WordNet as a base Lexicon Model for the Computation of Verbal Predicates

Raquel Amaro
Group for the Computation of Lexical and Grammatical Knowledge
Center of Linguistics of the University of Lisbon
ramaro@clul.ul.pt

Abstract

Assuming that the lexicon is a complex and dynamic knowledge system, the lexicon model and the information that is stated in it become crucial for the computation of meaning. In this paper we discuss how the WordNet model can support a decompositional approach to troponymy, the enrichment of the lexical entries with Qualia information, and the establishment of a lexical inheritance device, without much additional work. More specifically, we intend to show how the choice of the model and its further enrichment is motivated by the reflex of the semantic content on the syntactic behaviour of the lexical items, namely through co-occurrence restrictions, and address issues on argument structure, Aktionsart shifts, co-troponym's incompatibilities and shadow arguments' realization. We present a level of inheritance structure in the semantic representation of the lexical items and propose a semantic representation for prepositional arguments, using the Generative Lexicon notation. We also address the ability of the WN model to be used as a semantic type hierarchy.

1 Introduction

The lexicon can be seen as a complex and dynamic knowledge system that determines structure and meaning. In this perspective, the lexicon model and the information that is stated in it are crucial for the computation of meaning.

The organization of the lexical entries within the lexicon and the type of information that is introduced in the lexical entries should also be motivated by processing issues, namely restrictions on the (co-)occurrence of the lexical items. For instance, the model for the verbal lexicon must necessarily deal with the predicates' valence, reflected in restrictions on argument selection: intransitive verbs are expected to form a group, transitive verbs another group, ditransitive verbs yet another group, and so on. More still, it is expected that the verbs in these groups have the same resctrictions, reflected in the syntax. However, the analysis of Portuguese verbs within WordNet.PT[1], focusing mainly on verbs of movement, allowed us to notice that verbs conceptually and semantically connected – in a hypernymy / troponymy relation – have different selection restrictions, and that the predicates' valence is not enough to determine groups:

(1) a. *tirar* (take) is *mover* (move) from a given location
tirar is TROPONYM of *mover*

b. Ele moveu o caixote.
'he moved the box'

c. Ele tirou o caixote da rua.
'he took the box from-the street'

The examples in (1) show that the transitive verb *mover* and the ditransitive verb *tirar* are part of the same semantic domain (in (1a)), although displaying a different number of arguments (cf. (1b) and (1c)).

The Aktionsart properties of the verbs are also desirably mirrored in the organization of the lexicon, being state denoting verbs or activity denoting verbs naturally close to each other. Nonetheless, there are cases such as the ones in (2), where the troponym verb does not share the Aktionsart values of its hypernym: the verb *voltar* (to return) is an accomplishment denoting verb, hyponym of the activity denoting verb *mover-se* (move oneself). These Aktionsart differences reflect on the co-occurrence restrictions with certain adverbials. Accomplishment denoting verbs can occur with *in*-adverbials but

This research was supported by Fundação para a Ciência e a Tecnologia (grant SFRH/BD/13875/2003).

[1] See Marrafa (2002).

cannot occur with *for*-adverbials (see (2c)), whereas activity denoting verbs can occur with *for*-adverbials but not with *in*-adverbials (see (2b)).

(2) a. *voltar* (return) is *mover-se* (move oneself) back to the start location.
voltar is TROPONYM of *mover-se*

b. Ele moveu-se *em meia hora.
'he moved *in half an hour'
Ele moveu-se durante meia hora.
'he moved for half an hour'

c. Ele voltou em meia hora.
'he returned in half an hour'
Ele voltou *durante meia hora.
'he returned *for half an hour'

Also, it was possible to observe that co-troponym verbs – verbs that have the same direct hypernym – are not necessarily incompatible, i.e. co-troponym co-occurrence is not always ruled out:

(3) a. Ele subiu a rua correndo.
'he moved up the street running'

b. *Ele subiu a rua descendo.
'he moved up the street moving down'

c. Ele desceu a rua correndo.
'he moved down the street running'

The verbs *subir* (move up), *descer* (move down) and *correr* (run) are all troponyms of the verb *mover-se* (move oneself), but are not uniformely incompatible. The verbs *subir* and *descer* can co-occur with the verb *correr* ((3a) and (3c)) but cannot co-occur with each other (see (3b)).

In this paper we propose to account for these data by extending the information in the lexical entries in a WordNet (WN) model[2] lexicon with Generative Lexicon[3] (GL) structures.

Within a strict WN context – and to determine the base relation of troponymy – it is not necessary to consider or distinguish between different types of manner (Fellbaum, 1998):

(4) *To V1 is to V2 in some particular manner.* (Fellbaum, 1998, 79)

[2]The extension of the information in lexical entries, particularly of verbs, is the focus of several works, namely the recent Koenig et al. (2002), Agirre and Martinez (2002), Dang et al. (2000), among others.

[3]See Pustejovsky (1995).

However, these different *particular manners* can be reflected on the argument structure of verbs (number and / or type of arguments selected, as exemplified in (1)), motivating a decompositional approach to troponymy.

The hierarchichal organization of a relational net seems also to be able to account for phenomena such as *lexical shadowing*, namely the restrictions on the syntactic realization of *shadow arguments*, as exemplified bellow:

(5) a. *Ele serrou a árvore com a serra.
'he sawed the tree with the saw'

b. Ele serrou a árvore com a serra de ferro.
'he sawed the tree with the iron saw'

c. Ele serrou a árvore com a motosserra.
'he sawed the tree with the chainsaw'

The construction of a relational lexicon already points out most of the crucial similarities and differences that are established in a decompositional analysis, being the enrichment of the WN model not highly costly. Moreover, the ontological nature of the WN model, besides accounting for already known phenomena such as *lexical shadowing*, seems to allow for the semantic representation of lexical items without recurring to a hierarchy of semantic types.

Section 2 describes how troponymy reflects the predicate's meaning and structure. Section 3 focuses on the decompositional analysis of troponymy, namely on how it is reflected on argument structure and Aktionsart properties. Section 4 discusses the relevance of introducing Qualia information in the lexical entries, considering verbal Qualia unification and co-troponym compatibility. Section 5 is dedicated to the semantic representation of the verbal predicates, specifically in what concerns lexical inheritance and the semantic content of prepositional arguments. Section 6 refers how the WN model can be used as a semantic type hierarchy, and how the organization of this model can account for the restrictions on phonetic realization of shadow arguments. Finally, Section 5 presents our conclusions.

2 Troponymy and predicate's meaning

It is well known that the WN model shares some aspects with decompositional models. For instance, the correspondence between Lexical

Conceptual Structures[4] (LCS) primitives (GO in (6a)) and verbs that are top nodes in the hierarchy (*mover-se* in (6b)) is quite transparent:

(6) a. LCS for *andar* (walk)
V
_ NP_j
$[_{Event}$ GO $[_{Thing}\ ]_j\ ]$

b. *andar* (walk) is ***mover-se*** (move oneself), slowly, resting one foot after the other on a surface
andar is TROPONYM of *mover-se*

In this way, the hypernym relation already determines the core meaning of the troponym, playing a role similar to that of semantic primitives in the event structure representation, as informally exemplified in (7).

(7) andar(e_1, x, z):
e_1 = move_oneself,
x = animate_entity,
z = slowly, resting one foot after the other on a surface

The troponymy relation definition, stated in (4), allows us to explicitly "translate" the meaning specificity denoted by the troponym into semantic components, since the core meaning is already identified.

3 Decompositional analysis of troponymy

In Amaro (2005), a new set of semantic elements, using the notion of incorporation as defined in Talmy (1985, 60), is proposed in order to describe the meaning specificities that distinguish hypernyms from troponyms in a relational net. The resulting set is presented in (8):

(8) a. MANNER, how the event develops:
sussurrar (to whisper) = speak **softly**.

b. INTENTION, purpose/ intended goal of the event:
sacrificar (to sacrifice) = kill **to please/ honor divine entities**.

c. FIGURE, object that undergoes the event:
anestesiar (to anesthetize) = give **anesthesia**.

d. GROUND, object with respect to which the FIGURE undergoes the event:
aumentar (to increase) = make bigger **with respect to a previous dimension**.

e. CAUSE, what brings about the event:
sufocar (to suffocate) = die **due to lack of access to air or oxygen**.

f. SOURCE, FIGURE initial location:
remover (to take from place) = to take **from the usual place**.

g. GOAL, FIGURE final location:
encaixotar (to box) = to put **in a box**.

h. PATH, medium location(s) between the SOURCE and the GOAL:
retroceder (to backtrack) = to move back **through the same path**.

i. DIRECTION, way in which the motion event occurs:
descer(to move down) = to move (in which direction?) **down**.

This decompositional analysis identifies what is specific of each troponym. The consequences are more or less visible, i.e. the lexicalization of the semantic elements affects the predicate's inherited content in different degrees, mainly in what concerns argument structure, but also regarding Aktionsart properties.

3.1 Argument structure

The argument structure is the structure for the representation of logical arguments, independently of their exact syntactical realization. Here, we assume the types of arguments defined in Pustejovsky (1995, 63-64):

(9) a. True arguments (ARG$_n$) are necessarily realized syntactically.

b. Default arguments (D-ARG$_n$) pertain to the semantic content of a lexical item, but are not necessarily realized syntactically.

c. Shadow arguments (S-ARG$_n$) are semantically incorporated in the meaning of the lexical item, and can only be realized syntactically through discourse specification devices.

According to these types of arguments, the incorporation of GOAL in the synset {*pôr, colocar*} (put), represented *a la* GL in (11), is re-

[4] See Jackendoff (1983).

flected by an increase of the list of true arguments inherited from the hypernym, {*mover, deslocar*} (move, change location), in (10). The expression of the final location, the GOAL – the ARG_3 introduced by a preposition – becomes obligatory.

(10) {*mover, deslocar*}(move):

$$\begin{bmatrix} \dots \\ \text{ARG-STR} = \begin{bmatrix} \text{ARG}_1 = \boxed{1}\ x\text{:}\mathbf{entity} \\ \text{ARG}_2 = \boxed{2}\ y\text{:}\mathbf{entity} \end{bmatrix} \end{bmatrix}$$

(11) {*pôr, colocar*}(put):

$$\begin{bmatrix} \dots \\ \text{ARG-STR} = \begin{bmatrix} \text{ARG}_1 = \boxed{1} \\ \text{ARG}_2 = \boxed{2} \\ \text{ARG}_3 = \boxed{3}\ z\text{:}\mathbf{goal} \end{bmatrix} \end{bmatrix}$$

Again, this argument structure is inherited by the immediate troponym, {*meter*} (put inside of), in (12), that expresses a more specific DIRECTION of the event, replacing the preposition[5]. The lowest troponym, {*enjaular*}, incorporates a specific GOAL location, a cage, through lexical shadowing, thus changing the predicate type of argument from true argument to shadow argument.

(12) {*meter*}(put inside of):

$$\begin{bmatrix} \dots \\ \text{ARG-STR} = \begin{bmatrix} \text{ARG}_1 = \boxed{1} \\ \text{ARG}_2 = \boxed{2} \\ \text{ARG}_3 = \boxed{3} \end{bmatrix} \end{bmatrix}$$

(13) {*enjaular*}(cage):

$$\begin{bmatrix} \dots \\ \text{ARG-STR} = \begin{bmatrix} \text{ARG}_1 = \boxed{1} \\ \text{ARG}_2 = \boxed{2} \\ \text{S-ARG}_3 = z\text{:}\mathbf{cage} \end{bmatrix} \end{bmatrix}$$

In this fashion it is possible to account for different argument structures within a troponymy tree, without artificially isolating lexical items conceptually related for not sharing the same argument structure. Also, it becomes possible to characterize lexical inheritance using the lexical semantic structures available through the hypernymy relation.

3.2 Aktionsart shifting

The analysis of semantic incorporation also allows us to account for Aktionsart differences between hypernyms and troponyms.

We use the four Aktionsart classes defined in Vendler (1967), following Móia (2000)'s set of criteria:

(14) **state**: non-punctual (i.e. temporally unbounded), homogeneous, and simple (non-complex) event.

activity: non-punctual, relatively homogeneous (i.e. homogeneous considering subintervals), and simple event.

accomplishment: non-punctual, heterogeneous, and complex event (preparatory process, culmination point and consequent state).

achievement: punctual, heterogeneous, and complex event (culmination point and consequent state).

These ontological properties are reflected (and therefore testable) in terms of distributional properties. As follows,

(15) **state denoting predicates** can not occur in the present continuous tense form (*is V-ing*)

**John is being tall*;

activity denoting predicates occur in the present continuous tense form (*is V-ing*), entailing that *if X is (now) V-ing than X has already V*

John is running → John has already ran,

occur with *for*-adverbials, but not with *in*-adverbials

*John run for 10 minutes /*in 10 minutes*;

accomplishment denoting predicates occur in the present continuous tense form (*is V-ing*), entailing that *if X is (now) V-ing than X has not yet already V*

John is building the house → John has not yet built the house,

[5] Prepositions, although not the focus of the current paper, will eventually be treated within the same framework, in order to account for these data. Since the prepositions at stake have clear semantic content, i.e., they are more than theta markers, it is possible to assume a relational net in which the concept denoted by *dentro de* (inside of) is in a hyponymy-type relation with the concept denoted by *em* (in) – if x *is inside of* z then x is *in* z. In this way, it would not be necessary to consider a preposition replacement in (12), but rather a specification through a hyponym.

occur with *in*-adverbials, but not with *for*-adverbials

*John built the house in 10 months /*for 10 months*;

achievement denoting predicates can not occur with *in*-adverbials, nor with *for*-adverbials

*John died *in 10 months /*for 10 months.*

The cases observed within the class of verbs of movement involve the incorporation of SOURCE, GOAL, and MANNER.

Intuitively, the incorporation of SOURCE and GOAL establishes a limit to the event, shifting an activity type to an accomplishment type event, in (16) and (17), whereas the incorporation of MANNER, in (18), may be described as the addition of a process to an achievement type event[6].

(16) a. {*tirar*$_{accomplish.}$} (take) [move$_{activity}$ from source]

b. Ele tirou a caixa da rua
'he took the box from-the street
em 10 min. / *durante 10 min.
in 10 min. / *for 10 min.'

c. Ele moveu/deslocou a caixa
'he moved the box
*em 10 min. / durante 10 min.
*in 10 min. / for 10 min.'

(17) a. {*voltar, regressar*$_{accomplish.}$} (come back / return) [move oneself$_{activity}$ to the start point]

b. Ele voltou / regressou
'he came back / returned
em 10 min. / *durante 10 min.
in 10 min. / *for 10 min.'

c. Ele moveu-se *em 10 min. /
'he moved *in 10 min. /
durante 10 min.
for 10 min.'

(18) a. {*esgueirar-se*$_{accomplish.}$} (sneak out) [move-out-of$_{activity}$ furtively]

b. Ele esgueirou-se da casa
'he sneaked out of-the house
em 10 min. / *durante 10 min.
in 10 min. / *for 10 min.'

c. Ele saiu da casa
'he moved out of-the house
*durante 10 min. / *em 10 min.
*for 10 min. / *in 10 min.'

The regularity exhibited in the cases of semantic incorporation of SOURCE, GOAL and MANNER and Aktionsart shifting, as well as the argument structures' differences presented, provide strong motivation for some level of decompositional analysis in the representation of the lexicon.

Moreover, it is clear that this level of decompositional analysis is, for the most part, recoverable from a relational lexicon construction process.

4 Qualia information

The association of Qualia information (Pustejovsky, 1995) to other lexicon models is nowadays a commonly accepted strategy to enhance the lexicon model expressiveness, namely in what concerns the construction of lexica for computational purposes.[7]

According to Mendes and Chaves (2001), the association of Qualia information to synsets enables the model to account for issues such as co-hyponyms co-occurrence. By recognizing the different aspects of the meaning specification involved in the hypernymy/hyponymy relation through the Qualia values of each synset, it is possible to predict the compatibility in (19a) as well as the incompatibility in (19b):

(19) a. This dog is a pitbull$_{[FORMAL]}$ and a police-dog $_{[TELIC]}$.

b. *This dog is a pitbull$_{[FORMAL]}$ and a Saint-Bernard$_{[FORMAL]}$.

This contrast refutes the assumption that all co-hyponyms, as specifications of the same concept, are incompatible, since the data show that this incompatibility depends on the type of specific information each co-hyponym denotes.

4.1 Qualia unification

To account for this phenomenon, Mendes and Chaves (2001) assume a unification operation. Incompatibility is expressed considering that co-hyponyms are incompatible if their Qualia structures do not unify:

[6] These Aktionsart shifts are predicted in Moens (1987, 45).

[7] See, for instance, Fabre and Sébillot (1999), Busa et al. (2001), De Boni and Manandhar (2002), O'Hara and Wiebe (2003), Veale (2003), among others.

(20) Two qualia structures do not unify if there is a quale Q from two nominal qualia structures [Q=R_1] and [Q=R_2]
where roles R_1 and R_2 exist such that
$\neg(R_1{=}R_2) \wedge \neg$ subsumes(R_1,R_2).

The second conjunct in the last restriction ($\neg$ subsumes(R_1,R_2)) assures that cases of two co-hyponyms with different values for the same Qualia role are compatible, since one of these values is a subtype of the other, such as the ones in (21):

(21) a. This animal is a predator and a mammal.

b. {predator}:
$$Q=\begin{bmatrix}\text{FORMAL}=\textbf{animal}\\ \text{TELIC}=\textbf{hunting}\end{bmatrix}$$

c. {mammal}:
$$Q=\left[\text{FORMAL}=\textbf{mammal}\right]$$

d. {mammal} @-> {animal}.

Here the values of R_1 and of R_2, *animal* and *mammal* of the FORMAL quale are not equal, but R_1 is a subtype R_2, accounting for the compatibility of these co-hyponyms.

According to the decompositional analysis presented in section 2, it is possible to consider a version of this operation referring the semantic elements incorporated in the verbs. Thus, two co-troponym verbs are incompatible if some value of the semantic elements incorporated in their semantic structures do not unify:

(22) a. Ele saiu da casa, correndo
'he exited the house, running'
{sair}$_{\text{DIRECTION:}out}$, {correr}$_{\text{MANNER:...}}$

b. *Ele saiu da casa, entrando
'he exited the house, entering'
{sair}$_{\text{DIRECTION:}out}$,
{entrar}$_{\text{DIRECTION:}in}$

At a first glance, this hypothesis seems to cover the same kind of co-occurrence restrictions on verbs co-hyponyms. However, cases such as (23b) are not accounted for, since the values of the element incorporated refer to DIRECTION but are, all the same, compatible.

(23) a. Ele saiu da casa, recuando.
'he exited the house, moving back'
{sair}(exit)$_{\text{SOURCE}}$,
{recuar}(move back)$_{\text{DIRECTION}}$

b. Ele subiu a rua,
'he moved up the street
afastando-se da casa.
moving away from the house'
{subir}(move up)$_{\text{DIRECTION:}up}$,
{afastar-se}(move away)$_{\text{DIRECTION:}away}$

Note that, as stated before, the semantic incorporation is reflected in the argument structure of the verbs, and not directly in the Qualia role's values. Verbal Qualia structure's values are somewhat different from those of nouns and the verbal semantic structure seems to introduce several difficulties.

4.2 Verbal Qualia structure

The first issues to consider relate to the quale roles values licensed for verbs. Verbal Qualia structure is fulfilled with semantic predicates that establish the relations between the arguments of a verb.

Typically, only two Qualia roles are used in this description: the FORMAL role, for activity and state type events; and the AGENTIVE (bearing the causal chain) and FORMAL (stating the final state) roles for accomplishment and achievement type events[8], as exemplified in (24):

(24) a. eat$_{activity}$:
$$Q=\left[\text{FORMAL}=\textbf{eat\_act}(\mathbf{e}_1,\mathbf{x},\mathbf{y})\right]$$

b. kill$_{accomplishment}$:
$$Q=\begin{bmatrix}\text{FORMAL}=\textbf{kill\_result}(\mathbf{e}_2,\mathbf{y})\\ \text{AGENTIVE}=\textbf{kill\_act}(\mathbf{e}_1,\mathbf{x},\mathbf{y})\end{bmatrix}$$

In this way, we have the following representations of the verbs in question:

(25) a. correr (run):
$$\begin{bmatrix}\text{ARG-STR}=\begin{bmatrix}\text{ARG}_1=\textbf{x:ind}\\ \text{S-ARG}_1=\textbf{z:manner}\end{bmatrix}\\ Q=\left[\text{FORMAL}=\textbf{run\_act}(\mathbf{e}_1,\mathbf{x},\mathbf{z})\right]\end{bmatrix}$$

[8] See Pustejovsky (1995, 183-188). Note that, however, some authors also use other Qualia roles, namely the Telic, as in Marrafa and Moura (2005).

b. sair (exit, move out):

$$\begin{bmatrix} \text{ARG-STR}=\begin{bmatrix} \text{ARG}_1=\textbf{x:ind} \\ \text{ARG}_2=\textbf{y:loc} \\ \text{S-ARG}_1=\textbf{z:out} \end{bmatrix} \\ \text{Q}=\begin{bmatrix} \text{FORMAL}=\neg\textbf{at}(\textbf{e}_2,\textbf{x},\textbf{y}) \\ \text{AG.}=\textbf{exit\_act}(\textbf{e}_1,\textbf{x},\textbf{z}) \end{bmatrix} \end{bmatrix}$$

c. entrar (enter, move in):

$$\begin{bmatrix} \text{ARG-STR}=\begin{bmatrix} \text{ARG}_1=\textbf{x:ind} \\ \text{ARG}_2=\textbf{y:loc} \\ \text{S-ARG}_1=\textbf{z:in} \end{bmatrix} \\ \text{Q}=\begin{bmatrix} \text{FORMAL}=\textbf{at}(\textbf{e}_2,\textbf{x},\textbf{y}) \\ \text{AG.}=\textbf{enter\_act}(\textbf{e}_1,\textbf{x},\textbf{z}) \end{bmatrix} \end{bmatrix}$$

Although the incompatibility of the co-troponyms *sair* (exit) and *entrar* (enter), in (25b) and (25c) is covered, since these verbs have comparable values in their Qualia roles, to assume the unification operation as directly applying to verb Qualia structures is to predict that *correr* (run) is incompatible with *sair* (exit) or *entrar* (enter) since the values of their FORMAL role are different and none subsumes the other – the FORMAL quale role value is actually contrary, since the verb *sair* (exit) incorporates SOURCE, denoting a final state of *not being at a given location*, ($\neg$at(e_2,x,y)), whereas the verb *entrar* (enter) incorporates GOAL, denoting a final state of *being at a given location*, (at(e_2,x,y)). On the other hand, to allow for unification disregarding the Qualia roles' values is to ignore evidence relating the Qualia roles' values with the different meaning specifications among co-troponyms. This seems to show that Qualia structure unification is not feasible for the verbal lexicon.

4.3 Indirect Qualia unification

The data presented so far regarding compatibility issues among co-troponyms, and the fact that it is possible to relate this with the semantic elements incorporated in the troponyms, constitute evidence that Qualia structure unification is occurring at some level.

Considering that it is in the ARGUMENT STRUCTURE that the logical arguments of a predicate are listed, and that verbs' ARGUMENT STRUCTURE necessarily reflect the incorporation of the semantic elements responsible for the meaning specificities between hypernyms and troponyms (see section 2), we propose to account for co-troponym compatibility via indirect Qualia unification:

(26) Two co-troponym verbs are incompatible iff the arguments in their argument structures refer to incompatible co-hyponyms, i.e. if the Qualia structures of these arguments do not unify.

Let us reconsider the verbs in (19), here in (27):

(27) a. *sair* (exit):

$$\text{ARG-STR}=\begin{bmatrix} \text{ARG}_1 = \textbf{x: ind} \\ \text{ARG}_2 = \textbf{y: location} \\ \text{S-ARG}_1 = \textbf{z: out} \end{bmatrix}$$

b. *entrar* (enter):

$$\text{ARG-STR}=\begin{bmatrix} \text{ARG}_1 = \textbf{x: ind} \\ \text{ARG}_2 = \textbf{y:location} \\ \text{S-ARG}_1 = \textbf{z:in} \end{bmatrix}$$

c. *correr* (run):

$$\text{ARG-STR}=\begin{bmatrix} \text{ARG}_1 = \textbf{x: ind} \\ \text{S-ARG}_1 = \textbf{z:manner} \end{bmatrix}$$

Indirect Qualia structure unification enables us to predict that the co-troponym verbs *sair* (exit) and *entrar* (enter) are incompatible since the Qualia structure of *out* and *in* - which are co-hyponyms - do not unify. Conversely, co-troponym verbs *sair* (exit) and *correr* (run) are compatible since there is no co-hyponym arguments in their structure.

To consider Qualia information in a relational model of the lexicon proved once again of relevance. Also, the determination of the values for the Qualia roles does not require any additional work, since it only involves the statement of the meaning specifications that lead us to consider a troponymy relation between two items.

5 Semantic representation

5.1 Lexical Inheritance

In the previous section we focused on how to use Qualia information to determine which co-troponyms are compatible and which are not. And yet, we have not considered the hierarchical nature of the WN model and it's major relation – the hypernymy relation.

The compatibility issues arise among co-troponym verbs, i.e. verbs that have the same hypernym. In fact, co-troponyms are predicted to be incompatible since they can denote contradictory specifications of a shared core meaning

– that of their hypernym. In other terms, it is possible to consider that the core meaning of a troponym verb is inherited from its hypernym.

Following the four-level semantic representation presented in the GL, we propose to consider an INHERITANCE STRUCTURE (corresponding roughly to the LEXICAL INHERITANCE STRUCTURE of Pustejovsky (1995)). The hypernym lexical items are referred to in this structure, their semantic content becoming available, and being only necessary to add the specific information that concerns the meaning specifications of the troponym. In this way, we can represent the verb *andar* (walk), informally stated in (7), or the verb *sair* (exit), as follows:

(28) a. andar(walk)

$$\left[\begin{array}{l}\text{INH-ST}=\left[\text{HYP.}=\left[\begin{array}{l}\textbf{\textit{mover}}\\ \text{EV-ST}=\left[\text{E}_1=\boxed{1}\ \textbf{e}_1\textbf{:act.}\right]\\ \text{ARG-ST}=\left[\text{ARG}_1=\boxed{2}\ \textbf{x:ind}\right]\\ \text{Q}=\left[\begin{array}{l}\text{FORMAL}=\\ \textbf{move\_act(e}_1\textbf{,x)}\end{array}\right]\end{array}\right]\right]\\ \text{EV-ST}=\left[\text{E}_1=\boxed{1}\right]\\ \text{ARG-ST}=\left[\begin{array}{l}\text{ARG}_1=\boxed{2}\\ \text{S-ARG}_1=\textbf{z:manner}\end{array}\right]\\ \text{Q}=\left[\text{FORMAL}=\textbf{walk\_act(}\boxed{1}\textbf{, }\boxed{2}\textbf{, z)}\right]\end{array}\right]$$

b. sair(exit)

$$\left[\begin{array}{l}\text{INH-ST}=\left[\text{HYP.}=\left[\begin{array}{l}\textbf{\textit{mover}}\\ \text{EV-ST}=\left[\text{E}_1=\boxed{1}\ \textbf{e}_1\textbf{:act.}\right]\\ \text{ARG-ST}=\left[\text{ARG}_1=\boxed{2}\ \textbf{x:ind}\right]\\ \text{Q}=\left[\begin{array}{l}\text{FORMAL}=\\ \textbf{move\_act(e}_1\textbf{,x)}\end{array}\right]\end{array}\right]\right]\\ \text{EV-ST}=\left[\begin{array}{l}\text{E}_1=\boxed{1}\\ \text{E}_2=\textbf{e}_2\textbf{:state}\end{array}\right]\\ \text{ARG-ST}=\left[\begin{array}{l}\text{ARG}_1=\boxed{2}\\ \text{ARG}_2=\textbf{y:location}\\ \text{S-ARG}_1=\textbf{z:out}\end{array}\right]\\ \text{Q}=\left[\begin{array}{l}\text{FORMAL}=\neg\textbf{at(e}_2\textbf{,}\boxed{2}\textbf{, y)}\\ \text{AG.}=\textbf{exit\_act(}\boxed{1}\textbf{, }\boxed{2}\textbf{, z)}\end{array}\right]\end{array}\right]$$

This level of representation – the INHERITANCE-STRUCTURE – enables us to reformulate the indirect Qualia unification clause, since inheritance prevents the shared arguments, recovered from the hypernym lexical semantic structure, from being checked for incompatibility:

(29) Two co-troponym verbs are incompatible iff the **non-inherited arguments** in their argument structure refer to incompatible co-hyponyms.

5.2 Prepositional arguments

It is clear in Pustejovsky (1995) that the type of arguments (see section 3.1) in the argument structure also accounts for some level of syntactic mapping. Moreover, arguments are listed from the less oblique one (corresponding, roughly, to the subject position) to the more oblique one (corresponding, most of the times, to arguments expressed by prepositional phrases) (Pustejovsky, 1995, 62-67). However, the information of which prepositional argument is selected by which verb is not visible in the GL lexical entries.

Observe, for instance, the sentence in (30):

(30) Ele tirou o livro da mesa.
'he took the book from-the table'

The true arguments in the argument structure of this verb (ARG-STR=[ARG_1= x:ind, ARG_2= y:object, ARG_3= z:loc]) do not reveal that the ARG_3 is of the prepositional phrase form, necessarily headed by the preposition *de* (from), SOURCE marker. More still, it should be possible to associate this prepositional phrase – also present in the argument structure of the verb *sair* in (28b)– to the resulting state e_2:

(31) tirar (take)

$$\left[\begin{array}{l}\ldots\\ \text{Q}=\left[\begin{array}{l}\text{FORMAL}=\neg\textbf{at(e}_2\textbf{, y, z)}\\ \text{AG.}=\textbf{take\_act(e}_1\textbf{, x, y, w)}\end{array}\right]\end{array}\right]$$

It can be assumed that in the cases of the verbs *sair* (exit) and *tirar* (take) the semantic content attributed to the SOURCE denoting preposition *de* is not individuated since it is part of the semantic content of the verb itself. However, although consisting of true arguments in the previous cases, these prepositional phrases can occur as adjuncts with other verbs – namely within a directional construction – maintaining its meaning:

(32) Ele empurrou a cadeira daqui.
'he pushed the chair from-here'

Following the approach in Verspoor (1997), we assume that the semantic contribution of prepositional phrases is consistent across uses, regardless of their status as complements or adjuncts[9], and propose the following lexical entry for the preposition *de* (from):

(33) de (from)

$$\begin{bmatrix} \text{EV-ST}=\left[\text{E}_1=\mathbf{e_1:state}\right] \\ \text{ARG-ST}=\begin{bmatrix}\text{ARG}_1=\mathbf{x:obj}\\ \text{ARG}_2=\mathbf{y:loc}\end{bmatrix} \\ \text{Q}=\left[\text{FORMAL}=\neg\mathbf{at(e_1,\ x,\ y)}\right] \end{bmatrix}$$

The association of an event to the semantic content of a preposition is proposed in Pustejovsky (1995, 126), for the representation of the PP *into the cave*. The author considers that the semantic content of the preposition *into* includes, besides a state, a motion event that causes that state. Here we feel that this association is not motivated enough and that it is awkward to propose a transition representation for a preposition. Note, however, that the consequences of this view on the treatment of directional constructions are yet to be studied.

With this representation, it is possible to account for the semantic contribution of the prepositional phrase in sentences such as the one in (32), as well as in the semantic content of verbs as *sair* (exit) or *tirar* (take), stating at the lexical level the exact prepositional complement selected by the verb:

(34) sair(exit)

$$\begin{bmatrix} \text{INH-ST}=\begin{bmatrix}\text{HYP.}=\begin{bmatrix}\textit{mover}\\ \text{EV-ST}=\left[\text{E}_1=\boxed{1}\ \mathbf{e_1:act.}\right]\\ \text{ARG-ST}=\left[\text{ARG}_1=\boxed{2}\ \mathbf{x:ind}\right]\\ \text{Q}=\begin{bmatrix}\text{FORMAL}=\\ \mathbf{move\_act(e_1,x)}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{EV-ST}=\begin{bmatrix}\text{E}_1=\boxed{1}\\ \text{E}_2=\boxed{3}\end{bmatrix} \\ \text{ARG-ST}=\begin{bmatrix}\text{ARG}_1=\boxed{2} \\ \text{ARG}_2=\begin{bmatrix}\textit{de}\\ \text{EV-ST}=\left[\text{E}_1=\boxed{3}\ \mathbf{e_1:state}\right]\\ \text{ARG-ST}=\begin{bmatrix}\text{ARG}_1=\boxed{2}\\ \text{ARG}_2=\mathbf{y:loc}\end{bmatrix}\\ \text{Q}=\begin{bmatrix}\text{FORMAL}=\\ \boxed{4}\neg\mathbf{at(e_1,\ \boxed{2},\ y)}\end{bmatrix}\end{bmatrix} \\ \text{S-ARG}_1=\mathbf{z:out}\end{bmatrix} \\ \text{Q}=\begin{bmatrix}\text{FORMAL}=\boxed{4}\\ \text{AG.}=\mathbf{exit\_act(\boxed{1},\ \boxed{2},\ z)}\end{bmatrix}\end{bmatrix}$$

6 WN and semantic types' hierarchies

6.1 Semantic types

One of the interesting consequences of considering a relational model of the lexicon together with generative representations is, perhaps, the possibility of not having to build a semantic type hierarchy to describe the lexical items semantics, since it is possible to use the available lexical semantic structures in the hierarchy.

Our proposal is to consider the nodes in the lexicon as the semantic types themselves. In fact, as a model of the mental lexicon, the WN can be seen as a reflex of the relation between lexical knowledge and our ability to organize and classify the concepts in the world (cf. Pustejovsky (2001)).

As a consequence we would have that, when referring to an argument of type `human`, we would access not the type and the relations established between this type and the other types in the hierarchy of types, but the synset {*human*} and the relations established between this node and all the other nodes in the lexicon. Phenomena such as regular polysemy for instance, covered by complex types such as `information•physical_object`, are al-

[9] The analysis in Verspoor (1997) concerns mainly dative prepositional phrases, divided in three types of prepositional datives: complement PPs, pseudo-complement PPs, and adjunct PPs. According to this account, the prepositional phrases considered in this paper fall under the definitions of complement PPs or pseudo-complement PPs. Note, however, that the semantic content of all the types of PPs is the same, being the type of prepositional modification determined via lexical rules. If this proposal accounts for the constructions in which these particular PPs enter is an issue for further development.

ready accounted for by considering that synsets like {*book*} have two hypernyms, {*text*} and {*object*}.

Also, the hypernym relation enables lexical inheritance through the use of available lexical structures. All that is required is that the semantic specification of the arguments and the values of the Qualia roles on a lexical entry do not refer to types pertaining to an extra-lexicon hierarchy, but to other lexical items. Note that lexical gaps – concepts that are not lexicalized in a given language – might raise important issues to this proposal. Yet, it seems reasonable to consider that the ontological and hierarchical nature of a hand-build relational lexicon should be enough to ensure the conceptual categorization needed to describe the lexical items meaning. Nonetheless, we are naturally working to test further this hypothesis.

6.2 Shadow arguments

The ontological nature of a WN seems also to shed light on lexical shadowing phenomena. Fellbaum (1998, 81) states that the variety of nouns that a verb can select as arguments is contrary to the level in which it is embedded in the hierarchy, i.e., as the meaning of verbs gets more and more specific, their arguments are also selected in lower levels of the hierarchy. Let us consider the following hierarchy fragment:

(35) {*serrar*}(saw) @-> {*cortar*}(cut) @-> {*dividir*}(separate)

It is possible to verify that the lowest node has more selection restrictions, namely, that it incorporates an instrument argument through lexical shadowing, i.e. it incorporates an argument that can only be syntactically realized through discourse specification devices (see Pustejovsky (1995)). Note, however, that according to the hypothesis that there is a straight connection between what is semantically determined and what is syntactically realized, it is possible to account for the syntactic occurrence restrictions of shadow arguments by using its hypernymy relations: shadow arguments can be syntactically realized with their hyponyms.

(36) a. *The lumberjack sawed the trees with a saw.

b. The lumberjack sawed the trees with a chainsaw.

(37) {*chainsaw*} @-> {*saw*}

In this way, the model seems able to account for the syntactic behavior of a set of lexical items without resorting to external devices – namely considering also function relations –, although further study on this issue is in order.

7 Conclusion

In this paper, we showed that the construction of a WN, adding semantic structure to the lexical entries, is a non-costly task that greatly enhances the computational potentialities of the model.

The use of the troponym relation to determine both the core and specific meaning of the lexical items, through the decompositional analysis of this relation, identifies essential elements of the semantic representation of verbal predicates in the GL framework: the argument and Qualia structure values. The different argument types in the argument structure account for different selection restrictions of verbs, whereas the association of Qualia information to the lexical structures and the assumption of an indirect Qualia unification operation predict compatibility among co-troponym verbs. Also, the WN as base model for the lexicon allows for lexical inheritance without resorting to external resources. Finally, the ontological nature of this relational model seems to allow for the replacement of semantic types by referring to other lexical items in the structure.

This paper refers on-going work that we intend to continue, namely in what concerns further exploring the verbs' properties that account for differentiated syntactic behavior.

8 Acknowledgements

My thanks to Palmira Marrafa, Rui Chaves and Sara Mendes for their comments, reviews and suggestions.

References

E. Agirre and D. Martinez. 2002. Integrating selectional preferences in wordnet. In *Proceedings of the First International WordNet Conference*. Mysore, India.

R. Amaro. 2005. Semantic incorporation in a portuguese wordnet of verbs of movement: on aktionsart shifting. In *Third International Workshop on Generative Approaches to the Lexicon*, pages 1–9. École de Traduction et d'Interpretation, University of Genéve.

F. Busa, N. Calzolari, A. Lenci, and J. Pustejovsky. 2001. Building a semantic lexicon: Structuring and generating concepts. In H. Bunt, R. Muskens, and E. Thijsse, editors, *Computing Meaning*, volume vol. II, pages 29–51. Dordrecht: Kluwer.

H. Dang, K. Kipper, and M. Palmer. 2000. Integrating compositional semantics into a verb lexicon. In *Proceedings of the 17th Conference on Computational Linguistics*, pages 1–9. Saarbrücken, Germany.

M. De Boni and S. Manandhar. 2002. Automated discovery of telic relations for wordnet. In *Proceedings of the first International WordNet conference*. India, `url://www-users.cs.york.ac.uk/ mdeboni/publications.html`.

C. Fabre and P. Sébillot. 1999. Semantic interpretation of binominal sequences and information retrieval. In *Proceedings of International ICSC Congress on Computational Intelligence: Methods and Applications*. CIMA'99, Symposium on Advances in Intelligent Data Analysis,AIDA'99, New York, `url://www.irisa.fr/texmex/Papers/1999/ sebillotFS99.pdf`.

C. Fellbaum. 1998. A semantic network of english verbs. In C. Fellbaum, editor, *WordNet: an electronic lexical database*, pages 69–104. MA: The MIT Press.

R. S. Jackendoff. 1983. *Semantics and Cognition*. Cambridge: The MIT Press.

J. Koenig, G. Mauner, and B. Bienvenu. 2002. Class specificity and the lexical encoding of participant information. *Brain & Language*, 81:224–235.

P. Marrafa and Moura. 2005. Portuguese telic causative verbs. In *Proceedings of the Third International Workshop on Generative Approaches to the Lexicon*, pages 144–148. École de Traduction et d'Interpretation, University of Genéve.

P. Marrafa. 2002. *The Portuguese WordNet: General Architecture and Semantic Internal Relations*. DELTA.

S. Mendes and R. P. Chaves. 2001. Enriching wordnet with qualia information. In *Proceedings of the Workshop on WordNets and Other Lexical Resources*, pages 108–112. NAACL 2001 Conference, Pittsburgh.

M. Moens. 1987. *Tense, Aspect and Temporal Reference*. Ph.D Thesis, University of Edinburgh.

T. Móia. 2000. *Identifying and computing temporal locating adverbials: with a particular focus on Portuguese and English*. Ph.D Thesis, University of Lisbon.

T. O'Hara and J. Wiebe. 2003. Classifying functional relations in factotum via wordnet hypernym associations. In *Proceedings of CICLing-2003*. CIC, IPN, Mexico City, `url://www.cs.nmsu.edu/ tomohara/ ohara_factotum_Roles_cicling03.pdf`.

J. Pustejovsky. 1995. *The Generative Lexicon*. MA: The MIT Press.

J. Pustejovsky. 2001. Type construction and the logic of concepts. In P. Bouillon and F. Busa, editors, *The Syntax of Word Meanings*. Cambridge University Press, `url: www.cs.brandeis.edu/ jamesp/articles/ type-construction.pdf`.

L. Talmy. 1985. Lexicalization patterns: semantic structure in lexical forms. In T. Shopen, editor, *Language typology and syntactic description*, volume vol. III, pages 57–149. Cambridge University Press.

T. Veale. 2003. Qualia extraction and creative metaphor in wordnet. In *Proceedings of the 3rd International Workshop on Creative Systems*. IJCAI 03, `url://afflatus.ucd.ie/papers/ ijcai2003.pdf`.

Z. Vendler. 1967. *Linguistics in Philosophy*. New York: Cornell University Press.

C.M̃. Verspoor. 1997. *Contextually-Dependent Lexical Semantics*. Ph.D Thesis, University of Edinburgh.