

• 14 00191269  
Copia 1

## **Towards a VRQS Representation**

Mariona Taulé Delor

Report LSI-94-29-R

 UPC  
Facultat d'Informàtica  
de Barcelona - Biblioteca

20 SET. 1994

## TOWARDS A "VRQS" REPRESENTATION

Mariona Taulé Delor  
marionat@lsi.upc.es  
Universitat de Barcelona

June 1993.

### Abstract

In this paper we would like to present a proposal for the representation of verbal lexical meaning in a Lexical Knowledge Base. We start from the hypothesis that the verbal lexical meaning is the sum of different levels of semantic description: Argument Structure, Event Structure, Selectional Restrictions and Compositional Semantics. Our talk is based on the representation of the Compositional Semantic level (VRQS), the decomposition of the meaning into semantic components and the way in which they are saturated in Spanish. We will present three different types of saturation (argumental, morphological and "understood") and we will try to demonstrate how these sorts of saturation have different syntactic consequences.

### 1.- Introduction

This work is part of a more general extended lexical study developed in the framework of the Acquilex Project<sup>1</sup>. One of its basic purposes is the construction of a Computational Lexicon where the lexical entries are represented in a multilingual Lexical Knowledge Base (LKB) (Copestake, A. 1992)<sup>2</sup> which uses a lexical representation language (LRL) based on unification. The objects of the LKB are represented as typed feature structures which are ordered hierarchically to allow the inheritance of information.

The lexical entries are treated like lexical signs following the proposal of "Head-Driven Phrase Structure Grammar" (HPSG) (Pollard, C. & Sag, I. 1987-1992) where, basically, the morpho-syntactic and semantic information of the lexical items are represented.

---

<sup>1</sup> This work is developed in Acquilex-II Project (Esprit 7315) "The Acquisition of Lexical Knowledge for Natural Language Processing Systems" at the Universitat Politècnica de Catalunya.

This work has also been supported by PB91-0854 DGICYT.

<sup>2</sup> Copestake, A. (1992) shows a complete treatment of the functionality of the LKB and its language representation (LRL). Other interesting documents about the LKB are Ageno, A. et al. (1992), Copestake, A. et al. (1991), Sanfilippo, A. (1991), Castellón, I. (1992).

The main purpose of this work is the representation of a specific type of semantic information, that which refers to the semantic components or entities which constitute the verbal lexical meaning, and the way in which this sort of information (labeled <vrqs>) is encoded in the LKB. At this level of representation we propose a semantic classification of verbs according to the different semantic components which they are made up. We start from the hypothesis that the verbal lexical meaning is the sum of different levels of semantic description that contribute to the whole verbal lexical meaning in a different way (Pustejovsky, J. 1991): Argument Structure, Event Structure, Selectional Restrictions and Compositional Semantics<sup>3</sup>. Each one of these semantic levels brings different information to the verbal meaning and, therefore, in order to give a complete semantic characterization of verbs it is necessary to represent all of them (See section 2).

We will also explain, briefly, how the information referring to the Argument Structure, Event Structure and Selectional Restrictions is represented in the LKB (See section 3).

We have followed the methodology developed in Levin, B. (1993) in order to establish the classification of verbs in different semantic classes. This methodology is based on the "assumption that the behaviour of a verb, particularly with respect to the expression and interpretation of its arguments, is to a large extent determined by its meaning" (Levin, B. 1993: 1-19). Therefore, we consider, following this perspective, that diathesis alternations can be a useful approach to distinguishing the different semantic classes of verbs. It means that verbs belonging to the same semantic class would seem to share the same set of alternances in the patterns of subcategorization.

Firstly, we will present the way we carried out the decomposition of the verbal meaning into semantic components (See section 4.1) and the way in which they are saturated lexically (in an argumental, morphological or "understood" way) in Spanish (See section 4.2). Secondly, we will show how the information regarding semantic components and its particular saturation (specified in VRQS) is closely related to diathesis alternations (See section 5). Finally, we will try to encode all this information in the feature <vrqs> of the Lexical Knowledge Base (See section 6).

From a lexicographic point of view, the information included in the VRQS can be also used to identify or detect the different verbal senses, because we consider that if the same verbal form has distinct VRQs, then that verb has different senses.

In this work we basically deal with the semantic subsets of motion, cooking, ingesting and weather verbs in Spanish in order to illustrate the treatment and further

---

<sup>3</sup> We will not discuss the information referring to the argument structure, event structure and selectional restrictions because they are treated accurately in Sanfilippo, A. (1991).

encoding of the Compositional Semantic level (VRQS). This work is still in progress and, therefore, it does not seek to be an exhaustive study of the different verbal classes.

## 2.- Levels of Verbal Semantic Representation

This work starts with the hypothesis that the meaning of a verbal lexical item is composed of the sum of different levels of semantic expressivity. Basically, we can distinguish four types of semantic representation: Argument Structure, Event Structure, Selectional Restrictions and Compositional Semantics.

Information about the arity and the semantic relation between the predicate and its arguments is given in the Argument Structure. This relation is usually formulated by means of thematic roles, proto-roles or variables over arguments (Grimshaw, J. 1992, Rappaport, M. & Levin, B. 1986 and 1988, Dowty, D. 1989, Sanfilippo, A. 1991, and others). Argument Structure shows the way semantic arguments are related to their syntactic (grammatical) expressions. This level of representation is in clear correspondence with the syntactic information concerning the subcategorization and grammatical information of the predicate (Argument selection or Linking).

In the event structure, it is considered that verbs denote a type of “event” with respect to nouns denoting “entities”. If we use the generally accepted classification of Vendler, Z. (1967) and Dowty, D. (1986), “events” can be classified into three types: states, processes and events (the last one can be also subclassified into accomplishments and achievements). Therefore, encoded in the Event Structure is the information referring to the temporal relationships expressed in the verb: that is, the information about aspectual behavior. Verbal aspect is determined by the inherent aspectual information denoted in the verbal stem, by the specific morphological features and by the context in which the verb appears (the presence or absence of a particular argument, modifier, etc. can modify aspectual behaviour). (Dowty, D. R. 1986, Tenny, L. 1988)<sup>4</sup>

Selectional Restrictions are necessary to determine the sort of objects (or “entities”) which can flesh out the arguments requested by the verb (‘human’, ‘animal’, ‘inanimate’, ‘abstract’, etc.).

Finally, a level of Compositional Semantics is also necessary where the different semantic components or entities which make up the verbal lexical meaning are included. At this level of representation we propose a semantic classification of verbs according to the different semantic components which they are composed of.

---

<sup>4</sup>Other interesting works about the event structure are Sanfilippo, A. (1991) and Alonge, A. (1991).

This level should not be confused with that of the argument structure. The basic difference between them is that argument structure specifies those semantic components of meaning that have an expression in the arguments which are subcategorized syntactically by the verb, whereas compositional representation includes the necessary components for the characterization of a verb as belonging to a specific semantic class. Further on we will expose how this level of semantic representation also interferes with the information of a syntactic sort (See section 5).

Each one of these semantic levels brings different information to the verbal meaning and, therefore, in order to give a complete semantic characterization of verbs it is necessary to represent every one of them. In our opinion, the lexical meaning of a verb can not be exclusively reduced to the representation of argument structure or whichever of the other levels mentioned, even though the representations may be very sophisticated. At the moment of representing the verbal lexical meaning the different semantic levels must be taken into account. It is also necessary to note that the lexical meaning of verbs can not be understood independently from the syntactic information.

### 3.- Verbs in the Lexical Knowledge Base

Verbal lexical entries are represented in the LKB as lexical signs where morpho-syntactic and semantic information is encoded together (See Figure 1). Verbs are treated as 'head' elements of their sentences, that is, verbs provide the relational and semantic structure for their sentences.

<b>lex-sign</b> (sign)	<b>verb-sign</b> (lex-sign complex-sign)
<orth> = orth	<orth> = orth
<cat> = cat	<cat> = complex-cat
<sem> = sem	<sem> = verb-sem
<rqs> = rqs	<rqs> = vrqs.
<sense-id> = sense-id <sup>5</sup> .	

**Figura 1:** *Lexical and Verbal Sign.*

The morpho-syntactic information referring to the lexical category and the inherent morphological properties of verbs is included in the feature <cat>. Also included are the patterns of subcategorization and the sort of syntactic alternances<sup>6</sup> or diatheses shown by verbs.

---

<sup>5</sup> The feature 'sense-id' specifies the information about the source dictionary, language, sense, etc. of the entry.

<sup>6</sup> In Sanfilippo (1992), Taulé et al. (1993, 1994) there is an adequate treatment of this kind of information.

Bearing in mind the semantic information and the different levels of verbal semantic representation mentioned earlier, Argument Structure, Event Structure and the specification of Selectional Restrictions are encoded in the feature <verb-sem>, while the information concerning the different semantic components is defined in the <vrqs> feature.

Semantic information included in <verb-sem> is represented as a conjunctive logical form where verbal predicate is characterized as a type of event and the semantic relation between the verbal predicate and its arguments is expressed by means of thematic proto-roles. In the logical form, also specified are the selectional restrictions that can flesh out the arguments of the verbal predicate. This representation has been proposed by A. Sanfilippo (1990, 1991), where a neodavidsonian approach to verbal semantics (Parsons, T. 1990, Carlson, G. 1984, Dowty, D. 1989) is combined with a characterization of the thematic relations as proto-roles (Dowty, D. 1988), in a framework of a unification-based Categorical Grammar (Zeevat et al. 1987)<sup>7</sup>.

The last level of semantic representation, encoded in the LKB as <vrqs>, includes the information concerning the semantic class to which the verb belongs and also the information about the verbal arguments saturated by the verb.

Semantic information included in <verb-sem> is closely connected with the syntactic information concerning verbal subcategorization, in the sense that the arguments of the logical form correspond with the subcategorized arguments of the verb<sup>8</sup>, that is to say, if two arguments are subcategorized by the verb, its logical form also has two semantic arguments. In the verbal sign, syntactic and semantic information is related by means of the coindexation of the subcategorized arguments specified in <cat> and the arguments of the logical form specified in <verb-sem>, in the way that the verbal argument selection is shown.

We now present, in more detail, the nature and necessity of a level of compositional representation.

---

<sup>7</sup> In A. Sanfilippo (1991) the kind of semantic information included in <verb-sem> is expressed accurately.

<sup>8</sup> We assume, following the proposal in "HPSG", that the NP subject is also subcategorized by the verb.

## 4.- “VROS”: A Level of Compositional Semantics

### 4.1.- “VRQS” and Decomposition

The level of compositional representation includes the abstract specification of the meaning of a concrete verb by means of its decomposition into semantic components or entities (Talmy, L. 1985, Jackendoff, R. 1983, 1990, etc.). Each particular semantic domain is characterized by a defined subset of semantic components.

Following the methodology proposed in Levin, B. (1993), we have used diathesis alternations as a criterion to establish the semantic classes of verbs. Moreover, the study of the different alternances admitted by a concrete verb class can also be useful to detect or identify the linguistically relevant semantic components. That is, if we consider that verb behaviour is to a large extent determined by its lexical meaning, then the study of this behaviour (i.e. diathesis alternations) can be used, at the same time, to define coherent semantic classes of verbs. For instance, motion verbs participate, generally, in “oblique” alternations (i.e.: omission of argument Path, measure phrase or preposition drop alternation, etc.) and they do not admit causative diatheses or unspecified object alternation. Verbs belonging to the semantic domains of cooking and ingesting show transitive alternations, even though they have peculiar behaviour (for instance, cooking verbs participate in causative constructions whereas ingesting verbs do not and prefer the unspecified object alternation,...) etc.

Semantic components are, it would seem, a discrete set, but the (noun) objects that can flesh them out (i.e. selectional restrictions) differ according to the semantic class of verb. The sort of object of a verb that belongs to the semantic domain of cooking will be quite different to the object Patient of a motion verb; an Agent of a cooking verb will always be human, an Agent of an ingesting verb will be animate whereas an Agent of a motion verb will be animate or inanimate, depending on the verb, and a weather verb will refer to a temporal noun, etc.

Hence, verbal semantic classes are defined according to the semantic components which constitute the verbal meaning, according to the possible combinations of these components and according to the possible restrictions in the values of the components.

Relying on the four subsets of verbs treated throughout this work, we next illustrate the way we carried out the decomposition of meaning. We will also present a sample of verbs belonging to these semantic classes and the kind of diathesis alternations in which they can participate.

#### 4.1.1.- Cooking Verbs

The abstract representation of the meaning of a cooking verb (P.e.: 'asar' (to roast), 'cocinar' (to cook), 'guisar' (to stew), 'emparrillar' (to grill), 'brasear' (to barbecue), 'sofreír' (to fry lightly), 'rehogar' (to cook slowly), 'salpimentar' (to season), 'especiar' (to spice), etc.) involves, basically, the Agent, Patient, Medium, Manner and Cook components. That means, a predicate which belongs to the semantic class of cooking needs a human Agent (always causative in this subset) who carries out a cooking action (Cook), in a specific medium (Medium) (fire, grill, burning-coals, etc.), over a Patient, in principle, edible which will change its state. Moreover, verbs belonging to this semantic class can also give information about manner (Manner) of cooking (lightly, quickly, slowly, etc.).

Cooking verbs describe, basically, the different ways of cooking food. They fall into four subclasses according to the most outstanding component. There follows a sample of these subclasses of cooking verbs.

a) Verbs which describe the basic methods of cooking:

- asar (to roast), calentar (to heat), cocer (to cook), cocinar (to cook), escalfar (poach), freír (to fry), guisar (to stew), hervir (to boil), tostar (to toast), etc.

b) Verbs which express the manner of cooking, making a special point of the degree or intensity of cooking (i.e. too-much, little, not-much, slightly, etc.) or cooking speed (i.e. quickly, slowly, on a high or low flame, etc.):

- achicharrar (to fry crisp), aderezar (to season or garnish), asurar (to burn), dorar (to brown or cook lightly), recocer (to overcook), rehogar (to cook slowly, braise), requemar (to overdo), retostar (to toast too much), saltear (to sauté), sofreír (to fry lightly), etc.

c) Verbs which specify a particular ingredient (i.e. salt, pepper, tuna, etc.). Most of these verbs are zero-related to the names of the main ingredient involved:

- acecinar (to salt the meat), adobar (to season the meat), aliñar (to dress the salad), condimentar (to flavour), confitar (to preserve in syrup), escarchar (to candy), especiar (to spice), granizar (to make an iced drink), perdigar (to half-cook), resalar (to oversalt?), salar (to salt), salpimentar (to add salt and pepper to), sazonar (to season), etc.



d) Finally, verbs which indicate the medium (or utensil) used or implied by the cooking process (i.e. oven, live-coals, grill, fire, etc.). Verb members of this class are also zero-related to the names of mediums or utensils used by the process.

- brasear (to barbecue), emparrillar (to grill, broil), gratinar (to cook in the oven), hornear (to bake), etc.

Cooking verbs are characterized as transitive verbs and, consequently, the kind of diathesis alternances in which they can participate are transitive<sup>9</sup> too: “reflexive passive”<sup>10</sup>, passive, pronominal transitive and unspecified object.

- (1) a. Clara tuesta el pan. (Transitive)  
(‘Clara toasts the bread’)
- b. El pan se tuesta. (“Reflexive Passive”)  
(‘The bread is toasted’)
- c. ??Clara tuesta. (Unspecified Object)  
(‘Clara toasts’)
- d. El pan es tostado [por Clara]<sup>11</sup>. (Passive)  
(‘The bread is toasted [by Clara]’)
- e. \*Clara<sub>i</sub> se<sub>i</sub> tuesta (a si misma). (Reflexive)  
(‘Clara toasts herself’)
- f. Clara se tuesta el pan. (Pronominal Transitive)  
(‘Clara toasts the bread’)

When cooking verbs are used in the passive construction and in the pronominal transitive use, the “reflexive passive” is, at times, quite forced, in the same way as intransitive use results in unspecified object alternation. In fact, only those verbs describing the basic methods of cooking (‘cocinar’ (to cook), ‘asar’ (to roast), etc.) participate clearly in the unspecified object alternation.

This kind of verb can never admit reflexive alternations because the arguments Agent and Patient involved in the cooking action always have to be referentially different, and reflexive alternation requires the semantic identification between these two semantic components.

---

<sup>9</sup> In Taulé, M. (1994) (forthcoming) a full treatment of Spanish diathesis alternations is included.

<sup>10</sup> We use “Reflexive Passive” as the translation of the Spanish term Pasiva Refleja.

<sup>11</sup> Square brackets are used to indicate the optionality of an argument.

#### 4.1.2.- Ingesting Verbs

In the case of ingesting verbs (P.e.: 'comer' (to eat), 'beber' (to drink), 'tragar' (to swallow), 'pacer' (to graze or pasture), 'bellotear' (to feed acorns), 'adaguar' (to drink), 'cenar' (to dine), 'malcomer' (to eat badly), 'glotonear' (to be greedy or gluttonous), etc.), the semantic components involved are the Agent, necessarily animate, a Patient, preferably edible which is ingested (Ingest) by the agent in a particular way (Manner) (slowly, quickly, without chewing, etc.).

This kind of verb can also be classified in further subclasses —a sample is shown below— according to whether their meanings stick out a particular semantic component.

a) Verbs which describe the simple action of ingesting:

- beber (to drink), comer (to eat), ingerir (to ingest).

b) Verbs related to the time of ingesting a particular meal (i.e. morning, afternoon, noon, etc.):

- almorzar (to lunch), cenar (to dinner, dine), desayunar (to have breakfast), merendar (to have tea).

c) Verbs which express the way of ingesting, that is, the manner of eating or drinking something (i.e. slowly, quickly, without masticate/chew, etc.):

- comiscar (to nibble), chascar (to swallow), chupar (to suck), deglutir (to swallow), devorar (to devour), embeber (to imbibe), embocar (to wolf), embuchar (to wolf, bolt), engullir (to gobble, guzzle), lamer (to lick), mamar (to breastfeed, suckle), mascar (to chew), masticar (to masticate), papar (to gulp), picar or picotear (to nibble, pick), ronzar (to munch, crunch), rumiar (to chew the cud), sorber (to sip), tragar (to swallow), tragonear (to guzzle?), etc.

- atracarse (to gorge), glotonear (to be greedy or gluttonous), malcomer (to eat badly), etc.

d) Verbs which express the manner of ingesting something and which also produce a change in the physical state of the agent, or whoever experiences the action. These kind of verbs are included in the subclass of "Emborrachar". Their members

participate in a particular kind of diathesis alternations, the Reflexive<sup>12</sup> ones. They show different syntactic behaviour and that is why we treat them in a specific subclass.

- ahitar (to stuff), atragantar (to swallow the wrong way), emborrachar or embriagar (to intoxicate, to make or get drunk), empachar (to get indigestion, get stuffed), hinchar (to swell up?), etc.

e) In the last class, we have included those verbs which describe the specific food or drink ingested (i.e. water, wine, grass, pasture, etc.), that is, those verbs with the patient argument incorporated into the verbal stem.

- adaguar (to drink water (animals)), ahojar (to eat leaves (animals)), bellotear (to eat acorns (animals)), cuscurrear (to eat croutons), escanciar (to drink wine), frezar (to spawn), herbajar (to graze), pimplar (to drink wine), etc.

- apacentar (to graze), apastar (to graze), montanear (to graze in a hill), pacer, pastar or pastear (to graze, pasture), etc.

These subclasses of verbs show different sets of diatheses depending on the outstanding semantic component. The four first subclasses take part in transitive alternations whereas the last one —those verbs with the patient component incorporated in the verb— participate in the intransitive alternances.

In contrast to cooking verbs, the majority of ingesting transitive verbs can occur in intransitive constructions, that is, they can admit the unspecified object alternation. They can also appear in the passive, whereas the “reflexive passive” use is very limited, not to say inexistent. The “reflexive passive” in ingesting verbs is very forced if not assisted by a modal or instrumental complement (2b) and (2b’)) and it becomes a statement of a general kind implying a change of meaning (it has an impersonal reading more than a reflexive passive reading).

- (2) a. El niño come un huevo frito. (Transitive)  
      (‘The child eats a fried egg’)  
      b. ??El huevo frito se come con tenedor. (“Reflexive Passive”)  
      (‘The fried egg is eaten with a fork’)  
      b’.??El huevo frito no se come con las manos. (“Reflexive Passive”)

---

<sup>12</sup> Taulé, M. (1994) (Forthcoming).

- (‘The fried egg is not eaten with the hands’)
- c. El niño come. (Unspecified Object)  
(‘The child eats’)
- d. Un huevo frito es comido [por el niño]. (Passive)  
(‘A fried egg is eaten [by the child]’)
- e. \*El niño; se; come (a si mismo). (Reflexive)  
(‘The child eats himself’)
- f. El niño se come un huevo frito. (Pronominal Transitive)  
(‘The child eats a fried egg’)

With respect to the reflexive alternances, they are not accepted in a general way, only a special subclass of these ingesting verbs —those included in the “Emborrachar” class— admit the reflexive use and the so called source reflexive alternation.

- (3) a. Manuel emborracha a Pedro. (Transitive)  
(‘Manuel gets drunk Pedro’)
- a’. El vino emborracha a Pedro. (Transitive)  
(‘Wine makes drunk Pedro’)
- b. \*Manuel emborracha. (Unspecified Object)  
(‘Manuel inebriates’)
- b’. El vino emborracha. (Unspecified Object)  
(‘Wine inebriates’)
- c. Pedro es emborrachado [por Manue/el vino]. (Passive)  
(‘Pedro is got drunk [by Manuel/wine].’)
- d. Manuel; se; emborracha. (Reflexive).  
(‘Manuel gets drunk’)
- e. Manuel; se; emborracha con vino. (Source Reflexive)  
(‘Manuel gets drunk on wine’)
- f. \*Manuel se emborracha a Pedro. (Pronominal)  
(‘Manuel gets drunk Pedro’)

In these verbs, the source reflexive and the unspecified object alternations can only be accepted in the case in which the subject of the transitive use is not a human Agent (See the contrast between (40a), (40b) and (40a’), (40b’)).

#### 4.1.3.- Motion Verbs

The semantic domain of motion verbs (P.e.: ‘mover’ (to move), ‘andar’ (to walk), ‘correr’ (to run), ‘salir’ (to come or go out), ‘entrar’ (to come in), ‘subir’ (to go up), ‘bajar’ (to go down), ‘zigzaguear’ (to zigzag), ‘bordear’ (to skirt), ‘gatear’ (to crawl or to

go on all fours), 'aletear' (to wing), etc.) involves, basically, Agent, Patient, Path, Manner and Motion as semantic components. In this domain, it is necessary to distinguish between verbs characterized by the presence of an Agent who causes the motion of Patient (the object which is move), from those motion verbs, called 'unaccusatives', where it seems that there is an identification between Agent and Patient (P.e.: *Clara entra las sillas* vs. *Clara siempre entra en casa con los pies sucios* o *Clara anda deprisa* (Clara brings in the chairs vs. Clara always goes into the house with dirty feet or Clara walks fast)). Motion also expresses the Path performed by the moved or displaced element. Path components can be of different types: basically, a Path can be bounded (i.e. including source-paths or goal-paths) or unbounded (i.e. describing directions or routes) (Jackendoff, R. 1983), we can also distinguish between verbs involving a displacement and verbs involving movement without displacement (contained motion), etc. Moreover, verbs belonging to the motion class can also describe the manner, or medium, (Manner) in which the action is carried out and the part of the body involved in the movement.

Taking into account the main semantic component involved in motion verbs, we have subclassified this domain, basically, into the next four subclasses<sup>13</sup>.

a) Verbs which specify the direction of motion (i.e. bounded or unbounded Path):

- acercarse (to approach), abanzar or adelantar (to move forward, advance), aproximarse (to come near), arribar (to arrive, reach), ascender (to ascend, go up), atravesar (to go across), bajar (to get, let, or take down), caer (to fall down), cruzar (to cross), descender (to descend), entrar (to go or come in, enter), escalar (to climb), escapar (to escape, run away), huir or fugarse (to run away from, avoid, flee), ir (to go), llegar (to arrive), marchar (to go, leave), partir (to depart), pasar (to pass), regresar (to come back), retroceder (to go back), salir (to go out, exit), subir (to go up), tornar (to return), trepar (to climb, scale), venir (to come), volver (to go or turn back),

b) Verbs which include a notion of manner (or means) of motion:

- anadear (to waddle), andar (to walk), arrear (to hurry), atajar or atrochar (to take a short cut), balsear (to raft), bambolear or bambonear (to swing, sway), barquear (to canoe), bogar (to row), bornear (to twist), brincar (to bounce, jump, hop), cabalgar (to

---

<sup>13</sup> We could find more subclasses in this semantic domain and, possibly, we could make further distinctions in those described.

ride), *caminar* (to walk), *circular* (to circulate), *cojear* (to limp, hobble), *conducir* (to drive), *correr* (to run), *costear* (to sail along the coast, to coast), *culebrear* (to slither, wriggle), *chorrear* (to gush, spout out), *deambular* (to wander, saunter), *desbordarse* (to flood), *desfilar* (to march, parade), *errar* (to wander), *escabullirse* (to slip away), *escarabajear* (to wriggle), *esquiar* (to ski), *gatear* (to crawl), *jinetear* (to ride), *mudarse* (to move), *nadar* (to swim), *navegar* (to sail), *pasear* (to go for a walk), *patinar* (to skate), *peregrinar* (to go on a pilgrimage), *planear* (to glide), *recorrer* (to go over), *remar* (to row), *renquear* (to limp, hobble), *reptar* (to creep), *revolotear* (to flutter), *rondar* (to roam, prowl about)), *serpear* or *serpentear* (to wriggle, creep), *talonear* (to walk briskly, to hurry along), *titubar* or *titubear* (to stagger), *trafagar* or *trajinar* (to bustle about), *transitar* (to go along, travel), *trotar* or *trotear* (to trot), *vagar* (to wander, roam), *viajar* (to travel), *volar* (to fly), *zapatear* (to tap, to tap-dance), *zarpar* (to weigh anchor), *zigzaguear* (to zigzag), etc.

c) Verbs which involve internal body motion. We could also distinguish those sort of verbs which are zero-related to names of a body part (i.e. wing, eyelash, tail, etc.) from those in which the body part which is moved is not specified (in italics):

- *alear* or *aletear* (to flutter, flap), *alzar* (to lift up, to raise up, to rise), *aupar* (to help up), *befar* (to jeer at, taunt), *bracear* (to swing one's arms), *cabecear* (to bind), *codear* (to elbow), *colear* (to wag, tail), *contonearse* (to swagger, swing, wiggle), *chapotear* (to splash about), *erguir* (to raise, straighten up), *hocicar* (to root), *levantar* (to raise, lift up, arouse), *menear* (to shake), *mover* (to move), *nalguear* (to shake the buttocks), *patalear* (to stamp), *pestañear* (to blink, wink), *rabear* (to wag), etc.

d) Verbs which describe the performance or manner of dancing, included in the "Dance" subclass. In this subclass, we can also find those verbs zero-related to names of dances (i.e. rumba, salsa, waltz, etc.):

- *bailar*, *bailotear*, *danzar* or *tripudiar* (to dance)
- *polcar* (to polka), *rumbear* (to rumba), *salsear* (to salsa), *valsar* or *valsear* (to waltz), etc.

Verbs belonging to semantic motion class are usually characterized by the

predominance of the so called “oblique” and intransitive alternations, like the omission of Path argument, preposition drop alternation, measure phrase alternation and directional phrase alternation. A specific sort of reciprocal alternation is extended, in principle, to the whole class.

- (4) a. El caballo llegó a la meta sin jinete. (Oblique Intransitive)  
 ('The horse arrived at the winning post without a rider')
- b. El caballo llegó sin jinete. (Omission of Path)  
 ('The horse arrived without a rider')
- c. El caballo cruzó [por] la meta sin jinete. (Preposition drop)  
 ('The horse crossed the winning post without a rider')
- d. El caballo cabalgó cien metros sin jinete. (Measure Phrase)  
 ('The horse galloped a hundred metres without a rider')
- e. El caballo cabalgó hacia la meta sin jinete. (Directional Phrase)  
 ('The horse galloped to the winning post without a rider')
- f. El caballo y el jinete llegaron juntos a la meta. (Reciprocal Alt.)  
 ('The horse and the rider arrived together at the winning post')

The occurrence of these verbs in transitive alternations is more restrictive and, basically, affects three subclasses: 1) verbs which involve internal body motion, which admit the reflexive and passive uses (5); 2) verbs which describe the performance or manner of dancing, which participate also in the passive and in the unspecified object alternation (6) and 3) verbs which indicate, usually, the manner of motion that can be used transitively (7) and, therefore, which can take part in the passive construction and, sometimes, in the “reflexive passive” use too.

- (5) a. Celia Cruz mueve las caderas con mucha gracia. (Transitive)  
 ('Celia Cruz moves/shakes/swings her hips gracefully')
- b. Celia Cruz; se; mueve con mucha gracia. (Reflexive)  
 ('Celia Cruz moves/shakes/swings gracefully')
- (6) a. Celia Cruz bailó todas las rumbas con su esposo. (Transitive)  
 ('Celia Cruz danced all the rumbas with her husband')
- b. Celia Cruz bailó con su esposo. (Unspecified Object)  
 ('Celia Cruz danced with her husband')
- c. Celia Cruz y su esposo bailaron juntos. (Reciprocal alt.)  
 ('Celia Cruz and her husband danced together')
- (7) a. El viento movía las hojas de los árboles. (Transitive)  
 ('The wind moves the leaves on the trees')

- b. Las hojas de los árboles se movían. (“Reflexive Passive”)  
 (“The leaves of the trees moved”)

Pronominal alternations are only accepted by those verbs indicating the direction of motion (8).

- (8) a. El acusado marchó de la sala furioso. (Oblique Intransitive)  
 (“The accused left the law court furiously”)  
 b. El acusado se marchó de la sala furioso. (Pronominal)  
 (“The accused left the law court furiously”)

#### 4.1.4.- Weather Verbs

Verbs related to the semantic domain of weather (P.e.: ‘llover’ (to rain), ‘nevar’ (to snow), ‘amanecer’ (to dawn), ‘oscurecer’ (to get dark), ‘neviscar’ (to snow lightly), ‘lloviznar’ (to drizzle), ‘chispear’ (to drizzle), ‘diluviar’ (to pour with rain), etc.) express the components Agent, Manner and Weather. These kind of verbs indicate a weather phenomenon (Weather) by means of the specification of a temporal Agent (rain, snow, wind, stone, etc.) and the way (Manner) it happens (intensively, lightly, etc.).

Here, we list a sample of verbs belonging to this class. We have also included in it verbs like ‘amanecer’ (to dawn), ‘oscurecer’ (to get dark), etc. (in italics), in which a change of “temporal” state is also involved, therefore showing a different behaviour from the rest of weather verbs.

- *amanecer* (to dawn), *anocheecer* (to get dark), *apedrear* (to hail), *clarear* (to clear, brighten up), *diluviar* (to pour with rain), *chaparrear* (to pour in torrents), *chispear* (to drizzle), *chubasquear* (to storm), *granizar* (to hail), *llover* (to rain), *lloviznar* (to drizzle), *nevar* (to snow), *neviscar* (to snow lightly), *nublarse* (to mist), *oscurecer* (to get dark), *pintear*, *rachear* (to gust), *relampaguear* (to lightning), *tronar* (to thunder), *ventear* (to blow), *ventiscar* (to drift), etc.

The kind of alternances shown by these verbs are reduced to the presence of prepositional phrases (9b) or nominal phrases related to verbs (9b), that is, they can admit the cognate object alternation.

- (9) a. LLovía mucho. (Impersonal)  
 (“It was raining”)



- b. LLuvia a cantaros/sin parar. (Prepositional Phrase)  
(‘It was raining cats and dogs’)
- c. LLuvia barro. (Cognate Object)  
(‘It was raining mud’)

#### 4.2. “VRQS” and Lexical Saturation

Once a verbal lexical meaning is decomposed into semantic components, each of them must be satisfied or saturated lexically so that the meaning of that verb might be completed. The saturation of these arguments or components can be carried out in three different ways: structurally, morphologically or by means of the so called ‘understood’ arguments.

a.) In the first type of saturation, called structural or argumental saturation, the semantic components are saturated syntactically and independently from the verbal form. This type of saturation is related to the semantic arguments subcategorized by the verb and, therefore, is a kind of information included in the Argument Structure.

Here, we give examples of structural or argumental saturation:

(10a) Manuel asa/cocina/emparrilla/brasea/soasa/rehoga la carne.  
Agent Cook+Medium+Manner Patient

(‘Manuel roasts/cooks/grills/barbecues/fries\_lightly/cooks\_slowly the meat.’)

(10b) El niño come/bebe/traga/ingere algo con glotonería.  
Agent Ingest Patient Manner

(‘The child eats/drinks/swallows/consumes something with gluttony/greediness.’)

(10c) Clara entra/sube/baja/anda/corre/salta/gatea/nalguea.  
Agent Motion+Path+Manner+(Patient)

(‘Clara goes\_in/goes\_up/goes\_down/walks/runs/jumps/crawls or goes\_on\_all\_fours/shakes\_the\_buttocks.’)

(10d) Nunca nieva/llueve/graniza en Barcelona.  
Weather+Agent+Manner

(‘It never snows/rains/hails in Barcelona.’)

In the above examples, transitive use of a cooking verb (10a) or an ingesting verb (10b) entails the structural or argumental saturation of the Agent and Patient components: these arguments are expressed syntactically by two nominal phrases (10a) and (10b). In example (10c), the intransitive use of a motion verb entails the structural or argumental saturation of the Agent component. This argument is expressed syntactically by a nominal phrase. Finally, in the case of the impersonal verbs of weather (10d) its semantic components are not saturated in a structural way.

b.) In morphological saturation, the semantic components are carried out in the verbal form by means of a morphological process of lexical derivation. The expression of an argument which is morphologically saturated can be done either by means of prefixes and suffixes at the verbal form (11), or by means of the nominal or adjectival stem from which the verb is derived (12) (Baker, M. 1988, Hale, K. & Keyser, S.J. 1993).

(11) Soasar, requemar, brasear, saltear, picotear, glotonear, nalguear, gatear, chispear, lloviznar, etc.

The underlined prefixes and suffixes in verbs of (11) express, in these cases, the semantic component of Manner.

“so-” --> lightly  
“re-” --> lightly  
“-ear” --> manner of

In the examples of (12), we can observe that the argument which is morphologically saturated in the verbal form can be related to different semantic components<sup>14</sup>: Agent, Patient, Medium, Manner, etc. Further on, we will see what kind of syntactic repercussion this type of saturation creates.

(12) “brasear” --> ‘cocinar en brasas’ (Medium).  
“emparrillar” --> ‘cocinar en parrilla’ (Medium).  
“adaguar” --> ‘beber agua el ganado’ (Patient).  
“pajear” --> ‘comer paja el ganado’ (Patient).  
“nalguear” --> ‘mover las nalgas’ (Patient).  
“gatear” --> ‘andar a gatas’ (Manner).  
“nevar” --> ‘caer nieve’ (Agent).  
“llueva” --> ‘caer lluvia’ (Agent).

---

<sup>14</sup> Verbs presenting this type of saturation are zero-related to the name from which they are derived.

c.) Finally, in the last type of lexical saturation, called ‘understood’ saturation, the semantic components, although part of the verbal meaning, are not expressed in an explicit way, either syntactically or morphologically, in the verbal form. For this reason, they are called ‘understood’ arguments. We can explain, for instance, the intransitive use presented in some ingesting verbs (13), like ‘beber’ (to drink), ‘comer’ (to eat), ‘tragar’ (to swallow), etc., or in some cooking verbs (14), like ‘asar’ (to roast), ‘cocinar’ (to cook), ‘emparrillar’ (to grill), etc., where the syntactic omission of the Patient component is allowed but is still perfectly inferred. In the same way, this type of saturation also reports on the syntactic alternation of the Path argument presented in some motion verbs (15), and it may be structurally omitted or not.

(13a) Los monos comen con las manos.  
 Agent Ingest+Patient Manner

(‘Monkeys eat with their hands.’)

(13b) Los monos comen la fruta con las manos.  
 Agent Ingest Patient Manner

(‘Monkeys eats fruit with their hands.’)

(14a) Manuel cocina muy bien.  
 Agent Cook+Medium+Manner+Patient

(‘Manuel cooks very well.’)

(14b) Manuel cocina la carne muy bien.  
 Agent Cook+Medium+Manner Patient

(‘Manuel cooks the meat very well.’)

(15a) Los estudiantes bajan de forma ruidosa.  
 Agent Motion+Manner+Patient+Path

(‘The students go\_down noisily’)

(15a) Los estudiantes bajan por las escaleras de forma ruidosa.  
 Agent Motion+Manner+Patient Path

(‘The students go down the stairs noisily’)

## 5.- “VRQS” and Diatheses

In the previous section, we have already noted the important link between semantic and syntactic information. Next, we will try to show how information included in VRQS interacts with that referring to diatheses. Firstly, we will show the different syntactic consequences of lexical saturation and then we will see how semantic components are related or sensitive to specific diathesis alternations.

### **5.1.- Saturation and Diatheses**

In this section we will focus on the syntactic consequences of the different types of lexical saturation. Semantic components can only be saturated lexically once: structurally, morphologically or in the ‘understood’ way. The difference between them is that the first two types of saturation can block alternations in the patterns of subcategorization of verbs, whereas the ‘understood’ saturation allows alternances. Therefore, we can establish an essential distinction between the morphological saturation carried out by a lexical derivation on one hand, and the ‘understood’ saturation, on the other.

a.) In the first type of saturation, the presence of the morpheme (underlined in (16)) and the noun or adjective (underlined in (17)), in the verbal form, can block the syntactic realization of the arguments represented by the suffix or the noun or adjective.

(16a) Manuel soasa/rehoga la carne en la parrilla/brasas/sartén/cazuela.  
Agent Cook+Manner Patient Medium

(‘Manuel fries\_lightly/cooks\_slowly the meat in the  
grill/barbecue/frying\_pan/casserole.’)

(16b) Manuel soasa/rehoga la carne ligeramente/a fuego lento.  
Agent Cook+Manner+Medium Patient Manner

(‘Manuel fries\_lightly/cooks\_slowly the meat lightly/slowly.’)

(16c) Manuel soasa/rehoga la carne con toda tranquilidad.  
Agent Cook+Manner+Medium Patient Manner

(‘Very calmly, Manuel fries\_lightly/cooks\_slowly the meat.’)

In example (16), we can observe how the morphological saturation of the Manner

component at the verbal form, emphasizes the syntactic expression of the component with the prefix “so-” or “re-” (16b). Probably, the meaning of these prefixes is that the action of cooking is carried out “a fuego lento” (slowly) or “ligeramente” (lightly) respectively. In these cases, we can talk about an emphatic use of the Manner component which is also expressed syntactically, because, in fact, this semantic component has already been saturated and its syntactic expression is simply redundant.

On the other hand, this does not imply that in sentences with verbs like ‘soasar’ or ‘rehogar’ another complement of manner might appear which would modify the whole predicate (16c); in this case the prepositional phrase “con toda tranquilidad” modifies or qualifies the Agent action but not the manner of cooking the Patient.

Now we can look over the syntactic consequences of the morphological saturation which result from a lexical derivative process by a noun or adjective:

(17a)	Manuel	<u>emparrilla/brasea</u>	la carne	ligeramente/a fuego lento.
	Agent	Cook+ <u>Medium</u>	Patient	Manner

(‘Manuel grills/barbacues the meat lightly/slowly.’)

(17b)*	Manuel	<u>emparrilla/brasea</u>	la carne	<u>en la parrilla/brasas.</u>
	Agent	Cook+Manner+ <u>Medium</u>	Patient	<u>Medium</u>

(\*‘Manuel grills/barbacues the meat in the grill/barbacue.’)

(17c)	Manuel	<u>asa/cocina/guisa</u>	la carne	<u>en la parrilla/brasas.</u>
	Agent	Cook+Manner	Patient	<u>Medium</u>

(‘Manuel roasts/cooks/stews the meat in the grill/barbacue.’)

(18a)	Todos	<u>nalguean</u>	cuando bailan salsa.
	Agent	Motion+Manner+ <u>Patient</u> +Path	

(‘Everyone ‘shakes\_buttocks’ when they dance salsa.’)

(18b)	*Todos	<u>nalguean</u>	las <u>nalgas</u>	cuando bailan salsa.
	Agent	Motion+Manner+ <u>Patient</u> +Path	<u>Patient</u>	

(\*‘Everyone ‘move\_buttocks’ the buttocks when they dance salsa.’)

(18c) Todos mueven las nalgas cuando bailan salsa.  
 Agent Motion+Manner+Path Patient

(‘Everyone moves their buttocks when they dance salsa.’)

(19a) Nunca nieva en Barcelona.  
 Weather+Agent+Manner

(‘It never snows in Barcelona.’)

(19b) \*Nunca la nieve nieva en Barcelona.  
Agent Weather+Agent+Manner

(\*‘The snow never snows in Barcelona.’)

(19c) \*Nunca nieva nieve en Barcelona.  
 Weather+Agent+Manner Agent

(\*‘It never snows snow in Barcelona.’)

In these examples (17-19), we can observe clearly how the morphological presence of the noun, from which the verb is derived blocks, in the verbal form, the syntactic realization of the semantic component. Therefore, this kind of information can be very useful in the treatment of the syntactic behavior of these verbs, concretely in order to treat alternances in the patterns of subcategorization or diatheses.

Following the above examples, in the case of cooking verbs like ‘emparrillar’ or ‘brasear’ (17), the morphologically saturated semantic component is the Medium which blocks the possibility of expressing syntactically another medium of cooking (17b). The argument saturated at the verbal form and the argument expressed by the prepositional phrase (17b) are incompatible because they present the same degree of specification: in this sense, the expression of this component syntactically will also be redundant.

In the case of the intransitive verb ‘nalguear’ (18), the morphological saturation of the Patient component in the verbal form by the presence of the noun (‘nalga’) from which the verb is derived, blocks the syntactic realization of the semantic component and, therefore, any transitive alternation<sup>15</sup>. The only alternances allowed here are the intransitive ones, basically, the presence or absence of PP arguments. Furthermore, this fact allows us to categorize this kind of verb as strictly intransitive (‘strict-intransitive-sign’) and we can specify the attribute

<sup>15</sup> In Taulé, M. & Castellón, I. (1993), Taulé, M. et al. 1993/4?, Taulé, M. (1994) (in prep.), a treatment of the most important diathesis alternances in Spanish is presented, where a classification of it is established.

<diatheses> of them the value 'strict-intrans-diatheses'.

This kind of phenomena is widespread in the semantic domain of Spanish motion verbs, where the argument Patient is incorporated into the verbal form in order to specify the body part which is moved ('rabear' (to wag the tail), 'alear' (to flutter), 'aletear' (to wing), 'befar' (to jeer or taunt), 'lomear' (to move the back (for an animal)), 'pestañear' (to blink or wink), etc.). This fact allows for the possibility of distinguishing these intransitive motion verbs from those, also intransitive, so called unaccusative verbs where the Patient it is not a specific body part but a whole body. For this reason we say that from a compositional point of view, there is an identification with Agent and Patient ('andar' (to walk), 'correr' (to run), 'moverse' (to move oneself), 'ir(se)' (to go away), 'entrar' (to come in), 'salir' (to go out), 'bajar' (to go down), 'subir' (to go up), etc.).

In impersonal verbs like 'nevar' (19), we can consider the Agent as the morphological saturated argument, which blocks the presence of a nominal phrase subject (19b) and (19c). Therefore, we could explain the impersonality of most 'weather' verbs in the same way that we explain the intransitivity verbs with the component Patient incorporated at the verbal form. In these cases, we can also obtain the lexical category automatically and encode them like impersonal verbs.

In this sense, morphological saturation with nouns incorporated at the verbal stem is stronger than the saturation presented by morphemes (Baker, M. 1988, Hale, K.L. & Keyser, S.J. 1993).

We can also note that this sort of morphological saturation does not have the same lexical consequences in the verbal argument structure. Semantic components like Agent, Patient or Path when morphologically saturated have a clear consequence in argument structure: for instance, a verb with a morphologically saturated Patient presents a monadic argument structure whereas a verb with a structurally saturated Patient has a diadic argument structure. On the other hand, the morphological saturation of components like Manner or Medium does not seem to involve the argument structure of that verb.

b.) Finally, 'understood' saturation does not involve a lexical mark in the verbal form and does not show any block in the syntactic level. So, the 'understood' argument in the verbal form is like an open position that may or may not be carried out syntactically. Therefore, this kind of saturation is richer at subcategorization alternances.

Now we can look over the examples of this kind of saturation:

(20a) Los monos comen la fruta con las manos.  
Agent Ingest Patient

(‘Monkeys eat fruit with their hands.’)

(20b) Los monos comen con las manos.  
Agent Ingest+Patient

(‘Monkeys eat with their hands.’)

(21a) Clara baja por las escaleras.  
Agent Motion Path

(‘Clara goes down the stairs.’)

(21b) Clara baja ahora.  
Agent Motio+Path

(‘Clara goes\_down now.’)

In example (20), the Patient semantic component of an ingesting verb like ‘comer’ can be expressed syntactically (20a) or not (20b). But in the case in which the argument is not expressed syntactically (20b) it is perfectly inferred (or understood) from its verbal form, although it will be, obviously, less specific (they can eat anything) than in (20a) where the object of eat is expressed (‘fruta’).

The same phenomena happens with the semantic component Path in the motion verb ‘bajar’ (21). This argument can be carried out syntactically by a prepositional phrase (21a) or it can be ‘understood’ from the verbal form (21b). When we use a motion verb like ‘bajar’ we expect a Path so that the lexical verbal meaning might be completed, but when the argument is not syntactically expressed it can be inferred from the verbal form and, obviously, in (21b) the Path is less specific.

‘Understood’ saturation, like the morphological saturation mentioned before, also has repercussions in the argument structure. The fact that a verb like ‘comer’ might saturate the argument Patient structurally or morphologically affects its argument structure, in one case it would be diadic and in the other case monadic. The same is valid for verbs like ‘bajar’.



## 5.2.- Semantic Components and Diatheses

The major aim of this section is focused on the relation between diathesis alternations and semantic components (i.e. the interaction between information included in <vrqs> and <diatheses>). In showing this relation we demonstrate the initial hypothesis in which we consider that verbal behaviour, specifically that referring to diatheses, is determined by lexical meaning and, therefore, verbs belonging to the same semantic class share the same set of diathesis alternations (Levin, B. 1993).

Causative alternations —causative-incoative, “reflexive passive”— can only be admitted by those verbs which include the semantic component Cause as part of its lexical meaning and, more concretely, a direct or extern Cause ( Dowty, D. 1979, Jackendoff, R. 1983).

From the point of view of diathesis unspecified or cognate objects are clearly related to the Patient (or Theme) semantic component. Unspecified object alternation is related to verbs in which this Patient component is ‘understood’ saturated, whereas cognate object alternation is presented by verbs where the Patient component is morphologically saturated, that is, the Patient is incorporated into the verbal form.

Reflexive alternances, just like the Passive alternation, are related to Agent and Patient semantic components. In principle, all verbs expressing these two components syntactically can participate in a Passive construction, but not all of them can appear in the reflexive use because this alternation requires the semantic identification between both Agent and Patient components.

Intransitive and Oblique alternations shown in this work —omission of Path, directional phrase, preposition drop and measure phrase— are associated with the notion of motion and, specifically, with the semantic component of Path. For instance, omission of Path alternation is shown in verbs which express the component Path ‘understood’ saturated, preposition drop alternation is related to Paths indicating the Route carried out in the displacement (i.e. verbs like ‘cruzar’ (to cross), ‘atravesar’ (to go across), ...), etc.

In figure 2, we show schematically the different alternations presented by the semantic classes of cooking, ingesting and motion verbs<sup>16</sup>.

---

<sup>16</sup> Taulé, M. (1994) (Forthcoming) includes an accurate exposition of all of these alternations and the relation of them with semantic components. Also included in detail is the different behaviour in the subclasses of mentioned verbs.

Alternations	Cooking Verbs	Ingesting Verbs	Motion Verbs
Reflexive Passive	+	-(*)	-(*)
Causative-Incohesive	-(*)	-	-
Unspecified Object	+/-	+/-	-(*)
Passive	+	+/-	+/-
Reflexive	-	+/-	-
Body-Part Reflexive	-	-	+/-
Source Reflexive	-	+/-	-
Path Omission	-	-	+/-
Reciprocal Omission	+	+?	+
Preposition Drop	-	-	+/-
Cognate Object	-	+/-	+/-
Measure Phrase	-	-	+/-
Directional Phrase	-	-	+/-
Pronominal	+	+/-	+/-

**Figure 2: Diatheses and Semantic Classes<sup>17</sup>**

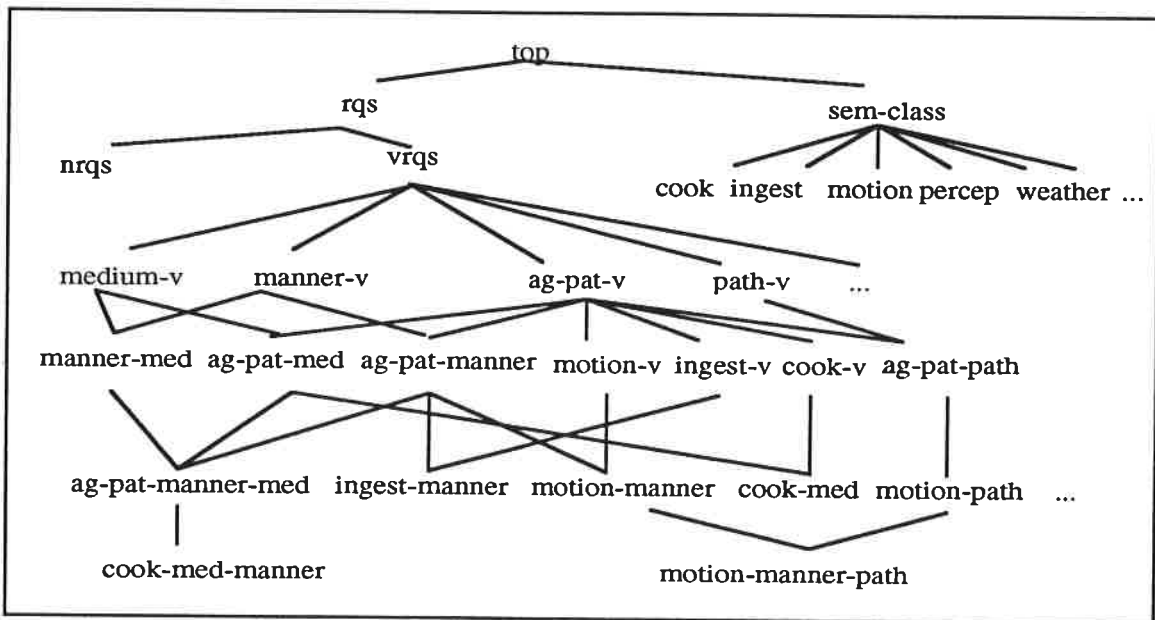
<sup>17</sup> Value of symbols used in Figure 2:

- + Alternation accepted by all verbs of the semantic class
- Alternation does not accepted
- +/- Alternation does not accepted by all verbs of the semantic class, but is admitted in a subset of them.
- (\*) Alternation accepted in a exceptional case.

## 6.- “VRQS” in the LKB

The semantic information concerning VRQS is represented by the feature <rqs> in the verbal entry. Its value is always <vrqs> for the verbal signs (See figure 1). We have adopted this term in order to unify the information contained in lexical entries: in nominal signs we express the semantic information of nouns by the feature <nrqs>, following the initial proposal of J. Pustejovsky (1991). VRQS is a complex typed feature structure where two kinds of information are expressed: the semantic class and the way in which the semantic components of the verb are saturated.

Figure 3 shows a partial view of the semantic classes included in the VRQS of the LKB's type system. The semantic components like 'medium-v', 'manner-v', 'ag-pat-v', 'path-v', etc. are specified at the first level of the hierarchy. In the successive levels, the combination of these semantic components is expressed giving as a result a semantic classification of verbs: 'cook-v', 'cook-manner', 'cook-med', 'ingest-v', 'ingest-manner', 'motion-manner', 'motion-path', 'motion-manner-path', 'weather-v', etc.



**Figure 3:** *Type Hierarchy: Semantic components*

For now, we represent the semantic components Agent and Patient together for descriptive reasons. We assume that all of the verbs can be analyzed according to the combination of these two semantic components (See figure 4): Agent and Patient (both can be morphological or 'understood' saturated)<sup>18</sup>. In this sense, we define an impersonal verb as a verb with a morphological saturated (or incorporated) agent (i.e.: 'nevar' (to snow) has

<sup>18</sup> When a component is argumental saturated it is encoded in the Argument Structure.

the agent 'nieve' (snow) incorporated at the verbal stem, and this fact blocks the syntactic realization of a NP subject), a strict intransitive verb is defined as a verb with an 'understood' agent and an incorporated patient (i.e.: 'alear' (to flutter) where the 'understood' agent is an animal and the incorporated patient is 'alas' (wings), blocking the syntactic realization of this patient) or simply with only an 'understood' agent (i.e.: 'andar' (to walk) where the 'understood' agent is an animate object) and, finally, we can also define a transitive verb as a verb with both 'understood' agent and patient (i.e.: 'comer' (to eat) where the agent is animate and the patient is inanimate) or with the patient saturated argumentally (i.e.: 'emparrillar' (to grill) with an 'understood' animate agent and inanimate argumental saturated patient).

For this reason, the semantic components that provide the labels for the verbal classes ('motion-v', 'ingest-v', 'cook-v', etc.) are 'daughters' of this 'ag-pat-v' semantic component.

Verbs	"Incorporated" agent	"understood" agent	"incorporated" patient	"understood" patient
Impersonal	+	-	-	-
Intransitive	-	+	+ / -	-
Transitive	-	+	-	+

**Figure 4: Definition of verbs.**

Every semantic component has a complex structure in which the information about the type of lexical saturation presented by verbs is specified, for instance:

motion-path (ag-pat-v)                      Path1 (top) (OR bounded unbounded).  
 <sem-class> = motion  
 <morph-saturation-pt> = string  
 <understood-saturation-pt> = path1.

**Figure 5: Type hierarchy, Semantic Components**

## 7.-Conclusions

We have presented a first approach to VRQS representation in a Lexical Knowledge Base, which describes the verbal lexical meaning in terms of its semantic components. We have shown how VRQS can encode some phenomena of lexical saturation —argumental, morphological and ‘understood’— and how this information has important syntactic consequences: we can block a specific subcategorization alternance according to the type of argument saturation which is carried out. We have also shown that the lexical meaning of verbs can not be understood independently from the syntactic information.

On the other hand, information included in the VRQS can be also used to distinguish the different senses of the same verbal form. Therefore, a particular verb (sense) can occur in different patterns of subcategorization (or diatheses) but its VRQS will always be the same.

Until now, we have worked with four different semantic classes of verbs: cooking, ingestion, motion and weather verbs. The future aim of this research is to develop more semantic classes in order to have enough complete representation from different verbs.

## References

Acquilex-II, The Acquisition of Lexical Knowledge, Basic Research Project 7315 ACQUILEX-II, Technical Annexe.

Agno, A., I. Castellón, M.A. Martí, F. Ribas, G. Rigau, H. Rodríguez, M. Taulé, M.F. Verdejo, (1992) “SEISD: An Enviroment for Extraction of Semantic Information from On-line Dictionaries” in *Proceedings of the 3rd Conference on Applied Natural Language Processing*.

Alonge, A. (1991) “Extraction of information on Aktionsart from verb definitions in machine-readable dictionaries” in 11th International Workshop on Expert Systems and their Applications, Avignon.

Baker, M. (1988) Incorporation: A theory of Grammatical Function Changing, University of Chicago Press, Chicago.

Castellón, I. (1992), Lexicografía Computacional: Adquisición de Conocimiento Léxico, Tesis Doctoral, Barcelona.

Copestake, A. (1992) “The ACQUILEX LKB: Representation Issues in Semi-Automatic Acquisition of Large Lexicons”. *Proceedings of the 3rd Conference on Applied Natural Language Processing*.

Dowty, D. (1986), “The effects of Aspectual Classes on the temporal structure of discourse: Semantics or Pragmatics?” in Linguistics and Philosophy, 9.

\_\_\_\_\_ (1988), “Thematic Proto-Roles, Subject Selection and Lexical Semantic Defaults. LSA Colloquium paper.

- \_\_\_\_\_ (1989), "On the Semantic content of the notion of 'Thematic Role'" in G. Chierchia, B.H. Partee and R. Turner, eds. Properties, Types and Meaning, Kluwer, Dordrecht.
- Grimshaw, J. (1992), Argument Structure, MIT Press, Cambridge, MA.
- Hale, K. & Keyser, S.J. (1993), The view from Building 20: Essays in Linguistics in honor of Sylvain Bromberger, MIT Press, Cambridge, MA.
- Jackendoff, R. (1983) Semantic and Cognition, MIT Press, Cambridge, MA.
- Jackendoff, R. (1990) Semantic Structures, MIT Press, Cambridge, MA.
- Levin, B. (1989) Towards a Lexical Organization of English Verbs, Lexicon Project Working Papers, Center for Cognitive Science, MIT, Cambridge, MA.
- \_\_\_\_\_ (1993), English Verb Classes and Alternations, University of Chicago Press, Chicago.
- Pollard, C. & Sag, I. (1987) An Information-Based Approach to Syntax and Semantics: Fundamentals, Vol. I. CSLI, Stanford.
- \_\_\_\_\_ (1992) Agreement, Binding and Control: Information-Based Syntax and Semantics, Vol.2. CSLI, Stanford.
- Pustejovsky, J. (1991), "Generative Lexicon" in *Computational Linguistics*, 17.
- Rappaport, M. & Levin, B. (1988), "What to Do with Theta-Roles?" in Wilkins, ed. (1988a).
- Sanfilippo, A. (1991a) "LKB Encoding of Lexical Knowledge". Briscoe, T., Copestake, A., & de Paiva, V. (eds.).
- \_\_\_\_\_ (1991b) "Thematic and Aspectual Information in Verb Semantics" in Perspectives on Aspect and Aktionsart, ed. C. Vetters and W. Vandeweghe, Belgian Journal of Linguistics, 6.
- \_\_\_\_\_ (1992) "Verbal Diatheses: Knowledge Acquisition, Lexicon Construction and Dictionary Compilation" (forthcoming).
- Talmy, L., (1985) "Lexicalization patterns" in Shopen, T. (ed) Language Typology and Syntactic Description vol. 3 (Grammatical Categories and the Lexicon) CUP.
- Taulé, M, Castellón, I. (1993) "Generación de Alternancias de Subcategorización via Reglas Léxicas" Actas del IX Congreso anual de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN) Santiago de Compostela, Septiembre.
- Taulé, M., Martí, M.A. & Castellón, I. (1994), "Subcategorization Alternances Generation via Lexical Rules", ESPRIT BRA-7315 ACQUILEX-II WP n° ??, Barcelona.
- Taulé, M. (1994) Representación Léxica de los verbos en una Base de Conocimiento, Tesis Doctoral, Barcelona, (in prep.).
- Vendler, Z. (1967) Linguistics in Philosophy, Cornell University Press, Ithaca.

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)