ParTUT
Annotation Guidelines

Last update:
March, 2014

1 Introduction

This is an extended version, for English and French, of the linguistic notes
developed for the Italian monolingual treebank created at the Department of
Computer Science of Turin (the *Turin University Treebank*, **TUT**).
Despite other parallel treebanks, where monolingual corpora were processed
independently with different tools, the texts of our collection were analyzed
from scratch with the same tool, the Turin University Lilnguistic Environment,
TULE. The annotation scheme is thus the same used within the TUT project,
whose main features are fully described in the project web page, at:

`http://www.di.unito.it/ tutreeb/#documents`

The following guidelines are meant to provide the appropriate representation for
those linguistic phenomena that distinguish French and English from Italian.

2 English

2.1 Modal auxiliaries: *do, will, shall, would*

Modal auxiliaries *do, will, would* and *shall* are attached to the main verb with
the generic label AUX (Figure 1).

3 Progressives

The auxiliary verb *be*, whenever used in progressive forms, is labbelled as
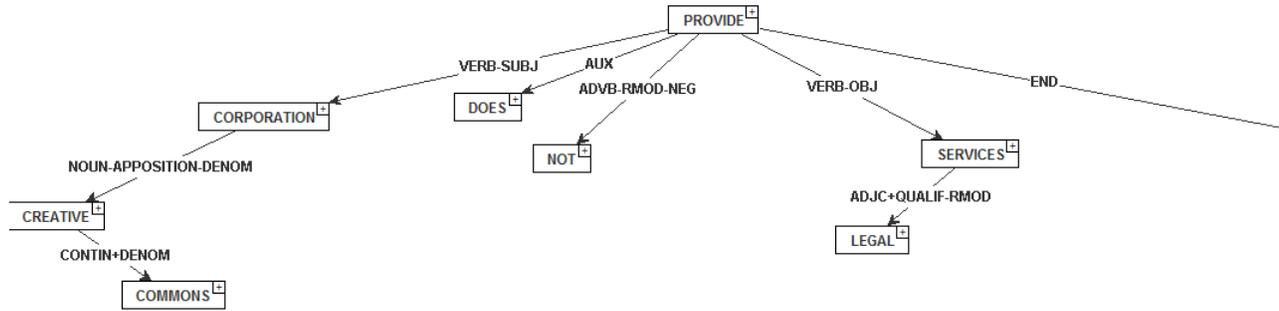AUX+PROGRESSIVE (Figure 2).

Figure 1: *"CREATIVE COMMONS CORPORATION DOES NOT PROVIDE LEGAL SERVICES."*
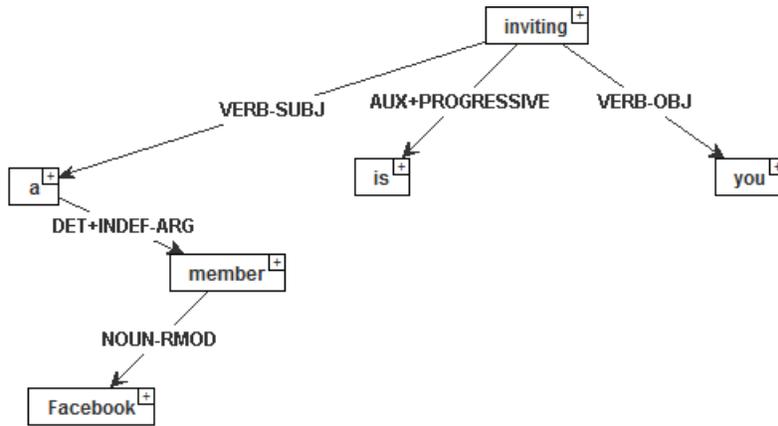

Figure 2: *"a Facebook member is inviting you"*

4 Double accusative constructions

Verbs such as *ask* or *give* may take two objects; the first one will then be labelled with both deep (INDOBJ) and surface (OBJ) roles, such as in Figure 3.
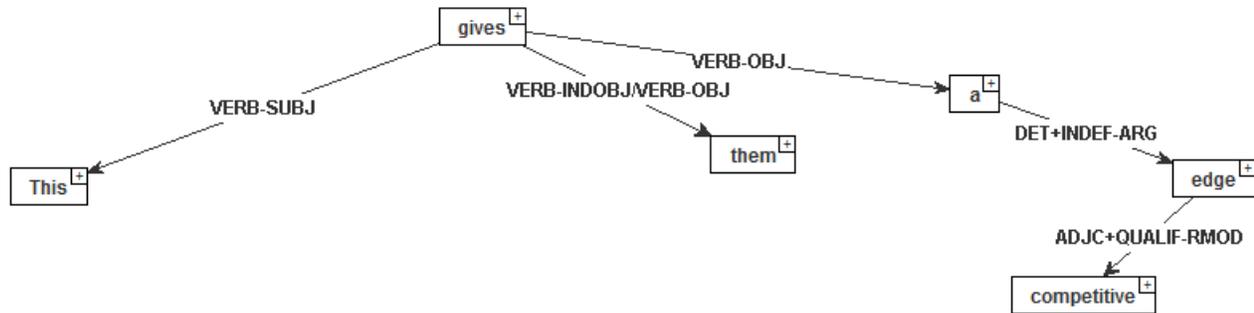

Figure 3: *"This gives them a competitive edge"*

The double role of these verbs' arguments is highlighted in their passive forms as well, where the surface subject is annotated with both INDOBJ and SUBJ relational labels, as in Figure 4.
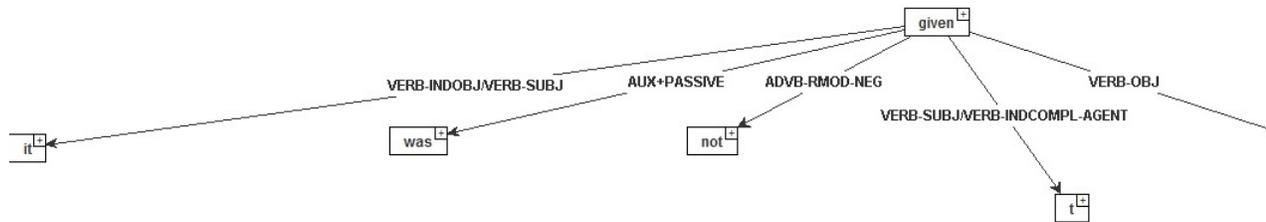

Figure 4: *"It was not given the opportunity to [...]"*

5 French

5.1 Verbal locutions: *il y a*

The verbal locution *il y a* is considered as a flexible locution and it is therefore annotated as such (see TUT linguistic notes for details on fixed and flexible locutions). As a result, each element is assigned its own syntactic role (verb, impersonal subject, modifier and direct object) together with the additional label * LOCUT. In case the expression *il y a lieu de + verb* is encountered, the preposition *de* which introduces the infinitive verb will be attached to the verb of the locution with the label VERB+INF-RMOD.
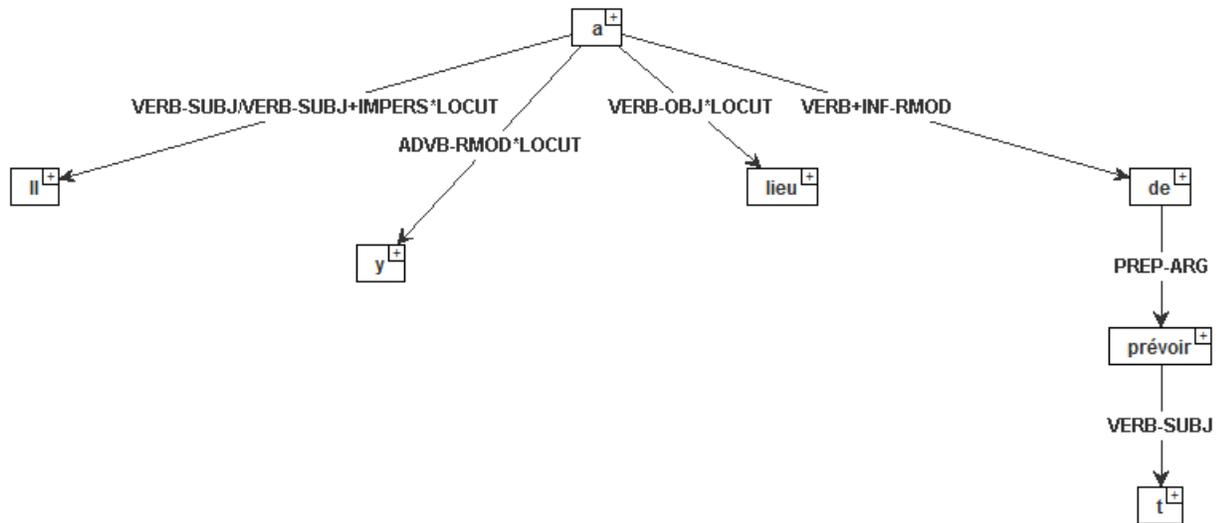
Figure 5: *"Il y a lieu de prévoir [...]"*

5.2 Superlatives

In Frenchsuperlative constructions, the definite article introduces not only the adjective (see the nominal group B in the example below), but also the noun to which it refers (group A). Therefore, in order to preserve the projectivity constraint, the first nominal group is attached to its head as VISITOR[1], while the second one is annotated according to the scheme already provided in TUT, but with the addition of a trace t that refers to the noun modified by the adjective.
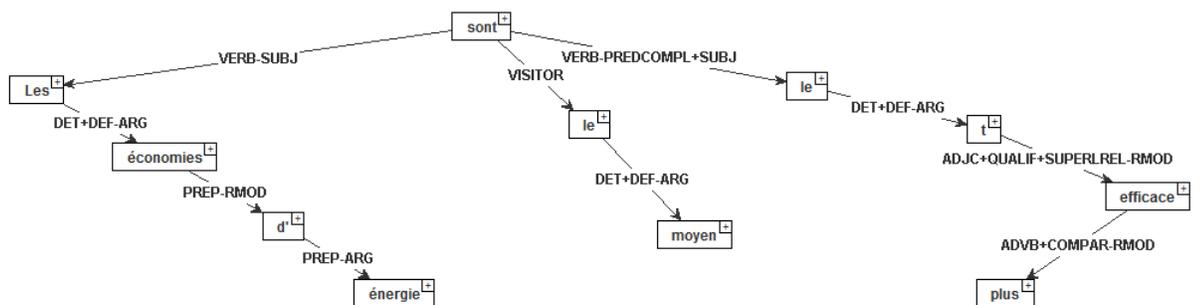

Figure 6: *"Les économies d'énergie sont (A) le moyen (B) le plus efficace..."*

---

[1]See TUT linguistic notes for details on the use of this particular label.

6   Common features: Expletives

As regards the treatment of cases where the element placed in the subject position plays a purely functional (not semantical) function: such element is therefore considered as an expletive subject. In those cases, the element with expletivefunction is labelled as VERB-EXPLETIVE/VERB-SUBJ, while the actual subject of the sentence is annotated as VERB-SUBJ/VERB -SUBJ+EXPLETIVE.
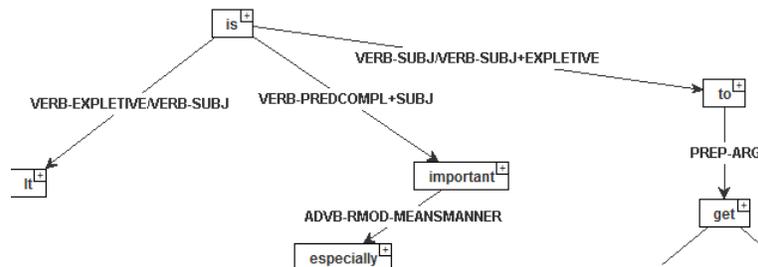
Figure 7: *"It is especially important to get [...]"*