
Guidelines

Bàrbara Soriano
Oriol Borrega
Mariona Taulé
M^a Antònia Martí

3LB-WP 03-02

Universitat de Barcelona
Departament de Lingüística

2008

Contents

1	Introduction	1
1.1	Preliminary questions	1
1.2	Technical issues	1
1.3	Elliptic subjects and word ordering	2
2	Constituents	7
2.1	Introduction	7
2.2	Preliminary questions	7
2.3	Stratification levels and bracketing	7
2.3.1	Structure stratification	7
2.3.2	Coordination	8
2.3.2.1	Some terminological issues on adjunction	9
2.4	Phrase level	10
2.4.1	Noun phrase (sn)	11
2.4.1.1	The head	12
2.4.1.2	Determiners	15
2.4.1.3	Complements of the noun phrase	19
2.4.1.4	Adjunction to the noun phrase	25
2.4.1.5	Nominalization	27
2.4.1.6	Other elements treated as noun phrases	30
2.4.1.7	Coordination within the noun phrase	31
2.4.1.8	Noun locutions	32
2.4.2	Adjective phrase and participle clause	32
2.4.2.1	Argumental adjective phrase (<sa>)	33
2.4.2.2	Adjective phrase as a noun complement (<s.a>)	33
2.4.2.3	Complements of the adjective phrase	34
2.4.2.4	Adjectivisation of nouns	37
2.4.2.5	Adjective locutions	38
2.4.3	Adverbial phrase (<sadv>)	38
2.4.3.1	<i>Més</i> and <i>menys</i>	41
2.4.3.2	Adverbial locutions	43
2.4.3.3	Attaching the adverbial phrase	44
2.4.3.4	Adverbs belonging to a specifier	45

2.4.4	Prepositional phrase (<sp>)	46
2.4.4.1	Prepositions (<s>)	49
2.4.4.2	Prepositional locutions	56
2.4.5	Verb group (<grup.verb> or <gv>)	59
2.4.5.1	Verbal periphrasis	61
2.4.5.2	Coordinating periphrases	63
2.4.5.3	Encrusted elements within the periphrases	64
2.4.5.4	Verb locutions	65
2.4.6	The morpheme 'es'	66
2.5	Sentence level	69
2.5.1	Non-finite sentences	69
2.5.1.1	Non-finite completive clause	70
2.5.1.2	Non-finite adjective clause and absolute clause	70
2.5.1.3	Non-finite adverbial clause	71
2.5.1.4	Non-finite relative clause	72
2.5.2	Finite sentences	72
2.5.2.1	Finite completive clause	73
2.5.2.2	Finite relative clause	75
2.5.2.3	Finite adverbial clause	80
2.5.2.4	Other adverbial clauses ('impròpies')	81
2.5.2.5	The superlative	84
2.6	Coordination	84
2.6.1	Coordinating elements	85
2.6.2	Distributed coordination	87
2.6.3	Distributed-like structures	88
2.6.4	Coordination of different elements	90
2.6.4.1	Sentence structures	91
2.6.4.2	Phrase structures	92
2.6.4.3	Phrases + sentence structures	93
2.6.5	Complementizer ellipsis at the coordination	95
2.6.5.1	Particular cases	95
2.6.6	Verbal partial ellipsis at the coordination	96
2.6.7	Adjunction at the coordination	96
2.6.7.1	Special cases in which there is no adjunction	96
2.7	The adjunction	97
2.7.1	Complements of a coordinated structure	97
2.7.2	Comparative structures	98
2.7.3	Adverbs preceding the verb	99
2.7.4	Complements of the noun phrase	100
2.7.5	Particular cases	101
2.8	Discontinuity	101
2.8.1	General case	101
2.8.2	Particular cases	102
2.9	Miscellany	103

2.9.1	Numbers	103
2.9.2	Dates	104
2.9.2.1	Time intervals	106
2.9.3	Currency	107
3	Functions	109
3.1	Introduction	109
3.2	Preliminary questions	109
3.2.1	Elements with no syntactic function	110
3.3	Tag set	111
3.4	Syntactic function, coordination and adjunction	113
3.5	Subject	114
3.5.1	Impersonal sentences	115
3.5.2	Subject of attributive verbs	117
3.5.3	The subject and the clitics	118
3.6	Direct object	118
3.6.1	Quantitative direct object	119
3.7	Indirect object	120
3.7.1	Ethical dative	120
3.8	Attribute	121
3.8.1	Particular cases	122
3.9	Predicative complement	123
3.9.1	Special cases	126
3.9.2	Predicative with <i>ser</i> ('to be')	126
3.10	Prepositional object	126
3.10.1	Verbs requiring either direct or prepositional object	128
3.10.2	Verbs with both direct and prepositional object	128
3.10.2.1	Second prepositional object	129
3.10.2.2	Verbs with two prepositional phrases	129
3.10.2.3	Verbs alternating a non-pronominal transitive structure with a pronominal one with prepositional object	129
3.10.3	Particular cases	130
3.10.3.1	The verb <i>canviar</i> 'to change'	130
3.10.3.2	The verb <i>servir</i> 'to serve, work as'	131
3.10.3.3	Obligatory complements expressing place or manner	131
3.10.3.4	Motion verbs with two complements	132
3.10.4	Preposition Swap and Drop	133
3.11	Agent complement	134
3.12	Adjunct	135
3.12.1	Particular cases	136
3.12.1.1	Adverbs <i>ja</i> and <i>encara</i>	136
3.12.1.2	The preposition <i>sobre</i>	136
3.12.1.3	Adverbs <i>més</i> and <i>menys</i> and other quantifiers	136
3.12.1.4	Participle clauses and absolute participle clauses	137

3.12.1.5	The verb <i>arribar</i> ‘to arrive’	137
3.12.1.6	The verb <i>servir</i>	137
3.12.1.7	The verb <i>canviar</i>	137
3.12.1.8	Verbs always cooccurring with modifiers	137
3.12.1.9	The verbs <i>passar</i> , <i>anar</i> i <i>venir</i>	138
3.12.1.10	Other complex structures as modifiers	138
3.13	Textual element	139
3.14	Modality element	139
3.15	Passive marker	140
3.16	Impersonality marker	141
3.17	Vocative	141
3.18	Sentence Adjunct	141
3.18.1	The preposition <i>sobre</i>	143
3.19	Clitic pronouns and syntactic functions	143
3.19.1	The pronoun <i>EN</i>	144
3.19.2	The pronoun <i>HI</i>	144
3.19.3	The pronoun <i>HO</i>	145
3.20	Repeated functions	146

Chapter 1

Introduction

This document presents the adopted methodology for the syntactic annotation (both constituents and functions) of AnCora Corpora. In this work there are no indications about semantic or anaphoric annotation.

This section gives a brief description of the annotation process and introduces the general directions to carry it out. Section 2 is devoted to the first stage of the syntactic annotation; that is, the constituents annotation. It includes the stratification and bracketing conventions we take for both phrase and sentence level annotation. In section 3, function tagset and annotation criteria are discussed. First, we present which nodes are to be annotated and then, we move on to provide information on each of the different syntactic functions which have been defined.

1.1 Preliminary questions

Regarding sentence delimitation, orthographic conventions are followed: It is considered to be a sentence every string of words between two full stops or strong punctuation marks (question and exclamation marks as well as certain suspension dots). The system delimits only unambiguous sentences, so it is possible to find non-splitted sentences. In such a case, the splitting must be done manually.

The annotation is carried out in two different stages: During the first one, only constituents are annotated: The bracketing given by TACAT is partial *-chunking-*, therefore this first stage of the annotation consists in marking, bracketing and labeling the constituents. It is usually required to enlarge the constituents by adding branches to those nodes of the tree that appear to be sisters instead of daughters. There are also few cases in which some nodes are grouped and they should be splitted up.

1.2 Technical issues

The annotation we carry out is based on a single XML data format. Items are stored in UTF-8 encoded XML format. XML allows portability and takes advantage of the several

tools and libraries available in a variety of platforms and programming languages. Besides, UTF-8 allows the format to be cross-lingual. XML has a tree structure itself, so it maps easily to the syntactic constituent structure. The following are some of the reasons for using this format:

- Small set of node names: only 15 node names are possible. Thus, nodes are only generic and specificity is reached through attributes.
- Attributes are atomic: each attribute labels one and only one feature of the node. This reduces the number of possible values and makes the annotation levels independent.
- Attributes describe only their node. This makes moving, deleting and creating nodes very simple tasks, and so coherence is guaranteed.
- No redundant data.
- Easy to add new annotation levels: only the design of a new attribute and its possible values is needed.

The following are some examples of the XML attribute-value pairs¹.

```
<nodename name="n" description="Noun" onlyterminal="yes" />
<nodename name="v" description="Verb" onlyterminal="yes" />
<nodename name="a" description="Adjective" onlyterminal="yes" />
<nodename name="d" description="Determiner" onlyterminal="yes" />
<nodename name="r" description="Adverb" onlyterminal="yes" />
<nodename name="f" description="Punctuation mark" onlyterminal="yes" />
<nodename name="s" description="Preposition" onlyterminal="yes" />
<nodename name="p" description="Pronoun" onlyterminal="yes" />
<nodename name="c" description="Conjunction" onlyterminal="yes" />
```

The second stage of the annotation consists in labeling the syntactic functions. The annotators of this stage of the process will work over the constituents annotation. This imply a revision of the splitting-up process and grouping carried out during the previous annotation task.

1.3 Elliptic subjects and word ordering

- The surface word ordering within the sentence should not be altered.
- It is worth saying that there is no constituent labeled as *verbal phrase*. The reason is found in the previous point: Since superficial word ordering should not be altered, the verbal phrase cannot be taken into account because Catalan and Spanish are free-ordering languages².

¹See appendix for a complete list.

²In many cases, such as that of the reflexive passives, for instance, the subject holds postverbal position. As a consequence, the verbal phrase sequence brakes up.

- Both Catalan and Spanish are *pro-drop* languages and for this reason the subject may not be explicit. In such a case, the subject should be manually added, by inserting a new node

```
<sn/>
```

and XML-features indicating that the constituent is elliptic:

```
- <attribute name="elliptic" onlyterminal="yes"
description="Is the constituent ellipted?">
  <value name="yes" />
  <value name="no" />
</attribute>
```

The final appearance of the node is the following:

```
<sn elliptic="yes" />
```

This new element has to be placed immediately before the conjugated verb or the set consisting of [modal adverbs, negative adverbs and/or clitics + verb] (1.1 and 1.2). We consider to be modal adverbs those of the kind *ja* ‘already’, *encara* ‘yet’, *potser* ‘maybe’, *també* ‘also’... (*ya, todavía, quizá, también* in Spanish); we include in the group of negative adverbs not only the form *no* but also others like *mai* ‘never’, *tampoc* ‘neither’, etc. (*nunca, tampoco* in Spanish).

- (1.1) *A Eivissa Ø tenim un patrimoni cultural únic* (Cat 06010100-49)
‘In Eivissa (we) have a unique cultural heritage’
- (1.2) *Ø ofreció un total de 114 millones de euros* (Sp 1925_20001103)
‘(s/he) offered a total amount of 114 million euros’

The only cases in which the covert subject should be placed after the verb are those of direct or indirect questions (1.7), imperatives (1.8) and those introducing direct speech 1.10.

The following are examples of elliptic subjects:

1. With clitics

- (1.3) *Ø les facin constar en els respectius pressupostos* (Cat 211nou)
‘(they) put them on record in the correspondent budgets’

2. With negative adverbs

- (1.4) *Ø no podem cedir terreny a Badia* (Cat 211nou)
‘(we) cannot give up ground to Badia’

3. With modal adverbs

- (1.5) *Ø També inclou els terrenys* (Cat 06019900-18)
 ‘(it) Also includes the grounds’

4. With modal and negative adverbs³

- (1.6) ... *quizás Ø decida continuar con los juegos de guerra* (Sp 15776_20000320_0)
 ‘... (s/he) maybe decides to continue with the war games’

5. In questions

- (1.7) *decidir en quina direcció enfoquem Ø els fets* (Cat 10716)
 ‘to decide in which direction (we) focus the facts’

6. In imperative sentences

- (1.8) *Párate Ø un momento, mira Ø hacia atrás y reconócese Ø* (Sp d16)
 ‘Stop for a while, look back and recognize you’

7. Before a temporal adverb

- (1.9) *Ø sempre està disponible* (Cat 211nou)
 ‘(it) is always available’

8. Introducing direct speech

- (1.10) *‘A Eivissa tenim un patrimoni únic al món i veient aquesta exposició queda demostrat’, va dir Ø* (Cat 06010100-49.4)
 ‘In Eivissa (we) have a unique cultural heritage and it is shown by seeing this exhibition, (s/he) said’

In some cases, as in *ahí estamos* ‘here we go’, it seems to be more natural to place the subject after the verb. This is probably due to the fact that the circumstantial complement of the verb *estar* ‘to be’ is dislocated. It is important to make a clear distinction between covert subject sentences and impersonal sentences. In the latter, the covert subject should not be marked.

We considered the following sentences to be impersonal:

1. **Sentences which contain predicates indicating atmospheric phenomena**, as long as they are used in a literal sense⁴.

³When *quizá* ‘maybe’ is a modal adverb, the subject must be placed in preverbal position.

⁴When the use is metaphoric the subject should be overt and marked: *li van ploure crítiques* (literally: critiques rained over him/her ‘(s/he) received many critiques’).

2. Non-finite sentences⁵

(1.11) *il·luminar aquest diumenge el puig* (Cat 06010300-28)
 ‘to illuminate the mountain next Sunday’

(1.12) *documents confiscats a la Generalitat* (Cat 211.7)
 ‘documents requisited to Generalitat’

(1.13) *incomplint així un acord del Parlament* (Cat 211nou)
 ‘no fulfilling, this way, a Parliament agreement’

3. Sentences containing the verbal morpheme *es*⁶

(1.14) *Es tracta d’una mostra* (Cat 06019900-19)
 ‘It is a sample’

4. Structures such as those of *sembla que* ‘it seems that...’; *cal que* ‘it is necessary...’; *hi ha* ‘there is/are’... (Catalan); and *parece que* ‘it seems that...’; *es que* ‘what happens is that...’ (Spanish):

(1.15) *No hi ha cap previsió* (Cat 06010100-13)
 ‘There is no forecast’

⁵Non-finite sentences are described in more detail in section 2.5.1.

⁶See section 2.4.6. for more details on verbal morphemes.

Chapter 2

Constituents

2.1 Introduction

This section describes the criteria to carry out the syntactic constituent annotation within the AnCora corpora (see <http://clic.ub.edu/ancora/>).

2.2 Preliminary questions

1. Labels of syntactic constituent are XML-like attribute-value pairs. These attributes give the morphological and syntactic information to different nodes within the sentence, from the part-of-speech (PoS) level to the highest phrase or sentence node.
2. During the constituency annotation stage, it is built the hierarchial distribution of every element within the main sentence.
3. There are specific marks to indicate whether or not a sentence is verbless, which is also annotated during this stage.
4. Specific attribute-value also indicates the different coordinated structures. Coordinated phrases have no head.

2.3 Stratification levels and bracketing

2.3.1 Structure stratification

The representation of tree structures we propose is quite flat. This is the solution to assume in those cases in which ambiguity poses difficulties in choosing between a plane structure and another with some stratification levels. To establish constituency bracketing we have to resort to any available source of information: The context itself, encyclopaedic knowledge, etc. Sometimes, however, it is still difficult to split the constituents of the sentence (2.1), (2.2):

(2.1) $[[ha\ ido][de\ la\ insatisfacci3n][a\ la\ violencia]]$ ‘(s/he) went from dissatisfaction to violence’

(2.2) $[[ha\ ido][de\ la\ insatisfacci3n\ [a\ la\ violencia]]]$

When this happens and the context does not help in disambiguating, the constituent will be attached to the highest possible node.

Regarding the noun phrase with more than one complement, the representation has to follow the criteria stated above. The cases in which, although at the semantic level it is arguable that the second complement refers to the group formed by the noun and its first complement, our representation shows both complements as if they were sisters 2.3:

(2.3) $punt\ [de\ vista]\ [de\ la\ mobilitat]$ (Cat 107-5.9)
‘point of view of the mobility’

2.3.2 Coordination

The coordinated elements and the coordinating conjunction are sisters and there is no hierarchical relationship among them. Unlike what happens to noun complements, the representation of coordination may be more complex, and we have to resort to semantic interpretation to disambiguate and decide which structure is the more suitable. The following are the feature-value XML-pairs to mark coordination:

```
<nodename name="coord" description="Coordinating element"
  onlyterminal="no" />
- <attribute name="conjunctiontype" description="Conjunction type"
parents="conj">
  <value name="subordinating" description="Subordinating conjunction" />
  <value name="coordinating" description="Coordinating conjunction" />
</attribute>
- <attribute name="coord" onlynonterminal="yes"
description="Is it a coordinate constituent?">
  <value name="yes" />
  <value name="no" />
</attribute>
- <attribute name="complex" onlynonterminal="yes"
description="Is it complex?">
  <value name="yes" />
  <value name="no" />
</attribute>
```

Example 2.4 shows the resulting tree of a noun phrase coordination. Note that an adjective phrase complements both noun phrases. Complements of a coordinated structure are adjoined to the whole structure.

(2.4) (Cat 06010100-12)

```

<sentence>
  <sn>
    <grup.nom coord="yes">
      <grup.nom coord="yes">
        <grup.nom gen="m" num="p">
          <n [...] wd="Diputats"/>
        </grup.nom>
      <coord>
        <c lem="i" pos="cc" postype="coordinating" wd="i"/>
      </coord>
      <grup.nom gen="m" num="p">
        <n [...] wd="regidors"/>
      </grup.nom>
    </grup.nom>
    <s.a adjunct="yes" gen="m" num="p">
      <grup.a gen="m" num="p">
        <a [...] wd="republicans"/>
      </grup.a>
    </s.a>
  </grup.nom>
</sn>

```

‘Republican deputies and councillors’

Although the basic criterion is to coordinate at the lowest possible level, sometimes this is not possible, as it happens when there are complements affecting in a different way the coordinated elements. However, when it is the subject the only argument shared by coordinated sentences, there is no adjunction: The overt subject remains in its correspondent sentence and a covert subject must be added in the other one (2.5):

(2.5) (Cat 204.2)

[*L'atur masculí es va situar ... en 64.248*] i [*∅ va disminuir ... en 2.493*]

‘Masculine unemployment arrived ... at 64.248 and decreased ... in 2.493’

2.3.2.1 Some terminological issues on adjunction

From a syntactic point of view, we do not make a distinction between argumental and non-argumental complements. However, some explanations about the usage of certain technical terms are needed to avoid ambiguity as regards this topic:

- **Adjunction:** By this term, we mean the act or process of *adjoining*; that is, a constituent or structure that depends externally of another one. In our representation, this implies that the head node is doubled (unless it is not possible due to linear reasons.) By adjoining, thus, we get structures such as the following:

(2.6) *[[Jordi_Portabella] [(ERC)]]*

- **Attachment**: It has no adjoining or doubling consequences, as it just deals with assign, associate, bind or make a node depend on another.
- **Adjunct**: This term describes two different concepts. On the one hand, it refers to the concrete XML feature value that indicates *adjunction*. On the other hand, it refers to a certain verbal complement indicating time, manner, reason, etc., which is optional or structurally dispensable. That is the way it is used in chapter 3, when referring to this kind of complements (see section 3.12).

2.4 Phrase level

We follow the X-bar formalism. Each phrase head consists on several nodes, which are manually added when they are not given by the parser. In addition to the phrase level node, there is a group node, to which the complements are attached. A noun that is the head of a noun phrase, therefore, has a morphological node (in which information on category, gender and number is contained), an intermediate node (`<grup.nom>`), and the phrase node (`<sn>`). Other phrases (adverbial, adjective, etc.) are labeled the same way. An example is given in 2.7:

(2.7) `<sn>`

```

<spec gen="m" num="s">
  <d gen="m" lem="el" num="s" postype="article" wd="El"/>
</spec>
<grup.nom gen="m" num="s">
  <n [...] num="s" postype="common" wd="secretari"/>
  <s.a gen="m" num="s">
    <grup.a gen="m" num="s">
      <a [...] num="s" postype="qualificative" wd="general"/>
    </grup.a>
  </s.a>
</grup.nom>
<sp>
  <prep>
    <s gen="c" lem="de" num="c" postype="preposition" wd="d'"/>
  </prep>
  <sn>
    <grup.nom>
      <n gen="c" lem="ERC" num="c" postype="proper" wd="ERC"/>
    </grup.nom>
  </sn>
</sp>
<sn>
  <f lem="," punct="comma" wd=","/>

```



```

    <grup.nom gen="m" num="s">
      <n [...] num="c" postype="proper" wd="Carod-Rovira"/>
    </grup.nom>
    <f lem="," punct="comma" wd=","/>
  </sn>
</grup.nom>
</sn>
‘The ERC party general secretary, Carod-Rovira’

```

2.4.1 Noun phrase (sn)

The structure of the noun phrase is the one shown below:

```

<sn>
  <spec gen="X" num="Y">
    <d [...] />
  </spec>
  <grup.nom gen="X" num="Z">
    <complements [...]>
      <[head]> />
    <complments [...]>
  </grup.nom>
</sn>

```

Specifiers and complements are optional elements. Complements can appear before or after the head and the way to represent them is completely flat: When a head has one or more complements, these are sisters of the head, even though they could be given a different semantic interpretation¹ 2.8

(2.8) *punt [de vista] [de la mobilitat]* (Catalan 107-5)
 ‘point of view of mobility’

The head of the noun phrase may be a noun, a pronoun, a number or date, or any other noun-like element. The phrase can vary as regards complexity, from very simple structures consisting of one element (the head) to a complex one with several complements or coordinated structures).

A particular case of noun phrase is that of the covert subject, used in those finite sentences in which the subject is not explicit. The correspondent XML marking is the following:

```

<attribute name="elliptic" onlyterminal="yes"
description="Is the constituent ellipted?">
  <value name="yes" />
  <value name="no" />
</attribute>

```

¹See section 2.3.1 for more details on stratification at the noun phrase.

2.4.1.1 The head

1. Common and proper nouns

Regarding the morphological features (gender and number) of proper nouns:

For names of persons, the label corresponding to morphological features must be specified (`<grup.nom gen='f' num='s'>`, `<grup.nom gen='m' num='s'>`).

For the rest of proper names, we add gender and number when it is possible to infer this information from the context. Otherwise we add to the `<grup.nom>` node no gender and number features.

Regarding common nouns, the `<grup.nom>` node must contain gender and number information, which is derived from the morphological (`<n>`) node :

```
<grup.nom gen="m" num="s">
  <n gen="c" lem="cap" num="s" postype="common" wd="cap"/>
```

Some nouns that are common at gender (`<n gen='c' lem='fiscal' num='s' postype='common' wd='fiscal'/>`) are assigned gender when it comes with specifiers or complements which disambiguate this feature. When there are no complements or they do not give information on gender, both complement and noun are assigned the masculine gender by default.

2. The pronoun

In the case of the pronouns, the node `<grup.nom>` must always be added. When complemented by specifiers or adjectives, it incorporates the features for gender and number of these complements. Otherwise:

When the information on gender and number is given by the morphological node, it should be added to the group node:

```
<grup.nom gen="f" num="p">
  <p gen="f" lem="algun" num="p" postype="indefinite" wd="algunes"/>
```

When the pronoun does not imply gender, it is specified as neuter:

```
<p gen="c" lem="en" num="c" person="3" postype="personal" wd="n"/>
```

Those forms such as *algú* ‘someone’, *allò* ‘that, it’, *ho* (neutral accusative pronoun), *res* ‘nothing’, *ningú* ‘nobody’, *això* ‘this’, and all neutral pronouns must be labeled as `<grup.nom gen='m' num='s'>` (*alguien* ‘someone’, *algo* ‘something’, *nada* ‘nothing’, *nadie* ‘nobody’ for Spanish).

3. Numbers and dates

Morphological features on gender and number should be added depending on the context:

```

<spec gen="f" num="p">
  <z lem="8" wd="8"/>
</spec>
<grup.nom gen="f" num="p">
  <n gen="f" lem="zona" num="p" postype="common" wd="zones"/>

```

‘8 zones’

In case of unknown gender, it is marked as common (<gen=‘c’>).

In dates², the days of the week, months and years are labeled as <grup.nom gen=‘f’ num=‘p’> although there are no specifiers or complements, but if there are any, the morphological features should be those given by the specifier or complement.

The hours of the day also indicate gender and number only when going along with specifiers or complements. Hours and days, months and/or years must be splitted up when appearing together. The analysis given by the parser is:

```

<spec gen="f" num="p">
  <d [...] wd="les"/>
</spec>
<grup.nom gen="f" num="p">
  <w [...] wd="sis_de_la_matinada_del_dilluns_4_de_juny"/>

```

and the correct representation, which is done manually by the annotator, should be:

```

<spec gen="f" num="p">
  <d [...] wd="les"/>
</spec>
<grup.nom gen="f" num="p">
  <w [...] wd="sis_de_la_matinada"/>
  <sp>
    <prep>
      <s [...] wd="del"/>
    </prep>
  <sn>
    <grup.nom gen="m" num="s">
      <w [...] wd="dilluns_4_de_juny"/>

```

‘six o’clock in the morning Monday June the 4th’

²Dates are marked as <w> at the morphological node. See section 2.9.2 for further details on time expressions.

4. Nominalized element

A nominalized element (adjective, participle, prepositional phrase or relative clause) can also constitute a noun phrase.

When the nominalized element is an adjective, the morphological node is maintained, while the group label has to be replaced by `<grup.nom>` (2.9):

(2.9) (cat 06019900-4)

```

<sp>
  <prep>
    <s [...] postype="preposition" wd="dels"/>
  </prep>
  <sn>
    <spec gen="m" num="p">
      <d gen="m" lem="dos" num="p" postype="numeral" wd="dos"/>
    </spec>
    <grup.nom gen="m" num="p">
      <a [...]num="p" postype="ordinal" wd="primers"/>
    </grup.nom>
  </sn>
</sp>
‘of the two first ones’

```

When the nominalized element corresponds to a participle (2.10), relative clause or prepositional phrase, both nodes (morphological and group) are maintained. In addition to this, we should add the nodes corresponding to a noun phrase; that is `<grup.nom>` and `<sn>`³.

(2.10) (Cat 111nou)

```

<sn>
  <spec gen="f" num="s">
    <d gen="c" lem="el" num="s" postype="article" wd="l'"/>
  </spec>
  <grup.nom gen="f" num="s">
    <S clausetype="participle" impersonal="yes">
      <participi>
        <a [...] posfunction="participle" wd="aprovada"/>
      </participi>
    <sadv>
      <grup.adv>
        <r lem="avui" wd="avui"/>
      </grup.adv>
    </S>
  </grup.nom>
</sn>

```

³See section 2.4.1.5 for more details on nominalization.

```

    </sadv>
  </S>
</grup.nom>
</sn>
‘the one aproved today’

```

2.4.1.2 Determiners

Determiners are marked as `<espec>` and usually occur before the head (2.11), although they can also appear after it (2.12):

(2.11) (Cat 14770_20000907)

```

<sn>
  <spec gen="f" num="s">
    <d [...] postype="possessive" wd="la_nostra"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n gen="f" lem="comunitat" num="s" postype="common" wd="comunitat"/>
  </grup.nom>
</sn>
‘our community’

```

(2.12) (Cat 73_20010801)

```

<sn>
  <grup.nom gen="f" num="s">
    <n gen="f" lem="casa" num="s" postype="common" wd="casa"/>
  </grup.nom>
  <spec gen="f" num="s">
    <d [...] num="s" person="3" postype="possessive" wd="seva"/>
  </spec>
</sn>
‘his/her house’

```

When two or more determiners appear together, we arrange them in a flat structure (2.13). When determiners are coordinated, however, the representation has a complete structure (2.14):

(2.13) <sn>

```

  <spec gen="f" num="p">
    <d gen="f" lem="el" num="p" postype="article" wd="les"/>
    <d [...] num="p" postype="indefinite" wd="diferents"/>

```

```

</spec>
<grup.nom gen="f" num="p">
  <n [... num="p" postype="common" wd="comunitats"/>
‘the different communities’

```

(2.14) (Cat 211nou)

```

<spec coord="yes" gen="m" num="s">
  <spec gen="m" num="s">
    <d gen="c" lem="el" num="s" postype="article" wd="l'"/>
    <d gen="m" lem="un" num="s" postype="indefinite" wd="un"/>
  </spec>
  <coord>
    <c lem="o" postype="coordinating" wd="o"/>
  </coord>
  <spec gen="m" num="s">
    <d gen="c" lem="el" num="s" postype="article" wd="l'"/>
    <d gen="m" lem="altre" num="s" postype="indefinite" wd="altre"/>
  </spec>
</spec>
‘the one or the other’

```

There exist some special directions as regards determiners for Spanish annotation. Some quantifiers, for instance, have a special lexicalized structure that must be joined as a single string of text:

(2.15) (Sp a1-1)
hacia [[Dios_sabe_qué <spec>] [reinos de figurones]]
‘to who-knows-what figure kingdoms’

Masculine determiners introduce feminine nouns starting with stressed ‘a’⁴:

(2.16) (Sp 5_20020202)
el [hacha]
‘the axe’

1. Flat representation of complex determiners

Structures such as those of *més_de* ‘more than’ + number, *gairebé* ‘almost’ + number, etc. should be analyzed in a flat representation, although maintaining their morphological nodes.

These structures are given the same gender and number features the noun group has, as in 2.17, although there are some exceptions, as happens in 2.18:

⁴Because of phonological reasons, those feminine nouns with a word-initial tonic [a] are specified by masculine determiners.

(2.17) (Cat 06010100-39)

```

<sn>
  <spec gen="f" num="s">
    <d gen="f" lem="el" num="s" postype="article" wd="la"/>
    <n gen="f" lem="resta" num="s" postype="common" wd="resta"/>
    <s gen="c" lem="de" num="c" postype="preposition" wd="de"/>
    <d gen="f" lem="el" num="s" postype="article" wd="la"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n gen="f" lem="família" num="s" postype="common" wd="família"/>
  'the rest of the family'

```

(2.18) (Cat 06010300-27)

```

<sn>
  <spec gen="f" num="s">
    <d gen="f" lem="el" num="s" postype="article" wd="la"/>
    <n gen="f" lem="resta" num="s" postype="common" wd="resta"/>
    <s gen="c" lem="de" num="c" postype="preposition" wd="de"/>
  </spec>
  <grup.nom>
    <n [...] num="c" postype="proper" wd="Catalunya"/>
  'the rest of Catalonia'

```

The following are some of the complex determiners that are found in the corpus:

- *més_de* ‘more than’, *menys_de* ‘less than’ + number (*más_de*, *menos_de* in Spanish)
- *gairebé* ‘almost’, *fins_a* ‘to, until’, *al_voltant_de* ‘arround’, *prop_de* ‘near’... + number
- *un total de* ‘a total amount of’, *un mínim de* ‘a minimum of’, *un màxim de* ‘a maximum of’... + number
- *els últims* ‘the latter’, *els primers* ‘the former’, *els pròxims* ‘the following’, *els propers* ‘the next’... + number

When the adjective occurs immediately before the noun group, it is not part of the specifier but of the nominal group 2.19:

(2.19) (Cat 207-7) *[[els cinc_ <spec>][primers_ <adj>] mesos_ <sn>]*
 ‘the first five months’

- number + *per* + number

(2.20) (Cat 103-20) *[[5 per 3,5_ <spec>] metres]* ‘5 meters long by 3,5 meters wide’

- *una mica de* ‘a little of’, *la resta de* ‘the rest of’, *la majoria de* ‘most of’... (2.21):

(2.21) (Cat 211nou) *[[una mica de_ <spec>] retard]*
‘a little delay’

2. Més/menys; más/menos

The adverbs *més* ‘more’ and *menys* ‘less’ (*más* and *menos* in Spanish) should be marked as specifiers (<spec>) when followed by a noun group and therefore they are part of a noun phrase (2.22):

(2.22) (Cat 06010100-8) *[més [competències normatives]]*
‘more regulatory competences’

But they can also function as adverbial phrases when followed by an adverb or an adjective (2.23):

(2.23) (Cat 104nou) *[[[més] a_prop] [de les zones]]*
‘closer to the zones’

Both cases are given the morphological node <rg>.

When *menys* and *menos* stand for *excepte* ‘except’ and they are followed by a noun phrase, they function as a preposition <s> (2.24):

(2.24) (Cat 111nou) *Tots els grups del Parlament, [[menys] el PP]]*
‘All parliamentary groups but the PP’

3. Tot

The adverb *tot* ‘everything’ functions as a specifier when it agrees with the noun it introduces, as shown in 2.25. It works as an adverb when there is no agreement (2.26):

(2.25) (Cat 19202_20001110)

```
<s gen="c" lem="de" num="c" postype="preposition" wd="de"/>
</prep>
<sn>
  <spec gen="f" num="s">
    <d gen="f" lem="tot" num="s" postype="indefinite" wd="tota"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n [...] num="c" postype="proper" wd="Catalunya"/>
  </grup.nom>
‘of whole Catalonia’
```


(2.26) (Cat 06010300-27)

```

    <s gen="c" lem="a" num="c" postype="preposition" wd="a"/>
  </prep>
  <sn>
    <grup.nom>
      <sadv>
        <grup.adv>
          <r lem="tot" wd="tot"/>
        </grup.adv>
      </sadv>
    <n [...] num="c" postype="proper" wd="Catalunya"/>
  ‘in all Catalonia’

```

4. Numbers

When working out as specifiers, numbers are given gender and number values depending on the noun they introduce (2.27):

(2.27) (Sp a2-0)

30 flexiones
 ‘30 press-ups’

Percentages (<Zp>) can also work out as specifiers (2.28):

(2.28) (Cat 06019900-21)

70.7/100.000 habitants
 ‘70.7/100.000 inhabitants’

See Appendix for a full list of locutions and complex determiners.

2.4.1.3 Complements of the noun phrase

The elements that can occur as complements within the noun phrase are the adjective phrase, the prepositional phrase, the adverbial phrase, other noun phrases, relative clauses, adjectival non-finite clauses, complement clauses, adverbial clauses, sentences, or the combination of all of them.

These elements are represented in a flat way instead of hierarchical: every complement is the daughter of the node <grup.nom> and sisters of the head, even though there could exist a different semantic interpretation (2.29):

(2.29) (Cat 107-5)

punt [de vista] [de la mobilitat]
 ‘point of view of the mobility’

Only in those cases in which a complement belongs to two or more coordinated nouns, the complement is adjoined to the coordinated element⁵ (2.30):

⁵See section 2.6 for more details on coordination.

(2.30) (Cat 111nou)

[[[retorn i dipòsit<coord>]<coord>]<complex>] [a l'Arxiu_Nacional de Catalunya<sp> <adjunct>]

'return and deposit to Catalonia National Archive'

1. Adjective phrase and non-finite participle clause

Within a nominal phrase, the label of an adjective phrase must include the specifications on gender and number⁶. Adjectives (2.31) and participle clauses (2.130), in their turn, can have their own complements:

(2.31) (Cat 204nou)

```
<sn>
  <grup.nom gen="f" num="s">
    <s.a gen="f" num="s">
      <grup.a gen="f" num="s">
        <a [...] num="s" postype="qualificative" wd="màxima"/>
      </grup.a>
    </s.a>
    <n [...] num="s" postype="common" wd="afluència"/>
    <s.a gen="f" num="s">
      <grup.a gen="f" num="s">
        <a [...] num="s" postype="qualificative" wd="turística"/>
      </grup.a>
    </s.a>
  </grup.nom>
</sn>
```

'maximum turistic inflow'

(2.32) (Cat 06010100-14)

```
<grup.nom gen="m" num="s">
  <n [...] num="s" postype="common" wd="informe"/>
  <S clausetype="participle" impersonal="yes">
    <participi>
      <a [...] posfunction="participle" wd="encarregat"/>
    </participi>
  <sp>
    <prep>
      <s [...] postype="preposition" wd="per"/>
    </prep>
  <sn>
    <spec gen="f" num="s">
```

⁶See section 2.4.2 for more details on adjective phrase (<sa>).

```

    <d [...] num="s" postype="article" wd="la"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n [...] postype="common" wd="corporació"/>

```

‘report entrusted by the corporation’

2. Adjectivised noun

Sometimes the head of the adjective phrase can be a noun⁷, which should maintain its morphological node (2.33 and 2.34). We treat this way all nouns when there is no appositive or identity relationship between both elements but the second noun specifies the meaning of the first one, usually restricting its type (in the examples below, for instance, *el tipus de conveni* ‘the type of agreement’, *el tipus de pàgina* ‘the type of page’, etc.) or identifying a particular case of this noun. Moreover, none of these structures can turn into copulative structures (**la pàgina és web* ‘the page is web’, **el benefici és rècord* ‘the benefit is record’).

(2.33) (Cat 211-nou)

```

<sn>
  <spec gen="m" num="s">
    <d gen="m" lem="el" num="s" postype="article" wd="el"/>
  </spec>
  <grup.nom gen="m" num="s">
    <n [...] num="s" postype="common" wd="conveni"/>
    <s.a gen="m" num="s">
      <grup.a gen="m" num="s">
        <n gen="m" lem="marc" num="s" postype="common" wd="marc"/>
      </grup.a>
    </s.a>

```

‘the agreement framework’

(2.34) (Cat 111nou)

```

<grup.nom gen="f" num="s">
  <n gen="f" lem="pàgina" num="s" postype="common" wd="pàgina"/>
  <s.a gen="f" num="s">
    <grup.a gen="f" num="s">
      <n gen="c" lem="web" num="s" postype="common" wd="Web"/>
    </grup.a>
  </s.a>

```

‘website (lit., *page web*)’

⁷See section 2.4.2.4 for further details on adjectivisation

Sometimes, a number or date can also function as the head of an adjective phrase (2.35):

(2.35) (Cat 06010100-23)

```
<n gen="m" lem="curs" num="s" postype="common" wd="curs"/>
<s.a gen="m" num="s">
  <grup.a gen="m" num="s">
    <a [...] num="s" postype="qualificative" wd="escolar"/>
  </grup.a>
</s.a>
<s.a gen="m" num="s">
  <grup.a gen="m" num="s">
    <w lem="[?:?:??:??:2000/2001:?:?.??:]" wd="2000/01"/>
  </grup.a>
</s.a>
'scholar course 2000/01'
```

3. Prepositional phrase (<sp>)

(2.36) (Cat 211nou) *agents antiavalots [de la policia municipal]*
 'riot agents from municipal police'

4. Noun phrase (<sn>)

A noun directly complements another noun in the case of appositions. In appositive structures there is a referential identity between the head and the apposition; if there is no identity, there is no apposition (2.37):

(2.37) [*L'alcalde de Barcelona, [Joan_Clos]*] (Cat 211nou)
 'The major of Barcelona, Joan_Clos'

A noun phrase of the kind *cadascun* 'each' can also complement another noun phrase, although there is no appositive relation 2.38.

(2.38) [*3 euros [la unitat]*] (Cat 06030200-1)
 '3 euros each'

The forms *senyor/a* 'mister, miss', *sant/a* 'saint' (Catalan); and *señor/a*, *santo/a* and *don/doña* 'mister, miss' (Spanish) are nouns, but they always occur before the noun they complement:

(2.39) (Sp t5-1))

```

<sn>
  <grup.nom gen="m" num="s">
    <sn>
      <grup.nom gen="m" num="s">
        <n gen="m" lem="don" num="s" postype="common" wd="don"/>
      </grup.nom>
    </sn>
    <n [...] num="c" postype="proper" wd="Antonio"/>
  </grup.nom>
</sn>
‘Mr. Antonio’

```

5. Adverbial phrase (<sadv>)

(2.40) (Cat 111nou)

```

<sn>
  <grup.nom gen="m" num="s">
    <p gen="c" lem="res" num="s" postype="indefinite" wd="res"/>
  <sadv>
    <grup.adv>
      <r lem="més" wd="més"/>
    </grup.adv>
  </sadv>
</grup.nom>
</sn>
‘nothing else’

```

6. Finite and non-finite relative clause

(2.41) (Cat 211nou)

```

<grup.nom gen="f" num="s">
  <n gen="f" lem="setmana" num="s" postype="common" wd="setmana"/>
  <S clausetype="relative" impersonal="no">
    <relatiu>
      <p gen="c" lem="que" num="c" postype="relative" wd="que"/>
    </relatiu>
  <grup.verb>
    <v [...] postype="main" tense="present" wd="ve"/>
  </grup.verb>
</S>
‘next week’

```

7. Finite completive clause

(2.42) (Cat 111nou)

```

<n gen="m" lem="fet" num="s" postype="common" wd="fet"/>
  <S clausetype="completive" impersonal="no">
    <conj conjunctiontype="subordinating">
      <c lem="que" postype="subordinating" wd="que"/>
    </conj>
    <sn/>
    <grup.verb>
      <v [...] postype="main" tense="present" wd="parteixi"/>
    </grup.verb>
    <sp>
      <prep>
        <s [...] postype="preposition" wd="del"/>
      </prep>
      <sn>
        <grup.nom gen="m" num="s">
          <n gen="c" lem="PSC" num="c" postype="proper" wd="PSC"/>

```

‘the fact that it comes from the PSC party’

8. Finite and non-finite adverbial clause

(2.43) (Cat 211-8) *la inhalació de gasos tòxics [quan estaven a l'interior d'un pou]*

‘the inhalation of toxic fumes when (they) were inside a shaft’

9. Sentence (<S>)

A whole sentence structure can complement a noun as an inserted element (<inc>)⁸
(2.44):

(2.44) (Cat 06010100-34)

Davant unes mil persones, [entre elles la ministra de Ciència i Teconologia, Anna_Birulés, l'alcalde de Barcelona, Joan_Clos, la Delegada del Govern, Julia_Garcia_Valdecasas i una nodrida representació del govern català <inc>], Pujol va donar un toc d'alerta...

‘In front of about a thousand people, among them the minister for Science and Technology, Anna_Birulés, the major of Barcelona, Joan_Clos, the delegate of the Government, Julia_Garcia_Valdecasas and a nourished representation of the catalan Government’

10. Combination of complements

⁸See section 3.2.1 for further details on inserted elements

(2.45) (Cat 103-12)

el primer premi Pentagrama [de composició <sp>], [obert a tots els estats i per a músics...] <participle>

‘the first Pentagrama award on composition, open to all states and for musicians...’

2.4.1.4 Adjunction to the noun phrase

There are other noun complements with the shape of a noun phrase that cannot be treated as appositions and cannot be adjectivised. This is so because the coreference condition is not satisfied or because they do not restrict the meaning of the head. In such cases, we should adjoin the second noun phrase to the first (2.46):

(2.46) </sn>

```

<sn>
  <grup.nom>
    <n [...] num="c" postype="proper" wd="Amposta"/>
  </grup.nom>
</sn>
<sn adjunct="yes">
  <grup.nom>
    <f [...] punctenclose="open" wd="("/>
    <n [...] postype="proper" wd="El_Montsià"/>
    <f [...] punct="bracket" punctenclose="close" wd=")"/>
  </grup.nom>
</sn>

```

When more than one of the above-mentioned noun phrase coocur, we group them and label the upper node as <verbless>. This node should be adjoined to the noun phrase node (<sn>) which becomes doubled by doing so, as shown in 2.47:

(2.47) (Cat 103-1)

```

<sn>
  <sn>
    <grup.nom gen="m" num="s">
      <n [...] postype="proper" wd="Robert_Brufau"/>
    </grup.nom>
  </sn>
  <S adjunct="yes" verbless="yes">
    <f lem="(" punct="bracket" punctenclose="open" wd="("/>
    <sn>
      <grup.nom>
        <n [...] postype="proper" wd="Mollerussa"/>
      </grup.nom>
    </sn>
  </S>
</sn>

```

```

</sn>
<f lem="," punct="comma" wd=","/>
<sn>
  <grup.nom>
    <w lem="[??:??/?/1946:??.??]" wd="1946"/>
  </grup.nom>

```

Addresses are also adjoined to noun phrases (2.48):

(2.48) (Cat 06030200-25)

```

</sn>
<sn adjunct="yes">
  <f lem="(" punct="bracket" punctenclose="open"wd="("/>
  <sn>
    <grup.nom>
      <n [...] postype="proper" wd="Casa_Montaner"/>
    </grup.nom>
  </sn>
  <sn>
    <f lem="," punct="comma" wd=","/>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="plaça"/>
      <sp>
        <prep>
          <s [...] gen="m" lem="del" num="p" wd="dels"/>
        </prep>
      <sn>
        <grup.nom gen="m" num="p">
          <n [...] postype="proper" wd="Països_Catalans"/>
        </grup.nom>
      </sn>
    </sp>
  </sn>
  <grup.nom>
    <n [...] num="c" postype="common" wd="s/n"/>

```

Prepositional phrases like the following, are also adjoined to the noun phrase they belong to (2.49):

(2.49) (Cat 111nou)

[tots els grups del Parlament [[llevat _ del PP_ <adjunct>]] ‘all groups of the Parliament except the PP’

We also adjoin those adverbs that do not depend on the whole sentence but affect just the noun phrase. They can only occur in that specific position (2.50) or before the verb provided that they appear between commas (2.51):

(2.50) (Cat 204-5)
 [[*la ciutat de Berga*] i [[*la comarca del Berguedà*] [*en_ general_ <adjunct>*]]]
 ‘the city of Berga and the Berguedà region in general’

(2.51) (Cat 06010100-9)
 ...*la poca repercussió que tenen els mitjans [...] i*[[, *especialment_ <adjunct>*],
l'escassa repercussió directa que tenen les fires turístiques]
 ‘...the little repercussion that media has [...] and, specially, the little direct
 repercussion that tourism fairs have’

2.4.1.5 Nominalization

A nominalized element can also constitute a noun phrase.

When the nominalized element is an adjective, the morphological node is maintained, but the group node corresponds to a noun one (<grup.nom>) as well as the phrase one (<sn>), as shown in (2.52):

(2.52) (Cat 06029900-9)

```
<sn>
  <spec gen="m" num="s">
    <d gen="m" lem="el" num="s" postype="article" wd="el"/>
  </spec>
  <grup.nom gen="m" num="s">
    <a [...] num="s" postype="ordinal" wd="primer"/>
  </grup.nom>
</sn>
```

‘the first one’

A particular case is that of nominalized participles: In such a case, the morphological node is maintained, but the nodes <grup.nom> and <sn> should be added (2.53). In this way, we keep the structure within the participle clause.

(2.53) (Cat 111nou)

```
<c lem="com" postype="subordinating" wd="com"/>
</prep>
<sn>
  <spec gen="f" num="s">
    <d gen="c" lem="el" num="s" postype="article" wd="l'"/>
```

```

</spec>
<grup.nom gen="f" num="s">
  <S clausetype="participle" impersonal="yes">
    <participi>
      <a [...] posfunction="participle" wd="aprovada"/>
    </participi>
    <sadv>
      <grup.adv>
        <r lem="avui" wd="avui"/>
      </grup.adv>
    </sadv>
  </S>
</grup.nom>
</sn>

```

‘as the one approved today’

When the nominalized element is a prepositional phrase, we add the node `<grup.nom>` as well as `<sn>` so we can represent its complements (2.54):

(2.54) (104nou)

```

<sn>
  <spec gen="m" num="s">
    <d gen="m" lem="el" num="s" postype="article" wd="el"/>
  </spec>
  <grup.nom gen="m" num="s">
    <sp>
      <prep>
        <s [...] num="c" postype="preposition" wd="de"/>
      </prep>
      <sn ">
        <spec gen="f" num="s">
          <d gen="c" lem="el" num="s" postype="article" wd="l'"/>
        </spec>
        <grup.nom gen="f" num="s">
          <n [...] postype="common" wd="avinguda"/>
        <sn>
          <grup.nom>
            <n [...] postype="proper" wd="Santa_Coloma"/>

```

‘the one at Santa_Coloma avenue’

When the nominalized element is any other phrase or structure, we maintain the label, as shown in example 2.55.

(2.55) (Cat 211nou)

```

<sn>
  <spec gen="m" num="s">
    <d gen="m" lem="el" num="s" postype="article" wd="el"/>
  </spec>
  <grup.nom gen="m" num="s">
    <S clausetype="relative" impersonal="no">
      <relatiu>
        <p gen="c" lem="que" num="c" postype="relative" wd="que"/>
      </relatiu>
      <sadv>
        <grup.adv>
          <r lem="menys" pos="rg" wd="menys"/>
        </grup.adv>
      </sadv>
    </S>
  </grup.nom>
</sn>
  <grup.nom num="s">
    <p [...] postype="personal" wd="em"/>
  </grup.nom>
</sn>
  <grup.verb>
    <v gen="c" [...] tense="present" wd="preocupa"/>
  </grup.verb>

```

‘what worries me the least’

Other cases of nominalization are those structures which usually involve emphatic expressions with adjectives but also with adverbs. Any other kind of nominalization will keep its morphological node.

When nominalization is introduced by the contracted forms (preposition + determiner) *al* ‘to the’, *del* ‘from the’, *pel* ‘to the’, the node corresponding to the specifier, which indicates nominalization, is lost (2.56). In such cases we should add the node <sn>:

(2.56) (Cat 104nou)

```

  <d gen="f" lem="el" num="s" postype="article" wd="la"/>
</spec>
<grup.nom gen="f" num="s">
  <n [...] num="s" postype="common" wd="qualitat"/>
  <sp>
    <prep>
      <s [...] postype="preposition" wd="del"/>
    </prep>
  </grup.nom>

```

```

<sn>
  <sp>
    <prep>
      <s [...] num="c" posttype="preposition" wd="de"/>
    </prep>
    <sn>
      <spec gen="f" num="s">
        <d [...] num="s" posttype="article" wd="la"/>
      </spec>
      <grup.nom gen="f" num="s">
        <n [...] num="s" posttype="common" wd="capital"/>
      </grup.nom>

```

‘the quality of the capital’

2.4.1.6 Other elements treated as noun phrases

In certain temporal expressions, there is a third person singular verb which does not function as the sentence head (*fa dos anys...* ‘two years ago’ (*hace dos años* in Spanish). The whole structure is analyzed as a single noun phrase representation without structure (2.57):

```

(2.57) <sn>
  <v [...] tense="present" wd="fa"/>
  <d [...] posttype="numeral" wd="sis"/>
  <n [...] posttype="common" wd="mesos"/>
</sn>

```

‘six months ago’

When this structure is followed by *que* ‘that’, this subordinate conjunction is not part of the noun and does not introduce a clause; it directly depends on the upper sentence node (2.58).

(2.58) (Cat 06010300-13)

```

<sn>
  <r lem="ara" wd="Ara"/>
  <v [...] posttype="main" tense="present" wd="fa"/>
  <d [...] posttype="numeral" wd="un"/>
  <n [...] posttype="common" wd="any"/>
</sn>
  <conj conjunctiontype="subordinating">
    <c lem="que" posttype="subordinating" wd="que"/>
  </conj>

```

‘It is one year ago since ...’

Expressions such as those of number + *km/h* are represented as if they were specifiers followed by a noun group including *km/h* (2.59):

(2.59) (Cat 104nou)

```
<sn>
  <spec gen="m" num="p">
    <z lem="350" wd="350"/>
  </spec>
  <grup.nom gen="m" num="p">
    <n gen="m" lem="quilòmetre" num="c" postype="common" wd="km"/>
    <f lem="/" punct="slash" wd="/" />
    <n gen="f" lem="hora" num="c" postype="common" wd="h"/>
  </grup.nom>
</sn>
```

2.4.1.7 Coordination within the noun phrase

The criterion is to coordinate at the lower possible level, so we should coordinate at noun group when possible. When noun group coordination is possible, the upper phrase node (<sn>) is given marked as <complex>, indicating that it contains a coordination (2.60). See section 2.6 for more details on coordination.:

(2.60) (Cat 107-22)

```
<sn complex="yes">
  <grup.nom coord="yes">
    <grup.nom gen="f" num="p">
      <n [...] num="p" postype="common" wd="mesures"/>
    </grup.nom>
    <coord>
      <c lem="i" postype="coordinating" wd="i"/>
    </coord>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="claredat"/>
    </grup.nom>
  </grup.nom>
</sn>
```

‘measures and clarity’

When there are complements or specifiers affecting differently the coordinated groups, the coordination is marked at the noun phrase node. In example 2.61, for instance, each noun phrase is introduced by a specifier, for the coordination at the <grup.nom> node is not possible:

(2.61) (Cat 111nou)

```

<sn coord="yes">
  <sn>
    <spec gen="m" num="s">
      <d gen="m" lem="el" num="s" postype="article" wd="el"/>
    </spec>
    <grup.nom gen="m" num="s">
      <n gen="c" lem="PP" num="c" postype="proper" wd="PP"/>
    </grup.nom>
  </sn>
  <coord>
    <c lem="i" postype="coordinating" wd="i"/>
  </coord>
  <sn>
    <spec gen="m" num="s">
      <d gen="m" lem="el" num="s" postype="article" wd="el"/>
    </spec>
    <grup.nom gen="m" num="s">
      <n gen="c" lem="PNB" num="c" postype="proper" wd="PNB"/>
    </grup.nom>
  </sn>
</sn>

```

‘the PP and the PNB’

2.4.1.8 Noun locutions

Noun expressions derived from verbal locutions are treated as noun locutions: *posada_en_marxa*, *posada_en_funcionament* ‘set up’ (*puesta_en_marcha* and *puesta_en_funcionamiento* in Spanish).

We also consider as noun locutions those expressions such as: *web_cam*, *tot_terreny* ‘all-terrain’, *dirt_track*, *benji_jumping*, *medi_ambient* ‘environment’, *mosso_d’esquadra* (Catalan police), *alts_i_baixos* ‘ups and downs’, etc.

See Appendix for a complete list of noun locutions.

2.4.2 Adjective phrase and participle clause

The adjective phrase has an adjective in its head which, in its turn, can be complemented by other elements. We distinguish among <sa>, <s.a>, and non-finite adjective phrases (participle and absolute participle clauses⁹).

⁹See section 2.4.2 and section 2.5.1.2 for further details on participle clauses.

2.4.2.1 Argumental adjective phrase (<sa>)

The node (<sa>) does not include gender or number specifications and it is used in those cases of adjective phrases which directly depend on a sentence node. Only this kind of adjective phrase can correspond to an attribute as well as a predicative complement.

(2.62) (Cat 204nou)

```

<grup.verb>
  <v [...] postype="main" tense="present" wd="considera"/>
</grup.verb>
<sa func="cpred" predicate="cd" tem="atr">
  <grup.a>
    <a [...] num="s" postype="qualificative" wd="necessari"/>
  </grup.a>
</sa>
<S clausetype="completive" impersonal="yes" tem="pat">
  <infinitiu>
    <v [...] postype="main" wd="ampliar"/>
  </infinitiu>
  <sn>
    <spec gen="f" num="s">
      <d gen="f" lem="el" num="s" postype="article" wd="la"/>
    </spec>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="plantilla"/>

```

‘(s/he) considers it necessary to enlarge the staff’

2.4.2.2 Adjective phrase as a noun complement (<s.a>)

Unlike the node <sa>, the adjective phrase <s.a> depends on a noun group (although there are some exceptions) and it is therefore assigned gender and number values:

(2.63) (Cat 211nou)

```

  <d gen="c" lem="el" num="s" postype="article" wd="l'"/>
</spec>
<grup.nom gen="m" num="s">
  <n gen="m" lem="any" num="s" postype="common" wd="any"/>
  <s.a gen="m" num="s">
    <grup.a gen="m" num="s">
      <a [...] num="s" postype="qualificative" wd="vinent"/>
    </grup.a>
  </s.a>

```

‘next year’

In addition to that, deverbal adjectives (participles), which are the head of a non-finite participle clause, may not be correctly marked by the automatic parsing. In such a case, the corresponding nodes must be added manually.

2.4.2.3 Complements of the adjective phrase

Adjective phrases can have the following complements:

1. Adverbial phrase (<sadv>)¹⁰

Although they keep their morphological node (<r>), adverbs affecting adjective phrases are marked as specifiers, as shown in 2.64. When affecting non-finite adjective clauses (participles), however, they belong to the <S> node (2.65)

(2.64) (Cat 111nou)

```
<grup.a gen="f" num="p">
  <a [...] postype="qualificative" wd="desafortunades"/>
</grup.a>
</s.a>
<coord>
  <c lem="i" postype="coordinating" wd="i"/>
</coord>
<s.a gen="f" num="p">
  <spec>
    <r lem="fins_i_tot" wd="fins_i_tot"/>
  </spec>
  <grup.a gen="f" num="p">
    <a [...] postype="qualificative" wd="desmesurades"/>
```

‘unfortunated and even excessive’

(2.65) (Cat 111nou)

```
<S clausetype="participle" impersonal="yes">
  <sadv>
    <grup.adv>
      <r lem="ben" wd="ben"/>
    </grup.adv>
  </sadv>
  <participi>
    <a [...] posfunction="participle" wd="lligat"/>
  </participi>
</S>
```

¹⁰See section 2.4.3 for more details on adverbial phrase.

‘well binded’

2. The adverb ‘no’ (neg)

Instead of being adjoined, the negative adverb *no* depends directly on the <sa> / <s.a> node (2.247):

(2.66) (Cat 104nou)

```
<n [...] num="p" postype="common" wd="autocars"/>
<s.a gen="m" num="p">
  <neg>
    <r lem="no" postype="negative" wd="no"/>
  </neg>
  <a [...] num="p" postype="qualificative" wd="urbans"/>
</s.a>
‘non-urban buses’
```

3. Prepositional phrase

(2.67) (Cat 111nou)

```
<r lem="més" pos="rg" wd="més"/>
</spec>
<grup.a gen="f" num="s">
  <a [...] num="s" postype="qualificative" wd="gran"/>
  <sp>
    <prep>
      <s [...] num="c" postype="preposition" wd="de"/>
    </prep>
    <sn>
      <spec gen="f" num="s">
        <d gen="f" lem="el" num="s" postype="article" wd="la"/>
      </spec>
      <S clausetype="relative" impersonal="no">
        <relatiu>
          <p [...] num="c" postype="relative" wd="que"/>
        </relatiu>
      </S>
    </sn>
  </sp>
</grup.a>
‘bigger than’
```

4. Adjective phrase

(2.68) (Cat 06010300-5)

```
<S clausetype="participle" impersonal="yes">
  <participi>
```

```

    <a [...] posfunction="participle" wd="ferits"/>
  </participi>
  <sa complex="yes">
    <grup.a coord="yes">
      <grup.a>
        <a [...] postype="qualificative" wd="greu"/>
      </grup.a>
      <coord>
        <c lem="i" postype="coordinating" wd="i"/>
      </coord>
      <grup.a>
        <a [...] postype="qualificative" wd="lleu"/>
      </grup.a>
    </grup.a>
  </sa>
  <sadv>
    <grup.adv>
      <r lem="respectivament" wd="respectivament"/>

```

‘serious and slight injureds, respectively’

5. Adjoined noun phrases

(2.69) (Cat 06019900-14)

```

  <sn adjunct="yes" coord="yes">
    <coord>
      <s [...] postype="preposition" wd="entre"/>
    </coord>
    <sn>
      <spec gen="m" num="s">
        <d [...] num="s" postype="indefinite" wd="un"/>
      </spec>
      <grup.nom gen="m" num="s">
        <z lem="5_%" pos="zp" postype="percentage" wd="5_%" />
      </grup.nom>
    </sn>
    <coord>
      <c lem="i" postype="coordinating" wd="i"/>
    </coord>
    <sn>
      <spec gen="m" num="s">
        <d gen="m" lem="un" num="s" postype="indefinite" wd="un"/>
      </spec>
      <grup.nom gen="m" num="s">
        <z lem="8_%" postype="percentage" wd="8_%" />

```

```

    </grup.nom>
  </sn>
</sn>
<sa>
  <grup.a>
    <a [...] postype="qualificative" wd="superior"/>
    ‘between a higher 5% and 8%’

```

2.4.2.4 Adjectivisation of nouns

Sometimes the head of the adjective phrase can be a noun. We treat this way all nouns when there is no appositive or identity relationship between both elements but the second noun specifies the meaning of the first one, usually restricting its type (as in *el tipus de conveni* ‘the type of agreement’, *el tipus de pàgina* ‘the type of page’, etc.) or identifying a particular case of this noun (as in *una dinàmica* ‘a dynamics’, *un problema* ‘a problem’, etc.). Moreover, none of these structures can turn into copulative structures (**la pàgina es web* ‘the page is web’, **un benefici rècord* ‘the benefit is record’).

The morphological features of this adjective phrase have to be the same as those of the noun (2.70).

(2.70) (Cat 211nou)

```

<sn>
  <spec gen="f" num="s">
    <d gen="f" lem="un" num="s" postype="indefinite" wd="una"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n [...]lem="dinàmica" num="s" postype="common" wd="dinàmica"/>
    <s.a gen="f" num="s">
      <grup.a gen="f" num="s">
        <n [...] postype="proper" wd="Catalunya-Espanya"/>
      </grup.a>
    </s.a>
  </grup.nom>
</sn>

```

‘a Catalonia-Spain dynamics’

Sometimes, a number or date can function as the head of an adjective phrase (2.71):

```

(2.71) <n gen="m" lem="curs" num="s" postype="common" wd="curs"/>
  <s.a gen="m" num="s">
    <grup.a gen="m" num="s">
      <a [...] num="s" postype="qualificative" wd="escolar"/>

```

```

    </grup.a>
  </s.a>
  <s.a gen="m" num="s">
    <grup.a gen="m" num="s">
      <w lem="[?:?/?/2000/2001:?.?]" wd="2000/01"/>
    </grup.a>
  </s.a>
  'scholar course 2000/01'

```

2.4.2.5 Adjective locutions

The adjective expressions derived from verbal locutions are treated as adjective locutions: *posat_en_marxa* ‘set up’, *dut_a_terme* ‘carried out’, etc. Also sequences like *en_viu* ‘life’ are considered to be adjective locutions.

See Appendix for a complete list of adjective locutions.

2.4.3 Adverbial phrase (<sadv>)

The head of an adverbial phrase is always an adverb that, in its turn, can have the following complements:

1. Other adverbs

(2.72) (Cat 104-2)

```

<sadv>
  <spec>
    <r lem="més" wd="més"/>
  </spec>
  <grup.adv>
    <r lem="a_prop" wd="a_prop"/>
  <sp>
    <prep>
      <s [...] num="c" postype="preposition" wd="de"/>
    </prep>
    <sn>
      <spec gen="f" num="p">
        <d [...] num="p" postype="article" wd="les"/>
      </spec>
      <grup.nom gen="f" num="p">
        <n [...] num="p" postype="common" wd="zones"/>
      </grup.nom>
    </sn>
  </grup.adv>
  'closer to the zones'

```

2. Prepositional phrase (postposed)

(2.73) (Cat 104nou)

```

<sadv>
  <grup.adv>
    <r lem="després" wd="Després"/>
    <sp>
      <prep>
        <s [...] num="c" postype="preposition" wd="de"/>
      </prep>
      <sn>
        <spec gen="f" num="s">
          <d [...] num="s" postype="article" wd="la"/>
        </spec>
        <grup.nom gen="f" num="s">
          <n [...] num="s" postype="common" wd="trobada"/>

```

‘After the meeting’

3. Finite completive or complement clause (postposed)

(2.74) (Cat 111nou)

```

<grup.adv>
  <r lem="després" wd="després"/>
  <S clausetype="completive" impersonal="no">
    <conj conjunctiontype="subordinating">
      <c lem="que" postype="subordinating" wd="que"/>
    </conj>
    <sn>
      <grup.nom>
        <n [...] num="c" postype="proper" wd="Deutsche_Bank"/>
      </grup.nom>
    </sn>
    <grup.verb>
      <v [...] postype="main" tense="imperfect" wd="presentés"/>
  [...]

```

‘after the Deutsche Bank presented...’

4. Noun phrase (postposed)

(2.75) (Cat 06010100-42)

```

<sadv>
  <r lem="ahir" wd="ahir"/>
  <sn>
    <grup.nom gen="m" num="s">

```

```

    <w [...] wd="dijous_,_31_de_maig"/>
  </grup.nom>
</sn>
</sadv>
‘yesterday, Thursday, May 31st’

```

5. Attached noun phrase (anteposed)

When a noun phrase is quantifying an adverbial phrase, the former is not a daughter of the latter, but it is adjoined to it (2.76):

(2.76) (Cat 104nou)

```

<sadv>
  <sn adjunct="yes">
    <spec gen="m" num="p">
      <d [...] num="p" postype="indefinite" wd="pocs"/>
    </spec>
    <grup.nom gen="m" num="p">
      <n [...] num="p" postype="common" wd="mesos"/>
    </grup.nom>
  </sn>
</sadv>
<sadv>
  <grup.adv>
    <r lem="després" wd="després"/>
    <sp>
    <prep>
      <s [...] num="c" postype="preposition" wd="de"/>
    </sadv>
‘few months after ...’

```

In a few cases, however, the noun is the head of the construction, as in (2.77):

(2.77) (Cat 104nou)

```

<sn>
  <spec gen="m" num="s">
    <d gen="m" lem="un" num="s" postype="indefinite" wd="un"/>
  </spec>
  <grup.nom gen="m" num="s">
    <n gen="m" lem="pas" num="s" postype="common" wd="pas"/>
  <sadv>
    <grup.adv>
      <r lem="endavant" wd="endavant"/>
    </grup.adv>
  </sadv>

```

```

    </grup.nom>
  </sn>

  ‘one step forward’

```

2.4.3.1 *Més* and *menys*

The adverbs *més* ‘more’ and *menys* ‘less’ should be labeled as specifiers when they are followed by a noun group and therefore they are part of a noun phrase (2.78):

(2.78) (Cat 06010100-8)

```

<sn>
  <spec gen="f" num="p">
    <r lem="més" pos="rg" wd="més"/>
  </spec>
  <grup.nom gen="f" num="p">
    <n [...] num="p" postype="common" wd="competències"/>
    <s.a gen="f" num="p">
      <grup.a gen="f" num="p">
        <a [...] postype="qualificative" wd="normatives"/>
      </grup.a>
    </s.a>
  </grup.nom>
</sn>

‘more normative competences’

```

They are part of an adverbial phrase in the rest of the cases (2.79):

(2.79) (Cat 104nou)

```

<sadv>
  <spec>
    <r lem="més" wd="més"/>
  </spec>
  <grup.adv>
    <r lem="a_prop" wd="a_prop"/>
    <sp>
      <prep>
        <s [...] num="c" postype="preposition" wd="de"/>
      </prep>
    <sn>
      <spec gen="f" num="p">
        <d gen="f" lem="el" num="p" postype="article" wd="les"/>

```

```

</spec>
<grup.nom gen="f" num="p">
  <n [...] num="p" postype="common" wd="zones"/>

```

‘closer to the zones’

In both cases, the morphological node is marked as an adverb (<r>).

When *menys* stands for *excepte* ‘except’ and is followed by a noun phrase it functions as a preposition (<s>) (2.80):

(2.80) (Cat 111nou)

```

<d gen="m" lem="tot" num="p" wd="Tots"/>
<d gen="m" lem="el" num="p" postype="article" wd="els"/>
</spec>
<grup.nom gen="m" num="p">
  <n gen="m" lem="grup" num="p" postype="common" wd="grups"/>
  <sp>
    <prep>
      <s [...] num="s" postype="preposition" wd="del"/>
    </prep>
    <sn>
      <grup.nom gen="m" num="s">
        <n [...] num="c" postype="proper" wd="Parlament"/>
      </grup.nom>
    </sn>
  </sp>
</grup.nom>
</sn>
<sp adjunct="yes">
  <f lem="," punct="comma" wd=","/>
  <prep>
    <s [...] num="c" postype="preposition" wd="menys"/>
  </prep>
  <sn>
    <spec gen="m" num="s">
      <d gen="m" lem="el" num="s" postype="article" wd="el"/>
    </spec>
    <grup.nom gen="m" num="s">
      <n gen="c" lem="PP" num="c" postype="proper" wd="PP"/>
    </grup.nom>

```

‘All parliamentary groups but the PP’

2.4.3.2 Adverbial locutions

We consider some complex expressions to be adverbial locutions.

1. *fins_a* ‘as many as’ is an adverbial locution in sequences like *van venir fins_a 40 persones* ‘as many as 40 people came’, where *fins_a* is a part of the specifier¹¹, as shown in 2.81. In the rest of the cases it works as a sequence of two prepositions.

(2.81) (Cat 06010100-4)

```

<sn arg="arg1" func="cd" tem="pat">
  <spec gen="f" num="p">
    <r lem="més" pos="rg" wd="més"/>
  </spec>
  <grup.nom gen="f" num="p">
    <n [...] num="p" postype="common" wd="competències"/>
    <s.a gen="f" num="p">
      <grup.a gen="f" num="p">
        <a [...] postype="qualificative" wd="normatives"/>
      </grup.a>
    </s.a>
  </grup.nom>
</sn>

```

‘more normative competences’

2. *en_canvi*, ‘on the other hand’, *no_obstant* ‘nevertheless’, and *per_contra* ‘on the contrary’, can function as well as coordinating conjunctions (<cc>), when they do not precede another coordinating conjunction.
3. *no_només* ‘not only’ is an adverbial locution (<rg>) but it works as a coordinating conjunction (<coord>).
4. *sí_que* (emphatic particle) introduces emphatic structures.
5. *un_cop* and *una_vegada* ‘once’, can work as subordinate conjunctions (<cs>), when followed by a finite clause.

Sequences consisting of an adverb followed by the conjunction *que* (*abans que* ‘before’, *després que* ‘after’) are analyzed as adverbial phrases instead of locutions. The adverbial phrase, in its turn, is complemented by a completive finite clause. One exception is that of *mentre que* ‘while’, which is analyzed as a coordinating locution.

6. Some of these complex expressions and locutions are ambiguous (as happens in *sempre que*): When this expression has a conditional meaning (‘provided that, as long as’) and rules the subjunctive mode, we should analyze it as a conjunctive locution. When it indicates time values and rules indicative mode, we should split it up and analyze it as an adverbial phrase (<sadv> followed by a relative clause).

¹¹See section 2.4.1.2 for more details on determiners.

7. Sequences like *a_la_vegada* ‘at the same time’, *al_cap_i_a_la_fí* ‘after all’, *a_més_a_més* ‘moreover’, and *tot_arreu* ‘everywhere’, are some examples of adverbial locutions. See Appendix for a complete list.

2.4.3.3 Attaching the adverbial phrase

Those adverbial phrases affecting the verb (that is, the whole sentence), belong to the sentence to which this verb belongs. Therefore, it is not analyzed as an adjunction (2.82):

(2.82) (Cat 111nou)

```
<S clausetype="completive" impersonal="yes">
  <infinitiu>
    <v [...] mood="infinitive" postype="main" wd="buscar"/>
  </infinitiu>
  <sadv>
    <grup.adv>
      <r lem="també" wd="també"/>
    </grup.adv>
  </sadv>
  <sn>
    <spec gen="m" num="s">
      <d gen="c" lem="el" num="s" postype="article" wd="l'"/>
    </spec>
    <grup.nom gen="m" num="s">
      <n gen="m" lem="acord" num="s" postype="common" wd="acord"/>
      <s.a gen="m" num="s">
        <grup.a gen="m" num="s">
          <a [...] num="s" postype="qualificative" wd="polític"/>
        </grup.a>
      </s.a>
    </grup.nom>
  </sn>
</S>
```

‘to search for a political agreement’

We adjoin those adverbs (*només* ‘only, just’, *potser* ‘maybe’, *fins_i_tot* ‘even’, *quasi* ‘almost’, *no*, ‘no’ etc.) that do not belong to the sentence but affect a single phrase structure, and:

1. This phrase structure is not an adjective phrase (or participle clause) or an adverbial phrase.
2. The adverb can only occur in that concrete position or even before the verb provided that it appears between commas (2.83):

(2.83) (Cat 111nou)

```

<n [...] num="s" postype="common" wd="proposició"/>
<sp>
  <neg adjunct="yes">
    <r lem="no" postype="negative" wd="no"/>
  </neg>
<sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="de"/>
  </prep>
<sn>
  <grup.nom gen="f" num="s">
    <n [...] num="s" postype="common" wd="lleï"/>
  </grup.nom>
‘Private Member’s Bill’

```

2.4.3.4 Adverbs belonging to a specifier

We represent as flat structures those expressions in which an adverb is a part of a complex specifier:

1. *més de* ‘more than’, *menys de* ‘less than’ + number (2.84):

(2.84) (Cat 211nou)

```

<sn>
  <spec gen="m" num="p">
    <d gen="m" lem="el" num="p" postype="article" wd="els"/>
    <r lem="més" wd="més"/>
    <s [...] num="c" postype="preposition" wd="de"/>
    <z lem="22" wd="22"/>
  </spec>
  <grup.nom gen="m" num="p">
    <n [...] num="p" postype="common" wd="milions"/>
  <sp>
    <prep>
      <s [...] num="c" postype="preposition" wd="de"/>
    </prep>
  <sn>
    <spec gen="f" num="s">
      <d [...] num="s" postype="article" wd="la"/>
    </spec>
    <grup.nom gen="f" num="s">
      <n [...] num="c" postype="proper" wd="JNC"/>

```

‘those more than 22 milions of the JNC’

2. *gairebé* ‘almost’, *fins_a* ‘to, untill’, *al_voltant_de* ‘arround’, *prop_de* ‘near’... + number (2.85):

(2.85) (Cat 203-14)

```
<sn>
  <spec gen="m" num="p">
    <r lem="gairebé" wd="gairebé"/>
    <z lem="14" wd="14"/>
  </spec>
  <grup.nom gen="m" num="p">
    <n [...] num="p" postype="common" wd="milions"/>
  <sp>
    <prep>
      <s [...] num="c" postype="preposition" wd="de"/>
    </prep>
    <sn>
      <grup.nom gen="f" num="p">
        <z lem="pesseta" postype="currency" wd="pessetes"/>

```

‘almost 14 milion pesetas’

2.4.4 Prepositional phrase (<sp>)

Prepositional phrases consist of a preposition and the term it introduces. This term, in its turn, can consist of one of the following elements:

1. **Noun phrase** (<sn>)

```
(2.86) <sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="de"/>
  </prep>
  <sn>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="longitud"/>
    </grup.nom>
  </sn>
</sp>
‘of lenght’
```

2. **Adverbial phrase** (<sadv>)

```
(2.87) <sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="fins"/>
```

```

    </prep>
    <sadv>
      <grup.adv>
        <r lem="ahir" wd="ahir"/>
    ‘until yesterday’

```

3. Finite completive clause

(2.88) (104nou)

```

    <r lem="després" wd="després"/>
    <S clausetype="completive" impersonal="no">
      <conj conjunctiontype="subordinating">
        <c lem="que" postype="subordinating" wd="que"/>
      </conj>
      <sn>
        <grup.nom>
          <n [...] num="c" postype="proper" wd="Deutsche_Bank"/>
        </grup.nom>
      </sn>
      <grup.verb>
        <v [...] tense="imperfect" wd="presentés"/>
      </grup.verb>

```

‘after Deutsche Bank presented ...’

4. Non-finite completive clause

(2.89) (Cat 104nou)

```

    <sp>
      <prep>
        <s [...] num="c" postype="preposition" wd="per"/>
      </prep>
      <S clausetype="completive" impersonal="yes">
        <infinitiu>
          <v [...] postype="main" wd="donar"/>
        </infinitiu>
        <sn>
          <spec gen="m" num="s">
            <d [...] num="s" postype="indefinite" wd="un"/>
          </spec>
          <grup.nom gen="m" num="s">
            <n [...] num="s" postype="common" wd="permís"/>

```

‘in order to give a permission’

5. Adjective phrase

(2.90) (Cat 104nou)

```

<n [...] num="s" postype="common" wd="aeroport"/>
<sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="de"/>
  </prep>
  <s.a gen="f" num="s">
    <grup.a gen="f" num="s">
      <a [...] num="s" postype="ordinal" wd="segona"/>
    </grup.a>
  </s.a>
</sp>
'minor airport'

```

6. Non-finite adjective clause

(2.91) (Cat 06010300-24)

```

<sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="per"/>
  </prep>
  <S clausetype="participle" impersonal="yes">
    <participi>
      <a [...] posfunction="participle" wd="finalitzada"/>
    </participi>
  </S>
</sp>
'to (consider it) done'

```

7. Absolute clause

(2.92) (Cat 06030200-13)

```

<sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="fins"/>
  </prep>
  <S [...] clausetype="participle" impersonal="yes">
    <grup.verb>
      <v [...] num="s" postype="main" wd="passat"/>
    </grup.verb>
  </S>
</sp>

```

```

    <sn>
      <spec gen="m" num="s">
        <d [...] num="s" postype="article" wd="l"/>
      </spec>
      <grup.nom gen="m" num="s">
        <n [...] num="s" postype="common" wd="estiu"/>
      </grup.nom>
    </sn>
  </S>
</sp>
‘until the end of the summer’

```

2.4.4.1 Prepositions (<s>)

We consider the following prepositions¹²: *a* ‘to, in, on, at’, *amb* ‘with’, *contra* ‘against’, *davant* ‘in front of’, *de* ‘of’, *després* ‘after’, *dins* ‘in, within’, *entre* ‘between, among’, *envers* ‘towards’, *fins* ‘till, until’, *malgrat* ‘despite’, *menys* ‘except’, *mijançant* ‘by means of’, *per* ‘for, to’, *segons* ‘according to’, *sense* ‘without’, *sobre* ‘about, on’, and *via* ‘via’.

Dins, *davant* etc. are prepositions (<s>) when they are not followed by another preposition (2.93). Otherwise, they are adverbs (<r>), as in 2.94.

(2.93) (Cat 06010100-33)

```

<sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="davant"/>
  </prep>
  <sn>
    <spec gen="f" num="s">
      <d gen="f" lem="el" num="s" postype="article" wd="la"/>
    </spec>
    <grup.nom gen="f" num="s">
      <n [...] num="c" postype="proper" wd="Cabra"/>
    </grup.nom>
  </sn>
</sp>

```

‘in front of the Chamber’

(2.94) (Cat 06010100-33)

```

<sadv>
  <grup.adv>

```

¹²Prepositional locutions are given in section 2.4.4.2.

```

<r lem="dins" pos="rg" wd="dins"/>
<sp>
  <prep>
    <s [...] postype="preposition" wd="del"/>
  </prep>
  <sn>
    <grup.nom gen="m" num="s">
      <n [...] wd="Departament_de_Medi_Ambient"/>
    </grup.nom>
  </sn>
</sp>
</grup.adv>
</sadv>

```

‘within the Environmental department’

Those contracted forms of preposition plus determiner, come without the node corresponding to the phrase noun from the automatic parser. This node is to be added manually (2.95):

(2.95) (Cat 104nou)

```

<sn>
  <spec gen="m" num="s">
    <d gen="f" lem="el" num="s" postype="article" wd="La"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n gen="f" lem="resta" num="s" postype="common" wd="resta"/>
    <sp>
      <prep>
        <s [...] lem="del" num="s" postype="preposition" wd="del"/>
      </prep>
      <sn>
        <grup.nom gen="m" num="s">
          <n [...] num="s" postype="common" wd="capital"/>
        </grup.nom>
      </sn>
    </sp>
  </grup.nom>
</sn>

```

‘The rest of the capital’

The following are special cases of prepositions and the directions for their analysis.

1. Preposition + preposition

Two consecutive prepositions (*per a, fins a, d'entre* etc.) should be represented as shown in 2.96:

(2.96) (Cat 06010300-24)

```

<sp>
  <prep>
    <s gen="c" lem="per" num="c" postype="preposition" wd="per"/>
  </prep>
  <sp>
    <prep>
      <s gen="c" lem="a" num="c" postype="preposition" wd="a"/>
    </prep>
    <S clausetype="completive" impersonal="yes">
      <infinitiu>
        <v [...] num="c" postype="main" wd="disputar"/>
      </infinitiu>
      <sn>
        <spec gen="m" num="s">
          <d gen="m" lem="el" num="s" postype="article" wd="el"/>
        </spec>
        <grup.nom gen="m" num="s">
          <n [...] num="s" postype="common" wd="campionat"/>

```

‘for the championship’

2. Com/como

Com is an ambiguous form. It functions as a preposition when followed by a noun phrase that does not belong to a sentence structure. The corresponding morphological node, however, corresponds to a subordinate conjunction (<cs>) (2.97):

(2.97) (Cat 104nou)

```

<n [...] num="p" postype="common" wd="empreses"/>
<sp>
  <prep>
    <s gen="c" lem="de" num="c" postype="preposition" wd="de"/>
  </prep>
  <sn>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="comunicació"/>
    </grup.nom>
  </sn>
</sp>

```

```

<sp>
  <prep>
    <c lem="com" postype="subordinating" wd="com"/>
  </prep>
  <sn>
    <grup.nom>
      <n [...] num="c" postype="proper" wd="Premsa_Ibèrica"/>
    </grup.nom>
  </sn>
</sp>

```

‘communication companies like Premsa Ibèrica’

It is also analyzed as a preposition when belonging to the sequence *com si* (‘as if’). We should represent the sequence as to be a preposition followed by a finite completive clause, which is introduced by the subordinate conjunction *si* (2.98)

(2.98) (Cat 124_19990202)

```

<sp>
  <prep>
    <c lem="com" postype="subordinating" wd="com"/>
  </prep>
  <S clausetype="completive" impersonal="no">
    <conj conjunctiontype="subordinating">
      <c lem="si" postype="subordinating" wd="si"/>
    </conj>
    <sn>
      <grup.nom num="s">
        <p [...] num="s" person="2" postype="personal" wd="et"/>
      </grup.nom>
    </sn>
    <grup.verb>
      <v [...] postype="main" tense="imperfect" wd="caigués"/>
    </grup.verb>
    <sn>
      <spec gen="f" num="s">
        <d [...] num="s" postype="indefinite" wd="una"/>
      </spec>
      <grup.nom gen="f" num="s">
        <n [...] num="s" postype="common" wd="llosa"/>
      </grup.nom>
    </sn>
  </S>
</sp>

```

‘as if a stone falled over you’

Com works as a subordinate conjunction (<cs>) when introducing a modal clause (2.99):

(2.99) (Cat 211nou)

```

<S clausetype="adverbial" impersonal="no">
  <conj conjunctiontype="subordinating">
    <c lem="com" postype="subordinating" wd="com"/>
  </conj>
  <sn elliptic="yes"/>
  <grup.verb>
    <v [...] postype="auxiliary" tense="present" wd="va"/>
    <v [...] mood="infinitive" postype="main" wd="passar"/>
  </grup.verb>
  ‘as it happened’

```

It functions as a relative pronoun (<p,relative>) when it introduces a relative clause 2.100:

(2.100) (Cat 211nou)

```

<sn>
  <grup.nom num="s">
    <p [...] num="s" postype="demonstrative" wd="tal"/>
    <S clausetype="relative" impersonal="no">
      <relatiu>
        <p [...] num="c" postype="relative" wd="com"/>
      </relatiu>
    <grup.verb>
      <v [...] tense="imperfect" wd="havia"/>
      <v [...] mood="pastparticiple" wd="demanat"/>
    ‘according to what (s/he) asked for’

```

It can also function as an interrogative pronoun (<p,interrogative>) when it introduces a completive clause (2.101):

(2.101) (Cat 111nou)

```

  <v [...] num="c" postype="main" wd="discutir"/>
</infinitiu>
</grup.verb>
<morfema.verbal>
  <p gen="c" lem="es" num="c" wd="-se"/>
</morfema.verbal>
<S clausetype="completive" impersonal="no" >
  <sadv>
    <grup.adv>
      <p [...] num="c" postype="interrogative" wd="com"/>

```

```

    </grup.adv>
  </sadv>
  <morfema.verbal func="pass">
    <p gen="c" lem="es" num="c" wd="s'"/>
  </morfema.verbal>
  <grup.verb>
    <v [...] tense="conditional" wd="articularia"/>
  </grup.verb>
  <sn>
    <spec gen="f" num="s">
      <d gen="f" lem="el" num="s" postype="article" wd="la"/>
    </spec>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="participació"/>

```

‘to discuss how the participation is to be articulated’

Finally, it can introduce a predicative complement. In such a case, it is followed by the preposition *a* (in Catalan) and the whole set is analyzed as a prepositional locution, for what its morphological node corresponds to a preposition (<s>)¹³:

(2.102) (Cat 06019900-20)

```

  <grup.verb>
    <v [...] lem="reconèixer" tense="present" wd="reconegui"/>
  </grup.verb>
  <sn>
    <spec gen="f" num="s">
      <d gen="f" lem="el" num="s" postype="article" wd="la"/>
    </spec>
    <grup.nom gen="f" num="s">
      <n gen="f" lem="dansa" num="s" postype="common" wd="dansa"/>
      <s.a gen="f" num="s">
        <grup.a gen="f" num="s">
          <a [...] postype="qualificative" wd="catalana"/>
        </grup.a>
      </s.a>
    </grup.nom>
  </sn>
  <sp>
    <prep>
      <s gen="c" lem="com_a" num="c" postype="preposition" wd="com_a"/>

```

¹³In Spanish the preposition *como* can introduce a predicative complement itself, and does not belong to a prepositional locution. The context and both syntax and semantics will help on disambiguating its category.

```

</prep>
<sn entityref="ne" ne="other">
  <grup.nom>
    <n [...] num="c" postype="proper" wd="Patrimoni_de_la_Humanitat"/>

```

‘(so that it) recognizes the Catalan danse as a World Heritage’

3. Segons / según

The form *segons* ‘according to’, can also work as a subordinate conjunction (2.103) or as a preposition (2.104). Unlike what happens in the form *com*, however, the morphological node should be <c,subordinating> (subordinating conjunction) in the former case, and <s> (preposition) in the latter.

(2.103) (Cat 104nou)

```

<S clausetype="adverbial" impersonal="no">
  <conj conjunctiontype="subordinating">
    <c lem="segons" postype="subordinating" wd="Segons"/>
  </conj>
  <grup.verb>
    <v [...] postype="auxiliary" tense="present" wd="ha"/>
    <v [...] num="s" postype="main" wd="explicat"/>
  </grup.verb>
  <sadv>
    <grup.adv>
      <r lem="avui" wd="avui"/>
    </grup.adv>
  </sadv>
  <sn>
    <spec gen="m" num="s">
      <d gen="m" lem="el" num="s" postype="article" wd="el"/>
    </spec>
    <grup.nom coord="yes">
      <grup.nom coord="yes">
        <grup.nom gen="m" num="s">
          <n [...] postype="common" wd="vicesecretari"/>

```

‘According to what the assistant secretary explained today’

(2.104) (Cat 211nou)

```

<sp adjunct="yes">
  <prep>
    <s [...] num="c" postype="preposition" wd="segons"/>
  </prep>
  <sn>

```

```

<spec gen="m" num="s">
  <d gen="m" lem="el" num="s" postype="article" wd="el"/>
</spec>
<grup.nom gen="m" num="s">
  <n [...] num="s" postype="common" wd="ministre"/>
  <s.a gen="m" num="s">
    <grup.a gen="m" num="s">
      <n [...] num="s" postype="common" wd="portaveu"/>

```

‘According to the ministry spokesperson’

2.4.4.2 Prepositional locutions

Some complex expressions are considered to be prepositional. Sequences consisting of a preposition followed by the conjunction *que* (as in *fins que* ‘till, until’), do not form a conjunction locution but a prepositional phrase followed by a finite completive clause introduced by *que*¹⁴.

Complex expressions such as *per_sobre de* ‘over’, *per_sota de* ‘below’ (adverbial locution + prepositional phrase); and those such as *a principis de* ‘at the beginning of’, *en qüestió de* ‘in terms of’, *relatiu a* ‘regarding’, *encara mes* ‘even more’, etc., are not prepositional locutions.

1. Preposition vs. adverb

no_obstant (*no_obstante* in Spanish) ‘nevertheless’ can function as an adverb when it occurs between commas and when it follows a coordinating conjunction.

When *lluny de* ‘far from’ has a locative meaning, it is not a prepositional locution but an adverb followed by a preposition.

2. Preposition vs. coordinating conjunction

conjuntament_amb ‘together with’, *en_col·laboració_amb* ‘in collaboration with’, *en_coordinació_amb* ‘in coordination with’ and *juntament_amb* ‘together with’ are marked as prepositions (<s>). However, they can work as coordinating conjunctions (<c,coordination>). They are prepositions when occurring before their antecedent, separated from the series to which they belong.

On the other hand, they are coordinating conjunctions when introducing the last member of the coordination (2.105).

(2.105) (Cat 06030200-5)

[[el gir dretà i pragmàtic d’UCD], juntament_amb <coord> [la defensa de la bondat del pacte CIU-PP]], contribueix encara més a alimentar la sensació de dependència

‘the pragmatic change to the Right of the UCD party, together with the defense of the godness of the CIU-PP pact, contributes even more to nurture the dependency feeling’

¹⁴See appendix 5 for a complete list of prepositional locutions

en_lloc_de ‘instead of’ and *en_comptes_de* ‘instead of’ are always labeled as prepositions (<s>) as well, although they can function as preposition or coordinating conjunctions (<c,coord>). They work as preposition when they occur before their antecedent, separated from the series to which they belong.

They are coordinating conjunctions when introducing the last member of the coordination (2.106):

(2.106) (Cat 211nou)

enfortir el partit, renovar-lo, en_lloc_d' anar-lo desfent
 ‘to make the party stronger, to renew it, instead of break it up’

al_marge_de ‘regardless of, despite’, *a_més_de* (or *a_més_a_més_de*), ‘as well as, besides, in addition to’, and *a_banda_de* ‘apart from’, have always the morphological node corresponding to prepositions (<s>) and can also work as prepositions as well as coordinating conjunctions (<c,coord>). They are prepositions when they occur before their antecedent and separated from the series to which they belong (2.107). On the other hand, they are coordinating conjunctions when introducing the last member of the coordination (2.108):

(2.107) (Cat 204nou)

els majors descensos, a_més_a_més_de Les Balears, es van registrar a Cantàbria [...]
 ‘the higher falls, in addition to that of Les Balears, were registered in Cantàbria [...]’

(2.108) (Cat 204nou)

Castella_Lleó (-2,09), Astúries (-2,08), Múrcia (-0,58), Extremadura (-0,10), a_més_a_més_de Melilla (-0,16)
 ‘Castella_Lleó (-2,09), Astúries (-2,08), Múrcia (-0,58), Extremadura (-0,10), as well as Melilla (-0,16)’

3. Coordinating prepositional phrases

When two prepositional phrases are in coordination, the first one is introduced by a prepositional locution, and the second just by a part of this locution (2.109):

(2.109) (Cat 06010100-8)

```
<sp>
  <prep>
    <s [...] postype="preposition" wd="en_funció_de"/>
  </prep>
  <sn>
    <spec gen="f" num="s">
      <d [...] num="s" postype="article" wd="la"/>
    </spec>
```

```

<grup.nom gen="f" num="s">
  <n [...] num="s" postype="common" wd="piràmide"/>
  <sp>
    <prep>
      <s [...] num="c" postype="preposition" wd="d'"/>
    </prep>
    <sn>
      <grup.nom gen="f" num="s">
        <n [...] num="s" postype="common" wd="edat"/>
      </grup.nom>
    </sn>
  </sp>
</grup.nom>
</sn>
</sp>
‘depending on the age pyramid’

```

4. Encrusted elements within the prepositional locution

Sometimes there are encrusted elements among the components of a prepositional locution (2.110):

```

(2.110) <sp>
  <prep>
    <s [...] postype="preposition" wd="a_conseqüència"/>
    <sadv>
      <f lem="," punct="comma" wd=","/>
      <grup.adv>
        <r lem="presumptament" wd="presumptament"/>
      </grup.adv>
      <f lem="," punct="comma" wd=","/>
    </sadv>
    <s [...] num="c" postype="preposition" wd="d'"/>
  </prep>
  <sn>
    <spec gen="f" num="s">
      <d [...] num="s" postype="indefinite" wd="una"/>
    </spec>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="intoxicació"/>
    </grup.nom>
  </sn>
</sp>
‘as a result, supposedly, of an intoxication’

```


2.4.5 Verb group (<grup.verb> or <gv>)

This node only includes conjugated verbal forms, both simple and complex¹⁵.

1. Simple verbal form:

(2.111) (Cat 06010100-52)

```
<grup.verb>
  <v [...] postype="main" tense="present" wd="és"/>
</grup.verb>
‘(s/he) is’
```

2. Complex verbal form:

(2.112) (Cat 06029900-3)

```
<grup.verb>
  <v [...] postype="auxiliary" tense="present" wd="han"/>
  <v [...] mood="pastparticiple" wd="decretat"/>
</grup.verb>
‘(they) have decreed’
```

3. Passive voice:

(2.113) (Cat 06029900-3)

```
  <d gen="f" lem="el" num="s" postype="article" wd="la"/>
</spec>
  <grup.nom gen="f" num="s">
  <n [...] num="s" postype="common" wd="decisió"/>
  </grup.nom>
</sn>
<grup.verb>
  <v [...] postype="auxiliary" tense="present" wd="ha"/>
  <v [...] postype="semiauxiliary" wd="estat"/>
  <v [...] postype="main" wd="adoptada"/>
</grup.verb>
‘the decision has been taken’
```

4. Verbal periphrasis:

(2.114) (Cat 06029900-15)

¹⁵In the Catalan version of the corpus, it is labeled as (<grup.verb>) while in the Spanish one the label is (<gv>).

```

<grup.verb>
  <v [...] postype="main" tense="future" wd="tornarà"/>
  <s [...] lem="a" num="c" postype="preposition" wd="a"/>
  <infinitiu>
    <v [...] mood="infinitive" postype="main" wd="atracar"/>
  </infinitiu>
</grup.verb>

```

‘(it) will dock again’

(2.115) (Cat 06029900-15)

```

<grup.verb>
  <v [...] postype="main" tense="present" wd="estan"/>
  <gerundi>
    <v [...] mood="gerund" postype="main" wd="construint"/>
  </gerundi>
</grup.verb>

```

‘(they) are building’

5. Verbal periphrasis within another periphrasis:

(2.116) (Cat 06019900-14)

```

<grup.verb>
  <v [...] postype="auxiliary" tense="present" wd="ha"/>
  <s [...] num="c" postype="preposition" wd="d'"/>
  <infinitiu>
    <v [...] mood="infinitive" postype="main" wd="acabar"/>
    <s [...] num="c" postype="preposition" wd="de"/>
    <infinitiu>
      <v [...] postype="main" wd="reomplir"/>
    </infinitiu>
  </infinitiu>
</grup.verb>

```

‘(s/he) has to finish refilling’

In conjugated forms of the impersonal verb *haver/haber* (*hi ha*, etc.) ‘there is/are’, the clitic pronoun *hi* should be marked as pronominal morpheme (<morfema.pronominal>) instead of a noun phrase (<sn>), as shown in 2.117:

(2.117) (Cat 06010100-13)

```

<neg>
  <r lem="no" postype="negative" wd="no"/>
</neg>

```

```

<morfema.pronominal>
  <p [...] num="c" person="3" postype="personal" wd="hi"/>
</morfema.pronominal>
<grup.verb>
  <v [...] postype="main" tense="present" wd="ha"/>
</grup.verb>
<sn>
  <spec gen="f" num="s">
    <d [...] num="s" postype="indefinite" wd="cap"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n [...] num="s" postype="common" wd="previsió"/>
  'there is no prediction'

```

2.4.5.1 Verbal periphrasis

1. Infinitive periphrases

We consider to be infinitive periphrases any structure that fulfils two fundamental criteria. On the one hand, the non-finite verbal form cannot be replaced by a noun or nominal element. On the other hand, the clitic pronouns can occur before or after the whole group. Both criteria should be obeyed, otherwise there is no verbal periphrasis.

- <acabar de + infinitive> ‘to have just done sth’
- <acostumar a + infinitive> ‘to use to do sth’
- <arrencar a + infinitive> ‘to burst into sth’
- <arribar a + infinitive> ‘to go so far as doing sth’
- <començar a + infinitive> ‘to start doing sth’
- <deixar + infinitive> ‘to let/cause something to be done’¹⁶
- <deixar de + infinitive> ‘to give up sth’
- <deure + infinitive> ‘may/might be’
- <entrar a + infinitive> ‘to start/begin doing sth’
- <esclatar a + infinitive> ‘to burst into sth’
- <estar a punt de + infinitive> ‘to be about to do sth’
- <fer + infinitive> ‘to make/cause something to be done’¹⁷
- <gosar + infinitive> ‘to dare to do sth’
- <haure de + infinitive> ‘to have to do sth’

¹⁶This structure stands for a causative alternation.

¹⁷This structure stands for a causative alternation.

- <*haver de* + infinitive> ‘to have to do sth’
- <*heure de* + infinitive> ‘to have to do sth’
- <*mirar de* + infinitive> ‘to try to do sth’
- <*parar de* + infinitive> ‘to stop doing sth’
- <*passar a* + infinitive> ‘to proceed to do sth’
- <*pensar* + infinitive> ‘to intend to do sth’
- <*poder* + infinitive> ‘to be able to do sth’
- <*posar-se a* + infinitive> ‘to start/begin doing sth’
- <*resultar* + infinitive> ‘to turn out to be/do sth’
- <*solet* + infinitive> ‘to do sth usually’
- <*tornar a* + infinitive> ‘to do sth again’

The representation of these periphrases should be like in 2.118:

Voler + infinitive, *pensar en* + infinitive, *saber* + infinitiu, etc. are not periphrases.

(2.118) (Cat 06029900-15)

```

<grup.verb>
  <v [...] postype="main" tense="future" wd="tornarà"/>
  <s [...] lem="a" num="c" postype="preposition" wd="a"/>
  <infinitiu>
    <v [...] mood="infinitive" postype="main" wd="atracar"/>
  </infinitiu>
</grup.verb>
‘(it) will dock again’

```

2. Gerund periphrases

- <*acabar* + gerund> ‘to end up doing sth’
- <*anar* + gerund> ‘to keep on doing sth’
- <*continuar* + gerund> ‘to continue doing sth’
- <*estar* + gerund> ‘to be doing sth’
- <*seguir* + gerund> ‘to continue doing sth’

The proper representation of gerund periphrases is the one shown in 2.119:

(2.119) (Cat 06029900-15)

```

<grup.verb>
  <v [...] postype="main" tense="present" wd="estan"/>
  <gerundi>
    <v [...] mood="gerund" postype="main" wd="construint"/>
  </gerundi>
</grup.verb>

```

‘(they) are building’

3. Participle periphrases

We only consider those complex tenses of the verbal conjugation, as well as the following:

<*deixar* + participle> ‘to leave sth in a certain way’

There are, however, some exceptions (2.120):

(2.120) (Cat 06030200-2)

```

<grup.verb>
  <v [...] postype="auxiliary" tense="present" wd="va"/>
  <v [...] num="c" postype="main" wd="deixar"/>
</grup.verb>
<S clausetype="participle" impersonal="yes">
  <participi>
    <a [...] posfunction="participle" wd="redactat"/>
  </participi>

```

‘(s/he) left it written’

Other sequences are considered to be verbal groups followed by a non-finite participle clause, rather than participle periphrases:

- *venen potenciats* ‘(they) come favoured’
- *van dirigides* ‘(they) are aimed at’
- *van destinades* ‘(they) are aimed at’
- *resultar* + participle ‘to turn out to be + participle’

2.4.5.2 Coordinating periphrases

As it happens in complex verbal tenses, sentence coordination may cause the lack of the finite form of the verb within the periphrasis of the second part of the coordination. In such a case, the non-finite verbal form should have its corresponding label (<infinitiu>, <gerundi> or <participi>) and the sentence to which it belongs must be marked as verbless (2.121):

(2.121) (Cat 111nou)

```

<grup.verb>
  <v [...] postype="auxiliary" tense="present" wd="han"/>
  <s [...] num="c" postype="preposition" wd="de"/>
  <infinitiu>
    <v [...] postype="main" wd="condemnar"/>
  </infinitiu>
</grup.verb>
<sn coord="yes">
  <sn>
    <spec gen="f" num="s">
      <d [...] num="s" postype="demonstrative" wd="aquesta"/>
    </spec>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="iniciativa"/>
    </grup.nom>
  </sn>
  [...]
</S>
<coord>
  <c lem="i" postype="coordinating" wd="i"/>
</coord>
<S verbless="yes">
  <infinitiu>
    <f lem="&quot;" punct="quotation" wd="&quot;"/>
    <v [...] num="c" postype="main" wd="explicar"/>
    <f lem="&quot;" punct="quotation" wd="&quot;"/>
  </infinitiu>

```

‘(they) have to condemn this initiative ... and explain ...’

2.4.5.3 Encrusted elements within the periphrases

Periphrases can also have encrusted elements, being these generally adverbial and noun phrases. They are usually placed between the auxiliary and the non-finite form of the verb. Since we do not alter the surface order, we must encrust those elements (which maintain their phrase label) within the verb group. The representation is the one shown in 2.122:

(2.122) (Cat 06029900-8)

```

<S clausetype="completive" impersonal="yes">
  <infinitiu>
    <v [...] mood="infinitive" postype="main" wd="fer"/>
    <sn>

```

```

    <grup.nom gen="m" num="s">
      <p [...] case="accusative" postype="personal" wd="-lo"/>
    </grup.nom>
  </sn>
  <infinitiu>
    <v [...] mood="infinitive" postype="main" wd="desistir"/>
  </infinitiu>
</infinitiu>
</S>

```

‘make him desist’

The reflexive pronoun *es* is another example of encrusted element. When it occurs in verb locutions we should split it up and analyze it as an encrusted element.

2.4.5.4 Verb locutions

We consider to be verb locutions those structures fulfilling the following criteria: *a)* there is no part belonging to the structure that can be replaced by a clitic pronoun; *b)* no other elements can be introduced among the components; and *c)* variation is limited when combining words. Some examples are listed below:

- *deixar_de_banda* ‘to leave sth aside or to one side’
- *donar_a_conèixer* ‘to announce, reveal’
- *donar_lloc* ‘to provoke, give rise to’
- *fer_front* ‘to face up to sth’
- *fer_veure* ‘to feign, fake’
- *posar_de_manifest* ‘to highlight sth’
- *tenir_a_veure* ‘to deal with, to have sth to do with’
- *voler_dir* ‘to mean’

The complete list of verb locutions is given in Appendix.

Noun and adjective expressions deriving from verbal locutions are also considered to be locutions:

posada_en_marxa, *posada_en_funcionament* ‘set up’; *posat_en_marxa* ‘tal’, *dut_a_terme* ‘carried out’.

The following expressions are NOT verb locutions:

- *continuar_endavant* ‘to carry on’
- *donar_resposta* ‘to give answer’

- *donar suport* ‘to give support’
- *estar d’acord* ‘to agree’
- *fer public* ‘to make (sth) public’
- *posar en perill* ‘to endanger’
- *posar-se d’acord* ‘to agree’
- *prendre posicio* ‘to take position’

2.4.6 The morpheme ‘es’

1. ‘es’ as a pronominal morpheme (<morfema.pronominal>)

When *es* is a pronominal morpheme¹⁸ its function is that of going along with the verb (2.123).

(2.123) (Cat 06029900-20)

```

<grup.verb>
  <v [...] postype="auxiliary" tense="present" wd="va"/>
  <v [...] mood="infinitive" postype="main" wd="acostar"/>
</grup.verb>
<morfema.pronominal>
  <p gen="c" lem="ell" num="c" person="3" wd="-se"/>
</morfema.pronominal>
<sp>
  <prep>
    <s [...] contracted="yes" postype="preposition" wd="al"/>
  </prep>
  <sn>
    <grup.nom gen="m" num="s">
      <n [...] postype="common" wd="taxista"/>

```

‘(s/he) went closer to the taxi driver’

The inherent morpheme *hi* in the verb *haver-hi* (*hi ha*) is also a pronominal morpheme and it should be marked as <morfema.pronominal> as well (2.124).

(2.124) (Cat 06010100-13)

¹⁸This morpheme is inherent in the verb, and agrees in gender and number with the subject. The following are the forms it can take: *em*, *et*, *es*, *ens*, *us*, etc. The morphological node is <p gen=‘c’ lem=‘ell’ num=‘c’ person=‘3’ wd=‘-se’/>.


```

<neg>
  <r lem="no" postype="negative" wd="no"/>
</neg>
<morfema.pronominal>
  <p [...] num="c" person="3" ostype="personal" wd="hi"/>
</morfema.pronominal>
<grup.verb>
  <v [...] postype="main" tense="present" wd="ha"/>
</grup.verb>
<sn>
  <spec gen="f" num="s">
    <d [...] num="s" postype="indefinite" wd="cap"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n [...] num="s" postype="common" wd="previsió"/>
  </grup.nom>

```

‘there is no prediction’

2. ‘es’ as a verb morpheme (morfema.verbal)

es can have two different functions as a verb morpheme:

It is a reflexive passive marker, that has a subject (usually after the verb), with which it agrees (2.125).

(2.125) (Cat 06019900-29)

```

<morfema.verbal func="pass">
  <p gen="c" lem="es" num="c" wd="es"/>
</morfema.verbal>
<grup.verb>
  <v [...] postype="main" tense="future" wd="durà_a_terme"/>
</grup.verb>

```

‘(it) will be carried out’

It is the marker of an impersonal sentence. In such a case, the verb does not agree with any noun phrase (2.126).

(2.126) (Cat 06019900-19)

```

<morfema.verbal func="impers">
  <p gen="c" lem="es" num="c" wd="Es"/>
</morfema.verbal>
<grup.verb>
  <v [...] postype="main" tense="present" wd="tracta"/>
</grup.verb>
<sp>

```

```

<prep>
  <s [...] num="c" postype="preposition" wd="d'"/>
</prep>
<sn>
  <spec gen="f" num="s">
    <d [...] num="s" postype="indefinite" wd="una"/>
  </spec>
  <grup.nom gen="f" num="s">
    <n [...] num="s" postype="common" wd="mostra"/>

```

‘it is an exhibition’

3. ‘es’ as a noun phrase (sn)

es is a pronoun which is the head of a noun phrase and functions as the direct or indirect object of the sentence (2.127):

(2.127) (Cat 104nou)

```

<sn>
  <grup.nom>
    <p [...] num="c" person="3" postype="personal" wd="s'"/>
  </grup.nom>
</sn>
<grup.verb>
  <v [...] postype="main" tense="present" wd="estan"/>
  <gerundi>
    <v [...] mood="gerund" postype="main" wd="disputant"/>
  </gerundi>
</grup.verb>
<sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="amb"/>
  </prep>
  <sn>
    <spec gen="f" num="p">
      <d [...] num="p" postype="indefinite" wd="sengles"/>
    </spec>
    <grup.nom gen="f" num="p">
      <n [...] num="p" postype="common" wd="OPES"/>
    </grup.nom>

```

‘(they) are competing for each of their takeover bids’

2.5 Sentence level

Sentence delimitation is carried out on the basis of orthographic conventions. Any string of words between two full stops or strong punctuation marks (question and exclamation marks as well as certain suspension dots) is considered to be a sentence. The system delimits only unambiguous sentences, so it is possible to find non-splitting passages. In such cases we should do it manually. The XML strings below show the feature-value pairs for sentence nodes. In addition, they can be given other values, such as <verbless>, <coord>, etc. In case of embedded clauses, they are given the corresponding values for <clausetype> or <clausesubtype> features (See the following sections for further details on each kind of sentence). No root node establishes a distinction as regards the traditional terminology about grammatical modality. That is to say, interrogative, declarative, exclamative, desiderative and dubitative mood are considered and tagged the same way.

```
<nodename name="S" description="Sentence/clause" onlyterminal="no" />
```

```
<nodename name="S" description="Sentence
/ clause" onlyterminal="no" />
<attribute name="verbless" parents="S,sentence"
description="Does clause / sentence have a main verb?">
<value name="yes" description="Clause /
sentence does not have a main verb" />
<value name="no" description="Clause
/ sentence has a main verb" />
</attribute>
```

```
<attribute name="impersonal" parents="S,sentence"
description="Is it an impersonal clause / sentence?">
<value name="yes" />
<value name="no" />
</attribute>
```

2.5.1 Non-finite sentences

Non-finite embedded sentences include either the verbal form alone (infinitive, gerund or participle) or the verbal form and its complements. Under no circumstance an empty node for the subject is added. Non-finite clauses can depend both on the main sentence (verbal complement) or on some other element within the sentence (noun, adjective or adverb complement).

```
<attribute name="clausetype" parents="S"
description="Clause type">
<value name="adverbial" description="Adverbial" />
<value name="completive" description="Completive" />
<value name="relative"
```

```

    description="Relative clause" />
<value name="participle"
  description="Participle clause" />
</attribute>

```

2.5.1.1 Non-finite completive clause

Non-finite completive sentences have an infinitive form¹⁹ as a head, and may have other complements. It is important to distinguish between those infinitives that constitute the head of a non-finite completive clause and those which are a part of an infinitive periphrasis²⁰. In order to check whether or not an infinitive form is the head of a non-finite clause, we should be able to replace the whole clause by a noun element. This element, in its turn, should be able to be placed before or after the whole group. These two criteria should be fulfilled, otherwise the structure is an infinitive periphrasis instead of a non-finite completive clause. Figure 2.128 shows the representation of an infinitive clause and figure 2.129 illustrates an infinitive periphrasis:

```
<S clausetype="completive" impersonal="yes" tem="pat">
```

(2.128) (Cat 06010300-28)

...pretén il·luminar aquest diumenge el puig de Sant_Martirià
 ‘... pretends to illuminate next Sunday Sant_Martirià’s mount’

(2.129) (Cat 06010100-4)

Aquest mateix vaixell tornarà a atracar al port de Tarragona
 ‘This very boat will tie up again at the seaport of Tarragona’

2.5.1.2 Non-finite adjective clause and absolute clause

The head of a non-finite adjective sentence is a deverbal adjective (<participle>), which may have other complements²¹. The following is the string marked at the sentence node:

```
<S clausetype="participle" impersonal="yes">
```

(2.130) (Cat 211nou)

documents confiscats a la Generalitat per les tropes de Franco
 ‘documents confiscated to the Generalitat by Franco’s troops’

The head of an absolute clause is a participle which is marked as a verbal form instead of an adjective. At the sentence node, moreover, we add a feature that indicates that the clause has an overt subject:

```
<S clausesubtype="suj" clausetype="participle" impersonal="yes">
```

¹⁹ Infinitives are always marked as <infinitiu>, both in completive clauses and in verbal periphrases.

²⁰ See section 2.4.5.1 for more details on verbal periphrases.

²¹ See section 2.4.5.1 for more details on possible complements of participle sentences.

Apart from the head, these clauses have their own overt subject (usually postponed); otherwise the structure is a simple non-finite adjective phrase. They can also be complemented by other elements. The internal node for absolute clauses is <grup.verb>²², as shown in 2.131.

(2.131) (Cat 211nou)

```
<S clausesubtype="subj" clausetype="participle" impersonal="yes">
  <grup.verb>
    <v [...] mood="pastparticiple" postype="main" wd="Donada"/>
  </grup.verb>
  <sn>
    <spec gen="f" num="s">
      <d [...] postype="demonstrative" wd="aquesta"/>
    </spec>
    <grup.nom gen="f" num="s">
      <n [...] postype="common" sense="16:10035141" wd="situació"/>
    </grup.nom>
  </sn>
</S>
```

‘Given this situation, ...’

2.5.1.3 Non-finite adverbial clause

The head of a non-finite adverbial sentence is a gerund²³, which may have other complements. They can be either a verb complement (2.132) or a noun complement (2.133):

```
<S clausetype="adverbial" impersonal="yes">
```

(2.132) (Cat 06010100-22)

*... no van ser avaluades per la comissió ... **incomplint així un acord del Parlament***

‘... (they) were not evaluated by the committee ... failing to fulfil, therefore, a Parliament’s agreement’

(2.133) (Cat 203-12)

*la pràctica totalitat d’alts càrrecs del departament, **començant** pel propi conseller*
‘almost all the high ranking officials of the department, starting with the very councillor’

²²This is one of the few cases in which a participle is labeled this way. Another exception is that of coordination of compound verbal tenses, where the auxiliary form can drop in one or more coordinated structures.

²³Gerunds are always marked as <gerundi>, both in non-finite embedded clauses and in verb periphrases.

This kind of clauses can occur with the adverb *tot* (emphazising particle) anteposed. In such a case, the adverb depends on the sentence node, as shown in 2.134:

(2.134) (Cat 107-3)

..., *tot seguint la línia de descentralització que s'ha imposat*
 '... following the decentralization line that has been imposed'

2.5.1.4 Non-finite relative clause

Non-finite relative clauses consist of a relative pronoun followed by an infinitive. They are always a noun complement and may have other complements in its turn.

(2.135) (Cat 06019900-11)

llocs on aprendre català
 'places where to learn Catalan'

2.5.2 Finite sentences

Regarding finite subordinate clauses, we establish the following classification: Completive, relative and adverbial finite clauses. Adverbial clauses, in their turn, are divided into two groups: The first group includes those functioning as circumstantial complements (place, time, manner, cause and finality); the second group is formed by conditional, concessive, consecutive and comparative clauses. This latter group does not work as circumstantial complements but as a complement affecting the whole sentence to which they belong.

```
- <attribute name="clausesubtype" parents="S"
description="Clause subtype">
  <value name="comparative" implies="clausetype:adverbial"
description="Adverbial comparative clause" />
  <value name="concessive" description="Adverbial
concessive clause" implies="clausetype:adverbial" />
  <value name="conditional" description="Adverbial
conditional clause" implies="clausetype:adverbial" />
  <value name="consecutive" description="Adverbial
consecutive clause" implies="clausetype:adverbial" />
  <value name="suj" implies="clausetype:participle"
description="Participle clause with subject" />
</attribute>
```

Finite clauses can also coordinate or may not have an overt verb. In such cases, they are marked as <coord> or <verbless>, respectively. Regarding coordination of finite clauses, the complementizer (relative pronoun or conjunction) usually occurs just once, introducing only the first coordinated clause. In such a case, the complementizer should depend on the whole coordinating structure; that is, as a sister node of the coordinated elements²⁴.

²⁴See section 2.6.5 for more details on the elipsis of the complementizer in coordination.

2.5.2.1 Finite completive clause

We include in this group those subordinate clauses that function as nouns and those relative clauses with no antecedent. The former are introduced by a conjunction (*que* ‘that’, *com* ‘how’ or *si*²⁵ ‘if, whether’) or by interrogative pronouns (*quan* ‘when’, *com* ‘as’, *on* ‘where’, *què*, ‘what’ *quin* ‘which’) or interrogative determiniers (*quin*, *-a* ‘which’) in indirect questions. Relative clauses without antecedent are introduced by the relative pronoun *qui* ‘who’, *quan* ‘when’, *com* ‘how’, etc.

1. Finite completive clauses introduced by *que*

They can complement verbs (2.136), nouns (2.137) and adverbs (2.138):

(2.136) (Cat 211nou)

L'ajuntament de Sabadell demanarà a la Generalitat [que s'obri un debat comarcal]

‘The city hall of Sabadell will ask the Generalitat for a regional debate’

(2.137) (Cat 111nou)

la possibilitat [que el Govern presenti aquest projecte a l'ONU]

‘the possibility that the Government presents this project to the UN’

(2.138) (111nou)

després [que el grup socialista bloquegés la seva elecció la setmana passada]

‘after the socialist party blocked their election last week’

2. Finite completive clauses introduced by *com*

(2.139) (Cat 06019900-5)

...van observar [com diversos camions es dirigien cap a...]

‘...(they) observed how some trucks were going to...’

1. Finite completive clauses introduced by *si*

(2.140) (Cat 211-8)

... conèixer [si el departament de Medi_Ambient de la Generalitat pot garantir la seguretat]

‘... to know whether the Generalitat Environmental Department is able to guarantee the security’

3. Indirect questions

(2.141) (Cat 211nou)

hauria de discutir-se [com s'articulària la participació d'Obiols]

‘it should be discussed how the participation of Obiols would be articulated’

²⁵In indirect questions or in dubitative sentences, like: *no sé si...* ‘(I) don’t know whether...’

(2.142) (Cat 111nou)

Vull saber [què en pensa Rodríguez_Zapatero]

‘I want to know what does Rodríguez_Zapatero think about it’

4. Finite completive clauses introduced by interrogative determiners

(2.143) (Cat 111nou)

L’únic dubte ... que em queda en aquesta qüestió és [quina posició mantindrà el Partit_Socialista]

‘The only doubt ... I have as regards this issue is which position the Socialist Party will maintain’

5. Finite completive clauses introduced by a relative pronoun without antecedent

(2.144) (Cat 111nou)

[qui ha canviat de criteri] és Maragall

‘the one who has changed his mind is Maragall’

6. Finite completive clauses introduced by a preposition

Prepositions can only introduce interrogative clauses or completive clauses introduced by *si* (2.145):

(2.145) (Cat 104nou)

des [que el PP va aconseguir la majoria absoluta]

‘since the PP party achieved the absolute majority’

6. Special cases

- (a) The subordinating conjunction *perquè* ‘because’ does not always introduce an adverbial clause but a finite completive clause, because it is a complement ruled by the verb²⁶, as shown in 2.146.

(2.146) (Cat 111nou)

ha advocat ... [perquè el partit no prescindeixi de persones com Serra i Obiols]

‘he pleaded ... for the party to no do without people like Serra and Obiols’

- (b) In structures such as the one in 2.147, the finite clause has neither covert subject nor other kind of complement because this subject or complement is the relative pronoun of the relative clause that, in its turn, has no syntactic function within the relative clause²⁷

²⁶ See section 2.4.4 for more details on prepositional phrase attachment.

²⁷ This case of discontinuity is solved at the function annotation stage: The relative pronoun is given the label corresponding to the syntactic function it has within the completive clause.

(2.147) (Cat 06020300-24)

*els ajuts **que** el Departament_d'_Agricultura preveu [concedir a aquests productors]*²⁸

'the incentives that the Ministry of Agriculture expects to concede'

- (c) Some time expressions²⁹ in which a third person singular verbal form occurs are treated as completely flat-structured noun phrases (*fa dos anys...* 'two years ago...'). These expressions can be followed by the subordinate conjunction *que* 'that'; in such cases, the conjunction directly depends on the sentence root node rather than introducing a completive clause, as showed in 2.148.

(2.148) (Cat 06010300-23)

***Fa anys que** Jaume_Mestres recorre països exòtics*

'It is years since Jaume Mestres travels all around exotic countries'

2.5.2.2 Finite relative clause

We include within this group those sentences introduced by a relative pronoun or adverb with an overt antecedent, this being usually a noun, although it can be any other nominalized element as well as an adverb. The relative pronoun is marked as <relatiu> (never as <sn> or <sadv>). Both restrictive -or identifying- (2.149) and non-restrictive -or non-identifying- (2.149) relative clauses are given the same representation. They differ just in the fact that non-restrictive clauses appear between commas:

(2.149) (Cat 211nou)

*les trobades [**que** està mantenint amb plataformes d'opinió]*

'the meetings (s/he) is mantaining with some opinion platforms'

(2.150) (Cat 06010300-28)

*La iniciativa, [**que** dependrà de les condicions meteorològiques, forma part de ...]*

'The initiative, which will depend on the weather conditions, belongs to ...'

1. Relative clause introduced by *que*

(2.151) (Cat 06010100-13)

```
<grup.nom gen="f" num="s">
  <n gen="f" lem="anàlisi" num="s" postype="common" wd="anàlisi"/>
  <S clausetype="relative" impersonal="no">
    <relatiu>
      <p gen="c" lem="que" num="c" postype="relative" wd="que"/>
    </relatiu>
    <morfema.verbal func="pass">
```

²⁸Here the relative pronoun has no syntactic function within the relative clause, while it is the direct object of the rightmost verb, that is, the main verb of the completive clause: *The Ministry of Agriculture expects to concede that [the incentives].*

²⁹See section 2.4.1.6 for more details on other elements we treat as noun phrases.

```

    <p gen="c" lem="es" num="c" wd="s'"/>
  </morfema.verbal>
  <grup.verb>
    <v [...] postype="auxiliary" tense="present" wd="ha"/>
    <v [...] mood="pastparticiple" postype="main" wd="fet"/>

```

‘the analysis that has been made’

2. Relative clause introduced by a stressed relative pronoun

(2.152) (Cat 211nou)

*al congrés de novembre [en **què** es preveu elevar a Mas a la secretaria general del partit]*

‘at the November congress, in which it is predicted to promote Mas to the party general secretary’

3. Relative clause introduced by *el qual, els quals, la qual, les quals* ‘which’

(2.153) (Cat 211nou)

*el 1999, any [en el **qual** es van celebrar eleccions municipals i autonòmiques]*

‘in 1999, year in which there were municipal and regional elections’

4. Relative clause introduced by *on* (‘where’)

(2.154) (Cat 111nou)

*una societat [**on** l'accés a un habitatge digne ha deixat de ser un dret]*

‘a society in which the access to a decent housing is not a right anymore’

5. Relative clause introduced by *quan* (‘when’)

(2.155) (Cat 111nou)

*... el 1996, [**quan** es va ajornar sense data concreta el trasllat...]*

‘... in 1996, when the transfer was postponed without a concrete date ...’

In those cases of ambiguity between relative pronoun and conjunction, we should choose the flattest one, which is usually the finite adverbial clause (2.156):

(2.156) (Cat 211nou)

*les [que ja es van obrir fa uns dies [**quan** un altre grup de funcionaris va cremar pneumàtics <adverbial>]<relatiu>]*

‘the ones that were already opened some days ago when another group of public officials burned pneumatics’

6. Relative clause introduced by *com* (‘as, since’)

(2.157) (Cat 211nou)

... *cedir terrenys municipals a Badia tal [com havia demanat l'alcade
...]*<relatiu>

'... to cede municipal grounds to Badia, as the major had asked for ...'

*Relative clause introduced by qui ('who')*³⁰

(2.158) (Cat 111nou)

... *president del club ...*, [*qui en una entrevista radiofònica va criticar
JordiPujol*]

'... the president of the club..., who in a radio interview criticized JordiPujol'

7. Relative clause introduced by a relative determiner *la qual cosa* ('what')

We consider those sentences introduced by relative determiners such as that of *la qual cosa* to be relative clauses. Since they affect the whole sentence, they belong to the sentence node <S> or <sentence>.

(2.159) (Cat 104nou)

... *les condicions són diferents de les de fa uns anys, [per la qual cosa és més
necessària la col·laboració ...]*

'... the conditions are different from those of some years ago, for what it is more necessary the collaboration ...'

8. Superlatives in relative clauses

Those comparative structures with relative superlatives are considered to be a plain relative clause (<relative>) rather than a comparative clause (<adverbial,comparative>):

(2.160) (Cat 06019900-28)

[la festa [més important] [que se celebra en aquesta població]<relative>]<sn>

'the most important party that is celebrated in this town'

9. Special cases

(a) Nominalization

Relative clauses without antecedent and introduced by an article³¹ are considered as relatives rather than completives; the article is the mark of the syntactic nominalization³²:

(2.161) (Cat 211nou)

[el [que menys em preocupa en aquest moment]<relative>]<sn>

'what worries me the least in this moment'

³⁰The relative pronoun *qui* having an overt antecedent; otherwise it introduces a completive clause. See section 2.5.2.1. for more details on completives.

³¹When there is an antecedent the relative pronoun that is used is *el qual*. See point 3 in this section in the previous section for more details on this issue.

³²See section 2.4.1.5 for more details on nominalization.

When the nominalization is carried out through the contracted form of an article plus a preposition, the node for the specifier (the one indicating the nominalization) is lost. In such cases we should add the noun phrase node <sn>:

(2.162) (Cat 111nou)

les conductes [dels qui sense violència ocupen un immoble]
 ‘the behavior of those who without violence occupy a building’

(b) Preposition + relative pronoun

In some cases a preposition precedes the relative pronoun. When there is an overt antecedent, the preposition is placed within the embedded clause and introduces a prepositional phrase which includes the relative pronoun:

(2.163) (Cat 111nou)

Alberto_Ruiz_Gallardón, [de qui recollirà informació sobre el finançament]
 ‘Alberto_Ruiz_Gallardón, from who (s/he) will get information on the financing’

Sometimes this relative pronoun belongs to a more complex structure and therefore it is placed further from the mother node (2.164).

(2.164) (Cat 211nou)

Els dirigents territorials, [el poder dels quals és clau]
 ‘The regional leaders, whose power is a key factor’

(c) *TAL COM* (‘as’)

When co-occurring as in example 2.165 below, *tal* is a demonstrative pronoun while *com* introduces an embedded relative clause and therefore the demonstrative is its antecedent:

(2.165) (Cat 211nou)

...cedir terrenys municipals a Badia tal [com havia demanat l’alcalde]
 ‘to cede municipal grounds to Badia as the major had asked for’

(d) Relative pronoun playing no function within the clause

In example 2.166 below, the relative pronoun plays no syntactic function within the relative clause, but it is a complement of the embedded completive clause³³. In the example, the pronoun is the direct object of the infinitive clause:

(2.166) (Cat 06020300)

els ajuts [que el Departament_d’_Agricultura preveu [concedir]]
 ‘the incentives that the Ministry of Agriculture expects to concede’

(e) Ellipsis within the relative clause

Sometimes, we can find finite relative clauses without relative pronoun and without a verb:

³³This case of discontinuity is solved at the function annotation stage: The relative pronoun is given the label corresponding to the syntactic function it has within the completive clause.

(2.167) (Cat 203-14)

```

<sn>
  <grup.nom gen="m" num="s">
    <n gen="c" lem="Fons" num="c" postype="proper" wd="Fons"/>
    <S clausetype="relative" impersonal="no" verbless="yes">
      <f lem="," punct="comma" wd=","/>
    <sp>
      <prep>
        <s [...] num="c" postype="preposition" wd="sota"/>
      </prep>
    <sn>
      <grup.nom gen="f" num="s">
        <n [...] num="s" postype="common" wd="reserva"/>
  'Funds under reserve'

```

(f) Complements of a relative pronoun

In principle, the relative pronoun cannot be complemented. However, there are special cases, like the one in example 2.168, in which this is the only possible analysis.

(2.168) (Cat 104nou)

un dels consorcis [que, integrat per Premsa_Ibèrica, Corporació_Accionarial_Lara ..., opten als dos canals]
 'one of the consortiums that, integrated by 'Premsa_Ibèrica, Corporació_Accionarial_Lara..., opt for the two channels'

(g) Coordination of relative clauses

In some cases, although two or more relative clauses are in coordination, the pronoun appears just once. We should represent this pronoun as a daughter of the coordinated node and sister of the coordinated structures³⁴. Sometimes the complementizer is placed within a phrase (2.169). Other times, there are attached complements between the complementizer and the coordinated relative clauses (2.170):

(2.169) (Cat 204nou)

accions [de les quals 144 pugen, 72 baixen i 28 no varien]
 'shares from which 144 are rising, 72 are dropping and 28 have no change'

(2.170) (Cat 111nou)

Obiols, [que en el pròxim congrés del PSC, al juny, debatrà el seu càrrec de president]
 'Obiols, who in the next PSC party congress, next June, will debate his president office'

³⁴See section 2.6.5 for more details on ellipsis and coordination.

2.5.2.3 Finite adverbial clause

Whithin this set of sentences we include time, place, modal, cause and final clauses. All of them are marked as follows at the sentence node:

```
<S adjunct="yes" clausetype="adverbial" impersonal="no">
```

1. Place clauses

On introduces an adverbial clause when it is a subordinate conjunction (<c,subordinating>). When it has an overt antecedent it is a relative pronoun, introducing therefore a relative clause. When it has no overt antecedent it is an adverb (<r>) and introduces a completive clause.

2. Time clauses

Abans ‘before’; *així que* ‘so that’; *a mesura que* ‘insofar, to the extent’; *a penes* ‘as soon as’; *cada vegada que* ‘every time that’; *de seguida que* ‘as soon as’; *mentre* ‘while’; *quan* ‘when’; *que* ‘that’; *tan aviat com* ‘as soon as’; *un cop* ‘once’.

Un cop can also work as an adverbial locution when followed by a participle or a prepositional phrase.

Time clauses can be both verbal (2.171) and noun complements (2.172).

(2.171) (Cat 111nou)

va ser agredit, segons la carta enviada per Molins, [quan intentava evitar els incidents]

‘(he) was attacked, according to the letter sent by Molins, when (he) was trying to prevent the incidents’

(2.172) (Cat 211nou)

la inhalació de gasos tòxics [quan estaven a l'interior d'un pou]

‘the inhalation of toxic fumes when (they) were inside a shaft’

When there is ambiguity between the relative pronoun and the conjunction, we should take the most straightforward representation; that is, the flattest one is usually the adverbial clause.

3. Modal clauses

Com ‘as’; *que* ‘that’; *segons (que)* (2.173) ‘according to’; *tal i com* ‘just as’. We also treat as adverbial clauses those structures equivalent to *com per exemple* ‘as for example’, when they go along with a verb.

(2.173) (Cat 104nou)

[Segons ha explicat avui el vicesecretari general i senador d'ERC, Carles Bonet], el seu partit enviarà ...

‘According to what explained today the ERC assistant secretary and senator, Carles Bonet, his party will send...’

4. Cause clauses

A_causa_que ‘because of’; *atès_que* ‘given that, since’; *car* ‘since’; *com_que* ‘as, since’; *en_la_mesura_que* ‘provided that, as long as’; *en_tant_que* ‘provided that’; *ja_que* ‘because’; *perquè* ‘because’; *per_tal_com* ‘as long as’; *puix (que)* ‘given that, since’; *que* ‘that’; *vist_que* ‘since’. The subordinate conjunction *perquè* ‘because’ can introduce both a cause or final clause.

(2.174) (Cat 06010100-25)

perquè no disposaven en aquell moment d'un PGOU o bé de Normes Subsidiàries de planejament municipal

‘because (they) did not have at that moment any PGOU neither subsidiary regulations on municipal planning’

5. Final clauses

A_fi_que ‘in order that’; *perquè* ‘so that’; *de_forma_que* ‘so that, in order to’; *per_tal_que* ‘in order to’; *que* ‘that’.

Final clauses can be both verbal (2.175) and noun (2.176) complements:

(2.175) (Cat 104nou)

el_seu partit enviarà un informe ... a associacions d'empresaris ... [perquè es sumin a treballar en aquesta entitat]

‘(his/her) party will send a report ... to employers associations ... so that they join this entity to work’

(2.176) (Cat 06010100-36)

una moció de CiU [per_tal_que la UE reconegui el català com_a idioma oficial]

‘a motion of CiU so that the EU recongnizes Catalan as an official language’

Sequences consisting of an adverb followed by the particle *que* ‘that’ (*abans que*, *després que*, etc.) are treated as adverbial phrases (where the adverb introduces a completive finite clause) instead of adverbial locutions. An exception is that of *mentre que* ‘whereas’, which is treated as a coordinate conjunction.

Some of these sequences show ambiguity. *Sempre que*, for instance, can have a conditional value, meaning ‘as long as, provided that’, therefore ruling subjunctive mood. In such a case we treat the sequence as a conjunctive locution. When, on the contrary, it has a time value (‘always that, every time that’), it rules indicative mood. We should split the sequence up and treat the group as an adverbial phrase followed by a relative clause.

2.5.2.4 Other adverbial clauses (‘impròpies’)

Even though there is no difference regarding their representation, we do make a distinction between place, time, modal, cause and final clauses, on the one hand, and conditional, concessive, consecutive and comparative clauses, on the other hand. The following are the corresponding XML-feature values:

```

<value name="comparative" implies="clausetype:adverbial"
  description="Adverbial comparative clause" />
<value name="concessive" description="Adverbial
  concessive clause" implies="clausetype:adverbial" />
<value name="conditional" description="Adverbial
  conditional clause" implies="clausetype:adverbial" />
<value name="consecutive" description="Adverbial
  consecutive clause" implies="clausetype:adverbial" />

```

1. Conditional clause

Conditional clauses are introduced by the following conjunctions and conjunctive locutions³⁵: *a condició que* ‘on the condition that’; *amb que* ‘provided that’; *si* ‘if’; (*en/en el*) *cas que* ‘if, in the event of’; *mentre* (+ subjunctive) ‘as long as’; *només que* ‘just by’; *posat que* ‘provided that’; *sempre i quan* ‘as long as, provided that’; *sempre que* ‘provided that’. The sentence node is marked as follows:

```

<S adjunct="yes" clausesubtype="conditional"
  clausetype="adverbial" impersonal="no">

```

(2.177) (Cat 111nou)

*[mentre no rebutgi la iniciativa del PSC] haurà d'explicar-ho a
Castella_i_Lleó*

‘while (s/he) does not reject the PSC’s initiative (s/he) will have to explain
that to Castella_i_Lleó’

Those structures containing the sequence *si no* ‘if not, otherwise’ (2.178) are always analyzed as conditional clauses:

(2.178) (Cat 111nou)

[perquè si no], no aconseguirem el clima de serenitat necessari

‘because, otherwise, we will not achieve the necessary atmosphere of serenity’

2. Concessive clause

The following are conjunctions and conjunctive locutions that, among others, can introduce a concessive clause³⁶: *al marge que* ‘beyond, outside’; *bé que* ‘in spite of’; *fora que* ‘although’; *llevat que* ‘unless’; *malgrat que* ‘although, in spite of’; *per més que* ‘even though, although’; *quan* ‘when’; *si bé* ‘even though, although’; *tot i que* ‘even though, although’. These clauses are marked as the following string at the sentence node:

```

<S clausesubtype="concessive"
  clausetype="adverbial" impersonal="no">

```

³⁵See appendix for a complete list of locutions.

³⁶See appendix for a complete list of locutions.

(2.179) (Cat 204nou)

*[Malgrat_ que en la passada Setmana_ Santa no es va produir cap incident],
la demanda del servei va superar ...*

‘Although last Easter there was no incident at all, the service’s demand exceeded ...’

3. Consecutive clause

Consecutive clauses are introduced by the following conjunctions and conjunctive locutions, among others³⁷: *a_ conseqüència_ que* ‘as a consequence of, as a result of’; *així_ que* ‘so that’; *de_ manera_ que* ‘so (that)’; *que* ‘that’; *per_ tant* ‘therefore’; *fins_ al_ punt_ que* ‘to the point that’; *de_ forma_ que* ‘so (that)’. The XML string is the following:

```
<S clausesubtype="consecutive"
clausetype="adverbial" impersonal="no">
```

(2.180) (Cat 06010300-25)

per_ tant els problemes continuen
‘therefore the problems are still there’

Another type of consecutive clause is that containing a correlative connection; that is, a quantifier as a first part and a conjunction as a second one: *tan* + Adj/Adv + *que* ‘so + Adj/Adv + that’; *tant* + Noun/Verb + *que* ‘so + Noun/Verb + that’... In this case, the subordinate clause is adjoined to the node containing the quantifier.

4. Comparative clause

Comparative clauses always contain correlative connectors: *menys ... que* ‘less ... than’; *menor que* ‘less/smaller than’; *pitjor que* ‘worse than’; *tan ... com* ‘as ... as’; *igual que* ‘the same as’; *més ... que* ‘more ... than’; *major que* ‘more/bigger than’; *millor que* ‘better than’, etc. All comparative clauses are adjoined to the node containing the first part of the connector. The following string is the one at the sentence node of comparative clauses:

```
<S adjunct="yes" clausesubtype="comparative"
clausetype="adverbial" impersonal="no" verbless="yes">
```

(2.181) (Cat 06029900-6)

ser igual que al carrer
‘to be the same as in the street’

(2.182) (Cat 06029900-11)

les mateixes que nosaltres detectem
‘the same that we detect’

³⁷See appendix for a complete list of locutions.

Sometimes, and due to the free constituent order of the language, it is impossible to adjoin a comparative clause to the node containing the adverb. In such a case, we solve the discontinuity by marking both head and complement³⁸

2.5.2.5 The superlative

There are two types of superlative structures. The first one contains a relative clause and it complements a noun, as shown in example 2.183 below. The second type is expressed by means of a prepositional phrase: *menys . . . de* ‘(the) least . . . of’; *més . . . de* ‘(the) most . . . of’, etc. They are also noun complements, as shown in 2.184:

(2.183) (Cat 06019900-28)

la festa més important que se celebra en aquesta població
‘the most important party that is celebrated in this town’

(2.184) (Cat 06010300-24)

la xarxa més extensa de Catalunya
‘the widest net in Catalonia’

2.6 Coordination

Coordination is marked by the XML feature values explained in section 3.4. The coordinated elements and the element indicating the coordination are sisters and there is no hierarchical relationship between them. All sentence structures can be coordinated, as well as all kind of phrases and specifiers.

Coordination must be done at the lowest possible level. Certain structures, however, pose difficulties in doing so:

- Complements and/or specifiers affecting the noun in different ways make impossible to coordinate them at the group level (<grup.nom>, <grup.a>, etc.). In such cases, the coordination is at the noun phrase (<sn>) level³⁹, as shown in 2.185:

(2.185) (Cat 06020300-18)

el PHN garantirà [[aigua de qualitat per a l'àrea metropolitana de Barcelona] i [la recuperació d'aqüífers]]<sn,coord>
‘the PHN will guarantee water of quality for the Barcelona metropolitan area and the aquifers recovery’

- When the negative adverb ‘no’ affects the verb; that is, the whole sentence instead of a phrase, makes the coordination to be placed at the highest level, as illustrated in 2.186:

³⁸See section 2.8 for further details on discontinuity.

³⁹See section 2.4.1.7 for more details on coordination within the noun phrase.

(2.186) (Cat 06010100-27)

*[[incloïa el tractament paisatgístic] i [no el cobriment del torrent]]*_{<S,coord>}
 ‘it included the landscape treatment but not the coating of the torrent’

2.6.1 Coordinating elements

Coordinating elements are daughters of the coordinated structure and sisters of the structures that they coordinate. We make no distinction between coordination and juxtaposition. Therefore, coordinating elements can be conjunctions, conjunctive locutions (labeled as <coord>) as well as punctuation marks.

```
- <attribute name="conjunctiontype"
description="Conjunction type" parents="conj">
  <value name="coordinating" description="Coordinating conjunction" />
</attribute>
```

1. **Coordinating conjunctions:** *i* ‘and’; *mentre* ‘while’; *més* (rg) ‘more’; *ni* ‘neither,nor’; *o* ‘or’; *però* ‘but’; *sinó* ‘(no...) but’; *tanmateix* ‘as well’.

(2.187) (Cat 06010300-23)

*la companyia de [[dues càmeres fotogràfiques], [objectius], [munts de rodets] i [un equipatge de subsistència mínim]]*_{<sn,coord>}
 ‘the company of two photograph cameras, lenses, a lot of films and a minimal survival equipment’

Però and *tanmateix* can be coordinating conjunctions (<c,coordinating>) as well as adverbs (<r>) when they occur between commas:

(2.188) (Cat 211nou)

*una condició [[necessària] però no suficient]]*_{<s.a,coord>}
 ‘a necessary but non-sufficient condition’

Although *mentre* usually introduces a time clause, it can function as a coordinating conjunction (equivalent to *mentre_ que* ‘while, whereas’) when it lacks of the time nuance (2.189):

(2.189) (Cat 06010300-2)

*[[Les víctimes són dos joves de nacionalitat colombiana, de 25 i 31 anys,] mentre [els detinguts són tres homes de nacionalitat colombiana, marroquina i dominicana]]*_{<S,coord>}
 ‘The victims are two young people of Colombian citizenship, of 25 and 31 years old, while the detainees are three men of Colombian, Moroccan and Dominican citizenship’

Another case is that of *no... sinó...* , which does not coordinate phrases but clauses because the negative adverb *no* affects just one of the clauses and *sinó* coordinates both of them. In such a case the negative adverb is labeled as <r,neg> (2.190):

(2.190) (Cat 06010100-21)

tres van morir però [[no a_causa_del bacteri] sinó [pels motius que ...]]<S,coord>

‘three of them died but not because of the bacterium but because of the reasons that...’

Likewise, *no... sinó_que* does not coordinate phrases but sentences, because the negative adverb *no* affects just one of the sentences and *sinó_que* coordinates both of them. In such a case the negative adverb is labeled as <r,neg>, as shown in 2.191:

(2.191) (Cat 06030200-6)

els militants [[no resten expulsats] sinó_que [temporalment s'adscriuran a una altra agrupació]]<S,coord>

‘the militants are not expelled, but they temporarily will adscribe to another grouping’

2. **Conjunctive locutions:** *a_banda_de* ‘apart from’; *així_com* ‘as well as’; *a_part_de* ‘apart from’; *juntament_amb* ‘together with’ *mentre_que* ‘while’; *no_obstant* ‘nevertheless’.⁴⁰

(2.192) (Cat 111nou)

[[les anteriors iniciatives ... havien partit de CiU], mentre_que [aquesta neix del PSC]]<S,coord>

‘the previous initiatives ... started from CiU, while this one arises from the PSC party’

amb_tot and *així_i_tot* can be coordinating conjunctions (<c,coordinating>) as well as adverbs (<r>) when they occur between commas, as it happens to *però* and *tanmateix*.

en_canvi, *no_obstant* and *per_contra* function as coordinating conjunctions (<c,coordinating>) when not preceding another coordinating conjunction. Otherwise they work out as adverbial locutions (<r>).

Although *en_lloc_de* and *en_comptes_de* are always marked at the morphological node as a preposition (<s>) they can function as coordinating conjunctions (<c,coord>) when introducing the last member of the coordination (they could be replaced by *i no* ‘and/but not’). Otherwise they are prepositions:

(2.193) (Cat 211nou)

[[enfortir el partit], [renovar-lo], en_lloc_d'[anar-lo desfent]]<S,coord>

‘to make the party stronger, renew it, instead break it up’

conjuntament_amb, *en_col.laboració_amb*, *en_coordinació_amb* and *juntament_amb* are always marked as prepositions (<s>) although they function as coordinating

⁴⁰See Appendix for a complete list of conjunctive locutions.

conjunctions (<c,coord>) when introducing the last member of the coordination. Otherwise, they work as prepositions:

(2.194) (Cat 06030200-5)

[[el gir dretà i pragmàtic d'UCD], juntament_amb [la defensa de la bondat del pacte CiU-PP]]<sn,coord>, contribueix a...

'the pragmatic change to the Right of the UCD party, together with the defense of the godness of the CIU-PP pact, contributes to ...'

al_marge_de, *a_més_de*, *a_més_a_més_de*, *a_banda_de* and *a_part_de* are always marked as prepositions at the morphological node (<s>) although they can function both as prepositions or coordinating conjunctions (as in 2.196). They are prepositions when preceding their antecedent and therefore separated from the series to which they belong:

(2.195) (Cat 204nou)

[[els majors descensos], a_més_a_més_de [Les_Balears]]<sn,coord>, es van registrar a ...

'the biggest falls, as well as the Balearic one, were registered at ...'

On the other hand, these elements function as coordinating conjunctions when introducing the last element of a series (and they can be replaced by *i_a_més* or *i_a_més_a_més*:

(2.196) (Cat 06010100-32)

[[participar en els òrgans de direcció del departament ...] a_més_de [vetllar pel compliment de les disposicions legals]]<S,coord>

'to participate in the directive organs of the department as well as to look after the performance of the legal regulations'

A complete list of coordinating elements is found in Appendix.

2.6.2 Distributed coordination

In distributed coordination there are two connectors instead of one. Both of them are daughters of the coordinated structure and therefore sisters of the coordinated elements. Distributed coordinating elements are, among others⁴¹: *ni...ni...* 'nor...neither...'; *o...o...* 'or...or...'; *o...o_bé...* 'or...or...'; *bé...o...* '(or...) or...'

(2.197) (Cat 06010100-27)

[[ni per_part_de_l'Ajuntament] ni [de l'Institut_Català_del_Sòl]]<sp,coord>

'neither from the City Hall nor from the Institut Català del Sòl'

⁴¹See Appendix for a complete list of coordinating elements and coordinating locutions.

2.6.3 Distributed-like structures

We treat as distributed coordination those structures containing two connectors, these being sisters of the coordinated elements. The morphological node usually corresponds to a preposition or an adverb and the syntactic one is *coord* for both connectors. The following are some examples⁴².

1. **(Des_)de... a... ; (des_)de... fins... ; (des_)de... fins_ a... ‘from... to...’**

When they lack of a strict place nuance so that they just point out the first and the last members of a series (2.198):

(2.198) (Cat 06030200-19)

s'ha previst [[des_ del disseny d'[[un joc] i [una pàgina]]<sn,coord>] fins_ a [la creació d'una base de dades]]<sn,coord>
 ‘it is planned from the design of a game and a web to the creation of a database’

(2.199) (Cat 06020300-21)

[[des_ de l'1_ de_ gener] fins [el 31_ de_ maig_ 2003]]<sn,coord>
 ‘from January the 1st to May the 31st 2003’

2. **tant... com... ‘both... and...’**

(2.200) (Cat 111nou)

[[tant Artur_ Mas] com [Josep_ Antoni_ Duran_ Lleida]]<sn,coord>
 ‘both Artur Mas and Josep Antoni Duran Lleida’

3. **entre... i... ‘between... and...’**

If it is the case, *entre* can be removed from the sentence and, although there is a slightly difference in meaning, the whole structure is still grammatical:

(2.201) (Cat 06010100-3)

els fets van passar [[entre 1998] i [l'any_ 2000]]
 ‘the facts happened between 1998 and 2000’

4. **no_ només... sinó_ també; no_ tan_ sols... sinó_ també ‘not only... but also...’**

(a) General case

Each connector precedes its corresponding coordinated element, as in (2.202):

⁴²See Appendix for a complete list of coordinating elements.

(2.202) (Cat 111nou)

[[no_ només per militants del partit] sinó_ també [per independents i membres de sindicats]]<*sp,coord*>

‘not only for the party militants but also for independents and trade union’s members’

Other connectors meaning the same as *no_ només... sinó_ també* are treated the same way, as shown in 2.203:

(2.203) (Cat 111nou)

[[no perquè tinguem que obtenir resultats concrets] sinó [perquè s’ha de preparar el terreny]]<*S,coord*>

‘not because we have to get concrete results but because (someone) has to pave the way’

(b) Particular cases

The first connector occurs within the first coordinated element. In such a case, the first connector (*no_ només*) does not work as a coordinating element but as an adverbial phrase (<*sadv*>). Therefore, its representation does not correspond to that of a distributed coordination. Instead of that, it is coordinated to the second connector (*sinó_ també*):

(2.204) (Cat 06010100-6)

[[que contempli no_ només el nombre d’habitants] sinó_ també [que recuperi el dèficit d’inversió]]<*S,coord*>

‘that considers not only the number of inhabitants but also that recovers the inversion deficit’

```
<sadv>
  <grup.adv>
    <r lem="no_només" wd="no_només"/>
  </grup.adv>
</sadv>
<sn>
  <spec gen="m" num="s">
    <d [...] num="s" postype="article" wd="el"/>
  </spec>
  <grup.nom gen="m" num="s">
    <n [...] num="s" postype="common" wd="nombre"/>
    [...]
  </grup.nom>
</sn>
</S>
  <coord>
    <c [...] lem="sinó_també" wd="sinó_també"/>
  </coord>
[...]
```

The first connector is placed to the left of the first coordinated element. In such a case, we attach the first connector to the node that is on its right and we coordinate it to the second connector. Example in 2.205 illustrates this point:

(2.205) (Cat 111nou)

[[no_ només des_ d' un punt de vista econòmic], sinó_ també [social i territorial]]<sp,coord>]<sp,coord>

‘not only from an economic point of view, but also from a social and territorial one’

In some cases, the possibility of coordinating structures at the lowest possible level depends on whether or not the negative adverb *no* occurs within the first coordinated element. We should coordinate sentences instead of phrases:

no només vol X sinó també Y ‘(s/he) does not only want X but also Y’

We would coordinate phrases only in cases like the following:

vol no només X sinó també Y ‘(s/he) wants not only X but also Y’

5. *no_ només... sinó_ que_ també* ‘not only... but also...’

(a) General case

The first connector always occur within the first coordinated element. Therefore, the first connector (*no_ només*) does not function as a coordinating element but as an adverb. Instead of corresponding to a distributed coordination, its representation is equivalent to a simple coordination; that is, it is the second connector (*sinó_ que*) the one coordinating two elements, as shown in 2.206:

(2.206) (Cat 06030200-13)

[[no_ només afectarà al territori], sinó_ que [perpetua un model enèrgic contaminant que tard o d' hora s'haurà de canviar]]<S,coord>

‘not only (it) will affect the territory, but it also perpetuates an pollutant energetic model that must be changed’

(b) Particular cases

In those structures in which both connectors cannot be analyzed as sisters of the coordinated elements, the first coordinating conjunction must be attached to the structure which it introduces, so that this first structure can be coordinated with the second, as it is illustrated in example 2.207 below:

(2.207) (Cat 104)

[[tant des_ del punt de vista global ...] com [de la situació a Catalunya]]<sp,coord>

‘both from a global point of view ... and the Catalan situation’

2.6.4 Coordination of different elements

There are cases in which coordinated elements may belong to different categories. In such cases, the upper coordinating node is marked as being the same category of one of

its daughter nodes. The rest of the structure is given the corresponding values, as it is already explained in section 2.6. Examples of this kind of coordination are listed below.

2.6.4.1 Sentence structures

1. $\langle \mathbf{S} \rangle + \langle \mathbf{S}, \text{verbless} \rangle \Rightarrow \langle \mathbf{S}, \text{coord} \rangle$
(or $\langle \text{sentence} \rangle + \langle \text{sentence, verbless} \rangle \Rightarrow \langle \text{sentence, coord} \rangle$):

(2.208) (Cat 104nou)

[El primer consistiria en utilitzar el llit del riu] i [el segon a utilitzar vies de comunicació]

‘The first would consist on using the river-bed ... and the second on using communication paths’

2. $\langle \mathbf{S} \rangle + \langle \mathbf{S}, \text{completive, nonfinite} \rangle^{43} \Rightarrow \langle \mathbf{S}, \text{coord} \rangle$:

(2.209) (Cat 06030200-25)

[... ha provocat la banalització del paisatge litoral] a_més_de [generar una reducció de l'extensió]

‘... it caused the devaluation of the coast landscape as well as to generate the reduction of its extension’

3. $\langle \mathbf{S} \rangle + \langle \mathbf{S} \rangle^{44} \Rightarrow \langle \mathbf{S}, \text{coord} \rangle$:

(2.210) (Cat 104nou)

[que opera a Catalunya amb la marca Kampió] i [a Madrid i Andalusia amb l'ensenyà Cobreros]

‘...which operates in Catalonia with the brand Kampió and in Madrid and Andalusia with the brand Cobreros’

4. $\langle \mathbf{S}, \text{nonfinite} \rangle + \langle \mathbf{S}, \text{nonfinite} \rangle \Rightarrow \langle \mathbf{S}, \text{nonpersonal, coord} \rangle$:

(2.211) (Cat 104nou)

ma d'obra, tant [qualificada] com [no]

‘labour, both qualified and not’

5. $\langle \mathbf{S} \rangle + \langle \mathbf{S}, \text{nonfinite} \rangle \Rightarrow \langle \mathbf{S}, \text{coord} \rangle^{45}$:

(2.212) (Cat 211nou)

[perquè s'instal·lessin indústries] i [d'aquesta manera aconseguir recursos econòmics]

‘so some industries settle there and therefore to get economic resources’

⁴³Nonfinite completive clause

⁴⁴Any kind of finite clause (completive, adverbial, relative).

⁴⁵Any finite clause (completive, adverbial, relative) plus any nonfinite clause (completive, adverbial, relative, participle) results in a coordinated finite clause

6. $\langle \mathbf{S}, \text{nonfinite, participle} \rangle + \langle \mathbf{S}, \text{relative} \rangle \Rightarrow \langle \mathbf{S}, \text{nonfinite, participle, coord} \rangle$:

(2.213) (Cat 204nou)

capital de risc [dirigit a empreses catalanes ...] i [que compta amb un patrimoni inicial de 4000 milions]

‘a risk capital addressed to Catalan companies ... and which has an initial patrimony of 4000 millions’

7. $\langle \mathbf{S}, \text{nonfinite, participle} \rangle + \langle \mathbf{S}, \text{nonfinite, adverbial} \rangle \Rightarrow \langle \mathbf{S}, \text{nonfinite, participle, coord} \rangle$:

(2.214) (Cat 06010100-39)

[abatuda], [caminant sense rumb] i [demanant ajut als cotxes que passaven]

‘dejected, wandering aimlessly and asking the cars passing by for help’

2.6.4.2 Phrase structures

1. $\langle \mathbf{sp} \rangle + \langle \mathbf{sadv} \rangle \Rightarrow \langle \mathbf{sadv, coord} \rangle$ (when it is a verb complement):

(2.215) (Cat 06010100-30)

es produiran [de forma tranquil·la] i [de mica_en_mica]

‘(they) will be produced in a quiet way and step by step’

2. $\langle \mathbf{sp} \rangle + \langle \mathbf{sadv} \rangle \Rightarrow \langle \mathbf{sp, coord} \rangle$ (when it is a noun complement):

(2.216) (Cat 203-13)

referent [a tot l'Estat] i fins_i_tot [fora]

‘as regards the whole State and even abroad’

3. $\langle \mathbf{sp} \rangle + \langle \mathbf{sa} \rangle$ or $\langle \mathbf{s.a} \rangle \Rightarrow \langle \mathbf{sa, coord} \rangle$ or $\langle \mathbf{s.a, coord} \rangle$:

(2.217) (Cat 204nou)

les empreses [industrials] i [de serveis]

‘industrial and service-oriented companies’

4. $\langle \mathbf{sp} \rangle + \langle \mathbf{sn} \rangle \Rightarrow \langle \mathbf{sn, coord} \rangle$:

(2.218) (Cat 207-4)

Luis_Eduardo_V.D, [de 30 anys] i [veí de Lloret]

‘Luis Eduardo V.D, aged 30 and resident in Lloret’

5. $\langle \mathbf{s.a} \rangle + \langle \mathbf{sadv} \rangle = \langle \mathbf{s.a, coord} \rangle$:

(2.219) (Cat 111nou)

un perfil [més independent que un diputat], [més per_ sobre de la política activa]

‘a profile more independent than a member of parliament’s, more above than active politics’

6. $\langle \text{sadv} \rangle + \langle \text{sn} \rangle \Rightarrow \langle \text{sn, coord} \rangle$:

(2.220) (Cat 207.3)

[demà] i [diumenge]

‘tomorrow and Sunday’

2.6.4.3 Phrases + sentence structures1. $\langle \text{sn} \rangle + \langle \text{S, finite, completive} \rangle \Rightarrow \langle \text{sn, coord} \rangle$:

(2.221) (Cat 204nou)

[l’alta qualitat d’imatge i de so] i [que no precisa descodificar]

‘the high image and sound quality and that (it) does not require to decode’

2. $\langle \text{sp} \rangle + \langle \text{S, finite, completive} \rangle \Rightarrow \langle \text{sp, coord} \rangle$ (when it is a noun complement):

(2.222) (Cat 211nou)

el dret de totes les autonomies [que la_ seva posició sigui considerada ...] i [a estar presents en les delegacions]

‘the right of all self-governments for their position to be considered ... and to be present at the delegations’

3. $\langle \text{sn} \rangle + \langle \text{S, nonfinite, completive} \rangle \Rightarrow \langle \text{sn, coord} \rangle$:

(2.223) (Cat 104nou)

[la inauguració abans de finals d’any a Holanda d’una central de compres per a tot Europa] i [desenvolupar el negoci electrònic]

‘the opening before the end of the year in Holland of a shopping centre for all Europe and to develop the electronic business’

4. $\langle \text{s.a} \rangle$ or $\langle \text{sa} \rangle + \langle \text{S, nonfinite, participle} \rangle \Rightarrow \langle \text{s.a, coord} \rangle$ or $\langle \text{sa, coord} \rangle$:

(2.224) (Cat 06010300-6)

una edificabilitat [ordenada] i [sostenible]

‘an arranged and sustainable architecture’

5. $\langle \text{sp} \rangle + \langle \text{S, nonfinite, participle} \rangle \Rightarrow \langle \text{S, nonfinite, participle, coord} \rangle$:

(2.225) (Cat 103-12)

[obert a tots els estils] i [per a músics]
 ‘open to all styles and for musicians’

6. $\langle \text{s.a} \rangle + \langle \text{S, nonfinite, participle} \rangle + \langle \text{sp} \rangle \Rightarrow \langle \text{s.a, coord} \rangle$:

(2.226) (Cat 204nou)

una programació [generalitzada], però [enfocada al segment dels joves] i [amb un important pes dels serveis interactius]
 ‘a generalized program, although addressed to young people and with an important value of the interactive services’

7. $\langle \text{sadv} \rangle + \langle \text{S, nonfinite, participle} \rangle \Rightarrow \langle \text{S, nonfinite, participle, coord} \rangle$:

(2.227) (Cat 06010100-41)

[amagats en una bossa de mà] i [dins de la guantera d’un vehicle]
 ‘hidden in a handbag and inside a car’s glove compartment’

8. $\langle \text{sp} \rangle + \langle \text{S, finite, adverbial} \rangle \Rightarrow \langle \text{S, finite, adverbial, coord} \rangle$:

(2.228) (Cat 06010100-15)

[ja_ que una quarta part de l’edifici està ocupada ...] i [per obrir una nova oficina del DARP]
 ‘because a fourth of the building is occupied ... and to open a new DARP office’

9. $\langle \text{sp} \rangle + \langle \text{S, nonfinite, adverbial} \rangle \Rightarrow \langle \text{S, nonfinite, adverbial, coord} \rangle$
 (when it is a verb complement; otherwise the combination results in $\langle \text{sp, coord} \rangle$):

(2.229) (Cat 103-12)

obrir la convocatòria [a tots els gèneres] i [prioritzant la qualitat de la composició]
 ‘to open the call to all genres and giving priority to the composition quality’

10. $\langle \text{sp} \rangle + \langle \text{S, finite, relative} \rangle \Rightarrow \langle \text{S, finite, relative, coord} \rangle$:

(2.230) (Cat 207-8)

una obra [d’aquestes característiques], i [que cap altre país té en la_ seva historiografia]
 ‘a piece of these characteristics and that any other country has in its historiography’

11. $\langle \text{s.a} \rangle$ or $\langle \text{sa} \rangle + \langle \text{S, finite, relative} \rangle \Rightarrow \langle \text{s.a, coord} \rangle$ or $\langle \text{sa, coord} \rangle$:

(2.231) (Cat 104nou)

*que no porta unitat tractora, [capaç d'arribar a una velocitat de 350 km/h],
[que serà fabricat íntegrament a Espanya]*

‘which has no tractor unit, able to reach a speed of 350 km/h, which will be wholly made in Spain’

2.6.5 Complementizer ellipsis at the coordination

In coordination of subordinate clauses it is frequent to find ellipsis of the complementizer; that is, the relative or interrogative pronoun, or the conjunction only occurs once (usually with the first of the coordinated clauses). In such cases, we place it as a daughter node of the coordinated structure and sister of the subordinate clauses, as shown in examples 2.232 and 2.233:

(2.232) (Cat 211nou)

[perquè [són dues persones complementàries] i [Ø aquesta fórmula pot funcionar molt bé]]<coord>

‘because (they) are two complementary people and this formula can work out well’

(2.233) (Cat 06010100-33)

[[quines [són] i [Ø seran]]<coord>] [les_ seves funcions<sn,adjunct>]]

‘which are and will be their functions’

2.6.5.1 Particular cases

Sometimes, we find complements which are attached between the complementizer and the coordinated structures. Example 2.234 shows the correct analysis:

(2.234) (Cat 111nou)

[que [si no s'aconsegueix ara aquest retorn<S,conditional,adjunct>] 'es mantindran les condicions objectives i continuarà havent-hi vencedors i vençuts']

‘that if this return is not achieved, the objective conditions will keep on and there will still be winners and losers’

We can also find a complementizer affecting more than two subordinate clauses that, in their turn, coordinate each other at different levels (??):

(2.235) (Cat 06029900-19)

[quan [[[el processat va amenaçar] i [Ø intimidat]]<coord>] [la víctima]<sn,adjunct>] i [Ø va aconseguir així que li lliurés un total 534000 pessetes]]<coord>

‘when the accused threatened and intimidated the victim and made (her/him) give a total amount of 534000 pesetas’

2.6.6 Verbal partial ellipsis at the coordination

Sometimes the verb of the second clause of the coordination occurs only in part. It is possible, for instance, the occurrence of the participle instead of the whole conjugated verb, as well as the infinitive or gerund forms of a periphrasis. In such cases, the non-finite form is marked as *infinitiu*, *gerundi* or *participi* and the sentence to which this form belongs is marked as coordinated.

(2.236) (Cat 111nou)

[han de condemnar aquesta iniciativa] i [Ø explicar a Castella _ i _ Lleó la proposta del PSC]

‘(they have to condemn this initiative and explain to Castella i Lleó the PSC’s proposal)’

2.6.7 Adjunction at the coordination

Frequently, when coordinating two structures (phrases or clauses), complements use to occur just once, although they affect both structures. We should adjoin these complements to the whole coordinated structure, as shown in 2.237:

(2.237) (Cat 111nou)

[[[retorn] i [dipòsit]]<sn,coord> [a l’Arxiu _ Nacional de Catalunya]]

‘the return and deposit to the Catalonia National Archive’

2.6.7.1 Special cases in which there is no adjunction

When two sentences share the same subject (the overt subject belonging to the first sentence, while there is a covert subject for the second sentence):

(2.238) (Cat 204nou)

[L’atur masculí es va situar a finals del mes passat en 64.248] i [Ø va disminuir sobre el mes anterior en 2.493]

‘the men’s unemployment settled at the end of last month at 64.248 and decreased from the previous month in 2.493’

There is an exception, however, in which the subject can only be attached (2.239):

(2.239) (Cat 06019900-4)

[[[Per aquestes injúries va ser jutjat] i [condemnat]<coord>] [el policia local al Jutjat _ de _ Sant _ Feliu el 15 _ de _ Maig _ de 1995]<sn,elliptic,adjunct>]

‘Because of these insults the local policeman was judged and sentenced at the Sant Feliu court on May the 15th, 1995’

When two coordinated subordinate clauses are sharing the complementizer (which is sister of the coordinated structures)⁴⁶

⁴⁶See section 2.6.5 for more details on the ellipsis of the complementizer.

Sometimes a discontinuous constituent affects two or more coordinated constituents. In such a case the discontinuous element is a sister node of the coordinated structures (2.240):

(2.240) (Cat 207-7)

... es van produir per sortida de la via i un 22% per col·lisions frontals, com a causes més destacades

‘... (they) were caused by runnig off the road and a 22% by head-on crashes, as the most outstanding reasons’

By representing this way cases such as the one above, they cannot be confused with other structures in which a constituent has to be attached to an element, but it is not, due to linear and ordering reasons. We mark this structure as an ****adjoined constituent (<adjunct>):

(2.241) (Cat 211nou)

se li pot donar una major transcendència a unes trobades que a altres

‘some meetings can be given more importance than others’

2.7 The adjunction

Attachment consists on making a structure depend on another by doubling one of the nodes. The node to which it is attached, is the one that is doubled. All adjoined nodes are marked by the corresponding XML feature-value pair:

```
</attribute>
<attribute name="adjunct" onlynonterminal="yes"
description="Constituent is adjoined?">
  <value name="yes" />
  <value name="no" />
</attribute>
```

2.7.1 Complements of a coordinated structure

We adjoin those complements belonging to a coordinated (phrase or sentence) structure, as in 2.242:

(2.242) (Cat 111nou)

```
<sn complex="yes">
  <grup.nom coord="yes">
    <grup.nom coord="yes">
      <grup.nom gen="m" num="s">
        <n [...] num="s" postype="common" wd="retorn"/>
      </grup.nom>
```

```

<coord>
  <c lem="i" postype="coordinating" wd="i"/>
</coord>
<grup.nom gen="m" num="s">
  <n [...] num="s" postype="common" wd="dipòsit"/>
</grup.nom>
</grup.nom>
<sp adjunct="yes">
  <prep>
    <s [...] num="c" postype="preposition" wd="a"/>
  </prep>
  <sn>
    <spec gen="m" num="s">
      <d [...] num="s" postype="article" wd="l'"/>
    </spec>
    <grup.nom gen="m" num="s">
      <n [...] num="c" postype="proper"wd="Arxiu_Nacional"/>

```

‘the return and deposit to the Catalonia National Archive’

For exceptional cases in which there is no adjunction, see section 2.6.7.1.

2.7.2 Comparative structures

Comparative⁴⁷ subordinated clauses are treated as structures adjoined to the phrase node (<sn>, <sa>, <s.a> or <sadv>) which to the adverb introducing the comparison belongs:

(2.243) (Cat 204nou)

```

  <v [...] num="c" postype="main" wd="ser"/>
</infinitiu>
<sa>
  <sa>
    <grup.a>
      <a [...] num="s" postype="qualificative" wd="igual"/>
    </grup.a>
  </sa>
  <S adjunct="yes" clausesubtype="comparative" [...]>
    <conj conjunctiontype="subordinating">
      <c lem="que" postype="subordinating" wd="que"/>
    </conj>
    <sp>

```

⁴⁷See section 2.5.2.4 for more details on finite adverbial comparative clauses.


```

<prep>
  <s [...] num="s" postype="preposition" wd="al"/>
</prep>
<sn>
  <grup.nom gen="m" num="s">
    <n [...] num="s" postype="common" wd="carrer"/>

```

‘to be as (you) are on the street’

2.7.3 Adverbs preceding the verb

Adverbs affecting the verb (that is, the whole sentence) belong to the sentence that contains this verb, instead of being adjoined (2.244):

(2.244) (Cat 111nou)

[ara es pretén buscar també l'acord polític parlamentari]

‘now it is also expected to look for the parliamentary political agreement’

We do adjoin adverbs (*només* ‘only’, *potser* ‘maybe’, *fins_i_tot* ‘even’, *quasi* ‘almost’, *no*, ‘no’) when they do not affect the whole sentence but certain phrase structure and:

1. That phrase structure is not an adjective phrase (or participle clause) or an adverbial phrase.
2. The adverb can only occur in that position (or can precede the verb provided that it occurs between commas).

(2.245) (Cat 111nou)

```

<n [...] num="s" postype="common" wd="proposició"/>
<sp>
  <neg adjunct="yes">
    <r lem="no" postype="negative" wd="no"/>
  </neg>
<sp>
  <prep>
    <s [...] num="c" postype="preposition" wd="de"/>
  </prep>
  <sn>
    <grup.nom gen="f" num="s">
      <n [...] num="s" postype="common" wd="lleï"/>
    </grup.nom>

```

‘Private Member’s Bill’

- (2.246) (Cat 06010100-9) ...*la poca repercussió que tenen els mitjans [...] i[[, especialment,*_{<adjunct>}] *l'escassa repercussió directa que tenen les fires turístiques]*
 ‘...the little repercussion that media has [...] and, specially, the little direct repercussion that tourism fairs have’

When the adverb affects either an adjective phrase (or participle clause) or adverbial phrase, we do not adjoin them:

- (2.247) (Cat 104nou)

```
<sn>
  <grup.nom gen="m" num="p">
    <f lem="-" punct="hyphen" wd="-"/>
    <n [...] num="p" postype="common" wd="autocars"/>
    <s.a gen="m" num="p">
      <neg>
        <r lem="no" postype="negative" wd="no"/>
      </neg>
      <a [...] postype="qualificative" wd="urbans"/>
    </s.a>
    <f lem="-" punct="hyphen" wd="-"/>
  </grup.nom>
</sn>
```

‘non-urban buses’

2.7.4 Complements of the noun phrase

Some noun complements that are noun phrase-like are not treated as apositions when the coreference condition does not apply. Since they are not adjectivised neither, we adjoin the second noun phrase to the first (2.248):

- (2.248) (Cat 111nou)

```
<sn>
  <sn>
    <grup.nom gen="m" num="s">
      <n [...] num="c" postype="proper" wd="Jordi_Portabella"/>
    </grup.nom>
  </sn>
  <sn adjunct="yes" entityref="ne" ne="organization">
    <grup.nom>
      <f lem="(" punct="bracket" punctenclose="open" wd="("/>
      <n [...] num="c" postype="proper" wd="ERC"/>
      <f lem=")" punct="bracket" punctenclose="close" wd=")"/>
    </grup.nom>
  </sn>
```

```

    </grup.nom>
  </sn>
</sn>

```

‘Jordi Portabella (ERC)’

When more than one noun phrase cooccur, we group them and marked the whole set as a verbless sentence so we can adjoin the sentence node to the noun phrase (2.249):

(2.249) (Cat 103-1)
Robert_Brufau [(Mollerussa, 1946)]<S,verbless,adjunct>

We also adjoin those prepositional phrases like the one shown in (2.250):

(2.250) (Cat 111nou)
Tots els grups del Parlament, [menys el PP]<sp,adjunct>
 ‘All groups of the Parliament, except the PP party’

Those adverbs depending on a noun phrase instead of affecting the whole sentence are also adjoined (2.251):

(2.251) (Cat 06010100-9)
*...la poca repercussió que tenen els mitjans [...] i[[, especialment,<adjunct>]
 l'escassa repercussió directa que tenen les fires turístiques]*
 ‘...the little repercussion that media has [...] and, specially, the little direct
 repercussion that tourism fairs have’

2.7.5 Particular cases

Sometimes we can find a double adjunction, as in 2.252:

(2.252) (Cat 107)
*dels productes locals, [[especialment]<adjunct> els que permeten diversificar el sector
 agrari]<adjunct>*
 ‘from local products, specially those that allow to diversify the agrarian sector’

2.8 Discontinuity

2.8.1 General case

Sometimes, and due to ordering reasons, a complement cannot depend directly on its head. In such a case, we mark the discontinuity by adding an XML feature value both to the head and the separated complement:

```
- <attribute name="discontinuous" onlyterminal="yes"
  description="Is it discontinuous?">
  <value name="yes" />
  <value name="no" />
</attribute>
<attribute name="discid" open="yes" method="discontinuu"
  description="Discontinuity identifier" />

- <attribute name="head" open="no" onlynonterminal="yes"
description="Is it the 'head' of the dependency structure?">
  <value name="yes" />
  <value name="no" />
</attribute>
```

The open value for <discid> can be fulfilled with <1n> for the head and <1c> for any complement of that head. If there is more than one discontinuity within a sentence, we can fulfil the value with <2n> and <2c>, etc.

(2.253) (Cat 06010300-18)

construir [una dessaladora<discid,1n>] a Singapur, [amb una capacitat de 50 hectòmetres...<discid,1c>]
 ‘to build up a desalination plant in Singapur, with a capacity of 50 hectometres...’

The complement belongs to the lowest *S* (whatever kind it is) within the structure, bearing in mind that the head must be within this <S>. Thus, this complement depends on the same <S> to which the head belongs. Otherwise, it should depend on a higher <S>. However, this is not always possible, because sometimes the complement can occur in a lower level than its head, as shown in (2.254):

(2.254) (Cat 06030200-14)

[agents<discid,1n>] socials implicats en l'atenció ... [(pediatre, psicòloga infantil, centres educatius, etc.)<discid,1c>] i que ...
 ‘social agents who are implied in the attention ... (pediatrician, children’s psychologist, educative centres, etc) and that...’

These kind of structures are usually complementing the noun phrase, but also finite and non-finite clauses, as well as adverbial and adjective phrases.

2.8.2 Particular cases

Sometimes, a complement (as shown in 2.255), can refer to more than one head at a time. In this example, it complements the noun phrases *111.000 pessetes*, *106.000 pessetes* and *121.000 pessetes*, which are all marked as <1n>:

(2.255) (Cat 107-15)

En escorcollar Noureddine_Hassoun li van trobar a sobre [111.000 pessetes<sn,1n>] ...; a Solimane_Boubouki [105.000 pessetes<sn,1n>]; i a Minoun_El_Douri [121.000 pessetes<sn,1n>], [diners que ... procedien de...<sn,1c>]
 ‘When (they) registered Noureddine_Hassoun, (they) found he had 111.000 pesetas; Solimane_Boubouki had 105.000 pesetas; and Minoun_El_Douri had 121.000 pesetas, money that... came from... ’

When the complement does not depend on a constituent but it is attached to it, both of them (head and complement) are added the corresponding suffix (<1n> and <1c>, respectively). In addition to that, the adjoined complement is marked as <adjunct> as well, even though the node is not doubled this time. This is the case of comparative clauses, since they cannot be attached to the node they belong due to ordering reasons. In such cases, the constituent to which they belong is marked as <1n>:

(2.256) (Cat 211nou)
se li pot donar una [major transcendència<sn,1n>] a unes trobades [que a altres<S,verbless,adverbial,comparative,1c>]
 ‘some meetings can be given more importance than others’

By adopting this representation, this kind of discontinuity cannot be confused with that of those cases in which the complement is not adjoined to its head because of real ordering reasons. Thus, both features <adjunct,1c> are only marked in the latter, as happens in example 2.256 above.

Some noun phrases should also be adjoined to other noun phrases (2.257):

(2.257) (107-7)
[l’espai web<sn,1c>] sempre actualitzat, [[especialment<sadv,adjunct>] els apartats d’Agenda i Notícies<sn,adjunct,1c>]
 ‘the always updated website, especially News and Events sections’

When a complement affects two or more coordinated constituents, it belongs to the upper node so that it is the sister of the coordinated structures (2.258):

(2.258) (Cat 207-7)
... es van produir per [sortida de la via<grup.nom,1n>] i un 22% per [col·lisions frontals<grup.nom,1n>], [com a causes més destacades<sp,1c>]
 ‘... (they) were caused by running off the road and a 22% by head-on crashes, as the most outstanding reasons’

2.9 Miscellany

2.9.1 Numbers

Numeric expressions⁴⁸ are part of a noun phrase. What comes from the automatic parsing, however, is usually <numero>, which should be changed manually either by <espec> or <grup.nom>. In addition, the proper gender and number values are marked as

⁴⁸All numbers are morphologically marked as <z>.

well, when they can be inferred from the head. In those cases of unknown gender, only number information is added, this being plural from number 2 onwards. The following examples show the final representation:

(2.259) (Cat 06010300-26)

```
<sn>
  <spec gen="m" num="p">
    <z lem="35" wd="35"/>
  </spec>
  <grup.nom gen="m" num="p">
    <n gen="m" lem="treball" num="p" postype="common" wd="treballs"/>
  </grup.nom>
</sn>

'35 works'
```

(2.260) (Cat 204nou)

```
<sn>
  <grup.nom>
    <z lem="609" wd="609"/>
  </grup.nom>
</sn>

'609'
```

2.9.2 Dates

Dates⁴⁹ are automatically marked as <data>, which should be replaced by <sn>. The node <grup.nom> should also be added:

(2.261) (Cat 06010300-26)

```
<sn>
  <spec gen="m" num="s">
    <d gen="m" lem="el" num="s" postype="article" wd="el"/>
  </spec>
  <grup.nom gen="m" num="s">
    <w lem="[??:??/8/?:?:??.?]" wd="mes_d'_agost"/>
  </grup.nom>
</sn>

'the month of August'
```

⁴⁹All dates are morphologically labeled as <w>.

Days of the week, months and years are labeled as *grup.nom.ms* even though there are no specifiers or complements indicating gender⁵⁰:

(2.262) (Cat 203-9)

```
<sn>
  <grup.nom gen="m" num="s">
    <w lem="[diumenge:??/??/?:?.?]" wd="diumenge"/>
  </grup.nom>
</sn>
```

‘Sunday’

(2.263) (Cat 203-14)

```
<sn>
  <spec gen="m" num="s">
    <d gen="c" lem="el" num="s" postype="article" wd="l'"/>
  </spec>
  <grup.nom gen="m" num="s">
    <w lem="[?:??/??/1999:?.?]" wd="any_1999"/>
  </grup.nom>
</sn>
```

‘the year 1999’

Hours of the day are added gender and number features only when going along with specifiers or complements:

(2.264) (Cat 107-17)

```
<sn>
  <grup.nom>
    <w lem="[dilluns:??/??/?:12.30]" wd="dos_quarts_d'una_del_migdia_de_dilluns"/>
  </grup.nom>
</sn>
```

‘Monday, half past twelve in the afternoon’

(2.265) (Cat 107-17)

⁵⁰If there are specifiers, however, dates are assigned their morphological features, except for those cases of coordination.

```

<sn>
  <spec gen="f" num="p">
    <d gen="f" lem="el" num="p" postype="article" wd="les"/>
  </spec>
  <grup.nom gen="f" num="p">
    <w lem="[??:??/??/?:09.00]" wd="nou_del_matí"/>
  </grup.nom>
</sn>

```

'9 in the morning'

When hours occur together with days, months or years, we split them up, as shown in (2.266):

(2.266) (Cat 06029900-8)

```

<sn>
  <spec gen="f" num="p">
    <d gen="f" lem="el" num="p" postype="article" wd="les"/>
  </spec>
  <grup.nom gen="f" num="p">
    <w lem="[??:??/??/?:05.00]" wd="5_del_matí"/>
  <sp>
    <prep>
      <s [...] lem="del" postype="preposition" wd="del"/>
    </prep>
    <sn>
      <grup.nom gen="m" num="s">
        <w lem="[??:23/08/1997:??.??]" wd="23_d'agost_de_1997"/>
      </grup.nom>
    </sn>
  </sp>
</grup.nom>
</sn>

```

'5 in the morning, 23rd of August of 1997'

2.9.2.1 Time intervals

In those cases in which we need to keep the structure of the prepositional phrase, as it happens to the hours of the day, the second prepositional phrase depends on the the <sn> of the first <sp> (2.267). Otherwise, we can keep the structure by coordinating both noun phrases (or noun groups):

(2.267) (Cat 06010300-23)


```

<sp>
  <prep>
    <s gen="c" lem="de" num="c" postype="preposition" wd="de"/>
  <prep>
  <sn>
    <grup.nom>
      <w lem="[?:9/?:?:?:??.?]" wd="9"/>
    </grup.nom>
    <sp>
      <prep>
        <s gen="c" lem="a" num="c" postype="preposition" wd="a"/>
      </prep>
      <grup.nom>
        <w lem="[?:?:?:?:?:10.00]" wd="10_del_matí"/>
      </grup.nom>
    </prep>
  </sn>

```

‘from 9 to 10 in the morning’

2.9.3 Currency

Numeric expressions indicating currency values are automatically parsed as locutions. We should split them up to apply the syntactic analysis, and change the value <zm> by the proper one. Actually, <zm> is only assigned to the word indicating the specific kind of currency, that is, *dollars*, *pounds*, *euros*, etc. The following examples illustrate the correct analysis:

(2.268) (Cat 111nou)

```

<sn>
  <spec gen="m" num="p">
    <z lem="2500" wd="2.500"/>
  </spec>
  <grup.nom gen="m" num="p">
    <n [...] num="p" postype="common" wd="milions"/>
    <sp>
      <prep>
        <s [...] postype="preposition" wd="de"/>
      </prep>
      <sn entityref="ne" ne="number">
        <grup.nom gen="f" num="p">
          <z [...] postype="currency" wd="pessetes"/>
        </grup.nom>
      </sn>
    </sp>
  </grup.nom>

```

```
</sp>  
</grup.nom>  
</sn>
```

'2500 millions of pesetas'

Chapter 3

Functions

3.1 Introduction

This section describes the criteria to carry out the syntactic function annotation within the AnCora corpora (see <http://clic.ub.edu/ancora/>).

3.2 Preliminary questions

1. Labels of syntactic function are XML-like attribute-value pairs, such as `<func='subj'>`, `<func='creg'>` and some other more specific, such as `<functype='quantitative'>`.
2. Functions are only added to those constituent nodes which depend on a sentence structure (be it finite or non-finite¹). Therefore, we only annotate what we consider 'basic sentence functions'; that is, those nodes which are directly connected to the main verb (subject, object, etc.). However, we also annotate other elements that occur within the sentence domain.
3. Nodes which do not depend directly on the root node S (complements of noun, adjective or adverb phrases) are not annotated with functions.
4. Some verbal complements occur within periphrastical forms. We do annotate this elements, since they directly depend on the main verb.
5. Only one syntactic function per element is annotated. This means that in a sentence like *li agrada cantar* 'she likes singing'², the pronoun *li* is assigned the indirect object label (`<ci>`), but it is not indicated that it is also the subject of the infinitive verb *cantar*.

¹See the constituency guidelines for the 3LB and AnCora corpora annotation.

²For both Catalan and Spanish, this kind of verbs rule the dative case for the experiencer of the action, while it is the infinitive, in this sentence, the one working as the subject.

3.2.1 Elements with no syntactic function

Apart from the elements that do not depend on a sentence structure, the following nodes are neither assigned any syntactic function:

1. Verbal forms. All elements labeled as *grup.verb*, *infinitiu*, *gerundi* and *participi* are the head of the sentence, but have not a function label.
2. Interjections, since they are discursive elements rather than syntactic constituents.
3. Nodes labeled as `<inc>` ('inserted element')³, even though they could depend on the higher node `<sentence>`. These nodes represent inserted elements within the text. They are parenthetical structures or digressions that do not belong to the syntactic sentence structure, and for this reason they are not assigned any syntactic function label, as shown in 3.1:

```
<nodename name="inc" description="Inserted element" onlyterminal="no" />
```

(3.1) (Cat 06010100-34)

A més, assenyala la proposició no de llei conjunta, l'accés als serveis genera problemes greus

'Moreover, the Private Member's Bill indicates, the access to the services generates serious problems'

4. Elements labeled as *morfema.pronominal*. Their function is just to give support to the verb (both strictly pronominal verbs, like *queixar-se* 'to complain') and those verbs having a pronominal version. It seems that pronouns like *en* and *hi* can also have these values⁴.
5. Relative clauses whose antecedent is a whole clause⁵. These relative clauses are not labeled as having syntactic function⁶, as it is shown in 3.2:

(3.2) (Cat 104-1.20)

que les condicions són diferents de les de fa uns anys, per la qual cosa és més necessària la col·laboració i coordinació...

'that the conditions are different from some years ago, and for this reason it is necessary the collaboration and coordination...

6. Constituents labeled as *S.F.AComp* (adverbial comparative clauses) and *S.F.R* (relative clauses). Relative clauses because they depend on a noun; and comparative clauses because they depend on phrases and noun groups or adjectives.

³Obviously, this issue affects only the node `<inc>`, but not its content. That is, any sentence structure below this node, as well as its constituents, are properly annotated with syntactic functions.

⁴See section 3.19 for more details on these pronouns

⁵However, when the relative clause relates to the whole main sentence, it is labeled as a sentence adjunct (`<ao>`).

⁶Elements that are adjoined to the whole sentence because of coordination (i.e. a noun phrase depending on two verbs at a time), are given function label.

7. Due to the hierarchical representation we carry out, word ordering is not altered and crossing branches are not possible. For this reason discontinuity poses a specific problem that in 3LB and AnCora is solved through specific attribute-value pairs. Section 3.20 in this document is devoted to this issue.
8. In coordinated structures we only assign one function to the upper node (the coordination node). However, when the two coordinated elements have different syntactic functions (which is quite unfrequent), each one is assigned its function, and it is the upper node the one without syntactic label (3.3).

(3.3) (Cat 06019900-3)

que van al monestir [[per inquietuds individuals<cc>] i [a buscar el recés individual<creg>]]

‘that go to the monastery due to individual interests and to find the individual recess’

3.3 Tag set

The following are the functions that are indicated within the Cat3LB/AnCora frame, in our XML node names document⁷:

```
<attribute description="Syntactic function" name="func"
  onlyterminal="no">
  <value name="suj" description="Subject" />
  <value name="cc" description="Adjunct" />
  <value name="cd" description="Direct object" />
  <value name="ci" description="Indirect object" />
  <value name="atr" description="Attribute" />
  <value name="cpred" description="Predicative complement" />
  <value name="creg" description="Prepositional object" />
  <value name="cag" description="Agent complement" />
  <value name="ao" description="Orational adjunct" />
  <value name="et" description="Textual element" />
  <value name="mod" description="Non-argumental verb modifier" />
  <value name="pass" description="Passive mark" />
  <value name="impers" description="Impersonality mark" />
  <value name="voc" decription="Vocative" />
</attribute>
```

Other attribute-value pairs that are relevant in the function annotation process are the following:

⁷You can find the complete node names document in http://clic.ub.edu/ancora/lng/en/constituents_tagset.pdf and http://clic.ub.edu/ancora/lng/en/func_tagset.pdf

1. Specific information related to a syntactic function

Some functions have extra information: Direct object, for instance, can be labeled as ‘quantitative’, meaning that it complements a measure verb or a state verb, being impossible to alternate with a passive subject. Adjuncts (<cc>) are specified only when they are locative or temporal. A few indirect objects can be ethical⁸, and we also specify this information by adding a description of the syntactic function:

```
- <attribute description="Syntactic function
    subclassification" name="functype"
    parents=".*:func:(ci|cd|cc)">
  <value name="dative" parents=".*:func:ci"
    description="Ethical/interest dative" />
  <value name="quantitative" parents=".*:func:cd"
    description="Quantitative direct object" />
  <value name="locative" parents=".*:func:cc"
    description="Locative adjunct" />
  <value name="temporal" parents=".*:func:cc"
    description="Temporal adjunct" />
</attribute>
```

2. Repeated functions

Sometimes, we find repeated functions due to either dislocations, pronominalizations or pleonasm. We mark this by adding an attribute value, so we avoid the same function to be doubled in the same sentence:

```
- <attribute name="repetition" onlynonterminal="yes"
description="Is it repeated?">
  <value name="yes" />
  <value name="no" />
</attribute>
```

3. Discontinuous functions

Other functions are splitted up due to linearity: When a noun complement, for instance, appears far from its head, it is coindexed by a specific attribute which allows the annotator to indicate a number (to both the head and the complement) as well as to mark whether the element is the head or the complement:

```
- <attribute name="discontinuous" onlyterminal="yes"
description="Is it discontinuous?">
  <value name="yes" />
  <value name="no" />
```

⁸See section 3.7.1 for further details on ethical datives

```

</attribute>
<attribute name="discid" open="yes" method="discontinuu"
  description="Discontinuity identifier" />

- <attribute name="head" open="no" onlynonterminal="yes"
description="Is it the 'head' of the dependency structure?">
  <value name="yes" />
  <value name="no" />
</attribute>

```

3.4 Syntactic function, coordination and adjunction

As mentioned above, in the case of coordinated structures that need function label, this should be assigned to the highest node. It is worth reminding that coordination must be carried out at the lowest possible level, but it is the whole coordinated group the one holding a syntactic function. Moreover, two or more coordinated elements are supposed to have the same syntactic function (3.4):

(3.4) (Cat 104-1.4)

va subratllar que aquesta falta de mà d'obra ... va començar a les comarques de Lleida però ja ha arribat a totes les províncies gironines
 '(s/he) emphasized that this staff shortage ... started at the Lleida regions but (it) has already reached all Girona regions as well'

When the coordinated elements have different syntactic functions (which is quite infrequent), the label is assigned to their corresponding phrase nodes in spite of being assigned to the upper coordinated node. This is illustrated in 3.5:

(3.5) (Cat 06019900-4.5) *El secretari, [[tot i conèixer la sentència, <ao>] i [sense fer cas dels fets considerats provats pel jutge <cc>]], va emetre...*

'The secretary, even knowing the judgement, and ignoring the proved facts, delivered...'

Another special case is that of what is called *zeugma*⁹. In a sentence like *parla anglès i molt ràpid* 'He speaks English and (he does it) very fast', there is a direct object (*anglès*) coordinated with an adjunct (*molt ràpid*).

In case of coordinated noun phrases or noun groups, there are two specific node names we add, depending on the kind of coordination:

1. Coordinated phrases and groups:

⁹A figure of speech describing the joining of two or more parts of a sentence with a single common verb or noun.

```
- <attribute name="coord" onlynonterminal="yes"
description="Is it a coordinate constituent?">
  <value name="yes" />
  <value name="no" />
```

2. Phrases containing coordinated groups (that is, complex phrases):

```
- <attribute name="complex" onlynonterminal="yes"
description="Is it complex?">
  <value name="yes" />
  <value name="no" />
</attribute>
```

As mentioned before, elements adjoined to a coordinated sentence structure are given function label when the reason of being adjoined is that they modify two or more coordinated elements at a time. Apart from the corresponding syntactic function label, it is given a specific label indicating the adjunction:

```
- <attribute name="adjunct" onlynonterminal="yes"
description="Constituent is adjoined?">
  <value name="yes" />
  <value name="no" />
</attribute>
```

When a constituent holding a syntactic function is given an adjoined element, the function label is to be placed at the highest node (the one containing the feature string for <adjunct>).

3.5 Subject

1. The subject is a noun phrase or equivalent structure (completive clause)¹⁰ that agrees in number and person with the main verb of the sentence to which it belongs¹¹.
2. Another characteristic of the subject is that, generally, it can be replaced by a strong pronoun in nominative case.
3. The subject can hold both preverbal and postverbal position.

¹⁰The only exception is that of a prepositional phrase in subject position: in infinitive clauses the subject can correspond to a prepositional phrase introduced by *de*: *no li va donar la gana de mantenir...* 'He did not want to maintain...'

¹¹In case of *S.NF.PA* (absolute clauses), the subject agrees in gender and number with the participle form of the sentence.

4. All ellipited subjects (or covert subjects) should be given the function label corresponding to a subject. The ellipsis is annotated by a specific node name feature:

```
- <attribute name="elliptic" onlyterminal="yes"
description="Is the constituent ellipited?">
  <value name="yes" />
  <value name="no" />
</attribute>
```

Sometimes, **agreement is made *ad sensum***. This means that it is the meaning instead of morphological features the one determining the agreement. There are some other cases, however, in which there is no agreement at all. Examples 3.6 to 3.9 show this *ad sensum* agreement:

(3.6) (Cat 207-1.6)

... en què hi prenen_part una quinzena de paradistes...
 ‘... in which fifteen shopkeepers participate...’

(3.7) (Cat 211-11.16)

... el gruix de les campanyes [...] corresponien a línies de comunicació...
 ‘... the bulk of the [...] campaigns corresponded to communication lines...’

(3.8) (Cat 211-6.4)

... cap dels dirigents [...] estaven en la direcció del partit...
 ‘... none of the leaders [...] were in the direction of the party...’

(3.9) (Cat 204-5-13)

... una sèrie d’institucions que tenen com_a obligació...
 ‘... a series of institutions having as their obligation...’

We consider those expressions containing *és_que ...* ‘is that ...’ not to be impersonal constructions but expressions having a covert subject, which corresponds semantically to nouns such as *la causa/motiu* ‘the cause/reason’ and the like, and the completive clause following the verb works as the attribute. The analysis of this kind of constructions is still under discussion, and this formal interpretation is to make the automatic learning easier.

3.5.1 Impersonal sentences

Impersonal sentences have no subject. We consider the following structures to be impersonal:

1. meteorological verbs (*ploure* ‘to rain’; *nevar* ‘to snow’; *tronar* ‘to thunder’) in their literal meaning.
2. other constructions, such as: *sembla que...* ‘it seems that...’; *hi ha...* ‘there is...’, as in 3.10.

(3.10) (Cat 111)

... a la terra hi ha 1300 milions de persones amb ingressos inferiors a un dòlar diari

‘... there are 1300 million people on Earth which get as incoming less than a dollar per day’

3. constructions such as: (*no*) *pot ser que...* ‘it can/cannot be that...’; when meaning *és possible que...* ‘it is possible that...’ and *és impossible que...* ‘it is impossible that...’ (3.11).

(3.11) (Cat 104-1.3)

no pot ser que es tardí vuit mesos per donar un permís de residència...

‘it cannot be that eight months are needed to give a residence license...’

In the example above there are two types of impersonality: the first is due to the construction *pot ser que*; the second is because of the verbal morpheme *es*.

4. ‘*ES*’ impersonal constructions. The following list includes some of the verbs of the corpus occurring in an ‘*ES*’ impersonal construction:

(3.12) **actuar**: *s’ ha actuat amb eficàcia en els casos concrets...* (Cat 111-6)

‘(it) has been done efficiently in the concrete cases...’

(3.13) **arribar**: *es podria arribar a les 20000 tones de pa...* (Cat 06010100-5)

‘(it) could reach 20000 tones of bread...’

(3.14) **avançar**: *... no s’ ha avançat i el tema està parat* (Cat 104-5)

‘... no advance has been made and the topic is at a standstill’

(3.15) **comptar**: *... s’ ha comptat amb la col.laboració d’ un equip tècnic* (Cat 107-6)

‘... the collaboration of a technical team has been taken into account’

(3.16) **desprendre**: *... pateix lesions físiques [...], com es desprèn de l’ informe...* (Cat 111-9)

‘... (s/he) suffers from physical injuries [...], as inferred from the report...’

(3.17) **disposar**: *... per al qual ja es disposa de terreny públic* (Cat 107-19)

‘... for which (it) is provided a plot of public land’

(3.18) **optar**: *... com si s’ opta per tirar endavant el projecte* (Cat 107-5)

‘... as if it is decided to carry on with the project’

(3.19) **treballar**: *... ja es treballa en el de l’ any vinent* (Cat 103-12)

‘... (it) is been working on the one of next year’

3.5.2 Subject of attributive verbs

We consider as attributive only the verbs *ser* ‘to be’; *estar* ‘to be’ and *semblar* ‘to seem, to look like’. In essence, only a noun phrase or equivalent construction can be the subject of an attributive verb¹², whereas the attribute can be a noun phrase, an adjective phrase and even a prepositional or adverbial phrase (although these cases are much less frequent). Sometimes, both complements of an attributive verb are of the same structure (noun phrases or equivalent constructions). Then, we apply the following criteria to disambiguate:

1. any elliptic noun phrase is the subject.
2. the constituent showing agreement with the verb, is the subject, and the other one, the attribute.
3. the constituent that can be replaced by the clitic *ho*, is the attribute.
4. when the agreement features are the same for both constituents, the one placed before the verb is the subject, and the one after it is the attribute.
5. when both constituents are placed after the verb, then the closest to the verb is the subject.

The following are some examples¹³:

- *El teu pare és el director del banc* ‘Your father is the bank director’. **subject**: el teu pare; **attribute**: el director del banc.
- *El director del banc és el teu pare* ‘The bank director is your father’. **subject**: el teu pare; **attribute**: el director del banc.

In both cases the pronominalization results in *el teu pare l’és*. Never **el director del banc l’és*.

- *Jo sóc Napoleó* ‘I am Napoleon’. **subject**: jo; **attribute**: Napoleó.
- *Napoleó sóc jo* ‘Napoleon it’s me’ **subject**: jo; **attribute**: Napoleó.

Here, the subject is the one agreeing with the verb. We find some examples in the corpus (3.20):

(3.20) (Cat 204-3.5) *el conseller-delegat i director general és Miquel Alabern*
 ‘the managing director and general director is Miquel Alabern’

¹²Except for cases such as *Avui és dimecres* ‘Today is Wednesday’, in which an adverb works as the subject.

¹³GCC (S2).

Other test to identify an attribute is by a *què*-question (‘what’). By contrast, a *qui*-question (‘who’) is not a good test, since it allows both the subject and the attribut to be the answer.

Example 3.21 shows a case in which the subject label is assigned because of the agreement:

- (3.21) (Cat 107-9.2) *Conèixer la gastronomia i els jocs d’altres països del món són alguns altres dels temes escollits en el programari de tallers*
 ‘To know the gastronomy and games of other countries are some other issues chosen for the workshop program’

3.5.3 The subject and the clitics

Only strong personal pronouns (*jo, tu, ell, ...*) can replace the subject of a sentence. However, the clitic pronoun *en*, as a partitive pronoun, can also replace a part of the subject (3.22):

- (3.22) *han dut els entrepans, però només en queda un*¹⁴
 ‘(they) brought the sandwiches, but there is only one left’

In such a case, we treat the subject as a discontinuous function¹⁵.

3.6 Direct object

Generally, the direct object (<cd>) corresponds to a noun phrase or equivalent structure with no identifying mark (In Catalan). However, strong pronouns and some indefinite pronouns are exceptions, since they can be introduced by the preposition *a*:

- *et prefereix a tu* ‘(s/he) prefers you’
- *he vist a tothom*¹⁶ ‘(I) saw everybody’

Direct objects typically hold postverbal position. When occurring before the verb, it usually appears a clitic repetition¹⁷. The definite direct object can be replaced by the clitics *el, la, los, les, em, ens, et, us* (and all their variants); the indefinite one, on the other hand, is replaced by the clitic *en*; when it is a sentence structure or a neuter phrase, it is replaced by the clitic *ho*. Although it is not a definitive test, the direct object usually becomes the subject of passive sentences. A direct object can correspond to:

- (3.23) (Cat 111-4) **a relative**
la nova etapa que dirigiran amb tota probabilitat ell mateix i Pascual Maragall
 ‘the new stage **that** he himself and Pascual Maragall will probably manage’

¹⁴GCC (pp. 1374-75)

¹⁵See section 2.8 for further details on discontinuity.

¹⁶GCC (p.1098).

¹⁷See section 2.8 for further details on double functions and repeated functions.

(3.24) (Cat 111-4) **a finite completive clause**

ha advertit els últimes dies que el govern de José María Aznar podria qüestionar la col·laboració amb CiU

‘(s/he) recently warned that the govern of José María Aznar could question the collaboration with CiU’

(3.25) (Cat 111-4) **a noun phrase**

una situació que genera confiança

‘a situation that generates confidence’

(3.26) (Cat 111-10) **a prepositional phrase**

1. *Maragall també ha citat com una altra de les persones ... a Jordi Porta*

‘Maragall has also mentioned as another of the persons ... **Jordi Porta**’

2. *un ciutadà iugoslau: a qui acusa d’haver disparat* (Cat 060299-7)

‘a Yugoslavian citizen **to whom** (s/he) accuses of having fired’

(3.27) (Cat 111-4) **a non-finite completive clause**

**0* no es planteja deixar l’alcaldia de Cornellà*

‘(s/he) is not considering **leaving the mayoralty of Cornellà**’

There are some verbs showing direct objects introduced by a preposition *de*, *a*. This is the case of the verbs like *provar* ‘to try’ or *aprendre* ‘to learn’ when the direct object is an infinitive clause. However, we analyse these prepositions in a way that they do not head any phrase (that is, they are not proper prepositions):

(3.28) (Cat 06029900-8.7)

(a) *provar de descordar-li els pantalons*

‘**try to** unbutton his/her pants’

(b) *han après a elaborar diferents productes artesans*

‘(they) **learnt to** elaborate different handmade products’

3.6.1 Quantitative direct object

We label as quantitative (<functype=‘quantitative’>) those direct objects depending on measure verbs such as *valer* ‘to cost, to be worth’, *fer* ‘to measure, to cost’, *pesar* ‘to weigh’, *costar* ‘to cost’, followed by complements expressing a measure; verbs such as *córrer* ‘to run’, *caminar* ‘to walk’, followed by complements expressing a distance and state verbs (*tenir* ‘to have’, *haver-hi* ‘there is/are’). In all these cases, clitics can replace the object although the object cannot become the subject of a passive structure (at least with a periphrastic passive form of the verb). In this situation, some adverbs can also hold a direct object function:

- *pesa molt* ‘it is heavy’ (lit., it weighs a lot)
- *dura massa* ‘it lasts too much’¹⁸

¹⁸GCC (p.1102).

3.7 Indirect object

Indirect objects are prepositional phrases introduced by the preposition *a*, which can be replaced by the clitics *em*, *et*, *es*, *li*, *els*, *ens*, *et*, *us* (and all their variants). When the indirect object is a clitic, its constituent label corresponds to a noun phrase. We do not consider those prepositional phrases introduced by the preposition *per* to be indirect objects (we treat this kind of structures as adjuncts).

An indirect object can correspond to:

(3.29) **a prepositional phrase**

Segons han indicat a Efe fonts del PSC i CiU (111-10)

‘According to what sources of PSC and CiU indicated **to Efe**’

(3.30) **a noun phrase**

li ha comunicat que prefereix ... una persona amb un perfil més independent
(111-10)

‘(s/he) communicated **him/her** that (s/he) prefers ... a person with a more independent profile’

Sometimes the indirect object is expressed by a noun that does not refer to a person but to an entity, institution, etc., for which the pronominal substitution with the clitic *li/els* is also possible¹⁹:

(3.31) (104-1.11) ... *el Deutsche Bank presentés a aquesta institució la informació sobre ...*

... el Deutsche Bank **li** presentés la informació sobre ...

‘Deutsche Bank presented **to this insitution/to it** the information about...’

(3.32) (Cat 111-1.7) *La proposició no de llei ha estat presentada a la comissió de política cultural del Parlament*

‘The Private Member’s Bill was presented to the Parliament’s political comission’

By contrast, when the replaced pronoun is *hi*, we have a predicative, adjunct or prepositional object, instead of an indirect object²⁰.

3.7.1 Ethical dative

A few cases of the ethical²¹ use of the dative (also dative of interest) can be found in the corpus:

¹⁹Sometimes the limit is not clear and it is not easy to decide whether we could replace the complement by either the clitic *hi* or *li*. When the prepositional phrase refers to a person or institution and personified thing, we should assign it the <ci>; otherwise, it can be an adjunct (<cc>). This happens, for instance, in: *el retret que la societat dirigeix a aquestes accions* ‘the reproach the society give to these actions’ (111.2)

²⁰Compare:- *he telefonat a la Maria* ‘I phoned Maria’ > **Li** he telefonat
- *he telefonat a la Biblioteca* ‘I phoned to the library’ > **Hi** he telefonat

²¹In some Romance languages, but rare in English: is the use of the dative case of a pronoun to signify that the person (or thing) being referred to is regarded with interest.

(3.33) *el nen li ha aprovat la Selectivitat*
 ‘the child CL-DAT passed the university entrance examination’

(3.34) *la meva filla se'm menja les ungles*
 ‘my daughter REFLX CL-DAT eats the fingernails’ (lit., ‘my daughter eats to me the fingernails’ or ‘my daughter eats my fingernails’)²²

3.8 Attribute

The attribute appears only in attributive sentences (that is, with the verbs *ser* ‘to be’, *estar* ‘to be’ and *semblar* ‘to seem, to look like’). It is an obligatory element of this kind of sentences. Attributes can be replaced either by *ho* or *el*. *Ho* replaces the so called characterizing attributes

- *en Carles és de bon conformar* ‘Carles is easy to please’ > *en Carles ho és*

while *el* replaces identifying attributes

- *el teu pare és el director del banc* ‘your father is the director of the bank’ > *el teu pare l'és*²³

An attribute can correspond to:

(3.35) **a finite completive clause**

no pot ser que es tardi vuit mesos per donar un permís de residència (Cat 104-1)

‘it cannot be that eight months are needed to give a residence license’

(3.36) **a noun phrase**

les petites i mitjanes empreses són el futur del país (Cat 104-1)

‘small and medium size companies are the future of the country’

(3.37) **an adjective phrase**

les condicions són diferents des de fa uns anys (Cat 104-1)

‘the conditions are different since a few years ago’

(3.38) **a non-finite participle clause**

Segons l'avantprojecte està previst que ... (Cat 104-2)

‘According to the preliminary plan, it is expected that ...’

(3.39) **a prepositional phrase**

el parc de vehicles ... és de 16 conjunts de vagonets (Cat 104-2)

‘the vehicle fleet amounts to 16 sets of couches’

²²GCC (p.1110).

²³GCC (pp.1122-1123)

(3.40) a non-finite completive clause

el seu objectiu és aconseguir que l'aeroport de l'Estat tingui caràcter intercontinental (Cat 104-3)

‘his/her aim is to get the State airport to be intercontinental’

Sometimes, both the attribute and the subject are noun phrases. See section 3.5.2 for further details and the criteria used to distinguish between them.

3.8.1 Particular cases

1. We do not consider those structures such as *és que ...*²⁴ to be impersonal. Instead, we analyze them as having an elliptic subject (which would correspond semantically to nouns such as *la causa*, *el motiu* ‘the reason’). The completive clause following the verb works as the attribute.
2. Structures such as *són les tres* ‘it is three o’clock’; *és dimecres* ‘it is Wednesday’; *és de nit* ‘it is nighttime’; *era prop de la una* ‘it was almost one o’clock’; *ja és tard* ‘it is already late’; *és aviat per saber-ho* ‘it is too early to know it’, in particular, usually show just one element in postverbal position²⁵.
3. Another particular structure that we consider to be attributive is that of emphasizing with *ser* ‘to be’:

- *és ell qui* ‘it is him who...’
- *és això que* ‘it is that what...’
- *és aquí on/que* ‘it is here where...’
- *és ara quan/que* ‘it is now when...’
- *és així com/que* ‘it is that way how...’

Since we do not have a more exhaustive and pragmatic analysis yet, we consider these structures in the following way:

- (a) the phrase holding the immediate postverbal position (whatever its type is) is given the **attribute** role;
- (b) the structure introduced by a relative pronoun is analysed as a relative clause without an antecedent (that is, a finite completive clause, for it is given the **subject** role).

²⁴According to GCC (S14), in these cases the completive clause following the copulative verb is not the attribute but the subject, because sentences like *és que són les dues* ‘it is that it is two o’clock’ can paraphrase with *la causa és que són les dues* ‘the reason is that it is two o’clock’. The issue of the pronominalization is still unclear, and for this reason we decided to take the ‘optional’ element (the reason) as the subject in the basis of the fact that, in Catalan and Spanish, the subject is optional, while the attribute is not. Thus, the completive, as obligatory, is the attribute of the sentence.

²⁵This issue is still unclear: According to GCC, since some adverbs can hold subject position (as in *ara són les tres* ‘Now (it) is three o’clock’), the element after the verb should be considered to be the attribute. However, it is this element the one agreeing with the verb. Moreover, the pronominalization test confirms this fact: *ara són les tres* > *ara les són*.

3.9 Predicative complement

The predicative complement is a structure which depends on the verb but at the same time agrees with either the subject or the object of the sentence. Some of them are obligatory (*han declarat culpable el testimoni principal* ‘(they) found guilty the main witness’), while others are optional (*la dona va mirar nerviosa els seus fills* ‘the woman looked nervously at her children’). Generally, a predicative complement corresponds to a noun or adjective phrase, although sometimes a prepositional phrase can hold this syntactic function as well (see below), as long as its term shows the required agreement with the subject or the object of the sentence²⁶. In Catalan, in addition, the clitic *hi* may replace the predicative complement²⁷, as shown in examples 3.41 to 3.47:

- (3.41) *Els meus pares s’han posat contents > s’hi han posat* ‘My parents were very happy’
- (3.42) *La Carme ha arribat cansada > hi ha arribat* ‘Carme arrived tired’
- (3.43) *Encara va poder menjar-se el sopar calent > va poder menjar-se-l’hi*²⁸ ‘(s/he) was still able to eat the hot supper’
- (3.44) *La dona va mirar nerviosa els seus fills > va mirar-los-hi* ‘The woman looked nervously at her children’
- (3.45) *La dona va veure nerviosos els seus fills > va veure’ls-hi* ‘The woman saw her children nervous’²⁹
- (3.46) *Han designat candidat en Josep > hi han designat en Josep; l’hi han designat* ‘(They) designated Josep candidate’
- (3.47) *Han declarat culpable el testimoni principal > hi han declarat el testimoni principal; l’hi han declarat* ‘(They) found the main witness to be guilty’

Certain prepositional phrases (generally those headed by the preposition *com* or the prepositional locution *com_a*) followed by a noun or adjective phrase that predicates the subject or the object of the sentence, introduce a predicative complement. However, there is a constraint on these kind of structures: The predication is an essential condition, otherwise it would be annotated as an adjunct. In 3.48, for instance

- (3.48) (Cat 104-1.23) *com_a primera aproximació hem coincidit en ...*
‘as a first approximation we agreed on ...’

²⁶In addition to subjects and direct objects, a predicative complement can also agree with a prepositional object, although this is infrequent.

²⁷Be that as it may, this is not always true: the pronoun *hi* cannot replace the predicative in the sentence *La Lluïsa fa de directora*.

²⁸These three first examples are taken from GCC (pp.1994-1995).

²⁹Examples 3.44 and 3.45 are taken from GCC (p.1119). The adjective ‘nervous’ predicates the subject in example 3.44, while in 3.45 it predicates the object.

com_a does not predicate the subject nor the object. A quite reliable test would be to check whether the structure is semantically equivalent to *així* ‘this/that way’.

In addition to prepositional phrases, we can also find a few adverbial phrases and completive clauses holding a predicative role. Some examples (3.49 to 3.56) are taken from the corpus:

- (3.49) (Cat 5009_20000320) *no es troba bé*
‘(s/he) is not feeling well’
- (3.50) (Sp 9155_20000712) *no me siento bien*
‘I’m not feeling well’
- (3.51) (Sp 137_20020501) *teníamos mucho en común*
‘(we) had many things in common’
- (3.52) (Sp t6-4) *dejándole a solas*
‘(s/he) leaving her/him alone’
- (3.53) (Sp 121_19991202) *ha terminado fichando por Ford*
‘(s/he) finally signed up for Ford’
- (3.54) (Sp 79_20020401) *no me veo haciendo siempre lo mismo*
‘(I) can’t see myself doing always the same’
- (3.55) (Cat 74_19981101_i) *conegut com el violador de la careta*
‘known as the rapist of the mask’
- (3.56) (Sp 13150_20000316) *al cual no dudó en calificar como el mejor equipo*
‘to which (s/he) did not hesitate to qualify as the best team’

According to GCC³⁰, not only *hi* but also the pronoun *en* can replace the predicative pronoun:

- (3.57) *Abans de l’operació mirava cansat, però ara ja no hi mira*
‘Before the operation (s/he) looked tired, but now (s/he) doesn’t’
- (3.58) *A Lluís el van nomenar president la setmana passada i a mi me’n nomenaran la setmana vinent*
‘(They) appointed Lluís as president last week, and (they) will appoint me -as president- next week’

This is related with the fact that, according to GCC³¹, there are ‘residual’ cases which predicate the dative (indirect object) or the prepositional object. The following are examples from that text:

³⁰S6, pp. 1371 and the following ones.

³¹S14.

- (3.59) *Li hem posat Miquel* ‘(We) called him Miquel’ > obligatory predication on the dative
- (3.60) *Li va agafar l’atac al cor asseguda al cotxe* ‘(s/he) got the heart attack sitting inside the car’ > non-obligatory predication over the dative
- (3.61) *Es queixa de la Maria com a metgessa, no com a amiga* ‘(S/he) complains about Maria as a doctor, not as a friend’ > obligatory predication over the prepositional object
- (3.62) *Cada dia que passava pensava encara més en la Carme [alliberada/fora de la presó]* ‘As days went by, (s/he) was thinking even more on Carme (being) [liberated/out of prison]’ > non-obligatory predication over the prepositional object

Apart from the corresponding label (<func=‘cpred’>), we complete the information adding a value for the attribute <predicate> in order to specify whether the complement predicates the subject, the direct object or the prepositional object:

- <... func=‘cpred’ predicate=‘suj’ ...>
- <... func=‘cpred’ predicate=‘cd’ ...>
- <... func=‘cpred’ predicate=‘creg’ ...>

When the predicative complement occurs within a non-personal sentence (participle clauses, mainly) or in cases of double predication, the additional attribute <predicate> is irrelevant, as shown in 3.63 and 3.64:

1. participle clauses, in which there is no grammatical subject nor object, although showing agreement:

- (3.63) (104-4.17) *un nou model de tren dissenyat en_base_al TGV, denominat ETR_550*
 ‘a new model train designed in basis of the TGV, called ETR_550’

2. double predication structures, in which a complement predicates two different complements at a time:

- (3.64) (06010100-49.5) *fer compatibles l’activitat turística amb la conservació del patrimoni natural i cultural*
 ‘to make compatible the turist activity with the preservation of the natural and cultural heritage’

3.9.1 Special cases

There are some cases in which it is not clear whether the complement predicates the subject or the object. This happens when the main verb is reflexive. Since structures like

- *ell ha mostrat la seva obra molt satisfet* ‘he showed his work with satisfaction (lit., very satisfied)’

are possible, and there would be a complement predicating the subject; we consider that, in the case of the reflexive version of the verb *mostrar-se* ‘to show oneself’, it should be analyzed as a predicative complement of the subject as well, giving it the <subj> value for the attribute <predicate>:

- *Guinovart s’ha mostrat satisfet del resultat* ‘Guinovart showed himself as being satisfied with the result’

However, if when undoing the reflexive construction there were a predication of the direct object, then it should be assigned the <cd> value for the attribute <predicate>. This means that, a sentence like *Antonio.N. va declarar-se responsable de ...* ‘Antonio.N. declared himself responsible for...’, could also be:

- *Antonio va declarar-los responsables de...* ‘Antonio declared them responsible for...’

3.9.2 Predicative with *ser* (‘to be’)

As mentioned in section 3.8, the complement agreeing with the subject and occurring with the verb *ser* ‘to be’ has to be the attribute. However, as we will see in section 3.10.3.3, *ser* has a predicative (locative) use: *En Joan és a Girona* ‘Joan is in Girona’. In such cases, there can cooccur another complement showing agreement with the subject as well. This complement, is analysed as <func=‘cpred’ predicative=‘subj’> (3.65). Another case is that of *ser* as the main verb of a periphrasis in which the auxiliary is *fer* ‘make’ (in its causative sense), as in 3.66:

(3.65) (Cat 06010300-15.4) *Ravelo és a Girona convidat per...*
 ‘Ravelo is in Girona invited by...’

(3.66) (Cat 06019900-14.12) *...els fa ser optimistes*
 ‘...(it) makes them be optimist’

3.10 Prepositional object

A prepositional object (in the Romance tradition *complement de règim verbal*) is generally a prepositional phrase, in which the head is a preposition ruled by the verb. This complement allows pronominalization by either *hi* and *en*. The clitic *en* replaces those

prepositional objects introduced by the preposition *de*, while *hi* replaces the rest of the cases (prepositions *a*, *amb*, *en*, *per*).

A prepositional object can correspond to the following structures³²:

(3.67) **a prepositional phrase**

1. *perquè l'ajuntament es personi en el procediment judicial* (Cat 111-9)
'for the city council to appear before the court proceedings'
2. *per interessar-se pel seu estat* (Cat 111-9)
'to be interested/to ask for his/her health'
3. *Pasqual Maragall carrega contra els dirigents que opten a succeir Jordi Pujol* (Cat 111-9)
'Pasqual Maragall charges at the leaders that choose to succeed Jordi Pujol'

(3.68) **a finite completive clause** *ha advocat en conferència de premsa perquè el partit no prescindeixi de persones com...* (Cat 111-4)

'(s/he) plead in press conference for the party not to do without people like...'

(3.69) **a noun phrase**

1. *Vull saber què en pensa Rodríguez Zapatero* (Cat 111-10)
'I want to know what Rodríguez Zapatero thinks (about it)'
2. *en sumar-s'hi el líder socialista* (Cat 111-3)
'when the socialist leader joined it'

(3.70) **an adverbial phrase**

1. *les vendes van bé ... però aquesta va més enllà* (Cat 104-5)
'the sales go well ... but this one is going beyond'

(3.71) **a non-finite adverbial clause**

1. *que es troba vorejant el mar* (Cat 103-21.5)
'which goes along the sea'

Usually, each verb selects a particular preposition to introduce the prepositional object. However, some verbs admit different prepositions at a time, implying a difference in meaning only in some of the cases:

- *parlar de / sobre* 'to talk about'
- *coincidir amb / en* 'to agree with / on'

³²There is a phenomenon, *Canvi i caiguda de preposició* 'Preposition swap and drop', affecting this complement. See 3.10.4 for further details.

- *interessar-se per / en* ‘to get interested in’
- *lluitar per / contra / a favor* ‘to fight for / against / in favor of’

A special case is that of the verb *participar*:

(3.72) *participar a / en* ‘to take part’

- *participar a la reunió* ‘to take part of the meeting’

(3.73) *participar de* ‘to have in common’

- *aquestes dues teories participen dels mateixos axiomes* ‘these two theories share the same axioms’

3.10.1 Verbs requiring either direct or prepositional object

Some verbs can occur with or without a ruled preposition. This can imply a difference in meaning, although in many cases both variants can be taken as synonymous:

(3.74) *pensar / pensar en*

1. *pensem que és necessari enfortir el partit* (Cat 211nou)
‘(we) think that it is necessary to make the party stronger’
2. *...al marge de pensar en Alemanya...* (Cat 9121_20000519)
‘...besides thinking about Germany...’

(3.75) *tractar / tractar de*

1. *les principals temàtiques que ha tractat en l'últim any* (Cat 06010300-15)
‘the main issues that (s/he) worked on during last year’
2. *‘Rius_per_viure’ tracta de la història dels moviments socials entorn del riu* (Cat 06019900-19)
‘Rius_per_viure’ deals with the history of social movements about the river’
3. *tratando de no caerse ni hacer el ridículo* (Sp a1-1)
‘trying not to fall down and make a fool of (him/her)self’

3.10.2 Verbs with both direct and prepositional object

(3.76) *basar quelcom en*

El col·legi basa [aquesta iniciativa<cd>] [en un estudi elaborat<creg>] (Cat 4_19991201)
‘The school bases this initiative on an elaborated study’

(3.77) *informar algú de*

la web informa [als visitants<cd>] [de les activitats que es poden fer<creg>]
(Cat 103-22)
‘the site informs the visitors on the available activities’

3.10.2.1 Second prepositional object

The second prepositional object (*complement de règim segon*) is a prepositional phrase which is not a dative nor a locative complement that cooccurs with a direct object. There are two kinds of second prepositional object:

(3.78) *Sentence structures introduced by a preposition which is required by the main verb*

1. *La van obligar a estudiar anglès* ‘(they) forced her to study English’

(3.79) *Prepositional phrases that are the obligatory alternative to a plural direct object*³³

1. *comparar una cosa amb una altra* > *comparar dues coses* ‘to compare a thing to another > to compare two things’
2. *dissociar quelcom de* > *dissociar dues coses* ‘to dissociate something from something else > to dissociate two things’
3. *relacionar algú amb* > *relacionar dues persones* ‘to relate someone to someone else > to relate two people’

3.10.2.2 Verbs with two prepositional phrases

There is a small set of verbs that usually occur with two prepositional phrases, only one of them being argumental. The following scheme shows how we analyze their complements³⁴:

1. *col·laborar [amb algú<creg>] [en alguna cosa<cc>]* ‘to collaborate with someone in something’
2. *cooperar [amb algú<creg>] [en alguna cosa<cc>]* ‘to cooperate with someone in something’
3. *tractar [amb algú<cc>] [sobre alguna cosa<creg>]* ‘to talk with someone about something’
4. *discutir [amb algú<cc>] [sobre alguna cosa<creg>]* ‘to discuss with someone about something’

3.10.2.3 Verbs alternating a non-pronominal transitive structure with a pronominal one with prepositional object

Some transitive verbs require a prepositional object only in their pronominal version. This implies a slight difference in meaning:

³³In some cases, the second prepositional object is optional.

³⁴According to GCC.

(3.80) **reunir / reunir-se amb**³⁵

1. *La mostra també reuneix una selecció dels millors elapés de tots els gèneres musicals* (Cat 203-5)
‘The exhibition also gathers a selection of the best lp’s of all musical genres’
2. *els alcaldes i el president del Consell s’havien reunit amb els representants dels treballadors* (Cat 107-16)
‘the majors and the Council president had a meeting with the employers’ representative’

(3.81) **tractar / tractar-se de**³⁶

1. *el servei terapèutic va tractar 101 casos* (Cat 06020300-28)
‘the therapeutic services treated 101 cases’
2. *Es tracta d’una contrarrelotge de 3000 quilòmetres* (Cat 06020300-4)
‘it is a 3000 kilometers timed event’

3.10.3 Particular cases**3.10.3.1 The verb *canviar* ‘to change’**

The verb *canviar* ‘to change’ followed by the preposition *de* occurs in two different structures: The first one refers to intrinsic (both physical and psychological) properties of the subject:

- *canviar de pentinat* ‘to change one’s hairstyle’
- *canviar de pensament* ‘to change one’s mind’

while the second possible structure deals with external elements:

- *canviar de feina* ‘to change one’s job’
- *canviar de casa* ‘to change one’s house’
- *canviar de roba* ‘to change one’s clothes’

³⁵Notice that some of these verbs can also show an alternation between a plural direct object construction and a singular direct object plus a prepositional object. According to what it is said above, then, these verbs show a threefold alternation; this is: *reunir (dues coses)* ‘to join, to gather (two things)’, *reunir una cosa amb una altra* ‘to join something with something else’ and *reunir-se amb* ‘to join oneself with somebody else / to meet / to have a meeting with’.

³⁶The non-pronominal version of this verb also admits prepositional object and implies a slight difference in meaning: *el metge tractarà la malaltia* ‘the doctor will treat the disease’ / *la conferència tractarà de/sobre les malalties* ‘the conference will deal with the diseases’ (see section above). The pronominal version of the verb, apart from requiring prepositional object as well, is an impersonal structure.

The prepositional phrase introduced by *de* is unnecessary in the first group, while in the second is obligatory: The sentence *ha canviat* '(s/he) has changed' cannot be interpreted as *ha canviat de casa* '(s/he) has changed his/her house', whereas it can replace *ha canviat d'idea* '(s/he) has changed his/her mind'. For this reason, we analyze the first group as adjuncts (<cc>) and the second as argumental (this is, prepositional objects, <creg>).

3.10.3.2 The verb *servir* 'to serve, work as'

Servir can occur in two different structures, depending on the head of the complement: When the prepositional phrase is introduced by *per* (3.82), this complement is an adjunct, while it is an argumental prepositional object when the complement is headed by *de* (3.83):

(3.82) (Cat 06029900-13.3) *la festa servirà per acomiadar el camió de menjar*
'the party will serve to say goodbye to the food-truck'

(3.83) (107-13.12) *una canonada de gran diàmetre que serveix de drenatge de l'aqüífer*
'a big-diameter piping that works as the aquifer drain'

3.10.3.3 Obligatory complements expressing place or manner

Those complements expressing place or manner that are obligatory and can be replaced by an adverbial clitic are analyzed as prepositional objects³⁷:

(3.84) *En Lluís vesteix elegantment / amb elegància*
'Lluís dresses elegantly / with elegance'

(3.85) *Els nostres fills es porten prou bé*
'Our children behave quite well'

(3.86) *Han anat a Cardedeu / cap allà / ben lluny / allà*
'(They) went to Cardedeu / that way / far from here / there'

(3.87) *El meu germà resideix a Gandia / al costat de Gandia / prop de Gandia / allí*
'My brother lives in Gandia / next to Gandia / close to Gandia / there'

(3.88) *El llibre es troba a la taula / damunt la taula / dins el calaix / allà*
'The book is on the table / on the table / in the drawer / there'

We also include in this analysis those locative complements in *ser / estar* 'to be' structures:

(3.89) *persones que han estat en primera línia de l'organització* (Cat 111-4.12)
'people that have been in the head of the organization'

³⁷GCC (pp. 1113-1114)

- (3.90) *perquè estem en la bona direcció* (Cat 111-6.2)
‘because (we) are in the right direction’
- (3.91) *a tres quarts d’una seran a la Fundació* (Cat 107-17.7)
‘(they) will be at the Foundation at a quarter to one’
- (3.92) *el seu cost estaria entre 300 i 600 milions per quilòmetre* (Cat 204-4.1)
‘its cost would be between 300 and 600 milions per kilometer’
- (3.93) *el funicular, que estava a l’estació inferior* (06010100-24)
‘the cable car, that was in the lower station’
- (3.94) *Noureddine_Hassoun esatava a la zona de la platja de Lloret de Mar* (Cat 107-15.2)
‘Noureddine_Hassoun was in the area of the beach at Lloret de Mar’
- (3.95) *per estar-s’hi una setmana* (06019900-22.4)
‘to be there during a week’

The generalization we can infer from these examples is that any obligatory complement that is not a noun phrase should be analyzed as a prepositional object (<creg>). This is what happens in 3.96, for instance, in which the verb *passar* ‘to pass, go through’ cannot appear alone:

- (3.96) *una prioritat que passa per una ampliació de recursos* (Cat 104-1)
‘a priority that goes through a resources increase’

3.10.3.4 Motion verbs with two complements

Some motion verbs require two complements at a time. When this happens, only one of them is obligatory: *passar (de A) a B* ‘to pass (from A) to B’; *anar (de A) a B* ‘to go (from A) to B’ (3.97 and 3.98). In such cases, we consider the optional complement (the one introduced by *de*) to be an adjunct (<cc>), and the obligatory one (introduced by *a* or *fins a*) to be an argumental prepositional object (<creg>). Unlike these two verbs, *venir* does not require any of the complements, and for this reason we label two adjuncts:

- (3.97) *una millora que pot anar [des_ de la reducció de costos <cc>] [fins a la disminució del cicle <creg>]* (Cat 107-1.11)
‘an improvement that can go from costs’ reduction to the cicle decrease’
- (3.98) *En aquestes condicions se’n van anar [al Pla dels Socs <creg>]* (Cat 207-10)
‘In this conditions they went to Pla dels Socs’

Regarding the verb *passar* ‘to pass, go through’, as mentioned above, only those complements introduced by *per* are analyzed as prepositional object (3.99). This preposition, however, can also introduce a predicative complement (predicating the subject) when the noun or adjective introduced by *per* agrees with the subject, as in 3.100. Otherwise, the complement is an adjunct (<cc>).

(3.99) *la prova ... passarà ... pel Pont de la Palanca* (06010300-24)
 ‘the race ... will go ... through the Pont de la Palanca’

(3.100) *passar per tonto* ‘to pass for silly’

Another particular case is that of the verb *arribar* ‘to arrive’. When it has a strict locative meaning (to get to a physical place, city, etc.), the complement is an adjunct, since it is not obligatory (3.101):

(3.101) *L’expedició arribarà a l’aeroport del Prat a dos quarts d’una del migdia de dilluns* (107-17.6)
 ‘The expedition will get the airport at half past twelve on Monday’

However, when *arribar* holds a non-locative meaning (‘to reach’, for instance), the complement should be analyzed as an argumental prepositional phrase:

(3.102) *espera arribar a un acord* (Cat 06010100-6)
 ‘(s/he) hopes to reach an agreement’

(3.103) *arribant als 360 km/h* (Cat 104-4)
 ‘reaching 360 km/h’

3.10.4 Preposition Swap and Drop

Catalan (unstressed) prepositions can be deleted or changed due to their phonetic properties by undergoing a process which is known as *Canvi i Caiguda de Preposició* (Preposition Swap and Drop), being (supposedly) affected only those prepositions which are both unstressed and governed by a verb. The pattern for this phenomenon is described by the norm as follows:

(3.104) Introducing non-finite clauses

1. Prepositions *a* and *de* do not undergo swap or drop
2. Prepositions *en* and *amb* change into *a* or *de* with certain variability depending on the verb

(3.105) Introducing finite clause

1. Unstressed prepositions are deleted

The following are some examples taken from the corpus:

(3.106) *... va insistir en els arguments que va anar esgrimint el seu cap de files* (Cat 74_20000201)
 ‘(s/he) insisted on the arguments her/his rank file boss was putting forward’ >
 noun phrase, no drop

- (3.107) ... *insisteixin a portar els temes lingüístics als tribunals* (Cat 24209_20010126)
 ‘(they) should insist on taking linguistic issues to court’ > non-finite clause; swap
 (*a* replaces *en*)
- (3.108) *Carod ha insistit Ø que ERC ja no pot fer més* (Cat 39886_20011003)
 ‘Carod insisted that ERC cannot do anything else’ > finite clause; drop

Counterexamples are found in the corpus, however, showing certain hesitation as regards this process. Prepositions in the following examples are supposed to change into *a* since they precede non-finite clauses. Unlike what is expected from the norm, the preposition *en* or *a* introducing the non-finite clause in 3.111 undergoes deletion³⁸:

- (3.109) ... *consistirà en fer les obres d’urbanització del carrer de la Font del Gat* (Cat 10588_20000609)
 ‘... will consist on the urbanization works at Font del Gat street’
- (3.110) *La iniciativa consisteix en repartir la guia del Tax_Free* (Cat 13218_20000728)
 ‘The initiative consists on sharing the Tax_Free guide’
- (3.111) *La companyia confia Ø poder augmentar la freqüència d’aquest vol ...* (Cat 10482_20000608)
 ‘The company hopes to be able to increase the frequency of this flight...’

3.11 Agent complement

The agent complement only occurs in passive structures (periphrastic, reflexive and participle passives) and is always a prepositional phrase introduced by *per* (*por* in Spanish)³⁹. It corresponds to the subject of its active and transitive equivalent structure:

- *sopar organitzat ahir a la nit per la patronal de la petita i mitjana empresa* (Cat 104-1.2)
 ‘a dinner organized by the management of the small and medium-sized enterprises’

Actually, the fact that the agent complement corresponds to the subject in a transitive structure is more relevant than whether or not this complement has human or animated features. In this sense, in *es troba emmarcada per edificis* ‘it is surrounded by buildings’, *per edificis* is the agent complement since the equivalent active structure would be *els edificis emmarquen ‘el que sigui’* ‘the buildings surround whatever’.

These are some examples taken from the corpus:

³⁸The phonetic properties of preposition *a* can be the motivation for the deletion: The contiguity of two unstressed vowels in Catalan leads to the deletion of one of them, specially when involving exactly the same features, as it happens in the example (*confia a* becomes *confia*).

³⁹There is a small set of verbs, however, that admits the agent complement introduced by *de* and similar prepositions and prepositional locutions such as *a_càrrec_de*, *per_part_de*, *entre*.

- (3.112) *La proposició no de llei ha estat presentada ... per la diputada del grup PSC-CpC* (Cat 111-1)
 ‘The Private Member’s Bill was presented ... by the Member of the Parliament from the party group PSC-CpC’
- (3.113) *es presenta als assistents el projecte de l’abocador a_càrrec_dels portaveus de la plataforma* (Cat 104-13.16)
 ‘it is presented to those present the project of the rubbish dump by the spokespersons of the platform’
- (3.114) *que siguin coberts a parts iguals entre el Govern estatal, la Generalitat i els Ajuntaments* ‘that (they) are covered in equal parts between (by) the state Government, the Generalitat and the Councils’ (Cat 06019900-8)

3.12 Adjunct

Adjuncts⁴⁰ (within the romance tradition, *complement circumstancial*) are complements which are not selected by the verb; this is, they are optional elements and have no incidence in sentence well formedness. Locative, manner, comitative or instrument adjuncts admit pronominalization with *hi* and *en*.

Generally, adjuncts correspond to adverbial and prepositional phrases, and sometimes to other equivalent structures, such as finite and non-finite adverbial clauses. When ambiguity between adjunct (<cc>) and prepositional object (<creg>) occur, the criteria must be the obligatoriness of the second one:

(3.115) **adjunct** (optional)

1. *...ha arribat a totes les províncies catalanes* (Cat 104-1)
 ‘(it) arrived to all Catalan provinces’

(3.116) **prepositional object** (obligatory)

1. *acudeixin a aquesta cita* (Cat 104-1)
 ‘(they) assist to this appointment’
2. *participar en un sopar amb empresaris* (Cat 111-1)
 ‘to take part in a business dinner’

For the time being, we do not make a formal distinction among the different kinds of adjuncts, apart from locative and temporal ones⁴¹.

⁴⁰See section 2.3.2.1 for a terminological usage explanation of *Adjunct* within this document.

⁴¹See section 3.3 for further details.

3.12.1 Particular cases

3.12.1.1 Adverbs *ja* and *encara*

Adverbs *ja* and *encara* usually hold preverbal position and can work as modality elements (<mod>) or adjuncts (<cc>). This is basically a semantic difference, however, and whether or not the complement indicates temporality can help us to disambiguate. They are analyzed as modality elements (<mod>)⁴² when they just give emphasis or are equivalent to *potser* ‘maybe’ (3.117 and 3.118), while we annotate the rest of the cases as adjuncts (<cc>) (3.119 and 3.120):

(3.117) (Cat 06010100-39)

*que permet provar **encara** amb més contundència*
‘what allows (us) to prove even more forcefully’

(3.118) (Cat 122_20010801)

***ja** és suficient per saber que Saviola s’ha apoderat del 7*
‘(it) is enough to know that Saviola got the 7’

(3.119) (Cat 111nou)

*institucions culturals que **encara** figuren a nom de l’estat*
‘cultural institutions that are still a property of the State’

(3.120) (Cat 111nou)

*Rajoy **ja** coneixia aquests dos projectes*
‘Rajoy already knew these two projects’

3.12.1.2 The preposition *sobre*

The preposition *sobre* introduces modifiers only when its meaning is strictly locative. Otherwise, when introducing the topic of the sentence, it is a sentence adjunct (<ao>)⁴³.

3.12.1.3 Adverbs *més* and *menys* and other quantifiers

We consider quantifiers like *més* and *menys* (among others) to be modifiers when they intensify a verbal form or a participle adjective:

(3.121) (Cat 104nou)

*han desenvolupat projectes **més** destacats...*
‘(they) developed more distinguished projects...’

(3.122) (Cat 107-9)

*el programa de cursos, però, és **força** variat*
‘the courses program, however, is quite varied’

⁴²See section 3.14 for further details on modality elements.

⁴³See section 3.18 for further details on sentence adjunct

3.12.1.4 Participle clauses and absolute participle clauses

Both participle and absolute⁴⁴ participle clauses are analyzed as modifiers (<cc>) when occurring at the beginning of the sentence.

3.12.1.5 The verb *arribar* ‘to arrive’

As mentioned before in this document, the verb *arribar* ‘to arrive’ can also mean ‘to reach’. In such a case, the complement is analyzed as a prepositional object (<creg>) and, otherwise, as and adjunct (<cc>):

1. *arribar a Barcelona* ‘to arrive in Barcelona’ (adjunct)
2. *arribar al 10 per cent* ‘to reach the 10 per cent’ (prepositional object)

3.12.1.6 The verb *servir*

This verb, as already mentioned in section 3.10.3.2, can cooccur either with a modifier or with a prepositional object, depending on the preposition introducing the complement:

(3.123) (107-13)

una canonada de gran diàmetre que serveix de drenatge de l'aqüífer
‘a big-diameter piping that works as the aquifer drain’

(3.124) (Cat 06029900-13)

la festa servirà per acomiadar el camió de menjar
‘the party will serve to say goodbye to the food-truck’

3.12.1.7 The verb *canviar*

The verb *canviar* ‘to change’ followed by the preposition *de* occurs in two different structures: The first one refers to intrinsic (both physical and psychological) properties of the subject. In such a case, the complement is a modifier (<cc>). The second possible structure deals with external elements and is analyzed as a prepositional object (<creg>)⁴⁵.

3.12.1.8 Verbs always cooccurring with modifiers

There are some verbs (*partir* ‘to leave’, *néixer* ‘to be born’, *sortir* ‘to come out’, *entrar* ‘to come in’) that are always complemented by an adjunct:

(3.125) (Cat 107-17)

els expedicionaris sortiran al balcó de l'Ajuntament
‘the expedition members will come out to the City Hall balcony’

⁴⁴These two structures only differ in explicit subject, which appears only in the absolute participle clauses. See the constituency guidelines for further details.

⁴⁵See section 3.10.3.1 for further details on this difference.

3.12.1.9 The verbs *passar*, *anar* i *venir*

Passar ‘to pass, go through’, *anar* ‘to go’ and *venir* ‘to come’, among other motion verbs, require two complements, although only one is obligatory. The optional one should be analyzed as a modifier⁴⁶.

3.12.1.10 Other complex structures as modifiers

Constructions such *mar endins* ‘out to sea’ show a particular structure. According to GCC⁴⁷, they are *highly grammaticalized bimember structures, which tend to get fixed*. The following are some of their characteristics:

- semantically, the noun element indicates a trajectory and the adverb indicates the direction;
- they are semiproductive structures: They do not admit all adverbs or nouns;
- they are highly grammaticalized, since they cannot be expanded by inserting elements and do not admit specifiers or complements;
- the noun element does not admit variation in number;
- they show syntactic defectiveness:
 1. the adverb does not admit specifiers (*carrer (*molt) amunt* ‘street very above’)
 2. some specifiers that affect the whole structure are allowed (*navegava [molt/més] mar endins* ‘(s/he) sailed very/more out to sea’)
 3. it is only possible to coordinate complete structures or to coordinate just the adverbs (*carrer amunt i carrer avall* ‘street up and street down’ or *carrer amunt i avall* ‘street up and down’)
- the deletion of any of the elements implies a change in meaning.

The criteria we adopt to analyze these structures within the AnCora framework is the following:

1. the head of the structure is the adverb
2. the noun phrase is adjoined to the adverb
3. the syntactic function of the whole set is modifier (3.126)

(3.126) *que recorrerà uns deu quilòmetres [mar endins <cc>]* (Cat 107-18.4)
 ‘that will cover about ten kilometers out to sea’

⁴⁶See section 3.10.3.4 for further details on this kind of verbs and their complements.

⁴⁷S 11.3.4.

3.13 Textual element

We consider to be textual elements (<et>) all those structures (adverbial and conjunctive, mainly) that have no syntactic function within the sentence. However, these elements link sentences in the text. Generally, they occur at the beginning of the sentence or after coordinating conjunctions. They are expressions such as *de fet* ‘actually, in fact’, *no obstant* ‘nevertheless’ and conjunctions occurring after a full stop, such as *Però ...* ‘But...’. We also include in this set discourse-structuring elements like *en primer lloc* ‘first’, *en segon lloc* ‘second’, etc.

The difference between a textual element (<et>) and a sentence adjunct (<ao>)⁴⁸ is that textual elements do not contribute semantically to the sentence to which they belong. *En primer lloc*, ‘First’, only puts in order the discourses, for instance. On the other hand, a phrase like *en l'àmbit del finançament* ‘within the funding sphere’ gives some content to the sentence.

(3.127) *Per la seva banda, el PSOE madrileny va responsabilitzar el Govern regional...*
(Cat 104nou)

‘On the other hand, the Madrid PSOE hold the regional Government responsible for ...’

(3.128) *No obstant això, González ha deixat clar que Pimec-Sefes continuarà treballant...* (Cat 104nou)

‘Nevertheless, González said clearly that Pimec-Sefes will continue working...’

Some other structures we analyse as textual elements (<et>):

- *a més a més* ‘moreover’
- *en aquest sentit* ‘in this sense’
- *de moment* ‘at the moment’
- *d'altra banda* ‘on the other hand’
- *així mateix* ‘in this way’
- *per tant* ‘therefore’

3.14 Modality element

Modality elements (<mod>) are mainly adverbs holding the immediate preverbal position: *no*, *també* ‘also’, *fins i tot* ‘even’, *encara* ‘maybe’. They do not correspond to any adjunct and they deal with sentence modality:

⁴⁸See section 3.18.

(3.129) (Cat 104nou)

*en el qual **també** va advertir els empresaris, sindicats i administracions que compleixin amb la normativa*

‘in which (s/he) also warned the businessmen, unions and administrations to fulfill the rules’

(3.130) (Cat 104nou)

*En el primer cas, **per exemple**, s’hauria de salvar...*

‘In the first case, for example, it should be saved...’

(3.131) (Cat 104nou)

*que **gairebé** s’ha duplicat en pocs mesos*

‘that has been almost doubled in a few months’

(3.132) (Cat 104nou)

*...en aquest moment **només** mantinc una vinculació de tipus mercantil*

‘... in this moment I **just** keep up a commercial connection’

Some other structures we analyse as modality elements (<mod>):

- *ja*⁴⁹ (emphatic particle)
- *amb tota probabilitat* ‘in all probability, likelyhood’
- *únicament* ‘just, only’
- *com a mínim* ‘at least’

3.15 Passive marker

The verbal morpheme *es* is analysed as <pass> when working as a passive marker. All reflexive passive sentences have an obligatory subject (usually holding postverbal position) that agrees with the verb and have no direct object. The main verb, in its active version, must be transitive and can be expressed either in singular or plural. Passive subjects are generally explicit, although sometimes they can also be elliptical (probably because they appear in the previous sentence):

(3.133) (Cat 103-12.8)

Ø s’han de presentar en cinta magnètica o CD

‘(they) should be presented either in magnetic tape or CD’

(3.134) (Cat 107-1)

aquests guardons s’atorquen a les empreses i persones

‘these awards are given to companies and people’

⁴⁹The adverb *ja* can also indicate temporality (‘already’). In such a case, it is analysed as a temporal adjunct. See section 3.12.1.1 for further details on adverbs *ja* and *encara*.

(3.135) (Cat 107-16)

aquestza nova protesta es considera clau en la política de...

‘this new protest is considered to be a key factor in the politics of...’

(3.136) (Cat 103-21)

l'argila que es fa servir per confeccionar estris

‘the clay that is used to make tools’

(3.137) (Cat 107-5)

un parc natural que s'ha de preservar

‘a natural park that must be preserved’

3.16 Impersonality marker

The verbal morpheme *es* can also indicate impersonality (<impers>). The main verb of an impersonal sentence can only occur in singular and there is no element within the sentence agreeing with it, because the presence of this marker blocks the occurrence of the subject of the verb.

(3.138) (Cat 107-19) *per al qual ja es disposa de terreny public*

‘for which (it) is provided a plot of public land’

(3.139) (Cat 104nou) *no s'ha avançat i el tema està parat*

‘no advance has been made and the topic is at a standstill’

(3.140) (Cat 103-12) *ja es treballa en el de l'any vinent*

‘(someone) is working already on the one of next year’

3.17 Vocative

This label (<voc>) only appears when there is direct speech within the text and the hearer is mentioned. There are few cases in the corpus, since most of it is journalistic style. The following example is taken from a written interview:

(3.141) *Benvingut al club, senyor Duran* (Cat 06030200-1)

‘Welcome to the club, mister Duran’

3.18 Sentence Adjunct

We assign this label (<ao>) to those elements which modify the whole sentence structure. There are some different constructions defined:

1. Adverbs showing the speaker's attitude in view of the statement

(3.142) *Desgraciadament, tot ha acabat* (Cat 111nou)
 ‘Unfortunately, everything is finished’

This adverbs usually appear at the beginning of the sentence and followed by a comma.

2. Constructions such as *segons ell* ‘according to him’, *segons ha dit ell* ‘according to what he said’, *per a ell* ‘in his opinion’... This structures hold a variable position within the sentence, and their grammatical category corresponds either to a prepositional phrase or an finite adverbial clause:

(3.143) (Cat 104nou)
que actúa, segons el seu parer, com a fre al creixement econòmic
 ‘which, in his opinion, curbs the economical growht’

3. Sentence elements that refer (or, better, correfer) to the topic of discussion. The structure heading this elements are usually similar to the following:

- *en_relació_a/amb...* ‘in relation to/with...’
- *quant_a...* ‘regarding to...’
- *pel_que_fa_a...* ‘regarding to...’
- *respecte_a/de...* ‘with regard to...’
- *en_referència_a...* ‘in reference to...’
- *sobre...* ‘about...’
- *al_voltant_de...* ‘within’
- *en_comparació_amb...* ‘compared to...’

4. Conditional, concessive and consecutive adverbial clauses and prepositional phrases having a similar value:

(3.144) (Cat 111nou)
mentre no rebutgi la iniciativa del Partit Socialista de Catalunya
haurà d’explicar-ho a Castella i Lleó
 ‘while (s/he) does not refuse the initiative of the Partit Socialista de Catalunya, (s/he) will have to explain it to Castella i Lleó’

(3.145) (Cat 111nou)
després que el grup del PSC-CpC bloquegés la votació a l’últim ple, malgrat
que ja se n’havia pactat prèviament la composició
 ‘after the group PSC-CpC blocked the votation during the last plenary, even though the composition had already been negotiated’

(3.146) (Cat 104nou)

ja que, segons ell, no anem tan bé respecte_a la sinistralitat

‘because, according to him, we are not in the right way regarding occupational injuries’

5. We also consider the adverb *tot* to be a sentence adjunct when it goes with gerunds in structures such as the following:

(3.147) (Cat 107-3)

L'assemblea tindrà lloc a la sala de plens ..., tot seguint la línia de descentralització que s'ha imposat a la Federació

‘The assembly will take place at the plenary room..., following the line of decentralization imposed by the Federation’

The main difference between sentence adjuncts <ao> and textual elements <et> is that textual elements give no semantic content at all to the sentence. *En primer lloc*, for example (‘First, in first place’), only puts in order the discourse (<et>), while *en l'àmbit del finançament* (‘within the funding sphere’) gives some content to the sentence (<ao>).

3.18.1 The preposition *sobre*

The preposition *sobre* has two basic values: When it holds a locative meaning, it introduces an adjunct (<cc>); when it indicates the topic of the verb, then it introduces a sentence adjunct (<ao>):

1. *va donar un toc d'alerta sobre l'augment dels accidents* ‘(s/he) alerted about the increase of accidents’ (Cat 104nou)
2. *sobre la reforma laboral que vol dur a terme el Govern, ha afirmat que ...* ‘regarding the reform that the Government wants to carry on, (s/he) confirmed that...’ (Cat 104nou)

3.19 Clitic pronouns and syntactic functions

Some syntactic functions can be inferred thanks to the pronominalization by clitics. The clitic pronouns can be conflictive, specially *en*, *hi* and *ho*. This section is devoted to show the different values they can get⁵⁰:

⁵⁰All the examples here, as well as the list of functions each pronoun can hold, have been taken from GCC, pp. 1374-76.

3.19.1 The pronoun *EN*

(3.148) **Partitive subject**, corresponding to an indefinite noun phrase

Han dut els entrepans, però només en queda un
 ‘(they) brought the sandwiches, but there is only one left’

(3.149) **Partitive direct object**, corresponding to an indefinite noun phrase

Sabia tots els noms, però no en recordava cap
 ‘(s/he) knew all the names, but (s/he) did not remember any’

(3.150) **Prepositional object**, corresponding to a prepositional phrase headed by *de*

Era el seu aniversari, però tothom se’n va oblidar
 ‘It was (his/her) birthday, but everybody forgot it’

(3.151) **Adjunct (locative)**, corresponding to a prepositional phrase headed by *de*

Van marxar a València, i encara no n’han tornat
 ‘(They) went to Valencia and haven’t come back yet’

(3.152) **Attribute**

De feliç me’n sento molt
 ‘I feel very happy’ (lit. Happy, I feel very much)

(3.153) **Predicative complement**, corresponding to a noun or adjective phrase

A lluís el van nomenar president la setmana passada i a mi me’n nomenaran la setmana vinent
 ‘(They) appointed Lluís as president last week, and (they) will appointed me -as president- next week’

(3.154) **Inherent** (pronominal pronoun in pronominal verbs)

Se n’han anat de casa
 ‘(They) have left home’

3.19.2 The pronoun *HI*

(3.155) **Prepositional object**, corresponding to a prepositional phrase headed by *a*, *en* or *per*

1. *No m’hi puc acostumar, a aquesta ciutat*

‘I cannot get used to this city’

2. *Van anar a la fira del llibre però no hi van trobar el llibre que buscaven*

‘(They) went to the book fair but (they) did not find the book (they) were looking for’

(3.156) **Adjunct (locative)**, corresponding to a prepositional phrase headed by *a*, *en* or *per* or equivalent prepositional locutions

1. *Avui passarem per aquest pont, però demà ja no hi passarem*
‘Today (we) will go through this bridge, but tomorrow we won’t’
2. *Para taula i posa-hi les flors al damunt*
‘Set (you) the table and place the flowers there’

(3.157) **Adjunct** (comitative, instrumental, causal)

1. *Amb Joan no hi tornaré a viatjar mai més*
‘I will never travel with Joan again’
2. *Amb aquest mètode tothom hi aprèn molt*
‘Everybody learns a lot with this method’
3. *De tuberculosi encara hi mor gent*
‘There are still people dying from tuberculosis’

(3.158) **Attribute**, corresponding to any kind of phrase

No eren rics, però hi han esdevingut
‘(They) were not rich, but (they) have become (rich)’

(3.159) **Predicative complement**, corresponding to a noun or adjective phrase

Abans de l’operació mirava cansat, però ara ja no hi mira
‘Before the operation (s/he) looked tired, but now (s/he) doesn’t’

(3.160) **Adjunct** (manner), corresponding either to a prepositional or adverbial phrase

Els nens de Marta es van portar bé, però els de l’Albert no s’hi van portar gaire
‘Marta’s children behaved well, but Albert’s children did not behave that well’

(3.161) **Inherent** *Amb aquestes ulleres no hi veig gaire bé* ‘I cannot see very well with these glasses’**3.19.3 The pronoun HO**(3.162) **Direct object** (neutral), corresponding to a sentence or to the pronouns *això*

‘this’ and *allò* ‘that’
Sabia que hi hauria vaga però no ho volia dir
‘I knew there was going to be a strike but I didn’t want to say it’

(3.163) **Attribute**

No eren d’argent però ho semblaven ‘They were not made from silver but they seemed to’

(3.164) **Inherent**

S’ho va passar molt bé ‘(S/he) had a lot of fun’

3.20 Repeated functions

We have repeated functions when two elements hold the same syntactic function within the sentence. This happens especially when a clitic pronoun reproduces the content of a preverbal constituent, which is frequently found with direct (3.165) and indirect complements, but also with adjuncts and prepositional objects (3.166) since some of them also allow clitic pronominalization.

(3.165) (Cat 207-2)
 ...[*això*] [*ho*] *havia fet anteriorment*
 ‘...(s/he) this it had done before’

(3.166) (Cat 06010100-18) *Cal remarcar que [en aquesta nova manifestació] [hi] participaran organitzacions agràries*
 ‘It is worth emphasizing that in this new demonstration in it will participate agrarian organizations’

The attribute marking of this phenomenon must occur at least twice in the sentence, as many times as the syntactic function appears.

```
<attribute name="repetition" onlynonterminal="yes"
description="Is it repeated?">
<value name="yes" />
<value name="no" />
</attribute>
```