

CORPUS ARBORE POUR LE FRANÇAIS (FTB) ANNOTATION MORPHO-SYNTAXIQUE

Les mots simples - Les mots composés
Corpus Le Monde

Anne Abeillé, Lionel Clément
LLF, Université Paris Diderot
anne.abeille@linguist.univ-paris-diderot.fr



1999; dernière mise à jour 2015
<http://ftb.linguist.univ-paris-diderot.fr/>

SOMMAIRE

1	LES MOTS SIMPLES	4
1.1	Introduction	4
1.1.1	Les Coquilles	4
1.1.2	La segmentation	4
1.1.3	Le codage XML	5
1.2	Les choix généraux	5
1.2.1	Les catégories employées	5
1.2.2	Les critères de choix	7
1.2.3	L'annotation des lemmes	8
1.3	Les mots les plus difficiles.....	8
1.3.1	2.1. CE.....	8
1.3.2	2.2. COMME.....	9
1.3.3	DE- D – DU - DES.....	10
1.3.3.1	DE- D.....	10
1.3.3.2	DU- DES.....	11
1.3.4	EN.....	11
1.3.5	LE – LA – LES - L'.....	12
1.3.6	LEUR	12
1.3.7	LUI	12
1.3.8	MEME (S).....	13
1.3.9	PLUS	13
1.3.10	QUE- QU'.....	14
1.3.11	S'.....	15
1.3.12	SI.....	15
1.3.13	TEL (LE) (S).....	15
1.3.14	TOUT (E) (S) - TOUS	16
1.4	Les ambiguïtés les plus fréquentes.....	16
1.4.1	Les ambiguïtés catégorielles les plus fréquentes.....	17
1.4.1.1	Adjectif(A) / Participe passé (VK).....	17
1.4.1.2	Adjectif (A) / Participe présent(VG).....	17
1.4.1.3	Adjectif (A) / Nom commun (NC).....	18
1.4.1.4	Adjectif (A) / Adverbe (ADV)	18
1.4.1.5	Préposition (P) / Adverbe(ADV)	19
1.4.1.6	Préfixes(PREF) / Adverbe (ADV)	20
1.4.2	Les ambiguïtés entre sous-catégories les plus fréquentes	21
1.4.2.1	Adjectifs qualificatif, cardinaux, indéfinis	21
1.4.2.2	Conjonctions subordination/ coordination	22
1.4.2.3	Noms communs / Noms propres	22
1.4.3	Pronoms personnels, Pronoms Clitiques.....	24
1.4.3.1	Pronoms personnels	24
1.4.3.2	Pronoms Clitiques Sujet, Objet ou Réflexif.....	25
1.4.4	Les relatifs et interrogatifs.....	26
1.4.4.1	Les mots relatifs et interrogatifs.....	26
1.4.4.2	Pronoms relatifs et interrogatifs	26
1.4.5	Les mots démonstratifs.....	27
1.4.6	Les mots négatifs.....	28
1.4.7	Les mots indéfinis	29
1.4.8	Les mots possessifs	29
1.4.9	Les quantifieurs : beaucoup, trop, peu, assez, bien, tant, tellement, moins	30
1.4.10	Les nombres	31

1.4.10.1	<i>Les dates</i>	32
1.4.10.2	<i>Les heures</i>	32
1.4.10.3	<i>Les adresses</i>	33
1.4.10.4	<i>Les numéros de téléphone</i>	33
1.4.10.5	<i>Les scores</i>	33
1.4.10.6	<i>Les notes et numéros de chapitres</i>	33
1.4.10.7	<i>Les monnaies</i>	33
1.4.10.8	<i>Les partitions</i>	34
1.4.10.9	<i>Cas particuliers</i>	34
1.4.11	Les mots étrangers.....	34
1.4.12	Les signes de ponctuation	34
1.5	Récapitulatifs	36
1.5.1	Récapitulatif Pronoms personnels (et Clitiques).....	36
1.5.2	Récapitulatif Mots démonstratifs	36
1.5.3	Récapitulatif mots indéfinis	37
1.5.4	Récapitulatif Mots négatifs	38
1.5.5	Récapitulatif Mots relatifs, interrogatifs, exclamatifs.....	38
1.5.6	Récapitulatif Mots coordonnants	39
1.5.7	Récapitulatif signes brachygraphiques.....	39
2	INDEX des mots simples	40
3	LES MOTS COMPOSES	40
3.1	Introduction	43
3.1.1	Le codage XML	43
3.1.2	Le Lemme	44
3.1.3	Catégorie, traits morphologiques	44
3.1.4	Les Composants	44
3.2	Les critères pour déterminer les mots composés	45
3.2.1	Critères morphologique, graphique.....	45
3.2.2	Critères syntaxiques	45
3.2.3	Critères sémantiques	46
3.2.3.1	<i>Préfixe ou mot composé ?</i>	46
3.3	Les différents types de mots composés	46
3.3.1	Les noms composés.....	47
3.3.1.1	<i>Les noms propres composés</i>	47
3.3.1.2	<i>Les noms communs</i>	48
3.3.2	Les pronoms composés	48
3.3.2.1	<i>Pronoms indéfinis</i>	48
3.3.2.2	<i>Pronoms personnels</i>	49
3.3.2.3	<i>Pronoms démonstratifs</i>	49
3.3.3	Les déterminants composés.....	49
3.3.3.1	<i>Les indéfinis composés</i>	49
3.3.3.2	<i>Les nombres</i>	49
3.3.4	Les adjectifs composés.....	50
3.3.4.1	<i>Les adjectifs cardinaux et ordinaux composés</i>	50
3.3.4.2	<i>Les adjectifs qualificatifs composés</i>	50
3.3.5	Les prépositions composées	50
3.3.6	Les adverbes composés	52
3.3.7	Les conjonctions composées	53
3.3.7.1	<i>Les conjonctions de subordination</i>	53
3.3.7.2	<i>Les conjonctions de coordination</i>	54
3.3.8	Les verbes composés.....	54

4 Annexe : Liste des étiquettes 55

1 LES MOTS SIMPLES

1.1 Introduction

Le corpus annoté de Paris 7 est segmenté pour les mots simples et certains composés, avec une étiquette par mot. Une étiquette se compose d'une catégorie (partie du discours : nom, adjectif etc.) plus éventuellement une sous-catégorie (ex cardinal ou qualificatif pour les adjectifs, interrogatif ou relatif pour les pronoms, sujet, objet ou reflexif pour les clitiques) plus des traits morphologiques (mode, genre, nombre, personne...) .

Tous les mots ont une étiquette ainsi que certaines unités graphiques :

- les signes de ponctuation : , . ; : " ' ? ! ... - () [] /
- le clitique *on* accompagné de *l'*, l'unité graphique est *l'on*
- le *t* euphonique avec ou sans tiret accompagné du clitique, l'unité graphique est *-t-il, -moi..*

Les mots composés peuvent être séparés par un blanc. (Voir partie II pour les mots composés)

1.1.1 Les Coquilles

On a corrigé les coquilles manifestes (aboutissant à un mot inexistant) ou les fautes d'orthographe aboutissant à une incompréhension (faute d'accord) .

Mitterand => Mitterrand

devcrait => devrait

On n'a pas corrigé les fautes d'accord, de style (concordance des temps...) ni les erreurs sur les homophones (*prêt / près* etc).

1.1.2 La segmentation

Les noms de famille peuvent être composés, par exemple:

Giscard d':Estaing:NPms

Les nombres séparés par des blancs sont traités comme des composés:

20_000

Les partitifs sont des déterminants composés :

de la:D

On a normalement corrigé les mots composés non pertinents en contexte, par exemple :

C'est ainsi que:CC la Russie devint ... => ainsi:ADV que:CS

(il ne s'agit pas de la conjonction *ainsi que*)

à:P coups_de_barre:NCmp de:P fer:NCms

=> à:P coups:NCmp de:P barre_de_fer:NCfs

Il ne s'agit pas du nom familier (*avoir un*) *coup de barre* mais du nom composé *barre de fer*

Les amalgames $\{des, du, au, au(x)quel(le)(s)\}$ et $\{duquel, desquel(le)(s)\}$ ont une analyse différente selon qu'ils forment une ou deux catégories :

- ils sont décomposés en deux catégories quand il s'agit de la Préposition (à ou de) suivie par un Déterminant ou un Pronom (relatif). Par exemple le déterminant {du} est un composé (lemme = du). La seconde catégorie a une graphie vide.
- ils ne sont pas décomposés quand il s'agit de l'article indéfini (*des*=> lemme=un) ou de l'article partitif (*du* => lemme=du).

exemple= au

```
<w lemma="à" cat="P"> au </w>
```

```
<w lemma="le" cat="D" subcat="def" mph="ms"></w>
```

exemple= du

```
<w lemma="de" cat="P"> du</w>
```

```
<w lemma="le" cat="D" subcat="def" mph="ms"></w>
```

ou bien

```
<w lemma="du" cat="D" subcat="part" mph=ms> du </w>
```

1.1.3 Le codage XML

On utilise la balise <w> (word) et les attributs suivants:

lemma = lemme

cat = partie du discours

subcat= sous-catégorie

mph= morphologie

catint= catégories internes (pour les composants de composés)

Dans les mots composés, on a des balises <w> enchâssées (pour les composants).

On a aussi les attributs ei (étiquette interne) et ee (étiquette externe) qui reprennent les champs ci-dessus :

EI = catsubcatmph (pour conj, pronoms, clitiques, et noms) exemples NCms, CC, CLS1ms

EI = catmph (pour les autres catégories) exemples VKms, Ams

EE = cat-subcat-mph exemples N-C-ms , V--Kms, C-C, A-qual-ms

Attention, les catégories suivantes n'ont pas d'attribut mph : C, PONCT, ADV, P, ET, I

Les catégories suivantes n'ont pas d'attribut subcat : P, ET, I

Les V ont un champ subcat=vide

L'attribut subcat est optionnel (généralement absent) pour les ADV

1.2 Les choix généraux

1.2.1 Les catégories employées

Une étiquette comporte les parties suivantes:

- catégorie (partie du discours) : cat

- sous-catégorie : subcat

-morphologie : mph

- lemme : lemma

- catégories internes (pour les composés) : catint

On a retenu une catégorie Clitique pour les pronoms personnels faibles (liste fermée et position contrainte), mais pas une catégorie Complémenteur.

On n'a pas retenu une catégorie Prédéterminant (car trop hétérogène) ni Connecteur (liste mal définie) ni Auxiliaire (notée au niveau syntaxique comme valence verbale)

Nous avons distingué des sous-catégories pour tenir compte de classes distributionnelles plus fines :

Catégorie	Sous-catégorie	Morphologie	Description
N	C, P	f,m + s,p	Noms
A	card, ord, poss, qual, indefi, inter, exclam	1,2,3+ f,m + s,p	Adjectifs
Adv	-, inter, exclam, neg	-	Adverbes
P	-	-	Prépositions
D	card, dem, def, indef, exclam, neg, poss, inter, part	1,2,3+ f,m + s,p	Déterminants
CL	S, R, O	1,2,3+ f,m + s,p	Pronoms clitiques
PRO	card, inter, pers, neg, poss, rel, indef	1,2,3+ f,m + s,p	Autres pronoms
PREF			préfixe
C	S, C	-	Conjonctions
I	-	-	Interjections
V	-	W, G, K, P, I, J, F, T, C, S, Y + 1,2,3 + f,m + s,p	Verbes
ET	-	-	Mots étrangers
PONCT	S, W	-	Ponctuation

Tableau 1. Liste des catégories et sous-catégories du FTB.

Notons quelques différences marquantes avec les projets Multext et GRACE:

- Nous n'avons pas de catégorie résiduelle pour les éléments qui n'entreraient dans aucune classification. Tous les mots ont été étiquetés. Ceux qui sont étrangers et ne peuvent être étiquetés en contexte reçoivent l'étiquette ET.
- Les auxiliaires verbaux, y compris les auxiliaires de temps, ne sont pas distingués des autres verbes, car il nous semble que leur identification fait partie du calcul de la valence verbale, qui doit se faire au niveau proprement syntaxique (compte tenu par exemple des différents types de discontinuité entre auxiliaire et participe ou infinitif suivant).
- Les nombres (cardinaux) sont étiquetés comme Déterminant, Adjectif, Pronom et exceptionnellement Nom selon leur position syntaxique. Nous n'avons pas retenu une catégorie distinguée, peu légitime du point de vue morphosyntaxique, pour ces termes.
- Tous les mots composés sont assortis d'une étiquette marquant la catégorie des ses composants.

On tente d'assigner une catégorie aussi précise que possible à tous les mots.

Pour les mots étrangers, on leur donne la catégorie la plus plausible en contexte, en se basant sur l'article qui précède par exemple:

le bargaining:NCms

On ne note ET que si rien dans le contexte ne permet de se prononcer sur la catégorie du mot, dans une citation par exemple:

"lasci:ET darem:ET la:ET mane:ET"

On note de même une flexion en se basant sur les règles d'accord, même si le mot est invariable:

une robe marron:Afs
les roues avant:Afp

Pour les pronoms, on se base sur leur antécédent :

elles se:CL3fp lavent

il se:CL3ms change

Moi qui (Pronom relatif féminin singulier) *lui* (Pronom fort masculin singulier) *ai donné un fils*.

Pour les pronoms figés (*en, y*) on note par défaut masc. sing.

Pour les noms propres, on note le genre s'il est clair dans le contexte, sinon on ne le note pas:

Paris:NPms libéré
Il va à Marseille:NPs

1.2.2 Les critères de choix

On a privilégié les critères distributionnels aux critères morphologiques. On a ainsi des Verbes ou Adjectifs invariables, des Adverbes variables:

Voilà:VP3s Paul (car on a : *Le voilà* et la position postclitique est réservée aux V)

La roue avant:Afs droite:Afs (car avant est dans une position interdite aux Prep)

Une pomme toute:ADV rouge:A (car prémodifieur d'Adjectif est une position réservée aux Adv)

Etiquetage contextuel

On a minimisé les cas de recatégorisation contextuelle car :

- on veut que le corpus annoté puisse servir à extraire un dictionnaire fiable et réutilisable
- on dispose d'autres niveaux d'annotation pour noter les syntagmes et les fonctions

Un N peut être épithète ou attribut (c'est un usage productif, et la recatégorisation Adj serait arbitrairement limitée aux cas rencontrés dans le corpus):

Paul est très famille:NCfs
une recette maison:NCfs

Un Adj peut être tête de groupe nominal (avec ellipse du nom) :

la:Dfs petite:Afs rouge:Afs
les:Dmp meilleurs:Amp

Un Adj peut être employé comme modifieur de V (usage productif et distribution différente des Adverbes correspondants) : *voter utile, acheter malin* etc.

refuser net:Ams (Il a nettement refusé vs * Il a net refusé)

1.2.3 *L'annotation des lemmes*

Nous entendons par lemme une abstraction sur l'ensemble des formes fléchies ou élidées d'un même terme lexical (pour une catégorie donnée). Nous ne distinguons pas les lexèmes ; nous notons les polysèmes homographes partageant les mêmes flexions par un même lemme. Par exemple, le lemme {fraise} correspond à la classe des formes fléchies {fraise, fraises} qu'il s'agisse de l'outil ou du fruit. Mais la forme {cuisinière} peut correspondre à deux lemmes différents "cuisinier" ou "cuisinière " (=fourneau)

Le lemme est parfois, mais rarement, ambigu pour une même forme et une même catégorie, c'est pourquoi il est important de le noter dans le corpus. Le lemme de la forme {suis} est, par exemple ambigu entre {être} et {suivre}, les lemmes {fil} et {fils} conviennent pour la forme { fils}.

Pour les pronoms, le lemme est la forme du masc. sing de la personne correspondante:

je, te, me, m' : lemme=je
il, elle, l', le, la, les, lui, leur : lemme=il

Pour les amalgames on a un lemme sans graphie:

duquel

<w lemma="de" cat="P"> duquel </w>

<w lemma="lequel" cat="PRO" sous-cat="rel" mph="ms"> </w>

1.3 *Les mots les plus difficiles*

On indique ici les principes de choix entre catégories pour des mots grammaticaux très fréquents.

1.3.1 *2.1. CE*

C' est toujours pronom faible (CL3ms) :
c'est la vie, c'était un brave

On doit choisir pour CE entre D, PROms et CL3ms. Il est Dms devant N ou A quand il commute avec un autre article, sa forme devant voyelle est *cet* :

devant:P ce:Dms spectacle:NCms
selon:P ce:Dms dernier:Ams,

Il est CL3ms sujet devant un autre clitique, (*ne*) ou un Verbe conjugué (*être*) ou objet devant un participe présent (*disant, faisant*) :

ce:CLS3ms me:CLO1ms semble:VP3s
ce:CLO3ms disant:VG
ce:CLO3ms faisant:VG
ce:CLS3ms serait:VC3s bien:ADV
est:VP3s -ce:CLS3ms fini:VKms ?
ce:CLS3ms ne:ADV serait:VC3s pas:ADV plus:ADV mal:ADV

Il est PROms devant une relative, un adverbe ou une conjonction de subordination :

ce:PROms qui:PROR3ms serait:VC3s bien:ADV
 ce:PROms pourquoi:ADV je:CLS1ms vous:CLO2mp ai:VP1s appelé:VKms
 il cherche:VP3s à:P ce:PROms que:CS tu:CLS2ms viennes:VS2s

Ce CL est toujours singulier, quand il est suivi d'un V pluriel on analyse ça comme un cas d'accord avec l'attribut (cf en anglais *there is... there are...*)

ce:CL3ms sont:VP3p des:Dfp fleurs:NCfp

1.3.2 2.2. COMME

Comme est annoté ADV, ADVE (exclamatif), P, CS et CC.

Comme est CS dans les interrogatives indirectes, les subordonnées causales et les comparatives non réduites :

Je n'aime pas:ADV comme:CS il:CLS3ms joue:VP3s
 Comme:CS il:CLS3ms faisait:Vp3s beau:Ams, je suis sortie
 Il:CLS3ms joue:VP3s comme:CS il:CLS3ms chante:VP3s

Comme est P quand il introduit une comparative réduite (sans verbe), un attribut nominal ou adjectival et quand il sert à coordonner :

Je:CLS1ms le:CL3ms considère:VP3s comme:P un:Dms frère:NCms
 Il est:VP3s comme:P un:Dms père:NCms pour:P moi:PRO1ms
 Jean:NPms comme:P son:Dms père:NCms aime:VP3s la musique
 Il:CLS3ms a:VP3s été:VKms engagé:VKms comme:P cuisinier:NCms
 Comme:P violoniste:NCms on:CLS3ms ne:ADV trouve:VP3s pas mieux
 je le considère comme:P fou:Ams
 Elle aime son père comme:P toi:PRO2ms
 Elle chante comme:P un:Dms rossignol:NCms
 Comme:P toujours:ADV il:CLS3ms est:VP3s en:P retard:NCms
 Comme:P prévu:VKms , il:CLS3ms ne:ADV viendra:VF3s qu':ADV après:P le:Dms spectacle:NCms

Comme est ADV quand il prémodifie un adjectif (sens de *presque, quasi*) ; il est ADVE en début d'exclamative :

Il était:VI3s comme:ADV mort:Ams
 Comme :ADVE il fait:VP3s beau:Ams !

Comme est CC quand on peut lui substituer une autre conjonction de coordination. Le groupe nominal coordonné est au pluriel.

Jean:NPms comme:CC Pierre:NPms sont:VP3p malades:Amp (comme =et)
 Jean:NPms comme:P Pierre:NPms est:VP3s malade:Ams (comme=de même que)

1.3.3 DE- D – DU - DES

1.3.3.1 DE- D

De (ou *d'*) peut être Préposition ou Déterminant. Il peut former un déterminant simple (pluriel ou singulier, masc. ou fem) ou complexe (*de la*, avec un déterminant défini qui suit). Pour trancher il faut voir si *de* commute avec un Déterminant ou avec une Préposition.

Le déterminant *de* (ou *d'* singulier) est indéfini ou partitif. Il faut noter cette différence de sous-catégorie. Pour trancher, on fait commuter *de* avec *un* ou avec *du/de l'/de la*.

Nous notons DI les Det indéfinis et DP les déterminants partitifs.

Je ne vois pas de:DIms chien (je ne vois pas un chien - *du chien)
 Je n'ai pas d':DPms argent:NCms (* un argent, de l'argent)
 Je n'ai plus:ADV de:DPfs monnaie:NCfs (de la monnaie - *une monnaie)

Jean:NPms vend:VP3s beaucoup:ADV de:DPms beurre:NCms (= il vend du beurre)
 Personne:PRO3ms ne:ADV prendra:VF3s de:DIfs décision:NCfs (= une décision précise)

De est déterminant indéfini ou partitif après une préposition (sauf cas de prep complexe) , après un verbe transitif direct, comme introducteur d'un groupe nominal sujet:
 Il apparaît au pluriel avant un adjectif épithète antéposé, et au singulier dans des tours négatifs.

Il:CLS3ms a:VP2s réussi:VKms malgré:P de:DIfp sérieuses:Afp difficultés:NCfp
 Je:CLS1ms vois:VP1s de:DImp beaux:Amp enfants:NCmp
 Je:CLS1ms ne:ADV veux:VP1s pas:ADV de:DPms beurre:NCms
 Vous:CLS2p m':CLO1ms en:CLO3fp apprenez:VP2p de:DIfp belles:Afp
 Personne:PROms ne:ADV prendra:VP3s de:DIfs décision:NCfs avant:P lundi:NCms
 De:DImp nouveaux:Amp problèmes:NCmp sont:VP3p à:P l':Dms horizon:NCms
 Il:CLS3ms n':ADV y:CLO3ms a:VP3s pas:ADV de:DIfs relève:NCfs
 De:DImp trop:ADV gros:Amp cadeaux:NCmp
 Ils:CLS3mp n':ADV auront:VF3p plus:ADV d':DIms effet:NCms
 On:CLS3ms en:CLO3fp annonce:VP3s de:DIfp nouvelles:Afp (cf. on en annonce une nouvelle)

De fait partie d'un déterminant composé partitif *de l'* ou *de la* :

Gagner:VW de l':DPms argent:NCms
 Je:CLS1ms veux:VP1s de la:DPfs farine:NCfs (cf je veux de cette farine)

De est préposition avant un Vinf, avant ou après un pronom, avant un déterminant, après un verbe non transitif direct, comme introducteur de complément de Nom ou d'Adjectif, comme introducteur d'un complément circonstanciel et après certaines prépositions complexes:

Je:CLS1ms viens:VP1s de:P finir:VW
 Il:CLS3ms essaie:VP3s de:P les:CLO3mp encourager:VW
 J':CLS1ms en:CLO3fs ai:VP1s une:PROfs de:P cassée:VKfs
 Il:CLS3ms travaille:VP3s de:P ses:Dfp mains:NCfp
 Je:CLS1ms rêve:VP1s de:P beaux:Amp enfants:NCmp
 La:Dfs maison:NCfs de:P Paul:NPms

Il:CL3ms est:VP3s fier:Ams de:P lui:PRO3ms
 Il vient de:P chez:P lui:PRO3ms
 Une pomme:NCfs de:P trop:ADV
 J'ai un fils:NCms de :P malade:Ams
 Je veux:VP1s plus:ADV de:P profits:NCmp

De est préposition composée dans les mêmes conditions:

En dépit de:P sérieuses:AQfp difficultés:Nfp
 Au-delà de:P la:Dfs mer:NCfs
 fermer:VW le:Dms gaz:NCms avant de:P partir:VW

D'entre est une préposition composée, variante de *de* :

Certains:PRO3mp de:P nous:PRO1ms
 Certains:PRO3mp d'entre:P nous:PRO1ms

1.3.3.2 DU- DES

- Tous les deux peuvent être agglutinés Prep+Det def (*du = de le, des = de les*.)

Il:CLS3ms tombe:VP3s du:P+Dms train:NCms (on peut dire : de ce train)
 Du:P+Dms balcon:NCms on:CLS3ms voit:VP3s la:Dfs lune:NCfs
 La:Dfs nomination:NCfs du:P+Dms Président:NCms
 Il:CLS3ms est:VP3s fier:Ams du:P+Dms résultat:NCms
 Jouer du:P+Dms piano:NCms (car : jouer de cet instrument)

Je:CLS1ms ne:ADV me:CL1ms souviens:VP1s pas:ADV des:P+Dfp vacances:NCfp (on peut dire : de ces vacances)
 Un communiqué des:P+Dmp plus:ADV étonnants:Amp

- Ou bien on a le déterminant composé partitif *du* (lemme=*du*),

Jean:NPms fait:VP3s du:DIms vélo:NCms
 Je:CLS1ms veux:VP1s du:DPms beurre:NCms

- Ou enfin on a le déterminant indéfini *des* (lemme=*un*) .

Malgré:P des:DIfp difficultés:NCfp énormes:Afp
 Je:CLS1ms vois:VP1s des:DImp enfants:NCmp
 Des:DImp problèmes:NCmp sont:VP3p à:P l':Dms horizon:NCms
 Bien:ADV des:DIfp erreurs:NCfp

1.3.4 EN

En est soit Prep soit Pronom faible (CL, Objet, 3ieme pers). *En* est P quand il introduit un groupe prépositionnel circonstanciel, complément de V, de N... Il est CL03 (ms, fs, fp ou mp selon antécédent) devant un V (sauf part présent) et Prep partout ailleurs:

En:P le:CL03ms voyant:VG
 Une:Dfs bague:NCfs en:P or:NCms
 En:P en:CL03ms ajoutant:VG
 Paul:NPms en:CL03ms parlera:VF3s demain:ADV

Il:CLS3ms est:VP3s mort:Ams en:P travaillant:VG

On a déterminé le genre et le nombre du CL selon le contexte (le coréférent) , on a mis masc. sing par défaut:

Paul vend des pommes:NCfp et:CC Max:NPms en:CLO3fp achète:VP3s

1.3.5 LE – LA – LES - L'

le, la, les sont annotés D ou CL. Ils sont pronoms faibles (CL, 3ième pers) devant V, *ne* ou devant un autre pronom faible, ou après un V impératif. Ils sont D partout ailleurs (devant N ou Adj...):

J':CLS1ms essaie:VP1s de:P les:CLO3mp encourager:VW
 Je:CLS1ms ne:ADV les:CLO3mp lui:CLO3ms donnerai:VP1s pas:ADV
 Pour:P ne:ADV pas:ADV les:CLO3mp y:CLO3ms pousser:VW
 En:P les:CLO3mp voyant:VG
 Les:Dmp trois:Amp brigands:NCmp
 De la:Dfs salade:Nfs

Les infinitifs substantivés, précédés d'un déterminant, sont catégorisés NC:

le:Dms boire:NCms et le:Dms manger:NCms
 le:Dms souvenir:NCms

1.3.6 LEUR

Leur est soit D (masc. ou fem, plur), soit Pronom personnel faible (masc. ou fem, plur), noté CL(O) soit Pronom possessif (masc. ou fem, sing), noté PRO. Pour trancher, on a regardé s'il commute avec un autre D, un autre CL (*lui*) ou un autre Pro possessif (*sien*) . *Leur* est D devant A ou N. *Leur* est CL devant V, *ne* ou un autre CL (*y*, *en*) et après V impératif.

Leur:Dms enfant:NCms est:VP3s malade:Ams
 Il:CLS3ms faut:VP3s tenir:VW compte:NCms de:P leur:Dfs performance:NCfs
 Je:CL1ms leur:CLO3mp parlerai:VP1s
 Parlons:VY1p -leur:CLO3mp
 Avant de:P leur:CLO3mp en:CLO3ms parler:VW
 Je:CLS1ms préfère:VP1s le:Dms leur:PRO3ms

Pour le Det, on a renseigné le genre.

Pour le pronom faible, on a indiqué le genre (selon l'antécédent, par défaut masc)

1.3.7 LUI

Lui est soit Pronom faible (CLO:masc. ou fem, sing, pluriel=leur) soit Pronom fort (PRO:masc. sing, pluriel=eux) soit VK= part passé (*lui*, RARE). Pour savoir si c'est le CL ou le PRO fort, on le met au pluriel (CL => *leur*, PRO => *eux*).

Lui est faible devant V, *ne* ou un autre CL (*en*, *y*) , faible aussi après un V impératif, fort sinon (après Prep, avant relative) :

Je:CLS1ms lui:CLO3ms en:CLO3ms parlerai:VP1s
 En:P lui:CLO3ms parlant:VG

Parle:VY2s -lui:CLO3ms !
 Pour:P lui:CLO3ms parler:VW
 Je:CLS1ms travaille:VP3s pour:P lui:PRO3ms
 une:Dfs photo:NCfs de:P lui:PRO3ms
 lui:PRO3ms qui:PROR3ms parle:VP3s si:ADV bien:ADV

1.3.8 MEME (S)

Même(s) est soit AQ (qualificatif), soit ADV soit AI (indéfini).

Même est AI (adj indéfini) comme épithète antéposé ou avec ellipse de nom:

le:Dms même:AIm homme:NCms
 ce:CL3ms sont:VP3p les:Dmp mêmes:AImp
 avec:P même:Aifs fortune:NCfs

Même est ADV (invariable) devant A, ADV, CS, P etc. et en prédéterminant :

même:ADV si:CS ...
 même:ADV mort:AQms
 même:ADV avec:P des:Dfp augmentations:NCfp
 même:ADV Jean:NPms a compris
 même:ADV les:Dmp chiens:NCmp sont plus propres
 même:ADV pas:ADV de:DPms blé:NCms

Même en suffixe ou en épithète postposé est A (masc. ou fem, sing ou plur) sauf *moi-même*, *lui-même* etc. qui sont des PRO composés

cet:Dms homme:NCms -même:AQms qui:PROR3ms etc.
 c'est la prudence:NCfs même:Afs
 aujourd'hui:ADV même:AQms

1.3.9 PLUS

Plus est soit ADV négatif soit ADV positif (quantifieur), soit P marque de l'addition soit participe passé pluriel (verbe *plaire* ou *pleuvoir*, rare !). Nous le notons adverbe ou préposition.

Il est ADV dans tous ses emplois négatifs, comme adverbe de quantité et dans les tours superlatifs ou comparatifs :

Il mange plus:ADV que:CS toi:PRO2ms
 Le:Dms plus:ADV grand:AQms des:P+Dmp trois:PROmp
 le:Dms succès:NCms le:Dms plus:ADV grand:AQms
 plus:ADV gentiment:ADV
 +:ADV 3:D %:NC

Il est P dans des contextes d'addition :

deux:NCms plus:P deux:NCms égalent:VP3p quatre:NCms

Il entre souvent dans la composition d'un adverbe : *d'autant plus*, *de plus en plus*, *qui plus est*, *plus ou moins*, etc.

le plus est un adverbe composé sauf en tour superlatif.

plus que est adverbe composé seulement comme prémodifieur d'adjectif.

plus de est adverbe simple suivi d'une préposition quel que soit le contexte.

plus:ADV de:P farine:NCfs
 ce que j'aime le plus:ADV
 une:Dfs situation:NCfs plus que:ADV florissante:Afs
 avec:P plus:ADV de:P trois-cents: personnes
 avec:P plus:ADV de:P facilité:NCfs

1.3.10 QUE- QU'

QUE est ambigu entre Adverbe simple (ADV) ou exclamatif (ADVE), Pronom relatif (PROR) ou interrogatif (PROI3ms) et Conjonction de subordination (CS).

Que est pronom objet introduisant une relative ou une interrogative directe ou indirecte. Il est pronom relatif dans les clivées:

la fille que:PROR3fs je vois
 l'homme que:PROR3ms tu aimes
 les:Dfp idées:NCfp que:PROR3fp tu:CLS2ms as:VP2s approuvées:VKfp
 je demande:VP1ms ce:PROms que:PROR3ms tu:CLS2ms vois:VP2s
 j'aime:VP1ms ce:PROms que:PROR3ms je:CLS1ms vois:VP1s
 Que:PROIms vois:VP2s -tu:CLS2ms ?
 je:CLR1ms ne:ADV sais:VP1ms que:PROIms faire:VW
 C'est Jean:NPms que:PROR3ms je vois
 C'est à Jean que:PROR3ms je pense
 quoi:PROIms qu':PROR3ms en:CLO3ms pensent:VP3p les:Dmp collègues:NCmp

On annote QUE PRO relatif pour la personne, le genre, le nombre de son antécédent, QUE interrogatif est toujours 3ème pers masc. sing.

QUE est ADV dans *ne...que* (tour restrictif) :

Je:CLS1ms ne:ADV vois:VP1s que:ADV toi:PRO2ms
 Ils:CLS3mp n':ADV ont:VP3p plus:ADV que:ADV la:Dfs haine:NCfs

QUE est ADVE dans les exclamatives :

Que:ADVE de:P bonbons:NCmp !

QUE est conjonction de subordination après un verbe (ou un nom ou un adjectif) à complétive, après une Prep, dans les comparatives ou les corrélatives (mêmes réduites) et dans les impératives.

Je veux:VP1s que:CS tu:CLS2ms viennes:VS2s
 Comme:CS il:CLS3ms pleuvait:VI3s et:CC que:CS tu:CLS2ms venais:VI2s ...
 Voilà:VP3s qu':CS il:CLS3ms pleut:VP3s
 Il:CLS3ms cherche:VP3s à:P ce:PROms que:CS vous:CLS2mp veniez:VS2p
 Il:CLS3ms s':CLR3ms étonne:VP3s de:P ce:PRO3ms que:CS vous:CLS2mp partiez:VS2p
 L':Dfs idée:NCfs que:CS tu:CLS2ms approuves:VS2s notre:Dms projet:NCms
 me:CLO1ms réjouit:VP3s

plus:ADV grande:Afs que:CS vous:PRO2mp
 plus:ADV grande:Afs que:CS vous:CLS2mp ne:ADV l':CLO3ms imaginiez:VI2p
 Qu':CS il:CLS3ms aille:VS3s se:CLR3ms faire:VW voir:VW !
 Une capacité:NCfs plus:ADV que:CS restreinte:Afs
 un homme:NCms tel:Ams que:CS lui:PRO3ms
 la regle:NCfs est:VP3s que:CS les députés soient élus ...

1.3.11 S'

Forme élidée de SE (CL 3ème pers) ou de SI (Conj Sub) :

Ils:CLS3mp veulent:VP3p s':CLR3mp amuser:VW
 S':CS il:CLS3ms pleut...

On met le nombre et le genre de S(E) selon son antécédent, par défaut masc. sing.

1.3.12 Si

Si est soit NC (note de musique ou affirmation) soit Adverbe (intensif, pouvant introduire une consécutive) soit Conj de subordination (introduisant une subordonnée conditionnelle ou une interrogative indirecte).

Si est toujours CS lorsqu'il introduit une phrase.

Il est Adv devant un Adj ou un autre Adv susceptibles de degrés ou un groupe prépositionnel (ou un Nom) employé comme attribut (*il est si médecin ! RARE*)

si:ADV gentil:Ams
 si:ADV rarement:ADV
 Il:CLS3ms est:VP3s toujours:ADV si:ADV en:P retard:NC:ms que:CS...
 Si:CS tu:CLS2ms viens:VP2s
 Je:CLS1ms me:CLR1ms demande:VP1s si:CS tu:CLS2ms viendras:VP2s
 Il m'a répondu que:CS si:NCms
 Il m'a répondu que:CS oui:NCms

1.3.13 TEL (LE) (S)

Tel est D, PRO ou A (indéfini).

TEL est A en prédéterminant, en épithète et en attribut (souvent avec inversion du sujet)

tel:Ams un:Dms lion:NCms
 un:Dms tel:Ams homme:NCms
 tel:Ams était:VI3s cet:Dms homme:NCms
 Il:CLS3ms était:VI3s tel:Ams que:CS tu:CLS2ms l':CLO3ms avais:VI2s
 laissé:VKms

TEL est D devant un NC sans déterminant:

tel:Dms père:NCms tel:Dms fils:NCms

TEL est PRO comme groupe nominal:

tel:PROms qui:PROR3ms rit:VP3s vendredi:NCms ...
 avec:P tel:PROms ou:CC tel:PROms de:P vos:Dmp partenaires:NCmp

1.3.14 TOUT (E) (S) - TOUS

Tout est D, PRO, ADV ou Adjectif.

Tout (e) est D singulier devant un N ou un Adj en tête de groupe nominal (rare) :

Tout:Dms homme:NCms a:VP3s le:Dms droit:NCms de:P...

Toute:Dfs nouvelle:Afs étudiante:NCfs se:CLR3fs verra:VF3s remettre:VW...

Tou(t) (e) (s) est Adj en position de prédéterminant (comme l'adjectif *seul*) :

Tout le monde est un pronom composé.

Tout le monde:PROms

Toutes:Afp les:Dfp femmes:NCfp

Tous:Amp les:Dmp deux:Amp

Tout:Ams cela:PROms est:VP3s faux:Ams

Tou (t) (e) (s) est PRO employé seul comme sujet ou complément et comme quantifieur flottant :

Tous:PROmp viendront:VF3p

Elles:CL3fp viendront:VF3p toutes:PROfp

Je veux tous:PROmp les:CL03mp voir:VW

Je les:CL03mp verrai:VF1s tous:PROmp

J'ai:VP1s tout:PROms lu:VKms

Il:CLS3ms pense:VP3s à:P tout:PROms

Tout:PROms est:VP3s possible:Ams

Tout:PROms ce:PRO3ms qui: PROR3ms évoque la mer

Tous:Amp deux:Amp ont:VP3p raison:NCfs

Tout (e) (s) est Adverbe devant un adjectif (ou un participe passé employé comme adjectif) , un adverbe ou une préposition :

Il est tout:ADV rouge:Ams

J'ai vu des filles toutes:ADV rouges:Afp

Une:Dfs toute:ADV nouvelle:Afs étudiante:NCfs

tout:ADV récemment:ADV

Tout:ADV en:P conduisant:VG

il est tout:ADV bouleversé:VKms

Tout:ADV juste:ADV.

Tout:ADV à:P l':Dms honneur:NCms

Tout peut être Nom commun (NCms) précédé d'un déterminant (rare) :

le:Dms tout:NCms est:VP3s de:P participer:VW

le tout:NCms pour le tout:NCms

1.4 Les ambiguïtés les plus fréquentes

On note ici les cas les plus systématiques d'ambiguïtés potentielles entre catégories ou sous-catégories, qu'on a dû trancher en contexte.

1.4.1 Les ambiguïtés catégorielles les plus fréquentes

1.4.1.1 Adjectif(A) / Participe passé (VK)

Après un auxiliaire de temps, au passif, avec un complément en *par* (ou un complément d'agent en *de*), on a un participe passé.

Quand on a les mêmes compléments que le verbe conjugué on a un participe passé.

En position épithète ou attribut on a généralement un adjectif.

Quand la forme est modifiée par *très* c'est un adjectif, quand elle est modifiée par *mal* ou *non* c'est un participe.

Quand le préfixe est *in* (*inconnu*, *insatisfait*...) c'est un adjectif.

Si l'on peut paraphraser avec une relative à l'actif, il s'agit du verbe au participe.

Un:Dms jugement:NCms prononcé:VKms (qu'on a prononcé)

Un:Dms recul:NCms prononcé:AQms de:P les:Dmp prix:NCmp (*qu'on a prononcé)

Il:CLS3ms a:VP3s parlé:VKms

Je:CLS1fs me:CLR1fs suis:VP1s trompée:VKfs

Elle:CLS3fs est:VP3s appréciée:VKfs de:P tous:PROImp

une canalisation bouchée:AQfs

laissé:VKms pour:P mort:AQms

rester:VW bouleversé:AQms

la porte est fermée:AQfs (résultatatif ou état permanent)

la porte est fermée:VKfs tous les soirs à 8 heures (par le gardien) (sens passif)

Quand le verbe et l'adjectif ont une différence importante de sens, le choix est fait en fonction du sens.

Un:Dms parking:NCms privé:AQms

Un:Dms enfant:NCms privé:VKms de:P dessert:NCms

1.4.1.2 Adjectif (A) / Participe présent(VG)

Le participe présent est invariable et garde les compléments du verbe.

L'adjectif s'accorde avec le nom et ne prend pas de complément d'objet direct.

En:P lisant:VG le:Dms journal:NCms...

Les:Dfp personnes:NCfp résidant:VG en:P France:NPfs depuis:P moins:ADV de:P 3:Dmp mois:NCmp

Les:Dfp erreurs:NCfp existantes:AQfp

Un:Dms déséquilibre:NCms persistant:AQms

1.4.1.3 Adjectif (A) / Nom commun (NC)

On ne recatégorise pas les Adjectifs en cas d'ellipse du nom sauf changement de sens ou de morphologie (invariable en genre, seulement déterminant défini etc.) :

une:Dfs grande:AQfs bleue:AQfs
 je veux les:Dmp deux:Amp
 il a perdu le:Dms gauche:AQms
 ce:Dms dernier:Ams
 le:Dms premier:Ams
 c':CLS3ms était:VI3s le:Dms soir:NCms de:P la:Dfs première:NCfs (sens théâtre)
 la:Dfs gauche:NCfs a:VP3s gagné:VKms
 il:CL3ms sera:VF3s le:Dms meilleur:Ams
 le:Dms rouge:NCms est une belle couleur
 10:Dmp mètres:NCmp de:P haut:NCms
 faire:VW son:Dms possible:NCms
 faire:VW l':Dms impossible:NCms
 aller:VW à:P l':Dms essentiel:NCms
 le:Dms mieux:NCms serait de ... (autrement, mieux est ADV et non A)
 à:P la:Dfs française:Afs
 le:Dms français:NCms Bull:NPms (les adjectifs de nationalité ne peuvent pas être antéposés)
 le:Dms japonais:NCms Sony:NPms
 Dans:P l':Dms immobilier:NCms
 Le:Dms secteur:NCms immobilier:AQms
 Le:Dms vide:NCms
 Une:Dfs chaise:NCfs de:P vide:AQfs

Dans les tours superlatifs, on a l'étiquette A:

l'homme:NCms le:Dms plus:ADV grand:AQms
 le:Dms plus:ADV grand:AQms des:P+Dmp hommes:NCmp

On ne recatégorise pas les noms employés comme épithètes ou comme attributs :

une veste:NCfs sport:NCms
 Il est très famille:NCfs
 le fichier:NCms clients:NCmp
 Fille d'une mère:NCfs écrivain:NCms.
 Un député:NCms RPR:NPms

1.4.1.4 Adjectif (A) / Adverbe (ADV)

On ne recatégorise pas les adjectifs en fonction adverbiale.

chanter faux, crier fort, tomber juste => AQms.

haut, bas sont Ams ou NCms, jamais ADV:

en:P haut:NCms de:P

vers:P le:Dms haut:NCms
 tout:ADV en:P haut:NCms
 un meuble:NCms haut:AQms
 rêver:VW tout:ADV haut:AQms
 parler:VW bas:Ams
 manger:VW léger:AQms
 rouler:VW propre:AQms
 Le:Dms plus:ADV vite:ADV possible:AQms

fort et *juste* sont ADV seulement comme prémodificateurs d'un adjectif, d'un adverbe ou d'une préposition.

il était:VI3s fort:ADV triste:AQms
 fort:ADV justement:ADV
 un homme fort:AQms
 parler:VW trop:ADV fort:AQms
 Ils sont venus juste:ADV à temps:ADV
 Nous avons payé la juste:AQfs somme:NCfs

bien, mal ne sont jamais Adjectifs, mais seulement ADV ou NCms:

Le:Dms bien:NCms
 bien:ADV des:Dfp années:NCfp plus:ADV tard:ADV
 tout:PROIm est:VP3s bien:ADV
 quelqu'un:PROIm de:P bien:ADV
 bien:ADV mûr:AQms
 bien:ADV évidemment:ADV
 bien-sûr:ADV
 bien-entendu:ADV
 bien-que:CS
 Il:CLS3ms aime:VP3s bien:ADV qu':CS on:CLS3ms le:CLO3ms flatte:VS3s

loin, près sont toujours des adverbes ou font partie de prépositions composées.

loin de:P chez:P toi:PROms
 près de:P 10:Dmp milliards:NCmp
 voir:VW loin:ADV
 il:CLS3ms est:VP3s loin:ADV le:Dms temps:NCms où:PRO-rel
 il:CLS3ms est:VP3s de loin:ADV le:Dms meilleur:NCms

sauf est adjectif ou préposition, mais jamais adverbe. *Sauf à* est une préposition composée seulement devant un verbe à l'infinitif.

l':Dms honneur:NCms est:VP3s sauf:AQms
 Il:CLS3ms est:VP3s sain:AQms et:CC sauf:AQms
 Tous:Amp les:Dmp fruits:NCmp sauf:P les:Dfp pommes:NCfp
 Sauf à:P prétendre:VW y:CLO3ms comprendre:VW quelque chose
 Sauf:P à:P Paris:NPs

1.4.1.5 Préposition (P) / Adverbe(ADV)

Une préposition prend toujours un complément ; sans complément, elle est Adverbe:

je vote pour:ADV
 je vote pour:P la:Dfs gauche:NCfs
 Avant:ADV il:CLS3ms y:CL3ms avait:VI3s ici:ADV un:Dms restaurant:NCms
 tu:CLS2ms partiras:VF2s avant:P la:Dfs fin:NCfs
 Depuis:ADV il:CLS3ms a:VP3s progressé:VKms

près_de:P, loin_de:P, **mais** plus:ADV de:P

durant est toujours Prep même quand son complément est antéposé:

durant:P trois:Dfp heures:NCfp
 trois:Dfp heures:NCfp durant:P

Pour les autres Prep de temps, on peut avoir un prémodifieur nominal compatible avec un complément après la Prep. En l'absence de ce dernier, l'étiquette est ADV:

trois:Dfp heures:NCfp avant:P la:Dfs fin:NCfs
 trois:Dfp heures:NCfp avant:ADV

On a regroupé les prépositions avec QUE qui suit => CS (composée) :
 pour que:CS, depuis que:CS, pendant que:CS etc.

On a regroupé les prépositions avec DE qui suit => P (composée):
 avant de:P, afin de:P

Voici, voilà sont P seulement comme tête de complément circonstanciel, sinon VP3s:

nous:CL01mp voici:VP3s en pleine tragédie
 son:Dfs arrivée:NCfs voici:P 3:Dmp ans:NCmp
 Voilà:VP3s sans doute:ADV le seul sens donné à notre existence.

il y a est P (composée) comme tête de complément circonstanciel, analyse verbale sinon:

ils:CLS3mp sont:VP3p partis:VKmp il y a:P 3:Dmp ans:NCmp
 il:CLS3ms y:CL3ms a:VP3s 3:Dmp ans:NCmp qu':CS ils sont partis

hors est toujours préposition simple ou composée.

Les:Dmp prix:NCmp hors:P énergie:NCfs
 hors de:P ma:Dfs vue:NCfs

1.4.1.6 Préfixes(PREF) / Adverbe (ADV)

L'étiquette PREF est employée pour catégoriser des mots du type : méta-, anti-, franco-, auto-, ex-, super- puis d'une liste ouverte de termes. Le préfixe ne doit pas être composant d'un terme lexical. Le tiret fait partie du préfixe.

l':Dfs ex-:PREF Union soviétique:NPfs
 l':Dmsex-:PREF mari:NCms de:P son:Dms client:NCms
 Cet:Dms appareil:NCms est:VP3s auto-:PREF régulé:VKms

Le:Dms super-:PREF PDG:NCms et:CC son:Dms directeur général:NCms

quasi et *non* sont ambigus entre PREF et ADV.

quasi est préfixe devant un nom, adverbe devant un adjectif ou un adverbe.

non est préfixe devant un nom, adverbe devant un adverbe, une préposition ou un participe.

La:Dfs quasi-:PREF totalité:NCfs
 Une:Dfs situation:NCfs quasi-:ADV parfaite:AQfs
 La:Dfs situation:NCfs de:P non:PREF reprise:NCfs
 Les:Dmp lots:NCmp non:ADV attribués:VKmp
 Non:ADV par:P méchanceté:NCfs
 Non:ADV plus:ADV

outré est toujours une préposition ou une partie d'un adverbe composé.

Outre-Manche:ADV
 Il:CLS3ms est:VP3s parti:VKms outre-Rhin:ADV
 Outre:P les:Dfp mesures:NCfp d':P usage:NCms
 Outre-mesure:ADV
 l'Outre-mer: NCms

1.4.2 Les ambiguïtés entre sous-catégories les plus fréquentes

1.4.2.1 Adjectifs qualificatif, cardinaux, indéfinis

Adjectif qualificatif ou cardinal : neuf

NEUF est Adjectif qualificatif (masc. sing) ou A cardinal (invariable, fem. ou masc, plur ou sing). Il est toujours cardinal en épithète antéposé, et généralement qualificatif en épithète (sauf dans tome neuf:ACms).

un:Dms livre:NCms neuf:AQms
 mes:Dmp neuf:ACmp livres:NCmp
 il est neuf:AQms
 ils sont neuf:ACmp
 ils sont neufs:AQmp

Adjectifs qualificatifs ou indéfinis :

SEUL (e) (s) est A qualificatif (rare) ou A indéfini (prédéterminant ou épithète gauche) :

un homme seul:AQms
 un seul:AIm homme
 seule:AIfs cette:Dfs femme:NCfs
 avec:P pour:P seul:AIm but:NCms de:P l':CL3ms entendre:VW

divers(e)(s) et *différents* sont ambigus entre D (initiale de syntagme nominal) , A qualificatifs (attribut ou épithète droit) ou A indéfinis (épithètes gauche après D) :

divers:Dmp conseils:NCmp

des:Dfp aventures:NCfp diverses:AQfp et:CC variées:AQfp
 ces:Dfp diverses:AIfp péripéties:NCfp
 tes:Dmp différents:AImp problèmes:NCmp
 des:Dmp problèmes:NCmp différents:AQmp
 ils:CLS3mp sont:VP3p différents:AQmp
 elles:CLS3fp sont:VP3p diverses:AQfp

1.4.2.2 Conjonctions subordination/ coordination

Les conjonctions de coordination (CC) sont les suivantes:

et, ou, mais, donc, or, ni, car, soit, c'est-à-dire, voire, sinon, tantôt...tantôt, puis

ou, et, mais, ni, c'est-à-dire ne sont pas ambigus

car peut être NCms (un car)

donc est plus souvent ADV (adverbe de phrase)

Il:CLS3ms est:VF3s donc:ADV venu:VKms
 donc:ADV il:CLS3ms viendra:VF3s
 il est honnête:Ams donc:CC pauvre:Ams

soit est CC sauf quand il est VS3s:

où:ADV qu':PROR3ms il:CL3ms soit:VS3s
 soit:CC chance:NCfs soit:CC hasard:NCms
 100:Dmp dollars:NCmp soit:CC 580:Dmp francs:NCmp

Sinon est le plus souvent coord, parfois adverbe :

il est consciencieux, sinon:CC intelligent
 aide-moi, sinon:CC va faire tes devoirs

Lorsque deux mots coordonnants se suivent, le second est analysé comme un adverbe

et:CC sinon:ADV
 et:CC donc:ADV

1.4.2.3 Noms communs / Noms propres

Plusieurs critères coexistent: la graphie (majuscule aux noms propres) , la syntaxe (Det devant NC), la sémantique (NP comme désignateurs rigides = référent unique).

Nous avons privilégié la sémantique, c'est-à-dire que nous admettons des NC avec majuscules et des NP avec Det (mais seulement Det défini figé pour NP si Det indef possible => NC)

Prénoms, Noms: NP

Jean:NPms de Broglie:NPms
 Les:Dmp Dupont:NPmp
 tous:AmP les:Dmp Martins:NPmp de la terre
 E.:NPms P.:NPms Thompson :NPms
 Le Luther:NP de Lucien.
 Faire un dessin à:P la:Dfs Picasso: NPms.
 l'administration Bush:NPms

On colle la particule, on ne colle pas prénom et nom. On accorde le nom de famille:

Mme:NCfs Dupont:NPfs

Mr:NCms Dupont:NPms

On garde NP pour les emplois comme NC des prénoms, sauf si changement de sens (son Jules, un César)

tous:AmP les:Dmp Martins:NPmp

On a regroupé les prénoms composés: Jean-Jacques:NPms, Marie-Pierre:NPfs...

Titres: NC

Mr:NCms le:Dms Président:NCms Chirac:NPms (un Président de la République)

Soeur:NCfs Emmanuelle:NPfs (une Soeur)

le Père:NCms de Foucault:NPms (un Père)

le Pape:NCms (un Pape)

le Général:NCms de Gaulle:NPms (un général)

Noms de société, d'institution unique, de parti politique : NP

IBM: NPs

la société:NCfs Coca Cola:NPs

l'Assemblée Nationale:NPfs

le Quai d'Orsay: NPms

Le RPR:NPms

Air France:NPs

Une: Dfs machine:NCfs Honda:NPs.

L'équipe Yamaha:NPs.

On n'a pas renseigné le genre des sociétés quand on ne le connaît pas et qu'il n'est pas clair dans le contexte.

Sigles : NP ou NC

les sigles sont NC ou NP selon les mêmes critères (désignation unique => NP, possibilité d'un Det indéfini => NC) :

la SNCF:NPfs

un HLM:NCms

les PME : NCfp

Noms de planète, de produit nom de fête: NC

Les noms d'astres sont des noms propres à l'exception du soleil et de la lune.

le soleil:NCms (Un soleil)

la lune:NCfs (une lune)

un Coca Cola:NCms

un:Dms IBM:NCms

une:Dfs twingo:NCfs

tous:Amp les:Dm No l:NCmp
 Une:Dfs 205:NCfs.
 l'avion F-14:NCms.
 le M-X:NCms
 le ELF-5: NCms.
 Le Loto - sportif:NCms.
 A:P Noël:NCms dernier:Ams (un Noël sans neige)

Points cardinaux, noms de ville, de pays, de mer, d'île, de région: NP

vers:P l':Dms ouest:NPms
 Paris:NPs, la Charente:NPfs, les Etats Unis:NPmp
 l':Dfs île de Ré:NPfs
 la:Dfs Méditerranée:NPfs
 la:Dfs Mer morte:NPfs
 Rio de Janeiro:NPs
 L'Europe:NPfs
 Berlin - ouest:NPs.ms.
 Malte:NPfs

On n'a pas renseigné le genre des villes, sauf lorsqu'il y a un déterminant non ambigu ou un contexte clair. Le genre des noms d'îles est féminin.

Noms de langue, de nationalité (sauf épithète ou attribut => A) ou de devises : NC

Il:CL3ms est:VP3s français:Ams
 la production:NCfs française:Afs
 Il:CL3ms parle:VP3s français:NCms
 le:Dms français:NCms est la langue universelle
 Les:Dmp français:NCmp votent:VP3p à:P gauche:NCfs
 Les:Dfp italiennes:NCfp sont:VP3p élégantes:Afp
 Un franc-français:NCms
 Un mark-allemand:NCms

1.4.3 Pronoms personnels, Pronoms Clitiques

1.4.3.1 Pronoms personnels

faible (CL) ou fort (PRO) : *moi, toi, lui, elle, nous, vous, eux, elles*

On a affaire au pronom faible (CL) quand il est sujet ou complément de verbe.

On a affaire au pronom fort (PRO) s'il est coordonné, après Prep ou Adv ou modifié par une relative ou un syntagme apposé:

Nous:PRO1mp, nous:CLS1mp nous:CLR1mp en:CL3ms moquons:VP1p
 Regardez:VY2p -nous:CLS1mp
 Entre:P nous:PRO1mp
 Lui:PRO3ms, le:Dms diplômé:NCms de:P ...
 Ils:CLS3mp n':ADV écoutent:VP3pl qu':ADV eux:PRO3mp et:CC nous:PRO1mp
 Elles:CLS3fp ne:ADV regardent:VP3p que:ADV nous:PRO1mp

Ecoute:VY2s -moi:CL01ms

moi-même, lui-même ... : PRO composé

On a gardé le pluriel pour le *vous* de politesse:

vous:CL22pl êtes:VP2p belle:AQfs

1.4.3.2 Pronoms Clitiques Sujet, Objet ou Réflexif

Les pronoms faibles (ou pronoms clitiques) CL se distinguent en morphosyntaxe par une sous-catégorie indiquant la fonction.

Nous indiquons la fonction des pronoms clitiques (sujet, objet ou réfléchi) en neutralisant les différences qui dépendent de la sous-catégorisation des verbes. Ainsi nous notons, *objet* pour datif et accusatif, *sujet* pour les clitiques nominaux et *refl.* pour les clitiques réfléchis.

Certains clitiques ont une seule sous-cat: *je* est toujours sujet; *en* toujours objet; d'autres en ont plusieurs (hors contexte): *nous, vous* peuvent être sujet, objet ou réfléchi.

Nous noterons CLS, CLO, CLR, les clitiques Sujet, Objet et Réflexifs.

nous et vous sont a priori ambigus entre CLS, CLO et CLR:

nous:CLS1mp sommes partis ensemble.
 nous:CLS1mp nous:CLR1mp sommes rencontrés.
 Jean nous:CL01mp donne des coups de pied.
 Jean nous:CL01mp frappe.

te, me sont a priori ambigus entre CLO et CLR:

Jean te:CL02ms donne des coups de pied
 Tu te:CLR2ms regarde grandir.
 Dis -moi:CL02ms si nous en reparlerons
 Regarde -toi:CLR2ms dans l'eau.

Nous notons le *t* euphonique et l'éventuel tiret collés au pronom. Le *l'* est également collé au clitique *on*.

pense:VPs3 -t-il:CLS3ms
 doit:VP3s -on:CLS3ms
 dis:VP1s -je:CLS1ms
 l'on:CLS3ms n':ADV en:CL3ms pense:VP3s pas:ADV moins:ADV

On est toujours singulier (masc. ou fem), à cause de l'accord avec le verbe:

On:CLS3fs est contente:Afs ?
 On:CLS3ms était contents:Amp

1.4.4 Les relatifs et interrogatifs

On doit choisir la catégorie (Adv, Pro, Det) et la sous-catégorie (inter ou rel ou exclam).

1.4.4.1 Les mots relatifs et interrogatifs

que: voir supra

qui, lequel, quoi : voir infra

Où est soit Adverbe interrogatif (noté ADV) et dénote le lieu, soit PRO relatif (PROR3) et dénote le lieu ou le temps:

l'époque:NCfs où:PROR3fs les:Dfp bêtes:NCfp parlaient:VI3p
 le lit:NCms où:PROR3ms dort:VP3s Jean:NPms
 où:ADV vas:VP2s -tu:CLS2ms ?
 je sais:VP1s où:ADV tu:CLS2ms vas:VP2s

Quand est Adverbe interrogatif (noté ADV) en interrogative directe ou indirecte, sinon Conj de subordination (CS) :

quand:CS il pleut, il mouille:VP3s
 je sais:VP1s quand:ADV il:CLS3ms arrive:VP3s
 elle part:VP3s quand:CS il:CLS3ms arrive:VP3s

Quel est Det interrogatif (noté DI) ou Det exclamatif (noté DE) devant N, ou Adj interrogatif (attribut) noté AI:

je sais quelle:DI fs tasse:NCfs il:CLS3ms veut:VP3s
 quelle:DE fs audace:NCfs !
 Quel:AIm est:VP3s cet:Dms homme:NCms ?
 Quel:AIm qu':PROR3ms il:CLS3ms soit:VS3s
 quel:AIm que:PROR3ms soit:VS3s son:Dms talent:NCms

Les formes en *n'importe* sont analysées comme ADV, PRO ou D composés, ce sont des formes indéfinies ni relatives ni interrogatives:

n'importe qui:PROms
 n'importe quoi:PROms
 n'importe quel:Dms
 n'importe lequel:PROms
 n'importe où:ADV
 n'importe quand:ADV
 n'importe comment:ADV

1.4.4.2 Pronoms relatifs et interrogatifs

Les pronoms relatifs introduisent des relatives, les pronoms interrogatifs introduisent une interrogative indirecte, ou ont des places variées dans une interrogative directe.

Qui PRO relatif est sujet ou après Prep, *qui* PRO interrogatif est seul, sujet, complément ou après Prep:

Moi:PRO1ms qui:PROR1ms ai:VP1s fait:VKms ..
 Qui:PROIms a:VP3s fait:VKms cela:PRO3ms ?
 Qui:PROIms as:VP2s -tu:CLS2ms vu:VKms ?
 Je sais:VP1s à:P qui:PROIms tu:CLS2s penses:VP2s

Lequel PRO relatif est rare comme sujet, fréquent après Prep, *lequel* PRO interrogatif est seul, sujet, complément, ou après Prep:

l'homme avec:P lequel:PROR3ms tu:CLS2ms parles:VP2s
 Lequel :PROIms veux:VP2s -tu:CLS2ms ?
 Lesquelles:PROIfp viendront:VF3p ?

Quoi est rare comme PRO relatif (après Prep), plus fréquent comme PRO interrogatif:

A:P quoi:PROIms penses-tu ?
 il ne sait pas quoi:PROIms faire:VW
 je cherche un outil:NCms avec:P quoi:PROR3ms ouvrir:VW cette boîte
 quoi:PROIms que:PROR3ms tu:CLS2s fasses:VS2s

Attention

Un PRO relatif a la personne, le genre, le nombre de son antécédent. Un Pronom interrogatif invariable (qui, que, quoi) est toujours 3ième pers masc. sing. Les pronoms interrogatifs sont toujours 3ème pers: on ne note pas la personne.

Les pronoms introduisant des « relatives » sans antécédent doivent être étiquetés PRO interrogatif (ou ADV) :

je pense à:P qui:PROIms tu penses
 j'aime qui:PROIms tu:CL2ms sais:VP2s
 je vais où:ADV tu:CLS2ms vas:VP2s

1.4.5 Les mots démonstratifs

Ils sont PRO, CLIT, ADV ou D et ont la sous-catégorie dem.

ce, c' : voir supra

ceci, cela: toujours PRO3ms

celui-ci, celle-là etc. : toujours PROms, PROfs etc.

Les particules *-ci* et *-là* collées à un nom sont ADV:

cette:Dfs fois:NCfs -ci:ADV

Les mots *voici* et *voilà* ne sont P que comme tête de syntagme prépositionnel (circonstanciel) . Comme tête de phrase (principale ou enchâssée) ils sont VP3s, et peuvent prendre un Clitique complément:

Nous:CLO1mp voici:VP3s !
 l'homme:NCms que:PROR3ms voici:VP3s
 je pense:VP1s que:CS voici:VP3s une:Dfs occasion:NCfs unique:Afs
 un premier jugement voici:P trois:Dmp ans:NCms l':CLO3ms avait:VI3s
 débotté:VKms

1.4.6 Les mots négatifs

Ils sont PRO, ADV, ou D et ont la sous-catégorie neg.

n' et *ne* sont analysés comme ADV. NON est ADV ou I (interjection) Les forclusifs sont soit des ADV (pas, plus, guère, jamais) soit des PRO3ms (*rien, personne, aucun, nul*) soit des D (*nul, aucun*) :

je ne:ADV vois:VP1s personne:PRO3ms
 Personne:PRO3ms n':ADV est:VP3s venu:VKms
Une:Dfs personne:NCfs viendra
 jamais:ADV Paul:NPms ne:ADV viendra:VF3s
 Paul:NPms n':ADV a:VP3s pas:ADV de:Dfs chance:NCfs
 faire:VW un:Dms pas:NCms
 Rien:PRO3ms ne:ADV l':CLO3ms arrêtera:VF3s
 On:CLS3ms ne:ADV voit:VP3s rien:PRO3ms de:P bien:ADV
 je n':ADV en:CL3ms vois:VP1s aucun:PRO3ms
 je ne:ADV vois:VP1s aucun:Dms chien:NCms

Personne en emploi négatif est PRO3ms, sinon c'est un NCfs (*une personne*) .

Rien est PROms

il:CL3ms n':ADV a:VP3s rien:PROms vu:VKms
 sans:P rien:PROms dire:VW
 compter:VW pour:P rien:PROms

Emploi substantivé (rare) de RIEN:

les:Dmp petits:AQmp riens:NCmp de:P la:Dfs vie:NCfs

Emploi archaïque (rare) de RIEN comme ADV:

c'est rien bien ! (= un peu)

NUL peut être PRO négatif, D négatif ou A qualificatif:

une partie:NCfs nulle:AQfs
 Nul:PRO3ms n':ADV est:VP3s censé:Ams ignorer:VW la:Dfs loi:NCfs
 sans:P nul:Dms doute:NCms

PLUS est ambigu entre ADV négatif et ADV positif:

plus:ADV ou:CC moins:ADV
 je n'ai plus:ADV-neg de:Dfs farine

AUCUN(E) est D ou PRO:

aucun:PRO3ms ne:ADV viendra:VF3s

sans:P aucun:Dms doute:NCms

ne pas, ne plus, ne jamais, placés devant un infinitif, forment des adverbes composés

1.4.7 Les mots indéfinis

Ils sont PRO, A, ADV ou D et ont la sous-catégorie Indef.

tout (e) (s): voir supra

Adjectifs indéfinis: *seul, divers, différents*: voir infra

Autre(s) est soit PRO (employé comme tête de groupe nominal), soit A qualificatif (rare : attribut ou épithète droit) soit A indéfini (épithète gauche) :

un autre:AImms homme:NCms
ils sont autres:AQmp
D':Dmp autres:PRO3mp viendront:VFmp
j':CLS1ms en:CLO3ms vois:VP1s un:Dms autre:PRO3ms
Les:Dmp uns:PRO3mp les:Dmp autres:PRO3mp

Certain (e) (s) est soit PRO (employé comme tête de groupe nominal) , soit A qualificatif (rare : attribut ou épithète droit) soit A indéfini (épithète gauche) soit D:

Certains:PRO3mp pensent:VP3p que:CS ...
certaines:Dfp femmes:NCfp
une:Dfs certaine:AIfs femme:NCfs
je:CLS1fs suis:VP1s certaine:AQfs que:CS ...

Quelque (s) est soit D soit A indéfini.
quelque (au singulier) est également un ADV.

quelques:Dmp restes:NCmp
ces:Dfp quelques:AIfp fleurs:NCfp
quelque:Dms talent:NCms qu':PROR3ms on:CLS3ms lui:CLO3ms trouve:VP3s
quelques:Dmp uns:PRO3mp
quelque:ADV riche:Ams qu':PROR3ms il:CLS3ms soit:VP3s
quelque:ADV trois:Dmp millions:NCmp de:P francs:NCmp
si:CS vous:CLS2mp avez:VP2p quelque:Dms soucis:NCms que:PROR3s ce:CLS3ms
soit:VS3s

Quelqu'un, quelque chose: toujours PRO3ms composé, sauf emploi particulier avec déterminant => NCms:

un petit quelquechose:NCms

1.4.8 Les mots possessifs

On doit choisir entre les catégories D, PRO, A. Ils ont la sous-catégorie poss.

Les formes *mon, ton, son* etc. sont toujours D.

Pour *leur*: voir supra

Leurs est soit D (plur) soit Pro possessif (après Det défini)

On a regardé si *leurs* commute avec un autre Det (*mon, ton*) ou un autre Pro (*sien, tien*) .
Après un déterminant défini, il est Pro.

Leurs:Dfp premières:Afp armes:NCfp
Je:CL1ms me:CL1ms souviens:VP1s de:P les:Dmp leurs:PRO3mp.

On a indiqué le genre (selon l'accord pour le Déterminant, selon l'antécédent pour le Pronom).
On ajoute la personne et le nombre du possesseur selon l'ordre:

Pers-Genre-Nombre-Nombredupossesseur :

mon : D-poss-1mss
notre : D-poss-1msp

Les formes *mien, tien, siennes* sont A (rares: attribut ou épithète gauche) ou PRO (après D défini)

un mien:Ams cousin:NCms
il est mien:Ams
le:Dms mien:PRO3ms, le:Dms nôtre:PROms

1.4.9 Les quantifieurs : beaucoup, trop, peu, assez, bien, tant, tellement, moins

Ils sont analysés comme adverbes (*dormir beaucoup, beaucoup de gens*) ou comme pronoms (*j'en veux beaucoup*) .

Ils sont ADV en tête de groupe nominal indéfini (tours quantitatifs), ou comme adverbe de fréquence ou d'intensité:

beaucoup:ADV de:P gens:NCmp
beaucoup:ADV de:P vin:NCms
beaucoup:ADV d':P autres:PROmp
il aime beaucoup:ADV la:Dfs musique:Nfs
bien:ADV des:Dmp travaux:NCmp
Il mange moins:ADV que:CS toi:PRO2ms
moins:ADV de:P farine:NCfs
ce que j'aime le:Dms moins:ADV

Ils sont PRO (plur, masc. ou fem) en tête de groupe nominal défini (tours partitifs), comme attributs et employés seuls comme sujet ou complément (avec en) :

j':CLS1ms en:CL03fp veux:VP1s beaucoup:PRO3fp (de fleurs)
 beaucoup:PRO3mp de:P ces:Dmp gens:NCmp
 beaucoup:PRO3mp considèrent que...
 beaucoup:PRO3mp d':P entre:P eux:PRO3mp
 peu:PRO3mp demandent:VP3p leurs:Dmp droits:NCmp
 Ils sont beaucoup:PRO3mp

peu

le peu de:Dfs vin qu'il boit ne peut lui faire du mal.
 Il a bu un peu:ADV de:P bière:NCfs.
 Il a bu peu:ADV de:P vin.
 Il est un peu:ADV malade.
 très:ADV peu:ADV d:P affaires:NCfp
 très:ADV peu:PRO3mp sont diffusés
 il viendra sous peu:ADV
 L'eau coule peu à peu:ADV
 Un peu:ADV plus:ADV vite:ADV

tant

il:CLS3ms travaille:VP3s tant:ADV
 tant:ADV de:P fleurs:NCfp
 tant qu':CS on:CLS3ms le:CL03ms paie:VP3s
 tant:CS il:CLS3ms paraît:VP3s en:CL03ms avoir:VW assez:ADV

1.4.10 Les nombres

Les nombres sont écrits en lettres (*vingt-sept*) ou en chiffres (27) . Ils ont normalement été regroupés (*100 000, quatre vingt-sept*) .

4 catégories sont possibles: D (plur), A (sing ou plur), AC (pour neuf), N ou PRO.

Les critères de choix sont les suivants:

En tête de groupe nominal, le nombre est D (pluriel, masculin ou féminin) :

une hausse de 40:Dmp %:NCmp
 Trois:Dfp fois:Nfp
 Tu:CL2ms as:VY2s 3:Dfp gomme:NCfp et:CC moi:PROms 4:Dfp
 à:P zéro:Dfs heure:NCfs
 le:Dms zéro:Ams défaut:NCms n':ADV existe:VP3s pas:ADV
 c':CL3ms est:VP3s un:Dms zéro:NCms

En emploi épithète ou attribut, le nombre est A (normalement pluriel, masculin ou féminin) :

Cas d'ellipse du nom, on assigne le genre selon le contexte.

Les:Dmp trois:Amp brigands:NCmp
 Il n'y avait que deux:Dfp solutions:NCfp , Les:Dfp deux:Afp ont échoué.
 Ils:CLS3mp sont:VP3p tous:AImp deux:ACmp venus:VKmp me:CL01ms voir:VW
 Ils:CLS3mp sont:VP3p quatre:ACmp

Il:CLS3ms est:VP3s numéro:NCms deux:ACms
 Un:Dms produit:NCms de:P 1,6:Dms quintal:NCms l':Dms hectare:NCms
 Dans:P les:Dmp 1,4:Dms %:NCms

Employé seul (sujet ou objet ou complément de Prep) , le nombre est PRO (pluriel, masc. ou fem. selon le contexte) :

(il y a 24 Présidents) Trois:PROC3mp viendront:VF3p.
 (Il y a 4 livres sur la table) J':CLS1ms en:CL3mp veux:VPls trois:PROC3mp.
 Il:CLS3ms mange:VP3s comme:P quatre:PROC3mp
 Tous:AImp trois:PROC3mp viendront
 50:PROC3mp de:P les:Dmp 60:ACmp producteurs:NCmp
 un:Dms litige:NCms sur:P dix:PROC3mp

Employé seul en interprétation métalinguistique, le nombre est NC (sing ou plur, masc) :
 C'est RARE !

Il:CL3ms y:CL3ms avait:VI3s un:Dms quatre:NCms écrit:VKms sur le mur

1.4.10.1 Les dates

Les noms de jour et de mois sont NC (même avec une majuscule) , les nombres sont normalement adjectifs sauf les nombres désignant une année qui sont NC (fs par défaut sinon ms). Les nombres adjectifs épithètes s'accordent en genre et en nombre avec le nom tête.

demain, aujourd'hui, hier sont des ADV.

un jour, ce matin ne sont pas regroupés en adverbe composé, de même que *cette fois, une fois* sauf *une fois* dans une participiale, qui est CS composée :

une fois:CS Jean:NPms arrivé:VKms

le:Dms 1er:Ams juillet:NCms 1924:NCfs
 le:Dms deux:Ams est:VP3s un:Dms dimanche:NCms.
 en:P 1924:NCfs
 tous:Amp les:Dmp trois:Amp juillet:NCms
 le:Dms deux:Ams de:P chaque:Dms mois:NCms
 en:P juillet:NCms 1924:NCfs
 l':Dfs année:NCfs 1956:NCfs
 l':Dms an:NCms 2000:NCms
 vers:P les:Dfp une:Dfs heure:NCfs
 les:Dfp années:NCfp 80:NCfs
 Samedi:NCms 3:Ams juillet:NCms
 courant:P décembre:NCms

1.4.10.2 Les heures

Les heures notées en toutes lettres sont décomposées. Les heures notées en chiffres sont regroupées (2h15, 3h15mn13s, 3h15mn13'20") .

Les noms *demi-heure* et *quart d'heure* sont regroupés.

Deux:Dfp heures:NCfp et:CC demie:NCfs plus:ADV tard:ADV il:CLS3ms
 revenait:VI3s
 Il est deux:Dfp heures:NCfp et:CC demie:NCfs
 Vers:P 2h30:NCfp
 ça:PROms fait:VP3s une:Dfs demi-heure:NCfs que:CS je:CLS1ms l':CLO3ms
 attends:VP1s
 L':Dfs horloge:NCfs sonne:VP3s les:Dmp quarts:NCmp et:CC les:Dfp demi-
 heures:NCfp après:P 20:Dfp heures:NCfp.

1.4.10.3 Les adresses

Comme pour les dates, on se rapproche de la structure habituelle des groupes nominaux sauf que le numéro de la rue est un NC (on dit "au 1" et non "au 1er"):

le:Dms 10:NCms , rue:NCfs Santeuil:NPms
 à:P le:Dms 10:NCms de:P ta:Dfs rue:NCfs
 18:NCms , rue:NCfs de:P Paris:NPms 34000:NCms Montpellier:NPms
 rue:NCfs de:P Charenton:NPs
 Paris:NPms 5ième:Ams (ellipse de arrondissement:NCms)
 dans:P le:Dms 5ième:Ams

1.4.10.4 Les numéros de téléphone

Les numéros de téléphone sont des noms communs :

j':CLS1s ai:VP1s appelé:VKms le:Dms 01-44-23-56:NCms
 Tél.:NCms : (514) 234 1234:NCms
 Tél.:NCms : 42 12 34 56 78:NCms

1.4.10.5 Les scores

Ce sont normalement des noms communs, tenir compte du contexte syntaxique:

3-5:NCms est un score minable.
 Il a été battu 5-7:NCms, 3-0:NCms, 7-0:NCms.

1.4.10.6 Les notes et numéros de chapitres

Ils sont NC, sauf en épithète droit:

voir:VW page:NCfs 5:Afs
 1.2:NCms De:P la:Dfs linguistique:NCfs

Les appels de note sont NCms:

(1):NCms

1.4.10.7 Les monnaies

Ils sont NC.

Livre-sterling, escudo-portugais, franc-suisse, dollars-canadien, etc. sont des noms composés, même en abrégé.

200:Dmp FF:NCmp
 trois:Dfp Livres-sterling:Dfp font:VP3p environ:ADV trente:Dmp francs
 français:NCmp

30:Dmp Kg:NCmp
 il:CLS3ms est:VP3s marqué:VKms \$:NCmp 30:NCmp
 Cet:Dms avion:NCms a:VP3s coûté:VKms environ:ADV 36:Dmp MFF:NCmp

1.4.10.8 Les partitions

quart, tiers, moitié, centième etc. sont toujours NC.
 demi (e) est NC en coordination, A avant le Nom:

une demie:Afs douzaine:NCfs
 une :Dfs heure:NCfs et:CC demie:NCfs
 une heure:NCfs et:CC quart:NCms

1.4.10.9 Cas particuliers

UN (E) (S)

D (sing), A (sing) ou PRO (sing ou plur).

Quand *un (e) (s)* est employé seul après un déterminant, il est PRO:

Les:Dmp uns:PROI3mp ont:VP3p tort:NCms
 L':Dfs une:PROI3fs de:P mes:Dfp amies:NCfp
 une:Dfs belle:Afs fille:NCfs
 le:Dms tome:NCms 1:ACms
 Dieu:NCms est:VP3s un:Ams

attention: *la une, une une* (en presse) => NCfs

Neuf: A ou AC voir infra

Nombres en épithète droit

A sing ou pl, masc. or fem. selon accord avec le N

Louis:NPms XVI:Ams, Elizabeth:NPfs II:ACfs
 le:Dms volume:NCms II:Ams, le:Dms tome:NCms 3:ACms

1.4.11 Les mots étrangers

Nous étiquetons « ET » les mots étrangers qui ne sont pas entrés dans la langue française et qui sont employés dans un contexte syntaxique peu clair. Sinon, ils ont les mêmes étiquettes que les mots français.

Un:Dms match:NCms de:P football:NCms.
 Un:Dms building:NCms
 La:Dfs condition:NCfs sine qua non:Afs
 El:ET contemporaneo:ET.
 The:ET history:ET

1.4.12 Les signes de ponctuation

Les signes de ponctuation ont l'étiquette PONCT avec la sous-cat S ou W (pour « strong » et « weak »).

Les ponctuations fortes (ou PONCTS) sont :

Le point, le point d'interrogation, le point d'exclamation, les points de suspension.

Les autres ponctuations sont des ponctuations faibles

La virgule, le point-virgule, le deux-points, le tiret d'incise, les guillemets, parenthèses et crochets.

« + » est une préposition comme opérateur arithmétiques de même que le « moins ». Placés devant un nombre pour en indiquer la valence (positive ou négative), ce sont des adverbes.

- :ADV 20 :D % :NCmp

% est un nom commun masculin.

« - » est soit une ponctuation faible, soit une partie de mot composé qu'il convient de ne pas détacher.

La virgule est toujours une ponctuation faible, même lorsqu'elle prend le sens d'une conjonction de coordination.

L'apostrophe est toujours rattachée au mot qui la précède, le tiret et le *t* euphonique sont toujours rattachés au clitique inversé.

C'est-à-dire:P -:PONCTW comme:P le:CL3ms pense:VP3s Mr:NCms Bliote:NPms -
:PONCTW les affaires ont changées .:PONCTS

L':CL3ms a:VP3s -t-il:CL3ms rencontré:VKms ?:PONCTS

Ici:ADV l'on:CL3ms dit:VP3s «:PONCTW requinqué:VKms »:PONCTW .:PONCTS

Il:CL3ms offrira:VF3s 120:Dmp \\$:NCmp +:P 21:Dmp %:NCmp de:P taxes:NCfp
.:PONCS

(:PONCTW 9,75:Dmp %:NCmp ,:PONCTW soit:CC +:ADV 1,8:Dms point:NCms)
:PONCTW

1.5 Récapitulatifs

1.5.1 Récapitulatif Pronoms personnels (et Clitiques)

Forme	Etiquette 1	Etiquette 2	Etiquette 3	Etiquette 4
c'	CLS3ms			
ce, -ce	CLS3ms	PRO3ms	Ddefms	
elle	CLS3fs	PRO3fs		
-elle, -t-elle	CLS3fs			
elles	CLS3fp	PRO3fp		
-elles, -t-elles	CLS3fp			
en	CL03fs	CL03ms	CL03mp	CL03fp
eux	PRO3mp			
ils, -ils, -t-ils	CLS3mp			
je, -je, j'	CLS1fs	CLS1ms		
me, m'	CLO1fs	CLO1fs	CLR1ms	CLR1fs
-leur	CLO3fp	CLO3mp		
leur	CLO3fp	CLO3mp	PRO3mp	PRO3fp
lui	CLO3ms	CLO3fs	PRO3ms	VKms
-lui	CLO3ms	CLO3fs		
moi	CLO1fs	CLO1ms	PRO1ms	PRO1fs
-moi	CLO1fs	CLO1ms	NCms	
nous	CLS1fp ou mp	CLO1fp ou mp	CLR1fp ou mp	PRO1mp ou fp
-nous	CLS1fp ou mp	CLO1fp ou mp	CLR1fp ou mp	
l'on, on, (-t)-on	CLS3ms	CLS3fs		
s'	CLR3ms	CLR3fs	CLR3mp	CLR3fp
se	CLR3fs	CLR3ms	CLR3fp	CLR3mp
soi	PROR3ms	PROR3fs		
toi	PRO2ms	PRO2fs		
-toi	CLR2fs	CLR2ms		
te, t'	CLO2fs	CLO2ms	CLR2fs	CLR2ms
tu	CLS2fs	CLS2ms	VKms	
-tu	CLS2fs	CLS2ms		
vous	CLS2fp ou mp	CLO2fp ou mp	CLR2fp ou mp	PRO2mp ou fp
-vous	CLS2fp ou mp	CLO2fp ou mp	CLR2fp ou mp	
y	CLO3ms	CLO3fs	CLO3fp	CLO3mp

1.5.2 Récapitulatif Mots démonstratifs

Forme	Etiquette 1	Etiquette 2	Etiquette 3
ça	PRO-dem-3ms		
-ci, -là	ADV		
c'	CLS3ms		
ce	CLS3ms ou CLO3ms	PRO-dem-3ms	D-dem-ms
ceci, cela, ça	PRO-dem-3ms		
celle-ci, celle-la	PRO-dem-3fs		
celles-ci, celles-la	PRO-dem-3fp		
celui-ci, celui-la	PRO-dem-3ms		
ces	D-dem-mp		
cet	D-dem-ms		
cette	D-dem-fs		

ceux-ci, ceux-là	PRO-dem-3mp		
voici, voilà	P	VP3s	

1.5.3 Récapitulatif mots indéfinis

Forme	Etiquette 1	Etiquette 2	Etiquette 3	Etiquette 4	Etiquette 5	Etiquette 6
autre	PRO-indef-ms	PRO-indef-fs	A-indef-fs	A-indef-ms	A-qual-ms	A-qual-fs
autres	PRO-indef-mp	PRO-indef-fp	A-indef-mp	A-indef-fp	A-qual-mp	A-qual-fp
certain	A-qual-ms	A-indef-ms				
certaine	A-qual-fs	A-indef-fs				
certains	A-qual-mp	A-indef-mp	D-indef-mp	PRO-indef-3mp		
certaines	A-qual-fp	A-indef-fp	D-indef-fp	PRO-indef-3fp		
chacun	PRO-indef-3ms					
chacune	PRO-indef-3fs					
chaque	D-indef-ms	D-indef-fs				
de	P	D-indef-ms	D-indef-fs	D-part-ms	D-part-fs	D-indef mp ou fp
de la	D-part-fs					
du	D-part-ms	P+D-def-ms				
des	D-indef-mp	D-indef-fp	P+D-def-mp	P+D-def-fp	P+D-indef-mp	P+D-indef-fp
différent	A-qual-ms	A-indef-ms	NCms			
différente	A-qual-fs	A-indef-fs				
différents	A-qual-mp	A-indef-mp	D-indef-mp	NCmp		
différentes	A-qual-fp	A-indef-fp	D-indef-fp			
divers	A-qual-mp	A-indef-mp	D-indef-mp			
diverses	A-qual-fp	A-indef-fp	D-indef-fp			
même	A-indef-ms	A-indef-fs	A-qual-ms	A-qual-fs	ADV	
mêmes	A-qual-mp	A-qual-fp	A-indef-mp	A-indef-fp		
quelqu'un	PRO-indef-3ms					
quelque chose	PRO-indef-3ms	NCms				
quelque	D-indef-ms	D-indef-fs	ADV			
quelques	D-indef-mp	D-indef-fp	A-indef-mp	A-indef-fp		
seul	A-qual-ms	A-indef-ms				
seule	A-qual-fs	A-indef-fs				
seuls	A-qual-mp	A-indef-mp				
seules	A-qual-fp	A-indef-fp				
tel	D-indef-ms	A-indef-ms	PRO-indef-3ms			
telle	D-indef-fs	A-indef-fs	PRO-indef-3fs			
telles	D-indef-fp	A-indef-fp				
tels	D-indef-mp	A-indef-mp				
tout	PRO-indef-3ms	D-indef-ms	ADV	NCms	A-indef-ms	
toute	D-indef-fs	A-indef-fs	ADV			
tous	PRO-indef-3mp	D-indef-mp	A-indef-mp	ADV		
toutes	PRO-indef-3fp	D-indef-fp	A-indef-fp	ADV		

1.5.4 Récapitulatif Mots négatifs

Forme	Etiquette 1	Etiquette 2	Etiquette 3	Etiquette 4
aucun	D-neg-ms	PRO-neg-3ms	A-neg-ms	
aucune	D-neg-fs	PRO-neg-3fs	A-neg-fs	
guère	ADV-neg			
jamais	ADV-neg			
n', ne	ADV-neg			
non	ADV-neg	I		
nul	PRO-neg-3ms	A-qual-ms	D-neg-ms	
nulle	PRO-neg-3fs	A-qual-fs	D-neg-fs	
nulles	A-qual-fp	D-neg-fp	PRO-neg-fp	
nuls	A-qual-mp	D-neg-mp	PRO-neg-mp	
pas	ADV-neg	NCms		
personne	PRO-neg-3ms	NCfs		
plus	ADV-neg	ADV	VKmp	
point	ADV-neg	NCms	VP3s	
que	ADV-exclam	CS	PRO-inter-3ms	PRO-rel-ms ou fs ou fp ou mp
rien	PRO-neg-3ms	ADV	NCms	

1.5.5 Récapitulatif Mots relatifs, interrogatifs, exclamatifs

Forme	Etiquette 1	Etiquette 2	Etiquette 3	Etiquette 4	Etiquette 5
combien	ADV-inter	ADV-exclam			
dont	PRO-rel-1ms ou mp ou fs ou fp	PRO-rel-2ms ou mp ou fs ou fp		PRO-rel-3ms ou mp ou fs ou fp	
duquel	P+PRO-rel-3ms	P+PRO-inter-3ms			
desquels	P+PRO-rel-3mp	P+PRO-inter-3mp			
desquelles	P+PRO-rel-3fp	P+PRO-inter-3fp			
auquel	P+PRO-rel-3ms	P+PRO-inter-3ms			
auxquels	P+PRO-rel-3mp	P+PRO-inter-3mp			
auxquelles	P+PRO-rel-3fp	P+PRO-inter-3fp			
lequel	PRO-rel-3ms	PRO-inter-3ms			
laquelle	PRO-rel-3fs	PRO-inter-3fs			
lesquelles	PRO-rel-3fp	PRO-inter-3fp			
lesquels	PRO-rel-3mp	PRO-inter-3mp			
que	PRO-rel-3ms ou fs ou mp ou fp	PRO-inter-3ms	PRO-rel-2ms ou fs ou mp ou fp	PRO-rel-1ms ou fs ou mp ou fp	ADV-neg ou CS ou ADV-neg
qui	PRO-rel-3ms ou fs ou mp ou fp	PRO-rel-2-ms ou fs ou mp ou fp	PRO-rel-1-ms ou fs ou mp ou fp	PRO-int-3ms	
quel	D-inter-ms	D-exclam-ms	A-inter-ms		
quelle	D-inter-fs	D-exclam-fs	A-inter-fs		
quels	D-inter-mp	D-exclam-mp	A-inter-mp		
quelles	D-inter-fp	D-exclam-fp	A-inter-fp		
quoi	PRO-rel-3ms ou fs	PRO-rel-3mp ou fp	PRO-inter-3ms		
quand	ADV-inter	CS			
où	ADV-inter	PRO-rel-3ms ou fs	PRO-rel-3fp ou mp		
comme	ADV-exclam	CS	P	CC	

1.5.6 Récapitulatif Mots coordonnants

Forme	Etiquette 1	Etiquette 2	Etiquette 3
car	NCms	CC	
donc	ADV	CC	
et	CC		
ni	CC		
ou	CC		
or	CC	NCms	
soit	CC	VS3s	
c'est-à-dire	CC		
voire	CC		
sinon	CC	ADV	
mais	CC		
comme	CC	CS	P
tantôt	CC	ADV	
y compris	CC	P	
puis	CC	ADV	

1.5.7 Récapitulatif signes brachygraphiques

Forme	Etiquette 1	Etiquette 2
" « »	PONCTW	
()[]{}	PONCTW	
(...) [...]	PONCTS	
/ &	PONCTW	CC
? !	PONCTS	
. ...	PONCTS	
, ; - :	PONCTW	
% § \$ (1)	NCms ou mp	
£	NCfs ou fp	
= / * ≠ < >	P (opérations)	
+ -	P (opérations)	ADV (signe négatif ou positif)

INDEX DES MOTS SIMPLES

-	33		
%	33		
+	33		
adjectif indéfini	12		
adjectif qualificatif	12		
adresse	31		
adverbe exclamatif	13		
adverbe interrogatif	25		
adverbe négatif	13		
adverbe quantifieur	13		
adverbe simple	13		
afin_de			
Prep	19		
ainsi			
Adv	4		
amalgame	4		
lemme sans graphie	8		
ambiguïté	16		
adjectif / adverbe	18		
adjectif / nom commun	17		
adjectif / participe passé	16		
adjectif / participe présent	17		
adjectifs qualificatif / cardinaux / indéfinis	20		
clitiques sujet / objets / réfléchis	24		
conjonctions subordination / coordination	21		
noms communs / noms propres	21		
préfixe / adverbe	20		
préposition / adverbe	19		
pronoms relatifs / pronoms interrogatifs	25		
ambiguïté entre sous-catégories	20		
aucun	27		
aujourd'hui	31		
autre	28		
auxiliaire	5		
avant			
Adj	6, 7		
Adv	19		
Prep	19		
avant_de			
Prep	19		
bas	18		
beaucoup	29		
bien	18, 29		
ce / c' / cet	8		
ce / c' / ceci / cela	26		
celui-ci / celui-là	26		
certain (e) (s)	28		
cet (te)	26		
-ci	26		
clitique	5, 23, 34		
codage XML	5		
comme	9		
complémenteur	5		
conjonction de coordination	21		
connecteur	5		
courant			
Prep	31		
date	31		
de / d'	9		
de_la			
Det	4		
demain	31		
demi (e)	32		
demi-heure	31		
démonstratif	26		
depuis_que			
CS	19		
des	9, 11		
déterminant composé partitif	10		
déterminant indéfini	10		
déterminant interrogatif	25		
déterminant partitif	10		
du	9, 11		
amalgame de le	4		
Det	4		
durant			
Prep	19		
en	11		
pronom figé	7		
ex-			
Pref	20		
faux	18		
fois	31		
fort	18		
haut	18		
heure	31		
hier	31		
hors			
Prep	19		
hors_de			
Prep	19		
il y a	19		
il-y-a			
Prep	19		
indéfini	25		

juste	18	personne	27
le / la / les / l'	11	peu	29
le_plus		plus	13, 27
adv composé.....	13	plus de	13, 19
lemme.....	7	plus_que	
leur	12	adv composé.....	13
loin	18	polysème homographe.....	7
loin_de	19	ponctuation.....	33
l'on	4	possessif.....	28
lui	12	pour	19
mal	18	pour_que	
marque de l'addition.....	13	CS.....	19
marron		prédéterminant.....	5
Adj.....	6	préfixe.....	20
même	12	prémodifieur d'adjectif.....	7
mêmes	12	prénom composé.....	22
mien / tien / sien	29	près	18
-moi	4	près_de	19
moins	29	pronom figé.....	7
monnaie.....	32	pronom interrogatif.....	25
mot.....	37	pronom personnel.....	23, 34
mot coordonnant.....	38	pronom personnel faible.....	5
mot démonstratif.....	26, 34	pronom relatif.....	25
mot étranger.....	6, 33	qu'	13
mot indéfini.....	25, 27, 35	quand	25
mot interrogatif.....	25, 37	quantifieur.....	29
mot invariable.....	6	quart	32
mot négatif.....	27, 37	quart_d'heure	31
mot possessif.....	28	quasi	20
mot relatif.....	25, 37	que	13
n'importe + qui/ que / quoi / quel /		CS.....	4
lequel	25	quel	25
net		quelqu'un	28
Adj.....	7	quelque (s)	28
neuf	33	quelque chose	28
nom de famille.....	4	qui / que / quoi	25
nom propre.....	21	recatégorisation.....	7
genre.....	7	rien	27
nombre.....	30	s' / se	14
nombre cardinal.....	6	s' / si	14
nombres séparés par des blancs.....	4	sauf	18
non	20	sauf_à	18
nôtre / vôtre	29	score.....	32
nul	27	si	
numéro de téléphone.....	31	Adv.....	14
où	25	CS.....	14
outré	20	NC.....	14
outré-	20	sigle.....	22
parties du discours.....	5	signe brachygraphique.....	38
partitif.....	4	sous-catégorie.....	5
pendant_que		super-	
CS.....	19	Pref.....	20

t euphonique	4, 24, 33	Adv	7
tant	30	un (e)	32
tard		voici/voilà	
Adv	31	Prep	19, 26
tel(s)	15	VP3s	19
telle(s)	15	vous de politesse	24
-t-il	4	y	
tous	15	pronom figé	7
tout(e)(s)	15		
toute			

2 LES MOTS COMPOSES

2.1 Introduction

Les mots composés ont été annotés automatiquement dans le texte à l'aide :

- d'un programme basé sur des expressions régulières pour reconnaître les nombres écrits en chiffres (*mille deux cent quatre vingt-seize*),
- d'un dictionnaire pour les autres mots composés (incluant notamment des extraits des dictionnaires accessibles sous INTEX).

Dans le cas où plusieurs mots composés étaient candidats (exemple le nom composé "face à face" ou la préposition composés "face à") c'est le plus long qui a été privilégié.

Le travail de validation a consisté à :

- décomposer les candidats non pertinents en contexte, par exemple on n'a pas l'adverbe composé "sur ce" dans la séquence "sur ce point", on n'a pas la préposition composée "il y a" dans "il y a beaucoup d'actions à vendre".
- ajouter les mots composés absents du dictionnaire (surtout les noms propres)
- modifier la catégorie de certains composés qui étaient mal catégorisés dans le dictionnaire (par exemple, nous nous sommes aperçu que le dictionnaire avait très peu d'adjectifs composés car ils étaient marqués à tort Adverbe).

Les mots composés sont notés par la suite de leur composants séparés par des blancs, des tirets ou des apostrophes: *c'est-à-dire, aujourd'hui, pomme de terre*.

2.1.1 Le codage XML

La syntaxe d'un mot composé est la suivante (un seul lemme et plusieurs mots):

[lemme, Catégorie, Sous-catégorie, traits morphologiques]+[Catégorie des composants]

Un attribut `compound="yes"` est au début du mot composé, un attribut `catint` note la catégorie interne de chaque composant.

Exemples :

Côte-d'Ivoire

```
<w compound="yes" lemma="Côte-d'ivoire" cat="N" subcat="P" mph=fs">
<w catint="N">Côte</w>
<w catint="PONCT">-</w>
<w catint="P">d'</w>
<w catint=N>Ivoire</w> </w>
```

faute de

```
<w compound="yes" lemma="faute de" cat="P">
<w catint="N">faute</w>
<w catint="P">d'</w> </w>
```

Il y a de rares cas d'insertion dans un mot composé. Ces insertions sont notées avec leur étiquettes, comme pour les mots simples. Chaque partie de mot composé ainsi séparée a un identifiant (id) et un attribut `next` (pointant vers la partie suivante) ou `prev` (pointant vers la partie précédente). Les identifiants codent le nombre de parties et le numéro de partie.

faute justement de...

```
<w compound="yes" id="21" lemma="faute de" cat="P" subcat="" next="22">
<w catint="N">faute</w></w>
<w cat="ADV" lemma="justement" subcat="">justement</w>
<w compound="yes" id="22" lemma="faute de" cat="P" subcat="" prev="21">
<w catint="P">de</w></w>
```

A trop bon compte

```
<w compound="yes" id="21" lemma="à bon compte" cat="ADV" subcat="" next="22">
<w catint="P">à</w></w>
<w cat="ADV" lemma="trop" subcat="">trop</w>
<w compound="yes" id="22" lemma="à bon compte" cat="ADV" subcat="" prev="21">
<w catint="A">bon</w>
<w catint="N">compte</w></w>
```

2.1.2 Le Lemme

Le lemme du mot composé est la forme morphologiquement non marquée du mot composé (Les noms et adjectifs sont au masculin singulier, les verbes à l'infinitif).

2.1.3 Catégorie, traits morphologiques

Les notations des catégories et traits morphologiques des composés sont les mêmes que pour les mots simples (voir supra).

2.1.4 Les Composants

Les catégories des composants sont notées par les abréviations suivantes :

- A : Adjectif
- ADV : Adverbe
- C : Conjonction
- D : Déterminant
- N : Nom
- P : Préposition
- PRO : Pronom
- PONCT : ponctuation
- V : Verbe
- CL : Clitique
- ET : Mot étranger

Exemples:

```
douche écossaise:NCms+NA (Nom Adjectif)
belle-fille:NCfs+AN
premier ministre:NC+AN (Adjectif Nom)
pomme de terre:NCms+NPN
chair à pâté:NCfs+NPN (Nom Préposition Nom)
homme-grenouille:NCms+NN (Nom Nom)
garde-chasse:NCms+VN (Verbe Nom)
poil à gratter:NCms+NPV (Nom Préposition Verbe)
autant que possible:ADV+ADVCA (Adverbe Conjonction Adjectif)
en effet:ADV+PN contre attaque:NCfs+PN (Préposition Nom)
à l'insu de:P+PDNP (Préposition Déterminant Nom Préposition)
```

2.2 Les critères pour déterminer les mots composés

Nous avons tenté une synthèse des travaux existants sur le sujet (cf Gross 1996, Gross 1989). Ces critères ne sont ni nécessaires ni suffisants. Ils permettent simplement d'établir un faisceau de caractères du mot composé.

Nous avons choisi d'être extensifs sur les mots composés grammaticaux (qui facilitent les analyses ultérieures), et restrictifs sur les mots composés lexicaux (qui sont plus controversés).

2.2.1 Critères morphologique, graphique

L'un des composants n'apparaît que dans l'expression complexe.

au fur et à mesure
a priori
aujourd' hui
à l'insu de
faire fi de
par monts et par vaux
la plupart

Le genre ou le nombre d'un composant change dans l'expression complexe.

une deux-chevaux
un peau-rouge

Présence d'un tiret ou d'une apostrophe.

entr'ouvert
garde-barrière

2.2.2 Critères syntaxiques

Le mot composé a la même distribution qu'une catégorie terminale.

En vertu de, près de sont des prépositions tout comme *à, de*.
un tantinet, un peu sont des adverbes comme *environ*

La catégorie syntaxique de l'ensemble ne correspond pas à la séquence de catégories des constituants:

peut-être:ADV+VV
c'est-à-dire:CC+CVPV
rendez-vous:N+VCL
garde-côtes:N+VN

La séquence des catégories des constituants d'un mot composé n'existe pas par ailleurs.

y compris:CC = Clitique + Verbe au participe passé
à la va-vite:ADV = Prep + Det + Vconj + Adv

Les composants sont contigus. Seule quelques petites insertions sont possibles (en général un petit adverbe ou adjectif).

à force de
 un maillot <doré> deux-pièces
 ?? un maillot <de ma soeur >deux pièces
 à l'insu <justement>de ses voisins
 ?? à l'insu <comme on dit > de ses voisins

Les règles d'accord ne sont pas respectées :

Une grand-mère
 *Une grande-mère

2.2.3 Critères sémantiques

La sémantique du mot composé n'est pas la composition sémantique du sens de ses composants.

une sage-femme n'est pas une femme qui est sage.
 carte bleue:NCfs+NA → différent de « une carte qui est bleue »

On ne peut pas remplacer un des constituants par un synonyme ou un antonyme.

à bas *à haut
 à la va vite *à la va rapidement

Le mot composé peut correspondre à une seule unité dans d'autres langues.

Exemples :

une pomme de terre => a potato / una patata
 en arrière => behind / backward / indietro
 y compris => including

2.2.3.1 Préfixe ou mot composé ?

L'étiquette PREF est employée pour catégoriser des mots du type : *méta*, *anti*, *archi*, suivis d'un tiret ou non puis d'une liste ouverte de termes. Le préfixe n'est pas composant d'un terme lexical, il est productif.

Sont des mots composés:

La géo-histoire:NCfs+NN
 Italo-belge:Ams+AA
 Un sous-marin:NCms+PN
 Un trop-plein:NCms+PN

Ne sont pas des mots composés:

archi:PREF - sûr:Ams
 Outre:PREF - Atlantique:NPms.
 Outre:PREF - Rhin:NPms.
 La peinture:NCfs méta:PREF - classique:Afs

2.3 Les différents types de mots composés

Toutes les catégories (sauf les ponctuations, les préfixes et les interjections) peuvent donner lieu à des composés.

2.3.1 Les noms composés

Attention: les noms propres composés n'ont pas forcément de morphologie renseignée.

Les noms composés sont nombreux dans le corpus, ceux qui ne doivent pas être considérés comme composés (en contexte) ont normalement été ôtés.

le:Dms Ministère-de-l'intérieur:NCms+NPN

mais pas forcément de mot composé dans:

Il connaît le:Dms Ministère:NCms de:P l':Dms intérieur:NCms (il peut s'agir d'un autre ministère par exemple)

Il y a quelques occurrences avec interposition, par exemple:

maillot deux pièces:NCms+NDN

maillot <doré:Am>deux pièces:NCms

le Ministre des Finances:NCms

le Ministre<belge:Ams>des Finances:NCms

On n'a pas forcément factorisé les mots composés coordonnés:

les ministres des finances et de l'agriculture

Seul *ministres des finances* est noté comme composé.

2.3.1.1 Les noms propres composés

Prénoms composés avec ou sans tiret, noms de famille composés avec ou sans particule, noms de journaux avec ou sans article.

Paul:NPms Dupont de Nemour:NPms

Valéry:NPms Giscard d'Estaing:NPms

le journal *Le Monde*:NPms

Titres de film, de livre, etc

Quand le titre complet forme un syntagme long, on ne fait pas de mot composé. La limite est parfois délicate à fixer.

pas de mot composé:

« Tout:PROms ce:PRO3ms que:PROR3ms vous:CLS2p avez:VP2p voulu:VYms
savoir:VW sur:P le:Dms sexe:NCms sans:ADV jamais:ADV oser:VW le:PRO3ms
demander:VW »

mot composé:

le petit Poucet:NPms

Noms de villes, de pays, de sociétés, d'institutions

On rassemble l'ensemble du nom, dans une certaine limite (lorsqu'il s'agit sûrement d'une entrée lexicale).

Renseigner le genre s'il est clair dans le contexte (ex. : *Paris libéré*). Ne pas renseigner le genre si vous l'ignorez. Grevisse indique masculin comme genre par défaut pour les villes et féminin pour les îles. Le genre des sociétés peut rester non marqué (NPs).

exemples:

Aix-en-Provence:NPfs+NPN
 Europe de l'Est:NPfs+NPDN
 île de Ré:NPfs+NPN
 les Etats-Unis:NPmp+NA
 Berlin-Est:NPs+NN
 Le Monde:NPms+DN
 Le Figaro:NPms+DN
 Air-France:NPs+NN
 Université:NCfs Paris VII:NPfs+NA
 le:Dms Parti Socialiste:NPms+NA
 l'Association:NCfs Bulgare:Afs de:P Libération:NCfs des:P+Dmp Nains:NCmp
 de:P Jardin:NCms (ABLNJ:NPfs)

2.3.1.2 Les noms communs

Les critères sont plus délicats à observer. On a essayé d'être cohérents sur les séries.

L'opacité du nom est un critère :
 un *point-d'orgue* n'est pas un point.

On se fie également au caractère figé de l'adjectif :

mauvaise langue:NCfs+AN *bonne langue:NCfs+AN

mais :

mauvaise:Afs humeur:NCfs bonne:Afs humeur:NCfs

Caisse d'épargne:NCfs+NPN

Caisse de retraite:NCfs+NPN

Caisse de crédit:NCfs+NPN

Les noms de métiers, de monnaie sont des noms communs composés. Les types d'élections, les types d'économie, les types d'écoles, de commissions, d'impôts, de syndicats, de journaux également.

Les références complètes de produits sont des noms communs composés.

Le:Dms ministère des finances:NCms

Le:Dms franc belge:NCms

De les:Dfp plaques:NCfp de:P PB 13:NCmp

une:Dfs renault:NCfs R10 turbo 16S:NCfs

le:Dms formulaire:NCms BB 102:NCms

2.3.2 Les pronoms composés

2.3.2.1 Pronoms indéfinis composés

Exemples

quelqu'un:PROms

quelque chose:PROms
 d'aucuns:PROmp
 la plupart : PROmp

2.3.2.2 Pronoms personnels composés

Même en suffixe ou en épithète droit est A (masc. ou fem, sing ou plur) sauf *moi-même*, *lui-même* etc. qui sont des PRO composés (voir supra).

Exemples

cet:Dms homme:NCms - même:Ams qui:PROR3ms etc
 c'est la prudence:NCfs même:Afs
 aujourd'hui:ADV même:Ams
 j':CLS1ms y:CLO3ms vais:VP1s moi-même:PRO1ms

2.3.2.3 Pronoms démonstratifs composés

Ci en suffixe est ADV. sauf *celui-ci*, *celle-ci* etc qui sont des PRO composés.

exemples

cet:Dms homme:NCms - ci:ADV qui:PROR3ms me:CL1ms regarde:VP3s
 celui-ci:PRO3ms me:CL1ms regarde:VP3s

2.3.3 Les déterminants composés

2.3.3.1 Les déterminants indéfinis composés

Exemple

n'importe quel:Dms+VD

Les formes en *n'importe* sont analysées comme ADV, PRO ou D composés, ce sont des formes indéfinies ni relatives ni interrogatives:

Exemples

n'importe qui:PROms
 n'importe quoi:PROms
 n'importe quel:Dms
 n'importe lequel:PROms
 n'importe où:ADV
 n'importe quand:ADV
 n'importe comment:ADV

2.3.3.2 Les nombres composés

Tous les nombres écrits en lettres ou en chiffres sont des A, D, PRO composés.
 Attention : *Million* et *milliard* sont toujours des noms communs.

Exemples

Il me doit quatre vingt cinq mille six cent soixante-dix:Dmp+D*
 francs:NCmp
 quatre cent:Dmp+DD millions:NCmp de:P francs:NCmp
 quatre cent:Dmp+DD millions:NCmp sont:VP3s venus:VKmp
 ça m'a couté quatre cent:Dmp+DD millions:NCmp

(D* signifie une séquence de D)

Les composants ont la même catégorie (interne) que le composé :

Vingt-deux personnes : D+D
 Les vingt-deux personnes : A+A

2.3.4 Les adjectifs composés

2.3.4.1 Les adjectifs cardinaux et ordinaux composés

Exemple

une:Dfs vingt-et-unième:Afs robe:NCfs
 les:Dfp quatre-vingt-trois:Afp robes:NCfp

2.3.4.2 Les adjectifs qualificatifs composés

Pour savoir si certaines expressions épithètes sont des Adj composés, on teste si elles ont d'autres contextes : emploi adverbial avec un V, emploi comme attribut, emploi nominal avec un D...

Si elles ont un emploi adverbial, on a mis l'étiquette Adv,

Si elles ont un emploi nominal (sans ellipse), on a mis l'étiquette N :

Exemples

une robe bon marché:Afs+AN
 un article haut de gamme:Ams+NPN
 un schéma:NCms au point:ADV+PDN (cf mettre quelque chose au point)
 mardi:NCms après-midi:NCms+PN (cf cet après-midi)
 un emprunt à long terme:ADV+PAN (cf raisonner à long terme)

Les adjectifs composés de deux adjectifs:

Mis à part quelques préfixes très productifs (*auto, anti, contre, ex, pseudo...*), on note la séquence comme Adj composé décomposable en deux adjectifs:

Exemples

anti:PREF - français:A
 syntaxico-sémantique:Ams+AA
 politico-financier:Ams+AA
 italo-helvetique:Ams+AA

2.3.5 Les prépositions composées

(voir Grevisse)

Elles sont nombreuses dans le corpus. On n'a normalement laissé que celles qui sont pertinentes dans le contexte considéré:

Exemples

ils:CLS3mp sont:VP3p partis:VKmp il y a:P 3:Dmp ans:NCmp
 Il:CLS3ms y:CLO3ms a:VP3s vingt:Dmp ans:NCmp que:CS...
 (pas la Prep *il y a*)

A la différence des Adv composés, elles attendent un complément obligatoire, qui est généralement un NP mais peut aussi être un PP ou un Adv.

Exemples

de part et d'autre:Adv
 de part et d'autre de:P
 en face:Adv
 en face de:P
 au-dessus:Adv
 au-dessus de:P

Elle se compose d'une séquence PNP ou PDNP ou ADVP sauf:

Exemples

y compris:P+PV (y compris:P à:P New York:NPms)
 à bas:P+PN (à bas:P le:D communisme:NCms)
 il y a:P+PPV (c:CL3ms était:VI3s il y a:P de les:Dfp années:NCfp)
 compte tenu de:P+NVP
 d'ici:P (d'ici:P trois:Dmp ans:NCmp, d'ici:P peu:Adv)

Les critères de figement sont la présence d'un constituant qui n'existe pas seul (*insu, afin*), l'équivalence avec une Prep simple (*en dépit de* = malgré; *en face de* = devant), la non compositionnalité sémantique (pas de dépit dans *en dépit de*, pas de vertu dans *en vertu de*), le caractère figé du N composant (pas de Det ou pas d'Adj insérables):

Exemples

à cause de:P *à cette cause, *à la cause de, *à une grande cause de...
 à côté de:P *à ce côté, *au côté de

mais pas de prep composée:

aux:P+Dmp côtés:NCmp de:P
 à:P ses:Dmp côtés:NCmp

L'alternance entre les séquences « P le N de » et « P son N » n'empêche pas d'y voir un composé (si les autres critères sont ok):

exemples

à l'insu de:P, à son insu:Adv
 à l'égard de:P, à leur égard:Adv, à cet égard:Adv

Certaines P composées ont un Det ou un Adj figés :

sans commune mesure avec:P (*sans mesure avec, *sans une commune mesure avec)

mais pas:

sans:P mesure:NCfs avec:P mesure:NCfs

Les locutions introduisant des infinitives sont notées Prep (comme pour les mots simples):

exemples

avant de:P
 avant que:CS
 afin de:P
 afin que:CS

2.3.6 Les adverbes composés

voir M. Gross 1986, Syntaxe transformationnelle de l'adverbe

Ils peuvent comporter une séquence de catégories particulière (peut-être:ADV+VV, par contre:ADV+PP) mais ont souvent une structure interne régulière de syntagme prépositionnel (PN ou PDN):

exemples

en revanche:ADV+PN

au contraire:ADV+PDN

La non compositionnalité sémantique est un critère de figement (pas de revanche dans *en revanche*), tout comme la présence d'un composant inexistant par ailleurs (aujourd'hui).

Le figement du Det (ou de l'absence de Det) est aussi un critère, tout comme la non substituabilité du N composant:

exemples:

en effet:ADV (**en un/cet effet*)

en l'occurrence:ADV (**en cette/une occurrence*)

à terre (**à sol...*)

à genoux (**à coudes, *à mains...*)

mais pas :

en:P partie:NCfs car on a : *en grande partie, en totalité...*

en:P avance:NCfs car on a : *en retard*

en:P ce:Dms cas:NCms, car on a : *en aucun cas, dans certains cas...*

par:P hasard:NCms (*par un hasard extraordinaire, par pur hasard, par chance*)

Attention : l'absence de Det est normale après <<en>> et ne constitue pas à soi seul un critère de figement (*en France, en linguistique, en or, en argent...*)

Parfois l'insertion d'un Adj ou d'un ADV (restreint) est possible:

Exemples

au complet:ADV+PDN

au <grand:Ams> complet:ADV+PDN

sans <grand:Ams> complet:ADV+PDN

La présence d'un Adj figé est aussi un critère:

Exemples

de longue date:ADV+PAN (**de courte date, *de date*)

sans commune mesure:ADV+PAN (**sans mesure*)

La plupart des ADV composés ont été corrigés dans le corpus, car ils n'étaient pas forcément pertinents dans le contexte considéré:

Exemple :

d'abord:ADV+PN

mais:

Il:CL3ms a:VP3s un:Dms drôle:Ams d':P abord:NCms

Exemple :

de plus:ADV+PP, en plus:ADV+PP

mais:

de:Dmp plus:Amp grands:Amp espoirs:NCmp

en:P plus:ADV grand:Ams

de plus en plus:ADV+PADVPADV

Exemples

au plus:ADV+PDN (ca vaut au plus:ADV 500:Dmp F:NCmp)

mais :

au:P+Dms plus:ADV haut:Ams niveau:NCms

Exemples

en tout:ADV+PN (*ça fait en tout 500 F*)

mais :

en tout cas:ADV (pas de composé interne à un composé)

au point:ADV+PDN (*c'est au point*)

mais:

au:P+Dms point de vue:NCms des:P+Dmp résultats:NCmp...

2.3.7 Les conjonctions composées

voir Grevisse et M Piot (thèse d'Etat Paris 7 et article *Langue française* Syntaxe des connecteurs 1988)

2.3.7.1 Les conjonctions de subordination

A la différence des ADV ou des Prep composées, elles introduisent une phrase subordonnée (avec un V conjugué):

Exemples

au fur et à mesure:ADV+PDNCPN

au fur et à mesure que:CS+PDNCPNC le vase se vide

Elles se comportent comme une conjonction simple (*parce que = puisque*), elles peuvent inclure un constituant n'existant pas par ailleurs (*parce, fur, tandis*), elles sont non décomposables sémantiquement.

Le figement du Det (ou de l'absence de Det) interne est aussi un critère:

Exemples

à mesure que:CS+PNC (**à la mesure que*)

au cas où:CS+PDNC (**à cas où, *à ce cas, *au cas extrême où*)

mais pas :

à:P l':D instant:NCms où:PROR3ms

car :

productivité du N interne: *au moment où, à la minute où, à l'heure où...*

non figement du Det interne: *à cet instant ;*

insertion d'un Adj: *à l'instant précis où...*

2.3.7.2 Les conjonctions de coordination

Comme pour les mots simples, on ne se limite pas à *mais, ou, et, donc, or, ni, car*:

Exemples :

ainsi que:CC

c'est-à-dire:CC

2.3.8 Les verbes composés

Attention : la plupart sont notés V+, préciser la flexion, ex VW+ ou VP3s+ etc

On a choisi d'en retenir très peu dans le corpus, car la plupart sont discontinus, et suivent une syntaxe régulière.

On a retenu les expressions verbales qui mettent en jeu un composant n'existant pas par ailleurs (*faire fi de*) ou celles qui sont très figées (V N sans déterminant possible: *faire partie de*) et non compositionnelles.

Exemples

avoir beau jeu:VW+VAN de:P (**avoir un beau jeu de*)

mais pas:

avoir peur (avoir une grande peur), avoir faim, avoir soif...

On ne fige pas la Prep à la fin de l'expression, car le V figé ne se comporte pas comme un V transitif direct, mais comme un V prenant un complément prépositionnel (cf possibilité d'un pronom fort comme complément:

faire fi de lui comme *rêver de lui*, et non comme **voir lui*) :

Exemple :

faire fi:VW+VN de:P lui:PRO3ms

Bien noter aussi les insertions:

ne:ADV fait <pas:ADV> fi:VP3s+VN de:P

On ne considère pas les combinaisons clitiques-verbes comme formant un V composé, même pour les clitiques intrinsèques (*s'apercevoir de, en avoir assez de...*)

Références

M Grevisse, A Goosse 1992. Le Bon usage, Duculot, Liège

G Gross 1996. Les expressions figées, Ophrys, Gap

M Gross 1986, Syntaxe transformationnelle de l'adverbe, Cantilène.

M. Piot 1985. Prépositions et conjonctions de subordination

3 ANNEXE : LISTE DES ETIQUETTES

212 étiquettes externes

ADJECTIF

qualificatifs

A-qual-ms	il est INTELLIGENT le PETIT chien de la voisine du fait que le candidat est déclaré SEUL sur la liste Un stylo NEUF
A-qual-mp	Ces enfants sont déjà GRANDS De GRANDS enfants
A-qual-fs	une défaite EVIDENTE une IMPORTANTE victoire
A-qual-fp	trois disquettes ENDOMMAGEES

Cardinal

A-card-fp	les NEUF différences
A-card-mp	les NEUF stylos
A-card-ms	le livre TROIS
A-card-fs	la porte TROIS

Indéfini

A-ind-ms	le MEME cabinet
A-ind-fs	Une SEULE fleur
A-ind-mp	TOUS les ans
A-ind-fp	TOUTES les fleurs

Interrogatif

A-int-ms	QUEL est ce vainqueur ?
A-int-fs	QUELLE est cette fille ?
A-int-fp	QUELLES sont nos chances ?
A-int-mp	QUELS sont les candidats ?

Ordinal

A-ord-ms	le XIXe siècle
A-ord-fs	une SECONDE chance
A-ord-mp	les TROISIEMES trimestres
A-ord-fp	les DIXIEMES saisons

Possessif

A-poss-1fss	Cette pipe est MIENNE
A-poss-2mss	un TIEN cousin
A-poss-1fps	Ces pipes sont MIENNES
A-poss-1fps	Ces cigares sont MIENS
etc	

ADVERBE

ADV Exclamatif

ADV-excl QUE oui ! COMME elle est belle !

ADV négatif

ADV-neg JAMAIS
Jean ne boit PLUS

ADV Interrogatif

ADV-inter COMBIEN sont ils ?
il me demande COMBIEN ils sont
OU vas-tu ? QUAND pars-tu ?

ADV démonstratif

ADV-dem ce garçon-CI

ADV

il est PRESQUE midi
le ministre est ENSUITE nommé
il s'avance LENTEMENT DESSOUS
BIEN des années ont passé
BEAUCOUP de ces étudiants sont diplômés
il en veut DAVANTAGE
DEPUIS, il est parti

CLITIQUESsujet

CL-suj-1ms je, j'
CL-suj-2ms tu
CL-suj-1mp nous
CL-suj-2mp vous
CL-suj-3ms on, ce
CL-suj-3mp ils
CL-suj-3fs elle
CL-suj3fp elles

Objet

CL-obj-1ms me, m' -moi
CL-obj-1fs me, m' -moi
CL-obj-2fs te, -toi, t'
CL-obj-2ms te, -toi, t'
CL-obj-3ms en, y, le, l', lui
CL-obj-3fs la, l', lui
CL-obj-3fp les, leur
CL-obj-3mp les, leur

Réfléchi

CL-refl-1ms me, m' -moi,
CL-refl-1fs me, -moi, m'
CL-refl-2ms te, -toi, t'
CL-refl-2fs te, -toi, t'
CL-refl-1mp nous, -nous
CL-refl-1fp nous, -nous
CL-refl-2mp vous, -vous
CL-refl-2fp vous , -vous

CL-refl-3fp se, s'
 CL-refl-3mp se, s'

CONJONCTION

Conjonction de coordination

C-C La table AINSI QUE les chaises sont en orme.
 Le combustible est mélangé à l'air ET injecté dans la chambre

Conjonction de subordination

C-S QUAND Marie travaille, elle n'écoute personne
 Le député n'a pas fait de déclaration LORSQU'il a été battu
 Jean est gros ALORS QUE Marie ne l'est pas.
 Il est aussi petit QUE Marie.
 Je pense QU' elle a tort.
 Je me demande SI elle m'écoute.

DETERMINANT

D démonstratif

D-dem-ms CE vélo est-il a vendre ?
 D-dem-mp CES ustensiles sont pratiques
 D-dem-fs Je regarde CETTE peinture
 D-dem-fp CES peintures sont laides

D Défini

D-def-ms LE ministre a pris la parole
 D-def-mp Marie garde LES enfants
 D-def-fs LA linguistique, c'est passionnant
 D-def-fp LES ordinateurs sont allumés

D Indéfini

D-ind-fs UNE action rentable
 D-ind-ms UN ordinateur est en panne
 D-ind-mp la grogne parmi CERTAINS anciens cadres
 D-ind-fp Tu achètera DES prunes

D cardinal

D-card-fp TROIS amies sont venues
 D-card-mp TROIS éléphants sont traqués par les chasseurs

D exclamatif

D-excl-ms QUEL homme que cet homme là !
 D-excl-mp QUELS individus sinistres !
 D-excl-fs QUELLE chance !
 D-excl-fp QUELLES belles fleurs !

D Interrogatif

D-int-fs QUELLE meilleure méthode que celle du puzzle ?
 D-int-fp Jean demande QUELLES filles sont candidates
 D-int-ms QUEL pain veut-tu ?
 D-int-mp QUELS enfants cherches-tu ?

D Négatif

D-neg-fs	NULLE fille
D-neg-ms	AUCUN garçon
D-neg-ms	NULS soucis
D-neg-fp	NULLES filles

D Partitif

D-part-ms	j'ai acheté DU pain
D-part-fs	Jean boit DE LA bière

D possessif

D-poss-1mss	MON chien
D-poss-3fss	SA voiture
D-poss-1mps	MES amis
D-poss-2fps	TES fleurs
D-poss-1msp	NOTRE chien
D-poss-2fpp	VOS fleurs
etc	

INTERJECTION

I	HELAS !
---	---------

NOMNom cardinal

N-card-fs	en 1990
N-card-ms	le 12

Nom Commun

N-C-fp	Trois VACHES Les PME
N-C-fs	Une VACHE
N-C-mp	TAUREAUX des MILLIONS de Français
N-C-ms	TAUREAU Un MILLIARD de fois le FRANCAIS

Nom propre

N-P-fp	les ALPES
N-P-fs	MARIE la FRANCE la SNCF
N-P-mp	les GOBELINS Les ETATS-UNIS
N-P-ms	PAUL le JAPON

PREFIXE

PREF EX-ministre,
AUTO-censure

PREPOSITION

P DEPUIS midi
je viens DE sortir
QUANT A vous, vous restez
elle est partie VOICI 3 ans
Le sel est SUR la table
Il jette un oeil AVANT DE partir

PRONOMSPronoms démonstratifs

PRO-dem-ms je veux CECI
PRO-dem-mp CEUX que tu sais
PRO-dem-fp CELLES que tu aimes
PRO-dem-fs CELLE que tu veux

Pronoms indéfinis

PRO-ind-3ms L'UN des notres, TOUT va bien
PRO-ind-3fs L'UNE des nôtres
PRO-ind-3mp TOUS sont partis, BEAUCOUP de ces étudiants, PLUSIEURS sont arrivés
PRO-ind-3fp Elles sont TOUTES venues, CERTAINES sont là

Pronoms interrogatifs

PRO-int-3ms LEQUEL est parti ?
PRO-int-3fs LAQUELLE veux-tu ?
PRO-int-3mp On ne sait pas LESQUELS viendront
PRO-int-3fp On se demande LESQUELLES choisir

Pronoms cardinaux

PRO-card-3mp j'en veux DEUX (des livres)
PRO-card-3fp j'en veux DEUX (des pommes), DEUX sont arrivées

Pronoms possessifs

PRO-poss-3ms le MIEN, le TIEN, le SIEN, le NOTRE, le VOTRE, le LEUR
PRO-poss-3mp les AUTRES, , les MIENS, les TIENS, les SIENS,
PRO-poss-3fp les NOTRES, les VOTRES, les LEURS
PRO-poss-3fs j'en veux UNE, la MIENNE, la TIENNE, la NOTRE, la VOTRE, la LEUR

Pronoms personnels (forts)

PRO-pers-1ms moi
PRO-pers-2ms toi
PRO-pers-3ms lui
PRO-pers-1mp nous
PRO-pers-2mp vous
PRO-pers-3mp eux
PRO-pers-1fp nous
PRO-pers-2fp vous

PRO-pers-3fp	elles
PRO-pers-1fs	moi
PRO-pers-2fs	toi
PRO-pers-3fs	elle

Pronom réfléchi (fort)

PRO-refl-3ms	soi
--------------	-----

Pronoms relatifs

PRO-rel-1ms	moi QUI suis parti
PRO-rel-2ms	toi QUE l'on cherche
PRO-rel-3ms	l'endroit OU on va
PRO-rel-1mp	nous QUI sommes ignorants
PRO-rel-2mp	vous QUE tous admirent
PRO-rel-3mp	des bons QU'il faut acheter
PRO-rel-1fs	moi QUI suis partie
PRO-rel-2fs	toi QUE l'on a arrêtée
PRO-rel-3fs	la femme à LAQUELLE on écrira
PRO-rel-1fp	nous à QUI personne ne pense
PRO-rel-2fp	vous QUI êtes belles
PRO-rel-3fp	les femmes DONT on parle

Pronoms négatifs

PRO-neg-ms	PERSONNE ne viendra, RIEN ne bouge, NUL n'est censé ignorer la loi
PRO-neg-mp	AUCUNS ne viendront
PRO-neg-fs	AUCUNE n'est partie
PRO-neg-fp	AUCUNES ne seront interrogées

MOTS ETRANGERS

ET	così fan tutte
----	----------------

VERBES

Pour les formes composées, nous notons les traits morphologiques du ou des auxiliaires puis du participe.

Exemple :

il:CL3ms a:VP3s été:VKms mangé:VKms

j:CL1ms eusse:VT1s mangé:VKms

j:CL1ms ai:VP1s eu:VKms mangé:VKms

présent conditionnel

V--C1s	MANGERAIS
V--C2s	MANGERAIS
V--C3s	MANGERAIT
V--C1p	MANGERIONS
V--C2p	MANGERIEZ
V--C3p	MANGERAIENT

futur indicatif

V--F1s	MANGERAI
V--F2s	MANGERAS

V--F3s	MANGERA
V--F1p	MANGERONS
V--F2p	MANGEREZ
V--F3p	MANGERONT

participe présent

V--G	MANGEANT
------	----------

imparfait indicatif

V--I1s	MANGEAIS
V--I2s	MANGEAIS
V--I3s	MANGEAIT
V--I1p	MANGIONS
V--I2p	MANGIEZ
V--I3p	MANGEAIENT

passé simple indicatif

V--J1s	MANGEAI
V--J2s	MANGEAS
V--J3s	MANGEA
V--J3p	MANGÈRENT

participe passé

V--Kms	MANGÉ
V--Kmp	MANGÉS
V--Kfs	MANGÉE
V--Kfp	MANGÉES

présent indicatif

V--P1s	MANGE
V--P2s	MANGES
V--P3s	MANGE, VOICI
V--P1p	MANGEONS
V--P2p	MANGÉZ
V--P3p	MANGENT

présent subjonctif

V--S1s	MANGE
V--S2s	MANGES
V--S3s	MANGE
V--S1p	MANGIONS
V--S2p	MANGIEZ
V--S3p	MANGENT

imparfait subjonctif

V--T1s	MANGEASSE, EUSSE
V--T3s	MANGEAT
V--T1p	MANGEASSIONS, CHOISSISSIONS
V--T3p	MANGEASSENT

infinitif

V--W MANGER

impératif

V--Y1p MANGEONS

V--Y2s MANGE

V--Y2p MANGEZ

PREFIXE

PREF ANTI, ARCHI, PRE, POST...

PONCTUATION

forte

PONCT-S . ? ! ...

faible

PONCT-W - , ; : / " ' !

ANNEXE 2

Les différents types de composés

Les adjectifs composés:

Catégorie composé	Catégories composants	Exemples
Ams	NPN	haut de gamme
Ams	AA	politico-financier
Ams	AN	bon marché
Ams	DA	vingt-deuxième
Ams	NA	Sud-américain

Les adverbes composés:

Catégorie composé	Catégories composants	Exemples
ADV	ETET	grosso modo, a fortiori
ADV	PN	en revanche
ADV	PDN	au contraire
ADV	PAN	de longue date
ADV	PNCN	de part et d'autre
ADV	ADVADV	aujourd'hui
ADV	PADV	par ailleurs
ADV	PP	par contre
ADV	PDNCPN	au fur et à mesure
ADV	PDVADV	à la va vite
ADV	DN	nulle part
ADV	PDNPDN	au bout du compte
ADV	PADVADV	de plus en plus
ADV	ADVCADV	tôt ou tard
ADV	ADVADV	peu à peu
ADV	PPRO	en soi

Les conjonctions composées:

Catégorie composé	Catégories composants	Exemples
CS	ADVC	avant que
CS	PDNC	au cas où
CS	PNC	à mesure que
CS	PDNCPNC	au fur et à mesure que
CS	ADVADVC	si peu que
CC	ADVC	ainsi que
CC	CLVAV	c'est-à-dire
CC	CLV	y compris

Les déterminants composés:

Catégorie composé	Catégories composants	Exemples
Dms	VD	n'importe quel
Dmp	DD	vingt-quatre
Dmp ou Dfp	DAD	la plupart des

Les mots étrangers composés :

ET ET ET Scala mobile

Les interjections composées:

Catégorie composé	Catégories composants	Exemples
I	DN	Mille sabords
I	PN	A bas

Les noms propres composés:

Catégorie composé	Catégories composants	Exemples
NPms	NPN	Giscard d'Estaing
NPms	NN	Jean Pierre
NPmp	NA	Etats-Unis
NPmp	NAA	Emirats Arabes Unis
NPfs	AN	Nouvelle Zélande
NPfs	NA	Colombie-britannique
NPfs	NPN	île-de-Ré

Les noms communs composés:

Catégorie composé	Catégories composants	Exemples
NCms	NA	téléphone rouge
NCms	VN	garde frontières
NCms	NN	homme-grenouille
NCms	NDN	maillot deux-pièces
NCms	NPN	fil de fer
NCms	NPV	fer à repasser
NCms	PN	après-midi
NCfs	NPV	machine à coudre
NCfs	NN	idée force
NCfs	NA	carte bleue
NCfs	NPN	pomme de terre, boîte à malices
NCfs	AN	belle-mère
NCms	NPRO	essuie-tout
NCms	VAN	va-nu-pied

Les pronoms composés:

Catégorie composé	Catégories composants	Exemples
PROmp	DPRO	d'aucuns
PROmp	DN	deux cents
PROmp	PROPRO	vingt-quatre
PROms	PROADV	celui-ci
PROms	PROA	lui-même
PROms	ADN	tout le monde

Les prépositions composées:

Catégorie composé	Catégories composants	Exemples
-------------------	-----------------------	----------

P	PADV	d'ici
P	ADVP	avant de
P	PNP	en face de
P	PDNP	à l'insu de
P	AP	quitte à
P	PP	quant à
P	PNCNP	de part et d'autre de
P	NP	face à
P	PVP	à partir de

Les verbes composés

Catégorie composé	Catégories composants	Exemples
V	VN	faire fi
V	VAN	avoir beau jeu
V	VPN	mettre à sac
V	VA	avoir beau