

Guide d'annotation Demonext

Typage sémantique de lexèmes nominaux hors contexte

Projet Demonext *Dérivation en extension* ANR-17-CE23-0005

Délivrable 1.1 *Typage sémantique pertinent en morphologie* - Lot 1 Traitements linguistiques

Mathilde Huguin, IGR CNRS Université de Lorraine, ATILF UMR 7118

Lucie Barque, MCF Université Paris 13, LLF UMR 7110

Pauline Haas, MCF Université Paris 13, LaTTiCe UMR 8094 et PLEIADE EA 7338

Fiammetta Namer, Prof. Université de Lorraine, ATILF UMR 7118 (Directrice scientifique du projet Demonext)

Delphine Tribout, MCF Université de Lille, STL UMR 8163

Résumé du projet Demonext

Demonext consiste en la construction d'une base de données morphologiques (BDM) du français qui décrit les propriétés dérivationnelles des mots de manière systématique. La BDM répondra à des besoins multiples, comme la confirmation empirique et l'élaboration d'hypothèses en morphologie, le développement d'outils en traitement automatique des langues (TAL), l'enseignement du vocabulaire et le traitement des troubles du langage développementaux ou acquis.

Demonext est un projet financé par l'Agence Nationale de la Recherche (ANR), sous la référence **ANR-17-CE23-0005**.



Table des matières

1	Avant propos	1
2	Démarche	1
3	Présentation de la typologie	2
4	Étiquette superordonnée : Tops	3
5	Entités concrètes	4
5.1	Entités animées	4
5.1.1	Animal	5
5.1.2	Person	5
5.2	Entités inanimées	5
5.2.1	Artifact	5
5.2.2	Body	6
5.2.3	Food	6
5.2.4	Object	6
5.2.5	Plant	6
5.2.6	Substance	7
6	Entités abstraites	7
6.1	Cognition	7
6.2	Communication	8
6.3	Group	8
6.4	Institution	8
6.5	Part	8
6.6	Phenomenon	9
6.7	Possession	9
6.8	Quantity	10
6.9	Time	11
7	Situations dynamiques	11
7.1	Act	11
7.2	Event	12
8	Situations non dynamiques ou propriétés	12
8.1	Attribute	12
8.2	Feeling	13
8.3	State	13
9	Catégories hybrides	14
9.1	Act+Cognition	14
9.2	Act+Possession	15
9.3	Act+State	15
9.4	Artifact+Cognition	15
9.5	Event+Object	16
9.6	Event+Phenomenon	16
9.7	Event+State	16
	Références	17

1 Avant propos

Le présent guide est un outil destiné aux linguistes et aux étudiants permettant l'annotation sémantique manuelle de lexèmes nominaux hors contexte en prenant en compte la polysémie (cf. §3).

Plus précisément, ce document récapitule la démarche adoptée pour l'annotation sémantique des noms de la table des lexèmes de Demonext (Namer et al., 2019). Il est essentiellement basé sur les travaux de Barque et al. (2020) et Haas et al. (à paraître) (voir aussi les travaux de sémantique nominale en français comme Flaux and Velde, 2000, Kleiber and Lammert, 2012, Tribout et al., 2014, Fasciolo and Gross, 2014, Huyghe, 2015). Il regroupe ainsi des définitions, des exemples et des tests permettant d'assigner l'une ou l'autre des catégories sémantiques de la section 3 à un nom. Chaque catégorie fait l'objet d'une section dans la suite (cf. §4 à §9).

2 Démarche

La première étape du travail a consisté à définir le jeu d'étiquettes pour l'annotation sémantique des noms de la base. Nous sommes parties d'un jeu d'étiquettes existant, adapté des *Unique Beginners* de Wordnet (Miller et al., 1990, Fellbaum, 1998) et utilisé pour le français dans le cadre d'un projet d'annotation en corpus (Barque et al., 2020). Ce jeu, finalement constitué de 22 étiquettes simples et de 21 étiquettes complexes, présente un grain suffisant pour décrire les sens construits par la morphologie et permet de capter l'essentiel des alternances de sens observées au sein du lexique construit (par ex. l'alternance `PERSON`; `ARTIFACT` pour les dérivés en *-eur* comme *codeur*, ou encore l'alternance `ACT`; `STATE` pour dérivés en *-ment*, comme *abaissement*).

Une fois stabilisé le jeu d'étiquette, un travail d'appariement avec les bases morphologiques existantes incluant des informations sémantiques (Demonette-1¹, Dimoc², Denom³, Convert⁴ et Mordan⁵) a été effectué. Par exemple, les noms typés AGF (agent féminin) ou AGM (agent masculin) dans Demonette-1 se sont vus attribuer automatiquement l'étiquette `PERSON`. Au total, 28 204 noms issus des ressources disposant d'informations sémantiques se sont vus attribuer, sur la base de cet appariement, une ou plusieurs étiquettes normalisées. Certains noms n'ont reçu aucune étiquette, le type proposé dans le lexique source regroupant des noms sémantiquement trop hétérogènes pour être appariés de manière automatique. Parmi les noms ayant reçu plusieurs étiquettes, on distingue les cas où la forme est effectivement polysémique et apparaît donc avec des types distincts dans les lexiques sources (par ex. *administrariat*, typé "activité" et "collectif humain" dans Dimoc), des cas où l'étiquette proposée dans le lexique source est sous-spécifiée, donnant lieu à une polysémie le plus souvent artificielle (par ex. comme les types AGM et AGF dans Demonette-1 regroupent les agents et les instruments, les noms typés AGM et AGF reçoivent les deux étiquettes `PERSON` et `ARTIFACT`). 7 437 appariements automatiques ont été validés et / ou corrigés manuellement.

La troisième et dernière étape a consisté à annoter manuellement un ensemble de noms de la base. Une première annotation-test, effectuée en double aveugle sur un échantillon de 145 noms (intersection des 5 bases morphologiques citées ci-dessus), a permis d'identifier les difficultés inhérentes à l'annotation lexicale (notamment la question de savoir quels sens du nom décrire) et d'affiner le guide d'annotation : les frontières entre classes sémantiques ont été précisées afin d'éviter des doubles étiquetages qui ne relèveraient pas de polysémies réelles, des tests d'identification des types et des exemples illustratifs permettant les raisonnements par analogie ont été ajoutés pour guider l'annotation. Après cette annotation test, les annotations des 2 674 noms du corpus FrSemcor figurant dans la base Demonette ont été reprises, révisées et complétées pour décrire extensivement leur polysémie. Enfin, 5 000 noms du lexique Échantinom (Bonami and Tribout, 2021) ont été annotés.

En tout, 15 112 (7 437 + 145 + 2 674 + 5 000) noms de la base auront donc été annotés manuellement.

1. Voir le travail de Hathout and Namer (2014).

2. Voir les travaux de Roché (1997, 2006, 2007, 2011), entre autres.

3. Voir la thèse de Strnadová (2014).

4. Voir le thèse de Tribout (2010).

5. Voir la thèse de Koehl (2012).

3 Présentation de la typologie

La typologie comporte 22 catégories *simples* et 21 catégories *complexes*, cf. figure 1 et tableau 1. Les catégories simples s'organisent en une hiérarchie ontologique. Les catégories complexes sont celles faisant intervenir deux étiquettes simples et un opérateur x ou +, dont la signification est donnée dans la suite.

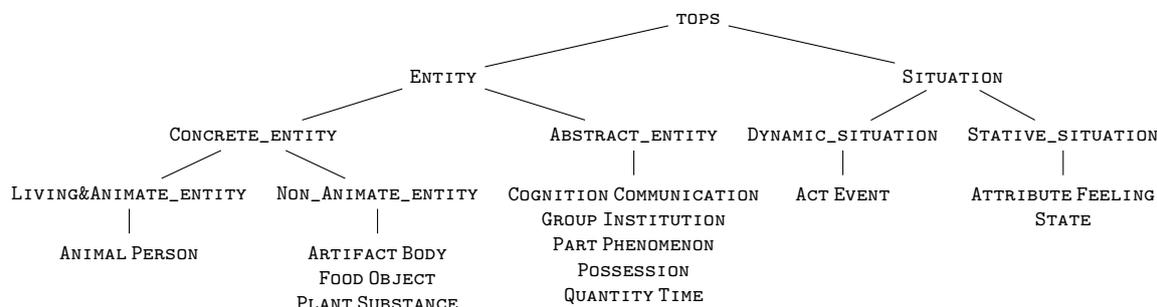


Figure 1 - Hiérarchie des catégories simples

ACT+COGNITION, ACT+POSSESSION, ACT+STATE, ARTIFACT+COGNITION, EVENT+OBJECT, EVENT+PHENOMENON, EVENT+STATE
GROUPXABSTRACT_ENTITY, GROUPXACT, GROUPXANIMAL, GROUPXARTIFACT, GROUPXCONCRETE_ENTITY, GROUPXENTITY, GROUPXPERSON, GROUPXPLANT, PARTXARTIFACT, PARTXCOGNITION, PARTXCOMMUNICATION, PARTXOBJECT, PARTXPHENOMENON, PARTXSUBSTANCE

Tableau 1 - Étiquettes complexes

Chaque nom se voit attribuer une ou plusieurs de ces étiquettes en fonction de son (ou ses) sens. Seuls les sens institutionnalisés et non vieillis sont annotés. Par exemple, le nom *verre* se voit attribuer les étiquettes *ARTIFACT* et *SUBSTANCE* mais aussi *FOOD*, car la métonymie est signalée comme lexicalisée dans le dictionnaire TLFi. Voici les configurations possibles selon le(s) sens du nom et les opérateurs utilisés.

Le nom a un seul sens

Si le nom est monosémique, il se voit attribuer une étiquette simple ou une étiquette complexe manipulant l'opérateur x. Par exemple *bavard* est annoté *PERSON*. Les étiquettes *GROUP* ('un ensemble de') et *PART* ('une partie de') peuvent être associées à une autre étiquette simple avec l'opérateur x qui exprime la distribution (A de B). Ainsi, le lexème *fôret* est annoté *GROUPXPLANT*. Certains lexèmes, avec des sens plus génériques, nécessitent de remonter dans la hiérarchie ontologique, cf. figure 1. Par exemple, le sens de *livraison* 'chose livrée' correspond à l'étiquette *ENTITY* qui regroupe les entités concrètes animées, non animées et les entités abstraites.

Le nom a plusieurs sens

Si le nom possède plusieurs sens, nous distinguons les cas où les sens sont compatibles des cas où les sens sont incompatibles ou antagonistes en discours.

Les noms à sens multiples compatibles, dits *hybrides*, admettent la co-prédication, cf. (1). Ils correspondent aux *noms à facettes* (Cruse, 1995, 2004) ou à *types pointés* (Pustejovsky, 1996) et se voient attribuer une étiquette complexe intégrant l'opérateur + qui exprime la conjonction (A et B). En (1a), le locuteur se focalise sur le contenu du livre (A), en (1b) sur sa matérialité (B), et (1c) concilie les deux lectures (A et B), le prédicat aimer pouvant viser les deux facettes (Abrard and Stosic, 2021). Le lexème nominal *livre* est donc annoté *ARTIFACT+COGNITION*.

(1) Sens compatibles

- a. Ce **livre** est passionnant. [COGNITION]
b. Ma mère m'a acheté un **livre** magnifiquement relié. [ARTIFACT]
c. J'**aime** ce livre. [ARTIFACT+COGNITION]

Les noms à sens multiples incompatibles sont ceux dont la co-prédication provoque un effet zeugmatique (noté *), ils sont qualifiés de *polysémiques* (Apresjan, 1974, Kleiber, 1999, Barque, 2008), cf. (2). Les noms polysémiques se voient attribuer plusieurs étiquettes simples ou complexes, rangées par ordre alphabétique, séparées par le symbole ;. Par exemple, *manucure* (2) est annoté ACT ; PERSON. Si les sens polysémiques correspondent à la même étiquette, la polysémie est invisible dans la base. C'est par exemple le cas pour la polysémie du nom *bureau*, cf. (3). Il en est de même pour une éventuelle homonymie : si deux lexèmes partagent la même catégorie syntaxique et la même étiquette sémantique, l'homonymie est invisible.

(2) Sens incompatibles, polysémie visible dans la base

- a. La **manucure** réalisée par Charline a duré moins de 15 minutes. [ACT]
b. Marie est **manucure**. [PERSON]
c. *La **manucure**, qui est Marie, a été effectuée par Charline. [incompatibilité]

(3) Sens incompatibles, polysémie invisible dans la base (Haas et al., à paraître)

- a. Il a acheté un **bureau** en acajou avec plusieurs tiroirs. [ARTIFACT]
b. Son **bureau** dispose d'une fenêtre. [ARTIFACT]
c. *Son **bureau**, qui est en acajou avec plusieurs tiroirs, dispose d'une fenêtre. [incompatibilité]

Dans la suite, nous présentons chaque catégorie sémantique en donnant :

- une définition générale de la catégorie et quelques précisions sur les frontières ontologiques du domaine ciblé ;
- le cas échéant, un ou plusieurs test permettant d'attribuer l'étiquette correspondante à un lexème nominal ; il est à noter que certains tests sont conjonctifs (*et*), c'est-à-dire que pour être considérés comme positifs, le nom testé doit entrer dans l'ensemble des structures distributionnelles, et que d'autres sont disjonctifs (*ou*) ;
- un exemple d'application du ou des tests proposés ;
- des exemples de lexèmes nominaux ayant reçu l'étiquette sémantique associée ; il faut souligner que les noms choisis peuvent être polysémiques.

4 Étiquette superordonnée : Tops

TOPS : nom très général, très vague sémantiquement (voir à ce propos Huyghe, 2021), ne relevant pas *a priori* d'un domaine ontologique ciblé par les autres classes. Sur la base du principe de l'annotation hors contexte, cette étiquette est plus fréquemment attribuée que lors d'une annotation contextualisée (cf. le travail de Barque et al., 2020). Par exemple, en dehors d'un énoncé, on ne peut pas déterminer le type référentiel ou ontologique du lexème *truc*.

Exemples. *affaire, aide, antipode, archétype, base, chose, doublon, découverte, entité, fait, machin, objet, paramètre, parangon, phénomène, prototype, surprise, symbole, truc, type, variante.*

5 Entités concrètes

CONCRETE_ENTITY : entité accessible aux sens, dotée d'une extension temporelle et spatiale (Flaux and Velde, 2000). Les étiquettes utilisées pour les noms dénotant des entités concrètes sont ANIMAL, ARTIFACT, BODY, FOOD, OBJECT, PERSON, PLANT ou SUBSTANCE mais certains lexèmes nominaux, dont le sens est plus général, sont annotés CONCRETE_ENTITY (*détritus*).

👉 Les entités concrètes valident les tests 1.

Tests 1 : Concrete_entity

- 1.1. **Dét N** se trouve + complément de localisation spatiale.^a
et
- 1.2. **Dét N** de + mesure.^b [pour les N comptables]
- 1.3. mesure^c + **de N**. [pour les N massifs]
- 1.4. **Dét N** de Dét.card N2.^d [pour les N collectifs]

a. ex. : **Dét N** se trouve sur la table, à côté du sac, dans le jardin, à Paris, entre la table et le mur, en Europe, près du poumon droit, etc..

b. ex. : **Dét N** de deux mètres de large/vingt mètres carrés/trois mètres cubes/cent hectares/cinq cents grammes/dix kilos/trente tonnes.

c. ex. : deux mètres/vingt mètres carrés/trois mètres cubes/cent hectares/cinq cents grammes/dix kilos/trente tonnes **de N**.

d. ex. : **Dét N** de quinze chèvres/loups/personnes/timbres/pierres.

Exemple d'application des tests 1

- 1.1. **La briquette** se trouve dans le jardin./ **La farine** se trouve dans la remise./ **Le troupeau** se trouve dans la forêt.
et
- 1.2. **Une briquette** de cinq cents grammes. [pour les N comptables]
- 1.3. 10 kg **de farine**. [pour les N massifs]
- 1.4. **Un troupeau** de 15 loups. [pour les N collectifs]

5.1 Entités animées

👉 Les entités animées (ANIMAL et PERSON) valident les tests 2.

Tests 2 : Living&Animate_entity

- 2.1. **Dét N** décide/choisit de V.
et
- 2.2. **Dét N** + SV volontairement/délibérément/consciemment.
et
- 2.3. **Dét N** a beaucoup mangé.

Exemple d'application des tests 2

- 2.1. **Le bûcheron** décide de rentrer chez lui.
et
- 2.2. **Le bûcheron** rentre volontairement.
et
- 2.3. **Le bûcheron** a beaucoup mangé.

5.1.1 Animal

ANIMAL : être animé non humain. Inclut les insectes (*gendarme*). N'inclut pas les bactéries et les virus (annotés OBJECT), les parties du corps des animaux (annotés BODY) la viande ou chair de l'animal (annotée FOOD).

☞ Répond positivement aux tests 1 et 2.

Exemples. *agnelet, aiguillette, bichon, bourriquet, bovin, bête, canard, castor, chardonneret, cousin, dauphin, gendarme, habitant, hérisson, lapinet, lapinet, pacha, pape, poussinnet, primate.*

5.1.2 Person

PERSON : humain et personnage imaginaire ayant des attributs humains saillants comme les dieux et les personnages mythologiques (à propos des noms d'humains voir Gross, 1995a, Schnedecker and Mihatsch, 2018). Les métonymies lexicalisées sont également annotées PERSON (*baryton*).

☞ Répond positivement aux tests 1 et 2.

Exemples. *Tonton, adolescent, anonyme, avocat, catéchiste, flemmard, noctambule, parent, paysan, philosophe, pochard, préfet, sacristain, sous-commissaire, spectateur, traînée, vice-doyen.*

5.2 Entités inanimées

Les entités inanimées sont annotées ARTIFACT, BODY, FOOD, OBJECT, PLANT, SUBSTANCE. Une partie d'entre elles valide les tests 3.

Tests 3 : Non_Animate_entity

- 3.1. N[humain] a fabriqué/déchiré **Dét N** à tel endroit/à tel moment.
ou
- 3.2. N[humain] a construit **Dét N** + Compl de matière + Adj de couleur.
ou
- 3.3. N[humain] a confectionné **Dét N** + Compl. d'arôme.

Exemple d'application des tests 3

- 3.1. Paul a déchiré **l'affiche** rue de Paris.
ou
- 3.2. Paul a construit **un aviron** en bois noir.
ou
- 3.3. Paul a confectionné **un gâteau** au chocolat.

5.2.1 Artifact

ARTIFACT : objet physique non animé et fabriqué. Inclut les bâtiments ou partie de bâtiment (*immeuble, vestibule*), les territoires construits (*banlieue*) sauf les zones géographiques délimitées administrativement (*commune*) qui reçoivent l'étiquette INSTITUTION. Inclut les formes (*décagone, octogone*). N'inclut pas les médicaments (SUBSTANCE).

☞ Répond positivement aux tests 1 et 3.1. et 3.2. mais négativement aux tests 2 et 3.3.

Exemples. *audiomètre, automobile, aviron, baron, caniveau, cantine, cavalier, correcteur, dent, écran, fibroscope, harpe, orgues, pipette, plafond, protection, rosace, roue, sculpture, téléphone.*

5.2.2 Body

BODY : corps ou élément du corps (*dos*) ou organe des humains (*estomac*) ou des animaux (*épine*) (voir à ce propos Simatos, 2015, Abeillé and Godard, 2021). N'inclut pas les fluides corporels (annotés **SUBSTANCE**).

☞ Répond positivement aux tests 1 mais négativement aux tests 2 et 3.

Exemples. *aile, aine, articulation, bassin, bois, bouche, boule, bras, coeur, colonne, cou, dent, doigt, dos, estomac, face, fémur, hippocampe, hémisphère, index, lune, lune, menotte, muscle, poitrine.*

5.2.3 Food

FOOD : aliment naturel ou préparé (*buffet, civet, dîner, sauce*). Inclut les boissons, les aliments ou tout ce qui entre dans la composition d'un aliment, les fruits, les légumes, les céréales, les viandes et chairs animales. Les métonymies lexicalisées, du type *ballon* ou *verre*, sont également annotées **FOOD**.

☞ Répond positivement aux tests 1 et parfois 3.3. mais négativement aux tests 2.

Exemples. *absinthe, agrume, alcool, aliment, ananas, arachide, avocat, baguette, biscuit, boisson, canon, chinois, citron, escargot, fraise, fève, opéra, piment, tartare, tartine, tomate.*

5.2.4 Object

OBJECT : objet naturel, non créé par l'homme. Inclut les corps célestes (*soleil*), les cours d'eau (*rivière*), les objets minéraux (*rocher, fossile*), les microbes et les bactéries (*miasme*), les territoires naturels et les milieux environnementaux (*antarctique, jungle, steppe, tundra*). Inclut aussi les résultats ou conséquences de certaines maladies (*tumeur, lésion*). N'inclut pas les noms de substances naturelles (*sable, pluie*), classées dans **SUBSTANCE**.

☞ Répond positivement aux tests 1 mais négativement aux tests 2 et 3.

Exemples. *anion, antre, astre, atmosphère, biotope, caillou, continent, cosmos, désert, ecchymose, embryon, fibrose, galaxie, ganglion, globe, golfe, récif, satellite, tropique, équateur.*

5.2.5 Plant

PLANT : entité du règne végétal ou une de ses parties. Inclut les champignons, les fruits, légumes et céréales lorsqu'ils désignent la plante (*maïs*). Ces derniers sont également annotées **FOOD** lorsque le même lexème désigne la plante et l'aliment comestible (*cèpe*). N'inclut pas les substances issues d'une entité végétale (*sève, pollen*), classées dans **SUBSTANCE**. Les arbres dont le bois est utilisé comme matériau sont également annotés **SUBSTANCE**.

☞ Répond positivement aux tests 1 et éventuellement 3.1. mais négativement aux tests 2.

Exemples. *algue, anémone, arbre, aromate, baie, brin, broussaille, caoutchouc, champignon, chrysanthème, chêne, citronnelle, cosmos, feuille, fleur, gentiane, pelouse, plante, racine, tige.*

5.2.6 Substance

SUBSTANCE : matière physique, liquide, solide, gazeuse ou composé chimique fabriqué ou non. Inclut les médicaments (*sirop, aspirine*), les fluides corporels (*sang, larme*), les substance végétales (*sève, bois*). N'inclut pas les substances consommables ou alimentaires annotées **FOOD** et les microbes annotés **OBJECT**.

☞ Répond positivement aux tests 1 mais négativement aux tests 2 et 3.

Exemples. *acide, acrylique, agent, air, aluminium, ammoniac, anticoagulant, antidote, argent, arsenic, aspirine, bain, bois, calcium, caoutchouc, carburant, chromosome, essence, fluide, synovie.*

6 Entités abstraites

Les entités abstraites, annotées **COGNITION**, **COMMUNICATION**, **GROUP**, **INSTITUTION**, **PART**, **PHENOMENON**, **POSSESSION**, **TIME**, répondent négativement aux tests 1.

6.1 Cognition

COGNITION : objet mental (*idée*) ou informationnel. Inclut les genres littéraires ou musicaux (*théâtre, métal*). Inclut les domaines de connaissance (*astrologie*).

☞ Répond positivement aux tests 4.

Tests 4 : Cognition

- 4.1. **Dét N** selon lequel P est difficile à admettre.
ou
- 4.2. N[humain] écoute **Dét[plur] N** décousus de Vincent.
ou
- 4.3. N[humain] a écrit **Dét N** intéressant/pertinent.
ou
- 4.4. N[humain] a reçu une formation dans le domaine **du/de la N**.

Exemple d'application des tests 4

- 4.1. **L'idée** selon laquelle P est difficile à admettre.
ou
- 4.2. Paul écoute **les propos** décousus de Vincent.
ou
- 4.3. Paul a écrit **un livre** intéressant.
ou
- 4.4. Paul a reçu une formation dans le domaine **de la lexicologie**.

Exemples. *architecture, doctrine, iconographie, idéalisme, idée, idéocratie, idéologie, intrapsychologie, islam, jésuitisme, lexicologie, loi, macroéconomie, médecine, prélude, théologie.*

6.2 Communication

COMMUNICATION : objets linguistiques (*langue*), métalinguistiques (*prédictat*) et médium de communication (*internet*).

Exemples. *anglicisme, dialecte, dialogue, hiatus, idiome, idiotisme, idéographie, imparisyllabe, leit-motif, lemme, lexème, lipogramme, litote, logogramme, nom, occitan, préfixe, quodlibet.*

6.3 Group

GROUP : nom au singulier dénotant un ensemble d'entités (*famille*) ou nom toujours au pluriel dénotant un groupe (*gens*). N'inclut pas les groupes de personnes d'une INSTITUTION, n'inclut pas la numération qui reçoit l'étiquette QUANTITY.

Exemples. *amas, bande, batterie, bloc, brigade, catégorie, cellule, cercle, combinaison, communauté, cortège, dyade, ensemble, famille, file, galaxie, gamme, groupe, système, taxinomie.*

6.4 Institution

INSTITUTION : nom dénotant une institution (*rectorat*), un pays (*Mongolie*), un organisme (*banque, université*). Inclut aussi les membres de l'institution et le lieu constitutif de l'institution : il n'y a pas de double typage, ou de type complexe, incluant INSTITUTION, ARTIFACT et GROUPXPERSON.

☞ Répond positivement aux tests 5.

Tests 5 : Institution

- 5.1. **Dét N** a été fondé en telle année.
ou
5.2. N[humain] a été nommé à la tête de **(Dét) N**.

Exemple d'application des tests 5

- 5.1. **Le sénat** a été fondé en 1958.
ou
5.2. Mr Dupont a été nommé à la tête **du sénat**.

Exemples. *académie, collège, confrérie, confrérie, émirat, fisc, Grande-Bretagne, Guatemala, Italie antique, Mongolie, paroisse, patrie, Prusse, rectorat, Sodome, sénat, thalassocratie, tétarchie.*

6.5 Part

PART : nom qui dénote une partie de quelque chose. Inclut les noms de localisation interne, les noms de partie positionnée (*côté, extrémité, milieu, sommet*), les point cardinaux (*est*), les parties non spécifiées, non pré-délimitées et non fonctionnelles d'un tout (*rondelle, morceau, tranche, partie*). N'inclut pas les méronymes, les expressions de quantité (*proportion*) et les parties du temps (*minute*) qui sont annotées TIME.

Exemples. *bande, bas, base, bord, bouche, bout, brin, cellule, centre, chapitre, coeur, composant, contrebas, coupe, côté, dessous, droite, début, échantillon, segment, vestige, zone.*

6.6 Phenomenon

PHENOMENON : entité qui s'observe à un moment donné, qui est soumise à une forme de temporalité mais qui n'a pas lieu. Inclut les forces naturelles, les forces non intentionnelles et les manifestations sensorielles comme les odeurs (*fragrance*), les sons (*cliquetis*) ou les phénomènes lumineux (*lueur*).

👉 Répond positivement aux tests 6.

Tests 6 : Phenomenon

- 6.1. On a entendu **Dét N** tonitruant/strident/mélodieux/aigu/grave.
et
- 6.2. **Dét N** a retenti/a résonné/s'est propagé au fond du couloir/dans la cour.
ou
- 6.3. On a vu **Dét N** aveuglant/éblouissant/blafard/pâle.
et
- 6.4. **Dét N** a resplendi/s'est propagé/a scintillé/a jailli au fond du couloir.
ou
- 6.5. On a senti **Dét N** nauséabond/âcre/enivrant/tenace/entêtant.
et
- 6.6. **Dét N** s'est répandu/persiste/embaume (dans) la pièce.

Exemple d'application des tests 6

- 6.1. On a entendu **un bruit** strident.
et
- 6.2. **Un bruit** s'est propagé dans la cour.
ou
- 6.3. On a vu **une lueur** aveuglante.
et
- 6.4. **Une lueur** a jailli au fond du couloir.
ou
- 6.5. On a senti **un arôme** entêtant.
et
- 6.6. **Un arôme** embaume la pièce.

Exemples. *arôme, bruit, brume, chance, climat, destin, épiphénomène, feu, flux, hasard, lueur, lumière, musique, nuisance, obscurité, phénomène, pression, pénombre, signe, tension, ton.*

6.7 Possession

POSSESSION : entité abstraite de type bien immatériel ou financier. Inclut les sommes d'argent (*gain*), les biens financiers (*dote*), les propriétés intellectuelles (*licence*), les valeurs financières dont les monnaies (*euro*). Le double typage pour les monnaies est systématique (ARTIFACT).

👉 Répond positivement aux tests 7.

Tests 7 : Possession

- 7.1. Il a versé **Dét N** en euros.
ou
- 7.2. Il a obtenu **Dét[indéf, sing] N** modique.
ou
- 7.3. **Dét N** est la monnaie de tel pays.
et
- 7.4. Quel est le taux de change de **Dét N** ?

Exemple d'application des tests 7

- 7.1. Il a versé **son loyer** en euros.
ou
- 7.2. Il a obtenu **un gain** modique.
ou
- 7.3. **L'euro** est la monnaie du Luxembourg.
et
- 7.4. Quel est le taux de change de **du dollar** ?

Exemples. *brevet, copyright, dollar, dot, dotation, dépense, euro, finance, frais, franc, gain, impôt, indemnité, licence, loyer, marge, montant, moyen, participation, patrimoine, pension, prix.*

6.8 Quantity

QUANTITY : objet abstrait numéral. Inclut les quantités pures, nombres (*pluralité, nombre, score*), les unités de mesure (*tonne, mètre*), les contenants employés comme unité de mesure (*verre, cuillère*), les proportions (*majorité, quart, moitié, pourcent*), les ordinaux (*premier, second*), les quotients et produits, les marqueurs de quantité (*coups*). N'inclut pas les unités de mesure de temps (*minute, heure*) qui reçoivent l'étiquette **TIME**.

☞ Répond positivement aux tests 8.

Tests 8 : Quantity

- 8.1. Dét courant/poids/lumière/intensité/tension/distance/période/vitesse de **Dét.card N**, c'est peu/beaucoup.
et
- 8.2. Dét N_0 mesure/fait/pèse/dure **Dét.card N**.
ou
- 8.3. **Dét N** de N_2^a , c'est peu/beaucoup.

a. ex. : Dét N de farine, huile, sable, messages, images, cheminées, projectiles, juristes.

Exemple d'application des tests 8

- 8.1. Une distance de **120 kilomètres**, c'est beaucoup.
et
- 8.2. Les cendres de moteurs Diesel mesurent **5 microns**.
ou
- 8.3. **Une mégatone** de farine, c'est beaucoup.

Exemples. *antilogarithme, barème, calorie, centimètre, farad, intervalle, kilomètre, logarithme, micron, millimètre, mégatonne, nanomètre, once, plasmon, pléthore, période, quanta, un.*

6.9 Time

TIME : espace ou repère temporel. Inclut les unités de temps (*minute*), les périodes (*préavis*), les dates (*avril*). Inclut aussi les mesures musicales (*tempo*).

☞ Répond positivement aux tests 9.

Tests 9 : Time

- 9.1. Ils se sont rencontrés **N/courant N/Dét N/Dét.card N durant/à N.**
ou
9.2 . Attendez Dét[indéf, sing] bref **N.**

Exemple d'application des tests 9

- 9.1. Ils se sont rencontrés **courant septembre.**
ou
9.2 . Attendez un bref **moment.**

Exemples. *an, anniversaire, année, août, apogée, aurore, avenir, avril, crétacé, date, dimanche, début, décennie, fin, holocène, Mars, matin, moment, passé, saison, septembre, shabbat, solstice, vie.*

7 Situations dynamiques

Les situations dynamiques, annotées ACT ou EVENT, répondent positivement à un des tests 10 (voir par exemple Haas and Gréa, 2015, Haas, 2017).

Tests 10 : Dynamic situation

- 10.1. **Dét N** a eu lieu/s'est produit à tel moment/à tel endroit.
ou
10.2. N[humain] a procédé à/a accompli/a effectué **Dét N** (+ expansion facultative).

Exemple d'application des tests 10

- 10.1. **L'amélioration** s'est produit au milieu du mois.
ou
10.2. Paul a effectué **un ajustement.**

7.1 Act

ACT : situation dynamique agentive intentionnelle ou non, incluant les actions agentives (*achat*) et les activités (*football*). Inclut aussi les domaines ou secteurs d'activités (*tourisme*).

☞ Répond positivement aux tests 10 et 11.

Tests 11 : Act

- 11.1. Paul a effectué/a procédé à/a accompli/a commis (volontairement, insidieusement) **Dét N** (+ exp.).
ou
- 11.2. **Dét N** délibéré.
ou
- 11.3. La poursuite de **Dét N** (avec le sens de qqn poursuit le N).
ou
- 11.4. N[humain] fait **du/de la N**.

Exemple d'application des tests 11

- 11.1. Paul a effectué **une angiographie**.
ou
- 11.2. **Une annulation** délibérée.
ou
- 11.3. La poursuite de **l'acheminement**.
ou
- 11.4. Paul fait **du tourisme**.

Exemples. *acheminement, adoption, affrontement, agissement, agriculture, allaitement, amélioration, annulation, applaudissement, arbitrage, artisanat, assassinat, inquisition, procédure.*

7.2 Event

EVENT : situation dynamique sans agent. Inclut les événements naturels (*avalanche*), les autres événements fortuits non agentifs (*rupture*), les noms dénotant un processus naturel de changement d'état (*coagulation, développement, diminution, gonflement*).

☞ Répond positivement aux tests 10.1. mais négativement à 10.2.

Exemples. *aboutissement, aurore, automne, avènement, catastrophe, cause, coagulation, controverse, croissance, cérémonie, diminution, dysfonctionnement, débâcle, décès, défaite, expiration.*

8 Situations non dynamiques ou propriétés

Les situations non dynamiques, annotées **ATTRIBUTE**, **FEELING** ou **STATE**, dénotent des états ou des propriétés (voir à ce propos Koehl, 2012, Kleiber, 2014).

8.1 Attribute

ATTRIBUTE : propriété constitutive d'une entité (individu ou objet) ou d'une situation. Cette catégorie s'apparente aux *individual level predicates* (ILP) de Carlson (1977), c'est-à-dire à des propriétés intrinsèques qui s'appliquent à l'individu en tant que tel, par opposition à celles qui ne s'appliquent qu'à une portion spatio-temporelle de son existence (SLP pour *stage level predicates*) annotées **STATE**.

☞ Répond positivement aux tests 12.

Tests 12 : Attribute

- 12.1. NO est d'un(e) grand(e) **N**.
ou
- 12.2. N[humain] a fait preuve de **N**/d'un(e) **N** + expansion.
ou
- 12.3. Cela m'a altéré **Dét N**^a.

a. Si le sujet est un animé, doit être paraphrasable par *avoir du N*.

Exemple d'application des tests 12

- 12.1. Ce programme est d'une grande **efficacité**.
ou
- 12.2. Paul a fait preuve de **flexibilité**.
ou
- 12.3. Cela m'a altéré **la mémoire**.

Exemples. *allure, comportement, dimension, dévouement, efficacité, faisabilité, flexibilité, force, format, grandeur, gravité, générosité, honneur, importance, mémoire, tempérament, timbre.*

8.2 Feeling

FEELING : état psychologique ou physiologique transitoire (*peur*). Le sujet est alors l'expérimenteur de l'affect dénoté (voir par exemple Anscombe, 1995, Gross, 1995b, Goossens, 2015). N'inclut pas les maladies annotées **STATE**.

☞ Répond positivement aux tests 13.

Tests 13 : Feeling

13. N[humain] ressent/éprouve **Dét N** (+expansion).

Exemple d'application des tests 13

13. Paul ressent **de la peine**.

Exemples. *douleur, désir, engourdissement, enthousiasme, envie, euphorie, fatigue, frayeur, frisson, haine, impatience, indifférence, inquiétude, joie, oppression, peine, plaisir, quiétude, sentiment.*

8.3 State

STATE : propriété non constitutive d'une entité (*paternité*), état résultant (*irradiation*), état des choses (*silence*). Cette catégorie s'apparente aux *stage level predicates* (SLP) de Carlson (1977), c'est-à-dire à des propriétés qui ne s'appliquent qu'à une portion spatio-temporelle de l'existence de l'individu (*concubinage, ivresse*). Inclut aussi les maladies comme états subis (*hypertension*).

☞ Répond positivement aux tests 14.

Tests 14 : State

- 14.1. N[humain] ressent/éprouve **Dét N** (+expansion).
ou
- 14.2. N[inanimé] est dans un état de **N** (+expansion).
pour les maladies
- 14.3. **Dét N** est/n'est pas une maladie contagieuse.
et
- 14.4. N[humain] a guéri de **Dét N**.
et
- 14.5. Le médecin m'a diagnostiqué **Dét[indef, sing] N**.

Exemple d'application des tests 14

- 14.1. Paul ressent **une grande liberté**.
ou
- 14.2. Paul est dans un état d'**ivresse**.
pour les maladies
- 14.3. **La grippe** est une maladie contagieuse.
et
- 14.4. Paul a guéri de **sa grippe**.
et
- 14.5. Le médecin m'a diagnostiqué **une grippe**.

Exemples. arthralgie, aseptie, besoin, captivité, concubinage, grippe, hypersensibilité, hypertension, hypocalcémie, hypotension, hémorroïde, illégalité, ivresse, liberté, malaria, manque, priorité.

9 Catégories hybrides

Les catégories hybrides englobent les noms à sens dits *hybrides*, c'est-à-dire dont les différentes facettes sémantiques ne provoquent pas d'effet zeugmatique lors d'une co-prédication (cf. §3).

9.1 Act+Cognition

ACT+COGNITION : nom qui dénote conjointement un acte, souvent communicationnel, et le contenu informationnel associé à cet acte (*déposition, traduction, témoignage*).

☞ Répond positivement au test 15.

Tests 15 : Act+Cognition

16. [Elle a lu/écouté]Pred_{COGNITION} **dét N** (+exp) [effectué(e) par le jeune homme]Pred_{ACT}.

Exemple d'application des tests 15

16. [Elle a écouté **la prophétie** effectué(e) par le jeune homme.

Exemples. analogie, analyse, énigme, prophétie, prosopographie, psychographie, sarcasme, satire, scène, simplisme, sophisme, stichomythie, sténotypie, tétralogie.

9.2 Act+Possession

ACT+POSSESSION : nom qui dénote conjointement une action impliquant un transfert d'argent et la somme transférée.

☞ Répond positivement au test 16.

Tests 16 : Act+Possession

17. [Elle a effectué]Pred_{ACT} **det N** (+exp) [modique]Pred_{POSSESSION}.

Exemple d'application des tests 16

17. Elle a effectué **une dépense** modique.

Exemples. économie, promotion, report, retenue, règlement.

9.3 Act+State

ACT+STATE : nom qui dénote conjointement une action agentive et l'état résultant (*prédation*)

☞ Répond positivement au test 17.

Tests 17 : Act+State

18. **Dét. N**, [effectuée par N]Pred_{ACT}, [a duré x temps]Pred_{STATE}.

Exemple d'application des tests 17

18. **La prédation**, effectuée par le lion, a duré 2 heures.

Exemples. avance, mentorat, prédation.

9.4 Artifact+Cognition

ARTIFACT+COGNITION : nom qui dénote conjointement un contenu informationnel et son support physique (*livre*).

☞ Répond positivement au test 18.

Tests 18 : Artifact+Cognition

19. **Dét N** [qui se trouve sur la table]Pred_{ARTIFACT} [a été écrit/rédigé/lu par ma soeur]Pred_{COGNITION}.

Exemple d'application des tests 18

19. **Le coran** qui se trouve sur la table a été lu par ma soeur.

Exemples. anthologie, archive, autobiographie, avesta, bible, bibliographie, calligramme, coran, diaporama, koran, midrash, monographie, script, scénario, synopsis, Talmud, téléfilm, évangile.

9.5 Event+Object

EVENT+OBJECT : nom qui dénote conjointement un événement naturel et les conséquences physiques de cet événement, qui prennent la forme d'un objet naturel (*éboulement*).

☞ Répond positivement au test 19.

Tests 19 : Event+Object

20. **Dét N** [qui s'est produit à tel endroit]Pred_{EVENT} [s'étend sur + dimension]Pred_{OBJECT}.

Exemple d'application des tests 19

20. **La gelée** qui s'est produit en Bourgogne s'étend sur plus de 15 hectares.

Exemples. *coulée, gelée.*

9.6 Event+Phenomenon

EVENT+PHENOMENON : nom qui dénote conjointement un événement et un phénomène (*éclair, sonnerie*).

☞ Répond positivement au test 20.

Tests 20 : Event+Phenomenon

21. **Dét N** [qui s'est produit à tel moment]Pred_{EVENT} [a résonné dans/illuminé la salle]Pred_{PHENOMENON}.

Exemple d'application des tests 20

21. **L'éclair** qui s'est produit hier matin a illuminé la salle.

Exemples. *accalmie, anticyclone, crue, cyclone, éclair, éclaircie, éclipse, infrason, séisme, tornade.*

9.7 Event+State

EVENT+STATE : nom qui dénote conjointement un événement et l'état résultant (*disparition*).

☞ Répond positivement au test 21.

Tests 21 : Event+State

22. **Dét.poss N**, [qui a eu lieu à tel moment]Pred_{EVENT}, [a duré x temps]Pred_{STATE}.

Exemple d'application des tests 21

22. **Sa paralysie**, qui a eu lieu pendant l'été, a duré 1 mois.

Exemples. *agonie, allergie, amnésie, métaplasie, névralgie, oligurie, orchite, paralysie, phlébite, pleurésie, pneumonie, psychose, rachialgie, streptococcie, syphilis, thrombose, trauma.*

Références

- Abeillé, A. and D. Godard (Eds.) (2021). *La Grande grammaire du français*. Arles : Actes Sud Littérature. Google-Books-ID : gndEAAAQBAJ.
- Abrard, O. and D. Stosic (2021). Polysémie standard et facettes de sens vues au travers du prisme de la diversité linguistique : entre la variation et la régularité. *Lexique* 28, 9–33.
- Anscombre, J.-C. (1995). Morphologie et représentation événementielle : le cas des noms de sentiment et d'attitude. *Langue française* 105(1), 40–54.
- Apresjan, J. D. (1974). Regular polysemy. *Linguistics* 12(142), 5–32. Published Online : 2009.
- Barque, L. (2008). *Description et formalisation de la polysémie régulière du français*. Thèse de doctorat, Paris 7, Paris.
- Barque, L., P. Haas, R. Huyghe, D. Tribout, M. Candito, B. Crabbé, and V. Segonne (2020). FrSemCor : Annotating a French corpus with supersenses. In *LREC-2020*, Marseille, France.
- Bonami, O. and D. Tribout (2021). Echantinom : a hand-annotated morphological lexicon of French nouns. In *Proceedings of Third International Workshop on Resources and Tools for Derivational Morphology (DeriMo 2021)*, pp. 42\textendash51.
- Carlson, G. N. (1977). *Reference to Kinds in English*. Thèse de doctorat, University of Massachusetts, Amherst.
- Cruse, D. A. (1995). Polysemy and related phenomena from a cognitive linguistic viewpoint. In P. Saint-Dizier and E. Viegas (Eds.), *Computational Lexical Semantics*, pp. 33–49. Cambridge : Cambridge University Press.
- Cruse, D. A. (2004). Lexical facets and metonymy. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies Journal of English Language, Literatures in English and Cultural Studies*(47), 73–96.
- Fasciolo, M. and G. Gross (2014). Classifications linguistiques vs classifications ontologiques? *Travaux de linguistique* 69(2), 129–144.
- Fellbaum, C. (1998). *WordNet : An Electronic Lexical Database*. Language, Speech, and Communication. Cambridge, London : The MIT Press.
- Flaux, N. and D. V. d. Velde (2000). *Les noms en français : esquisse de classement*. Paris : Ophrys. Google-Books-ID : BIG1uULWjmwC.
- Goossens, V. (2015). Les noms d'affect parmi les noms abstraits intensifs : nouvelles perspectives typologiques. *Langue française* 185, 59–72.
- Gross, G. (1995a). À propos de la notion d'humain. *Linguisticae investigationes* 17, 71–91.
- Gross, M. (1995b). Une grammaire locale de l'expression des sentiments. *Langue française* 105(1), 70–87.
- Haas, P. (2017). La fortuité des événements nominaux. In N. Flaux, P. Haas, V. Mostrov, K. Paykin, and F. Tayalati (Eds.), *La passion du sens en Linguistique*, pp. 377–396. Valenciennes : PUV.
- Haas, P. and P. Gréa (2015). Action et événement, deux types nominaux distincts? *Langue française* 185, 85–98.
- Hathout, N. and F. Namer (2014). Demonette, a French derivational morpho-semantic network. *Linguistic Issues in Language Technology* 11(5), 125–168.
- Huyghe, R. (2015). Les typologies nominales : présentation. *Langue française* 185(1), 5–27.
- Huyghe, R. (2021, December). Noms généraux et noms sous-spécifiés : des relations à préciser. *Corela HS-34*, [En ligne].
- Kleiber, G. (1999). *Problèmes de sémantique la polysémie en questions*. Sens et structures. Paris : Presses Universitaires du Septentrion.

- Kleiber, G. (2014). Massif/comptable et noms de propriété. *Langue française* 183(3), 71-86.
- Kleiber, G. and M. Lammert (Eds.) (2012). *SCOLIA n°26/2012, Questions de sémantique nominale*. Strasbourg : Presses universitaires de Strasbourg.
- Koehl, A. (2012). *La construction morphologique des noms désadjectivaux suffixés en français*. Thèse de doctorat, Université de Lorraine.
- Miller, G. A., R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller (1990). Introduction to WordNet : An On-line Lexical Database. *International Journal of Lexicography* 3(4), 235-244.
- Namer, F., L. Barque, O. Bonami, P. Haas, N. Hathout, and D. Tribout (2019). Demonette2 - Une base de données dérivationnelles du français à grande échelle : premiers résultats. In *Actes de la Conférence sur le Traitement Automatique des Langues Naturelles (TALN) PFIA*, Toulouse.
- Pustejovsky, J. (1996). *The Generative Lexicon*. Cambridge, London : The MIT Press.
- Roché, M. (1997). Briard, bougeoir et camionneur : dérivés aberrants, dérivés possibles. *Sillexicales* 1, 241-250.
- Roché, M. (2006). La dérivation en -ier(e) en ancien français. *Lexique* 17, 44.
- Roché, M. (2007). Logique lexicale et morphologie : la dérivation en -isme. In *Selected Proceedings of the 5th Décebrettes : Morphology in Toulouse*, Somerville, MA, pp. 45-58. Cascadilla Proceedings Project.
- Roché, M. (2011). Quelle morphologie? In M. Roché, G. Boyé, N. Hathout, S. Lignon, and M. Plénat (Eds.), *Des unités morphologiques au lexique*, Collection langues et syntaxe, pp. 15-37. Paris : Lavoisier Hermes science.
- Schnedecker, C. and W. Mihatsch (Eds.) (2018). *Les noms d'humains - théorie, méthodologie, classification, Nouvelles approches en sémantique lexicale*. Number 4 in *Études de linguistique française*. Tübingen : De Gruyter.
- Simatos, I. (2015). Les noms de parties du corps et la possession inaliénable : le cas des constructions datives. *Langue française* 185, 127-140.
- Strnadová, J. (2014). *Les réseaux adjectivaux. Sur la grammaire des adjectifs dénominaux en français*. Thèse de doctorat, Université Paris Diderot (Paris 7) & Univerzita Karlova, Paris, Prague.
- Tribout, D. (2010). *Les conversions de nom à verbe et de verbe à nom en français*. Thèse de doctorat, Université Paris Diderot (Paris 7), Paris.
- Tribout, D., L. Barque, P. Haas, and R. Huyghe (2014). De la simplicité en morphologie. In *SHS Web of Conferences*, Volume 8, pp. 1879-1890. EDP Sciences.



Accéder aux ressources du projet Demonext :

La table des lexèmes : https://demonext.llf-paris.fr/demonext/vues/table_lexemes.php

La table des relations : https://demonext.llf-paris.fr/demonext/vues/table_relations.php

Le graphe des relations : https://demonext.llf-paris.fr/demonext/vues/graph_relations.php