named with an *abrelatas* or *scarecrow* noun is culturally more
peripheral, to the point of being non-existent, in the community of those who speak the other
language. Most trucks in Mexico until relatively recently sported a protective grillwork in front
of the front bumper. Such a *tumbaburros* 'donkey-toppler' is not nearly so common in the United
States or (I presume) other English-speaking countries, and I do not know what to call it in
English: there seems to be no established term for it. (It is functionally similar to a *cowcatcher*
on a train, but is not really the same thing). Similarly the noun *sacadólares* [takes.out-dollars]
arose from a cultural situation in which people wanted to refer to rich people who would convert
their wealth into foreign currencies and bank it outside the country, protecting it from anticipated
devaluations or governmental freezing of assets. I don't know a really good English translation
for the noun, certainly not a standard one.

    Speakers, then, have the option of using a V + **O = S/I** structure, especially if their

language already provides a pattern or set of patterns for such forms, but there is nothing that forces them to do so. Often (and certainly for both Spanish and English in this case) a language will have a number of different patterns available to accomplish very similar communicative purposes, and speakers are free to choose among them what will best serve their purposes or suit their fancy.


## IV. CLUMPING AND SPOTTY PRODUCTIVITY
### IV.1 Clumps or Families of Forms
In both languages the attested examples occur in clumps or sub-families rather than just at random. English, for instance, has an inordinate number of scarecrow nouns in which the object is all: *catchall*, *cure-all*, carryall, *coverall*, be-all and end-all (but note that the first compound in this last example has an intransitive verb), and a raft of cornmercially-coined names such as *clean-all*, *copy-all*, *dispose-all*, *dust-all*, farm-all, *fix-all*, *hide-all*, *haul-all*, *lift-all*, *saws-all*, *sticks-all*, *store-all*, tote-all, etc. These forms constitute a robust clump or subfamily of *scarecrow* nouns.

This sort of clumping is cspecially evident for English if historical data are included. The OED lists no forms built on the verbs *add*, or  drag, or *flip*, for instance, but a number with stretch (e.g. *stretchgut* 'glutton', stretchneck 'pillory', *stretchrope* 'bell ringer', stretchhalter or stretchhemp 'gallows bird, one who deserves to be hung', stretchleg 'Death'). There were many forms with lack (e.g. lackbeard 'callow youth', *lackland* 'younger son', and lackall 'deficient person'), includinga half-dozen of the pattem *lack* + **noun.of.intellect** = **stupid person:** *lackwit*, *lackthought*, lackbrain, *lackmind*, *lacksense*, lacklearning, and (O intolerable slur!) *lacklatin*. Many forms meant 'miser', includingpinclipenny, *skinflint*, *clutchfist*, and (using nip) nipcake, *nipcrumb*, nipcheese, *nipfarthing*, and *niptoast*. Many terms (some still in use) meant 'criminal', including cutthroat, *cutpurse*, *pickpocket*, and turncoat, and, in a somewhat different social sphcre, *killjoy*, *spoilsport*, and tattle-tale, along with such picturesque obsolete forms as stretchhaltcr or stretchhemp (mentioned above) or *thatchgallows*.

Similarly, I know of no Spanish formations with compra 'buys' (and could find no clear examples in standard dictionaries)'"; for vende 'sells' I have found only two: *vendehúmos* [sells-smokes] 'one who claims to be able to dispense the favors of a powerful person', and vendepatrias [sells-fatherlands] 'traitor'.[21] There are, however, many forms with mata 'kills', whether listed in the dictionaries or not: e.g. matacandelas [kills-candles] 'snuffer' matamoscas [kills-flies] 'flyswatter', matamoros [kills-Moors] 'bravo, cocksure blusterer', *matarratas* 'rat poison, rotgut' etc.—my Pequeño *Larousse* lists 19, and there are plenty more that are not in the dictionary (e.g. *matahombres* [kills-men] 'knock-out (movie star or other much-desired female)' or *matahambres* [kills-hungers] 'hunger-buster, oversized meal'). The subpattem mata + **pests** = **poison** is quite freely productive; one can buy matabichos [kills-bugs] or mata-avispas or matacucarachas or *matahormigas* or *matalombrices* (wasp, cockroach, ant or worm poisons)

along with the (non-alcoholic) matarratas at many stores (of these only matarratas is listed in the Pequeño *Larousse*).

There is also a large family with para 'halts': e.g. paraguas [halts.waters] 'umbrella', paracaídas [halts-falls] 'parachute', pararrayos [halts-lightning.bolts] 'lightning rod', parasol [halts-sun] 'parasol'. A sub-family of the latter pattern names parts of an automobile that shield people or sensitive parts from dirt, wind, or more substantial obstacles: *parabrisas* [halts.breezes] 'windshield', *parachoques* [halts-crashes] 'bumper', paragolpes [halts-blows] 'bumper. especially for off-road vehicles, sometimes including a plate to protect the oilpan or other parts on the bottom of a car', orparafango [halts-mire] 'fender'. An interesting feature of this family is that for a good many speakers the preposition para 'for' is discerned instead of (or occasionally in addition to) the verb 'halts'. This makes these into $P + O = S/I$ structures parallel toparabien [for-good] 'best wish'. This is no problem in CG: the analyst simply, "bottom-up", records the fact and extracts the appropriate schema(s).

Cars have other parts which are named with abrelatas nouns, including *quemacocos* [bums-coconuts] 'sunroof', *descansabrazos* [rests-arrns] 'armrest', limpiaparabrisas 'windshield wiper' and *lavaparabrisas* 'windshield washer' (here incorporating one abrelatas noun within another), and so forth.

A number of other forms use lava or limpia and name laundry or cleaning products, including lavatrastes [washes-dishes] or lavaplatos [washes-plates] or lavavajilla [washes-china] 'dish soap' (or, 'dish-washing machine'), lavarropa 'laundry soap', *lavamanchas* [washes-stains] 'stain remover'; other lava-object compounds also occur, such as lavamanos [washes-hands] 'bathroom sink', *lavadinero* [washes-rnoney] 'money launderer', lavacoches and lava autos, both 'car wash', and so forth. Other 'detergent' forrns include arrancagrasa [yanks-grease], matamanchas [kills-stains], and so forth.

The saca-O family (*saca* = 'takes out, takes away, extracts, produces') is another robust one: sacabotas [takes.off-boots] 'bootjack', sacapuntas [produces-points] 'pencil sharpener', sacamuelas [takes.out-molars] '(inept) dentist', *sacacorchos* [takes.out-corks] 'corkscrew' sacamanchas 'detergent' (again), and many others.

Like English, Spanish has a 'criminal' family (e.g. asaltabancos [assaults-banks] 'bank robber', sacamantecas [takes.out-lards] 'disemboweler', *matapolicías* [kills-policernen] 'murderer of a policeman', *matoniños* [kills-children] 'infanticide' lavadinero [washes-money] 'money launderer' and others) including a family built on roba 'steals', including *robaniños* [steals-children], 'kidnapper', robacoches or robautos 'car thief', robaganado 'cattle thief'.

Many terms for birds fit the *abrelatas* pattern, including *picamadero* [perforates-tree] 'woodpecker' trepatroncos [scales-trunks] 'woodcreeper', saltaparedes [leaps-walls] 'wren', correcaminos [runs-roads] 'roadrunner', *tapacamino* [blocks-road] 'nighthawk', saltapalo [leaps-stick/tree] 'nuthatch'. The list includes at least three with chupa 'sucks' anda singular object, all meaning 'hummingbird': *chupaflor* [sucks-flower], *chupamirto* [sucks-myrtle], and chuparrosa [sucks-rose]. Relatively few animal names follow the pattem, though some do (e.g. *pasarríos*

[passes-rivers] 'crested lizard', *chupamiel* [sucks-honey] 'ant-eater', *tragavenado* [swallows-deer(sg.)] 'anaconda') (Schoenhals 1988).

In sum, it is clearly not the case that *scarecrow* and *abrelatas* nouns occur at random, with any and every transitive verb and possible object noun equally likely to be pressed into service or any kind of designatum for the composite structure. Rather they clump in families of different types, some of them cross-cutting others.

In the "bottom-up" spirit of CG we record such clumps where we find them, representing their commonalities in schemas. That is, we posit that speakers (at some cognitive level) recognize the similarities which hold within these sub-families, and extract schemas which embody them. Those schemas which have many subcases are more likely to be solidly entrenched, whereas those which have only a few (like V +*parabrisas* = **car part)** may be expected to be relatively tenuous. A few of these clumps in Spanish are represented in Figure 6; something similar for English may be found in Tuggy (1987). Almost all of the clumps should have a box marked "etc." in them—this representation is far from exhaustive even of the families it represents; it does not represent all the clumps discussed in the text, and those do not even begin to represent all the families or clumps there are. Also there are doubtless relationships of partial schematicity among the subcases of the clumps (see Figures 7 and 8) which are not represented in the diagram.
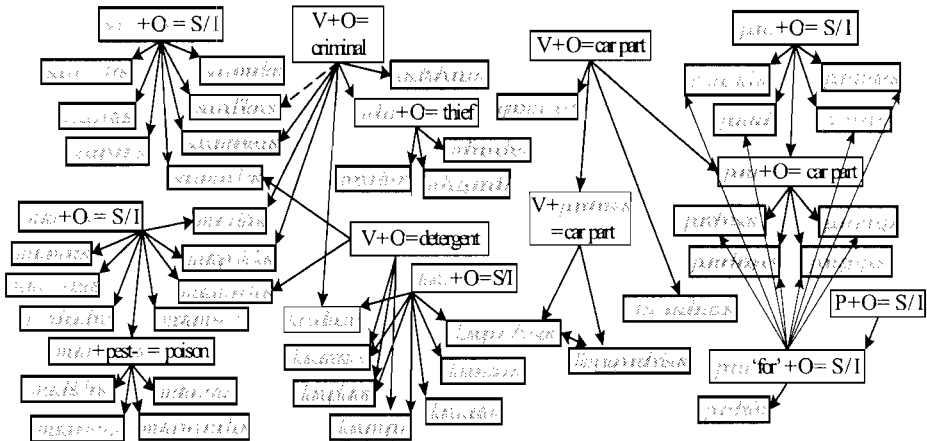


*Figure 6:* Sub-families (clumps) of *abrelatas* nouns

In Section III.3 we mentioned the striking fact of non-coincidence between the specific forms in the two languages. This is true of these clumps as well. Both languages have clumps, but the clumps do not necessarily or even usually coincide. Both Spanish and English have a V + **O =criminal** family (which may well be due to borrowing or calquing or other cross-linguistic

influences), but Spanish does not have as prominent a cluster of V + *all* forms, nor does English have V + O = **detergent** or V + O = **car part** or (despite the existence of *stopgap*) *stop* + N = I families. Such non-coincidence is to be explained in the same way the non-coincidence of the specific examples was explained; and in fact given non-coincidence of specific examples and the bottom-up character of language as viewed by CG, it is completely to be expected. Note however, that even when specific examples do not coincide a schema may. Thus most of the *scarecrow* nouns meaning 'criminal' do not have *ahrelatus* counterparts (*cutpurse* / *cortabolsas* being the exception), yet the two schemas correspond quite closely.

## IV.2 Spotty Productivity

We noted in Sections II.3 and III.2 that most *scarecrow* and *ahrelatus* nouns have meanings that are idiosyncratic or unpredictable in some degree. Whether or not they exhibit such semantic idiosyncrasies (and most do), the vast majority of the *ahrelatas* and *scarecrow* forms that one encounters are clearly already established. We call a *breakwater* by that name for the same reason we might call it a *jetty:* because that is an established word for that kind of thing. And a *parabrisas* [halts-breezes] is so called for the same reason it is called a *windshield* in English: because that is the established name for it.

Yet in both languages new forms are at least occasionally coined. The product name *CompactAll* (for a trash compactor) is a recent formation, and the coining of *toca-DVDs* [plays-DVDs] 'DVD player' clearly postdates the invention and naming of the DVD. In the May 2003 issue of the magazine *Muy Interesante* there is an article on *bacterias coniepiedras,* 'stone-eating bacteria', for which the word *comepiedras* [eats-stones], was most likely coined.

However, in neither language can just any and every transitive verb-object combination be used to form an automatically acceptable new compound. *??Drivetruck* and *??manejacamiones* seem about equally improbable to me: *Munge-All* for a computer program to systematically munge data, or *quemadiscos* [burns-discs] for a computer disk burner, both sound much more reasonable to me, and in fact a couple of examples of each showed up in an Internet search.

Of course, what counts as a productive or novel usage is not a cut-and-dried matter. A usage may be old hat to the speaker yet be unknown to the hearer, and thus effectively novel for the hearer. A speaker may re-invent a usage that others had invented before him; as long as he has not heard their coinages, or has completely forgotten them, the usage is still effectively novel for him. But of course conscious forgetting need not mean forgetting at all cognitive levels. And the same mechanisms of sanction that are important for a completely novel usage continue to be important for the establishment of a structure as it goes through the process of recurrent usage which entrenches it as part of the language.

It is a striking fact in English that only the V + *all* pattern, and in fact the V + *all* = **commercially advertised product** sub-pattern, seems to be currently productive. The relationship of such productivity to the "clumping" described in the previous section (IV.1)

would seem obvious. The reason there is a clump of V + *all* = **commercially advertised product** nouns in English is precisely because that pattern is being used to form new nouns. The same, we may be sure, happened in the past to produce the other "clumps".
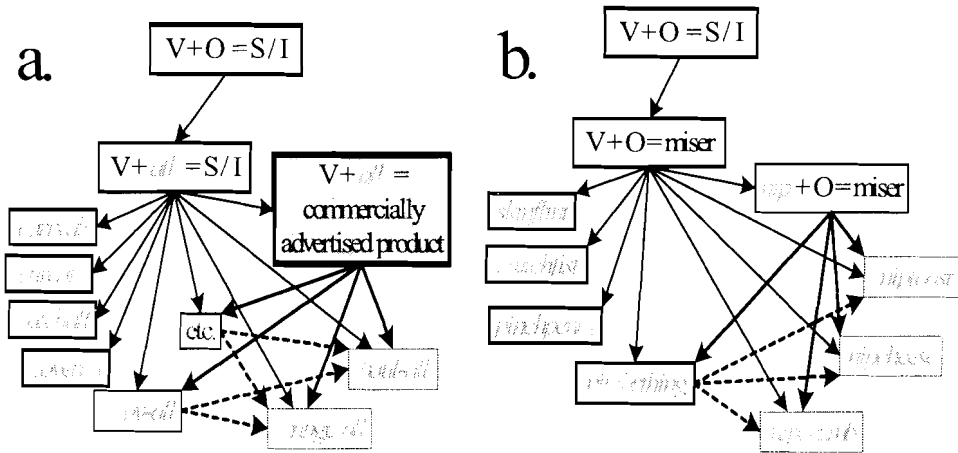


*Figure* 7: Productivity of V + *all* and *nip* + O = miser nouns

In Figure 7.a something like the current state of affairs for English, as I understand it, is represented. V + *all* = **commercially advertised product,** and secondarily **V + *all* = S/I**, are used to bcing uscd to categorize, and thus to sanction, novel structures (so their productivity is symbolized by the convention of a gray background). The forms *carryall*, *cureall*, *catchall* and *coverall* are very well-established, and so, to a lesser degree, is *copy-all* (along with other previously coined commercial product names summarized as "etc."). All of thesc receive sanction from V + *all* = S/I and (more distantly) V + O = S/I; and *copy-all* and **etc.** reccive further sanction from the lower-level schema V + *all* = **commercially advertised product.** The sanction from the latter schema is enhanced (as indicated by boldfacing the arrow), in the case of *copy-all* and etc., by the close, detailed fit of the schema to the target structure. This same sanction pattern obtains for the putatively novel forms *munge-all* and *haul-all* (both attested, but not widely known), and explains why they are likely candidates, relatively easily accepted as good English formations. Note too the partial but still appreciable sanction these novel forms receive from their already-established sisters *copy-all* and etc. The result of all this is the "clump" of V + *all* forms as we have it today.

Figure 7.b represents a plausible historical scenario for the V + O = **miser** category in English. Assuming that *skinflint*, *clutchfist*, *pinchpenny* and *nipjarthing* were established forms

and that *nipfarthing* was the first such formation with the verb nip, we can suppose that *nipfarthing* was first extended (in a partial schematicity relationship) to one of the novel forms niptoast, nipcheese, or *nipcrumb*, which was also sanctioned directly by V **+ O = rniser.** In the process, however, the extraction of the new schema nip **+ O = rniser** was facilitated, and it was involved productively in sanctioning the later novel forms. (This would be what CG would view as the very natural case of analogy coexisting with and mutually enhancing a nascent *rule*-governeo'productivity). In any case, the net result is the "clump" that examination of historical data reveals. It appears that the general pattern of V **+ O = S/I** noun formation was morc productive in that time period, which is why I have represented it as lightly grayed, but clearly even at that time most novel formations were formed with the primary sanction of productivc low-level schemas, hence the "clumping" we observe.

The abrelatas category is more like the English category was in centuries past. Many new nouns are being productively formed, usually in clumps of one kind or another. Some clumps are more strongly established and highly productive than others. Higher-level schemas are used productively as well at times, when novel formations not belonging to any particular clump occur (e.g. montacargas [mounts-loads] 'forklift' would be sanctioned by 10.a.i, and *trotamundo* [trots-world] 'globetrotter' by 10.a.iv), but usually their sanction is mediated and enhanced by the lowcr-level schema or schemas associated with a clump, and extension from neighboring specific examples is probably active as well.
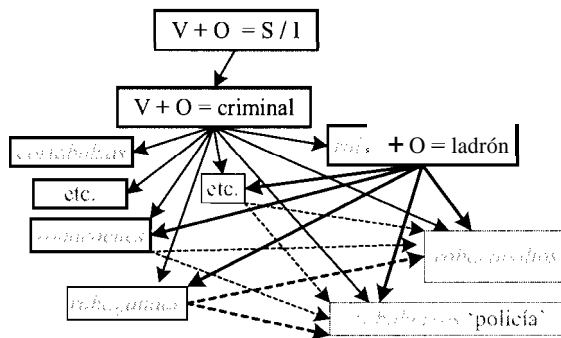


Figure 8: Productivity of *roba* **+** O nouns

Figure 8 represents attested (but not yet common) extensions of the roba **+ O = thief** pattern. Given the existence of that pattern, and supposing that someone wants to express the concept HORSE-THIEF, what would be more natural than to say robacaballos [steals-horses]? If no one ever happened to have invented the word before, this would be a fully productive usage, but the chances would be pretty good that someone, in fact numbers of people over the centuries, would have thought of it before, despite the availability of the word cuatrero 'horse thief, eattle thief. If it were I, I would by no means be certain that I had never, nor that I had ever, heard the

word before. In any case, the sanction the word would receive from the roba-Ns pattem and other abrelatas patterns would make it perfectly usable and understandable. Note that in Figure 8 the sanction from robaganado is represented as enhanced. This is because its meaning matches those of the novel forms exceptionally well: horses (caballos) and donkeys (burros) are kinds of GANADO 'cattle', and so at the semantic pole robaganado directly and closely sanctions robacaballos and robaburros;" it is only at the phonological pole that there is conflict of specifications. The somewhat surprising specialization of the meaning of robaburros represented in the diagram is apparently attested in certain parts of Ecuador.

## V. DIFFERENT PROTOTYPES FOR THE CATEGORIES
Just as differing usage has established different prototypes for scarecrow and espantapájaros (Section 111.2, Figure 5.a.i and 5.b.i), differing usage has established different prototypes for the scarecrow and abrelatas categories as wholes.

We observed of scarecrow and espantapájaros (Figure 3) that the Spanish word specifies a third person singular indicative verb, a plural object, and masculine gender of the composite, while the English word lacks those specifications. These differences also hold true of the prototypes." Both nouns in Figure 3 designated inanimate things; that specification holds prototypically in Spanish, but not particularly in English. A related difference is that the Spanish prototype is an instrument while in English it is a subject (though it may be an inanimate, instrument-like subject).
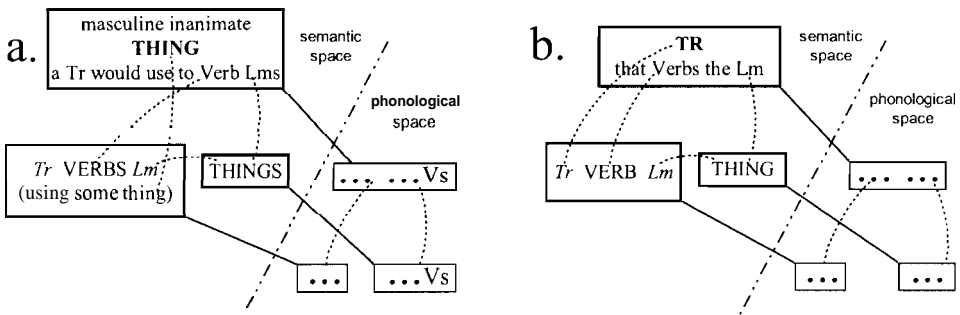


*Figure 9:* Protoiypes of the *abrelatas* and *scarecrow* categories

Figure 9 represents the two prototypes. Like the specific examples of Figure 3, they are bipolar (symbolic) structures, with the phonological specification that the schematic phonological form of the verb precedes the schematic form of the object noun. The Spanish construction also specifies (because it is common to the prototypical cases) that phonological forms of the noun

component and of the compound as a whole end in a Vowel-s combination, the –*s* being in fact
the plural marker. The component noun in each case is again specified, by linkage to the primary
landmark, to be the object of the verb. Again, the composite structure is not an elaboration
(semantic subcase) of either component, (as TOWTRUCK is of TRUCK in Figure 2.a), so the
constructions are exocentric. The composite Spanish structure in 9.a has a gray background
because it is productive, in contrast to the nonproductive English structure in 9.b.

Figure 10 represents the prototypes in relation to some other subcases of the two
categories. 10.a.i. is the prototype for the Spanish category, an abbreviation of 9.a, and similarly
10.b.i. is an abbreviation of 9.b. Note that the schemas based on semantic and lexical
commonalities which we discussed in Section IV (diagrams 6-8) are not represented here: they
are to be thought of as coexisting with the constructionally or syntactically based ones mentioned
here, not supplanting them, and coinciding to sanction specific forms. Under CG they are not
different in kind, though there are differences of degree that allow one to distinguish the two
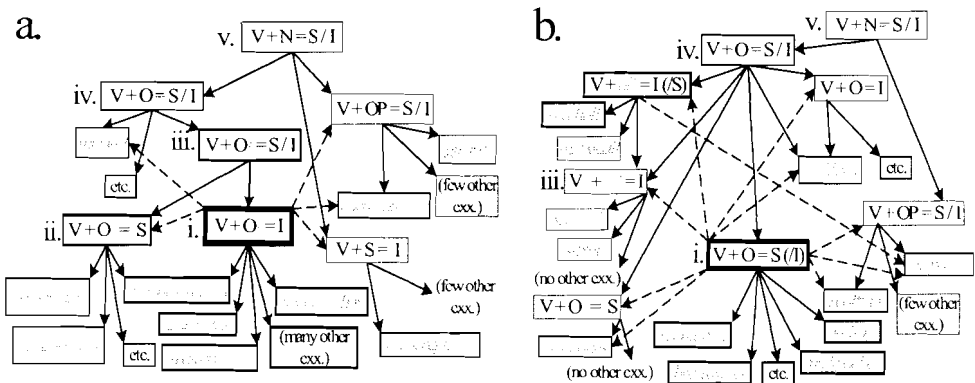categories.



*Figure 10:* The *abrelatas* and *scarecrow* noun categories

The plurality of the object in Spanish and the inflection of its verb will be revisited in
Section VII.

## VI. OTHER PROTOTYPICAL CHARACTERISTICS

There are other characteristics which one might consider including in a prototype for the
categories: they are at least prototypical in the sense of being rather common and helping a
proposed *abrelatas* or *scarecrow* noun "sound" or "feel" better than the alternatives. But they
vary quite independently of the characteristics represented in Figure 9 and of each other, and thus
are here considered separately.

## VI.1 Deprecation and Jocularity

Many *scarecrow* and ahrelatas compounds exhibit a couple of characteristics having to do in part with the attitudes the speakers using the words are likely to have towards the things the words designate.

Especially when a scarecrow or *abrelatas* noun designates a human being there is usually a perceptible, sometimes a strong tinge of deprecation, condescension, contempt, or ridicule conveyed. You do not often get either *scarecrow* or ahrelatas nouns designating respected persons, or if you do the term is pretty sure to be disparaging or poking fun at them. *Pickpockets*, *turnkeys*, *spitfires* and *skinflints* are not held up for our admiration; nor is a *matahombres* 'mankiller (attractive woman)', a rohacoches 'car thief', a buscapleitos [seeks-quarrels] 'quarrel-monger', a *sacadólares* 'person who transfers money from the national economy to overseas bank accounts', a *sacamuelas* 'inept dentist', or a *matasanos* [kills-healthy.ones] 'quack'. The fact that both languages have a V **+ O = criminal** pattern is obviously related to this tendency. There are exceptions, of course: a .salvavidas [saves-lives] 'lifeguard' (also 'life preserver') is not denigrated by being so named, and a guardameta [guards-goal] 'goalie', if he does his job well, is likely to be positively looked up to. There may be more awed admiration than anything else attendant on the naming of a *tragavenado* [swallows-deer] 'anaconda'. But these are rather exceptional cases. The note of contempt or ridicule can show up strongly even in a term for an inanimate item, as in *taparrabo* [covers-handle/end] 'butt-cover, i.e. loincloth, tight shorts, over-short skirt'.

The use of these nouns to ridicule and poke fun fits in with a note of jocularity that also shows up very frequently. Many of the forms are just plain fun, often fun with a sting to it, but fun nevertheless. There is a kind of whimsical or fantastic sense of poetry in them, an exuberance which is sometimes reflected in an prodigality of words with the same meaning, colorful terms proliferated for no apparent reason other than just for the fun of it. It makes prosaic sense to comment on the fact when someone spends less money than might be expected or wished, and there is a certain satisfaction in describing him as "stingy", but how much more fun to call him a *skinflint*, apinchpenny or a *nipfarthing*.[21] (Cf. the other terms mentioned in 0 for 'miser', or for 'stupid person'). It is all right, but rather boring, to call a framework of pipes protecting the front of your truck aprotector 'protector' or a reja 'grill', but there is more *chiste* 'fun, point, spark' to it when you can call it a *tumbaburros* [knocks.over-donkeys] 'donkey-toppler' or a *mataperros* [kills-dogs] 'dog-killer'. You can call a speed bump a tope 'bump', but it sounds more colorful if you call it a rompemuelles [breaks-leafsprings] 'spring-buster' (or a policía muerto 'dead policeman'. Contrast this with the bureaucratic ungainliness and insipidity of the governmentally preferred *reductor* de velocidad [reducer of velocity] 'speed-reducer'). Or why would you call a party noiseniaker that uncurls and sticks out when you blow on it anything else, once you've thought of calling it an espantasuegras [frightens-mothers.in.law] 'mother-in-law scarer'? When my daughter's high-school friend called her photo credential which had fallen on the floor an *espantacucarachas* [frightens-cockroaches], it was quite in the cards that his coinage should elicit

    

visible and audible manifestations of delight froni several bystanders.

This kind of color and "sting" naturally enhances the salience of a form. CG predicts that forms nianifesting it would be more quickly and easily entrenched and conventionalized, even in the absence of much repetition, than more humdrum expressions. That indeed is the case: these words tend to be memorable.

Figure 11 illustrates, for Spanish, how these prototypical specifications can interrelate and coincide in specific examples or lower-level schemas with the other types we have been examining. It should be clear that similar structures could (and should) be posited for English as well. It is also clear that many Spanish examples, particularly of the highly prototypical classes of everyday implements and household substances for common purposes such as *matamoscas* 'flyswatter', *matahormigas* 'ant poison' and *abrelatas* 'can-opener', are neither particularly despised or particularly laughable.
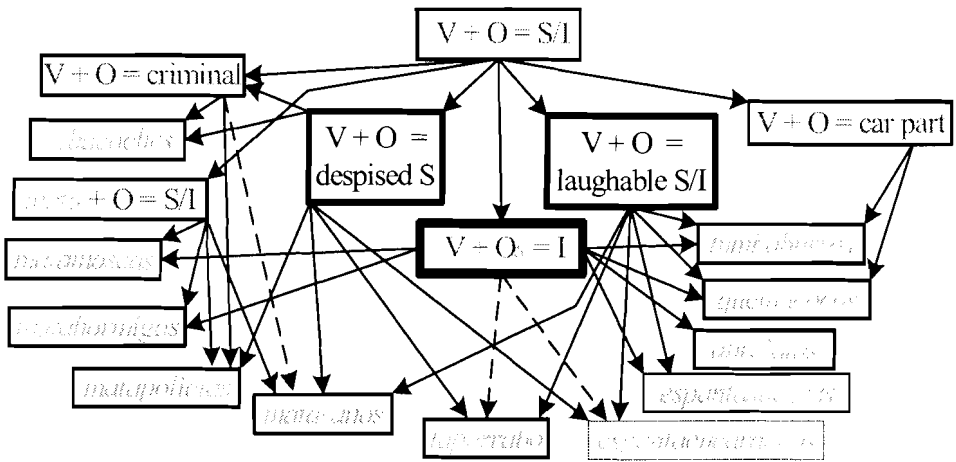


*Figure 11*:  Deprecation and jocularity in the *abrelatas* category

It is noteworthy that both languages conventionally associate these kinds of attitudes with their $V + O = S/I$ pattems more than with other pattems. I suspect it is in some degree a product of common origin and/or of language-contact phenomena; the association of this sort of "aura" with this sort of construction may be a kind of European areal feature. On the other hand, I would not rule out the possibility that there is something about this particular kind of structure that lends itself to such construals, perhaps a kind of intrinsic humor in the abruptness of the headless construction, a seniantic voltage induced by juxtaposing verb and noun without the customary clausal trappings to insulate and mediate their valence, a pithiness and pungency of short specific lexical items which is an antidote to the mealy-mouthedness that lurks in derivational

morphology. Or perhaps it results from the incongruity of applying to humans a construction associated with instruments, in effect identifying them by what they're good for rather than who they are. It would be an interesting area for further investigation.

## VI.2 Prototypical Phonological Specifications

We have seen that having the object end in s is a typical phonological specification of the Spanish construction. Both languages have their favorite syllabic structures as well. The Spanish norm is clearly to combine iwo disyllables, each with penultimate stress, with the primary stress on the rightmost of the two. That is, the pattern is 'SS-'SS for Spanish. In English monosyllabic components are preferred, and the stress is on the leftmost of the two: 'S-S. Thus *abrelatas, matamoscas, tumbaburros, rompeolas,* and dozens of other examples are prototypical for Spanish (they are pronounced [abrelátas, màtamóscas] and so forth), and *scarecrow, killjoy, rotgut, picklock, spitfire, breakfast* and a number of others fit the English prototypical pattern. Substituting polysyllabic synonyms makes these forms sound very improbable: *\*frighten-raven, \*eliminate-happiness,* or *\*decay-intestines* would never do.

Both of these patterns characterize other kinds of compounds as well (e.g. *pelirrojo* [hair-and?-red] 'redhead', *compraventa* [buys-sells] 'commercial transaction' *agridulce* [sour-and?-sweet] 'bittersweet'; *redhead, bigshot, blackboard,* etc). This does not mean there is not a specification attached to the *abrelatas* or *scarecrow* construction itself, but that if there is, it is a subcase of a schema which applies other types as well.

In any case, exceptions are tolerated, and are not even uncommon. For instance *espantasuegras* 'mother-in-law-scarer' and *limpiaparabrisas* 'windshield-wiper' have, respectively, a polysyllabic verb and noun, and *espantapájaros* and even more the nonce formation *espantacucarachas* have both; *mirasol* 'sunflower', *parasol* 'parasol, umbrella for shadc', and *chupaflor* 'hummingbird' have monosyllabic objects. Similarly, English forms like *pickpocket* or *tattletale* have disyllabic rather than monosyllabic components.

It should be clear that a structure similar to that in Figure 11 would allow schemas embodying these prototypical patterns to apply to the appropriate specific examples in the category.

## VII. FUNCTIONAL MOTIVATION AND PERSISTENCE OF PATTERNS

The plurality of the object in the prototypical Spanish patterns, and the non-plurality of the object in the corresponding English patterns, form an interesting contrast. There are motivations for both patterns, contradictory as they are. The plurality in Spanish corresponds, native-speaker intuitions persistently indicate, with the plurality of actual or expected objects. Why is *latas* 'cans' plural in *abrelatas* 'can-opener', or *ratas* 'rats' in *matarratas* 'rat poison'? Because one expects a can-opener to open multiple cans, and hopes that rat poison will do away with whole colonies of rats. It makes sense.

It also makes sense that the verb is in its third-person singular form. When you are naming a can-opener or rat poison, the thing you are talking about is neither speaker ($1^{st}$ person) nor hearer ($2^{nd}$ person), but is known to both speaker and hearer, thus third person. It also makes sensc to use the simple prcsent indicative form, since that is the standard form for expressing unbounded, habitual action.

However, the English patterns make their own kind of sense. English has a preference for using the basic forms of stems for forming compounds. These are generally the shortest and most characteristic forms of the stems in question, which adds a kind of motivation of economy of effort to using them. So despite the motivations that produce the plural and third-person present indicative markings in Spanish, English generally lacks them, except in a few peripheral cases such as *sawbones* and *sawzall*.[25] It is not that the birds frightened by a *scarecrow* are not just as likely to be plural as those frightened by an *espantapájaros*, or that English speakers fail to realize this probable plurality: it is just that the strongly predominant (but not absolute) English pattern refrains from overtly marking the fact.

The motivation for pluralizing the object breaks down in a couple of well-defincd cases, and this naturally results in systematic exceptions to the pattern in Spanish. One case is wherc the object is a mass rather than a count noun: mass nouns generally do not have plurals except under special, non-mass construals. It is for this reason that *pasatiempo* 'pastime' is not *\*pasatiempos*:[26] the time that is caused to pass is construed as a mass. Similarly a picahielo [picks-ice] 'ice-pick' is not a *\*picahielos* i    f it were it would mean 'icecube-pick'—; and a guardapolvo [protects-dust] is not a *guardapolvos, and a tragaluz [swallows-light] 'skylight' is not a *tragaluces, as if it swallowed many little lights. The second case is where there is a single unique object. This explains why a mirasol [looks.at-sun] or girasol [tums-sun] 'sunflower' is not a *\*mirasoles* or a *girasoles, and (perhaps) why a lighthouse-keeper is a guardafaro 'guards-beacon' rather than a ?*guardafaros* 'guards-beacons/headlights'.

Sometimes it is not clear whether a mass or a count construal is more appropriate, and some nouns appear in both forms. Is a papenveight a *pisapapeles* [tread/press.on-papers] or a *pisapapel* [tread/press.on-paper]? Either construal makes sense, and both forms exist, each favored in different varieties of Spanish." Does a roadrunner run the roads (correcaminos [runs-roads]) or is it only the one road that 'goes ever on and on' (*correcamino* [runs-road])? Is a lightning rod apararrayo [halts-lightning] or a *pararrayos* [halts-lightning.bolts]? It can be said either way.

This is the type of motivation dear to many linguists who accept the label "Functionalist". It offers real, but not absolute, explanations of why speakers have established the pattems they have.

But an important limitation on it is that a pattem, once established, can be applied beyond the scope of its original motivation. As LaPolla (to appear) puts it, "our language use is a set of habits we form, […] and once we have a habit, it is hard to change, including habits of language and even thought". The persistence of English non-plural forms for the object of *scarecrow* nouns

even when plurality is clearly understood is a case in point. And Spanish similarly shows plural morphology in a number of cases where by the functional motivation one would expect singular forms.

I would have expected a snow-fence to be a \**paranieve* [halts-snow], since nieve 'snow' is a mass noun in most usages, but the dictionary listsparanieves instead. I would have expected an icebreaker (boat) to be called a \**rompehielo* [brcaks-ice], but the standard form is, instead, *rompehielos* [breaks-ices/ice.cubes]. An automatic door closer will (at least normally) only close one door in its lifetime, but it is nevertheless a cierrapuertas [closes-doors] and not a \**cierrapuerta*. A vendepatrias [sells-fatherlands] 'traitor' will normally have only one fatherland that he can sell out, and although a globe-trotter can be called a *trotamundo* [trots-world] —naturally so, given that there is only one world that he is ever likely to trot—, the form trotamundos is also established, and in fact is considerably more common.'"

Given this pcrsistence of the –*s* where it is not strongly motivated or even expected, one might expect it to appear on the word for 'bedspread'. However, the word is cubrecama [covers-bed], not \**cubrecamas*. Similarly the dictionary lists the word for a loincloth (or tight shorts or a short skirt) as taparrabo [covers-handlelend], not taparrabos (though the latter showed up four timcs as often in an Internet search). Is the reasoning that only one bed or rear end (or other "handlc") at a timc is expected to be covered by the designatum, and usually the same onc on repeatcd occasions? Similar reasoning could explain why a goalie is a *guardameta* [guards-goal] and not (usually) a *guardametas*: the goalie defends one goal during a game and quite emphatically not the other. Even in different games it is still his team's goal, describable as a single cntity, that he defends. But by the same reasoning you would expect \**guardaespalda* with a singular rather than the established form *guardaespaldas*: after all, a bodyguard generally guards only one person's back at a time and usually the same one consistently, and most other examples are similar in some degree. Even a 'can-opener' abrelatas only opens one can ata time. When you call a hummingbird a *chupamirto* [sucks-myrtle], the singular object could be a generic noun denoting the species, but such a construal, at least for me, is considerably harder with chuparrosa [sucks-rose] and even more difficult with *chupaflor* [sucks-flower]. The plurality of the flowers that the hummingbird sucks is salient, yet these forms all take the singular noun.

The conclusion is that (1) the way a form is conventionalized can override thc influence of rules or motivations that might have led one to expect something different, (2) the motivations themselves can be contradictory to each other.

Persistence of the well-established pattern shows up clearly when the normal, statistically predominant 3$^{rd}$ person singular present indicative construal is contextually contradicted. The most common case is when the subject is contextually plural: the verb continues to be marked as for a singular subject. The forms bacterias *comepiedras* 'rock-eating bacteria' and cables *pasacorriente* 'jumper cables' simply cannot be \**comenpiedras* [they.eat-rocks] or \**pasancorriente* [they.pass-current]. Similarly with changes of person, tense or mood: if I say

    

that I am a bodyguard or that you might be a lifeguard, I cannot say I am a *\*guardoespaldas* [I.guard-backs] or that you would be a *\*salvariasvidas* [you.would.save-lives]. The forms were coined as third person singular present indicatives, and they stay that way.

It shows up extra spectacularly in the exceptional cases where instead of an object a subject follows the verb. In *escurre-platos* [drains-dishes] 'dish drainer' it is the dishes that *escurrir* 'have liquid run down [self] in rivulets', but the form is neither *\*escurrenplatos* [they.drain-dishes] nor *escurre-plato* [drains.dish]. Similarly for *trepamonos* [clambers-monkeys] 'jungle gym, monkey-bars' you say neither *\*trepanmonos* [they.clamber-monkeys] nor *\*trepamono* [clambers-monkey]; and *Cantarranas* 'Frog-croak Creek' is neither *\*Cantanrranas* [they.sing-frogs] nor *\*Cantarrana* [sings-frog]. The verb remains singular despite having a plural subject, and the post-verbal noun remains plural despite being understood as subject of a singular verb.

The moral of the story is that functional motivations, useful and relevant though they may be to explanation, do not tell the whole story: sheer conventionality of patterns also comes to bear. Perhaps the best way to say it is to say that there is a functional pressure to persist in using previously-established patterns even when they don't fit the case at hand very well."

## VIII. THE POSITION OF THE *ABRELATAS* AND *SCARECROW* PATTERNS IN SPANISH AND ENGLISH

The *abrelatas* and *scarecrow* patterns have somewhat different statuses in their respective languages.

The *scarecrow* pattern is (currently — the picture was rather different a few centuries ago) a relatively minor one, with only residual productivity (largely confined, as we have seen, to the **V + all = commercially advertised product** subcase); it is even somewhat quaint. One might without great difficulty speak English for days or weeks without ever using a *scarecrow* noun. The pattern's function, that of naming a Thing by an activity characteristic of it, is usually fulfilled, especially in productive usage, by another construction, the **(O +) V + -er** construction. The *abrelatas* pattern, by contrast, is robust and quite highly productive, and is one of the major patterns for naming certain large classes of culturally important items which are usefully identified by a characteristic activity."

We have considered in some detail the internal complexity of the constructions (as depicted in Figures 3-11), in both languages the constructions fit together with other, similar constructions, into yet more complex categories or families of constructions." In both languages, for instance, there are related constructions in which a **V + O** combination designates an action or occasion rather than a subject or instrument (e.g. *breakfast* or *ceasefire* or *cumpleaños* [completes-years] 'birthday'), or where it names something more like a place than like an instrument (e.g. *wardrobe, descansa-pies* [rests-feet] 'footrest'). English has **V + O = adjective** structures (e.g. *lackluster, catchpenny);* Spanish nouns and adjectives generally overlap more

than English ones do, and many of the forms we have seen can be used adjectivally (e.g. bacterias *comepiedras* 'rock-eating bacteria', *polvo* lavatrastes 'dish-detergent powder'). English also has forms that are used both ways, e.g. *cutthroat (competition)*, stopgap *(measures)*, *lickspittle (followers)*. Both languages have structures where a preposition occurs instead of the transitive verb ($P + O = S$) (e.g. *afternoon* or *overall(s)* or dounstairs, or anteojos [before-eyes] 'glasses' or parabién [for-good] 'felicitation, best wish') or sinvergüenza [without-shanie] 'brazen-faced scoundrel'. (We noted in 0 that the para-O [halts-O] family belongs to this category for some speakers). Both languages have structures in which more than two words are combined in a $V + O$ structure, such as *know-it-all* and sabelotodo [knows-the/it-all] 'know-it-all' or curalotodo [cures-the/it-all] 'cure-all'. Both, as previously mentioned, have well-established patterns for naming things by a characteristic activity using an agentive suffix like d o r or *-ero* or *−er*. And so forth.

But the inventories of such structures, and their relative prominences in the two languages. do not coincide in all cases. English has a greater range of relatively common compound types than Spanish does. For instance, there are a number of forms like *scarecrow* nouns except that the $V + O$ order is reversed (windbreak contrasts nicely with the *scarecrow* noun *breakwater*: other examples would be c a r *wash*, *chimneysweep*, watershed, and *doorstop*). Such examples are rare and marginal at best in Spanish." English also has $V + P = S$ structures (*turn-off*, *go-between*), and $V = S$ structures (cheat, *hore*, cook, *sneak*), and (headed) $V + O = O$ and $V + S = S$ structures (push-pin, draw-bridge, *pull-toy*; towtruck, *scrubwoman*, *prybar*, cf. Figure 2.b), even a few $V + S = O/Loc$ structures (*Godsend*, cowlick); these all are rare if not nonexistent in Spanish."

Figure 12 represents the position of the scarecrow pattern vis a vis some of these related constructions." The scarecrow pattern, as this network indicates, is by no means the most prominent kind of English compound; those honors go (among the structures represented here, which of course do not represent all types in the language) to the right-headed compounds and to the V-er constructions. *Scarecrow* is nonetheless well-connected, being a first cousin to the following compounds. each representing a different type: lackluster, windbreak, *know-it-all*, *overall(s)*, *shuteye*, *picklock*, pushpin, and towtruck. Another way to talk about the same configurations is to note that the $V + O = S$ pattern represenis the confluence of at least six slightly more abstract patterns. Scarecrow is related somewhat less closely, but still definitely related, to all the rest of the specific forms in the diagram, each again a representative of a different type. Once again, a network structure of schemas and their subcases helps us explicate the relationships.

The relationships of the *abrelatas* construction to the other Spanish types would of course fit in a similar diagram; similar in that it would have the same general character of schema-mediated relationships, though somewhat simpler in that the abrelatas pattern does not have quite so many relatives as does the *scarecrow* pattern.
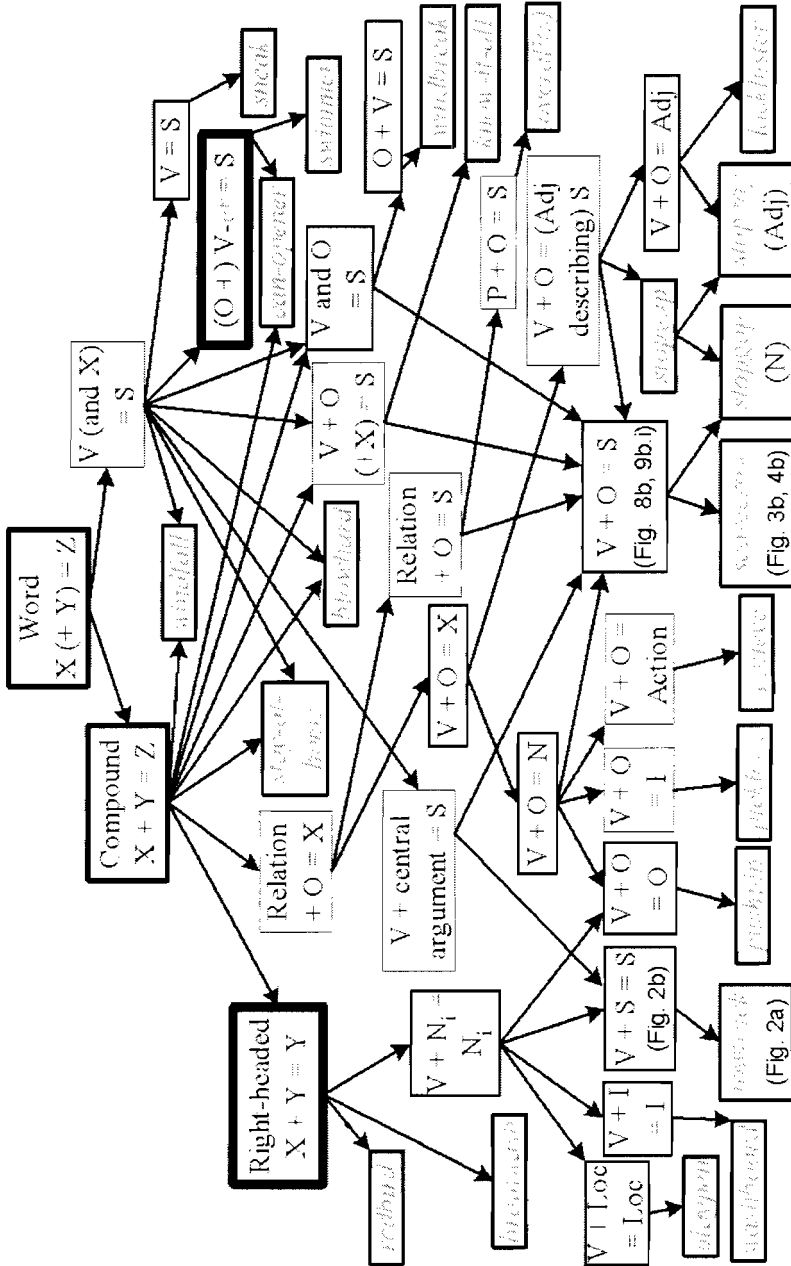
Figure 12: *Scarecrow nouns related to other construction types*

## IX. SUMMARY AND **CONCLUSION**

Spanish and English both have exocentric V **+** O **= S/I** compound constructions, of which *abrelatas* and scarecrow are typical examples. The categories are similar in their general constructional pattern, in consisting of "clumps" or subfamilies of forms, and in often having a negative andlor jocular and colorful tone to their meanings, but they differ in a number of ways. The individual items fail to coincide across the languages more often than not, and so do the clumps or subfamilies. The prototypical structure in Spanish has a plural object and a $3^{rd}$-person singular present indicative verb; English typically has neither. The Spanish construction is widely, though spottily, productive, and constitutes one of the major mechanisms of the language for naniing instruments, whereas the English construction does not favor an instrumental interpretation, is a minor pattern in the language, and is currently productive only in one of its subfaniilies (V **+** *all* = commercially advertisable product). The Spanish prototype consists of two disyllabic components of which the rightmost is stressed, whereas the English pattern prefers monosyllables and stresses the leftmost. Each language has exceptional or peripheral forms which contrast wiih its prototypical specifications, in some cases approaching the other languagc's prototype. And in both languages the category as a whole fits into wider families or catcgories of constructions, but those wider families are not the same in their membership.

Many of thcse insights have been facilitated by the basic structure that *CG* assumes is common to languages, and would be difficult to achieve or express under certain other models of grammar. Particularly models that posit innate, absolute, relatively simple and rclatively few rules can fit data of this sort uncomfortably at best. (1) The bottom-up, usage-based character of language as *CG* sees it makes the non-coincidences natural and acceptable, and encourages thc analyst to capturc all thc generalizations ihat are there to be captured. (2) The multitudinous, cross-cutting patterns do not supplant or contradict each other, but coexist peacefully. (3) Low level patterns tend to bc more important than morc abstract pattcrns in sanctioning novcl formations, and thus tend to form ihe ccnters of subfamilies. (4) The choices people madc in establishing those patterns are reasonablc, and reflect different kinds of functional motivations. Ncverihcless, (5) oncc a pattern is established it can bc adhered to cven when that functional motivation is no longer present. (6) The same data may support contradictory analyses, as expressed by incompatible patterns being used to categorize them. All thcse features of languagc are natural undcr *CG,* which makes it a useful framework under which to considcr data of this sort.

As I conteniplatc the abrelatas and scarecrow nouns, I am moved by a sense of gratitude, of apprcciation for the beauty of these constructions, for the creative imagination and quirky humor that they express. It is data like these that make linguistics a constant source of refrcshment. It is fun and profitable to think through how to analyze ihem, and the products of analysis may achieve their own kinds of beauty, but it is the data themselves that give us greatest cause to rejoicc. They are a gift from generations of speakers past and present to thosc of us privilcged to speak the languages of which they are a part.

## NOTES:

[1] The phenomenon is probably cognate in the two languages. It seerns to have become a prorninent feature of Romance during the ninth century, though isolated exarnples occur as early as the fourth century (Jirnénez Ríos 2001:326). It likely entered English with the French influx during the period of Norman hegernony, and seerns to be rare in German and Dutch, though forms like *Habe-nichts* [have-nothing] 'have-not' and *veetveel* [know-much] 'knowledgeable person' or *doe-al* [do-all] and archaic forms like *Traugott* [trust-God] '(old-fashioned given name, surname)' do occur (Petra Schroeder, Oliver Stegen, Jaap Feenstra and Arie Verhagen, personal cornrnunications), and Germanic roots are very cornrnon in the English forrns. But given the history of bilingualism and intercultural exchange between English and French, and French and Spanish, and Italian with all of the above, later intercultural aiid interlingual influences rnay have been equally important.

   Sanskrit, as described by Panini, had an exocentric compound *bahuvrihi* [niuch-rice] 'place where ricc abounds', and *scarecrow* nouns have sometirnes been said to be a kind of *bahuvrihi* cornpound. The structures are similar in their norninality and exocentricity. but are rather different in rnost other respects. In the shorthand ernployed beginning in Section II.4, *baliuvrihis* are $A + N = $ **Possr** cornpounds. Their existence would certainly not justify positing $V + O = S/I$ for proto-Indo-European.

[2] Spanish examples are generally followed, at least at their first occurrence, by a literal two-word gloss between square brackets and a word-level gloss in single quotes.

[3] Speakers "know" their language in the sense that they *know how to* "do" it, not, by and large, in the sense that they *have knowledge of* it. "All this is old hat, and is exactly why linguists traditionally have chosen to spcak of habits or skills [r]ather thaii of knowledgc. Chornsky inveighs against 'habit' on the grounds that it has no established sense in which it can explain language cornpetence [...] Perhaps the point is cogent on the assurnption that a language is a well-defined systern [...] Of course, rncrely to say that a language is a systern of habits is not to say nearly enough: we rnust spell out what kinds of habits, and what kind of system". (Hockett 1968:63-64). Newineyer (1998), writing from the generative tradition, also allows usage considerable intluence over grarnrnatical systerns.

[4] The number of specific example types (i.e. how rnany differcnt specific exarnples thcre are) inattcrs, but the token-iiumbcrs for those specific examplcs (how many times the sarne specific exarnples occur) inay rnatter evcii more. Iii any case, as we have said, any irnbalance of saliencc of the exarnples when used enters iiito thc equation as well as nurncrical predominance.

[5] Linguistically-ininded speakers rnight well come up with such a structure, but, until they used sornc specific exarnples of it, it could only be entrenched in their rninds by solitary usage, not yet conventionalized by communicativc usage.

[6] Tliere is probably a kind of priining or facilitation to be recognized in that the word *tow* rnight bring the word *truck* to rnind. The reverse association leading frorn *truck* to *tow* is probably less strong (Harley 2001: 84-85).

[7] If the references to Lrn in 2.b are understood as not necessarily requiring the existcnce of a particular prirnary landmark or expectation of its being claborated syntagrnatically (i.e. if the verb stern is not specified as being transitivc) the scherna will also cover cases with intransitive verbs, such as *cry-baby*, *hop-toad*, *play-boy*, or *pop-corn*.

[8] A different kind of opacity is exemplified by the *scarecrow* adjective *jerkwater* 'small, remote and insignificant', as in *jerkwater town*. Many speakers who use the word recognize that it is built of jerk and water, but have no idea what 'jerking' or 'water' have to do with the meaning

[9] Many patterns ofteii cited as non-productive, such as English "strong" past tenses, tum out to be at least marginally productive.

[10] It is related to the cross-hatching often used to mark elaboration-sites (Langacker 1987:312), and in fact elaboration-sites can be usefully thought of as productive substructures within a semantic or phonological structure.

[1] The cranium or top of the head is often called, by a conventionally jocular metaphor, a coco 'coconut'; a bald man may be called a coco pelado 'peeled coconut'.

[12] Many languages do not have such structures, however. (For instance, they are rare or non-existent in many Mexican indigenous languages). The fact that English and Spanish do is presumably due to areal influences including a common origin in early Romance. Yet this is a natural kind of structure, which shows up at least occasionally around the world in unrelated language families. It is attested in Bantu, for instance, and central Malayo-Polyncsian (Oliver Stegen, Rick Nivens, p.c.).

[13] An analytical possibility I will not pursue further here would be to combine the verb and its object into a sort of verb phrase with the verb as profile determinant (thus head), and then invoke a metonymy, perhaps ACCOMPLISHMENT FOR ITS CAUSER, to account for the fact that the subject or instrument is in fact designated in the eiid. This might (or might not) be accompanied by positing a similar coinplexity at the phonological pole, with a zcro morpheme syinbolizing the metonymic shift. A similar metonymy, with or without a zero morpheme, may be posited for the related V = S nouns such as *cook* or hore. I do not see such analyses as contradictory to what is posited iii the text, aiid nor do I doubt that some spcakcrs may conceive tlie forms in that way. On the other hand, neither do I scc such analyses as necessarily saliently activated by all speakers. The analysis in the text is relatively straightforward, aiid allows *scarecrow* nouiis to parallel similar, headcd compounds (c.g. V + S = S nouns likc iowiruck [Fig. 2] or V + O = O nouns likc *pushpin*) more directly than insisting on invariant metonymy would. Many aspccts of the differences between tlie analyses may turn out to be notational rather than substantive (cf. Langacker's discussion of the equivalence of a zero-morpheme or semantic [inetonymic] extension analyses for V = S nouns, 1987:470-474).

[14] For a discussion of what grammatical gender looks like in CG, with Spanish as the prime example, see Langackcr (1991a:180-189, 1991b:309-313). Note that the masculinity of the compound is not coded by the masculine −o of pujar-o-S; cf. *abre-lat-a-s* and many other compounds where the component noun is feminine (and plural) yet the compound is inasculine (and singular, at least in most usages).

[15] The Spanish form can be used either as a singular or plural. and indeed the form has no other plural. (I.e. you say los *espantapájaros* 'the scarecrows', and cannot say los *espantapájaroses*). Several analyses, including a zero plural suffix or a portmanteau in which tlie final −s on the plural fonns does double pluralizing duty. are possiblc, and come out to be near notational variants in thc CG analysis. This characteristic holds truc for all the Spanish formations with a plural component noun, and for some other kinds of nouns ending in *s*, e.g. *(el/los)* análisis, *(la/las) crisis*. I will not discuss it further in this paper.

[16] A reviewer of this paper objccted to uniting a 'grainmatical relation' (subject) with a 'semantic role' (instrument) evcn informally (as with the notation "S/I"). However, under CG grammatical relations and semantic roles are not differcnt in kind: thcy are both semantic (cognitive) constructs. They differ largely in degree of schematicity; 'scinantic roles' such as agent and instrument being rclated to 'conceptual archetypes' while 'grammatical relations'

relate more schematically to choice of which participant in a scene will be accorded primary or secondary prominence (salience) (Langacker 1999:41). Unquestionably agent and instrument roles are associated with subject status: that is, agents and instruments (as opposed to patients) tend to be accorded primary prominence for most usages of most verbs. In the Spanish prototype (section V) the designaturn is an instrument, whether or not that is the typical subject oftlie verb in questioii. In the English prototype thedesignatum is an acceptable subject (trajector) for tlie verb. whether it be an agent, an instrument, or whatever. The notation involves a mismatch with regard to level of sc_hernaticity because the specifications of the two constructions do not match in that regard.

[17] This last cxample is so opaquc as to probably not be a *scarecrow* noun at all for most English speakers. The same is cvcn inore true in other cases borrowed from Romance languages, where some of the morphemcs are of doubtful English occurreiice. Sorne examples are *portfolio*, which corresponds to *portafolios* [carrics-folios] in Spanish, or *parasol* aiid *parachute*, which correspond to *parasol* [halts-sun] and *paracaídas* [halts-falls], respectively.

[18] *Lava autos* is usually written as two words. Many other *abrelatas* nouns are often written so as well; in this case the awkwardness of the two a's together is probably a factor. Noun stems beginning with r are pronounced with the trilled pronunciation normal for word-initial position, and are accordingly written rr when written as a single word (e.g. *matarratas* or *mata* ralas). Similarly scarecrow nouns are sometimes written with two words or with hyphens in English (e.g. Haul All, Guard-All, etc.). I will not make anything of these; I see them as basically orthographical matters, tangentially related (but not much more) to the linguistic question of what the structure of these nouns is.

[19] Sometimes *cortabolsas*, which as we have noted corresponds more directly with *cutpurse*, is used with this ineaniiig.

[20] Tlie *Real Academia* lists *comprachilla* as a kind of Guatemalan bird, but it is not clear to me if *chilla* is eveii a noun, inuch less what it would mean in this context.

[21] A reviewer of the article reports *vendemotos* [sells-motorcycles], apparently used in Spain to mean something like 'persoii who eiigages in worthless commerce' – "like calling soinebody a used-car salesman"

[22] This comment ignores the differences in salience of individuation between the mass-noun construal of GANADO and the overt plurality of CABALLOS and BURROS.

[23] Jiménez Rios (2001:325-337) summarizes different proposals that have been niade in regard to the Spanish construction, agrccing with the conclusions that the Spanish verb is in fact a third pcrson singular indicative rather than an imperative or some kind of neutral stem fom, that the coinponent noun is its direct object and that thc construction is exocentric.

[24] Or a *tightwad*, to use a different kind of exocentric cornpound (a truc hahuvrihi).

[25] An intcresting obsolete adjectival *scarecrow* fo m is *breakteeth*, as in *breakteeth* words. It is aiiother case of a plural object, but witliout the –*s* suffix.

[26] More accurately, *pasatiempos* means 'pastiines' and its final –*s* rnarks plurality of the composite rather than of the cornponeiit object. The same goes for the other examples starred or question-rnarked in this paragraph; they are perfectly well-formed as plurals of the composite. Note as well that singular *pasatiempos, picahielos*, and *guardapolvos*, though not standard, can be easily attested by an Internet search; not so 'girasoles or *mi>-asoles.

Yct soinetiines even when the coinposite is plural the final –*s* is lacking after a inass noun. For instaiice, a set of jumper cables were labeled cables *pasacorriente* [cables passes.current], iiot *cables pasacorrientes.*

[27] My impression, though I do not have this docurnentcd, is that the plural-noun pattem is generally strongcr in Mexico than it was in rny nativc Vcnczuela.

[28] The forms deprecated with an asterisk in this paragraph can nevertheless be fouiid, usually rnuch less cornrnonly than their alternatives, in Internet searches. The sarne goes for starred forrns in thc following paragraph, but not in the rest of thc scction.

[29] Such facts have been used by sorne (e.g. Newrncyer 1983:7-10) to argue for the autonorny of grarnrnar, i.e. its independence frorn scmantics/pragmatics and phonology. Langacker rightly points out that this argurnent "confuses two issues that are in principle quite distinct: thc types of structures to be found in language. and the predictability of their behavior. Absolute predictability cannot in general be expected for natural language, and any assurnption that a ccrtain level of predictability is criterial for a particular typc of structure is essentially gratuitous" (1987:421).

[30] Spanish has cognates of the English V-*er/-(t)or* constructions, plus an **O-*er(o)*** construction designating the subject or instrument of a proccss involving the 0 ; and the phrasal V -*(d)or de* O(-pl) [V-cr of objects] is common and productive. For instancc. a disk burner on a coinputer is normally called a *quemador de discos* [buriier of disks] rather thaii, following tlie *abrelatas* pattern, a ?!*quemadiscos*.

[31] Whether or not the cornplcxity is scen as internal to the category or external to it is of coursc a rnattcr of dcfiiiition. For instance, if in Figure 10.b wc wish to talk about the category to which the cases with a non-prototypical objcct are external rather than internal, our dcfining scherna will be at the levcl of 10.b.iv, and if we takc the prototypical scherna 10.b.i as definitioiial. only cascs of dcsignating the subject will be included and not thosc dcsignating thc iiistruinciit. On the other liaiid, if the catcgory is dcfined sufficiently abstractly (e.g. at thc levcl of thc schema V + O − X in Figure 12), thc cases of designating an action or an objcct rathcr than the subject will be included within tlie category. Eithcr way the point rernains: we are dealing with a complex category here, not a rnonolithic type of construction.

[32] In Mcxican youth radio announcer patter, the word *radioescuchas* (radio-listcn-pl) is used to rnean 'radio audiencc'. Thc singular forrn would presurnably be ?*radioescucha* 'radio-listener', but I do not have it docurnented. Tlie V + O ordcr has systernic rnotivation in both languages frorn the fact that thc standard clausal order is V + 0 . That clearly docs not mean, howevcr, that the opposite order has to be proscribed in cornpounds. It seerns to be nearly so in Spanish, but it is iiot so in English.

[11] Some relatively coinrnon Spanish cascs could be thought of as V = N structures, but there is usually, at least potentially, a chaiige froni a verb-tliciiiatic vowel to a hornophonous gendcr vowel: e.g. *cocin-a* can be [cook-present.indicative.3$^{rd}$.person.singular] or [kitchcn-fcininine]; similarly *des-ayun-o* can be [un-fast-present.indicativc.1$^{st}$.person.singular] 'I (eat) brcakfast' or [un-fast-masculinc] 'breakfast'. In sorne cases the verbal ending sense may bc more active: e.g. *i-ecih-o* 'reccipt' may in sorne degree rnean 'a paper on which the recipieiit has writteii "I (hercby) rcccive (the rnerchandise)"'.

[34] Figurc 12 is adapted from a slightly more complete diagrain in Tuggy 1987.

## REFERENCES

### Dictionaries cited:

Larousse (2002). *El Pequeño Larousse Ilustrado.* México: Ediciones Larousse

Real Academia Española (1992). *Diccionario de la lengua española: Vigésimaprimera edición.* Madrid: Editorial Espasa Calpe.

Oxford English Dictionary (1971). *The compact edition of the Oxford English Dictionary.* Oxford: University of Oxford Press.

Schoenhals, Louise C. (1988). *A Spanish-English glossary of Mexican flora and fauna.* México: Instituto Lingüístico de Verano.

### Other works:

Harley, T. (2001). *The psychology of language.* London and New York: Psychology Press

Hocketi. Ch. F. (1968). *The state of the art.* The Hague: Mouton

Jiinénez Ríos, E. (2001). ¿Existe asignación y inarca de caso en los compuestos nominales de verbo+nombre? *Nueva Revista de Filología Hispánica* XLIX Número 2. México: El Colegio de México.

Langacker, R. W. (1987) *Foundations of Cognitive grammar.* Vol. I, *Theoretical prerequisites.* Stanford: Stanford University Prcss.

Langacker, R. W. (1991a). *Foundations of Cognitive grammar.* Vol. II, *Descriptive application.* Stanford: Stanford University Press.

Langaeker, R. W. (1991b). *Concept, image, and symbol: The cognitive basis of grammar.* Cognitive Linguistics Research 1. Berlin: Mouton de Gruyter.

Langacker, R. W. (1999). "Assessing the cognitive linguistic enterprise." In Theo Janssen and Gisela Redeker, eds., *Cognitive linguistics: Foundations, scope and methodology*, 13-59. Cognitive linguistics research Vol. 15. Berlin: Mouton de Gruyter.

Langacker, R. W. (2000). *Grammar and conceptualization.* Cognitive Linguistics Research 14. Berlin: Mouton de Gruyter.

LaPolla, R. J. (*to appear*). Why languages differ: Variation in the conventionalization of constraints on inference. In a collection honoring Jatnes Matisoff.

Newmeyer, F. J. (1983). *Grammatical theory: Its limits and possibilities.* Chicago: University of Chicago Press.

Newineyer, F. J. (1998). *Language form and language function.* Cambridge, Mass.: MIT Press.

Tuggy, *D. (1987). Scarecrow* nouns, generalizations, and Cognitive grammar. *Pacific Linguistics Conference* 3. 307-320. Availableon the Internet: www.sil.org~tuggyd scarecrow scarecro htm.