

# Guide d'annotation FR-SemCor

L. Barque, P. Haas, R. huyghe et D. Tribout

<b>Introduction</b>	<b>2</b>
<b>I. Présentation des classes sémantiques</b>	<b>3</b>
1. N dénotant des entités concrètes	3
Animal	3
Person	3
Plant	4
Artifact	4
Body	5
Food	5
Object	5
Substance	6
2. N dénotant des situations dynamiques	6
Act	6
Event	7
3. N dénotant des situations non dynamiques ou des propriétés	8
Attribute	8
Feeling	8
State	9
4. N dénotant des groupes, des parties, des quantités	9
Group	9
Institution	10
Part	11
Quantity	11
5. Autres étiquettes	12
Cognition	12
Communication	13
Relation	13
Possession	13
Phenomenon	14
Time	14
Tops	14
<b>II. Les étiquettes multiples liées par un opérateur</b>	<b>14</b>
L'opérateur de distribution	15
L'opérateur x	15
Les cas de typages multiples	15
L'opérateur /	15
L'opérateur +	16
<b>III. Annotation des expressions polylexicales</b>	<b>16</b>

# Introduction

Ce guide donne la marche à suivre pour l’annotation des noms du corpus Sequoia (Candito & Seddah 2012) dans le cadre du projet Fr\_SemCor<sup>1</sup>.

Le tagset utilisé pour annoter les noms du corpus est composé de 24 étiquettes issues des *First Beginners* de WordNet, qui peuvent se répartir de la façon suivante :

1. Etiquettes pour les N dénotant des entités concrètes, tangibles, c’est-à-dire des objets du monde qui peuvent être appréhendés par les sens : Animal, Person, Plant, Artifact, Object, Body, Substance, Food
2. Etiquettes pour les N dénotant des situations dynamiques, c’est-à-dire des choses qui ont lieu, qui se produisent ou encore que l’on peut effectuer : Act, Event
3. Etiquettes pour les N dénotant des situations non dynamiques ou des propriétés : Attribute, Feeling, State
4. Etiquettes pour les N dénotant un groupe, une partie ou une quantité de quelque chose : Group, Part, Quantity, Institution
5. Les étiquettes restantes correspondent à des types sémantiques “isolés”, qui n’entrent pas, ou pas clairement, dans les 4 grandes catégories citées précédemment : Cognition, Communication, Relation, Possession, Phenomenon, Time, Tops

NotaBene 1 : Une occurrence de N reçoit généralement une seule des 24 étiquettes. Dans certains cas toutefois, il est possible d’annoter une occurrence de N à l’aide d’une étiquette complexe composée de deux étiquettes simples liées par un opérateur. Par exemple, dans la phrase (1) ci-dessous, le nom *groupe* dénote un groupe de personnes et sera donc étiqueté Group x Person (pour plus de détails, voir section II)

(1) *En effet, parmi les adhérents à cette activité, plusieurs manient déjà avec brio différentes techniques artistiques, qu'ils pourront faire partager à l'ensemble du **groupe***<sub>Group x Person</sub>

NotaBene 2 : les noms à annoter peuvent faire partie d’une expression polylexicale (EP), pré annotée dans le cadre du projet Parseme-fr. Dans ces cas là, seuls les noms correspondant à la tête d’une EP nominale sont à annoter et l’étiquette choisie doit représenter le sens de l’EP dans son ensemble (pour plus de détails, voir section III).

(2) *Par le biais de l’activité théâtre, ils se sont intéressés à l’expression orale et à la **[mise***<sub>Act</sub> *en scène]*<sub>EP</sub> *avec un intervenant, Julien Georges, maître de théâtre.*

---

<sup>1</sup> Financé par le labex EFL (reference: ANR-10-LABX-0083)

# I. Présentation des classes sémantiques

Cette section du guide présente les 24 étiquettes utilisées pour l'annotation des noms, regroupées selon qu'elles dénotent des entités concrètes (section 1), des situations dynamiques (section 2), des situations statives ou des propriétés (section 3), des groupes, des parties ou des quantités (section 4) ou autres (section 5). Chaque étiquette est définie et les contours de la classe qu'elle représente sont précisés par la mention de sous-groupes inclus dans la classe (notifiés par le symbole  $\subseteq$ ) ou au contraire non inclus dans la classe (notifiés par le symbole  $\not\subseteq$ ). Des tests linguistiques sont parfois proposés pour évaluer l'appartenance d'un nom à une classe. Ces tests, lorsqu'ils sont disponibles, ne sont toutefois qu'indicatifs (suffisants mais non nécessaires). Enfin, un ensemble d'exemples, provenant du du journal *Le Monde* ou du corpus Sequoia (ajoutés après la campagne d'annotation), illustrent chacune des classes.

## 1. N dénotant des entités concrètes

Les étiquettes utilisées pour les noms dénotant des entités concrètes sont : Animal, Person, Plant, Artifact, Object, Body, Substance, Food

### Tests indicatifs :

- *{Det N} se trouve sur la table / au fond du jardin / dans le prolongement du bras*
- *Paul a serré {det N} dans ses doigts / ses mains / ses bras*

### Animal

**Définition** : être animé non humain

- $\not\subseteq$  les bactéries et les virus (micro-organismes), classés dans Object
- $\not\subseteq$  la viande/chaire de l'animal, classés dans Food

NotaBene 1 : les N comme *bébé*, *jeune*, etc. sont étiquetés Animal s'ils ont dans la phrase un référent de type animal

NotaBene 2 : les animaux dans les histoires sont annotés Animal même s'ils ont des comportements humains

### Exemples :

- *A la grande époque, le **chien**<sub>Animal</sub> vedette passait son temps à se castagner avec son compère, le **chat**<sub>Animal</sub> Hercule.*
- *La destruction de l'habitat naturel de cette **espèce**<sub>Group x Animal</sub> endémique de l'île-continent a causé une hécatombe : 53 % des **marsupiaux**<sub>Animal</sub> ont disparu dans le Queensland.*

- Les premiers rayons du soleil percent le ciel quand Olivier Fiadana finit de préparer ses belaroa, ses « balances à **crabes**<sub>Animal</sub> ».
- À l'époque, ce cher "Jo" élevait jusqu'à 120 **pigeons**<sub>Animal</sub> voyageurs, et les spécialistes n'ont pas oublié que l'un de ses **champions**<sub>Animal</sub> fut de loin le meilleur sur 500 kilomètres lors d'un retour d'Evreux et de Bayeux ! [Sequoia-annodis.er\_00091]

## Person

**Définition :** humains et personnages imaginaires ayant des attributs humains saillants (dieux, personnages mythologiques)

⊖ les métiers (ex. *boucher, avocat*)

⊖ les fonctions/statuts (ex. *mère, maire, actionnaire*). Attention : si le statut s'applique à une Institution (ex. *La martinique, plus grand **producteur** de ...*), choisir l'étiquette Institution

### Exemples :

- De mémoire de **météorologues**<sub>Person</sub> seuls l'ont surpassé des typhons extrêmes dans le Pacifique.
- La chouha ("humiliation" en arabe dialectal) c'est ce que subissent les **Marocaines**<sub>Person</sub> qui ont eu des rapports sexuels avant le mariage.
- Donald Trump s'est érigé en **défenseur**<sub>Person</sub> des symboles des Etats-Unis.
- Les **clients**<sub>Person</sub> fragiles continuent d'alimenter largement les banques en commissions diverses.
- **M.**<sub>Person</sub> Dumez, **président**<sub>Person</sub> du conseil général, a parlé d'une "réalisation apte à l'épanouissement des **petits**<sub>Person</sub>", non sans rappeler le rôle primordial de l'éducation familiale. [Sequoia-annodis.er\_00035]

## Plant

**Définition :** N dénotant une entité du règne végétal ou une de ses parties

⊖ végétaux (ex. *arbre, plante, algue, chêne*)

⊖ parties de plantes (ex. *tige, feuille, racine*)

⊖ fruits, légumes, céréales, classés dans Food

⊖ substances issues d'une entité végétale (ex. *sève, pollen*), classées dans Substance

### Exemples :

- Dans ce labyrinthe de **palétuviers**<sub>Plant</sub> qui forment la **mangrove**<sub>Group x Plant</sub>, Olivier Fiadana connaît les recoins de tous les canaux.
- La société Spread produit en hydroponie (hors-sol) sur 2 900 m<sup>2</sup> quatre variétés de salades, de la plus tendre à la plus croquante, disposées en rangs serrés sur de vastes étagères, dans un environnement à 25 degrés et à 60 % d'humidité, sous la lumière de simples néons et avec les **racines**<sub>Plant</sub> trempant dans une solution riche en nutriments comme le potassium.

- *En parfaite femme d'intérieur, elle a pu, lors de ce repos bien mérité, s'occuper de sa maison, de ses **fleurs**<sub>plant</sub>.* [Sequoia-annodis.er\_00398]

## Artifact

**Définition** : Objets physiques non animés fabriqués

- ☞ entité utilisée dans un but précis (ex. *radio-réveil, décoration*)
- ☞ bâtiments ou partie d'un bâtiment (ex. Immeuble, vestibule)
- ☞ territoires construits (ex. *banlieue*) (sauf ceux pouvant avoir l'étiquette Institution)
- ☞ les supports d'objet informationnel (ex. *partition, document*), qui recevront un multi-typage (voir l'exemple d'*acte* ci-dessous, qui relève du domaine du droit et renvoient donc à des documents écrits). Attention, les objets informationnels situables à l'intérieur d'un autre objet informationnel ne sont pas Artifact (ex. *Le **paragraphe** situé en haut de la page 3.* > Cognition seulement)
- ☞ les plats préparés, classés dans Food
- ☞ les médicaments, classés dans Substance
- ☞ les zones géographiques délimitées administrativement (ex. *commune, département*), classées dans Institution

**Tests indicatifs** : Pierre a {fabriqué / confectionné} ce N {à tel endroit / à tel moment}

**Exemples** :

- *Depuis le **pont**<sub>Artifact</sub> supérieur de ces **navires**<sub>Artifact</sub>, on domine les **toits**<sub>Artifact</sub> de la **ville**<sub>Artifact</sub>.*
- *Réunir les éléments dans une **casserole**<sub>Artifact</sub> et porter à ébullition.*
- *Les parties ont deux jours pour demander des **actes**<sub>Artifact / Cognition</sub> supplémentaires.*
- *Bref, en quittant le **fornil**<sub>Artifact</sub>, chacun ne rêvait que de remettre la main à la pâte avec Sylvie et Daniel.* [Sequoia-annodis.er\_00053]
- *Au mois de mars dernier, les enfants de CE2 de Mme Philippe avaient montré leur **travail**<sub>Artifact + Cognition</sub> réalisé d'après "Les contes du chat perché" de Marcel Aimé.* [Sequoia-annodis.er\_00146]
- *Depuis, elle vivait seule dans son petit **pavillon**<sub>Artifact</sub> à Pilviteuil, entourée de toute l'affection de sa famille.* [Sequoia-annodis.er\_00169]

## Body

**Définition** : corps ou élément du corps ou organe des humains ou des animaux

- ☞ les fluides corporels, classés dans Substance

**Exemples** :

- *"À l'époque, je ne pensais qu'à prendre du plaisir et à découvrir mon **corps**<sub>Body</sub>", se souvient la jeune femme, silhouette fine et **yeux**<sub>Body</sub> sombres.*
- *Une étude pointe le lien entre ces produits industriels et des troubles digestifs, notamment le syndrome du **côlon**<sub>Body</sub> irritable.*

- *Angiox est prévu pour un usage par voie intraveineuse (dans une **veine**<sub>Body</sub>).*  
[Sequoia-emea-fr-dev\_00020]
- *À la ménopause, les **ovaires**<sub>Body</sub> arrêtent leur production d'hormones féminines, les estrogènes, qui permet de maintenir les **os**<sub>Body</sub> en bonne santé.*  
[Sequoia-emea-fr-dev\_00408]

## Food

**Définition** : aliment naturel (ex. gluten, sel, viandes) ou préparé (ex. buffet, civet, dîner, sauce)

- ⊆ substances ingérables, incluant boisson, aliment ou tout ce qui entre dans la composition d'un aliment
- ⊆ les fruits, légumes et céréales
- ⊆ les viandes/chaînes animales (ex. (du) *canard*)

**Exemples** :

- *Faites revenir les **oignons**<sub>Food</sub> à allure modérée et à couvert.*
- *Placer les **croûtons**<sub>Food</sub> en surface afin de bien couvrir la **soupe**<sub>Food</sub> et recouvrir le **pain**<sub>Food</sub> avec le **fromage**<sub>Food</sub> râpé.*
- *Réunir les éléments (**eau**<sub>Food</sub> ou **lait**<sub>Food</sub> - **sel**<sub>Food</sub> - **beurre**<sub>Food</sub> en petites parcelles).*
- *La société Spread produit en hydroponie (hors-sol) sur 2 900 m<sup>2</sup> quatre variétés de **salades**<sub>Food</sub>, de la plus tendre à la plus croquante, disposées en rangs serrés sur de vastes étagères, dans un environnement à 25 degrés et à 60 % d'humidité, sous la lumière de simples néons et avec les racines trempant dans une solution riche en nutriments comme le potassium.*
- *La culture verticale de **légumes**<sub>Food</sub> en usine connaît un véritable boom au Japon.*
- *Ils ont été reçus à la boulangerie Leroy pour visiter le fournil et surtout pétrir la **pâte**<sub>Food</sub> afin de confectionner de délicieux **pains**<sub>Food</sub> **au chocolat** qu'ils ont dégustés à l'heure du goûter avec un verre de **jus**<sub>Food</sub> de **fruit**<sub>Food</sub>. [Sequoia-annodis.er\_00050]*

## Object

**Définition** : objets naturels, non créés par l'homme

- ⊆ corps célestes (soleil)
- ⊆ cours d'eau (rivière)
- ⊆ objet minéral (rocher, fossile)
- ⊆ microbes (micro-organismes) : bactéries, virus, amibes, etc.
- ⊆ territoires naturels, milieu environnemental (antarctique, steppe, jungle, steppe, toundra, forêt)
- ⊕ les N de plantes, classés Plant
- ⊕ les N de substance naturelle (ex. sable, pluie), classés Substance

**Exemples** :

- *Irma a frappé mercredi 6 septembre l'**arc**<sub>Object</sub> des Antilles.*
- *Les navires de moins de 130 000 tonnes resteront autorisés à entrer dans la **lagune**<sub>Object</sub>*

- Ils devront emprunter le **chemin**<sub>Object</sub> des pétroliers et navires de commerces rejoignant le **site**<sub>Object</sub> de Porto Marghera, sur le **continent**<sub>Object</sub>.
- Vingt-six tableaux pour expliquer ce qu'est une éclipse de soleil avec une quinzaine de définitions des mots-clefs tels : **amas**<sub>Group x Object</sub> globulaire, **amas**<sub>Group x Object</sub> ouvert, **comètes**<sub>Object</sub>, **constellations**<sub>Object</sub>, **galaxie**<sub>Object</sub>, **mama** (machine automatique à mesurer l'astronomie). [Sequoia - annodis.er\_00125]
- Il convient que l'utilisateur prenne conscience des coûts réels des transports aériens, y compris de toutes les dépenses pour la protection de **l'environnement**<sub>Object</sub>. [Sequoia-Europar.550\_00420]

## Substance

**Définition** : matière physique, liquide, solide, gazeux + composés chimiques / fabriqué ou non

- ☞ les médicaments (ex. *sirop*, *aspirine*)
- ☞ les fluides corporels (ex. *sang*, *larme*)
- ☞ les substance végétales (ex. *sève*, *bois*)
- ☞ les substances consommables, alimentaires (cf Food)
- ☞ les microbes (cf. Object)

### Exemples :

- Les pailles en **plastique**<sub>Substance</sub> seront remplacées par des pailles biodégradables fournies par cette fondation, gratuitement jusqu'à la fin de 2019.
- Dimanche en milieu d'après-midi, quelques enfants s'adonnaient à des glissades alors que **l'eau**<sub>Substance</sub> cachée, recouvrait encore tout le secteur. [Sequoia-annodis.er\_00205]
- Angiox est une poudre blanche à reconstituer en **solution**<sub>Substance</sub> injectable ou pour perfusion (goutte à goutte dans une veine). [Sequoia-emea-fr-dev\_00011]
- Aucune accumulation de la **substance**<sub>Substance</sub> **active** n'a été observée dans le **plasma**<sub>Substance</sub> après l'administration de doses multiples tous les 28 jours. [Sequoia-emea-fr-test\_00295]

## 2. N dénotant des situations dynamiques

Les étiquettes utilisées pour les noms dénotant des situations dynamiques sont Act pour les situations dynamiques agentives et Event pour les situations dynamiques non agentives.

**Test indicatif** : les N dénotant des situations dynamiques (Act ou Event) passent le test *le N a eu lieu à tel moment*.

### Act

**Définition** : Noms dynamiques agentifs. Nb l'agentivité n'est pas forcément intentionnelle

- ☞ N qui dénotent des actions agentives pures
- ☞ N qui dénotent des actions causées (effectuation de l'action par une force)

- ⊆ N qui dénotent des activités (ex. tourisme, football),
- ⊆ les N qui dénotent des domaines d'activité ou des secteurs professionnels (ex. *Il travaille dans le tourisme*)
- ⊆ les N d'événement collectif (ex. spectacle, conférence, ballet)

**Tests indicatifs** (pour les N dynamiques agentif) :

- X {a effectué / a procédé à / a accompli / a commis} un N (+ exp.)
- le N effectué par X<sub>agent</sub>
- un N délibéré
- volontairement, insidieusement
- la poursuite du / de la N (avec le sens de qqn poursuit le N)

**Exemples :**

- Ces scènes sont devenues le symbole des **ravages**<sub>Act</sub> du **tourisme**<sub>Act</sub> de masse sur Venise.
- La chouha, c'est ce que subissent les Marocaines qui ont eu des **rapports**<sub>Act</sub> sexuels avant le mariage.
- La répétition du **passage**<sub>Act</sub> de ces monstres de mers entraîne des dommages irréparables sur les fondations des palais.
- A la belle saison, le **ballet**<sub>Act</sub> peut se répéter cinq ou six fois par jour.
- La **descente**<sub>Act</sub> du canal de la Giudecca restera autorisée, mais seulement aux bateaux de taille plus modeste.
- Les "grandi navi" permettent le **maintien**<sub>Act</sub> de 5000 emplois.
- Une **réflexion**<sub>Act / Cognition</sub> commune est menée avec les enseignants et les délégués de parents d'élèves, sous la conduite du CAUE. [Sequoia-annodis.er\_00030]

## Event

**Définition** : N dynamiques non agentifs / sans agent

- ⊆ Événement naturels (ex. *avalanche*)
- ⊆ Autres événements fortuits non agentifs (ex. *La rupture de la canalisation, la panne de la voiture, la mort de Jean, la scène s'est produite sous mes fenêtres*)
- ⊆ N dénotant un processus naturel de changement d'état (ex. *coagulation, développement, diminution, gonflement*)
- ⊆ Certaines maladie (ex. *peste*)

**Tests indicatifs** : un nom appartenant à la classe Event ne passe pas le test d'agentivité X {a effectué / a procédé à / a accompli / a commis} un N (+ exp.)

**Exemples :**

- De mémoire de météorologues, seuls l'ont surpassé des **typhons**<sub>Event</sub> extrêmes dans le Pacifique.
- Irma s'accompagne de **rafales**<sub>Event</sub> de vent dépassant 300 km/h.
- Il vient de provoquer une **tempête**<sub>Event</sub> médiatique.



- Un consommateur en difficulté sur cinq se voit même prélever plus de 500 euros de frais d'**incidents**<sub>Event</sub>
- La **répétition**<sub>Event</sub> du passage de ces monstres de mers entraîne des dommages irréparables sur les fondations des palais.
- Il semble qu'un malaise soit à l'origine de cet **accident**<sub>Event</sub>
- **Hémorragie**<sub>Event</sub> grave (ce qui peut être due à une **réduction**<sub>Event / State</sub> de votre nombre de plaquettes et empêche la **coagulation**<sub>Event</sub> nécessaire de votre sang)  
[Sequoia-emea-fr-dev\_00544]

### 3. N dénotant des situations statives ou des propriétés

Les étiquettes pour les noms dénotant des situations statives (=non dynamiques) ou des propriétés sont Attribute pour les propriétés constitutives et State et Feeling pour les états transitoires

#### Attribute

**Définition** : propriété constitutive d'une entité (cf +/- ILP) ou d'une situation ou propriété en soi (considérées de manière abstraite, ex. Aimer la beauté)

⊆ les noms de propriétés mesurables (ex. *taille*), y compris lorsque ceux-ci sont employés métonymiquement pour désigner la mesure elle-même (ex. *Indiquez-moi votre taille*).

**Tests indicatifs** : *Pierre/ce mur est d'un grand N, Pierre a fait preuve de N/d'un N exp*

#### Exemples :

- De **mémoire**<sub>Attribute</sub> de météorologues, seuls l'ont surpassé des typhons extrêmes dans le Pacifique.
- L'ouragan menace aussi la Floride, mais de **façon**<sub>Attribute</sub> moins prévisible.
- Ils s'éloignent lentement en direction de la passe du Lido, pour retrouver l'**immensité**<sub>Attribute</sub> de la mer Adriatique.
- "... " se souvient la jeune femme, **silhouette**<sub>Attribute</sub> fine et yeux sombres.
- Ils s'agenouillent pendant l'hymne américain pour dénoncer le **racisme**<sub>Attribute</sub>.
- La descente du canal de la Giudecca restera autorisée, mais seulement aux bateaux de **taille**<sub>Attribute</sub> plus modeste.
- Venise se dépeuple à un **rythme**<sub>Attribute</sub> accéléré.
- Guy Hosneld était une personne très connue dans la commune et aux alentours, pour son grand **dévouement**<sub>Attribute</sub> aux autres et sa participation à la vie associative du village.  
[Sequoia-annodis.er\_00075]
- Ceux d'Ancerville ont participé à l'événement en faisant découvrir leur **savoir-faire**<sub>Attribute / Cognition</sub> et leurs recettes. [Sequoia-annodis.er\_00048]

## Feeling

**Définition** : état psychologique ou physiologique transitoire

- ⊆ Affect : sentiment et émotion
- ⊆ les états physiologiques (ce qu'on ressent) (ex. *plaisir, fatigue*)

**Tests indicatifs** : Pierre {ressent / éprouve} du N / un N (+exp.) | Le sujet est l'expérimenteur de l'affect dénoté

### Exemples :

- “A l'époque, je ne pensais qu'à prendre du **plaisir**<sub>Feeling</sub> et à découvrir mon corps”, se souvient la jeune femme, silhouette fine et yeux sombres.
- Les sportifs expriment leur **mécontentement**<sub>Feeling</sub> en s'agenouillant lorsque l'hymne américain est joué dans les stades pour dénoncer le racisme.
- Le **sentiment**<sub>Feeling</sub> qui domine est surtout la **peur**<sub>Feeling</sub> de l'accident.
- Les bisphosphonates (groupe de substances à laquelle appartient Aclasta) peuvent provoquer des gonflements, rougeurs, **douleurs**<sub>Event / Feeling</sub>, **démangeaisons**<sub>Event / Feeling</sub> aux yeux ou une sensibilité des yeux à la lumière. [Sequoia-emea-fr-test\_00489]
- L'après-midi, rencontre entre Lorrains pour la 3e place contre les voisins de Joeuf, mais la **fatigue**<sub>Feeling</sub> commençait à se faire sentir. [Sequoia-annodis.er\_00226]

## State

**Définition** : propriété non constitutive d'une entité (cf +/- SLP) (ex. *paternité, séropositivité*) / états résultants (ex. *irradiation*) / état des choses (ex. *silence*)

- ⊆ les maladies comme état subi (ex. *rhume, grippe, cancer, peste*). **Attention** les maladies comme épisodes qui arrivent / se produisent (concerne certains noms de maladie seulement) sont Event > multi-typage possible
- ⊈ les états psychologiques ou physiologiques transitoires > Feeling

### Exemples :

- Il y a quatre ans, elle a perdu sa **virginité**<sub>State</sub> avec un camarade de l'université.
- La **situation**<sub>State</sub> troublée de l'époque ne facilitait pas les contrôles de l'OIC et encore moins les inculpations.
- Un consommateur en **difficulté**<sub>State</sub> sur cinq se voit même prélever plus de 500 euros de frais d'incidents.
- Ils les frôlent, moteurs à l'**arrêt**<sub>State</sub>.
- Cette solution répond à la **volonté**<sub>State</sub> d'apaiser les habitants de la lagune mais aussi à l'impératif de ne pas nuire au tourisme.
- Les “grandi navi” permettent le maintien de 5000 **emplois**<sub>State</sub>.
- La **présence**<sub>State</sub> d'une abondante forêt, fournissant l'indispensable combustible, liée à l'existence du minerai de fer exploitable à ciel ouvert, a incité les hommes à créer". [Sequoia-annodis.er\_00017]

## 4. N dénotant des groupes, des parties, des quantités

Les trois classes Group, Part et Quantity incluent un nombre important de noms relationnels, c'est-à-dire des noms impliquant la présence d'un autre élément (groupe *de qqch*, quantité *de qqch*), qui peuvent dans certains cas nécessiter le recours à un multi-typage (voir section II infra). La classe Institution se distingue des trois autres classes sur ce point.

### Group

**Définition :** N dénotant (au singulier) un ensemble d'entités ou N toujours au pluriel dénotant un groupe (gens)

- ⊆ ensemble d'entités (ex. *meute*, *classification*, *batterie (de cuisine)*)
- ⊆ "lecture collective" (ex. ennemi dans *Intelligence avec l'ennemi*)
- ⊆ quantité (vague) de qqch (*amas*, *tas*, *pile*)
- ⚡ les groupes de personnes appartenant à une organisation collective : division administrative employée comme collectif de personnes (pays, région, département (de N)), partis politiques, etc. (cf. Institution)
- ⚡ la numération (*dizaine* cf. Quantity)

### Exemples :

- Les *biphosphonates* (**groupe**<sub>Group</sub> de substances à laquelle appartient Aclasta) peuvent provoquer des gonflements.
- Notre **cheptel**<sub>Group x Animal</sub> est sous perfusion et chaque **ruche**<sub>Group x Animal</sub> morte est un nouveau coup de poignard qui nous laisse désespérés et en colère.
- Une banque prélève chaque année en moyenne 34 euros de frais pour incidents de paiement sur l'ensemble de sa **clientèle**<sub>Group x Person</sub>.
- Une juste récompense pour les efforts de qualité entrepris par Jeanine et Willy Schaer, jeune **couple**<sub>Group</sub> d'hôteliers originaires de Haute-Saône et installés à Arc-et-Senans depuis maintenant treize ans. [Sequoia-annodis.er\_00157]
- Sinon, on pouvait la rencontrer bien souvent en **forêt**<sub>Group x Plant</sub> accompagnée de ses enfants, à la recherche des champignons ou au ramassage d'escargots qu'elle préparait elle-même, pour des repas de famille. [Sequoia-annodis.er\_00171]

### Remarques pour l'annotation (voir aussi section II infra) :

- Dans une structure det N1 **de** N2 où N1 dénote un ensemble (ex. *une meute de chiens*, *le groupe d'étudiants*, *une ribambelle d'enfants*) et où, N1 est annoté Group
- Dans une structure det N où N dénote un groupe de quelque chose, que le "quelque chose" soit inclus dans le sens du nom ou qu'il soit récupérable en contexte, l'occurrence doit être annotée avec une étiquette complexe du type Group x Etiquette2. Exemple
  - *Pierre leur montra sa collection* [avec un contexte dans lequel on sait qu'il s'agit d'une collection de timbre] : le N "collection" sera annoté Group x Artifact

## Institution

**Définition :** N dénotant une institution, un organisme (ex. *banque, université*)

- ⊆ le groupe de personnes qui le constituent (ex. *le gouvernement s'est fait photographier sur les marches de l'Elysée*)
- ⊆ le lieu constitutif de l'institution (ex. *Il est allé à la banque*)
- ⊆ les N qui ont un aspect Institution sans que cet aspect soit nécessairement le plus saillant, comme les zones géographiques administrées (ex. *commune, pays*), les groupe de personnes organisés (ex. *comité, assemblée*)
- ⊆ les fonctions/statuts qui s'appliquent à l'institution (ex. *la Martinique, le plus grand producteur de ...*)

**Tests indicatifs :** *N a été fondée en / A a été nommé à la tête de N*

**Exemples :**

- *Il y a quatre ans, elle a perdu sa virginité avec un camarade de l'**université**<sub>Institution</sub>*
- *La volonté des pouvoirs publics de limiter les frais facturés par les **banques**<sub>Institution</sub> n'y a rien changé.*
- *Le **comité**<sub>Institution</sub> interministériel pour Venise, réunissant cinq **ministères**<sub>Institution</sub> et les principaux élus locaux, a tranché.*
- *Guy Hosneld était une personne très connue dans la **commune**<sub>Institution</sub> et aux alentours, pour son grand dévouement aux autres et sa participation à la vie associative du **village**<sub>Institution</sub> ; [Sequoia-annodis.er\_00075]*
- *Par téléphone, il contactait divers **hôtels**<sub>Institution</sub> du secteur pouvant accueillir les éventuels sinistrés. [Sequoia-annodis.er\_00326]*

## Part

**Définition:** N qui dénotent une partie de qqch.

- ⊆ N de localisation interne / noms de partie positionnée (*côté, extrémité, milieu, sommet*)
- ⊆ point cardinaux en position "N1 + de" (le nord de l'île)
- ⊆ partie non spécifiée (non pré-délimité et non fonctionnel) d'un tout (ex. *rondelle, morceau, tranche, partie*)
- ⊈ les méronymes. Ex. *volant* ne sera pas étiqueté Part (bien que méronyme de *voiture*) mais Artifact
- ⊈ les expressions de quantités numériques (proportions, fractions, etc) > dans Quantity
- ⊈ les parties d'une période ou d'une situation > dans Time

**Exemples :**

- *Il menace le **nord**<sub>Part</sub> d'Haïti.*

- Le sentiment qui domine est la colère à l'idée que la répétition du passage de ces monstres des mers entraîne des dommages irréparables sur les **fondations**<sub>Part</sub> des palais qu'ils frôlent.
- Après avoir coupé le ruban qui en marque symboliquement l'**entrée**<sub>Part x Artifact</sub> le maire et la directrice ont conduit la visite commentée des locaux. [Sequoia-annodis.er\_00026]
- Frôlant le **début**<sub>Part</sub> d'une glissière de sécurité, le véhicule a gravi le talus, basculé de l'autre côté, traversé un champ, est entré dans un secteur boisé, pour finalement plonger vers le centre Leclerc dans une zone à pic. [Sequoia-annodis.er\_00111]
- Seule une **bande**<sub>Part</sub> du territoire d'environ 110 km de large comprenant la Lorraine, bénéficiera de l'éclipse totale. [Sequoia-annodis.er\_00121]

## Quantity

**Définition** : objet abstrait numéral

- € quantité pure, nombres (ex. pluralité, nombre, score)
- € unité de mesure (ex. *tonne, mètre*), contenant employé comme unité de mesure (*verre, cuillère*)
- € proportion (ex. *majorité, quart, moitié, pourcent*)
- € ordinaux (*premier, second*)
- € quotient et produit
- € marqueur de quantité (*coups*)
- ⚡ les unités de mesure de temps (ex. *minute, heure*) sont dans Time
- ⚡ les noms de propriétés mesurables employés métonymiquement pour désigner la mesure elle-même (ex.: donne-moi ta **pointure** : ce qui est la mesure mais annoté comme Attribute, autres ex : poids, volume)

**Exemples** :

- Irma s'accompagne de rafales de vent dépassant 300 **km**<sub>Quantity</sub>/h.
- Une banque prélève chaque année en moyenne 34 euros de frais pour incident de paiement sur l'**ensemble**<sub>Quantity</sub> de sa clientèle.
- Je ne l'ai fait qu'une **fois**<sub>Quantity</sub>.
- Ces navires transportent plusieurs **milliers**<sub>Quantity</sub> de passagers.
- 55 000 **tonnes**<sub>Quantity</sub>, trois **fois**<sub>Quantity</sub> moins que la **limite**<sub>Quantity</sub> actuelle.
- Il a dépensé des **millions**<sub>Quantity x Possession</sub> pour ça.
- Il y avait une **dizaine**<sub>Quantity</sub> de participants.
- Quand elle ne vaquait pas à ses occupations ménagères, elle utilisait les transports en commun et descendait seule faire ses courses en centre ville, et prenait alors plaisir à faire un **brin**<sub>Quantity</sub> de causerie avec ses copines de rencontre. [Sequoia-annodis.er\_00170]
- Il rappelle néanmoins un certain **nombre**<sub>Quantity</sub> d'actions entreprises par le gouvernement, et plus particulièrement la mise en place du groupe de travail, réuni pour la première fois il y a quelques jours. [Sequoia-annodis.er\_00284]

## 5. Autres étiquettes

### Cognition

**Définition** : objets mentaux / objets informationnels

- ⇒ Produits de l'activité mentale / objet info (idées, domaines, chimère, rêverie, fantaisie, symbole) / objet iconiques
  - **Tests indicatifs** : le N selon lequel, *Pierre écoute les N décousus de Vincent, Pierre a écrit un N {intéressant / pertinent / contenant des erreurs}*, un N1 de N2 avec N2 qui est l'élément représenté (*un portrait de son père*)
- ⇒ Domaines de connaissances (Astrologie, Histoire)
- ⇒ Doctrines
- ⇒ Méthodes
- ⇒ “les mots énoncés” (ex. non, oui, merci) + facette Act pour ces noms

**Exemples :**

- *Donald Trump s'est érigé en défenseur des **symboles**<sub>Cognition</sub> des Etats-Unis.*
- *Les sportifs expriment leur mécontentement en s'agenouillant lorsque l'**hymne**<sub>Cognition</sub> américain est joué dans les stades pour dénoncer le racisme.*
- *Le sentiment qui domine est la colère à l'**idée**<sub>Cognition</sub> que la répétition du passage de ces monstres des mers entraîne des dommages irréparables sur les fondations des palais qu'ils frôlent.*
- *Cette **solution**<sub>Cognition</sub> répond à la volonté d'apaiser les habitants de la lagune mais aussi à l'**impératif**<sub>Cognition</sub> de ne pas nuire au tourisme.*
- *Je remercie le président en exercice pour sa **réponse**<sub>Act / Cognition</sub>*

Remarque sur les act à “contenu” : si le nom passe, dans le sens qu'il a dans l'occurrence à annoter, un des tests d'objet informationnel (cf. ci-dessus), alors Act/Cognition (ou Act+Cognition)

*intervention* (langagière) (Act/Cog) vs *enquête (de police)* (Act pas Cog)

### Communication

**Définition** : Objets linguistiques et médium de communication

- ⇒ Médium de communication (ex. Radio, internet)
- ⇒ Objets linguistiques (ex. idiome, parenthèse, préposition, passé simple)
- ⇒ Langues (arabe, chinois)

**Exemples :**

- *La chouha (“humiliation” en **arabe**<sub>Communication</sub> dialectal) c'est ce que subissent les Marocaines qui ont eu des rapports sexuels avant le mariage.*

- Et je viens juste d'entendre à la **radio**<sub>Communication</sub> que les pêcheurs du coin utilisaient les moustiquaires qu'on leur avait distribuées.
- La justification pour ne pas inclure du **braille**<sub>Communication</sub> a été acceptée. [Sequoia-emea-fr-dev\_00443]
- À deux reprises sont inscrites deux **initiales**<sub>Communication</sub>, avec, en face, des sommes. [Sequoia-frwiki\_50.1000\_00631]

## Relation

**Définition :** N qui dénote une relation non événementielle entre X et Y (ex. *hyponymie, amitié, incompatibilité, différence, équivalence*) ou du point de vue de X par rapport à Y (ex. *direction*)

### Exemples :

- L'ouragan Irma a ensuite pris la **direction**<sub>Relation</sub> d'Anguilla.
- Si les réformes envisagées par le gouvernement pour les européennes et les régionales sont adoptées telles qu'envisagées, la **parité**<sub>Relation</sub> hommes-femmes en politique risque de régresser. [Sequoia-annodis.er\_00351]
- La valeur ACT présente une **corrélation**<sub>Relation</sub> positive avec la dose et la concentration plasmatique de la bivalirudine. [Sequoia-emea-fr-dev\_00305]
- Quel est le point de **divergence**<sub>Relation</sub> entre la Commission et le Conseil ? [Sequoia-Euopar.550\_00216]

## Possession

**Définition :** N qui dénote des entités abstraites de type biens immatériels ou biens financiers

- € sommes d'argent / biens financiers (ex. salaire, bourse d'étude, budget, assurance-vie, (5) dollars, montant, somme (d'argent))
- € propriété intellectuelle (ex. royalties, copyright)
- € valeurs financières (taux de change, taux de crédit) dont les monnaies (dollar, yen)

### Exemples :

- Les clients fragiles continuent d'alimenter largement les banques en **commissions**<sub>Possession</sub> diverses.
- Une banque prélève chaque année en moyenne 34 euros de **frais**<sub>Possession</sub> pour incident de paiement sur l'ensemble de sa clientèle.
- Les propriétaires de **franchises**<sub>Possession</sub> de football américain (la NFL) sont invités par Trump à licencier les sportifs qui s'agenouillent pour manifester contre le racisme.
- Ce **montant**<sub>Possession</sub> passe à 296 **euros**<sub>Possession</sub> pour les clients fragiles.
- Une **manne**<sub>Possession</sub> à laquelle il est difficile de renoncer.
- Quant au **chèque**<sub>Artifact / Possession</sub> de 76.000 euros, la banque a apparemment refusé de l'honorer. [Sequoia-annodis.er\_00271]

## Phenomenon

**Définition :** N non dynamiques dénotant une entité qui s'observe à un moment donné, qui est soumise à une forme de temporalité mais qui n'a pas lieu (entre l'événement et l'objet)

- ⊆ forces naturelles (ex. gravité terrestre)
- ⊆ forces non intentionnelles (ex. chance)
- ⊆ manifestations sensorielles (ex. bruit)

### Exemples :

- *Irma s'accompagne de rafales de **vent**<sub>Phenomenon</sub> dépassant 300 km/h.*
- *À l'époque, je ne mesurais pas les **risques**<sub>Phenomenon</sub>.*
- *Puis, le jour J, dans le cadre de l'opération "ville, vie, vacances", les jeunes Maxévillois accompagnés des animateurs se rendront à Mont-Saint-Martin près de Metz où ils seront dans l'**obscurité**<sub>Phenomenon</sub> totale durant 2 mn 30. [Sequoia-annodis.er\_00132]*
- *Pour éviter une reprise du **feu**<sub>Phenomenon</sub> mais aussi pour prévenir des actes malveillants" constatait Martial Bourquin. [Sequoia-annodis.er\_00336]*
- *Le **bruit**<sub>Phenomenon</sub> des armes a trop longtemps couvert les cris de ceux qui ont faim. [Sequoia-Europar.550\_00333]*

## Time

**Définition :** N dénotant un espace temporel ou un repère

- ⊆ unités de temps (ex. minute, heure, jour, lustre, siècle, cycle, fois)
- ⊆ périodes (ex. époque, éternité, enfance, adolescence, vieillesse, jurassique, durée, vie, stade, phase, années 60, octobre, juin, mars, ramadan, moment)
- ⊆ date (ex. date, équinoxe, solstice, lundi, mardi)
- ⊆ repérage temporel (ex. lendemain, veille, soir, matin)
- ⊆ les parties d'une période ou d'une situation (début, fin, etc)

### Exemples :

- *Irma s'accompagne de rafales de vent dépassant 300 **km/h**<sub>Time</sub>.*
- *Selma, 27 **ans**<sub>Time</sub> le sait : son mariage est basé sur un mensonge.*
- *À l'**époque**<sub>Time</sub> je ne mesurais pas les risques.*
- *C'est l'un des **moments**<sub>Time</sub> de l'année où est publiée la plus grosse quantité de données.*
- *Tout au long des **années**<sub>Time</sub> de nouveaux sites furent créés, notamment à Hairoville, puis plus tard à Ancerville. [Sequoia-annodis.er\_00006]*
- *Elle occupait depuis cette **date**<sub>Time</sub> un bâtiment préfabriqué dans le parc Verneau. [Sequoia-annodis.er\_00028]*



## Tops

**Définition :** noms très généraux, très vagues sémantiquement, ne relevant pas a priori d'un domaine ontologique ciblé par les autres classes (entité, chose, fait, truc, machin, ...)

NB : sur la base du principe de l'annotation en contexte, c'est la référence qui prévaut dans l'annotation. L'étiquette Tops est donc en pratique rarement appliquée.

### Exemples :

- *Monsieur le Président, je voudrais souligner rapidement deux passages qui partent de la conscience du **fait**<sub>Tops</sub> que, souvent, l'Occident n'est pas seulement l'acteur, mais aussi le protagoniste des violations des droits de l'homme.*

## II. Les étiquettes multiples liées par un opérateur

Il est possible d'attribuer dans certains cas plusieurs étiquettes reliées par un opérateur lorsque le nom à annoter a un sens complexe. Trois cas de figures sont à prendre en compte, encodés par les différents opérateurs.

### L'opérateur de distribution

#### L'opérateur x

Cet opérateur exprime la distribution (A de B). Ici, A correspond à un nom en général relationnel -- de type Group (groupe de qqch), Part (partie de qqch) ou Quantity (quantité de qqch) -- dont l'élément lié (de B) est soit inclus dans le sens du nom cible (ex. *clientèle*), soit non exprimé mais présent dans le contexte (ex. *haut* ou *millions*).

- *Le **groupe**<sub>Group x Person</sub> se dispersa sans un mot. (vs le groupe<sub>Group</sub> des participants se dispersa)*
- *Une banque prélève chaque année en moyenne 34 euros de frais pour incident de paiement sur l'ensemble de sa **clientèle**<sub>Group x Person</sub>.*
- *Le **haut**<sub>Part x Artifact</sub> est tout tâché depuis que le chat a marché dessus. (vs Le haut<sub>Part</sub> du meuble est tout tâché depuis que le chat a marché dessus)*
- *Il a finalement annoncé la somme de 530 **millions**<sub>Quantity x Possession</sub>. (vs Il a finalement annoncé la somme de 530 millions<sub>Quantity</sub> d'euros)*

Tel que défini, cet opérateur encode donc deux phénomènes linguistiques différents : un phénomène lexical (ex. Le nom *clientèle* dénote toujours un ensemble de personnes) et un phénomène discursif, avec l'encodage du complément du nom (ou de l'adjectif relationnel) non

exprimé dans le contexte droit (ex. *Le groupe se dispersa sans un mot* = le groupe de personnes; *la somme de 530 millions* = millions de dollars/euros/...)

## Les cas de typages multiples

### L'opérateur /

Cet opérateur exprime la disjonction (A et/ou B). Il s'applique dans les cas où le contexte ne permet pas de décider si le nom renvoie à A, à B, ou à A et B. Il faut une hésitation entre au moins 2 des 3 étiquettes pour recourir à cet opérateur.

- *Son mariage*<sub>Act / State</sub> *repose sur un mensonge*<sub>Act / Cognition</sub>
- *La chouha*<sub>Act / Feeling</sub> (*"humiliation*<sub>Act / Feeling</sub> *" en arabe dialectal) c'est ce que subissent les Marocaines qui ont eu des rapports sexuels avant le mariage.*
- *La coagulation sanguine peut poser un problème en cas de perturbation*<sub>Act / Event</sub> *du flux sanguin.* [Sequoia-emea-fr-dev\_00030]
- *Les événements indésirables sont listés par classe de système d'organes dans le tableau*<sub>Artifact / Cognition</sub> 4. [Sequoia-emea-fr-dev\_00237]

### L'opérateur +

Cet opérateur exprime la conjonction (A et B). Il s'applique au cas où, dans le contexte, l'occurrence du nom renvoie explicitement à A et à B. Cet opérateur encode donc l'emploi contextuel dans lequel les différentes facettes du N sont explicitement mobilisées (coprédication). Par exemple pour *étude* ci-dessous, le sens Act est explicité par la séquence *menée par* et le sens Cognition est explicité par *selon*. Pour *humiliation*, le sens Act est convoqué par *subir*, le sens Feeling est convoqué par *ressentir*.

- *Selon une étude*<sub>Act + Cognition</sub> *menée par l'UNAF, une banque prélève chaque année en moyenne 34 euros de frais pour incident de paiement.*
- *Le livre*<sub>Artifact + Cognition</sub> *qui est sur la table est très bien écrit.*
- *Cette humiliation*<sub>Act + Feeling</sub> *qu'il a subi à trois ans, il la ressent encore.*
- *Au mois de mars dernier, les enfants de CE2 de Mme Philippe avaient montré leur travail*<sub>Artifact + Cognition</sub> *réalisé d'après "Les contes du chat perché" de Marcel Aimé.* [Sequoia-annodis.er\_00146]

**Nota Bene 1** : pour les opérateurs + et / (et pour eux uniquement) les étiquettes doivent être classées par ordre alphabétique. Dans WebAnno, l'opérateur est sélectionné en 3e. On aura ainsi Event|State|/ ou encore Act|Substance|+

**Nota Bene 2** : l'opérateur / ne doit pas être utilisé dans les cas où l'on hésite entre plusieurs étiquettes, parce que le N est difficile à annoter ou que le guide n'est pas explicite sur un point. Il

est réservé aux cas où le N peut avoir deux sens ou deux facettes et que le contexte ne permet pas de déterminer s'il s'agit de l'une, de l'autre ou des deux.

**Nota Bene 3 :** l'étiquette Cognition (pour objet informationnel) correspond souvent à l'une des deux étiquettes en cas de typage multiple (*étude, conférence, livre, accusation*, etc). La frontière entre les noms ayant dans leur sens une facette Cognition et ceux qui n'en ont pas n'étant pas toujours facile à tracer, les noms pour lesquels on soupçonne une facette Cognition doivent passer au moins l'un des trois tests indicatifs de la classe (voir supra Cognition), sans quoi on optera pour un typage simple.

### III. Annotation des expressions polylexicales

Les éléments du corpus correspondant à des expressions polylexicales (EP) ou à des entités nommées (EN) ont fait l'objet d'un pré-repérage dans le cadre du projet du projet Parseme-fr (<http://parsemefr.lif.univ-mrs.fr/doku.php?id=start>). Les EN ont déjà reçu une étiquette sémantique dans le cadre de ce projet et les étiquettes utilisées étant de même granularité que celle utilisées dans Fr-Semcor (Person, Event, etc), les EN ne doivent pas être annotées ici.

Pour FR-Semcor, seuls les EP dont la tête correspond au nom cible (le nom à annoter) doivent être annotés. Comme on le voit dans les exemples ci-dessous, l'étiquette associée au nom tête correspond à la classe sémantique de l'EP prise dans son ensemble.

- *Par le biais de l'activité théâtre, ils se sont intéressés à l'expression orale et à la **[mise<sub>Act</sub> en scène]<sub>EP</sub>** avec un intervenant, Julien Georges, maître de théâtre.*
- *On continue à optimiser nos **[points<sub>Attribute</sub> forts]<sub>EP</sub>** et à travailler nos faiblesses.*

**En résumé, ne doivent pas être annotés :**

1. Les noms qui apparaissent dans des EN;
2. Les noms qui apparaissent dans des EP non nominales, tels que *effet* dans *en effet* ou *action* dans *passer à l'action*;
  - *L'automobile, **[en effet]<sub>EP</sub>** est invisible depuis la rue de Danjoutin.*
  - *Je voudrais souligner, comme l'a fait le précédent orateur, la nécessité de **[passer à l'action]<sub>EP</sub>** de développer la coopération entre les polices et d'augmenter l'information dans ce domaine.*
3. Les noms qui ne correspondent pas à la tête dans les EP nominales, tels que *affaire* dans *homme d'affaire* ou *administration* dans *mode d'administration*;

- *Alfred Sirven, bras droit de Loïk Le Floch-Prigent (président du groupe pétrolier Elf), propose aux dirigeants de Thomson les services du "réseau Elf" pour jouer les intermédiaires avec Pékin par le biais d'Edmond Kwan, un [homme d'affaires]<sub>EP</sub> chinois.*
- *Posologie et [mode d'administration]<sub>EP</sub>.*