

• 14 00191271  
cepia l'

**Lexical Mismatches: A Semantic Representation  
of Adjectives in the LKB**

Salvador Climent  
Carme Soler

Report LSI-94-31-R

# LEXICAL MISMATCHES: A SEMANTIC REPRESENTATION OF ADJECTIVES IN THE LKB

Salvador Climent  
Carme Soler

Departament de Llenguatges i Sistemes Informàtics.  
Universitat Politècnica de Catalunya.  
Pau Gargallo 5, 08028 Barcelona

## ABSTRACT

A semantic representation of adjectives in the LKB, taking into account the semantic relationships which they can hold with the head in a noun group (N + Adj.), is presented as a first step to improve the treatment of nominal lexical mismatches due to different lexicalization patterns.

## INTRODUCTION

Languages are known to exhibit distinct preferences in lexicalization patterns in such a way that some word senses in one language correspond to a phrase in the other, giving rise to what is known as a lexical mismatch. Such differences in lexicalization arise specially when translating. This is why bilingual dictionaries appear to be an important source of mismatches, from the (apparently) simplest ones such as "crayon -> lápiz pastel", to the most complex ones such as "commute -> viajar diariamente de casa al lugar de trabajo".

The observation of such mismatches in the bilingual "English-Spanish Spanish-English dictionary VOX-HARRAP'S ESENCIAL leads us to study adjective-noun combinations in Spanish and the way of representing the semantic relationships which can hold between both noun group constituents, N and Adj.

A first approach to nominal lexical mismatches (Martí & Soler, 93) revealed that there is some correspondence between lexical mismatches in one language and morphological processes of word formation (such as derivation and composition) and semantic processes (such as metaphoric and metonymic sense extensions) in another language. According to this, some subsets could be established and, therefore, be treated together and generalised by means of a TLINK rule. However, this treatment of nominal lexical mismatches appears to be insufficient for two main reasons.

First, not all lexical mismatches due to different lexicalization patterns can be included in morphological or semantic processes. Therefore, establishing such correspondences between mismatches and processes enables us to create some subsets within the dictionary, but it is by no means exhaustive.

The second reason is related to the mismatch phrase itself and it arose when trying to establish any kind of semantic relationship which could or could not appear lexicalized in one language. Semantic relationships which hold between elements of a mismatch phrase appeared to be many and various. Its representation involved different criteria.

This is why we have reconsidered the treatment of mismatches and, starting from Spanish noun groups (N + Adj.) which constitute a mismatch, we have concentrated our study on the semantic relationship between the head (N) and the modifier (Adj.).

The relationship which holds between the constituents (N and Adj.) of a noun group is more complex than it seems to be.

1) On the one hand, it implies selectional restrictions; therefore if we take as an example "chip -> patata frita", the mismatch resulting is not just the sum of "patata + frita". In fact "frita", as a modifier, selects a kind of head which can support being "fried", and at the same time the head "patata" selects a modifier according to its semantic features.

2) On the other hand, from a semantic point of view, adjectives are not always "modifiers of nouns". Certainly in "heavy car" the adjective is expressing a quality of a noun, since "car", as concrete physical object, has the feature "weight" associated to its type. The adjective is filling this feature with a positive value. But could we establish the same kind of relationship between the components of the group "angry letter"? "Angry" does not directly modify "letter", the meaning of "angry letter" may be paraphrased by "a letter manifesting the anger experienced by the writer" (Aarts & Calbert, 1979). In "rotating knives" is the adjective really expressing a quality of a noun or is it an action? Does "musical instrument" refer to "an instrument that is musical" or to "an instrument used in music"? These are just some examples which show how complex and various can be such semantic relationships.

3) Finally, we have to take into account adjective's complex nature both from a syntactic and a semantic point of view. Syntactically it has been related to verbs, being both "predicators". Different semantic studies have revealed that most adjectives are not semantic primitives but are semantically based on noun and verbs (Talmy Givón, 1970) or other adjectives (Stati, 1973):

the **long** rope -> the rope has **length**  
rotating knives -> knives which **rotate**

Although literature is rather rich in classifications of adjectives, each one different from another and based on different criteria, our main purpose here is not classifying adjectives. ACQUILEX is aimed to obtain a (theoretically-motivated but) theory-independent representation of lexical knowledge (LKB) which could be reusable by different Natural Language Processing (NLP) systems -and extract it from Machine-Readable Dictionaries (MRDs). What we are concerned about is the fact that when combined with a noun -which owes certain (semantic) properties- an adjective produces a bar-one noun whose properties have been somehow modified or increased. Our task has to be to detect which kind of information do different sorts of adjectives carry to the noun -and represent it within the LKB.

## REPRESENTATION OF SOME SORTS OF ATTRIBUTIVE ADJECTIVES IN ACQUILEX LKB.

ACQUILEX lexical representation language (LRL) is based on unification and use of feature structures, which are typed in the manner of CARPENTER(1992). LRL is intended to be sufficiently general and theory-independent to encode a wide range of possible approaches to linguistic information.

An LKB entry has three main feature structures for representing linguistic information: CAT, SEM and QUALIA-STRUCTURE. CAT encodes syntactic knowledge -adopting a categorial grammar approach. Argument structure and formal semantics are represented in SEM. QUALIA STRUCTURE (following PUSTEJOVSKY(1989)'s proposal) reflects lexical semantics, that is, intrinsic properties of (the referent of) the word.

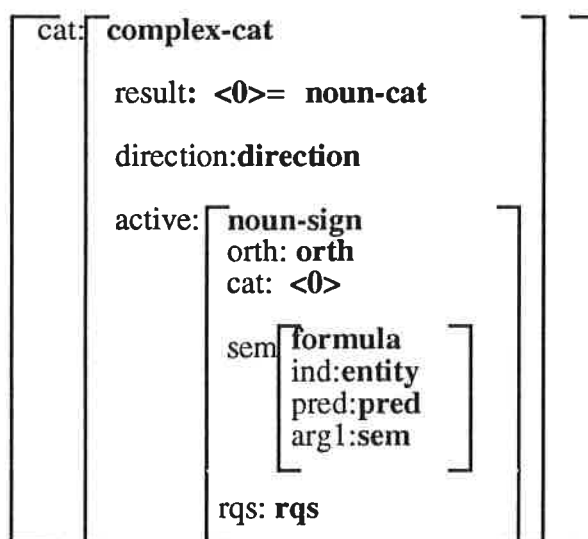
Semantic information that adjectives carry may be partially represented in SEM and ADJECTIVAL QUALIA STRUCTURE (AQS) feature structures, and as well in CAT, whose structure allows us to encode in it selectional restrictions and word-order sensitive sense disambiguation.

Each representation we are going to propose below is not to be understood as complete for every kind of adjective; otherwise, each one is a representation of a part of the information an adjective conveys. Depending on its sort, an adjective is going to be represented by means of several of those partial representations -one CAT and one SEM structures, plus zero or more features in AQS.

### 1.- Information representable in CAT: selectional restrictions and sense disambiguation by word order.

An adjective is defined syntactically as a complex category which cancelling a noun-sign results in a noun category, (N/N), as shown in (1):

(1)



**Selectional Restrictions.** Undoubtedly, specifying the lexical semantic types of nouns with whom an adjective can combine is a partial semantic description of that adjective. The mechanism adopted in LRL to describe syntactically words allows us to do that by declaring the <active:rqs> sort and the <active:sem:ind> type of the noun with whom the adjective can combine.

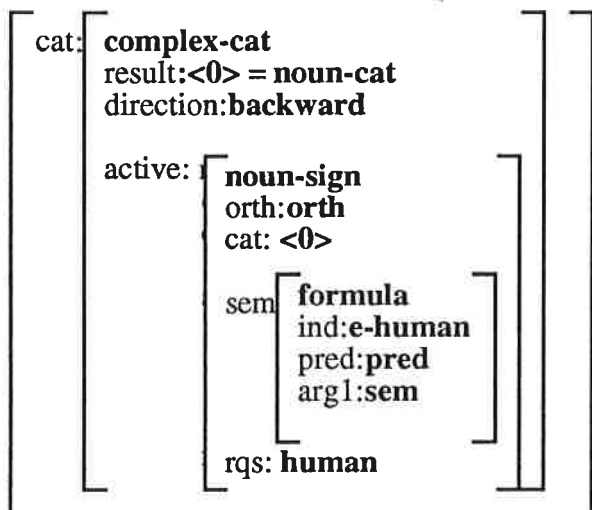
**Word Order.** Else, in Spanish the position of an adjective, before or after the noun, usually

implies a change on its meaning. E.g., the adjective “pobre” (poor), referring to a human, means ‘unlucky’ when placed before the noun; and ‘having very little money’ when placed after it. Therefore the value of <cat:direction> feature will also supply partial semantic information of an adjective.

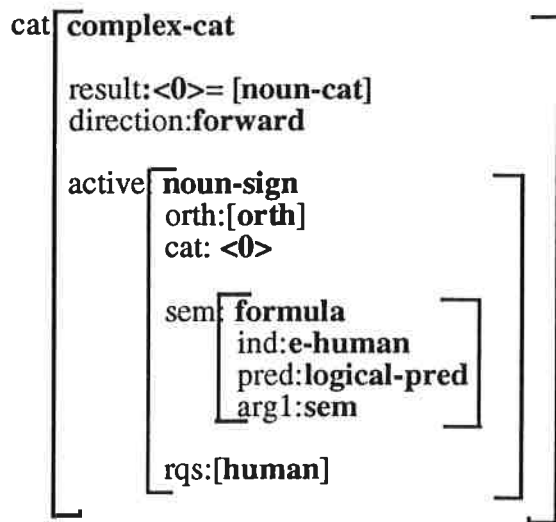
Let us take “pobre” to exemplify how information in CAT disambiguates among three different senses of that entry. In the VOX (1987) dictionary -the source we use as MRD for Spanish- sense I\_1 is that meaning ‘without money’; sense I\_5, ‘unlucky’ (both referring to humans); and sense I\_3 means roughly ‘lacking of something to be complete’ (e.g. ‘food lacking vitamins’). They would be (syntactically) represented like, respectively (2), (3) and (4).

(2)

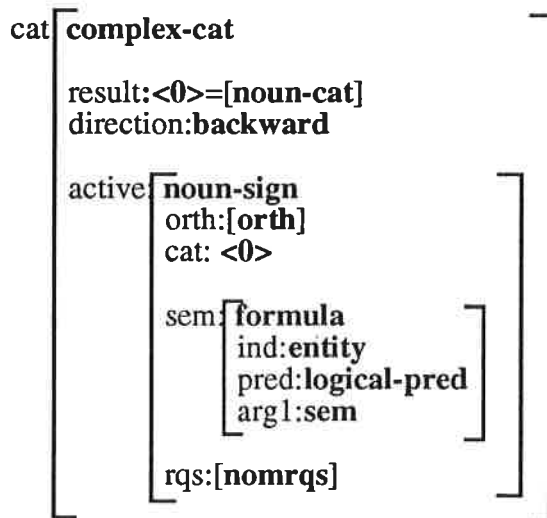
**pobre\_O\_I\_1**



(3)  
pobre\_O\_I\_5



(4)  
pobre\_O\_I\_3



Notice that while in (2) and (3) the type and the qualia sort selected is **human**, in (4) the selection is for a more general range, allowing the adjective to modify all kind of entities. Besides, while in (2) and (4) the adjective is postponed to the noun (<direction:backward>), in (3) it is placed before it.

The empirical data obtained when elucidating which sorts of nouns every kind of adjectives can combine to, will result as a collateral benefit in an enrichment and a progressive precision of the lexical semantic topology of nouns of the LKB.

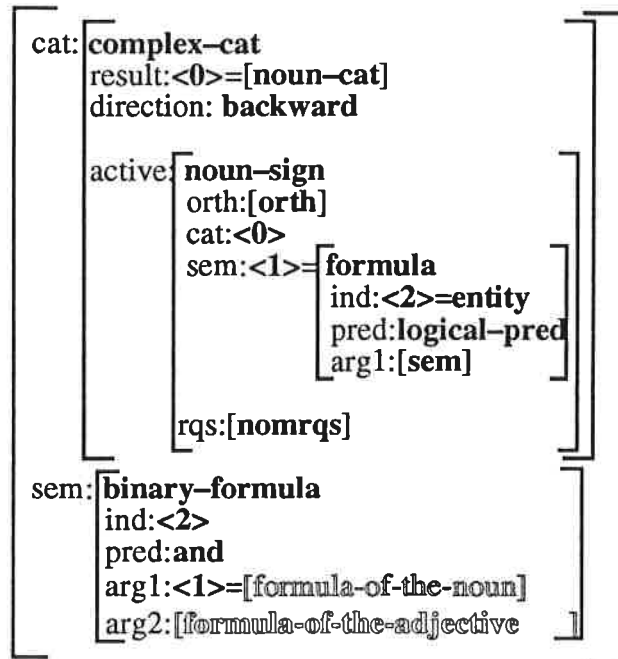
## 2.- Semantic formula of Intersective Adjectives.

The basic formula we are going to take to represent adjunct adjectives is that showing an intersective behaviour. Following ARNOLD & SADLER (1992) and other authors, we assume that in many NOUN+ADJECTIVE groups, the properties denoted by the adjective intersect with that of the noun. E.g. a "green hat" is something that "it is a hat and it is green"; so it responds to the formula:

hat(x)  $\wedge$  green(x)

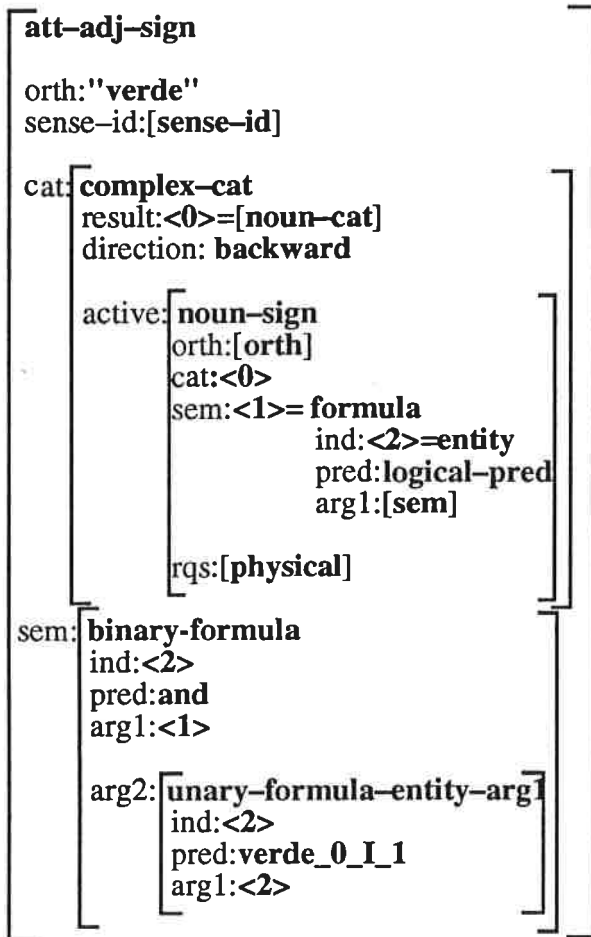
It will be represented in the LKB as in (5), exemplified for verde\_O\_I\_1 ('green') in (6):

(5)



(6)

verde\_0\_I\_1 :



Logical conjunction is pointed out here in SEM feature structure by <pred:and> applied on both arguments <arg1> and <arg2>. <arg1> is the semantics of the noun (notice the structure-sharing with the SEM feature structure of the noun selected in CAT), and <arg2> is the semantics of the adjective. The logical type resultant is <2> (that is, that of the noun).

### 3.- Qualia-intersection Adjectives.

ACQUILEX LKB system has assumed PUSTEJOVSKY (1989)'s proposal of Qualia Structure (RQS), a level of representation of intrinsic semantic properties of nominals. Qualia Structure is a powerful (world) knowledge representation frame. In an analogous way, intrinsic semantic properties of adjectives may be represented in a similar structure, namely Adjectival Qualia Structure. Whereas in SEM we represent formal effects of adjective predication, AQS is intended to encode somehow intrinsic properties of adjectives and its effects either on noun's RQS or on other non-formal aspects of nominal meaning.



**Noun-RQS Instantiation.** Inside that framework, different kinds of adjectives placed under different classes and getting different nouns in different classifications may be thought as instantiating values of attributes of their adjunct noun's RQS.

Typically, adjectives denoting colours will fill <RQS:QUAL:COLOUR\_SPEC:COLOUR> feature of objects; we can also think in:

- "catalán" ('Catalan') or "inglés" ('English') attributing 'CATALONIA' and 'ENGLAND' to <RQS:ORIGIN-AREA>;
- "muerto" ('dead') attributing FALSE to <RQS:ANIMACY>;
- "viscoso" ('viscous'), "etéreo" ('ethereal') or "sólido" ('solid') instantiating <RQS:PHYSICAL\_STATE>;
- "opaco" ('opaque') or "translúcido" ('translucent'), <RQS:QUAL:TRANSPARENCY>;
- "fétido" ('stinking'), <RQS:QUAL:SMELL>;
- "caliente" ('hot') or "hirviendo" ('boiling'), <RQS:QUAL:TEMPERATURE>;
- "grande" ('big'), <RQS:QUAL:SIZE>;
- "atigrado" ('marked like a tiger skin'), <RQS:SIMILAR>;
- "ligero" ('light'), <RQS:FORM:WEIGHT>;
- "cuadrado" ('square'), <RQS:FORM:SHAPE>;

and so forth.

Once again, this semantic characterisation is not to be understood as exclusive. As an adjective may carry complex information, its representation can share one or more noun-RQS-instantiations as those shown above with other partial representations of different status.

**Evaluation.** Some adjectives, for instance, besides attributing qualities to noun referents provide some kind of evaluation about it. We can think on "viejo" ('old'), "anciano" (a polite term for 'old') or "decrépito" ('decrepit'); all of them refer to a same age range of a person, but whilst "anciano" shows a positive evaluation, "viejo" is quite neutral (or maybe negative, depending on the context) and "decrépito" is strongly pejorative. Therefore it would not be precise to represent all of them only by attributing some value like 'high' or 'old' to the <RQS:AGE> feature of the noun. Igor Mel'čuk's *fonctions lexicales* (LFs) (or LRLs type system's 'evaluative' types, which we propose to redefine by using Mel'čuk's LFs to get it simpler) will be useful here.

MEL'ČUK(1984) posits FLs to express co-occurrence, collocation and meaning relationship in a lexicon. FLs structure is as follows:

function(Word1)→Word2

where Word1 is the sense of the entry to be characterised and Word2 the result of applying function to Word1. Word1 and Word2 may belong or not to the same category.

eg: SON(dog)→(bark)

where the function SON means the characteristic sound emitted by an entity or related to an action or process.

Mel'čuk uses few dozens of FLs. We are interested here in those which actually are or can be used to generate adjectives:

MAGN:"intensity"; VER:"correct, how it has to be"; POS:"positive evaluation"; CULM : "culmination of"; DEGRAD:"go worst"; EXCESS:"work in an abnormally excessive way"

eg.:

(7)

MAGN(memory)—>prodigious  
VER(reward)—>just  
POS(behaviour)—>good  
CULM(anger)—>paroxysm  
DEGRAD(milk)—>to go off  
EXCESS(drink)—>to get drunk

FLs may be combined with other FLs, especially in a very productive way with ANTI (meaning 'the opposite'):

ANTI\_POS\_MAGN(age)—>decrepitude

and may as well refer to thematic arguments ('actants' in Mel'cuk's terminology):  
S1(crime)—>criminal; S2(crime)—>victim.

Here we will only extract from Mel'cuk's FLs the idea that sometimes the meaning of adjectives corresponds to a subjective or social evaluation (just, good, drunk)

Actually we have in ACQUILEX LKB the type 'evaluative', used to build part of the qualia structure of nominals:

**type: evaluative (OR common delicate dressed exquisite good healthy inferior pleasant popular pure splendid superior typical unpleasant)<sup>1</sup>.**

Our idea is that instead of declaring as a type every evaluative adjective we come across, we can express that piece of meaning by combining a smaller set of evaluative FLs (commencing by those in (7)) with qualia structure features.

Then, coming back to the example of "anciano", "viejo" and "decrépito", in addition to the scalar-like value for <RQS:AGE> which all three should share, "anciano" would be attributed evaluation POS, "decrépito" ANTI\_POS\_MAGN and "viejo" maybe ANTI\_POS or no evaluative attribute at all.

Besides, some adjectives provide evaluation not to a single qualia feature but to (the referent of) the noun understood as a whole. E.g. "malo" ('bad'), "estupendo" (aprox. 'wonderful'), and so forth. Evaluative types should as well be used here to partially represent adjective's meaning by attributing them to the whole noun semantics instead of to a single RQS feature.

**Combining NRQS-Instantiation and Evaluation.** Considering all this we put forward that the AQS of some sorts of adjectives bear the features NRQS\_VALUES and/or EVALUATION, defined as follows:

**nrqs\_values**  
**nrqs\_feat:top**  
**value:top**

where NRQS\_FEAT is a feature structure belonging to the RQS of the noun the adjective combines with; and VALUE a value which could actually fill that precise NRQS\_FEAT;

---

<sup>1</sup> CASTELLON (1993a)

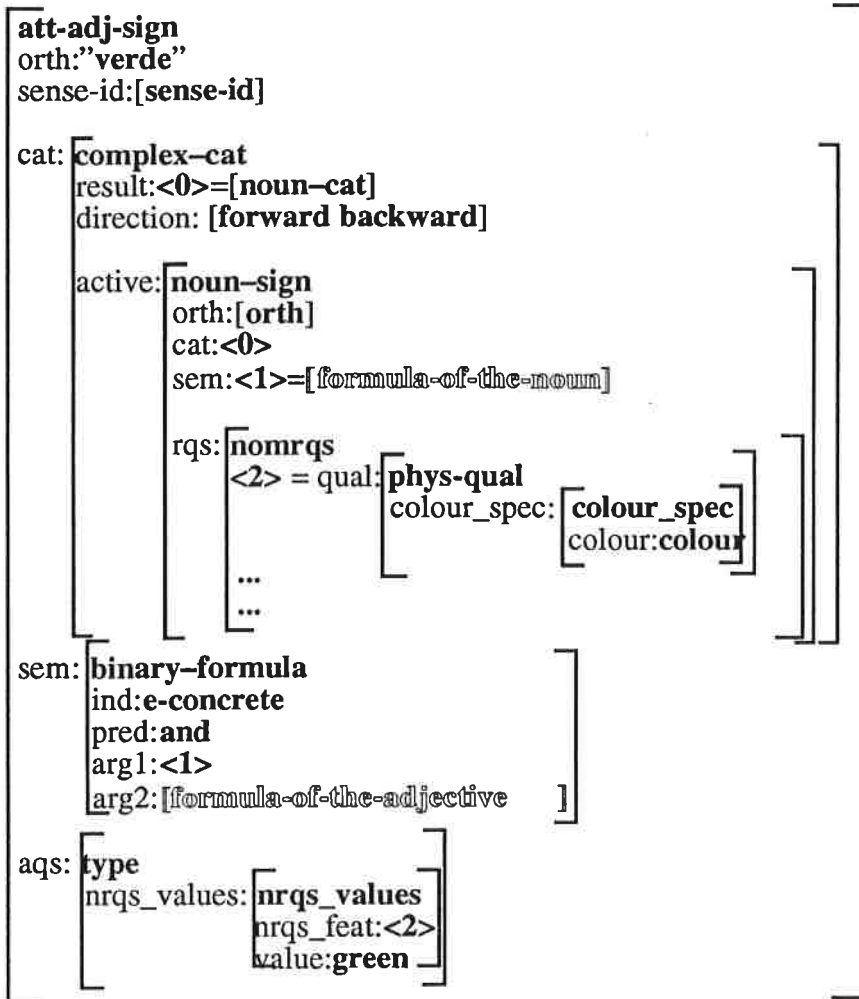
**evaluation**  
**valuee:top**  
**evaluative:evaluative**

where **EVALUÉE** is either the **NRQS\_FEAT** or the feature structure **SEM** of the noun the adjective combines with, and **EVALUATIVE** an atom of type **evaluative**.

Let us exemplify it representing the adjectives “verde”, “anciano” and “malo” in (8), (9) and (10):

(8)

verde\_O\_I\_1



(9)

anciano\_O\_I\_1

att-adj-sign  
orth:"anciano"  
sense-id:[sense-id]

cat:[complex-cat  
result:<0>=[noun-cat]  
direction:[forward backward]

active:[noun-sign  
orth:[orth]  
cat:<0>  
sem:[formula-of-the-noun]

rqs:[human  
<1>=age:age  
...]

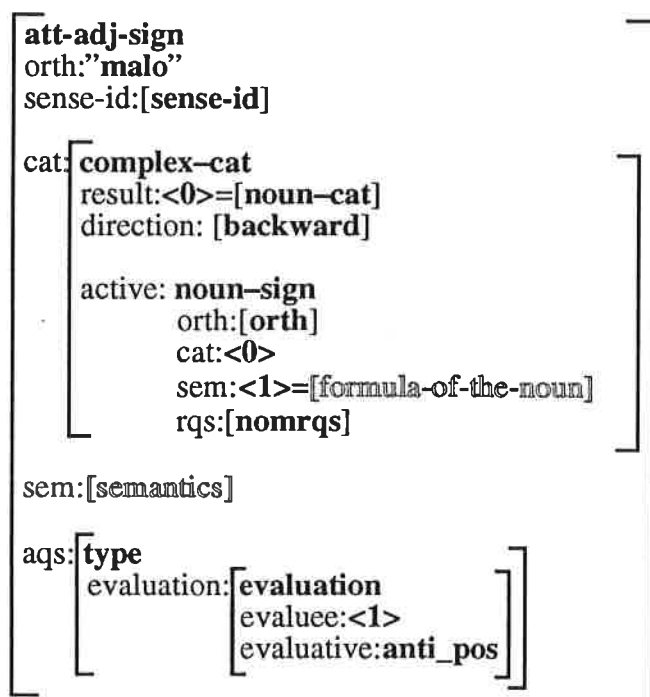
sem:[semantics]

aqs:[type  
nrqs\_values:[nrqs\_values  
nrqs\_feat:<1>  
value:old]

evaluation:[evaluation  
evaluee:<1>  
evaluative:pos]

(10)

**malo\_O\_I\_1**



Note that in (8) ('green') there is instantiation of the RQS feature of the noun -showed by the structure-sharing of <cat:active:rqs:qual:colour\_spec:colour> and <aqs:nrqs\_values:nqs\_feat>- but no evaluation.

In (9) (polite 'old') there is instantiation of <age> and evaluation of that feature.

And in (10) there is evaluation (ANTI\_POS) of the whole noun -marked by structure sharing with <1>: the whole semantics of the noun- not only of one of their features or properties.

**Re-utilisation of LKB features.** Another resource not to be forgotten when representing qualia-intersective adjectives is the re-utilisation of features already represented in other LKB entries. Sometimes, noun-derived adjectives are defined in the MRD by means of the noun from which they derive. E.g.:

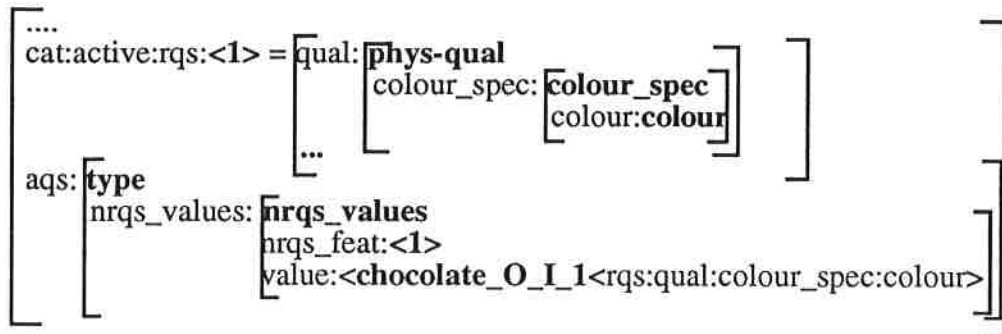
**achocolatado:** Que tiene color de chocolate ('of the colour of chocolate').

**acopado:** De figura de copa de árbol ('tree-crown-shaped').

Instead of declaring new types, those adjectives should be represented having the following AQS feature structures (relevant portion of CAT is also displayed to show structure-sharing):

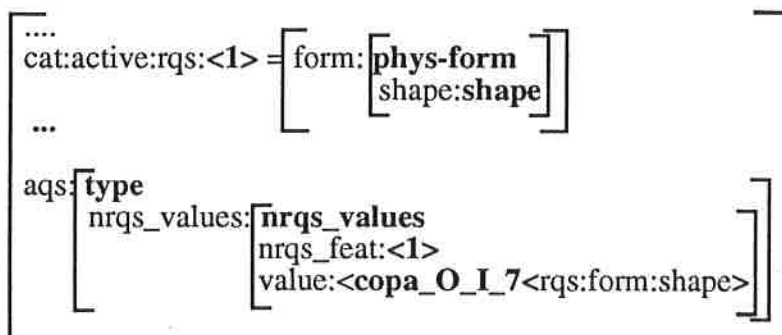
(11)

**achocolatado\_O\_I\_1**



(12)

**acopado\_O\_I\_1**

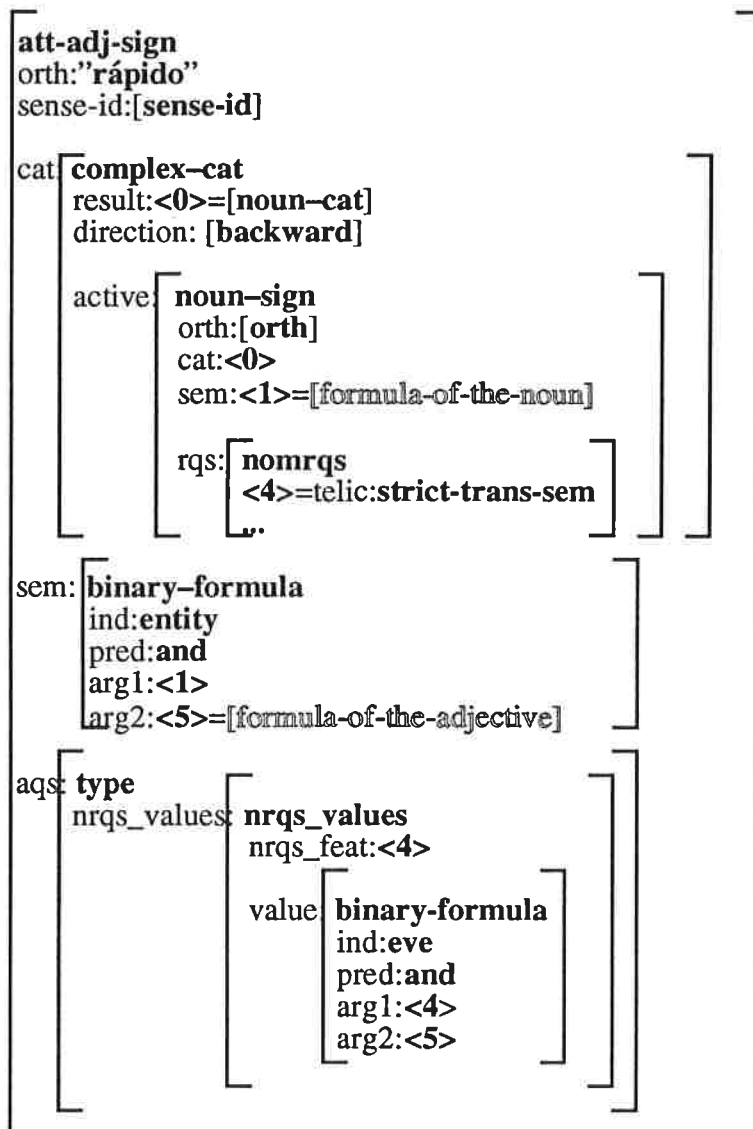


**Telic Role.** A distinguished qualia-structure feature is the telic role. PUSTEJOVSKY(1991) shows that sometimes predicators do not act as an attribute over the whole noun denotation but as modifying its 'telic role' (the prototypical action attributed to it). E.g., in 'fast car' what is said to be fast is not exactly the object 'car' but the action he is to do.

This behaviour may also be represented in the AQS, as in (13).

(13)

**rápido\_O\_I\_1** (“fast”)



What <aqs:nrqs\_values> is doing here is instantiating the feature <telic> of the noun rqs (<4>) -which would be a formula equivalent to “run”- with a formula which is the conjunction of <4> and <5> -the semantics of the adjective-. Thus, the telic role of the bar-one noun will be ‘fast(e)  $\wedge$  run(e)’ -instead of ‘run(e)’.

In the same direction it has to be noticed that, although being the ‘telic role’ the main-purpose action of an object, there are other not so privileged purposes an object may be used for -and those may also be modified by adjectives.

For instance, one can talk about a “coche cómodo” (‘comfortable car’). Indeed, sitting inside it is not the main purpose (=telic role) a car is for; but it is something to do with it. In the example



given, "cómodo" is evidently a property related to the action of 'sitting' or 'be placed' in the car in the same way that 'fast' is related to the action of 'running'. The enrichment of the qualia structure of nouns with that kind of 'underlying' actions (other than the telic role) is a way that we suggest may be explored<sup>2</sup>.

#### 4.- Deverbal Adjectives.

Some adjectives derive from verbs or deverbal nouns, usually by addition to the root of typical suffixes as [-oso/a], [-able], [-ado/a].

E.g.: from "envidia" or "envidiar" (envy, to envy):

"envidioso/a" ('someone who feels or denotes envy'),  
"envidiado/a" ('someone who is envied by someone'),  
"envidiable" ('someone deserving to be envied').

The phenomenon is much similar, in fact parallel, to that of nominalization<sup>3</sup>, what leads us to assume that, in the same way as deverbal nouns do, there are deverbal adjectives and their meaning is tightly related to theta-role assignation.

Thus, the noun combining with "envidioso" can be considered the p-agent of 'feeling envy', while that combining with "envidiado" will be its p-patient.<sup>4</sup>

This is representable in the LKB taking account of the internal verb from which the adjective derives instead of the adjective itself. Then, <arg2> in SEM would not be the semantics of the event denoted by the adjective, but the semantic formula of the underlying verb. E.g., for "hombre envidioso" ('envious man'):

$\text{hombre}(x) \wedge \text{envidiar}(e) \wedge \text{agent}(e,x)$  [where "envidiar"='to envy'].

it will be represented in LRL as in :

---

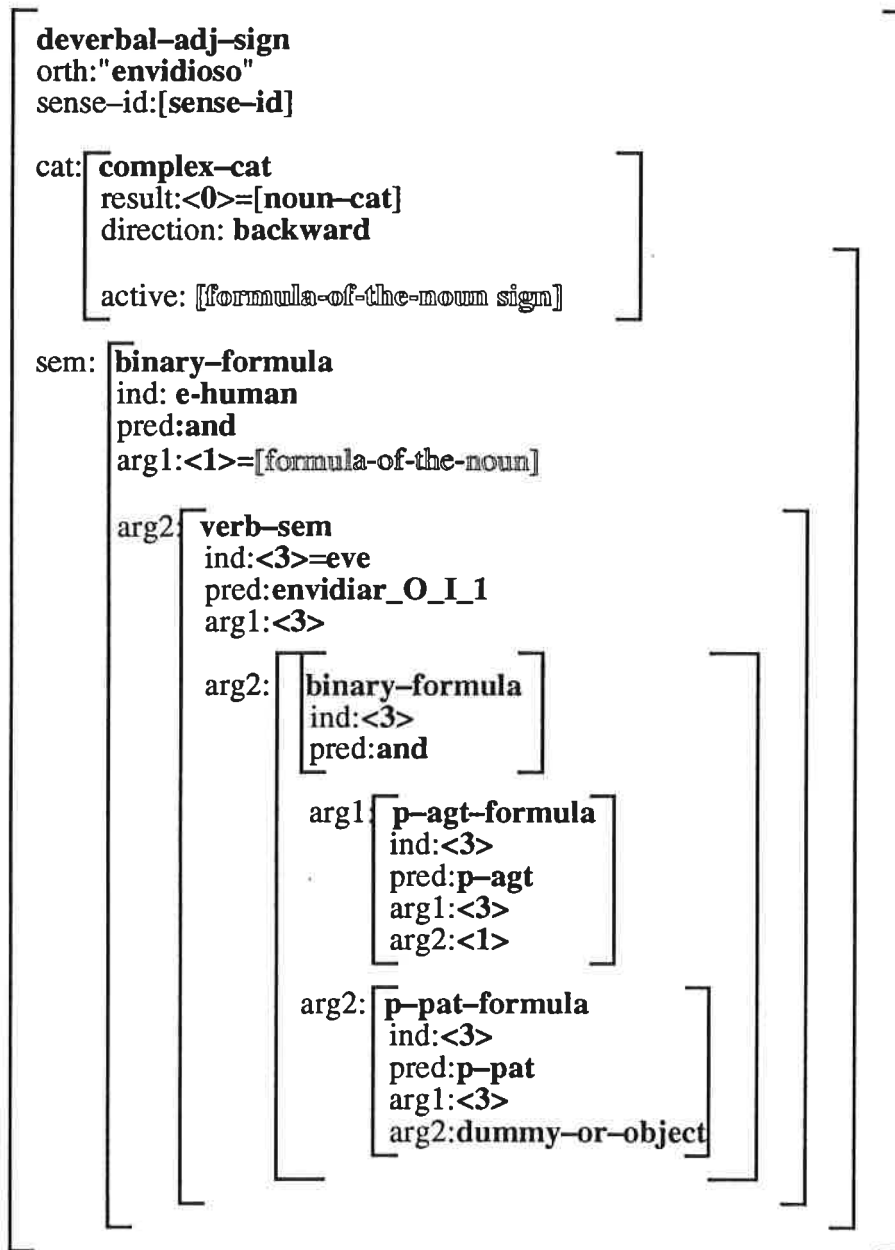
<sup>2</sup> One could think that 'cómodo' here is somehow modifying the Constitutive Role of 'car', as its seats are a part of a car, and 'cómodo' refers to them; but one can also think that a car may be 'comfortable' not just because its seats are comfortable -but maybe because there is enough room in the inside.

<sup>3</sup> For an accurate analysis of nominalization in Spanish and its representation in the LKB *vid* CASTELLON (1993b).

<sup>4</sup> In the strategy adopted in ACQUILEX for representing thematic roles, a theta role may be either a p-agt ('proto-agent') or a p-pat ('proto-patient'), which are defined to be, respectively, the least and the most agentive argument of a verb. Further p-pat and p-agt sorts provide a finer granularity of specifying role contents.

(14)

**envidioso\_O\_I\_1**



Structure-sharing between <sem:arg1> and <sem:arg2:arg2:arg1:arg2>, shows here that the noun is the agent of the event denoted by the adjective. If the entry to be represented would be "envidiado", then the noun would have to fill <arg2>, i.e. the p-pat formula.

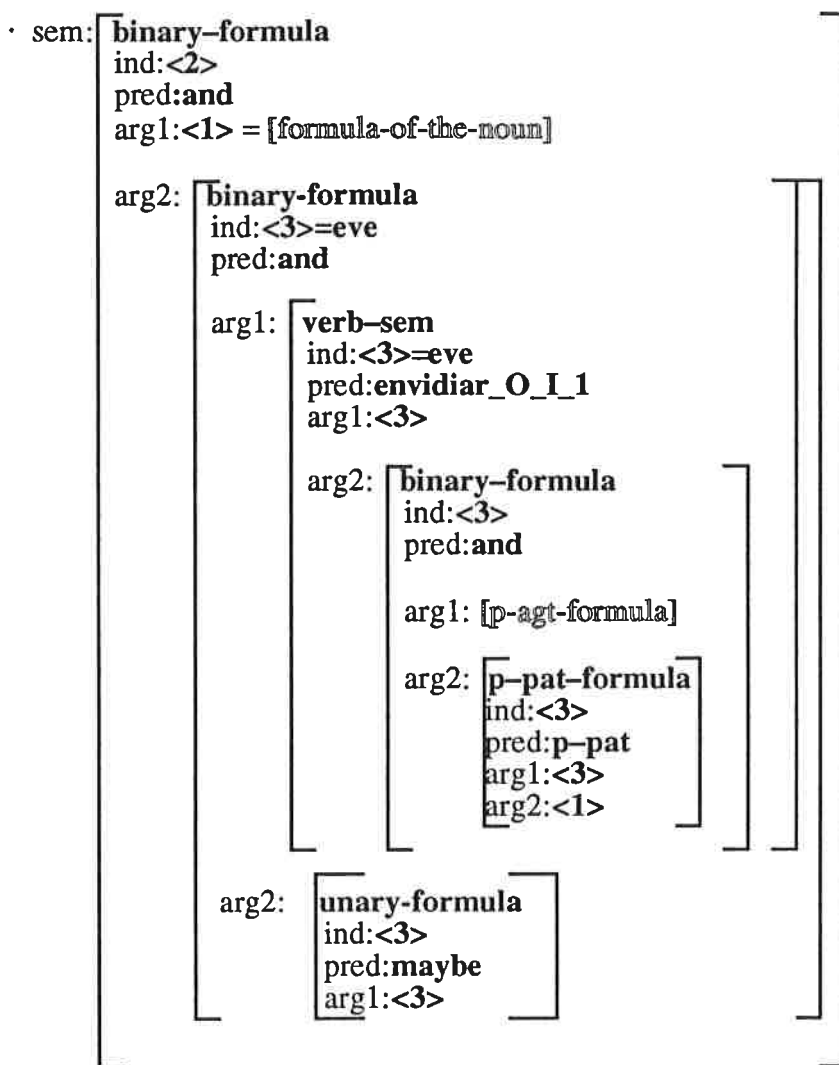
**Type redefinition.** Current definition of theta-formula types does not allow having a formula as <arg2> but demands an atomic type subsumed by the type **dummy-or-object**. We think either this definition should be changed in order to accept any type subsumed by **sem**

as <arg2> of theta-formulas or we should assume <sem:arg2:arg2:arg1> and <sem:arg2:arg2:arg2> of this sort of adjectives not being of type **theta-formula** but just of type **binary-formula**; because what is interesting here to represent is that a thematic argument of the verb underlying to the adjective is filled just by the precise noun it is combining with (not simply by any other object of the same type).

**Introducing operators.** A new problem arises when one thinks about the representation of adjectives like "envidiable". Spanish suffix '-ble' adds to a verbal root the idea of possibility or potentiality. Therefore maybe the person referred by the noun is not actually envied by anyone, as the action of 'envying' is not actual but potential.

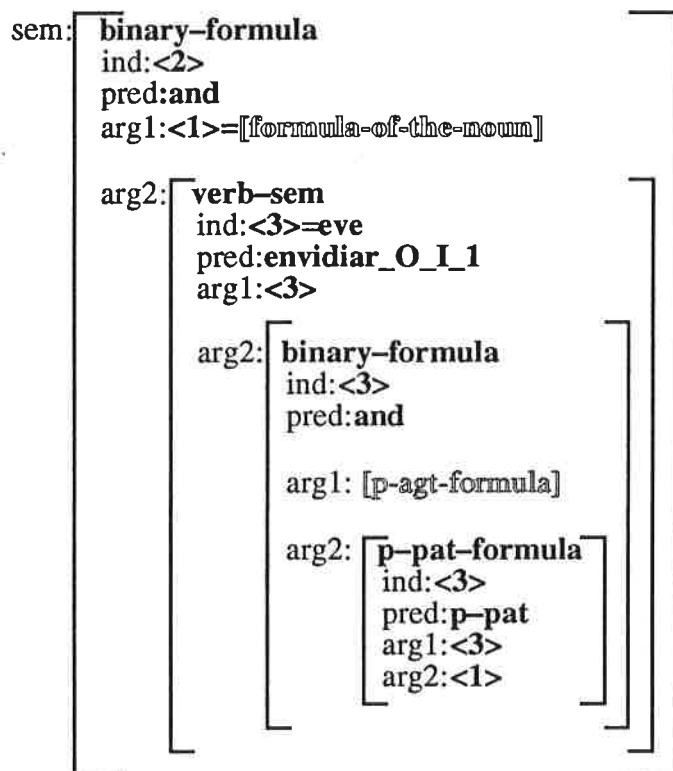
One solution for this and similar cases would be introducing in the type system operators coming from logics other than first-order. In the case of "envidiable" the modal operator **maybe** could be one solution. Then <sem> of "envidiable" would be as in (15):

(15)



Instead of that corresponding to "envidiado", in (16):

(16)



So while the logical formula for "envidiado" would be:  
 $(\text{hombre}(x) \wedge \text{envidiar}(e) \wedge \text{p-pat}(e,x))$ ,  
that for "envidiable" would be:  
 $(\text{hombre}(x) \wedge ((\text{envidiar}(e) \wedge \text{p-pat}(e,x)) \wedge \text{maybe}(e)))$ .

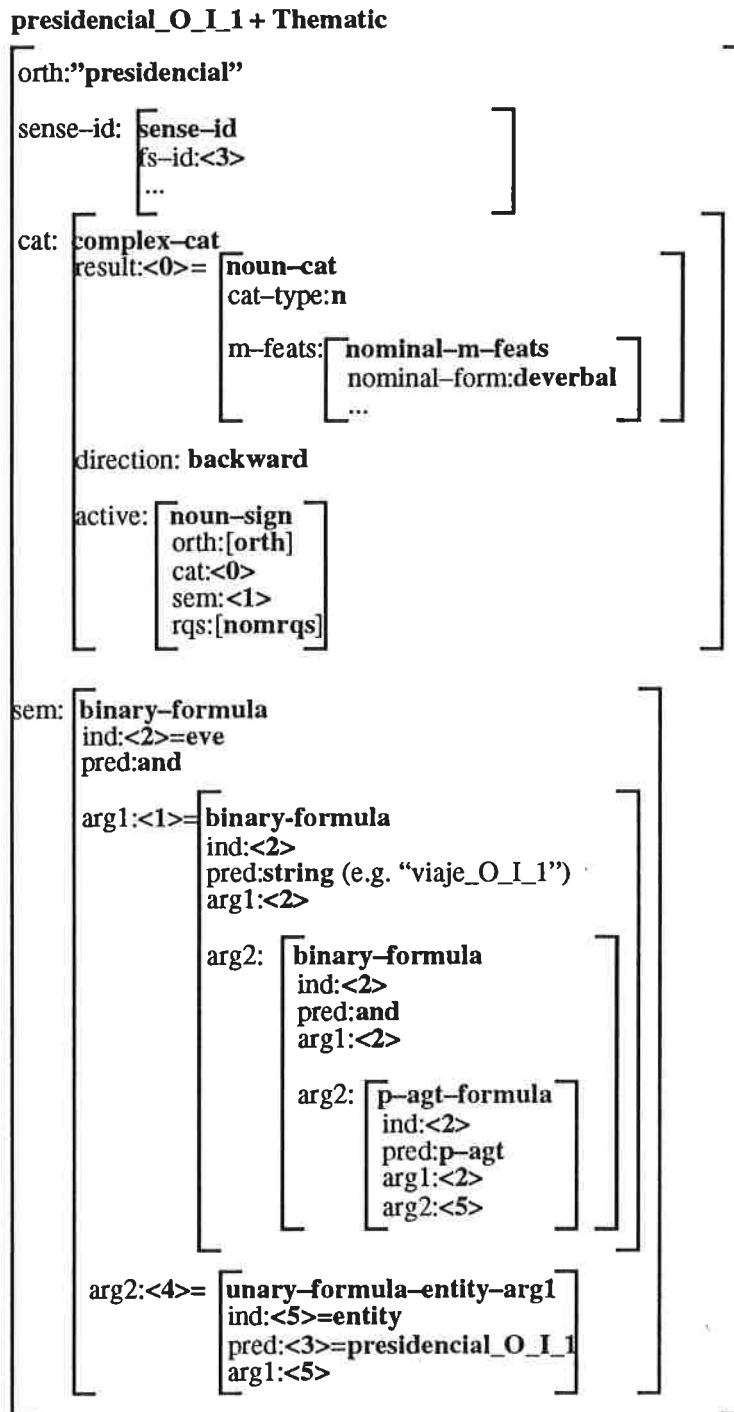
## 5.- Thematic Adjectives.

Bosque (1993) argues that there is a class of adjectives so called "temáticos" ('thematic') which are arguments of the noun. As he exemplifies it, "viaje presidencial" ('presidential trip') has the meaning of a journey (or maybe a psychedelic experience) made by the president, thus 'the president' is the agent of that action (travelling). Bosque says that kind of adjectives usually show ambiguities with what he calls classificative adjectives ("clasificativos"): "elecciones presidenciales" ('presidential elections') does not mean that the president be the agent of electing, but that those elections are of class 'presidential'.

We don't think there actually exist a class of thematic adjectives, otherwise we think that many relational adjectives behave 'thematically' when combining with action nouns. Anyway, this is a semantic phenomenon that ought to be represented in the LKB. We suggest doing it this way

(e.g. “presidencial”)<sup>5</sup>:

(17)



<sup>5</sup> Here we suppose the representation of action nouns as bearing an argument structure in the same way than verbs bear. Thus an intransitive action noun would bear a formula like  $\text{action-noun}(x) \wedge \text{p-agt}(x,y)$ .

The structure above stands for the formula (example "viaje presidencial"):

$\text{viaje}(e) \wedge \text{agent}(e,x) \wedge \text{presidencial}(x)$ .

Structure-sharing of  $\langle \text{sem:arg1:arg2:arg2} \rangle$  and  $\langle \text{sem:arg2} \rangle$  shows that the p-argt of the action noun is just the entity  $x$  which is "presidencial( $x$ )".

## 6.- Complex-Predication Adjectives.

Sometimes adjectives, when combining with certain kind of nouns –maybe that kind of nouns which the adjective does not select 'naturally', usually *fig* senses on dictionaries– show some kinds of predication more complex than simple intersection.

For instance, "sad book", which we could represent as something like

' $\text{book}(x) \wedge \text{sadness}(y) \wedge \# \text{DENOTES} \#(x,y)$ ',

better than just

' $\text{book}(x) \wedge \text{sad}(x)$ '.

It would be better representing that kind of predication in terms of the Qualia Structure, as in (3) above, but probably not every adjective will admit it. Thus, where the Qualia solution prove not to be available, it will be interesting adopting a relational approach which sets in sem feature formulae of the kind:

$\text{noun}(x) \wedge \text{adjective/underlying-noun-or-verb}(y) \wedge \text{relation}(x,y)$

instead of simple intersection ' $\text{noun}(x) \wedge \text{adjective}(x)$ '.

AARTS & CALBERT(1979) proposed a mechanism intended to express the semantics of the group Noun+Adjective in a relational way, positing the following conceptual primitives which they call Predicational Relators (PRs):

HAS\_PROPERTY(N,A)

eg: heavy thing

paraphrase:N has as a temporary or permanent property being A

EXPERIENCES(N,A)

eg: raging man

N experiences the mental state A [rage]

SHOWS(N, EXPERIENCES(A))

eg: sad letter

N shows that someone experiences A [sadness]

TIME\_WHEN(N, EXPERIENCES(A))

eg: happy days

N is the moment when someone experiences A [happiness]

PLACE\_WHERE(N, EXPERIENCES(A))

eg: happy home  
N is the place where someone experiences A [happiness]

CAUSE(N, EXPERIENCE(A))  
eg: sad event  
N causes that someone experiences A [sadness]

MANNER\_HOW(A,N)  
eg: quick computation  
A is the manner how the action N is brought about

DEGREE\_OF(A,N)  
eg: complete fool  
A is the degree of intensity of the quality N

FREQUENCY\_OF(A,N)  
eg: occasional rains  
A is the frequency of occurrence of the action or event N

TEMPORAL\_SEQUENCE\_OF(A,N)  
eg: next meeting  
A is the relative point in time when the action or event N occurs

OBJECTIVITY\_OF(A,N)  
eg: possible attack  
A is the degree of objectivity of the action or event N

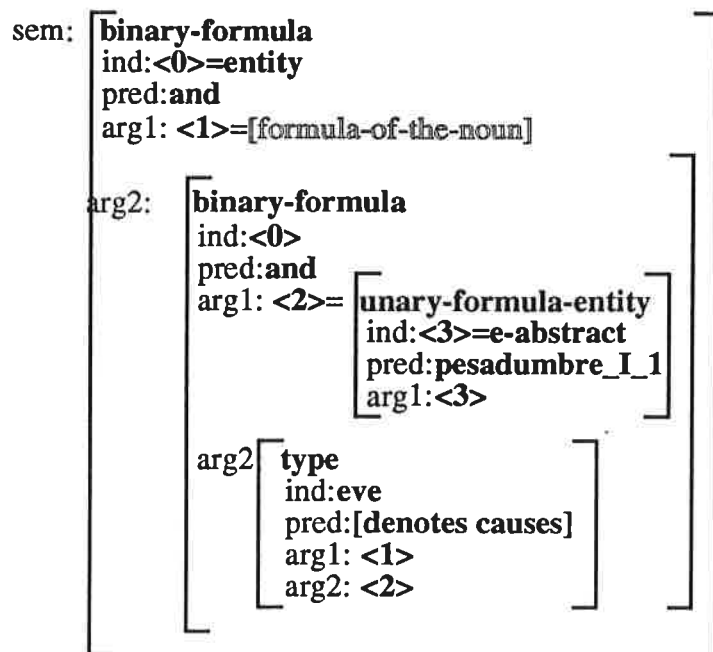
ACTS(N,A)  
eg: rotating knives  
N realises the action A

EXPERIENCES\_2(N,A)  
eg: fallen leaves  
N has experienced the action A

Predicational primitives should be declared as types in the type system and we expect them to be a limited set. Besides, we trust in finding patterns on MRDs definitions which systematically make manifest conceptual primitives.

As an example, **triste\_O\_I\_3** [= 'sad'] (defined in the VOX dictionary as "que denota u ocasiona pesadumbre: cara ~, noticia ~, habitación ~" [= 'denoting or causing sorrow: ~ face, ~ news, ~ place']) could be represented in the LKB owing the following SEM feature structure:

(18)



The hypothesis about nouns with whom such adjectives are related (**pesadumbre** in the example above) is that they could be retrieved from MRDs definitions of every adjective.

## 7.- Modal Adjectives.

Modal adjectives, like "supuesto" ('alleged') or "probable" ('probable') does not intersect with the noun but they act as modal operators.

"Alleged athlete", for instance, can not be represented as:

\*alleged(x)  $\wedge$  athlete(x)

but:

alleged(athlete(x)).

The obvious solution for treating that kind of predication would be making second order treatment, as in (19) below:

(19)

**supuesto\_O\_I\_2**



sem [ type  
ind:ind  
pred:supuesto\_O\_I\_2  
arg1: [formula-of-the-noun] ]

The effect here is representing in SEM the formula 'supuesto\_O\_I\_2(noun(x))'

But second order treatments are computationally unattractive, so it would be necessary to develop a first order approach for representing modal adjectives.

## FUTURE RESEARCH

Once at this point the next step is:

(i).- go on both detecting other phenomena existing in adjectival predication and proposing how to represent them in the LKB,

(ii).- to develop strategies to both detecting the different sorts of adjectival predication behaviour from MRD definitions and as well retrieving from them the appropriate information to build each sort of representation structures.

(iii).- to find in which precise cases do complex predication occur, and then stablish what primitives will be needed, starting from Aarts & Calbert (1979) or other authors, as Hobbs or Mel'cuk.

(iv).- to develop a first-order strategy representation for modal adjectives.

(v).- to generate deverbal adjective entries from verbs via lexical rules.

## **REFERENCES:**

- AARTS & CALBERT (1979). Metaphor and Non-Metaphor. The semantics of Adjective-Noun combinations.** Max Niemeyer Verlag, Tübingen
- ACQUILEX (1992). The acquisition of Lexical Knowledge.** Esprit Basic Research Project 7315.
- ARNOLD, D & SADLER, L (1992). Noun-Modifying Adjectives in HPSG.** Working Papers in Language Processing 35. University of Essex.
- BOSQUE, I(1993). Sobre las diferencias entre los adjetivos relacionales y los calificativos.** Revista Argentina de Lingüística, vol. 1993
- CASTELLON, I (1993a). Lexicografía Computacional: Adquisición Automática del Conocimiento Léxico.** Report LSI-93-28-R. Universitat Politècnica de Catalunya.
- CASTELLON, I (1993b). Nominalization in the LKB framework.** ACQUILEX II WP.
- CARPENTER, B(1992). The Logic of Typed Feature Structures.** Cambridge University Press.
- COPESTAKE, A (199?). The representation of group denoting nouns in a lexical knowledge base.** Acquilex-II WP.
- GIVON T. (1970), Notes on the semantic structure of English Adjectives,** Language, 46 (1970), pp. 1-45.
- MARTI, M.A. & SOLER, C. (1993), Dealing with Lexical Mismatches,** ACQUILEX II WP.
- MEL'CUK, I (1984). Dictionnaire explicatif et combinatoire du français contemporain.-** Les presses de l'Université de Montréal.
- PUSTEJOVSKY, J (1991).- The generative lexicon.** Computational Linguistics vol 17 nr. 4, pp. 409-441.
- STATI, S. (1973), Les traits sémantiques de l'adjectif,** Cahier de Lexicologie, XXIII, 2, pp.51-61.
- VOX (1987) Diccionario General Ilustrado de la Lengua Española VOX.** Ed. Bibliograf S.A. Barcelona.
- VOX HARRAP'S (1992), Diccionario Esencial Inglés-Español Español-Inglés.** Ed. Bibliograf S.A. Barcelona

**Departament de Llenguatges i Sistemes Informàtics**  
**Universitat Politècnica de Catalunya**

**Research Reports – 1994**

- LSI-94-1-R “Logspace and logtime leaf languages”, Birgit Jenner, Pierre McKenzie, and Denis Thérien.
- LSI-94-2-R “Degrees and reducibilities of easy tally sets”, Montserrat Hermo.
- LSI-94-3-R “Isothetic polyhedra and monotone boolean formulae”, Robert Juan-Arinyo.
- LSI-94-4-R “Una modelización de la incompletitud en los programas” (written in Spanish), Javier Pérez Campo.
- LSI-94-5-R “A multiple shooting vectorial algorithm for progressive radiosity”, Blanca Garcia and Xavier Pueyo.
- LSI-94-6-R “Construction of the Face Octree model”, Núria Pla-Garcia.
- LSI-94-7-R “On the expected depth of boolean circuits”, Josep Díaz, María J. Serna, Paul Spirakis, and Jacobo Torán.
- LSI-94-8-R “A transformation scheme for double recursion”, José L. Balcázar.
- LSI-94-9-R “On architectures for federated DB systems”, Fèlix Saltor, Benet Campderrich, and Manuel García-Solaco.
- LSI-94-10-R “Relative knowledge and belief: SKL preferred model frames”, Matías Alvarado.
- LSI-94-11-R “A top-down design of a parallel dictionary using skip lists”, Joaquim Gabarró, Conrado Martínez, and Xavier Messeguer.
- LSI-94-12-R “Analysis of an optimized search algorithm for skip lists”, Peter Kirschenhofer, Conrado Martínez, and Helmut Prodinger.
- LSI-94-13-R “Bases de dades bitemporals” (written in Catalan), Carme Martín and Jaume Sistac.
- LSI-94-14-R “A volume visualization algorithm using a coherent extended weight matrix”, Daniela Tost, Anna Puig, and Isabel Navazo.
- LSI-94-15-R “Deriving transaction specifications from deductive conceptual models of information systems”, María Ribera Sancho and Antoni Olivé.
- LSI-94-16-R “Some remarks on the approximability of graph layout problems”, Josep Díaz, María J. Serna, and Paul Spirakis.

- LSI-94-17-R "SAREL: An assistance system for writing software specifications in natural language", Núria Castell and Àngels Hernández.
- LSI-94-18-R "Medición del factor modificabilidad en el proyecto LESD" (written in Spanish), Núria Castell and Olga Slávkova
- LSI-94-19-R "Algorismes paral·lels SIMD d'extracció de models de fronteres a partir d'arbres octals no exactes" (written in Catalan), Jaume Solé and Robert Juan-Arinyo.
- LSI-94-20-R "Una paral·lelització SIMD de la conversió d'objectes codificats segons el model de fronteres al mdel d'octrees clàssics" (written in Catalan), Jaume Solé and Robert Juan-Arinyo.
- LSI-94-21-R "Clausal proof nets and discontinuity", Glyn Morrill.
- LSI-94-22-R "A formal method for the synthesis of update transactions in deductive databases without existential rules", Joan A. Pastor.
- LSI-94-23-R "On the sparse set conjecture for sets with low density", Harry Buhrman and Montserrat Hermo.
- LSI-94-24-R "On infinite sequences (almost) as easy as  $\pi$ ", José L. Balcázar, Ricard Gavaldà, and Montserrat Hermo.
- LSI-94-25-R "Updating knowledge bases while maintaining their consistency", Ernest Teniente and Antoni Olivé.
- LSI-94-26-R "Structural facilitation and structural inhibition", Glyn Morrill.
- LSI-94-27-R "Formalising existential rule treatment in the automatic synthesis of update transactions in deductive databases", Joan A. Pastor.
- LSI-94-28-R "Proceedings of the Fifth International Workshop on the Deductive Approach to Information Systems and Databases" (Aiguablava, 1994), Antoni Olivé (editor).
- LSI-94-29-R "Towards a VRQS representation", Mariona Taulé Delor.
- LSI-94-30-R "MACO: Morphological Analyzer Corpus-Oriented", Sofia Acebo, Alicia Ageno, Salvador Climent, Javier Farreres, Lluís Padró, Francesc Ribas, Horacio Rodríguez, and Oscar Soler.
- LSI-94-31-R "Lexical mismatches: a semantic representation of adjectives in the LKB", Salvador Climent and Carme Soler.
- LSI-94-32-R "LDB/LKB integration", German Rigau, Horacio Rodríguez, and Jordi Turmo.
- LSI-94-33-R "Learning more appropriate selectional restrictions", Francesc Ribas.
- LSI-94-34-R "Oracles and queries that are sufficient for exact learning", Nader H. Bshouty, Richard Cleve, Ricard Gavaldà, Sampath Kannan, and Christino Tamon.

---

Copies of reports can be ordered from:

Nuria Sánchez  
Departament de Llenguatges i Sistemes Informàtics  
Universitat Politècnica de Catalunya  
Pau Gargallo, 5  
08028 Barcelona, Spain  
[secrelsi@lsi.upc.es](mailto:secrelsi@lsi.upc.es)