

# ASFALDA French FrameNet - Guide d'annotation

Marie Candito et Marianne Djemaa

Dernière révision le 26 février 2017  
<https://sites.google.com/site/anrasfalda/>

## Contents

<b>1</b>	<b>Contexte et vocabulaire</b>	<b>3</b>
<b>2</b>	<b>Cycle d'annotation</b>	<b>4</b>
2.1	Pré-annotation automatique . . . . .	4
2.1.1	Pré-annotation des déclencheurs composés . . . . .	5
2.1.2	Frame spécial Other_sense . . . . .	5
2.1.3	Préannotation des verbes pronominaux . . . . .	5
2.2	Double annotation . . . . .	6
2.3	Adjudication . . . . .	7
2.4	Post-traitement automatique . . . . .	7
2.4.1	Tête syntaxique . . . . .	7
2.4.2	Tête sémantique . . . . .	8
2.5	Demande d'évolution du lexique . . . . .	8
2.6	Accès au contexte de la phrase à annoter . . . . .	8
<b>3</b>	<b>Annotation d'un frame: principes généraux</b>	<b>9</b>
3.1	Dans quel cas annoter une instance de frame? . . . . .	9
3.1.1	Annotation indépendante de la factivité . . . . .	9
3.1.2	Déclencheur avec sens générique / déclencheur dans entité nommée . . . . .	9
3.2	Alternances syntaxiques . . . . .	9
3.3	Emplois métaphoriques / figurés . . . . .	11
3.4	Métonymie . . . . .	11
3.5	Redondance de déclencheurs ou pléonasmes . . . . .	12
<b>4</b>	<b>Annotation des remplisseurs: principes généraux</b>	<b>13</b>
4.1	Quels sont les rôles à annoter ? . . . . .	13
4.2	Utilisation de l'annotation en syntaxe . . . . .	14
4.2.1	Gestion des erreurs dans l'annotation syntaxique / divergences syntaxe-sémantique . . . . .	14
4.3	Mots fonctionnels introduisant des remplisseurs . . . . .	14
4.3.1	Prépositions et complémenteurs vides . . . . .	14
4.3.2	Coordination / énumération de prépositions régies . . . . .	15
4.3.3	Coordination de préposition composée . . . . .	15
4.3.4	Coordination de prépositions qui correspond sémantiquement à une ellipse de prédicat . . . . .	16
4.4	Localité des remplisseurs . . . . .	16
4.4.1	NLI: remplisseurs non-locaux . . . . .	16

4.4.2	Imbrication de remplisseurs . . . . .	18
4.4.3	DNI : Definite null instantiation . . . . .	18
4.4.4	ENI : Extra-linguistic null instantiation . . . . .	19
4.4.5	UNI : Unknown null instantiation . . . . .	19
4.5	Remplisseurs anaphoriques . . . . .	20
4.5.1	Relatives . . . . .	21
4.5.2	Réfléchis . . . . .	21
4.5.3	Possessifs . . . . .	22
4.5.4	Cas particuliers: "en" quantitatif / démonstratif "celui" . . . . .	22
4.5.5	Aspects techniques : 2 instances de remplisseurs, versus un remplisseur "multi-pattes" . . . . .	23
<b>5</b>	<b>Points spécifiques aux déclencheurs verbaux</b>	<b>24</b>
5.1	Ellipses de verbes . . . . .	24
5.1.1	Head Gapping : coordination, où le verbe n'est pas répété (mais 2 sujets différents) . . . . .	24
5.1.2	Argument cluster (i.e. "coordination de séquences de constituants") . . . . .	24
5.1.3	Comparatives réduites . . . . .	25
<b>6</b>	<b>Points spécifiques aux déclencheurs nominaux</b>	<b>26</b>
6.1	Nominalisation référant à un évènement versus nominalisation référant à un résultat de l'évènement . . . . .	26
6.1.1	Critères de distinction entre référence à une éventualité versus référence à un participant ou résultat . . . . .	27
6.2	Noms relationnels et frames représentant une relation . . . . .	29
6.3	Usage prédicatif versus usage référentiel d'un déclencheur nominal . . . . .	31
6.3.1	Liste non exhaustive des types d'emplois prédicatifs (où le nom déclencheur n'est pas en outre remplisseur) . . . . .	31
6.3.2	Exemples d'emplois référentiels . . . . .	33
6.4	Noms quantificateurs . . . . .	33
6.5	Constructions à verbe support . . . . .	34
<b>7</b>	<b>Annotation des remplisseurs: points divers</b>	<b>34</b>
7.1	Remplisseurs collectifs versus rôle _1 / rôle _2 . . . . .	34
7.2	Astuce dans le cas de déclencheurs qui dépendent syntaxiquement d'un de leur arguments sémantiques . . . . .	34
7.2.1	Adjectifs et participes épithètes . . . . .	34
7.2.2	Incises . . . . .	35
7.3	Remplisseur adjectival synonyme d'un PP . . . . .	36
<b>8</b>	<b>Phénomènes de portée</b>	<b>36</b>
8.1	remplisseurs phrastiques/infinitivaux: gestion des verbes introducteurs de complétives/d'infinitives . . . . .	36
8.1.1	Portée sémantique sur verbe enchâssé versus portée sur verbe introducteur . . . . .	37
8.1.2	Difficultés de détermination de portée . . . . .	37
8.1.3	Gestion dans Asfalda : classes de verbes introducteurs . . . . .	38
8.2	Portée de la négation et des modaux dans un remplisseur . . . . .	40

<b>9 Saturation des remplisseurs prédictifs et divergence entre tête syntaxique et tête sémantique</b>	<b>41</b>
9.1 Saturation des remplisseurs infinitivaux . . . . .	41
9.2 Saturation des remplisseurs adjectivaux . . . . .	41
9.2.1 Remplisseur adjectif ou participe épithète et divergence entre tête syntaxique et sémantique . . . . .	42
<b>10 Bibliographie</b>	<b>44</b>

## 1 Contexte et vocabulaire

Le projet ASFALDA (projet ANR-12-CORD-023) vise à produire un "FrameNet du français", sur le modèle de FrameNet (Baker et al., 1998) développé pour l'anglais. Ce guide suppose une certaine familiarité avec les concepts de base de FrameNet, voir en particulier (Fillmore, 2007).

- **frame** : un frame est une représentation schématique d'une éventualité (au sens événement ou état) ou d'une entité. Plus précisément, une représentation schématique de la connaissance qu'ont les locuteurs d'une éventualité ou entité, sous-tendant le sens des unités lexicales (Fillmore, 2007). Un frame est constitué d'une définition fixant son périmètre sémantique du frame, d'une liste de **déclencheurs** (des unités lexicales pouvant évoquer le frame), et d'une liste de **rôles**. Dans l'exemple précédent, on considère que
- **rôle<sup>1</sup>** : les rôles sont des composants nommés d'un frame, et correspondent aux participants et propriétés des situations représentées par le frame. Un rôle a un nom (en anglais), une définition, et est spécifique à un frame. Voir section 4.1 pour une précision du lien entre les rôles et la valence syntaxique des déclencheurs.
- **déclencheur<sup>2</sup>**: un déclencheur est un mot (simple ou composé) pouvant évoquer un frame donné. Par exemple le verbe *convaincre* peut évoquer le frame FR\_Cognizer\_affecting.veracity : "un Cognizer est persuadé de la véracité d'un certain Content, soit du fait d'un State\_of\_affairs soit parce qu'un Persuader humain l'en persuade par la parole".
- **remplisseur (de rôle)**: il s'agit en contexte d'une expression linguistique qui instancie un rôle donné. Par exemple dans *M. Fitz a convaincu les députés que l'identité du témoin ne devait pas être révélée.*, le verbe *convaincre* déclenche le frame FR\_Cognizer\_affecting.veracity, et les expressions "*M. Fitz*", "*les députés*" et "*l'identité du témoin ne devait pas être révélée*" remplissent respectivement les rôles Persuader, Cognizer et Content.

Le FrameNet du français visé comprendra :

- **une structure de frames**: un ensemble de frames, reliés par des relations entre frames. Un frame décrit une éventualité (état ou événement) prototypique ou une entité. Il est constitué d'une définition fixant son périmètre sémantique du frame, et d'une liste de rôles correspondant
- **un lexique**: pour chacun des frames, les lemmes (français) pouvant évoquer ces frames.

<sup>1</sup>Appelé "frame element" dans la littérature framenet.

<sup>2</sup>Appelé "frame evoking element" dans la littérature framenet.

- des **annotations** de frames et rôles en corpus

Dans une première phase du projet (Candito et al., 2014), les 2 premières parties ont été constituées:

- des domaines notionnels ont été retenus, et les frames relevant de ces domaines ont été définis (en partant des frames existant pour l'anglais). On parlera ci-après des "frames Asfalda".
- le lexique pour les frames Asfalda a été mis au point, domaine par domaine. Le travail a été manuellement réalisé par des "experts" par domaine, en partant d'une souche automatiquement obtenue du lexique anglais.

Le présent guide concerne la tâche subséquente d'annotation de ces frames en corpus. Même si le travail manuel sur le lexique a été fait en ayant accès à un concordancier pour visualiser les occurrences dans le corpus visé, il est prévu que le lexique évolue au cours de la phase d'annotation, mais toujours après validation par les "experts" de chaque domaine.

Le corpus retenu est la concaténation de deux corpus arborés (i.e. ayant fait l'objet d'une annotation syntaxique):

- le Sequoia treebank (Candito and Seddah, 2012) (dans sa version en dépendances, version 7.0, 3099 phrases)
- le French Treebank (Abeillé and Barrier, 2004) (dans sa version en dépendances, telle que livrée à l'occasion de la shared task SPMRL 2013, 18535 phrases (Seddah et al., 2013))

Dans tout le projet, l'**identifiant d'une phrase** est son rang dans la concaténation de ces deux corpus.

## 2 Cycle d'annotation

Le cycle d'annotation comprend une pré-annotation automatique, une double annotation manuelle, une adjudication, et un post-traitement automatique.

### 2.1 Pré-annotation automatique

Les occurrences de déclencheur à annoter sont repérées par un pré-annotateur automatique, qui assigne à chaque occurrence tous les frames possibles pour le déclencheur en question. La tâche manuelle restante est de choisir le bon frame et d'annoter ses rôles.

Etant donnée la difficulté à maîtriser le contenu des différents frames, et de certaines frontières entre frames, on a choisi de ne pas annoter linéairement le corpus: des groupes de déclencheurs ont été constitués, et l'outil produit des fichiers pré-annotés pour un groupe de déclencheurs bien précis. Les groupes de déclencheurs sont définis automatiquement, en utilisant plusieurs heuristiques:

- on groupe en priorité les déclencheurs ayant la même ambiguïté de frames (ils sont dans le lexique associés au même ensemble de frames)
- on groupe un verbe pronominal et le verbe non pronominal correspondant
- on groupe un verbe et ses nominalisations éventuelles

- pour un déclencheur composé, si un de ses composants est également un déclencheur, il est ajouté au groupe.

Pour un groupe de déclencheurs ainsi automatiquement constitué, le pré-annotateur produit un ou plusieurs fichiers pré-annotés, en limitant la taille à 100 phrases. Pour un déclencheur ayant plus de 100 occurrences, on n'annotera que le premier fichier produit (extension "PRIMO", les autres fichiers auront une extension "EXTRA", à ne pas annoter).

**Important:** lorsque l'on travaille sur un fichier pré-annoté, il faut avoir en tête le groupe de déclencheurs "cibles" de ce fichier (cf. tableur), qui sont ceux automatiquement pré-annotés dans ce fichier. Seuls ces déclencheurs doivent recevoir un frame lors de l'annotation du fichier. Il se peut que des déclencheurs synonymes apparaissent dans les phrases du fichier, mais ne soient pas pré-annotés s'il ne sont pas un des déclencheurs cibles du fichier. Ces synonymes, s'ils sont bien dans le lexique courant, donneront lieu à leurs propres fichiers pré-annotés (→ les occurrences du même frame mais pour d'autres déclencheurs sont proposées dans un autre fichier).

### 2.1.1 Pré-annotation des déclencheurs composés

La pré-annotation des mots composés est conçue pour être tolérante aux incohérences d'annotation de composés dans les treebanks sous-jacents. Ainsi, on pré-annoté systématiquement "en effet" avec le frame Evidence, que la séquence "en"+"effet" soit ou pas annotée comme un composé dans le treebank.

En outre, des heuristiques déclenchent la pré-annotation d'occurrences potentielles de composés présentant des discontinuités : la pré-annotation permet des discontinuités avant certains composants de composés (à , de, où ...)

Cela permet par exemple de capturer le déclencheur "au moment où", dans des occurrences comme:

- (1) ... ce qu' on pourrait attendre du parti au pouvoir en France , **au moment où** la révolution de 1989 ...
- (2) Cela **au moment où** même **où** la " révolution " des pays de l' est ...

Cela occasionne certainement du bruit (→ il faut dans ces cas simplement enlever les frames proposés, cf. section 2.2), sans doute un peu de silence.

### 2.1.2 Frame spécial Other \_sense

On utilise un "frame" spécial: Other \_sense afin de pouvoir faire la différence entre

- les occurrences de déclencheur examinées lors de l'annotation manuelle, mais pour lesquelles il a été jugé que l'occurrence ne correspondait à aucun frame Asfalda → (l'occurrence est alors annotée avec le frame Other \_sense)
- et les occurrences de déclencheurs sur lesquelles aucun annotateur n'est passé (et donc aucun frame annoté)

Les déclencheurs préannotés le sont aussi systématiquement avec ce frame Other \_sense.

### 2.1.3 Préannotation des verbes pronominaux

Le clitique réfléchi *se* peut correspondre à différents processus syntaxiques ou lexicaux, ce qui fait l'objet d'une abondante littérature car les distinctions de statut sont souvent difficiles à faire. On distingue schématiquement:

- le vrai réfléchi (*Jean se parle*)
- le vrai réciproque (*Paul et Marie se parlent*)
- le se-moyen (*Ce genre de livre, ça se vend dans les gares.*), dont la caractéristique est que l'agent n'est pas réalisable localement, mais est forcément interprété comme existant (dans l'exemple, on ne peut pas interpréter la phrase sans faire intervenir un agent qui vend ce genre de livre)
- le se-neutre (*Tout d'un coup, le vase s'est fendu*), ici au contraire l'agent n'est pas forcément existant
- le verbe dit intrinsèquement pronominal, i.e. soit pour lequel une version non pronominale n'existe pas (ex: *se suicider*) ou bien pour lequel on ne repère pas en diachronie de relation régulière à une version non pronominale (ex: *s'apercevoir*)
- les verbes dits coopératifs (*Paul et Virginie se marient*), où la version Xcollectif se V alterne avec "X se V à/avec Y".

Cette catégorisation est importante pour l'annotation, car selon les cas, on considère que le déclencheur est le verbe avec ou sans le clitique *se*.

- dans les cas vrai réfléchi, vrai réciproque, se-moyen, on considère que le *se* marque un changement de diathèse régulier, sans modification du nombre des arguments sémantiques. Le déclencheur à annoter est le verbe sans le *se*.
- Alors que pour les cas se-neutre et intrinsèquement pronominaux, le déclencheur à annoter inclut le clitique *se*. Les frames possibles pour le se+V, recensés dans le lexique, ne sont en général pas les mêmes que les frames du V (voir aussi la section 3.2).

Parce que la distinction entre ces différentes catégories de *se* est souvent difficile à faire, elle a fait l'objet d'une annotation manuelle en amont du projet ASFALDA. La pré-annotation automatique utilise la catégorisation des occurrences de *se* pour préannoter les frames pertinents pour selon les cas le se+V ou le V seul.

## 2.2 Double annotation

Pour un fichier pré-annoté donné, on vise de le faire annoter par 2 personnes indépendamment, puis de faire une adjudication des conflits entre les 2 résultats, via l'outil Salto (Burchardt et al., 2006) qui repère automatiquement les différences d'annotation.

La pré-annotation génère des tableaux recensant des fichiers à annoter, accessibles en ligne. Un tableau est assigné à un ensemble d'annotateurs. On produit les fichiers de façon à traiter:

- d'abord les déclencheurs associés à un seul frame,
- puis les déclencheurs associés à plusieurs frames mais tous du même domaine
- enfin les déclencheurs ambigus entre frames de différents domaines

Les annotateurs choisissent leur fichier à annoter n'ayant pas déjà 2 annotateurs signalés et notent tout de suite la date de début de travail sur ce fichier, dans la colonne prévue à cet effet (pour éviter d'annoter un fichier déjà pris par 2 annotateurs). Un annotateur AAA travaille ensuite sur un fichier XXX à annoter en réalisant des sauvegardes sous le svn, sous

\$ASFALDASVN/annotation/out/AAA/XXX.salto.xml. Lorsqu'il a terminé un fichier, il le signale dans le tableau, dans la colonne prévue.

Pour une occurrence de déclencheur (simple ou composé, et si composé, continu ou discontinu) pré-annotée comme évoquant un ou plusieurs frames, 3 possibilités:

1. Il s'agit d'une erreur de pré-annotation, i.e. il ne s'agit pas d'une occurrence de déclencheur cible
  - il peut s'agir d'une séquence pré-annotée comme déclencheur composé mais ne correspondant pas au composé en contexte (par exemple dans "*les assurances du gouvernement sur l'emploi ne trouvent pas grâce à leurs yeux*", "*grâce à*" sera pré-annoté à tort
  - ou inversement, un composant de composé peut avoir été pré-annoté avec ses propres frames: par exemple *rapport* dans *par rapport à*
  - → dans ces deux cas, retirer tous les frames proposés (y compris `Other_sense`)
2. Il s'agit bien d'une occurrence de ce déclencheur qui déclenche bien un des frames proposés → garder uniquement ce frame et annoter ses rôles
3. Il s'agit bien d'une occurrence du déclencheur mais ce dernier n'évoque aucun des frames proposés → enlever tous les frames proposés sauf `Other_sense`

**Important:** Avant de commencer l'annotation d'un fichier, lisez bien:

- la documentation des frames pré-annotés dans le fichier, et les indications de frontières entre frames
- les éventuels guides de polysémie pour les triggers du fichier: des guides de polysémie ont été écrits pour une bonne partie des déclencheurs ambigus (cf. fichier en ligne Polysémie). Si un déclencheur n'a pas un tel guide, et que son annotation se révèle trop difficile, suspendez l'annotation et provoquez une discussion sur le wiki pour mettre au point un guide de polysémie pour ce déclencheur.

## 2.3 Adjudication

L'adjudication est faite via Salto. Voir le guide technique d'adjudication sous Salto.

## 2.4 Post-traitement automatique

Pour accélérer l'annotation des remplisseurs, en particulier permettre de n'avoir à pointer qu'un seul noeud syntaxique pour un remplisseur, on utilise un post-traitement automatique pour supprimer d'un remplisseur certaines parties. Le traitement n'est pas détaillé ici car il nécessite d'avoir introduit d'autres notions. Reportez vous à la section 7.2.

### 2.4.1 Tête syntaxique

Le post-traitement se charge également de calculer la tête syntaxique des remplisseurs de rôles. On utilise l'annotation syntaxique sous-jacente pour cela. Dans le cas d'un remplisseur formé d'un seul arbre syntaxique, on marque comme tête syntaxique du remplisseur la racine de cet arbre. Dans le cas d'un remplisseur formé de plusieurs sous-arbres syntaxique non connectés, on considère les noeuds racines de ces sous-arbres, la fonction grammaticale qu'ils portent, et leur catégorie. On choisit en priorité l'attribut le plus à gauche s'il existe,

à défaut l'adjectif le plus à gauche, etc... (l'ordre de priorité est attribut, adjectif, verbe, nom, adverbe, pronom).

Ainsi par exemple dans:

(3) Paul **trouve** être parti à 23h

On considère que le rôle Message est rempli par "Paul" + "être parti à 23h" (cf. section 9). Le remplisseur est formé de 2 sous-arbres syntaxiques déconnectés, de tête nom et verbe → On choisit verbe comme tête syntaxique du tout.

Dans:

(4) Paul **trouve** scandaleux que Pierre soit parti.

On considère que le rôle Content est rempli par "scandaleux" + "Pierre soit parti". D'après le schéma d'annotation syntaxique du FTB en dépendances (et si l'annotation syntaxique sous-jacente est correcte), les têtes des sous-arbres syntaxiques sont "scandaleux" et "parti", avec "scandaleux" qui porte la fonction attribut de l'objet, c'est lui qui est choisi comme tête.

#### 2.4.2 Tête sémantique

Par défaut, la tête syntaxique sera interprétable dans les données comme étant la tête sémantique. On a cependant repéré des cas de divergence, avec annotation manuelle de la tête sémantique (section 9.2.1). On obtient après post-traitement que la tête sémantique d'un remplisseur est celle marquée si une telle marque existe, et sinon est la tête syntaxique.

### 2.5 Demande d'évolution du lexique

Le lexique a été construit en général en étudiant les premières occurrences des déclencheurs (utilisation du concordancier sur le corpus), mais il se peut que l'annotation effective révèle des problèmes de lexique (déclencheur associé à tort à un frame, ou bien frame manquant pour un déclencheur donné).

Si pour un déclencheur pré-annoté, vous voudriez annoter avec un frame Asfalda existant, mais qui n'a pas été prévu pour ce déclencheur → provoquez une discussion à ce sujet dans le wiki. S'il s'avère après discussion qu'il s'agit d'un oubli, le frame sera ajouté à ce déclencheur dans le lexique, mais vous devrez manuellement ajouter le frame dans le fichier en cours d'annotation, avec tous les rôles adéquats.

### 2.6 Accès au contexte de la phrase à annoter

Les phrases à annoter dans les fichiers proposés sont regroupées par déclencheur ou groupe de déclencheurs et ne sont donc pas du texte continu.

Si la phrase est insuffisante pour une bonne compréhension, allez voir le texte dans le fichier \$ASFALDASVN/corpus/sequoia+ftb.id\_raw\_articlefrontiers, en utilisant les id de phrases pour retrouver la phrase et son contexte. Le fichier contient des marques de découpage du corpus en articles, ce qui permet de mieux cerner le contexte d'une phrase.



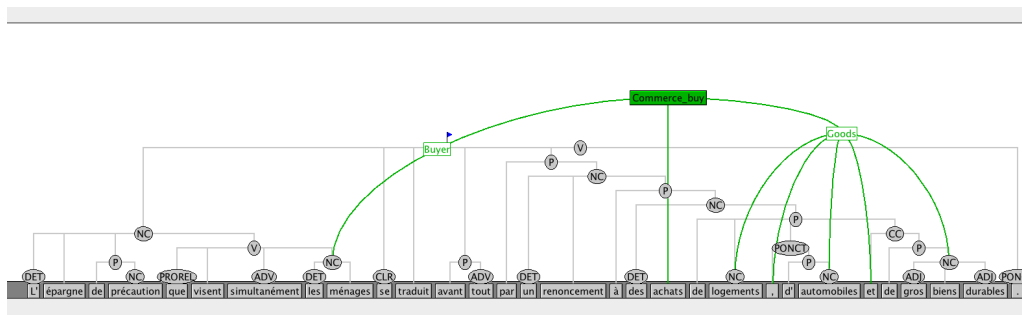


Figure 1:

### 3 Annotation d'un frame: principes généraux

#### 3.1 Dans quel cas annoter une instance de frame?

##### 3.1.1 Annotation indépendante de la factivité

On annote indépendamment de la réalisation effective de l'éventualité évoquée : on annote une instance de frame dès lors que l'idée du frame est évoquée, même dans le cas d'une éventualité hypothétique, future, niée etc... Il faut donc annoter indifféremment selon que l'on a du passé/présent/futur, du conditionnel, ou des prédicats sous négation.

Exemple : dans l'exemple suivant, les achats ne sont pas réalisés (cf. "renoncement") mais l'idée d'achat est bien évoquée (rem: les ménages sont NLI (section 4.4.1); à noter aussi l'annotation d'un remplisseur coordonné avec prépositions régies exclues du remplisseur, cf. section 4.3.2)

##### 3.1.2 Déclencheur avec sens générique / déclencheur dans entité nommée

Pour décider si on annote une instance d'un frame X sur une occurrence de déclencheur Y, il faut se demander si Y dans ce contexte évoque bien l'idée de X.

A noter, qu'il peut s'agir d'une évocation générique: dans l'exemple ci-dessous, on a un acte d'achat générique, mais qui évoque bien l'idée d'un Commerce\_buy.

- (5) Un **achat** est toujours source de stress.

Ce raisonnement amène à annoter des instances de frame y compris au sein d'entités nommées, pour peu qu'elles aient un sens au moins partiellement compositionnel: dans "le ministre du commerce et de l'industrie", "commerce" évoque bien l'idée de Commercial\_scenario.

### 3.2 Alternances syntaxiques

On distingue les changements de diathèse qui ne modifient pas le nombre d'arguments sémantiques du verbe et les autres. En effet, un principe FrameNet fondamental est que pour être regroupés sous le même frame, des déclencheurs doivent avoir le même nombre et type de rôles sémantiques (ou le même nombre de paquets de rôles, lorsqu'il y a des paquets de rôles en exclusion mutuelle). Les alternances qui n'enlèvent ni n'ajoutent d'argument sémantique sont considérées comme des variantes dans la correspondance entre arguments sémantiques et arguments syntaxiques, et annotées avec le même frame.

C'est le cas par exemple pour le passif, alternance la plus fréquente. Il est important de noter que l'on parle ici de l'interprétabilité de l'existence d'un argument, et pas de sa

réalisation effective. On traite donc de la même manière une phrase à l'actif, au passif avec agent et au passif sans agent. Le cas sans agent est traité comme les autres cas d'omission (cf. les flags de "null instantiation" section 4.4).

- (6) [La France]<sub>Seller</sub> a **vendu** [les frégates]<sub>Goods</sub> à [Taiwan]<sub>Buyer</sub>.  
 (7) [Les frégates]<sub>Goods</sub> ont été **vendues** à [Taiwan]<sub>Buyer</sub> (par la [France]<sub>Seller</sub>).

Le clitique réfléchi se a des fonctions diverses, certaines relevant de changements de diathèse. On rappelle qu'à ce stade de l'annotation, la désambiguïsation du se a déjà été faite, et les frames pertinents sont proposés en fonction des cas (cf. section 2.1.3).

Pour le se-moyen, le réfléchi, le réciproque, on utilise les mêmes frames pour la version avec ou sans se. La caractéristique du se-moyen est en effet que l'actant sujet de la version transitive n'est pas réalisable dans la version se-moyen, mais est interprétable. Dans l'exemple ci-dessous, il faut bien interpréter un Cognizer anticipant la spéculation, que l'on marque comme UNI (unknown null instantiation).

- (8) 12969. [La spéculation]<sub>Phenomenon</sub>, il est vrai, ne s' **anticipe** pas. Cognizer=UNI  
 (9) 12969'. \*La spéculation, il est vrai, ne s' **anticipe** pas par les agents.

Pour le réfléchi ou réciproque, on considère schématiquement que le sujet du verbe remplit deux rôles. Le clitique se est considéré comme un marqueur grammatical et ne porte pas de rôle.

- (10) 4688. [Le marché]<sub>Cognizer+Persuader</sub> commençait à se **convaincre** d' [une tendance haussière durable]<sub>Content</sub>.

Les verbes "intrinsèquement pronominaux" (soit que la forme sans se n'existe pas (se suicider), ou ait un sens très différent de la forme pronominal (s'apercevoir de qqch versus apercevoir qqch)) constituent des entrées de lexique se+V distinctes du V, avec des frames en général différents. On a la même situation dans le cas du se-neutre (en effet, la caractéristique du se-neutre est que l'actant sujet de la version transitive n'est pas sémantiquement interprétable dans la version neutre, on a bien un actant en moins et donc un frame différent<sup>3</sup>

D'autres alternances moins productives ne modifient pas le nombre d'arguments (comme l'alternance locative) mais ne sont pas pertinentes pour les domaines notionnels annotés. Le causatif est syntaxiquement représenté dans les treebanks que nous utilisons comme une alternance syntaxique introduisant un actant (l'argument causateur). Le verbe *faire* est traité comme un auxiliaire. Par exemple:

- (11) analyse syntaxique de 1774 modifié: [Les inspecteurs de l'ONU pour le désarmement de l'Irak]<sub>sujet</sub> ont [fait]<sub>aux\_caus</sub> détruire [les stocks d'armement]<sub>objet</sub> [par des casques bleus]<sub>par\_obj</sub>.

Etant donnée la régularité et productivité de la construction causative, il serait aberrant de définir pour tout frame un équivalent frame-causatif, avec un argument causateur supplémentaire. Pour la représentation sémantique, on choisit de traiter le causatif de manière compositionnelle, avec le verbe faire déclenchant le frame Causation. Le rôle Effect est réalisé par l'infinitif introduit par faire et ses dépendants autres que le sujet. Le sujet remplit le rôle Cause ou Actor.

<sup>3</sup>La relation sémantique entre la version transitive et la version neutre est capturée par une relation entre frames, la relation is\_inchoative\_of.

- (12) annotation sémantique de 1774 modifié: [Les inspecteurs de l'ONU pour le désarmement]<sub>Actor</sub> de l'Irak ont **fait** [détruire les stocks d'armement par des casques bleus]<sub>Effect</sub>.

Du fait de la divergence entre analyse syntaxique et sémantique, on peut avoir fréquemment discontinuité du remplisseur du rôle Effect, en particulier avec des clitiques sur le verbe *faire*:

- (13) 2252: [Il]<sub>Actor</sub> [le]<sub>Effect</sub> **fera** [condamner à un franc de dommages et intérêts pour diffamation]<sub>Effect</sub>.

### 3.3 Emplois métaphoriques / figurés

En cas d'emploi métaphorique ou figuré, on annote uniquement la lecture littérale, mais en utilisant un flag "METAPHOR" sur l'instance de frame. Dans l'exemple suivant, il ne s'agit pas à proprement parler d'un achat de biens ou de services, on utilise un flag "METAPHOR" sur l'instance de frame Commerce\_buy.

- (14) l' **achat** d' une clientèle politique majoritairement liée à l' Occident.

Il est cependant parfois délicat de trancher entre emploi métaphorique ou considérer qu'il s'agit d'un autre sens du déclencheur. Pour trancher, on reprend d'une part les critères de (Ruppenhofer et al., 2006, p. 90) pour distinguer polysémie et métaphore : dans une métaphore, le déclencheur doit être substituable par ses synonymes / il doit y avoir correspondance parfaite (bijection) entre les rôles du frame littéral, et les rôles de l'emploi métaphorique.

- (15) Pour eux , l' économie britannique , en renonçant au " grand large " , **acquitte** décidément un bien lourd tribut à l' Europe .

On ajoute le critère qu'à l'inverse, si il y sens distinct, celui-ci doit pouvoir s'utiliser ds d'autres contextes. Dans l'exemple suivant, pour "revendre", on constate qu'il n'est pas substituable par synonyme, donc il n'y pas métaphore. Mais le verbe n'admet pas les constructions syntaxiques classiques (#Elle revend des idées), donc pas de sens distinct pour "revendre" mais plutôt une construction figée (hors champ) "avoir des X à revendre".

- (16) 15050 Cette fille d' immigrés italiens installés à Brighton a des idées à revendre.

### 3.4 Métonymie

Les noms de lieu peuvent classiquement être employés avec un sens correspondant aux personnes habitant dans le lieu ou bien appartenant à l'institution représentant le lieu. On les annote alors comme remplisseur d'autres types sémantiques :

- (17) Au total , [le Japon]<sub>Exporter</sub> a **exporté** , en 1991 , 9 500 véhicules vers l' ex- URSS .

- (18) l'objectif étant ... d' **inciter** [ces pays]<sub>Agent</sub> à passer par la porte insulaire réunionnaise.

Attention cependant, on n'applique cela qu'aux cas où le nom de lieu est dans une position commutant effectivement avec un nom d'animé. On ne considèrera pas qu'il y a métonymie dans le cas de noms de lieu introduits par une préposition locative, même si c'est ce PP qui donne clairement l'indication sur la sémantique d'un remplisseur animé. Dans l'exemple

suyvant, le complément en *chez* introduit bien le rôle Agent du frame Reason (ici l'animé ayant les doutes).

(19) ce réseau **suscitait** plus de doutes que de certitudes chez [les dirigeants d' Elf]<sub>Agent</sub>.

mais dans les deux exemples suivants, le PP souligné donne l'indication de qui remplit le rôle Agent (qui a l'espoir / qui a l'opposition), mais ne commutent pas avec un animé, donc on considère que le rôle Agent est non instancié (ici à annoter comme "extra-linguistic null instantiation", cf. section 4.4).

(20) les espoirs que **suscite** dans les Etats membres une future union économique et monétaire.

(21) Ce projet ... a **suscité** une telle opposition , de part et d' autre de la frontière.

Autre exemple de ce type :

(22) "Dans tel pays, les **importations** de blabla" → rien ne permet de dire que "tel pays" est l'Importer, seulement l'Importing\_area: on annote "tel pays" en Importing\_area mais pas Importer.

### 3.5 Redondance de déclencheurs ou pléonasmes

On a parfois "redondance de déclencheurs". Par exemple dans :

(23) Les mécanismes déflationnistes souvent évoqués ces derniers temps sont -ils les seuls **à l'origine de** ce **résultat** ?

On a deux déclencheurs du frame Causation, référant à la même instance de relation de cause à effet:

- *à l'origine de* déclenche Causation, Cause="les seuls", Effect="ce résultat"
- *résultat* déclenche Causation, avec Effect="ce résultat", cf. usage référentiel de "résultat", cf. section 6.3)

Malgré la redondance, on annote bien l'exemple (23) avec deux instances du frame Causation, non reliées entre elles (le fait de repérer quelles instances de frames sont "coréférentes" dépasse le cadre de notre annotation).

(NB: l'annotation du déclencheur "à l'origine de" et du déclencheur "résultat" ne se fait pas forcément dans le même fichier pré-annoté : il ne faut pas ajouter de frame sur un déclencheur qui n'est pas préannoté).

Autre exemple similaire :

(24) 1699: J' en suis satisfait pour deux **raisons** , tout d' abord **parce que** la Commission relie la Charte des droits fondamentaux à la conférence intergouvernementale.

NB: les deux instances de frame peuvent, selon les déclencheurs, ne pas avoir le même niveau de granularité (mais être plutôt en relation de subsomption). Par exemple:

- "*susciter une réaction*" : susciter → frame Reason + réaction → frame plus spécifique Response
- "*dus à des motifs personnels*" : motif → frame Reason + dû → frame Causation

## 4 Annotation des remplisseurs: principes généraux

### 4.1 Quels sont les rôles à annoter ?

Le lien entre les rôles et la valence syntaxique des déclencheurs n'est pas complètement précisé dans la littérature framenet. Nous prenons le parti de restreindre les rôles définis dans les frames et annotés en corpus: seront définis et annotés uniquement des rôles remplissables par

- des éléments sous-catégorisés par les déclencheurs du frame (cela correspond à la majorité des cas)
- dans le cas de déclencheurs adjectivaux, prépositionnels ou conjonctions, on considère en outre le rôle qui dans l'emploi standard (canonique) du déclencheur, est réalisé par le gouverneur syntaxique du déclencheur. Par exemple pour un adjectif, on peut considérer l'emploi épithète comme emploi canonique, et le nom modifié par l'adjectif est typiquement un remplisseur de rôle. Par exemple le frame FR\_Awareness-certainty-opinion a en gros deux rôles principaux: un Cognizer a dans son modèle cognitif un certain Content. L'adjectif *conscient* peut déclencher ce frame. Il sous-catégorise un complément en *de*, remplissant le rôle Content, et le rôle Cognizer est réalisé typiquement par le nom modifié lorsque l'adjectif est employé comme épithète, comme dans *les partenaires sociaux conscients des sacrifices à accomplir*.
- dans certains cas le déclencheur est lui-même remplisseur de rôle (plus précisément est la tête du constituant remplissant le rôle). C'est typiquement le cas de noms relationnels comme par exemple *conséquence*, qui évoque une relation de cause à effet, représentée par le frame Causation, avec un rôle Effect et un rôle Cause. Dans *Les assurances doivent chiffrer les conséquences du sinistre.*, on considère que *conséquences* déclenche le frame Causation, avec *le sinistre* comme rôle Cause, et le rôle Effect est tout le NP dont la tête est *conséquences*. Voir la section 6 pour plus de détails sur ces cas.

Pour chaque frame a été définie en amont la liste des rôles à annoter (cette liste apparaît en colonne 6 du fichier décrivant la sous-structure asfalda : \$ASFALDASVN/framenet/fndata-asfalda/substructure\_asfalda.txt), et pour chaque rôle s'ils sont Core ou pas:

- Les rôles Core sont sémantiquement obligatoires étant donnée la sémantique du frame et sont soit sous-catégorisés par au moins un déclencheur du frame, soit réalisables comme le gouverneur syntaxique du déclencheur, soit exprimés par le déclencheur lui-même (cf. les 3 cas supra).
- Les autres rôles à annoter sont les rôles sous-catégorisés par au moins un déclencheur du frame, mais pas sémantiquement obligatoires. Par exemple dans *Paul a crié au dentiste "j'ai mal"*, l'interlocuteur est réalisable par un complément sous-catégorisé (un PP en *à*), mais qui n'est pas sémantiquement obligatoire: lorsqu'il est absent comme dans *Paul a crié "j'ai mal"*, c'est soit qu'il est omis mais interprétable en contexte, soit qu'il est inexistant.

Sous Salto, apparaissent par défaut uniquement les rôles Core à annoter. Les autres rôles déclarés comme "à annoter", mais non Core, sont accessibles par clic droit "Add element".

## 4.2 Utilisation de l'annotation en syntaxe

L'annotation des frames et rôles se fait sur des arbres en dépendances syntaxiques, qui ont été annotés manuellement indépendamment du projet ASFALDA.

Vous avez toute liberté pour indiquer le contenu exact d'un remplisseur (i.e. l'ensemble des mots-feuilles constituant le remplisseur). La syntaxe peut permettre d'accélérer les choses, en particulier, dans la majorité des cas, un remplisseur est constitué de tous les mots dominés directement ou indirectement par une tête. Pointer vers le noeud dominant accélère l'annotation.

Attention à bien prendre dans un remplisseur pas seulement la tête, mais également tous les compléments et modificateurs (en particulier les relatives).

### 4.2.1 Gestion des erreurs dans l'annotation syntaxique / divergences syntaxe-sémantique

Il se peut cependant qu'il y ait des erreurs d'annotation syntaxiques, ou plus rarement, qu'il y ait des divergences syntaxe-sémantique telles que l'on ne trouve pas de noeuds dominant exactement tous les mots du remplisseur, et seulement ceux-ci. C'est le cas pour un remplisseur réalisé via deux compléments syntaxiques, dans le cas d'un attribut de l'objet par exemple. C'est le cas également pour un remplisseur discontinu.

Dans ces cas, on peut sous Salto rattacher un rôle à plusieurs noeuds (annotation "multi-pattes, voir la section 4.5.5 également), quitte à rattacher directement au niveau des tokens si on ne trouve pas de noeud syntaxique adéquat.

Attention du coup, cela veut dire que l'on a toujours plusieurs façons d'annoter le même ensemble de mots (pour former un remplisseur) : on peut utiliser les noeuds syntaxiques, les noeuds feuilles tokens et un mélange des deux. On prend la convention d'annoter au plus haut possible : i.e. d'annoter le plus possible en utilisant des noeuds syntaxiques. Mais les différences possibilités sont de toutes façons neutralisées lors du post-traitement automatique des annotations.

( Voir également infra section 7.2: Cas de déclencheurs qui dépendent syntaxiquement d'un de leur arguments sémantiques )

On se sert effectivement de l'analyse syntaxique pour annoter (en particulier en prenant un noeud NC pour annoter tout un NP) MAIS ceci est modulo les erreurs d'analyse syntaxique: si possible il faut annoter en fonction de l'analyse syntaxique correcte (que l'on pense correcte) qui peut diverger de celle qui est proposée.

**NB:** On ne note pas les erreurs repérées d'analyse syntaxique : c'est hors de portée du projet d'annotation sémantique.

## 4.3 Mots fonctionnels introduisant des remplisseurs

### 4.3.1 Prépositions et complémenteurs vides

Contrairement au FrameNet anglais, on choisit de ne pas inclure dans les remplisseurs les prépositions régies (sémantiquement vides) ni les complémenteurs sémantiquement vides:

(25) J'ai **vendu** ma voiture à [Paul]<sub>Buyer</sub>  
→ le Buyer est "Paul" et pas "à Paul"

On exclut le "à" introduisant un A\_OBJ, le "de" introduisant un DE\_OBJ, le "par" introduisant un complément d'agent, le "que" introduisant une complétive. On exclut le "de" introduisant une source (*je suis parti de Paris*), le "à" introduisant une destination (*je vais à Paris*).

**Attention** : ceci vaut aussi pour les contractions *au / aux* de catégorie P+D, pourvu que le à soit un a-obj, et idem pour les P+D *du / des* dans le cas de DE\_OBJ (ou de complément de nom, auquel cas la fonction n'est pas spécifiée). Cela amène malheureusement à des remplisseurs auxquels manque le déterminant:

(26) l' **acquisition** du [yacht]<sub>Goods</sub> par Deviers

Plus généralement, on n'inclut la préposition que si elle ajoute une information qui n'est pas encodée dans le nom du rôle, comme par exemple "à propos de" pour le rôle Topic:

(27) 4334 Une regrettable erreur nous a fait **écrire** à propos de [Wadi , film d' Amos Gitai [...]]<sub>Topic</sub> , que [Haïfa se trouvait dans les territoires occupés]<sub>Message</sub>

**Attention** : on garde le complémenteur "si" à l'intérieur du remplisseur.

(28) 3109 Rien ne permet encore de **savoir** [si l' élimination touchera davantage les sortants que les nouveaux venus]<sub>Content</sub>

versus

(29) 3109 transformé: Rien ne permettait alors de **savoir que** [l' élimination toucherait davantage les sortants que les nouveaux venus]<sub>Content</sub>

#### 4.3.2 Coordination / énumération de prépositions régies

En cas de coordination avec répétition de la préposition régie, on n'inclut dans le remplisseur aucune des prepositions, mais on inclut bien les conjonctions de coordination et virgules (ce qui donne un rôle "multi-pattes"):

(30) 6816 L'acquisition par la firme d' [un courtier et]<sub>Goods</sub> d'[un contrepartiste]<sub>Goods</sub> :

Idem pour le complémenteur que : *Paul dit que blabla et que blibli* → rôle Message = blabla + et + blibli.

Lors de coordinations avec plus de 2 coordonnés, on prend la virgule comme coordonnant. Par exemple dans "*je vends aux grands, aux petits et aux moyens*", le Buyer est "grands , petits et moyens".

Idem également dans le cas d'une énumération, i.e. sans coordonnant : "*je vends aux grands, aux petits, aux moyens*" : on ne prend pas les prépositions régies dans le remplisseur, mais on prend les virgules : "grands , petits , moyens".

#### 4.3.3 Coordination de préposition composée

(31) La balle **au-dessus de1** la table et de2 la chaise  
→ le déclencheur retenu est "au-dessus de1" (frame Locative\_relation), le rôle Ground est constitué de "la table" "et" "la chaise".

#### 4.3.4 Coordination de prépositions qui correspond sémantiquement à une ellipse de prédicat

- (32) J'ai **acheté** [une chemise]<sub>Goods</sub> à [Paul et]<sub>Beneficiary</sub> à [Pierre]<sub>Beneficiary</sub>.  
→ même si on comprend qu'il y a 2 achats (a priori 2 chemises distinctes), étant donné que les 2 actes pointeraient vers le même remplisseur pour le rôle Goods ("une chemise", même si cela correspondrait sémantiquement à 2 instances distinctes de chemises, mais une seule formulation), on n'utilise pas l'annotation "ARG CLUSTER" pour ce cas. On aurait un arg cluster dans le cas où les 2 actes auraient 2 jeux de 2 ou plus remplisseurs distincts :
- (33) J'ai **acheté** une chemise à Paul, et une chemise à Pierre → arg cluster (voir la section 5.1.2).

#### 4.4 Localité des remplisseurs

Pour une occurrence de déclencheur donnée, le cas standard le plus typique est le cas où les rôles, en particulier les rôles Core, sont exprimés dans la même phrase, qui plus est "localement" au déclencheur, i.e. dans le domaine de localité syntaxique du déclencheur (voir infra pour sa définition).

On peut cependant avoir des cas où un rôle n'est pas exprimé localement au déclencheur, soit qu'il n'est pas du tout exprimé dans la phrase courante, soit qu'il est exprimé dans une position non locale. On note ces deux cas par des "flag": soit sur le rôle non exprimé, flag de type "null instantiation" (dans le même esprit que les flags CNI / DNI / INI du FrameNet anglais), soit un flag "NLI" sur un rôle exprimé dans la phrase du déclencheur, mais pas localement à celui-ci.

Techniquement sous Salto un "flag" sur le rôle s'obtient par clic droit sur le rôle (les flags pré-définis apparaissent).

**Attention:** on ne met ces flags que sur les rôle "Core" (ceux qui sont affichés par défaut sous Salto).

**NB:** dans tous les cas, on annote toujours le remplisseur le plus local possible (cette annotation étant éventuellement doublée, dans le cas d'un remplisseur anaphorique, de l'annotation de l'antécédent, si présent dans la phrase, cf. infra section 4.5).

##### 4.4.1 NLI: remplisseurs non-locaux

Le flag à utiliser pour les remplisseurs non locaux est NLI (Non-local instantiation).

Pour un déclencheur **verbal**, les remplisseurs locaux sont les dépendants syntaxiques directs (ou indirects via prépositions) du verbe/du nom. On ajoute aussi le nom modifié en cas de participiale épithète.

Donc dans ces 3 exemples, tous les remplisseurs sont locaux, et donc sans flag:

- (34) Paul a **vendu** une chemise à Anna
- (35) La **vente** d'Adidas par Tapie à Pentland
- (36) les voitures **achetées** par Paul ne durent pas → "les voitures" = rôle Goods, local



Pour un déclencheur adjectival, on considère comme local :

- les dépendants (prépositionnels ou autres) de l'adjectif (*Paul est conscient du problème*)
- le nom modifié en cas d'adjectif épithète (*les gens conscients du problème*)
- le sujet en cas d'adjectif attribut du sujet (*les observateurs sont sceptiques*)

On considère NLI l'objet direct pour tous les cas d'attribut de l'objet:

- attribut de l'objet adjectival (*le plan laisse les observateurs sceptiques*)
- attribut de l'objet prépositionnel introduit par comme (*ils considèrent ce départ comme probable* : "ce départ" NLI par rapport à "probable")
- attribut de l'objet infinitive (*elle entend Paul parler à Pierre* : "Paul" NLI par rapport à "parler")

Pour un déclencheur nominal, on considère comme local les dépendants du nom (sous forme de SP ou d'adjectifs).

On aura un NLI aussi par exemple:

- pour le sujet des infinitifs (constructions avec verbe à contrôle, verbe à montée...)

(37) Jean veut vendre sa moto  
→ le Seller "Jean" est NLI

- pour l'objet des infinitifs dans le cas d'un tough movement

(38) un livre difficile à vendre → le rôle "Goods"=le NP de tête "livre" est NLI par rapport à "vendre")

- dans le cas d'extraction longue distance, et donc potentiellement en cas de relatives, interrogatives, clitique "en", clitique "dont" (à juger selon les occurrences)

(39) le livre [que]<sub>Goods/ANAPH+NLI</sub> Paul veut **vendre**

(40) Le collectionneur [en]<sub>Goods/NLI</sub> a fait récemment

**NB:** un remplisseur avec flag FULLANT sera par définition non local, donc pas besoin de remettre explicitement un flag NLI en plus de FULLANT.

En dehors de ces cas de non-localité syntaxiquement régulières, on a aussi des cas de mentions hors domaine de localité, mais qui ne présentent pas de régularité. Exemples :

(41) 20089 [...] [il]<sub>Interlocuteur<sub>1</sub></sub> n' avait pas **discuté** de la baisse du dollar contre le yen lors de sa rencontre avec [le secrétaire américain au Trésor]<sub>Interlocuteur<sub>2</sub></sub> [...] → le deuxième interlocuteur de la discussion est clairement le deuxième participant de la rencontre.

(42) 21231: La terrible pression de [la grande distribution]<sub>Buyer/NLI</sub> pour **acheter** le plus bas possible asphyxiait [les fabricants]<sub>Seller/NLI</sub> .

NB: lorsqu'un rôle n'est pas exprimé localement au déclencheur, on peut avoir des cas où plusieurs segments sont possibles pour être remplisseur non local (NLI). On choisit dans ce cas le remplisseur le plus local possible, au sens syntaxique. C'est une notion difficile à expliquer précisément (la définition formelle est simple, mais difficile à appliquer manuellement). On doit juger de la longueur approximative du chemin syntaxique entre 2 têtes, et également du type de chemin : on considère toujours comme plus local des dépendances de compléments sous-catégorisés par rapport à des modificateurs. Exemple :

- (43) **le Crédit lyonnais** souscrira directement au capital de la Bf.G. , par ailleurs , un apport fusion à la Bf.G de Crédit lyonnais Deutschland pour 190 millions de deutschemarks ainsi que des **acquisitions** directes de titres à hauteur de 160 millions de deutschemarks permettront à **l' établissement de M. Jean - Yves Haberer** de détenir 26 % du capital de la Bf.G pour environ 3 milliards de francs .  
 → les 2 mentions en orange sont concurrentes pour être le rôle Buyer/NLI du déclencheur "acquisitions"  
 → la seconde mention est plus proche syntaxiquement du déclencheur
- (44) L' **acquisition** d' Adidas , dont **il** possédait déjà 20 , 05 % du capital , permet à **ce brasseur d' affaires sans complexe** d' ajouter ce joyau des stades à son empire familial .  
 → la seconde mention, bien que plus éloignée en termes de nombre de mots, est plus locale syntaxiquement (on préfère remonter à un dépendant du verbe "permettre", dont le déclencheur est le sujet, plutôt que de descendre dans la relative modifiant le déclencheur.)

#### 4.4.2 Imbrication de remplisseurs

Un autre cas de non localité est celui d'un remplisseur apparaissant au sein d'un autre remplisseur. Ces cas sont assez rares, mais surviennent en particulier lorsqu'un des rôles peut être exprimé localement ou bien comme sujet d'un autre rôle. On annote alors une imbrication de remplisseurs.

Exemple:

- (45) 2081 Il serait utile que je fasse quelques brefs commentaires sur les **raisons** de  $[[\text{nos}]_{\text{Agent/NLI}} \text{recommandations}]_{\text{Action}}$  .

Une expression locale alternative de l'agent pour l'exemple ci-dessus serait:

- (46) Il serait utile que je fasse quelques brefs commentaires sur les **raisons** pour  $[\text{nous}]_{\text{Agent}}$  de  $[\text{faire ces recommandations}]_{\text{Action}}$  .

#### 4.4.3 DNI : Definite null instantiation

A utiliser si le remplisseur n'est pas exprimé dans la phrase du déclencheur, mais correspond clairement à un élément mentionné dans les phrases précédentes (contexte linguistique, voir le fichier \$ASFALDASVN/corpus/sequoia+ftb.id\_raw pour avoir accès aux phrases de contexte).

→ le rôle ne pointe sur aucun élément de la phrase, mais porte un flag DNI

#### 4.4.4 ENI : Extra-linguistic null instantiation

A utiliser si le remplisseur n'est pas exprimé dans la phrase du déclencheur, ne correspond pas à un élément mentionné clairement et explicitement dans les phrases du contexte, mais est tout de même interprétable, en utilisant des connaissances extra-linguistiques et/ou le contexte).

→ le rôle ne pointe sur aucun élément de la phrase, mais porte un flag ENI

- (47) 443: D'ores et déjà, il faut envisager la phase de négociation avec les riverains concernant la procédure, qui peut prendre d'ailleurs des formes diverses, pour l'**acquisition** des emprises nécessaires au passage des réseaux d'assainissement.  
→ d'après les phrases précédentes, on comprend que l'acquisition serait faite par la mairie, mais pas de mention précise "la mairie" dans le contexte → ENI
- (48) 3825: D'autre part, la grande conférence internationale qui se tiendra à Washington pour **discuter** de l'aide à apporter aux nouveaux Etats indépendants pourrait avoir lieu vers le 22 janvier.  
→ les interlocuteurs de la discussion seront les personnes qui participeront à la conférence mentionnée dans la phrase. Elles ne sont pas citées donc pas DNI, mais clairement identifiées donc ENI.
- (49) Dans ces conditions, loin de s'améliorer, la situation se dégrade encore plus vite et ceux qui ont **acheté** cherchent à **revendre** sans y arriver.  
→ on comprend qu'il s'agit d'achat / vente de logement (et pas de n'importe quoi), mais pas de mention précise préalable → ENI
- (50) Les constructeurs (automobiles) concentrent leurs **ventes**  
→ les Goods sont des voitures, donc ENI → les Buyer sont ... n'importe qui pouvant acheter une voiture, on ne sait rien de plus, donc UNI

#### 4.4.5 UNI : Unknown null instantiation

A utiliser pour tous les autres cas : i.e. cas où le remplisseur est totalement sous-spécifié, et tout ce que l'on sait c'est qu'il satisfait aux restrictions de sélection imposées par le prédicat et le contexte. Exemples :

- (51) J'étais en train de manger quand Paul a appelé  
→ tout ce que l'on sait c'est que "je" mange qqch de comestible / un repas → UNI
- (52) 3161 les argentins **achètent** déjà des tirelires parce qu'ils ne jetteront plus les pièces de monnaie  
→ le Seller est totalement sous-spécifié (à part qu'il faut que ce soit qqun qui ait des tirelires à vendre) → UNI
- (53) 19201: ...les crédits affectés dans le budget 1990 de [Paris]<sub>Buyer/NLI</sub> aux **acquisitions** [foncières]<sub>Goods/ELLIPTIC</sub> ont subi une coupe ...  
(pour le rôle Goods : voir section 7.3, pour le rôle Buyer : "Paris" est une mention métonymique référant à l'organisation la ville de Paris)

**NB:** les flags UNI, ENI et DNI ne sont utilisables uniquement s'il n'y a pas de remplisseur exprimé dans la phrase. Si on a une expression linguistique pour un remplisseur, pas de flag UNI / ENI / DNI, même si l'expression est vague ou anaphorique. Exemple:

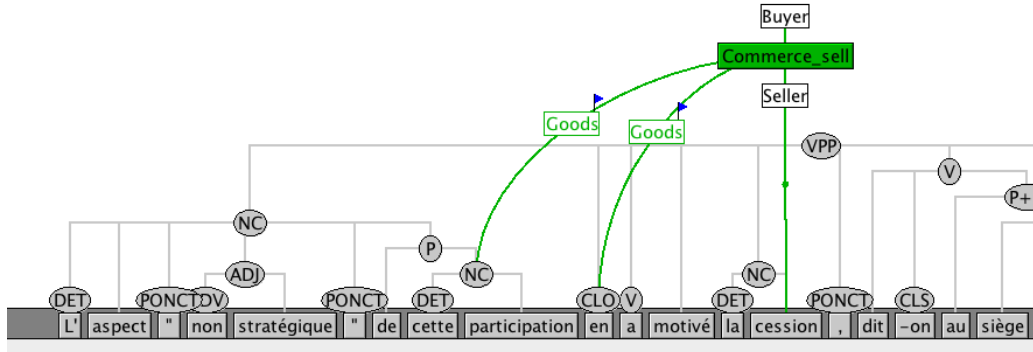


Figure 2: Exemple d'annotation d'un remplisseur anaphorique (en outre non local) et de son antécédent.

- (54) la **cession** d'Adidas à [un tiers]<sub>Buyer</sub> → on ne connaît pas l'identité de ce tiers, mais on ne met pas de flag UNI

#### 4.5 Remplisseurs anaphoriques

Dans le cas d'un remplisseur anaphorique, d'un côté on voudrait annoter l'anaphore comme remplisseur. En effet cela donne des informations relatives à la distribution des fonctions grammaticales pour remplir tel ou tel rôle. D'un autre côté, d'un point de vue sémantique, on voudrait récupérer l'antécédent, pour avoir des informations sur la classe sémantique du remplisseur.

Pour cette raison, on annote à la fois l'élément anaphorique (pronom, clitique, possessif (anaphore du possesseur)) et l'antécédent sémantiquement plein le plus proche (comme fait également dans le projet FrameNet original).

Mais on distingue les 2 remplisseurs en utilisant le flag "FULLANT" pour l'antécédent sémantiquement plein le plus proche, et "ANAPH" pour le remplisseur anaphorique.

**NB:** On n'utilise le flag "ANAPH" que s'il y a annotation d'un "FULLANT".

Exemple de remplisseur de type clitique :

- (55) On pensait que les plus - values compenseraient l'insuffisance du rendement de [ces actifs]<sub>Goods/FULLANT</sub> par rapport aux charges financières assumées pour [les]<sub>Goods/ANAPH</sub> **acheter**.

**NB:** un remplisseur marqué FULLANT est par définition non local, donc inutile d'ajouter le flag NLI. En revanche, un remplisseur marqué ANAPH peut être en outre non local : on doit alors cumuler les 2 flags ANAPH et NLI, comme par exemple le clitique "en" dans l'exemple de la figure 2.

**NB:** on n'utilise le doublage ANAPH / FULLANT que dans le cas d'un élément anaphorique

qui n'est pas un nom (pronom, clitique, déterminant possessif, formulations comme "ce dernier"). On ne l'utilise pas dans le cas de deux NP définis qui réfèrent à la même entité. Ainsi dans l'exemple ci-dessous, il y a deux expressions qui réfèrent à l'acheteur : "le groupe français" et "Thomson" mais la mention la plus locale étant nominale ("le groupe français") et pas pronominale, on l'annote sans annoter par ailleurs Thomson.

- (56) Thomson a calmé ses ardeurs, en effet le **rachat** de LTV par [le groupe français]<sub>Buyer</sub> a été mal perçu.

**NB:** Par ailleurs, on n'utilise le doublage ANAPH / FULLANT que si l'on dispose dans la phrase d'un antécédent comportant une tête lexicale sémantiquement pleine (donc pas un pronom). Dans l'exemple suivant, le remplisseur "qui" a pour antécédent "ce", qui a lui-même pour antécédent "La thrombine active également les facteurs V et VIII" → c'est cette dernière qui est annotée comme FULLANT (le "ce" étant lui-même pronominal).

- (57) [La thrombine active également les facteurs V et VIII]<sub>Cause/FULLANT</sub>, ce [qui]<sub>Cause/ANAPH</sub> **entraîne** une plus grande production de thrombine.

#### 4.5.1 Relatives

Ce traitement s'applique en particulier dans le cas d'un déclencheur inclut dans une relative.

- (58) [les gens]<sub>Buyer/FULLANT</sub> [qui]<sub>Buyer/ANAPH</sub> ont **acheté** du vin.  
 (59) ceux [qui]<sub>Buyer</sub> ont **acheté** du vin → "qui"=Buyer, pas d'antécédent sémantiquement plein, donc pas d'annotation de "ceux" et pas de flag  
 (60) des consommateurs, qui ... veulent vérifier avant d' **acheter** et en avoir pour leur argent → le Buyer le plus local est "qui" (→flag ANAPH) + "des consommateurs" =Buyer/FULLANT  
 (61) [le type]<sub>Buyer/FULLANT</sub> à [qui]<sub>Buyer/ANAPH</sub> Paul a **vendu** sa moto.  
 (62) Et les parents vont soupirer en pensant aux [piles]<sub>Goods/FULLANT</sub> [qu']<sub>Goods/ANAPH</sub> il faudra inlassablement **acheter**.

**NB:** pour le FULLANT d'un pronom relatif, c'est en général le nom modifié par la relative.

**POUR ALLER PLUS VITE**, on note comme remplisseur FULLANT le noeud NC dominant le nom modifié : ce noeud NC inclut lui-même la relative.

- (63) 12824: ... pour lui permettre d' acheter [les francs [qu']<sub>Goods/ANAPH</sub> on lui **vend**]<sub>Goods/FULLANT</sub> libéralement. → le Goods/FULLANT est noté sur le noeud NC dominant "francs", et couvre donc "les francs qu'on lui vend libéralement".

**ATTENTION:** si on a un syntagme relatif (comme "prep+qui") et que la préposition est sémantiquement pleine (et pas régie), alors l'antécédent du pronom relatif ne correspond pas au remplisseur dans son ensemble. Par exemple :

- (64) le champ à travers lequel Paul a **couru** → "à travers lequel" = le rôle Path, "champ" est l'antécédent de "lequel", mais n'est pas le remplisseur entier → dans ce cas on annote uniquement "à travers lequel"=Path, sans flag.

#### 4.5.2 Réfléchis

Le "se" n'est jamais utilisé comme remplisseur anaphorique : même pour les occurrences du clitique *se* qui sont considérées comme des occurrences du "vrai réfléchi" ou "vrai ré-

ciproque" (cf. section 2.1.3, on considère le "se" comme un marqueur a-sémantique, et c'est le sujet du verbe réfléchi ou réciproque qui remplit deux rôles à la fois. Par exemple :

- (65) Et tous se **disaient** que , tout bien réfléchi , une âme , cela n' a pas de prix . → le sujet "tous" remplit à la fois le rôle Speaker et l'Addressee.

### 4.5.3 Possessifs

Les déterminants possessifs sont anaphoriques d'un point de vue sémantique, et donc peuvent constituer la mention la plus locale d'un remplisseur d'un déclencheur nominal. Le possessif correspond à la pronominalisation du complément en "de" d'un nom, et est sémantiquement ambigu quand au rôle joué: dans *l'acquisition de XXX*, XXX peut être l'acheteur ou les biens vendus. Exemples:

- (66) Des intermédiaires ont facilité l' **acquisition** des [frégates de Taiwan]Goods  
 (67) Concernant [les frégates de Taïwan]Goods/FULLANT, il est à noter que des intermédiaires ont facilité [leur]Goods/ANAPH **acquisition**.  
 (68) Concernant [les frégates de Taïwan]Goods/FULLANT, il est à noter que des intermédiaires [en]Goods/ANAPH ont facilité l' **acquisition**.  
 (69) [Des consommateurs]Buyer/FULLANT qui retardent [leur]Buyer/ANAPH **achat** → il s'agit d'un achat fait par les consommateurs. La mention la plus locale est "leur"=Buyer/ANAPH, ensuite "qui", mais on remonte jusqu'au premier antécédent sémantiquement plein : "des consommateurs"=Buyer/FULLANT.

**ATTENTION:** le déterminant doit être un possessif. Dans le cas d'un déterminant démonstratif, c'est tout le NP qui est une description définie anaphorique, faisant référence à un évènement préalablement évoqué dans le contexte. Par exemple :

- (70) La **vente** d'Adidas par Tapie a été rapide. Cette **cession** intervient à un moment critique. → pour le Commerce\_sell évoqué par "cession", on aura des DNI sur les rôles, cf. ils sont remplis dans la phrase précédente, avec le déclencheur "vente".

### 4.5.4 Cas particuliers: "en" quantitatif / démonstratif "celui"

Le pronom "celui" est en général anaphorique avec une classe de référents, mais pas avec le référent précis lui-même. Dans l'exemple suivant, "celui" reprend la classe sémantique "sac" mais ne renvoie pas à l'entité dénotée par le NP "deux sacs". :

- (71) J'avais repéré deux sacs, j'ai **acheté** finalement celui avec une bandoulière.

Dans ce cas, on note quand même un double remplisseur, avec flag ANAPH / FULLANT, mais pour le FULLANT, on pointe sur le ou les tokens de la classe (ici le token "sacs"), et pas sur le NC englobant tout le NP ("deux sacs").

**Important:** La double annotation ANAPH / FULLANT n'est à utiliser que si c'est l'anaphore est la tête du remplisseur. Donc par exemple dans *"Paul a vendu sa voiture"*, on a Goods="sa voiture", et on ne note pas l'antécédent de "sa", qui n'est qu'une partie du remplisseur.

Avec le "en" quantitatif (*Paul en a vendu trois* → Goods="en"+"trois" / *Paul en vend une autre* → Goods="en"+"une autre"), c'est le "en" qui fournit la classe sémantique du remplisseur, via sa relation anaphorique avec un nom du contexte. Dans ce cas, on utilise le

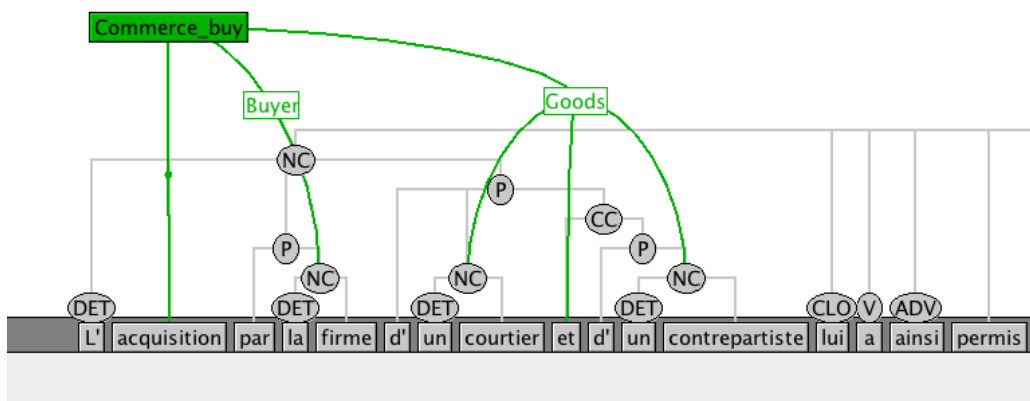


Figure 3: Exemple d'annotation "multi-pattes": une instance de rôle (ici Goods) contient plusieurs "pattes".

même type d'annotation qu'avec "*celui*", le FULLANT est le nom et pas tout le NP. Donc par exemple :

- (72) Une législation peut en **amener** une autre plus restrictive. → Cause="une législation", Effect/ANAPH="en"+"une autre", Effect/FULLANT="législation".

#### 4.5.5 Aspects techniques : 2 instances de remplisseurs, versus un remplisseur "multi-pattes"

La solution ci-dessus dans le cas d'un remplisseur anaphorique, doublé d'un remplisseur "FULLANT", est à distinguer techniquement d'un remplisseur contenant plusieurs noeuds (dit "multi-pattes"):

- Pour le cas ANAPH / FULLANT : après avoir créé le premier remplisseur (par exemple celui avec flag ANAPH, il faut pour créer le second faire un clic droit sur le FRAME, puis Add element : répéter le même rôle. Par exemple dans la figure 2, on a bien deux instances de rôles Goods.
- En revanche, pour un remplisseur constitué de plusieurs noeuds : il faut sous salto annoter le rôle sur un premier noeud, puis faire un clic sur le rectangle donnant le nom du rôle et en maintenant le clic aller jusqu'au prochain noeud à inclure dans le remplisseur. Ce qui donne graphiquement par exemple dans la figure 3 pour le rôle Goods : on a un seul lien Goods, divisé en 3 pattes.

## 5 Points spécifiques aux déclencheurs verbaux

### 5.1 Ellipses de verbes

Certaines coordinations mettent en jeu une ellipse du prédicat, et la partie elliptique peut correspondre sémantiquement à une autre instance d'éventualité. On choisit dans ces cas d'annoter deux instances du même frame, sur la même occurrence de déclencheur, et de marquer par un flag la deuxième instance de frame<sup>4</sup>. On se limite aux 3 cas suivants:

#### 5.1.1 Head Gapping : coordination, où le verbe n'est pas répété (mais 2 sujets différents)

- (73) Exemple construit: Paul **achète** une chemise et Pauline des chaussures  
→ il y a bien 2 actes d'achat, avec 2 jeux de remplisseurs différents.
- (74) 6604: ... dont 6,4 milliards **achetés** par les Français en vacances et 4,2 milliards par les touristes étrangers.  
→ on a 2 actes d'achat : les 6,4 milliards achetés par les français, et les 4,2 milliards achetés par les touristes étrangers  
→ on crée deux instances du même frame Commerce\_buy, sur le même token "achetés", avec chacun les remplisseurs adéquats.

En outre, on marque l'instance correspondant à l'ellipse (la deuxième : l'achat par les touristes) au moyen d'un flag "HEADGAP" sur l'instance de frame.

#### 5.1.2 Argument cluster (i.e. "coordination de séquences de constituants")

Dans ce cas, le deuxième conjoint de la coordination ne contient ni le verbe (non répété), ni le sujet (non répété).

De la même manière, on utilise deux instances de frames, sur la même occurrence de déclencheur, avec un flag ARGCLUSTER sur la deuxième instance.

- (75) Exemple construit: Paul achète une chemise à Marie et des chaussures à Pierre.
- (76) 10047: Steven Ross avait **acheté** une agence de jeunes talents en 1967 et , deux ans plus tard , la société cinématographique Warner - Seven Arts.  
→ une instance de Commerce\_buy avec Buyer="Steven Ross", Goods="une agence de jeunes talents"  
→ plus une autre instance, flag ARGCLUSTER, avec Buyer="Steven Ross", Goods="la société cinématographique Warner - Seven Arts"

**Important:** attention à distinguer arg cluster et remplisseur coordonné. Pour choisir d'annoter via arg cluster et donc avec plusieurs instances de frames, il faut que les deux instances de frame aient au moins un remplisseur différent, pour un des rôles à annoter: si le arg cluster porte sur des éléments qui ne sont pas à annoter (par exemple des modificateurs temporels), alors on ne fait qu'une seule instance de frame :

- (77) Exemple construit: Paul a **acheté** [une fleur]<sub>Goods</sub> le lundi à Paris, et le mardi à Tokyo.  
→ une seule instance de Commerce\_buy

---

<sup>4</sup>flags accessibles via clic droit sur l'instance de frame "Add flags".



Il faut en outre que les 2 remplisseurs distincts ne forment pas une coordination, sauf si on a des modificateurs de temps ou de lieu différents pour les différents coordonnés:

- (78) Paul **achète** [une chemise et un maillot]<sub>Goods</sub>.  
 → la coordination porte sur un seul remplisseur (Goods) → on ne fait qu'une seule instance de Commerce\_buy, avec un remplisseur coordonné.
- (79) Paul **achète** une chemise à [Marie et]<sub>Beneficiary</sub> à [Pierre]<sub>Beneficiary</sub> → idem la coordination porte sur un seul remplisseur → on ne fait qu'une seule instance de Commerce\_buy, avec remplisseur coordonné.
- (80) 10141: ... les **rachats** successifs de la CFAO puis de Conforama et de Le Printemps ....  
 → on ne fait pas trois instances de Commerce\_buy pour le déclencheur "rachats", mais un seul, avec Goods="la CFAO puis Conforama et Le Printemps".
- (81) 16928: - Accor a **acquis**, depuis octobre 1989, non pas 1 %, comme nous l' avons écrit [...], MAIS bien 2 % du capital du club méditerranée.  
 → ici "mais" est le coordonnant du NP coordonné "non pas 1 %, comme nous l' avons écrit [...], MAIS bien 2 % du capital du club méditerranée".  
 → tout ce NP coordonné est le remplisseur Goods, pour le déclencheur "acquis".

Mais si les 2 coordonnées correspondent à des lieux / des temps explicitement différents, alors on fait un arg cluster :

- (82) Exemple construit: Paul a **versé** 3 millions en 1992 et 4 millions en 1993.  
 → arg cluster remplisseur 1 = "3 millions", et remplisseur 2 = "4 millions"

A fortiori, lorsque les 2 remplisseurs ne forment pas de coordination, comme par exemple avec la préposition "contre" :

- (83) 12688: Pour l' heure, l' Etat a **versé**, en 1992, 13,7 milliards de francs de prestations contre 12,1 milliards en 1991.  
 → instance 1 : (en 1992) l'état a versé 13,7 milliards  
 → instance 2 : (en 1991) l'état a versé 12,1 milliards

### 5.1.3 Comparatives réduites

- (84) Exemple construit: Paul **félicite** plus Pierre que Nathan.  
 → ambiguïté à résoudre en contexte entre "Paul félicite plus Pierre que Paul ne félicite Nathan" et "Paul félicite plus Pierre que Nathan ne félicite Pierre".

Dans les 2 cas on fait 2 instances du frame évoqué par "félicite", la deuxième instance avec flag "REDUCEDCOMP".

- (85) 9361: Chipie, [...], reconnaît être plus **acheté** par les filles que les garçons.<sup>5</sup>

On annote pour ce dernier exemple deux instances de Commerce\_buy:

- une avec Goods="Chipie, [...]" / Buyer="les filles"
- l'autre, avec flag "REDUCEDCOMP" avec Goods="Chipie, [...]" / Buyer="les garçons"

<sup>5</sup>On interprète que ce qui est acheté c'est "Chipie, [...]", avec métonymie entre le nom de marque et les produits produits par la marque.

**NB:** la situation est différente lorsque l'on annote un déclencheur présent dans le deuxième terme de la comparative et pas le premier (ds X adv\_comparatif que Y : qd le déclencheur est dans Y). Dans ce cas, le Y a tendance à être elliptique : attention, on n'utilise pas des morceaux du premier terme de la comparative pour annoter le second terme elliptique :

- (86) 5653: ... faire passer la pilule des réformes nécessaires ... que beaucoup d' Italiens ... risquent ... de trouver plus amère que **prévu** → le déclencheur "prévu" est dans le deuxième terme de la comparative → on note le frame Expectation avec rôle Phenomenon = ENI

## 6 Points spécifiques aux déclencheurs nominaux

Voir à ce sujet également la page "focus: déclencheurs nominaux" <https://sites.google.com/site/wikiasfalinks/en-corpus-via-salto/d%C3%A9clencheurs-nominaux>

→ on annote non seulement les occurrences de sens processus, mais également les occurrences de sens correspondant à un participant du processus et/ou au résultat du processus

### 6.1 Nominalisation référant à un évènement versus nominalisation référant à un résultat de l'évènement

Certains noms correspondant à des nominalisations de verbes sont ambigus, dans la mesure où ils peuvent référer à l'éventualité décrite par le verbe ou bien référer à un "participant" de cette éventualité, et/ou au résultat de l'éventualité. Par exemple pour le nom "achat":

- (87) l'achat a duré 3 heures / l'achat a été annulé → "achat" réfère bien au processus d'achat, qui serait évocable avec le verbe "acheter" → achat déclenche le frame Commerce\_buy
- (88) j'ai montré mes achats à Vanessa  
→ ≠ processus, on parle des objets achetés → les objets achetés sont considérés comme un "participant" de l'évènement d'achat, ce qui se traduit par le fait que le frame Commerce\_buy a un rôle Goods pour y référer.

Pour le nom "reproduction", on peut avoir le sens évènementiel "reproduire une oeuvre" ou bien le sens du résultat de l'évènement de reproduction :

- (89) La reproduction du tableau a été effectuée avec soin, par un faussaire professionnel
- (90) La reproduction du tableau a été placée à la place habituelle du tableau.  
→ "reproduction" réfère ici à l'objet physique, qui est le résultat d'un processus de reproduction.

Dans toute la suite on parle des occurrences comme (88) et (90) comme les occurrences REFTOPARTICIPANT (on ne distingue pas participant ou résultat). Dans un emploi de type REFTOPARTICIPANT, est tout de même évoquée l'éventualité sous-jacente, même si le nom réfère à un participant ou résultat, et pas à l'éventualité.

Pour cette raison, on choisit d'annoter ces éventualités sous-jacentes, s'il s'agit d'éventualités évoquant des frames de nos domaines. Par définition, pour ce type d'emploi, un des rôles du frame sera rempli par le constituant dont le déclencheur nominal est la tête.

Donc par exemple pour (88) : "achat" déclenche Commerce\_buy, et on note le rôle Goods = "mes achats" (noeud NC dominant "achat"). On se retrouve alors avec un remplisseur qui contient le déclencheur.

Rem: en théorie, dans certains cas de référence au résultat d'une éventualité (comme (90) ci-dessus), parfois le résultat ne correspond pas à un participant : la sous-cat du verbe "reproduire" ne fait pas intervenir le résultat de la reproduction, mais uniquement le reproducteur et l'objet reproduit. Dans ces cas-là, il manquerait un rôle à annoter. MAIS pour l'instant, pour les frames qui nous occupent, on a toujours trouvé un rôle déjà prévu, auquel réfère la nominalisation de type REFTOPARTICIPANT.

### 6.1.1 Critères de distinction entre référence à une éventualité versus référence à un participant ou résultat

Reste à décider en contexte si une nominalisation évoque l'éventualité ou bien un participant/résultat de l'éventualité.

**Comme principe général, on pose que lorsque les 2 lectures sont possibles, on privilégie la lecture événementielle, et donc avec annotation sans remplisseur qui contienne le déclencheur.** Par exemple :

- (91) 412: Du point de vue financier , l' association fonctionne essentiellement grâce à l' aide d' une subvention attribuée par le siège basé à Valenciennes ainsi qu' une **cotisation** de 17 euros par an et par adhérents . → pour **cotisation**, les 2 lectures semblent possibles (la somme d'argent, ou bien l'action de cotiser, par exemple, on peut forcer cette deuxième lecture avec "une cotisation par les adhérents de 17 euros") → on privilégie en cas d'ambiguïté, la lecture événementielle.
- (92) 14647: Une Europe dont les **importations** en provenance de pays en voie de développement ( hors OCDE pour simplifier ) augmentent inexorablement : 9 , 2 % du total en 1984 , 13 , 7 % en 1992 .  
→ d'un côté "en provenance de pays" fait plutôt opter pour importations=les biens importés  
→ d'un autre côté c'est plus les actes d'importation qui augmentent que les biens importés (#les biens augmentent ne veut pas dire grand chose) (cf. exemple supra avec "réduire")  
→ on privilégie l'annotation processus

**Indice: Indication de durée :** Une consigne générale est que lorsque l'on a / on peut modifier le nom avec une durée, ou bien que l'on a une information de durée, alors il s'agit forcément de l'évènement:

- (93) 4786 Enfin , le CERC rappelle qu' un allongement de la durée de **cotisation** [...] avantagerait relativement les non- cadres , dont la durée de cotisation est actuellement de quarante ans .  
→ COTISATION interprété comme l'évènement (cf. durée de cotisation).

**Indice: coordination:** Un autre indice est la coordination du nom. En effet, on ne coordonne en général que des noms de même type sémantique, donc par exemple :

- "achats et autres babioles" → "achats" réfère aux biens achetés (→ REFTOPARTICIPANT)

- "achats et autres activités" → "achats" réfère au processus d'achat (→ pas REFTOPARTICIPANT)

- (94) 6320 ... une nouvelle convention collective qui prévoit notamment des garanties de reclassification et de **rémunération**  
 → comme "reclassification" n'est pas ambigu (réfère forcément au processus), il permet de désambigüiser "rémunération"

**Indice: Verbe support d'action ou d'évènement** : si le N est un complément des verbes comme "effectuer", "faire", "commettre", "accomplir", "mener à bien", "mener à son terme", "procéder à" alors il s'agit d'un évènement. Donc par exemple :

- (95) 3555 Près\_de 50\_000 **importations** de ce type auraient été effectuées en 1990.  
 → importations = forcément un évnt, donc ça ne peut référer qu'à l'acte d'importer et pas aux biens importés

**Indice: restrictions de sélection du verbe ou nom enchâssant** : si le N est un complément d'un verbe ou nom posant des restrictions de sélection sur ce complément (i.e. des contraintes sur le type sémantique du complément) alors la restriction de sélection peut parfois désambigüiser. En effet, le sens "processus" et le sens "participant" n'ont en général pas le même type sémantique, cf. en général le sens "participant" réfère à un objet (par exemple les biens achetés dans le cas de "achat"). Idem si le N est complément d'un N posant lui-même des restrictions de sélection

- (96) 15209: ... le règlement des **achats** par les détaillants slovaques constitue la principale angoisse de Pavel Maly .  
 → on ne peut pas régler un acte d'achat / on règle une personne, un bien ou une somme → donc dans cet exemple "achats" réfère aux Goods.
- (97) 16223 Le groupe de Pennsylvanie affirme payer 154 millions de dollars ( 831 millions de francs ) sa nouvelle **acquisition** , prix d' achat et reprise des dettes compris .  
 → on paie un bien un certain prix, mais on ne peut pas "payer un acte d'achat un certain prix" → donc ici frame sur "acquisition" avec REFTOP (et frame également sur "achat")

**Indice: modification du nom par un modifieur référant au processus**

- (98) M. Dumas a rappelé la **décision** " rapide , claire , ferme " , prise par les\_Onze , la semaine dernière , à Oslo : pas de renégociation du Traité .  
 → même si la partie après les ":" réfère au rôle Decision, ici la modification par "rapide" du nom "décision" indique qu'il s'agit du processus (c'est le processus de décision qui est rapide, ça ne peut pas être l'absence de renégociation du traité qui est rapide)  
 → pas REFTOPARTICIPANT

**Indice: modification du nom par un modifieur référant à l'agent du processus**

- (99) le gouvernement avait pris , fin 1991 , une **décision** courageuse  
 → la modification par "courageuse" est un indice qu'il s'agit du fait de décider : "courageuse" s'interprète comme "en décidant le décideur a fait preuve de courage, on privilégie dans ce cas la lecture événementielle → pas REFTOPARTICIPANT

En l'absence de ce type d'indices (qui sont rares...), il faut se fier à son interprétation, en privilégiant l'annotation sans REFTOPARTICIPANT en cas d'hésitation. Ci-dessous

quelques exemples :

- (100) 9386 De plus , on constate à l' heure actuelle un changement de comportement des consommateurs , qui , d' abord , réduisent leurs **achats** , et ...  
→ on ne peut pas réduire les biens (\*les consommateurs réduisent leurs sacs) mais on réduit (un peu mieux) l'acte d'achat (les consommateurs réduisent leurs actes d'achat) → frame Commerce\_buy
- (101) Depuis dix - huit mois , il dirige la production et le personnel de Lafarge Zemment , la dernière **acquisition** du groupe Lafarge Coppée .  
→ L'apposition indique que Lafarge Zemment=acquisition , donc il s'agit forcément de l'objet acquis, pas du processus.

Dans le cas d'un nom évènementiel avec déterminant démonstratif, on a souvent affaire à une coréférence évènementielle (et donc pas à un cas REFTOPARTICIPANT), sans que ce soit systématique.

- (102) 3795 Cette **acquisition** permet surtout à Grand Metropolitan de se doter d' un réseau commercial en Amérique latine ...  
→, "cette acquisition" réfère à un acte d'achat évoqué dans le contexte précédent.
- (103) 11423 Ebauches SA , dont le siège est à Besançon , s' est engagé à reprendre 273 des 422 employés d' ici à 1996 et à investir 30 millions de deutschemarks ( 120 millions de francs ) dans sa nouvelle **acquisition** .  
→ Ebauches SA n'investit des millions dans un acte d'achat, mais investit des millions dans l'entreprise achetée → pas évènementiel → REFTOP
- (104) Carryfast est la seizième **acquisition** d' UPS en Europe depuis 1985 .  
→ la construction avec copule + attribut du sujet indique que l'on a coréférence entre "Carryfast" et "la seizième acquisition ..." → donc on n'a pas ici stricto sensu le processus → MAIS 27/05/2015 : On a ici une interférence avec l'emploi prédicatif : ici effectivement il s'agit du "acquisition" qui s'il était *en emploi référentiel* référerait aux biens achetés, mais il est ici en emploi prédicatif (cf. la section suivante 6.2). Si on adopte le même traitement que celui pour "cause" ou "conséquence" en emploi prédicatif, on doit considérer qu'*acquisition* ne réfère pas au bien acheté dans cet exemple!!! on a donc plusieurs manières pour "acquisition" de ne pas référer aux biens achetés : soit parce qu'il réfère à l'évènement, soit parce qu'il est en emploi prédicatif.

## 6.2 Noms relationnels et frames représentant une relation

On considère dans cette section qu'un nom est **prédicatif** s'il a une structure argumentale, i.e. au moins un argument syntaxique, correspondant à un argument sémantique du nom.

On a vu donc dans la section précédente des noms prédicatifs ambigus car pouvant référer à une éventualité ou bien à un participant ou au résultat d'une éventualité.

Par exemple "achat" réfère soit au processus d'acheter, ou bien aux biens effectivement achetés. Les deux sens sont annotés avec le même frame, mais avec ou sans annotation de type REFTOPARTICIPANT.

Nous abordons ici un autre caractéristique de certains noms prédicatifs, dits noms relationnels: leur sémantique met en relation deux entités (entités concrètes, ou bien elles-mêmes évènementielles) et est telle que lorsque le nom est employé comme tête d'un NP saturé,

i.e. avec tous ses compléments, et en emploi "référentiel", le NP réfère à un des arguments de la relation. (On précisera dans la prochaine section (section 6.3) ce que l'on entend par emploi référentiel par opposition à emploi prédicatif.)

Typiquement le nom *père* évoque la relation de "paternité" entre deux individus, mais un NP saturé de tête "père" réfère à un des 2 individus: dans *J'ai rencontré le père de Paul*, les deux individus en question sont *Paul* et *le père de Paul*, on a bien que le NP réfère à un des 2 individus mis en relation.

Par contraste, le nom *similitude* évoque lui-aussi une relation entre deux (ou plus) d'individus, mais en emploi référentiel, un NP de tête *similitude* réfère à la relation elle-même et pas à une des entités similaires.

- (105) la similitude entre Paul et Pierre est troublante  
→ le NP de tête similitude réfère à la relation et ni à Paul ni à Pierre.

Les noms d'évènements sont d'autres exemples typiques de noms prédicatifs qui ne sont pas relationnels. Par exemple le nom *cambriolage*, est prédicatif (cambriolage de X par Y), mais non relationnel cf. en emploi référentiel *cambriolage* réfère à un évènement, et ni au cambrioleur ni au cambriolé.

De fait, la modélisation en frames des noms relationnels pourrait suivre la syntaxe: i.e. considérer un frame avec un seul rôle pour "le-père-de-X", mais il est plus naturel de considérer un frame à deux rôles (pour "relation de paternité entre X et Y"). On a alors la particularité d'un frame évoqué par un nom, qui employé référentiellement réfère à un des rôles du frame.

Dans les domaines annotés dans ASFALDA, on a rencontré des noms relationnels en particulier dans le domaine de la causalité.

Par exemple le nom *cause* est relationnel en cela qu'il évoque une relation de cause à effet entre 2 entités (une cause et un effet), et réfère, en emploi référentiel, à l'entité jouant le rôle de la cause. On considère *cause* comme un déclencheur du frame à 2 rôles (Cause et Effect) Causation, et le NP de tête *cause* est le remplisseur du rôle Cause. Le complément en *de* du nom remplit le rôle Effect.<sup>6</sup> Exemple:

- (106) 330: Les policiers ont ouvert une enquête afin de déterminer les [ **causes** exactes de [cet incendie]<sub>Effect</sub> ]<sub>Cause</sub>.

De la même manière, le nom *conséquence* est modélisable avec le même frame Causation, mais réfère (en emploi référentiel) au rôle Effect:

- (107) 10220: Refusant de " faire supporter aux travailleurs les [ **conséquences** de [la politique de la direction]<sub>Cause</sub> ]<sub>Effect</sub> " ...

Il s'agit donc d'une situation proche du cas vu section précédente, i.e. de nominalisations qui référant à un des participants, sauf qu'ici on n'a pas polysémie (i.e. il y a polysémie entre achat1="l'acte d'achat" et achat2="les biens achetés", alors que l'on n'a pas cette polysémie pour "cause" ou "conséquence"). Aussi par cohérence, on utilise aussi une annotation de type REFTOPARTICIPANT pour les noms relationnels en emploi référentiel :

<sup>6</sup>Une alternative suivant la syntaxe aurait été de définir un frame "cause\_of\_situation\_X", à un seul rôle pour X.

ils déclenchent le frame, et sont tête d'un des remplisseurs.

### 6.3 Usage prédicatif versus usage référentiel d'un déclencheur nominal

Dans cette section on aborde une distinction typique dans l'interprétation d'un nom, qui complique le traitement des noms relationnels vu à la section 6.2.

- **USAGE PREDICATIF** : soit le nom est utilisé uniquement pour prédiquer une propriété sur un autre élément de la phrase
  - par exemple dans "*L'assassin a utilisé un outil contondant. Cet outil était peut-être un COUTEAU.*" on considère que "un couteau" ne réfère pas à un objet précis, mais indique des propriétés sur le référent de "Cet outil".
- **USAGE REFERENTIEL** : ou bien, le plus fréquemment, le nom réfère directement à une entité (et l'usage de ce nom indique que cette entité en a les propriétés)  
Par exemple dans "*L'assassin a utilisé un COUTEAU*" → "un couteau" sert à référer à un objet (en plus de dire que cet objet a les propriétés de couteau).

L'usage standard du nom est l'usage référentiel, mais l'usage prédicatif est relativement fréquent pour les noms relationnels.

**ANNOTATION des usages REFERENTIELS**: On considère 2 cas:

- si le déclencheur nominal (cas le plus fréquent) évoque un frame mais ne réfère pas à un des rôles du frame (comme par ex. *cambriolage* ou *similitude*), l'annotation n'a rien de spécial
- mais dans le cas d'un nom référant (en emploi référentiel) à un participant (par exemple un nom relationnel comme *cause* ou bien un des sens de *achat*), alors en emploi référentiel, le déclencheur en question est à la fois déclencheur du frame, et remplisseur d'un certain rôle du frame (plus précisément le déclencheur est la tête syntaxique d'un remplisseur du frame). On dit que le rôle est référé par le déclencheur. C'est une propriété du déclencheur que de déterminer à quel rôle il réfère dans le cas d'un usage référentiel (cf. les exemples (107) et (106) vu à la section précédente 6.2.

**ANNOTATION des usages PREDICATIFS**: le déclencheur en question est simplement déclencheur du frame. Le rôle qui serait référé dans le cas d'un emploi référentiel peut dans le cas prédicatif soit être présent dans la phrase (et annoté comme tel), soit omis (et il doit alors recevoir un des flags de NULL INSTANTIATION, voir infra). Exemple :

(108) [Cette fusion]<sub>Effect</sub> est une **conséquence** directe de [la fin de la guerre froide]<sub>Cause</sub>

On liste ci-dessous des exemples de cas typiques d'emplois prédicatifs, et quelques exemples d'emplois référentiels. Si vous trouvez un cas qui ne ressemble à aucun de ces exemples, signalez le via un commentaire wiki.

#### 6.3.1 Liste non exhaustive des types d'emplois prédicatifs (où le nom déclencheur n'est pas en outre remplisseur)

**Phrase copulative avec attribut nominal** : NP1 + verbe être + NP-déclencheur: (le verbe être peut avoir toute sorte d'auxiliaire ou de modal)

(109) [Cette situation]<sub>Effect</sub> est le **résultat** de [plusieurs phénomènes qui se télescopent]<sub>Cause</sub>

NB: il faut que le déclencheur soit directement tête de l'attribut. Dans le cas contraire, on considère que l'emploi est référentiel. Exemple:

(110) Cette situation est l'une des **raisons** de la conduite de la Bundesbank  
→ on considère le NP "(les) raisons de la conduite de la bundesbank" comme référentiel.

**Phrase copulative "inversée"** : dans les cas "NP1 + verbe être + NP2", c'est parfois en fait le NP1 qui prédique une propriété sur le NP2 référentiel (on parle de phrase copulative spécifique). Le test pour repérer lequel des 2 NPs est référentiel est l'utilisation de la clivée. Mais pour simplifier ici, au vu des exemples, si le NP1 contient un nom du domaine causalité → on considère qu'il s'agit forcément d'un emploi prédicatif:

(111) Le **résultat** est [un taux d'épargne élevé, mais qui ne s'accroît plus]<sub>Effect</sub>  
(Cause=DNI)

(112) Le **résultat** est qu'[un décalage sensible se creuse sur le marché]<sub>Effect</sub> (Cause=DNI)

**ATTENTION**: on ne fait cette analyse que pour une phrase copulative (verbe être), dans d'autres cas de type NP1 verbe NP2, même si l'interprétation revient à identifier que NP1 et NP2 réfèrent à la même entité, on va considérer les deux NPs comme référentiels. Exemple:

(113) Le principal **motif** d'inquiétude vient du Japon, actuellement le premier consommateur mondial de métal blanc ...  
→ même si la phrase peut s'interpréter comme "le principal motif d'inquiétude EST le Japon...", qui serait une phrase copulative inversée, on colle à ce qui a effectivement été écrit, et on considère que l'emploi est référentiel → il est le remplisseur du rôle State\_of\_affairs de Reason (la cause), et on ne capture pas que c'est coréférent à "Japon".

**Apposition sans déterminant** : le nom déclencheur peut apparaître comme apposition sans déterminant, au sein d'un NP :

(114) Que ce soit l'impossibilité de loger les plus démunis ou, à l'autre bout de la chaîne, [l'importance des encours bancaires dans l'immobilier]<sub>Effect</sub>, **résultat** d'[imprudences et]<sub>Cause</sub> de [spéculations]<sub>Cause</sub>, rien ne va plus.

**Phrase averbale (ou construction de type "connecteur")** : nom déclencheur sans déterminant suivi de ":" ou bien ", "

Dans les exemples suivants, c'est toute la phrase qui suit le ":" ou la ", " qui est le remplisseur:

(115) **Résultat**: [leurs dirigeants forçaient les doses pour pousser les rendements]<sub>Effect</sub>

(116) **Résultat**, [faute de pluies suffisantes pour les renouveler, leur niveau baisse]<sub>Effect</sub>

(117) **A l'origine de** [ce désarroi]<sub>Effect</sub> : [les premiers effets de l'application de la nouvelle PAC]<sub>Cause</sub>



- (118) **Conséquence** : [les services de Mme Véronique Neiertz ... ont décidé de renouveler l'expérience à intervalles réguliers]<sub>Effect</sub>

#### Construction de type "avoir pour N":

- (119) [Cette solution]<sub>Cause</sub> devrait avoir pour **conséquence** probable de [stopper le harcèlement judiciaire]<sub>Effect</sub>
- (120) [les rumeurs]<sub>Effect</sub> avaient pour **origine** [des propos au Nouvel Observateur]<sub>Cause</sub>

#### Cas figé (mot composé de type préposition / adverbe / conjonction):

Le N déclencheur peut former un composé (comme à *l'origine de*, *en conséquence*) qui n'est pas un nom. On a alors forcément le cas standard où le déclencheur (ici composé) n'est pas en plus remplisseur:

- (121) 5704 **À l'origine de** [cette glissade du dollar]<sub>Effect</sub>, [...] on trouve [deux explications]<sub>Cause/NLI</sub>.

#### 6.3.2 Exemples d'emplois référentiels

En dehors des cas prédicatifs cités supra, on considère que l'emploi est référentiel, et donc si le nom est relationnel, on annote avec un des rôles rempli par le déclencheur lui-même (plus précisément le constituant dont il est tête). Exemples:

- (122) 17761 [Cet abandon de compétence]<sub>Cause/NLI</sub> aurait des [**conséquences** très importantes]<sub>Effect</sub>
- (123) Les ventes françaises [en]<sub>Cause/NLI</sub> subiront [les **conséquences**]<sub>Effect</sub>
- (124) Et le [**résultat**]<sub>Effect</sub> a été conforme à leur attente : c' est devenu un frein à la mobilité .  
 → attention, par cohérence le remplisseur du rôle Effect est bien "le résultat" ici, et on ne capture pas le lien d'identité avec la phrase derrière les ":".

**Attention** : dans le cas d'une phrase copulative avec attribut du sujet de type adjectival, ou bien être + locatif, le sujet est forcément référentiel (et donc annoté avec incorporation):

- (125) 19466 [Les **conséquences** génétiques de [la technique du transfert d'embryon]<sub>Cause</sub>]<sub>Effect</sub> sont potentiellement immenses.

#### 6.4 Noms quantificateurs

Il existe des NP dont la tête syntaxique est un N qui correspond sémantiquement à une quantification, ne donnant pas le type sémantique du référent du NP global. Dans ces cas on annote quand même comme remplisseur tout le NP, incluant le N quantificateur (que l'accord soit au singulier ou au pluriel) :

- (126) [(une partie / la majorité / la plupart ...) des 300 manteaux produits]<sub>Goods</sub> (a / ont) été **achetée** par ...
- (127) [La bagatelle de 52 , 4 milliards de francs]<sub>Money</sub> ont été **versés** au titre de l' une des trois aides.

## 6.5 Constructions à verbe support

Il s'agit de constructions où un nom prédicatif est employé comme dépendant d'un verbe, et d'une part le verbe n'est pas employé dans un de ses sens habituels (comme "prêter" dans "prêter attention") et/ou c'est un verbe à sens très général (faire, avoir...). D'autre part le sujet du verbe est forcément un dépendant du nom prédicatif.

On a envisagé d'annoter les verbes support des noms prédicatifs, mais finalement dans un premier temps, on se concentre sur l'annotation des arguments du nom prédicatif (comme remplisseurs) sans noter le verbe support.

Une seule exception: lors de l'annotation des verbes qui sont déclencheurs de frames de nos domaines, si certaines occurrences sont hors-domaine mais correspondent clairement à un verbe support, on utilise le frame FR\_Support (au lieu de Other\_sense).

## 7 Annotation des remplisseurs: points divers

### 7.1 Remplisseurs collectifs versus rôle \_1 / rôle \_2

Pour les déclencheurs admettant l'alternance entre 2 rôles réalisés comme 2 constituants différents versus les 2 rôles réalisés comme un seul NP collectif, on reprend la solution du FrameNet anglais, où l'on dispose dans le frame de deux rôles distincts XXX\_1 et XXX\_2 et d'un rôle pour XXXs pour le collectif. Par exemple, dans Be\_in\_agreement\_on\_assessment, on a les rôles Cognizer\_1, Cognizer\_2 et Cognizers:

(128) réalisation comme NP collectif : [Les enfants]<sub>Cognizers</sub> sont d'accord (entre eux)

(129) réalisation comme 2 NP : [Paul]<sub>Cognizer\_1</sub> est d'accord avec [Marie ]<sub>Cognizer\_2</sub>

(130) [Les enfants]<sub>Cognizer\_1</sub> sont d'accord avec [le professeur\_2]<sub>Cognizer</sub>.

(131) réalisation avec un seul NP, l'autre étant omis car présent dans le contexte: Le professeur propose de faire une pause. Les [Les enfants]<sub>Cognizer\_1</sub> sont d'accord.

Pour choisir qui de X ou Y est XXX\_1 : en général, le rôle XXX\_1 est le rôle le plus saillant syntaxiquement (en particulier le sujet).

Attention, dans le cas d'un seul NP coordonné, on ne descend pas au sein de la coordination pour trouver les rôles XXX\_1 et XXX\_2, on utilise le rôle XXXs, cf. on peut avoir des coordinations n-aires (si compatible avec l'interprétation):

(132) 13267: Pendant la décennie 80 , le **commerce** entre [la colonie et Pékin]<sub>Participants</sub> a été multiplié par 8 ...

### 7.2 Astuce dans le cas de déclencheurs qui dépendent syntaxiquement d'un de leur arguments sémantiques

Dans le cas de déclencheurs qui dépendent syntaxiquement d'un de leurs remplisseurs, on est confronté au problème que ce remplisseur ne constitue pas un syntagme à lui tout seul. On utilise une astuce pour accélérer l'annotation, corrigée par post-traitement automatique.

#### 7.2.1 Adjectifs et participes épithètes

Un adjectif épithète ou un participe épithète dépend syntaxiquement d'un élément (en général un nom) qui correspond à un de leurs arguments sémantiques, et donc en général à un de leurs remplisseurs:

- (133) la **jolie** balle → *balle* est argument sémantique de *jolie*
- (134) l'appartement **acheté** par Deviers → *l'appartement* est argument sémantique de *acheté*
- (135) les personnes **voulant** annoter des corpus → *les personnes* argument sémantique de *voulant*

On a la même situation typiquement pour une préposition ou une conjonction :

- (136) il a fait cela **avant** la conférence → "*il a fait cela*" = le Trajector du frame FR\_Temporal\_relation déclenché par "avant"
- (137) il a fait cela **parce que** Paul le lui a demandé → "*il a fait cela*" remplit le rôle Effect du frame Causation déclenché par *parce que*

Mais en syntaxe, la préposition ou conjonction modifie le verbe principal, et la principale sans ce modifieur ne forme pas un constituant (*il a fait cela* ne forme pas un constituant. Ou bien dit en termes de dépendances: il n'y a pas de noeud qui domine le remplisseur effectif (i.e. la principale sans le modifieur introduit par la prep ou la conjonction).

Par exemple il n'y a pas de noeud pour "*il a fait cela*" → le noeud dominant la tête "fait" inclut le modifieur "avant la conférence" / "parce que Paul le lui a demandé".

→ Pour accélérer l'annotation manuelle, on code comme remplisseur le noeud dominant le nom modifié / le verbe modifié. I.e. dans les exemples ci-dessus : le noeud NC pour "balle" / appartement / personnes, ou le noeud VPP pour "fait".

**Post-traitement:** On appliquera un post-traitement automatique corrigeant cela: si le déclencheur (plus tout le sous-arbre du déclencheur) est contenu dans un remplisseur, sans en être la tête syntaxique<sup>7</sup>, alors on interprétera en réalité le remplisseur comme étant celui annoté moins le déclencheur et son sous-arbre.

### 7.2.2 Incises

On utilise également ce raccourci dans le cas de déclencheurs verbaux qui sont tête d'une incise :

- (138) 14023: Si un partenariat avec l' Afrique est nécessaire , **croient** les auteurs , c' est qu' elle est " la terre d' élection d' amitiés renforcées .

→ "croient les auteurs" est une incise, et "croient" est rattaché au verbe principal (ici le verbe "est"). Le verbe "croient" déclenche FR\_Awareness-certainty-opinion et le rôle Content précis est toute la phrase sauf "croient les auteurs" → on annote par commodité sur le noeud V dominant le verbe principal "est", c'est le post-traitement qui s'appliquera et retirera l'incise.

**ATTENTION:** Ce raccourci ne doit pas être utilisé pour les cas où le verbe de l'incise est la tête du noeud qu'on souhaite annoter (cas d'erreurs dans l'annotation syntaxique sous-jacente). En effet, lorsqu'un rôle est annoté comme rempli par le noeud dominant directement le déclencheur de son frame, on interprétera qu'il s'agit d'un cas d'incorporation de rôle au déclencheur. L'heuristique de suppression du déclencheur (et de ses dépendants)

<sup>7</sup>Restriction nécessaire, cf. la représentation des noms relationnels en emploi référentiel section 6.2).

du span du remplisseur ne sera donc pas appliquée.

**ATTENTION:** Ce raccourci ne doit pas non plus être utilisé si l'heuristique donne un remplisseur mal formé. Par exemple:

- (139) 12995: celle - ci est loin d' être une réalité , comme en **témoignent** notamment les difficultés de la négociation sur le GATT .

On veut coder "celle - ci est loin d' être une réalité" comme rôle Proposition/FULLANT (et "en" est le remplisseur ANAPH). MAIS on ne peut pas utiliser le noeud V dominant le verbe principal "est" cf. cela couvrirait toute la phrase et si on enlève (automatiquement) le déclencheur et son sous-arbre on obtiendra au final un remplisseur mal formé (car incluant le "comme" : on aurait Proposition="celle - ci est loin d' être une réalité , comme"). Il faut alors annoter précisément le rôle Proposition (annotation "multi-pattes") pour obtenir exactement "celle - ci est loin d' être une réalité".

### 7.3 Remplisseur adjectival synonyme d'un PP

On a parfois la mention d'un remplisseur sous une forme elliptique, au travers d'un adjectif qui caractérise le remplisseur. → contrairement au choix de départ, on décide d'annoter ces adjectifs comme remplisseur, avec un flag "ELLIPTIC". Par exemple :

- (140) les **achats** [chinois]<sub>Buyer/ELLIPTIC</sub> (→ on comprend "les achats des Chinois")  
(141) 19201: ...les crédits affectés dans le budget 1990 de Paris aux **acquisitions** [foncières]<sub>Goods/ELLIPTIC</sub> ont subi une coupe ... (→ on comprend : les acquisitions de biens fonciers)  
(142) 5140 Tout au long de la **discussion** [parlementaire]<sub>Interlocutors/ELLIPTIC</sub>

Attention: le nom du flag ELLIPTIC est assez mal choisi. On ne l'utilise en fait pas pour les cas d'ellipse du nom comme:

- (143) 54 % des fonctionnaires ont un second emploi et près de 30 % se livrent à une forme de commerce sur leur lieu de travail  
→ pas de flag ELLIPTIC sur le participant "près de 30%"

On l'utilise bien en revanche pour des cas où le remplisseur a pour tête un nom qui a un type sémantique différent de celui attendu par le rôle, et que l'interprétation nécessite d'ajouter une tête au remplisseur:

- (144) 286 Il n'en **justifie** pas moins [le décret]<sub>Eventuality/ELLIPTIC</sub> ...  
→ "justifie" sélectionne une éventualité, ici on reconstruit "justifie (la prise du) décret"

## 8 Phénomènes de portée

### 8.1 remplisseurs phrastiques/infinittivaux: gestion des verbes introducteurs de complétives/d'infinittives

**NB:** cette section utilise des exemples contenant des déclencheurs du domaine Causalité. Certains (*parce que*) sont ambigus entre les frames Causation et Evidence (cf. le guide spécifique de désambiguïsation Causation / Evidence). Dans cette section, on laisse de côté cette ambiguïté, en ne prenant que des exemples où *parce que* sera annoté avec le frame

Causation (on utilise également des exemples avec le déclencheur *grâce à*, qui n'évoque que Causation et jamais Evidence).

On considère les cas de remplisseurs phrastiques ou infinitivaux contenant ou introduit par un verbe enchâssant. Le verbe enchâssant (souligné dans les exemples de cette section) est selon l'interprétation, inclus (exemple (145)) ou pas (exemple (146)) au remplisseur.

(145) [Paul pense que les enfants sont rentrés]<sub>Effect</sub> **parce qu'** [il a vu leur cartable dans l'entrée]<sub>Cause</sub>.

(146) Paul pense que [les enfants sont rentrés]<sub>Effect</sub> **parce qu'** [ils avaient faim]<sub>Cause</sub>.

En (145), c'est parce qu'il a vu leur cartable que Paul pense que les enfants sont rentrés, on inclut le verbe enchâssant *penser* au rôle Effect. En (146) au contraire, la portée est sur le verbe enchâssé uniquement: Paul pense que c'est parce qu'ils avaient faim que les enfants sont rentrés. Ce cas est le plus fréquent.

Dans ces deux exemples, l'ambiguïté est normalement levée par l'analyse syntaxique : *parce que* est normalement rattaché à *pense* en (145) et à *rentrés* en (146). Mais attention : on ne doit pas se fier aveuglément à la structure syntaxique proposée, qui peut comporter des erreurs: il vous faut annoter le remplisseur en fonction de votre interprétation de la phrase. En outre, on peut avoir structurellement un rattachement obligatoire du déclencheur sur le verbe introducteur, mais avec une portée du déclencheur sur le verbe enchâssé. C'est en particulier le cas en cas d'antéposition de la subordonnée introduite par le déclencheur:

### 8.1.1 Portée sémantique sur verbe enchâssé versus portée sur verbe introducteur

Portée sur verbe enchâssé :

(147) **Grâce à** [ces mesures]<sub>Cause</sub>, [les intervenants]<sub>Effect</sub> déclarent [avoir économisé du temps]<sub>Effect</sub>.

(148) **Grâce à** [ces mesures]<sub>Cause</sub>, les intervenants déclarent qu'ils ont économisé du temps]<sub>Effect</sub>. → l'interprétation est claire : les intervenants déclarent (qu'ils ont économisé grâce à ces mesures).

À contraster avec portée sur verbe introducteur :

(149) **Grâce à** [ces mesures]<sub>Cause</sub>, [les intervenants sont autorisés à économiser du temps]<sub>Effect</sub>. → l'interprétation est claire : c'est grâce à ces mesure que les intervenants ont le droit d'économiser.

On annote les cas de portée sur verbe introducteur de façon régulière (pas de problème particulier). En revanche, pour les cas de portée sur V enchâssé, **il faut veiller à exclure du remplisseur le verbe introducteur**. On annote comme remplisseur l'infinitive ou la complétive enchâssée, et en outre dans le cas de l'infinitive, on inclut le sujet de l'infinitif (cf. infra section 9).

### 8.1.2 Difficultés de détermination de portée

Pour certains verbes introducteurs, trancher entre portée sur verbe introducteur et portée sur verbe enchâssé peut s'avérer assez difficile. C'est en particulier le cas pour les verbes épistémiques (exprimant un degré de croyance, *penser*, *croire*...)

- (150) **Grâce à** ces mesures, les intervenants pensent avoir économisé du temps. → l'ambiguïté est soit ce sont les mesures qui causent qu'ils pensent avoir économisé, soit ils pensent que ce sont les mesures qui leur ont fait économiser du temps.

On peut construire des exemples forçant clairement la portée. Par exemple ci-dessous, sur verbe épistémique introducteur:

- (151) **Grâce aux** [mesures de propagande et de manipulation de la part des organisateurs]<sub>Cause</sub>, [les intervenants pensent avoir économisé]<sub>Effect</sub>.

Selon les différents types de verbe introducteur listés ci-dessous, on demande de désambigüiser en contexte la portée, ou bien on recommande une annotation systématique.

- (152) **Grâce à** [ces mesures]<sub>Cause</sub>, [les intervenants]<sub>Effect</sub> pensent [avoir économisé du temps]<sub>Effect</sub>.

### 8.1.3 Gestion dans Asfalda : classes de verbes introducteurs

#### Verbes de communication/d'influence sur autrui

Pour les verbes dénotant la communication verbale : il ne semble pas difficile de faire effectivement la désambigüisation de la portée (cf. supra l'annotation différente pour (148) versus (149)). On a en général plutôt un remplisseur n'incluant pas le verbe enchâssant de communication.

Liste non exhaustive de v de discours rapporté (cf. le guide annodis p.6) : clamer, proclamer, insister, soutenir, déclarer, affirmer, faire valoir, revendiquer, imposer, maintenir, témoigner, attester, confirmer, assurer, garantir, certifier, conclure, démontrer, montrer, souligner, informer, noter, révéler, annoncer, expliquer, dire, rapporter, raconter, narrer, relater, indiquer....

persuader, convaincre

Exemple:

- (153) 2813: [Un dirigeant de la CG2A]<sub>Effect</sub>, lors\_de l' attribution du marché, a ainsi déclaré au juge [avoir acquitté de semblables factures au mois de février 1992]<sub>Effect</sub> " **parce que** [la commission d' appel d' offres se réunissait la semaine suivante]<sub>Cause</sub> " .

#### Verbes aspectuels

Pour les verbes aspectuels (=indiquant le début, la fin, la continuation, le futur proche d'une éventualité : commencer, finir, continuer, cesser, aller etc...), on décide de ne pas désambigüiser entre portée sur verbe enchâssé ou sur verbe introducteur: on inclut systématiquement le verbe aspectuel dans le remplisseur (il en est la tête). Donc un remplisseur de tête verbe aspectuel sera en fait sous-spécifié quant à la portée.

Exemples:

- (154) 10936': **Grâce à** [cet accord]<sub>Cause</sub>, [les hongrois vont élargir leur activité en intégrant un réseau européen]<sub>Effect</sub> .
- (155) 15690: [L' accès grandissant des femmes ingénieurs au milieu professionnel continuera sans doute de leur profiter]<sub>Effect</sub> **grâce à** [la pénurie d' ingénieurs diplômés]<sub>Cause</sub> .
- (156) 15827: **Grâce au** [président Delaporte]<sub>Actor</sub>, [elle commence à s' ouvrir à l' extérieur]<sub>Effect</sub> .

### Verbes de tentative/ promesse / factifs

(essayer de , s'appliquer à ...)

(oser, avoir le toupet de, se permettre de)

→ il semble que l'on aura toujours portée sur verbe introducteur.

- (157) 1813: Cette demande a été rejetée et [il est en prison]<sub>Effect</sub> **parce qu'** [il a osé , il y a quelque temps , proposer la recherche d' une solution politique au problème kurde]<sub>Cause</sub> .

### Verbes modaux (pouvoir, devoir) / degré de croyance de la part de l'auteur (sembler, paraître)

X peut V<sub>inf</sub> (épistémique: il est probable que V<sub>inf</sub>/ déontique: X a la possibilité morale ou physique de V<sub>inf</sub>)

X doit V<sub>inf</sub> (épistémique: il est vraisemblable que V<sub>inf</sub> / déontique: X a l'obligation morale ou physique de V<sub>inf</sub>)

X semble V<sub>inf</sub> / il semble que P

Bien qu'on puisse souvent avoir portée sur le verbe enchassé, on inclut systématiquement ces verbes introducteurs au remplisseur, sans chercher à désambigüiser la portée, cf. cela s'avère trop difficile.

- (158) 4394: [Prague devrait devenir la plaque tournante aérienne d' Europe centrale]<sub>Effect</sub> **grâce à** [sa vocation touristique et à ses atouts industriels]<sub>Cause</sub>

- (159) 5974: **Grâce à** [l' appui technique de ces deux organismes]<sub>Cause</sub> , [les choses devraient aller plus vite]<sub>Effect</sub> → sur enchâssé : il est probable que ( grâce à l' appui technique ..., les choses aillent plus vite )

- (160) **Grâce à** [ces mesures]<sub>Cause</sub>, [les intervenants]<sub>Effect</sub> doivent [avoir économisé du temps]<sub>Effect</sub>.

Remarque: on peut faire le parallèle avec le cas d'un adverbe épistémique (indiquant un degré de croyance), où on suit la syntaxe, en incluant l'adverbe dès lors que le verbe qu'il modifie est inclus au remplisseur. On ne représente pas que sémantiquement l'adverbe porte sur toute la relation de causalité.

- (161) **Grâce à** [ces mesures]<sub>Cause</sub>, [les intervenants ont certainement économisé du temps]<sub>Effect</sub>.

### Verbes de degré de croyance d'un participant

X ( pense / croit / imagine / suppose / estime / juge / a l'impression / devine / suppose / s'attend à ce que / anticipe / sent / est convaincu / est persuadé / savoir / être sûr / être certain... ) que P

Il semble à X que P

X ( pense / croit / imagine / suppose / estime / juge / a l'impression de / devine / suppose / s'attend à / anticipe / sent / est convaincu de / est persuadé de ) V<sub>inf</sub>

Exemples:

- (162) 13725: Le groupe s'attend à [une hausse de 50 % de ses résultats]<sub>Effect</sub> **grâce à** [des plus - values de cession , notamment sur Sopharga]<sub>Cause</sub> .
- (163) 1701 modifié: Je pense qu' il s' agit d' un objectif important **parce que** ... nous n' avons d' autre choix que d' aller là où ils se trouvent . → le verbe épistémique introduit une opinion, ce cas est annoté avec le frame Evidence (voir guide de désamb Causation / Evidence)

### Verbes d'orientation

Sentiment d'un cognizer par rapport à la réalisation future / ou passée d'une éventualité (souhaiter, espérer, vouloir, désirer, craindre, redouter, regretter, apprécier ...)

On fait l'interprétation en contexte, en privilégiant l'annotation n'incluant pas le verbe introducteur au remplisseur.

- (164) 10132: **Grâce à** [ces mesures]<sub>Cause</sub> , les intervenants espèrent enfin revoir [l'économie repartir]<sub>Effect</sub> .

## 8.2 Portée de la négation et des modaux dans un remplisseur

On ne gère pas la portée de la négation au sein d'un remplisseur. Par exemple:

- (165) 15107: En cas de troubles , guérillas ou autres soulèvements , la production n' est pas perturbée **grâce au** stockage flottant  
→ = le stockage flottant cause la non perturbation de la production

versus

- (166) 15107 modifié : La production n'est pas perturbée **à cause du** stockage flottant mais à cause de l'invasion par les sardines  
→ = il n'est pas vrai que (le stockage flottant cause la perturbation de la production)

On ne traite pas cette subtilité de portée de la négation : dans les 2 cas, le remplisseur intègre la marque de négation (Effect= la production n'est pas perturbée).

D'une manière générale on ne gère pas non plus la portée d'un modal. Exemple interagissant avec un modal :

- (167) 1667: ... ils ne peuvent être punis **parce qu'** ils viennent volontairement en aide au Kosovo .  
lecture 1 : il n'est pas possible qu' ( ils soient punis parce qu'ils viennent en aide au Kosovo )  
versus  
lecture 2 (erronée) : (il n'est pas possible qu'il soient punis ) parce qu'ils viennent en aide au Kosovo

On ne gère pas la distinction de portée du modal, et on annote Effect="ils ne peuvent être punis".



## 9 Saturation des remplisseurs prédicatifs et divergence entre tête syntaxique et tête sémantique

Lorsqu'un rôle correspond sémantiquement à une éventualité (un évènement, un état), son remplisseur sera constitué d'une tête sémantique (le mot décrivant l'éventualité: un verbe, un nom prédicatif, un adjectif...), et des dépendants sémantiques de cette tête.

Le cas le plus simple est un remplisseur de type phrase (168) ou bien de type NP complètement saturé (169) (i.e. avec tous ses compléments).

(168) Paul a **appris** à 17h que [Pierre partirait bientôt]<sub>Phenomenon</sub>

(169) Paul a **appris** à 17h [le départ imminent de Pierre]<sub>Phenomenon</sub>

D'une part on a alors, sauf cas d'extraction, un remplisseur sans discontinuité, et d'autre part la tête *syntactique* du remplisseur (le verbe partirait en (168) ou le nom départ en (169)) correspond bien à la tête *sémantique* du remplisseur.<sup>8</sup> On a alors la situation simple où on peut annoter le remplisseur simplement en pointant le noeud de sa tête à la fois sémantique et syntaxique.

Mais nous détaillons maintenant:

- des cas où le remplisseur est discontinu: il faut aller récupérer d'autres constituants pour "saturer" le remplisseur de tous ses arguments,
- des cas où cette saturation amène à des situations où la tête syntaxique d'un remplisseur n'est pas la tête sémantique voulue, cas pour lesquels on introduit donc un marquage spécial.

### 9.1 Saturation des remplisseurs infinitivaux

Lorsque la tête sémantique du remplisseur est un verbe non conjugué, le sujet de cette tête sém ne peut pas être réalisé comme un dépendant syntaxique de la tête sém du remplisseur. C'est typiquement le cas pour un remplisseur infinitival : le sujet de l'infinitif est soit non réalisé, soit réalisé non localement. Si la sémantique du rôle correspond bien à la prédication complète, alors on choisit d'inclure au remplisseur le sujet de l'infinitif.

(170) Anna **dit** que [Paul est parti à 23h]<sub>Message</sub>

(171) Paul **dit** être parti à 23h → Message = "Paul" + "être parti à 23h"

Attention cependant, si la définition du rôle n'inclut pas le sujet, alors on n'applique pas cette règle d'inclusion du sujet au remplisseur. C'est typiquement le cas pour les rôles Action des frames XXX.to\_do. Par exemple dans Remembering\_to\_do, le frame prévoit 2 rôles Cognizer et Action, et pour tous les déclencheurs du frame, l'action a forcément pour sujet le Cognizer. Le périmètre du rôle Action est toute la prédication sauf le sujet:

(172) Pierre a **oublié** de [fermer le gaz]<sub>Action</sub> → Cognizer = Pierre / Action= fermer le gaz

### 9.2 Saturation des remplisseurs adjectivaux

Dans le cas d'un remplisseur dont la tête *sémantique* est un adjectif, on doit lister les 3 positions syntaxiques principales pour un adjectif

<sup>8</sup>Dans tous les exemples de cette section, la tête sémantique du remplisseur est soulignée.

- épithète : *une enfance heureuse*
- attribut du sujet : *mon enfance fut heureuse*
- attribut de l'objet : *il trouve son enfance heureuse*

Dans tous ces cas, l'argument sémantique premier de l'adjectif (qu'on appelle parfois le "sujet" de l'adjectif) n'est pas un dépendant syntaxique de l'adjectif (*enfance* dans les 3 exemples ci-dessus ne dépend pas syntaxiquement de *heureuse*).

Comme pour les infinitifs, on choisit de saturer les remplisseurs, en y incluant les sujets des adjectifs.

Plus précisément: pour un attribut du sujet, dépendant donc d'un verbe d'état, on annote comme remplisseur simplement toute la phrase du verbe d'état.

(173) Anna a **dit** que [Paul **était** content de la fête]<sub>Message</sub>.

Pour un attribut de l'objet, comme en (174), si l'analyse syntaxique ne comporte pas d'erreur, on aura dans ce cas un remplisseur constitué de 2 sous-arbres syntaxiques : un pour l'objet direct ("Pierre") et un pour l'attribut ("content"), qui est bien la tête sémantique de l'attribut. (Rem: dans ce cas le post-traitement (cf. section ??) donnera l'adjectif comme tête syntaxique et tête sémantique du remplisseur).

(174) Paul **trouve** Pierre intelligent  
 → Content = "Pierre" + "intelligent"

### 9.2.1 Remplisseur adjectif ou participe épithète et divergence entre tête syntaxique et sémantique

Le plus fréquemment l'adjectif est épithète. On se retrouve alors dans la situation où l'élément que l'adjectif modifie est le gouverneur syntaxique de l'adjectif, mais en est un argument sémantique. Si l'adjectif correspond à la tête sémantique d'un remplisseur, et que l'on sature ce remplisseur en récupérant le sujet de l'adjectif, alors le remplisseur sera tout le NP.

Ainsi pour un NP remplisseur, il faut distinguer le cas classique où c'est bien l'entité dénotée par le NP qui est sémantiquement le remplisseur, comme infra en (175), (tête sém = la tête synt = *enfance*) :

(175) Paul a longuement **évoqué** [son enfance heureuse]<sub>Topic</sub>.

du cas où la tête sémantique est une prédication interne au NP : un adjectif (ou un participe) épithète, et c'est en saturant celui-ci que l'on récupère tout un NP. Ainsi en (176), le rôle Effect est sémantiquement "son enfance a été heureuse", et rôle Cause = "son enfance a été paisible", ce que l'on résume par "la tête sémantique du remplisseur est "heureuse"":

(176) Paul a évoqué longuement [son enfance heureuse]<sub>Effect</sub> **car** paisible.

Autre exemple réel cette fois-ci: Dans l'exemple ci-dessous (6875), le rôle Effect est sémantiquement la prédication équivalente à "l'écart de taux (était) très important il y a quelques années", mais syntaxiquement en incluant le "sujet" de "important" on inclut le NP de tête "écart", dont "important" est syntaxiquement un dépendant modifieur. On utilise une astuce pour bien repérer les cas où le remplisseur a une tête syntaxique (ici "écart") qui ne correspond pas à la tête sémantique (ici "important") : on annote sous Salto un remplisseur à 2 pattes une sur le noeud dominant le remplisseur complet (ici le NC dominant "écart")

et en plus une patte sur le token qui est en fait la tête sémantique prédicative (ici le token "important").

- (177) 6875: ..., [l' écart de taux, très important il y a quelques années]<sub>Effect</sub> , **car** il traduisait les inquiétudes des milieux internationaux sur la santé du franc , s' est très fortement réduit ...  
→ le rôle Effect a une tête synt (écart) distincte de la tête sém → on doit marquer la tête sém. (une patte en plus sur le token "important").

On a des situations similaires avec des participes, présents ou passés. Dans l'exemple ci-dessous, le remplisseur complet est "un PDG de thomson CSA, M. Gomez, soupçonné... blancs", il domine syntaxiquement à la fois le déclencheur soupçonné, et la tête sém du remplisseur correspondant au contenu du soupçon (tête sém = s'intéresser)

- (178) 16016: Conjonction d'intérêts d'un PDG de Thomson SA , M. Gomez , **soupçonné** de ne s' être jamais intéressé à la vie de sa filiale de " produits blancs " (...)  
→ rôle Content = le noeud NC dominant PDG, plus 2 pattes sur les tokens "s'" et "intéressé"

On peut avoir combinaison de récupérations de sujet. Dans l'exemple suivant, on a ce phénomène pour le rôle Effect (remplisseur dont la tête synt est "ensemble", et la tête sém est "grave"), et en outre, pour le rôle Cause, qui est un participe ("alimentée"), on va récupérer le sujet ("ensemble"). 11348: L' ensemble du système financier mondial subit aussi une crise immobilière bien plus grave que les précédentes CAR alimentée par la déréglementation des années 80 . → Causation, Effect="une crise immobilière .... précédentes" + patte sur "grave" ET AUSSI, rôle Cause ="une crise immobilière ... précédentes" + une patte sur "alimentée"

- (179) 14758: Le chômage de longue durée , jusqu' alors éradiqué **grâce à** un suivi social systématique, ...  
→ Effect = noeud NC dominant "chômage" , plus une patte sur "éradiqué"

Dans l'exemple suivant, on va récupérer le sujet de "acteur" et de "courant" :

- (180) 5553: Acteur moyen de l' industrie des télécommunications et **donc** courant le risque d' être marginalisée , Matra Communication , filiale spécialisée du groupe Matra , a décidé de nouer une alliance  
→ Effect= "Matra communication filiale...Matra" + "acteur moyen de l'industrie de télécoms" / Cause="Matra communication filiale...Matra" + "courant le risque d' être marginalisée"

On a une situation différente avec l'exemple suivant.

- (181) 7158: De leur côté , les équipementiers européens ont fait des efforts pour adapter le système en réduisant son volume ( 30 litres au lieu de 60 ) , sa sophistication et **donc** son coût .

On peut reconstruire qu'une cause (les équipementiers ont réduit le volume et la sophistication du système) a causé un effet (les équipementiers ont réduit le coût de ce système). D'une part pour la cause, la tête sém ("réduisant") n'a pas son sujet localement (→ on va récupérer "les équipementiers" : rôle Cause = "les équipementiers" + "réduisant son volume ... sa sophistication"). Et d'autre part, pour l'effet, on a une coordination elliptique : l'objet direct du prédicat "réduisant" est coordonné, le prédicat n'est pas répété. On

doit par cohérence récupérer Effect = "les équipementiers" + "réduisant" + "son coût". C'est-à-dire que l'on doit récupérer non seulement le sujet du prédicat remplisseur, mais le prédicat lui-même ...

## 10 Bibliographie

### References

- Anne Abeillé and Nicolas Barrier. 2004. Enriching a french treebank. In *Proc. of LREC'04*.
- C. F. Baker, C. J. Fillmore, and J. B. Lowe. 1998. The Berkeley FrameNet project. In *COLING-ACL '98: Proc. of the Conference*, pages 86–90.
- A. Burchardt, K. Erk, A. Frank, A. Kowalski, S. Padó, and M. Pinkal. 2006. SALTO – A Versatile Multi-Level Annotation Tool. In *Proc. of LREC'06*.
- M. Candito and D. Seddah. 2012. Le corpus sequoia : annotation syntaxique et exploitation pour l'adaptation d'analyseur par pont lexical. In *Proc. of TALN 2012 (in French)*.
- M. Candito, P. Amsili, L. Barque, F. Benamara, Ga. de Chalendar, M. Djemaa, P. Haas, R. Huyghe, Y. Y. Mathieu, P. Muller, B. Sagot, and L. Vieu. 2014. Developing a French FrameNet: Methodology and first results. In *Proc. of LREC'14*.
- C. J. Fillmore. 2007. Valency issues in framenet. In Thomas Herbst and Katrin Gotz-Votteler, editors, *Valency: theoretical, descriptive and cognitive issues*, volume 187 of *Trends in Linguistics*, pages 129–162. Mouton de Gruyter.
- J. Ruppenhofer, M. Ellsworth, M. R.L. Petruck, C. R. Johnson, and J. Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice*. International Computer Science Institute, Berkeley, California. Distributed with the FrameNet data.
- D. Seddah, R. Tsarfaty, S. Kübler, M. Candito, J. Choi, R. Farkas, J. Foster, I. Goenaga, K. Gójenola, Y. Goldberg, S. Green, N. Habash, M. Kuhlmann, W. Maier, J. Nivre, A. Przepiorkowski, R. Roth, W. Seeker, Y. Versley, V. Vincze, M. Woliński, A. Wróblewska, and E. Villemonte de la Clérgerie. 2013. Overview of the SPMRL 2013 Shared Task: A Cross-Framework Evaluation of Parsing Morphologically Rich Languages. In *Proc. of the 4th Workshop on Statistical Parsing of Morphologically Rich Languages: Shared Task*.