

CINTIL DependencyBank PREMIUM Handbook: Design options for the representation of grammatical dependencies

António Branco, João Silva, Andreia Querido and
Rita de Carvalho

DI-FCUL-TR-2015-05

DOI:10451/20226

(<http://hdl.handle.net/10451/20226>)

October 2015



Published in the repository of the Department of Informatics of the University of Lisbon, Faculty of Sciences (<http://repositorio.ul.pt/handle/10451/12140>).

CINTIL DependencyBank PREMIUM Handbook:
Design options for the representation of grammatical dependencies

António Branco, João Silva, Andreia Querido, Rita de Carvalho

University of Lisbon

July 2015

1- INTRODUCTION	4
2- DEPENDENCY IN A NUTSHELL	5
2.1 semantic relations	5
3- TAG SET.....	6
3.1 inflection tags.....	6
3.2 lexical categories.....	7
3.3 grammatical functions.....	9
3.4 semantic functions.....	10

4- DEPENDENCY RELATIONS.....	10
5- PHONETICALLY NULL ITEMS.....	10
5.1 verb ellipsis.....	11
5.2 noun ellipsis.....	11
5.3 traces.....	13
6- SPECIFIC CONSTRUCTIONS.....	13
6.1 comparatives.....	13
6.2 consecutives.....	14
6.3 coordination.....	15
6.4 complex predicates: auxiliary, raising, modal and control verbs.....	17
6.4.1 notes on the subject relation in complex predicates.....	20
6.5 completive clauses.....	21
6.6 "tough" constructions.....	22
6.7 cleft constructions.....	22
7- LONG-DISTANCE RELATIONS.....	23
7.1 topicalization.....	23
7.2 relatives.....	24
7.3 interrogatives.....	25
7.3.1 indirect interrogatives.....	27
8- VALENCY ALTERNATIONS.....	27
8.1 passives.....	27
8.2 anticausatives.....	28
9- TOKENIZATION.....	28
9.1 sentence splitting.....	28

9.2 non verbal utterances.....	29
9.3 contractions.....	29
9.4 clitics.....	29
10- MULTI-WORD EXPRESSIONS.....	29
10.1 proper names.....	29
10.2 cardinals.....	29
11- TEXTUAL MARKING.....	30
11.1 punctuation.....	30
11.2 comma.....	30
11.3 quotation marks.....	31
12- OTHER PHENOMENA.....	32
12.1 predicative over the object.....	32
12.1.1 other predicatives.....	32
12.2 emphatic duplication.....	33
12.3 notes on the M-PRED relation.....	33
12.4 notes on the DEP relation.....	34
12.5 notes on the ERROR relation.....	35
13- REFERENCES.....	36
APPENDIX.....	37

1- Introduction

Treebanks are data sets of utmost importance for the study of natural languages and for their computational processing. They permit the training and evaluation of different processing tools, including taggers, chunkers, parsers, deep linguistic grammars, etc.

A treebank is an annotated corpus. It is a data set consisting of a collection of individual written utterances associated to the representation of their linguistic structure, which can be set to capture different degrees of linguistic information.

CINTIL DependencyBank PREMIUM is a corpus of Portuguese utterances manually annotated with the representation of grammatical dependency relations and the information of part-of-speech, inflection and lemmas. It is being developed and maintained at the University of Lisbon.

This document aims at supporting the utilization and exploitation of the CINTIL DependencyBank PREMIUM. It presents its major design options in what concerns the representation of syntactic relations.

The adopted design options were informed by advanced linguistic theorizing. The reader is referred to the literature for a thorough discussion and justification of them.

For the annotation methodology used see (Barreto et al., 2006).

These guidelines follow the same principles that guide the annotation of CINTIL-DepBank (Branco et al., 2011).

CINTIL-DepBank is produced semi-automatically. A text is annotated by LXGram (Branco & Costa, 2008), a computational grammar that generates a parse forest for each sentence. The correct analysis is then picked manually. The analysis that LXGram produces is quite complex, including, among other information, syntactic constituency, grammatical dependencies and a semantic representation of meaning. From this analysis we extract vistas, or subsets of the full linguistic information. CINTIL-DepBank corresponds to the grammatical dependency vista.

CINTIL-DepBank is extracted from the analyses produced by LXGram and, for that reason, it only includes those sentences that the grammar is able to parse. Given that the grammar has not full coverage, some sentences are not treebanked. This circumstance motivated the creation of a treebank with grammatical dependencies that does not depend on LXGram.

The current corpus, CINTIL DependencyBank PREMIUM, is created by manually correcting the output of the morpho-syntactic analysis provided by the LX-Suite (Branco & Silva, 2006) and of the dependency parsing produced by the LX-DepParser¹, both tools with full coverage. This fact enables 100% coverage of the sentences in the corpus.

The manual correction is done by two annotators under a double-blind scheme, that is followed by adjudication by a third annotator. This process is supported by a general purpose annotation tool.²

1 Available at <http://lxcenter.di.fc.ul.pt/services/en/LXServicesParserDep.html>

2 In our particular case, we use the WebAnno tool (<https://code.google.com/p/webanno/>).

2- Dependency in a nutshell

In an utterance, a lexeme B depends on a lexeme A when the occurrence of B in its specific position is made possible by the occurrence of A. In such case, it is considered to exist a grammatical dependency relation from the lexeme B, the dependent element, to the lexeme A, the governor element of the dependency.

Dependency relations can be depicted as graphs whose nodes are lexemes and whose directed arcs establish a connection from a governor to its dependent lexemes.

In the CINTIL DependencyBank PREMIUM, individual lexemes are further annotated with a feature bundle containing information on lexical category, lemma and inflection.

Dependency relations can be of a number of different types, which are mostly the usual grammatical functions, and with whose tags the arcs are decorated.

A grammatical function results from an abstraction over complements and modifiers of different predicates. It permits to categorize complements, or modifiers, with similar syntactic constraints on their realization, such as category, case, agreement, canonical word order, inflection paradigm, etc.

The possible values of lexical functions are listed in section 3.2 below.

The possible values of grammatical functions are listed in section 3.3 below.

2.1 semantic relations

The CINTIL DependencyBank PREMIUM was extended so that besides the tags for the different dependency relations, the arcs are further decorated with tags indicating the semantic relation at stake.

A semantic function, or semantic role, is also an abstraction over complements and modifiers of various syntactic predicates, but along a different, semantic, dimension. It permits to categorize complements, or modifiers, according to similar semantic constraints on their denotation, that is in terms of the similar contribution that the extra-linguistic elements they may denote bring for the characterization of the event being described.

The possible values of semantic functions are listed in section 3.4 below.

3- Tag set

3.1 inflection tags

Tag	Description
Tags for nominal categories	
m	Masculine
f	Feminine
s	Singular
p	Plural
dim	Diminutive
sup	Superlative
comp	Comparative
Tags for verbs	
1	First Person
2	Second Person
3	Third Person
pi	Presente do Indicativo
ppi	Pretérito Perfeito do Indicativo
ii	Pretérito Imperfeito do Indicativo
mpi	Pretérito Mais que Perfeito do Indicativo
fi	Futuro do Indicativo
c	Condicional
pc	Presente do Conjuntivo
ic	Pretérito Imperfeito do Conjuntivo
fc	Futuro do Conjuntivo
imp	Imperativo
Tags for infinitive verbs	
ifl	Inflected
nifl	Not Inflected

3.2 lexical categories

Tag	Category	Examples
ADJ	Adjectives	bom, brilhante, eficaz, ...
ADV	Adverbs	hoje, já, sim, felizmente, ...
CARD	Cardinals	zero, dez, cem, mil, ...
CJ	Conjunctions	e, ou, tal como, ...
CL	Clitics	o, lhe, se, ...
CN	Common Nouns	computador, cidade, ideia, ...
DA	Definite Articles	o, os, ...
DEM	Demonstratives	este, esses, aquele, ...
DFR	Denominators of Fractions	meio, terço, décimo, %, ...
DGTR	Roman Numerals	VI, LX, MMIII, MCMXCIX, ...
DGT	Arabic Numerals	0, 1, 42, 12345, 67890, ...
DM	Discourse Marker	olá, ...
EADR	Electronic Addresses	http://www.di.fc.ul.pt , ...
EOE	End of Enumeration	etc
EXC	Exclamation	ah, ei, ...
GER	Gerunds	sendo, afirmando, vivendo, ...
GERAUX	Gerund "ter"/"haver" in compound tenses	tendo, havendo
IA	Indefinite Articles	uns, umas, ...
IND	Indefinites	tudo, alguém, ninguém, ...
INF	Infinitive	ser, afirmar, viver, ...
INFAUX	Infinitive "ter"/"haver" in compound tenses	ter, haver, ...
INT	Interrogatives	quem, como, quando, ...

ITJ	Interjection	bolas, caramba, ...
LTR	Letters	a, b, c, ...
MGT	Magnitude Classes	unidade, dezena, dúzia, resma, ...
MTH	Months	Janeiro, Dezembro, ...
NP	Noun Phrases	idem, ...
ORD	Ordinals	primeiro, centésimo, penúltimo, ...
PADR	Part of Address	Rua, av., rot., ...
PNM	Part of Name	Lisboa, António, João, ...
PNT	Punctuation Marks	., ?, (, ...
POSS	Possessives	meu, teu, seu, ...
PPA	Past Participles not in compound tenses	sido, afirmados, vivida, ...
PP	Prepositional Phrases	algures, ...
PPT	Past Participle in compound tenses	sido, afirmado, vivido, ...
PREP	Prepositions	de, para, em redor de, ...
PRS	Personals	eu, tu, ele, ...
QNT	Quantifiers	todos, muitos, nenhum, ...
REL	Relatives	que, cujo, tal que, ...
STT	Social Titles	Presidente, dr ^a ., prof., ...
SYB	Symbols	@, #, &, ...
TERMN	Optional Terminations	(s), (as), ...
UM	"um" or "uma"	um, uma
UNIT	Abbreviated Measurement Unit	kg., km., ...
VAUX	Finite "ter" or "haver" in compound tenses	temos, haveriam, ...
V	Verbs (other than PPA, PPT, INF or GER)	falou, falaria, ...
WD	Week Days	segunda, terça-feira, sábado...

Tags for multi-word expressions		
LADV1... LADVn	Multi-Word Adverbs	de facto, em suma, um pouco, ...
LCJ1...LCJn	Multi-Word Conjunctions	assim como, já que, ...
LDEM1... LDEMn	Multi-Word Demonstratives	o mesmo, ...
LDFR1...LD- FRn	Multi-Word Denominators of Fractions	por cento
LDM1...LDMn	Multi-Word Discourse Markers	pois não, até logo, ...
LINT1...LINTn	Multi-Word Interrogatives	o que, ...
LITJ1...LITJn	Multi-Word Interjections	meu Deus
LPRS1...LPRSn	Multi-Word Personals	a gente, si mesmo, V. Exa., ...
LPREP1... LPREPn	Multi-Word Prepositions	através de, a partir de, ...
LQD1...LQDn	Multi-Word Quantifiers	uns quantos, ...
LREL1... LRELn	Multi-Word Relatives	tal como, ...

3.3 grammatical functions

Tag	Meaning
C	Complement
DO	Direct Object
IO	Indirect Object
M	Modifier
N	Relationship between words in named entities
OBL	Oblique Complement
PRD	Predicate
SJ	Subject
SP	Specifier

3.4 semantic functions

Tag	Meaning
ADV	Adverbial
ARG1	First Argument
ARG2	Second Argument
ARGA	Causative agent of verb with causative alternance
CAU	Cause
DIR	Direction
EXT	Extension
LOC	Localization
MNR	Mode
NULL	Null
PNC	Objective
POV	Viewpoint
PRD	Secondary predication
TMP	Time

4- Dependency relations

The head of a constituent that is a complement, a modifier or a specifier in a given predication, is dependent of the corresponding syntactic predicator.

The dependency relation of a specifier is SP, and of a Modifier is M.

The dependency relation of a complement is its grammatical function.

In the case of a complement of a Preposition, its complement is dependent under the grammatical relation of C (standing just for "Complement").

5- Phonetically null items

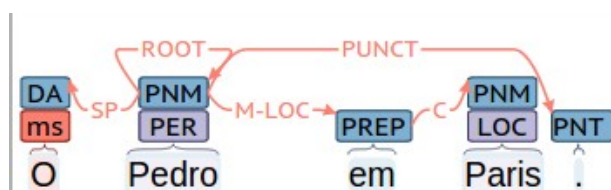
Phonetically null items signal positions related to other positions in the graph (in the case of traces), or signal elided elements whose context is rich enough to support the recovery of their interpretation (in case of null subjects or null heads). Phonetically null items and their arcs are not represented in the graph.

For example:

The diagram illustrates the morphological and syntactic analysis of the sentence "A a porta de casa, tudo em a mesma." Each word is represented by a box containing its morphological components (e.g., CN, fs, PORTA) and its syntactic role (e.g., PREP, DA, CN, PNT, IND, LADV1, LADV2, LADV3). Red arrows indicate dependencies between morphological and syntactic levels, such as M-LOC, DEP, C, SP, M-ADV, ROOT, M-MNR, PUNCT, and ADV.

The verb is the sentence ROOT.

When the verb ellipsis corresponds to a copulative verb, the arc has a M-*n* label, between the subject and the former predicative over the subject.



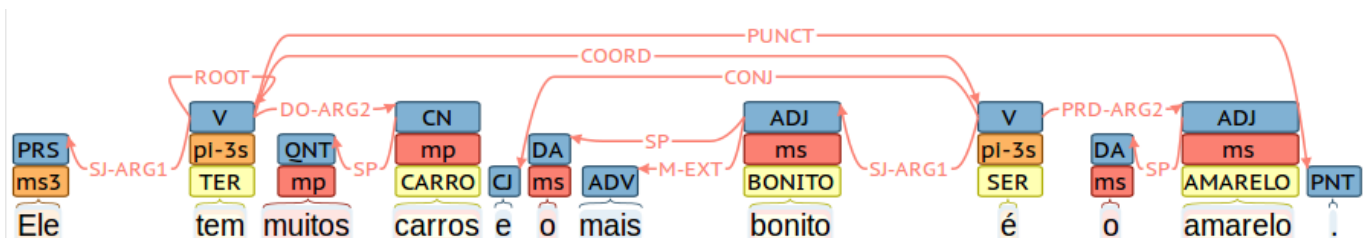
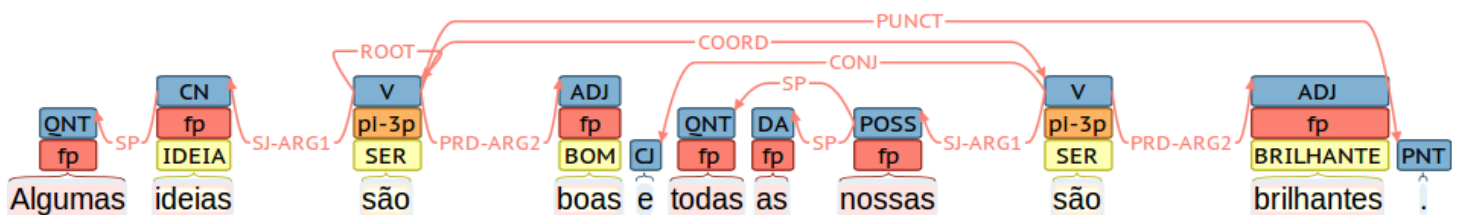
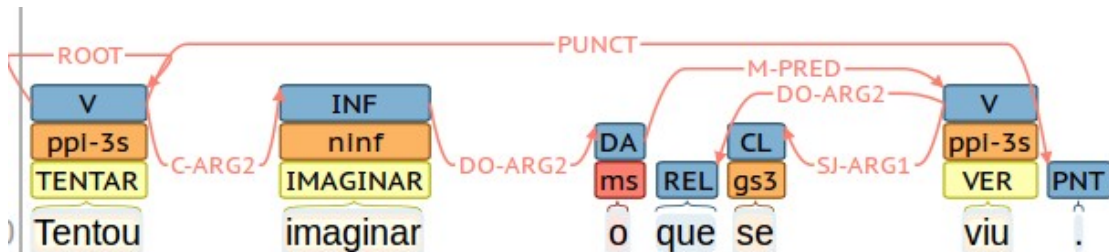
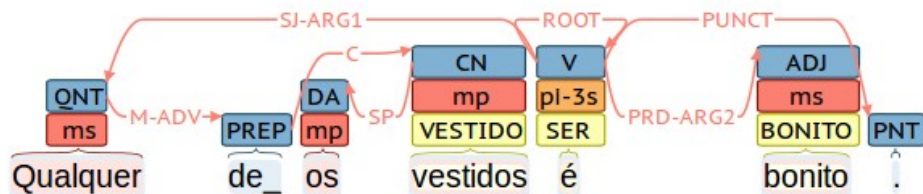
Here, the end point of that incoming edge is moved to the possible modifier(s) of the ellided item provided these modifiers are not PREP or V (the head of relative clauses).

Any other edges that should outgo of the ellided item are updated so that they start also from that modifier node.

If there would be multiple outgoing modifier edges, the one whose end point is superficially closest to the ellided item is chosen, while the other modifier edges are updated to start on that one.

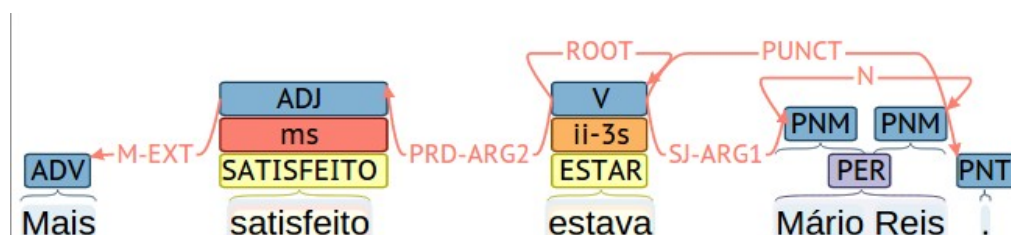
In case there are modifiers to the left and to the right of the ellided item, the left one is opted for.

If there are no modifiers (different from PREP or V), and thus not outgoing modifier edges from the putative ellided item to them, the specifier node is used instead (taking the superficially closest specifier if there are more than one).



5.3 traces

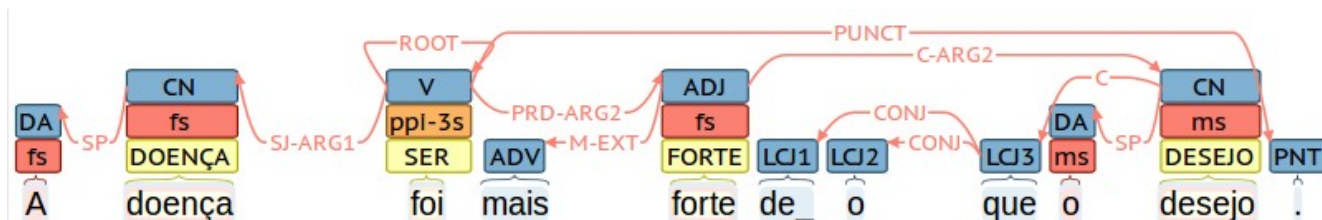
Differently from a constituency treebank, traces of constituents displaced are not represented here by a phonetically null category. Instead, the head of the displaced node is dominated by its governor.



6- Specific constructions

6.1 comparatives

A comparative construction is typically built around an adjective by two constituents, an adverbial of degree and a CONJP phrase, which have the following rendering in terms of dependency relations:



(Some adverbs may also support comparative constructions, as *perto*, in the example *mais perto do que a Maria*.)

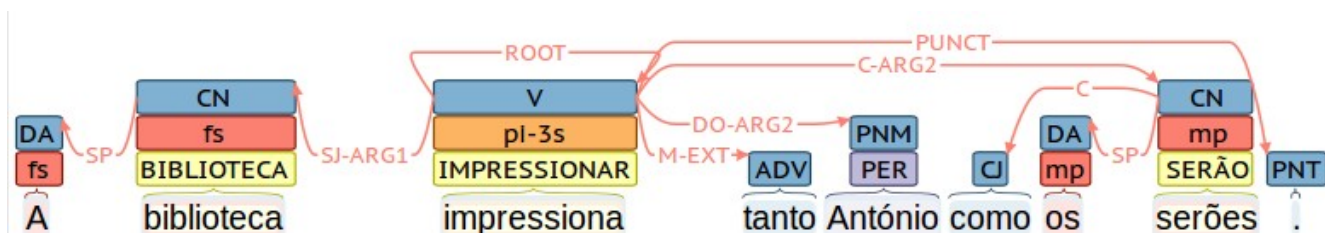
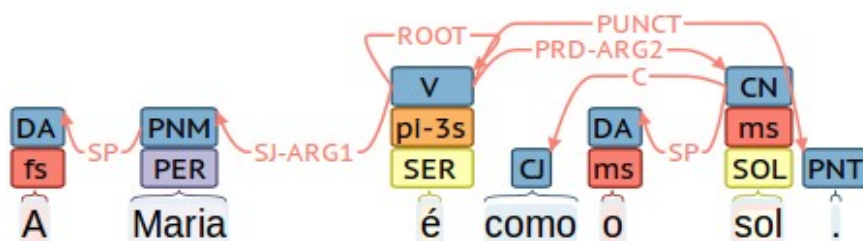
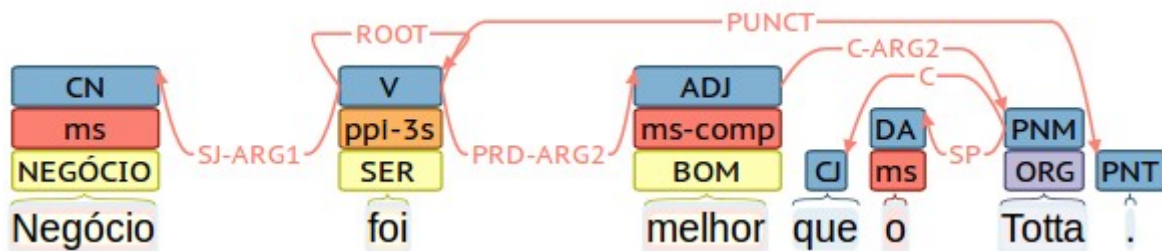
The adverbial of degree (e.g. *mais*, *menos*, *tão*) is dependent on the adjective.

The head of the CONJP phrase is dependent of the adjective, of which it is a complement.

The CONJP may be absent of the comparative construction. In this case, though it can be semantically recovered from the context, there is no phonetically null item inserted in the graph.

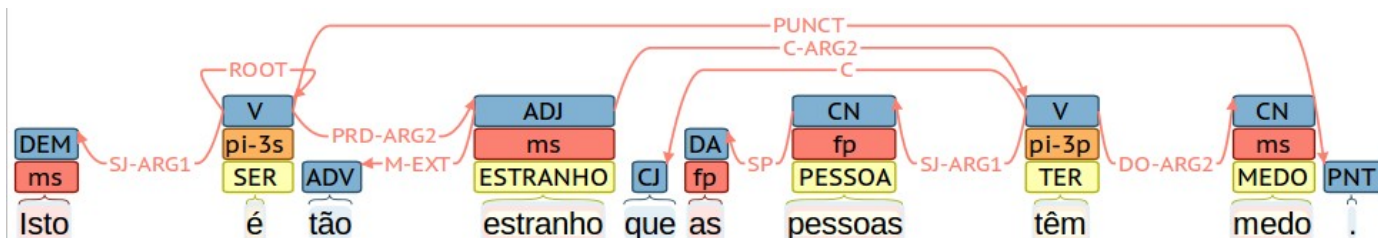
The exception happens with adjectives like *maior*, *menor*, *melhor*, *pior*, which also express the comparison, in which case the comparative construction is built around the adjective and the CONJP phrase:

There are some comparative constructions in which there are no adjectives. In those cases the analysis will respect the general rules of ellipsis (see section 5 above).



6.2 consecutives

Consecutive clauses denote the result (intended or actual) of the action indicated by the verb in the main clause. They have the following rendering in terms of dependency relations:



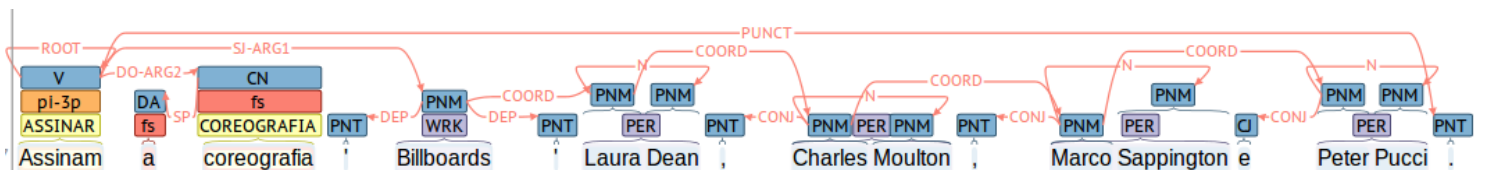
6.3 coordination

Each conjunct is dependent on the immediately preceding conjunct under a dependency relation of type COORD.

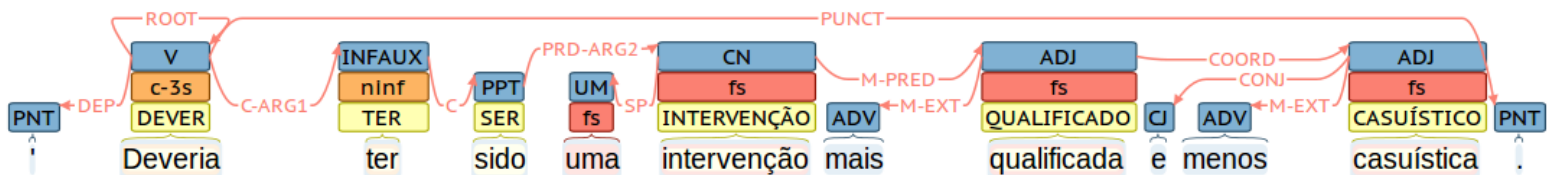
The commas or conjunctions are dependent on the subsequent conjunct under a dependency relation of type CONJ.

There are always only one ROOT, even in a coordination of sentences.

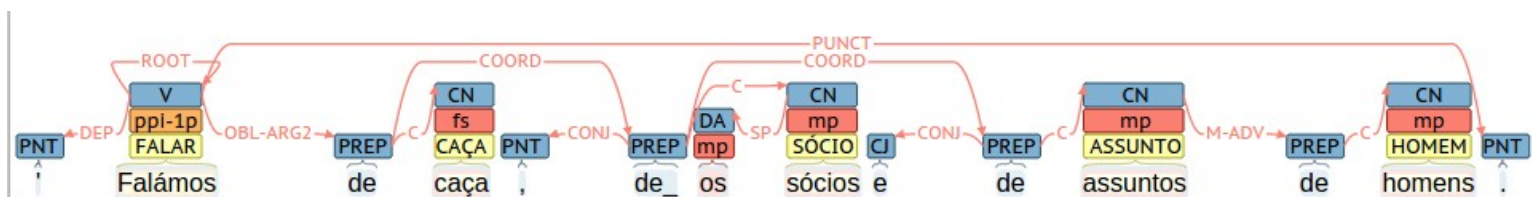
Coordination inside a NP:



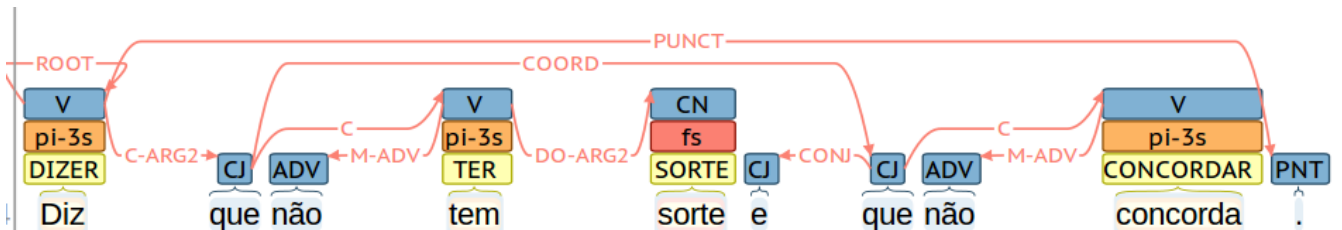
Coordination inside an ADJP:



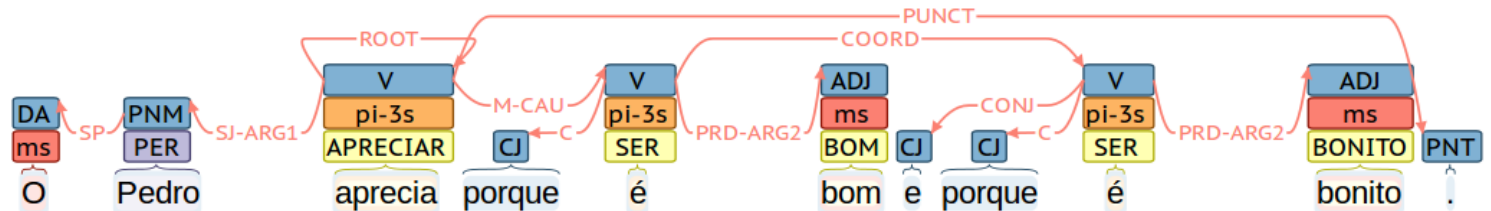
Coordination of Prepositional Phrases:



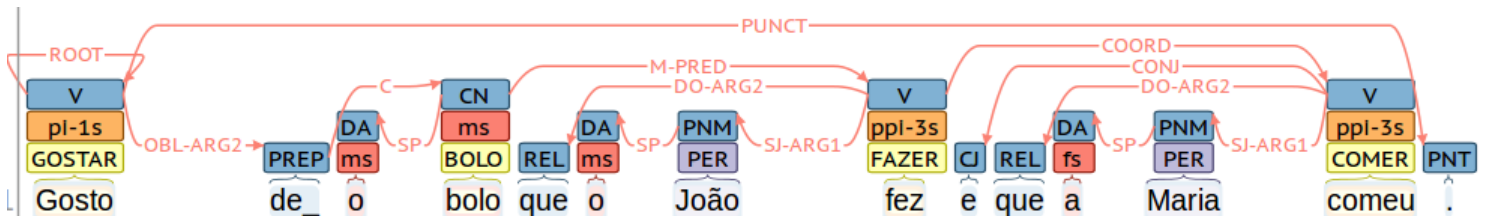
Coordination of Completive Clauses:



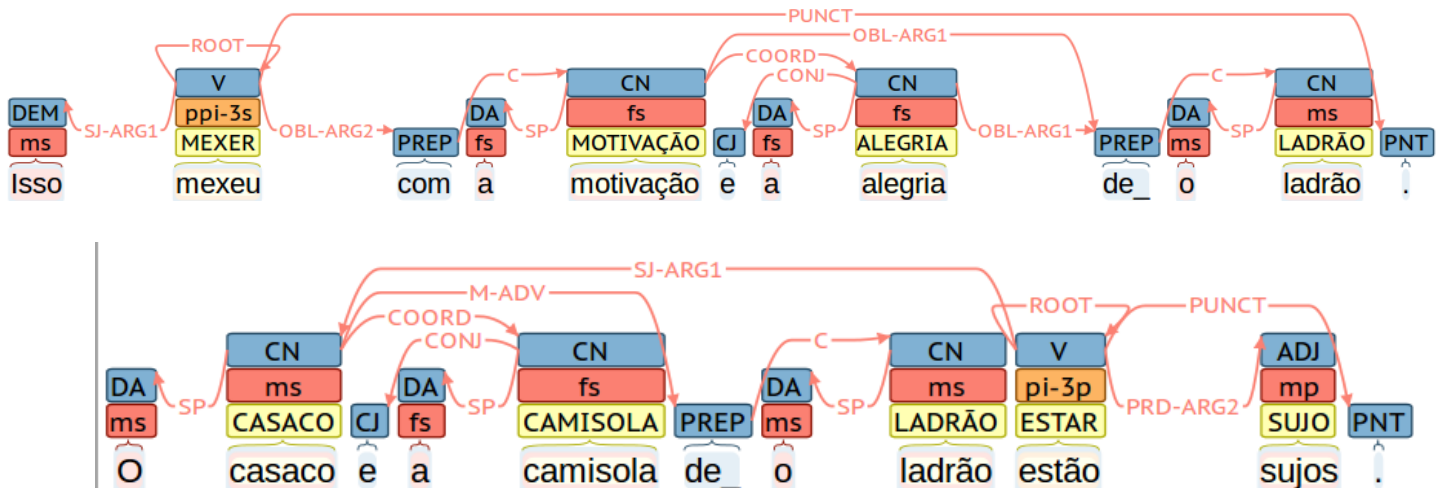
Coordination of Subordinate Clauses:

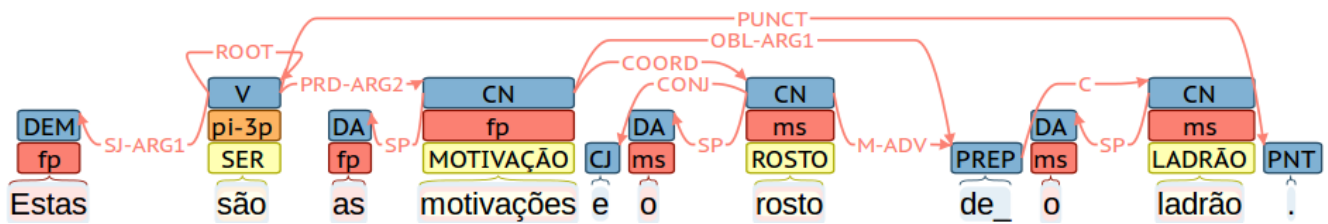


Coordination of Relative Clauses:

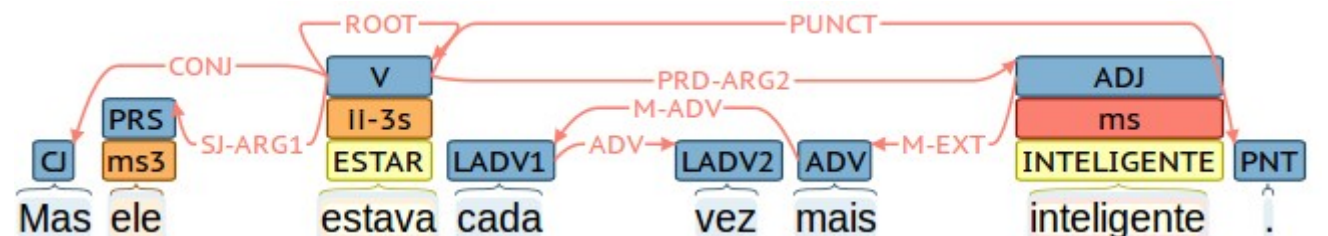
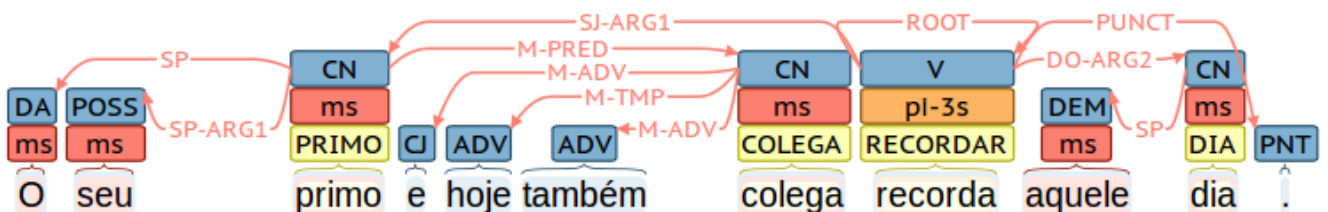


When two coordinated elements (eventually) share the same complements or modifiers, the annotation is done as is shown below:





There are cases where the coordinative conjunctions do not imply a relation of coordination. Below are two examples:

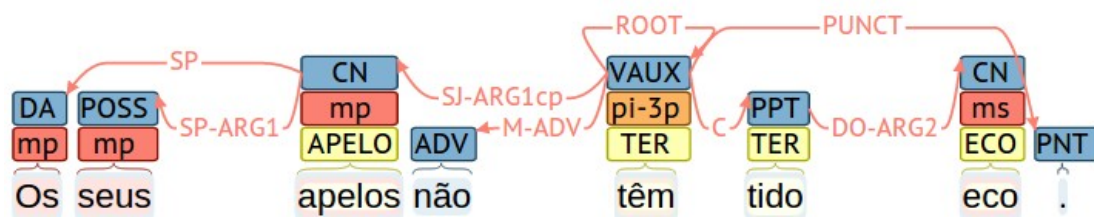


6.4 complex predicates: auxiliary, raising, modal and control verbs

In a complex predicate, the Subject relation is established with the leftmost verb and the other argument relations are established with the main verb.

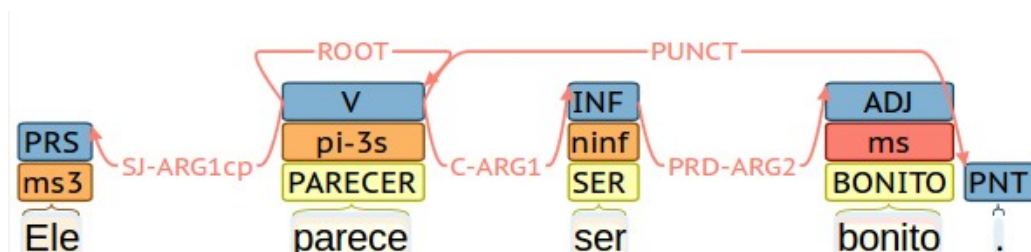
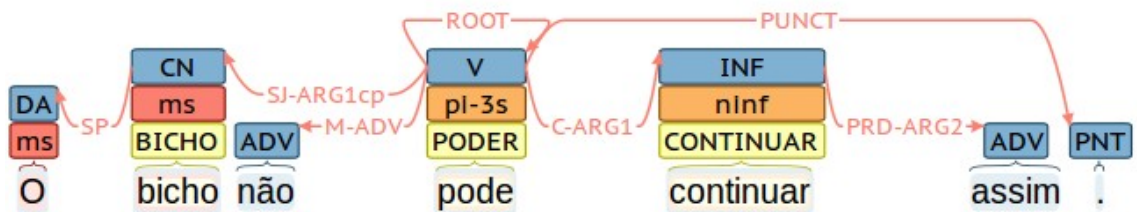
Auxiliary verbs:

The relation between the auxiliary verb (*ter*, *haver* or *ir*) and the main verb is tagged with C. The subject of the complex predicate is marked with "cp":

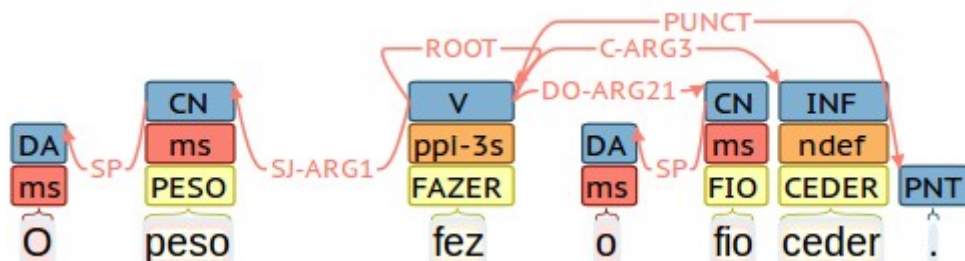


Raising and modal verbs:

The relation between the auxiliary verb and the main verb is tagged with C-ARG1. The subject of the complex predicate is marked with "cp":



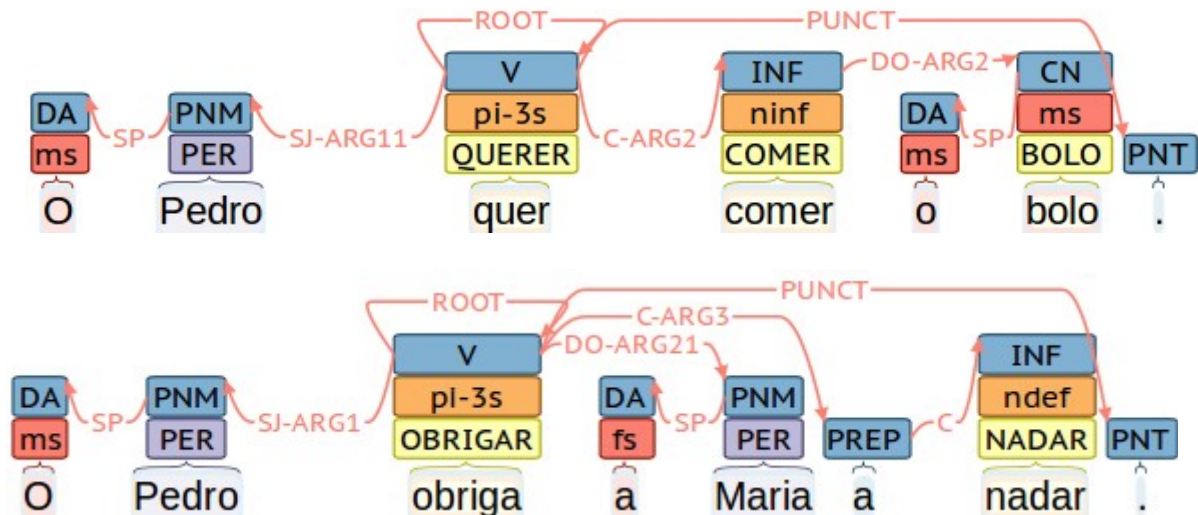
Raising-to-object predicates are annotated with the same tags as object control verbs.



Control verbs:

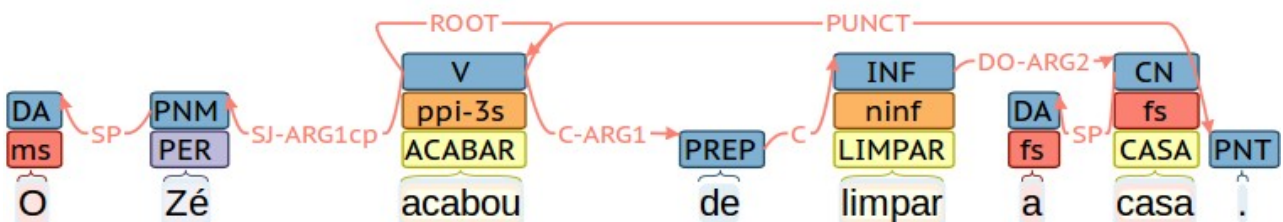
In subject control verbs the Subject relation is tagged with SJ-ARG11. The relation between the auxiliary verb and the main verb is tagged with C-ARG2.

In object control verbs the Direct Object relation is tagged with DO-ARG21 and the Indirect Object relation is tagged with IO-ARG31.

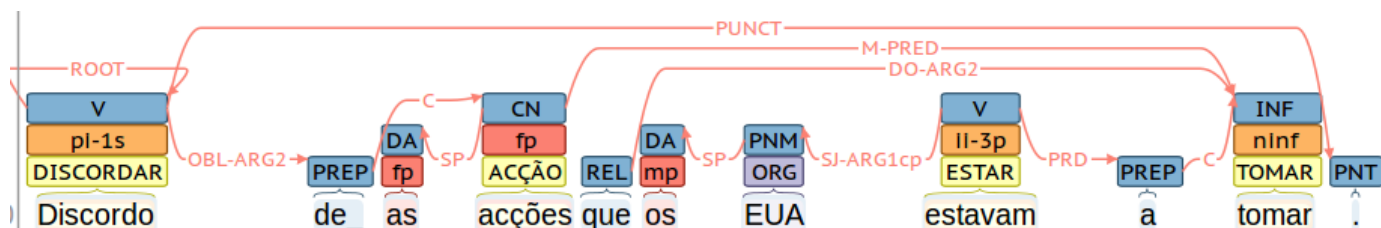


Aspectual auxiliary verbs:

The relation between the auxiliary verb and the preposition is tagged with C-ARG1. The subject of the complex predicate is marked with “cp”.

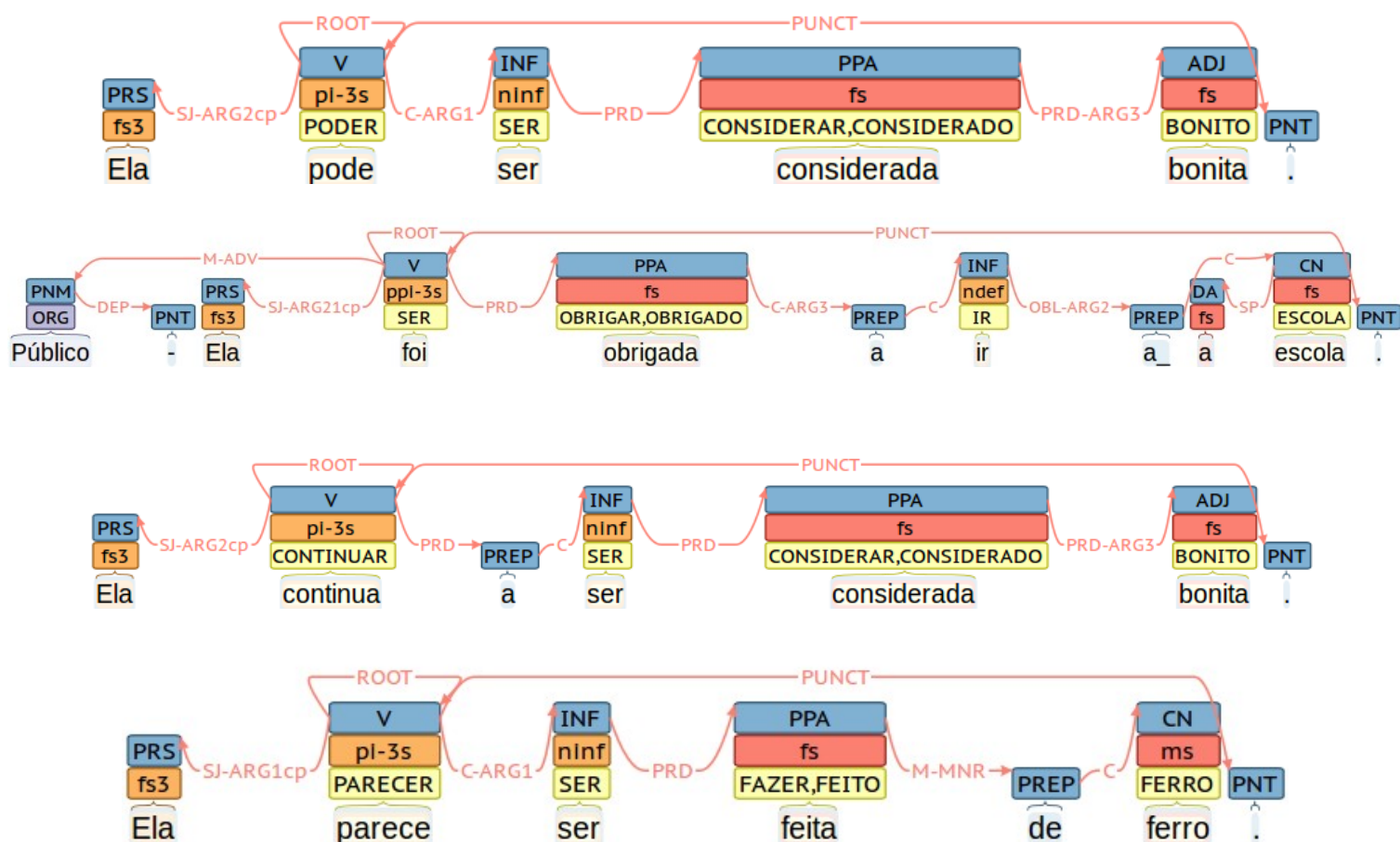


In the case of the auxiliary verbs “ficar a/por”, “continuar a” and “estar a/por” the relation between the auxiliary verb and the preposition is tagged with PRD.



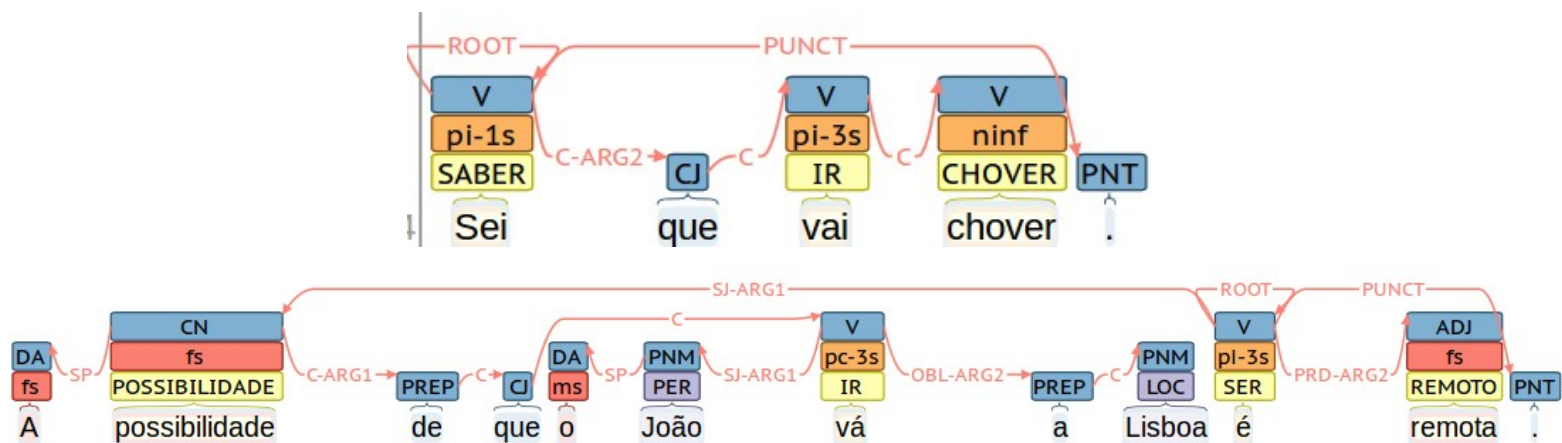
6.4.1 notes on the subject relation in complex predicates

In doubly complex predicates the tag chosen for the subject relation will be the one with more information.

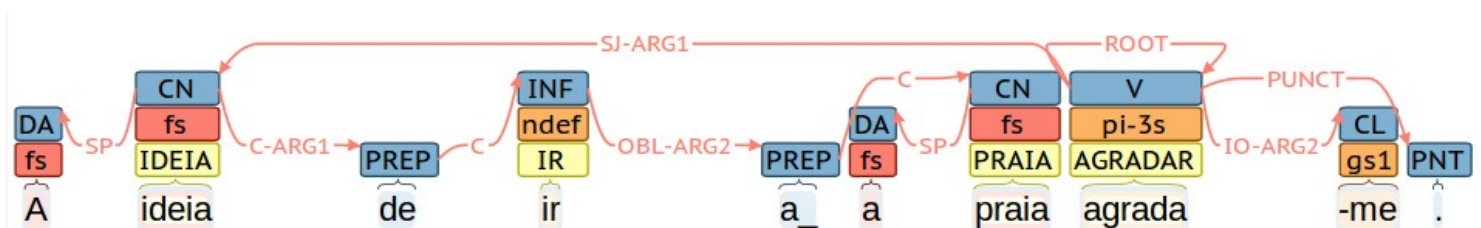


6.5 completive clauses

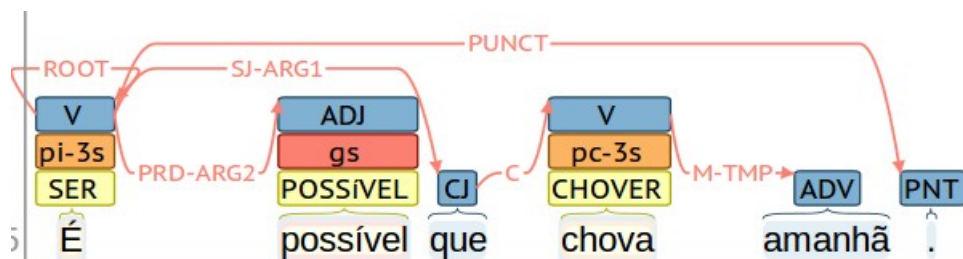
Completive clauses are subordinated clauses that are arguments of a verb, a noun or an adverb. There are finite and non-finite completive clauses. The finite completive clauses are introduced by the conjunctions *que*, *se*, *para* and *como*.

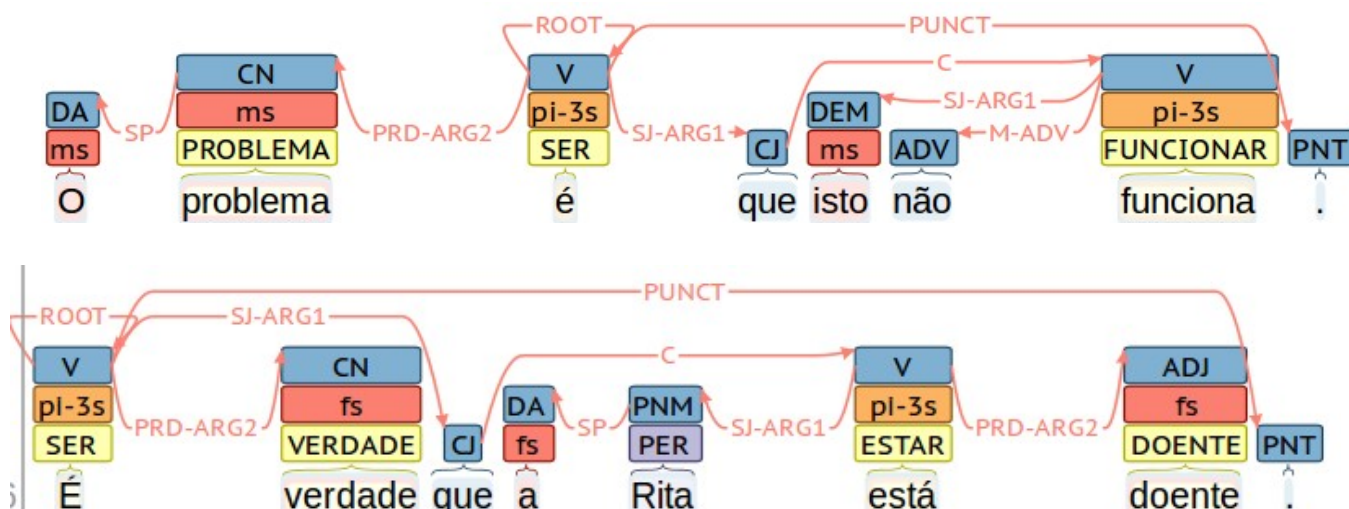


The non-finite completive clauses are not introduced by any conjunction:



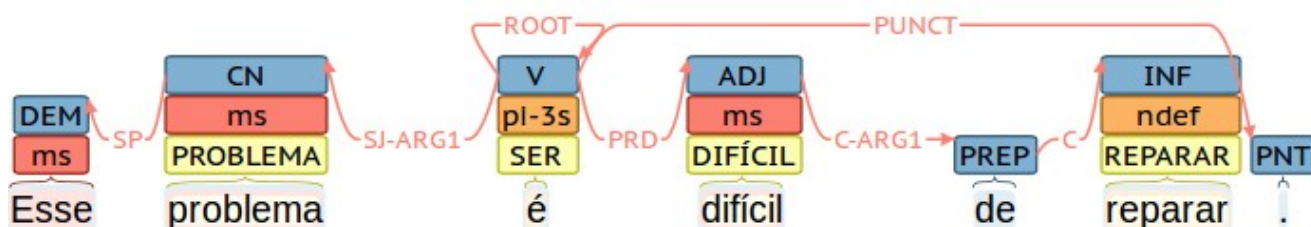
The completive clause can also be the subject of the sentence, as shown below:





6.6 "tough" constructions

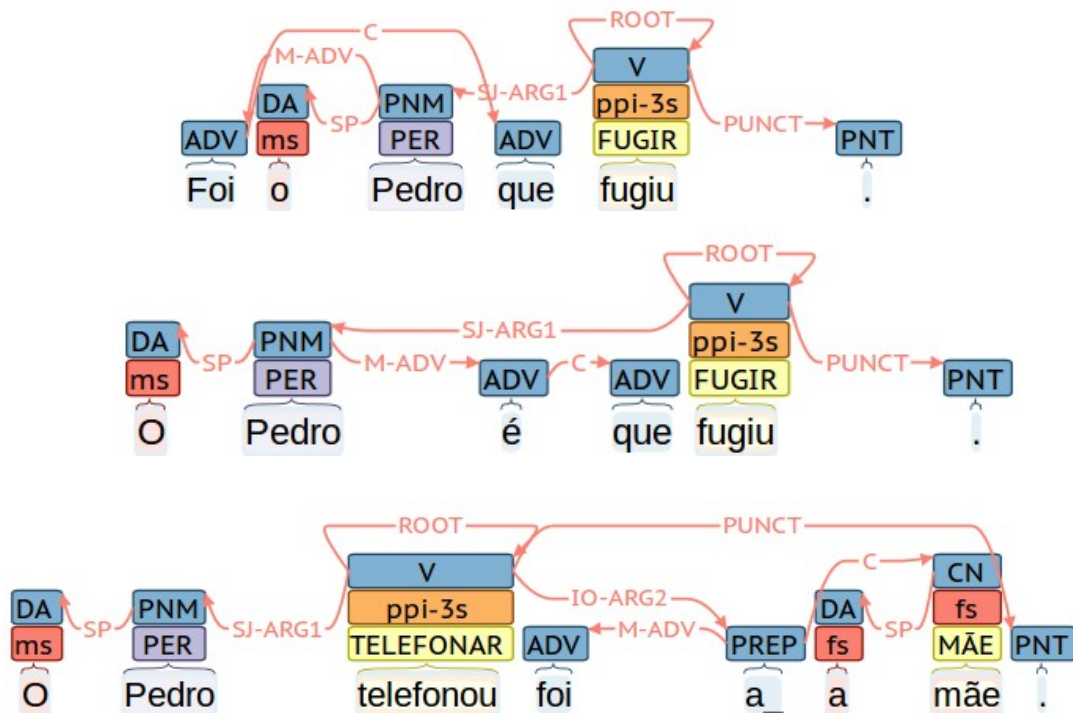
The sentential complement of the adjective, introduced by the preposition *de*, and projected by an inflected infinitive, has the tag C-ARG1:



6.7 cleft constructions

Cleft constructions are constructions with the verb *ser* and the pronoun *que*. They give focus to one of the sentence elements. The verb and the pronoun can be omitted and the result is a simple sentence.

Cleft constructions with *quem* are analysed as free relatives. In the other cleft constructions, the elements *ser* and *que* are analysed as ADV with a relation of C.

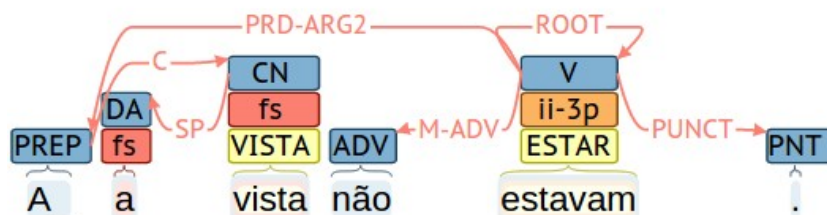


7- Long-distance relations

Long distance relations of dependency are established between a lexeme and a governor of the minimal predication where it typically occurs in (declarative) counterparts with canonical SVO word order. Constructions with long-distance relations include topicalizations, interrogatives and relatives.

7.1 topicalization

The head of the topicalized constituent is dependent on the governor of the minimal predication from which it was topicalized.

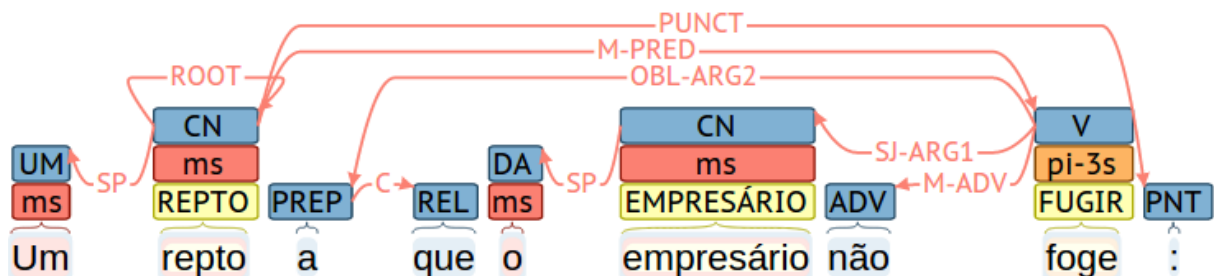
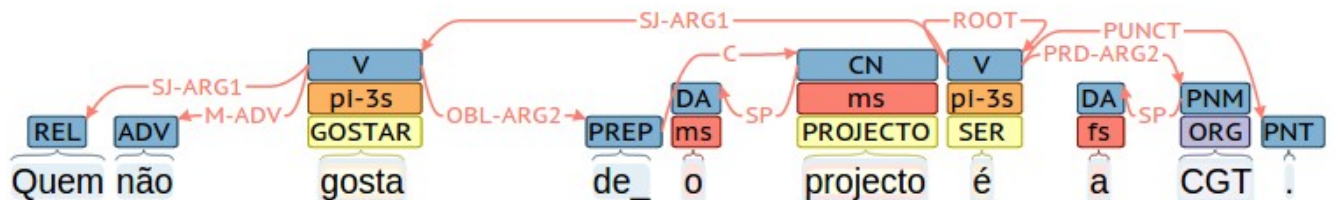
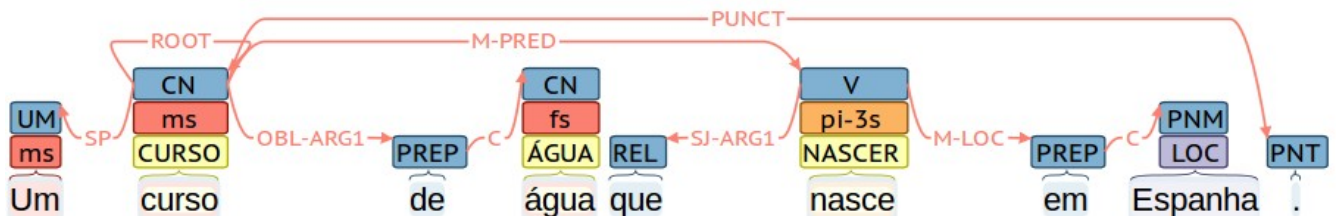
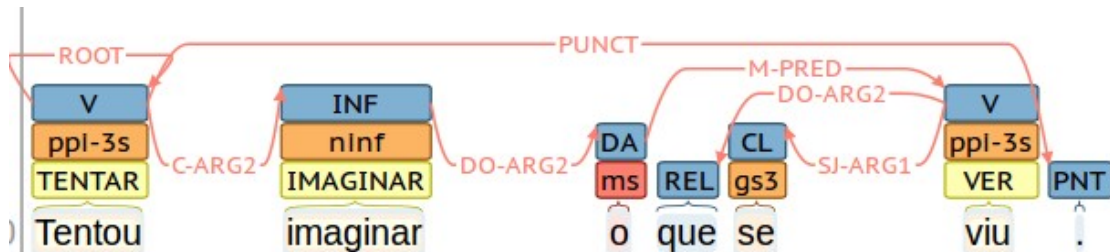


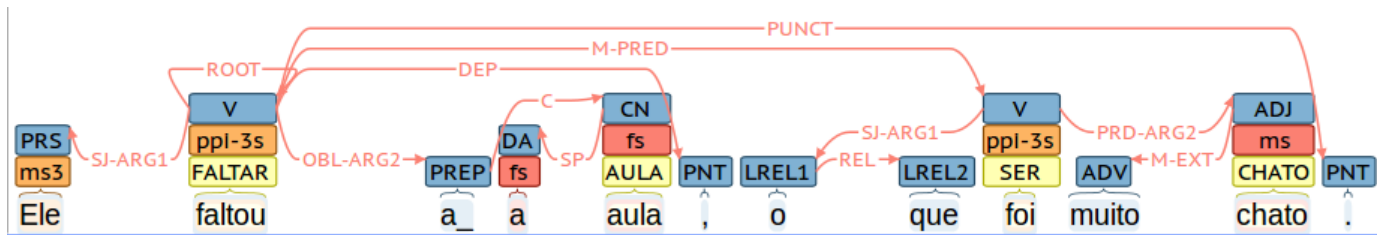
7.2 relatives

The head of a relative phrase is dependent on the governor of the minimal predication from which it was relativized.

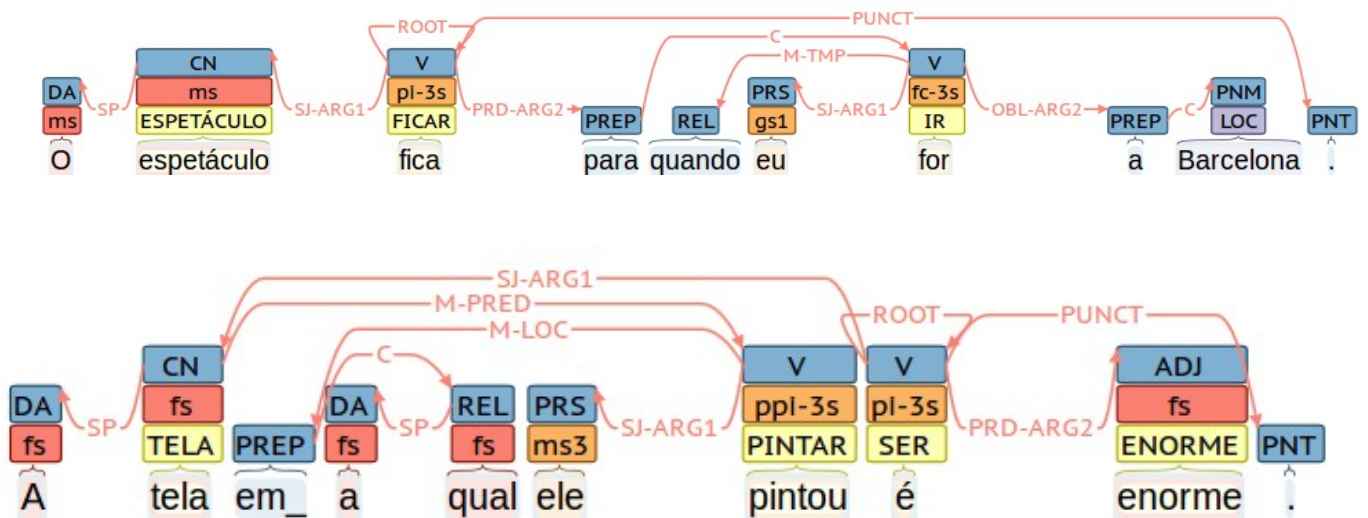
The relative pronoun can be:

- An argument:

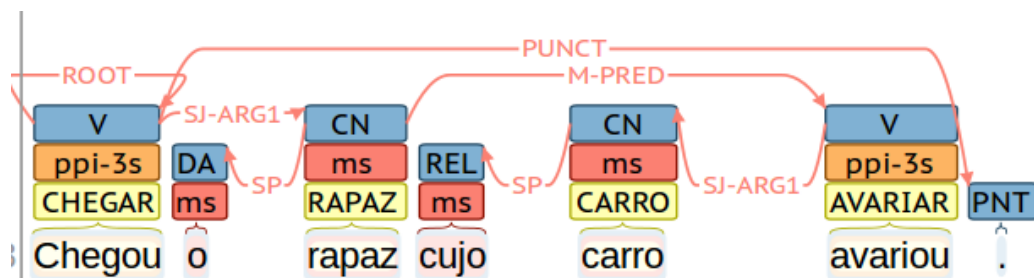




– A modifier:



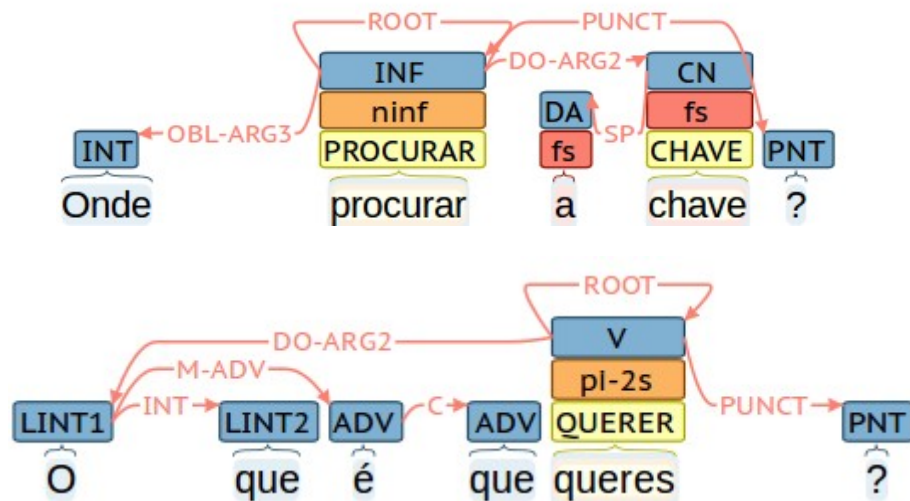
– A specifier:



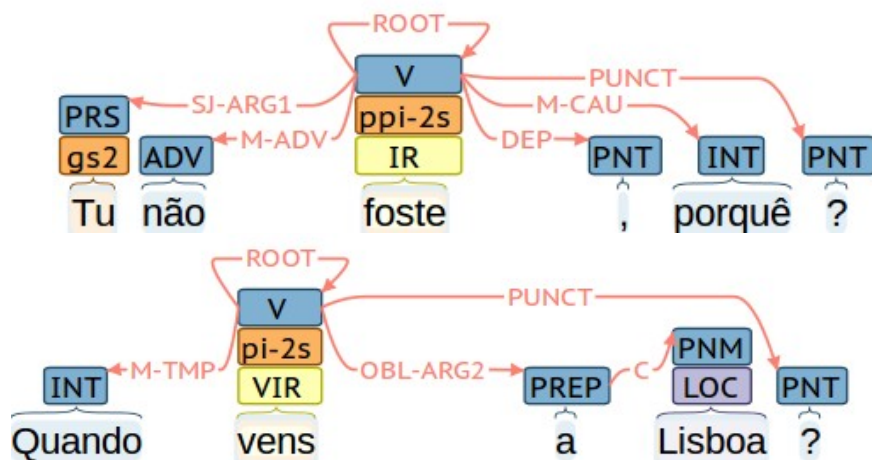
7.3 interrogatives

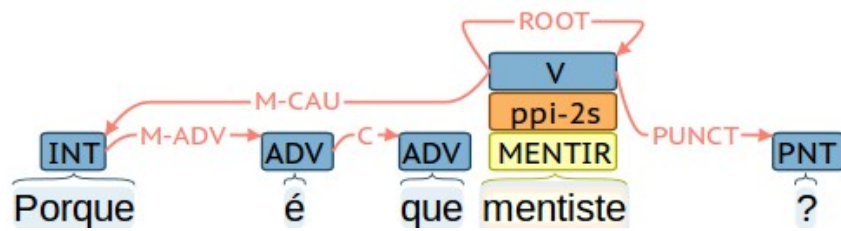
In partial interrogatives the interrogative pronoun is dependent on the verb and their relation is tagged with the grammatical function that the interrogative pronoun has in the sentence. The interrogative pronoun can be:

- An argument:

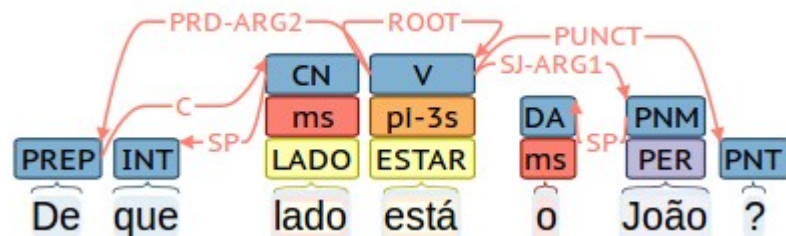


- A modifier:



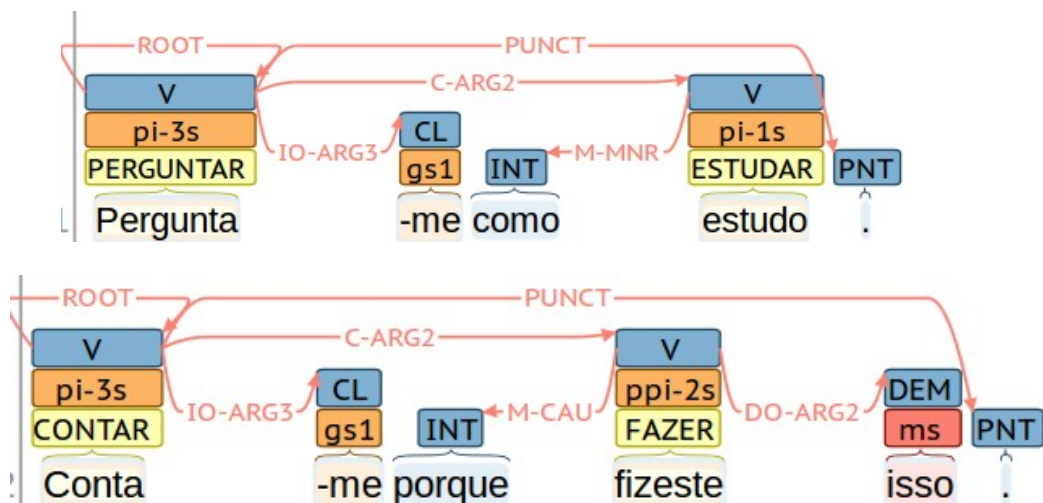


- A specifier:



7.3.1 indirect interrogatives

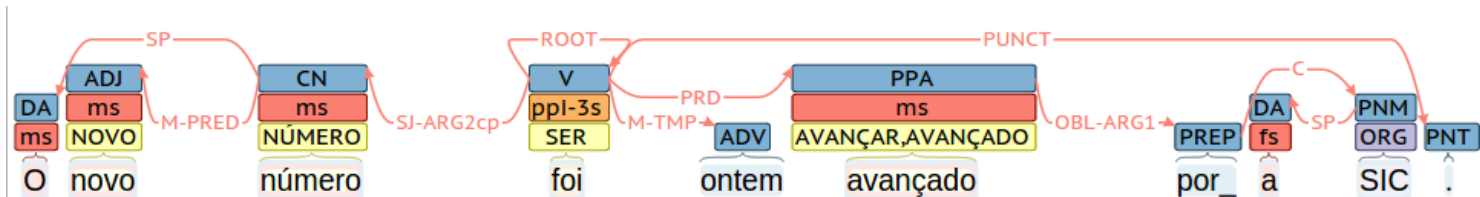
Partial interrogatives can also appear in a completive clause. The annotation has the following rendering in terms of dependency relations:



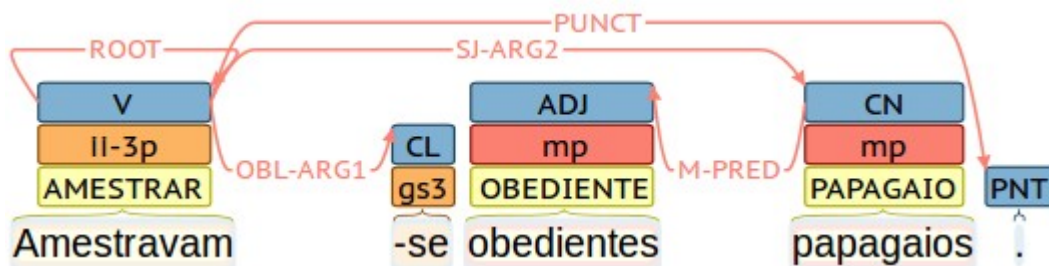
8- Valency alternations

8.1 passives

The *by*-phrase is an OBL-ARG1 and the Subject relation is tagged with SJ-ARG2cp, once there is a complex predicate.

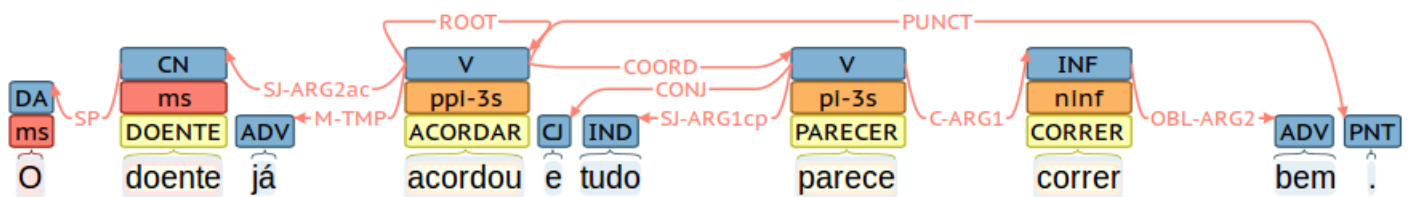


In passives with *se* there is no *by*-phrase: the OBL-ARG1 relation is established between the verb and the clitic *se*. The Subject is annotated with SJ-ARG2, since there is not a complex predicate.



8.2 anticausatives

The Subject of an anticausative is SJ-ARG2ac.



When the sentence is in the transitive form of the verb, the Subject is SJ-ARG1.

9- Tokenization

9.1 sentence splitting

Sentences are split at the expected points. It is worth of mention the case of utterances involving colon ":", which will be split into two separate entry sentences in the treebank, one preceding the colon and another following it.

9.2 non verbal utterances

Titles of newspaper articles, stretches around colons, etc. are cases of possible utterances in the corpus which are not structured around a corresponding verbal governor.

9.3 contractions

Contractions are expanded. The first element of an expanded contraction is marked with an "_" (underscore) symbol, for instance $do \rightarrow |de_o|$.

9.4 clitics

Clitics are detached from the verb. The detached clitic is marked with a "-" (hyphen) symbol, as for instance *dá-se-lho* → |*dá*|-*se*|-*lhe*|-*o*|.

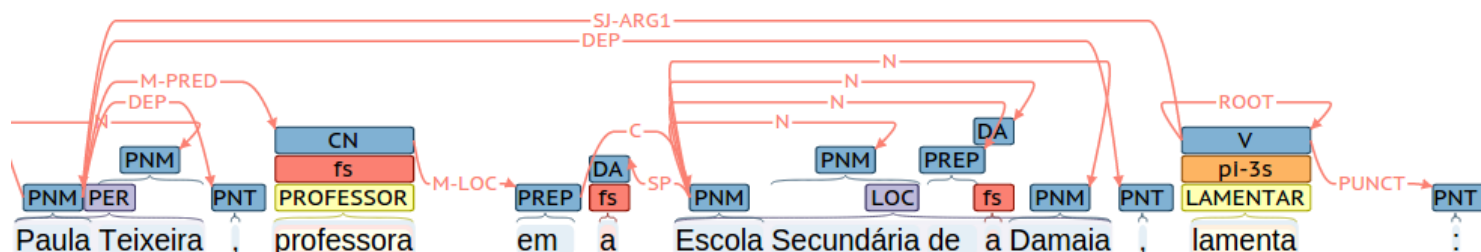
When in mesoclisys, a "-CL-" mark is used to signal the original position of the detached clitic: *afirmar-se-ia* → |*afirmar-CL-ia*| -se|.

Possible vocalic alterations of the verb form are marked with "#" (hash) symbol, as for instance in *vê-las* → |*vê#*|-*las*|.

10- Multi-word expressions

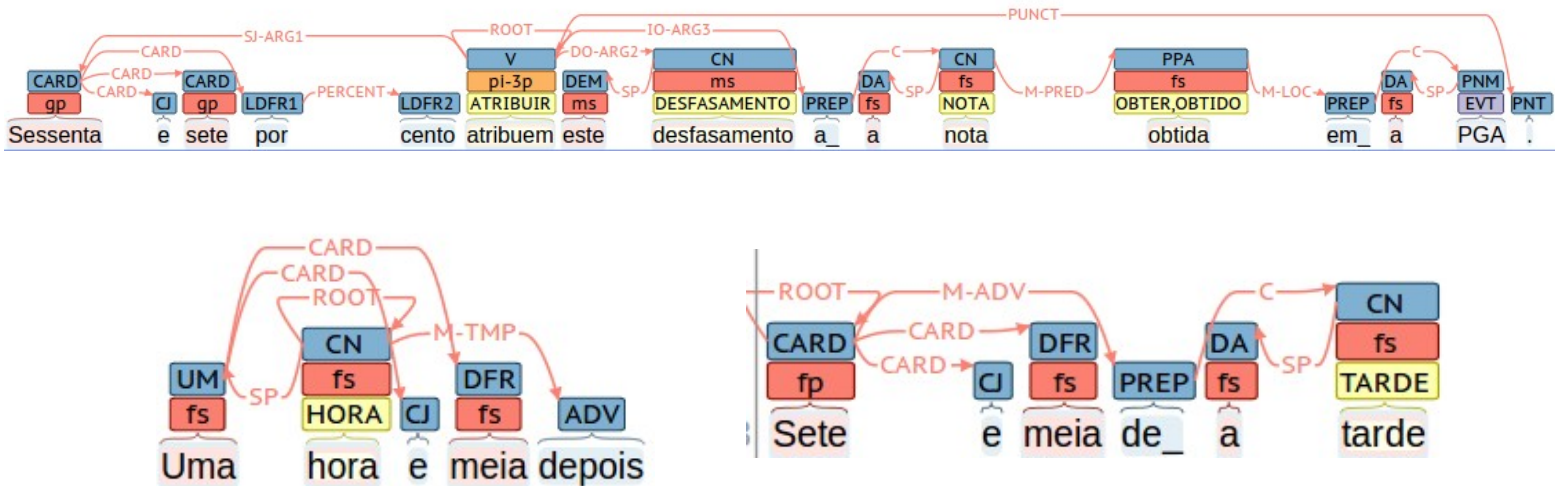
10.1 proper names

The first element (left to right) of a multi-word proper name is dependent on the governor subcategorizing for that proper name. The remaining lexemes of the multi-word proper name are dependent on that first element under dependency relations tagged with N.



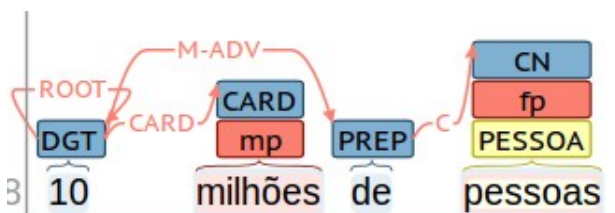
10.2 cardinals

Complex cardinals have a graph representation like a multi-word named-entity, whose arcs are labelled with CARD.



Important note:

“Milhões” is a CARD, not a MGT, and it is also annotated as a complex cardinal:



11- Textual marking

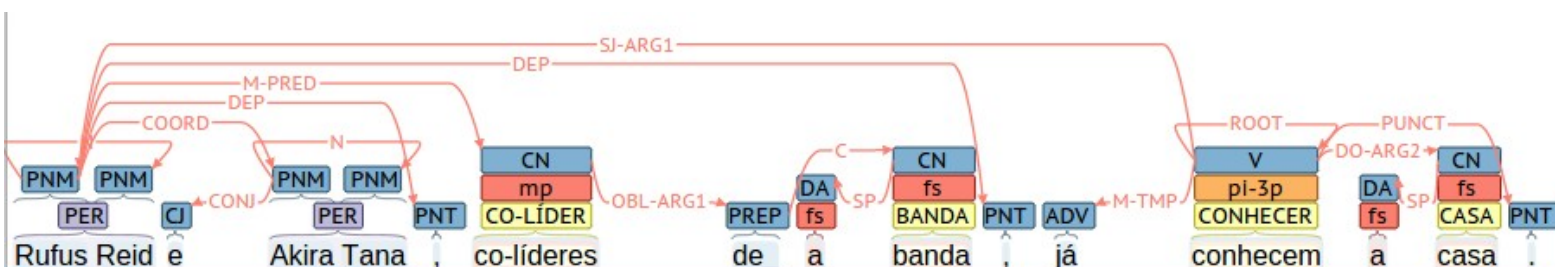
An end of sentence full stop is dependent under a dependency relation tagged with PUNCT. Every other textual mark is dependent under a dependency relation tagged with DEP.

11.1 punctuation

End of sentence markers are dependent from the main syntactic predicate of the utterance.

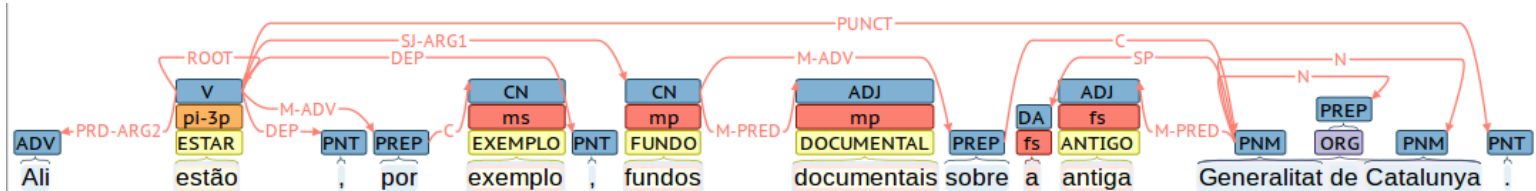
11.2 comma

Commas separating left periphery constituents are dependent on the head of these constituents. Commas surrounding appositions are dependent on the head of the NP being modified by the apposition.

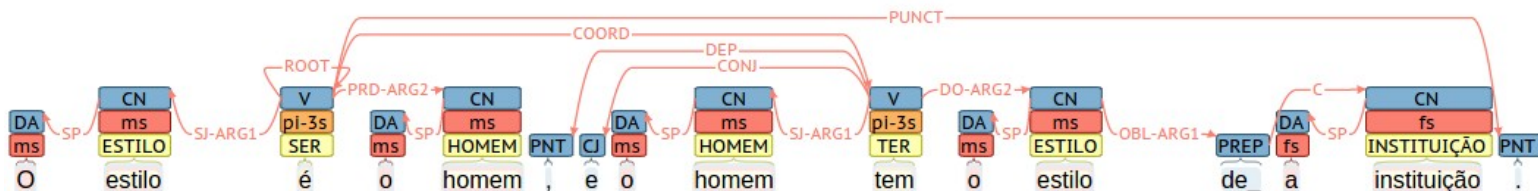


Commas with coordinative value are represented like lexical coordinative conjunctions: for details, see section 6.3 on Coordination.

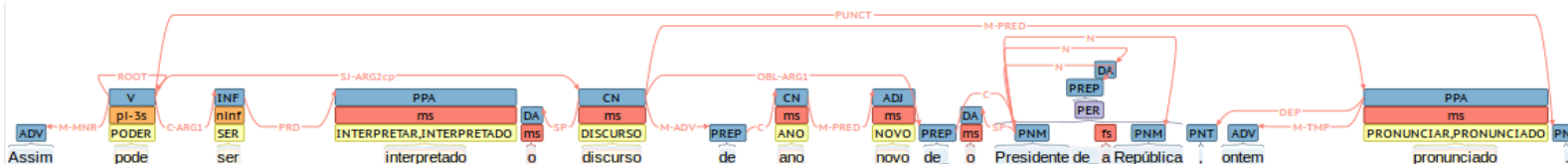
Commas surrounding parentheticals are dependent on the head of constituent being modified by the parenthetical.



A comma emphasizing a conjunction, thus immediately preceding it, is dependent on the same governor as that conjunction, but under the dependency tagged with DEP.

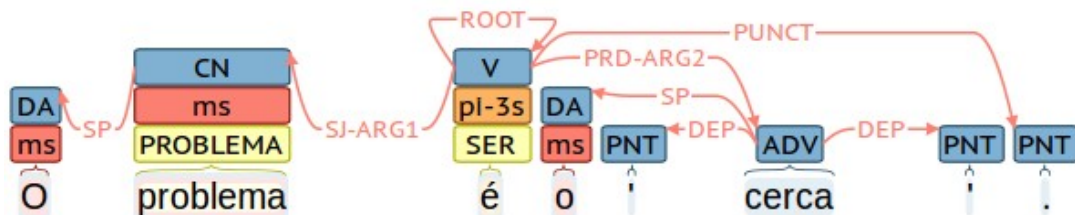


Other "pause" commas are dependent on the main predicate of the constituent immediately to its right.



11.3 quotation marks

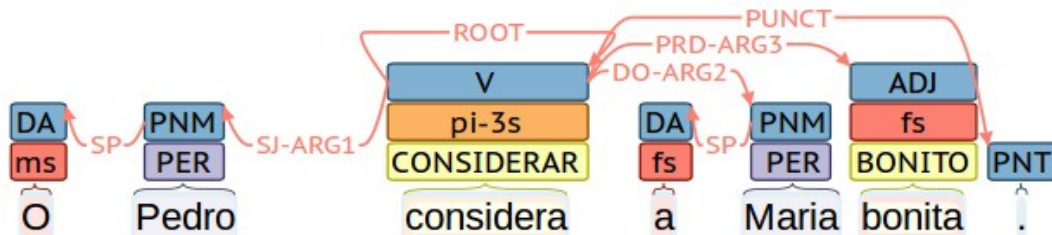
Quotation marks surrounding a constituent are dependent on the head of that constituent.



12- Other phenomena

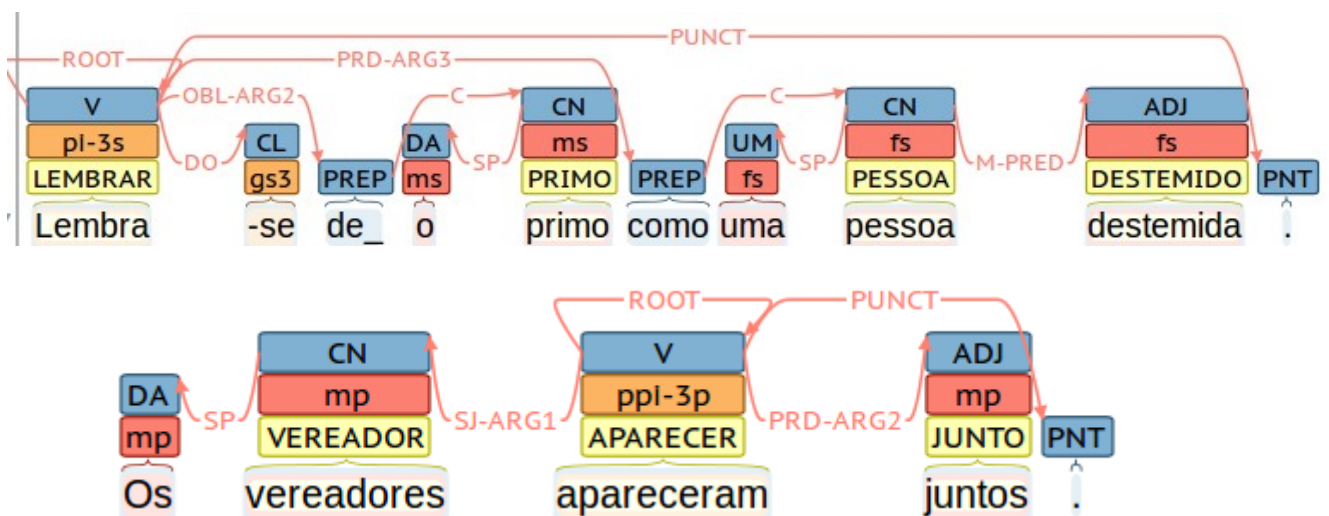
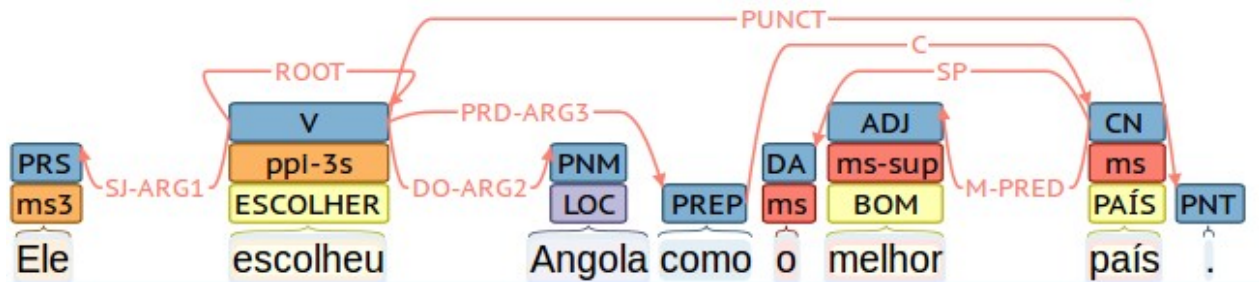
12.1 predicative over the object

The predicative over the object is dependent on the verb and is tagged as PRD-ARG3.



12.1.1 other predicatives

By analogy with the annotation of the predicative over the object, other predicative relations are tagged as PRD-ARG3, or even with PRD-ARG2.



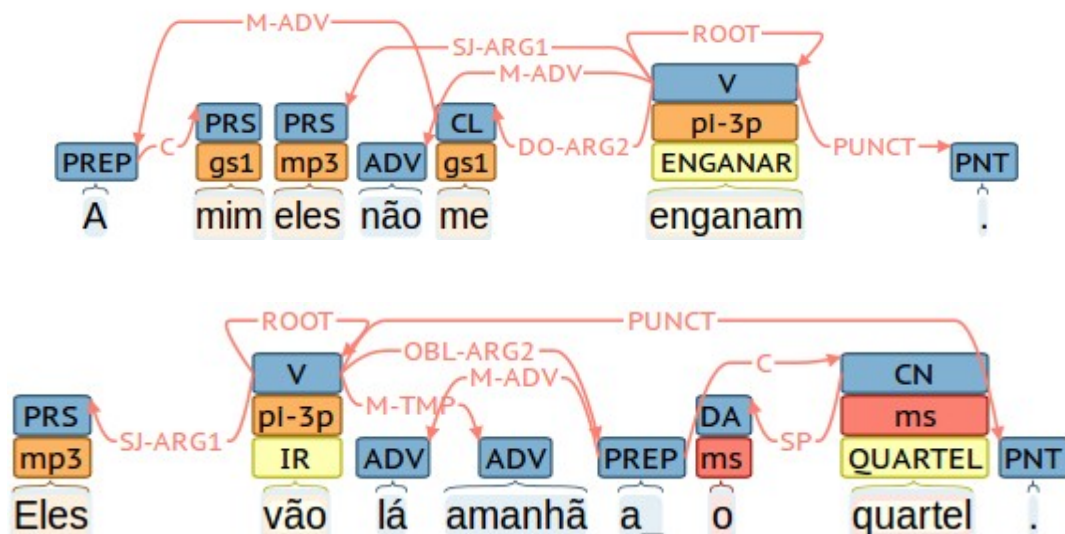
12.2 emphatic duplication

When the same position is filled with two elements, the relation with the verb is established with one of them:

i) the element which cannot be removed from the sentence without causing ungrammaticality (first example);

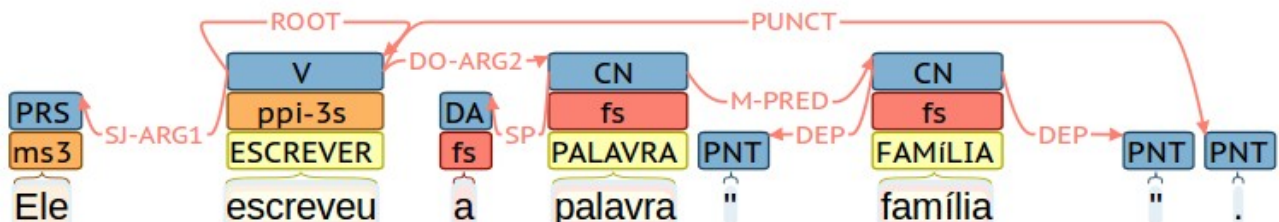
ii) the element which brings more information to the sentence (second example).

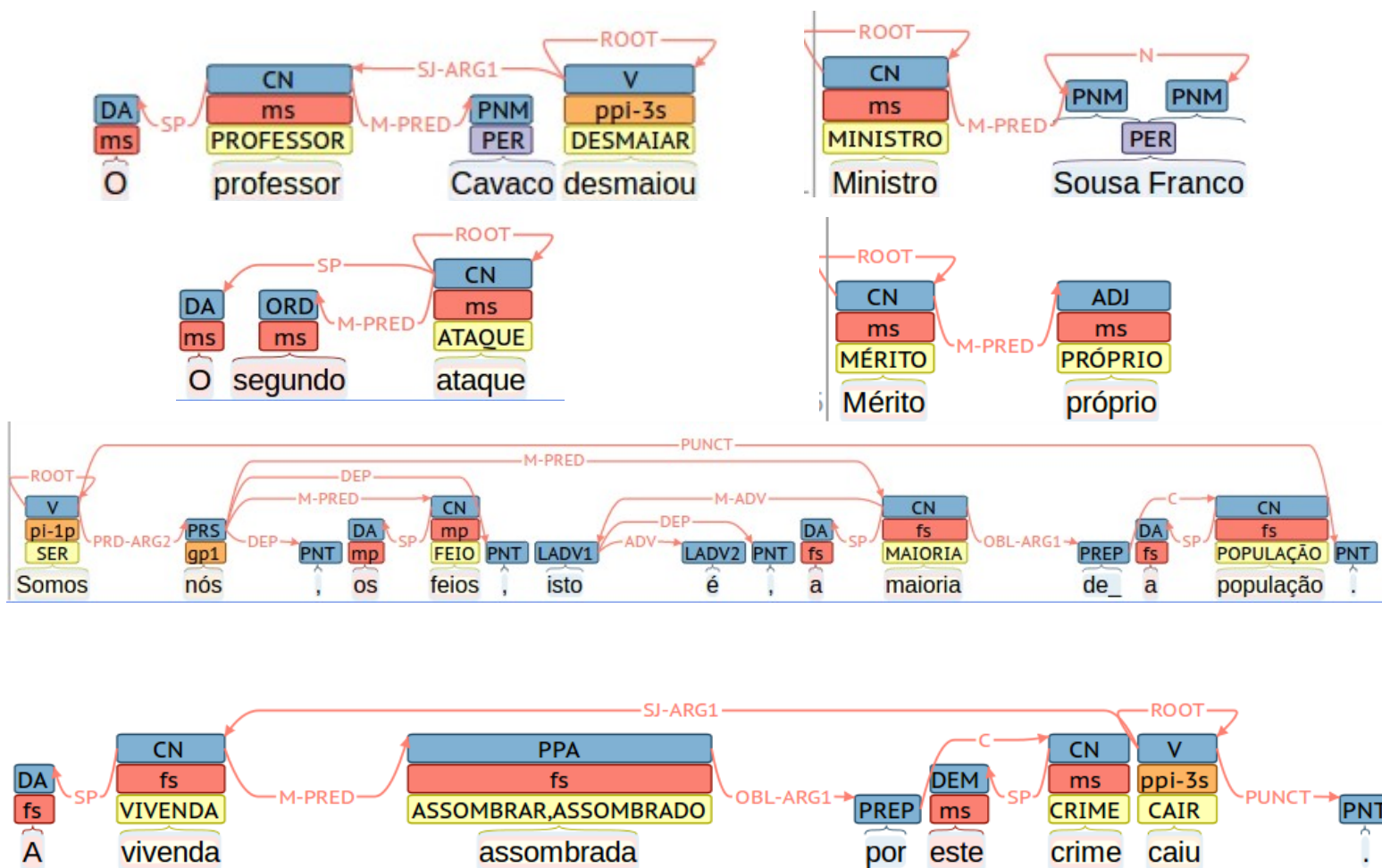
Among each other, the tag that decorates the arc is M-ADV.



12.3 notes on the M-PRED relation

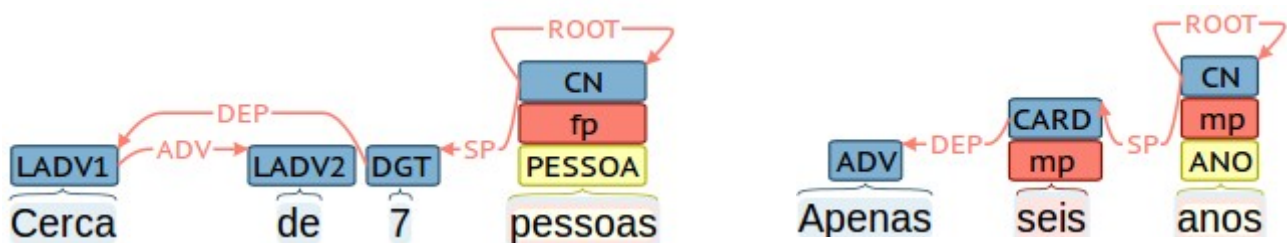
The M-PRED label is used in different cases, most of them related to adjectival modification or nouns in apposition. Below are some examples.

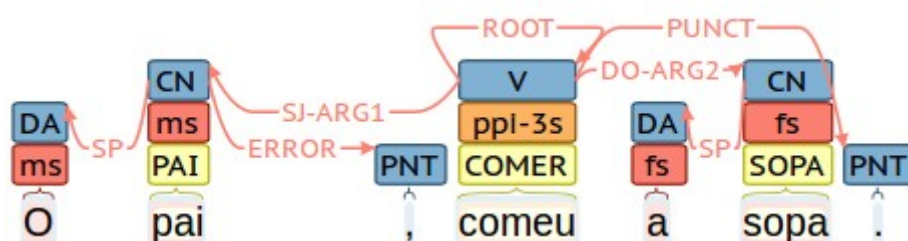
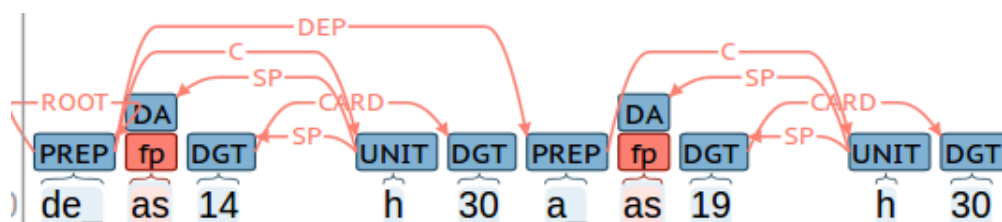


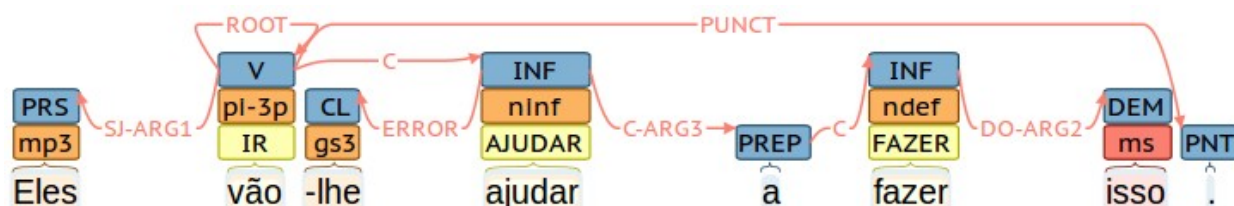


12.4 notes on the DEP relation

The DEP tag is a generic tag, referring to a generic dependency. Even so, there are some cases in which this label is used, like the case of commas (see section 11.2). Below are some other cases:







13- References

- Barreto, Florbela, António Branco, Eduardo Ferreira, Amália Mendes, Maria Fernanda Nascimento, Filipe Nunes and João Silva, 2006, "Open Resources and Tools for the Shallow Processing of Portuguese", *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC2006)*, Genoa, Italy.
- Branco, António and João Silva, 2006. *A Suite of Shallow Processing Tools for Portuguese: LX-Suite*. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL'06)*.
- Branco, António and Francisco Costa, 2008, *A Computational Grammar for Deep Linguistic Processing of Portuguese: LXGram, version A.4.1*, Technical Report, University of Lisbon, Department of Informatics.
- Branco, António, João Silva, Francisco Costa and Sérgio Castro, 2011, *CINTIL DepBank Handbook: Design options for the representation of grammatical dependencies*. Universidade de Lisboa, Faculdade de Ciências, Departamento de Informática, Technical Report TR 2011-3, <http://repositorio.ul.pt/handle/10451/14185>

Appendix

Some Cases Less Discussed in the Literature

