

Guide d'annotation en dépendances profondes pour le français

Guide d'annotation en dépendances profondes pour le français	1
0 Introduction	2
0.1 Résumé des différences par rapport à la syntaxe de surface	2
0.1.1 Neutralisation des changements de diathèse	2
0.1.2 Suppression des marqueurs grammaticaux sémantiquement vides	3
0.1.3 Ajout de dépendances	3
1 Représentation syntaxique profonde	3
1.1 Le schéma d'annotation en dépendances de surface	4
1.2 Choix théoriques pour les dépendances profondes	4
1.2.1 Cadre de sous-catégorisation final	5
1.2.2 Cadre de sous-catégorisation canonique	5
1.2.3 Une double distinction : final/canonique versus surface/profond	6
1.2.4 Extension aux adjectifs	6
1.3 Caractéristiques formelles	7
1.4 Notations et représentation graphique	8
1.5 Synthèse des différences par rapport à la représentation syntaxique de surface	9
1.5.1 Court-circuitage des marqueurs grammaticaux sémantiquement vides	9
2 Catégories, traits et fonctions grammaticales	10
2.1 Nœuds	10
2.1.1 Traitement des mots composés et unités polylexicales	10
2.1.2 Mots considérés comme sémantiquement vides	11
2.1.3 Catégories morpho-syntaxiques	11
2.1.4 Traits profonds	13
2.1.5 Autres traits	14
2.2 Fonctions grammaticales	15
3 Changements de diathèse	19
3.1 Passif	20
3.2 Vrai réfléchi et réciproque	21
3.3 Causatif	22
3.3.1 Analyse des constructions en	23
3.4 Changement de diathèse impersonnel versus tournure impersonnelle	25
3.5 Se Moyen	28
3.6 Se Neutre	29
4 Dépendants du verbe	30
4.1 Auxiliaires	30
4.2 Clitiques	32
4.2.1 Clitique il	32
4.2.2 Clitique le	33
4.2.3 Clitique lui	34
4.2.4 Clitique en	35
4.2.5 Clitique y	38
4.2.6 Clitique se	39
4.3 La fonction suj	42
4.3.1 Réalisation locale ou non locale	42
4.3.2 Inversion du sujet	43
4.4 La fonction obj	44
4.4.1 Objets post-verbaux	44
4.4.2 Objets pré-verbaux	45
4.4.3 Objets directs non standards	46
4.5 Les constructions attributives et les fonctions attribut ats, ato	48
4.5.1 La fonction ats attribut du sujet	48

4.5.2 La fonction ato attribut de l'objet	50
4.6 La fonction a_obj	52
4.7 La fonction de_obj	54
4.8 Les fonctions p_obj*	56
4.8.1 La fonction p_obj.agt	56
4.8.2 La fonction p_obj.o	57
4.9 Les fonctions mod*	59
5 Dépendants du nom	62
5.1 Déterminants	62
5.1.1 Cas de base	62
5.1.2 Déterminant contracté avec une préposition	62
5.1.3 Adverbes de quantité ou comparatifs :	62
5.1.4 Déterminants cardinaux	63
5.2 Groupes prépositionnels	64
5.3 Epithètes et appositions	65
5.4 Relatives	65
5.4.1 Relatives avec antécédents	66
5.4.2 Relatives sans antécédent	70
6 Dépendants de la préposition	70
6.1 Préposition liées	72
6.2 Modificateurs de préposition	72
7 Dépendants de l'adjectif, de l'adverbe	73
7.1 Dépendants de l'adjectif	73
7.1.1 Sujet profond d'un adjectif	73
7.1.2 Compléments requis et adjoints d'un adjectif	74
7.2 Dépendants de l'adverbe	75
7.2.1 Distinction argument / ajout pour les dépendants prépositionnels	75
7.2.2 Adverbes de quantité	76
7.2.3 Adverbes comparatifs	77
8 Dépendants participiaux et adjectivaux	77
8.1 Participes et adjectifs épithètes	77
8.2 Participes et adjectifs dépendants de verbes	79
8.3 Participiales absolues	80
8.4 Adjectivales et participiales	81
8.5 Gérondif	83
9 Dépendants phrastiques et infinitivaux	84
9.1 Complémenteurs / prépositions sémantiquement pleins versus sémantiquement vides	84
9.2 Dépendants infinitivaux	85
9.2.1 Explicitation du sujet final des infinitifs	85
9.2.2 Explicitation de l'objet des infinitifs (tough movement)	91
9.2.3 Fonctions portées par les infinitives	92
9.3 Dépendants phrastiques	95
9.3.1 Phrases arguments ou ajouts de verbes	95
9.3.2 Dépendants phrastiques d'adjectifs	98
9.3.3 Dépendants phrastiques d'autres catégories	99
10 Dépendants pronominaux	99
11 Dépendants adverbiaux	100
11.1 Adverbe comme modificateur	100
11.1.1 Comparatif et superlatif	101
11.2 Adverbe sous-catégorisé	103
12 Clivées, dislocation et incises	104

12.1 Clivées	104
12.1.1 Clivée avec foyer sujet ou objet du V de la subordonnée	104
12.1.2 Clivée avec autre type de foyer	105
12.2 Dislocation	106
12.3 Incises	107
13 Coordination, énumération	109
13.1 Représentation en dépendances de surface	109
13.2 Représentation en dépendances profondes : principe général	110
13.3 Détail par catégories de conjoints	111
13.3.1 Coordination de syntagmes prépositionnels	112
13.3.2 Coordination de prépositions seules	112
13.3.3 Coordination d'adjectifs	112
13.3.4 Coordination de participes	113
13.3.5 Coordination de nominaux (communs, propres ou pronoms)	113
13.3.6 Coordination de verbes conjugués	115
13.3.7 Coordination d'infinitifs	117
13.3.8 Coordination de déterminants et/ou de cardinaux	117
13.4 Cas plus complexes	118
13.4.1 Coordination hétérogène	118
13.4.2 Coordination d'arguments agrégés (ou	118
13.4.3 Coordination à trou (ou	119
13.4.4 Zeugme	119
13.4.5 Coordination double	119
13.5 Énumération	120
14 Comparatives	121
15 Consécutives	123
16 Références	124

Marie Candito* et Guy Perrier^

avec la participation de Eric de la Clergerie*, Karen Fört^, Bruno Guillaume^, Corentin Ribeyre*, Djamé Seddah*

Equipes * Alpage et ^Sémagramme

0 Introduction

Ce guide décrit le schéma d'annotation en dépendances profondes utilisé pour l'annotation en syntaxe profonde du corpus Sequoia initialement annoté en syntaxe de surface (Candito et Seddah, 2012a). Cette annotation profonde est un projet commun aux équipes INRIA Alpage et Sémagramme. Le corpus résultant est librement disponible, téléchargeable et visualisable en ligne (<http://deep-sequoia.inria.fr/fr/>).

Contact : sequoia@inria.fr

Les corpus annotés en dépendances syntaxiques présentent un intérêt croissant par rapport aux corpus annotés en syntagmes dans la mesure où ils permettent plus directement d'extraire les relations prédicat-argument constitutives d'une représentation sémantique. Cette extraction reste cependant non triviale, la syntaxe offrant une grande variabilité dans la façon d'exprimer ces relations.

Nous proposons un niveau de représentation syntaxique dit **profond**, qui puisse être considéré comme une représentation intermédiaire avant une représentation sémantique.

Dans la représentation profonde, nous ne retenons que les mots sémantiquement pleins (faisant ainsi une distinction d'ordre sémantique) mais les dépendances entre ceux-ci restent étiquetées par des fonctions grammaticales (et pas par exemple par des rôles sémantiques ou une simple numérotation des actants sémantiques).

Il s'agit de neutraliser certaines variations syntaxiques. On limite les distinctions de représentation qui ne relèvent que de la sémantique (en particulier, la distribution de dépendances dans le cas de prédicats coordonnés n'est faite que si elle est sémantiquement pertinente).

0.1 Résumé des différences par rapport à la syntaxe de surface

On peut schématiquement distinguer trois explicitations principales par rapport à la syntaxe de surface:

0.1.1 Neutralisation des changements de diathèse

Les changements de diathèse sont neutralisés : les sous-catégorisations verbales observées en surface sont considérées comme le résultat de redistributions à partir de sous-catégorisations considérées comme **canoniques** et ce sont les fonctions grammaticales canoniques qui sont représentées au niveau profond. Les arguments sous-catégorisés portent ainsi une fonction grammaticale finale et une fonction grammaticale canonique, et les deux peuvent différer (voir [Section 1.2](#)).

0.1.2 Suppression des marqueurs grammaticaux sémantiquement vides

On explicite le statut sémantique (vide / plein) de chaque forme fléchie de la phrase. Seules les formes sémantiquement pleines sont conservées dans la représentation syntaxique profonde (voir section [2.1.2](#) pour la liste exhaustive).

- si le mot sémantiquement vide n'a pas de dépendants en surface, alors il est simplement ignoré (et la dépendance qui le lie en surface à son gouverneur n'est pas conservé au niveau profond. C'est par exemple le cas pour les auxiliaires de temps.
- si le mot sémantiquement vide a des dépendants en surface, comme par exemple l'objet d'une préposition régie, on dit alors qu'il est "**court-circuité**" : ses dépendants en surface sont reportés en profond sur son gouverneur. Par exemple pour une préposition régie à dans "*Anna parle à Paul*", on a en surface une dépendance entre "*parle*" et "*à*", et une dépendance entre "*à*" et "*Paul*", alors qu'au niveau profond on ne conserve qu'une dépendance directement entre "*parle*" et "*Paul*".

0.1.3 Ajout de dépendances

- la dépendance entre un verbe non conjugué et son sujet final (non exprimé localement au verbe) est ajoutée
 - par ex. le sujet des infinitifs (cf. [Section 9.1.1](#)) : *Paul veut être photographié avec le maire.* => ajout "Paul" comme sujet final et objet canonique de "photographié"
 - ou le sujet des participes épithètes (section [Section 8.1](#)) :
 - *les personnes interrogées par la police* => ajout de "personnes" comme sujet final et objet canonique de "interrogées"
 - *les personnes ayant un badge* => ajout de "personnes" comme sujet final et canonique de "ayant"
- par extension, on considère le sujet d'un adjectif (l'élément que l'adjectif employé comme épithète modifie), et la dépendance entre l'adjectif et son sujet est explicitée
 - en épithète : *les longues phrases* => ajout de "phrases" comme sujet de "longues"
 - en attribut du sujet ou de l'objet (cf. [Section 4.5](#)), on fait le même ajout de dépendance pour
 - *ses phrases sont longues*
 - *Je trouve ses phrases très longues*
- dans le cas de dépendants d'éléments coordonnés, alors que le schéma de surface prévoit une dépendance sur le premier conjoint uniquement, la dépendance est distribuée sur les autres conjoints si cela est sémantiquement justifié
 - par ex. pour le sujet partagé par 2 verbes coordonnés comme dans *Anna a conçu le projet et a trouvé des investisseurs*, on n'a en surface que le lien entre le premier verbe et son sujet (entre *conçu* et *Anna*), on a en outre en profond le lien entre le 2eme verbe (*trouvé*)
 - par ex. pour l'objet partagé par 2 prépositions coordonnées : pour "*avec et sans paracétamol*", on a en surface "paracétamol" objet de la préposition "avec", et on ajoute en profond "paracétamol" objet de la préposition "sans"

1 Représentation syntaxique profonde

Nous avons défini notre schéma d'annotation en syntaxe profonde en partant du schéma d'annotation en dépendances de surface du corpus Sequoia, ce qui a eu une influence certaine sur le résultat. Nous commençons donc par détailler comment le schéma de surface a été lui-même défini, avant de détailler pour la représentation syntaxique profonde, les choix théoriques réalisés, ses caractéristiques formelles, les notations utilisées. Nous terminons cette section par une synthèse des

informations explicitées au niveau de la syntaxe profonde.

1.1 Le schéma d'annotation en dépendances de surface

- le corpus Sequoia a été annoté d'abord en arbres de constituants, **en suivant très largement le schéma d'annotation du French Treebank (ci-après FTB) (Abeillé et Barrier, 2004), en utilisant les guides d'annotation disponibles (Abeillé, 2004; Abeillé et al. 2004).**
- Le corpus en constituants a ensuite été converti automatiquement en dépendances de surface, en suivant la procédure décrite dans (Candito et al., 2010). Il s'agit d'une procédure fondée sur la définition pour chaque règle de réécriture syntagmatique $XP \rightarrow X_1 X_2 \dots X_n$ d'un élément tête dans la partie droite de la règle. Lors de la conversion en dépendances les têtes lexicales des constituants X_i non têtes sont rattachés à la tête lexicale du constituant tête. Par exemple, pour la règle $NP \rightarrow DET ADJ N PP$, l'élément tête dans la partie droite est le noeud N . Les têtes lexicales du DET , de l' ADJ et du PP sont ajoutées comme dépendant de la tête lexicale du N . Les labels de dépendance utilisés sont
 - repris des arbres en constituants, dans le cas où les fonctions grammaticales sont présentes. En effet le schéma en constituants du FTB prévoit pour les dépendants de verbes conjugués de noter les fonctions. Ainsi on aura une règle de la forme $Sint \rightarrow NP-SUJ VN NP-OBJ PP-A_OBJ$.
 - ou bien prédits par heuristiques dans tous les autres cas
- Il est important de noter que les arbres de dépendances résultants suivent ainsi très largement les choix linguistiques du FTB (Abeillé et al., 2004; Abeillé, 2004), dans la mesure où la majorité des phénomènes syntaxiques sont mécaniquement traduits en dépendances. Des informations supplémentaires sont cependant prédites par la procédure de conversion, en cas de sous-spécification dans la version en constituants (c'est le cas pour les étiquettes de dépendances pour les dépendants de gouverneurs non verbaux). (cf. le guide Candito, Crabbé et Falco, <http://alpage.inria.fr/statgram/frdep/Publications/FTB-GuideDepSurface.pdf>)
- Enfin, les dépendances longue distance ont été corrigées manuellement (Candito et Seddah, 2012a) dans les arbres de dépendance obtenus par conversion automatique, ce qui a introduit quelques arcs non projectifs. Par exemple, à la conversion de l'arbre syntagmatique de « ... le succès que la municipalité était en droit d'attendre », le pronom relatif objet que est mécaniquement rattaché à tort comme dépendant de était, et manuellement corrigé pour dépendre de attendre.

Le schéma d'annotation de surface résultant est notre point de départ. Nous avons cherché, pour des raisons pragmatiques, à minimiser les divergences entre les niveaux surfacique et profond, pour nous concentrer sur les phénomènes non directement représentables dans les arbres de surface. Aussi, par exemple, avons-nous conservé la représentation des coordinations avec le premier conjoint comme tête. La distinction entre les dépendants partagés par plusieurs conjoints et les dépendants du seul premier conjoint est faite en dépendances profondes, en "distribuant" (répétant) les arcs de dépendances sur les différents conjoints (voir [Section 13](#)).

1.2 Choix théoriques pour les dépendances profondes

L'objectif principal de nos représentations syntaxiques profondes (**REPRSINTPROF** dans la suite) est de généraliser sur la variation syntaxique autant que possible sans faire de distinctions ni de généralisations purement sémantiques.

Nous utilisons pour cela la notion de *sous-catégorisation canonique* et représentons les changements de diathèse comme des redistributions des fonctions grammaticales sous-catégorisées par un lexème. Nous inspirant de la Grammaire Relationnelle (Perlmutter, 1983), nous distinguons **fonction grammaticale canonique** et **fonction grammaticale finale**

d'une part, et **cadre de sous-catégorisation canonique** (CS canonique) et **cadre de sous-catégorisation final** (CS final). [footnote: La grammaire relationnelle utilise les termes de fonction *initiale* et *finale*].

1.2.1 Cadre de sous-catégorisation final

Définissons d'abord la notion de cadre de sous-catégorisation final pour un verbe : elle contient

- d'une part les fonctions observées associées aux arguments *exprimés* du verbe,
- et d'autre part, dans le cas d'ellipse et/ou de verbes non conjugués, les fonctions des éléments qui seraient des arguments du verbe si celui-ci était conjugué et utilisé sans ellipse.

Cette formulation permet de faire entrer dans le CS final par exemple le sujet (final) des infinitifs, le sujet (final) des participes épithètes, le sujet de verbes coordonnés ou plus généralement tout argument partagé par plusieurs prédicats. Par exemple, dans « *Anna veut dormir, mais devra peut-être veiller* », le CS final de *dormir* est [sujet] rempli par *Anna*, et le CS final de *devra* est [sujet, objet], rempli par *Anna* et *veiller*.

Entrent également dans le CS final l'élément modifié par un participe épithète : par exemple pour « *les personnes nées en 40* », le nom *personnes* est la tête du sujet final de *nées*.

1.2.2 Cadre de sous-catégorisation canonique

Passons maintenant à la définition précise de la sous-catégorisation canonique. Afin de neutraliser la variation syntaxique due aux changements de diathèse, nous considérons ceux-ci comme des redistributions des fonctions canoniques associées aux arguments syntaxiques.

Suivant la Grammaire Relationnelle (Perlmutter, 1983), le CS final est vu comme résultant de l'application de 0 à n redistributions sur un CS canonique. Etant donnée une occurrence de verbe, le CS canonique peut donc par définition être obtenu par application inverse des redistributions appropriées. Un exemple simple est le cas d'un verbe au passif dont le CS final est [SUJET, PAR-OBJET] et le CS canonique est [SUJET, OBJET].

Redistributions

Pour définir nos REPRSYNTPROF, nous n'avons considéré que les redistributions qui comportent un marquage morpho-syntaxique (typiquement l'auxiliaire pour le passif, ou le clitique sémantiquement vide *se* pour les alternances moyennes et neutres). Les alternances syntaxiques sans marquage morpho-syntaxique ne sont pas capturées au sein de nos REPRSYNTPROF, et donnent lieu à des CS canoniques différents. En effet, repérer ces alternances, en l'absence de marquage formel, relève pour nous de l'analyse sémantique. C'est le cas par exemple avec une alternance causative/inchative comme *X coule Y / Y coule* : on considère la sous-cat canonique [SUJET, OBJET] pour *X coule Y*, et la sous-cat canonique [SUJET] pour *Y coule* : l'argument sémantique « l'entité qui coule » est donc tantôt sujet canonique, tantôt objet canonique.

En revanche, pour l'alternance moyenne (par exemple dans « *On avale facilement ce médicament / Ce médicament s'avale facilement* ») ou bien l'alternance neutre (qui "efface" l'actant agentif ou causal, comme dans « *Cela dissout le médicament / Le médicament se dissout (de lui-même)* »), qui sont toutes deux marquées par le clitique "se", le lien entre les deux formulations est capturé par redistribution, et pour ces deux alternances, l'objet direct dans la version transitive (*médicament*) est, dans la version intransitive, le sujet final mais l'objet canonique.

Nous retenons comme redistributions : le passif, l'impersonnel, le moyen, le neutre et le causatif ([Section 3](#)), certaines pouvant interagir. Nous renvoyons à (Candito, 1999) pour une étude des interactions entre redistributions pour le français.

Arguments syntaxiques canoniques versus arguments sémantiques

On rajoute la contrainte que les éléments du CS canonique, appelés **arguments syntaxiques canoniques**, sont obligatoirement sémantiquement pleins. Typiquement le *il* explétif n'appartient qu'au CS final. Donc dans l'exemple « *Trois personnes arrivent* », à partir du CS canonique [SUJET], la redistribution de l'impersonnel rétrograde le sujet en objet direct, et un *il* explétif remplit la fonction sujet final : on considère que dans « *Il arrive trois personnes* » le verbe a pour CS canonique [SUJET] et pour CS final [SUJET,OBJET].

Mais les arguments syntaxiques canoniques, même s'ils sont sémantiquement pleins, ne sont pas forcément des arguments

sémantiques de leur gouverneur, ce qui constitue une différence avec une représentation sémantique.

Ainsi par exemple, on conserve dans nos REPRSYNTPROF les sujets des verbes à montée : dans "Paul semble dormir" on a en représentation profonde la sous-catégorisation [SU], [OBJ] pour *sembler* (remplie par *Paul* et *dormir*), alors que sémantiquement *sembler* n'a qu'un seul argument. De la même manière, les attributs restent rattachés à leur verbe (même si on explicite le lien prédicatif entre le sujet et l'attribut du sujet / entre l'objet et l'attribut de l'objet). On a également pour les constructions avec adjectifs à "tough movement" une divergence entre syntaxe profonde et sémantique (cf. [Section 9.1.2](#)).

1.2.3 Une double distinction : final/canonique versus surface/profond

Nous avons volontairement fait deux distinctions différentes : représentation profonde *versus* de surface, et fonction grammaticale finale *versus* canonique : le point ici est que l'**on évite de parler de « fonction de surface » versus « fonction profonde »**, car les changements de diathèse peuvent interférer avec le caractère profond / surfacique d'une dépendance.

Cela nous est utile pour capturer certaines régularités syntaxiques, comme par exemple celle concernant le contrôle des sujets des infinitifs : pour un infinitif, quelle que soit sa diathèse, c'est toujours le sujet *final* de l'infinitif qui est non réalisé localement à l'infinitif. Ainsi, dans (1) *Paul veut embaucher*, on veut capturer en dépendances profondes que *Paul* est sujet de *embaucher*. Et pour (2) *Paul veut être embauché*, idem, on a la régularité que *Paul* est sujet de (*être*) *embauché*, pour peu que l'on précise bien qu'il s'agit du sujet final. En interagissant avec le passif, on obtient de manière régulière que dans (2) le sujet final de *être embauché* est son objet canonique.

Or la relation entre un infinitif et son sujet (final) est l'exemple par excellence d'une dépendance *profonde*, en ce sens qu'elle ne peut pas être représentée dans un arbre de dépendances de surface (cf. ce sujet a alors 2 gouverneurs). Pour éviter de parler de fonction de surface pour une dépendance uniquement profonde, on utilise donc plutôt le vocabulaire fonction finale *versus* fonction canonique.

A noter que l'on retrouve la même régularité pour tous les verbes non conjugués. Ainsi pour un participe épithète, le nom modifié par le participe en est systématiquement son sujet final, pour peu que l'on traite les participes passés de transitifs comme des passifs. On a alors les cas :

- participe présent actif : *les enfants demeurant côté pair de la rue sont sectorisés au collège Bellevue* => "enfants" est sujet final (et canonique) de "demeurant"
- participe présent passif : *les enfants étant déjà inscrits au CP* => "enfants" est sujet final (et objet canonique) de "inscrits"
- participe passé transitif (considéré comme un passif) : *les enfants déjà inscrits au CP* => idem "enfants" est sujet final (et objet canonique) de "inscrits"
- participe passé intransitif (considéré comme un actif) : *les personnes nées en 62 ont aujourd'hui ...* => "personnes" est sujet final (et canonique) de "nées"

A l'inverse, certaines régularités d'ordre sémantique se trouvent naturellement représentées en utilisant les fonctions canoniques et pas finales. C'est le cas de la propriété **lexicale** pour un verbe à contrôle de définir quel est le dépendant contrôleur de l'infinitive (quel est le dépendant qui est le sujet final de l'infinitive). Par exemple, avec le verbe *contraindre*, c'est l'argument sémantique "celui qui est contraint" qui est le sujet final de l'infinitive, cet argument étant l'objet canonique de *contraindre*. Ainsi, quelle que soit la diathèse de *contraindre*, c'est son objet canonique qui est le sujet *final* de l'infinitive.

Paul est contraint à partir

Le vote contraint Paul à partir

1.2.4 Extension aux adjectifs

Outre les verbes, nous traitons dans nos annotations les adjectifs, auxquels la notion de sous-catégorisation peut être

étendue, mais sans distinction entre final et canonique (la sous-cat finale est toujours identique à la sous-cat canonique).

Pour un adjectif, on définit son argument syntaxique profond premier comme l'élément modifié par l'adjectif lorsque celui-ci est épithète. On parle alors de « sujet de l'adjectif ». Cette dénomination emprunte au vocabulaire syntaxique, pour faire une analogie avec les prédicats verbaux, bien que ce « sujet » ne soit jamais en surface un dépendant syntaxique de l'adjectif.

Donc par exemple pour l'adjectif *content*, on considère la sous-cat canonique [SUJET, DE_OBJ], et dans *Paul est content de toi*, le sujet canonique est "Paul", et le DE_OBJ est "toi".

A ce jour, le travail reste à faire pour les autres catégories de prédicats, en particulier les noms.

1.3 Caractéristiques formelles

Nous définissons une représentation *complète* comme un graphe de dépendances contenant à la fois la représentation de surface et la représentation profonde.

Les nœuds sont les mots de la phrase (ou des composants de composés réguliers), et sont typés comme sémantiquement vides (en rouge dans la figure) ou pleins (en noir) [footnote : une version avec ajout de nœuds non réalisés pour gérer certaines ellipses de prédicat est envisagée mais pas encore implémentée]

L'ordre linéaire de la phrase est encodé par un identifiant sur les nœuds, de type entier, local à chaque phrase (par convention le premier mot de chaque phrase reçoit l'identifiant 1).

Les arcs orientés entre les nœuds portent :

- une information sur leur appartenance à la REPRSYNTSURF ou pas, et leur appartenance à la REPRSYNTPROF ou pas : un arc peut être surfacique mais non profond (arcs rouges), profond mais non surfacique (arcs bleus), et à la fois profond et surfacique (arcs noirs) ;
- une étiquette qui est constituée soit d'une seule fonction, à la fois finale et canonique, pour les fonctions n'intervenant jamais dans des changements de diathèse (comme la fonction **mod**), soit de la fonction finale et de la fonction canonique (dans toute la suite, une étiquette notée "**fff:ccc**" correspond à la fonction finale **fff** et la fonction canonique **ccc**).

La REPRSYNTPROF pour une phrase donnée est formée des nœuds sémantiquement pleins et des arcs profonds, qu'ils soient par ailleurs aussi présents en surface (arcs noirs) ou pas (arcs bleus). En outre dans les étiquettes des dépendances, seules les fonctions canoniques sont conservées en représentation profonde.

Formellement, alors que la représentation en dépendances de surface est un arbre, la REPRSYNTPROF est un graphe orienté, cf. un même élément peut avoir plusieurs gouverneurs.

Le graphe peut contenir des cycles. En outre le graphe peut contenir des arcs multiples (un même couple gouverneur / dépendant peut être relié par plusieurs arcs de même orientation mais d'étiquettes différentes), ce qui apparaît essentiellement dans la gestion du réfléchi (cf. [Section 4.2.6](#)).

NB : étant donné le caractère intermédiaire de la représentation, entre syntaxe de surface et représentation sémantique, on obtient des dépendances de statuts assez différents. En particulier, on conserve toutes les dépendances de type modifieur, y compris dans le cas où on explicite la dépendance inverse (donc par exemple pour *les longues phrases*, on conserve **mod**(*phrases, longues*) et on ajoute **suj**(*longues, phrases*)). C'est la source principale de cycles dans la REPRSYNTPROF, avec également le court-circuitage des pronoms relatifs dans les relatives épithètes.

NB: Le schéma d'annotation n'impose pas la projectivité, i.e. la projection d'un nœud (l'ensemble formé du nœud lui-même et des nœuds qui en dépendent directement ou indirectement) peut correspondre à un segment discontinu de la phrase.

1.4 Notations et représentation graphique

On utilise dans les exemples soit une représentation graphique soit, plus rarement, une notation en triplet **étiquette**(gouverneur, dépendant)

Dans le cas des fonctions qui n'entrent jamais dans des changements de diathèse, **étiquette** correspond à une fonction simple.

Pour les fonctions entrant dans des changements de diathèse (**suj**, **obj**, **a_obj**, **de_obj**, **p_obj.agt**), l'étiquette a deux parties séparées par « : » **fct_finale:fct_canonique** (*gouverneur, dépendant*).

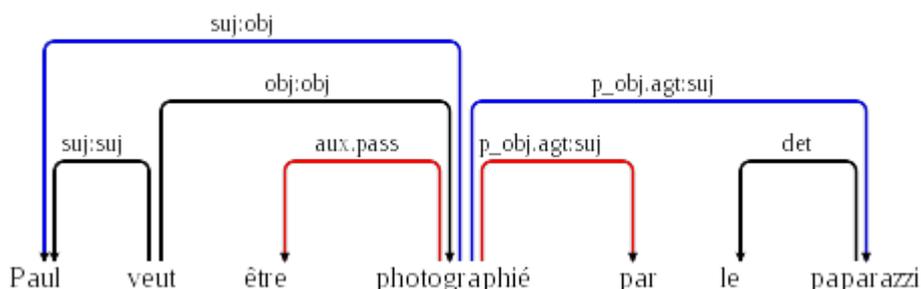
Le statut dépendance profonde / dépendance de surface est indiqué via une couleur :

- dépendance uniquement de surface en **rouge** (à ignorer pour la représentation profonde)
- dépendance surfacique et profonde en noir
- dépendance uniquement profonde en **bleu**

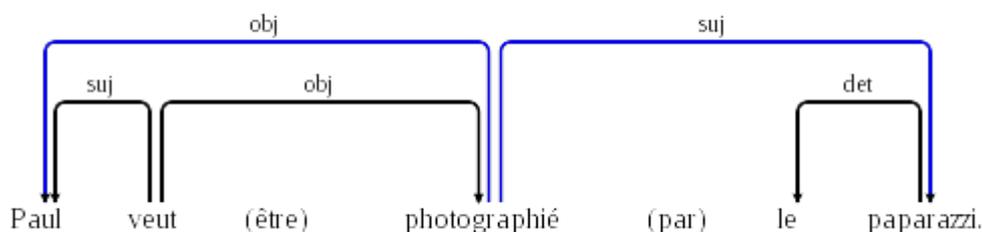
En outre, dans la représentation graphique, les mots sémantiquement vides (n'appartenant pas à la représentation profonde) sont en rouge.

Exemple représentation complète (i.e. surfacique + profonde) :

1. *Paul veut être photographié par le paparazzi*



La représentation syntaxique profonde s'obtient en ne conservant que les nœuds noirs, les arcs noirs et les arcs bleus, et en ne conservant que les fonctions canoniques en cas de double étiquette **fct_finale:fct_canonique**. Cela donne pour l'exemple 1 :



Le présent guide est centré sur la représentation syntaxique profonde, mais celle-ci étant définie en partant d'un schéma de représentation en dépendances de surface (cf. supra), une bonne part des exemples est fournie en donnant à la fois la représentation de surface et la représentation profonde (cf. supra "la représentation complète"). En outre, on fournit en plus des fonctions canoniques les fonctions finales lorsque c'est pertinent (bien que les fonctions finales n'appartiennent pas stricto sensu à la représentation profonde).

1.5 Synthèse des différences par rapport à la représentation syntaxique de surface

La représentation syntaxique profonde est définie par le fait de rendre explicite les informations suivantes :

- **statut sémantique** : on explicite le statut sémantique (vide / plein) de chaque forme fléchie de la phrase. Seules les formes sémantiquement pleines sont conservées dans la représentation syntaxique profonde
- **diathèse et sous-catégorisation canonique des verbes et des adjectifs** : on explicite les changements de diathèse : les verbes utilisés dans une diathèse non canonique porte un trait diat indiquant le type de redistribution (passif, impersonnel etc...), et les arguments syntaxiques profonds des verbes et des adjectifs sont identifiés et leur fonction canonique est explicitée.
 - Cela inclut l'explicitation du sujet (final) des infinitifs et des participes, lorsqu'il est interprétable au sein de la phrase (qu'il s'agisse de contrôle obligatoire ou arbitraire)
 - le « sujet » canonique d'un adjectif est l'élément que l'adjectif modifie lorsqu'il est épithète
 - on explicite donc la relation prédicative entre le sujet et l'attribut du sujet (respectivement l'objet et l'attribut de l'objet), dans le cas d'un attribut adjectival
 - par extension, on fait la même explicitation dans le cas d'attributs prépositionnels ou adverbiaux (pour les noms ce sera à intégrer lorsque la sous-catégorisation des noms sera traitée)
- la diathèse, le mode, le temps et la voie verbale sera explicitée sur le verbe final du nucléus verbal (le participe, ou l'infinitif en cas de causatif)
- les relations anaphoriques syntaxiquement déterminables (pronoms relatifs) sont résolues : pour les relatives épithètes, le pronom relatif est considéré comme sémantiquement vides, et ses dépendances sont reportées sur son antécédent

1.5.1 Court-circuitage des marqueurs grammaticaux sémantiquement vides

Dans la représentation profonde visée, les mots sémantiquement vides ne portent aucune dépendance. En particulier :

- la dépendances portée au niveau intermédiaire par une préposition régie pointe au niveau profond vers l'objet de la préposition
- la dépendances portée au niveau intermédiaire par le complémenteur *que* introduisant une phrase pointe au niveau profond vers la racine de cette phrase (v fini)
- la dépendances portée au niveau intermédiaire par le complémenteur *à* ou *de* introduisant une infinitive pointe au niveau profond vers l'infinitif (attention, ces complémenteurs sont codés comme des prep, ils sont repérés car ils portent une dépendance "obj" vers leur gouverneur, et pas **a_obj** ou **de_obj**)
- la dépendance entre le pronom relatif et le verbe d'une relative adnominale est déplacée au niveau profond sur le nom modifié par la relative

- dans une clivée, la dépendance portée par le *que* ou *qui* introduisant la subordonnée est déplacée au niveau profond vers le foyer de la clivée
- la dépendance du clitique réfléchi se (en cas de vrai réfléchi ou réciproque) est transférée au niveau profond vers :
 - dans le cas général, l'élément qui est sujet canonique
 - sauf dans le cas d'un clitique réfléchi sur auxiliaire causatif *faire*, auquel cas la dépendance pointant vers le clitique réfléchi est transférée sur l'argument causateur (par exemple pour *Paul fait se coiffer Pierre* : la dépendance objet portée par le réfléchi au niveau intermédiaire, est remplacée par une dépendance objet pointant sur *Pierre*, qui est objet final, mais sujet canonique de *coiffer*. En revanche pour *Paul se fait coiffer par Pierre*, la dépendance objet portée par le réfléchi est remplacée par une dépendance objet pointant sur *Paul* qui est sujet final, mais a pour fonction canonique **argc**).

On obtient ainsi une représentation relevant en partie de la sémantique, mais où sont absentes toutes les informations de portée, les informations de coréférence non déterminables syntaxiquement et les informations de structure informationnelle. L'intérêt d'une telle représentation est de préparer l'analyse sémantique en donnant accès aux structures argumentales.

2 Catégories, traits et fonctions grammaticales

On détaille ici

- le statut linguistique des nœuds de la REPRSYNTPROF, les catégories morpho-syntaxiques utilisées, les autres traits associés aux nœuds
- les fonctions grammaticales constituant les étiquettes des dépendances

2.1 Nœuds

Les nœuds correspondent à une forme fléchie et sont étiquetés par l'identifiant d'ordre linéaire dans la phrase, une forme fléchie, un lemme, et par la catégorie syntaxique de ce lemme.

2.1.1 Traitement des mots composés et unités polylexicales

Le présent schéma d'annotation a été défini grossièrement comme si on avait un nœud par forme fléchie.

Cela dit dans son application au corpus Sequoia, nous n'avons pas modifié la tokenisation présente dans le corpus original. Or ce corpus a un traitement actuellement hybride des mots composés : les mots composés grammaticaux sont codés sous la forme d'un seul token, avec un "_" séparant les composants du composé (par exemple : "*grâce_à*"), mais le corpus ne contient aucun composé nominal verbal ou adjectival. Ainsi "*pomme de terre*" est traité avec 3 nœuds.

On a donc actuellement dans le corpus Sequoia annoté en syntaxe profonde une situation où un

nœud correspond soit à une forme fléchie sémantiquement pleine (majorité des cas), soit un élément d'une unité sémantique plus large.

2.1.2 Mots considérés comme sémantiquement vides

Les mots considérés comme sémantiquement vides sont marqués comme tels (on utilise un trait « **void=y** »), et ne font pas partie de la ReprSyntProf. Il s'agit précisément des cas suivants :

- *il* explétif (dépendance **subj** uniquement finale)
- *se* moyen ou neutre (dépendance **add.demsuj**, trait **diat=demsuj** sur le verbe)
- *se* « réfléchi intrinsèque » (dépendance **aff**)
- *se* « vrai réfléchi ou réciproque » (qui est en surface **obj** ou **a_obj** du verbe, mais en profond cette dépendance est reportée sur le sujet du verbe)
- *se* pour les constructions "se faire + infinitif" (dépendance **aff**), **diat=caus_seobj**
- *à* et *de* prépositions régies (fonctions **a_obj** et **de_obj** respectivement)
- *par* introduisant un complément d'agent (fonction finale **p_obj.agt**)
- auxiliaires (relations **aux.***), y compris le verbe *faire* dans le complexe causatif
- le clitique *ce* dans les clivées
- *à* et *de* complémenteurs introduisant une infinitive objet (fonction **obj** : *Paul promet de se lever tôt / Paul cherche à gagner plus*). NB: ils sont codés comme des prépositions dans la représentation de surface originale du FTB, et cela n'est pas modifié dans la représentation profonde.
- *que* complémenteur introduisant un argument phrastique **subj**, **obj** ou **de_obj** (*que Paul puisse penser cela m'a indignée, Paul souhaite que la réunion soit avancée, Paul se réjouit que la réunion soit annulée*)
- *si* complémenteur introduisant une interrogative. (*Paul demande si le spectacle est terminé / Paul ne sait pas si le spectacle est terminé*) même si *si* marque le caractère interrogatif et n'est donc pas strictement vide. => on opte pour uniformiser le traitement de *si* et de *que* : on considère *si* vide, et en dépendances profondes, on a une dépendance directe entre le verbe enchâssant et le verbe de l'interrogative. **Cependant**, on note le caractère interrogatif de la subordonnée, pour différencier par exemple *Paul ne sait pas que / si le spectacle est terminé*.
- *le* dans les tours "le cas LE plus grave", "les cas LES plus longuement débattus"

Pour les prépositions régies par une dépendance **p_obj.o**, on a des cas de préposition totalement vide sémantiquement (*compter sur*) ou bien portant une sémantique (*voter pour / voter contre*) => on ne distingue pas ces cas, et conserve ces prépositions dans la représentation profonde.

Par ailleurs certains clitiques ou pronoms anaphoriques sont résolus, et supprimés de la représentation profonde : c'est le cas des pronoms relatifs dont on peut repérer structurellement l'antécédent (relatives épithètes, clivées), et c'est le cas des clitiques coréférents à un syntagme disloqué (cf. [Section 12.2](#)).

2.1.3 Catégories morpho-syntaxiques

Nous retenons deux niveaux de granularité pour les catégories morpho-syntaxiques.

Le jeu de catégories grossier est exactement constitué des 13 catégories de base du FTB.

Le jeu plus fin est celui défini par Crabbé et Candito (2008). Il est construit à partir des 13 catégories de base du FTB, croisées avec le mode des verbes, et certaines distinctions apparaissant dans le trait **subcat** dans le FTB, soit :

- la caractéristique interrogatif ou pas des adverbes, adjectifs et déterminants

- une indication du cas des clitiques
- le trait commun/propre pour les noms
- la distinction conjonction de coord versus conjonction de subordination
- des étiquettes particulières pour les amalgames préposition+déterminant (*au*), préposition+pronom (*auquel*)

La catégorie fine est donc une information redondante qui peut être obtenue à partir de la catégorie grossière et des traits.

Catégorie fine	Catégorie de base du FTB	Trait supplémentaire du FTB	Description / exemple
V	V	m=ind	verbe indicatif
VS	V	m=subj	verbe subjonctif
VINF	V	m=inf	verbe infinitif
VPP	V	m=part, t=past	participe passé
VPR	V	m=part, t=pst	participe présent
VIMP	V	m=imp	verbe impératif
NC	N	s=c ord card	nom commun
NPP	N	s=p	nom propre
CS	C	s=s	conjonction de subordination
CC	C	s=c	conjonction de coordination
CLS	CL	s=suj	clitique sujet
CLO	CL	s=obj	clitique objet
CLR	CL	s=refl	clitique réfléchi
P	P		préposition non amalgamée
P+D	P+D		préposition+dét : <i>le lutin des alpages</i>
P+PRO	P+PRO		préposition+prorel : <i>le lieu auquel on pense</i>
I	I		interjection
PONCT	PONCT		ponctuation
ET	ET		mots étrangers
ADJ	A	s=qual ind card ord	adjectifs non interrogatifs

ADJWH	A	s=int	adjectifs interrogatifs
ADV	ADV	s=ind neg ou pas de trait s	adverbes non interrogatifs
ADVWH	ADV	s=int	adverbes interrogatifs
PRO	PRO	s= card dem ind neg pers	pronoms ni interrogatifs ni relatifs
PROREL	PRO	s=rel	pronoms relatifs
PROWH	PRO	s=int	pronoms interrogatifs
DET	D	s= def dem ind neg part poss ou pas de trait s	déterminants non interrogatifs
DETH	D	s=int	déterminants interrogatifs
PREF	PREF		préfixe

2.1.4 Traits profonds

Il s'agit de traits explicités sur le verbe lexicalement plein d'un noyau verbal comprenant éventuellement des clitiques et des auxiliaires.

Le verbe lexicalement plein est toujours le dernier verbe du noyau verbal.

On explicite :

- **dm** = le mode profond : c'est le mode du premier auxiliaire (*a été mangé* => indicatif, *fasse manger* => subjonctif, *être allé* => infinitif, *se faisant dépasser* => participe présent)
- **dl** = le lemme profond : incluant les clitiques figés s'ils existent (en particulier le réfléchi intrinsèque)
- **diat** = la diathèse :
 - **passif**,
 - **impers**, actif impersonnel (*il est arrivé trois personnes, il est surprenant que Paul soit parti*),
 - **passif_impers**, passif impersonnel (*il a été décidé de maintenir la gratuité de ce service*),
 - **demsuj**, démotivation du sujet pour le moyen et le neutre ; on ne distingue pas entre les deux (*Ce plat se mange souvent froid / La branche s'est finalement cassée*),
 - **caust**, causatif d'un verbe transitif (*Paul fait soigner ses enfants par un rebouteux. Paul fait balayer la cour aux nouvelles recrues*),
 - **causi**, causatif d'un intransitif (*Paul fait venir un prestidigitateur*),
 - **caust_seobj**, se+faire+inf sans interprétation réfléchie (*Paul s'est fait huer*).

Pour un verbe plein, l'absence de trait *diat* signifie qu'il s'agit d'une diathèse canonique, forcément à l'actif.

- **cltype** = le type de la proposition indiqué sur sa tête :
 - **int** = interrogative (percolé du complémenteur *si* ou bien de l'élément interrogatif antéposé)

- **imp** = impérative (*Que la lumière soit !*)
- **excl** = exclamative (*Qu'il est beau!* ou *Comme ils sont bêtes!* ou *De quelle audace elle a fait preuve!*)

L'absence de trait *cltype* sur la tête d'une proposition signifie que celle-ci est déclarative.

2.1.5 Autres traits

Traits de genre, nombre, personne, mode (non profond), temps grammatical, personne + quelques autres informations :

Trait	Signification	Commentaire
g=f	genre=féminin	
g=m	genre=masculin	
m=imp	mode=impératif	redondant avec cat=VIMP
m=ind	mode=indicatif	redondant avec cat=V
m=inf	mode=infinitif	redondant avec cat=VINFINF
m=part	mode=participe	redondant avec cat=VPP VPR
m=subj	mode=subjonctif	redondant avec cat=VS
n=p	nombre=pluriel	
n=s	nombre=singulier	
p=1	personne=1	
p=2	personne=2	
p=3	personne=3	
s=c	nom commun	redondant avec cat=NC
s=card	cardinal NC,A,D or PRO	
s=def	defini DET or P+D	
s=dem	demonstratif DET or PRO	
s=ind	indefini DET, ADJ or PRO	
s=int	interrogatif PROWH, ADJWH, DETWH, ADVWH	
s=neg	negatif ADV, PRO, DET, NC (pour <i>non</i>)	

s=obj	clitique objet réfléchi ou pas	
s=ord	adjectif ordinal	
s=p	nom propre	redondant avec cat=NPP
s=part	déterminant partitif	
s=pers	pronom personnel	
s=poss	possessif DET, PRO	
s=qual	adjectif qualificatif	
s=refl	clitique réfléchi	redondant avec cat=CLR
s=rel	pronom relatif (PROREL ou P+PRO)	
s=s	conjonction de subordination et ponctuation de fin de phrase	systematique pour cat=CS
s=suj	clitique sujet	redondant avec cat=CLS
s=w	ponctuation interne à une phrase	
t=cond	temps=conditionnel	
t=fut	temps=futur	
t=impft	temps=imparfait	
t=past	temps=passé simple participe passé	
t=pst	temps=présent	

2.2 Fonctions grammaticales

Les fonctions grammaticales représentent un typage des réalisations syntaxiques des arguments, telles qu'observées dans des diathèses finales. Le fait d'utiliser la même étiquette SUJ pour un sujet canonique ou un sujet final est que *lorsque le sujet canonique est employé dans une diathèse canonique (i.e. sans redistribution), alors il a toutes les propriétés du sujet final*. Ainsi par exemple, le sujet canonique du passif, i.e. le p_obj.agt final, a toutes les propriétés d'un sujet lorsque le même verbe est employé à l'actif.

La plupart des fonctions peuvent être aussi bien canonique que finale, mais par exemple **p_obj.agt** ne peut être que finale (et correspond toujours à un sujet canonique), et **argc** (l'argument causateur du causatif) ne peut être que canonique (ce qui constitue un cas limite, cf. du coup on n'observe jamais **argc** dans une sous-catégorisation finale. Il est défini uniquement par la propriété de redistribution causative : c'est le sujet final dans une diathèse causative).

En outre seules quelques fonctions peuvent être liées par des redistributions : **suj**, **obj**, **a_obj**, **de_obj**, **p_obj.agt**, **argc**. Toutes les autres fonctions sont systématiquement finale et canonique (et notées

par commodité par un label simple).

Nous allons lister ici toutes les fonctions grammaticales utilisées en représentation syntaxique profonde, mais également celles utilisées uniquement en représentation de surface.

Nous utilisons premièrement les fonctions grammaticales à gouverneur **verbal** définies dans le FTB (Abeillé, 2004), avec cependant des distinctions au sein des fonctions **mod** et **p_obj**. **On reprend ainsi exactement les fonctions suj, obj, ats, ato, a_suj** et **de_obj**, telles que définies dans le FTB.

La fonction **p_obj** du FTB est subdivisée en :

- **p_obj.agt** pour un complément d'agent,
- **p_obj.o** pour les autres dépendants sous-catégorisés, dont la prep est fixée, et qui ne sont ni **a_obj**, ni **de_obj**

La fonction **mod** est subdivisée en diverses sous-fonctions énumérées par le tableau ci-dessous.

Nous étendons l'utilisation de ces fonctions dans deux cas non annotés dans le FTB :

- les dépendants dans une participiale passée
- les dépendants adverbiaux réduits à un seul mot comme dans *ils sont là*.

On ajoute par ailleurs les relations **aux.caus**, **aux.pass**, **aux.tps** pour les auxiliaires, et **aff** pour les clitiques figés.

La dépendance vers le clitique *se* a comme étiquette :

- soit une fonction **obj** ou **a_obj** quand le clitique a effectivement un sens réfléchi ou réciproque,
- soit la fonction **aff** pour les verbes essentiellement pronominaux,
- Soit la fonction **aff.demsuj** pour le moyen ou le neutre, qui ne sont pas distingués l'un de l'autre .

On ajoute également une fonction finale **dis** pour les dépendants disloqués (cf. manquant dans le guide FTB, cf. p.c. d'Anne Abeillé), qui portent par ailleurs comme étiquette canonique la fonction canonique de leur clitique coréférent (voir [Section 12.2](#)).

Pour les gouverneurs adjectivaux, on précise la fonction de leurs dépendants, cf. l'objectif d'obtenir la sous-catégorisation canonique complète des adjectifs.

Pour *Paul est enclin à la mélancolie*, l'adjectif *enclin* a *Paul* pour sujet canonique, et *la mélancolie* pour **a_obj** canonique.

Pour les dépendants d'autres catégories de gouverneur, on précise ou pas la fonction grammaticale du dépendant, selon que cette fonction est directement prédictible d'après la position syntaxique et la catégorie du dépendant et du gouverneur :

C'est le cas pour un déterminant (=> fonction *det*), un adjectif épithète (=> fonction *mod*), une relative adnominale (=> fonction *mod.rel*).

Par contre, on utilise une fonction sous-spécifiée **dep** pour les cas où il faudrait faire appel à une notion de sous-catégorisation du gouverneur, non directement accessible. C'est en particulier le cas pour les prépositions dépendant de gouverneurs ni verbaux ni adjectivaux ni adverbiaux: autrement dit, on ne gère pas la distinction argument/ajout pour les dépendants *prépositionnels* des noms.

Rem: les fonctions de la forme XXX.YYY sont interprétables comme des sous-types de la fonction XXX

Dans le tableau ci-dessous, on note :

- en **rouge** les fonctions qui n'apparaissent que sur des mots sémantiquement vide et/ou uniquement comme fonction finale, et donc uniquement dans la représentation de surface
- en **bleu** les fonctions qui n'apparaissent que dans la représentation profonde
- en **noir** les fonctions pouvant apparaître dans les deux

Par ailleurs, on indique également les quelques différences entre le schéma de représentation de surface du FTB et le schéma de représentation de surface du Sequoia : les fonctions mod et p_obj ont en effet été précisées. On note cela en fournissant entre crochets l'équivalent FTB.

<i>Fonction</i>	<i>Catégorie possible pour le gouverneur</i>	<i>Possible comme fonction FINALE ?</i>	<i>Possible comme fonction CANONIQUE ?</i>	Définition / commentaire
-	V	non	oui	fonction canonique vide portée par le sujet final <i>il</i> explétif d'une construction impersonnelle
aff	V	oui	non	pour un clitique figé et le clitique se ni réfléchi ni réciproque
aff.demsuj	V	oui	non	fonction du clitique <i>se</i> par rapport au verbe pour exprimer le neutre ou le moyen
arg	P	oui	oui	Utilisée dans le cas de prépositions « liées » : dans <i>de Charybde en Scylla</i> , <i>en Scylla</i> est dépendant de type <i>arg</i> de la première préposition <i>de</i>
argc	V	non	oui	fonction uniquement canonique : pour l'argument causateur dans les constructions causatives
arg.comp	ADJ, ADV	non	oui	relation d'une comparative avec son gouverneur qui est un adverbe ou adjectif exprimant la comparaison
arg.cons	ADJ, ADV	non	oui	relation d'une consécutive avec son gouverneur adverbial ou adjectival
ato	V	oui	oui	attribut de l'objet
ats	V	oui	oui	attribut du sujet

aux.caus	V	oui	non	auxiliaire du causatif ou du complexe <i>se faire + inf</i>
aux.pass	V	oui	non	auxiliaire du passif
aux.tps	V	oui	non	auxiliaire de temps
a_obj	V, ADJ	oui	oui	complément indirect non locatif introduit par <i>à</i>
coord	tout type	oui	oui	relation portée par un coordonnant, avec comme gouverneur le premier conjoint
dep	ADJ, D, ET, N, PRO	oui	oui	relation sous-spécifiée, pour les dépendants prépositionnels de gouverneurs non verbaux et non adjectivaux (pas de gestion de la distinction argument / ajout pour ces cas)
dep.coord	CC	oui	oui	relation portée par un coordonné (sauf le premier), avec comme gouverneur le coordonnant immédiatement précédent
det	ADJ, ADV, N, PRO	oui	oui	déterminant
de_obj	V, ADJ	oui	oui	complément indirect non locatif introduit par <i>de</i>
dis [n'existe pas dans le FTB]	V	oui	non	fonction uniquement finale, pour les dépendants disloqués (apparaissant en tête ou fin, avec un clitique de reprise : <i>Paul, sa moto, il l'a cassée</i>).
mod	tout type	oui	oui	modifieur de gouverneur verbal ou adjectival modifieur, autre qu'une relative, d'un autre type de gouverneur pour peu que l'on puisse repérer structurellement qu'il s'agit d'un modifieur (adjectif épithète par exemple)
mod.app [équivalent FTB: mod]	ADJ, N, PRO	oui	oui	apposition à un groupe nominal
mod.cleft [équivalent FTB: mod]	V	oui	oui	la subordonnée dans le cas d'une clivée

mod.comp	<i>ADJ, ADV, N, P, PRO, V</i>	<i>oui</i>	<i>oui</i>	modifieur introduisant une comparaison
mod.inc [équivalent FTB: mod]	<i>ADJ, CC, CS, N, P, V</i>	<i>oui</i>	<i>oui</i>	proposition incise
mod.rel	<i>ADJ, ADV, N, PRO</i>	<i>oui</i>	<i>oui</i>	dépendance d'une relative par rapport à son antécédent
mod.rel.part	<i>N</i>	<i>non</i>	<i>oui</i>	dépendance de la tête nominale d'une relative avec ellipse introduite par <i>dont</i> par rapport à l'antécédent
mod.super	<i>ADJ, ADV, N</i>	<i>non</i>	<i>oui</i>	modifieur exprimant le superlatif
mod.voc	<i>V</i>	<i>oui</i>	<i>oui</i>	vocatif
obj	<i>V</i>	<i>oui</i>	<i>oui</i>	objet de verbe
obj.cpl	<i>CS</i>	<i>oui</i>	<i>oui</i>	objet de complémenteur
obj.p	<i>P</i>	<i>oui</i>	<i>oui</i>	objet de préposition
ponct	<i>tout type</i>	<i>oui</i>	<i>oui</i>	relation portée par tout dépendant typographique, sauf pour les virgules jouant le rôle de coordonnant (qui porte la relation coord)
p_obj.agt [équivalent FTB: p_obj]	<i>V</i>	<i>oui</i>	<i>non</i>	fonction uniquement finale : complément d'agent en <i>par</i> ou <i>de</i> en cas de passif, ou en <i>par</i> dans le cas de causatif
p_obj.o [équivalent FTB: p_obj]	<i>V, ADJ</i>	<i>oui</i>	<i>oui</i>	complément indirect du verbe ou de l'adjectif autre
subj	<i>V, ADJ</i>	<i>oui</i>	<i>oui</i>	sujet d'un verbe ou d'un adjectif

3 Changements de diathèse

Dans cette section nous détaillons la représentation des différents changements de diathèse traités dans la REPRSYNTPROF comme des "redistributions" des fonctions canoniques.

On rappelle (cf. [Section 1.2.2](#)) que seuls les changements de diathèse marqués morpho-syntaxiquement (par un auxiliaire, un clitique *se* etc...) sont traités. Les alternances non marquées

comme l'alternance locative par exemple donnent lieu à 2 sous-catégorisations canoniques distinctes. Ainsi par exemple *Paul charge le camion d'oranges* et *Paul charge les oranges dans le camion* n'a pas une représentation unifiée dans le REPRSYNTPROF, la neutralisation étant à faire en représentation sémantique ou lexicale.

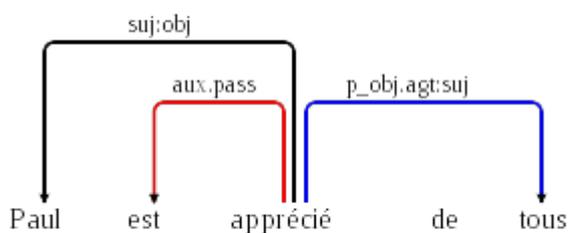
NB: dans cette section, par souci de lisibilité, on choisit parfois de ne pas noter les dépendances uniquement surfaciques (rouges). Les dépendances vers les déterminants sont également omises.

3.1 Passif

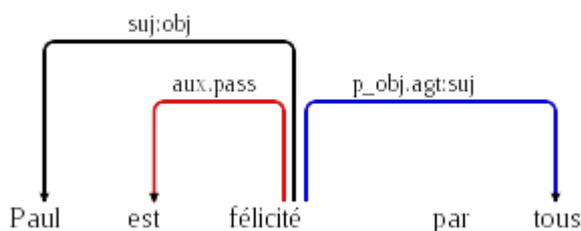
L'auxiliaire du passif est dépendant du participe (relation **aux.pass**) :

Le complément introduit par la préposition régie *par* ou *de* porte la fonction **p_obj.agt:suj**.

2. *Paul est apprécié de tous*

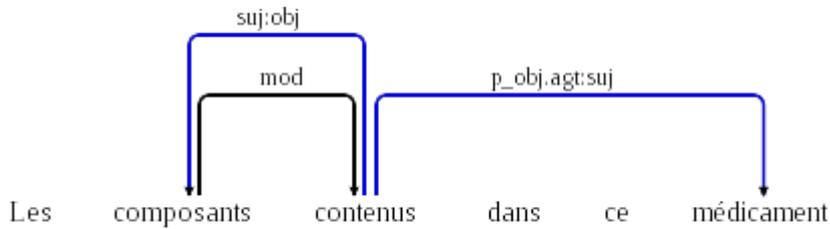


3. *Paul est félicité par tous*



On utilise également la fonction **p_obj.agt** pour le cas particulier du verbe *contenir* au passif : *Le médicament contient un excipient / l'excipient (est) contenu (??par / dans) le médicament.*

4. *Les composants contenus dans ce médicament*



Le passif impersonnel est traité dans [Section 3.4](#).

3.2 Vrai réfléchi et réciproque

Dans toute la suite on note clitique SE l'ensemble des formes du paradigme *me, te, se, nous, vous, se*, quelle que soit leur valeur grammaticale.

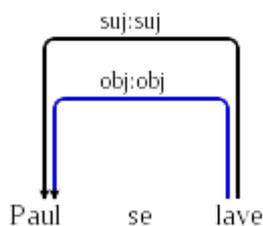
Cette section traite du cas où SE marque une construction à sens réfléchi ou réciproque (voir [Section 4.2.6](#) pour un récapitulatif des différentes valeurs de SE, et voir Candito, 2013 pour le guide spécifique pour l'annotation du SE). Dans la suite de cette section on parle de "réfléchi" pour le sens réfléchi ou réciproque.

Bien qu'il y ait des arguments pour considérer le réfléchi comme un changement de diathèse (en particulier l'interaction avec l'impersonnel ou le causatif indique que le réfléchi est intransitivant), **par souci de simplification on ne représente pas ici le réfléchi comme un changement de diathèse.**

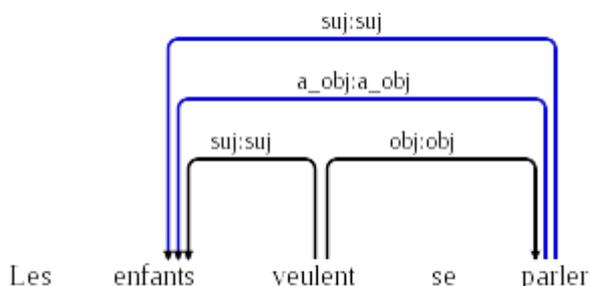
On a en syntaxe de surface le SE réfléchi ou réciproque qui porte une fonction finale **obj** ou **a_obj**. Dans la représentation profonde, la dépendance portée par le SE est déplacée pour pointer vers le sujet final (ce qui donne formellement un multi-graphe), sauf dans le cas d'interaction avec le causatif (traité dans la [Section 3.3](#))

Les distinctions sémantiques entre réfléchi et réciproque ne sont pas explicitées.

5. *Paul se lave*



6. *Les enfants veulent se parler*



Le réfléchi peut interagir avec l'impersonnel (comme dans *Il se lavait à cette fontaine un pauvre gars sans le sou*) mais on n'a aucune occurrence dans le corpus.

Le réfléchi peut aussi interagir avec le causatif.

3.3 Causatif

On suit l'annotation du FTB où le verbe *faire* plus infinitif est considéré comme un verbe complexe causatif. La conversion en dépendances de surface donne *faire* comme un auxiliaire (**aux.caus**) gouverné pour l'infinitif.

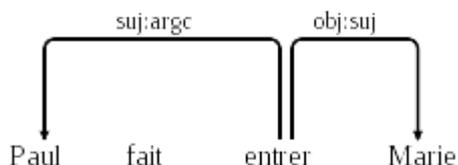
En dépendances profondes, l'auxiliaire est transformé en un trait de diathèse sur l'infinitif, et les fonctions canoniques sont explicitées. Le causatif a la particularité d'introduire un argument causateur supplémentaire : un verbe à n arguments à l'actif aura $n+1$ arguments au causatif.

Pour les redistributions causatives simples, l'argument causateur est toujours le sujet final du verbe au causatif. On note **argc** la fonction canonique de cet argument causateur.

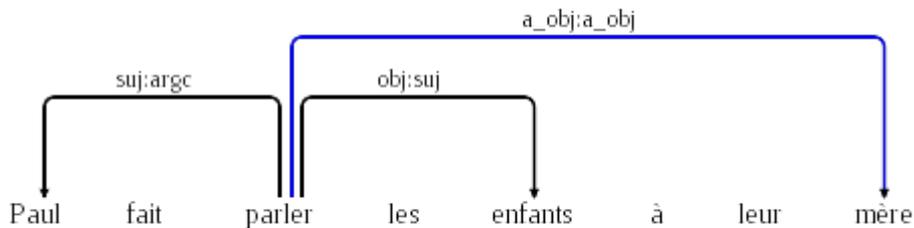
On a une redistribution possible pour un verbe employé intransitivement, et deux redistributions possibles pour un verbe employé transitivement :

- diat=causi : { **argc** ==> **su**j + **su**j ==> **obj** }, (exemples [7](#), [8](#))
- diat=caust : { **argc** ==> **su**j + **obj** ==> **obj** + **su**j ==> **par_obj** }, (exemple [9](#))
- diat=caust : { **argc** ==> **su**j + **obj** ==> **obj** + **su**j ==> **a_obj** }, (exemple [10](#))

7. *Paul fait entrer Marie*

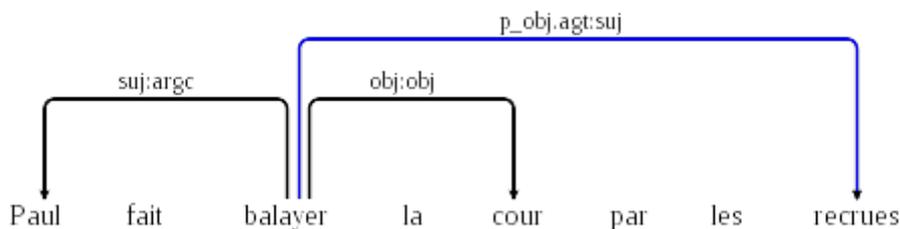


8. *Paul fait parler les enfants à leur mère*

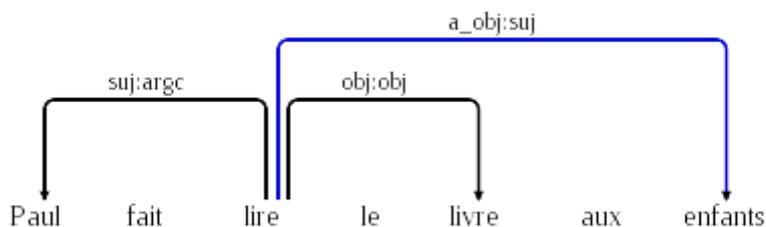


Dans le causatif d'un transitif, le sujet initial du transitif est soit **a_obj** soit **p_obj.agt** (trait diat=caust dans les 2 cas) :

9. *Paul fait balayer la cour par les recrues*



10. *Paul fait lire le livre aux enfants*



Il existe des contraintes complexes sur la combinatoire des différents arguments syntaxiques d'un verbe au causatif, sur le placement des clitiques, sur le choix du causatif en *par* versus le causatif en *à* etc... (voir Tasmowski, 1984 par exemple).

3.3.1 Analyse des constructions en "se faire" + infinitif

Le causatif peut ne peut pas interagir avec le passif en français, ni avec l'impersonnel. Il peut cela dit interagir avec le réfléchi, ou plus généralement avec un clitique *se*.

On peut avoir le placement du "se" devant l'infinitif, qu'il s'agisse d'une interaction du causatif sur un vrai réfléchi (11), sur un réfléchi intrinsèque (12), ou sur un se-moyen (13). Mais il s'agit de

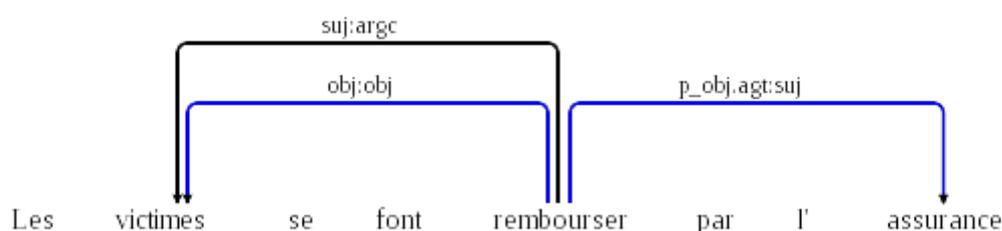
tournures pour le moins rares, et en tous cas absente du corpus visé.

11. *La direction a fait s'acheter le dernier macintosh à tous les élèves*
12. *La suite des évènements les a fait s'apercevoir de leur erreur*
13. *La crise fait se vendre ces kits comme des petits pains*

Une autre interaction est par contre relativement fréquente au sein des causatifs (6 occurrences sur les 34 occurrences de causatif dans le corpus Sequoia), avec le clitique *se* apparaissant sur l'auxiliaire causatif *faire*.

On a un cas analysable comme le réfléchi s'appliquant sur un complexe causatif, comme dans la paire suivante :

14. *Le juge fait rembourser les victimes par l'assurance*
15. *Les victimes se font rembourser par l'assurance*



On analyse [15](#) comme si un réfléchi s'appliquait sur un complexe causatif. Sémantiquement, on a bien le sens de la diathèse causative : le sujet final cause partiellement l'éventualité décrite par l'infinitif, qu'il s'agisse d'un sujet animé volitionnel (en [15](#) on interprète que les victimes agissent volontairement pour obtenir le remboursement), ou bien d'un cas non volitionnel mais qui reste causal (par exemple en [12](#) où l'on a un causatif sur un réfléchi intrinsèque, le sujet final *la crise* n'est pas volitionnel mais bien interprété comme la cause de la prise de conscience).

Mais la construction *se + faire + infinitif*, largement étudiée (voir Veacock, 2008 pour une revue) est également utilisée pour des cas où le sujet final du complexe causatif ne peut pas clairement être analysé *sémantiquement* comme ayant *causé* la situation, comme par exemple en ([16](#)).

16. *Il s'est fait voler son poste de radio (bien malgré lui)* (Blanche-Benveniste, 2007, citée par Veacock, 2008)

En [16](#), le sujet *il* n'est clairement pas volitionnel (mais on vient de voir que cela n'est pas une condition nécessaire de la construction causative). L'important ici est qu'il n'est pas clair qu'il soit considéré comme causal, i.e. comme un "responsable malgré lui".

C'est plus net avec des infinitifs sémantiquement négatifs pour leur sujet comme *frapper*, *violer* etc... où on n'a aucune interprétation de causalité.

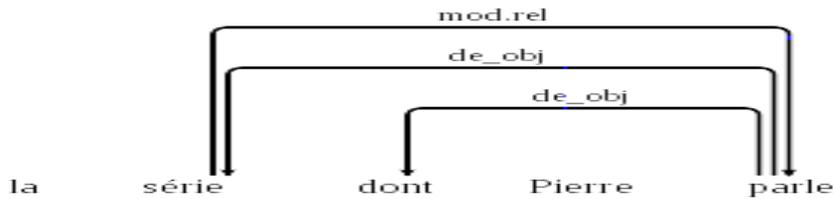
17. *Paul s'est fait tabasser en bas de chez lui*

On choisit de considérer ces cas non causaux comme une construction spécifique, ne mettant pas en jeu la redistribution causative (on retient comme critère pour le causatif qu'il y ait une sémantique

de causalité, qu'elle soit ou non volitionnelle).

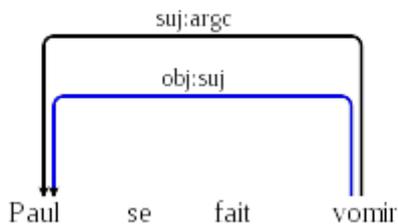
On fait une distinction binaire entre 15 où le sujet final est l'argument causateur (argc) et 17 où il ne l'est pas. Dans les cas d'un sujet (final) animé, on le marque argc uniquement si la causalité est indubitable. Ainsi en 18, on ne peut pas établir clairement que l'orateur est responsable. Idem pour 17 et pour 16, analysés sans argument causateur : le sujet final est directement considéré comme l'objet canonique de l'infinitif.

18. *L'orateur s'est fait huer par le public*

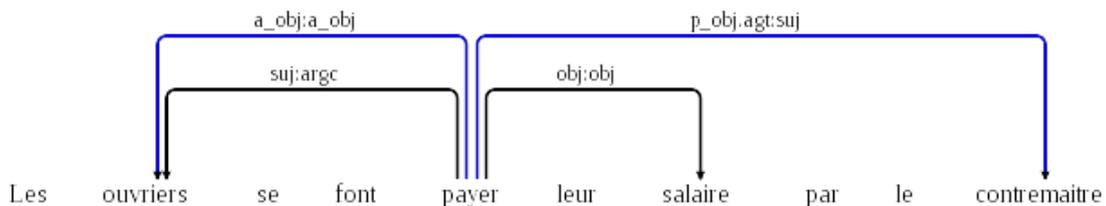


- On analyse par contre comme un causatif + réfléchi l'exemple 15, ainsi que les deux exemples suivants:

19. *Paul se fait vomir*



20. *Les ouvriers se font payer leur salaire par le contremaître*



3.4 Changement de diathèse impersonnel versus tournure impersonnelle

Un clitique *il* non référentiel est systématiquement considéré comme vide, et donc absent de la représentation profonde.

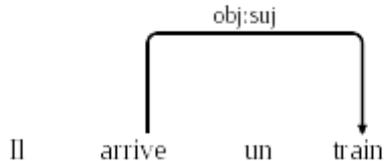
Mais on doit distinguer :

- un changement de diathèse "impersonnel" ou "impersonnel passif" : repéré parce que le verbe, dans le même sens, alterne avec une construction non impersonnelle
 - **impersonnel (actif)** : *il semble évident qu'il s'agit d'un faux => qu'il s'agit d'un faux semble évident*
 - **passif impersonnel** : *il vous est demandé de régler le problème => régler le problème vous est demandé*
- une "tournure impersonnelle" : n'alternant pas avec un équivalent non impersonnel
 - *il pleut*
 - *il s'agit d'un faux*
 - *il (me) semble que c'est un faux => *que c'est un faux (me) semble*

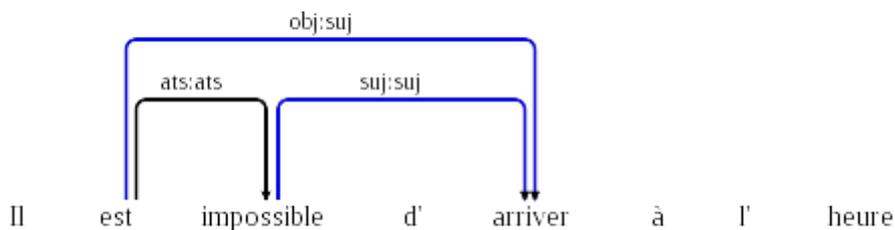
Dans les deux cas, le *il* ne fait pas partie de la sous-catégorisation canonique du verbe (il n'a pas de fonction canonique (en pratique elle est notée "_")).

Mais dans le premier cas, l'objet final est sujet canonique, et le V porte un trait de diathèse impers (cf. [21](#), [22](#), [23](#)) ou "passif_impers" [24](#).

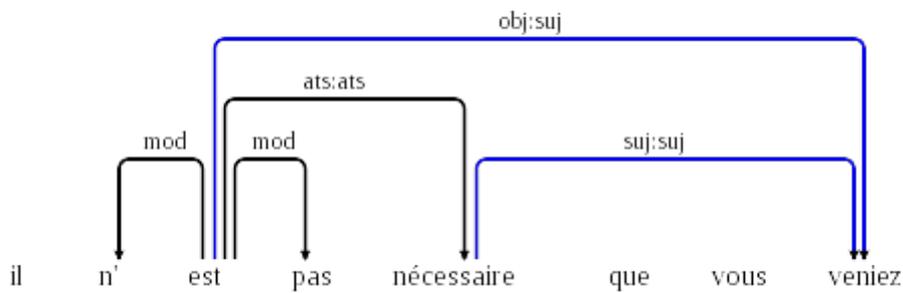
21. *Il arrive un train*



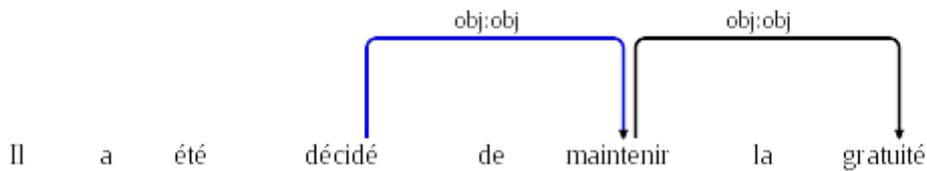
22. *Il est impossible d'arriver à l'heure*



23. *il n'est pas nécessaire que vous veniez*

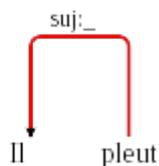


24. Il a été décidé de maintenir la gratuité

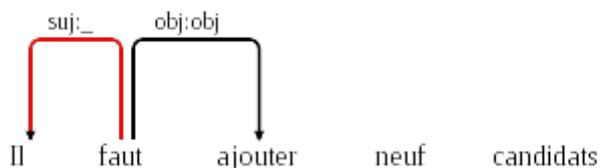


Pour les tournures impersonnelles au contraire, (par exemple avec *pleuvoir*, *falloir*, *s'agir*), on n'a aucune redistribution, l'objet final éventuel est également l'objet canonique. On a simplement suppression en syntaxe profonde du *il* explétif : on considère que le verbe n'a pas de sujet canonique (cf. les arguments syntaxiques profonds sont forcément sémantiquement plein).

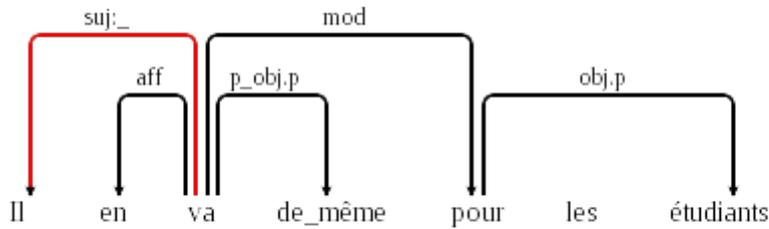
25. Il pleut



26. Il faut ajouter neuf candidats



27. Il en va de même pour les étudiants



3.5 Se Moyen

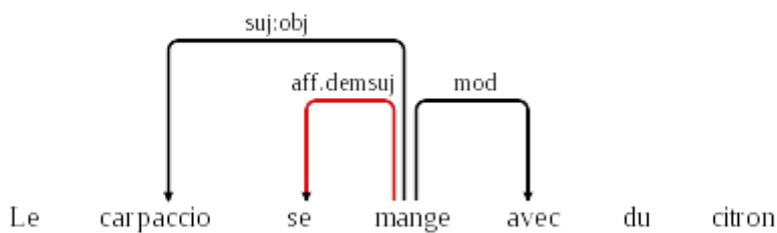
Il s'agit d'une construction productive, pour tout transitif à sujet canonique agentif.

Syntaxiquement : le sujet final correspond à l'objet canonique ; le sujet canonique n'est pas exprimable ; le verbe porte un *se*

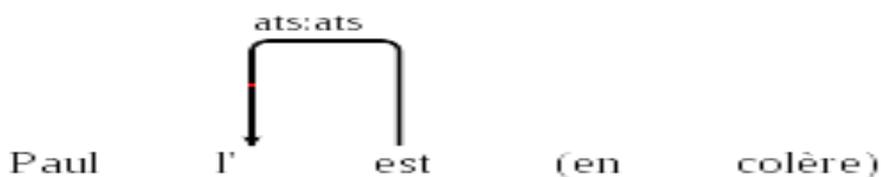
Sémantiquement : le sujet canonique est interprétable (il existe un participant même si non exprimé, dont l'interprétation est résolvable en contexte, ou la plupart du temps générique comme en (28) ou en (29) où il y a forcément des gens qui échangent des milliards). Zribi-Herz (82) repère deux sous-types sémantiques du moyen :

- propriété générique de l'objet canonique (sujet final)
- processus (événement non ponctuel)

28. *Le carpaccio se mange avec du citron*



29. *Des milliards s'échangent en ce moment même sur la planète*



On peut avoir un moyen impersonnel, en particulier avec la sémantique de processus :

30. *Il se dit bcp de choses là-dessus dans les ministères concernés*



3.6 Se Neutre

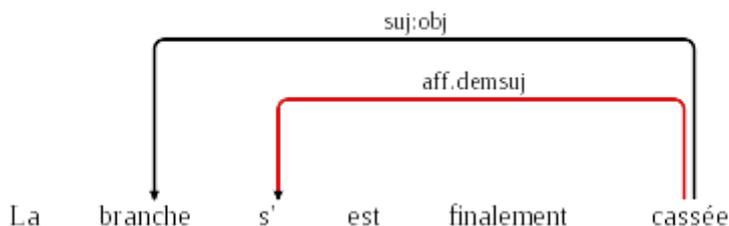
Comme dans la construction moyenne, un transitif apparaît avec son objet canonique en position sujet, et un clitique SE. Mais dans le cas du neutre, *l'agent (sujet canonique) n'est pas interprétable*, l'aspect ponctuel est possible. Par exemple en (31), on ne peut pas spécialement interpréter que quelqu'un ou quelque chose a cassé la branche.

Le neutre est possible pour une classe sémantique de verbes bcp plus restreinte que pour le moyen (aspect inchoatif).

On considère cependant que la différence entre se-neutre et se-moyen est d'ordre sémantique, et elle n'est pas désambiguïlée dans la REPRSYNTPROF.

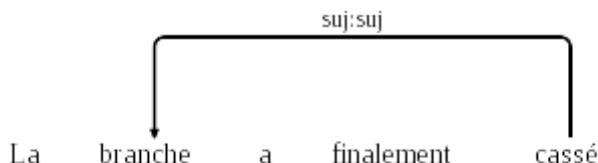
On a dans les deux cas le sujet final qui est objet canonique. Le sujet canonique est absent (soit qu'il ne soit pas interprétable (se-neutre) soit qu'il le soit (se-moyen), mais systématiquement non réalisé car non réalisable syntaxiquement.

31. *La branche s'est finalement cassée*



Parmi les verbes admettant le se-neutre, certains peuvent également se construire sans le SE comme en (32). Mais dans ce cas, comme nous ne retenons que les redistributions marquées, on considère que le verbe est dans une sous-catégorisation canonique, et son sujet final est son sujet canonique, le lien entre (32) et (31) n'est pas capturé, le linking entre sous-cat canonique et arguments sémantiques doit être spécifié.

32. *La branche a finalement cassé*



4 Dépendants du verbe

Cette section spécifie les relations gouverneur verbal et dépendants du verbe.

Pour la représentation de surface, on suit d'une manière générale le guide d'annotations fonctionnelles du FTB (Abeillé, 2004) définissant les fonctions des dépendants du verbe, modulo :

- des subdivisions des fonctions **p_obj*** et **mod***
- l'ajout de labels spécifiques pour les auxiliaires
- une gestion spécifique des comparatives et des superlatives
- en cas de dépendance longue distance, l'annotation du gouverneur non-local (correct), ce qui donne parfois lieu à de la non-projectivité (alors que ces gouverneurs non locaux ne peuvent pas être obtenus par conversion automatique des arbres syntagmatiques du FTB)

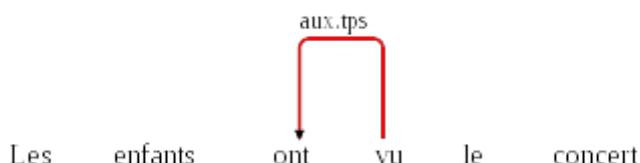
La représentation profonde utilise la même définition des fonctions grammaticales pour les dépendants de verbe, mais se charge de (i) neutraliser les changements de diathèse, et (ii) expliciter le sujet (final) des verbes non conjugués. Voir également les sections [Dépendants participiaux et adjectivaux](#) et [Dépendants phrastiques et infinitivaux](#).

4.1 Auxiliaires

Dans le cas d'un complexe verbal avec auxiliaire (de temps, du passif, causatif) + verbe infinitif ou participe passé, on considère l'auxiliaire comme dépendant du verbe, y compris pour le sujet. Toutes les relations **aux.*** sont des dépendances uniquement surfaciques.

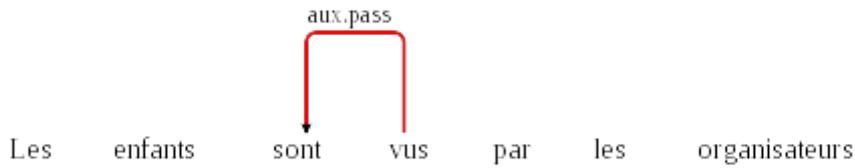
Les auxiliaires de temps sont annotés avec la relation finale **aux.tps** ([voir corpus](#)).

33. *Les enfants ont vu le concert*



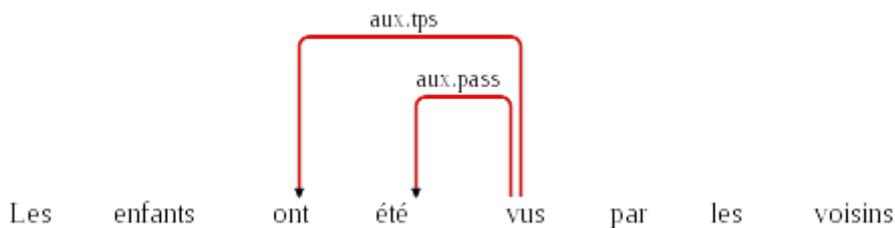
Dans les tours passifs, on annote l'auxiliaire passif avec la relation **aux.pass** ([voir corpus](#)).

34. *Les enfants sont vus par les organisateurs*



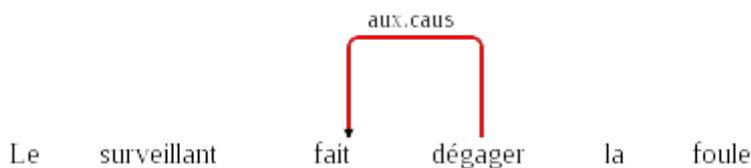
Tous les auxiliaires peuvent être eux-mêmes à un temps composé. MAIS, en surface, tous les auxiliaires sont codés comme dépendants du dernier verbe du nucleus verbal. Ce choix a été fait pour ne pas créer de la non-projectivité en surface, pour des phénomènes massifs mais triviaux comme ceux-ci.

35. *Les enfants ont été vus par les voisins*

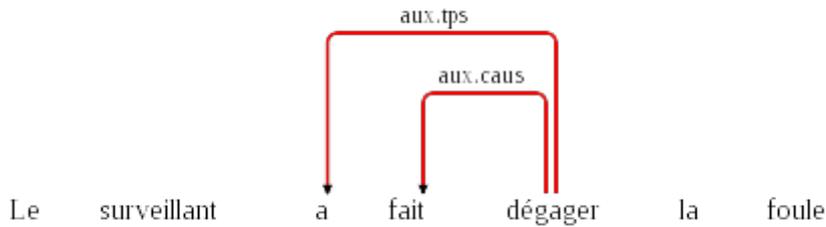


Dans les cas de causatif, on annote le causatif avec la relation **aux.caus** ([voir corpus](#)).

36. *Le surveillant fait dégager la foule*



37. *Le surveillant a fait dégager la foule*



4.2 Clitiques

D'une manière générale, les clitiques peuvent remplir trois fonctions :

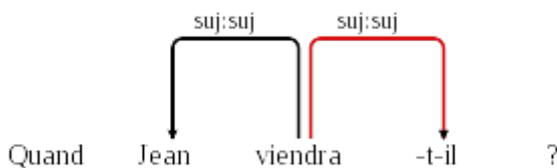
- une fonction "pleine", celle d'un argument syntaxique du verbe (*Paul le voit*),
- celle de former avec le verbe auquel ils se rattachent un tout sémantique comme clitiques figés (*Paul en veut à Marie*),
- celle d'exprimer la démotion du sujet pour la voie moyenne et neutre ; c'est la fonction exclusive du clitique *se* (*Ce livre se vend bien*).

Un même clitique peut remplir plusieurs de ces fonctions. C'est le cas de *se* qui peut aussi être figé ou avoir une fonction pleine.

4.2.1 Clitique *il*

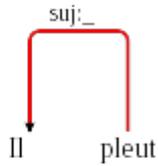
Le clitique *il* porte toujours la fonction pleine de sujet (**su**j). *Il* peut être la reprise d'un sujet plein ([voir corpus](#)), comme dans l'exemple suivant.

38. *Quand Jean viendra-t-il ?*



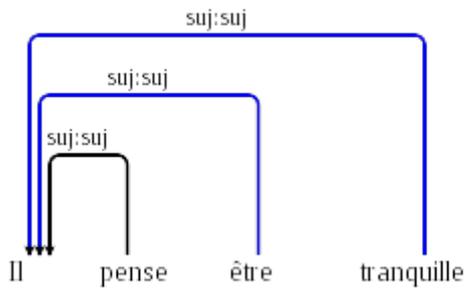
Le clitique *il* peut aussi représenter un sujet impersonnel (cf [Section 3.4](#), [voir corpus](#)).

39. *Il va pleuvoir*



Le clitique *il* peut enfin représenter le sujet profond d'un infinitif (cf [Section 9, voir corpus](#)) ou d'un adjectif (cf [Section 8, voir corpus](#)).

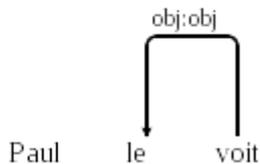
40. *Il pense être tranquille*



4.2.2 Clitique *le*

Le clitique *le*, quand il a une fonction pleine, est soit objet direct (**obj**) ([voir corpus](#)), soit attribut du sujet (**ats**) ([voir corpus](#)). Il varie en genre et nombre uniquement dans le cas de l'objet direct.

41. *Paul le voit*



42. *Paul l'est (en colère)*

Le clitique *le* peut aussi être figé ([voir corpus](#)) comme dans l'exemple suivant.

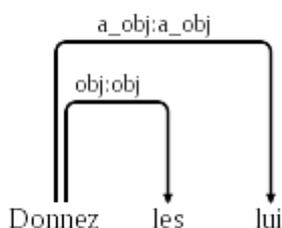
43. *Paul l'emporte sur son adversaire*



4.2.3 Clitique *lui*

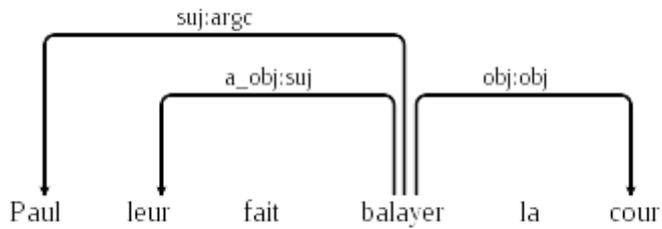
Le clitique *lui* a toujours la fonction pleine finale d'objet indirect (**a_obj**). En général, il a aussi la fonction canonique **a_obj** ([voir corpus](#)).

44. *Donnez-les lui*



Le clitique *lui*, dans une construction causative (cf [Section 3.3](#)), peut être l'objet indirect final et le sujet canonique du verbe causé ([voir corpus](#)).

45. *Paul leur fait balayer la cour*



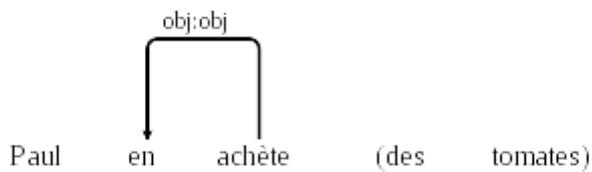
4.2.4 Clitique *en*

Le clitique *en*, quand il a une fonction syntaxique pleine par rapport à un verbe, est soit **de_obj**, **obj** ou **p_obj.o** ([voir corpus](#)).

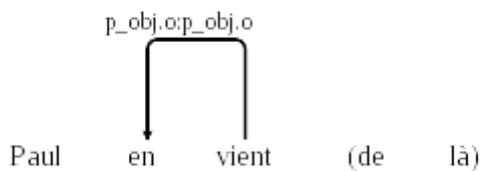
46. *Paul en parlera (du film)*



47. *Paul en achète (des tomates)*



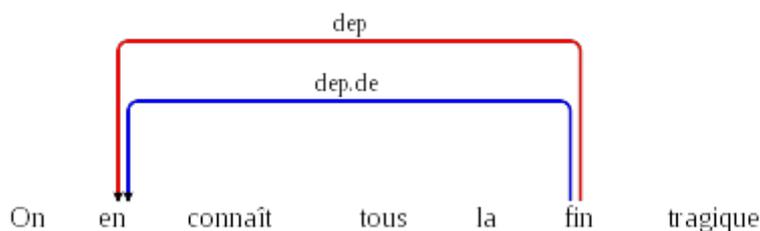
48. *Paul en vient (de là)*



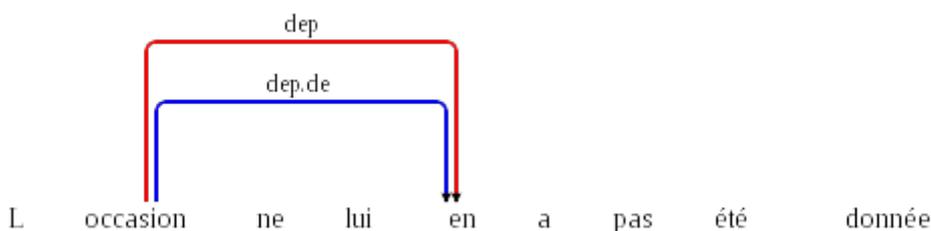
Le clitique *en* peut représenter un dépendant d'un dépendant du verbe hôte ([voir corpus](#)). Dans ce

cas, il a la fonction **dep** en surface et **dep.de** en profond, s'il dépend d'un nom ou d'un pronom et **de_obj**, s'il dépend d'un adjectif (il s'agit de cas de gouverneur non local). Ainsi par exemple le *en* peut dépendre de l'objet direct de son verbe hôte (49), de son sujet (50), ou de son attribut (51).

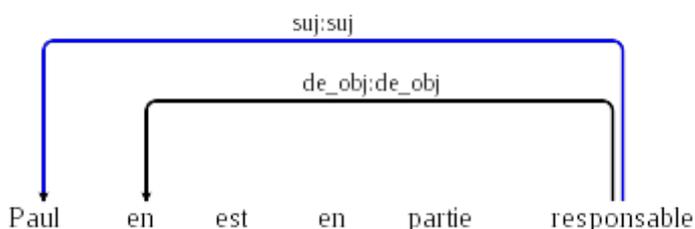
49. *On en connaît tous la fin tragique*



50. *L'occasion ne lui en a pas été donnée*



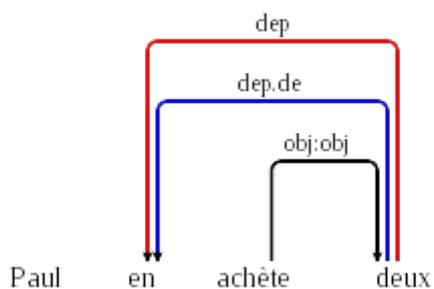
51. *Paul en est en partie responsable*



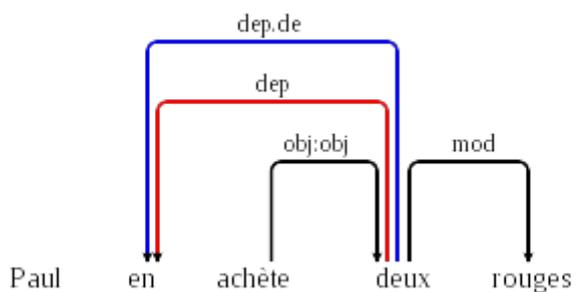
Pour le *en* quantitatif, il n'y a pas de précision du guide d'annotation du FTB. On a a priori deux solutions possibles : faire du *en* l'objet direct "principal", et le matériel restant postverbal dépendrait du *en*, ou bien l'inverse. La première solution capturerait bien la régularité qui est que le *en* reprend la partie nominale de l'objet direct. Mais les cardinaux post-verbaux sans nom sont traités comme des pronoms dans le schéma du FTB dont nous nous servons comme base, et donc il est alors plus cohérent de traiter le pronom post-verbal comme l'objet direct, et faire du *en* un dépendant de celui-ci : on choisit donc de prendre le pronom post-verbal comme tête, et le *en* en

dépend avec une relation **dep** en surface, et **dep.de** en profond.

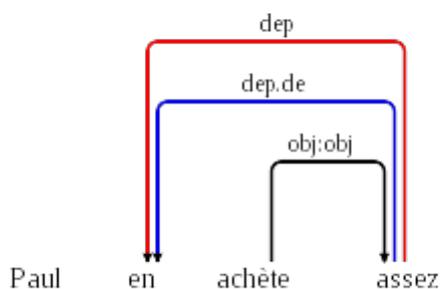
52. *Paul en achète deux*



53. *Paul en achète deux rouges*

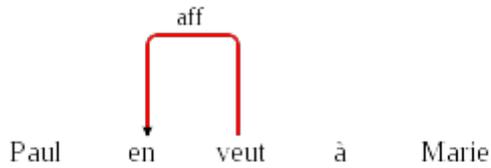


54. *Paul en achète assez*



Le clitique *en* peut également être figé et dans ce cas, on note sa dépendance au verbe **aff**.

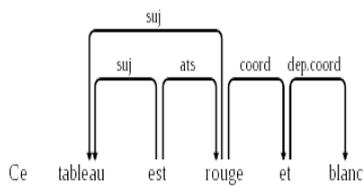
55. *Paul en veut à Marie*



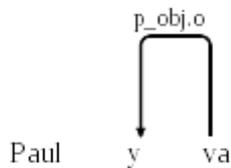
4.2.5 Clitique y

Le clitique y, quand il a une fonction syntaxique pleine par rapport à un verbe, est soit **a_obj**, **p_obj.o** dans le cas d'un locatif sous-catégorisé ou **mod** dans le cas d'un locatif non sous-catégorisé ([voir corpus](#)).

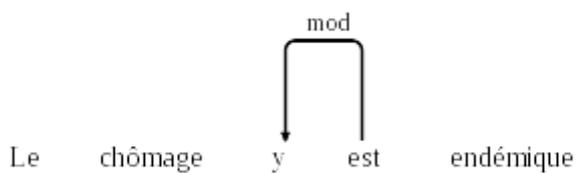
56. *Paul y pense*



57. *Paul y va*

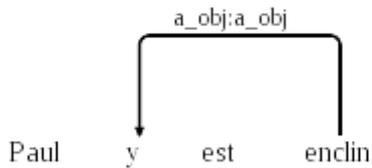


58. *Le chômage y est endémique*



Le clitique y peut être **a_obj** d'un adjectif qui est lui-même attribut du verbe auquel se rattache le clitique.

59. *Paul y est enclin*



Le clitique *y* peut également être figé ([voir corpus](#)).

60. *Il y avait trois statues dans le jardin*



4.2.6 Clitique *se*

Dans toute la suite on note clitique SE l'ensemble des formes du paradigme *me, te, se, nous, vous, se*, quelle que soit leur valeur grammaticale.

Le clitique SE est systématiquement considéré comme sémantiquement vide, et n'appartenant pas à la REPRSYNTPROF. On considère cependant 4 classes d'occurrences du SE, qui ont un impact sur la représentation profonde du verbe portant le SE et de ses arguments.

La classification du SE (la définition des critères, et la classification des occurrences sur le corpus Sequoi) ont fait l'objet d'un travail spécifique, à Alpage, avec un guide d'annotation spécifique (Candito, 2013). On donne ici une description succincte des 5 classes retenues et de leur représentation.

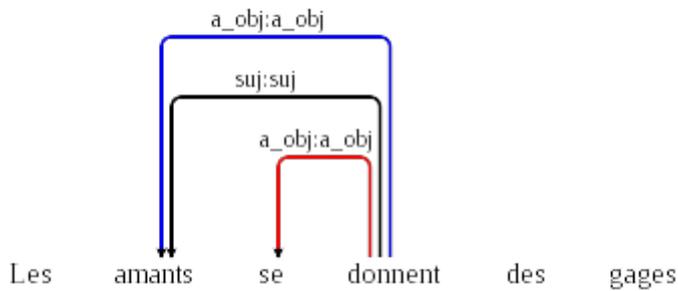
(voir également [Section 3](#) sur les changements de diathèse).

Vrai réfléchi ou réciproque

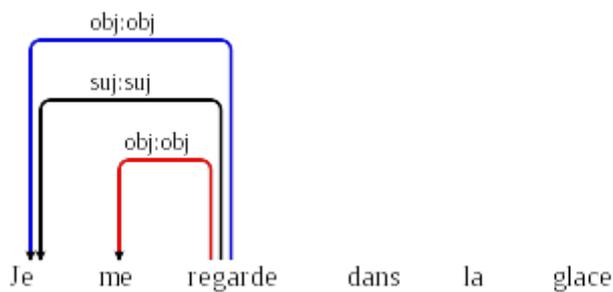
Le test principal est la possibilité $X \text{ SE } V \text{ W} \Rightarrow X \text{ ne } V \text{ que lui-même}$ ou $X \text{ ne } V \text{ qu'à lui-même}$ (Creissels, 2007), auquel on ajoute un test moins restrictif : la possibilité de forcer une lecture réciproque (en remarquant que si le réciproque est possible, alors le réfléchi doit l'être aussi).

Le clitique est alors **obj** ou **a_obj** en surface ([voir corpus](#)) et en représentation profonde, il est court-circuité au profit du sujet.

61. *Les amants se donnent des gages*



62. *Je me regarde dans la glace*



Réfléchi intrinsèque

On considère que le SE + V forme un seul lexème (utilisé pour les verbes dits "pronominaux", l'antipassif, l'autocausatif (Creissels, 2007)). En représentation de surface, le SE porte une dépendance **aff** ([voir corpus](#)). En représentation profonde, on supprime le SE et le V porte un lemme de la forme "se_XXX".

63. *Je me souviens de mes vacances*



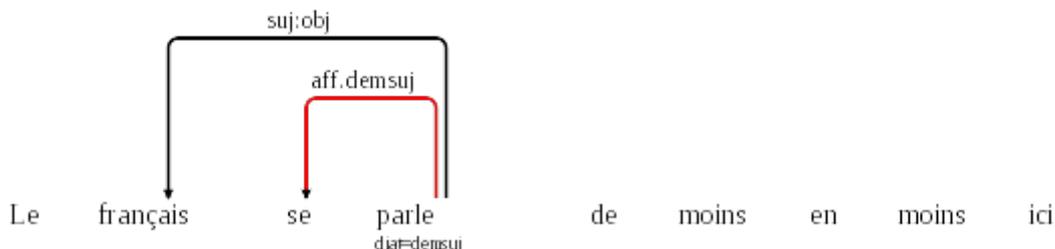
Se moyen ou se neutre

Le clitique SE peut être la marque du moyen ([Section 3.5](#)), du neutre ([Section 3.6](#)).

Le clitique SE moyen n'est pas distingué du SE neutre ([voir corpus](#)). Le SE porte la relation

aff.demsuj, et le verbe porte le trait *diat=demsuj* (redondance pour les repérer facilement...). Le sujet final du verbe est son objet canonique.

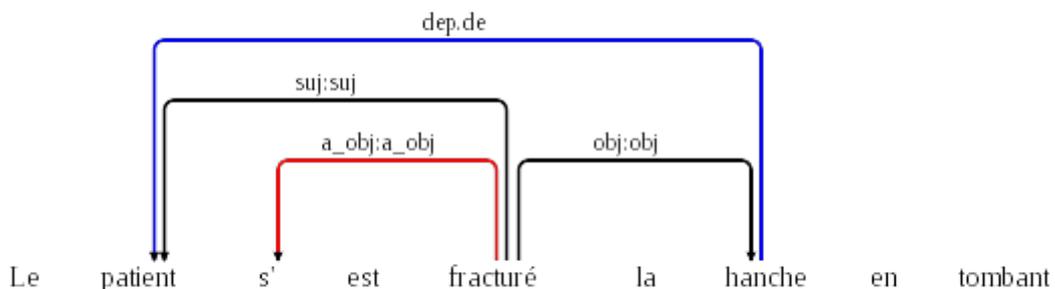
64. *Le français se parle de moins en moins ici*



Construction avec partie inaliénable

On a une 4ème classe. Les occurrences sont de la forme Xsuj SE V Yobj, avec le Yobj correspondant à "une partie inaliénable" de Xsuj, ce qui se traduit par l'équivalence avec "Xsuj V possessif Xobj". Le verbe n'admettant pas forcément en standard de **a_obj**, on ne peut pas analyser la construction comme un réfléchi a_objet. On explicite en profond un lien **dep.de** entre le Xsuj et le Yobj (on peut reconstruire *Yobj de Xsuj*).

65. *Le patient s'est fracturé la hanche en tombant* (=> *Le patient a fracturé sa hanche en tombant*)



Se faire Vinf

Enfin, on utilise une 5ème classe pour parmi les constructions en se + faire + Vinf, celles qui ne sont pas analysées comme un vrai réfléchi sur construction causative (cf. [Section 3.3.1](#))

On distingue ainsi :

66. *Les victimes se font rembourser par l'assurance* => classe vrai réfléchi

67. *Paul s'est fait tabasser en bas de chez lui.* => classe sefaire (repérable par le fait que le sujet ne peut pas être interprété comme un argument causateur).

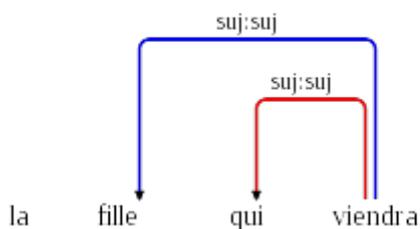
4.3 La fonction suj

4.3.1 Réalisation locale ou non locale

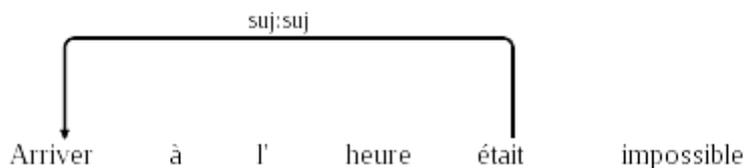
Le sujet (final) est obligatoirement réalisé localement pour un verbe conjugué, sauf à l'impératif. Il s'agit d'un groupe nominal, d'un pronom, d'un clitique ou d'une subordonnée (infinitive ou finie) ([voir corpus](#)).

Dans le cas de noms, de pronoms ou de clitiques il s'accorde en nombre et en personne avec le verbe. Il se cliticise en *il*, *on*, *ce* etc.

68. *la fille qui viendra*

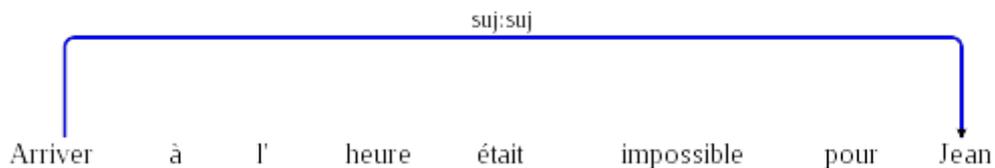


69. *Arriver à l'heure était impossible*



Les infinitifs n'ont pas de sujet final réalisé localement, mais il peut être réalisé ailleurs dans la phrase, de manière syntaxiquement contrôlée ou de manière arbitraire (voir la [Section 9.1.1](#)). Ce n'est pas le cas pour *arriver* dans la phrase *Arriver à l'heure était impossible*. Par contre, dans la phrase *Arriver à l'heure était impossible pour Jean*, *arriver* a pour sujet final *Jean*.

70. *Arriver à l'heure était impossible pour Jean*



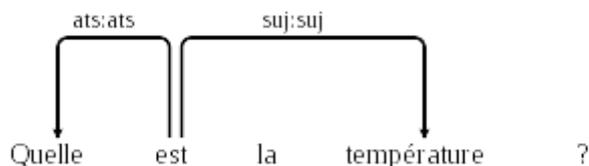
Par ailleurs, les changements de diathèse s'appliquent de manière régulière sur les infinitifs, ce qui

peut faire que le sujet final explicité pour un infinitif, ait une autre fonction canonique ([Section 9.1.1](#)). C'est ce qui a motivé le fait de ne pas parler de fonctions de surface / fonctions profondes, mais de fonctions canoniques / fonctions finales ([Section 1.2.3](#)).

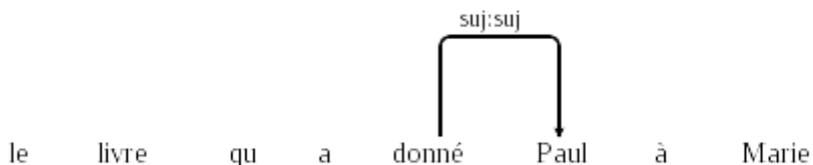
4.3.2 Inversion du sujet

On a des sujets inversés dans des phrases interrogatives, dans des relatives, et dans des contextes particuliers comme l'inversion locative, ou des incises de citation.

71. *Quelle est la température ?*



72. *le livre qu'a donné Paul à Marie*



73. *de cette attitude découle toute une série de malentendus*



74. *Monsieur est sorti, annonce le majordome*



Dans le cas de l'inversion complexe, on note deux fois la fonction sujet : cf. (38)

4.4 La fonction obj

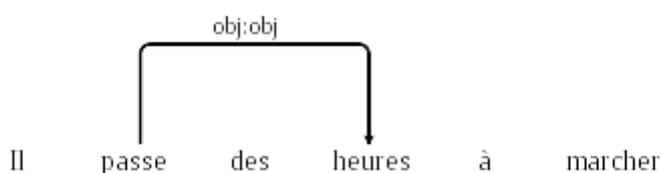
Les objets, notés **obj** (voir corpus), peuvent être nominaux ou pronominaux, clitiques (cf. Section 4.2), phrastiques ou infinitivaux (cf. Section 9).

Le test pour identifier l'objet est la réalisation clitique en *le/la/les* (avec accord contrairement à l'attribut), le *en* quantitatif, le caractère direct (sans préposition) dans le cas nominal.

4.4.1 Objets post-verbaux

Les objets se situent typiquement après le verbe dont ils dépendent.

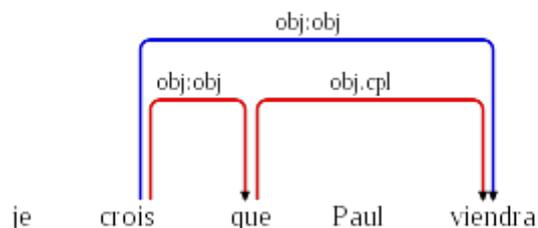
75. *Il passe des heures à marcher*



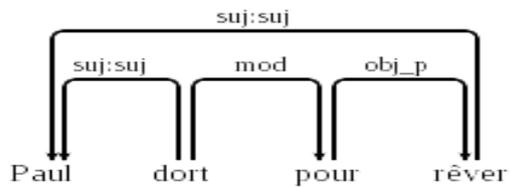
76. *Le président a pris position*



77. *je crois que Paul viendra*



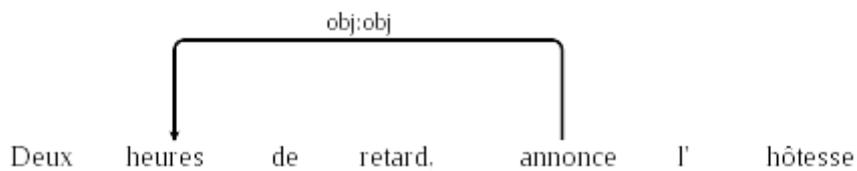
78. *Je sais à qui tu penses*



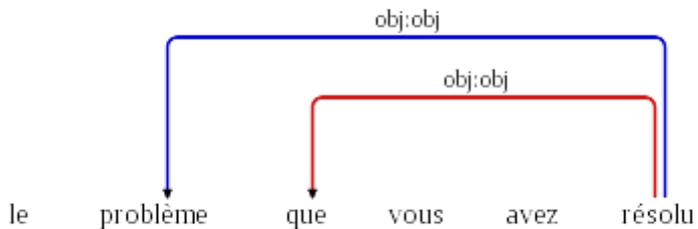
4.4.2 Objets pré-verbaux

Dans les cas de topicalisation, relativisation, cliticisation, exclamation, ils peuvent se situer avant le verbe dont ils dépendent.

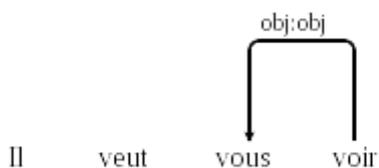
79. *Deux heures de retard, annonce l'hôtesse*



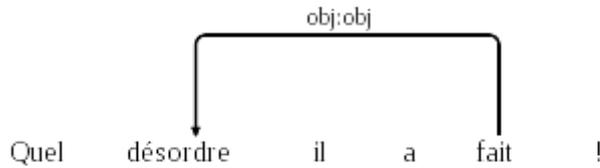
80. *le problème que vous avez résolu*



81. *Il veut vous voir*

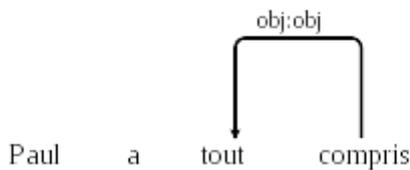


82. *Quel désordre il a fait !*



Certains pronoms exprimant la quantification sont, comme les clitiques, intégrés au noyau verbal. Leur position dépend du mode du verbe.

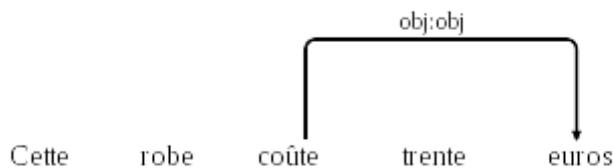
83. *Paul a tout compris*



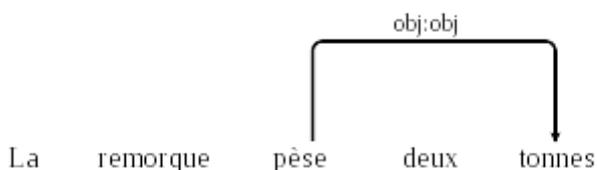
4.4.3 Objets directs non standards

Certains compléments nominaux obligatoires mais non cliticisables sont considérés comme objets directs, le test déterminant est alors le *en* quantitatif (cf. le guide du FTB).

84. *Cette robe coûte trente euros => cette robe en coûte trente (d'euros)*

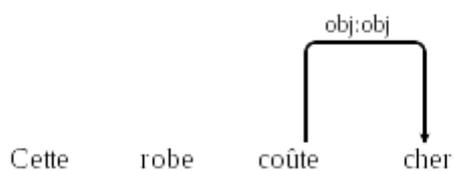


85. *La remorque pèse deux tonnes => La remorque en pèse deux (de tonnes)*



Du coup les adjectifs alternant avec ces objets directs sont également étiquetés objet direct :

86. *Cette robe coûte cher*

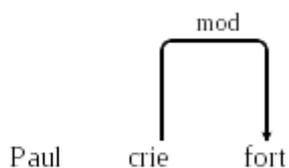


87. *La remorque pèse lourd*



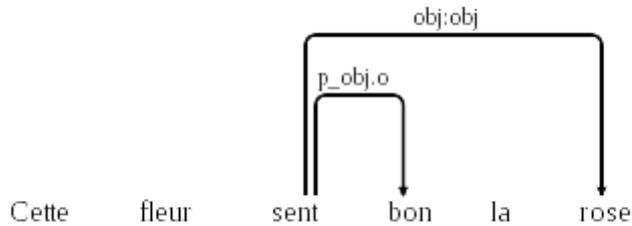
A ne pas confondre avec un adjectif modifieur : qui est non obligatoire, et peut être cooccurrent avec l'objet direct nominal (comparer *Marie crie fort sa joie* à **La remorque pèse lourd deux tonnes*) :

88. *Paul crie fort*



Enfin, on note **p_obj.o** un adjectif pouvant se substituer ou être cooccurrent avec l'objet direct nominal (cf. **mod** impossible du fait du caractère obligatoire, et **obj** impossible du fait de la cooccurrence possible avec l'objet nominal).

89. *Cette fleur sent bon la rose*

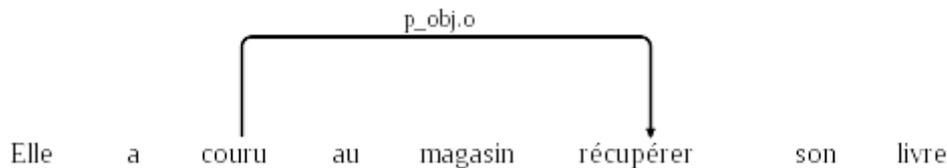


90. *Cette fleur sent bon*



Les **infinitives de but** des verbes de mouvement, lorsqu'elles sont directes, ne sont pas considérées comme **obj**, car elles ne remplissent pas tous les critères d'un objet direct. Elles sont considérées comme **p_obj.o**, cf. l'alternance avec une infinitive en *pour*.

91. *Elle a couru au magasin récupérer son livre*



4.5 Les constructions attributives et les fonctions attribut ats, ato

On suit complètement le guide du FTB pour les contours des fonctions **ats** et **ato**.

En représentation profonde, on explicite en outre le sujet des attributs verbaux et adjectivaux.

4.5.1 La fonction ats attribut du sujet

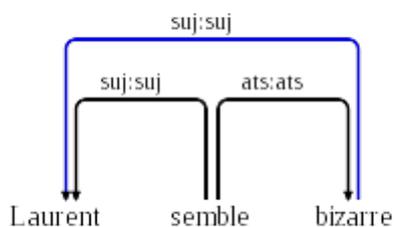
Pour l'**ats**, les caractéristiques sont la catégorie variable, la non mobilité et la cliticisation en *le*, sans accord avec l'antécédent :

92. *Ils sont restés capitaines toute leur vie => ils (le / *les) sont restés.*
 93. *Elles restent élégantes => elles (le / *les) restent.*
 94. *Ils ne paraissaient pas de bonne humeur => Ils ne le paraissaient pas*

Cette absence d'accord permet, dans le cas d'un attribut variant en genre et nombre, de faire la distinction avec l'objet direct, qui se cliticise aussi en *le*, mais avec accord (*J'ai vu les capitaines => je les ai vus*). Pour les attributs de catégories qui ne varient pas en genre et nombre (**ats** prépositionnel ou **ats** adverbial), la distinction se fait uniquement par la variabilité de la catégorie.

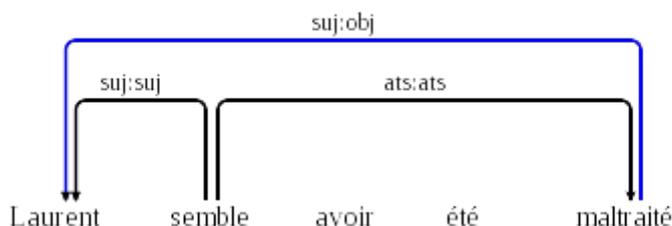
On explicite en représentation profonde la dépendance **su_j** (final) de l'attribut du sujet vers le sujet de cet attribut ([voir corpus](#)), dans le cas uniquement d'un attribut adjectival ou verbal. Le traitement des autres catégories n'est pas fait dans la version actuelle du schéma d'annotation, centré sur les verbes et les adjectifs.

95. *Laurent semble bizarre*



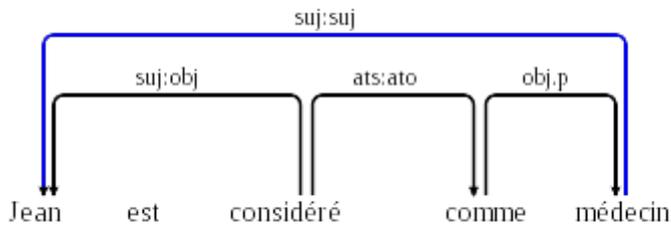
La régularité capturée est que l'attribut du sujet est toujours le sujet final du sujet du verbe hôte, mais si l'attribut a une diathèse non canonique, les changements de diathèse s'appliquent de manière régulière, comme par exemple ci-dessous :

96. *Laurent semble avoir été maltraité*



Le verbe ayant l'attribut du sujet peut être dans une diathèse non canonique : si le sujet final est objet canonique, alors l'**ats** final devient attribut de l'objet (**ato**) canonique ([voir corpus](#)).

97. *Jean est considéré comme responsable*



4.5.2 La fonction ato attribut de l'objet

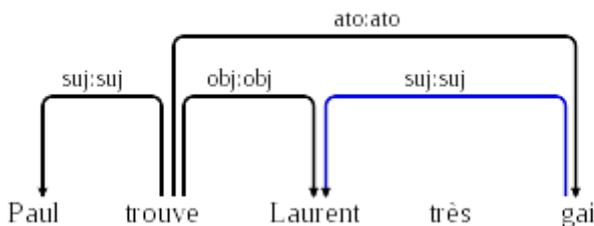
La fonction **ato** (attribut de l'objet) est utilisée pour les cas de prédication sur l'objet direct ([voir corpus](#)) : *X trouve/considère/nomme Y Z*, les infinitives introduites par un verbe de perception, ou bien par *laisser*.

Sa catégorie est variable (infinitival, prépositionnel, adjectival, nominal, adverbial...).

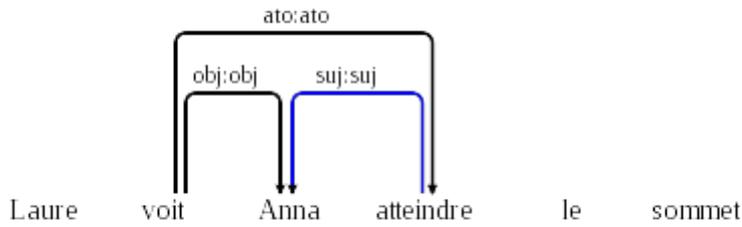
On explicite en représentation profonde la dépendance **suj** de l'attribut de l'attribut de l'objet vers l'objet, dans le cas d'un attribut verbal ou adjectival, et également dans le cas d'un attribut en *comme* + adjectif (les autres catégories ne sont pas traitées).

Concernant la dépendance entre le verbe et son objet direct, même si elle ne vaut pas forcément en sémantique, on garde l'objet direct comme objet canonique en dépendances profondes, cf. les exemples comme *Paul nomme Marie présidente*, où on a du mal à ne pas considérer *Marie* comme un argument de nommer. Cela s'applique également aux constructions avec verbe de perception : même s'il n'est pas clair qu'il existe une relation sémantique entre le verbe de perception et l'objet direct, on garde cette dépendance en dépendances profondes ([99](#)).¹

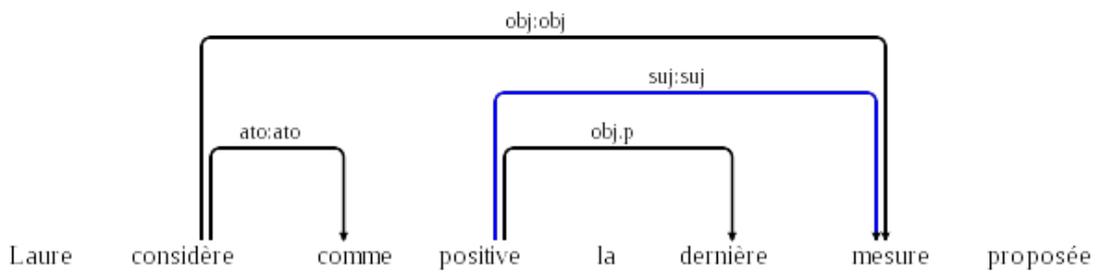
98. *Paul trouve Laurent fatigué par le voyage*



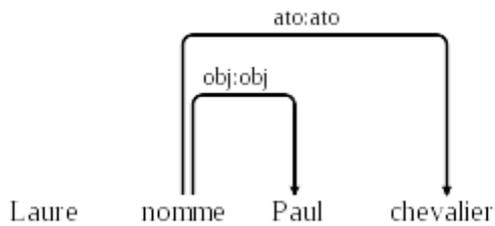
99. *Laure voit Anna atteindre le sommet*



100. *Laure considère comme positive la dernière mesure proposée*

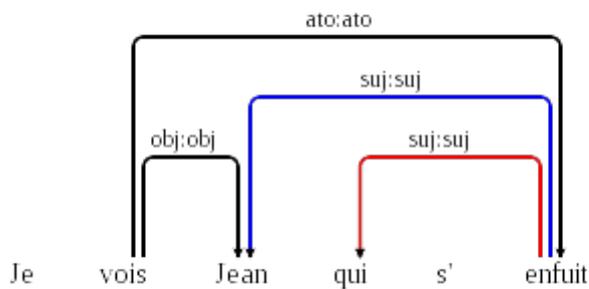


101. *Laure nomme Paul chevalier*



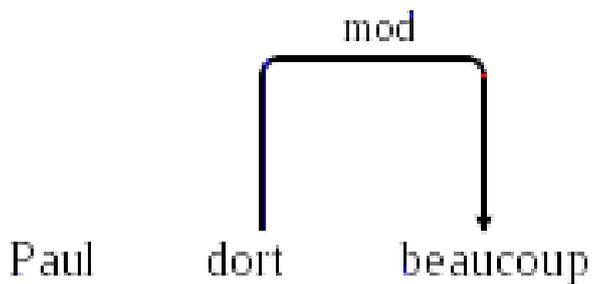
L'**ato** d'un verbe de perception peut être une relative introduite *par qui*. L'antécédent peut être pronominalisé (*Je le vois qui s'enfuit*). Dans ce cas, on court-circuite le relatif en représentation profonde, étant donné que son antécédent est syntaxiquement déterminé (c'est forcément l'objet du verbe principal).

102. *Je vois Jean qui s'enfuit*



Quand le verbe à attribut de l'objet est au passif, on a un **ats** en fonction finale, et un **ato** en fonction canonique (plus généralement quand le sujet final est objet canonique, l'**ats** final est l'objet canonique).

103. *Laurent est jugé coupable par tous*



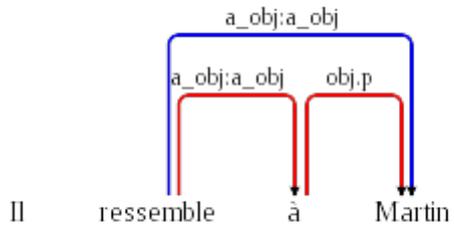
4.6 La fonction **a_obj**

Les objets indirects en *à*, notés **a_obj**, sont des compléments obligatoires introduits par la préposition fixe *à* ([voir corpus](#)), à distinguer des locatifs sous-catégorisés (pour lesquels la préposition peut varier, pourvu qu'elle soit locative).

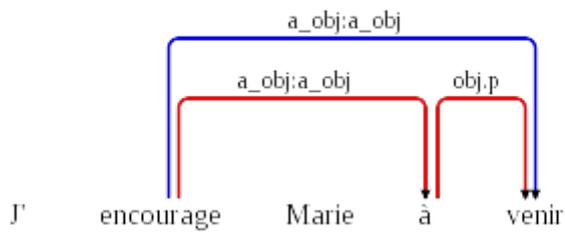
Le test pour identifier les **a_obj** est la cliticisation par *lui*, *leur* ou *y*, ET une relativisation en *où* impossible.

En représentation de surface, c'est le *à* qui porte la dépendance au verbe. En représentation profonde, on court-circuite le *à* :

104. *Il ressemble à Martin*

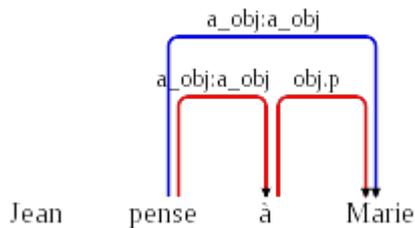


105. *J'encourage Marie à venir*

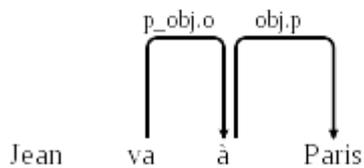


La cliticisation en y indique généralement un locatif sauf dans certains cas, repérés par l'impossibilité d'une question en où :

106. *Jean pense à Marie*



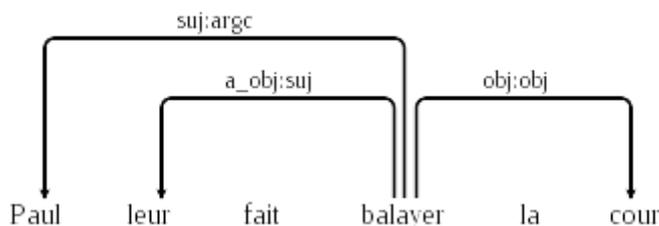
107. *Jean va à Paris*



Car on a pas *Où pense Jean ?* mais bien *Où va Jean ?*

Dans une construction causative (cf [Section 3.3](#)), un mot peut être l'objet indirect final et le sujet canonique du verbe causé.

108. *Paul leur fait balayer la cour*



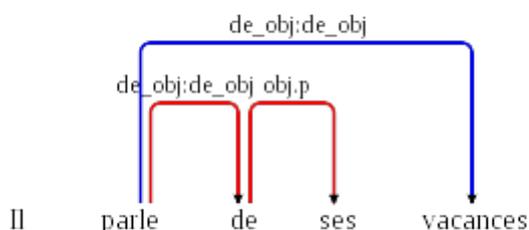
4.7 La fonction de_obj

Les objets indirects en *de*, notés **de_obj**, sont des compléments introduits par *de* (prep fixe), ou bien des compléments phrastiques ou infinitivaux alternant avec un complément nominal introduit par *de*.

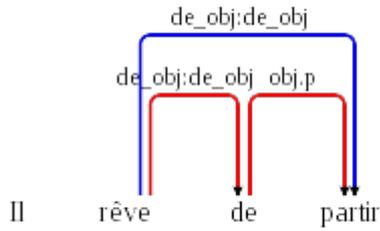
Le test pour identifier les compléments **de_obj** est la cliticisation en *en* ou la pronominalisation en *dont* ou en *d'où*.

En représentation de surface, le *de* porte la dépendance par rapport au verbe. En dépendances profondes, c'est l'objet de la prep qui porte la dépendance.

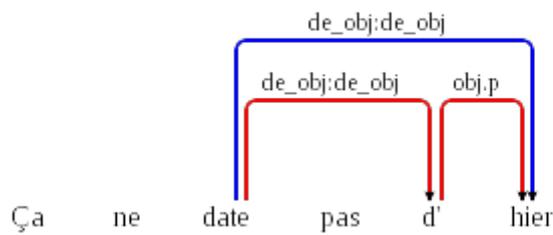
109. *Il parle de ses vacances*



110. *Il rêve de partir*



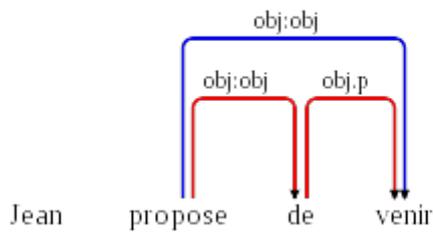
111. Ça ne date pas d'hier



Attention : les locatifs obligatoires introduits par *de* sont des **p_obj.o** pas des **de_obj**. On les repère par une pronominalisation en *d'où* (et clitique en *en*). Le *de* n'est pas supprimé (120).

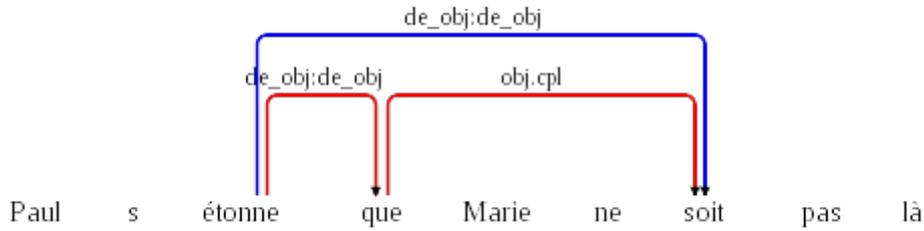
Attention : les infinitives introduites par *de* qui alternent avec *cela* sont notées **obj** (voir corpus).

112. Jean propose de venir



Les complétives introduites par *que* sont notées **de_obj** quand elles se pronominalisent en *en*, ou s'il peut y avoir une complétive en *de ce que*, ou une pronominalisation en *de cela*

113. Paul s'étonne que Marie ne soit pas là



4.8 Les fonctions p_obj*

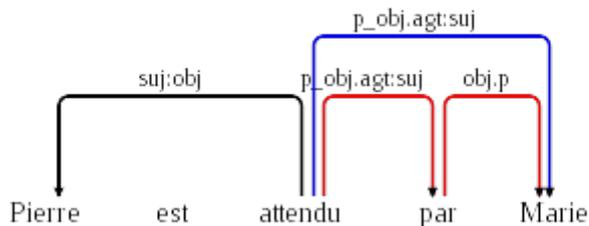
Les objets indirects sous-catégorisés ne satisfaisant ni **de_obj** ni **a_obj** sont notés **p_obj.***.

Par rapport à la fonction **p_obj** du FTB, on distingue **p_obj.agt** et **p_obj.o**.

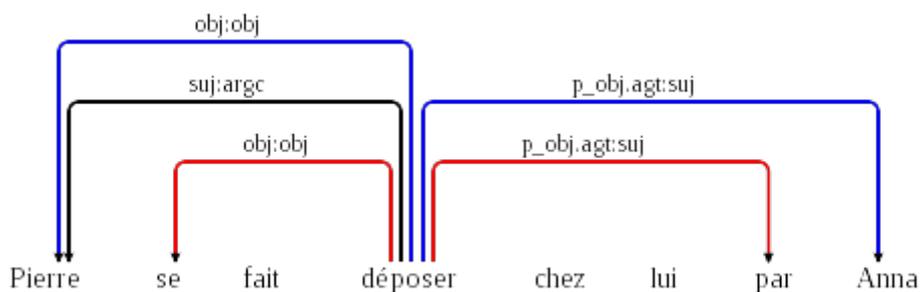
4.8.1 La fonction p_obj.agt

p_obj.agt est utilisé pour le complément d'agent du passif ou du causatif (prep *par* et plus rarement *de* comme dans *Il a été compris de tous les fidèles*).

114. *Pierre est attendu par Marie*



115. *Pierre se fait déposer chez lui par Anna*



4.8.2 La fonction **p_obj.o**

Parmi les compléments **p_obj.o**, on ne fait pas de distinction entre les compléments locatifs et les autres ([voir corpus](#)).

Par contre, les compléments **p_obj.o** sont requis par le verbe dont ils dépendent et se distinguent des modificateurs, notés **mod**. Pour effectuer la distinction, on utilise les tests suivants :

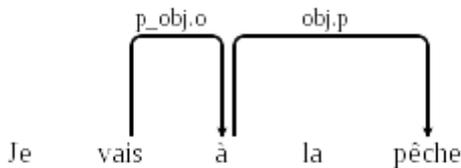
- Si obligatoire alors **p_obj.o**,
- Si non déplaçable en tête de phrase en gardant le même sens alors **p_obj.o**

A Paris, je mange des spaghettis => **mod**

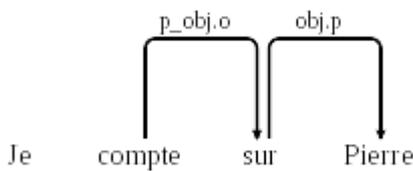
A paris, je vais* => **p_obj.o

Mais attention : *Je cours à Paris* => **p_obj.o** car *A Paris, je cours* est possible, mais le sens est modifié.

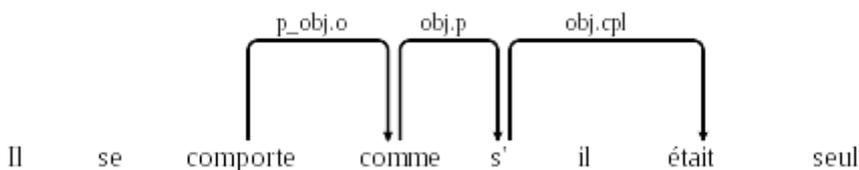
116. *Je vais à la pêche*



117. *Je compte sur Pierre*

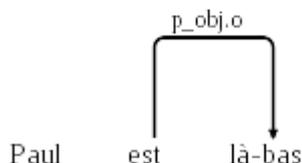


118. *Il se comporte comme s'il était seul* => ici *comme* est préposition, et prend une complétive comme objet



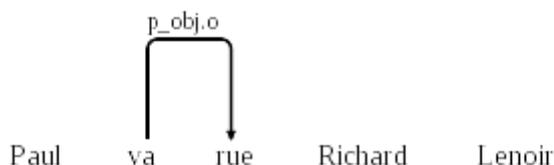
Les adverbes obligatoires et les NPs locatifs ont également la fonction **p_obj.o** lorsqu'ils sont obligatoires.

119. <ex>Paul est là-bas



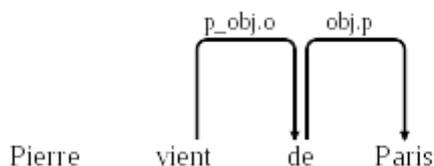
</ex>

120. Paul va rue Richard Lenoir



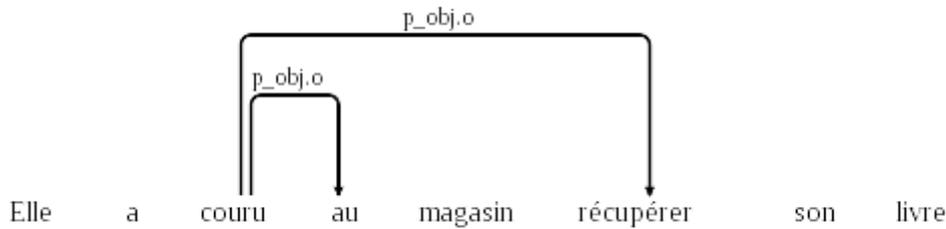
On utilise **p_obj.o** également pour les locatifs obligatoires introduits par *de*. On les repère par une pronominalisation en *d'où* (et clitique en *en*). Le *de* n'est pas supprimé.

120. Pierre vient de Paris

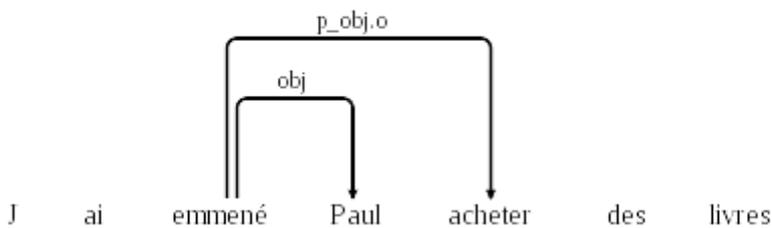


Les infinitives de but des verbes de mouvement (121), sont considérées comme **p_obj.o**. En effet, elles ne peuvent pas être **obj**, cf. le cas des transitifs de mouvement 122, et l'absence de cliticisation en *le*. Le guide du FTB les classe comme **mod**, mais nous privilégions leur caractère non mobile, leur unicité et l'alternance avec un pour-Vinf pour les classer plutôt **p_obj.o**.

121. Elle a couru au magasin récupérer son livre

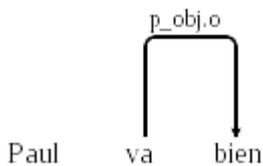


122. *J'ai emmené Paul acheter des livres*



Les adverbes de manière obligatoires et à place fixe sont également notés **p_obj.o**.

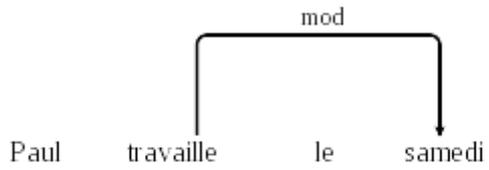
123. *Paul va bien*



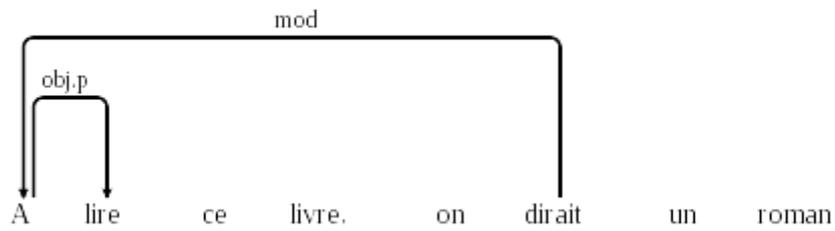
4.9 Les fonctions mod*

La relation **mod** couvre tous les dépendants non sous-catégorisés du verbe (et des autres catégories ce qui n'est pas détaillé dans cette section). Les modifieurs sont optionnels et généralement mobiles (en particulier déplaçables en tête de phrase). Leurs catégories sont variables : il peut s'agir de PP, d'adverbes, d'adjectifs, de NPs, d'infinitives, de relatives extraposées ou de subordonnées circonstancielles.

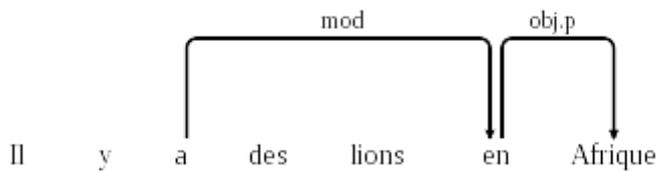
124. *Paul travaille le samedi*



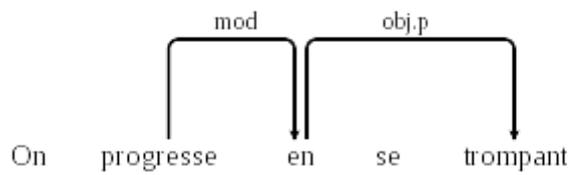
125. *A lire ce livre, on dirait un roman*



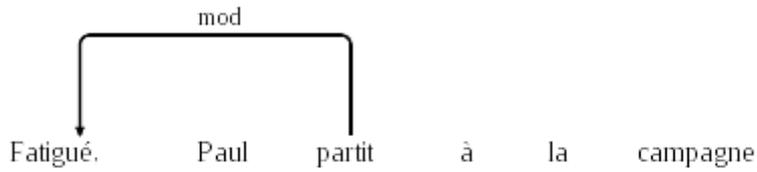
126. *Il y a des lions en Afrique*



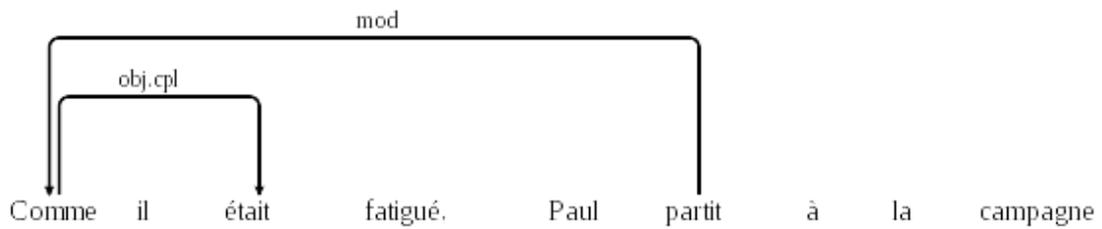
127. *On progresse en se trompant*



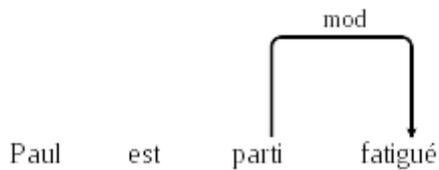
128. *Fatigué, Paul partit à la campagne*



129. *Comme il était fatigué, Paul partit à la campagne*

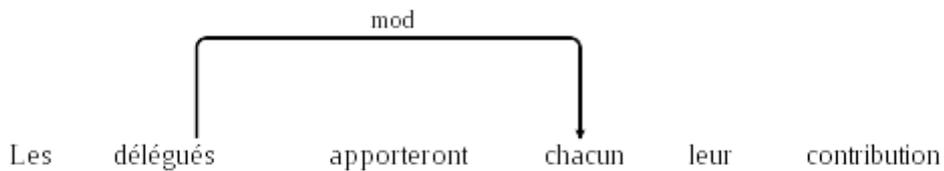


130. *Paul est parti fatigué*

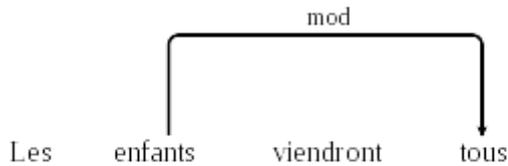


Les quantifieurs flottants sont également notés **mod** mais ils modifient le nom qui est quantifié !

131. *Les délégués apporteront chacun leur contribution*



132. *Les enfants viendront tous*



- 1. La question est par exemple *Paul* est-il argument sémantique de *voir* dans *Laure voit Paul dévaler les escaliers*. Des exemples comme *J'entends le prof se faire chahuter par ses élèves* (Muller, 95) où l'objet direct n'est pas à l'origine de la perception font douter de la relation sémantique entre l'objet et le verbe de perception. Egalement (Skytte et al. 91) citent comme acceptable l'équivalent italien de *Je vois le vent remuer les feuilles*. Comme argument contraire, on peut rapprocher la phrase *Je vois le vent remuer les feuilles* de *Je vois le vent qui remue les feuilles* où il n'y a pas d'ambiguïté. Dans le doute on conserve donc la relation objet au niveau profond (*vent* est objet de *voit* et sujet de *remue*).

5 Dépendants du nom

5.1 Déterminants

5.1.1 Cas de base

Les déterminants dépendent des noms qu'ils déterminent par une dépendance de type **det**.

En outre le caractère défini du déterminant est reporté sur le nom sous la forme d'un trait $def=y$.

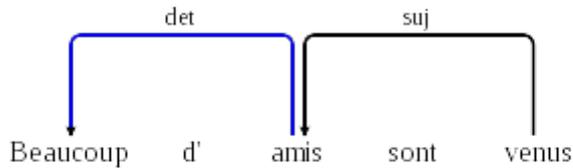
5.1.2 Déterminant contracté avec une préposition

La catégorie P+D est réservée aux prépositions contractées. En termes de dépendance, cette catégorie est traitée comme une préposition. Le caractère défini du déterminant est reporté sur le nom sous la forme d'un trait $def=y$.

5.1.3 Adverbes de quantité ou comparatifs : "plus / beaucoup de N"

On traite de la même manière le cas avec adverbe de quantité "beaucoup/peu/énormément... de N" et le cas avec adverbe comparatif "plus/moins/... de N" : l'adverbe porte est un dépendant de type **det** du nom. Le *de* est supprimé en représentation profonde.

133. *Beaucoup d'amis sont venus*

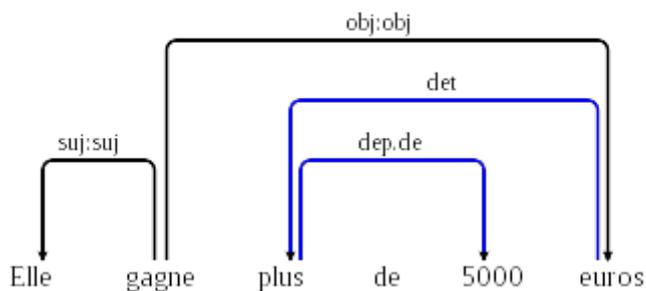


Attention, le cas "*beaucoup/peu/... de* DET N" est traité différemment (le quantifieur est traité comme un pronom, voir [Section 7.2.2](#)).

5.1.4 Déterminants cardinaux

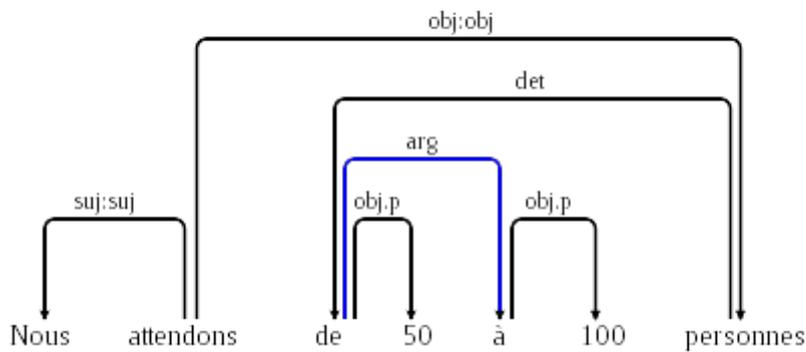
Pour les déterminants cardinaux avec adverbes comparatifs : *plus de CARD N*, on a en surface une structure plate, avec *plus*, *de*, et le card qui dépend du nom. En représentation profonde, on considère que c'est l'adverbe comparatif qui est tête de la détermination. Le *de* est court-circuité (de ce fait, on traite ici l'adverbe comparatif au sein du NP, comme un **det**, et pas comme un **mod.comp** comme dans les autres comparatives cf. [Section 14](#)) ([voir corpus](#)).

134. Elle gagne plus de 5000 euros



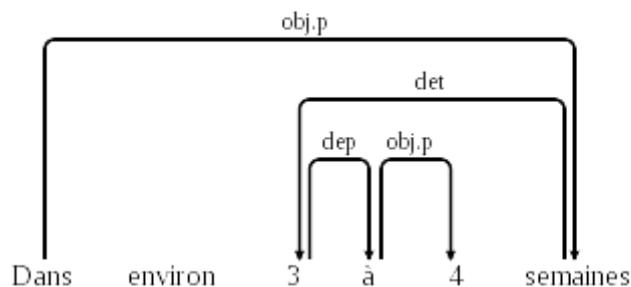
Concernant les "plages" : les NP de la forme "*de* CARD à CARD N" (*de 3 à 5 personnes*), en représentation profonde, on choisit de représenter un "déterminant complexe" pour "*de* X à Y", dont la tête est la préposition *de* (qui dépend du nom avec une dépendance **det**) et celle-ci gouverne la préposition *à* à l'aide d'une dépendance **arg**.

135. Nous attendons de 50 à 100 personnes => en représentation de surface on a une structure plate, où "*de*", "*50*", "*à*" et "*100*" dépendent de "*personnes*". En représentation profonde, c'est le "*de*" qui est considéré comme déterminant du nom (cf. le traitement des prépositions liées *de NP à NP*, [Section 6.1](#))



L'exemple précédent est à distinguer d'un NP de la forme "CARD1 à CARD2 N" (3 à 4 personnes), où on se contente de rattacher le "à CARD2" au premier CARD (cf. pas de préposition "de"):

136. Dans environ 3 à 4 semaines



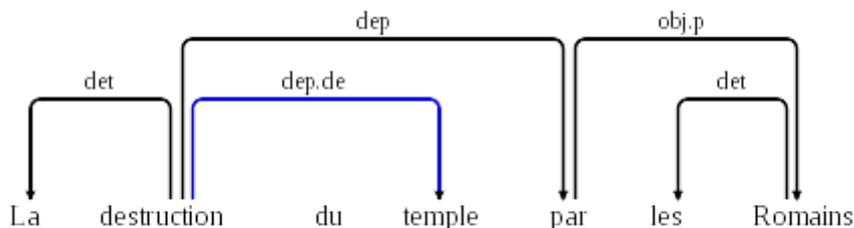
5.2 Groupes prépositionnels

Pour les groupes prépositionnels dépendant du nom, nous n'avons pas fait le travail de définition des critères pour distinguer entre arguments et ajouts. C'est pourquoi, en surface, on code systématiquement une relation générique **dep**.

En syntaxe profonde, on conserve la préposition et les dépendances sauf lorsque la préposition est *de* : elle est considérée systématiquement comme sémantiquement vide (même si elle peut marquer différents types syntaxiques de dépendant : le sujet du nom, l'objet du nom, le possesseur...), et court-circuitée. On doit cependant conserver l'information que le nom est introduit par *de*, et on spécialise pour cela la fonction **dep** en **dep.de**.

En outre, ce traitement est appliqué pour les prépositions contractées en *de* (*du*, *des*). (Le caractère défini du déterminant compris dans les prépositions contractées est reporté sous forme de trait *def=y* sur le nom, cf. [Section 2.1.4](#) sur les traits profonds)

137. La destruction du temple par les Romains a eu lieu à l'aube de ce millénaire.



Pour les dépendants *que + Phrase*, on utilise également **dep.de** (voir [Section 9.2](#) sur les dépendants phrastiques), du fait de l'alternance :

138. *L'idée de son départ m'est insupportable*

139. *L'idée qu'il parte m'est insupportable*

5.3 Epithètes et appositions

Les adjectifs, noms et participes épithètes sont systématiquement **mod** en surface. Cette dépendance est toujours conservée en représentation profonde, y compris dans le cas où on ajoute une dépendance inverse : pour les participes et adjectifs, le nom modifié est leur dépendant profond (voir [Section 8.1](#)).

Dans le cas d'un nom dépendant d'un nom, on distingue les noms épithètes et les appositions.

L'apposition est séparée du nom tête par une virgule (ou apparaît entre parenthèses, ou un « : »), et a sémantiquement la même référence que le nom qu'elle modifie (on n'a donc que des appositions nominales).

On marque l'apposition par une dépendance **mod.app**, versus **mod** pour le nom épithète.

Dans le cas d'un NP avec plusieurs noms n'étant pas des appositions, c'est le premier nom qui est choisi comme tête (*M. Paul Duchmuc* => la tête est « M. », les deux autres noms propres dépendent de *M.* avec une relation **mod**).

140. *Le général Pierre Duchmuc est parti* => **mod**(général, Pierre), **mod**(général, Duchmuc)

141. *Le général, Pierre Duchmuc, est parti* => **mod.app**(général, Pierre), **mod**(Pierre, Duchmuc)

Plus précisément, la tête nominale choisie est le premier nom après un éventuel déterminant : dans *M. le général Pierre Duchmuc*, tous les éléments dépendent de *général*

Distinction entre apposition et énumération : Attention à distinguer l'apposition d'une énumération : dans le premier cas, les 2 noms réfèrent à la même entité, alors que dans l'exemple suivant *professeurs* et *ministres* réfèrent à deux référents distincts.

142. *J'ai connu d'illustres professeurs, des ministres compétents* => **obj**(connu, professeurs), **obj**(connu, ministres) (=énumération)

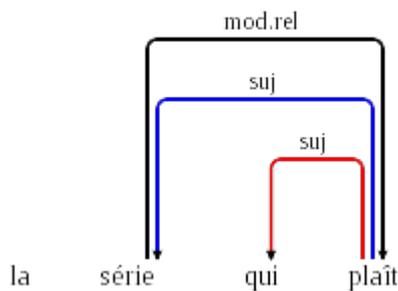
5.4 Relatives

5.4.1 Relatives avec antécédents

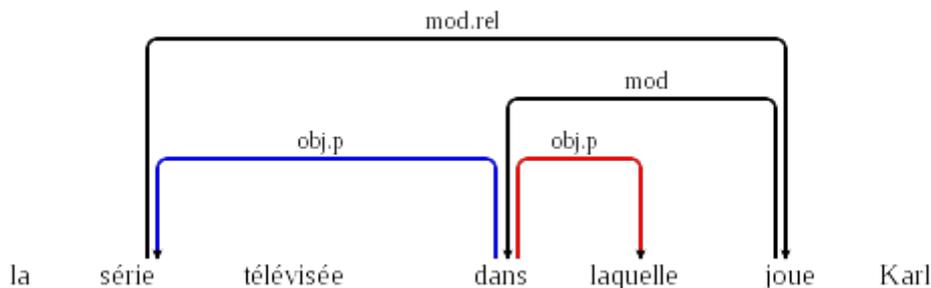
On part du schéma en dépendances de surface dans lequel les relatives épithètes dépendent avec une relation **mod.rel** de l'élément (en général le nom) qu'elles modifient. La tête de la relative est le verbe (y compris pour relatives en *qui* et *que*: les relatifs ne sont pas traités comme des compléments).

En représentation profonde, on a choisi de résoudre les anaphores structurellement résolubles, et le pronom relatif est donc dans une relative épithète systématiquement court-circuité : la ou les relations entrantes ou sortantes sont reportées sur l'antécédent.

143. *la série qui plaît*



144. *la série télévisée dans laquelle joue Karl*



Cas de dont

Dont peut correspondre en surface à un complément ou modifieur du verbe de la relative (145), (146) ou d'un verbe enchâssé au sein de la relative 147, ou bien correspondre à un dépendant d'un NP interne à la relative.

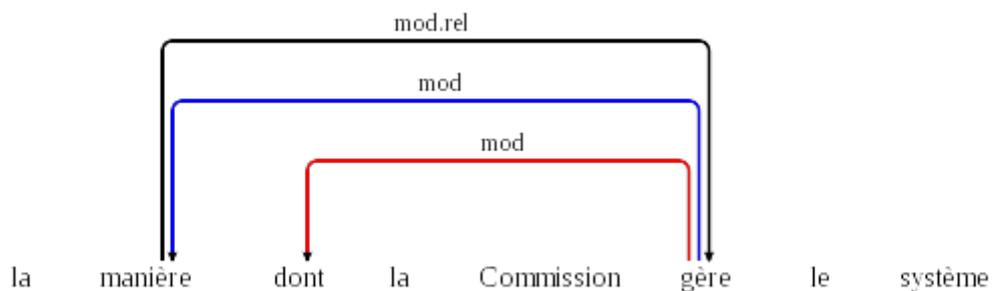
Comme pour toutes les relatives épithètes, la relation est reportée en dépendances profondes sur le nom modifié par la relative.

Lorsqu'il est dépendant d'un verbe, *dont* est soit **de_obj**, soit **p_obj.o** dans le cas où il est interrogeable par *d'où*, cf. exemple (148).

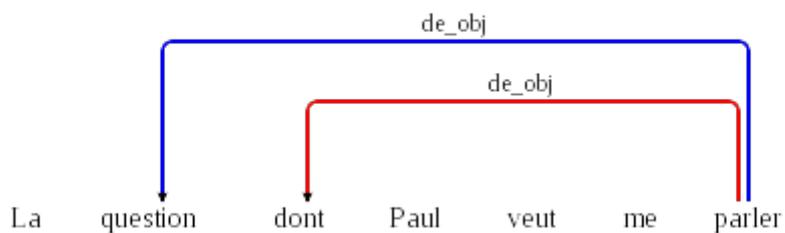
145. *la série dont Pierre parle*



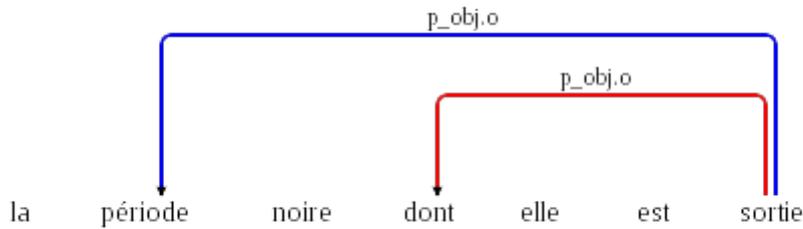
146. *la manière dont la Commission gère le système*



147. *La question dont Paul veut me parler*



148. *la période noire dont elle n'est pas sortie*



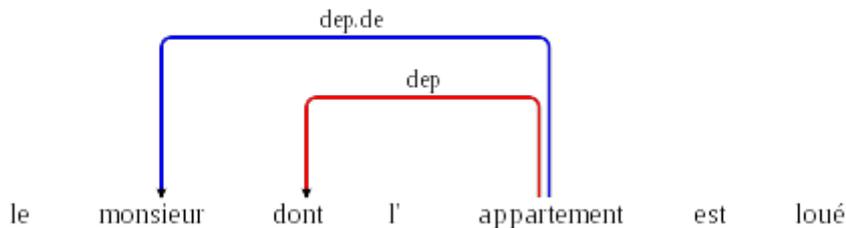
Dont est également fréquemment un élément extrait hors d'un NP sujet (149) ou objet (150).

On rappelle que dans ces cas d'extraction ont été manuellement annotés, en corrigeant la conversion automatique en dépendances de surface du corpus Sequoia (Candito et Seddah, 2012b).

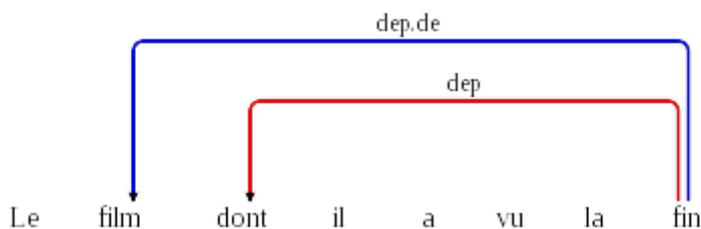
La dépendance en surface de *dont* est alors **dep** (cf. on ne gère pas la distinction argument / ajout pour les dépendants d'un nom).

En représentation profonde, cette dépendance est reportée sous forme d'un **dep.de** vers l'antécédent de la relative (cf. le *dont* correspond à un dépendant en *de* du nom).

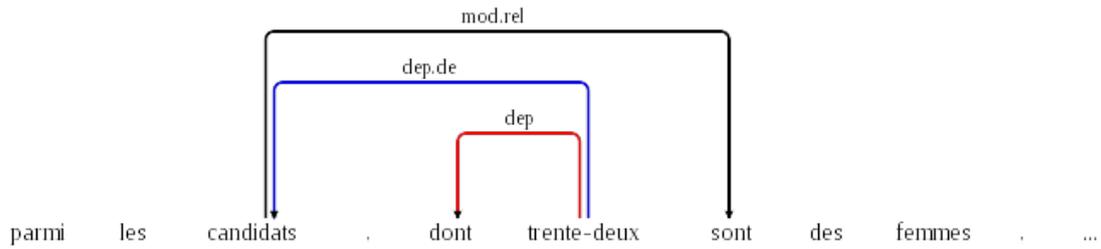
149. *le monsieur dont l'appartement est loué*



150. *Le film dont il a vu la fin*

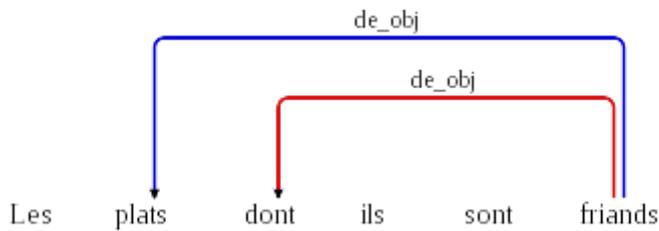


151. *parmi les candidats, dont trente-deux sont des femmes, ...⇒* (cf. on peut reconstruire "trente-deux des candidats sont des femmes")



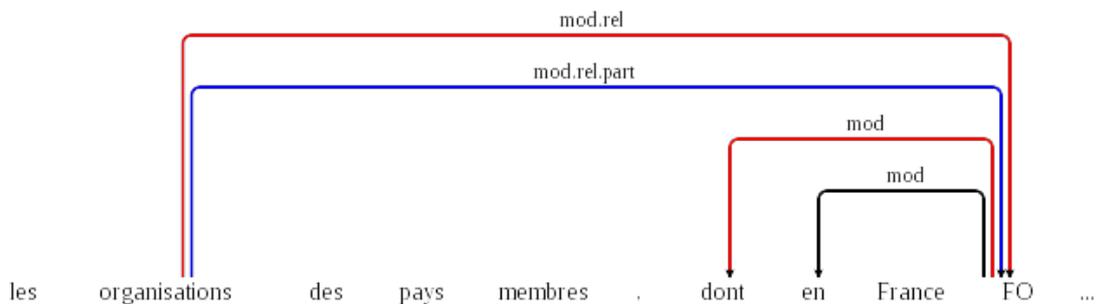
On peut aussi avoir extraction hors d'un attribut, en particulier adjectival, auquel cas on gère de manière régulière la distinction argument / ajout. Ainsi *dont* est **de_obj** de *friand* dans l'exemple suivant :

152. *Les plats dont ils sont friands*



Quand la relative en *dont* est sans verbe (tour elliptique), la tête de la relative est en surface le NP introduit par *dont*, et *dont* en est un modifieur (cf. les spécifications du schéma Easy). En surface, les éventuels modifieurs (par exemple "en France" pour l'exemple 153) sont rattachés à cette tête nominale (*FO*), ce qui ne rend pas compte du fait que la modification concerne la prédication sous-entendue "d'existence d'une sous-partie de l'ensemble dénoté par l'antécédent" :

153. *les organisations des pays membres, dont en France FO ...*



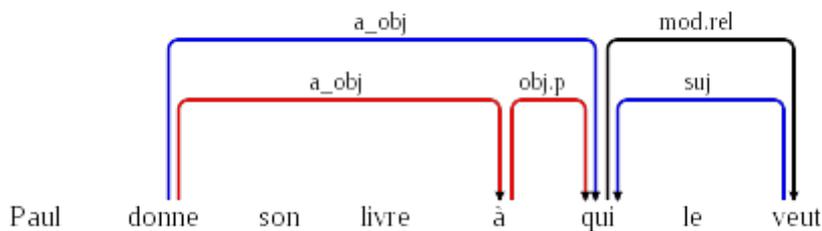
5.4.2 Relatives sans antécédent

Le pronom relatif est l'élément déterminant la distribution de NP des relatives sans antécédent, **il est donc considéré comme tête de la relative** (dans les exemples ci-dessous, *qui* est la tête de *qui le veut*, et *quiconque* est la tête de *quiconque entre ici*).

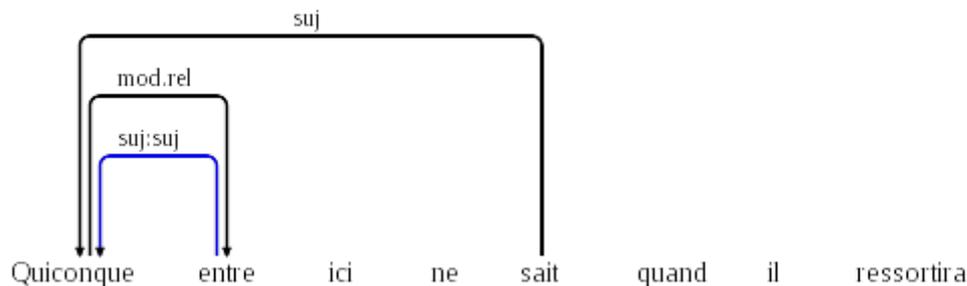
Le verbe de la relative en est un dépendant en surface, avec la relation **mod.rel**.

En dépendances profondes, on ajoute seulement la dépendance du pronom relatif envers le verbe de la relative :

154. *Paul donne son livre à qui le veut*



155. *Quiconque entre ici ne sait quand il ressortira*



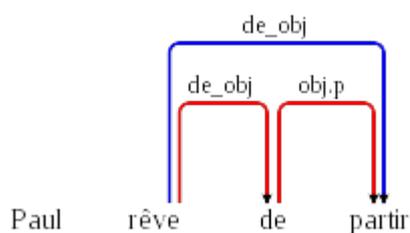
6 Dépendants de la préposition

En dépendances de surface, on note systématiquement la préposition comme gouverneur de l'élément qu'elle introduit, que la préposition soit sémantiquement pleine ou vide.

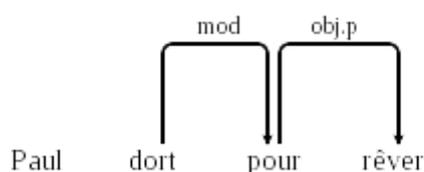
Le complément de la préposition est en général un objet nominal (*pour la gloire*), infinitival (*pour chanter*) ou phrastique (*pour que Paul chante*). Dans ces 3 cas on utilise la fonction **obj.p** pour "objet de préposition" pour le différencier de **obj** qui est un objet de verbe).

En dépendances profondes, on rappelle que les prépositions régies (ou complémenteurs *à* et *de* codés comme prépositions) sont court-circuitées, et n'apparaissent pas en représentation profonde (156), alors que les prépositions pleines gardent la même représentation qu'en surface (157).

156. *Paul rêve de partir*

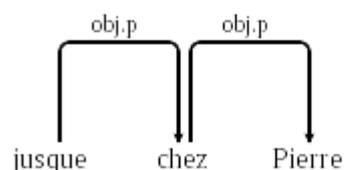


157. *Paul dort pour rêver*



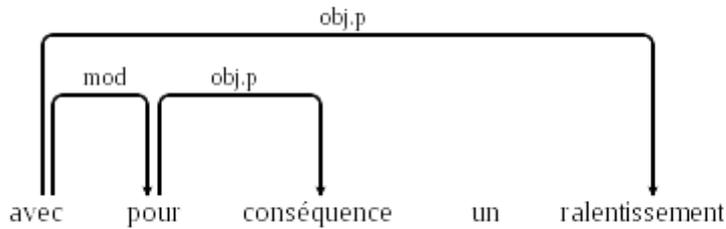
Par commodité, on utilise cette même fonction **obj.p** y compris dans le cas marginal où l'unique complément de la préposition est lui-même prépositionnel ou adverbial, comme c'est le cas pour la préposition *jusque*:

158. *jusque chez Pierre*



Mais attention, à l'exemple suivant : le complément **obj.p** de *avec* est le complément direct *un ralentissement*, l'autre PP (en *pour*) en est ici un modifieur:

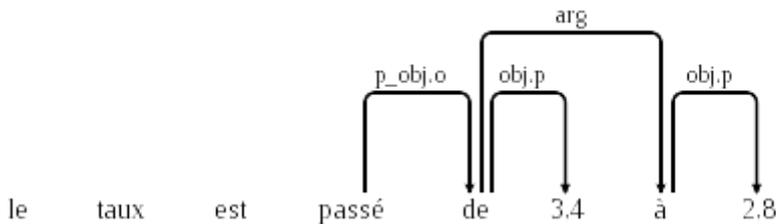
159. *avec pour conséquence un ralentissement*



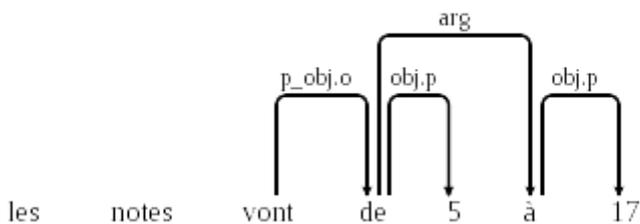
6.1 Préposition liées

Pour les séquences de la forme "de NP en NP" ou "de NP à NP", où l'une ou les 2 prépositions ne peuvent pas apparaître l'une sans l'autre, on note la deuxième préposition comme dépendant sous-catégorisé (**arg**) de la première.⁴

160. *le taux est passé de 3,4 à 2,8*



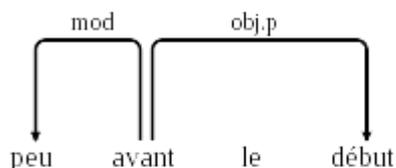
161. *les notes vont de 5 à 17*



6.2 Modifieurs de préposition

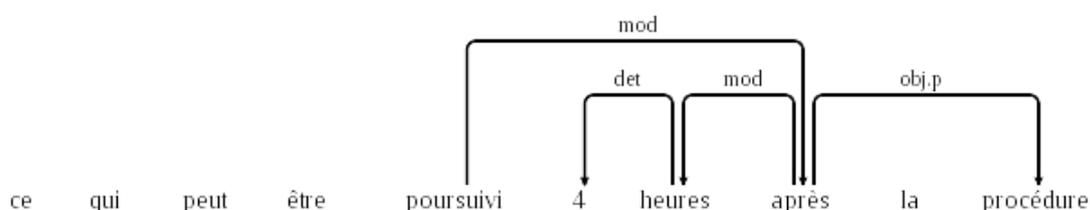
La préposition peut être modifiée par un adverbe :

162. *peu avant le début*



ou un NP temporel :

163. *ce qui peut être poursuivi 4 heures après la procédure*



- 4. cf. le guide d'annotation en constituants du FTB, p.24: *de X à Y* est prépositionnel uniquement si elle commute avec un "vrai" SP : *le pays a besoin de 100 à 200 millions => SP, le pays dépensera de 300 à 400 millions => SN*; l'équivalent en dépendances donne la prep *de* comme tête dans le premier cas, et *millions* dans le second).

7 Dépendants de l'adjectif, de l'adverbe

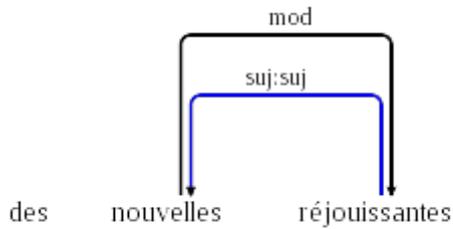
7.1 Dépendants de l'adjectif

7.1.1 Sujet profond d'un adjectif

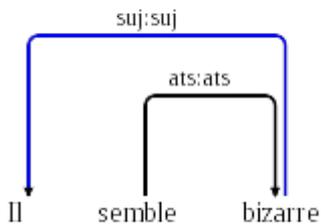
Quelle que soit la fonction d'un adjectif, il exprime une propriété s'appliquant à un nom ou un pronom, ce qui est marqué par une dépendance profonde de l'adjectif vers ce nom ou pronom. Cette dépendance est appelée "sujet" par analogie avec le sujet du verbe, notamment quand ce dernier est au participe passé, et elle est notée **suj** ([voir corpus](#)).

Dans les exemples suivants, l'adjectif a différentes fonctions mais toujours un sujet profond.

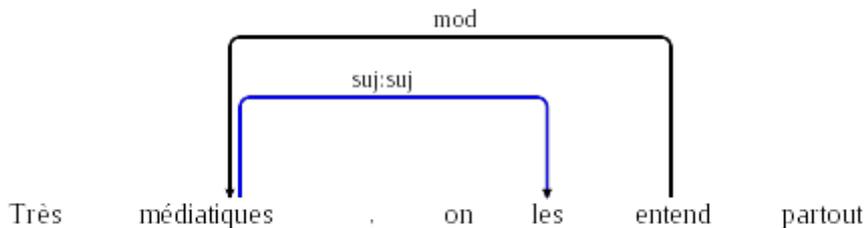
164. *des nouvelles réjouissantes*



165. *Il semble bizarre*



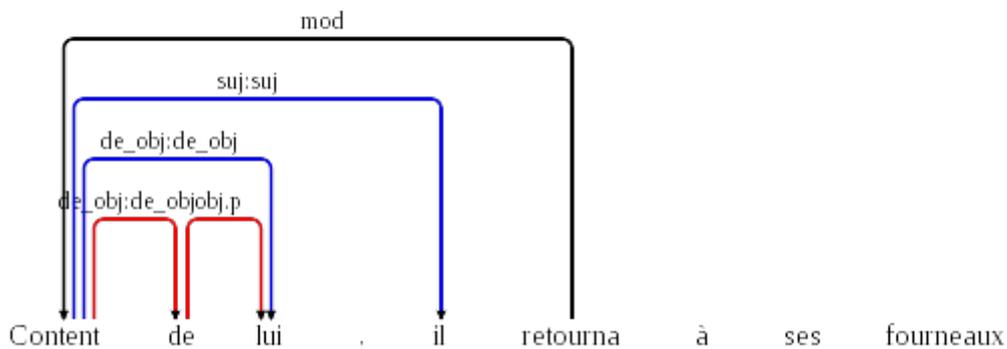
166. *Très médiatiques, on les entend partout*



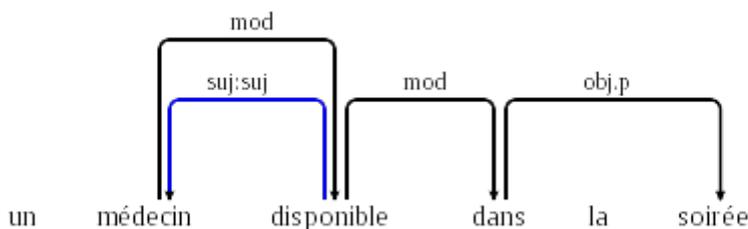
7.1.2 Compléments requis et adjoints d'un adjectif

Les compléments de l'adjectif sont catégorisés de la même façon que ceux du verbe en complément requis par l'adjectif ou compléments adjoints. Dans le premier cas, les dépendances sont étiquetées **ats** (cf. [corpus](#)), **a_obj** (cf. [corpus](#)), **de_obj** (cf. [corpus](#)) ou **p_obj.o** (cf. [corpus](#)). Dans le second cas, les dépendances sont étiquetées **mod** (cf. [corpus](#)). Celles-ci comprennent aussi les dépendances d'un adjectif vers un adverbe considéré comme modifieur.

167. Content de lui, il retourna à ses fourneaux



168. *un médecin disponible dans la soirée*



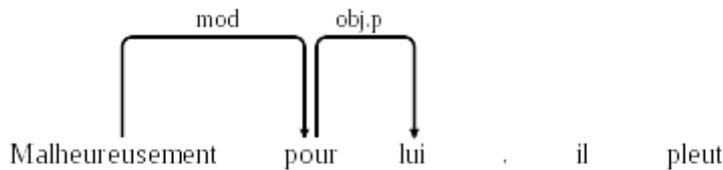
7.2 Dépendants de l'adverbe

7.2.1 Distinction argument / ajout pour les dépendants prépositionnels

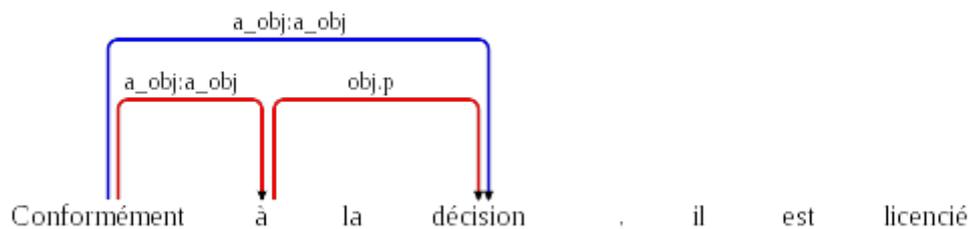
Les compléments de l'adverbe sont catégorisés de la même façon que ceux du verbe en complément requis par l'adverbe ou compléments adjoints. Dans le premier cas, les dépendances sont étiquetées **a_obj**, **de_obj** ou **p_obj.o**. Dans le second cas, les dépendances sont étiquetées **mod**.

En surface, les dépendances ont alors comme gouverneur l'adverbe lui-même et comme dépendant la préposition. En syntaxe profonde, les choses restent en l'état, sauf dans le cas des prépositions *à* ou *de*, qui sont effacées, les dépendances étant alors transférées sur la tête du syntagme suivant la préposition.

169. *Malheureusement pour lui, il pleut*

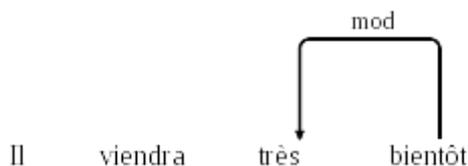


170. Conformément à la décision, il est licencié



Les adverbes peuvent être modifiés par d'autres adverbes, ce qui s'exprime par des dépendances de type **mod**.

171. Il viendra très bientôt

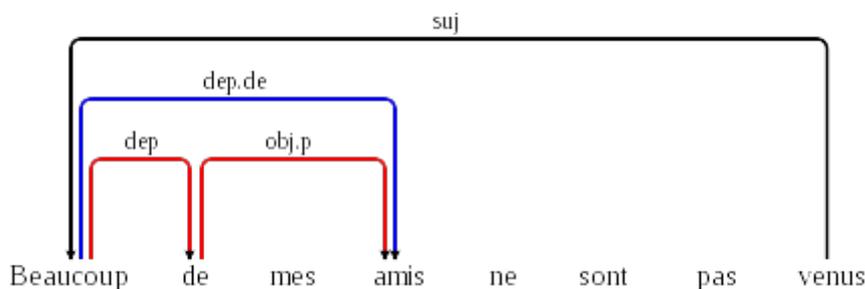


7.2.2 Adverbes de quantité

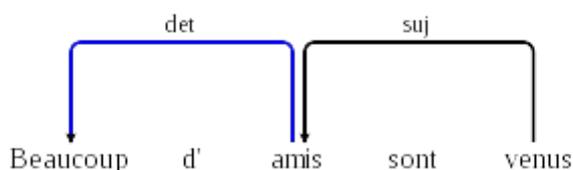
Certains quantifieurs (*beaucoup, plus, moins...*) peuvent être analysés comme des pronoms, quand ils sont employés seuls (*Beaucoup sont venus, peu sont repartis*), ou quand ils sont modifiés par un complément prépositionnel en *de*. Attention on distingue deux constructions syntaxiques différentes :

- Si le complément nominal introduit par *de* est déterminé (*beaucoup de MES amis*), le quantifieur (*beaucoup*) est analysé comme un pronom, il est la tête de l'ensemble (le pronom et le complément introduit par *de* réfèrent à deux entités différentes). Le groupe prépositionnel en *de* est traité de manière régulière (le *de* est court-circuité) ([172](#)).
- Si le nom n'est pas déterminé (*beaucoup de personnes*), on considère l'adverbe comme jouant le rôle d'un déterminant (avec le *de*). Il n'y a alors qu'une seule entité de référence et la tête du syntagme complet est le nom ([173](#)) voir également [Section 5.1](#) sur les déterminants).

172. *Beaucoup de mes amis ne sont pas venus*



173. *Beaucoup d'amis sont venus*



7.2.3 Adverbes comparatifs

Le cas *plus/moins/davantage... de N* est traité comme l'exemple (173) supra (voir également [Section 5.1](#)).

Les autres emplois des adverbes comparatifs *plus, moins, davantage, autant ...* sont traités [Section 14](#).

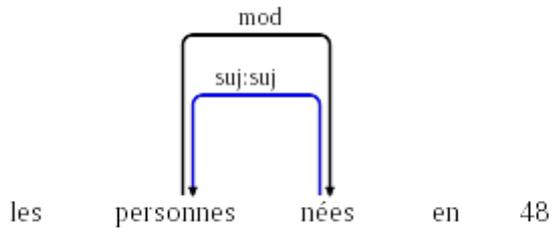
8 Dépendants participiaux et adjectivaux

8.1 Participes et adjectifs épithètes

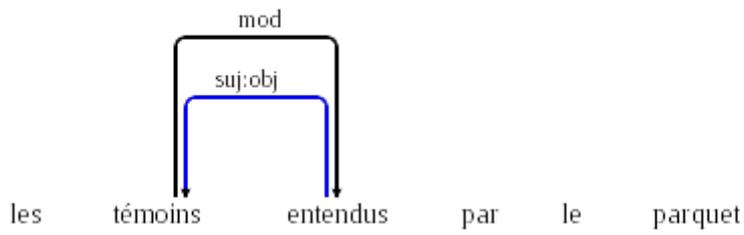
Le participe est modifieur du nom en surface, et le nom est un dépendant profond du participe : il s'agit toujours du sujet final du participe, mais selon les cas, du sujet ou de l'objet canonique ([voir corpus](#))

NB: on garde en représentation profonde la relation mod entre le nom et le participe ou l'adjectif, de manière à distinguer "les hommes bleus" et "les hommes sont bleus". Une représentation purement sémantique identifierait les 2 cas.

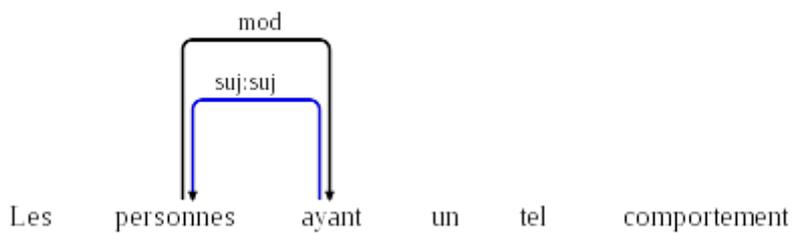
174. *les personnes nées en 48*



175. *les témoins entendus par le parquet*

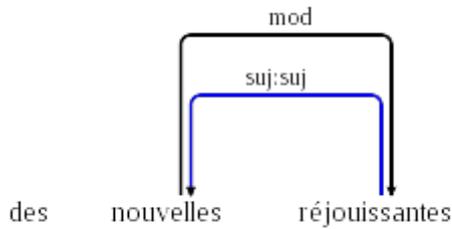


176. *Les personnes ayant un tel comportement*



On traite de la même façon l'adjectif épithète ([voir corpus](#)).

177. *des nouvelles réjouissantes*



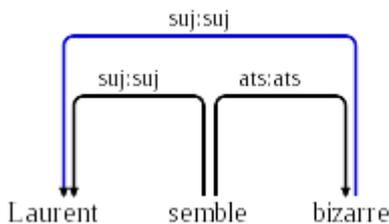
NB: on obtient un cycle en dépendances profondes.

8.2 Participes et adjectifs dépendants de verbes

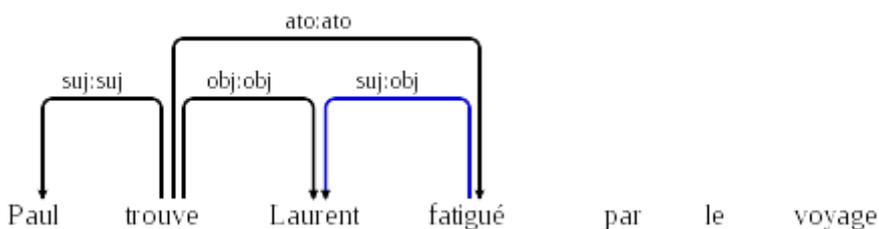
Les participes et adjectifs compléments d'un verbe prennent en général la fonction **ats**, **ato** (cf. [Section 4.5](#)) ou **mod** (cf. [Section 4.9](#)). Exceptionnellement, ils peuvent avoir la fonction **obj** (par exemple dans l'expression *plaider coupable*).

On conserve en représentation profonde le lien avec le verbe (ce faisant on reste à un niveau syntaxique), mais l'élément "sujet" de l'adjectif ou du participe est explicité en dépendances profondes.

178. *Laurent semble bizarre*



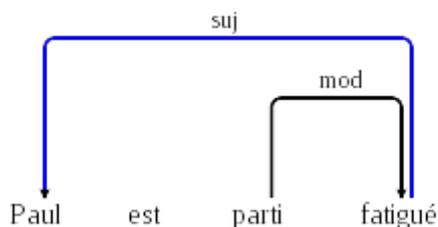
179. *Paul trouve Laurent fatigué par le voyage*



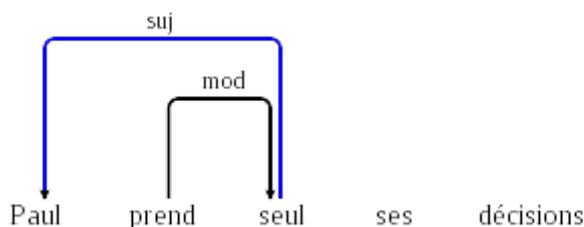
On distingue le cas ci-dessus où l'attribut de l'objet ou du sujet est un complément sous-catégorisé

par le verbe, du cas dit "de modifieur attributif", où l'adjectif bien que formant un syntagme distinct du NP auquel il se rapporte sémantiquement (sujet ou objet du verbe), n'est pas sous-catégorisé par le verbe. On a en surface une relation **mod** entre l'adjectif et le verbe (conservée en profond), et le sujet de l'adjectif est ajouté en dépendances profondes.

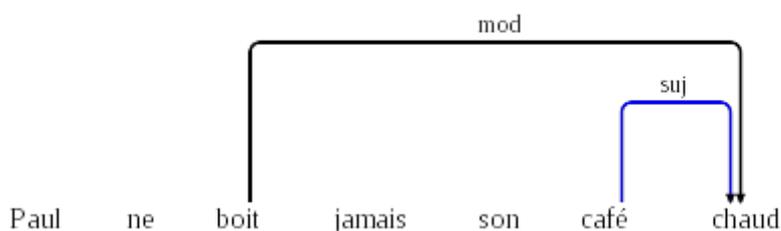
180. *Paul est parti fatigué*



181. *Paul prend seul ses décisions*



182. *Paul ne boit jamais son café chaud.*



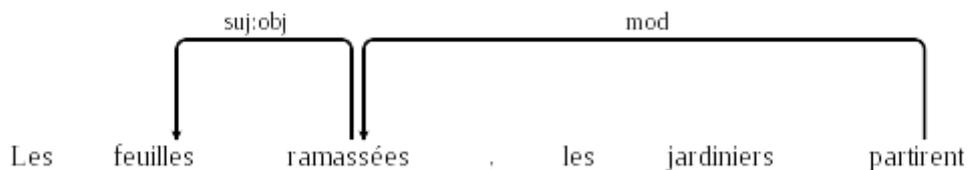
8.3 Participiales absolues

On parle ici de participiales passées ou présentes, dont le sujet est réalisé localement. La participiale est modifieur du verbe principal et le sujet final du participe est sujet ou objet canonique selon la diathèse.

183. *La nuit venant, les chats sortirent*

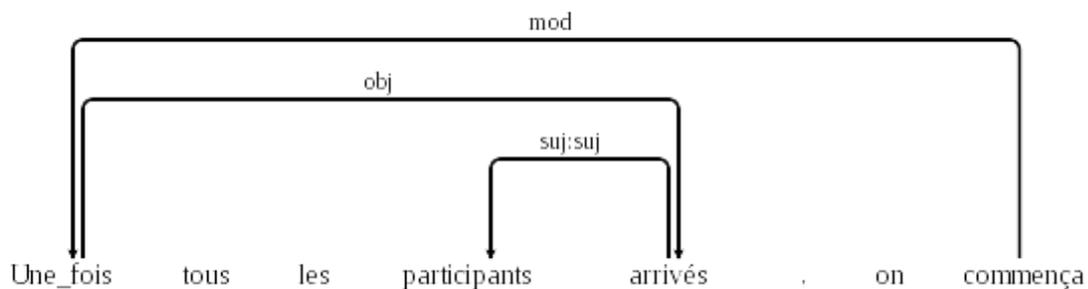


184. *Les feuilles ramassées, les jardiniers partirent*



Pour les participiales introduites par le composé *une fois*, on peut hésiter sur la nature de la tête : *une fois* ou bien le participe. Certes *une fois* est alors optionnel (*Une fois / Ø*) les participants arrivés, on commença), mais il est forcément en tête de la participiale (à ne pas confondre avec le cas du modifieur de la forme "Xcard fois" : *Tous les participants recalés plusieurs fois, le jury décida d'assouplir les critères*). On privilégie le parallélisme avec le cas où l'adverbe *une fois* introduit une complétive : *une fois que les participants sont arrivés, ...*).

185. *Une fois tous les participants arrivés, on commença*



8.4 Adjectivales et participiales "détachées"

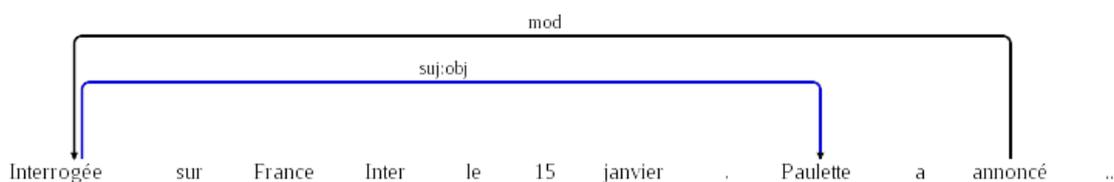
Il s'agit ici à l'inverse de cas où le participe n'a pas de sujet final réalisé localement. Un dépendant du verbe principal, en général le sujet, correspond au sujet final du participe (et à son sujet ou objet canonique, selon la diathèse du participe).

Dans tous les cas, dans la représentation de surface, la participiale est modifieur du verbe principal.

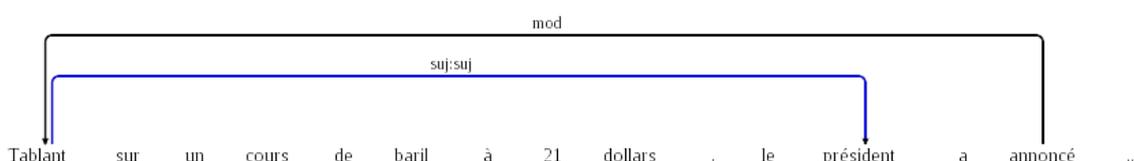
En dépendances profondes,

- d'une part, parallèlement au cas de participiale absolue (i.e. avec sujet propre réalisé), la relation **mod** entre le participe et le V principal reste également en dépendances profondes, elle note une relation discursive.
- d'autre part, on explicite quel est le sujet final du participe (et sa fonction canonique). C'est en général le sujet de la principale ([voir corpus](#)).

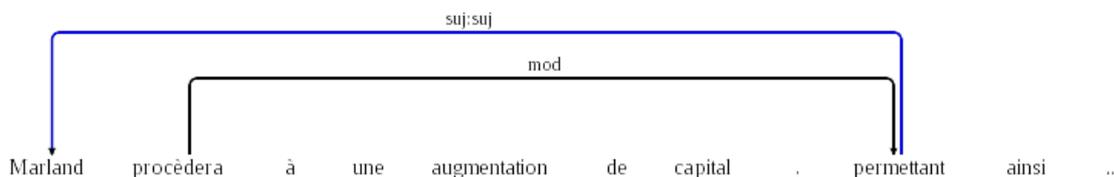
186. *Interrogée sur France Inter le 15 janvier, Paulette a annoncé ...*



187. *Tablant sur un cours de baril à 21 dollars, le président a annoncé ...*

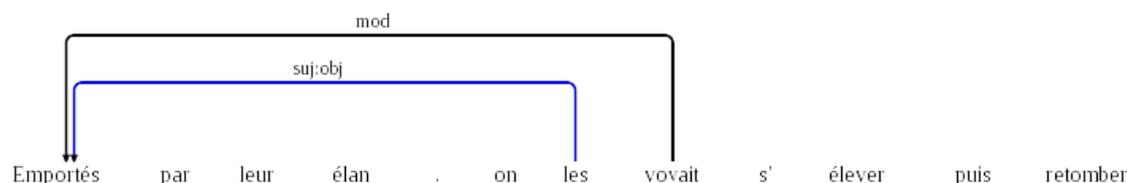


188. *Marland procèdera à une augmentation de capital, permettant ainsi ...*

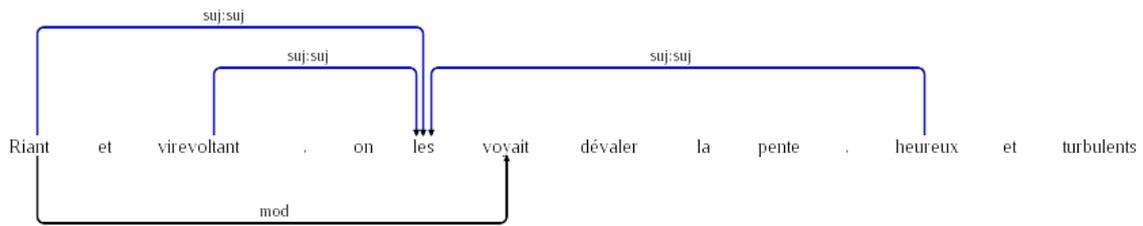


Mais, attention : le sujet du participe n'est pas obligatoirement (même si la plupart du temps) le sujet du verbe enchâssant.

189. *Emportés par leur élan, on les voyait s'élever puis retomber*

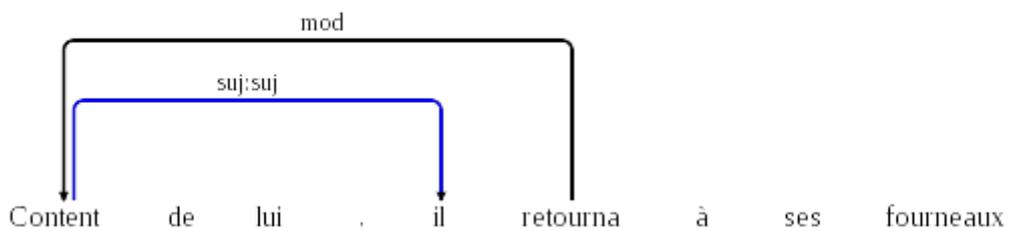


190. *Riant et virevoltant, on les voyait dévaler la pente, heureux et turbulents (voir la [Section 13](#) sur la coordination))*

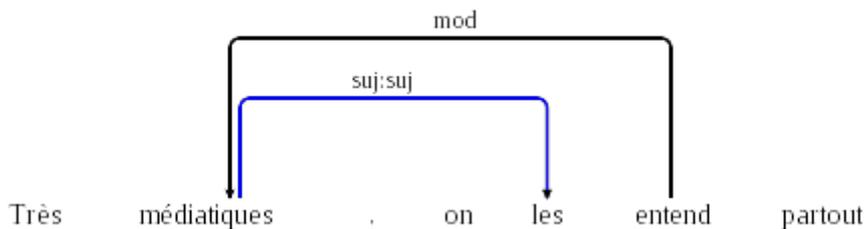


Les adjectivales détachées se comportent comme les participiales détachées.

191. *Content de lui, il retourna à ses fourneaux*



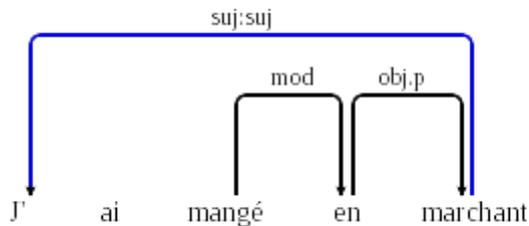
192. *Très médiatiques, on les entend partout*



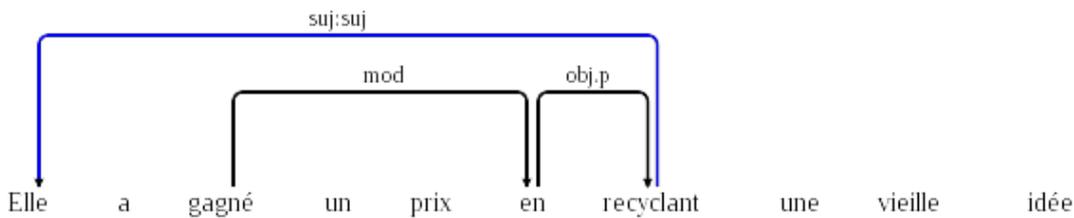
8.5 Gérondif

Pour le gérondif *en*+participe présent on note en surface le participe dépendant de la prep *en* (on ne considère pas la prep comme sémantiquement vide mais plutôt qu'elle porte, avec la marque du participe présent, la sémantique (simultanéité, causalité...)). Le sujet du participe est la plupart du temps le sujet du verbe enchassant, auquel cas on l'explicite en dépendances profondes :

193. *J'ai mangé en marchant*



194. *Elle a gagné un prix en recyclant une vieille idée*



Mais attention, ce n'est pas toujours le cas, en particulier si la principale a un sujet asémantique : exemple FTB:

195. *En attendant que les choses changent, il nous faut ... => le sujet de attendant n'est pas l'explétif il.*

On a ce phénomène en particulier avec des gérondifs presque figés : "*en tenant compte de ...*", "*en attendant*"

9 Dépendants phrastiques et infinitivaux

9.1 Complémenteurs / prépositions sémantiquement pleins versus sémantiquement vides

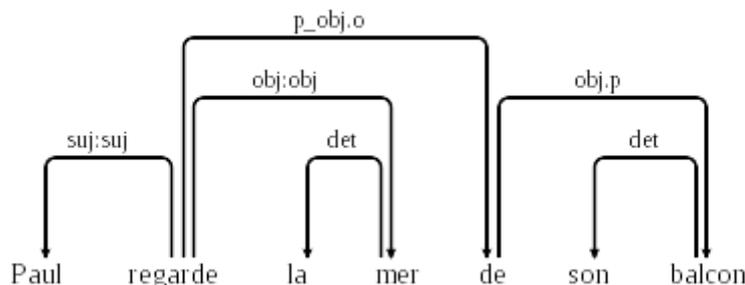
Contrairement au schéma d'annotation du French Treebank, le schéma en dépendances *de surface* uniformise le traitement des prépositions, que celles-ci introduisent un groupe nominal, une infinitive ou une phrase. Ainsi, en surface une préposition domine-t-elle toujours la tête du syntagme qu'elle introduit. De la même manière, un complémenteur est toujours en surface le gouverneur du verbe de la phrase ou de l'infinitive qu'il introduit.

En représentation profonde, on introduit une modification liée au statut sémantique des prépositions et complémenteurs : les complémenteurs et prépositions sémantiquement vides sont "court-circuités" (peu importe que leur complément soit nominal, phrastique ou infinitival). Cela concerne le complémenteur *que*, systématiquement considéré comme vide, et certaines occurrences de *à* et *de*.

Les compléments et prépositions sémantiquement pleins (comme par exemple *lorsque* ou *sous*) sont inchangés en représentation profonde.

Comparez par exemple le traitement de *de* dans l'exemple (197) où il est préposition régie, versus le traitement de *de* dans l'exemple suivant, où il est sémantiquement plein (196) :

196. *Paul regarde la mer de son balcon*



9.2 Dépendants infinitivaux

9.2.1 Explicitation du sujet final des infinitifs

Un des points centraux du schéma d'annotation en dépendances profondes est l'explicitation du sujet final des verbes non conjugués (infinitifs et participes). C'est bien une particularité des infinitifs et des participes qu'un de leurs arguments ne soient pas localement exprimé. Cet argument est systématiquement le sujet **final** de l'infinitif ou du participe, mais peut correspondre à une fonction canonique autre que le sujet, pour peu que l'infinitif ou le participe apparaisse dans une diathèse non canonique. En effet, les changements de diathèse s'appliquent de manière régulière que le verbe soit conjugué ou pas.

On trouve donc que l'argument non exprimé localement d'un infinitif peut correspondre, selon la diathèse de l'infinitif, à son sujet canonique, son objet canonique etc... Par exemple, l'argument manquant de l'infinitif correspond à un sujet à la fois final et canonique dans le cas d'un infinitif actif (197) mais il correspond à un sujet final et objet canonique dans le cas d'un infinitif passif (199).

La régularité sous-jacente est bien que cet argument non exprimé localement est le sujet final de l'infinitif. C'est pour capturer cette régularité, et la régularité des changements de diathèse que le verbe soit conjugué ou pas, que l'on a introduit les notions de fonctions finales versus canoniques, de manière orthogonale à la notion de dépendance de surface versus dépendance profonde.

Donc par exemple dans le cas d'un infinitif passif (199), la dépendance profonde entre l'infinitif et son argument non exprimé localement porte un label **suj** final et **obj** canonique.

Consignes pour l'explicitation du sujet final des infinitifs

On explicite le sujet final de l'infinitif, ainsi que sa fonction canonique, dans le cas où il est exprimé au sein de la phrase, sans ambiguïté d'après le contexte (éventuellement le contexte des phrases

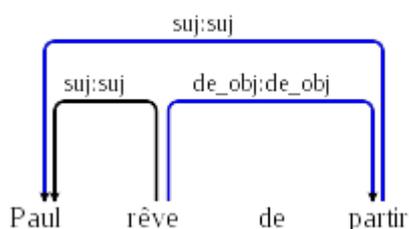
précédentes) : cela couvre les cas de contrôle obligatoire (comme en [197](#), [198](#), [199](#)) et de contrôle arbitraire (comme en [203](#), [204](#)).

Infinitives gouvernées par un verbe (via complémentateur à ou *de*, ou préposition régie à ou *de*)

Contrôle obligatoire : il s'agit des cas où le verbe (verbes à contrôle, verbes de perception, verbes à montée) impose lequel de ses compléments est le "contrôleur", i.e. le sujet final de l'infinitif.

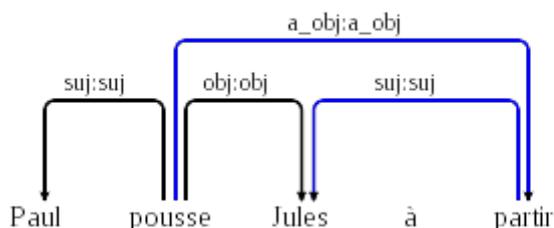
Par exemple, *rêver* est un verbe à contrôleur sujet : son sujet est le sujet final de l'infinitive.

197. *Paul rêve de partir*



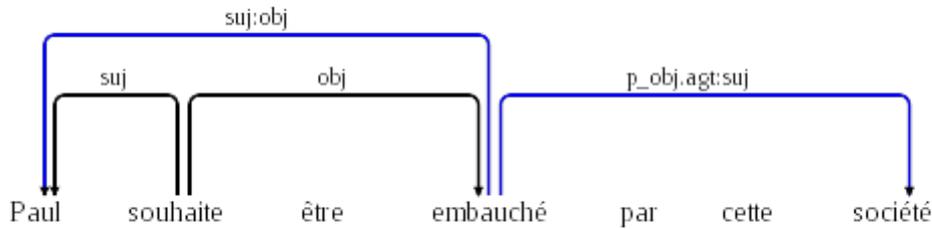
Alors que *pousser* est un verbe à contrôleur objet : son objet (ci-dessous *Jules*) est le sujet final de l'infinitive.

198. *Paul pousse Jules à partir*



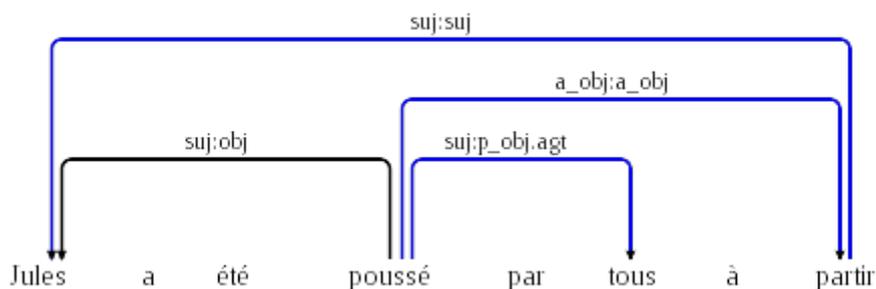
On a vu plus haut que l'infinitif peut être dans une diathèse non canonique, mais c'est toujours le sujet final de l'infinitif qui est non local en surface, et donc explicité en syntaxe profonde : ci-dessous *Paul* est sujet final et objet canonique de *embauché*.

199. *Paul souhaite être embauché par cette société*



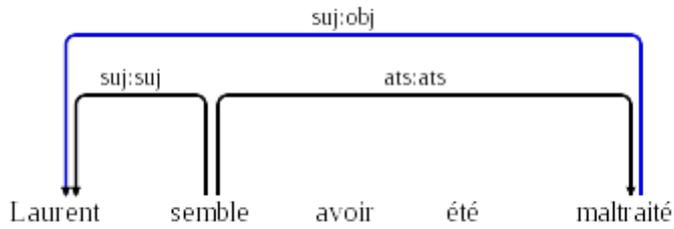
Mais il est important de noter que la propriété lexicale définissant quel complément du verbe à contrôle est le contrôleur s'exprime en revanche **en termes de fonction canonique et pas de fonction finale**: ainsi par exemple *pousser* est bien un verbe à contrôleur objet *mais* plus précisément "à contrôleur objet *canonique*". A l'actif, c'est l'objet final et canonique de *pousser* qui est le sujet final de l'infinitive (cf. 198), mais lorsque *pousser* est au passif, c'est son objet canonique mais sujet final qui est le sujet final de l'infinitive :

200. *Jules a été poussé par tous à partir*

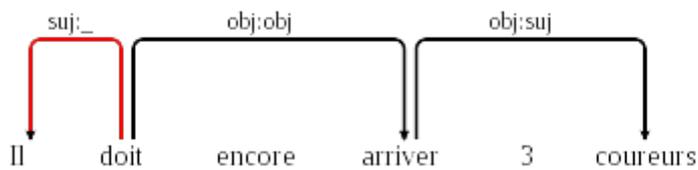


La distinction dans la REPRSYNTPROF entre verbe à contrôle, verbe à montée, verbe de perception tient à la fonction portée par l'infinitive. Mais à noter que le fait que le verbe à montée ne sélectionne pas son sujet (i.e. le sujet (ou objet) du verbe à montée n'est pas un argument sémantique du verbe à montée, alors qu'il l'est pour un verbe à contrôle) n'est pas capturé dans la REPRSYNTPROF, d'où la représentation en (201). Cela dit, on peut du coup avoir un verbe à montée sur une diathèse impersonnelle, alors le sujet explétif *il* n'est pas conservé en dépendances profondes (comme tous les *Il* explétifs) (202).

201. *Laurent semble avoir été maltraité*

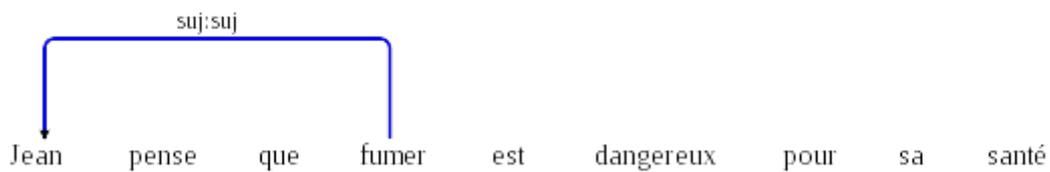


202. *Il doit encore arriver 3 coureurs*



Contrôle arbitraire (voir Bashung, 96 pour le français) : il s'agit de cas où le sujet de l'infinitif, s'il est exprimé, ne peut pas être déterminé structurellement ou fonctionnellement de manière systématique. C'est le cas en particulier pour les infinitives sujet (exemples de Bashung 96) :

203. *Jean₁ pense que ₁fumer est dangereux pour sa santé*



204. *Jean₁ a disputé Marie₂, il pense que ₂fumer est dangereux pour sa santé*



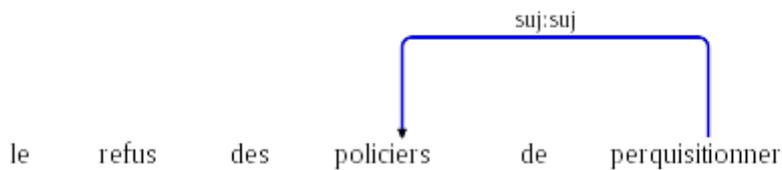
205. *Jean₁ pense que _xfumer est dangereux pour la santé.* => sujet non exprimé

Jean pense que fumer est dangereux pour la santé.

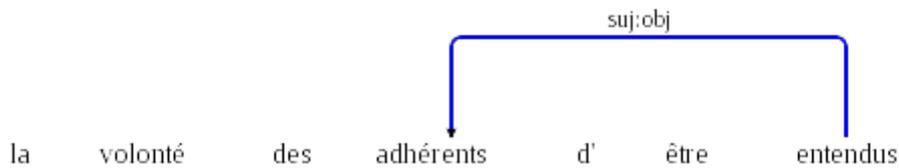
Infinitives gouvernées par un nom (via préposition régie à ou de)

Il s'agit de certains noms prédicatifs, où un des compléments du nom est syntaxiquement forcément le sujet (final) du VINF :

206. *le refus des policiers_i de perquisitionner_i*



207. *la volonté des adhérents_i d'être entendus*

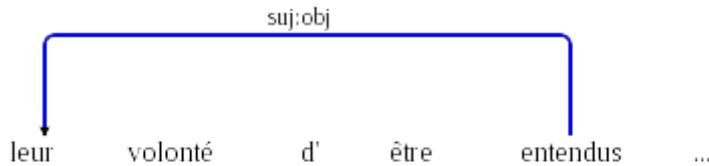


208. *L'obligation de/pour Paul_i de partir*



Dans le cas où le contrôleur est un complément du nom, introduit par *de*, on peut avoir une réalisation du complément via un déterminant possessif [209](#):

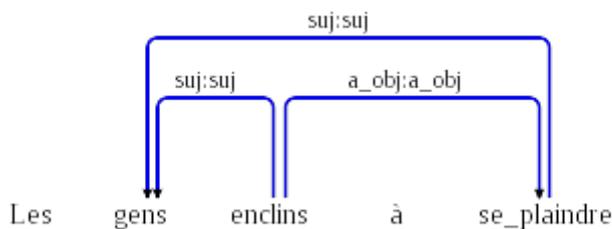
209. *leur_i volonté d'être entendus...*



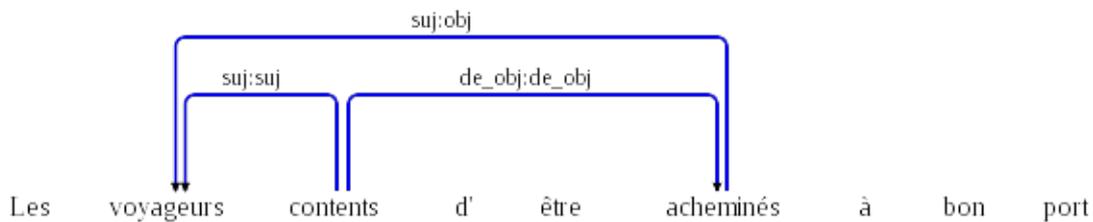
Infinitives gouvernées par un adjectif (via préposition régie à ou de)

Entrent dans le champ du contrôle obligatoire : certains adjectifs construits avec une infinitive ont la propriété lexicale que le sujet (final) de l'infinitif est forcément le sujet canonique de l'adjectif :

210. *Les gens_i enclins à se_i plaindre* => *gens* est le sujet de *se plaindre*, et le *à* est court-circuité :



211. *Les voyageurs_i contents d'_i être acheminés à bon port* => *voyageurs* est le sujet final et objet canonique de *acheminés*, et *d'* est court-circuité



Infinitives gouvernées par préposition non régie

A priori on n'a jamais de phénomène de contrôle syntaxique obligatoire au sens strict : c'est l'interprétation qui permet d'inférer le sujet de l'infinitif dépendant d'une préposition (forte).

Pour la préposition *pour*, la position du sujet de l'infinitif varie selon les emplois, cf. les exemples suivants (Cadiot, 1990):

Paul_i était sorti pour _jsouper.

Je vous_i paie pour vous _joccuper de mes enfants.

Elle a donné son enfant à une vieille dame_i pour le _jramener au pays.

J'indiquais la voiture à la femme_i de chambre pour s'y _jdébarrasser de ses paquets.

Mais pour certaines prépositions, comme "sans", "avant (de)", "après", "au lieu de", on a une tendance massive à avoir le sujet (final) du V modifié par la prep comme sujet (final). On a d'ailleurs une difficulté à forger des exemples où le sujet de l'infinitive ne serait pas le sujet du V que modifie la prep.

La lune_i a enchanté les badauds_j, avant de _{i/j}rentrer

**La lune a enchanté les badauds, sans la voir.*

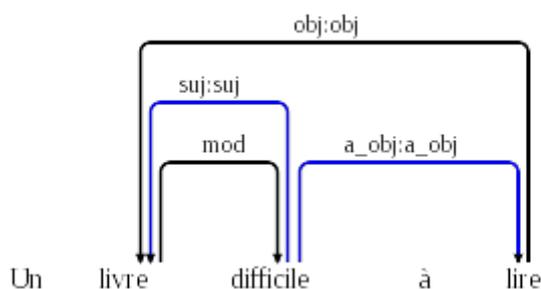
Mais on peut citer un beau contre-exemple tiré du FTB avec *après* :

Un petit nombre de produits verront leur prix_i plafonnés, après _javoir été multipliés par 2

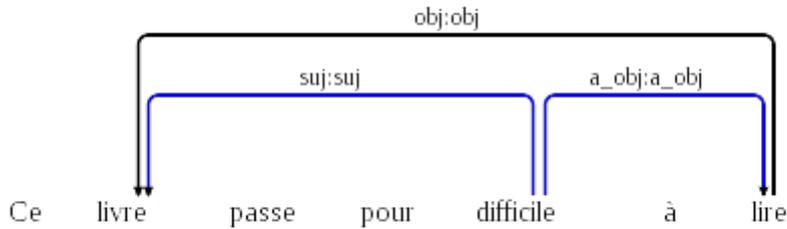
9.2.2 Explicitation de l'objet des infinitifs (tough movement)

On a également les constructions infinitives à objet manquant ("tough movement" (Kayne, 75), Huot, 81 pour le français), où c'est toujours l'argument "sujet" de l'adjectif qui est objet de l'infinitif, et ce quel que soit le type de réalisation de surface du sujet de l'adjectif (cf. les exemples de Bonfante et al., 2011):

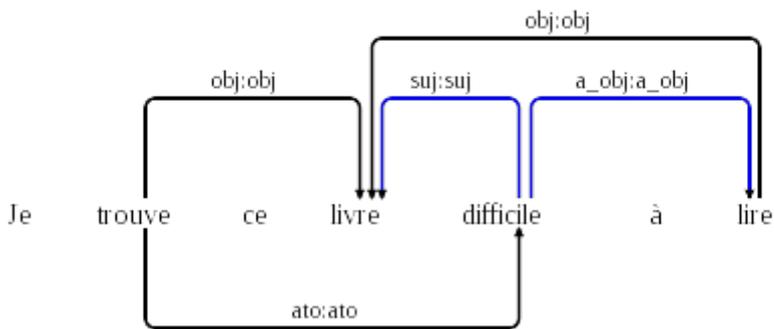
212. *Un livre difficile à lire*



213. *Ce livre passe pour difficile à lire*



214. *Je trouve ce livre difficile à lire*



215. *Difficile à lire, ce livre n'est pas à conseiller à tout le monde*

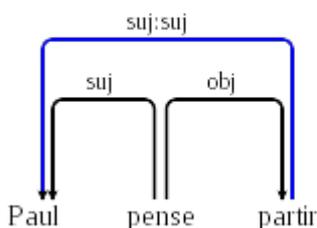
Il est cependant important de noter que la représentation en syntaxe profonde obtenue présente une divergence structurale avec la sémantique : en effet dans le cas de construction à tough adjective, on obtient deux arguments syntaxiques pour l'adjectif (son sujet, et un a_objet), alors que sémantiquement l'adjectif a un seul argument sémantique, dont la tête est l'infinitif, ce qui peut se représenter grossièrement par : **difficile(lire(livre))**

On a là comme pour les constructions à attribut de l'objet une divergence entre syntaxe (profonde) et sémantique.

9.2.3 Fonctions portées par les infinitives

Objets directs sans complémenteur

216. *Paul pense partir*



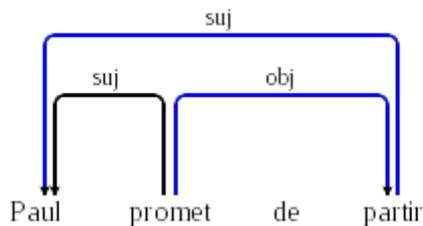
Objets directs avec complémenteur

Repérés car alternant avec un complément nominal direct (*Paul promet de partir / Paul promet (*de/ø) cela*).

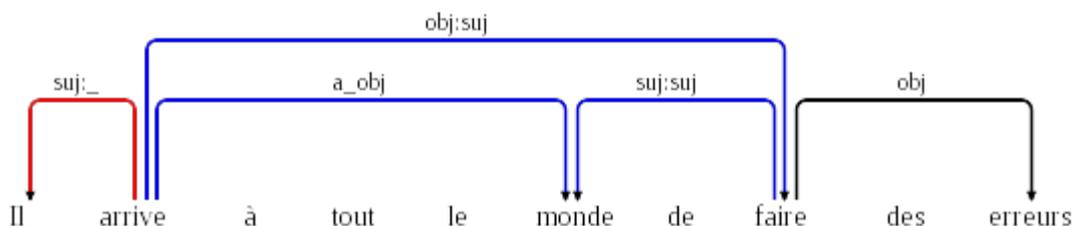
Rappel : Le FTB étiquette dans ces cas le complémenteur comme une préposition. La distinction entre le cas alternant avec *cela* et le cas alternant avec *de cela* ou *à cela* est faite au niveau de la fonction : **obj** (217) au lieu de **de_obj** (219) ou **a_obj** (220 et 221).

Pour la représentation profonde, on explicite le sujet de l'infinitif, et on court-circuite les complémenteurs *à* et le *de*.

217. *Paul promet de partir*

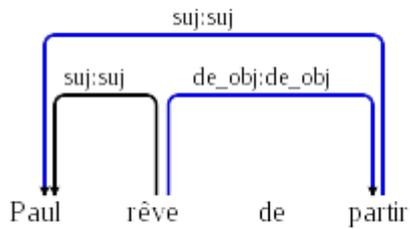


218. *Il arrive à tout le monde de faire des erreurs* => on a ici une diathèse impersonnelle (cf. section [3.4 Changement de diathèse impersonnel versus tournure impersonnelle](#)). L'infinitive est objet final mais sujet canonique (et le "de" l'introduisant est court-circuité). Le sujet de l'infinitif est ici le **a_obj** du verbe *arriver*.

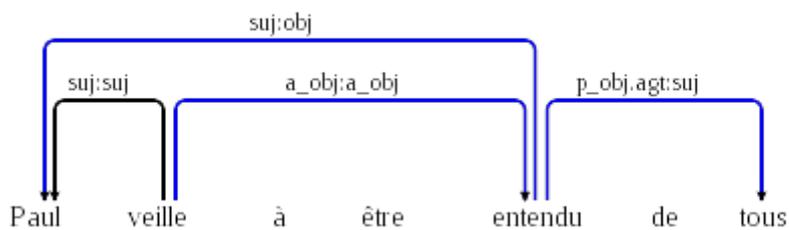


Infinitives complément indirects

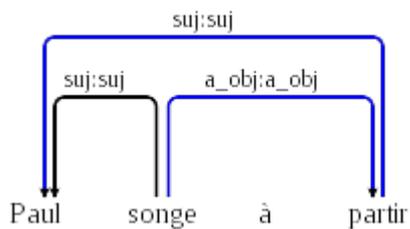
219. *Paul rêve de partir*



220. *Paul veille à être entendu de tous*



221. *Paul songe à partir*



Modifieurs prépositionnels infinitifs

NB: dans le cas où il n'y a pas contrôle obligatoire, **on choisit d'annoter le sujet de l'infinitif lorsqu'il est effectivement exprimé ailleurs dans la phrase, et s'il n'y a aucune ambiguïté quant à l'interprétation.**

222. *Paul dort pour rêver*



9.3 Dépendants phrastiques

Pour les phrases introduites par un complémenteur, le complémenteur est tête de la subordonnée en représentation de surface.

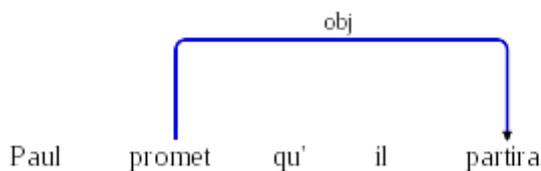
En représentation profonde en outre, les complémenteurs *que* et *si* sont systématiquement "court-circuités" : c'est le V de la subordonnée qui est tête. La différence sémantique introduite par *si* est gérée par un trait sur l'interrogative (cltype=int).

9.3.1 Phrases arguments ou ajouts de verbes

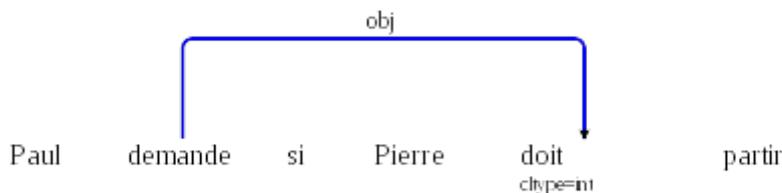
Phrase objet direct : complétives et interrogatives indirectes

Les objets directs phrastiques sont soit des complétives en *que*, soit des interrogatives indirectes.

223. *Paul promet qu'il partira*

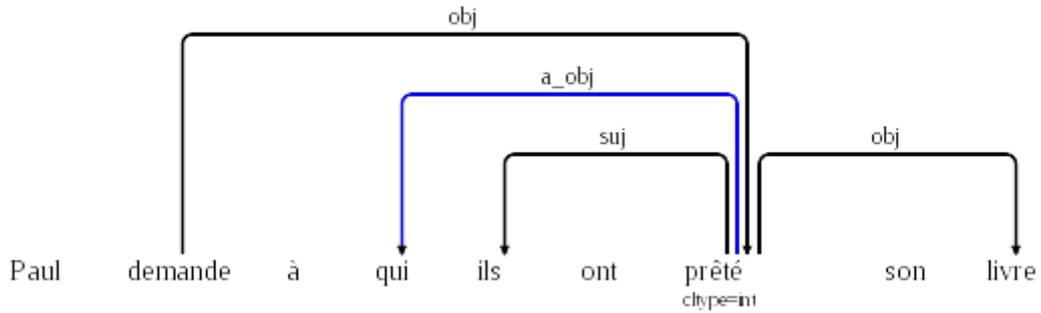


224. *Paul demande si Pierre doit partir*

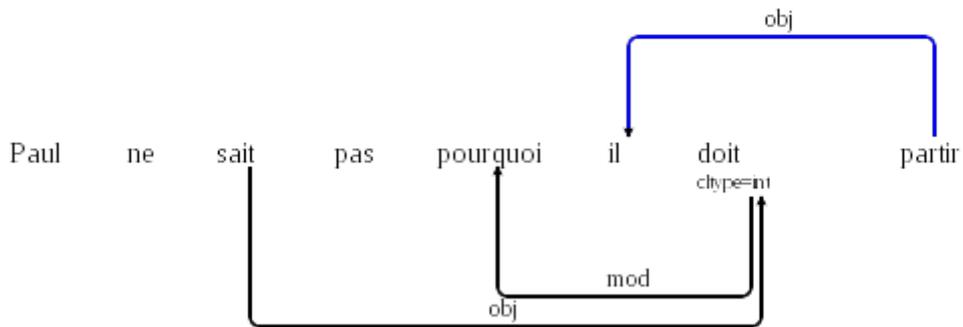


De la même manière, les interrogatives indirectes avec syntagme antéposé ont pour tête le verbe de l'interrogative, et celui-ci porte un trait cltype=int :

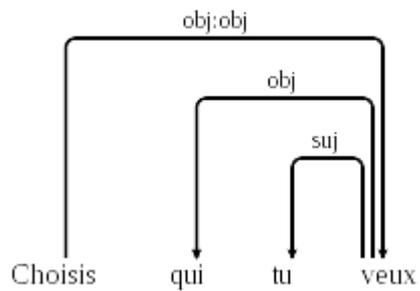
225. Paul demande à qui ils ont prêté son livre



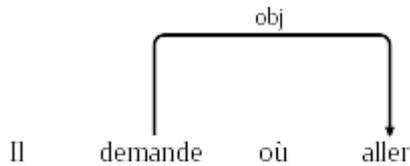
226. Paul ne sait pas pourquoi il doit partir



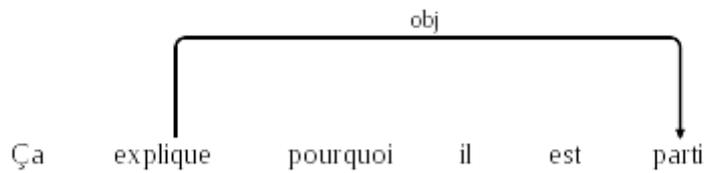
227. Choisis qui tu veux



228. Il demande où aller

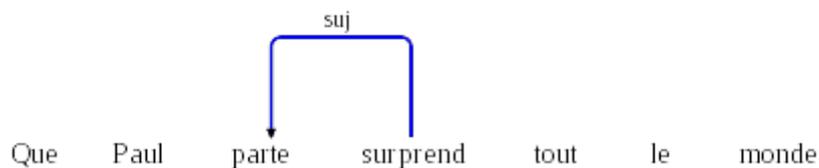


229. *Ça explique pourquoi il est parti*



Phrase sujet

230. *Que Paul parte surprend tout le monde.*

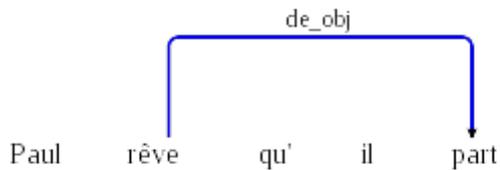


Compléments indirects phrastiques

Apparaissant soit avec un tour *de ce que*, soit sans préposition, mais repérables car alternant avec SP nominaux (cf. par exemple *Paul rêve qu'il part* / *Paul rêve de la lune*)

Attention, même dans le cas sans préposition, l'alternance avec *de cela* implique la fonction **de_obj** et pas **obj**

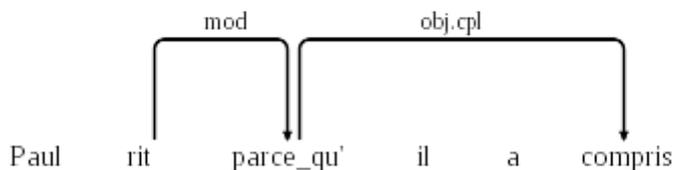
231. *Paul rêve qu'il part* => la complétive est indirecte (fonction **de_obj**), et le *qu'* est court-circuité



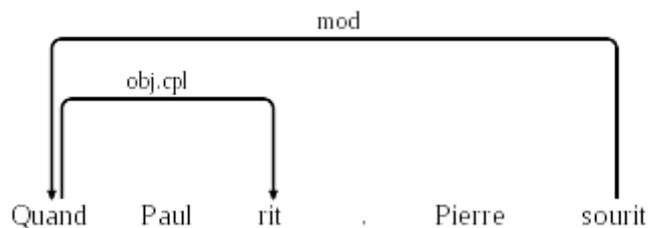
Ajouts phrastiques

Dans le cas d'ajouts phrastiques, c'est le complémenteur qui est la tête. Le verbe qu'il introduit en dépend avec l'étiquette **obj.cpl**.

232. *Paul rit parce qu'il a compris*



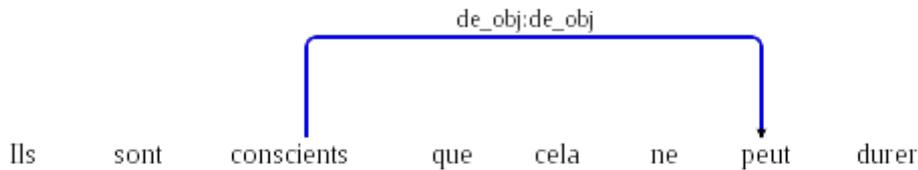
233. *Quand Paul rit, Pierre sourit*



9.3.2 Dépendants phrastiques d'adjectifs

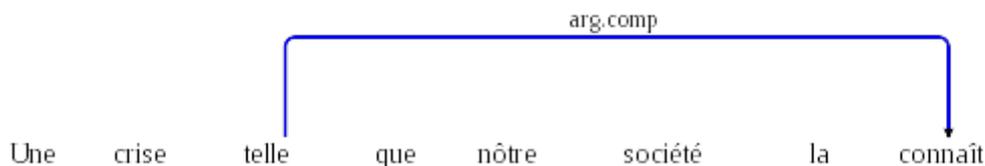
De la même manière que pour les verbes, une complétive en *que* est notée **de_obj** si elle alterne avec un de+nom : *Ils sont conscients (que cela a assez duré / de cela)*

234. *Ils sont conscients que cela ne peut durer*



Pour l'adjectif *tel*, dans le cas *tel que P*, le *que P* alterne avec un SN et pas un *de SN* (*il est tel qu'on le connaît / il est tel le phénix qui renaît de ses cendres*). On privilégie une annotation capturant la sémantique comparative : le *que P* est noté **arg.comp** (et le *que* est court-circuité) (voir également [Section 14](#)).

235. *Une crise telle que nôtre société la connaît*

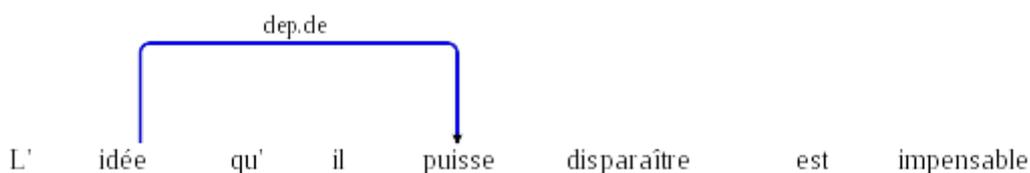


Et en représentation profonde, on court-circuite le complémenteur *que*.

9.3.3 Dépendants phrastiques d'autres catégories

On peut avoir des phrases complément de noms. Dans le cas où le *que P* alterne avec un *de SN*, on utilise la dépendance **dep.de** (de la même manière que pour les *de SN* dépendant d'un nom, cf. [Section 5.2](#)) :

236. *L'idée qu'il puisse disparaître est impensable*



On peut également avoir des phrases compléments d'adverbes, en particulier comparatifs (voir [Section 14](#)).

10 Dépendants pronominaux

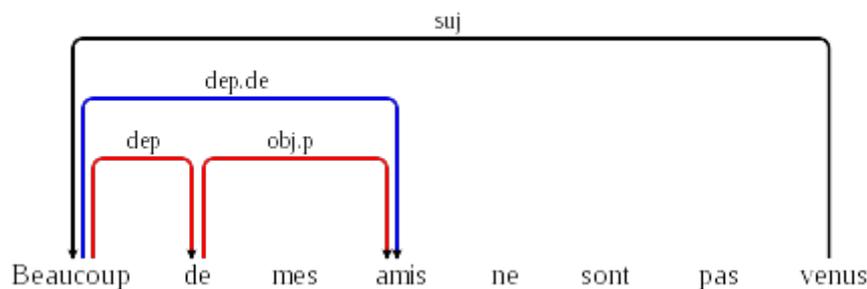
Voir [Section 4.2](#) sur les clitiques.

Voir [Section 5.4](#) pour les pronoms relatifs.

Rien de spécial pour les pronoms interrogatifs (Voir [Section 9.2.1](#) sur les interrogatives indirectes)

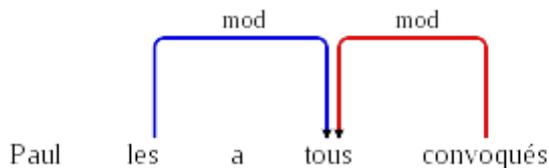
Les pronoms "de quantité" peuvent avoir un complément en *de* (cf. [section 7.2.2](#), à distinguer du cas *ADV de N, beaucoup d'amis*).

237. *Beaucoup de mes amis ne sont pas venus*



Les quantifieurs flottants sont "résolus" en dépendances profondes, c'est-à-dire qu'on les rattache au nom qu'ils quantifient sémantiquement, alors qu'en surface ils sont rattachés au verbe.

238. *Paul les a tous convoqués*



11 Dépendants adverbiaux

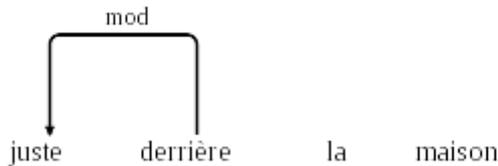
11.1 Adverbe comme modifieur

L'adverbe est en général utilisé comme modifieur d'un verbe, un adjectif, un autre adverbe, un nom ou un pronom, une préposition ou encore une conjonction ([voir corpus](#)).

239. *Paul dort beaucoup*



240. *juste derrière la maison*



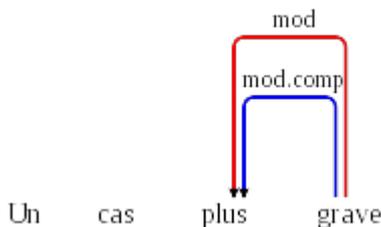
11.1.1 Comparatif et superlatif

On distingue la représentation du superlatif (*le cas le plus grave / le plus grave cas*) de la représentation du comparatif (*le cas plus grave*) en utilisant en syntaxe profonde deux étiquettes distinctes **mod.super** et **mod.comp** entre l'adverbe superlatif ou comparatif (*plus, moins, davantage...*) et l'élément qu'il modifie. En syntaxe de surface, dans les deux cas, ils ont la fonction **mod**.

Comparatif

En surface, l'adverbe exprimant une comparaison a la fonction **mod** et en profondeur sa fonction est spécifiée **mod.comp** ([voir corpus](#)).

241. *Un cas plus grave*



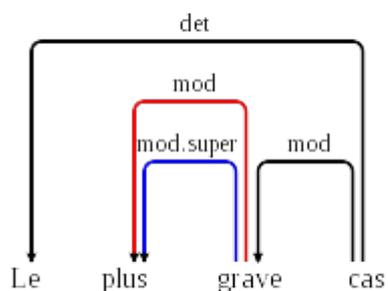
Voir par ailleurs la [Section 14](#) pour le traitement des comparatives (lorsque qu'un complément en *que* exprime l'autre terme de la comparaison : *un cas plus grave que celui-là*).

Superlatif

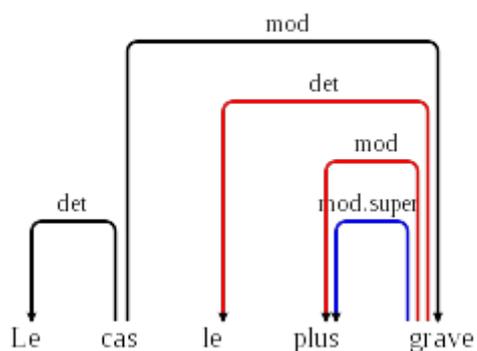
En surface, l'adverbe exprimant un superlatif a la fonction **mod** et en profondeur sa fonction est spécifiée **mod.super** ([voir corpus](#)). Le déterminant placé devant est conservé en profondeur comme

déterminant, seulement s'il remplit cette fonction (en tête d'un syntagme nominal), sinon il est effacé car n'apportant aucune information par rapport à l'étiquette **mod.super**.

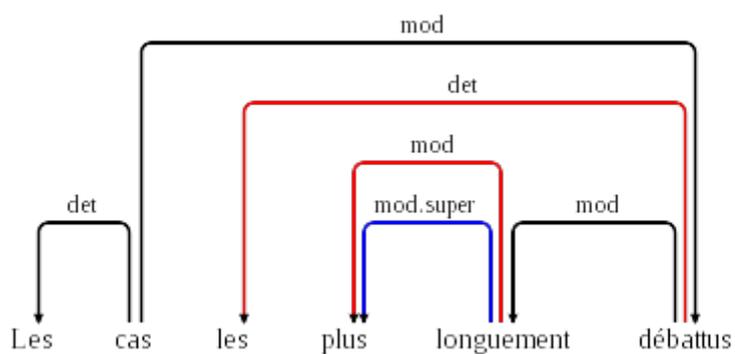
242. Le plus grave cas



243. *Le cas le plus grave*



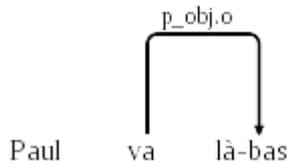
244. *Les cas les plus longuement débattus*



11.2 Adverbe sous-catégorisé

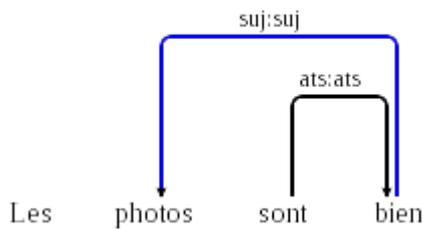
Un adverbe peut être sous-catégorisé par un verbe ([voir corpus](#)). Il reçoit alors la fonction **p_obj.o** ou **ats / ato**.

245. *Paul va là-bas*

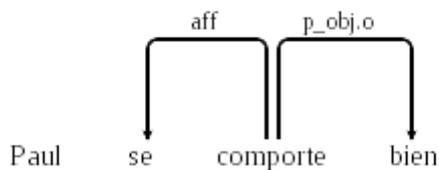


Lorsqu'un adverbe est attribut, on lui attribue un sujet en dépendances profondes (cf. le traitement des attributs [Section 4.5](#)).

246. *Les photos sont bien*

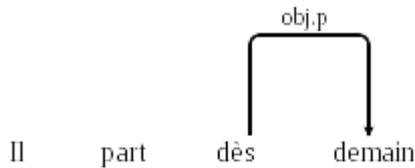


247. *Paul se comporte bien*



Les temporels *demain, aujourd'hui etc...* ont une distribution de nom, mais sont codés dans le corpus de manière traditionnelle comme des adverbes. Ils peuvent occuper les positions de noms, et donc en particulier être l'objet d'une préposition, avec alors la fonction **obj.p**.

248. *Il part dès demain*



12 Clivées, dislocation et incises

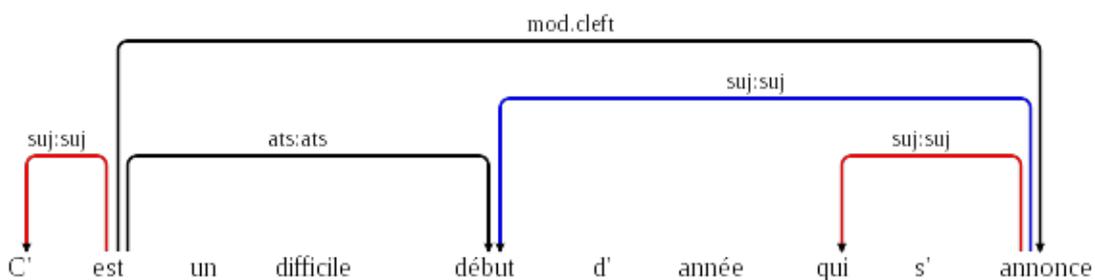
12.1 Clivées

En dépendances de surface, le schéma suit celui du French Treebank où la subordonnée en *que/qui* est codée comme une **relative**, non intégrée au foyer de clivée. Ainsi la subordonnée a comme tête le verbe. Le foyer de clivée est attribut du sujet. La subordonnée (en *que/qui*) est notée comme un dépendant **mod.cleft** du verbe être.

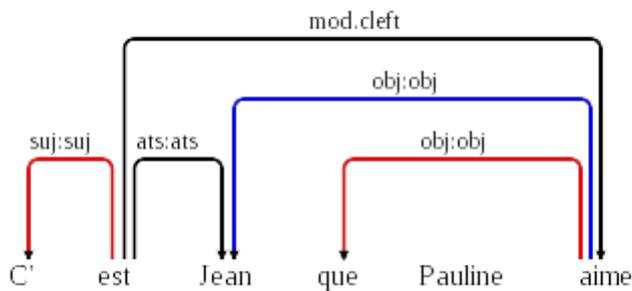
En dépendances profondes, on explicite la fonction que porte le foyer de la clivée par rapport au verbe de la subordonnée (et le pronom relatif est court-circuité). On peut distinguer deux cas :

12.1.1 Clivée avec foyer sujet ou objet du V de la subordonnée

249. *C'est un difficile début d'année qui s'annonce*



250. *C'est Jean que Pauline aime*

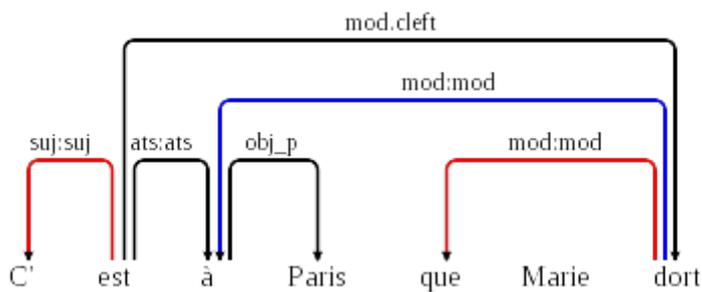


12.1.2 Clivée avec autre type de foyer

Quand le foyer de clivée est prépositionnel ou adverbial, la subordonnée est introduite par *que*, mais NB il est toujours noté comme un pronom relatif.

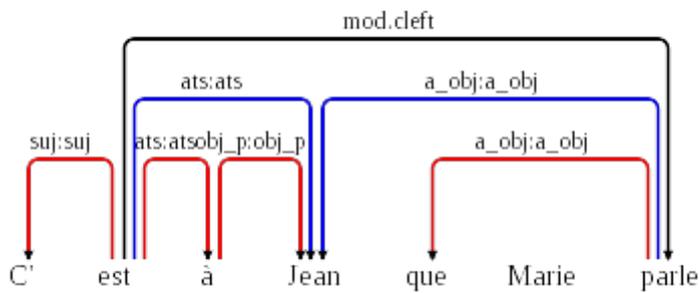
Le foyer de clivée peut correspondre à un modifieur du V de la subordonnée :

251. *C'est à Paris que Marie dort*



ou bien à un complément prépositionnel régi. Dans ce cas, en dépendances profondes, outre le *que* introduisant la clivée, on court-circuite la préposition régie :

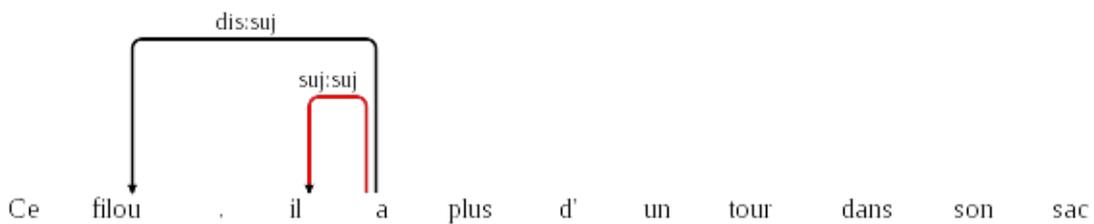
252. *C'est à Jean que Marie parle*



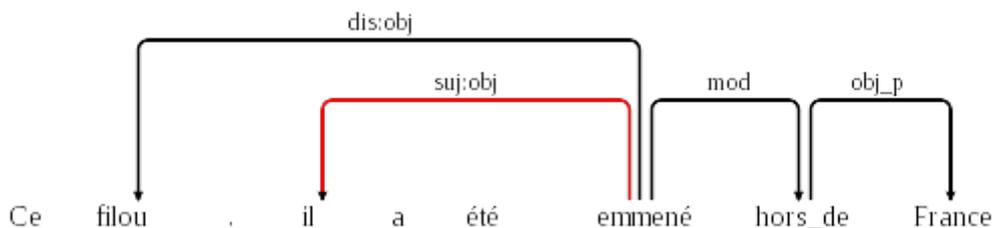
12.2 Dislocation

Le clitique redondant dans les dislocations est annoté également comme dépendant du verbe (et le disloqué porte une fonction finale **dis**, et comme fonction canonique, la fonction canonique de son clitique de reprise):

253. *Ce filou, il a plus d'un tour dans son sac*

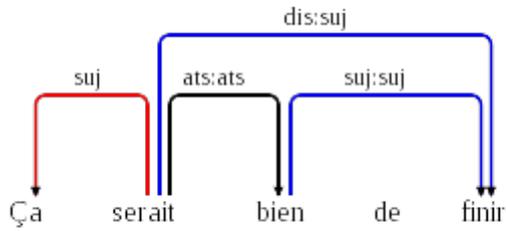


254. *Ce filou, il a été emmené hors_de France*

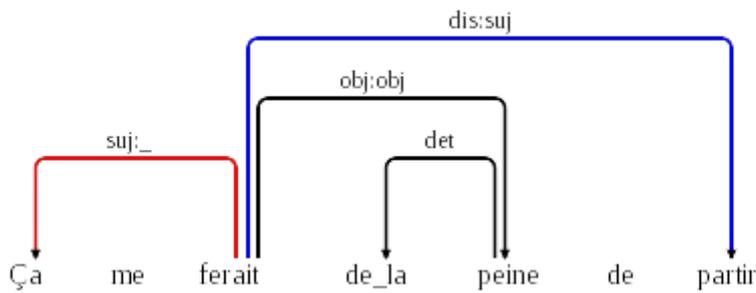


On applique la même représentation pour les constructions en *ça*

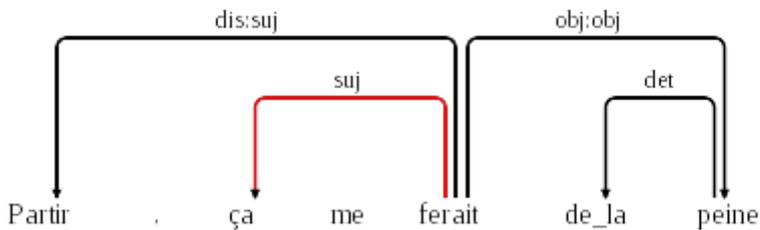
255. *Ça serait bien de finir*



256. Ça me ferait de_la peine de partir



257. Partir, ça me ferait de_la peine

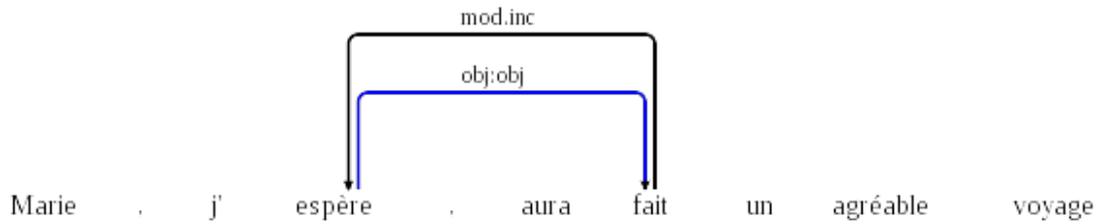


Rem : le FTB (arbres en constituants) ne note pas de fonction pour l'adverbe nu *bien*, alors qu'elles sont notées dans le schéma de dépendances de surface.

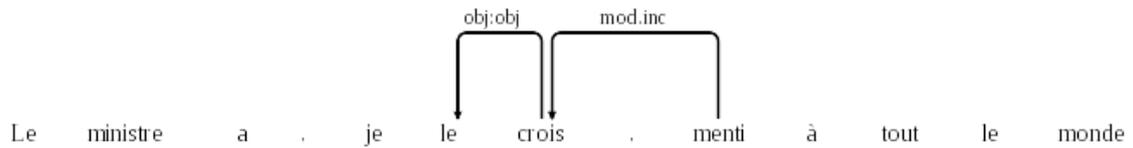
12.3 Incises

Au niveau surfacique, une incise apparaît comme modifieur du plus petit syntagme qui l'englobe. On le marque par une dépendance **mod.inc** de la tête du syntagme matrice vers la tête de l'incise. On ajoute au niveau profond et intermédiaire une dépendance **obj** du verbe de l'incise vers son objet, quand ce dernier est présent dans le syntagme matrice :

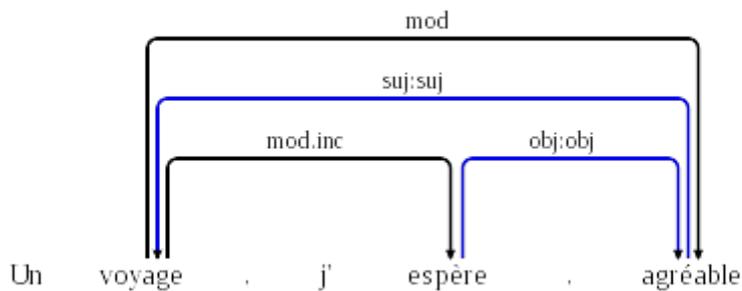
258. Marie, j'espère, aura fait un agréable voyage



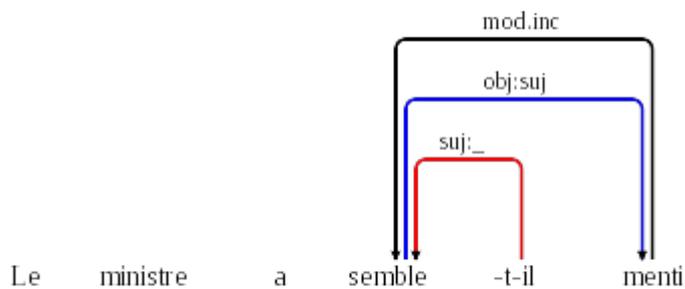
259. *Le ministre a, je le crois, menti à tout le monde*



260. *Un voyage, j'espère, agréable*



261. *Le ministre a semble-t-il menti* (cf. *il semble que le ministre a menti* a une diathèse impersonnelle le "que P" est objet final et sujet canonique du V à l'impersonnel)



Attention, on n'ajoute le complément du verbe de l'incise que si celui-ci admet ce complément dans une formulation hors incise :

262. *Le ministre a semble-t-il menti ... => il semble que le ministre a menti... => on ajoute*

263. *Marie, j'espère, aura fait ... => J'espère que Marie aura fait ... => on ajoute*

mais

264. *Le ministre a menti a tout le monde, ricana Paul => *Paul ricana que ... => on n'ajoute rien*

13 Coordination, énumération

NB: dans toute la suite, on ne parle que des fonctions finales, les changements de diathèse s'appliquant de manière régulière y compris dans le cas coordonné.

13.1 Représentation en dépendances de surface

Dans le cas d'une **coordination non double**, de la forme :

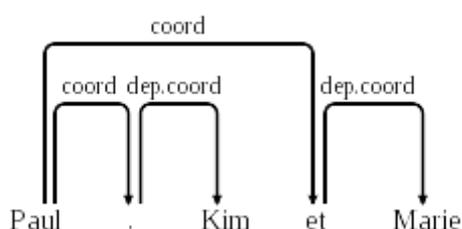
conjoint1 coord1 conjoint2 coord2 conjoint3 etc...

on a comme **représentation en surface**, le conjoint1 gouverne les coordonnants coord1 coord2 etc... et chaque conjoint_i est rattaché à coord_{i-1}

Lorsqu'une coordination comporte plus de deux conjoints, les virgules ainsi que la conjonction de coordination finale se rattachent toutes par la dépendance **coord** à la tête du premier conjoint.

Exemple :

265. *Paul, Kim et Marie*



Rem: cette représentation est directement reliée aux choix d'annotation du FTB. A noter que l'inconvénient est de ne pas pouvoir faire la différence en surface entre une dépendance sur le premier conjoint, valant en réalité pour tous les conjoints ou bien pour le premier conjoint uniquement.

On lève partiellement cette sous-spécification dans notre représentation profonde, selon qu'il s'agit d'une dépendance d'éléments coordonnés vers un dépendant ou bien d'un gouverneur vers des éléments coordonnés.

Ainsi on a la même topologie en surface entre "*Le jeune professeur et l'étudiant*" et "*Le professeur et l'étudiant jeunes*" (les dépendances seules ne permettent pas de trancher, il faut utiliser l'ordre linéaire). On lève cette sous-spécification en représentation profonde (voir infra distribution des dépendants d'éléments coordonnés).

Mais à l'inverse, on ne lève pas la sous-spécification dans le cas inverse : pour une dépendance entre un gouverneur et le premier conjoint d'une coordination, on n'indique pas si la dépendance vaut pour toute la coordination ou pour le premier conjoint seulement. Cette sous-spécification n'est pas résolue dans la représentation profonde, du fait du choix de ne pas distribuer les gouverneurs d'éléments coordonnés (voir infra).

13.2 Représentation en dépendances profondes : principe général

En représentation profonde, on "distribue" la coordination dans certains cas, c'est-à-dire que les dépendances portées par le premier conjoint sont explicitées sur tous les conjoints suivants.

Plus précisément : on utilise comme principe général que l'on ne distribue que les dépendances sortantes du premier conjoint (i.e. dont le conjoint est le gouverneur), et pas les dépendances entrantes, en outre la distribution n'est faite uniquement dans le cas où cela est sémantiquement correct (cf. 274 pour un exemple de cas où la distribution serait sémantiquement fausse).

Pour le dire plus précisément, on utilise deux grands principes pour décider des cas de "distribution" :

Principe 1 : On distribue les dépendances de coordonnés vers leur dépendant

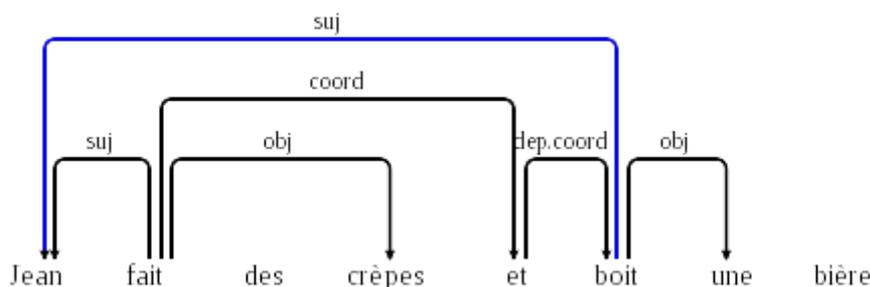
Exemples :

• *Paul dîne et dort* : "Paul" noté suj de "dîne" (déjà présent en surface) et de "dort" (à rajouter en représentation profonde)

• *Paul veut partir et revenir* : "partir et revenir" sont gouverneurs de "Paul" => "Paul" noté suj de "partir" et de "revenir"

Ce choix permet de compléter la structure argumentale des prédicats apparaissant comme xième conjoint (non premier) d'une coordination : par exemple pour 266 cela permet d'expliciter la structure argumentale de "boit" qui en surface apparaît sans sujet.

266. *Jean fait des crêpes et boit une bière*



On fait cependant des exceptions :

- **Exception 1** : on ne distribue pas si la distribution est clairement sémantiquement fausse
 - *les hommes et femmes amis* => on a les coordonnés "hommes et femmes" vers le dépendant modifieur "amis", mais il est faux de dire "les hommes amis" et "les femmes amis" => pas de distribution dans ce cas
- **Exception 2** : et on choisit de systématiquement ne pas distribuer pour les dépendants de type cardinal et/ou les déterminants, cf. la distribution est souvent fausse :
 - la distribution est par exemple fausse en cas de commentaire métalinguistique ("*l'électrocardiogramme ou ECG*") ou de coréférence ("*ce collègue et ami*") => dans tous les cas, le dépendant de type DET est noté déterminant du premier conjoint uniquement
 - (*les*) *CINQ hommes et femmes* => on ne distribue pas le cardinal "cinq", qui est noté modifieur du premier conjoint uniquement (hommes)

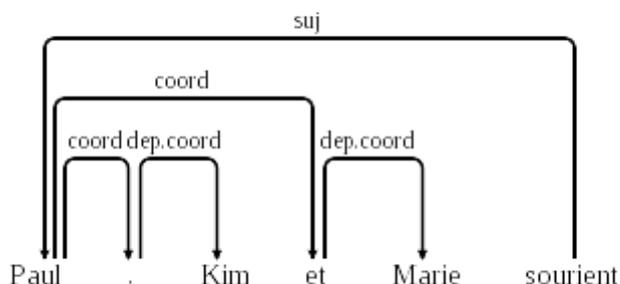
Principe 2 : On ne distribue pas les dépendances de coordonnés vers leur gouverneur :

Paul, Kim et Marie sourient => "Paul, Kim et Marie" dépendant profond (suj) de "sourient" => pas de distribution, on a en profond la même représentation qu'en surface, seul le premier conjoint (*Paul*) est suj de "sourient"

En effet, dans ce cas, le gouverneur a déjà tous ses dépendants (cf. le lien vers le premier conjoint). On reporte alors en

sémantique le fait de distinguer une lecture distributive vs une lecture collective de la coordination (dans "Paul et Marie sourient", on a équivalence sémantique avec "Paul sourit" et "Marie sourit", alors que dans "Paul et Marie partent pour Tokyo", on peut avoir une lecture collective où ils partent ensemble. De même, on ne peut pas réduire "Paul et Marie se disputent" à "Paul se dispute" et "Marie se dispute").

267. Jean voit Paul, Kim et Marie



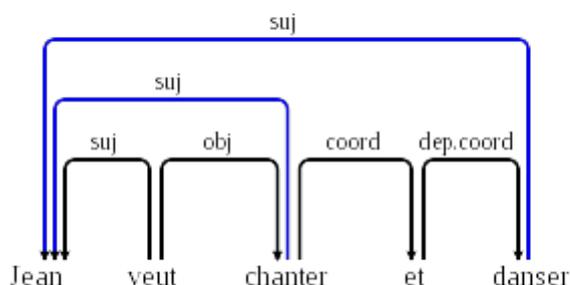
NB: ces principes s'appliquent également pour les dépendances n'apparaissant que dans la représentation profonde.

En particulier, on obtient une situation croisée dans le cas d'un **épithète** :

- pour un épithète coordonné : *un type content et fier*
 - la dépendance **mod** vers les épithètes coordonnés n'est pas distribuée
 - on distribue la dépendance **suj** (existante uniquement en représentation profonde) depuis les épithètes vers le nom modifié : "type" noté **suj** de "content" et de "fier"
- pour un épithète sur noms coordonnés, c'est l'inverse : *une fille et un gars contents*
 - la dépendance **mod** des coordonnés vers l'épithète est distribuée : "content" noté **mod** de "fille" et de "gars"
 - la dépendance **suj** depuis l'épithète vers les noms coordonnés n'est pas distribuée : "contents" noté **suj** de "fille" uniquement

Autre exemple : deux verbes coordonnés, non racine :

268. Jean veut chanter et danser => on distribue bien "Jean" comme sujet du deuxième conjoint "danser", mais "danser" n'est pas distribué comme objet de "veut".



13.3 Détail par catégories de conjoints

13.3.1 Coordination de syntagmes prépositionnels

Pas de distribution vers le gouverneur de la première préposition donc rien à faire pour :

- *de nouveaux sites furent créés à Hairoville, PUIS à Ancerville*
- ... *dans_le_cadre_d'* un projet école *ET des CATE*
- *restaurant bien connu des habitants ET des touristes*
- *pour éviter une reprise MAIS aussi pour prévenir des actes malveillants*

Distribution uniquement dans le cas où un dépendant vaut pour les deux prépositions coordonnées :

- *de Charybde et de Pluto en Scylla* => le "en Scylla" est distribué sur les deux "de"

Ce dernier cas est en fait inexistant, cf. on a plutôt un dépendant prépositionnel valant pour deux noms sous prep :
prep1 N1 coord prep2 N2 prep3 N3 : parfois "prep3 N3" se rattache à N1 et à N2

Exemple : *(des périodes) de paralysie ou de réduction de l'activité*

prep1=de N1=paralysie coord=ou prep2=de N2=réduction prep3=de l'activité

=> on veut bien comprendre "paralysie de l'activité" et "réduction de l'activité"

=> dans ce cas, on ne peut pas en surface noter que la prep3 se rattache au premier conjoint N1, car il est sous prep1

=> on obtient par conversion automatique que la prep3 dépend du N2, et **il faut noter en profond qu'elle dépend aussi du N1**

13.3.2 Coordination de prépositions seules

On a le cas différent où ce sont seulement les prépositions qui sont coordonnées, mais avec un seul objet de préposition pour les 2 prépositions :

- *avec ou sans paracétamol* => "paracétamol" est à noter comme dépendant **obj.p** de "avec" et de "sans"

13.3.3 Coordination d'adjectifs

Le principe général est que le nom auquel se rapporte la coordination d'adjectifs est noté comme sujet (profond) de tous les adjectifs coordonnés :

- adjectifs coordonnés épithètes :
 - *une longue et vieille histoire* => "histoire" est suj de "longue" et de "vieille"
 - NB: le lien inverse **mod** n'est pas distribué : "longue" est **mod** de "histoire"
- adjectifs coordonnés attributs :
 - *cette histoire est longue et vieille* => "histoire" est **suj** profond de "longue" et de "vieille"
 - NB: on garde un seul lien **ats** : "longue" est **ats** de "est"

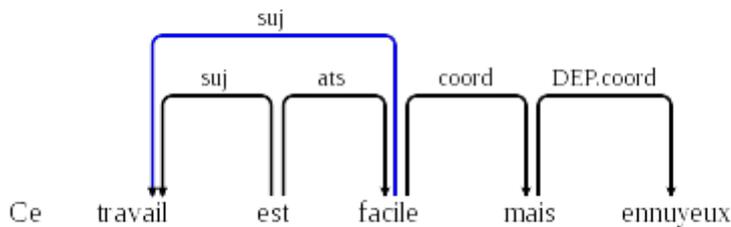
On fait cela sauf si c'est clairement sémantiquement faux :

- *un foulard noir et blanc* => il est faux de distribuer "foulard noir" et "foulard blanc"

269. *Ce tableau est rouge et blanc*



263. <ex>Ce travail est facile mais ennuyeux



</ex>

13.3.4 Coordination de participes

=> même principe que pour les adjectifs

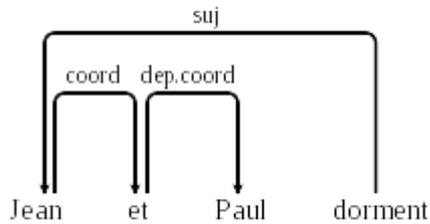
- participes coordonnés épithètes => le nom modifié (ou pronom, ou adj si ellipse...) est le sujet (final) profond du premier épithète et de tous les autres épithètes conjoints
 - *le produit reconstitué ET dilué* => "produit" est **suj** profond de "reconstitué" et de "dilué", mais on a un seul lien **mod** : "reconstitué" est **mod** de "produit"
- participes coordonnés attributs : rare cf. on a en général une analyse comme un passif

13.3.5 Coordination de nominaux (communs, propres ou pronoms)

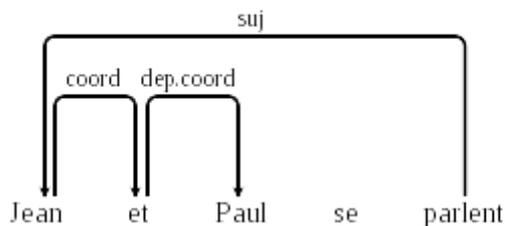
On ne distribue pas les noms coordonnés vers leur gouverneur, donc par exemple rien de spécial à faire pour :

- sujet nominal coordonné :
 - *le maire ET la directrice ont conduit ...*
 - *lorsque des utilisateurs VIRGULE associations OU particuliers les occupent ...*
- objet nominal de préposition, coordonné :
 - *le maire a remercié la municipalité ... pour la conception des locaux ET la qualité des matériaux*
- objet direct nominal coordonné :
 - *il ne reste plus aujourd'hui que trois clairons ET un tambour*

270. *Jean et Paul dorment*



271. *Jean et Paul se parlent*

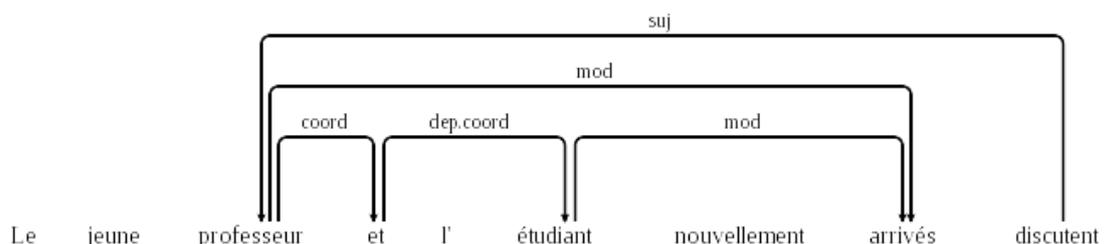


Mais on distribue les dépendants de noms coordonnés, sauf les relations *det* et les cardinaux, donc on distribue un dépendant prép, et un modifieur adjectival, participial, une relative, si la distribution est sémantiquement vraie. Le test est de répéter le dépendant sur tous les conjoints.

- la *directrice* ET le *sous-directeur de l'association* => "de" **dep** de "directrice" et de "sous-directeur" (cf. ok sémantiquement "la directrice de l'association et le sous-directeur de l'association")
- la *directrice* ET le *maire, convaincus* de ... => "convaincus" **mod** de "directrice" et de "maire"
- la *directrice* ET le *maire qui ont décidé* de => "décidés" **mod.rel** de "directrice" et de "maire"

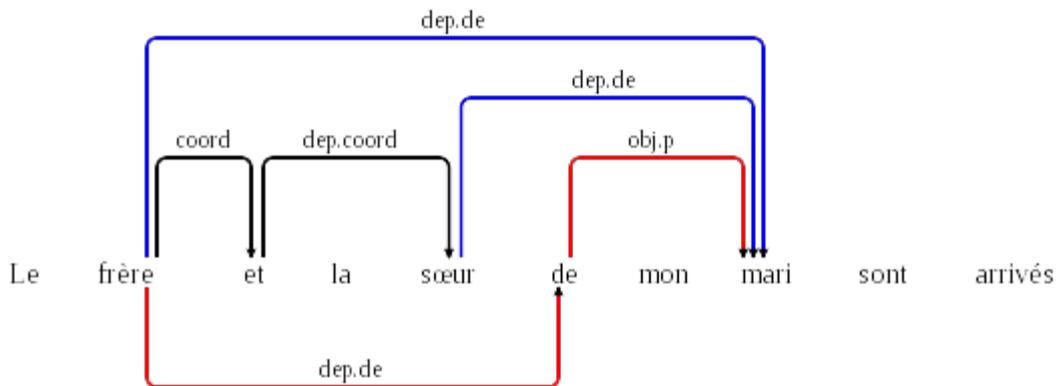
Concernant les dépendants de noms coordonnés, en particulier les modifieurs, leur distribution fait que contrairement à la représentation de surface, on obtient bien une différence topologique dans la représentation profonde entre un modifieur partagé par les différents conjoints et un modifieur d'un seul conjoint. Par exemple ci-dessous, "arrivés" est **mod** de "professeur" et de "étudiant", alors que "jeune" l'est seulement de "professeur". En revanche la non-distribution des gouverneurs d'éléments coordonnés fait que pour une dépendance entre un gouverneur et le premier conjoint d'une coordination, on laisse sous-spécifié si la dépendance porte sur tous les conjoints ou sur le premier uniquement.

272. *Le jeune professeur et l'étudiant nouvellement arrivés discutent*



273. *Le frère et la sœur de mon mari sont arrivés* => on a bien sémantiquement "le frère de mon

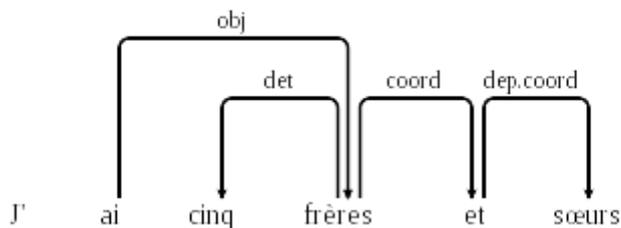
mari / la sœur de mon mari", dc on partage le **dep** prépositionnel sur les deux noms coordonnés



En revanche pour *les hommes et femmes amis*, la distribution n'est pas à faire, cf. elle serait sémantiquement fausse.

Les déterminants et cardinaux ne sont pas distribués sur les noms coordonnés :

274. *J'ai cinq frères et sœurs* (≠ *J'ai cinq frères et j'ai cinq sœurs*)



13.3.6 Coordination de verbes conjugués

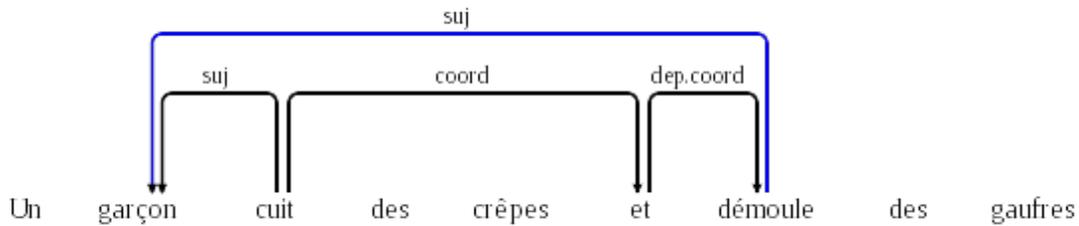
On peut avoir certains des dépendants qui sont en réalité partagés par les différents verbes. Le test est de voir si répéter ces dépendants sur les différents verbes coordonnés est bien équivalent sémantiquement.

On a en particulier le cas fréquent de "partage du sujet" : le sujet n'est pas répété sur les nièmes conjoints => on rajoute en profond la dépendance suj sur les nièmes verbes conjoints.

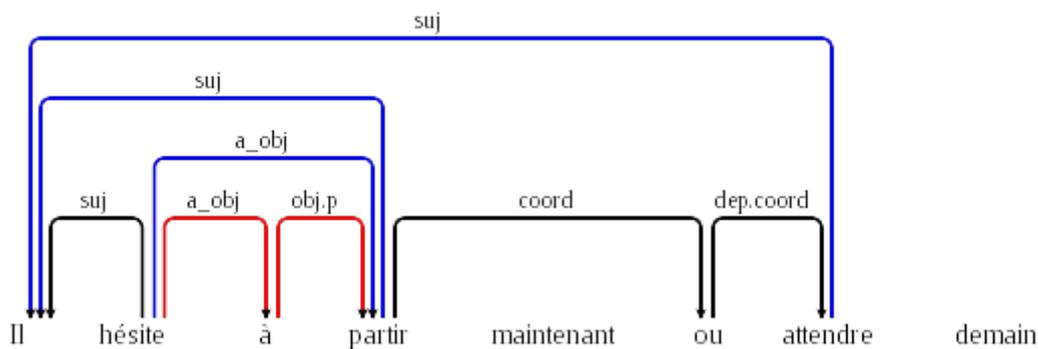
X V1 compléments1 et V2 compléments2 : si le deuxième verbe n'a pas lui-même son sujet, on ajoute que X est sujet de V2

- le **visiteur** *enrichit ses connaissances ET apprend que ...*
- elle *travailla quelques temps PUIS épousa Gaston*

275. *Un garçon cuit des crêpes et démoule des gaufres*



276. *Il hésite à partir maintenant ou attendre demain*



Attention, si on avait : *elle travailla quelques temps PUIS elle épousa Gaston*
=> on ne rajoute rien : le sujet de "épousa" est le deuxième "elle".

(correspond au cas coordination de phrases).

Plus généralement, dans ce cas de coordination de phrases, aucun dépendant n'est partagé.

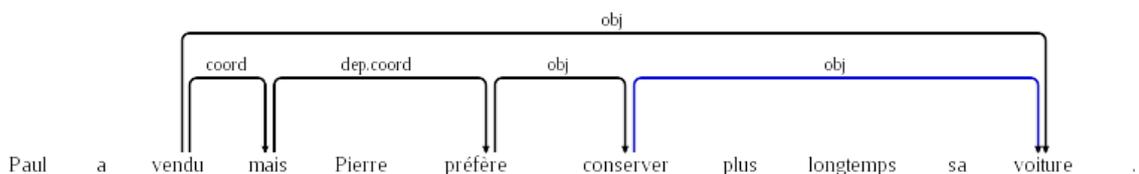
Nous avons bon espoir ..., CAR dans l'usine les machines se sont arrêtées

On a quelques cas où c'est un autre type de dépendant qui est partagé entre plusieurs verbes:

- Paul a cuit ET Marie a vendu ces **crêpes** => on distribue "crêpes" comme objet de "cuit" et "vendu" (= Right node raising)
- quand **nous constatons** ET **convenons** tous qu'**il y a** un problème => ici c'est à la fois le sujet (nous) et la complétive (qu'il y a un problème) qui sont partagés par les deux verbes coordonnés

Même type d'exemple avec partage du sujet, et partage de l'objet (Right node raising). NB: dans cet exemple, le NP "sa voiture" est partagé par les deux verbes, mais ne réfère pas à la même entité dans les 2 cas. Ceci n'est pas explicité dans la représentation syntaxique profonde.

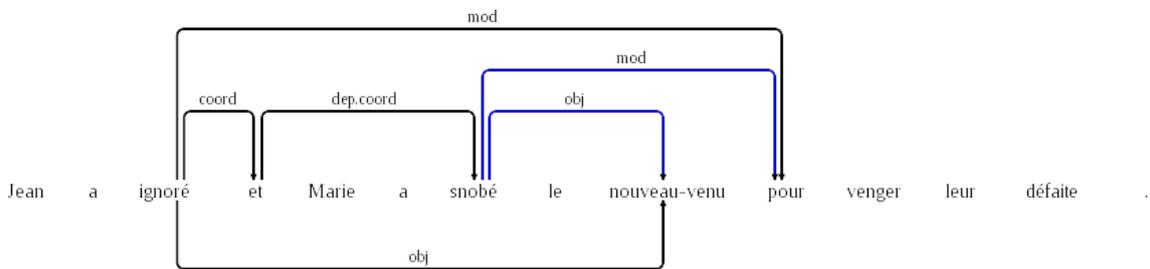
277. *Paul a vendu mais Pierre préfère conserver plus longtemps sa voiture.*



Dans l'exemple suivant, on a un partage de l'objet ("le nouveau-venu") et également du modifieur ("pour venger leur défaite"), sur les deux verbes coordonnés, qui ont cependant chacun leur sujet

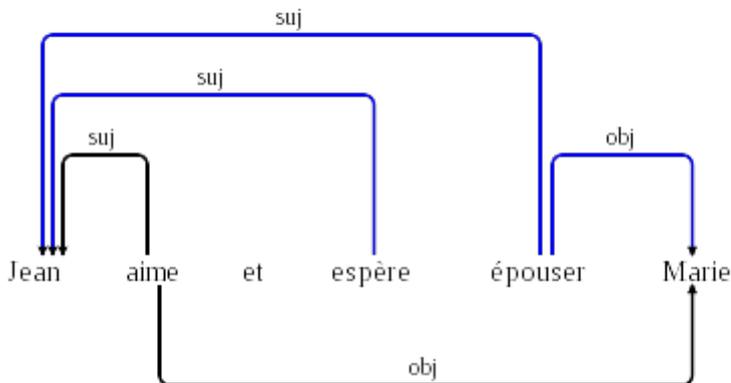
distinct :

278. *Jean a ignoré et Marie a snobé le nouveau-venu pour venger leur défaite.*



A noter que le partage de dépendants de verbes coordonnés peut interagir avec le traitement du sujet des infinitifs, comme par exemple ci-dessous:

279. *Jean aime et espère épouser Marie*



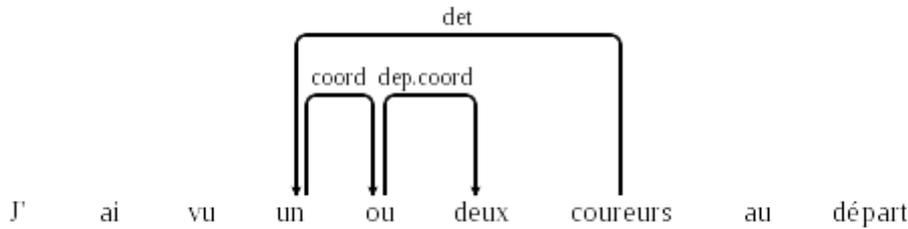
13.3.7 Coordination d'infinitifs

=> même principe, on partage les dépendants : en pratique il s'agit essentiellement du sujet, s'il est repérable
Cela permet à X d' échanger VIRGULE transmettre un savoir ET partager
=> X à noter comme sujet de "échanger", "transmettre" et "partager"

13.3.8 Coordination de déterminants et/ou de cardinaux

On n'a en général pas de dépendants des déterminants ou cardinaux, donc aucune dépendance n'est ajoutée en représentation profonde.

280. *J'ai vu un ou deux coureurs au départ*

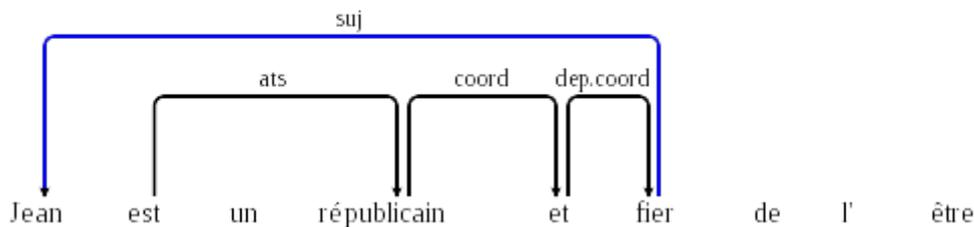


13.4 Cas plus complexes

13.4.1 Coordination hétérogène

Elle est traitée en suivant les principes généraux d'annotation de la coordination. Donc par exemple ci-dessous, pour une coordination d'un nom et d'un adjectif, on ajoute bien le ou les dépendants de l'adjectif deuxième conjoint :

281. *Jean est un républicain et fier de l'être*



13.4.2 Coordination d'arguments agrégés (ou "coordination de séquences de constituants" ou "arg cluster")

Voir par exemple : [frwiki_50.1000_00587](https://frwiki.50.1000.00587).

Il s'agit de cas où ce sont 2 groupes de dépendants qui sont coordonnés. Le groupe de dépendants peut comporter le sujet, il s'agit alors d'une "coordination à trous" ou "head gapping" (283) ou bien le groupe de dépendants ne comporte pas le sujet, il s'agit alors du cas dit "coordination de séquence de constituants" ou "arg cluster". Par exemple dans *Paul offre des baskets bleues à la petite fille et une casquette au garçon*, le premier conjoint est le groupe "des baskets bleues"+"à la petite fille", et le deuxième conjoint est le groupe "une casquette"+"au garçon".

Un des objectifs du schéma d'annotation défini ici est d'expliciter les structures argumentales de prédicats. Cet objectif nécessiterait d'expliciter plusieurs instances de prédicats pour une seule occurrence effective. Cependant dans cette version du schéma en dépendances profondes, on s'interdit d'ajouter des nœuds.

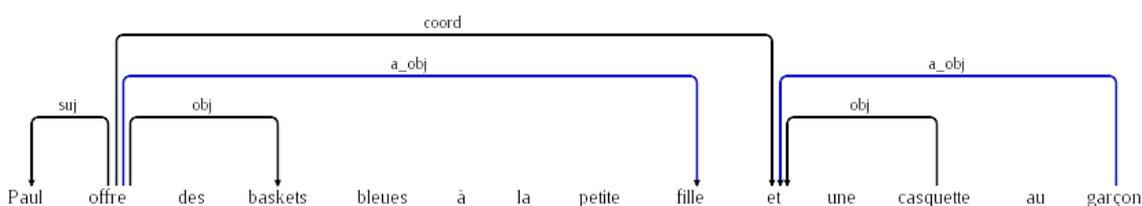
On utilise alors en représentation profonde simplement le traitement fait en représentation de

surface, qui ne répète pas de prédicat, mais qui permet tout de même un repérage de ces cas. On hérite en effet du schéma d'annotation du FTB le fait que le coordonnant introduisant le deuxième conjoint est rattaché sur le verbe, et pas comme habituellement sur le premier conjoint (cf. celui-ci est un groupe de dépendants) : donc en (282) le "et" est rattaché au verbe "offre".

Ensuite, les éléments introduits par le coordonnant (le groupe conjoint "une casquette"+"au garçon") dépendent du coordonnant non pas comme habituellement avec une relation dep.coord, mais **en utilisant les fonctions grammaticales pertinentes par rapport au verbe**. Donc "casquette" est **obj** de "et", et "garçon" est **a_obj** de "et".

Les différences entre représentation de surface et représentation profonde tiennent à d'autres aspects comme infra simplement au court-circuitage des prépositions régies:

282. *Paul offre des baskets bleues à la petite fille et une casquette au garçon*



13.4.3 Coordination à trou (ou "head gapping")

Voir par exemple : [frwiki_50.1000_00843](#).

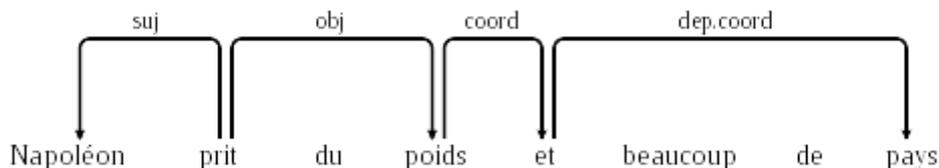
On utilise le même principe : les dépendants apparaissant dans la "clause à trous" sont rattachés au coordonnant, mais portent la fonction grammaticale pertinente par rapport au verbe. Donc en 283 "Paul" est **suj** et "Virginie" est **obj**.

283. *Jean aime Marie et Paul Virginie*

13.4.4 Zeugme

L'ambiguïté du prédicat n'est pas levée.

284. *Napoléon prit du poids et beaucoup de pays*

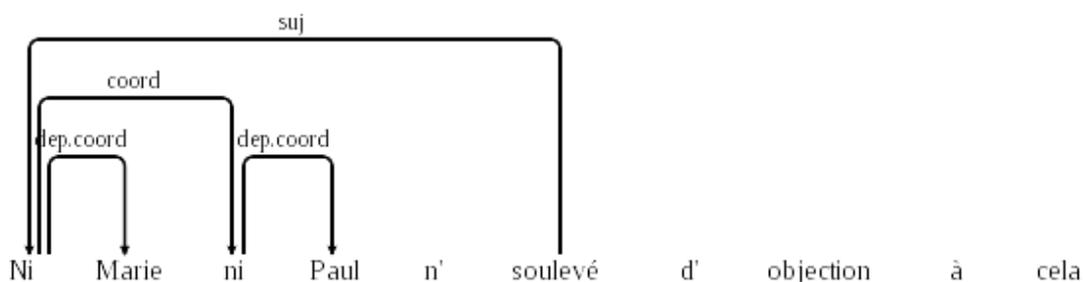


13.4.5 Coordination double

Il s'agit de coordinations de la forme « ni...ni... » ou « soit...soit... ». En surface comme en

profondeur, on essaie de se calquer sur une coordination classique en considérant la première occurrence de la conjonction double comme tête de la coordination. La seconde occurrence de la conjonction double est liée à la première par une dépendance de type **coord**. Les deux conjonctions de coordination sont les gouverneurs des têtes effectives des conjoints, avec une dépendance **dep.coord**.

285. Ni Marie ni Paul n'ont soulevé d'objection à cela

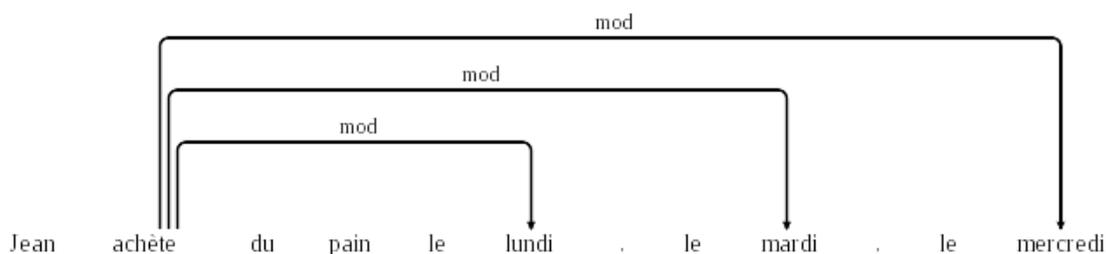


13.5 Enumération

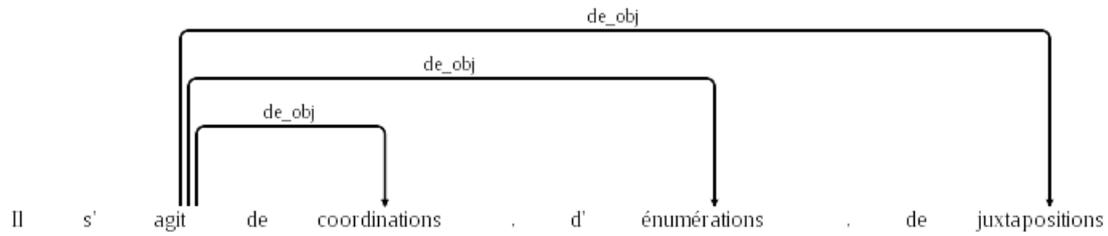
Pour les énumérations, on suit le FTB en faisant de chaque membre de l'énumération un argument à part entière, portant chacun le même label par rapport à leur gouverneur. On a la même représentation en surface et en profond (modulo le shuntage des mots vides)

Remarque : une raison pour laquelle cette "distribution" des dépendances n'est pas faite pour les coordinations est que la distribution n'est pas toujours vraie sémantiquement (par exemple *Pierre et Marie se regardent* \neq *Pierre se regarde ET Marie se regarde*). Pour une énumération, au contraire les éléments énumérés ne forment pas un référent collectif, sur lequel puisse porter collectivement une prédication.

286. Jean achète du pain le lundi, le mardi, le mercredi



287. Il s'agit de coordinations, d'énumérations, de juxtapositions



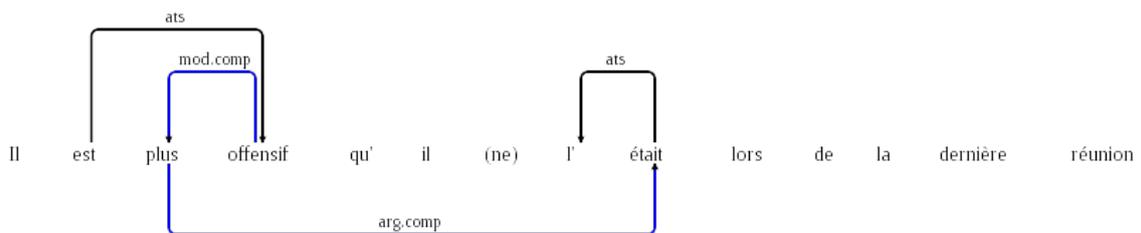
14 Comparatives

Les adverbes comparatifs (*plus, moins, aussi, davantage* ...) dépendent en surface de leur gouverneur (le terme de la comparaison) par une relation **mod**, qui est précisée en représentation profonde comme un **mod.comp**. (voir la distinction entre **superlatif** et **comparatif** [Section 11.1.1](#)).

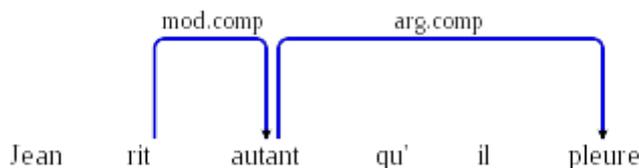
Dans le cas d'une construction comparative avec subordonnée, la subordonnée dépend en surface de l'élément modifié par l'adverbe. En dépendances profondes, on prend l'adverbe comparatif comme gouverneur de la subordonnée, avec une dépendance **arg.comp** (dont le complémenteur *que* est court-circuité).

La subordonnée peut contenir ou pas une ellipse, le cas le plus simple étant celui sans ellipse.

288. *Il est plus offensif qu'il (ne) l'était lors de la dernière réunion*



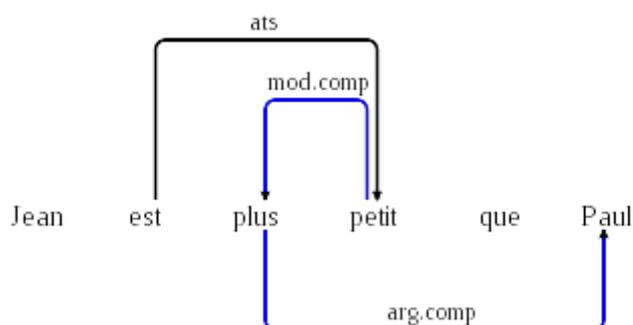
289. *Jean rit autant qu'il pleure*



Toutefois, le cas le plus courant est celui des comparatives avec ellipse de la tête prédicative. Dans ce cas, un traitement complet requerrait de rajouter un nœud pour la tête prédicative, pour que les

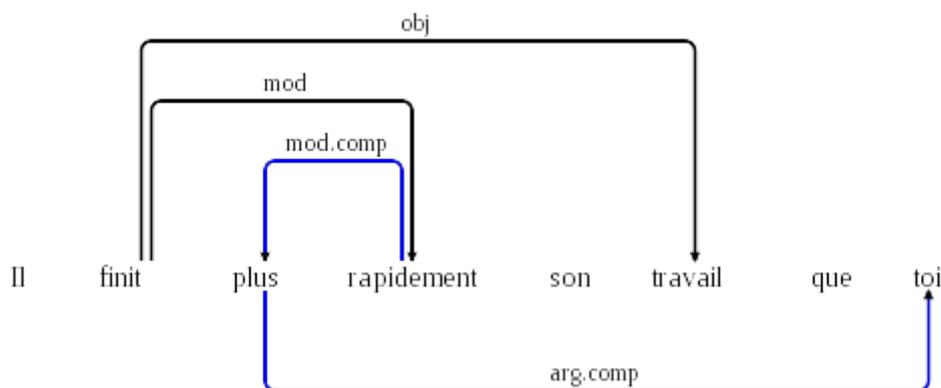
différents dépendants dans la subordonnée puissent y être attachés. Mais en l'absence d'ajout de nœuds dans nos REPRSYNTPROF, une tête est sélectionnée dans la subordonnée (on utilise comme tête de la subordonnée la tête du dernier syntagme). NB: On sous-spécifie le rôle que joue cette tête par rapport au prédicat éliidé.

290. *Jean est plus petit que Paul*



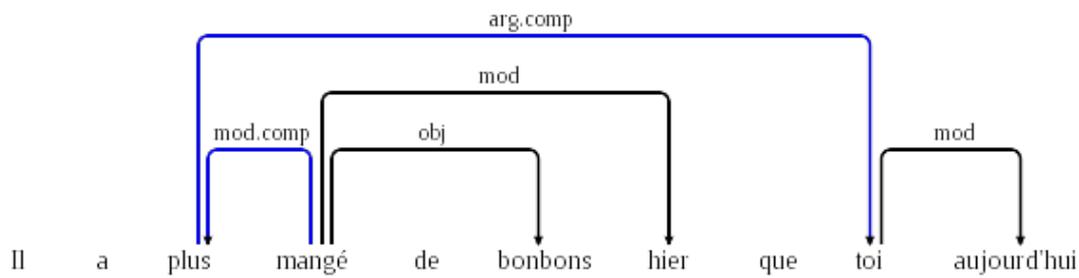
On peut avoir une discontinuité entre l'élément modifié par l'adverbe comparatif et la subordonnée comparative.

291. *Il finit plus rapidement son travail que toi*



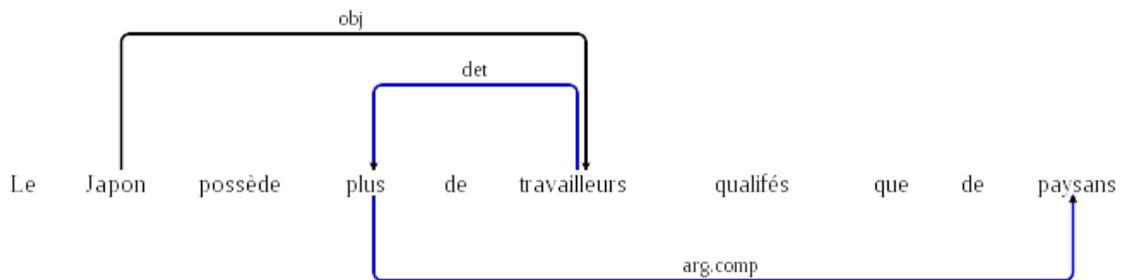
Cas de quantifieur flottant : lorsque l'adverbe est placé sur le verbe, on garde une représentation surfacique, où l'adverbe comparatif modifie le verbe.

292. *Il a plus mangé de bonbons hier que toi aujourd'hui*



Exceptionnellement, dans le cas d'une comparative avec adverbe comparatif modifiant un nom, l'adverbe comparatif est **det** et pas **mod.comp**, comme en [Section 5.1.3](#).

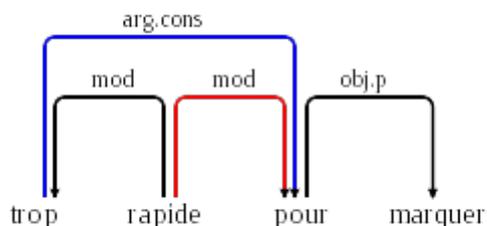
293. *Le Japon possède plus de travailleurs qualifiés que de paysans*



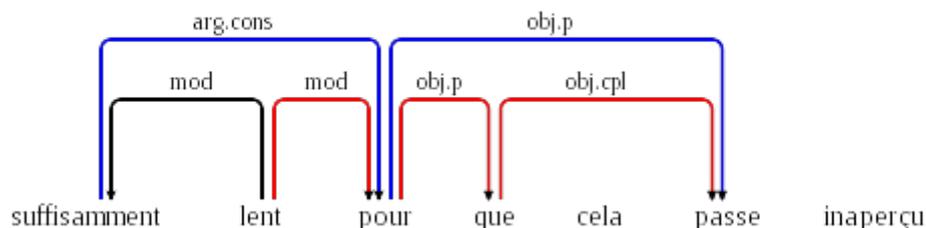
15 Consécutives

Les constructions consécutives sont introduites par un adverbe (*trop, assez, tant, tellement, si ...*) ou un adjectif (*tel*) qui est suivi d'une complétive introduite par *que* ou d'une infinitive introduite par *pour* ([voir corpus](#)). En surface, on exprime la dépendance de *que* ou de *pour* par rapport à la tête du syntagme introduisant l'adverbe ou l'adjectif introduisant la construction en la marquant **mod**. En syntaxe profonde, on remplace cette dépendance par une dépendance **arg.cons** qui lie directement l'adverbe ou l'adjectif introduisant la construction à *pour* ou au verbe gouverné par *que*.

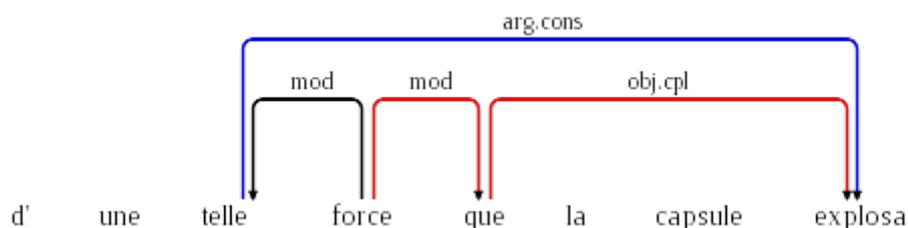
294. *trop rapide pour marquer*



295. *suffisamment lent pour que cela passe inaperçu*



296. *d'une telle force que la capsule explosa*



16 Références

ABEILLÉ A., TOUSSENEL F. & CHÉRADAME M., (2004). [Corpus Le Monde, Annotation en constituants, guide pour les correcteurs](#)

ABEILLÉ A. (2004). [Guide des annotateurs : ANNOTATION FONCTIONNELLE](#)

ABEILLÉ A. & CLÉMENT L. (2003): [Corpus le Monde — Annotation morpho-syntaxique — Les mots simples - Les mots composés](#)

ABEILLÉ A. & BARRIER N. (2004). Enriching a french treebank. In LREC 2004, Lisbon.

ABEILLÉ A. , CLÉMENT L. & TOUSSENEL F. (2003). Building a treebank for French, In Treebanks. Kluwer : Dordrecht.

- BASHUNG K. (1996) : Une approche lexicalisée des phénomènes de contrôle. *Revue Langages*, n°122, 1996, pp. 96-122.
- BONFANTE G., GUILLAUME B., MOREY M. et PERRIER G. (2011) : Enrichissement de structures en dépendances par réécriture de graphes. *Actes de TALN 2011*, Montpellier, 2011.
- CADIOT. (1990) : Contrôle anaphorique et prépositions. *Revue Langages*, n° 97, 1990, pp. 8-23.
- CARROLL J. , BRISCOE T. & SANFILIPPO A. (1998). Parser evaluation : a survey and a new proposal. In *LREC 1998*, Granada.
- CANDITO M. (1999). Organisation modulaire et paramétrable de grammaires électroniques lexicalisées. Thèse de doctorat. Université Paris 7.
- CANDITO M., CRABBÉ B., DENIS P. & GUÉRIN F. (2009). Analyse syntaxique du français: des constituants aux dépendances. In *TALN 2009*, Senlis.
- CANDITO M. et SEDDAH D. (2012a) Le corpus Sequoia : annotation syntaxique et exploitation pour l'adaptation d'analyseur par pont lexical. In *TALN'2012*, Grenoble, France
- CANDITO M. et SEDDAH D. (2012b) Effectively long-distance dependencies in French : annotation and parsing evaluation, Marie Candito and Djamé Seddah. In *TLT'11*, Lisbon, Portugal
- CANDITO M., PERRIER G., GUILLAUME B., RIBEYRE C., FORT K., SEDDAH D., de la CLERGERIE E. (2014) Deep Syntax Annotation of the Sequoia French Treebank. In *LREC 2014*, Reykjavik, Island, pages 2298-2305.
- CANDITO M. (2013) Guide d'annotation du clitique réfléchi. Rapport technique <http://www.linguist.univ-paris-diderot.fr/~mcandito/Publications/Guide-annotation-se.pdf>
- CRABBÉ B. & CANDITO M. (2008). Expériences d'analyse syntaxique statistique du français. In *TALN 2008*, p. 45-54, Avignon.
- CREISSELS D. (2007), Réflexivisation, transitivité et agent affecté, in Rousseau, A., D. Bottineau & D. Roulland (éds.), 2007, *L'énoncé réfléchi*, 83-106, Rennes, Presses Universitaires de Rennes.
- KAHANE S. (2001). Grammaires de dépendances formelles et théorie sens-texte. In *TALN 2001*, Tours, France.
- KING T. , CROUCH R. , RIEZLER S. , DALRYMPLE M. & KAPLAN R. (2003). The parc 700 dependency bank. In *EACL workshop on Linguistically Interpreted Corpora*, Budapest.
- DE MARNEFFE M. , MACCARTNEY B. & MANNING C. D. (2006). Generating typed dependency parses from phrase structure parses. In *LREC 2006*.
- PAROUBEK P., POUILLOT L.-G. , ROBBA I. & VILNAT A. (2005). Easy : campagne d'évaluation des analyseurs syntaxiques. In *Actes de TALN'05, Atelier EASy : campagne d'évaluation des analyseurs syntaxiques*, Dourdan.
- PERRIER G., CANDITO M., GUILLAUME B., RIBEYRE C., FORT K., SEDDAH D. (2014) Un schéma d'annotation en dépendances syntaxiques profondes pour le français. In *TALN 2014*, pages 574-579, Marseille.

TASMOVSKI L. (1984) *Lui faire téléphoner quelqu'un d'autre: une stratégie? In *Linguisticae Investigationes VIII:2*, pp. 403-427.

VEELOCK C. (2008) Se + faire + infinitif : valeurs pragmatico-énonciatives d'une construction agentive. In *CMLF 2008*, Paris, juillet 2008.