

Semantic Classes in CESS-LEX: Semantic Annotation of CESS-ECE¹

Taulé, M., Martí M.A., Castellví, J.

CLiC-UB (Centre de Llenguatge i Computació, Universitat de Barcelona)
{mtaule, amarti, joan.castellvi}@ub.edu

Abstract

Annotated corpora constitute a crucial resource to acquire or induce linguistic knowledge about how languages are used. In this sense, it is widely admitted that tagged corpora appear to be a very useful resource for computational and linguistic analysis of languages. The more explicit linguistic information they contain, the more interesting and useful they are. In this paper we present the theoretical basis for semantic annotation of two treebanks, CESS-ESP and CESS-CAT, focusing specially on the verbal semantic classes that determine the mapping between syntactic functions and semantic roles.

1 Introduction

The aim of this paper is to present the theoretical basis for semantic annotation of CESS-ESP and CESS-CAT corpora,² to be more specific, we will explain the semantic classes that determine the mapping between syntactic functions and semantic roles. Before annotating the corpora, a verbal lexicon (CESS-LEX)³ has been created, in which each verb is related to one or more semantic class, depending on its senses. For each sense the mapping between functions and thematic roles is declared, as well as the corresponding Lexical Semantic Structure (LSS). This lexicon will be used for the semiautomatic tagging of the treebank with thematic roles and verb senses. Further, each verb sense will be linked to one or more WordNet synsets.

Semantic properties used in our project have been defined assuming lexical decomposition (Levin & Rappaport Hovav [9], and Rappaport Hovav & Levin [12]), from which the concept of Lexical Semantic Structure is

¹ This research has been supported by the projects CESS-ECE (Corpus Etiquetados sintáctica y semánticamente para el Español, Catalán y Euskera) (HUM 2004-21127-E) and Lang2World (TIN2006-15265-C06-06).

² Syntactic annotation (constituents and functions) of these corpora began with the 3LB project, with the annotation of two corpora of 100.000 words, for Catalan and Spanish languages. Currently, the number of words is being extended by 400.000 words with CESS-ECE project, in CESS-CAT for Catalan and CESS-ESP for Spanish.

³ (Taulé *et al.* 2005).

taken. LSS determines the number of arguments that a verbal predicate requires and the thematic role of these arguments. In this direction, we follow the lines laid down by Kipper *et al.* [6] and Kingsbury *et al.* [5] in the construction of *VerbNet*. In order to adapt it to Catalan and Spanish languages we have also taken into account Demonte [3] and Mateu [10]. Moreover, we follow the *PropBank* annotation system for thematic roles (Palmer *et al.* [11], and Vázquez *et al.* [16] for the diathesis alternation treatment in both languages.

Semantic annotation of CESS-ECE corpora is being carried out in two phases: in the first stage, a corpus subset of 100.000 words is annotated in a semiautomatic way Civit *et al.* [2] and, in a second stage, the rest of the corpus, 400.000 will be automatically annotated applying machine learning techniques (Surdeanu [13]) and incremental methods in automatic annotation (Busser & Morante [1]).

In order to guarantee the coherence and quality and to ensure the correct mapping between thematic roles, syntactic functions and LSS, annotator agreement tests have been carried out. The aim of this annotator agreement task has been twofold: First, to improve the annotator agreement guidelines, and, second, to guarantee the consistency among the annotator's team.

This paper is organized as follows. Section 2 introduces the theoretical fundamentals, section 3 presents and discusses Lexical Semantic classes and, finally, section 4, contains our conclusions and focuses on further works.

2 Theoretical fundamentals in predicate analysis

The correspondence between syntactic functions and thematic roles is carried out following the predicate analysis presented in Levin & Rappaport-Hovav [9] and Rappaport-Hovav & Levin [12]. We consider that their proposal is appropriated for our work for a number of reasons. First, because in this model converge lexical semantic, event and argument information and diathesis alternations. And second, because similar works in corpus and computational linguistics have been carried out following this approach, such as the construction of *VerbNet* (Kipper *et al.* [7]), a lexicon with lexical semantic, argument and diathesis information for English predicates. *VerbNet* follows Levin's semantic classification (Levin [8]) and adopts *PropBank* semantic annotation (Palmer *et al.* [11]).

We characterize predicates by means of a limited number of LSS and Event Structure Patterns, according to the four basic event classes: states, activities, accomplishments, and achievements (Vendler [17], Dowty [4]). These general classes can be split in subclasses, as we will see in section 3. Thematic roles are determined by the event class that the predicate belongs to and the type of diathesis alternation that the predicate presents. Thus, not only thematic roles are assigned, but also predicates are characterized both from the aspectual and argument perspective. This enriches the linguistic

NLP applications. This information is currently being stored in the lexicon CESS-LEX for both languages.

2.1 Lexical Semantic Structures (LSS)

In contrast to Kipper *et al.* ([7]), which develop very complex LSSs, we have chosen more general classes that can be subclassified depending on thematic roles and diatheses. In the definition of these main classes we have taken into consideration only Argument0 and Argument1 because they are the basic arguments in defining the predicate structure. This gives rise to a very coarse grained classification that can be further split into subclasses. This subclassification has not been developed as much detailed as the thematic role assignment, since, although it can be very useful, it is not the main goal of this methodology (mapping thematic roles into syntactic functions).

We take four main structures that correspond to the four ontological event classes (Vendler [17], Dowty [4]), *states*, *activities* (or *processes*), *accomplishments* and *achievements*:

- (1) [x <STATE>]
- (2) [x ACT <MANNER/ INSTRUMENT> y]
- (3) [x CAUSE [BECOME [y <STATE/THING/PLACE>]]]
- (4) [BECOME [y <STATE>]]]

The LSS in (1) corresponds to the ontological class *state*, with just one entity involved in the event, and focuses in the state. The LSS (2) corresponds to *activities* or *processes*, it usually presents agentive subjects and allows passive constructions⁴. The LSS (3) corresponds to *accomplishments* that refer to resulting states in external cause processes, usually with a causative subject. Finally, the LSS of (4) corresponds to *achievements* that refer to a resulting state in processes without external cause.

2.2 Diathesis alternations and thematic roles

In our proposal each LSS restricts the set of all possible diatheses.⁵ Each verb sense is associated to one LSS, and the diatheses that each sense allows are the result of focusing different components of the LSS they belong to. In

⁴ In Catalan and Spanish there are two types of passive constructions: passives with the participle verb form and passives with ‘se’ (Esta mañana han sido vendidos cinco libros – Esta mañana se han vendido cinco libros ‘Five books have been sold this morning’).

⁵ We follow in essence the diathesis classification of Vázquez *et. al* (2000).

other words, diatheses are surface structures that result from focusing different components of the predicate LSS (CAUSE, ACT and BECOME).

For example, predicates defined as *accomplishments* (LSS (3)) allow causative, anticausative and resultative diatheses, which focus on the primitives CAUSE, BECOME, STATE:

- (5) *abrir*: [x CAUSE [BECOME [y <STATE>]]]
- (5a) Juan abre la puerta: [Juan CAUSE [BECOME [puerta <OPEN>]]]
(Causative)
'Juan opens the door'
- (5b) La puerta se abre: [x CAUSE [BECOME [puerta <OPEN>]]]
(Anticausative)
'The door opens'
- (5c) La puerta está abierta: [x CAUSE [BECOME [puerta <OPEN>]]]
(Resultative)
'The door is open'

(5) shows that the LSS - *x* causes *y* to change its state – associated to the predicate *abrir* 'open' surfaces in three different argument structures: causative (5a), anticausative (5b) and resultative (5c), as a result of focusing different aspects of the LSS (the component CAUSE, CHANGE (BECOME) or STATE, respectively).

Predicates allowed by these LSS hold the thematic role *Cause* in the variable *x* and *Patient* in *y*. Therefore, the relationship between thematic roles and syntactic functions is defined by the argument positions and by the diathesis alternations, ignoring the semantic type of arguments, i.e. [+/-human], [+/-volitional], [+/-affected], etc.

3 Spanish and Catalan Semantic Classes

In this section, we present de basic Lexical Semantic Classes derived from the LSS mentioned above. These classes are the result of combining the LSS with the Argument Structure and the thematic roles that can fulfill each argument. Each verbal class is also characterized for admitting specific diathesis alternations. All this information is captured in the verbal lexicon CESS-LEX where the syntactic-semantic interface is expressed. For each verbal sense a semantic class is established and the mapping between their syntactic functions⁶ with the corresponding argument structure and thematic roles is declared.

⁶ We extracted the verbal syntactic frames from the corpus as it has been described in

The semantic classes used to characterize verbal predicates are hierarchically arranged in two levels. The first level contains information about the LSS structure, which is closely related to the event structure. The second level contains information about argument structure and thematic roles. Thus, if a verb is related to a semantic class, it will provide access to syntactic and semantic information, and it will be possible to infer its event structure.

In the next section we present the 11 semantic classes that we so far have compiled, which are grouped around the 4 main LSS types. These classes are the result of the analysis of 470 verbs in the corpora with high-medium frequency of occurrence (with an occurrence rate between 40 and 4).⁷

On the basis of a draft of the annotation guide, annotator agreement tests have been carried out. In a first step, 70 verbs have been studied and tagged by five annotators in parallel, and in three phases (10, 30 and 30 verbs in each phase). After annotating each group an agreement discussion was carried out in order to revise and update the annotators guide. Once the guidelines were established, in a second step, 400 verbs were annotated by two pairs of annotators, each pair working in parallel with the same set of verbs. For these pairs of annotators the agreement rate was of 95% and 96%, respectively. This agreement rate has been obtained by confronting the results of the mapping between functions and thematic roles of one member of the pair against the other. The remaining 4% and 5% of disagreement has been discussed and the annotator guide modified when necessary. Almost all cases of disagreement are related to sense discrimination (assignment of LSS) and the identification of verbal forms, for instance, when it is necessary to decide if a given structure corresponds to a verb and its complements or to an idiom (dar + un susto vs. dar_un_susto, 'to fright'). We are currently working on the analysis of the remaining 1000 verbs, which comprises the 36 most frequent verbs (5096 occurrences) and 988 verbs with a frequency between 3 and 1 (1448). In this last step, annotators work independently.

3.1. LSS1: [x CAUSE [BECOME [y <STATE/THING/PLACE>]]]

LSS1 usually corresponds to the event structure of the accomplishments⁸ and shares the resultative alternation. In this LSS, we distinguish two main classes, the transitive-causative class and the causative agent class.

Transitive-Causative class:

The transitive-causative class is characterized by the fact that the verbs belonging to this class accept, as a specific characteristic, the anticausative

⁷ The 470 selected verbs correspond to 4585 verb occurrences. The total number of verbs in the corpus is 1495, which corresponds to 11708 occurrences.

alternation⁹. These verbs can be characterized as verbs of change of state, where the object is always affected. The subject of these verbs in the causative alternation maps into the Argument-0 with the thematic role *Cause*, and the direct object maps to Argument-1 and *Patient* role. In this case, we are dealing with a direct cause.

LSS1.1

[x CAUSE [BECOME [y <STATE >]]]

SUJ Arg0##CAU

OD Arg1##PAT

Diatheses: [+Anticausative] [+Resultative] [+/- Passive]

Example: Juan_{Arg0-CAU} abre la ventana
'Juan opens the window'

Spanish verbs: *abrir, aclarar, agotar, alegrarse, asustarse, babosear, balancear, cerrar, construir, emocionar(se), enamorar, freír, hervir, hundir, mejorar, oscurecer, purificar, romper, sacralizar, tintar...*¹⁰

Catalan verbs: *aclarir, bullir, construir, emocionar, enfonsar, esgotar, espantar, fregir, millorar, obrir, purificar, tancar, trencar...*

Causative Agent class:

The causative agent class includes basically those verbs implying a change of location, where an acting agent, the syntactic subject, causes the direct object to become in another location or position. Thus, in these cases, we are dealing with indirect causes in which the Argument-0 is represented as an *Agent* and the Argument-1 as a *Patient*. Most of these verbs allow the passive alternation and not the anticausative one. We consider that all these facts support the treatment of the subject as an *Agent*.

LSS1.2

[x CAUSE [BECOME [y <PLACE >]]] or [x CAUSE [BECOME [<THING> IN y]]]

SUJ Arg0##AGT

OD Arg1##PAT

Diatheses: [- Anticausative] [+/-Resultative] [+Passive]

Example: El médico_{Arg0##AGT} hospitalizó al paciente_{Arg1##PAT}
'The doctor hospitalized the patient'
Juan_{Arg0##AGT} ensilla el caballo_{Arg1##PAT}
'Juan saddles up the horse'

⁹ Anti-causative alternation is also known as ergative or inchoative alternation.

¹⁰ We understand that it is one of the possible senses of these verbs. Obviously, we

Spanish verbs: *almacenar, bajar (an object), colocar, encarcelar, empaquetar, empapelar, enharinar, ensillar, ensobrar, hospitalizar, meter (las sillas), poner, subir (an object)*...

Catalan verbs: *baixar (an object), col·locar, emmagatzemar, emmantegar, empaquetar, empresonar, ensobrar, hospitalitzar, posar, pujar (an object), ...*

3.2. LSS2: [BECOME [*y* <STATE >/<PLACE>]]

Verbs belonging to the LSS2 correspond to the event structure of achievements, and they are basically unaccusative verbs. Currently, we include in this class the verbs of inherent directed motion and verbs of appearance and disappearance. These verbs neither participate in the passive, the anticausative nor the resultative alternation. The subject maps into the Argument-1 with the thematic role *Theme*.

Unaccusative class:

This class includes intransitive verbs whose subject behaves as an internal argument. In some languages, such as Catalan¹¹, this subject is characterized by the fact that allows the clitization with the pronoun ‘*en*’; for example: ‘Han arribat quinze turistes’ vs. ‘**N**han arribat quinze’ (‘Fifteen tourists have arrived’). The subject of these verbs usually appears in the postverbal position. This last characteristic is also found in Spanish. Most of these verbs are included in the Levin’s *inherently directed motion class*, which is a subgroup of verbs that can express the Origin (Argument-1) and the Goal (Argument-2), as in ‘Ana viene de París_{CREG-Arg1-ORI}’ (‘Ana comes from Paris’); ‘Ana sale de casa_{CREG-Arg2-DES}’ (‘Ana leaves home’).

LSS2.1

[BECOME [*y* <PLACE>]] or [BECOME [*y* <STATE>]]

SUJ Arg1##TEM

Diatheses: [-Passive]

Example: Los niños_{Arg1##TEM} llegaron tarde

‘The kids arrived late’

Los ladrones_{Arg1##TEM} desaparecieron sin dejar rastro

‘The thieves vanished without a trace’

Spanish verbs: *aparecer, caer, crecer, desaparecer, emerger?, entrar, llegar, morir, salir, venir,...*

Catalan verbs: *aparèixer, arribar, caure, créixer, desaparèixer, entrar, morir, sortir, venir, ...*

3.3. LSS3: [*x/y* <STATE >]

The verbal classes related to LSS3 denote states, and typically they can not be controlled by an *Agent*. We have distinguished three classes depending on the argument structure and the type of subject allowed by the verbal predicates. We basically differentiate between *state unaccusative*, *state unergative* and *state transitive* classes.

State Unaccusative class:

All the members of this class have intransitive uses and they are specifically treated as unaccusative. We represent their subject as an Argument-1 mapping the thematic role *Theme*. We also include aspectual verbs in this class, that is to say, verbs that basically describe the initiation and termination of an activity.

LSS3.1

[y <STATE >]

SUJ Arg1##TEM

Diatheses: [-Passive] [-Cognate Object]

Example: Hay cuatro personas Arg1##TEM esperando

‘There are four people waiting’

El año Arg1##TEM acaba el 31 de diciembre ArgM##TMP

‘The year ends on December 31’

El 31 de diciembre ArgM##TMP acaba el año Arg1##TEM’

‘The 31st of December finishes the year’ (*literal translation*)

Spanish verbs: *acabar, comenzar, empezar, existir, haber, terminar*,...

Catalan verbs: *acabar, començar, existir, haver-hi*,...

State Unergative class:

This class comprises unergative verbs denoting a state. Though intransitive, they are different from LSS2.1 in their thematic role assignment, since subjects of *State Unergative* verbs are Argument-0 mapping the role *Experiencer*.

LSS3.2

[x <STATE >]

SUJ Arg0##EXP

Diatheses: [-Passive], [+Cognate Object]¹²

Example: Juan Arg0##EXP sueña

‘Juan dreams’

Spanish verbs: *babear, brillar, burbujear, centellar, crecer (niño), chorrear, destellar, dormir, llorar, oler, parpadear, roncar, soñar, sudar, temblar, vivir*,...

Catalan verbs: *brillar, dormir, roncar, somiar, suar, tremolar, viure*...

State Transitive class:

This class is mainly integrated by copulative verbs, but also by those verbs that describe the value of some attribute of an entity along a scale (*Measure verbs* in the Levin's classification (Levin B., [8]). Passive alternation occurs rarely in this class of verbs.

LSS3.3

[*x* <STATE> *y*]

SUJ Arg1##TEM

ATR Arg2##ATR

OD Arg2##ATR

Diatheses: [-Passive]

Example: Juan_{Arg1##TEM} tiene dos hipotecas_{Arg2##ATR}

'Juan has got two loans'

La película_{Arg1##TEM} es interesante_{Arg2##ATR}

'The film is interesting'

Spanish verbs: *estar, medir, parecer, pesar, poseer, ser, tener ...*

Catalan verbs: *estar, medir, pesar, posseir, semblar, ser, tenir...*

3.4. LSS4: [*x* ACT <MANNER/INSTRUMENT> *y*]

Most verbal classes related to LSS4 denote activities and, consequently, the verbs involved share an agentive subject. That is, Argument-0 always maps into thematic role *Agent*, while the Argument-1, if there is any, always fits with *Patient*. If there is a *Patient*, the passive alternation is necessarily possible. By the moment, we have distinguished four different semantic classes, mainly depending on the predicate's arity: *agentive inergative class, agentive transitive class, agentive ditransitive class and locative ditransitive class*.

Agentive Unergative class:

All the verbs in this class have intransitive uses and most of them typically describe manner of motion, involving or not displacement. Most verbs involving movement (i.e. *nadar, correr*, etc.) can display the extension object alternation, that is, they can be used in a transitive way expressing an extension or a measure phrase.

LSS4.1

[*x* ACT <MANNER/INSTRUMENT> *y*]

SUJ Arg0##AGT

Diatheses: [-Passive], [+/-Extension Object]

Example: Juan_{Arg0##AGT} corre

'Juan is running'

Juan_{Arg0##AGT} caminó tres kilómetros_{Arg1##EXT}

‘Juan walked three kilometers’

Spanish verbs: *caminar, contonearse, correr, establecerse, ir, menear(se), nadar...*

Catalan verbs: *anar, caminar, córrer, cridar, nadar, ...*

Agentive Transitive class:

This class comprises verbs typically transitive that present Argument-0 with the thematic role *Agent* and Argument-1 with *Patient*. It is the largest class in Catalan and Spanish languages.

LSS4.2

[x ACT <MANNER/INSTRUMENT> y]

SUJ Arg0##AGT

OD Arg1##PAT

Diatheses: [+Passive]

Example: Juan_{Arg0##AGT} lee una novela histórica
‘Juan reads historical novel’

Spanish verbs: *amar, barrer, beber, cantar, cepillar, comer, desear, fregar, leer, odiar, peinar, silbar...*

Catalan verbs: *beure, cantar, desitjar, escombrar, estimar, fregar, llegir, odiar, pentinar, raspallar, xiular...*

Agentive Ditransitive class:

The verbs of this semantic class are characterized by presenting a double object, one expressing the *Patient* (Argument-1) and another referring to the *Beneficiary* (Argument-2). For example, verbs expressing change of possession and communication verbs can fit this class. That is, when any kind of transfer of possession, information or ideas is carried out.

LSS4.3

[x ACT <MANNER/INSTRUMENT> y]

SUJ Arg0##AGT

OD Arg1##PAT

OI Arg2##BEN

Diatheses: [+Passive], [-Subject Locative]

Example: Juan_{Arg0##AGT} da un caramelo_{Arg1##PAT} a la niña_{Arg2##BEN}
‘Juan gives a candy to the child’

Spanish verbs: *cantar, contar, dar, decir, entregar, enviar, explicar...*

Catalan verbs: *cantar, contar, dir, donar, enviar, explicar, lliurar...*

Locative Ditransitive class:

This class is characterized by admitting the subject locative alternation, that is, Argument-2 with thematic role *Locative* can occur in a subject

position, for example: ‘El autor aborda la discriminación de género en su ensayo_{CC-Arg2-LOC}’ vs. ‘El ensayo_{SUJ-Arg2-LOC} aborda la discriminación de género’ (‘The author tackles gender discrimination in his essay’ vs ‘The essay tackles gender discrimination’).

LSS4.4

[x ACT <MANNER/INSTRUMENT> y]

Arg0##AGT

Arg1##PAT

Arg2##LOC

Diatheses: [+Passive], [+Subject Locative]

Example: El autor_{Arg0##AGT} aborda esa temática_{Arg1##PAT} en la novela

Arg2##LOC

‘The author tackles this subject in his novel’

Spanish verbs: *abordar, acoger, registra, tratar ...*

Catalan verbs: *abordar, recollir, registrar, tractar...*

4 Conclusions and further work

In this paper we have presented the theoretical and methodological basis for semantic annotation of two treebanks, CESS-ESP and CESS-CAT, focusing specially on the verbal semantic classes that determine the mapping between syntactic functions and semantic roles. We have set four main Lexical Semantic Structures (LSS) that define the diatheses alternations that a predicate accepts. For each LSS several subclasses are defined taking into account the argument structure and thematic roles. All this information is represented in the verbal lexicon CESS-LEX, where verbal predicates are semantically characterized. This lexicon will be used for the semiautomatic semantic tagging of CESS-ECE. As a result of the tagging process it will be available a corpus annotated, not only with thematic roles, but also with event and argument structure for each sentence.

Since these corpora are being annotated with WordNet synsets, both resources, CESS-LEX and WordNet, will be linked and mutually enriched.

CESS-ECE will be used as training and test corpora in SemEval-1, in the multilevel semantic annotation of Catalan and Spanish Tasks.

Currently, a consulting interface for a small part of these corpora is available at <http://www.lsi.upc.edu/~mbertran/cess-ece/index.php>.

References

- [1] B. Busser, R. Morante (2005). “Designing an active learning based system for corpus annotation”. *SEPLN*, Granada.
- [2] M. Civit, I. Aldezabal, E. Pociello, M. Taulé, J. Aparicio, L. Màrquez, B. Navarro, J. Castellví, M.A. Martí (2005a). “3LB-LEX: léxico verbal con frames sintáctico-semánticos”. *SEPLN*, Granada.

- [3] V. Demonte (2003). "Preliminares de una clasificación léxico-sintáctica de los predicados verbales del español". En Sybille Grosse y Axel Schönberger (eds.) *Ex oriente lux: Festschrift für Eberhard Gärtner zu seinem 60. Geburtstag*. Frankfurt am Main: Valentia.
- [4] D. Dowty (1991). 'Thematic proto-roles and argument selection'. *Language*, 67.
- [5] P. Kingsbury, M. Palmer, M. Marcus (2002). "Adding semantic annotation to Penn TreeBank". 2002. En *Proceedings of the 2002 Conference on Human Language Technology*, San Diego, CA.
- [6] K. Kipper, H T. Dang, M. Palmer (2000). "Class-Based Construction of a Verb Lexicon". *AAAI-2000 Seventeenth National Conference on Artificial Intelligence*, Austin, TX.
- [7] K. Kipper, M. Palmer, O. Rambow (2002). "Extending PropBank with VerbNet Semantic Predicates". *Workshop on Applied Interlinguas, held in conjunction with AMTA-2002*. Tiburon, CA.
- [8] B. Levin (1993). *English Verb Classes and Alternations: A preliminary investigation*. University of Chicago Press, Chicago.
- [9] B. Levin, M. Rappaport-Hovav (1995). *Unaccusativity. At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.
- [10] J. Mateu (2004). "Relational Construal at the Interface", *Anuari de Filologia*, xxVI, Barcelona: Universitat de Barcelona.
- [11] M. Palmer, P. Kingsbury, D. Gildea (2005). "The Proposition Bank: An Annotated Corpus of Semantic Roles", *Computational Linguistics*, 21 (1). USA: MIT Press.
- [12] M. Rappaport Hovav, B. Levin (1998). "Building Verb Meanings", in M. Butt and W. Geuder, eds., *The Projection of Arguments: Lexical and Compositional Factors*, CSLI Publications, Stanford, CA, 97-134.
- [13] M. Surdeanu (2006): <http://www.lsi.upc.edu/~surdeanu/swirl.html>. Etiquetador automático de roles semánticos de software libre.
- [14] M. Taulé, J. Aparicio, J. Castellví, M.A. Martí, (2005). "Mapping syntactic functions into semantic roles", *Proceedings of the Fourth Workshop on Treebanks and Linguistic Theories (TLT 2005)*". Barcelona: Universitat de Barcelona.
- [15] M. Taulé, J. Castellví (2006a). "La interfaz sintaxis-semántica: anotación de corpus con papeles temáticos", *Actas del VII Congreso de Lingüística General*. Barcelona: Universitat de Barcelona.
- [16] G. Vázquez, A. Fernández, Martí, M.A. (2000). *Clasificación verbal. Alternancias de diátesis*. Edicions de la Universitat de Lleida: Lleida.
- [17] Z. Vendler (1967). *Linguistics in Philosophy*. Cornell University Press: New York.