Deliverable 3.13

Documentation of the Italian Grammar

The DeepThought Consortium
October 2004
<table>
<thead>
<tr>
<th>Project ref. no.</th>
<th>IST-2001-37836</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project title</td>
<td>DeepThought - Hybrid Deep and Shallow Methods for Knowledge-Intensive Information Extraction</td>
</tr>
<tr>
<td>Deliverable status</td>
<td>Public</td>
</tr>
<tr>
<td>Contractual date of delivery</td>
<td>M24: October 2004</td>
</tr>
<tr>
<td>Actual date of delivery</td>
<td></td>
</tr>
<tr>
<td>Deliverable number</td>
<td>D3.13</td>
</tr>
<tr>
<td>Deliverable title</td>
<td>Documentation of Italian Grammar</td>
</tr>
<tr>
<td>Type</td>
<td>Lingware</td>
</tr>
<tr>
<td>Status &amp; version</td>
<td>Draft 0.3</td>
</tr>
<tr>
<td>Number of pages</td>
<td></td>
</tr>
<tr>
<td>WP contributing to the deliverable</td>
<td>WP3b</td>
</tr>
<tr>
<td>WP / Task responsible</td>
<td></td>
</tr>
<tr>
<td>Other contributors</td>
<td></td>
</tr>
<tr>
<td>Author(s)</td>
<td>Dario Gonella, Giampaolo Mazzini</td>
</tr>
<tr>
<td>EC Project Officer</td>
<td></td>
</tr>
<tr>
<td>Keywords</td>
<td>HPSG Grammar Matrix, Italian grammar, documentation, agreement, restructuring, raising and control, argument structure, passivation, cliticization, clitic climbing, absolute phrases, perception verbs</td>
</tr>
<tr>
<td>Abstract (for dissemination)</td>
<td>Description of the Italian grammar v. 1.0, based on the HPSG Grammar Matrix v. 0.6.</td>
</tr>
</tbody>
</table>
# INDEX

0. INTRODUCTION .......................................................................................................................... 5
  0.1. Pre-processing Protocol...................................................................................................... 5

1. STATUS OF THE ITALIAN LEXICON ..................................................................................... 7
  1.1 Lexical entries ...................................................................................................................... 7
    1.1.1 Verbs: ............................................................................................................................ 8
    1.1.2 Nouns ............................................................................................................................ 8
    1.1.3 Adjectives ...................................................................................................................... 9
    1.1.4 Adverbs ........................................................................................................................ 9
  1.1.5 An example of the inheritance mechanism ................................................................... 10

2. AGREEMENT ............................................................................................................................ 11
  2.1 PNG Agreement: Determiners and nouns ....................................................................... 11
  2.2 PNG Agreement: Verbs and nouns (subjects and objects) .............................................. 12
  2.3 PNG Agreement: Adjectives and nouns (modified) ........................................................... 13

3. SELECTION CRITERIA .............................................................................................................. 14
  3.1 Auxiliaire selection ............................................................................................................ 14
  3.2 CAT Selection & Predicative Structures ......................................................................... 15
    3.2.1 Notes ........................................................................................................................... 17
  3.3 Selection of prepositional complements .......................................................................... 17
  3.4 Selection of clausal complements (with complementizers) .............................................. 19

4. ARGUMENT STRUCTURE ....................................................................................................... 20
  4.1 Optionality ......................................................................................................................... 20
  4.2 Free Order (subj inversion, NP-PP and NP-AP inversion) ............................................... 20

5. PASSIVATION ........................................................................................................................... 22

6. RESTRUCTURING VERBS ....................................................................................................... 24

7. RAISING AND CONTROL VERBS ....................................................................................... 25

8. CLAUSAL COMPLEMENTS ..................................................................................................... 27

9. CLITICS .................................................................................................................................... 28
  9.1 Preliminary overview .......................................................................................................... 28
    9.1.1 Proclitics vs Enclitics ................................................................................................... 28
    9.1.2 Clitics “combining” ..................................................................................................... 28
    9.1.3 Clitics “compounding” ................................................................................................ 28
    9.1.4 Clitic climbing .............................................................................................................. 29
    9.1.5 Clitic doubling .............................................................................................................. 29
    9.1.6 Gender, number, person, case ..................................................................................... 29
    9.1.7 Argument vs Modifier role ......................................................................................... 31
  9.2 Lexical information in the Italian grammar ....................................................................... 31
    9.2.1 Information in “lexicon.tdl” and “italian-grammar” files ........................................ 31
9.3. A lexical approach to cliticization ................................................................. 33
9.4. The current treatment of cliticization in the Italian grammar .................. 35
9.5. The head-comp rules .................................................................................. 35
  9.5.1 A new comp-rule (“head-final”) ................................................................. 35
  9.5.2 Specific schemata 1 ................................................................................. 36
  9.5.3. Specific schemata 2 .............................................................................. 37
  9.5.4 Surface order ......................................................................................... 37
9.6. Interactions between auxiliary selection, agreement and cliticization ........ 38
  9.6.1 Examples: ............................................................................................. 38
  9.6.2 Final remarks on auxiliary selection ...................................................... 40
9.7. Notes .......................................................................................................... 41
10 SEMANTIC ROLES - SYNTACTIC FUNCTIONS LINKING ................................. 42
11 PERCEPTION VERBS IN ITALIAN.................................................................... 44
  11.1. PDS verbs complementation ................................................................... 44
  11.1.1. Infinitive complementation: cases $j$ - $k$ - $l$ ....................................... 46
  11.1.2. Finite complementation: cases $m$ - $n$ ................................................. 53
  11.1.3. Predicative structure complementation: case $o$ ................................... 55
  11.2. Implementing perception verbs in the italian grammar......................... 56
12 MODIFIERS AND ABSOLUTE PHRASES .................................................. 58
  12.1. Introduction ............................................................................................ 58
  12.2. The treatment of AAPs and APPs in the italian grammar ...................... 59
  12.3. The absolute phrase sentence position .................................................. 63
  12.4. Notes ..................................................................................................... 64
13 MODIFIERS: ADVERBIAL PHRASES ......................................................... 65
  13.1 The adv_type hierarchy ........................................................................... 65
  13.2 The “negation” adv_type ......................................................................... 66
14 MODIFIERS: RELATIVE CLAUSES .......................................................... 68
15 COORDINATION ....................................................................................... 69
0. INTRODUCTION

This document represents the Italian contribution to the Deliverable D3.13 *Description of the Grammars*. It is based on the results of the workpackage WP3a (and its Deliverable document D3.3 *Very basic grammar for Italian*, which has been largely reused mainly in chapters from 2 to 10) and on the following domain- and task-specific extensions to the grammar (mainly described in chapters 1, 12, 13, 14 and 15), according to the workpackage WP3b.

The Italian Grammar described in this document (v. 1.0, updated to 30/09/2004) is the one used in the *IE for Business Intelligence* application (see WP4a). The version of the Matrix used by the italian grammar is the Matrix 0.6.

Types and rules quoted in this report have been partially modified, in order to stress only their relevant elements, accordingly to the context (not relevant ones are simply omitted).

0.1. Pre-processing Protocol

From the very beginning of the Italian Grammar development phase, the integration of already existing pre-processing components for morphological analysis has been considered.

The component, that is actually a sub-component of CELI's Sophia2.1™ platform, is called MorphTool. The integration within the LKB grammar development platform is more in detail described in the document of deliverables D3.8 *Inclusion of morphology for Norwegian and Italian* and D3.9 *Pre-processing for all Grammars*. From March 2003, a protocol (SPPP, “Simple Pre-Processing Protocol”) developed by Stephan Oepen, interfaces the LKB system and the pre-processing module:

- the Italian grammar points to an external executable that communicates with the LKB through standard in- and output. As part of the grammar loading, that process is created and connected to a Lisp stream.
- `preprocess-sentence-string()` is extended to use the external engine if "sppp-stream" is non-nil.
- all communication is in XML, the LKB sends a string as an XML document and reads the preprocessor output as another XML document from the stream. Doing XML in both directions, if nothing else, has the advantage of declaring which encoding gets used (the default is UTF-8).
• .user-input. in parse() now is a list of morphologically analyzed and POS-tagged tokens; modifying add-morphs-to-morphs() for that to work was the only change in the parser proper.

An example of the Sophia-LKB interface is the following (for the sentence “il letto cigola”, “the bed is squeaking”):

→(input to Sophia)
<?xml version="1.0" encoding="utf-8"?>
<text>il letto cigola</text>

← (preprocessing result returned to the LKB)
<?xml version="1.0" encoding="utf-8"?>
<segment>
  <token form="il" from="3" path="1" to="5">
    <analysis cat="ARTICOLO" format="" inflection="MASC_SING"
      stem="det_art" tag="ARTICOLO"/>
  </token>
  <token form="letto" from="6" path="1" to="11">
    <analysis cat="NOME-MASCHILE" format="" inflection="MASC_SING"
      stem="letto_nm" tag="NOME-MASC-SING"/>
    <analysis cat="VERBO" format="" inflection="PARTIC_PASSATO_MASC_SING"
      stem="leggere_v" tag="VERBO-PARTPASS"/>
  </token>
  <token form="cigola" from="12" path="1" to="18">
    <analysis cat="VERBO" format="" inflection="INDIC_PRES_THIRD_SING"
      stem="cigolare_v" tag="VERBO-FIN"/>
  </token>
</segment>

Each token has one or more morphological analyses with information concerning mainly:
• stem (the string used for lexical look-up)
• inflection (the identifier of inflectional rule)

In the previous example, the stem value “cigolare_v” corresponds to (=is the name of) the lexical entry in the “verb_lexicon.tdl” file

\[
\text{cigolare_v} := \text{strict-intrans-verb-le} \& \text{[STEM.FIRST "cigolare_v"]}
\]

and in the “morph-lrules.tdl” file the inflectional information of a token is supplied referring to the inflection value “INDIC_PRES_THIRD_SING” by means of a correspondent inflectional type:

\[
\text{INDIC_PRES_THIRD_SING.irule} := \text{finite-verb-word} \& \text{[SYNSEM.LOCAL [ AGR.PNG 3sing, CONT.HOOK.INDEX.E [ TENSE pres, MOOD Indicative ] ] ]}.
\]
1. STATUS OF THE ITALIAN LEXICON

One of the most significant enhancement of the Italian grammar (in respect to the D3.3 deliverable) has been the growth of the lexicon, that currently consists of

- Number of lemmas: ca. 5,000
  - verbs: ca 1250
  - nouns: ca 3000
  - adverbs: ca 200
  - adjectives: ca 400
  - closed-class entries: ca 150

- Number of lexical types: 121
  - for verbs: 82
  - for adverbs: 31
  - for adjectives: 5
  - for nouns: 3

- Number of inflectional types: 88

The lexicon is represented/build up in some tdl files (somehow distinguished for practical reasons), herewith enclosed:

- verb_lexicon.tdl
- noun_lexicon.tdl
- adv_adj_lexicon.tdl
- closed_class_lexicon.tdl
- domain_noun_lexicon.tdl
- mwu_lexicon.tdl
- letypes.tdl
- categories.tdl

1.1 Lexical entries

Most of the lexical entries\(^1\) (lemmas) inherit from a lexical-entry-type ("_le" suffixed) assigning a value to STEM and PRED

\(^1\) Citics, pronouns, determiners (generally all the entries belonging to closed-classes) inherit directly from a corresponding lexeme or word subtype, e.g.:

\[
\begin{align*}
\text{miCliDat} & := \text{clt-dative-word} \\
\text{tuPron} & := \text{personal-pronoun-word}
\end{align*}
\]
lemma-name := type_le &
    [ STEM.FIRST "lemma_pos",
      SYNSEM.LKEYS.KEYREL.PRED "_orth_pos_sense_rel" ].

The STEM value is the name of the lemma as returned by the morphological pre-processing with Sophia (see above, the previous chapter about the pre-processing).

The PRED value conforms to the template

"_orth_pos_sense_rel"

as described in the document of the deliverable *D3.5 Documentation of Grammar Matrix* (chapter 3.2.2).

Some examples from the "_lexicon.tdl" files:

1.1.1 Verbs:
For the verbs, the ERG value is used mainly for the selection of the right auxiliaire: “avere letto” vs “essere appartenuto”:

leggere_v := trans-np-verb-le &
    [STEM.FIRST "leggere_v",
     SYNSEM.LKEYS.KEYREL.PRED "_leggere_v_rel",
     SYNSEM.LOCAL.CAT.HEAD.ERG - ].

appartenere_v := intrans-a-pnoun-verb-le &
    [ STEM.FIRST "appartenere_v",
      SYNSEM.LKEYS.KEYREL.PRED "_appartenere_v_rel",
      SYNSEM.LOCAL.CAT.HEAD.ERG + ].

1.1.2 Nouns
Nouns are currently encoded following a quite rough distinction between refential nouns without comps (which inherit directly from a lexeme type) and nouns with an ARG1:
1.1.3 Adjectives

All the adjectives inherit from specific lexical types, according their “category” (the capitalized letter) and their argumental frame:

\[
\text{successivo_adj} := \text{adj}_B\_\text{a\_pp\_le} & \\
\text{[ STEM.FIRST "successivo_adj",} \\
\text{SYNSEM.LKEYS.KEYREL.PRED "successivo_j\_rel" ]}. \\
\text{idoneo_adj} := \text{adj}_B\_\text{a\_inf\_or\_pp\_le} & \\
\text{[ STEM.FIRST "idoneo_adj",} \\
\text{SYNSEM.LKEYS.KEYREL.PRED "idoneo_j\_rel" ]}. \\
\]

1.1.4 Adverbs

Also the adverbs inherit from specific lexical types according their “category”:

\[
\text{bene_adv} := \text{pure\_modal\_adv\_le} & \\
\text{[ STEM.FIRST "bene_adv",} \\
\text{SYNSEM.LKEYS.KEYREL.PRED "bene_r\_rel" ]}. \\
\]

1.1 Lexical Types

Any lexical-type carries on information mainly about:

- the argument optionality
- the preposition/complementizer expected as governing an argument
- the “synsem type” corresponding to a given lexical item

following an “inheritance mechanism” as shown in the following examples:
1.1.5 An example of the inheritance mechanism

(“appartenere” [to belong]) appartenere_v := intrans-a-ppnoun-verb-le.

intrans-a-ppnoun-verb-le := intrans-a-select &
[ SYNSEM.LOCAL.CAT.VAL.COMPS < [OPT -] > ]].

intrans-a-select := verb-lxm &
[SYNSEM obl1-a-select-synsem].

obl1-a-select-synsem := intrans-ppnoun-synsem &
[LKEYS.--COMPKEY #ckey, LOCAL.CAT intrans-ppnoun-verb-cat &
[ VAL.COMPS < [LOCAL.CAT.HEAD.KEYS.KEY #ckey], ... > ]].

intrans-ppnoun-synsem := subj-arg1-verb & obl1-handle &
[LKEYS.--COMPKEY #ckey, LOCAL.CAT intrans-ppnoun-verb-cat &
[ LOCAL.CAT.HEAD.KEYS.KEY #ckey, ... ]].

intrans-verb-cat := verb-cat &
[HEAD.TRANS -.].

Some linking rules deal with the syntactic/semantic mapping: e.g.

subj-arg1-verb := lex-synsem &
[ LKEYS.KEYREL.ARG1 #1, LOCAL.CAT.QVAL.SUBJECT synsem &
[LOCAL.CONT.HOOK.INDEX #1] ].

obl1-handle := lex-synsem &
[ LKEYS.KEYREL.ARG2 #1, LOCAL [ CAT.QVAL.OBL1 synsem &
[ LOCAL.CONT.HOOK.INDEX #1] ]].

Nouns, adjectives and adverbs follow more or less the same mechanism.
2. AGREEMENT

In general, the agreement value is carried out by PNG (for referentials) and AGR (for event). According to a more general strategy, which uses multiple inheritance, agreement values are "calculated" by means of a type hierarchy. Values can be also "underspecified" at various degrees.

The pre-processing stream gives as a result an inflection value, pointing to a specific inflectional rules (in the “morph-lrules.tdl” file).

Any i-rule inherits from a word-type and contains the agreement information, carried by the PNG/TENSE/MOOD feature

\[
\text{MASC\_SING\_irule} := \text{word} & \\
\quad \left[ \text{SYNSEM} \#\text{synsem} & [\text{LOCAL}.\text{CONT}.\text{HOOK}.\text{INDEX}.\text{PNG} \text{ masc-sing-3rd}], \right. \\
\quad \left. \text{ARGS} < [\text{SYNSEM} \#\text{synsem}] > \right].
\]

\[
\text{INDIC\_PRES\_THIRD\_SING\_irule} := \text{finite-verb-word} \\
\quad [\text{SYNSEM}.\text{LOCAL} [\text{AGR}.\text{PNG} 3\text{sing}, \\
\quad \text{CONT}.\text{HOOK}.\text{INDEX}.\text{E} [\text{TENSE} \text{pres}, \text{MOOD} \text{indicative}]]].
\]

2.1 PNG Agreement: Determiners and nouns

The noun-lxm constrains the SPR (namely a determiner) and the head-daughter to share the PNG values via CONT.INDEX.PNG path :

\[
\text{noun\_lxm} := \text{infl\_lxm} & \\
\quad [\text{SYNSEM}.\text{LOCAL} [\text{CAT} [\text{HEAD} \text{noun}, \\
\quad \text{VAL} [\text{SPR} < \text{synsem} & \\
\quad \left. \left[ \text{LOCAL} [\text{CONT}.\text{HOOK}.\text{INDEX}.\text{PNG} \#\text{png}] > \right] \right], \\
\quad \text{CONT}.\text{HOOK}.\text{INDEX} [\text{PNG} \#\text{png}]]].
\]

Some subtypes of noun-lxm have been added, namely model-name-l xm (used for the model of cellular phones, e.g. “6600i”), brand-name-lxm (for the cellula phone brands, e.g. “Nokia”)
and entity-name-lxm (used for the specific parts or functions of cellular phones, e.g. “bluetooth” or “mms”). Model-names and brand-names have a PNG value underspecified as it regards the number.

2.2 PNG Agreement: Verbs and nouns (subjects and objects)

The finite-verb-word constrains the SUBJ and the head-daughter to share the PNG value via AGR.PNG path:

\[
\text{finite-verb-word} := \text{verb-word} \& \\
[ \text{SYNSEM.LOCAL} \ [\text{CAT} \ [\text{HEAD} \text{verb}, \\
\text{VAL} \ [\text{SUBJ} < [\text{LOCAL} \ [\text{CONT.HOOK.INDEX} \ [\text{PNG} #agr]] > ]], \\
\text{AGR.PNG} #agr] ].
\]

Note that a referential noun, with number=plural and being the head of a binary-headed-phrase, can have an n-bar projection with an empty SPR list, so that as an n-bar referential nouns without specifiers can play the role of the subject.

As it regards clitics, the PNG agreement is needed also for accusative pronouns in affixal position and the past participle in past compounded tenses, e.g.:

“Gli uomini hanno sempre amato le donne”
(Men have always loved women)

“Gli uomini le [femm] hanno sempre amate [femm]”
(Men them have always loved)

* “Gli uomini le [femm] hanno sempre amato [masc]”
* (Men them have always loved)

The PNG value of the clitic’s host-verb has to match with the (accusative) clitic’s one, and that’s carried out by the PSTP feature (see forward, the chapter about clitics).
2.3 PNG Agreement: Adjectives and nouns (modified)

The \textit{generic-adj-lx}m constrains the MOD (namely a noun) and the head-daughter to share the INDEX values via \textsc{CONT.HOOK.INDEX} path:

\[
generic-adj-lx := \text{infl-lx} \land \\
[ \text{SYNSEM.LOCAL} [ \text{CAT.HEAD adj} \land [\text{MOD} < \text{LOCAL}.\text{CONT.HOOK.INDEX \#index}] > ] }, \\
\text{CONT. HOOK.INDEX \#index } ]
\]
3. SELECTION CRITERIA

In this section are commented the criteria for selecting:

- the right auxiliaire type for a given verb
- the right category for the arguments of a given verb [included predicative structures] (categorial selection);
- the right preposition form for an argument of a given verb
- the right complementizer for a strongly selected clausal argument (syntactic strong argumental selection);

The mechanism of selection works at the LOCAL.CAT.VAL level (for verb arguments [prepositionals, complementizers, nominal, predicative]; see "letypes.tdl" and "categories.tdl").

3.1 Auxiliaire selection

As far as it concerns the auxiliaire selection in composing "aux+verb" verbal complex, at the VERB level two features are entailed, expressely AUX (distinguishing between auxiliaires and other verbs) and ERG (for distinguishing ergative verbs).

In italian language, while all passive forms selects the "essere" auxiliaire, in the active form we have to choice between two "tensal" auxiliaire: "essere" or "avere". We have three different cases:

i) all verbs with the ERG value "-" select the "avere" auxiliaire (e.g. "mordere" (bite), "lottare" (fight));
ii) all verbs with the ERG value "+" select the "essere" auxiliaire (e.g. "morire" (die) or reflexives forms like "muoversi" (move oneself);
iii) all verbs with the ERG value underspecified (bool) can select both the "essere" and the "avere" auxiliaire (e.g. "bruciare" (burn), "cuocere" (cook))

Consequently, is ungrammatical a sentence like:

* "Il cane E' morso il postino" (The dog IS bitten the postman)
Unfortunately, this assumption partially contrasts with the case of “SI” clitic pronoun (see below, in the clitics section). As a consequence, one cannot simply prevent a rule from assigning the ESSERE tensal auxiliaire to these kind of verbs, while a sentence like:

“il cane SI E’ morso” (The dog IS bitten HIMSELF)

is absolutely correct.

The same problem concerns the SI impersonal/subject clitic. A sentence like:

“Il cane è ucciso” (The dog is killed)

is clearly grammatical only in its passive acceptance, while a sentence like

* “Il cane E’ ucciso il postino” (The dog IS killed the postman)

should be rejected as ungrammatical. Nevertheless, a sentence like

“Il cane SI E’ ucciso” (The dog IS killed HIMSELF)

is once again correct. Therefore, some “aux+verb” phrases apparently ungrammatical are parsed, given the clitics treatment (see forward).

### 3.2 CAT Selection & Predicative Structures

Dealing with the valence of the verbs, the selection between PP and NP or COMPL-PHRASE is simply achieved by using the standard concept of "transitiveness", so that:

i) "transitive verbs" have a COMPS list containing at least one element (the "trans-verb-cat" in the "categories.tdl" plays this role) and subcategorize for both a noun phrase or a clausal complement, the right selection being operated (with a variable degree of (under)specification, see appendix II) by various subtypes of "strict-trans-synsem" (always in the "categories.tdl"), imposing given constraints on the "head" of the first complement, e.g.
trans-np-synsem := trans-np-or-complement-synsem &
[LOCAL.CAT.VAL.COMPS < [LOCAL.CAT.HEAD noun-or-pronoun] > ].

trans-np-or-complement-synsem := strict-trans-synsem &
[LOCAL.CAT.VAL.COMPS < [LOCAL.CAT.HEAD compl-or-noun] > ].

ii) "intransitive verbs" could have a list of COMPS
   o empty ("strict-intrans-verb-cat")
   o with a prepositional phrase ("intrans-ppverbnoun-verb-cat")
   o with an adverbial phrase ("itnrans-advcomp-verb-cat")

strci-intrans-verb-cat := intrans-verb-cat &
[ VAL.COMPS < > ].

intrans-ppverbnoun-verb-cat := intrans-verb-cat &
[ VAL.COMPS < [LOCAL.CAT.HEAD infincompl-or-prep] > ].

intrans-advcomp-verb-cat := intrans-ppall-verb-cat &
[ VAL.COMPS < [LOCAL.CAT.HEAD adv &
[ADV_TYPE pure] ] > ].

iii) raising verbs, restructuring verbs and auxiliaires are kept underspecified as far as
it concerns the TRANS value. The comps selection, in this case, is managed by
specific le-types or synsem-types.

Copulative verbs ("essere" [to be], "sembrare" [to seem]) and “ditrans-predicative” verbs
("considerare" [to consider]) subcategorize for a "predicative" argument (adjectival or
prepositional or else "locational adverbial" phrases and "undefined" noun or "defined and
modified" noun phrases). E.g.

copulative-verb-synsem := subj-arg1-verb &
[LOCAL.CAT [VAL.COMPS < [LOCAL.CAT.HEAD.PRED ST + ] ] ].

dittrans-predicative-synsem := ditrans-np-and-var-synsem &
[LOCAL.CAT [VAL.COMPS < synsem,
[LOCAL.CAT.HEAD.PRED ST + ] ] ].
In order to capture this kind of generalisation, the feature PRED-ST has been introduced at the HEAD level

\[
pred-st := head \& \ [ \text{PRED-ST bool} ].
\]

to assign the right (+) "predicative" value to adjectival and prepositional phrases:

\[
\begin{align*}
\text{adj} & := pred-st \& \ [ \text{PRED-ST +} ]. \\
\text{prep} & := pred-st \& \ [ \text{PRED-ST +} ]. \\
\end{align*}
\]

And the non-predicative value ( - ) e.g. to

\[
\begin{align*}
\text{adv} & := pred-st \& \ [ \text{PRED-ST -} ]. \\
\text{proper-noun} & := noun \& \ [ \text{PRED-ST -} ].
\end{align*}
\]

Being the PRED-ST at the HEAD level, his value is simply passed up given the Head Feature Principle.

### 3.2.1. Notes

The distinction between "undefined" or "defined and modified" noun phrases is not yet managed by the italian grammar, thus by now all common nouns are underspecified as it regard the PRED-ST value. The past participle currently has the PRED-ST value set to "-", since we don't have yet semantic constraints on verbs (and as far as we can't distinguish between "morire" (to die) and "partire" (to leave), we build up the same structure for "è morto" (he died) and "è partito" (he has left), but "è morto" could be rightly interpreted as "copulative verb + predicative argument", (e.g. "è vivo o morto?" (is alive or dead?) ) whereas it shouldn't be correct for a sentence like "è partito" ( *"è in casa o partito?" ( *is he at home or left?).

### 3.3 Selection of prepositional complements

In the Matrix version 0.6 the attributes KEY and ALTKEY can have as values subsorts of the type 'predsort' (the same kinds of values allowed for the attribute PREDSORT in semantic
relations). These attributes enable a word or phrase to be semantically selected by a predicate, and as head features they propagate up from the lexical head of the phrase. For example, a verb can select for a prepositional phrase headed by a particular preposition, as long as the preposition has lexically assigned a specific value (a subtype of predsort) to its SYNSEM.LOCAL.CAT.HEAD.KEYS.KEY attribute, and the verb similarly constrains the KEY value of its PP complement (accessed via the SYNSEM.LKEYS.--COMPKEY or --OCOMPKEY of the verb). For example, a verb like "appartenere" takes a PP object, governed by the "a" preposition: this verb inherits from the "intrans-a-select" which (via the "obl1-a-select-synsem") constrains the KEY value of the complement to be the "_a_p_rel" relation:

\[
a_{\text{prep}} := \text{adv-prep-word} \& \\
[\text{SYNSEM.LKEYS.KEYREL.PRED}_a_\text{p_rel}].
\]

\[
\text{appartenere}_v := \text{intrans-a-ppnoun-verb-le}.
\]

\[
\text{intrans-a-ppnoun-verb-le} := \text{intrans-a-select}.
\]

\[
\text{intrans-a-select} := \text{verb-lxm} \& \\
[\text{SYNSEM.LKEYS.--COMPKEY}_a_\text{p_rel}].
\]

The hierarchy of the predsort has been extended in order to deal with “supertypes” of specific relation regarding prepositions, used in the subcat of some verbs, e.g.

\[
\begin{array}{c}
\text{predsort} \\
\_\text{locmov}_p\_\text{rel} & \_\text{sociat}_p\_\text{rel} \\
\_\text{in}_p\_\text{rel} & \_a\_p\_\text{rel} & \_\text{con}_p\_\text{rel}
\end{array}
\]
3.4 Selection of clausal complements (with complementizers)

As for the case of preposition selection, also the right complementizer selection is pursued by means of synsem-types. E.g. a verb like "imparare" [to learn] take an NP as an argument ("imparare il latino" [to learn Latin] or an infinitive phrase with the preposition/complementizer "a" ("imparare a suonare la tromba" [to learn to play trumpet]. This verb inherits from the "trans-np-or-inchocomplement-synsem":

\[
\text{trans-np-or-inchocomplement-synsem} \equiv \text{strict-trans-synsem} \& \left[ \text{LOCAL.CAT.VAL.COMPS} < \left[ \text{LOCAL.CAT.HEAD inchocompl-or-noun} \right] > \right].
\]

"inchocompl-or-noun" is a supertype inherited both by nouns and the a-complementizer.
4. ARGUMENT STRUCTURE

4.1. Optionality

The feature OPT bool (introduced in the Matrix at the synsem-min level) is used by lexical entry types in the COMPS list, e.g.

\[(e.g. \text{“gustare” [to taste]})\]

\[\text{trans-np-verb-le} := \text{verb-lxm} \&\]
\[\{\text{SYNSEM.LOCAL.CAT.VAL.COMPS} < \{\text{OPT - } \} > \}.\]

\[(e.g. \text{“mangiare” [to eat]})\]

\[\text{trans-np*-verb-le} := \text{verb-lxm} \&\]
\[\{\text{SYNSEM.LOCAL.CAT.VAL.COMPS} < \{\text{OPT + } \} > \}.\]

The *star conventionally is used for marking up an optional complement in the lexical entry name. In the Matrix, the “basic-head-opt-comp-phrase” skips an optional complement (marked as OPT +)

The treatment of the optionality has been extended also to nominal, adjectival and adverbial arguments, by means of some specific schemata all inheriting (indirectly or directly) from the Matrix “basic-head-opt-comp-phrase”.

4.2 Free Order (subj inversion, NP-PP and NP-AP inversion)

For the time being, the Italian grammar doesn't take into account cases of topicalization, although they are quite frequent in the newsgroup messages domain.

However, some case of relative free order are taken into account, namely the subject inversion, the NP-PP and the NP-AP inversion in verbal argument structure. We adopt the strategy of using some lexical rules, which simply deal with the inversion of the elements in the COMPS list., e.g.:

\[\text{np-pp-inversion} := \text{d-word}\&\]
\[\{\text{SYNSEM.LOCAL.CAT.VAL.COMPS} < \{\text{#pp, #np }\},\]
\[\text{ARGS} < \{\text{SYNSEM.LOCAL.CAT.VAL.COMPS} < \{\text{#np }\& \{\text{LOCAL.CAT.HEAD noun}\},\]
\[\text{#pp }\& \{\text{LOCAL.CAT.HEAD prep}\}>\}>\}.\]
As it concerns the subject inversion, since we could always have (generalizing a little) a posthead subject, two rules inherits from the head-subj-phrase: the first inheriting from head-final (pre-head subject) and constraining the tense to be “finite”, the second inheriting from the head-initial (post-head subject); both the rules constrain the case of the NON-HEAD-DTR to be “nominative”:

\[
\text{head-pre-subj-phrase} := \text{head-subj-phrase} \; \& \; \text{head-final} \; \& \\
\quad [\text{NON-HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD.CASE nominative} \\
\quad \text{HEAD-DTR.SYNSEM.LOCAL [CONT.HOOK.INDEX.E.TENSE finite] }].
\]

\[
\text{head-post-subj-phrase} := \text{head-subj-phrase} \; \& \; \text{head-initial} \; \& \\
\quad [\text{NON-HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD.CASE nominative}].
\]
5. PASSIVATION

A lexical rule, *passive-lex-to-word*, according to the traditional approach (e.g. Sag & Wasow, “Syntactic Theory”), rearranges the elements of the COMPS list.

The rule applies to transitive & non-ergative verbs [TRANS +, ERG -]. Further information is concerning the semantic role ARG1 (coindexed with the ARG2 of the preposition "da" (by).

```
passive-lex-to-world := d-word &
SYNSEM [ LOCAL [ CAT #cat,
  SUBJ < #subj & [LOCAL.CONT.HOOK.INDEX #index >,
  COMPS < [OPT +,
  LOCAL.CONT.HOOK.INDEX #index,
  LOCAL.CAT.HEAD prep & [PFORM da -form],
  LOCAL.KEYS.KEY.ARG2 #1] . #rest > ]],
  AGR #png,
  CONT.HOOK.INDEX.E.DIATH passive,
  KEYS.KEY.ARG1 #1 ],
  NON-LOCAL #nonlocal ],
ARGS < pastpart-verb &
  SYNSEM.LOCAL.CAT
  [HEAD.ERG -, HEAD.TRANS +,
  VAL.COMPS < #subj . #rest >,
  VAL.SUBJ < [LOCAL.CONT.HOOK.INDEX #index >],
  SYNSEM.NON-LOCAL #nonlocal ]).
```

In correspondence with the *passive-lex-to-world* another lexical rule is dealing with the past participle, namely the *active-lex-to-world*:

```
active-lex-to-word := d-word &
SYNSEM [ LOCAL [ CAT #cat,
  AGR #png,
  CONT #cont & [HOOK.INDEX.E.DIATH active],
  NON-LOCAL #nonlocal ],
ARGS < pastpart-verb &
  SYNSEM.LOCAL #nonlocal & [ CAT #cat,
  AGR #png,
  CONT #cont ]).
```

All past participles have two alternative interpretations, one as "passive", the other as "active" past participle. Since in our grammar, in order to avoid the overgeneration, all participle verbs have the feature REAL - (whereas all non-participle verbs have the value REAL +) such a "duplication" of lexical rules concerning the participles seems to be needed:
past participle (e.g. "morso" [bitten])

}`pastpart-verb`

```
(REAL -)
```

|-- "passive-lex-to-word" --

```
(AUXTYPE pass-aux) |
```

|-- "è morso dal cane"

```
[d-word] (REAL +)---|
```

|-- "ha morso il cane"

|-- "active-pastpart-lex-to-word"

```
(AUXTYPE nonpass-aux)
```

The "d-word" projects a verb-headed structure as a word type, since lexical rules don't work on lexemes. STEM, HEAD, CONT, KEYS, AGR and all NON-LOCAL values are simply passed up, while INFLECTED and REAL become “+”.

```
  d-word := word & [STEM #stem,
            INFLECTED +,
            REAL +,
            SYNSEM.LOCAL [ CAT [ HEAD verb & #1 ],
                        CONT #cont,
                        KEYS #keys,
                        AGR #agr ],
            SYNSEM.NON-LOCAL #nonloc
            ARGS < [STEM #stem,
                    SYNSEM.LOCAL [ CAT [ HEAD verb & #1 ],
                                  CONT #cont,
                                  KEYS #keys,
                                  AGR #agr ],
                    SYNSEM.NON-LOCAL #nonloc ] > ].
```
6. RESTRUCTURING VERBS

Partially according to the most argumented approaches to the Italian restructuring verbs (e.g. Monachesi 1999, following Rizzi, 1982), both auxiliares and restructuring verbs (modals, temporal aspectuals and pure motion verbs) are subjected to the argument composition and give rise to a verbal complex (see also interactions with clitics in the correspondent chapter). The feature \( \text{RV +} \) (\( \text{Restructuring Verb} \)) (introduced at the head level) marks all restructuring verbs which inherit, at the lexical level, from the \( rv-lxm \) where the argument-composition mechanism applies:

\[
rv-lxm := srv-lxm \& \\
[ \text{SYNSEM} \ \top \text{LOCAL.CAT} [ \\
\quad \text{RV +}, \\
\quad \text{VAL} [ \text{COMPS} < [ \text{LOCAL.CAT} \ \top \text{VAL.COMPS} \ \#\text{comps}]. \ \#\text{comps} ] ] ].
\]
7. RAISING AND CONTROL VERBS

As far as it concerns subject raising structures, in the Italian grammar ALL auxiliaries, modals, motionals and copulatives are considered raising verbs (with structure sharing between the subject of the governed infinitive and the subject of the governor); ALL OTHERS (governing at least one more complement besides the infinitive one) are treated as equi verbs (by means of the coindexation, as described e.g. in Pollard & Sag, 1994). Informally:

**Raising verbs**

**Auxiliaries**: essere [be], avere [have] = srv-lxm [AUX +, RV -]

**Copulatives**: sembrare [seem], essere [be] = srv-lxm [AUX -, RV -]

**Modals**: volere [want], potere [can], dovere [must] = srv-lxm [AUX -, RV +]

**Motionals**: andare [go] = srv-lxm [AUX -, RV +]

\[
\text{srv-lxm} := \text{base-verb-lxm} \& \left[ \text{SYNSEM.LOCAL.CAT.VAL} \left[ \text{SUBJ} #1, \text{COMPS} < \left[ \text{LOCAL.CAT.VAL.SUBJ} #1 \right] > \right] \right].
\]

**Equi (control) verbs**

**Subj-control**: promettere [promise] = subj-to-subj-binding

**Obj-control**: mandare [send], avvisare [inform] = dirobj-to-subj-binding

**Iobj-control**: consigliare [advise], insegnare [teach] = iobj-to-subj-binding

\[
\text{subj-to-subj-binding} := \text{dirobj-arg2} \& \left[ \text{LOCAL.KEYS.KEY.ARG1} #1, \text{LOCAL.CAT.QVAL.DIROBJ.LOCAL.CONT.HOOK.XARG} #1 \right].
\]

\[
\text{iobj-to-subj-binding} := \text{indirobj-arg3} \& \text{dirobj-arg2} \& \left[ \text{LOCAL.KEYS.KEY.ARG3} #1, \text{LOCAL.CAT.QVAL.DIROBJ.LOCAL.CONT.HOOK.XARG} #1 \right].
\]

\[
\text{dirobj-to-subj-binding} := \text{indirobj-arg3} \& \left[ \text{LOCAL.KEYS.KEY.ARG2} #1, \text{LOCAL.CAT.QVAL.INDIROBJ.LOCAL.CONT.HOOK.XARG} #1 \right].
\]
Raising verbs inherit from the “srv-lxm”; control verbs inherit from the correspondant "binding" rule operating the coindexation between the semantic role in KEYS.KEY and the XARG of the controlled argument.

Actually, control verbs mostly seem to be precisely distinct, e.g. “ordinare” (to order) is an indirect-object-control verb (iobj-ctrl) and “promettere” (to promise) is unambiguously a subject-control-verb (subj-ctrl). Some other verbs (mainly a subclass of performative verbs: “chiedere” (to ask for), “urlare” (to shout) and so on) seem to allow for both the interpretation:

\[
\text{Giovanni ha chiesto a Paolo di usare il nuovo cellulare}
\]

Giovanni asked Paolo to use the new mobile-phone

could be ambiguous (either Paolo has been asked to use the phone [iobj-ctrl], or Giovanni asked Paolo permission to use the phone [subj-ctrl]).

Nevertheless, in order to avoid overgeneration, in the lexicon all these verbs have been assigned one ctrl value according their prevailing use.
8. CLAUSAL COMPLEMENTS

The grammar deals mainly with three kinds of clausal complement:

i) subordinate finite clauses, introduced by the "che" complementizer;
ii) subordinate infinitive clauses, introduced by the "di" complementizer;
iii) subordinate infinitive clauses, introduced by the "a" complementizer

Above the clausal complements introduced by these complementizers, the current grammar deals also with interrogative clauses playing an argumental role, e.g. in sentences like

"molti amici mi chiedono se acquistare un nuovo cellulare"

(many friends ask me whether to buy a new cellular)

No constraint on tense or mood is managed yet.
9. CLITICS

9.1. Preliminary overview

9.1.1. Proclitics vs Enclitics

The Italian clitics can appear both in "proclitical" (i) and in "enclitical" (ii) position:

i) "Giovanni LO dà a Maria"
   (John gives IT to Mary)

ii) "DalLO a Maria!"
   (Give IT to Mary!)

In both cases, clitics exhibit verbal affix behavior (the host is always a verb, clitics have a rigid ordering, don't have wide scope over a coordination of hosts, nothing can intervene neither between clitics nor between clitic(s) and host verb; see Monachesi 1999).

9.1.2 Clitics “combining”

Italian clitics can be "combined" (iii) up to 3 elements:

iii) "TE LO SI dirà domani"
    (THIS (lo) IT (si) will be said TO-YOU (te) tomorrow)

9.1.3 Clitics “compounding”

In some cases, Italian clitics can be "lexically" merged in one lexical unit (iv):

iv) "Giovanni GLIELO dà"
    (John gives IT (-lo) TO-HIM (glie-))
9.1.4. Clitic climbing

Italian clitics can occur in clitic climbing (a clitic - originally a dependent of a complement verb - can "climb" and attach to the trigger verb) (v):

v) "Giovanni LO vuole dare a Maria"
   (John wants to give IT to Mary)

9.1.5 Clitic doubling

Italian language allows (especially in non-literary context) for "clitic-doubling": a complement of a given verb can have contemporary both the cliticized and the normal realization (vi):

vi) "Giovanni GLIELO dà a Luca
   (John gives IT (-lo) to Luca (TO-HIM (glie-)))

Frequently the clitic-doubling originates from a topicalization of a complement (e.g. "Giovanni il libro [topic.] LO dà a Maria" (John [the book] gives IT to Mary)), some kind of topicalizations being almost "obligatory" also in written texts, especially in question-answer context (vii):

vii) [q] "LO hai letto il libro di Giovanni?"
    (IT Did you read the John's book?)
    [a] "Non LO ho letto il libro di Giovanni ma so che è noioso"
    (IT I didn't read the John's book, but I know that it is boring)

9.1.6 Gender, number, person, case

Italian clitics are distinguished as for gender (masc/femm), number (sing/plur), person (1st, 2nd, 3rd), case (in accusative, dative, reflexive, partitive/thematic and locative clitics), as in the following table:
The clitic "SI" is also used as a pseudo-reflexive (in a lexical realization of "inherent" cliticized complement) as in (viii):

viii) "Giovanni SI sveglia e tu TI addormenti"
     (John HIMSELF(inherent) wakes-up and you YOURSELF(inherent) go to sleep)

The clitic "SI" can also play the subject role in impersonal/generalized sentences (ix/x):

ix) "SI dice che Giovanni sia giovane"
     (ONE(subj) says that John is young)

x) "SI è arrivati a casa"
     (ALL-OF-US(subj) arrived home)

The clitic "SI" has also, in some cases, the role of a diathetical (passive) marker (xi):

xi) "SI sono dette molte bugie"
     (Many lies ("SI" passivizing) have been said)

In combining each others, clitics can have some phonological variation, e.g.:

- the datives "mi/ti/si/ci/vi" become "me/te/se/ce/ve" ["Giovanni MI dà il libro - Giovanni ME LO dà" (John gives TO-ME the book - John gives TO-ME IT)],
- the locatives "ci/vi" become "ce/ve" ["Giovanni CI mette il libro - Giovanni CE LO mette" (John puts the book ON-IT - John puts ON-IT IT)],
9.1.7. Argument vs Modifier role

Italian clitics can be argument/modifier of the host verb, as in (xii) or argument/modifier of a complement of the host verb, as in (xiii)

xii) "TI ucciderò"
(I will kill YOU)

xiii) "TI ucciderò il cane"
(I will kill YOUR dog)

As for argument role, clitics are alternative to the normal realization (up to now our grammar doesn't allow for "clitic-doubling", which is anyway forbidden with the "SI"(subj) and the reflexive clitics).

As for the modifier role, it seems that can be cliticized only PPs introduced by a small set of prepositions ("di" (of); locative "a"(to), "su"(on) and "in"(in), "da"(from))

9.2. Lexical information in the Italian grammar

9.2.1. Information in "lexicon.tdl" and “italian-grammar” files

At the lexical level, clitics carry on information about PNG and CASE (inheriting from specific subtypes of clt-word), e.g.:

\[ lo\_cli := clt\text{-}accusative\text{-}word \& \]
\[ \text{[ STEM.FIRST} "lo\_cli", SYNSEM.LOCAL.CONT.INDEX.PNG masc\text{-}sing ] \]

\[ mi\_cli\_acc := clt\text{-}accusative\text{-}png\text{1st\text{-}word} \& \]
\[ \text{[ STEM.FIRST} "mi\_cli", SYNSEM.LOCAL.CONT.INDEX.PNG 1sing ] \]
Clitics inherit form "clt-word" and his subtypes, e.g.

\[
\text{clt-word} := \text{word} & \left[\text{WORD -}, \right.
\text{SYNSEM.LOCAL} \left[ \text{CAT.HEAD prep-or-noun } & [\text{CLT +, } \text{MOD } \langle \rangle], \right.
\text{CAT.VAL} \left[ \text{SPR } \langle \rangle, \text{COMPS } \langle \rangle \right] \right].
\]

\[
\text{clt-accusative-word} := \text{clt-word} & \left[\text{SYNSEM.LOCAL.CAT.HEAD noun } & \left[\text{CASE accusative }\right]\right].
\]

\[
\text{clt-accusative-png1st-word} := \text{clt-accusative-word} & \left[\text{HOST} \left[ \text{AGR.PNG not1st} \right]\right].
\]

Notes:

a) the feature "HOST avm", introduced at the sign level in the Matrix, contains all useful information about the "host" verb of a clitic: LOCAL is simply copied in the HOST via the "head-comp-phrase"; constraints (e.g. on PNG) are specified by different types of clt-word. The HOST feature allows to have a good level of information at the lexical level.

b) clitics are "WORD -" (clitics, like determiners, prepositions, complementizers and conjunctions, cannot directly satisfy any element in a COMPS list)

c) clitics are CLT + (in order to distinguish clitics from other "prep-or-noun" types)

d) clitics have empty COMPS, SPR, MOD lists

f) "si_cli_pass" has the value "CASE nocase" (being the "passivizing" SI just a diathetical marker)

g) "ne_cli_varg" has the value "CASE v_argument" (a generic "verbal argument"), as opposite to the "NE" modifier

h) orthophonetical variants are explicitely distinguished in the lexicon

i) "compounded" clitics are recognized by the tokenizer and the input of the lexicon is constituted by two distinguished entries (e.g. "GLIELO" --> "GLI", "LO")

---

2 The grammar doesn't deal with some puzzling constructions that we are looking suspiciously at, like e.g. "Io non amo tutte le donne; LE amo CON I CAPELLI ROSSI" (I don't love every woman; I like THEM WITH RED HAIRS) or "I cellulari LI compro NUOVI" (The mobiles I buy THEM NEW); ...
9.3. A lexical approach to cliticization

The most convincing approach to the Italian cliticization (including the clitic climbing) is the one suggested by Paola Monachesi in several papers. A CLT-COMPS list contains the cliticized complements of a given verb; a lexical rule applies to the "normal" verb and consequently the cliticized complements are removed from the COMPS list and added as members of the CLT-COMPS list. For example, verbs like "amare" (to love) and "dare" (to give) would have two possible realizations:

a1) **not cliticized** ("Maria ama il marito", [Mary loves the husband])

   "amare" = [ COMPS < #1 [np] >, CLT-COMPS < > ] →

→ a2) **cliticized** ("Maria LO ama", [Mary loves him])

   "amare" = [ COMPS < >, CLT-COMPS < #1 [np] > ]

b1) **not cliticized** ("Maria dà il libro a Giovanni", [Mary gives the book to John])

   "dare" = [ COMPS < #1 [np], #2 [pp] >, CLT-COMPS < > ] →

→ b2) **cliticized** ("Maria LO dà a Giovanni", [Mary gives IT to John])

   "dare" = [ COMPS < #1 [np] >, CLT-COMPS < #1 [np] > ]

→ b3) **cliticized** ("Maria GLI dà il libro", [Mary gives the book TO-HIM])

   "dare" = [ COMPS < [np] >, CLT-COMPS < #2 [pp] > ]

→ b4) **cliticized** ("Maria GLIELO dà", [Mary gives IT TO-HIM])

   "dare" = [ COMPS < >, CLT-COMPS < #1 [np], #2 [pp] > ]

(Note that cases from b2) to b4) should be managed by specific lexical rules, in the lack of a "shuffle" operator).

In order to avoid too many alternative edges in the parsing, it could be possible to use (instead of several lexical rules) different schemata for combining accusative/reflexives and dative/locative/partitive clitics with the host verb: those schemata should play the role of removing the cliticized complement from the COMPS list and marking in some way the verb as a "cliticized" form).

As for clitic climbing, triggered in Italian by auxiliary, modal, aspectual and motion verbs
("restructuring verbs"), specific lexical rules should deal with the "argument composition" of restructuring verbs: the trigger verb inherits the complements of the embedded verb, including the ones which could be realized as clitics. As an effect of such a lexical rule, a verb like "volere" (to want) should have a COMPS list as in the following:

\[
\text{"volere" } = [\text{COMPS < [vp(infin), COMPS #1] >, CLT-COMPS < > } ] \rightarrow \rightarrow \text{"volere" } = [\text{COMPS < [vp(infin), COMPS #1] . #1 >, CLT-COMPS < > } ]
\]

and then, applying a lexical rule for cliticization,

\[
\rightarrow \text{"volere" } = [\text{COMPS < [vp(infin), COMPS #1] >, CLT-COMPS #1 } ]
\]

The combined action of the "cliticizing" lexical rules and the "argument composition" lexical rules seems to be adequate in several cases, but it seems to be inefficient in the case of "multiple restructuring". Being the auxiliary verb a restructuring one, a "double restructuring" is quite frequent, e.g.:

xiv) "GLI ha voluto dare il libro"
(TO-HIM he wanted to give the book"

but also a "triple" restructuring is not so rare, especially with motion restructuring verbs, e.g.:

xv) "LO ha voluto andare a comprare"
(IT he wanted to go to buy)

In those case, the cliticized complement could be pushed deeply into the bottom of the COMPS list ([COMPS.REST.REST.REST.FIRST] ) and the matching between the clitic and the correspondent complement couldn’t be so easy to perform, even if one would impose an arbitrary/pragmatic limit to the "restructuring" chain.
9.4. The current treatment of cliticization in the Italian grammar

We have tried to overcome the problems connected with the multiple restructuring by adopting an hybrid approach to cliticization and clitic climbing: we saved the argument composition mechanism for restructuring verbs, but we delay the attachment of the clitic until the (possible) restructuring chain has been completed.

The argument composition is performed at the lexeme level for the restructuring verbs (RV +). They inherit from rv-lxm (subtype of srv-lxm for subject raising verbs):

```
rv-lxm := srv-lxm &
[SYNSEM.LOCAL.CAT [ RV +,
   VAL.COMPS < [LOCAL.CAT.VAL.COMPS #comps]. #comps > ] ].
```

Auxiliaires are a subtype of rv-lxm (AUX +):

```
aux-lxm := rv-lxm &
[SYNSEM.LOCAL.CAT [HEAD.PSTP #png,
   AUX +,
   VAL.COMPS < [LOCAL.AGR.PNG #png], ... > ] ].
```

In the aux-lxm, the feature PSTP has the same value of the complement (the past-participle). PSTP and PNG unify (in the head-comp-rule) and produce as unification result some specific subtypes (e.g. masc-sing-3rd, femm-plur-3rd) which are useful for the agreement with composed tenses (e.g. “LA ho amatA [femm-sing]”(I loved HER), “LI ho amatI [masc-plur] (I loved THEM)).

9.5. The head-comp rules

The clitic attachment is performed by a small set of "comp-rules".

9.5.1 A new comp-rule (“head-final”)

Beside the previous "head-comp-rule" (expanding the head-initial), a new comp-rule has been introduced, inheriting from "head-final" (in order to deal with proclitics):

...
head-comp-phrase ::= base-head-comp-phrase & head-initial &

pre-head-comp-phrase ::= base-head-comp-phrase & head-final &
[HEAD-DTR [ SYNSEM.LOCAL.CAT.HEAD verb],
NON-HEAD-DTR [ SYNSEM.LOCAL.CAT [ HEAD.CLT +, POSTHEAD + ] ] ].

The latter rule applies (as for now) only to cliticized complements (the pre-head-comp rule
DOES NOT manage the left-dislocated full complements).

The CLT-COMPS bool (in the following rules) is intended to be a marker for cliticized verbs
(marked as CLT-COMPS +).

9.5.2 Specific schemata 1

A specific schema is devoted to manage the "SI" (passivizing)

base-clt-passive ::= head-final &
[INFLECTED +,
SYNSEM.LOCAL.CAT [ VAL [ SUBJ < >,
COMPS < >,
CLT-COMPS + ] ],
NON-HEAD-DTR [SYNSEM [LOCAL.CAT.HEAD noun & [CASE nocase, CLT + ] ] ],
HEAD-DTR [SYNSEM.LOCAL [ CAT [ VAL [ COMPS < >,
CLT-COMPS -,
SUBJ < > ],
QVAL.DIROBJ [LOCAL.CONT.INDEX.PNG #png ] ],
AGR.PNG #png,

The "SI" (passivizing) can occur only when
a) the complements in the COMPS list of the host verb have been realized (COMPS <>),
b) the SUBJ hasn't been realized (SUBJ <[->)
c) the passivizing "SI" can't be combined with other clitics ([HEAD-DTR.CLT-COMPS -] )
d) the host verb must be finite
The mother node of the schema has an empty SUBJ list.
9.5.3. Specific schemata 2

A specific schema is devoted to manage the impersonal "SI"(subj)

\[
\text{base-clt-nominative} := \text{head-final} \& \text{head-valence-phrase} \& \\
\quad \text{[INFLECTED +,} \\
\quad \text{SYNSEM.LOCAL [ CAT [ VAL [ SUBJ < >,} \\
\quad \text{COMPS #comps,} \\
\quad \text{CLT-COMPS + ]],} \\
\quad \text{NON-HEAD-DTR.SYNSEM #1} \& \text{LOCAL [ CAT.HEAD noun \& [CASE nominative,} \\
\quad \text{CLT + ],} \\
\quad \text{CONT_INDEX.PNG #png ]}, \\
\quad \text{HEAD-DTR.SYNSEM.LOCAL [ CAT [ VAL [ SUBJ < #1 >,} \\
\quad \text{COMPS #comps & olist ]],} \\
\quad \text{AGR.PNG #png & 3sing,} \\
\quad \text{CONT_INDEX [ E.MOOD finite, E.DIATH active ]].}
\]

The impersonal "SI" plays the role of a subject of the host verb. Nevertheless, the "base-clt-nominative" cannot inherit form the head-subj-phrase in the Matrix because an empty COMPS list cannot be a correct general constraint in this case. Actually, when other clitics are present, the impersonal "SI" is the latter in the clitics chain ("e.g.: "ME LO SI dice spesso" (TO-ME IT ONE says often)), so that when the SI attaches to the host verb, its COMPS list is really not empty.

The COMPS list of the Head-dtr is constrained to be "olist", so that optional not realized complements don't block the rule application (e.g. "SI ama" (ONE loves)).

The base-clt-nominative removes the element from the SUBJ list, coindexed with the clitic.

The host verb has to be singular, finite, active.

9.5.4 Surface order

As for enclitics, the surface order is inverted in respect of the proclitical situation, so that a lexical rule "np-pp-inversion" (a general lexical rule dealing with the free argument order for ditransitive verbs) is needed (and enough) to deal with such a "clitic-chain order" inversion.
9.6. Interactions between auxiliary selection, agreement and cliticization

9.6.1 Examples:

a) reflexives (accusative and dative)

(i) “Giovanni ha ucciso il gatto”
   (John killed the cat)

(ii) “Giovanni SI è ucciso”
    (John HIMSELF killed)

(iii) “Ho regalato un libro a Giovanni”
    (I gave-as-a-present a book to John)

(iv) “MI sono regalato un libro”
    (TO-MYSELF I gave-as-a-present a book)

(v) “GLI ho regalato un libro”
    (TO-HIM I gave-as-a-present a book)

b) “SI” (impersonal / subject)

(vi) “Gli amici sono arrivati”
    (Friends have arrived)

(vii) “SI è arrivati”
    (ALL-OF-US have arrived)

(viii) “Giovanni ha discusso di politica”
    (John has talked about policy)

(ix) “SI è discusso di politica”
    (IT has been talked about policy)
c) inherent/pseudo-reflexive

(x) “Mi sono innamorato”
(I felt in love)

(xi) “Mi sono voluto innamorare”
(I wanted to fall in love)

(xii) “Ho voluto innamorarMI”
(I wanted to fall in love)
9.6.2 Final remarks on auxiliary selection

Cliticization may influence the auxiliary selection. More precisely (as shown in the previous examples, sentences from (i) to (ix)), verbs which normally select, in the active form of composed tenses, the auxiliary “AVERE”, modify their behaviour in presence of a reflexive or impersonal clitic, so that their cliticized version select the auxiliary “ESSERE”. A different case involves the “inherent/pseudo-reflexive” clitic. It really gives rise to a different entry of the verb: such a “lexically cliticized” verb is always intransitive and selects always the auxiliary “ESSERE”. But (once again) some modification can occur: for instance, as shown in sentence (xii), when the inherent clitic get the enclitical position, the selected auxiliary is “AVERE”.

Given the fact that in the current approach clitics are attached once the verbal complex is completed, we are enforced to build up some redundant structures (and to leave in some cases the auxiliary selection being underspecified), waiting for a possible clitic attachment. For instance, we need to construct verbal complex as the following:

(a) “è ucciso” [DIATH active]
(b) “è regalato” [DIATH active]
(c) “è discusso” [DIATH active]
(d) “è partiti” [DIATH active, PNG masc-plur]

In the last case (d), a special type “PNG collective-3rd” has been introduced, which unifies only with an impersonal (“SI”) subject.

In former cases (a) to (c), transitive and ditransitive verbs are concerned.
9.7. Notes

Some linguistic phenomena related to cliticization are not covered by the grammar; mainly the cliticization of argument of nouns (e.g. “Il presidente NE ha apprezzato la decisione” [the president of-him appreciated the decision])

Given the highly informal character of the adopted domain (newsgroups about mobile phones), the grammar should be able to deal with some “ungrammatical” frequent phenomena: the most frequent in the corpus are related to the cliticization phenomena, namely the so-called “clitic doubling” (e.g. “il nokia gli piace molto a mio padre”) and the “ethical dative” (“la suoneria non mi funziona”). Dative clitic pronouns are therefore allowed to be verb modifiers, without restrictions, so that a sentence like “Il marito LE ha comprato un regalo” could be parsed even if with a wrong modifier attachment.
10 SEMANTIC ROLES - SYNTACTIC FUNCTIONS LINKING

In the italian-grammar (according to the strategy described and adopted by the Norsk group) we have introduced some functional categories (SUBJECT, DIROBJ, INDIROBJ, OBL1...). The functional categories are attributes of a type qval which is the value of a feature QVAL in qval-cat:

\[ qval-cat := \text{cat} & [\text{QVAL qval}] \].

The linking is done

i) in \textit{synsem types} between semantic arguments and indices of functional categories

ii) in \textit{cat types} between functional categories and syntactic arguments

E.g.:

i) \textbf{subj-arg1-verb} := \text{lex-synsem} &
   \[ \text{LOCAL} [\text{CAT.QVAL.SUBJECT synsem} &
   \text{LOCAL.CONT.HOOK.INDEX #1},
   \text{KEYS.KEY.ARG1 #1}] \].

ii) \textbf{subj-verb-cat} := \text{verb-cat} &
   \[ \text{VAL.SUBJ < #1 >,}
   \text{QVAL.SUBJECT #1} \].

The functional category OBL1 is used for oblique PPs like "a mio fratello" [to my brother] in "il libro appartiene a mio fratello" [the book belongs to my brother]. For verbs like "appartenere" (intransitive verbs whit PP complement) in italian-grammar we use the subj-arg1-verb above and following linking types:

\textbf{obl1-handle} := \text{lex-synsem} &
\[ \text{LOCAL} [\text{CAT.QVAL.OBL1 synsem} &
\text{LOCAL.CONT.HOOK.LTOP #1},
\text{KEYS.KEY.LBL #1}] \].

\textbf{intrans-ppnoun-verb-cat} := \text{intrans-verb-cat} &
\[ \text{VAL.COMPS < #1 >,}
\text{QVAL.intrans-qval & [OBL1 #1]} \].

\textbf{intrans-verb-cat} := \text{subj-verb-cat} & [\text{HEAD.TRANS -}].
As an example, the composed synsem for the verb "appartenere" [belong] contains the following information (as far as they concern the linking):

\[
\text{lex-synsem} := \text{subj-arg1-verb} \& \text{obl1-handle} \&
[ \text{LOCAL} [ \text{CAT intrans-ppnoun-verb-cat} \&
[ \text{HEAD \ verb,}
\text{VAL [ \text{SUBJ < #1 >,}
\text{COMPS < #2 > },}
\text{QVAL [ \text{SUBJECT #1 \& [ \text{LOCAL.CONT.HOOK.INDEX #3 },}
\text{OBL1 #2 \& [ \text{LOCAL.CONT.HOOK.LTOP #4 } ] ]}],}
\text{KEYS.KEY arg1-rel} \&
[ \text{LBL #4, PRED "appartenere_rel", ARG1 #3 } ] ]].
\]

As for infinitive clauses, see upwards (chapter 6)
11 PERCEPTION VERBS IN ITALIAN

Some verbs, which imply a direct sensorial perception (PDS, *Percezione Diretta Sensoriale*)

\[
\text{perceiver} \rightarrow \text{perception} \rightarrow \text{entity + process/state complements} \\
\text{subject} \quad PDS \text{ verb} \quad \text{complements}
\]

e.g. "vedere" [see] and "sentire" [hear], share some peculiar behaviours.
Ignoring by now the semantics of perception, PDS verbs from a syntactic point of view are
distinguished for some connected aspects of complementation and cliticization.

11.1. PDS verbs complementation

Perception, roughly speaking, can be subdivided in pure *entities* perception (e.g. “to see
something”) and *entities + process/state* perception.
Limiting us to the latter case, possible subcategorization examples are the following:

1. S vede mangiare G  (S sees eating G)
2. S vede G mangiare (S sees G eating)
3. S lo vede mangiare (S sees him [lo] eating)
4. S vede G mangiare T (S sees G eating T)
5. S lo vede mangiare T (S sees him [lo] eating T)
6. S gli vede mangiare T (S sees him [gli] eating T)
7. S glielo vede mangiare (S sees him [glielo] eating it [lo])
8. S vede mangiare G da Z (S sees G eaten by Z)
9. S lo vede mangiare da Z (S sees him [lo] eaten by Z)
10. S vede G partire (S sees G leaving)
11. S lo vede partire (S sees him [lo] leaving)
12. S vede G che mangia T (S sees G that eats T)
13. S vede che G mangia T (S sees that G eats T)
14. S lo vede che mangia T (S sees him [lo] that eats T)
15. S vede G stanco (S sees G tired)
16. S vede G inseguito dai fans (S sees G followed by fans)
17. S vede G nei guai (S sees G in troubles)
18. S lo vede stanco (S sees him [lo] tired)
19. S lo vede inseguito dai fans (S sees him [lo] followed by fans)
20. S lo vede nei guai (S sees him [lo] in troubles)
In this examples, $S$ is the subject (perceiver) and $G$ the entity (perceived) of the PDS verb “vedere” [to see] (enough representative to our goals).

A first analysis of the examples produces a draft synthesis of possible PDS verbs complementation scheme:

![Diagram of PDS verb complementation scheme]

3. The term "deactivation" (DEACT) indicates a “deep” passivation of an infinite verb without a correspondent morphological realization. E.g. the sentence

$G_{iovanni} \text{ ha visto } uccidere$ (John saw him to-kill)

can have a double interpretation, depending on the diathesis of the infinitive “uccidere” (kill). When other complements are expressed, the sentence is disambiguated: e.g. in the sentence

$G_{iovanni} \text{ ha visto } uccidere \text{ sua moglie}$ (John saw him kill (= killing) his wife)

the only possible interpretation of the infinitive verb is the “active” one and the sentence could be paraphrased as “Giovanni ha visto "lo" mentre uccideva la moglie” (John saw him while he was killing his wife); in the sentence

$G_{iovanni} \text{ ha visto } uccidere \text{ da sua moglie}$ (John saw him kill (= being killed) by his wife)

the verb “uccidere” is "deactivated" (not lexicalized passive) and the sentence could be paraphrased as "Giovanni ha visto "lo" essere ucciso dalla moglie” (John saw him being killed…). An analogous infinitive “deactivation” case could be found in causative verbs complementation (perception and causative verbs seem to share a lot of features), e.g.: “Giovanni lo ha fatto uccidere dalla moglie” (John made him to-be-killed by his wife).

4. By using the PRCPT acronym here I mean the “perceived” as a whole, that is a perceived “entity” and the “state/process” introduced by a predicative phrase. In the In the previous examples, “G stanco”, “G inseguito dai fans”, “G nei guai” are all PRCPT instances.
11.1.1. Infinitive complementation: cases ① - ② - ③

The most complex case is the one regarding the infinitive complementation. There is no ambiguity in role assignment when the verb is INTRANSITIVE: in the sentence “Vedo Giovanni arrivare” (I see John coming) Giovanni is undoubtfully the subject of the infinitive verb, and this is true also when an inversion occurs, e.g. “Vedo arrivare Giovanni” (I see coming John). Cliticization is allowed (“Lo vedo arrivare” (I see him coming)) and the PDS verb is behaved as an “obj_control” verb, e.g. “incaricare”, “to charge” (see also chapter 6):

“Maria incarica Giovanni(i) di (i) andare dal presidente”
(Mary charges John(i) to (i)go to the president)

“Maria lo(i) incarica di (i) andare dal presidente”
(Mary charges him(i) to (i) go to the president)

The lexical entries of the PDS verb ③ should be (informally) the following

\[ pds\_verb\_3 = (\text{COMPS} < \text{NP} \ , \ \text{VPinf} [\text{TRANS -}, \ \text{SUBJ} \ ] \ ) \]

and it could inherit from the ditrans-dirobj-ctrl-synsem type.

When the VPinf is TRANSITIVE, the situation is a little more complex and the “deactivation” (see note 4) phenomenon plays a significant role.

In sentence 1 the role assignment is ambiguous:

S vede mangiare G

G could be the inverted subject OR the direct object of the PDS verb “mangiare” (eat). If G is (interpreted as) the subject, no direct object can be expressed:

* S vede mangiare G la bistecca (* S sees eating G a beefsteak)
* S vede mangiare la bistecca G (* S sees eating a beefsteak G)

An agent (a PP introduced by “da” [by] preposition) is allowed, playing the role of a deep subject (ARG1), as in the sentence 9:

S vede mangiare G da Z
In this case it seems reasonable to assume that the verb “mangiare” is “deactivated” and exhibits a passive verb valence:

\[
\text{mangiare (act)} = (\text{SUBJ 1, COMPS < NP 2 >}) \rightarrow
\]
\[
\text{mangiare (deact)} = (\text{SUBJ 2, COMPS < PP_da 1 >})
\]

The surface subject of the deactivated verb is allowed, e.g. in the sentence 9:

\[S \ lo \ vede \ mangiare \ da \ Z\]

and also a doubled cliticization is allowed, as in the sentence 7:

\[S \ glielo \ vede \ mangiare\]

In the latter case, the sentence could be paraphrased as

\[S \ lo \ vede \ essere\_mangiato \ gli \ (S \ see \ him \ [lo] \ to\_be\_eaten \ by\_him \ [gli])\]

Has to be emphasized that only with PDS verbs (and deactivated infinitive complement) the agent could be cliticized, e.g.:

\[S \ vede \ mangiare \ una \ bistecca \ da \ G \ (i) \ (S \ sees \ eat \ a \ beefsteak \ by \ G)\]
\[S \ gli \ (i) \ vede \ mangiare \ una \ bistecca \ (S \ sees \ by\_him \ [gli] \ eat \ a \ beefsteak)\]

As a first conclusion, one could say that a PDS verb governing a deactivated VPinf

- Is a \textbf{monotransitive verb} ("deactivated" infinitive verb \textit{can never be preceded by an NP});
- Is possible the cliticization (and clitic climbing) of the surface subject, the agent and all other complements (for ditransitive and alternatively in respect of the agent) of the governed infinitive, e.g.

\[S \ glielo \ vede \ dare \ a \ Maria\]
\[(S \ sees \ it \ [_lo] \ to\_be\_given \ by\_him \ [glie_] \ to \ Mary)\]
The **PDS verb** has an *argument composition*: it inherits, in its COMPS list, both the SUBJ and the COMPS of the governed verb.

In its turn, a **PDS verb governing a not-deactivated (active) VPinf**

- Is a **ditransitive verb** (the infinitive verb can be preceded by an NP)
- As for an object-control verb, the SUBJ list of infinitive verb has to be not empty; in the active acceptation of the sentence

\[ S \text{ vede mangiare Giovanni} \]

*Giovanni* is NOT the expressed subject of the infinitive verb, but the NP (inverted) governed and controlled by the PDS verb:

\[
\text{Vedere} = \text{COMPS} (\text{NP} \, \text{VPinf})
\]

- Only the NO governed by the PDS verb can be **cliticized**
- There is no argument composition
We have to postulate two different lexical entries\textsuperscript{5} for the PDS\textsuperscript{1} (act) and PDS\textsuperscript{2} (deact) verbs:

\begin{align*}
\text{pds}_\text{verb}_\text{1} &= (\text{COMPS} < \text{NP} [], \text{VPinf} [\text{DEACT} - , \text{SUBJ} [] ] >) \\
\text{pds}_\text{verb}_\text{2} &= (\text{COMPS} < \text{VPinf} [\text{DEACT} + , \text{SUBJ} [], \text{COMPS} []] , [], [] >)
\end{align*}

Analizing the structure of the sample sentences (1 to 9), possible conclusions and generalization are the following:

I) sentences \textbf{without cliticization}:

a) \textbf{sentence 1} “S vede mangiare G” has two possible (isomorphic) interpretations:

i) \textbf{not deactivated} infinitive verb (\textbf{ditransitive} PDS verb)

\begin{tikzpicture}
    \node {vede} child {node {mangiare} child {node {G}}};
\end{tikzpicture}

[G is the infin Subj = “S vede che G mangia” (S sees that G eats)]

ii) \textbf{deactivated} infinitive verb (PDS verb \textbf{monotransitive}, with argument composition)

\begin{tikzpicture}
    \node {vede} child {node {mangiare} child {node {G}}};
\end{tikzpicture}

[G is the infin surface Subj = “S vede che G è mangiato” (S sees that G is eaten)]

\textsuperscript{5} Two lexical entries (one for a control structure, the other one “flat” with argument composition) are suggested also in Abeillé, Godard, Miller & Sag, \textit{French bounded dependencies} in “Romance in HPSG” (edit. Baltari & Dini, 1998, pp. 35 ss.)
b) **sentence 2** “S vede G mangiare” and **sentence 4** “S vede G mangiare T: just one interpretation is possible: G is the (deep and surface) subject of “mangiare”:

i) **not deactivated** infinitive verb

```
    vede  G  mangiare
```

```
    vede  G  mangiare  T
```

c) **sentence 8** “S vede mangiare G da Z”: only one interpretation is possible:

i) **deactivated** infinitive verb

```
    vede  mangiare  G  da Z
```
II) Cliticized sentences:

d) sentence 3 “S lo vede mangiare”: two possible (isomorphic) interpretations:

i) not deactivated infinitive verb

```
lo  vede  mangiare
```

[“lo” is the subject = “S vede che “lo” mangia (q.cosa)”]

ii) deactivated infinitive verb

```
lo  vede  mangiare
```

[“lo” is the surface subject = “S vede che “lo” è mangiato (da q.uno)”]

e) sentence 5 “S lo vede mangiare T”: one interpretation

i) not deactivated infinitive verb

```
lo  vede  mangiare  T
```

[“lo” is the subject = “S vede che “lo” mangia T”]
f) **sentence 6** “S gli vede mangiare T” one interpretation

   i) **deactivated** infinitive verb

   
   ![Diagram]

   [“lo” is the surface subject = “S vede che T è mangiato da ‘gli’”]

---

g) **sentence 7** “S gielo vede mangiare”: one interpretation:

   i) **deactivated** infinitive verb

   
   ![Diagram]

   [“lo” is the surface subject and “gli” the deep subject = “S vede che “lo” è mangiato da “gli””]

---

h) **sentence 9** “S lo vede mangiare da Z”: one interpretation:

   i) **deactivated** infinitive verb

   
   ![Diagram]

   [“lo” is the surface subject = “S vede che “lo” è mangiato da Z”]
11.1.2. Finite complementation: cases

In the sample sentence 13

\[ S \text{ vede che G mangia T} \]

the PDS verb governs a finite clause (introduced by “che” (that)). As for this kind of complementation, the perception could be either a **direct preception** of a process/state, as in the example, or an **indirect perception** or **cognition**, as in the following sentences:

- “Giovanni ha visto che Maria ha lavato i piatti”  
  [John saw that Mary has washed the plates]
- “Giovanni ha visto che Maria vuole bene al suo gatto”  
  [John saw that Mary loves her cat]
- “Giovanni ha visto che sarebbe meglio andare via”  
  [John saw that it should be better to go away]

In such cases the finite complementation is the only one admitted, so that is not possible to “translate” the sentences in an infinitival form, e.g.:

- *Giovanni ha visto Maria aver lavato i piatti*  
  [*John saw Mary to have washed the plates*]
- *Giovanni ha visto Maria voler bene al suo gatto*  
  [*John saw Mary to love her cat*]
- *Giovanni ha visto essere meglio andare via*  
  [*John saw to be better to go away*]

The perception verb could be paraphrased as a cognition or opinion verb (e.g. “notare”, “constatare” or “immaginare”) and the the perceived state/process could be not contemporary (that is a quite strong constraint for PDS verbs): it’s possible to “perceive” a “concluded event”, e.g. ”Vedo che hai lavato i piatti“ (I see that you have washed the plates). It’s also possible to “perceive” a “negated event”, e.g. ”Ho visto che Maria NON ha lavato i
piatti” (I saw that Mary didn’t wash the plates), that is in general not possible for the other cases of PDS verbs complementation (“Ho visto non arrivare Giovanni” (“I saw John not coming), but there are some counter-examples).

As a rough conclusion, one could say that a PDS verb governing a finite complement clause is actually a “cognitive/suppositive” verb. Therefore the lexical entry for the PDS verb \( \text{pds}_4 \) could be the following:

\[
\text{pds}_4 = (\text{COMPS} < \text{Sfin} [\text{compl} “che”], \text{SUBJ} <>, \text{CLT} - >)
\]

with an empty SUBJ list, “che” (that) complementizer and no cliticization allowed.

Quite different seems to be the complementation in the sentence 12:

\[
\text{S vede G che mangia T}
\]

In this case the process/event perceived has to be contemporary (“\( \text{S ha visto G che mangia T} \) and it’s possible to “translate” the sentence in an infinitival form (“\( \text{S vede G mangiare T} \)’); the perceived cannot be negated (“\( \text{S vede G che non mangia T} \)”).

A possible interpretation is that the “che” introduces a kind of relative clause, as for analogous construction in French (Miller & Lowrey, *La complémentation des verbes de perception en anglais et en français*).\(^6\)

On the other hand, if we consider the “che” as a normal complementizer introducing a complement finite clause, the sentence 12 could be a surface variant of an infinitival complementation (see the verb PDS\(^3\)). “\( \text{S vede G che mangia T} \)” and “\( \text{S vede G mangiare T} \)” should be equivalent and the lexical entry of the PDS verb \( \text{pds}_5 \) could be the following one:

\[
\text{pds}_5 = (\text{COMPS} < \text{NP} [], \text{Sfin} [\text{compl} “che”, \text{SUBJ} []]>)
\]

\(^6\) Such a relative phrase should have a lot of peculiar behaviours: the relative pronoun cannot be substituted by “il quale”; the introduced verb has to be contemporary to the PDS verb (A sentence like “\( \text{S vede G che ha mangiato T} \)” has a different meaning); the NP could be cliticized (\( \text{S lo vede che mangia T} \)).
with a not-empty SUBJ list in the complement clause. As for all the others complement finite clause, cliticization is allowed:

\[
\text{S} \quad \text{lo} \quad \text{vede} \quad \text{che mangia q.cosa}
\]

[“lo” is the subject = “S vede che “lo” mangia q.cosa”]

11.1.3. Predicative structure complementation: case ©

The sample sentences (15 to 20) seem to share the same complementation structure, that involves a generic PRCPT (“perceived”) constituted by an NP (the entity) and a predicative phrase (state or process). The lexical entry of the PDS verb © is the following:

\[
pds\_verb\_6 = (\text{COMPS} < \text{NP} \square, \text{phrase} [\text{PRED} -\text{ST} +, \text{SUBJ} \square])
\]

the feature PRED-ST “+” is used to mark the syntactic structures with a predicative function. 
PRCPTs are adjectival phrases, e.g.

\[
\text{S vede G stanco} \quad \text{S} \quad \text{lo} \quad \text{vede stanco}
\]

participial phrases, e.g.

\[
\text{S vede G inseguito dai fans} \quad \text{S} \quad \text{lo} \quad \text{vede inseguito dai fans}
\]

and prepositional phrases, e.g.

\[
\text{S vede G nei guai} \quad \text{S} \quad \text{lo} \quad \text{vede nei guai}
\]
PDS verb \( \odot \) is a **ditransitive** verb inheriting, in the italian grammar, from the *ditrans-predicative-synsem* type.

### 11.2. Implementing perception verbs in the italian grammar

All the PDS verbs have been grouped in 4 verb types:

- **a)** PDS **control** verbs (\( \odot, \odot, \odot, \odot \))
- **b)** PDS verbs with **argument composition** (\( \odot \))
- **c)** PDS **monotransitive** verbs (governing a “che” finite clause)(\( \odot \))
- **d)** PDS **predicative** verbs (\( \odot \))

As it concerns cases c) and d), lexical entries inherit from existing lexical types, namely:

\[
\begin{align*}
pds\_verb\_3 & := \text{trans-fin-complement-verb-le}. \\
pds\_verb\_4 & := \text{ditrans-predicative-verb-le}.
\end{align*}
\]

As for cases a) and b), two new lexical types have been introduced:

\[
\begin{align*}
(pds\_verb)_1 & := \text{perceptive-ctrl-verb-le}. \\
(pds\_verb)_2 & := \text{perceptive-ac-verb-le}.
\end{align*}
\]

**perceptive-ctrl-verb-le** := verb-lxm &  
[SYNSEM.LOCAL [CAT [  
RV +,  
VAL.COMPS \(< \#1 \& [LOCAL.CAT.HEAD noun],\) [ OPT -,  
LOCAL.CAT.HEAD base-or-fincompl \& [DEACT -],  
LOCAL.CAT.VAL.SUBJ \(< \#1 \& [LOCAL.CAT.H EAD prep], ... > ] ] ] ] ]

**perceptive-ac-verb-le** := verb-lxm &  
[SYNSEM.LOCAL [CAT [  
RV +,  
VAL.COMPS \(< \text{synsem} \& [OPT -,  
LOCAL.CAT [ HEAD verb \& [DEACT +],  
VAL.SUBJ \(< \#1 \& [LOCAL.CAT.HEAD noun] >,  
VAL.COMPS \#comps \& < \text{synsem} \& [LOCAL.CAT.HEAD prep], ... > ] ]  
< \#1 . \#comps > ] ] ] ]]
A specific “inversion rule” is used for perception verbs:

\[
\text{perceptive-verb-inversion} := \text{d-word} \& \\
[\text{SYNSEM subj-arg1-verb \& dirobj-arg2 \&} \\
\quad \text{LOCAL.CAT subj-verb-cat} \& \\
\quad \text{VAL.SUBJ #subj,} \\
\quad \text{VAL.COMPS < [LOCAL #2, OPT -],} \\
\quad \text{[LOCAL #1, OPT -] > }], \\
\text{SYNSEM.NON-LOCAL #nonlocal,} \\
\text{ARGS < verb-word \&} \\
\quad \text{SYNSEM.LOCAL.CAT.VAL [} \\
\quad \text{SUBJ #subj,} \\
\quad \text{COMPS < [LOCAL #1 \& [CAT.HEAD noun \& [CLT -]]],} \\
\quad [LOCAL #2 \& [CAT.HEAD verb \& [DEACT -]],} \\
\quad \text{CAT.VAL.CLT-COMPS -] >,} \\
\quad \text{SYNSEM.NON-LOCAL #nonlocal | >}.]
\]

A new feature \text{DEACT bool} (introduced in pred-st) is used for the “deactivation” phenomenon: the value “+” is assigned by a lexical rule for infinitive verbs:

\[
\text{passive-infin-lex-to-word} := \text{base-passive -lex-to-word} \& \\
[\text{ARGS < infinitive-verb \& [SYNSEM.LOCAL.CAT.HEADE.DEACT +] > }].
\]
12 MODIFIERS AND ABSOLUTE PHRASES

12.1. Introduction

Scopal and intersective modifiers are managed by a small set of construction rules, inheriting respectively from the Matrix types “head-adj-int-phrase”, “head-adj-scop-phrase”, “adj-head-int-phrase” and “adj-head-int-phrase”. The grammar allows for underconstrained structures (“spurious ambiguities”) as far as it concerns PP modifiers: for instance, in a sentence like "Il cane morde il gatto per gioco" [The dog bites the cat in fun] two structures are allowed, in which "per gioco" [in fun] is attached as a modifier of “gatto” and “orde il gatto”.

According to their MOD value, the grammar can deal with modifiers like the following ones:

- adjectives
- adverbs
- relative clauses
- prepositional phrases
- subordinate clauses (governed by a subordinative-conjunction)
- absolute phrases
  - participial
  - gerundives

Absolute phrases (both participial and gerundive phrases) in Italian usually occur in a sentence initial position and play the role of (mainly) temporal modifiers.

For examples:

1. Partita Maria, Giovanni pianse
2. Partita, Maria pianse
3. Baciato il marito, Maria è felice
4. Baciato dagli amici, Giovanni è commosso
5. Ucciso Cesare, Roma cadde nella confusione
6. Avendo distrutto la moto, Giovanni va a piedi
7. Partendo per le vacanze, Maria sorrideva
A first distinction can be operated between AAP (*Absolute Aspectual Phrase*) and APP (*Absolute Predicative Phrase*):  

- **AAPs**: gerundive phrases (for all types of verb) and participials with an “active” diath (for *transitive verbs* with a *direct object* or *intransitive verbs* (with or without subject) [see sentences 1, 2, 3, 5, 6, 7]

- **APPs**: participials with a “passive” diath (for *transitive verbs* neither preceded nor followed by an NP (the surface subject cannot be expressed) [see sentence 4]

The main difference between AAPs and APPs (not taking in account semantic or temporal and aspectual problems) is caused by the syntactic role: AAPs are *modifiers* of the governing phrase (matrix phrase); APPs are “predicatives” of the matrix phrase subject, in a kind of subj-control structure.

### 12.2. The treatment of AAPs and APPs in the Italian grammar

A possible approach could be to generate AAPs and APPS by using two distinct lexical rules, which could define the matrix phrase structure and deal with the indexation matter. In the Italian grammar there are already two lexical rules for past participles (one for the active diathesis, the other for the passive one), so that it seems more efficient to adopt a kind of “mixed” approach, making use of lexical constraints and specific schemata.

---

7 See Luca Dini, *Aspectual Constraints on Italian Absolute Phrases*, 1995
8 Sentence 5 has an ambiguous interpretation. While in the other sentences the absolute phrase subject is easily detected, in the sentence “Ucciso Cesare, Roma cadde nella confusione” the name “Roma” could or could not be the subject of “ucciso”. Anyway, the past participle “ucciso” has an active interpretation.
a) Lexical constraints

- **gerundive phrases**: a gerundive verb expands the type “gerund-verb-word”, (verb-word subtype). The head of the MOD list is constrained to be a verb and its SUBJ list has to be empty:

  \[
  \text{gerund-verb-word} := \text{verb-word} & \\
  & \text{[SYNSEM.LOCAL.CAT.HEAD} \\
  & \text{[MOD < [LOCAL.CAT.HEAD verb,} \\
  & \text{LOCAL.CAT.VAL.SUBJ < >] > ]].}
  \]

- **participial phrases**: a participle verb is the input of three lexical rules (two for the active diathesis, the third for the passive one) variously constraining the MOD list element:

  \[
  \text{active-lex-to-word} := d-word & \\
  \text{[SYNSEM [ LOCAL [ CAT [HEAD base-ppv & [} \\
  & \text{MOD < [LOCAL.CAT.HEAD verb,} \\
  & \text{LOCAL.CAT.VAL.SUBJ < >] > ],} \\
  & \text{CONT #cont & [HOOK.INDEX.E.DIATH active],} \\
  & \text{NON-LOCAL #nonlocal },} \\
  & \text{ARGS < pastpart-verb & [SYNSEM.LOCAL [CONT #cont,} \\
  & \text{SYNSEM.NON-LOCAL #nonlocal ] > ].}
  \]

  \[
  \text{tr-active-lex-to-word} := \text{active-lex-to-word} & \\
  \text{[SYNSEM.LOCAL.CAT.VAL.SUBJ < >,} \\
  \text{ARGS < [SYNSEM.LOCAL.CAT.HEAD} \text{[TRANS +,}} \\
  \text{TRANSACT + ] > ].}
  \]

  \[
  \text{intr-active-lex-to-word} := \text{active-lex-to-word} & \\
  \text{[ARGS < [SYNSEM.LOCAL.CAT.HEAD} \text{[TRANS -,}} \\
  \text{TRANSACT - ] > ].}
  \]

  \[
  \text{passive-lex-to-word} := d-word & \\
  \text{[SYNSEM[LOCAL][CAT][HEAD base-ppv & [} \\
  & \text{TRANS +, TRANSACT -,} \\
  & \text{MOD < [LOCAL.CAT.HEAD verb-or-noun,} \\
  & \text{LOCAL.CAT.VAL.SPR <>},} \\
  & \text{LOCAL.CAT.VAL.SUBJ < >] > ],} \\
  & \text{VAL [SUBJ < [LOCAL.CAT.POSTHEAD -] >,} \\
  & \text{COMPS < [LOCAL.CAT.VAL.COMPS < >} \\
  & \text{. #rest >]},} \\
  & \text{CONT.HOOK.INDEX.E.DIATH passive ],} \\
  & \text{NON-LOCAL #nonlocal ],} \\
  & \text{ARGS < pastpart-verb & [SYNSEM.LOCAL.CAT} \text{[HEAD.ERG -,}} \\
  & \text{HEAD.TRANS +,} \\
  & \text{VAL.COMPS < [ ] . #rest >,} \\
  & \text{VAL.SUBJ < [ ] > ]].}
  \]

More precisely:

- **Transitive verbs** are the input of the two *l-rules* `tr-active-lex-to-word` and `passive-lex-to-word`: as a result, two verb-words are generated with different values of `TRANSACT (+ or -)` and `DIATH (active or passive).
  - Active transitive participles `[TRANSACT +]` have an empty SUBJ list as well as their matrix phrase (MOD).
  - Passive transitive participles `[TRANSACT -]`, besides the modifications in the COMPS list caused by the passivation, can have only a POSTHEAD – subject and their matrix phrase needs an empty SUBJ list.

- **Intransitive verbs** are the input of the `intr-active-lex-to-word` rule, obtaining the values “TRANSACT –“ and “DIATH active”; the SUBJ list of their matrix phrase has to be empty.

**b) Specific schemata**

Participial absolute phrases are treated by two schemata (construction rules):

i) **Transitive active** verbs (TRANSACT +) (direct object always expressed, subject NOT expressed). This rule doesn’t treat sentences (highly stylistically marked) as “Ucciso Cesare dai nemici, Roma piombò nel caos” (Killed Caesar by the enemies, Rome falled into the chaos)

```
pre-head-sent-phrasal-modifier-phrase4 := pre-head-sent-modifier-phrase &
[ARGS < [SYNSEM.LOCAL.CAT.HEAD verb & [TRANSACT +],
  SYNSEM.LOCAL.CAT.VAL.SUBJ < [ ] >,
  SYNSEM.LOCAL.CAT.VAL.COMPS < >,
  SYNSEM.LOCAL.CONT.HOOK.INDEX.E.MOOD pastpart],
  [ ] > ].
```

ii) **Intransitive or Passive verbs** (TRANSACT -) (the subject has to be expressed in the matrix phrase)

```
pre-head-sent-phrasal-modifier-phrase5 := pre-head-sent-modifier-phrase &
[ARGS < [SYNSEM.LOCAL.CAT.HEAD verb & [TRANSACT -],
  CONT.HOOK.INDEX.E.MOOD pastpart],
  SYNSEM.LOCAL.CAT.VAL.SUBJ < > > ].
```
Gerundive absolute phrases have partially different constraints in respect of participial phrases. For example, in an active transitive gerundive phrase the direct object could be unexpressed:

Sbagliando, Giovanni ha imparato molte cose

And that’s impossible in participial phrases:

* sbagliato, Giovanni pianse

Sbagliata strada, Giovanni pianse

In a passive gerundive phrase, the surface subject can be expressed (in opposite to the participial phrase constraint):

Essendo stato ucciso Cesare dai congiurati, Roma piombò nel caos

* ?? Ucciso Cesare dai congiurati, Roma piombò nel caos

Pugnalato dai congiurati, Cesare morì

Three construction rules are needed for gerundive absolute phrases:

iii) Transitive active verbs (subject unexpressed)
(Note: the rule doesn’t deal with stylistically marked sentences with expressed subj, as e.g. “Pronunciando il Presidente una sola parola, la situazione si è normalizzata”)

\[
\text{pre-head-sent-phrasal-modifier-phrase}_1 := \text{pre-head-sent-modifier-phrase} & \left[\text{ARGS} < [\text{SYNSEM.LOCAL [CAT [HEAD verb & [TRANS +], \text{VAL.SUBJ} < ] ] \}, \text{CONT.HOOK.INDEX.E [ DIATH active, \text{MOOD gerund}]}, [\text{SYNSEM.LOCAL.CAT.VAL.SUBJ < >] > }] \right].
\]

iv) Passive verbs (with an optional subject)

\[
\text{pre-head-sent-phrasal-modifier-phrase}_2 := \text{pre-head-sent-modifier-phrase} & \left[\text{ARGS} < [\text{SYNSEM.LOCAL [CAT [HEAD verb & [TRANS +], AUX +], \text{CONT.HOOK.INDEX.E [ DIATH passive, \text{MOOD gerund}]}, [\text{SYNSEM.LOCAL.CAT.VAL.SUBJ < >] > }] \right].
\]
v) Intransitive verbs (no constraints)
(in this case the subject could be expressed: a sentence like “Essendo arrivato Giovanni, Maria partì” is absolutely grammatical and has no stylistic markers)

\[
\text{pre-head-sent-phrasal-modifier-phrase} := \text{pre-head-sent-modifier-phrase} \land \\
\text{ARGS} < [\text{SYNSEM.LOCAL} [ \text{CAT.HEAD verb} \land \text{TRANS -}], \\
\text{CONT.HOOK.INDEX.E.MOOD gerund}], \\
[ ] > .
\]

All the rules above mentioned inherit from the following more general schema:

\[
\text{pre-head-sent-modifier-phrase} := \text{head-mod-phrase-simple} \land \text{head-final} \land \\
\text{SYNSEM} [ \text{MODIFIED lmod}, \\
\text{LOCAL.CAT.VAL} [ \text{SUBJ} < >, \text{COMPS} < > ]], \\
\text{ARGS} < [\text{REAL} +], \\
\text{SYNSEM} [ \text{LOCAL.CAT} [ \text{VAL} [ \text{COMPS olist}], \\
\text{HEAD verb} \land [ \text{MOD null} ]], \text{LOCAL.CONT.HOOK.INDEX.E.MOOD finite}], \\
\text{REAL} + >, \\
\text{C-CONT.RELS} < ! ! >, \\
\text{C-CONT.HCONS} < ! ! > .
\]

12.3. The absolute phrase sentence position

Introducing the absolute phrases, has been wrote that “in italian usually occur in a sentence initial position”. Actually, they can occur as well in a sentence final position (with a different semantic value). Playing with the sample sentences, we can generate:

Giovanni pianse, partita Maria
Maria è felice, abbandonato il marito
Giovanni è triste, abbandonato dagli amici
Roma cadde nella confusione, ucciso Cesare
Giovanni va a piedi, avendo distrutto la moto
Maria sorrideva, partendo per le vacanze

The only case of a not-grammatical result of such an inversion is for the sentence 2 “Partita, Maria pianse”:

* Maria pianse, partita

This sentence could be grammatical only when a verb modifier occurs, e.g.

Maria pianse, partita per le vacanze
One generic schema dealt, in previous versions of the grammar, with all possible post head modifiers:

\[
\text{post-head-modifier-phrase} :\! = \\text{head-adj-phrase} \& \\
\quad \text{[SYNSEM.MODIFIED rmod,} \\
\quad \quad \text{ARGS \ {}< \{REAL +,} \\
\quad \quad \quad \quad \text{SYNSEM \ #1 \ & \ \{ LOCAL.CAT.AUX - \}} \}, \\
\quad \quad \{REAL +,} \\
\quad \quad \quad \quad \text{SYNSEM.LOCAL.CAT} \ \{ \text{HEAD} \ \{ MOD \ < \#1 > \} \} > \}].
\]

This rule didn’t impose any constraint. As for participial and gerundive phrases, there were obviously all the constraints imposed at the lexical level (see above, 2.a). A sentence like

\* Pianse partendo, Maria

had no parse result, given the absence of the matrix phrase subject. Only one possible grammatical interpretation was given to sentences like:

Giovanni pianse, baciato l’amico ("baciato" = active)
Giovanni pianse, baciato dagli amici ("baciato" = passive)

The generic post-head-modifier-phrase, anyway, constructed also some sentences grammatically dubious, as:

?? Giovanni pianse, baciato
?? Giovanni pianse, baciando

and caused some undesirable overgenerations, so that in the version 1.0 the rule has been dropped out.

12.4. Notes

The Italian grammar does not deal with absolute phrases in the middle of a sentence, e.g. “Giovanni, baciando la moglie, era felice” or “Il presidente, criticato da tutti, diede le dimissioni”. The latter case is parsed in its attributive acception (“criticato da tutti” as a modifier of the NP “Il presidente”).
13 MODIFIERS: ADVERBIAL PHRASES

13.1 The adv_type hierarchy

The Italian grammar deals with some phenomena concerning Adverbial Phrases, e.g.

- the role of the advp
- the collocation in the sentence
- some attachment restrictions

and more precisely:

a) some locative adverbs can play the role of a verbal (or a deverbal noun) argument as in the sentence “vado là” (I’m going there). Other classes of adverbs (e.g. quantifier adverbs as “troppo” or “poco” (too much / little) are sometimes supposed to have an argumental role (e.g.: “io mangio troppo” [I eat too-much]), but it seems that as a first step they could be managed as modifiers (“io mangio troppo” could be quite similar to “io corro poco” [I run little], given also the fact that if an argument is not optional, it can’t be satisfied by an adverb (e.g. “io ammiro poco” [I admire little], vs “io ammiro poco gli scrittori italiani” [I admire little the italian writers])

b) as it concerns their position with regard to the modified element, adverbs could belong to two different types:

- the adverbs that always follow the modified [POSTHEAD + , e.g. “parlo male” vs “male parlo” [“I speak badly” vs “I badly speak”]]

- the adverbs that could either follow or precede the modified [POSTHEAD +/- , e.g. “presto partirò” vs “partirò presto” [“soon I’ll leave” vs “I’ll leave soon”]]

c) regarding the modified element, it could be (depending on the adverb involved)

- a verb
- an adjective
- a prepositional phrase
- a noun
- an adverb
In the Italian lexicon, any adverb is given an \textit{adv\_type} value according the following hierarchy

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{diagram.png}
\caption{Diagram showing the hierarchy of \textit{adv\_type} values.}
\end{figure}

Any subtype of \textit{adv\_type} select as a possible MOD one of the following hypetypes

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{hypetypes.png}
\caption{Diagram showing the hypetypes for \textit{adv\_type}.}
\end{figure}

13.2 The “negation” \textit{adv\_type}

As far as it regards negation, only the negative adverb “non” (“not”) has been taken into account, both in the case in which it precedes the verbs in negated clauses, as in the following
i) "Luca non è andato in Francia" (Luca didn't go in France)

ii) "Non nevicherà" (It wont snow)

and in the case of other constituent negation, as

iii) "il non corretto funzionamento dipende da un difetto di fabbricazione"

iv) "Paolo ha parlato non molto bene"

v) "Luca è andato in Francia non per divertirsi"

vi) "Luca è andato in Francia non con la sua macchina"

Two different entries of "non" are in the lexicon: the first entry with a VP scope (cases i) and ii), the second one with a non-VP scope, both inheriting from an underspecified neg-adv-word

```
[STEM < "non" >
  LKEYS.KEYREL neg-adv-relation
  HEAD adv [MOD [LOCAL.CONT.HOOK.INDEX #arg0] >]
  POSTHEAD +
  VAL [SUBJ >
       COMPS < >]
  CONT.RELS ![
               arg1-rel
               PRED neg_rel !>
               ARG0 #arg0
]

neg-vp-scope-le := neg-adv-word & [SYNSEM.LOCAL.CAT.HEAD.MOD [LOCAL.CAT.HEAD verb,
                                LOCAL.CAT.VAL.COMPS < >] > ].

neg-novp-scope-le := neg-adv-word & [SYNSEM.LOCAL.CAT.HEAD.MOD [LOCAL.CAT.HEAD neg-scope] > ].
```

The neg-scope type as an HEAD value is needed in order to unify with PPs, some kind of ADVPs, some kind of ADJPs and so on.
14 MODIFIERS: RELATIVE CLAUSES

A specific subtype of pronoun-word, namely wh-pronoun-word, has been added (simply assigning the \(WH +\) value).

Two extraction rules (extracted-subj-phrase and extracted-comp-phrase) and two filler rules (wh-subj-filler-phrase and wh-obj-filler-phrase) deal with relative clauses (currently only with relative pronouns playing the role of subject or direct object).

Extraction rules inherit directly from the Matrix rules “basic-extracted-subj-phrase” and “basic-extracted-comp-phrase”, apply only on finite clauses and assign the SUBJGAP value:

\[
\text{extracted-subj-phrase} := \text{basic-extracted-subj-phrase} \& \\
\quad \text{[SYNSEM.LOCAL.CAT.VAL.SUBJGAP + ,} \\
\quad \text{HEAD-DTR.SYNSEM.LOCAL.CAT.INFINITENESS - ]}.
\]

\[
\text{extracted-comp-phrase} := \text{basic-extracted-comp-phrase} \& \\
\quad \text{[SYNSEM.LOCAL.CAT.VAL.SUBJGAP - ,} \\
\quad \text{HEAD-DTR.SYNSEM.LOCAL.CAT.INFINITENESS - ]}.
\]

The SUBJGAP feature ensures that the subj-filler-rule applies only on an extracted-subj-rule output (the same mechanism applies for the obj-filler-rule and extracted-comp-rule). Note that SUBJGAP value has to be inherited by restructuring verbs (including auxiliares) and modal verbs (respectively in the rv-lxm type and in the modal-verb type).

Filler rules inherit (indirectly, via the “wh-head-filler-phrase”) from the Matrix rule “basic-head-filler-phrase”

\[
\text{wh-subj-filler-phrase} := \text{wh-head-filler-phrase} \& \\
\quad \text{[ARGS < word, } \\
\quad \text{[SYNSEM.LOCAL.CAT.VAL.SUBJGAP + ] > ]}.
\]

\[
\text{wh-obj-filler-phrase} := \text{wh-head-filler-phrase} \& \\
\quad \text{[ARGS < word, } \\
\quad \text{[SYNSEM.LOCAL.CAT.VAL.SUBJGAP - ] > ]}.
\]
15 COORDINATION

In the current Italian Grammar a very rough and purely syntactic treatment of the coordination has been implemented, according to the main guidelines of the proposal of Ennio Calzone, a student from Pisa University who collaborated with CELI in making his thesis about the italian coordination in HPSG. Its proposal is quite well articulated (mainly as far as it concern the ellipsis in coordination) and complex, but we didn’t attempt a full implementation for reasons of time consuming. Some documentation about the approach proposed by E. Calzone is currently disposable only in an Italian version.

A first rule conjunction-phrase has the role of creating a binary structure (marked as CONJ +) with a conjunction as a non-head-daughter and any consistent (verbal phrase, noun phrase, adjectival phrase, …, marked as CONJ -) as the head-daughter (e.g. “e Giovanni” [and John], “e siamo partiti” [and we left] and so on):

```
conjunction-phrase := head-valence-phrase & head-final &
  [ SYNSEM canonical-synsem &
    [ REAL -, PHR +,
      LOCAL.CONT.HOOK #hook,
      LOCAL.CAT [ MC #mc,
        VAL #val,
        CONJ +,
        HEAD [CTYPE #ctype ] ],
      NON-HEAD-DTR.SYNSEM canonical-synsem &
      [ LOCAL.CAT.HEAD cong &
        [CTYPE #ctype ] ],
      HEAD-DTR [SYNSEM [ LOCAL.CONT.HOOK #hook,
        LOCAL.CAT [ MC #mc,
          HEAD v-j-p-r-n & [CLT -],
          VAL #val,
```

A second set of rules (all inheriting form the same coord-binary-phrase) builds up a coordinated binary phrase (e.g. “Michele e Giovanni” [Michael and John], “abbiamo salutato gli amici e siamo partiti” [we said goodbye to our friends and we left], …), the only constraint being the syntactic category of constituents:
coord-binary-phrase := head-valence-phrase & head-final &
[ SYNSEM canonical-synsem &
[ LOCAL.CAT [ MC #mc,
  VAL #val,
  HEAD.MOD #mod,
  HEAD.CASE #case,
  POSTHEAD #ph,
  CONJ na ],
  REAL +,
  LEX #lex],
  NON-HEAD-DTR.SYNSEM canonical-synsem &
[ LOCAL.CAT [ VAL #val,
  HEAD v j-p-r-n & [CASE #case],
  CONJ - ]],
  HEAD-DTR conjunction-phrase & [SYNSEM [ LOCAL.CAT [ MC #mc,
  VAL #val,
  HEAD.MOD #mod,
  POSTHEAD #ph],
  LEX #lex ]]] ]}. 

n_coord_n := coord-binary-phrase &
[NON-HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD noun-or-pronoun,
HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD noun-or-pronoun].

v_coord_v := coord-binary-phrase &
[NON-HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD verb,
NON-HEAD-DTR.SYNSEM.NON-LOCAL.SLASH 0-dlist,
HEAD-DTR.SYNSEM.NON-LOCAL.SLASH 0-dlist,
HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD verb].

adj_coord_adj := coord-binary-phrase &
[NON-HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD adj,
HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD adj].

adv_coord_adv := coord-binary-phrase &
[NON-HEAD-DTR.SYNSEM.LOCAL [CAT.HEAD adv,
  CONT.HOOK.INDEX.E.MOOD #mood],
  HEAD-DTR.SYNSEM.LOCAL [CAT.HEAD adv,
  CONT.HOOK.INDEX.E.MOOD #mood] ].

p_coord_p := coord-binary-phrase &
[NON-HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD compl-or-prep,
HEAD-DTR.SYNSEM.LOCAL.CAT.HEAD compl-or-prep].

The grammar doesn’t deal neither with conjunction reduction nor with unlike-constituents coordination.