



HAL
open science

De l'intrusion d'un lexique allogène : l'exemple des éléments néoclassiques

Marine Lasserre

► **To cite this version:**

Marine Lasserre. De l'intrusion d'un lexique allogène : l'exemple des éléments néoclassiques. Linguistique. Université Toulouse le Mirail - Toulouse II, 2016. Français. NNT : 2016TOU20012 . tel-01723307

HAL Id: tel-01723307

<https://theses.hal.science/tel-01723307>

Submitted on 5 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse - Jean Jaurès*

Présentée et soutenue le *26/02/2016* par :

MARINE LASSERRE

**De l'intrusion d'un lexique allogène.
L'exemple des éléments néoclassiques**

JURY

FABIO MONTERMINI	Directeur de Recherche, CNRS/U. Toulouse Jean Jaurès	Directeur
FIAMMETTA NAMER	Professeur, U. de Lorraine	Rapporteur
FLORENCE VILLOING	Professeur, U. de Paris Ouest Nanterre La Défense	Rapporteur
MARK ARONOFF	Distinguished Professor, Stony Brook University	Examineur
GILLES BOYÉ	Maître de Conférences, U. de Bordeaux Montaigne	Examineur
JESSE TSENG	Chargé de Recherche, CNRS/U. Toulouse Jean Jaurès	Examineur

École Doctorale et spécialité :

CLESCO : Sciences du langage

Unité de Recherche :

CLLE-ERSS (UMR 5263)

Remerciements

Ce travail long et pas toujours facile a été fortement allégé, facilité, égayé par de nombreuses personnes que je tiens à remercier.

Tout d’abord, je remercie Fabio Montermini, mon directeur de thèse, qui a accepté de diriger cette thèse sans même me connaître, et qui, tout au long de ce travail, a été particulièrement disponible et à l’écoute, indulgent et pédagogue. Il m’a fait confiance depuis le début et je tiens à lui exprimer toute ma gratitude.

Je tiens également à remercier les membres de mon jury de thèse : Fiammetta Namer et Florence Villoing pour avoir accepté d’en être les rapporteurs et Mark Aronoff, Gilles Boyé et Jesse Tseng d’en être les examinateurs. Je les remercie chacun également pour les conseils et avis éclairés dispensés tout au long de mon doctorat et les discussions partagées. En particulier, je remercie Gilles Boyé de m’avoir fait découvrir et aimer la morphologie en Master et de m’avoir, un jour, soufflé l’idée de ce sujet de thèse qui m’a passionnée. Je remercie également Mark Aronoff de m’avoir si bien accueillie à l’université de Stony Brook, m’avoir permis d’assister à ses cours et avoir pris du temps pour discuter autour de mes recherches.

Plus largement, je remercie la morphosphère et plus particulièrement, pour les discussions, les conseils, Michel Roché, Delphine Tribout, Stéphanie Lignon ou Olivier Bonami. J’ai une pensée particulière pour les filles de la doctorantomorphosphère, qui, même pour celles qui ont déjà soutenu, sont, souvent à distance, d’un soutien sans faille : Jana Strnadová, Aurore Koehl, Edwige Dugas, Aurélie Guerrero.

Mon doctorat n’aurait pas pu mieux se passer qu’au sein de l’ERSSopôle, qui m’a offert des conditions idéales, tant sur le point de l’intégration dans l’équipe que de soutien moral, technique et financier. Je remercie en particulier Nabil Hathout, qui m’a fait gagner beaucoup de temps dans la collecte de mes données, mais aussi Basilio Calderone, Fabio Del Prete, Cécile Fabre, Hélène Giraud, Mai Ho-Dac, Josette Rebeyrolle, Ludovic Tanguy, Jean-Michel Tarrier et tous les autres. Sans Laurence Lamy et Nathalie Moulic, toute la partie administrative aurait été bien

plus pénible, et je les remercie pour leur aide et leur bonne humeur communicative.

L'ERSSoctorantosphère a été au top également et a donné lieu à des moments de partage, de soutien, de solidarité et de franche rigolade, au feu le 6^{ème} étage. Plus particulièrement je dois beaucoup à :

- Caroline Atallah, qui mériterait une page de remerciements entière pour elle tant elle a été présente, du début jusqu'aux derniers instants et d'un soutien sans faille. Nos sessions gerçaises et nos aventurière JéToulogiques resteront longtemps dans ma mémoire...
- ... aventures dans lesquelles Cécile Viollain, ma Céçaïle, n'était pas en reste (merci pour le selfie-floc). Ses connaissances anglophones ont souvent eu raison de mon anglophobie. Et l'année passée avec elle en tant que représentante des doctorants restera, je pense, inoubliable (surtout la chasse aux œufs de Pâques dans les couloirs du 6^{ème}),
- Luce Lefeuvre, avec qui j'ai également partagé une année en tant que représentante, et avec qui j'ai aimé partager de longues discussions (et surtout de nombreux ragots!),
- mes compagnons bureauticoles du feu C616 : Fanny Lalleman, Florian Savreux, Willy Beaujean, Laure-Anne Johnsen, et (pendant deux ou trois jours!), Léa Courdès-Murphy,
- Marianne Vergez-Couret grâce à qui j'ai pu dé-crochet de la thèse et découvrir une nouvelle passion,
- Giusi Todaro et Matilde Accattoli desquelles je me suis très vite sentie proche et avec qui, j'espère, on ne rompra pas le lien, qu'il soit morphologique ou pas,
- Karla Orihuela pour sa bonne humeur au quotidien,
- François Morlane-Hondère et ses qualités de L^AT_EXomane,
- et tous les autres : Laury Garnier, Emilie Chorin, Manuel Perez, Marie-France Roquelaure, Stéphanie Lopez, Maxime Warnier, Aleksandra Miletic, etc.,
- et j'ai une pensée pour Simon Leva avec qui on a commencé la même année, mais qui est parti vers d'autres aventures moins linguistofères.

Je n'oublie pas ceux qui m'ont apporté une aide salvatrice à un moment donné : Matteo Pascoli qui, en deux-trois manipulations de données, m'a fait gagner un temps énorme, Marta Cagliani qui a eu la patience de traiter certaines données, Sébastien Garot qui m'a éclairée de ses connaissances mathologiques, la latinologue Olga Spevak pour m'avoir aidée dans mes recherches bibliographiques sur le latin, et Marie-Laure Maraval pour le temps qu'elle m'a accordé pour la mise en forme de ma base de données.

Je remercie les membres du département de linguistique de l'Université de Stony Brook (et pas que), pour leur accueil chaleureux, et en particulier Jiwon Yun, Seongyeon Ko, Donna Di Donato et Rob Pasternak.

Je remercie du fond du cœur ma famille et en particulier mes parents qui m'ont offert, bien plus qu'un abri Toulousicole, un soutien constant, des petites et grandes attentions et par-dessus tout leur confiance. Je remercie également ma sœur Anne-Sophie pour ses encouragements. Elle et Maman ont également usé de leurs dons fautedorthographovores et je les en remercie.

Je remercie mes amis, qui l'étaient avant cette aventure thésalogique et qui ont bien voulu le rester. Par un petit mot ou une distraction, par un intérêt pour mes recherches ou même grâce à quelques heures de baby-sitting, ils m'ont permis de tenir le cap. Merci à Laurie et Chloé (et à nos activités biérotaniques!), à Sarah, Mélanie, Perrine (qui est passée par là il y a peu!), Margaux, Cindy, Namu, Sophie, Cyril, Jean-Renaud, David, Vadim, Damien, Evelyne, Elysandre, Seb et les autres.

Le plus Marinophile d'entre tous, mon Julien, m'a prouvé, par son soutien, ses encouragements, ses blagues et son amour, que j'avais bien trouvé Saussure à mon pied.

Enfin, je remercie Charlie pour avoir bien voulu, sans trop rechigner, partager, lors de sa première année, sa Maman avec sa petite sœur Thésounette.

Et l'on constate ainsi que, dans une langue comme le français, la possibilité de créations sans cesse renouvelées s'allie parfaitement avec la fidélité aux racines mêmes. Comment ne pas s'émerveiller d'une telle vitalité ?

(Jacqueline de Romilly, *Dans le Jardin des Mots*)

Table des matières

Introduction	1
1 Des composés grecs aux composés néoclassiques	7
1.1 Un lexique allogène pour un vocabulaire savant	8
1.1.1 Le mécanisme de l'emprunt	9
1.1.2 Aperçu translinguistique	13
1.1.3 Pourquoi un lexique allogène ?	24
1.2 Du grec ancien aux langues modernes	30
1.2.1 La composition en grec ancien	30
1.2.2 La composition en latin	37
1.2.3 La pénétration d'éléments classiques en français moderne . .	43
1.3 Les composés néoclassiques dans la littérature	51
1.3.1 Qu'est-ce qu'un élément néoclassique ?	51
1.3.2 Éléments néoclassiques, affixes, lexèmes	55
1.3.3 Quelles sont les caractéristiques des constructions néoclas- siques ?	63
2 Questions méthodologiques	71
2.1 Choix du cadre théorique	72
2.1.1 Un objet d'étude peu homogène	73
2.1.2 Des schémas dans le lexique	80
2.1.3 La construction du sens des mots construits	90
2.2 L'utilisation de la Toile en morphologie	95
2.2.1 Une ressource inestimable	95
2.2.2 Problèmes qui se posent	96
2.2.3 Écueils à éviter et précautions à prendre lors de l'utilisation de la Toile	99
2.3 Création de la base de données NeoClassy	101

2.3.1	Choix des éléments néoclassiques à étudier	101
2.3.2	Présentation des éléments néoclassiques choisis	105
2.3.2.1	<i>-logie, -logue</i>	106
2.3.2.2	<i>-cratie, -crate</i>	106
2.3.2.3	<i>-phobe, -phobie</i>	108
2.3.2.4	<i>-phone</i>	108
2.3.2.5	<i>-phage</i>	109
2.3.2.6	<i>-vore</i>	110
2.3.2.7	<i>-cide</i>	111
2.3.2.8	<i>-cole</i>	111
2.3.3	Recueil des données et création de la base de données NeoClassy	112
2.3.3.1	Sur la question de la lexicalisation	112
2.3.3.2	NeoClassy : présentation	113
3	Description des données	119
3.1	Contraintes en jeu	120
3.1.1	Le constituant initial	121
3.1.2	La voyelle de liaison	138
3.1.3	Taille des mots construits	155
3.2	Organisation du lexique	162
3.2.1	Description sémantique	163
3.2.1.1	Du discours et des spécialistes	163
3.2.1.2	Du pouvoir et des coalitions	171
3.2.1.3	De la peur à la haine	175
3.2.1.4	De la voix qui sonne	179
3.2.1.5	Manger, dévorer, consommer	183
3.2.1.6	Manger, dévorer, consommer (bis)	188
3.2.1.7	Du meurtre à l'extermination	191
3.2.1.8	Habiter, cultiver, honorer	194
3.2.2	Le lexique s'organise	197
3.2.2.1	Organisation par mot leader	198
3.2.2.2	Organisation distribuée	202
3.2.2.3	Des liens qui se tissent	206
4	Les éléments néoclassiques dans la construction du sens	213
4.1	Construction du sens et interprétation	214
4.1.1	Proposition d'analyse distributionnelle	217
4.1.1.1	Fondements et méthode	217
4.1.1.2	Ce que le contexte des lexèmes <i>Xcide</i> nous apprend	219
4.1.1.3	Les <i>-phobie</i> en contexte	228
4.2	Les relations inter-schématiques singulières	237

4.2.1	Dérivés possibles et impossibles	237
4.2.2	Des liens privilégiés dans le réseau lexical	242
4.3	La charge sémantique des éléments néoclassiques	246
4.3.1	Les constructions néoclassiques comme adjectifs relationnels	246
4.3.2	Les éléments néoclassiques, des suffixes comme les autres? .	257
4.3.2.1	La construction de noms de spécialistes <i>Xlogue</i> . .	257
4.3.2.2	Les anti-étrangers sont-ils xénophobes?	263
5	Un modèle unifié	271
5.1	Sur la question du lexème	272
5.1.1	Le lexème, une unité multidimensionnelle	272
5.1.1.1	Forme(s) du lexème	277
5.1.1.2	Signification(s) du lexème	286
5.1.1.3	Catégorie lexicale	294
5.1.2	Lexème et affixe : porosité des classes	297
5.2	Les schémas constructionnels	302
5.2.1	La composition, un type de dérivation	302
5.2.2	Perméabilité des frontières	317
5.2.2.1	De l’idiome constructionnel à la grammaticalisation	317
5.2.2.2	Les lexèmes tronqués	330
5.2.2.3	Aux frontières des éléments néoclassiques	341
5.3	Quelle place pour les éléments néoclassiques?	343
5.3.1	Critères de classification	344
5.3.1.1	Les éléments finaux	344
5.3.1.2	Les constructions à élément final spécifié	352
5.3.2	Une typologie des éléments finaux	358
	Conclusion et perspectives	363
	Annexes	367
	La base de données NeoClassy	369
	<i>Xcide</i>	369
	<i>Xcole</i>	375
	<i>Xcrate</i>	379
	<i>Xcratie</i>	382
	<i>Xlogie</i>	387
	<i>Xlogue</i>	407
	<i>Xphage</i>	423
	<i>Xphobe</i>	426
	<i>Xphobie</i>	437

<i>Xphone</i>	447
<i>Xvore</i>	454
Doublets et triplets <i>-phage / -vore</i>	459
Bibliographie	478

Table des figures

2.1	Modèle selon Halle (1973)	82
2.2	Exemple de schéma et mot complexe mémorisés selon la théorie de Booij (2010)	86
2.3	Exemple de constructions hiérarchisées	87
2.4	La polysémie de la construction [Xphage]	89
2.5	Représentation de la fréquence de plusieurs lexèmes <i>Xphobie</i> dans Google Books, de 1830 à 2008	94
3.1	Origine de C1, pour tous les lexèmes de NeoClassy, selon l'EN final	122
3.2	Origine de C1, pour tous les lexèmes de NeoClassy, selon la dictionnarité des mots construits	123
3.3	Dictionnarité des lexèmes de NeoClassy, selon chaque EN	124
3.4	Dictionnarité et origine de C1, pour les lexèmes <i>Xcrate</i> et <i>Xcratie</i> .	125
3.5	Dictionnarité et origine de C1, pour les lexèmes <i>Xcole</i> et <i>Xphage</i> . .	125
3.6	Mesure de la distance de Levenshtein entre C1 et C1bis, C1bis et X et C1 et X, pour les C1 non classiques (n = 2862)	131
3.7	Types d'accidents morphophonologiques subis par le C1, selon la distance de Levenshtein mesurée entre le C1 et le C1bis (n = 2862)	133
3.8	Catégorie des C1 natifs des lexèmes de NeoClassy (n=2904)	137
3.9	Distribution du segment précédant l'EN, tous ENs de NeoClassy confondus (n=4459)	139
3.10	Comparaison des EN <i>-logie</i> et <i>-phone</i> du point de vue de l'origine de C1 (A.) et du segment de liaison (B.)	141
3.11	Voyelle de liaison selon l'origine du C1	142
3.12	Voyelle de liaison affichée pour les C1 non classiques	143
3.13	Distribution du segment précédant l'EN, selon l'origine de l'EN final, aperçu	145
3.14	Distribution du segment précédant l'EN, selon l'origine de l'EN final	146

3.15	Distribution du segment précédant l'EN, pour chacun des EN de NeoClassy	147
3.16	Distribution du segment précédant l'EN latin, selon l'origine de C1	148
3.17	Prévisions de distribution des probabilités pour chaque valeur possible de X selon la loi binomiale $\mathcal{B}(372; 0, 9311)$	151
3.18	Prévisions de distribution des probabilités pour chaque valeur possible de X selon la loi binomiale $\mathcal{B}(3083; 0, 7507)$	152
3.19	Proportion, en pourcentages, de chaque élément de liaison selon que les lexèmes sont dictionnarisés ou non	153
3.20	Taille des mots construits de NeoClassy (A.) et de la séquence à gauche de l'EN, soit X (B.) (boxplot)	156
3.21	Taille des mots construits de NeoClassy (A.) et de la séquence à gauche de l'EN, soit X (B.), en fonction de chaque EN final étudié (boxplot)	157
3.22	Taille des lexèmes, en syllabes, en fonction de chaque EN	158
3.23	Taille des C1, mis en perspective avec la taille des MC, pour les lexèmes ayant un seul constituant à gauche de l'EN final (n=4182)	159
3.24	Comparaison, pour les lexèmes <i>Xlogue</i> et <i>Xcrate</i> , de la taille des C1 et de la taille des mots construits	160
3.25	Taille des mots construits selon leur dictionnarité, en pourcentages : pour l'ensemble des lexèmes de NeoClassy (A.), pour les lexèmes <i>Xlogue</i> (B.) et les lexèmes <i>Xcrate</i> (C.)	160
3.26	Taille de X selon l'origine du C1 (A.) et la provenance de la voyelle de liaison (B.) (boxplot)	161
3.27	Taille des lexèmes de NeoClassy ayant un C1 natif et unique, en fonction des accidents morphophonologiques subis par le C1 (n=2862)	162
3.28	Différents schémas impliquant l'élément néoclassique <i>-logie</i>	170
3.29	Différents schémas impliquant l'élément néoclassique <i>-cratie</i>	176
3.30	Différents schémas impliquant l'élément néoclassique <i>-phage</i>	186
3.31	Aperçu d'une organisation lexicale avec des mots leader	203
3.32	Aperçu d'une organisation lexicale du type distribué	206
3.33	Gradabilité des ENs <i>-lâtre</i> , <i>-mane</i> , <i>-phage</i> , <i>-phile</i> et <i>-phobe</i>	209
3.34	Réseau formé par les constructions impliquant certains éléments néoclassiques	211
4.1	Type sémantique du C1, selon le type de schéma impliquant l'EN <i>-cide</i>	225
4.2	Nombre de résultats en fonction de la méthode employée	235
4.3	Origine du C1 en fonction du sens de <i>Xphobie</i> (n = 377)	236
4.4	Taux de <i>Xphobique</i> par rapport aux <i>Xphobie</i> étiquetés comme 'peur' ou 'hostilité'	240

4.5	Taux de <i>Xphile</i> par rapport aux <i>Xphobe</i> étiquetés comme ‘peur’ ou ‘hostilité’	243
4.6	Uni-, bi- et trigrammes précédant la séquence /ik/ (source : www.lexique.org)	256
4.7	Réseau formé par les constructions impliquant certains éléments néoclassiques ainsi que certains affixes	269
5.1	Polysémie régulière des constructions XCIDE	292
5.2	Types de composés proposés par Bisetto et Scalise (2009) illustrés par les exemples de Villoing (2012)	305
5.3	Sous-schémas [NN] _N	318
5.4	Typologie des constructions présentant un élément final récurrent	359

Liste des abréviations

A	Adjectif
Abl	Ablatif
Ang	Anglais
C	Consonne
C1	Constituant 1 (celui de gauche)
CN	Construction Néoclassique
DHLF	Dictionnaire Historique de la Langue Française
EN	Élément Néoclassique
Esp	Espagnol
Fr	Français
GR	Grand Robert
It	Italien
LDC	Linguistic Data Consortium
LGP	Language for general purposes
lit.	Littéralement
LSP	Language for special purposes
MC	Mot Construit
N	Nom
Nb	Nom Base
Nr	Nom Recteur
RCL	Règle de Construction de Lexème
SAE	Standard Average European
TLFi	Trésor de la Langue Française informatisé
V	Verbe
vdM	Les voisins de Le Monde
vdW	Les voisins de Wikipédia
WFR	Word Formation Rule

Introduction

Parmi les passages qui existent d'une langue à une autre, il y a ceux qui empruntent une voie linéaire, subissant des évolutions phonétiques et sémantiques continues, ce qui fait qu'un mot d'une langue moderne n'est plus tout à fait le même voire est éloigné du mot de la langue ancienne duquel il est le descendant. Et quand ces évolutions linéaires entraînent des différences significatives entre les ancêtres et les descendants, par exemple des différences phonétiques, on estime que la langue plus ancienne a cédé la place à une autre langue. Du latin, on est ainsi passé au bas latin, puis au gallo-roman qui a laissé la place à l'ancien français qui, passant par le moyen français et le français classique, a évolué jusqu'au français moderne, langue que nous parlons aujourd'hui en France et plus largement dans la francophonie. Mais d'autres passages sont bien moins linéaires. C'est le cas des emprunts qui, souvent pour pallier un besoin de dénomination d'un objet ou concept nouveau, volent, bien plus qu'ils n'empruntent, un mot à une autre langue. La langue volée possédait soit déjà ce mot pour cette désignation précise, soit un mot qui a un sens qui s'en rapproche et qui convient, modulo une variation sémantique par exemple, pour le besoin de dénomination de la langue emprunteuse. Le français a ainsi emprunté l'*opéra* à l'italien, en conservant le sens qu'il avait dans cette langue, mais a détourné le sens de *people* qui signifie 'personne, gens' en anglais mais désigne des stars en français. Les langues sont donc constituées d'une part d'un lexique qu'on peut appeler *indigène*, c'est-à-dire hérité des langues desquelles elles descendent, et d'autre part d'un lexique *allogène*, provenant d'une langue autre, contemporaine ou ancienne, et introduit à un moment donné dans la langue cible. C'est de l'introduction de ce lexique allogène dans une langue telle que le français, et en particulier d'éléments grecs ou latins, dont il est question dans cette thèse.

Dès le latin, des mots ont été empruntés au grec ancien pour satisfaire les besoins lexicaux, en particulier liés à des avancées scientifiques ou philoso-

phiques. Le grec possédait dès lors un prestige qui seyait déjà au caractère savant des concepts à dénommer. C'est à ce prestige des langues anciennes que les scientifiques, dès le XVI^{ème} mais surtout aux XVIII^{ème} et XIX^{ème} siècles, ont fait appel, pour répondre aux besoins onomasiologiques croissants dus à l'essor des sciences et des techniques. En 1516, est directement emprunté aux langues classiques le terme *anthropologie*. Le TLFi date de 1802 l'emploi de *philologie* avec le sens 'étude des langues', terme employé au grec *φιλολογία* où il signifiait 'amour des lettres'. Et quand le fonds classique n'a plus suffi à alimenter les avancées des savants, ceux-ci ont créé de nouveaux mots de toutes pièces, à partir d'éléments grecs ou latins et mimant les structures des mots de ces langues : puisque *anthropologie* et *philologie* existaient pour désigner des sciences, *morphologie* a pu être formé, en 1822, pour qualifier la « science qui étudie la forme et la structure des organismes » (TLFi).

Ces composés néoclassiques, ainsi nommés puisque forgés dans la langue moderne au moyen de constituants d'origine classique, ont proliféré d'abord dans la langue scientifique, s'éloignant peu à peu des structures classiques, par exemple en mêlant des constituants grecs et latins (*delphinologie*) ou en combinant des mots de la langue française à des constituants d'origine classique (*transplantologie*). Ensuite, la langue courante s'est emparée de ce type de mots construits : celui qui arrive à apparier des chaussettes devient expert en *chaussettologie*, un travail fastidieux est *chronophage* ou *tempslibricide* et un raisonnement difficile à croire est *capillotracté*. Ainsi, des locuteurs sans connaissance particulière des langues classiques ont accès à ce type de formation pour créer des mots qui sont destinés à être employés dans le langage courant. Ces formations appartiennent-elles donc aux langues classiques ou aux langues modernes ? Doit-on les distinguer, en français du moins, des composés dits natifs, du type *ouvre-boîte* ou *homme-grenouille* ou au contraire les considérer comme des composés obéissant à des contraintes particulières ?

D'autre part, quel statut donner à ces constituants d'origine classique ? Les travaux traitant de ce sujet dans la littérature se heurtent tous à une aporie : ce sont des éléments qui, comme les lexèmes, ont un sens de type lexical et, comme les affixes, ne se rencontrent pas de façon autonome. Cette aporie est résolue de manière très variée dans la littérature : les constituants néoclassiques sont soit classés parmi les lexèmes (ou considérés comme des unités infra-lexicales, c'est-à-dire des lexèmes non autonomes), soit parmi les affixes, soit parmi une classe d'éléments à part entière, par exemple les *combining forms* dans la littérature anglophone ou les *archéoconstituants* chez Danielle Corbin (1999, 2001). Nous verrons dans cette thèse qu'il est difficile de classer tous ces constituants néoclassiques dans une même catégorie d'éléments, quelle qu'elle soit. Certains ont en effet un sens lexical quand d'autres ont

perdu cette charge sémantique et se rapprochent ainsi des affixes.

Cette thèse se propose de répondre à plusieurs questions : tout d'abord, donc, celle du statut de ces éléments et des constructions dans lesquelles ils entrent en jeu. S'agit-il de lexèmes, ou bien d'affixes ? Par conséquent, les constructions sont-elles à classer parmi les composés ou les dérivés ? Si ce sont des composés, s'agit-il de composés à part ? Les règles de formation de ces lexèmes sont-elles des règles elles aussi allogènes, empruntées aux langues classiques ? D'autre part, sera abordée la question des contraintes morphophonologiques et sémantiques qui pèsent sur la formation de ces mots. Ces mots répondent-ils aux mêmes contraintes que les autres mots construits du français ? Ont-ils des particularités phonologiques, sémantiques ? Par ailleurs, dans une vision plus globale, je discuterai de la construction du sens des mots construits : en effet, le fait que les mots ici étudiés soient formés au moyen d'éléments allogènes par des locuteurs n'ayant pas de connaissance de ce fonds allogène nous éclaire sur la construction du sens des mots plus en général : comment peut-on former un lexème au moyen d'un constituant dont on ne connaît pas le sens premier ? Des généralisations sont-elles possibles à travers la connaissance que l'on a du lexique existant ? Comment un constituant allogène peut voir son sens varier au contact d'éléments indigènes ?

Pour répondre à ces questions, j'ai choisi de limiter l'étude des constructions néoclassiques à celles qui comprennent certains éléments finaux récurrents, particulièrement disponibles pour les locuteurs non spécialistes. Les constructions envisagées dans cette thèse sont donc les suivantes : *Xlogie*, *Xlogue*, *Xcratie*, *Xcrate*, *Xphobie*, *Xphobe*, *Xphage*, *Xphone*, *Xvore*, *Xcide* et *Xcole*. Les éléments finaux impliqués dans ces constructions semblent les plus à même de répondre à la question du statut des éléments néoclassiques, car leur récurrence peut leur faire perdre tout ou partie de leur charge sémantique et ainsi les situer, de prime abord, à la jonction de la catégorie des lexèmes et de celle des affixes. De par leur récurrence également, j'ai pu avoir accès à un grand nombre de mots construits selon ces schémas, ce qui a permis une analyse plus fine des contraintes auxquelles répond ce type de formation. De plus, ces éléments sont manipulés par bon nombre de locuteurs sans connaissances étymologiques, et très souvent combinés à des lexèmes français par ailleurs autonomes, ce qui permet une comparaison plus aisée avec les autres types de formations existant en français.

* * *

Le plan suivi par cette thèse est conçu comme un cheminement qui, d'une manière inductive, permet de passer de l'étude des données à l'enrichissement

de la théorie. Ainsi, certaines questions seront abordées au début sans a priori théorique, par exemple la notion de lexème, et précisées ultérieurement.

Le chapitre 1 sera l'occasion d'une description détaillée du sujet de cette thèse. J'aborderai initialement la question de l'emprunt et en particulier de l'emprunt à des fins scientifiques. Le français, ou même les langues européennes, ne sont pas les seules à puiser dans une autre langue pour construire un vocabulaire savant ; par exemple, le vietnamien a emprunté au chinois des éléments qui semblent fonctionner de la même manière que les constituants néoclassiques empruntés par les langues européennes au grec et au latin. Ensuite, dans une deuxième partie, je ferai une description de la composition en grec ancien puis en latin afin de pouvoir évaluer les différences qui existent entre les composés des langues classiques et les composés néoclassiques des langues modernes. Je ferai également un état des lieux des différents moyens pour des éléments allogènes de pénétrer le lexique d'une langue. Enfin, ce premier chapitre sera également l'occasion d'une revue des principaux travaux traitant de la question des composés néoclassiques dans la littérature.

Le chapitre 2 sera dévolu aux questions méthodologiques. En premier lieu, je présenterai le cadre théorique dans lequel j'ai choisi de m'inscrire, celui de la Morphologie Constructionnelle (Booij, 2010). Ce cadre ne nécessite pas des frontières imperméables entre des éléments comme les lexèmes et les affixes. En effet, les schémas de construction, qui sont des généralisations faites sur le lexique existant et qui agissent comme des recettes pour la création de nouveaux items, permettent d'intégrer en leur sein des éléments spécifiés, qui peuvent être des affixes ou des lexèmes. Cela permet de rendre compte d'éléments hétérogènes à l'intérieur d'un même cadre. Dans un second temps, j'argumenterai en faveur de l'utilisation de la Toile en morphologie : en effet, ce moyen, quoique non parfaitement adapté à l'étude de la langue, donne accès à un grand nombre de formations introuvables dans tous les corpus disponibles, aussi grands soient-ils, et permet ainsi une investigation plus précise des modes de formation de mots. Dans un troisième temps, je présenterai la base de données que j'ai constituée pour faire cette étude : baptisée NeoClassy, elle regroupe des lexèmes récoltés dans des dictionnaires et sur la Toile, tous construits au moyen des éléments néoclassiques finaux récurrents sélectionnés.

Les observations faites sur la base de données NeoClassy seront décrites dans le chapitre 3. Ces observations donneront lieu à une liste de contraintes auxquelles répondent les constructions néoclassiques. Je me concentrerai sur trois contraintes morphophonologiques : une contrainte pesant sur la sélection du constituant de gauche, une contrainte pesant sur la présence d'une voyelle de liaison dans ces constructions et enfin une contrainte de taille.

Ensuite, je ferai des observations d'ordre sémantique : les sens que peut recouvrir chaque type de construction seront présentés. Dans un second temps, je relierai ces sens entre eux et montrerai que le lexique peut s'organiser de différentes façons : parfois, le sens des mots se construit en référence à un mot leader qui sert ainsi de pôle d'attraction, parfois l'organisation lexicale est plus distribuée et l'éventail de sens possibles élargi.

La construction du sens des constructions néoclassiques sera abordée plus précisément dans le chapitre 4, à travers divers axes. J'y présenterai tout d'abord une analyse distributionnelle des constructions *Xcide* et *Xphobie* et montrerai ainsi qu'une étude systématisée du contexte peut aider à l'interprétation des mots construits selon des constructions pouvant présenter des sens variés. Ensuite, ces différents sens seront mis en évidence par l'observation des relations existant entre des schémas à l'intérieur du lexique. En effet, des constructions impliquant le même élément néoclassique final peuvent entretenir des relations avec des constructions différentes. Enfin, nous verrons que les éléments néoclassiques peuvent n'apporter aucune contribution sémantique à la construction, se comportant ainsi comme, par exemple, les affixes que l'on retrouve dans les adjectifs de relation.

Le chapitre 5 permettra de tirer des conclusions à partir de ce qui a été avancé dans les chapitres précédents. Dans le cadre théorique de la Morphologie Constructionnelle, je commencerai par définir les unités de l'analyse, le lexème et l'affixe, ainsi que les constructions dans lesquelles ils sont employés : composition et dérivation. Je montrerai ainsi que les frontières entre ces concepts sont tout à fait perméables et que les lexèmes peuvent avoir, tout comme les affixes, des positions spécifiées à l'intérieur des schémas de construction. J'argumenterai également en faveur de l'analyse de la composition comme un type de dérivation, étant par là sur le même plan que l'affixation. Ces réflexions m'amèneront à réfléchir sur la question de la grammaticalisation, c'est-à-dire ici le passage d'un statut de lexème à un statut d'affixe. Je décrirai ainsi plusieurs schémas de construction impliquant un même élément final récurrent et conclurait par une typologie de ces constructions à élément final récurrent sur la base d'un grand nombre de critères, typologie qui permettra d'englober les constructions qui sont le cœur de cette thèse et qui ont été listées précédemment.

Cette thèse offre ainsi un point de départ à une réflexion plus générale sur la langue et sur les éléments entrant en jeu dans la formation de mots complexes, que ces éléments soient indigènes ou allogènes.

Chapitre 1

Des composés grecs aux composés néoclassiques

Sommaire

1.1	Un lexique allogène pour un vocabulaire savant	8
1.1.1	Le mécanisme de l'emprunt	9
1.1.2	Aperçu translinguistique	13
1.1.3	Pourquoi un lexique allogène?	24
1.2	Du grec ancien aux langues modernes	30
1.2.1	La composition en grec ancien	30
1.2.2	La composition en latin	37
1.2.3	La pénétration d'éléments classiques en français moderne	43
1.3	Les composés néoclassiques dans la littérature .	51
1.3.1	Qu'est-ce qu'un élément néoclassique?	51
1.3.2	Éléments néoclassiques, affixes, lexèmes	55
1.3.3	Quelles sont les caractéristiques des constructions néoclassiques?	63

Un composé néoclassique en français, comme nous aurons l'occasion de le voir tout au long de cette thèse, a recours pour sa formation à des constituants d'origine grecque ou latine et est créé, à l'origine, dans un vocabulaire spécialisé : le vocabulaire technico-scientifique. Ce recours à un lexique que j'appellerai ici *allogène*, par opposition avec *indigène*, pour la création d'une langue spécialisée est un phénomène courant tant d'un point de vue diachronique que typologique. Si les langues de culture européenne empruntent au grec et au latin, d'autres langues affichent un même phénomène : nous

verrons par exemple que le vietnamien adopte des procédés de composition chinois. Socialement et culturellement, emprunter à une langue considérée comme prestigieuse semble valoriser l'objet du discours ; concernant les langues indo-européennes, les savants pouvaient afficher leurs connaissances étymologiques et ainsi mettre une distance entre le discours scientifique et le discours courant. Ce chapitre va poser les jalons de l'analyse conduite dans cette thèse en décrivant son objet d'étude. La partie 1.1 va s'attacher, d'un point de vue général, à se questionner sur les mécanismes et les processus en jeu dans l'emprunt qu'opère une langue à une autre langue pour former un lexique spécialisé ; je verrai aussi des exemples autres que les composés néo-classiques et essaierai de trouver des raisons de cette tendance à l'emprunt. Je recentrerai la discussion à mon objet d'étude dans la partie 1.2 : d'un point de vue diachronique, les processus de formation des composés grecs et latins vont éclairer les moyens et les raisons du recours à ces langues pour créer des composés savants en français. Enfin, la partie 1.3 exposera les différents traitements qui ont pu être faits de ces composés néoclassiques dans la littérature, et nous verrons qu'ils sont bien variés.

1.1 Un lexique allogène pour un vocabulaire savant

Le recours à un lexique allogène est un phénomène courant dans de nombreuses langues. Ce recours peut se faire de différentes manières. L'emprunt lexical est la forme la plus répandue : une langue emprunte à une autre un lexème, cet emprunt étant plus ou moins intégré dans la langue d'accueil. Cependant, il y a d'autres façons de recourir au lexique d'une autre langue. Par exemple, le calque permet de n'emprunter que le sens d'un mot ou la structure morphologique d'un mot complexe ; dans ces cas-là, un mot acquiert certaines propriétés allogènes, même s'il peut paraître tout à fait indigène. Cependant, l'apparence d'un mot, le fait qu'il soit phonologiquement apparenté à un mot étranger peut avoir son importance. Dans le cas du vocabulaire scientifique, les mots ont une consonance volontairement non indigène, cela pour des raisons de jargon, de vocabulaire voulu élitiste, ou même des raisons utilitaires, les chercheurs de langues différentes ayant besoin de forger un vocabulaire commun, et quoi de plus commun que des langues, comme le latin et le grec en Europe, dans lesquelles avaient été éduqués la plupart d'entre eux ?

Dans cette première partie, je m'attacherai tout d'abord à définir les notions d'emprunt et de calque. Dans un second temps, je regarderai ce qu'il peut

se passer dans certaines langues avant de chercher les raisons sociologiques et terminologiques qui poussent une langue à emprunter à une autre pour construire un vocabulaire spécialisé.

1.1.1 Le mécanisme de l'emprunt

Josette Rey-Debove (1973, p.109) définit l'emprunt de la façon suivante :

« On appelle emprunt lexical au sens strict le processus par lequel une langue L1, dont le lexique est fini et déterminé dans l'instant T, acquiert un mot M2 (expression et contenu) qu'elle n'avait pas, et qui appartient au lexique d'une langue L2 (également fini et déterminé dans l'instant T). Ce processus se déroule de l'instant T à l'instant T' ; le temps écoulé entre T et T' est très variable et correspond à la codification plus ou moins rapide d'un fait de discours dans la langue. »

À la « codification plus ou moins rapide » dont Josette Rey-Debove parle correspondent plusieurs stades d'intégration d'un mot dans la langue d'accueil. Trois stades peuvent être dégagés, qui ont été nommés, par Guilbert (1975, p.92-93) par exemple, xénisme, pérégrinisme et mot intégré. Michèle Fruyt (1987, p.228) considère ces trois stades comme « un *continuum*, comportant, à une extrémité, le *xénisme*, mot perçu comme étranger par les locuteurs de la langue emprunteuse et désignant des réalités étrangères, puis le *pérégrinisme*, mot qui est encore perçu comme étranger, mais désigne des réalités devenues familières à la langue emprunteuse et, avec des degrés intermédiaires, à l'extrémité opposée, le mot *complètement intégré*, dont l'origine étrangère n'est plus perçue par le sujet parlant ». Le premier stade est donc le xénisme, défini dans le *Trésor de la Langue Française informatisé* (TLFi) comme l'« introduction de mots étrangers dans une langue donnée, sans altération de la graphie, sans les marques de genre et de nombre de la langue-hôte ». Il correspond à un mot qui est toujours perçu dans la langue d'accueil comme étranger : il affiche des propriétés morphologiques, phonologiques et sémantiques propres à une autre langue. Guilbert (1975, p.92) dit que le xénisme « demeure effectivement étranger » et qu'il « produit un effet d'exotisme ». Pour lui, « ces emplois ne relèvent à aucun degré de l'emprunt ». Cependant, nous ne sommes pas ici dans l'autonyme, défini par Rey-Debove (1973) comme la citation d'un mot étranger (1) :

(1) Un emprunt se dit *loan* en anglais.

Le xénisme, s'il garde une consonance pleinement étrangère, conservant sinon sa phonie du moins sa graphie et ses propriétés morpho-syntaxiques, peut être

intégré dans une phrase sans faire référence à des propriétés métalinguistiques (traduction, définition) :

- (2) Donc, *bravissimo*, mon cher ami, et je vous serre la main fortement.
(Gustave Flaubert, *Correspondance : supplément (1879-1880)*, p. 254)

Nous sommes donc là bien dans un premier stade d'emprunt lexical. Le second stade est le pérégrinisme. Le TLFi donne comme définition « variété d'emprunt d'un mot senti comme étranger et en quelque sorte cité », ne le différenciant pas vraiment du xénisme. Or, un pérégrinisme est davantage intégré dans la langue d'accueil qu'un xénisme ; s'il est toujours perçu comme étranger, il peut acquérir certaines des propriétés morpho-syntaxiques de la langue emprunteuse, comme nous le verrons dans la partie 1.1.2 avec des mots grecs empruntés par le latin, qui conservent leur nominatif grec mais sont déclinés selon les règles du latin. En français, c'est le cas de nombre de mots empruntés qui perdent peu à peu les marques du pluriel de leur langue d'origine : on trouve aujourd'hui comme pluriel de *jazzman* aussi bien *jazzmen* que *jazzmans*. Le premier serait encore à considérer comme xénisme quand le deuxième montre le passage à un pérégrinisme. Ce passage se fait par rapport à la réalité extralinguistique, comme le fait remarquer la citation de Fruyt : le pérégrinisme désigne « des réalités devenues familières », comme l'est le jazz introduit dans les pays francophones. Cependant, on ne peut pas considérer ce mot comme totalement intégré, puisqu'il affiche un phonème qui n'est pas un phonème du français. Le troisième stade est donc celui de l'emprunt parfaitement intégré, pour lequel aucune différence n'est perçue avec un mot indigène. Pour Guilbert (1975, p.95), « un terme d'origine étrangère cesse d'être un néologisme à partir du moment où il est entré dans le système linguistique de la langue d'accueil, c'est-à-dire quand, précisément, il cesse d'être perçu comme terme étranger ». Il n'affiche aucune particularité phonique, graphique ou morpho-syntaxique. C'est le cas de nombre de lexèmes du français : *sofa*, *abricot*, *tennis*, *café*, etc. Si la distinction de ces trois stades d'intégration n'est pas primordiale à première vue, il est important de remarquer que deux emprunts peuvent ne pas être perçus de la même façon par les locuteurs d'une langue, certains étant toujours considérés comme étrangers. Les composés néoclassiques que nous étudions dans cette thèse sont-ils perçus comme étrangers ? Quelles conséquences cela a-t-il dans leur formation et leur utilisation ? C'est ce que nous essaierons de voir ultérieurement.

A côté de l'emprunt lexical, qui consiste à emprunter simultanément la forme et le sens d'un mot, d'autres types d'emprunt sont possibles. En effet, une langue peut emprunter à une autre uniquement des traits sémantiques, des

traits syntaxiques et des éléments infralexicaux (préfixes, suffixes, éléments de composition). L'emprunt de traits sémantiques est traité dans la littérature sous le nom de *calque sémantique*. Christian Nicolas (1994, p.75) le définit ainsi :

« Le calque sémantique est un processus d'un fonctionnement compliqué, à la fois linguistique et extra-linguistique, qui consiste à solidariser un concept nouveau dans la langue, non encore lexicalisé, et une unité lexicale pré-existante sur le modèle d'une combinaison concept-lexème telle qu'on peut la voir déjà réalisée dans une langue étrangère. Le calque sémantique est donc le procédé par lequel un locuteur transpose sur un signe de la langue A une relation sens-forme jusque là inédite, en se fondant sur l'existence de cette même relation à l'intérieur d'un signe de la langue B. »

L'exemple souvent donné est celui du verbe français *réaliser* qui a acquis le sens de 'se rendre compte' en calquant le sens du verbe anglais *realize*. Un autre exemple est celui de l'allemand *Ente*, qui n'avait que le sens de 'canard' (volatile). Il s'est enrichi du sens de 'canard' (rumeur de presse), en empruntant ce deuxième sens au nom français *canard*.

Cependant, le calque peut être autre que sémantique : Humbley (1974, p.62) définit le calque comme « la reproduction d'une structure lexicale étrangère avec des éléments de la langue I, qui a un sens différent de celui de la somme des éléments et qui, en principe, correspond au modèle ». La structure d'un mot complexe peut donc être copiée, calquée. Fruyt (1987, p.229) souligne ainsi que seuls des traits syntaxiques peuvent être empruntés : une langue peut donc emprunter à une autre l'ordre des constituants de composés ou locutions, ce que fait le français en copiant sur l'anglais l'ordre déterminant-déterminé dans les expressions contemporaines (*téléshopping*, *open space*). Pour André (1986, p.11), « le calque peut être considéré comme un effort intellectuel pour échapper à l'emprunt ». Ainsi, au lieu d'emprunter au grec le substantif *ποιότης*, 'qualité', formé sur l'adjectif *ποιός*, 'tel, de telle nature', et le suffixe *-της*, ou de le traduire, Cicéron a forgé le mot *qualitās* en prenant comme modèle la structure morphologique du mot grec : l'adjectif *qualis* et le suffixe *-itās*. Nous sommes ici dans un cas de calque morphologique, qualifié par Di Spaldro *et al.* (2010) d'« emprunt de construction ».

Si les calques ne sont pas à proprement parler des emprunts, il est important d'en parler ici, tant ce mode de recours à un lexique ou à un système allogène est fréquent. Dans le sujet de ce travail, le calque semble au cœur de l'analyse diachronique que nous pouvons mener sur les composés néoclassiques. Certains composés ont été empruntés au grec, souvent par l'intermédiaire du latin. *Démocratie* a ainsi été emprunté en 1370 au grec *δημοκρατία*, par l'in-

termédiaire des traductions latines d'Aristote (*democratia*)¹. D'autres sont formés en français mais calqués sur la structure des composés grecs : *gérontocratie* a été formé en 1825 à partir de deux éléments grecs et selon une structure morphologique identique à celle de *démocratie*. Cependant, nous ne sommes plus dans un cas d'emprunt lexical, uniquement de calque. Et Humbley (1974, p.63) précise que « la création lexicale sur un modèle étranger n'est généralement pas reconnue du tout comme un emprunt ». Cependant, une autre façon de considérer la formation de *gérontocratie* est de regarder l'élément *-cratie* comme ayant été emprunté au grec.

En effet, à côté des calques dont nous venons de parler, une langue peut emprunter un élément infralexical, c'est-à-dire un affixe ou un élément de composition. Pour Humbley (1974, p.48), « des éléments autres que des lexies peuvent être empruntés, mais les emprunts de plus petites unités passent obligatoirement par le stade lexical ». Ainsi, des mots doivent être empruntés, puis analysés dans la langue d'accueil comme complexes, décomposés et certains de leurs éléments réutilisés avec des bases indigènes. Pour Fruyt (1987), le critère d'intégration d'un affixe étranger est la possibilité qu'il a de s'ajouter à une base indigène ou sentie comme telle, c'est-à-dire, une base étrangère totalement intégrée, non perçue comme étrangère. Ce serait le cas par exemple du suffixe italien *-issimo*, lui-même hérité du latin, qui est aujourd'hui utilisé en français avec des bases indigènes ou intégrées : le TLFi cite ainsi *élégantissime*, *énormissime* ou même *béjartissime*. Par ailleurs, le suffixe *-issimo* dans sa forme italienne est lui aussi disponible en français pour de nouvelles créations (*colissimo*, *VTTissimo*, *Noëlissimo*, *orthophonissimo*²). Humbley (1974, p.54) ajoute qu' « il semble que les morphèmes les moins liés soient les plus facilement empruntés – ils sont déjà en partie lexicalisés. Il faut d'abord que le morphème soit connu dans une série paradigmatique d'emprunts dans laquelle il figure ; il peut, par analogie, s'il s'intègre dans le système morphologique de la langue I, se généraliser et devenir disponible pour d'autres créations ».

Maintenant que nous avons vu ce que l'on peut mettre sous l'étiquette d'emprunt, interrogeons-nous sur les motivations : pourquoi une langue emprunte-t-elle à une autre ? Pour Humbley (1974, p.47), il faut dégager deux types de causes. Les premières sont des causes externes : une langue emprunte pour des besoins extralinguistiques, c'est-à-dire quand elle a le besoin de dénommer un nouveau concept. C'est ce que précise Fruyt (1987, p.232) pour expliquer les nombreux emprunts faits par le latin au grec :

1. cf. Dictionnaire historique de la langue française (DHLF) Le Robert (Rey, 1992).

2. Ces exemples ont été trouvés sur le corpus en ligne FrWac (http://nl.ijs.si/noske/wacs.cgi/corp_info?corpname=frwac).

« les désignations furent empruntées en même temps que les choses désignées, le besoin d'une dénomination nouvelle se faisant sentir lors de l'adoption d'une réalité nouvelle : à l'époque ancienne, on emprunte le nom d'éléments du monde minéral, végétal, animal et le vocabulaire du commerce, de la navigation ; à une époque plus récente, surtout à partir du 1^{er} siècle avant JC, on forme le vocabulaire des lettres et des sciences (médecine, sciences naturelles, etc ?), des techniques, puis le vocabulaire chrétien, surtout à partir du II^{ème} siècle après JC ».

Par ailleurs, un emprunt peut se faire dans le cadre d'un processus plus généralisé, l'emprunt d'une nomenclature par exemple. Humbley (1974, p.51) souligne qu' « une terminologie forme une micro-structure dans le vocabulaire d'une langue : on ne peut en emprunter des éléments isolés, car, c'est le système, la hiérarchie des éléments, qui importe ». Un mot emprunté peut donc faire partie d'un ensemble plus large. C'est notamment le cas dans les domaines scientifiques et les vocabulaires techniques auxquels l'emprunt, selon Fruyt (1987, p.229), « fournit souvent des vocables d'une assez grande précision sémantique ». A côté de ces causes externes, des raisons internes, qui seraient plutôt des conditions, sont également invoquées : il faut que la structure de la langue d'accueil soit prédisposée à accueillir un élément allogène. C'est surtout vrai quand on parle d'emprunt d'unités infralexicales. Enfin, l'emprunt est à différencier d'autres processus, nommés par Cottez (1980, p.XVII) *adaptation* et *production*. Si l'emprunt suppose que la langue d'accueil importe la totalité d'une unité lexicale, par exemple *hémorragie* emprunté au grec *αἱμορραγία*, il y a adaptation lorsque c'est uniquement le signifiant qui est importé : c'est le cas de *hématose*, adapté du grec *αἱμάτωσις*, 'conversion en sang', qui signifie en français 'transformation du sang veineux en sang artériel par oxygénation du sang dans les poumons' (définition du TLFi). Cottez parle enfin de production lorsqu'un mot est entièrement nouveau, ni emprunté, ni adapté du grec ou du latin. On trouve de nombreux exemples en français, comme *hématologie* qui n'est pas attesté en grec.

1.1.2 Aperçu translinguistique

Le recours à un lexique allogène, que ce soit par emprunt, traduction ou calque, adaptation ou production, se produit dans certaines langues, parfois dans le but de construire un vocabulaire savant. Souvent, cela advient quand une langue, généralement une langue ancienne, possède un certain prestige dans la langue qui y a recours. Ce côté prestigieux permet non seulement aux scientifiques et autres lettrés de distinguer leur domaine en lui attribuant

une terminologie érudite, mais également de marquer leur différence en utilisant un jargon qu'eux seuls peuvent décrypter. Pour illustrer ce propos, je regarderai tout d'abord comment le latin a emprunté au grec. Ensuite, le vietnamien nous donnera un bon exemple d'un processus similaire typologiquement éloigné des langues indo-européennes. Enfin, il sera intéressant d'observer l'utilisation qu'a le grec moderne du grec ancien dans la constitution d'un vocabulaire savant³.

Le recours à une langue étrangère considérée comme prestigieuse est particulièrement fréquent, dans certaines langues, dans la formation du vocabulaire savant. Ce phénomène n'est pas récent. En effet, Bottéro (1987, p.118) rapporte qu'au IV^{ème} millénaire, deux peuples se sont retrouvés en Basse-Mésopotamie : les Sumériens et les Akkadiens, un peuple sémite. Chacun de ces peuples avait sa propre langue, constituée de systèmes typologiquement différents. Les phonèmes disponibles en sumérien étaient plus variés que ceux de l'akkadien. En sumérien, les mots étaient plutôt monosyllabiques et invariables quand l'akkadien, langue flexionnelle, préférait des mots polysyllabiques. Le sumérien a d'abord été l'unique langue officielle, celle de l'administration, de la religion et de la culture. Il a cependant été rapidement concurrencé par l'akkadien, qui est devenu la langue parlée par le peuple mésopotamien. Le sumérien a toutefois survécu parmi les lettrés et est devenu une langue savante.

Prenons encore un exemple, cette fois-ci plus proche de nous, le latin. Deux populations hellénophones demeuraient dans l'Empire romain. D'un côté, des esclaves grecs étaient « importés » ; le grec est la langue familière, celle parlée dans la rue, à l'époque de Plaute (I^{er} siècle avant Jésus-Christ) qui s'en est d'ailleurs inspiré dans ses pièces. Les commerçants parlaient également grec : c'était la langue internationale dans toute la Méditerranée (cette langue grecque, dérivée du grec attique par des changements dus à un usage massif, est appelée la *koinè*, κοινή, terme qui vient de l'adjectif κοινός, 'commun'). Whitehead (2011, p.222) note d'ailleurs que davantage d'emprunts au grec étaient faits dans les niveaux de langue les moins élevés, puisque beaucoup d'esclaves étaient grecs. D'un autre côté, le grec attique, considéré comme un modèle idéal, était la langue d'un niveau élevé, parlée par les lettrés, qui faisaient en général un séjour en Grèce pour parfaire leur bilinguisme. Surtout à partir du I^{er} siècle avant Jésus-Christ, parler grec est de la plus

3. Je choisis ici de me limiter à la question du recours à une langue étrangère pour la constitution d'un vocabulaire spécialisé ; cependant, il existe des langues qui ont également recours à d'autres langues et montrent des phénomènes similaires à ceux dont il est question dans cette partie, mais qui le font pour d'autres raisons. Aronoff et Sridhar (1988) montrent par exemple que le kannada a emprunté au sanskrit beaucoup de vocabulaire dans le registre courant, et a même intégré dans son système des préfixes empruntés à ce dernier.

grande coquetterie (Fruyt, 1987, p.243). Les hellénismes sont rentrés en latin par ces deux biais pour enrichir le lexique savant. C'est ce que note André (1986, p.12) :

« Le premier médecin établi à Rome était, selon la tradition, venu du Péloponnèse. Plus tard, outre les médecins grecs importés comme esclaves, accoururent en Italie beaucoup de praticiens qui ignoraient ou baragouinaient le latin, donnaient en grec un enseignement fondé sur des traités médicaux rédigés en grec, et qui, appelés en consultation au chevet des malades, s'entretenaient en grec. Ils constituaient une société fermée, dont la clientèle aisée savait le grec. A quoi bon dès lors traduire en latin la terminologie médicale? (...) On se contenta d'abord de transcrire les termes grecs, puis la langue les adopta moyennant certaines déformations (...). »

C'est parce que les esclaves, qui pouvaient être médecins, ne parlaient pas une autre langue que le grec que celui-ci a été utilisé dans la médecine. Les lettrés, de leur côté, ont recours au grec pour former des mots du vocabulaire technique ainsi que le vocabulaire chrétien. Ce besoin de dénomination fait écho au progrès des sciences d'une part et à l'arrivée du christianisme d'autre part, car « c'est l'apparition de la chose et de la notion représentées qui donne naissance au terme technique, et l'enrichissement lexical provient du progrès des connaissances scientifiques et techniques » (André, 1986, p.8). C'est en effet par les juifs de langue grecque, les Hellénistes, que le christianisme s'est implanté dans le monde gréco-romain.

D'autre part, le grec a pu apporter une certaine spécialisation que le latin n'avait pas. Par exemple, *ἀήρ*, 'l'air' en tant qu'élément va remplacer sur le plan scientifique les quasi-synonymes *spirites*, *caelum*, *anima*, *uentus* qui n'avaient pas de spécification sémantique. André (1986, p.15) note également que le latin nommait les maladies de façon simple et primitive quand le grec propose une dénomination plus spécifique. Ainsi, pour désigner l'écoulement des humeurs, le latin ne disposait que d'un mot, *destillatio* 'flux, écoulement', et usait de périphrases pour préciser la partie du corps concernée, ce qui donnait lieu à des noms de maladie : *destillatio pectoris*, *pulmonis*, *thoracis*, *faucium*, *capitis*, *narium*, *stomachi*, *ventri*. Le grec a quant à lui trois mots pour désigner et spécifier ces maladies, *κατάρρους*, *ῥυματισμός* et *διάρροια*. Le grec est donc une langue qui permet plus de précisions.

Les Romains bilingues ont non seulement emprunté des termes au grec mais ont également créé des néologismes, ce que Frédérique Biville (2002, p.364) appelle « le grec des Romains » : par un recours conscient ou inconscient à l'analogie, ils créent des formes inexistantes en grec quoique possibles : ce

peut être pour des besoins de dénomination scientifique mais il peut aussi s'agir de mots d'auteur. Plaute s'est d'ailleurs beaucoup moqué de cela en inventant lui-même bon nombre de mots grecs dans ses pièces : dans *Miles gloriosus* (vers 1055) il crée le composé *urbicapus* 'celui qui capture les cités', en s'appuyant sur le composé grec *πολίπορθος* 'qui pille les cités'.

Le grec a eu une telle influence sur le latin que Biville (2002, p.353) parle de l'installation en latin d'un « système alloglotte, véritable « état dans l'état » » : le latin a emprunté en masse des mots au grec, et a su y déceler les relations unissant les différents constituants, en analysant des séries dérivationnelles, identifiant la base et les affixes et mettant au jour les schémas dérivationnels les plus productifs en grec ; les Romains ont ainsi mis en évidence des « couples dérivationnels », par exemple un substantif et un adjectif en *-ικός* construit sur ce substantif. Le latin possédait le suffixe *-icus* et ne l'utilisait qu'avec des bases désignant des personnes. Les emprunts faits au grec, où le suffixe *-ικός* s'attachait à des bases avec une sémantique plus étendue, a fait que le suffixe *-icus*, dans les vocabulaires techniques, s'est retrouvé attaché à des bases plus diverses : des animaux (*cynicus*), des inanimés concrets (*hydraulicus*), des inanimés abstraits, des noms de maladie, de symptôme, de remèdes, etc (Fruyt, 1987, p.245).

Les mots grecs ne sont pas tous intégrés de la même manière dans le système de la langue latine. Biville (2002, p.367) cite les grammairiens latins anciens qui avaient défini une échelle d'intégration des hellénismes en latin : tout d'abord les « *nomina tota graeca* », qui appartiennent à la langue poétique et technique et sont marqués par une absence totale d'intégration. Ils sont soit indéclinables, soit fléchis selon les règles grecques (3). Ces *nomina tota graeca* correspondent, en termes modernes, aux xénismes que nous avons vus dans la partie 1.1.1.

(3) δελφίν, -ίνος > *delphin*, acc. *delphina*⁴

Ensuite, davantage intégrés, il y a les « *uocabula notha* » qui sont qualifiés de mots bâtards, semi-intégrés : ils conservent leur nominatif singulier grec mais sont déclinés selon les modèles latins (4). Ils correspondent ainsi aux pérégrinismes.

(4) σειρήν, -ήνος > *siren*, *-enis*⁵

4. Le nom δελφίν, -ίνος appartient à la troisième déclinaison des noms grecs, qui ont parmi leurs particularités d'afficher une désinence *-α* à l'accusatif singulier. Le nom latin *delphin*, *-inis* appartient à la troisième déclinaison des noms latins. Cependant, contrairement aux autres noms de cette classe qui ont comme désinence pour l'accusatif singulier *-em*, celui-ci aura la désinence *-a*, tout comme le nom grec.

5. Σειρήν, -ήνος appartient également à la troisième classe des noms grecs. Son accusatif porte

Enfin, les « nomina tota latina » sont entièrement latinisés, parfaitement intégrés aux structures latines .

(5) ῥήτωρ, -ορος > *rhetor*, *-oris*

Nous voyons donc que les mots des sciences et des techniques appartiennent aux xénismes, ces emprunts qui sont toujours perçus comme étrangers. Je fais ici l'hypothèse que cette consonance étrangère, qui plus est grecque, rend ces termes plus érudits encore qu'ils ne le sont par leur référence à un concept scientifique. Cette distance instaurée par rapport au vocabulaire courant est nécessaire dans la constitution d'un vocabulaire savant, comme nous le verrons dans la partie 1.1.3.

Le vietnamien⁶ a également eu recours à une langue plus ancienne et surtout plus prestigieuse pour construire un vocabulaire savant : le chinois. Si on reprend l'histoire de ces deux langues, exposée par Thompson (1965), on voit que de 207 avant Jésus-Christ à 939 après Jésus-Christ, le Vietnam était une province chinoise. Les Chinois, qui avaient une culture plus avancée, ont dominé le peuple vietnamien considéré comme « semi-civilisé ». Lorsque les Vietnamiens ont acquis leur indépendance, celle-ci n'a été que politique, tellement la culture chinoise les avait influencés et allait les influencer pendant longtemps. Par exemple, le chinois est resté pendant longtemps la langue parlée dans les tribunaux. Le système d'écriture chinois a également été adopté pour les documents officiels. Même si une adaptation en vietnamien natif des caractères chinois s'est développée, le chinois a connu un vrai prestige, plusieurs siècles après la fin de la domination chinoise (Thompson, 1965, p.52). Cette influence du chinois se fait aujourd'hui encore ressentir en vietnamien, marquant le prestige de cette langue. Par exemple, phonétiquement, la séquence [nw] est très marginale en vietnamien, et n'apparaît que dans des formes savantes, toutes empruntées au chinois :

(6) noãn [nwãn] 'œuf, testicule'

De plus, Thompson (1965) montre que des constructions qu'elle nomme « pseudo-composés » (« pseudo-compounds »), bien plus nombreux que les composés natifs en vietnamien, incluent deux constituants dont au moins un est un constituant non autonome, toujours emprunté au chinois. Ces pseudo-

donc la désinence *-α*. Cependant, le nom latin *siren*, *-enis* est décliné comme un nom standard de la troisième déclinaison latine, avec une désinence *-em* à l'accusatif singulier.

6. Si je m'attarde ici uniquement sur le vietnamien, Arcodia et Basciano (2012) font des analyses similaires concernant le mandarin moderne, le japonais et le coréen.

composés se trouvent dans un usage plus formel que les composés natifs. C'est pourquoi elle les compare aux composés anglais que nous appelons néoclassiques, c'est-à-dire construits à partir d'éléments non autonomes empruntés au fonds lexical grec. En vietnamien ils sont de quatre types, types dégagés par Thompson selon des critères sémantiques. Selon la classification de Bisetto et Scalise (2009), les deux premiers sont du type coordonné, le troisième du type attributif et le dernier du type subordonné. Les premiers sont appelés « generalizing pseudo-compounds » et ce sont des composés coordonnés dont les éléments dénotent différents concepts. Ces composés ont le sens général de 'ces deux items et d'autres similaires', définissant ainsi une classe générale⁷ :

(7) thi-văn 'littérature' (< *thi* 'poésie'; **văn** 'littérature')

(8) ca-tụng 'louer, faire l'éloge de' (< **ca** 'chanter'; *tụng* 'louer')

Ils sont à rapprocher des composés natifs nommés par Thompson « generalizing compounds » :

(9) bàn-ghế 'meuble' (< **bàn** 'table'; **ghế** 'chaise')

Le deuxième type regroupe les « reinforcing pseudo-compounds ». Ce sont toujours des composés coordonnés mais cette fois-ci les constituants dénotent des items identiques ou similaires. Les mots construits ont souvent un sens abstrait ou métaphorique.

(10) liên-hiệp 'unir, allier' (< *liên* 'unir, associer'; **hiệp** 'unir, se rejoindre')

(11) kỳ-dị 'être étrange, bizarre' (< **kỳ** 'être étrange'; *dị* 'être bizarre, différent')

Là encore, ce type de pseudo-composés peut être rapproché de composés natifs, nommés par Thompson « reinforcing compounds » :

(12) sợ-hoảng 'être terrifié' (< **sợ** 'être effrayé'; **hoảng** 'être pris de panique')

Sémantiquement, ces pseudo-composés et composés de type coordonné ont le même comportement.

7. Dans les exemples suivants, les constituants liés sont en italique et les constituants autonomes en caractères gras.

Ensuite, Thompson dégage des pseudo-composés de type attributif⁸, qui représentent la majorité des pseudo-composés. Le constituant de droite constitue la tête du composé, celui de gauche étant le complément. Si certains sont rentrés dans le langage courant, la plupart apporte une dimension formelle ou littéraire au contexte dans lequel ils apparaissent.

(13) quốc-ngữ ‘langue nationale’ (< quốc ‘nation’; ngữ ‘langue’)

(14) quốc-văn ‘littérature nationale’ (< quốc ‘nation’; văn ‘littérature’)

Il existe encore une fois une catégorie de composés natifs correspondante :

(15) học-trò ‘écolier, élève’ (< học ‘école’; trò ‘enfant (en âge d’aller à l’école)’)

(16) chung-cùng ‘ensemble’ (< chung ‘être en commun’; cùng ‘suivre, et, avec’)

Cependant, la grande majorité de ces composés prend en fait comme constituants des éléments empruntés au chinois qui ont acquis une autonomie en vietnamien. Dans l’exemple (15), le mot *học*, ‘école’, est ainsi emprunté au chinois. Les formations récentes semblent être construites par analogie avec ces formes empruntées au chinois. C’est le cas de l’exemple (16) qui contient deux constituants natifs.

Thompson mentionne un quatrième type, les « descriptive pseudo-compounds » qui correspondent à des composés de type subordonné selon Bisetto et Scalise (2009), avec un ordre tête-complément : le complément décrit une spécificité de la tête si celle-ci est nominale et désigne un argument interne de la tête si celle-ci est verbale. Ces composés sont peu nombreux, et sont à étudier séparément selon Thompson : « they represent a different kind of relationship in the Chinese language from which they were borrowed » (p.136). Même si ces pseudo-composés sont d’origine chinoise, ils n’apparaissent pas en chinois dans des constructions similaires et sont donc de purs produits vietnamiens.

(17) xuất-cả ‘exporter’ (< xuất ‘sortir’; cả ‘port’)

Ces pseudo-composés n’ont pas d’équivalent du côté des composés mais Thompson les rapproche, d’un point de vue sémantique, des « descriptive phrases » du type :

8. Si Thompson utilise le terme de « attributive » pour qualifier ces pseudo-composés, ceux-ci ressemblent davantage aux composés de type subordonné selon la classification de Bisetto et Scalise (2009).

(18) hôm nay ‘aujourd’hui’ (< **hôm** ‘jour’ ; **nay** ‘maintenant’)

Ces constructions sont communes en vietnamien, surtout lorsqu’il y a une tête nominale. À côté de ceux déjà cités, Thompson dégage deux autres types de composés qui n’ont pas leur équivalent avec des constituants liés : les « idiom compounds » et les « specializing compounds ».

Si les pseudo-composés sont différents des composés natifs en ce qu’ils mettent en jeu des constituants liés, Thompson rapproche tout de même ces deux types de constructions en dégageant quasiment les mêmes sous-types. Ces constructions semblent faire partie du même système, c’est-à-dire qu’il ne faut pas postuler un système différent pour les constructions issues du chinois, ou calquées sur le chinois. Cependant, ce petit tour d’horizon de la composition en vietnamien nous montre que le trait [+ chinois] apporte une dimension plus formelle, plus littéraire, aux pseudo-composés, ce qui les démarque des composés formés avec des constituants autonomes natifs. De plus, les formations impliquant des constituants liés obéissent aux règles de formation du chinois, qui sont différentes de celles du vietnamien : si un composé NN vietnamien natif affiche un ordre modifié-modifieur, cet ordre est inversé dans un pseudo-composé. Cependant, l’ordre natif tend actuellement à reprendre le dessus : *Âu Châu* ‘Europe’ (littéralement ‘Europe-continent’) est en train d’être remplacé par *Châu Âu* (Arcodia et Basciano, 2012).

Ces constituants liés empruntés au chinois classique ne sont pas réservés au vietnamien. Arcodia et Basciano (2012) notent qu’ils sont aussi présents en mandarin moderne, en japonais et en coréen. Durant la période Meiji (1868-1912), le japonais a créé de nombreux mots pour faire face à une modernisation du pays ainsi que pour retranscrire des idées provenant de l’Ouest. Ces mots, impliquant des constituants liés empruntés au chinois classique, ont ensuite été empruntés par les autres langues des alentours (19) :

- (19) ‘état’ (lit. ‘pays-famille’) :
- a. Japonais : *kokka*
 - b. Mandarin moderne : *guójiā*
 - c. Coréen : *gukga*
 - d. Vietnamien : *quốc gia*

Cela n’est pas sans rappeler les composés néoclassiques européens, empruntés au même fonds classique, que l’on retrouve dans différentes langues, après une adaptation phonologique (voir plus bas l’exemple (20)).

De la même façon, le grec moderne, directement issu du grec ancien qui a donné lieu à tant de formations savantes dans les langues européennes, a eu également besoin d’instaurer une sorte de distance pour créer son vocabulaire scientifique. Ralli (2013) explique que, au XIX^{ème} siècle, une querelle a opposé la langue écrite, appelée *καθαρεύουσα* ‘puriste’, et la langue orale, nommée *δημοτική* ‘démotique, populaire’. Le grec moderne standard, officiellement adopté en 1976, est basé sur la deuxième. Cependant, la *καθαρεύουσα* a été utilisée pour créer du vocabulaire scientifique. Ces formes ne sont aujourd’hui plus identifiables par les locuteurs non savants du grec moderne. Par ailleurs, une autre façon de construire le vocabulaire scientifique a été d’introduire des calques de termes créés dans d’autres langues européennes (surtout en anglais et en français) à partir de mots du grec ancien. C’est ainsi que l’on trouve en grec moderne des allomorphes tels que *xroma-*, calqués sur le français ou l’anglais, alors que le grec moderne a hérité du grec ancien le thème *xromat-*, utilisé pour les nouvelles créations. La plupart de ces composés ainsi formés ne sont pas analysables synchroniquement ; Ralli les considère comme des formations lexicalisées. Là encore, nous voyons donc qu’une distance est nécessaire pour caractériser un vocabulaire savant, scientifique.

C’est cette distance entre composition native et composition néoclassique que je vais étudier dans ce travail de thèse. Les composés néoclassiques, que je définirai de façon précise ultérieurement, mettent en jeu des constituants d’origine classique (grec ancien ou latin) non autonomes, de la même manière qu’en vietnamien avec le chinois. Certaines constructions peuvent être considérées comme des calques : par exemple, les composés néoclassiques des langues romanes ne présentent pas l’ordre ordinaire déterminé-déterminant mais un ordre inverse, déterminant-déterminé⁹, étant par là similaires aux composés grecs, qui seront présentés dans la partie 1.2.1. Ces composés néoclassiques se trouvent dans de nombreuses langues. Certains peuvent même être qualifiés d’internationalismes, puisque, après une adaptation morphologique et phonologique, ils se retrouvent dans une pluralité de langues. Ces internationalismes peuvent être pensés sous l’angle d’une caractéristique partagée par les langues européennes, au même titre que d’autres traits dégagés dans le Standard Average European (SAE), concept introduit par Whorf et Carroll (1956) : comparant les langues amérindiennes comme le hopi aux langues européennes, Whorf a constaté que ces dernières, même celles qui ne sont pas génétiquement reliées, partageaient des similarités grammaticales, sémantiques ou syntaxiques. Haspelmath (2001) a repris ce concept

9. J’évite volontairement ici de parler de tête. Cette notion complexe sera abordée ultérieurement, dans la partie 5.2.1.

de SAE en listant des universaux linguistiques qui s'observent au sein d'un Sprachbund (aire linguistique). Si ceux-ci sont essentiellement syntaxiques (par exemple la présence d'articles définis et indéfinis ou la construction passive formée avec un verbe copule et un participe passé), nous pourrions considérer que ces internationalismes scientifiques sont présents au sein de ce même Sprachbund¹⁰ et pourraient donc constituer une autre caractéristique du SAE. Une observation de la page Wikipédia de *psychology* nous permet ainsi de dire que cet internationalisme est présent dans au moins 91 langues¹¹ :

10. On remarque par ailleurs que l'aire linguistique dépasse ici les frontières européennes puisque dans l'exemple (20) des langues philippines (cebuano) ou africaines (oromo) construisent également du vocabulaire à partir de constituants grecs. On peut supposer toutefois que les périodes coloniales et donc l'influence des langues européennes expliquent l'expansion de l'aire linguistique comprenant des internationalismes scientifiques.

11. Toutes les langues n'ont pas recours à cet internationalisme pour dénommer cette discipline ; on trouve par exemple *sielkunde* en Afrikaans, *bredoniezh* en breton ou *apytu'ũkuaaty* en guarani.

	Psychologie	(Français)	Psihologija	(Letton)
	Psychology	(Anglais)	Психологија	(Macédonien)
	Psicologia	(Italien)	Psikologi	(Malais)
	Psychologie	(Allemand)	Psikologija	(Maltais)
	Psicología	(Espagnol)	Psicologie	(Mirandais)
	Psychologia	(Polonais)	Psicologgia	(Napolitain)
	Saekologyi	(Akan)	Psychologie	(Bas saxon)
	Psicolochía	(Aragonais)	Psychologie	(Hollandais)
	Psüchologii	(Alémanique)	Psykologi	(Norvégien nynorsk)
	Sicoloxía	(Asturien)	Psykologi	(Norvégien bokmål)
	Psixologiya	(Azéri)	Psikologia	(Novial)
	Psikoluogëj	(Samogitien)	Psychologie	(Normand)
	Psihologija	(Bosnien)	Psicologia	(Occitan)
(20)	Psicologia	(Catalan)	Saayikooloojii	(Oromo)
	Sikolohiya	(Cebuano)	Психологи	(Ossète)
	Psicologia	(Corse)	Psychologie	(Picard)
	Psychologijô	(Cachoube)	Psicologìa	(Piémontais)
	Psychologie	(Tchèque)	Psicologia	(Portugais)
	Seicoleg	(Gallois)	Psicologia	(Romanche)
	Psykologi	(Danois)	Psihologie	(Roumain)
	ψυχολογία	(Grec)	Психология	(Bulgare)
	Psikologio	(Esperanto)	Психология	(Russe)
	Psühholoogia	(Estonien)	Psicoluggìa	(Sicilien)
	Psikologia	(Basque)	Psychologie	(Scots)
	Sicologia	(Estrémègne)	Psicologia	(Sarde)
	Psükoloogia	(Võro)	Psihologija	(Serbo-croate)
	Psykologia	(Finnois)	Psychológia	(Slovaque)
	Psüchologii	(Frison septentrional)	Psihologija	(Slovène)

Psicologjie	(Frioulan)	Saikolosi	(Samoan)
Psychology	(Frison occidental)	Psikologjia	(Albanais)
Psicoloxía	(Galicien)	Психологија	(Serbe)
Psihologija	(Croate)	Psychologie	(Frison oriental)
Pszichológia	(Hongrois)	Saekholoji	(Sesotho / sotho du sud)
Psikologi	(Indonésien)	Psikologi	(Soudanais)
Sikolohia	(Ilokano)	Psykologi	(Suédois)
Psikologio	(Ido)	Sikolohiya	(Tagalog)
Psikologi	(Javanais)	Saikolosi	(Néo-malanésien)
Психология	(Kazakh)	Psikoloji	(Turc)
Psíkolojî	(Kurde)	Психология	(Tatar)
Psikolojiya	(Ladino)	Психологија	(Ukrainien)
Psychologia	(Latin)	Psixologiya	(Ouzbek)
Psychologie	(Luxembourgeois)	Sicologia	(Vénitien)
Психология	(Lezghien)	Psychologie	(Flamand occidental)
Psicologia	(Ligure)	Psikolohiya	(Winaray)
Psychologie	(Limbourgeois)	Psychohohie	(Zélandais)
Psichologija	(Lithuanien)		

Ce phénomène de recours à un lexique allogène pour constituer un vocabulaire savant est donc translinguistique. Cependant, pourquoi avoir besoin d'emprunter à une autre langue pour créer le vocabulaire d'un domaine particulier ? En me penchant sur des aspects davantage sociologiques et terminologiques, je vais maintenant tenter de répondre à cette question.

1.1.3 Pourquoi un lexique allogène ?

Par lexique savant, j'entends ici un lexique relevant du domaine technico-scientifique et qui s'inscrit, dans la dichotomie langue spécialisée / langue générale, du côté de la première. La langue spécialisée a des caractéristiques qui lui sont propres. Elle peut être définie selon deux points de vue : le(s) sujet(s) traité(s) et les locuteurs qui la pratiquent. Les sujets sont spécifiques à un domaine particulier : la médecine, l'astronomie, la chimie, la linguistique, etc. Les locuteurs sont les spécialistes de ce domaine ; si un spécialiste s'adresse à un non spécialiste, il aura besoin, pour se faire comprendre, de restreindre l'utilisation d'un jargon qui serait inintelligible pour son interlocuteur. La question des langues spécialisées a été très largement étudiée dans la littérature, et donc très largement décrite. Hoffmann (1979, p.16) affirme par exemple : « By Language for special purposes (LSP) we understand a complete set of linguistic phenomena occurring within a definite sphere of communication and limited by specific subjects, intentions, and conditions ».

Pour d'autres cependant, les différences pouvant être relevées entre langue spécialisée et langue générale ne se situent qu'au seul niveau lexical : « il faut noter que les expressions « langue de spécialité » et « langue commune » ne recouvrent qu'un sous-ensemble de la langue, celui des lexèmes » (Rondeau, 1981, p.23-24) ; « au sens propre il n'existe pas de langue du droit en soi mais seulement, à l'intérieur de la langue française, un vocabulaire du droit, et sans doute quelques tours syntaxiques spécifiques » (Mounin, 1979, p.13). Cependant, Bowker et Pearson (2002, p.26) soulignent qu'il ne faut pas limiter l'analyse des langues spécialisées par le seul angle du lexique : « although the specialized vocabulary of an LSP is often its most striking feature, it is important to note that LSP is not simply Language for general purposes (LGP) with a few terms thrown in. An LSP may also have special ways of combining terms or of arranging information tant differ from LGP ». La différence entre langue spécialisée et langue générale ne se limiterait donc pas au vocabulaire mais serait également active au niveau de la syntaxe par exemple. Je m'intéresse ici au vocabulaire qui peut être spécifique aux langues technico-scientifiques, en particulier aux constructions qui ont été empruntées ou sont adaptées d'autres langues. Quand il s'agit de créer de nouveaux mots, le but n'est donc pas le même : « il ne s'agit plus en effet de rechercher le terme neuf en soi et pour soi, par souci de variation de l'expression ou par fantaisie, mais il devient nécessaire d'inventer un vocable précis, concis et univoque, destiné à remplir un vide dans un champ lexical spécifique : une langue technique se doit de se doter d'un lexique *ad hoc* répondant aux exigences de rigueur fondamentales de toute discipline scientifique. » (Nicolas, 1996, p.35). En effet, langue spécialisée et langue générale divergent par le but qui leur est donné. Quand la seconde intègre ce qui relève de l'expressivité, de la fantaisie, du jeu de mots, cela est totalement exclu de la première. Le but d'une langue spécialisée, comme la langue scientifique, est de donner une information rigoureuse, qui nécessite une précision parfaite. En linguistique, *lexème* et *lemme* sont des termes qui, bien que pouvant être confondus par des non spécialistes, recouvrent des concepts bien différenciés par la communauté des linguistes. Un mot est donc créé dans la langue spécialisée afin de dénommer un concept précis et se place à une certaine distance sémantique des autres mots du domaine. Prenons à titre d'exemple des termes appartenant au domaine linguistique ; quand Mark Aronoff a créé le mot *morphome*, c'était pour évoquer une conception que ne recouvraient pas les mots *morphème* ou *radical* : la création lexicale est ici au service de l'avancée scientifique, qui fait des distinctions de plus en plus fines et donc nécessite autant de nouvelles dénominations. Pour Benveniste (1966), l'acte de dénomination est l'acte majeur dans la constitution d'une science :

« La constitution d'une terminologie propre marque dans toute science l'avènement ou le développement d'une conception nouvelle, et par là elle signale un moment décisif de son histoire. On pourrait même dire que l'histoire propre d'une science se résume en celle de ses termes propres. Une science ne commence d'exister ou ne peut s'imposer que dans la mesure où elle fait exister et où elle impose ses concepts dans leur dénomination. Elle n'a pas d'autre moyen d'établir sa légitimité que de spécifier en le dénommant son objet, celui-ci pouvant être un *ordre* de phénomènes, un *domaine* nouveau ou un mot nouveau de *relation* entre certaines données. L'outillage mental consiste d'abord en un inventaire de termes qui recensent, configurent ou analysent la réalité. Dénommer, c'est-à-dire créer un concept, est l'opération en même temps première et dernière d'une science. Nous tenons donc l'apparition ou la transformation des termes essentiels d'une science pour des événements majeurs de son évolution. Tous les trajets de la pensée sont jalonnés de ces termes qui retracent des progrès décisifs et qui, incorporés à la science, y suscitent à leur tour de nouveaux concepts. C'est que, étant par nature des inventions, ils stimulent l'inventivité. »

Un texte technique, souligne Cabré (1998, p.92), doit répondre donc à différentes exigences : la concision, la précision et l'adéquation à la situation de communication. La concision est nécessaire pour éviter le bruit que pourrait apporter une information non nécessaire. La précision est primordiale car le domaine technico-scientifique ne supporte pas des informations vagues. Une des conceptions dominantes de la pensée scientifique (qui remonte entre autres à Karl Popper (1973)) serait ainsi qu'il n'est de scientifique que ce qui est réfutable. Or plus une information est précise, plus elle est potentiellement réfutable. La précision est donc une caractéristique clé du raisonnement scientifique et donc d'un texte scientifique. Enfin, l'adéquation suppose qu'un texte technico-scientifique s'adapte à la situation de communication dans laquelle il est produit, aux caractéristiques des interlocuteurs et leur niveau de connaissance du domaine. Un texte comprendra donc plus ou moins de « jargon », plus ou moins de lexique spécialisé, et donc de lexique emprunté ou adapté d'autres langues.

Maintenant que nous avons vu les spécificités de la langue spécialisée, voyons en quoi le recours à un lexique allogène a pu aider à la constitution de langues de spécialité telles que les langues technico-scientifiques qui sont « un faire-connaître qui vise une transposition explicative du sens et suppose un méta-discours » (Gambier, 1991, p.9). Tout d'abord, comme nous venons de le

voir, toute science qui s'élabore, se transforme, a besoin d'une terminologie. « L'utilisation de termes normalisés contribue à rendre plus efficace la communication entre spécialistes », dit Cabré (1998, p.92). Cette terminologie doit être rigoureuse et homogène : le fait d'utiliser les mêmes suffixes pour former des concepts comme *lexème*, *morphème*, *monème*, *sémantème*, *phonème*, *synthème*, etc. offre une vision homogène des concepts utilisés en linguistique et permet, à chaque fois qu'un linguiste souhaite créer un nouveau terme, de l'inclure dans le paradigme des concepts utilisés dans le domaine. Comme nous le verrons dans la partie 1.2.3, les savants français, et plus largement européens, du XVII^{ème} et XVIII^{ème} siècles ont d'abord emprunté des mots au grec ancien et au latin, dans la mesure où ces langues étaient largement utilisées dans la communication scientifique. Il a donc été tout naturel, quand il a fallu créer de nouveaux mots pour désigner de nouvelles disciplines, méthodes, etc., de les construire sur les patrons classiques existant, afin de garder une homogénéité dans la terminologie. On crée des mots qui ressemblent aux mots déjà existants. C'est pourquoi Kocourek (1982, p.108) affirme que « les mots confixés¹² sont généralement terminogènes ».

Ce qui est donné comme primordial pour les sciences et les techniques est la diffusion du savoir. « La *fabrication* de la science ne va pas sans préoccupation de sa diffusion », affirme Gambier (1991, p.13), qui met l'accent sur « le besoin de reconnaissance, de crédit » qu'éprouvent les spécialistes d'un domaine. Or cette diffusion du savoir ne peut se faire que sur la base d'une terminologie commune. De plus, le savoir ne se diffuse pas qu'au sein d'une communauté linguistique : une dimension internationale est à prendre en compte. Avoir recours à un lexique allogène, qui plus est à un lexique hérité d'une langue ancienne que quelques-uns partagent comme langue mère, offre la possibilité de jouir d'internationalismes, c'est-à-dire de termes qui, modulo une adaptation phonologique, sont identiques dans toutes les langues. Le latin et le grec étaient par ailleurs les langues privilégiées pour la communication entre spécialistes et ont ainsi donné lieu aux composés néoclassiques auxquels je m'intéresse dans ce travail de thèse. Un exemple d'internationalisme a été donné dans la partie 1.1.2. Cabré (1998, p.159) affirme que « les normes internationales de terminologie vont jusqu'à recommander expressément le recours aux formes savantes pour les besoins de la néologie terminologique, souvent appelée « néonymie », afin de promouvoir l'intercompréhension internationale ».

12. Kocourek (1982) utilise *confixation* comme synonyme de *formation savante* ; pour elle, « il y a confixation si deux ou plusieurs morphèmes liés gréco-latins s'unissent, et donnent un mot confixé ». La question des différentes appellations données à ces formations sera par ailleurs traitée dans la partie 1.3.1.

Une deuxième raison que nous pouvons trouver est la volonté de se reconnaître dans un jargon, qui confère à ses pratiquants un statut particulier : « On peut définir la langue technique comme la langue spécifique d'un groupe social utilisateur d'une technique ou spécialiste d'une science, ayant une activité fermée et refusant même parfois de communiquer ses méthodes. En cela, elle se rapproche des argots par le désir de rendre plus secrète la communication entre individus d'un même groupe, secret qui met en valeur l'initié et lui donne un sentiment de supériorité » (André, 1986). Ce « besoin de reconnaissance » dont parlait Gambier se retrouve ici. Avoir une place au sein de la communauté des spécialistes d'un domaine, c'est partager un jargon avec les autres spécialistes, ce que confirment Sager *et al.* (1980, p.3) : « special languages, or more precisely special subject languages, are usually thought of as means of expression of highly qualified subjects specialists like engineers, physicians, lawyers, etc. and are often derogatively referred to as « jargon » ». Le recours à un lexique allogène, là encore, renforce cette impression de jargon : au XVII^{ème} siècle, les scientifiques, bénéficiant grâce à leur formation d'une maîtrise du grec et du latin, excluaient d'office les non-spécialistes en empruntant des termes aux langues classiques ou en créant des termes sur le modèle classique pour la constitution de la terminologie de leur domaine. Aujourd'hui, quand beaucoup de scientifiques ne maîtrisent plus les langues classiques, le jargon formé par des mots d'origine étrangère continue d'exclure les néophytes. Ainsi, tout médecin employant un mot se terminant en *-tomie* signifie à un confrère qu'il parle de l'incision de l'objet décrit par le premier constituant du mot. Or s'il veut s'adresser à un patient, il fera mieux d'employer la paraphrase *incision de* pour se faire comprendre.

Cependant, on peut assister à une détechnification (mot emprunté à André (1986)) de certains termes. En effet, certains mots construits à des fins scientifiques passent dans la langue commune. Bowker et Pearson (2002, p.26) utilisent un terme de Meyer et Mackintosh (2000), la déterminologisation, et expliquent ce passage de la langue spécialisée à la langue générale : « this usually happens when terms that once belonged exclusively to a specialized domain make their way into the everyday lives of ordinary people, either through the mass media or through direct impact. (...) Of course, a lay person's understanding of such terms may not be as deep as an expert's understanding which means that the terms may be used in a slightly different way in LGP from LSP ». Ce qui m'intéresse ici, plus que la détechnification ou la déterminologisation, c'est la diffusion dans la langue commune d'un processus de formation de mots, et non plus de mots en eux-mêmes. Nous avons déjà vu plus haut que la différence entre langue spécialisée et langue générale se situe au niveau du but qui leur est donné : l'expressivité, la fan-

taisie, ne sont réservées qu'à la seule langue générale, ce qui se ressent au niveau de la création lexicale : « du point de vue de leur fonction, les néologismes se classent en néologismes référentiels et expressifs. Les premiers font leur apparition parce qu'ils sont nécessaires, c'est-à-dire parce qu'il faut combler une lacune dénomminative dans un domaine de spécialité précis ; les seconds naissent simplement pour fournir de nouvelles formes expressives dans la communication. » (Cabré, 1998, p.256). La langue générale, que ce soit en littérature ou dans la vie quotidienne, use des néologismes à des fins expressives : il s'agira de faire des jeux de mots ou de colorer son propos d'une émotion particulière. C'est ce que souligne Cabré (1998, p.92) : « dans [les textes généraux], l'expressivité, la variété et l'originalité sont plus importantes que d'autres caractéristiques. En revanche, les textes spécialisés sont caractérisés par la concision, la précision et l'adéquation ». L'exemple (21), issu d'un ouvrage de médecine, comme son nom l'indique, affiche un mot utilisant le constituant *-tomie*, répondant au besoin d'une dénomination précise. Le mot apparaissant dans l'exemple (22) n'existe lui que par imitation du processus de formation de mots utilisé dans le langage médical.

- (21) Craignant déjà d'être forcé de recourir à la **trachéotomie**, j'ai pris ces précautions pour empêcher que la vésication ne s'étendît au-dessous de la région thyroïdienne. (Pierre Bretonneau, *Des Inflammations spéciales du tissu muqueux, et en particulier de la diphtérie*, 1826, p. 305)
- (22) deuxièmement, l'estropiation ou **pédotomie** volontaire à seule fin de non-belligérance est quasiment réprimandée par la justice locale (Georges Perec, *Quel petit vélo à guidon chromé au fond de la cour ?*, 1966, p. 27)

La littérature, d'autant plus chez des auteurs reconnus pour être particulièrement créatifs du point de vue lexical comme Georges Perec ou Raymond Queneau, regorge de termes pseudo-savants de la même sorte que celui que nous trouvons ici. Cependant, ces termes pseudo-savants ne sont pas dévolus à la seule littérature. Dans la vie quotidienne, il est tout à fait fréquent de trouver des mots formés grâce à des procédés savants, comme le montre l'exemple (23) trouvé dans un forum :

- (23) sa s'appel la **coeurotomie** il se contracte et a la fin il explose et sa t'explose le bide comme si tu t'etai pri une grenade de MGL dans la poitrine¹³

13. <http://www.jeuxvideo.com/forums/1-50-31414673-1-0-1-0-maladie-qui-comprim>

Dans cette thèse, comme je le montrerai dans la partie 2.3, les mots formés dans la langue courante grâce à ces procédés de formation de mots à l'origine réservés aux vocabulaires technico-scientifiques ont la part belle. Le nombre d'occurrences trouvées sur le Web n'appartenant pas à un vocabulaire spécialisé, qui sont le fait de locuteurs non spécialistes, surpasse en général celui du nombre d'occurrences créées pleinement dans un périmètre lexical technico-scientifique. Cela me permettra de formuler l'hypothèse que ce mode de formation ne dépend plus d'une possible connaissance étymologique mais est accessible par tout locuteur, initié ou non aux langues classiques. La création de ces items se fait cependant en général dans un contexte particulier, souvent humoristique, et le recours à ce mode de formation permet une imitation du discours scientifique.

1.2 Du grec ancien aux langues modernes

Après avoir vu de façon générale les processus du recours à une langue allo-gène pour la constitution d'un lexique savant, ainsi que quelques exemples, recentrons nous sur notre sujet : les mots créés dans les langues modernes, en particulier le français, à partir de constituants d'origine classique. Pour comprendre les ressorts de ces formations, il me semble important d'observer avant tout les moyens qu'avaient les langues classiques pour former des composés. Le but n'est pas d'en faire une description exhaustive ; je m'attacherai seulement à donner les principales caractéristiques des composés dans ces langues. Cela me permettra de différencier ou de rapprocher les processus de composition dans les langues classiques et dans les langues modernes. A ma connaissance, peu de travaux existent sur la composition en grec ancien, du moins en français ou en anglais, davantage sur le latin, ce qui peut paraître paradoxal, le latin étant bien moins friand de composition que le grec ancien. Les observations de la composition dans ces langues m'amèneront ensuite à une étude diachronique du passage d'éléments classiques dans la langue moderne et de la création des composés néoclassiques, principalement en français.

1.2.1 La composition en grec ancien

Pour cette section, je m'appuierai essentiellement sur les travaux de Smyth (1920), Chantraine (1933), Grandi et Pompei (2010) et Le Feuvre (2012) qui sont parmi les rares, à ma connaissance, à traiter de la composition en

grec ancien, du moins dans des langues qui me sont accessibles. Je m'aiderai également des remarques que fait Ralli (2013) sur le grec ancien afin d'éclairer le processus de composition en grec moderne. Tout d'abord, comme le rappellent entre autres Biville (1985) et Ralli (2013, p.273), il n'y avait pas une langue unique en Grèce antique. Ce qu'on appelle aujourd'hui grec ancien regroupe en fait une pluralité de dialectes. Cependant, la domination d'Athènes a fait que le grec qui y était parlé, l'attique, est considéré comme la langue de référence autrement appelé grec classique. C'est donc à celui-ci que je fais référence dans cette section quand je parle de « grec ancien ».

D'une manière générale, quand je parle de composition, et jusqu'à ce que cette notion soit définie (section 5.2.1), je réfère à un processus de formation de lexèmes qui s'appuie sur l'union de deux lexèmes, parfois davantage. Elle se distingue de la dérivation en ce que cette dernière implique un lexème, la base, et un exposant de construction, par exemple un affixe. Suivant Guevara et Scalise (2009), un composé peut donc être représenté comme en (24), où X et Y représentent les constituants, \mathfrak{R} la relation qui existe entre eux et Z la catégorie de l'output.

$$(24) \quad [X \mathfrak{R} Y]_Z$$

Comme Ralli (2013) l'explique, la composition en grec ancien serait basée sur le thème (*stem-based*) où le thème serait la concaténation d'une racine et d'une voyelle thématique. Dans le mot *ἄνθρωπος* 'homme', la racine serait donc dans cette perspective *ἄνθρωπ-* et la voyelle thématique *-ο*. Le ς final serait la désinence flexionnelle du nominatif singulier. Si définir avec exactitude ce qu'est un thème ou se poser la question de la pertinence de la notion de voyelle thématique serait intéressant, je ne retiendrai cependant pour le traitement des composés en grec ancien que le fait que le matériel employé est du matériel lexical dépourvu de désinences flexionnelles. Les données exposées par Smyth (1920) permettent de dégager différents types de composés.

Parmi les composés nominaux, plusieurs configurations sont possibles. Quand le second constituant est un nom qui fournit au composé ses propriétés syntaxiques, le premier constituant peut-être soit nominal (25) soit adjectival (26)¹⁴.

14. En grec ancien, selon la grammaire grecque de Allard (1971), trois classes flexionnelles se dégagent pour les noms ; ces classes sont dégagées selon la finale du thème des lexèmes – c'est-à-dire ici la finale de la séquence précédant la désinence –, et des désinences caractéristiques. La première déclinaison comprend des noms ayant un thème en *-α-* (*ἡμέρα*, gén. *ἡμέρας* 'jour') qui se réalise parfois en *-η-* (*κεφαλή*, gén. *κεφαλῆς* tête'). La deuxième déclinaison comprend les lexèmes, surtout masculins, ayant un thème en *-ο-* (*λόγος*, gén. *λόγου* 'parole'). Enfin, la troisième déclinaison comprend des lexèmes affichant soit un thème consonantique (*κόραξ*, gén.

- (25) [NN]_N σωματοφύλαξ ‘garde du corps’ < σωμα (gén. σωματος) ‘corps’ ;
φύλαξ ‘garde’
- (26) [AN]_N ψευδομαρτυρία ‘faux témoignage’ < ψευδής ‘faux’ ; μαρτυρία
‘témoignage’

Une autre construction qui se rencontre beaucoup est celle du type (27) :

- (27) [NV]_N λογογράφος ‘celui qui écrit les discours’ < λόγος ‘discours’ ;
γράφω ‘écrire’

Ces composés sont constitués d’un premier élément nominal qui est l’argument du second élément, qui lui est verbal. Le statut accordé à cet élément verbal diffère selon les auteurs. Pour Smyth (1920, p.251), il s’agit d’un « verbal element that is not used separately ». Le Feuvre (2012, p.182) parle d’un « nom d’agent déverbal ». Dans ces cas-là, la séquence verbale et le suffixe sont considérés comme ne formant qu’un seul élément, déverbal, qui ne se trouve que dans ce type de composés. Cependant je ne pense pas que le suffixe -ος soit un suffixe dérivationnel qui opérerait un changement catégoriel. Il s’agit plutôt d’une désinence permettant d’insérer le composé dans un paradigme flexionnel. La désinence -ος marque ainsi que le composé est un nom de deuxième déclinaison au nominatif singulier. De nombreux verbes grecs se retrouvent ainsi en composition, par exemple ceux cités en (28) :

- (28) a. φαγειν ‘manger’ > -φάγος ex. άρτοφάγος (< άρτος ‘pain’) ‘celui
qui mange du pain’
- b. γράφω ‘écrire’ > -γράφος ex. λογογράφος (< λόγος ‘discours’) ‘celui
qui écrit les discours’
- c. φερώ ‘apporter’ > -φόρος ex. άνθοφόρος (< άνθος ‘fleur’) ‘celui
qui porte des fleurs’
- d. δέμω ‘construire’ > -δόμος ex. οίκοδόμος (< οίκος ‘maison’) ‘constructeur,
architecte’
- e. τρέχω ‘courir’ > -δρόμος ex. όρειδρόμος (< όρος ‘colline’) ‘celui
qui court sur les collines’
- f. ποιέω ‘faire’ > -ποιός ex. άγαλματομοιός (< άγαλμα ‘statue’) ‘sculpteur’

Certains de ces constituants verbaux sont phonologiquement éloignés de la forme de citation du verbe car formés sur un autre thème du paradigme qui

κόρακος ‘corbeau’ ; σῶμα, gén. σώματος ‘corps’ ; δαίμων, gén. δαίμονος ‘démon’ par exemple), soit des lexèmes affichant d’autres voyelles thématiques , par exemple -ι- (πόλις, gén. πόλεως ‘cité’, lexème dans lequel le thème πολι- alterne avec πολε-).

semble spécifique à ce type de composés. Le verbe *τρέχω* au présent présente ainsi un thème différent de celui sélectionné pour le futur (*δραμουμαι*), l'aoriste (*ἔδραμον*) ou le parfait (*δεδράμηκα*). C'est certainement ce qui a pu amener certains auteurs à traiter ce type de composés de manière spécifique par rapport aux autres types. Ces constituants verbaux suffixés se rencontrent également dans les composés adjectivaux, comme nous le verrons en (31). Pour Grandi et Pompei (2010), il s'agit de constructions du type $[[X]_X[\text{RAD}]_N]_{A/N}$. Ils traitent les composés nominaux comme des endocentriques alors que les adjectivaux sont traités comme des exocentriques. Cependant, même si certaines séquences comme *φάγος*, *φόρος*, *δόμος* ou *δρόμος* peuvent se rencontrer comme lexèmes autonomes (avec respectivement les sens de 'glouton', 'tribut', 'construction' et 'course'), cela ne justifie pas une analyse de ces composés comme des composés endocentriques. D'une part, le sens du verbe dans le composé n'est pas exactement le même que celui du lexème autonome. D'autre part, ces composés sont pour moi à analyser comme des composés synthétiques, selon la définition donnée par exemple par Melloni et Bisetto (2010, p.217) : « a morphological merger of two stems (with compositional semantics) which is licensed by the presence of a category-changing affix ». Dans ce cas, les désinences suffixales *-ος*, *-ε*, *-ο*, *-ου*, etc. sont des désinences flexionnelles permettant l'intégration du nom constitué en syntaxe ; ce ne sont pas des suffixes permettant la nominalisation de l'élément verbal uniquement. Le deuxième élément ne peut donc pas être la tête du composé qui est donc exocentrique, au même titre que les composés adjectivaux que nous verrons en (31).

Par ailleurs, une autre construction nominale, qui semble peu courante, met en jeu un adjectif et un verbe pour la formation d'un composé nominal (29) :

(29) $[\text{AV}]_N$ *μόναρχος* 'monarque' < *μόνος* 'seul' ; *ἄρχος* 'diriger'

Les composés adjectivaux sont pour Smyth (1920) les plus fréquents en grec ancien. Parmi eux, nous trouvons des composés constitués de deux noms (30) ou des composés synthétiques (31), du même type que ceux que nous venons de voir pour les nominaux.

(30) $[\text{NN}]_A$ *κεφαλαλγής* 'qui a mal à la tête' < *κεφαλή* 'tête' ; *ἄλγος* 'douleur'

(31) $[\text{NV}]_A$ *μαντιπόλος* 'inspiré' < *μάντις* 'devin' ; *-πόλος* (< *πέλω* 'devenir')

Les constituants peuvent également être des adjectifs, combinés soit à des noms (32), soit à d'autres adjectifs (33) :

(32) [AN]_{A/N} μελάγχολος ‘plongé dans une bile noire’ < μέλας ‘noir’ ;
χολή ‘bile’

(33) [AA]_A λευκομέλας ‘gris’ < λευκός ‘blanc’ ; μέλας ‘noir’

Pour les noms [+humains], la distinction est souvent poreuse entre nom et adjectif : l'exemple (32) peut donc avoir l'une ou l'autre des catégories, comme c'est le cas pour grand nombre de constructions [AN] désignant une caractéristique que possède le référent du composé, comme en (34) :

(34) [AN]_{A/N} μακρό-χειρ ‘qui a de longs bras’ < μακρός ‘long’ ; χειρ ‘bras’

Nous trouvons par ailleurs des composés [VN] qui peuvent aussi être adjectivaux ou nominaux (35) :

(35) [VN]_{A/N} φέρρασπις ‘qui porte un bouclier’ < φέρω ‘porter’ ; άσπίς ‘bouclier’

Enfin, nous rencontrons en grec ancien des composés dont le premier élément est une préposition et le deuxième un nom, verbe, ou adjectif, le composé prenant les caractéristiques sémantico-syntaxiques de ce second élément (36)¹⁵ :

(36) a. [Prép N]_N εἴσοδος ‘entrée’ < εἴς ‘dans’ ; ὁδός ‘chemin’
b. [Prép V]_V ἀποφεύγω ‘fuir, échapper’ < ἀπό ‘de (depuis un lieu)’ ;
φεύγω ‘fuir’

En ce qui concerne le segment final du premier élément, autrement dit le segment qui se trouve à la jonction des deux constituants du composé, celui-ci varie. Quand le premier élément est un nom, la voyelle sera un *a* (37-a) ou un *o* (37-b), plus rarement un *η* (37-c), quelle que soit la classe flexionnelle du nom initial.

(37) a. ἀγορανόμος ‘vendeur sur le marché’ < ἀγορά (gén. ἀγοράς) ‘marché’ ; νομός ‘usage’
b. δικογράφος ‘qui compose les discours médico-légaux’ < δίκη (gén. δίκης) ‘droit, justice’ ; λογογράφος ‘qui écrit les discours’ < λόγος (gén. λόγου) ; -γράφος ‘celui qui écrit’
c. ἐλαφηβόλος ‘qui chasse le cerf’ < ἔλαφος (gén. ἔλαφου) ‘cerf’ ; -βολος ‘celui qui lance’

15. La différence entre composés [Prép N] et préfixés est souvent mince (voir par exemple sur ce sujet Amiot (2004)). Je ne m'attarderai pas dessus dans cette thèse, cependant, il faut noter que les exemples donnés en (36) pourraient tout aussi bien être considérés comme des préfixés.

Si le deuxième élément du composé commence par une voyelle, celle-ci servira de voyelle de jonction :

(38) κεφαλαλγής ‘mal de tête’ < κεφαλή (gén. κεφαλής) ‘tête’

Si le nom appartient à la troisième déclinaison, c’est-à-dire la déclinaison regroupant les thèmes consonantiques ou les thèmes se terminant en *-υ*, *-ι*, soit le thème vocalique est conservé (39-a), soit la forme du nominatif est sélectionnée (39-b), soit c’est le thème consonantique et dans ces cas-là, un *ο* (39-c) ou un *α* (39-d) est inséré :

- (39) a. ιχθυβόλος ‘qui pêche un poisson’ < ιχθύς (gén. ιχθύος) ‘poisson’ ; -βολος ‘celui qui lance’
 b. σελασφόρος ‘qui apporte la lumière’ < σέλας (gén. σέλατος) ‘lumière’ ; -φόρος ‘celui qui (ap)porte’
 c. σωματοφύλαξ ‘garde du corps’ < σωμα (gén. σωματος) ‘corps’ ; φύλαξ ‘garde’
 d. ποδάνιπτρον ‘eau qui sert à laver les pieds’ < πους (gén. ποδος) ‘pied’ ; νίπτρον ‘eau pour laver’

Quand le premier élément est un verbe, le thème du présent ou de l’aoriste sera sélectionné. Devant une voyelle, ce thème ne sera pas modifié (40-a) ou effacera sa voyelle finale (40-b). Devant une consonne, une voyelle sera ajoutée : *ε*, *ο* ou *ι* (40-c) :

- (40) a. φέρασπικς ‘qui porte un bouclier’ < φέρω ‘porter’ (thème du présent : φέρ-) ; ἄσπικς ‘bouclier’
 b. μισάνθρωπος ‘qui hait les hommes’ < μισέω ‘haïr’ (thème du présent : μισέ-) ; ἄνθρωπος
 c. νικόβουλος ‘qui prédomine à l’assemblée’ < νικάω ‘prédominer’ (thème du présent : νικά-) ; βουλεύς ‘assemblée’

Nous voyons donc que la voyelle *ο* n’est pas la seule voyelle employée pour la composition. Le grec moderne ne conservera cependant que celle-ci : ayant perdu son caractère thématique, elle a fait l’objet d’une réanalyse et elle n’est plus qu’un marqueur de composition (Ralli, 2013). Dans les autres langues modernes, comme nous le verrons tout au long de ce travail, le continuateur du *ο* grec demeure également la voyelle marquant une composition héritée ou adaptée du grec.

Concernant l’ordre des constituants, dans les composés grecs, l’élément déterminant précède en général celui déterminé. Si on suit la classification de Bisetto et Scalise (2009), on peut rendre compte des composés grecs selon le

Tableau 1.1.

Attributifs	Endocentriques	- [AN] _N : ἄκρο-πολις ‘ville haute, citadelle’ (< ἄκρος ‘le plus haut’; πόλις ‘cité’) - [NN] _N : μητρόπολις ‘métropole’ (< μήτηρ ‘mère’; πόλις ‘cité’)
	Exocentriques	- [AN] _{A/N} : μακρό-χειρ ‘qui a de longs bras’ (< μακρός ‘long’; χειρ ‘bras’)
Subordonnés	Endocentriques	- [NN] _N : ἔντερο-κήλη ‘hernie de l’intestin’ (< ἔντερα ‘entrailles’; κήλη ‘hernie’)
	Exocentriques	- [NV] _N : λογογράφος ‘celui qui écrit les discours’ (< λόγος ‘discours’; -γράφος (< γράφω ‘écrire’)) - [NV] _A : φλεβο-τόμος ‘qui entaille les veines’ (< φλέψ ‘veine’; -τομός < τέμω ‘couper’) - [NN] _A : μηνοειδής ‘qui a une forme de croissant’ (< μήνη ‘lune’; εἶδος ‘forme’) - [VN] _{A/N} : ἀρχέ-πολις ‘qui gouverne la cité’ (< ἀρχω ‘commander’; πόλις ‘cité’)
Coordonnés	Endocentriques	- [AA] _A : γλυκύ-πικρός ‘doux-amer’ (< γλυκύς ‘sucré’; πικρός ‘amer’) - [NN] _N : ἰατρόμαντις ‘physicien et devin’ (< ἰατρός ‘physicien’; μάντις ‘devin’)
	Exocentriques	- [NN] _N : ἀνδρόγυνος ‘hermaphrodite, personne efféminée’ (< ἀνήρ ‘homme’; γυνή ‘femme’)

Tableau 1.1 – Types de composés en grec ancien

Les composés attributifs, dans lesquels l’élément déterminant exprime un attribut de l’élément déterminé, sont en grec soit endocentriques soit exocentriques. Les premiers se trouvent dans les combinaisons [AN]_N ou [NN]_N. Parmi les composés attributifs exocentriques, nous trouvons principalement les composés *bahuvrīhi*, très courants en grec ancien, qui permettent de qualifier une personne selon une caractéristique, souvent physique, qu’elle possède. On trouve également des composés subordonnés endocentriques et exocentriques. Les composés endocentriques sont en général des [NN]_N avec l’élément déterminé dans la position droite. Les composés exocentriques présentent eux des constructions plus variées. Parmi celles-ci se trouvent les

composés synthétiques impliquant un deuxième élément verbal suffixé ainsi que des composés $[NN]_A$. Tous ceux-là présentent l'élément qui est déterminé à droite. Les composés $[VN]_{A/N}$ sont également classés parmi les subordinés exocentriques mais ils présentent eux l'élément syntaxiquement recteur à gauche. Enfin le grec ancien possède des composés coordonnés, mais ils sont très rares.

Après avoir donné un aperçu rapide de la composition en grec ancien, je vais maintenant observer la composition en latin.

1.2.2 La composition en latin

Pour traiter du latin, j'ai pu disposer d'une littérature plus conséquente. Je m'appuierai sur les travaux de Fruyt (2002, 2011), Oniga (1992), Brucalè (2012) et Biville (1985, 2002). Tous s'accordent à dire que la composition en latin est un processus de formation de mots bien moins courant qu'en grec. En fait, c'est un processus qui prend le plus souvent exemple sur le grec, en calquant des processus de formation, en « traduisant » des composés grecs, en intégrant des composés grecs avec peu ou pas d'adaptation dans la langue latine. C'est le principe des calques que j'ai abordé dans la partie 1.1.1 et que j'ai développé en 1.1.2 concernant le latin.

En latin, la composition est donc un processus de construction de mots peu courant, en comparaison avec des langues comme le sanskrit, le grec ou l'arménien. Ce processus, si admiré en grec, est moqué en latin, comme le montre cette citation de Quintilian :

« cum κυρτάχρα mirati simus, *incuruiceruicum* uix a risu defendimus¹⁶ »

'bien que nous admirions κυρτάχρα, nous pouvons difficilement nous arrêter de rire de *incuruiceruicum*'

(Quintilian, *Institution Oratoria*, 1.5.70, cité par Fruyt (2011))

Fruyt (2002) avance que le latin préfère éviter la composition et lui préfère la suffixation ou l'agglutination. Elle entend par agglutination la juxtaposition de deux lexèmes fléchis alors que la composition induit pour elle l'emploi de radicaux sans aucune marque flexionnelle. Ainsi, *agrī-cultura*, où *agrī* est la forme du génitif singulier du lexème *ager*, serait une agglutination alors

16. κυρτάχρα, 'qui a le cou voûté' est un composé *bahuvrīhi*, construit à partir de *κυρτός* 'courbé' et *ἀχρήν*, 'cou'. *incurui-ceruic-us* (*incuruus* 'courbé'; *ceruix* 'cou') est son calque morphologique latin, forgé par Pacuvius au sujet des dauphins.

que *agrīcola* serait un composé, le <i>¹⁷, lorsqu'il est bref, n'étant pas une marque flexionnelle. Alors que la composition implique une relation sous-jacente entre deux lexèmes, comme indiqué en (24), l'agglutination n'est que la juxtaposition de plusieurs membres d'un syntagme. Pour traduire des composés grecs techniques, le latin utilisera des lexèmes suffixés (41-a) ou des syntagmes (Fruyt (2002) parle de « lexies complexes »)(41-b) :

- (41) a. φιλογύνης (< φιλέω 'aimer'; γυνή 'femme') > *mulierōsitās* (< *mulierosus* 'dévoué aux femmes') 'penchant pour les femmes'
 b. φυσιολογία (< φύσις 'nature'; λογος 'étude') > *nātūrae ratio* 'étude de la nature'

Pour Brucale (2012, p.94), la grande majorité des composés en latin sont des hapax legomena : ils sont sémantiquement très contextualisés, dévolus essentiellement à des vocabulaires spécifiques (religion, justice) ou à la poésie archaïque et sont en général des calques de modèles grecs. Plusieurs raisons expliquent la faible disponibilité de ce processus de formation. C'est tout d'abord une question de taille : le mot prototypique latin est plus court que le mot prototypique grec ou sanskrit. En effet, les mots lexicaux latins font en général deux à trois syllabes au nominatif singulier ou à la première personne du présent de l'indicatif. Un composé, qui implique la présence de deux constituants lexicaux, serait donc plus long que les autres mots latins. Cette petite taille des mots latins est en grande partie due à des développements phonétiques qui ont affecté les voyelles, ce qui a engendré des réductions de syllabes : les frontières existantes ont pu disparaître, empêchant ainsi l'analyse de ces mots comme des composés et donc de servir de modèles pour la création de nouveaux composés. Cependant, certains éléments favorisent la création de composés. Fruyt (2002) cite trois sortes de formations latines qui demandent l'utilisation de composés : tout d'abord les composés poétiques, qui ont pour la plupart la forme d'un choriambre¹⁸. La frontière du composé se situe alors au milieu de celui-ci, entre les deux syllabes brèves, la première syllabe brève s'achevant par un -i- bref, marqueur de composition. Ensuite, on trouve en latin des composés qui ne sont dus qu'à l'influence du grec, en particulier les composés *bahuvrīhi*, que j'ai évoqués dans la partie 1.2.1. Enfin,

17. Dans cette thèse, pour désigner un segment se trouvant à la jonction de deux constituants, segment communément appelé « voyelle de liaison », j'ai choisi d'utiliser des chevrons simples. En effet, le <i> et le <o> sont communs à différentes langues mais ne sont pas prononcés de la même manière. Des caractères phonologiques ne me paraissent donc pas forcément pertinents.

18. Le choriambre est en poésie un pied de quatre syllabes constitué d'un chorée (ou trochée), qui comprend une syllabe longue suivie d'une syllabe brève, et d'un iambe, formé d'une syllabe brève et d'une syllabe longue. Le choriambre affiche donc deux syllabes brèves entre deux syllabes longues (˘˘˘˘).

beaucoup de composés latins sont des créations fantaisistes, en particulier inventées par Plaute : par exemple, dans *Persa* (vers 705), celui-ci crée un composé à partir d'une proposition relative et d'une proposition principale : *Quod-semel-arrip-ides-Numquam-erip-ides* 'ce qu'il a pris une fois ne lui sera jamais rendu'.

Si les composés ne sont pas très fréquents en latin, certains types montrent toutefois une certaine disponibilité. Comme je l'ai fait pour le grec dans la section précédente, je vais dans un premier temps observer les différentes constructions présentes en latin et pour cela, je m'appuie sur le travail de Brucale (2012). Elle met tout d'abord en évidence ce qu'elle appelle des « radical compounds » (RC), c'est-à-dire des composés qui ont comme structure « X+Y+SUFF, where X is normally a shortened noun or adjective, Y is typically a verbal root, and the suffixal ending can be overtly or covertly expressed » (Brucale, 2012, p.104). Il s'agit en fait du pendant latin des composés grecs¹⁹ que j'ai exemplifiés en (28) ou (31). Cependant, ici le suffixe permettant d'attribuer la catégorie au composé peut disparaître :

- (42) [NV]_N :
- a. *agricola*, 'agriculteur' (< *ager* 'champ' ; *colo* 'cultiver')
 - b. *aurifex*, 'orfèvre' (< *aurum* 'or' ; *facio* 'faire')
 - c. *causidicus*, 'avocat' (< *causa* 'cause' ; *dico* 'dire')

L'ensemble des éléments pouvant constituer le deuxième élément est restreint : les plus communs sont *-fex* ou *-ficus*, *-ficus* (< *facio* 'faire') ; *-ceps*, *-cipis* (< *capio* 'prendre') ; *-spex*, *-scipis* (< *specio* 'regarder') ; *-dex*, *-dicis* (< *dicere* 'dire') ; *-ex*, *-icis* (< *agio* 'agir') ou *-cida*, *-ae* (< *caedo* 'tuer'). Je ne pense pas qu'il faille statuer sur l'existence d'un suffixe zéro, comme le fait Brucale (2012, p.103) ; si, suite à une évolution phonologique, le nominatif de certaines constructions n'affiche pas de désinence flexionnelle à proprement parler, cela n'empêchera aucunement la déclinaison du nom, et le génitif de *aurifex* est *aurificis*.

Il en va de même pour les autres composés nominaux du latin ; certains sont formés par parasynthèse, d'autres par simple composition :

- (43) a. [NN]_N : *domuitio* 'trajet vers la maison' < *domus* 'maison' ; *itio* 'action d'aller'
- b. [NN+Suff]_N : *armilustrum* 'festival de la consécration des armes' < *arma* 'armes' ; *lustrum* 'sacrifice purificateur'

19. Ce type de composé existait également en sanskrit : *madhu-pa-s* 'lit. qui boit du miel, abeille'.

- (44) a. [AN]_N : *perenniservus* ‘esclave perpétuel’ < *perennis* ‘durable’ ;
servus ‘esclave’
 b. [AN+Suff]_N : *plenilunium* ‘pleine lune’ < *plenus* ‘plein’ ; *luna*
 ‘lune’

Du côté des composés adjectivaux, nous retrouvons les mêmes composés [NV] que pour les nominaux. Ici, les verbes les plus fréquents entrant dans ces constructions sont *facio* ‘faire’ (> *-ficus*)²⁰, *fero* ‘porter’ (> *-fer*) ou *gero* ‘porter’ (> *-ger*).

- (45) [NV]_A :
 a. *armiger* ‘armé’ < *arma* ‘armes’ ; *gero* ‘porter’
 b. *honorificus* ‘honorable’ < *honor* ‘honneur’ ; *facio* ‘faire’

Le latin possède également des composés adjectivaux du type *bahuvrīhi* :

- (46) [AN]_A : *magnanimus* ‘qui a une grande âme’ < *magnus* ‘grand’ ;
animus ‘âme’

D’autres constructions adjectivales se rencontrent également, combinant soit deux noms (47), soit deux adjectifs (48) :

- (47) [NN]_A : *nocticolor* ‘qui a la couleur de la nuit’ < *nox* ‘nuit’ ; *color*
 ‘couleur’
 (48) [AA]_A : *dulcacidus* ‘aigre-doux’ < *dulcis* ‘doux’ ; *acidus* ‘aigre’

Les composés verbaux sont très rares en latin ; seuls trois verbes se rencontrent comme seconds éléments de constructions [XV]_V : d’une part *-fero(r)* et *-gero(r)*, qui prennent comme premier élément toujours un nom, qui constitue leur argument interne (49), d’autre part *facio* ‘faire’ qui, selon Brucale (2012) rentre dans des constructions nombreuses et productives. *Facio* peut être présent sous la forme *-fico(r)*, il se combinera alors avec des noms ou des adjectifs (50), ou la forme *-facio* (parfois *-fio*) qui peut se combiner avec un adverbe, un adjectif ou un verbe et il aura dans ce dernier cas une valeur causative (51).

- (49) [NV]_V :
 a. *pacifero* ‘garder la paix’ < *pax* ‘paix’ ; *fero* ‘porter’
 b. *verbifero* ‘parler, contester’ < *verbum* ‘mot’ ; *gero* ‘porter’

20. *facio* donne deux thèmes pour les composés nominaux, *-fex* et *-ficus*, alors que les composés adjectivaux ne font intervenir que le second.

- (50) [X *fico(r)*]_V :
- a. [N *fico(r)*]_V : *aedifico* ‘ériger un bâtiment’ < *aedes* ‘maison’ ; *facio* ‘faire’
 - b. [A *fico(r)*]_V : *amplifico* ‘élargir’ < *amplus* ‘large’ ; *facio* ‘faire’
- (51) [X *facio*]_V :
- a. [Adv *facio*]_V : *satisfacio* ‘satisfaire’ < *satis* ‘suffisamment’ ; *facio* ‘faire’
 - b. [A *facio*]_V : *multifacio* ‘estimer’ < *multus* ‘beaucoup’ ; *facio* ‘faire’
 - c. [V *facio*]_V : *calefacio* ‘rendre chaud’ < *caleo* ‘être chaud’ ; *facio* ‘faire’

Concernant le segment à la jonction des deux constituants du composé, il s’agit le plus souvent d’un *-i-*, en alternance avec un *-u-* phonologiquement conditionné (quand le second constituant commence par une bilabiale ou une labiodentale). Brucalè (2012), à l’instar de Ralli concernant le grec moderne, analyse cette voyelle comme un marqueur de composition. Cette voyelle est en tout cas un élément structurel, qui apparaît toujours devant une consonne et n’est omise que devant un élément débutant par une voyelle. Elle semble donc davantage systématique qu’en grec où *o*, *a*, plus rarement *η*, étaient en concurrence.

Voyons maintenant, comme cela a été fait pour le grec dans la section précédente, les types de composés dans lesquels entrent les composés du latin (Tableau 1.2). Comme pour le grec, tous les types de composés définis par Bisetto et Scalise (2009) sont présents, même si certains sont anecdotiques, comme les coordonnés exocentriques, dont le seul exemple que j’ai pu trouver n’est pas moins qu’une traduction d’un composé grec. D’une manière générale, comme il a déjà été dit, beaucoup de composés latins sont inspirés du grec, je pense par exemple aux composés attributifs exocentriques [AN]_{A/N} qui ne sont autres que des composés *bahuvrīhi*.

Les composés latins sont comparés aux composés grecs dans le Tableau 1.3 au début de la section suivante dans lequel tous sont mis en perspective avec les composés français dits néoclassiques qui sont le sujet de cette thèse.

Attributifs	Endocentriques	[AN] _N : <i>perenniservus</i> ‘esclave perpétuel’ (< <i>perennis</i> ‘durable’; <i>servus</i> ‘esclave’) [NN] _N : <i>verbivelitatio</i> ‘dispute verbale’ (< <i>verbum</i> ‘mot’; <i>velitatio</i> ‘dispute’)
	Exocentriques	[AN] _{A/N} : <i>magnanimus</i> ‘qui a une grande âme’ (< <i>magnus</i> ‘grand’; <i>animus</i> ‘âme’) [NN] _N : <i>veriverbium</i> ‘le fait de dire la vérité’ (< <i>verum</i> ‘vérité’; <i>verbum</i> ‘mot’)
Subordonnés	Endocentriques	[NN] _N : <i>domuitio</i> ‘trajet vers la maison’ (< <i>domus</i> ‘maison’; <i>itio</i> ‘action d’aller’)
	Exocentriques	[NV] _N : <i>agricola</i> , ‘agriculteur’ (< <i>ager</i> ‘champ’ + <i>colo</i> ‘cultiver’) [NV] _A : <i>armiger</i> ‘armé’ (< <i>arma</i> ‘armes’; <i>gero</i> ‘porter’) [NN] _N : <i>armilustrum</i> ‘festival de la consécration des armes’ (< <i>arma</i> ‘armes’; <i>lustrum</i> ‘sacrifice purificateur’)
Coordonnés	Endocentriques	[AA] _A : <i>dulcacidus</i> ‘aigre-doux’ (< <i>dulcis</i> ‘doux’; <i>acidus</i> ‘aigre’)

Tableau 1.2 – Types de composés en latin

1.2.3 La pénétration d'éléments classiques en français moderne

Maintenant que nous avons eu un aperçu de la manière dont le grec et le latin construisaient des composés, voyons comment et pourquoi le français a puisé dans ce fonds classique pour former ses propres composés dans le domaine des sciences tout d'abord, puis dans la langue générale ensuite. Le Tableau 1.3 met en correspondance ce que nous venons de voir concernant les composés grecs et latins, et les types de composés néoclassiques productifs en français selon Villoing (2012). Si chaque type de composé grec ou latin n'a pas forcément donné lieu à des composés néoclassiques en français, nous pouvons tout de même remarquer que tous les types de composés néoclassiques du français sont des calques de composés grecs et/ou latins.

Mais comment et pourquoi ce passage des langues classiques au français²¹ s'est-il opéré? Cette partie fait écho à la partie 1.1.3 qui exposait d'une manière générale les spécificités de la langue spécialisée.

Le passage de mots du latin au français a pu se faire selon différentes voies, au travers des siècles. A côté de transmissions populaires, qui font qu'un mot, continuellement employé, sera transformé au fil du temps, des transmissions savantes sont opérées entre le XIII^{ème} et le XIV^{ème} siècle : des savants empruntent directement au fonds latin des mots, cet emprunt les préservant de toute modification orthographique ou phonologique. Ainsi, le nom *prehensio* 'action de prendre' a donné *prison* par la voie populaire et *préhension* par la voie savante. Dès le XIII^{ème} siècle, un essor culturel se met en place et le nombre de savants maîtrisant le latin classique s'accroît. Le vocabulaire français s'enrichit donc de formations savantes. Ce phénomène prendra de l'ampleur entre le XIV^{ème} et le XVI^{ème} siècles, accompagné d'une redécouverte du grec qui participera également à l'enrichissement du vocabulaire (Bouffartigue et Delrieu, 1996b). Historiquement, il était donc d'usage d'utiliser le latin pour la rédaction de terminologies scientifiques. Ainsi, tous les dictionnaires de médecine du XVI^{ème} et XVII^{ème} siècle étaient rédigés en latin, à l'exception de celui d'Ambroise Paré, au XVI^{ème} siècle, qui était écrit en français avec toutefois un grand nombre de mots grecs (Hatzfeld *et al.*, 1890). Jusqu'au XVIII^{ème} siècle, le français n'aura de rapport avec le grec que par le latin. On peut dater le début de la pénétration puis de la généralisation des termes technico-scientifiques empruntés ou calqués sur le grec à la fin XVIII^{ème} siècle. En effet, Brunot (1927a) dit qu'à la période

21. Ici, je m'attacherai à traiter uniquement le français, cette langue étant au cœur de cette thèse; il est toutefois à noter que les éléments classiques dont il est question ont également pénétré de nombreuses langues européennes, sensiblement à la même époque et pour les mêmes raisons que le français.

	Grec Ancien	Latin	Français
Attributifs	Endocentriques	[AN] _N : <i>perennis</i> ‘esclave perpétuel’ (< <i>perennis</i> ‘durable’; <i>servus</i> ‘esclave’) ; [NN] _N : <i>verbielatio</i> ‘dispute verbale’ (< <i>verbum</i> ‘mot’; <i>velitatio</i> ‘dispute’)	[AN] _N : <i>mégalthé</i>
	Exocentriques	[AN] _{A/N} : <i>magnanimus</i> ‘qui a une grande âme’ (< <i>magnus</i> ‘grand’; <i>animus</i> ‘âme’) ; [NN] _N : <i>ververbium</i> ‘le fait de dire la vérité’ (< <i>verum</i> ‘vérité’; <i>verbum</i> ‘mot’)	[AN] _A : <i>quadrupède</i>
Subordonnés	Endocentriques	[NN] _N : <i>domutio</i> ‘trajet vers la maison’ (< <i>domus</i> ‘maison’; <i>itio</i> ‘action d’aller’)	[NN] _N : <i>hippodrome</i>
	Exocentriques	[NV] _{A/N} : <i>agricola</i> ‘agriculteur’ (< <i>ager</i> ‘champ’ + <i>colo</i> ‘cultiver’) ; [NN] _N : <i>armilustrum</i> ‘festival de la consécration des armes’ (< <i>arma</i> ‘armes’; <i>lustrum</i> ‘sacrifice purificateur’)	[NA] _A : <i>photosensible</i>
Coordonnées	Endocentriques	[AA] _A : <i>dulcacidus</i> ‘aigre-doux’ (< <i>dulcis</i> ‘doux’; <i>acidus</i> ‘aigre’)	[AA] _A : <i>franco-allemand</i>
	Exocentriques	[NN] _N : <i>androgyné</i>	[NN] _N : <i>rhinopharynx</i>
			[NN] _N : <i>androgyné</i>

Tableau 1.3 – Évolution des types de composés, du grec ancien au français (pour le français, les exemples sont de Villouing (2012))

de la Révolution, la nomenclature industrielle par exemple usait encore de périphrases à rallonge pour attribuer des noms aux machines qui étaient inventées (par exemple *machine pour enrayer et désenrayer à volonté les roues d'une voiture*). La pénétration du grec commence, même si elle est encore timide : par exemple, le mot *panorama* apparaît en 1789 d'abord en anglais puis en français. Toujours selon Brunot, des hésitations demeurent ; quand un haut parleur est créé, son inventeur se pose la question de sa dénomination : *tour parlante* ou *télélogue* ? Sous l'Empire, la nomenclature industrielle utilise encore de nombreuses périphrases (par exemple, un *appareil de distillation propre à éclairer les ateliers et les appartements par le gaz hydrogène de la houille*), mais un changement est en train de s'opérer : on commence à emprunter à d'autres langues (ex : *machine aéro-fluviale*), à créer des composés en français (ex : *bateau-brouette*). Surtout, « on s'aventurait de plus en plus hors du domaine national sans craindre plus « l'opacification » des exposés que celle du verre » (Brunot, 1927a, p.1220) : on a de plus en plus recours au latin (le préfixe latin *in-* commence à surpasser la forme française *non*), mais c'est surtout au grec qu'est donnée la primauté. C'est à cette époque que commencent à être fréquemment utilisés en français des unités comme *-mètre* (au sens d'instrument de mesure généralisé : *alcalimètre*), *-scope* (d'après *télescope* on trouve *phantasmascope*), *-graphe* (à la suite de *télégraphe* : *anémographe*, *anthropographe*, etc.), *-logie*, *-technie*, *-ide*, etc. On trouve aussi des composés « véritables » tels que *hypothalattique* 'art de naviguer sous l'eau' (Brunot, 1927a, p.1224). Les échanges entre savants de pays différents ne se sont jamais interrompus malgré les conflits politiques et, souvent, les composés scientifiques que l'on trouve en français ont d'abord été inventés dans une autre langue (la première attestation de *toxicologie* se trouve en allemand en juin 1803, celle de *astromètre* en anglais, et celle de *bathomètre* en hollandais). Du côté des sciences, le développement de la botanique avec Linné et Jussieu ou de la chimie avec Lavoisier et Guyton de Morveau ne se fait qu'en parallèle de la mise en place d'une nomenclature. Si Linné utilise le latin pour établir sa nomenclature en 1735, Lavoisier, Guyton de Morveau, Bertholet et de Fourcroy publient en 1787 une *Méthode de nomenclature chimique* qui sera lue le 18 avril 1787 devant l'Académie Royale des Sciences et dans laquelle les savants argumentent en faveur d'une dénomination précise des composés chimiques, et utilisent le grec pour former nombre de termes de leur nomenclature. C'est dans cet ouvrage que le mot *oxygène*, formé sur l'adjectif grec *ὄξύς* 'acide' et le verbe grec *γείνομαι* 'engendrer' apparaît pour la première fois. A cette période, la généralisation du recours aux formations gréco-latines pour la constitution d'un vocabulaire technico-scientifique s'intensifie grâce aux mesures concrètes prises pour le

développement et la protection des inventions ainsi que grâce au rôle joué par l'Académie des Sciences (Guilbert, 1975, p.225). Le développement des sciences et de l'industrie de la première moitié XIX^{ème} siècle offrira une large diffusion et propagation de ces formations gréco-latines.

Mais pourquoi avoir recours de façon prédominante au grec ? Quelles sont les causes du triomphe de ce « monstre grec qui attire en surprenant » (Brunot, 1927a) ? Bouffartigue et Delrieu (1996a) donnent deux raisons principales à cela : tout d'abord, les Latins, qui n'étaient pas de grands scientifiques, quand ils ont eu connaissance du monde grec, ont tout importé : connaissances, modèles de connaissances, terminologies. En effet, la majorité des découvertes scientifiques ont été faites du côté de la Grèce antique et non de Rome. Donc les bases du discours scientifique ont été forgées sur le grec dès l'Antiquité. De plus, comme nous l'avons vu, le latin se prête beaucoup moins à la composition que le grec, ce qui a créé une difficulté dans la création de mots. Revenons maintenant aux caractéristiques requises par la langue spécialisée énumérées dans la partie 1.1.3 : la langue scientifique se doit d'être concise, doit favoriser l'intercompréhension internationale et doit permettre aux scientifiques de communiquer dans un jargon communautaire qui leur est propre. La concision est apportée par la langue grecque, qui permet la formation de composés exprimant en un seul mot plusieurs idées là où le français utilise des périphrases et des juxtapositions (Guilbert, 1975, p.226-227). Le grec a des qualités linguistiques reconnues de tous les savants : « des espèces innombrables d'insectes, de plantes de minéraux, de fossiles, sont découvertes et classées ; il faut les dénommer : où trouver, sinon dans le grec, assez de termes clairs, précis, bien faits ? La richesse de cette langue, ses remarquables qualités de précision et de netteté, son égale puissance de composition et de dérivation, la désignaient naturellement aux savants, qui y puisent à pleines mains. » soulignent Hatzfeld *et al.* (1890, p.103), qui précisent également que le roman était insuffisant pour évoquer des idées philosophiques. Le grec est également une langue qui est accessible à tous les savants européens, quelle que soit leur langue. Puiser dans ce réservoir classique permet donc une meilleure compréhension entre les différents scientifiques. Brunot (1927a, p.1221) rapporte une note au sujet du mot *gleucomètre* de Chevallier, extraite du *Journal d'Economie rurale et domestique* : « Quand un mot nouveau devient nécessaire, il faut l'admettre. Il est convenu dans l'Europe savante que, pour être entendu dans toutes les langues, un mot doit tirer son origine du grec qui, pour toutes les nations civilisées, est la langue génératrice des mots, la langue mère ; souscrivons à cette loi. Ce mot sera plus aisé à admettre que nombre de mots insignifiants que la mode introduit journallement ». La notion de « langue mère » est

mise en avant pour justifier et précipiter l'utilisation du grec dans la langue scientifique. Enfin et surtout, ce sont des raisons idéologiques qui ont poussé à l'emploi du grec pour forger une langue technico-scientifique : les formations gréco-latines participaient à l'élaboration d'un jargon dans lequel se retrouvaient les savants, ce que montre cette délicieuse citation de Brunot (1927a, p.1229) :

« Nous qui sommes à distance et qui savons l'impression que font, dans des cérémonies religieuses, les mots incompris psalmodiés dans des langues inconnues, les charmes qu'exercent les secrets des sorcières et les baracadabra des kabbalistes, les suggestions qu'ont données à des générations les recipés des vieux médecins latiniseurs, le prestige que gardent encore les formules retentissantes, même lorsqu'elles n'ont aucun sens, des politiciens, bref l'action éternelle du mot mystérieux, nous imaginerions volontiers qu'un homme de génie a conçu l'idée d'employer, en les enchâssant dans la banalité de la langue vulgaire, des vocables faits pour séduire lecteurs et acheteurs, et qu'il a choisi le grec non seulement parce qu'il fournissait d'inépuisables ressources à la formation, mais parce que, moins familier que le latin, il avait l'avantage d'être inconnu et incompréhensible à une génération peu versée dans l'étude de la Grèce antique et de sa langue. »

Les langues classiques, et en particulier le grec, répondent donc parfaitement aux exigences qu'a une langue spécialisée comme la langue technico-scientifique.

Mais comment ce passage du grec au français s'est-il déroulé ? Bouffartigue et Delrieu (1996a) exposent six formes de transmission du grec au français. Les trois premières ont pris comme étape intermédiaire le latin : ce sont des transmissions anciennes, qui ont été opérées dès l'antiquité.

- (52) a. $\pi\upsilon\zeta\acute{\iota}\varsigma > buxis > boîte$
 b. $\acute{\epsilon}\kappa\kappa\lambda\eta\sigma\acute{\iota}\alpha > ecclesia > \acute{e}glise$
 c. $\phi\acute{\iota}\lambda\acute{o}\sigma\omicron\phi\omicron\varsigma > philosophus > philosophe$

Il y a tout d'abord une transmission ancienne populaire (52-a), dans laquelle des mots grecs ont été empruntés par le latin, les deux langues étant alors des langues vivantes. Le passage du grec au latin a provoqué une première évolution phonétique et celui du latin aux langues romanes une deuxième : il y a donc deux « accidents ». Dans l'exemple (52-a), la consonne initiale grecque /p/ s'est transformée par voisement en /b/ en latin ; le mot latin, par une série de modifications phonologiques, s'est ensuite altéré en *boîte*.

Par ailleurs, des transmissions anciennes vulgarisées ont pu être préservées du premier accident : des lettrés latins ont eu la volonté d'introduire des mots grecs en limitant d'éventuelles modifications phonétiques ; cependant, des modifications ont eu lieu lors du passage de ces mots du latin aux langues romanes. C'est le cas dans l'exemple (52-b), dans lequel le mot latin *ecclesia* a été emprunté au mot grec *ἐκκλησία* sans modification majeure, alors que, dans le passage du latin au français, le mot a connu quelques évolutions phonétiques, par exemple le voisement de /k/.

Le troisième type de transmission qui a eu lieu par l'intermédiaire du latin est nommé par Bouffartigue et Delrieu transmission ancienne savante (52-c). Le passage du grec au latin a eu lieu de la même manière que la précédente, grâce à des lettrés. Cependant, ces mots ne sont pas, cette fois-ci, passés dans la langue courante, et leur présence en français n'est due qu'à l'effort de lettrés du Moyen-Âge et de la Renaissance qui ont importé en français un grand nombre de mots qu'ils ont trouvés dans les textes latins. Cela explique qu'il n'y ait eu aucune modification majeure entre le grec et le français.

La disparition du latin comme langue parlée n'a pas arrêté le passage de mots grecs en français. À côté des transmissions anciennes que nous venons de voir, Bouffartigue et Delrieu (1996a) relèvent trois types de transmissions récentes.

- (53) a. (i) μισογύνης > *misogyne*
 (ii) τέλε-σκοπέω 'qui voit loin' > *télescope*
 b. cosmonaute, dermatologue
 c. polyvalent, bureaucratie

Le premier type est constitué de réemplois récents : pour exprimer des notions nouvelles, les philosophes et savants européens sont allés chercher dans le grec des mots qui avaient le sens requis (53-a-i) ; il s'agit à proprement parler d'emprunts. Quand l'objet à dénommer était nouveau, ils ont utilisé des mots qui avaient un sens voisin (53-a-ii). Il s'agit là d'une adaptation, dans les termes de Cottez (1980). Puis, avec l'élargissement des nomenclatures, de plus en plus de mots ont dû être trouvés et le réservoir grec est devenu insuffisant. Les savants ont donc utilisé des procédés et des constituants grecs pour former en français (et dans les langues européennes) des mots imitant les mots grecs. C'est ce que Cottez (1980) a appelé des « productions » (53-b). Mais ces imitations ne se font souvent que par des hellénistes avisés, un composé de formation récente supposant, pour Hatzfeld *et al.* (1890), la création d'un composé grec qui lui sert de modèle. Enfin, Bouffartigue et Delrieu (1996a) dégagent un troisième type de transmission récente qu'ils nomment dérivation : c'est un phénomène qui apparaît à l'intérieur de la

langue moderne et qui permet des formations hybrides (53-c). Ce sont ces formations hybrides qui seront fortement critiquées. Par exemple, Darmesteter (1874, p.254) s'inquiète du grand nombre de créations et de l'invasion du vocabulaire grec dans la langue française : « C'est à se croire revenu aux beaux jours de la Pléiade ». Il trouve les mêmes raisons que celles que j'ai citées précédemment pour expliquer cette invasion, à savoir la possibilité pour des savants de langues différentes de se comprendre, la nécessité d'employer des termes d'une grande précision sémantique et la pauvreté de la composition française, mais ne trouve pas ces raisons valables pour légitimer l'emploi d'un si grand nombre de formations d'origine grecque. Mais ce qui le révolte par-dessus tout est la formation de termes hybrides : ceux créés conjointement avec des éléments latins et d'autres grecs dont l'usage s'étend (*-mètre*, *-game*, *-logie*, *-algie*, *etc.*) mais surtout ceux créés avec des constituants latins ou français qui affichent la voyelle *o* à leur jonction : pour lui, « ce sont autant de formes barbares qui menacent de pénétrer la langue commune » (Darmesteter, 1874, p.256). Hatzfeld *et al.* (1890, p.103) emploient également le terme de « barbare » pour évoquer ces formations hybrides :

« L'extension, le progrès des sciences, la vulgarisation, pour employer le terme consacré, l'action incessante de la presse, le développement de l'industrie, répandent dans l'usage général certains termes qui n'auraient pas dû sortir du laboratoire du chimiste ni du cabinet des philosophes. De plus, une foule de suffixes, de particules grecques, étant devenus usuels, chacun se croit autorisé à combiner ces éléments à sa guise, en bravant les lois du grec et celles du français. De là des formations hybrides et barbares qu'on rencontre dans plus d'un de nos composés modernes. De là aussi des composés mi-grecs, mi-français, qui montrent à quel point beaucoup d'éléments composants grecs sont devenus organiques et ont perdu leur cachet d'origine. »

Les temps modernes, en plus de créer des termes hybrides, n'ont pas non plus respecté les mécanismes propres au grec. Le Feuvre (2012) montre ainsi que, si *splanchnomégalie* ne correspond à aucun type de composé en grec ancien, il est conforme à la structure productive dans la terminologie médicale moderne. *Mégalosplanchnie*, au contraire, correspond à un type de composé du grec, le type possessif que j'ai évoqué en 1.2.1. Cependant, une inversion des deux constituants du composé a permis une conformité aux structures productives à l'époque moderne, avec un premier constituant dénotant un nom d'organe et le second un nom d'affection, comme *céphalgie* ou *ostéomalacie* : « la justification synchronique l'emporte sur la justification diachronique » (Le Feuvre, 2012, p.190).

De plus, quoiqu'en disent Darmesteter ou Hatzfeld, de nombreux mots formés grâce à des constituants grecs ou latins, hybrides ou non, ont passé la frontière de la langue générale et sont aujourd'hui employés par tout locuteur du français. Certains sont plus facilement candidats à ce passage : Iacobini (2004, p.83) souligne que les mots construits avec un élément classique (élément formant) à l'initiale et un mot de la langue moderne en deuxième position offrent la structure qui facilite le plus le passage de ces mots dans l'usage commun ; pour lui, c'est une structure plus facilement segmentable, les éléments sont aisément identifiables et donc interprétables sémantiquement. Une analyse statistique appuie son hypothèse : cette analyse porte sur des néologismes enregistrés dans des dictionnaires non spécialisés du XIX^{ème} siècle et montre que la diffusion des éléments formants dans la langue italienne va de pair avec leur utilisation dans des structures [élément formant + mot].

Si un locuteur actuel peut interpréter, imaginer, créer, analyser des termes dont la structure implique un ou plusieurs constituants classiques, doit-il nécessairement pour cela connaître le grec et le latin ? Les locuteurs d'aujourd'hui ont d'autres mécanismes qui leur permettent d'utiliser ces formations hybrides sans avoir de connaissance étymologique particulière. Cottez (1980, p. VIII) dit d'ailleurs que « la question de savoir si aujourd'hui n'importe quel locuteur est capable d'identifier les [formants] n'est pas linguistiquement pertinente ». C'est d'autant plus vrai que les formants peuvent changer de sens ; c'est le cas pour la plupart d'entre eux, notamment ceux fréquemment utilisés. Dans un premier temps, ce changement s'opère en discours, dans la *parole* au sens saussurien du terme, puis la nouvelle interprétation s'installe dans la langue, suite à de nombreuses répétitions chez de nombreux individus. Enfin, l'interprétation est institutionnalisée dans le lexique. Nous pouvons prendre pour exemple le formant *-phobie* dont je ferai une présentation détaillée en 3.2.1.3. Au sens originel de 'peur' s'ajoute un sens d'hostilité'. Deux sens s'opposent donc : un sens fonctionnel et un sens parallèle (ces termes sont empruntés à Fruyt (1996)). Le premier est le sens que le mot a dans une phrase, le deuxième la vision que le locuteur a du mot. Ainsi, un mot comme *xénophobie* a comme sens fonctionnel 'hostilité envers les étrangers' mais d'aucuns pourront lui attribuer le sens parallèle 'peur des étrangers' pour légitimer ce sentiment qui les habite. La polysémie se situe donc dans le sens parallèle. Il faudra par conséquent s'attacher à distinguer, dans ce travail, la conscience linguistique du locuteur, que Fruyt (1996) nomme « sentiment du sujet parlant », qui est l'interprétation que le sujet projette sur sa langue, un sens « naïf » et « spontané » et le savoir métalinguistique du savant. Le premier est un sentiment synchronique, le

deuxième un savoir diachronique. Même si je me suis attachée jusqu'à maintenant à retracer diachroniquement l'utilisation des constituants classiques dans la langue moderne, ce qui va m'intéresser à présent et pour tout ce travail de thèse est ce qui relève du synchronique : comment les locuteurs d'une langue moderne comme le français sont capables d'interpréter et même de créer des mots en utilisant des formants classiques ? Ces formations sont-elles toujours dépendantes du système de la composition grecque et latine ou sont-elles au contraire totalement intégrées au système de la composition française ? Avant de proposer, grâce à des analyses sur certains constituants (chapitres 3 et 4) une réponse à ces questions (chapitre 5), je vais au préalable détailler les analyses et les conclusions qui ont pu être tirées dans la littérature sur ces questions.

1.3 Les composés néoclassiques dans la littérature

Les composés néoclassiques ont connu un intérêt croissant qui s'est largement manifesté au cours des deux dernières décennies. Pourtant, Adams (1973, p.133) parlait déjà de « their proeminent role in word-formation today » ; elle ne leur consacre toutefois qu'un chapitre de six pages. Le traitement qui est fait de ces formations n'est cependant pas univoque. Les auteurs divergent sur de nombreux points : les dénominations de ces formations et des éléments qui entrent en jeu à l'intérieur, leur statut, leurs propriétés, etc. Cette partie va donner un aperçu sinon exhaustif, du moins représentatif, des différentes positions qui ont pu être prises pour traiter de ces formations. Je m'attacherai tout d'abord à donner les critères définitoires des éléments entrant dans la formation de ces mots construits souvent appelés composés néoclassiques, avant d'en étudier le statut, en les différenciant ou en les rapprochant des lexèmes ou des affixes. Enfin, je m'intéresserai au mot construit dans son ensemble, en l'étudiant au travers des différences et des ressemblances qu'il entretient avec les composés ordinaires.

1.3.1 Qu'est-ce qu'un élément néoclassique ?

Les éléments qui entrent en jeu dans les formations auxquelles cette thèse s'intéresse ont reçu un nombre important de dénominations différentes. Dans la littérature anglophone, le terme de *combining form* semble satisfaire la majorité (Bauer, 1983; Warren, 1990, par exemple). Kastovsky (2009) fait remonter l'apparition de ce terme à l'*Oxford English Dictionary* (OED) dans

sa version précédente appelée *New English Dictionary* (1884-1928) qui l'employait pour évoquer non seulement les parties d'emprunts grecs ou latins, mais également les constituants entrant dans des formations anglaises qui n'étaient pas des mots à proprement parler ni des affixes facilement identifiables (*astro-*, *bio-*, *neo-*, *micro-*, etc.). Cependant cette dénomination n'est pas unique dans la littérature anglophone : Lüdeling *et al.* (2002) parlent de *neoclassical elements* ou de *neoclassical items*, mettant l'accent sur l'origine classique de ces éléments, réutilisés dans la langue moderne d'une nouvelle façon. Amiot et Dal (2007), tout comme Lehrer (1998), mélangent les deux appellations et parlent de *neoclassical combining forms*. De façon presque similaire, Fradin (2000) utilise le terme de *(classical) combining form*. Scalise (1986) utilise quant à lui le terme *learned stems*, qu'il traduira en italien *semiparole*, 'semi-mots' (Scalise, 1990). Bauer (1998), s'il utilise en général l'expression *combining form*, distingue certains éléments comme *socio-* dans *sociolinguistics* avec la dénomination *neoclassical stem*. Le terme de *stem* est aussi préféré par Ralli (2013) qui parle de *bound stem*, mettant l'accent sur la non autonomie de ces constituants.

En français ou dans d'autres langues comme l'italien, les dénominations sont également variées. Par exemple, Amiot (2011) utilise la dénomination *constituant néoclassique*, Kocourek (1982) utilise le terme de *confixe*, dont le correspondant italien est défini par Sgroi (2003, p.82) comme :

« ogni parola/lessema (i) dotata di autonomo significato lessicale (cioè non-grammaticale), costituita quindi da 'radice + vocale tematica + desinenza', suscettibile di dar luogo a un composto, un derivato, un parasintetico, (ii) forma legata, non suscettibile cioè di apparire come parola libera, (iii) forma « non-nativa », storicamente spesso di origine greca o latina, che (iv) forma la testa in particolare nei composti endocentrici²² ».

On retrouve souvent le terme de *formant*, par exemple chez Cottez (1980, p.VIII) qui explique : « ce sont des signes, eux-mêmes inanalysables, qui *forment* des signes analysables et qui n'existent que parce qu'ils forment effectivement ces signes complexes ». C'est donc leur faculté à constituer des mots en se combinant qui est mise en avant, comme le font Villoing et Namer (2006) avec le terme *élément de formation* ou Amiot et Dal (2008) avec le terme *élément de construction néoclassique*. Cottez inclut sous la dénomination de *formant* aussi bien les traductions d'éléments grecs que les unités

22. Traduction : « Chaque mot/lexème (i) doté d'un sens lexical autonome (c'est-à-dire non grammatical), constitué donc de 'racine + voyelle thématique + désinence, susceptible de donner lieu à un composé, un dérivé, un parasynthétique, (ii) forme liée, c'est-à-dire non susceptible d'apparaître comme mot libre, (iii) forme "non-native", historiquement souvent d'origine grecque ou latine, qui (iv) forme la tête, en particulier dans les composés endocentriques »

représentant une autre unité, comme *cyano-* ou *cya-* qui représentent *cyanogène* (*cyanure*, *cyaméthine*) ou *-al-* qui représente *alcool* (*chloral*, *althionique*). Ce sont des unités différentes des apocopes comme *vélo* car elles ne peuvent pas être utilisées de façon autonome, et ne fonctionnent que comme formant, c'est-à-dire d'élément servant à la formation d'un mot. Certains auteurs différencient ces deux types d'unités dans leurs appellations. Tournier (1985) parle de *quasimorphèmes* quand il s'agit d'unités héritées du grec ou du latin (par exemple *-drome* dans *boulodrome*) et de *fractomorphèmes* quand il s'agit d'un lexème tronqué (par exemple *pétro-* dans *pétrodollar*), unités qu'il distingue des *paléomorphèmes*, éléments obtenus sur la base de séries paradigmatiques mais qui ne sont pas disponibles pour la création de nouveaux lexèmes (par exemple *-duire* dans *conduire*, *séduire* ou *déduire*). Corbin (2001) parle quant à elle d'*archéoconstituants* pour les éléments empruntés au grec et au latin et de *fractoconstituants* pour les éléments tronqués représentant des unités lexicales. Iacobini, qui utilise soit le terme de *semi-parole* (Iacobini, 1999), soit le terme d'*elementi formativi* (Iacobini, 2004, p.71), identifie en cette difficulté terminologique une recherche vaine d'un terme unique qui servirait à désigner un ensemble homogène. Afin de pallier à cette difficulté terminologique, j'ai fait le choix d'utiliser la dénomination la plus large possible, élément néoclassique (EN), le terme *néoclassique* servant à désigner non seulement des éléments hérités des langues classiques et utilisés dans la formation de mots dans la langue moderne, mais également des unités non héritées des langues classiques mais employées 'à la manière des éléments classiques', dans le même type de formations, les constructions néoclassiques, que je définirai ultérieurement.

A ces nombreuses dénominations correspondent des traitements différents de ces éléments, à commencer par leur définition et les frontières de l'ensemble. Les auteurs s'appuient en général sur plusieurs critères, qui ne sont d'une part pas les mêmes pour tous, d'autre part ne sont pas envisagés comme ayant le même poids.

Le premier critère souvent mis en avant est un critère étymologique : un élément néoclassique est un élément d'origine grecque ou latine (Iacobini, 2004; Amiot et Dal, 2008; Amiot, 2011, par exemple). Cependant, ce critère étymologique est parfois considéré comme insuffisant (Lüdeling *et al.*, 2002; Amiot et Dal, 2008) : tous les locuteurs n'ont pas une connaissance étymologique. Lüdeling *et al.* (2002) ajoutent un problème de datation : est-ce qu'un mot comme *International*, emprunté par l'allemand à l'anglais au XIX^{ème} siècle mais clairement d'origine latine, doit être considéré comme néoclassique? Pour eux, le critère étymologique n'est pas satisfaisant. Pour d'autres, ce critère empêche d'englober dans la catégorie des éléments néoclassiques des

éléments qui ont toute raison d'y être. Par exemple, Villoing (2012) intègre dans la catégorie des éléments néoclassiques des formes tronquées de lexèmes français (*anglo* dans *anglo-allemand* par exemple), ceux là même que Corbin nommait *fractoconstituants*. Iacobini (2004, p.76) rapproche aussi des éléments néoclassiques des éléments natifs ayant subi une transformation phonologique (par exemple l'ajout d'une voyelle *-o* ou *-i*) : il les nomme *parole modifiée* et fait l'hypothèse qu'ils peuvent devenir, selon des critères phonologiques ou combinatoires, de véritables éléments formants, la frontière étant instable (il cite par exemple *fungo* 'champignon' dans *fungicida* 'fongicide', *fungiforme* 'fongiforme' ou *fungicultura* 'fongiculture', affichant une différence phonologique entre le lexème autonome [fungo] et la *parola modifiée* [fundzi]).

Le deuxième critère, partagé plus ou moins implicitement par la majorité sinon la totalité des auteurs, est un critère lié à l'autonomie syntaxique : si les éléments néoclassiques étaient des lexèmes autonomes en grec ou en latin, ils ont perdu cette autonomie dans les langues modernes (Amiot et Dal, 2007, 2008; Amiot, 2011; Villoing, 2012, parmi d'autres).

Si certains (Amiot et Dal, 2008, par exemple) s'en tiennent à ces deux seules propriétés, d'autres peuvent être mises en avant par certains auteurs. Par exemple, un critère socio-linguistique ou terminologique est souvent donné : les éléments néoclassiques servent à former du vocabulaire savant. Même si cette propriété peut être nuancée, la construction du vocabulaire savant est tout de même considéré comme étant la première finalité de ces éléments (Amiot et Dal, 2007; Iacobini, 2004; Lüdeling, 2006; Villoing, 2012).

Ensuite, des critères combinatoires peuvent être avancés : les éléments néoclassiques servent à former des lexèmes (Amiot, 2011), ou plus particulièrement des composés : « a special combining form which appears only in compounds » (*Oxford English Dictionary* (OED)), « a form of a word used (only) in compounds » (*Shorter Oxford English Dictionary*)²³. Iacobini (2004, p.72), parmi bien d'autres, signale que certains ne se trouvent qu'à l'initiale des mots construits, d'autres seulement à la finale, d'autres enfin peuvent se trouver dans les deux positions. Lüdeling (2006) quant à elle appuie sur le fait que les éléments néoclassiques tendent à se combiner avec d'autres éléments néoclassiques. Les mots alliant éléments néoclassiques et éléments natifs ont souvent mauvaise presse, surtout dans les écrits plus anciens. On se souvient de Hatzfeld *et al.* (1890) qui utilisait les mots de « barbare », Wilmanns (1899) les nomme *Bastardbildungen* ('formations bâtardes').

Parfois, des critères phonologiques sont aussi invoqués : tout d'abord, la présence d'une voyelle de liaison <o> ou <i>, sur le statut de laquelle je revien-

23. Ces deux dictionnaires sont cités par Kastovsky (2009).

drai dans la partie 1.3.3. Pour des langues comme l'anglais, l'allemand ou le néerlandais, la question de l'accent est également mise en avant. Booij (2002) affirme que, en néerlandais, les affixes non natifs influencent la position de l'accent dans le mot construit, alors que les affixes natifs sont neutres de ce point de vue. Lüdeling (2006, p.580) fait une remarque similaire pour l'allemand, quoique moins catégorique : les affixes néoclassiques attireraient ou porteraient l'accent du mot, contrairement aux affixes germaniques. Concernant l'anglais, Plag (2003, p.157) affirme que « the stress behavior of neo-classical compounds differs considerably from that of other compounds » ; par exemple *biblio-thérapie* est accentué sur le constituant de droite, alors qu'un composé natif l'est sur celui de gauche. De même, les mots construits avec les éléments néoclassiques *-graphy*, *-cracy* ou *-logy* ont leur accent sur l'antépénultième, contrairement aux composés non-néoclassiques (*astrólogy*, *biócracy*, *bibliógraphy*). Cette question, controversée chez les auteurs, ne sera pas discutée ici, ne concernant pas le français qui est la langue d'étude principale.

Enfin, plus anecdotiquement, certains auteurs marquent des particularités orthographiques, avec la présence de certains graphèmes comme <ph> (Lüdeling, 2006) ou encore sémantiques : pour Iacobini (2004, p.73), un même concept peut être exprimé par plus d'un élément néoclassique, par exemple *sabuli-* et *aréni-*, 'sable'. Pour Corbin (1999, 2001), sémantiquement les archéoconstituants et fractoconstituants sont de nature descriptive et non instructionnelle ; ce sont des unités infralexicales à sens descriptif. Villoing (2012, p.42) note par ailleurs que les éléments néoclassiques permettent « the creation of a few lexemes that go beyond the scientific nomenclature », appuyant par là un critère de fréquence, comparativement à des éléments natifs. La question de la définition des éléments néoclassiques n'est donc pas une question à laquelle les auteurs répondent de façon univoque. Certains d'entre eux (Iacobini, 2004; Amiot et Dal, 2007; Bauer, 1998, par exemple) soulignent que ce n'est pas une classe homogène, avec des frontières nettes, bien distincte d'autres classes d'éléments. Tous les éléments n'ont pas les mêmes contraintes et ne répondent pas de la même façon à tous les critères.

1.3.2 Éléments néoclassiques, affixes, lexèmes

Les éléments néoclassiques sont donc des éléments particuliers, mettant en jeu différentes propriétés qui leur sont parfois propres ou qu'ils partagent avec d'autres objets morphologiques. Doit-on pour autant en faire une catégorie à part ? Ou faut-il les considérer comme des affixes ? Des lexèmes ? Les auteurs n'ont, là non plus, pas une réponse unique, chacun mettant en avant certaines

propriétés pour appuyer leurs hypothèses.

Pour certains, les éléments néoclassiques sont des lexèmes, ou des unités lexicales (appelés selon les auteurs et les langues thèmes, racines, stems ou encore roots). Le principal argument mis en avant pour étayer cette hypothèse est un argument sémantique. Selon Corbin (1999, 2001) et à sa suite Villoing et Namer (2006), Villoing (2012), entre autres, ils auraient la capacité référentielle de dénoter des entités, des propriétés ou des relations. Ce sens référentiel, descriptif, les oppose aux affixes qui ont, eux, un sens instructionnel. De plus, ces éléments présentent les mêmes qualités combinatoires que des lexèmes autonomes. Ils peuvent non seulement servir de base à la dérivation (par exemple *hydr-* > *hydrique*) mais se combinent également entre eux de la même manière que les constituants de composés (Iacobini, 2004, p.84). Ralli (2013, p.206) note également que les thèmes liés déverbaux en seconde position en grec moderne rentrent dans la composition de lexèmes présentant le marqueur de composition *-o-* qu'on ne trouve pas en dérivation. Ils imposent également une structure argumentale héritée de la base verbale sous-jacente, la même que celle observée dans les composés déverbaux typiques :

- (54) a. *karði-o-lóγ(os)* (lit. 'qui étudie le cœur')
 b. *karði-o-kataktití(s)* ('conquérant du cœur')

Selon Ralli, les exemples (54-a) et (54-b) affichent tous deux une tête déverbale, le premier un thème lié *-lóγ-* 'qui parle à propos de' et le deuxième le nom *kataktitís* 'conquérant' (du verbe *katakt(ó)* 'conquérir'); cette tête déverbale est dans les deux cas combinée avec le même constituant initial, *karði-*, 'cœur', qui constitue à chaque fois le thème de la base verbale sous-jacente, c'est-à-dire, respectivement, le verbe *λέγ(o)* 'parler' et le verbe *katakt(ó)* 'conquérir'. La différence entre les deux constituants déverbaux est que le second peut être autonome quand le premier ne le peut pas. Ces mots construits avec un thème déverbal lié en seconde position peuvent également présenter une structure récursive, tout comme les composés grecs typiques.

Un autre critère très souvent avancé pour rapprocher les éléments néoclassiques des lexèmes est un critère syntaxique : ces éléments appartiennent à une catégorie lexicale majeure. Cependant, la grande différence que les éléments néoclassiques présentent par rapport aux lexèmes est d'avoir la « particularité est de n'être jamais, en tant que tels, promus au statut de mots grammaticaux » (Amiot et Dal, 2008). Pour Montermini (2010b, p.82) ou Villoing (2012, p.33), suivant Corbin (1992), par exemple, ce sont des formes particulières de lexèmes, portant le trait [+lié] dans leur représentation lexicale. D'autres les appellent *bound roots* (Bauer, 1983; Booij, 1992; Plag, 2003, par exemple). Amiot et Dal (2007) démontrent que des éléments

comme *lud-*, *pyr-* ou *pétri-* sont des radicaux B de lexèmes à plusieurs radicaux, respectivement JEU, FEU et PIERRE : ils apparaissent en effet dans des distributions complémentaires avec les thèmes natifs de ces mêmes lexèmes et cette alternance est motivée phonologiquement. Par exemple, le thème /ʒø/ sera utilisé en seconde position de mots construits (*anti-jeu*) quand /lyd/ ne sera présent qu'en position initiale (*ludothèque*). Le fait que /ʒø/ soit formé d'une syllabe unique qui plus est ouverte l'empêche de précéder un affixe commençant par une voyelle, ce qui est possible pour /lyd/. De la même façon, Lüdeling *et al.* (2002, p.265) montrent qu'en allemand, *elektro-* est le *compounding stem* et *elektr-* le *derivation stem* de *elektric* qui, lui, est autonome. La nature liée de ces éléments semble incompatible avec leur appartenance à une catégorie lexicale. Cependant, certains auteurs déjouent ce problème et trouvent d'autres moyens que l'utilisation de ces unités dans une phrase comme éléments autonomes pour leur attribuer une catégorie. Villoing et Namer (2006) par exemple, proposent qu'un élément néoclassique ait la même catégorie que le lexème dont il est la traduction ou l'allomorphe ; de plus, s'il est utilisé comme base pour une affixation, les contraintes de sélection de la base qu'a l'affixe donnent une piste pour inférer la catégorie de l'élément néoclassique. Enfin, le type de construction dans lequel il apparaît est également un indice à prendre en compte. Lüdeling *et al.* (2002, p.270) arrivent à la conclusion que *-morph* serait un adjectif dans *antropomorph* puisque, selon eux, comme le composé a la même catégorie grammaticale que sa tête et que la tête du composé est ici l'élément néoclassique *-morph*, la catégorie du composé induit forcément la catégorie de l'élément néoclassique. Or, ici, *antropomorph* est un adjectif.

Cependant, d'autres caractéristiques peuvent être mises en avant pour distinguer les éléments néoclassiques des lexèmes. Si c'est en général leur non autonomie qui empêche certains auteurs de les considérer comme des lexèmes, d'autres critères sont également considérés. Iacobini (2004, p.86) apporte ainsi des arguments phonologiques : les éléments néoclassiques se terminent par un *-o* ou un *-i* et sont le plus souvent bi- ou trisyllabiques, restrictions qu'on ne retrouve pas dans les mots prototypiques. Ralli (2013) montre également qu'en grec moderne, les composés construits avec des thèmes liés, ce que nous appelons élément néoclassique, ne suivent pas les règles d'accentuation spécifiques aux composés typiques, et n'affichent de plus que des relations de dépendance, laissant de côté les relations de coordination utilisés par les composés. Iacobini (2004, p.86) souligne également une différence du point de vue de la productivité de ces éléments : les éléments néoclassiques peuvent s'insérer dans des listes associatives et syntagmatiques, ce que ne font pas les lexèmes.

Tous les critères qui viennent d'être énoncés éloignent les éléments néoclassiques de la classe des lexèmes. Faut-il en conclure que ce sont des affixes ? Cette position n'est que rarement proposée dans la littérature, alors qu'un élément comme *-ifier*, par exemple, est toujours considéré comme un suffixe, sans que cette question ne soit même questionnée (voir par exemple Lignon (2013)), bien qu'il soit à l'origine un lexème autonome, le verbe *facio* 'faire' en latin, et qu'il ait aujourd'hui perdu son autonomie, comme les éléments néoclassiques dont il est question ici. Episodiquement, seulement certains éléments néoclassiques particuliers sont élevés au rang d'affixes : Darmesteter (1874) dit ainsi que « *cida* de *Parricida* est devenu un suffixe français ». Pour Amiot et Dal (2008), *-forme* est un suffixe ou un constituant en cours de suffixation avancée ; il se comporte en effet comme *-oïde* et *-acé*, voire parfois comme *-ien* et *-ique*. Bauer (1979) avance que, en anglais, *-crat* et *-phil* sont des suffixes, *bio-* et *electro-* des préfixes, mais *-phob* n'est pas un suffixe puisque peut être mis en relation avec *phobia* et *phobic*, qui sont autonomes : c'est donc une racine liée. Cependant, dans le cas où deux affixes seraient combinés, comme dans *electrophil*, ces éléments devraient être requalifiés de racines. Warren (1990) estime que les éléments néoclassiques, étant liés, ont été plus résistants à la grammaticalisation que les affixes qui étaient à l'origine des mots indépendants devenus par la suite liés. Elle exclut toutefois de cette analyse des éléments comme *semi-* ou *pre-* qu'elle considère comme des affixes, alors que *pseudo-* et *neo-* resteraient des éléments néoclassiques. Corbin (1999, 2001) identifie deux éléments *micro-* distincts : si en (55-a) il s'agit d'une unité à sens descriptif ayant le sens de 'petit', donc plus proche du lexème, cette unité recouvre un sens instructionnel devant un nom d'unité de mesure (55-b), 'diviser par 10⁶', et peut donc être analysée comme un affixe :

- (55) a. microtracteur, micro-ordinateur
 b. microampère, microgramme

Amiot et Dal (2007) analysent quant à elles l'élément *micro-* comme un exposant d'une Règle de Construction de Lexème (RCL) et soulignent une similarité dans le comportement avec le suffixe *-et(te)*. Elles font une analyse similaire pour *-logue* qui serait un suffixe, au même titre que *-iste* : cet élément est pour elles très différent d'éléments comme *-vore*, *-phile*, *-fère* ou *-cide*, car il n'est pas perçu comme ayant un verbe dans sa structure, ni comme ayant une interprétation agentive.

A la différence de ceux que je viens de citer, beaucoup d'auteurs avancent nombre d'arguments pour distinguer les éléments néoclassiques des affixes. Prčić (2005, 2008) identifie ainsi huit critères permettant de différencier d'une

part préfixes et éléments néoclassiques initiaux, d'autre part suffixes et éléments néoclassiques finaux. Le premier est un critère de délimitation de la catégorie : les affixes appartiennent à une classe fermée alors que les éléments néoclassiques appartiennent à une classe ouverte. Cet argument est partagé par d'autres auteurs comme Fradin (2000) ou Iacobini (2004). Kastovsky (2009) rejette en partie cet argument, arguant qu'il est d'une part difficilement quantifiable et d'autre part, que des suffixes et préfixes ont fait, et font toujours, leur apparition dans l'histoire, ce qui ne fait pas des affixes une classe totalement fermée. Ralli (2013) fait quant à elle appartenir les thèmes liés grecs à une classe fermée, les rapprochant ainsi des affixes, alors que le même argument est avancé dans son sens opposé par la majorité des auteurs. Le deuxième critère est d'ordre phonologique : les éléments néoclassiques initiaux ont une forme distinctive, puisqu'ils affichent une voyelle *-o* ou *-i*. Warren (1990) avance le même argument. Le statut de cette voyelle est cependant loin d'être une question à laquelle les auteurs répondent de façon univoque. Par exemple, Kastovsky (2009) pense que dans un mot comme *hamburger-o-logy*, c'est l'élément *-logy* qui déclenche l'insertion d'un élément liant qui n'appartient à aucun des deux constituants. Je reviendrai sur le traitement de cette question dans la partie suivante (1.3.3). Le troisième critère, très souvent repris, tient compte de la combinabilité des éléments néoclassiques : il ne peut pas y avoir de mots construits uniquement avec un préfixe et un suffixe, tout mot construit nécessite au moins une base lexicale. Donc, comme dit Plag (2003), si *bio-* est un préfixe et *-logy* un suffixe, *biology* n'aurait pas de base, ce qui n'est pas tenable : nous ne pouvons pas être en présence d'éléments affixaux mais lexicaux. Pour Prčić (2005, 2008), les suffixes ne peuvent apparaître qu'avec des éléments autonomes, c'est-à-dire qu'ils ne peuvent pas être employés avec un autre affixe ou des éléments néoclassiques. Les éléments néoclassiques finaux peuvent être employés avec d'autres éléments non autonomes (préfixes, autres éléments néoclassiques), ce ne sont donc pas des suffixes. Le quatrième critère est syntaxique : un affixe peut avoir la fonction de maintenir la classe syntaxique du mot construit dans lequel il est employé (*possible_A* 'id.' > *impossible_A* 'id.'; *leader_N* 'chef' > *leadership_N* 'qualité de chef') ou de la changer, en particulier pour les suffixes (*friend_N* 'ami' > *befriend_V* 'devenir ami'; *broad_A* 'large' > *broaden_V* 'élargir'). Les éléments néoclassiques n'ont aucune de ces fonctions, étant eux-mêmes des bases nominales, verbales ou adjectivales sur lesquelles les fonctions syntaxiques interviennent, au même titre que des bases autonomes. Le cinquième critère met en avant la relation tête-modifieur. Prčić (2005, 2008) assume qu'un préfixe entre dans une relation endocentrique de subordination avec la tête puisqu'il porte un poids sémantique plus faible que

la tête. Un suffixe est quant à lui la tête du mot construit, agissant comme un hyperonyme conceptuel du mot construit. En revanche, les éléments néoclassiques entrent dans des relations de coordination conceptuelles, puisque la tête et le modifieur apportent à chaque fois un poids sémantique et structurel identique (*morpho-logy*). Le sixième critère avancé par Prčić est sémantique : comme je l'ai déjà évoqué, les éléments néoclassiques ont un sens descriptif quand les affixes ont un sens fonctionnel. Les éléments néoclassiques ont donc une densité sémantique équivalente à celle des noms ou des adjectifs. Warren (1990) avance aussi cet argument, soulignant que les suffixes ont une fonction dérivationnelle ou flexionnelle qui modifie le contenu de la base, ce que ne font pas les éléments néoclassiques finaux. Cependant, elle avance que les préfixes ont un contenu lexical, ce qui tendrait à réduire la distance avec les éléments néoclassiques initiaux. Iacobini (2004, p.85) souligne aussi que les éléments néoclassiques ont un sens plus large que les affixes et ajoute qu'ils sont non seulement sémantiquement plus autonomes, mais qu'ils n'ont pas non plus de restriction dans les éléments avec lesquels ils peuvent se combiner, au contraire des affixes. Prčić avance comme septième critère que, onomasiologiquement, les affixes entreraient dans des patrons de formation déjà constitués, et sont le seul choix offert pour exprimer un sens spécifique, quand les éléments néoclassiques rentrent dans des patrons non récurrents et peuvent présenter des alternatives pour exprimer une idée : *bibliophile* peut par exemple être remplacé par le composé ordinaire *book lover*. Ce critère est largement discutable ; en effet, ne peut-on pas faire commuter *pianiste* et *joueur de piano* ? De plus, un sens spécifique peut être traduit par différents suffixes et donc différents patrons de formation, comme le montre cet exemple, parmi d'autres, de Roché (1997) :

- (56) a. chameau > chamelier 'conducteur de chameau'
 b. mehari > mehariste 'conducteur de mehari (sorte de dromadaire)'

Le huitième et dernier critère avancé par Prčić (2005, 2008) concerne la notion de productivité : les affixes affichent une productivité « systématique » puisqu'ils sont employés dans des patrons de formation récurrents, de façon régulière et systématique, ce qui n'est pas le cas des éléments néoclassiques. Un affixe peut donc être positionné sur une échelle de productivité, ce qui n'est pas le cas des éléments néoclassiques. Là encore, nous verrons dans la partie 5.3 que certains éléments néoclassiques sont en tout point identiques aux affixes de ce point de vue là. Warren (1990) considère, dans la même mouvance, que les éléments néoclassiques n'ont pas à avoir une forte productivité, ils peuvent apparaître dans des formations uniques ou occasionnelles.

Cependant, des auteurs comme Iacobini (2004, p.86) mettent en avant que certains éléments néoclassiques forment de longues séries de mots, sans toutefois en calculer la réelle productivité. De plus, certains affixes peuvent être très peu sinon pas productifs (par exemple *-aume* dans *royaume*).

En plus de ces huit critères énoncés par Prčić, d'autres ont pu être avancés par d'autres auteurs pour distinguer affixes et éléments néoclassiques. Warren (1990) considère ainsi que les éléments néoclassiques (en particuliers les initiaux) peuvent être reliés à un mot modèle (*model word*), ce qui n'est pas le cas des affixes. Cependant, Roché (2011a) montrera le contraire puisque pour lui, les dérivés en *-iste* et *-isme* se construisent en relation avec un leader word (voir plus bas partie 2.1). Par ailleurs, Fradin (2000) et Iacobini (2004, p.85) arguent qu'un affixe a toujours une position fixe, ce qui n'est pas le cas de tous les éléments néoclassiques, puisque certains peuvent se retrouver à l'initiale ou à la finale, avec des possibilités de variations allomorphiques : si la position d'un élément peut changer, c'est que ce n'est pas un affixe. Ralli (2013) nous informe quant à elle que les thèmes liés grecs, à l'exception de certains qui requièrent uniquement un thème d'origine classique, n'ont pas de restriction de sélection, ce qui n'est pas le cas des affixes. Scalise (1986, 1990) ajoute comme argument que les éléments néoclassiques ont la faculté de pouvoir être factorisés, contrairement aux affixes :

(57) Scalise (1986, p.75) :

- a. it does not matter if they are philo- or anti-Soviet.
- b. *I do not know if he should be dis- or encouraged.

Enfin, Bauer (1983) affirme que les éléments néoclassiques initiaux ont tendance à produire davantage d'hyponymes de la base que ne le font les préfixes. Si certains, comme Warren (1990), tirent la conclusion que les éléments néoclassiques ne peuvent être ni des lexèmes ou allomorphes de lexèmes, ni des affixes, mais constituent une classe à part, d'autres auteurs préfèrent voir une classe hétérogène d'éléments qu'on regroupe sous la même appellation mais qui ne possèdent pas les mêmes caractéristiques. Kastovsky (2009) prône d'ailleurs l'abandon de cette catégorie qu'il trouve superflue. Amiot et Dal (2007) distinguent par exemple *lud-* de *-logue*, le premier étant un radical B de lexème quand le second s'apparente plus à un suffixe. Pour Amiot (2011), *poly-*, *multi-* ou *pluri-* sont des radicaux du grammème PLUSIEURS alors que *hyper-*, *archi-* ou *méga-* sont des préfixes et non pas des radicaux supplétifs de *sur-*. Iacobini (1999, p.391) avance six critères pour distinguer les affixes des constituants de composés :

1. la position : les affixes ont une position fixe alors que celle des constituants de composés peut varier

2. la capacité combinatoire : les affixes ne peuvent former un mot avec un autre affixe, les constituants de composés peuvent servir de base à une dérivation
3. la relation entre les constituants : alors que la dérivation ne permet qu'un rapport de subordination, la composition permet également un rapport de coordination
4. la position de la tête :
 - (a) à droite pour les dérivés, à gauche pour les composés ;
 - (b) la composition, à l'inverse de la dérivation, peut être exocentrique
5. sémantiquement, les affixes sont du type fonctionnel/réalisational, alors que les constituants de composés sont du type dénominatif/lexical
6. l'autonomie : un affixe est un élément lié, le constituant de composé est autonome.

A partir de ces critères, Iacobini démontre que les éléments néoclassiques (*semiparole*) ne peuvent être traités en tant qu'ensemble homogène. Un élément comme *-cida* se comporte comme un affixe selon tous les critères hormis le 5 pour lequel l'élément répond aux caractéristiques sémantiques et des affixes et des constituants de composés, avec tout de même une prévalence du côté des affixes. Un élément comme *radio* au contraire a les mêmes caractéristiques que les constituants de composés, à l'exception de la position de la tête (4a) pour lequel il se rapproche un peu plus des affixes. Des éléments comme *anglo-* ou *socio-* sont plus ambivalents : ils se rapprochent des constituants de composés pour les critères 3, 4b et 5, des affixes pour les critères 1, 4a et 6 ; pour le critère 2, ils répondent aux caractéristiques des affixes et des constituants de composés, même s'ils semblent plus proches de ces derniers. Dans Iacobini (2004), il regroupe les éléments néoclassiques en plusieurs types, refusant toujours de les considérer comme un ensemble homogène. Le premier type regroupe des unités ayant des caractéristiques lexicales ; il y a une grande variété à l'intérieur même de ce type, certaines peuvent être à la base de dérivés, certaines peuvent avoir deux positions quand d'autres n'en auront qu'une seule, etc. Le deuxième type avancé par Iacobini est le type *pseudoproblema, equidistante* : l'élément néoclassique est toujours en position initiale et a toujours la fonction de déterminant, que la construction soit endocentrique ou exocentrique. Cet élément est proche des préfixes, mais l'inventaire n'est pas fermé. Le fait que les éléments de ce type soient employés devant des mots de la langue commune ainsi que le type de signifié qu'ils expriment montre qu'ils sont en voie de grammaticalisation. A l'intérieur de ce type, un sous-groupe homogène se détache, constitué par les éléments à signification numérale. Iacobini présente un troisième type, représenté par *eurocrate* ou

politologo : les éléments néoclassiques sont seulement en position finale dans des composés nominaux ou adjectivaux, endocentriques ou exocentriques. Ils ont dans ces composés la fonction de tête sémantique et/ou syntaxique. Les éléments sont d'origine verbale et étaient déjà utilisés dans des composés dans les langues classiques ; ils forment des noms ou des adjectifs ayant un signifié de type agentif ou instrumental. Ils constituent des séries paradigmatiques : les non suffixés sont des noms d'agent, les suffixés en *-ia* désignent des actions, qualités, états, activités, résultats d'activités et les suffixés en *-ico* sont des adjectifs se référant aux noms en *-ia*.

Maintenant que nous avons vu les différentes positions prônées au sujet des éléments néoclassiques, voyons ce qu'il en est des constructions dans lesquelles ils interviennent.

1.3.3 Quelles sont les caractéristiques des constructions néoclassiques ?

Au même titre que pour les éléments néoclassiques, les constructions dans lesquelles ces éléments entrent en jeu répondent à des appellations différentes dans la littérature. Commençons par le terme « composition » : si la plupart des auteurs l'utilisent, que ce soit en français, en anglais (*compounding*), en italien (*composizione*) ou dans d'autres langues, certains lui préfèrent l'alternative de *formation* (Ralli, 2013) ou *word formation* (Marchand, 1969; Lüdeling *et al.*, 2002; Lüdeling, 2006; Fradin, 2000). Cette alternative offre l'avantage de ne pas catégoriser d'office ces formations comme des composés, alors que cette catégorisation peut ne pas être exacte, comme nous le verrons dans cette thèse. Pour qualifier cette composition ou cette formation, les auteurs déploient un large éventail de dénominations, selon ce qu'ils souhaitent mettre en avant. Certains veulent mettre l'accent sur l'origine étrangère des constituants : Guilbert (1975) parle de *composition allogène* (p.224 : « Le trait dominant de la composition dite savante, à savoir qu'elle fonctionne essentiellement avec des éléments linguistiques tirés de langues anciennes, latin ou grec, et aussi avec des éléments empruntés, pour l'essentiel, à l'anglo-américain, ne peut manquer de retenir l'attention. Nous proposons de la désigner sous le nom de composition allogène, la compréhension de cette notion étant susceptible de mieux recouvrir l'ensemble du phénomène ; »), Kastovsky (2009) de *non-native stem compound*, Marchand (1969) de *word formation on a Neo-Latin basis*, la notion de *Neo-Latin* comprenant les patrons de formation grecs et latins. Pour d'autres, c'est la fonction de ces mots construits qui est mise en avant, en particulier celle de dénommer des concepts scientifiques : Corbin (2004) parle de *composition savante*, Ralli

(2013) et Fradin (2000) de *learned compounding*, qu'Amiot et Dal (2008) traduisent par *composition érudite*. Iacobini (2004) quant à lui désigne ce processus de formation de mots par les éléments qu'il met en jeu : *composizione con elementi neoclassici*. Mais c'est la dénomination *néoclassique* ou *neoclassical* qui est la plus couramment utilisée. Bauer (1983) argumente ainsi : « the ancient Greeks never needed the word *telephone* although it is made up of Greek elements. It is because these elements are put together by speakers of English that it is possible for coiners to mix Greek and Latin as in *television*. Hence it is that such words are termed not 'classical' but neoclassical compounds ». Prčić (2005) a la même analyse : « 'neoclassical' rather than 'classical', because they have been, and are still are, created in modern times ». Il propose par ailleurs une plus grande variété dans les types de composition : il réserve le nom de *neoclassical composition* pour les mots dont les deux éléments sont d'origine classique (par exemple *morphology*), et introduit les termes de *semi-neoclassical composition* pour les formations impliquant un élément classique et un élément moderne (ex : *filmography*) et de *quasi-neoclassical composition* pour les formations impliquant deux éléments modernes mais qui sont forgés sur le modèle des composés néoclassiques, en mettant en jeu par exemple une voyelle de liaison (*semantico-pragmatic*). Par ailleurs, certains auteurs utilisent, dans un même article ou ouvrage, des dénominations différentes : Fradin (2000) emploie aussi bien *learned word formation* que *learned compounding*, Ralli (2013) se sert indifféremment des appellations *learned compounds*, *neoclassical compounds* ou *neoclassical formations*. Concernant les mots qui sont produits par ce processus de formation, si la plupart des auteurs les appellent *composés néoclassiques* ou *neoclassical compounds*, certains vont préférer des alternatives, par exemple Kastovsky (2009) qui les nomme *stem compounds*, sous-entendant par là que ce sont des composés formés grâce à des unités infralexicales, des *stems*.

Donc, que faire, que choisir dans cette multitude d'appellations ? Certaines d'entre elles délimitent trop étroitement l'objet d'étude : nous verrons tout au long de ce travail que ces formations ne sont pas dévolues au seul registre savant, même si c'était le cas à l'origine. Les appeler *formations savantes* n'est donc pas juste. Il en est de même pour une dénomination ne s'appuyant que sur l'origine étrangère des constituants : cela élimine d'office des formations comme *pétrodollar* ou *poulpe attitude* (voir partie 5.2.2). La première met en jeu un constituant français tronqué (*fractoconstituant* selon l'appellation de Corbin, voir section 1.3.1), la seconde deux éléments natifs mais combinés dans une configuration plus proche des composés néoclassiques que natifs (voir section 5.2.2.1). Le terme de néoclassique permet d'englober le plus grand nombre de formations qui entrent en jeu selon moi : les formations qui

sont effectivement empruntées au grec, comme *aristocratie* mais qui peuvent recouvrir aujourd'hui un sens différent, les formations qui imitent les composés grecs ou latins, impliquant des éléments d'origine classique (*dermatologue*), des éléments classiques et natifs (*cancérologue*) ou uniquement des éléments natifs (*pétrodollar*). Ces dernières peuvent contenir une voyelle de liaison (*cocaïnodépendant*), un ordre particulier, dit néoclassique (Amiot et Dal, 2008, p.89) (*la Gignac-dépendance*). Ce ne sont plus uniquement des formations classiques qui sont en jeu mais des formations « comme classiques », soit néoclassiques. Je les nommerai donc constructions néoclassiques (CN).

Quelles sont les différences entre composition ordinaire (ou native) et constructions néoclassiques ? Si l'on part du principe que les éléments qui entrent en jeu dans la deuxième sont des lexèmes, et donc que nous sommes bien face à un cas de composition, trois principales différences sont en général rapportées : premièrement, la composition ordinaire implique des éléments natifs, la composition néoclassique des éléments d'origine classique. Ensuite, dans les langues romanes, l'ordre des constituants est inversé : la composition native présente la tête des composés à droite (ordre XY, où X apporte les caractéristiques sémantico-syntaxiques au composé), la composition néoclassique l'inverse (ordre YX). Cet ordre est aussi nommé par Roché²⁴ ordre allogène. Amiot et Dal (2008) soulignent que « en français, l'ordre YX s'observe dans les composés dès lors que l'un des constituants qu'il met en jeu ne coïncide pas avec une forme de lexème du français ». Pour elles, cet ordre est marqué par rapport à l'ordre XY. Il permet également la récursivité.

La troisième différence est la présence, dans les composés néoclassiques, d'une voyelle de liaison. C'est en général un <o> quand il s'agit d'éléments grecs et un <i> quand il s'agit d'éléments latins. La question de cette voyelle prête à discussion dans la littérature : si c'était une voyelle thématique en grec ancien, quel est son statut aujourd'hui ? Auquel des constituants appartient-elle ? Bauer (1998) propose quatre hypothèses pour répondre à cette dernière question. La première hypothèse, souvent défendue par les lexicographes est que la voyelle de liaison n'appartient ni au premier, ni au second constituant ; elle est seulement un élément de liaison. Ainsi, *photograph* pourrait être décomposé : *phot-o-graph*. Le problème de cette hypothèse, nous dit Bauer, est que l'élément *phot* n'apparaît que très rarement sans cette voyelle dans d'autres constructions en anglais. Ce <o> ferait donc partie du premier élément : c'est la deuxième hypothèse. Quand le premier élément est attaché à un lexème, il prend ce <o> (*photoluminescence*). Cette hypothèse serait plus proche de l'intuition native, comme le montrent les apocopes ou abréviations qui conservent ce <o> à la fin du constituant tronqué. C'est l'hypothèse

24. Communication personnelle.

défendue par Warren (1990) pour qui la voyelle de liaison appartient aux éléments néoclassiques initiaux, ce qui les différencie des éléments finaux en ce qu'ils ont une forme phonologique caractéristique. La troisième hypothèse est selon Bauer la moins préférable : la voyelle de liaison appartiendrait au second élément. Quand le second élément est attaché à un constituant initial, on lui adjoint la voyelle (*phraseograph*). Scalise (1986, 1990) rejette en effet cette hypothèse : pour lui, postuler un élément *-ology* (*-ologia* dans la version italienne) parce que la voyelle est présente quand le second élément, d'origine grecque, est employé (*music + logy* > *musicology*) impliquerait que cette voyelle fasse également partie d'éléments comme *-ography* ou *-ophile*, ce qui empêcherait de mettre au jour des relations existant entre *historiography* et *graphology* ou *graphomaniac*, ou entre *Germanophile* et *philantropist* (Scalise, 1986, p.75). Enfin, une quatrième hypothèse envisagée par Bauer mais qui n'est soutenue par personne, à ma connaissance, serait que la voyelle appartiendrait à chacun des constituants : la séquence /oo/ qui apparaîtrait à la jonction des éléments serait morphophonémiquement simplifiée en /o/.

Ralli (2013) se rapprocherait de la première hypothèse quand elle analyse le statut de la voyelle <o> en grec moderne. Pour elle, cette voyelle n'est plus aujourd'hui qu'un marqueur de composition : « I assign to *-o-* the status of a compound marker since its presence is closely related to compounding and has no other function than connecting the two components of the compound structure ». Si la voyelle appartenait à un des constituants du composé, il faudrait considérer que ce constituant ait un allomorphe spécifique qui ne servirait qu'à la composition, ce qui n'est pas tenable pour Ralli. Le <o> est systématiquement présent dans les composés grecs, ses rares absences étant justifiées phonologiquement. Si le marqueur de composition est dans la majorité des cas un <o> , il arrive que ce soit un <i> : le composé est alors marqué dans le lexique comme [+ savant]. Pour Kastovsky (2009), la même analyse peut être faite en anglais. Dans *hamburger-o-logy* c'est le suffixe *-logy* qui déclenche l'insertion d'un élément liant qui n'appartient à aucun des deux constituants. Cet élément liant est de la même nature que les éléments apparaissant dans les composés allemands natifs : *Universitätsbibliothek* 'bibliothèque universitaire', *Kind-er-garten* 'jardin d'enfant'.

Cependant, cette voyelle <o> n'est pas toujours présente dans les composés néoclassiques ; pour Plag (2003), sa présence est déterminée phonologiquement. S'il y a une voyelle à la fin de l'élément néoclassique initial, le <o> n'apparaît pas (cf. *tele-scope* vs. *laryng-o-scope*). Bauer (1983) dégage même des règles pour expliquer la présence ou l'absence de la voyelle. Pour lui, les éléments néoclassiques initiaux se terminent souvent par un <o> et presque toujours par une voyelle. Une forme qui se combine avec un élément néoclas-

sique final se termine en général par un <o> ou à défaut par une voyelle. Si cet élément initial est un thème se terminant par une voyelle, il pourra se combiner directement avec l'élément néoclassique final. Si, au contraire, c'est un lexème se terminant par une consonne, un <o> lui sera ajouté, afin de le transformer en élément néoclassique initial. Les exceptions à ces règles peuvent être expliquées par des généralisations phonologiques. Bauer en déduit donc que les éléments néoclassiques finaux ne peuvent se combiner qu'avec des éléments néoclassiques initiaux, les lexèmes étant transformés en éléments néoclassiques avant leur combinaison.

Pour certains auteurs, il faudrait plutôt distinguer différents <o>. Plag (2003), par exemple, pense que ce segment n'est pas le même dans toutes les formations néoclassiques : il est parfois une voyelle de liaison, comme nous l'avons déjà vu, mais peut également faire partie de la représentation phonologique d'un élément néoclassique (c'est le cas par exemple de *bio-*). Amiot et Dal (2008) identifient également deux types de <o> : le premier constitue une voyelle qui alterne avec <i> dans la composition néoclassique. Le deuxième, présent en syntaxe et en morphologie, n'alterne avec aucun autre phonème. Dans *chiraco-gauchisme*, la voyelle appartient au premier élément, la forme *chiraco-* étant réservée à la composition²⁵. Cela expliquerait la présence majoritaire du <o> comparé à d'autres voyelles comme le <i>. Dans certains cas, le <o> peut avoir davantage de concurrence ; Saulnier (2012) montre ainsi que pour les formations néoclassiques construites avec des numéraux cardinaux, il peut y avoir d'autres voyelles puisque la voyelle originale est souvent conservée (<u> dans *quadrupède*, ou <a> dans *tétraèdre*, par exemple). L'utilisation du <i> semble quant à elle être la preuve d'une construction qui a été faite en français moderne et non pas empruntée aux langues classiques, puisque les emprunts conservent leur voyelle d'origine (cf. *quadrupède*, emprunté au latin et *quadrichromie*, créé en français).

En plus de ces trois différences mises en avant (l'origine des constituants, l'ordre des éléments et la voyelle de liaison), d'autres ont pu être invoquées pour distinguer la composition néoclassique de la composition native. Fradin (2000) met l'accent sur la finalité des formations savantes, qui sont utilisées pour construire les nomenclatures scientifiques modernes et qui ont pour but de forger des lexèmes capables de décrire de façon adéquate les propriétés de référents parfaitement identifiés à l'intérieur d'un domaine de connaissance restreint. Villoing (2012) pointe des différences formelles : la composition néoclassique donne plus de place à la formation d'adjectifs que la composition native (voir Tableau 1.4). La composition native construira plutôt des

25. Amiot et Dal (2008) ne précisent pas ici si, pour elles, il s'agit d'un thème ou d'un radical du lexème.

	Endocentriques		Exocentriques	
	Natifs	Néoclassiques	Natifs	Néoclassiques
Coordonnés	<u>NN</u> <i>histoire-géographie_N</i>	<u>NN</u> <i>rhinopharynx_N</i>	–	<u>NN</u> <i>androgyn_{A/N}</i>
	<u>AA</u> <i>aigre-doux_A</i>	<u>AA</u> <i>franco-allemand_A</i>	–	–
Attributifs	<u>NN</u> <i>poisson-chat_N</i>	<u>NN</u> <i>pithécantrope_N</i>		<u>NN</u> <i>anthropomorphe_A</i>
	<u>AN</u> <i>bleu-ciel_A</i>			
	–	<u>AN</u> <i>mégalithe_N</i>		<u>AN</u> <i>quadrupède_A</i>
Subordonnés	<u>NN</u> <i>pause-café_N</i>	<u>NN</u> <i>hippodrome_N</i>	<u>VN</u> <i>ouvre-boîte_N</i>	<u>NV</u> <i>biographe_{A/N}</i>
	–	<u>NA</u> <i>photosensible_A</i>		

Tableau 1.4 – Composés productifs en français (Villoing, 2012)

endocentriques, la composition néoclassique des exocentriques. La flexion est de plus plus simple pour les composés néoclassiques, qui ont des constituants graphiquement non séparés, que pour les composés natifs pour lesquels il peut être difficile de déterminer quel(s) constituant(s) doivent marquer la flexion. Elle identifie également des patrons de formation productifs différents pour les deux types de composition. Le Tableau 1.4 reporte le tableau récapitulatif qu'elle donne en conclusion de son article, et répertorie, pour le français, les types de composés productifs selon le type de composition. Certains de ces patrons sont plus productifs que d'autres. Par exemple, si les composés coordonnés NN sont fréquents pour la composition native, ils sont rares en composition néoclassique. De plus, parmi les composés VN / NV, une différence majeure sépare composés natifs et composés néoclassiques, à savoir que les verbes statifs sont impossibles pour les premiers, possibles pour les seconds (par exemple, *anglophone* (Villoing, 2012, p.49)).

Faut-il pour autant étudier séparément les deux types de composition ? Adams (1973) éprouve quelques difficultés à ne pas le faire : elle estime qu'il faudrait traduire les composés néoclassiques pour les paraphraser mais que cer-

tains éléments sont difficilement traduisibles. Cependant, pour Lüdeling *et al.* (2002, p.277), « there is no principled difference between neoclassical word formation and native word formation ». Pour eux, le trait [+classique] est un trait morphologique parmi d'autres et les éléments néoclassiques sont comme les éléments natifs à beaucoup d'égards. Plag (2003, p.74) a la même opinion : « the only difference between the neoclassical forms and native compounds is that the non-native elements are obligatory bound ». Concernant le français, Corbin (2004) ou Villoing (2012) traitent les deux types de composition de façon similaire, en se basant sur les mêmes points d'analyse, dans un essai d'uniformisation.

Bilan. Ce chapitre a montré que l'emprunt d'un fonds lexical à une langue ancienne pour la constitution d'un vocabulaire savant n'est pas réservé aux langues européennes. En effet, des langues comme le vietnamien puisent dans un lexique allogène pour la création des mots appartenant à un registre de langue soutenu. Cela est possible quand la langue à laquelle on emprunte a un statut prestigieux : ce prestige transparaît ainsi dans la langue emprunteuse. C'est cette raison qui a certainement poussé les savants européens à utiliser le fonds lexical gréco-latin pour constituer un vocabulaire scientifique : ces langues, étudiées alors par une élite, permettaient à ces savants de conférer un statut imposant à leurs disciplines. De plus, l'utilisation du grec et du latin leur offrait la possibilité de pratiquer un jargon qui n'était compréhensible que par leurs pairs.

Cependant c'est ce statut même de lexique allogène qui pose question aujourd'hui aux linguistes. Les constructions dans lesquelles ces éléments néoclassiques apparaissent reprennent certaines caractéristiques des composés grecs ou latins, par exemple le fait que l'élément sémantiquement et syntaxiquement recteur soit à droite, la présence d'une voyelle de liaison entre les deux constituants, tout en s'en éloignant : ainsi, les constructions néoclassiques du français peuvent afficher des formations hybrides, mêlant éléments grecs et éléments latins, voire français, elles ont également restreint la voyelle de liaison au <o> ou au <i> quand l'éventail possible était plus large dans les langues classiques. Dans la littérature, il y a donc une oscillation entre traiter ces formations à part ou les intégrer dans le système. De même, les éléments néoclassiques sont parfois traités comme des lexèmes, parfois, plus rarement, comme des affixes, ou bien alors sont répertoriés dans une classe à part : ce sont par exemple, dans la littérature anglophone, des *combining forms*.

Le chapitre suivant va me permettre de m'inscrire dans un cadre théorique, celui de la Morphologie Constructionnelle (Booij, 2010), qui permet de ne pas rendre compte de ces éléments comme relevant d'une classe homogène.

Par ailleurs, je présenterai la base de données que j'ai constituée afin de conduire une analyse sur les constructions néoclassiques qui ont pénétré la langue courante en français.

Questions méthodologiques

Sommaire

2.1	Choix du cadre théorique	72
2.1.1	Un objet d'étude peu homogène	73
2.1.2	Des schémas dans le lexique	80
2.1.3	La construction du sens des mots construits	90
2.2	L'utilisation de la Toile en morphologie	95
2.2.1	Une ressource inestimable	95
2.2.2	Problèmes qui se posent	96
2.2.3	Écueils à éviter et précautions à prendre lors de l'utilisation de la Toile	99
2.3	Création de la base de données NeoClassy	101
2.3.1	Choix des éléments néoclassiques à étudier	101
2.3.2	Présentation des éléments néoclassiques choisis	105
2.3.2.1	<i>-logie, -logue</i>	106
2.3.2.2	<i>-cratie, -crate</i>	106
2.3.2.3	<i>-phobe, -phobie</i>	108
2.3.2.4	<i>-phone</i>	108
2.3.2.5	<i>-phage</i>	109
2.3.2.6	<i>-vore</i>	110
2.3.2.7	<i>-cide</i>	111
2.3.2.8	<i>-cole</i>	111
2.3.3	Recueil des données et création de la base de don- nées NeoClassy	112
2.3.3.1	Sur la question de la lexicalisation	112
2.3.3.2	NeoClassy : présentation	113

Ce chapitre a pour objet de définir le cadre de l'étude proposée dans cette thèse. J'aborderai tout d'abord, en 2.1, le cadre théorique dans lequel j'ai choisi de m'inscrire : dans ce cadre, celui de la Morphologie Constructionnelle de Booij (2010), les lexèmes servent de base aux constructions lexicales ; les patrons de formation de lexèmes, appelés schémas, sont inscrits dans le lexique, au même titre que les lexèmes. Il permet d'intégrer des items disparates comme le sont les éléments néoclassiques. Dans un second temps, je présenterai le cadre méthodologique qui a permis l'analyse de l'intégration du lexique allogène représenté par les éléments néoclassiques dans la langue : en 2.2 je plaiderai en faveur d'une étude des créations relevées entre autres sur la Toile qui sont le témoin de la vitalité d'un processus de formation. Je présenterai enfin, en 2.3, la base de données NeoClassy que j'ai constituée, en justifiant le choix que j'ai fait de me concentrer sur les créations impliquant certains éléments néoclassiques finaux qui entrent dans la formation de longues listes de lexèmes, étant créées selon des schémas qui semblent disponibles pour les locuteurs.

2.1 Choix du cadre théorique

La question cruciale du choix d'un cadre théorique est abordée ici en fonction du sujet qui nous préoccupe. En effet, le cadre théorique qui sera choisi devra rendre compte au mieux des objets étudiés, à savoir les éléments néoclassiques. Comme nous venons de le voir dans la partie 1.3, les éléments néoclassiques, ainsi que les constructions dans lesquelles ils entrent en jeu, sont des objets qu'il est difficile de regrouper au sein d'une catégorie homogène, ayant certaines caractéristiques qui les rapprochent des lexèmes et d'autres qui les font ressembler à des affixes ; de plus, chacune de ces caractéristiques n'est pas forcément identique pour chacun des ENs. Il arrive aussi qu'un même élément, dans une même ressource, ne soit pas traité de façon uniforme. Ainsi, dans l'article *-logie* du Grand Robert (version en ligne¹), l'EN est d'abord désigné comme « élément tiré du grec entrant dans la composition de (...) » puis, quelques lignes plus bas, acquiert le statut de « suffixe *-logie* ». Le cadre théorique le plus pertinent pour traiter de ces objets sera donc un cadre qui permettra de dépasser la dichotomie lexical / grammatical afin de mettre en perspective les ENs avec les autres éléments servant à la construction de lexèmes et de comprendre comment se construit le sens de mots dont les éléments appartiennent à un fond classique qui n'est pas forcément disponible pour tous les locuteurs du français.

1. <http://www.lerobert.com/le-grand-robert/>.

2.1.1 Un objet d'étude peu homogène

Dans la littérature traditionnelle, les CNs sont généralement définies comme des composés répondant à des règles spécifiques. Ainsi, les CNs, contrairement aux composés natifs, impliqueraient des constituants non autonomes : ce sont des « signes complexes dont la caractéristique est d'être construits de signes simples qui n'ont pas d'existence indépendante de ces constructions (...) à l'inverse des composés du vocabulaire commun dont les unités de base sont autonomes » (Cottez, 1980, p.VII). L'ordre des constituants varie également par rapport aux composés natifs : « appelons X le constituant sémantiquement et syntaxiquement recteur du composé, Y le constituant régi. L'ordre XY correspond à une composition « non savante » : ex. *poisson-chat*, *gris-bleu*, *aide-mémoire* ; l'ordre inverse YX à une composition « savante » : ex. *anthropomorphe*, *microcéphale*, *afro-cubain*, *électro-aimant*, *calorifère*, *anthropophage* » (Corbin, 2004, p.1295). Dans la préface du *Grand Larousse*, Guilbert *et al.* (1986) voient dans les constructions dont la base verbale est le second élément et l'argument interne du verbe le premier un ordre contraire à celui de la syntaxe française, dans laquelle le verbe précède généralement ses compléments. Cet ordre qui suit la syntaxe française est, selon eux, productif dans les composés natifs du type *couvre-lit* mais pas dans des composés du types *philosophe* ou *misogyne* qui ne seraient que des emprunts directs au grec ou au latin. On pourrait faire la même remarque concernant les constructions n'impliquant pas d'élément verbal, par exemple *anthropomorphe* versus *homme-grenouille*. L'ordre inversé serait ainsi une des caractéristiques majeures des CNs : « ces composés se caractérisent à la fois par la morphologie savante des éléments composants tirés du grec et du latin et par l'ordre syntagmatique des éléments, emprunté aussi aux langues mères anciennes et devenu procédé motivé de création dans les vocabulaires scientifiques et techniques » (Guilbert *et al.*, 1986, p.LXX). Pour Riegel *et al.* (1994, p.550), si ces constructions sont différentes des composés natifs, c'est qu'elles suivent les règles de formation des langues classiques : « l'originalité de la composition savante tient à l'utilisation quasi exclusive d'éléments empruntés directement au latin et surtout au grec et à une formation qui respecte à peu près les règles de la composition dans ces deux langues ». Comme le soulignent ces deux dernières citations, l'origine des constituants des CNs est également très souvent mise en avant pour distinguer ces constructions des composés natifs. Cette caractéristique est souvent jugée comme essentielle. Cependant, pour Corbin (2004, p.1295) « la distinction entre les deux types de composition ne repose en effet que secondairement sur l'origine des constituants ». Pour elle, c'est davantage l'ordre des constituants ainsi que leur non-autonomie qui distingue les CNs des composés natifs. D'autres ca-

ractéristiques spécifiques aux CNs sont mises en avant, comme nous l'avions vu précédemment dans la partie 1.3 : la présence de la voyelle <o> qui n'est pas réservée aux seuls composés d'origine grecque (« sur le modèle des composés gréco-latins ou gréco-français dont le premier élément était terminé en *o* (...) on en est venu à créer des composés latins-français avec un premier élément terminé en *o*, qu'il eût ou non droit à cette terminaison » (Hatzfeld *et al.*, 1890, p.105)). Cette voyelle est d'ailleurs perçue parfois comme un élément permettant de marquer graphiquement la frontière entre les deux constituants du mot construit, puisque, contrairement aux composés natifs, les CNs ne présentent en général pas de séparation graphique du type trait d'union ou espace typographique : « la conjonction graphique étroite des éléments constituants dès leur formation, à l'inverse des composés [non savants] où la disjonction et le trait d'union sont de règle à la formation et où la soudure des éléments est le résultat d'un long usage (...) dans le modèle savant, la perception de l'autonomie morphologique et sémantique des termes se trouve maintenue par la soudure graphique constituée par la voyelle *o* » (Guilbert *et al.*, 1986, p.LIX). Si les ENs sont définis traditionnellement par opposition avec les constituants de composés natifs, ils n'en sont pas moins opposés aux affixes. La différence majeure mise en avant, particulièrement par Corbin, se situe au niveau du sens : « sémantiquement, les affixes constructionnels sont porteurs d'un sens instructionnel, c'est-à-dire qu'ils n'ont pas en eux-mêmes de capacité référentielle mais sont porteurs d'une instruction sémantique leur permettant, en combinaison avec tout ou une partie des propriétés sémantiques de leur base, de donner à voir d'une certaine façon le référent désigné par le mot construit. (...) archéconstituants et fractoconstituants ont un sens de nature descriptive et non instructionnelle » (Corbin, 2001, p.43-44).

Cependant, de nombreux exemples mettent à mal ces caractéristiques spécifiques qui ont été identifiées. Ce sont des exemples sur lesquels, pour la plupart, je reviendrai plus longuement dans les parties ultérieures afin d'en faire une description plus poussée, mais dont l'évocation va me permettre ici de démontrer l'hétérogénéité certaine des ENs et donc des CNs, afin d'argumenter en faveur d'un cadre théorique pertinent pour leur analyse. Concernant la non autonomie des ENs, voyons ces quelques exemples :

- (1) a. clownophobie, requinophobie, homophobie, russophobie
- b. truiticulture, maïsiculture, kiwiculture, lavandiculture

En (1), chacun des mots est formé à partir de lexèmes autonomes, même si l'ordre des constituants et la voyelle présente à la jonction des deux constituants les fait se rapprocher des CNs. De plus, chacun de ces mots entre dans

le réseau lexical de constructions impliquant des éléments non autonomes, qu'on peut schématiser :

- (2) a. Xphobe :
- (i) Xphobe : atteint de X-phobie
 ➤ *arachnophobe*
 - (ii) Xphobe : qui fait preuve de X-phobie
 ➤ *homophobe*
- b. Xcole :
- (i) Xcole : qui concerne la X-culture, relatif à la X-culture
 ➤ *truiticole*
 - (ii) Xcole : qui vit, grandit dans X
 ➤ *limicole*

En ce qui concerne l'élément *-phobie*, il entre dans le réseau lexical de *-phobe*. Toutefois, l'élément attaché n'a pas toujours le même sens que l'élément autonome. Si PHOBIE désigne une 'peur panique', *Xphobie* désigne soit la 'peur de X' (*clownophobie*, *requinophobie*), soit, par glissement de sens, une 'hostilité envers X' (*homophobie*, *russophobie*). Chacun de ces deux sens a son pendant en *-phobe* qui désignera alors soit une 'personne atteinte de X-phobie' (2-a-i), soit une 'personne qui fait preuve de X-phobie' (2-a-ii). *Xculture* est également un élément autonome qui se trouve dans le même réseau qu'un élément attaché. Contrairement à *-phobie*, il a sensiblement le même sens quand il est attaché que quand il est autonome (bien que *culture* ne s'utilise pas de façon autonome pour désigner un élevage). Cependant, à chaque lexème en *-cole* ne correspond pas un (potentiel²) lexème en *-culture*. En effet, aucun lexème *Xculture* ne peut correspondre à l'interprétation représentée en (2-b-ii). Cependant, quoi qu'en disent Amiot et Dal (2008), puisque les finales *-cole* et *-culture* peuvent être substituantes, selon leur termes, ou font partie du même réseau lexical, selon mes termes, *-culture* ne peut pas être analysé comme l'exact correspondant du lexème CULTURE qui serait employé dans des séries de lexèmes, comme peuvent l'être des lexèmes comme *poisson* ou

2. Certains lexèmes, s'ils sont tout à fait possibles, ne sont pas pour autant présents dans les ressources, comme la Toile, qui sont à notre disposition. Leur absence peut être fortuite car la construction, la sélection du thème adéquat sont possibles, les contraintes de bonne formation respectées, mais qu'aucun locuteur, à la connaissance du linguiste, n'a eu le besoin de créer ce mot. Ces lexèmes possibles sont rencontrés dans la littérature francophone souvent précédés de la pastille °. Ils sont à différencier des mots impossibles, qui violent des contraintes majeures ou dont l'interprétation sémantique est impossible, et qui sont traditionnellement représentés précédés d'une astérisque. Cette notion de mots possibles et mots existants a fait l'objet par exemple de nombreux travaux, parmi lesquels Corbin (1997) ou Rainer (2012). Je reviendrai sur cette notion dans la section 5.1.1.

fleuve (voir partie 5.2.2 pour des exemples). La caractérisation des ENs dépasse donc la simple opposition autonome / non autonome. Certains lexèmes autonomes se retrouvent dans des constructions qui ont un lien direct avec des constructions impliquant un élément non autonome. Le cadre théorique choisi devra donc être en mesure de dépasser cette dichotomie largement rencontrée dans la littérature.

Une autre dichotomie très fréquente, particulièrement dans les articles de Danièle Corbin, est celle qui oppose le sens grammatical ou instructionnel et le sens lexical ou référentiel, que j'ai citée plus haut. Le premier serait réservé aux affixes quand le deuxième serait dévolu aux lexèmes ; les ENs afficheraient ce sens référentiel et se rapprocheraient par là des lexèmes. Cependant, les exemples (3) et (4) montrent que des ENs peuvent rentrer dans les mêmes schèmes de nomination qu'un suffixe canonique.

- (3) a. Il existe aussi sur la toile des sites où vous pouvez concourir entre **sudokistes** : sudoku. fluctuat. net. (<http://www.lexpress.mu/article/la-fi%C3%A8vre-du-sudoku>)³
- b. Selon les estimations de Keesing France, un tiers des **sudoku-philes** sont de nouveaux joueurs, plutôt jeunes et masculins, venus de l'univers de l'informatique. (<http://frobert.over-blog.com/article-2157894.html>)
- (4) a. Cet « **antiquiste** » ne dédaignait pas la préhistoire et ne concevait pas de séparer la Grèce , Chypre et l' Orient anciens de l' Orient d' aujourd'hui (<http://www.mom.fr/La-mort-d-un-hum-aniste-Jean.html>)
- b. En l'année 20 000, un **antiquologue** trouvera un fossile en très mauvais état de ce qui s'avèrera avoir été une feuille de papier. (<http://maths-msf.site2.ac-strasbourg.fr/spip/spip.php?article421>)

En (3), les éléments *-iste* et *-phile*, attachés au même lexème SUDOKU construisent tous deux un nom d'agent signifiant 'qui pratique le sudoku'. En (4) les éléments *-iste* et *-logue* servent à la construction d'un nom de spécialiste. Les occurrences sur Google relevées en (5) montrent bien que les ENs ne sont pas les outils privilégiés dans ces schèmes de nomination, mais les exemples mentionnés ci-dessus montrent qu'ils peuvent toutefois être des moyens disponibles pour les locuteurs. La section 4.3.2 abordera ces questions plus en

3. Toutes les citations relevées sur Internet seront données en l'état tout au long de ce travail, c'est-à-dire telles que je les ai recueillies, en respectant l'orthographe, la syntaxe et la ponctuation originales.

détails.

	Occurrences sur Google (recherche effectuée le 16/06/15) :	
	sudokiste	8770
(5)	sudokuphile	8
	antiquiste	2373
	antiquologue	4

Par ailleurs, comme il est montré dans Lasserre et Montermini (2015) ainsi que dans les exemples reportés en 4.3.1, des ENs peuvent perdre leur sens référentiel et ne servir qu'à la formation d'adjectifs relationnels, n'opérant ainsi qu'une opération phonologique et catégorielle, et non sémantique.

- (6)
- a. Le tout autour et avec les excellentissimes choix **discologiques** de sebastien Tison, dont la playlist figure dans le même blog. (<http://patwhite.com/node/4314>)
 - b. Sur le salon EGAST 2012 à Strasbourg, ces mêmes acteurs avaient associé leur savoir-faire lors d'un cook-show de haute qualité gastronomique et **biérolgique**. (http://www.julienbinz.com/R elais-de-la-poste-Meteor-fait-mousser-les-mets-etoiles -de-Laurent-Huguet_a3970.html)
 - c. La victoire d'Hugo Chavez est en même temps une cinglante gifle envers cette dictature pétrolifère du Moyen-Orient , où les élections sont interdites, où il n'existe aucune forme de vie démocratique, qui pratique l'apartheid humain entre hommes et femmes, et où les ressources **pétrolifères** sont confisquées par une famille omnipotente qui a fait main basse sur ce pays depuis des dizaines d'années, avec le soutien actif des USA, de la Grande-Bretagne, de la France.... ([http://www.letelegramme.fr/ig/generales /france-monde/monde/venezuela-presidentielle-hugo-chavez-reelu-pour-six-ans-08-10-2012-1865056.php?xtor=EPR-3-\[quotidien\]-20121008-\[detailarticle\]](http://www.letelegramme.fr/ig/generales /france-monde/monde/venezuela-presidentielle-hugo-chavez-reelu-pour-six-ans-08-10-2012-1865056.php?xtor=EPR-3-[quotidien]-20121008-[detailarticle]))
 - d. Le village de Bouzigues fête l'huître à l'occasion de la "foire aux huîtres et produits du terroir" ou encore lors du "salon de la gastronomie **ostréicole**" (<http://www.adt-herault.fr/articles /-144-1.html>)
 - e. je ne vais pas vous souler avec mes connaissances **cinéphiles** et les multiples références que je pourrais mettre dans ma critique , faire lire deux pages de mots sur mon avis sur la manière dont le film a été fait , cela m'importe peu , ce que je dirais simplement c'est que ce film ma fait passer un bon moment , cela est important (<http://www.senscritique.com/film/Bric>

k_Mansions/critique/37424155)

Dans les exemples ci-dessus, il n'y a plus de sens de 'parole' ou 'étude' dans *discologique* et *biérogique*, le sens 'contenir' est effacé dans *pétrolifère*, *ostreicole* ne désigne plus 'relatif à la culture des huîtres' mais bien 'relatif aux huîtres' et *cinéphile* n'affiche aucun des sens habituels de *X-phile*, à savoir 'qui aime X', 'qui collectionne X', 'qui est sexuellement attiré par X', etc. La partie 4.3.1 sera réservée à l'étude de ces CNs qui fonctionnent comme des adjectifs de relation, mais ces exemples montrent déjà que l'attribution d'un sens référentiel aux ENs ne peut pas être faite de façon univoque.

Un autre exemple de perte du sens référentiel est donné en (7) :

- (7)
- a. Bien manger réduit non seulement le risque de développer un cancer mais l'adjonction d'aliments dits "**anti cancer**" fournit une aide extraordinaire aux cellules de notre corps pour combattre les inflammations et le développement de cellules cancéreuses. (<http://alimentation-anti-cancer.com/>)
 - b. j ai souvent pu lire que le curcuma, l'ail, la papaye déshydratée sont des bons aliments **anti cancéreux**. (http://www.ampo.cz/data/mg/201505/a_4559.html)
 - c. Y'a aussi l'ail, l'artichaut, le céleri, les épinards, les brocolis, les crevettes, la grenade, le maquereau, le miel... Y'en a plein d'aliments **anti-cancerigène** (<http://www.jeuxvideo.com/forums/1-50-74716319-1-0-1-0-a-chaque-fois-que-jvais-au-macdo.htm>)

Ces trois citations tirées de la Toile présentent trois adjectifs employés avec un nom recteur identique, qui semblent synonymes. Peu importe qu'il y ait une séquence après la base /kãser/, peu importe sa nature, les trois syntagmes *aliments anti cancer*, *aliments anti cancéreux* et *aliments anti cancerigènes* désignent tous des aliments qui préviennent du cancer. Comme Hathout (2011, p.303) par exemple l'analyse, les adjectifs préfixés en *anti-* répondent à une contrainte de transparence catégorielle : même si la préfixation porte sur le nom, l'emprunt d'un radical à un autre lexème permet au mot construit d'afficher une finale d'adjectif. L'EN *-gène* n'a donc ici aucun sens référentiel et a une fonction identique à celle de *-eux*. Hatzfeld *et al.* (1890, p.103) disaient déjà « des composés mi-grecs, mi-français, qui montrent à quel point beaucoup d'éléments composants grecs sont devenus organiques et ont perdu leur cachet d'origine ». Les ENs ne seraient donc pas uniquement des constituants conservant le sens qu'ils avaient dans les langues classiques pour former des composés dans les langues modernes, certes avec

des caractéristiques différentes de celles des composés natifs, mais qui le seraient de façon homogène. Le fait que certains soient également disponibles, au même titre que d'autres suffixes, pour former des lexèmes dans lesquels leur contribution sémantique est nulle montre que, au moins pour certains ENs, leur analyse comme constituants de composés ne peut pas tenir. Je reviendrai sur cette utilisation des ENs comme suffixes dans la partie 4.3.2, dans laquelle je développerai les exemples cités ci-dessus.

De plus, la question de l'appellation *composé* se pose également. La littérature traditionnelle répond souvent à cette question par le seul argument que ces éléments étaient des composés dans les langues classiques. Cependant, pourquoi considérer les constructions en *-ifier* comme suffixées alors qu'elles sont les descendantes des constructions en *-fico(r)* que j'ai évoquées dans la partie 1.2.2 et dont je reporte ici l'exemple donné à cet endroit-là :

(50-b) [A *fico(r)*]_V : *amplifico* 'élargir' < *amplus* 'large' ; *facio* 'faire'

Pour Brucale (2012, p.112), les adjectifs occupant la première position du composé latin ont une fonction prédicative et le composé exprime un sens causatif du type 'rendre Q' où Q est une qualité, un état ou une condition. Or Plag (1996) dégage également le sens causatif, parmi d'autres, que peut recouvrir *-ify*, par exemple dans ARIDIFY, ce que corrobore Lignon (2013) pour le français, avec l'exemple MOMIFIER. Dans les deux cas, les séquences *-ify* et *-ifier* sont caractérisées sans plus de questionnements comme des suffixes. Le TLFi et le *Grand Robert* traitent cet élément également comme un suffixe, tout en notant son étymologie, le verbe latin FACIO. Pourquoi alors *-ifier* serait sans nul doute un suffixe en français, alors que *-cide* serait un élément néoclassique, entendant par là qu'il aurait un sens lexical et non instructionnel, et inversement, pourquoi ne pas considérer *-ifier* comme un constituant de composé⁴ ? Le critère étymologique, c'est-à-dire le fait que les éléments néoclassiques étaient des constituants de composés dans les langues classiques, ne peut pas être tenu bien longtemps.

Ces exemples problématiques dont l'analyse a été amorcée ici nous montrent la nécessité d'étudier les éléments néoclassiques et les constructions dans lesquelles ils apparaissent dans un cadre théorique qui permette de dépasser les dichotomies traditionnellement inhérentes à leur analyse, celle du lexical contre le grammatical, l'autonome contre le non autonome. Ce cadre doit pouvoir nous autoriser à ne pas attribuer un même statut à des éléments

4. Je reviendrai sur la question de la grammaticalisation de la séquence *-ifier* dans la section 5.2.2.1.

qui, lorsqu'on observe leur utilisation effective, affichent une hétérogénéité certaine.

2.1.2 Des schémas dans le lexique

Le cadre théorique que je considère comme le plus pertinent pour l'analyse des CNs est celui de la Morphologie Constructionnelle. Avant de présenter en quoi cette théorie consiste, je vais m'attacher à définir dans un premier temps sur quelles unités les mots construits sont formés et dans un second temps pourquoi les règles, au sens traditionnel, ne sont pas adaptées à la description du lexique.

Tout d'abord, il convient de préciser que morphologie n'est pas « morceaulogie »⁵. En effet, la construction d'un lexème ne se fait pas par la simple concaténation de morceaux, les morphèmes, qui seraient, selon la définition de Hockett (1958, p.123) généralement donnée, « the smallest individually meaningful element in the utterance of a language ». L'opération est plus complexe et nombre de contre-exemples dont les partisans des théories *morpheme-based* se sortent difficilement le démontre. Comme le dit Aronoff (2007, p.804), « linguists like exoticizing the ordinary » : par exemple, postuler des morphèmes zéro là où une opération sémantique (8-a) ou catégorielle (8-b) se fait sans opération phonologique ou des morphèmes vides de sens là où il y a opération phonologique et catégorielle mais non sémantique (8-c) :

- (8) a. souris 'animal' > souris 'outil informatique'
 b. colle_N > coller_V
 c. président_N > présidentiel_A

Or une théorie *lexeme-based* dépasse ces artifices. Un lexème construit n'est pas un assemblage d'unités minimales douées de sens, comme les théories constructivistes l'avancent. Une approche plus abstractive, au sens de Blevins (2006), permet au contraire de faire des généralisations à partir de mots existants. Aucun « morceau » n'existe a priori, mais des thèmes ou des exposants peuvent être dégagés à partir de la description de l'ensemble des formes d'un paradigme : « an analysis may, in particular, depend on other forms in an inflectional paradigm. (...) a word-and-paradigm model may recognize roots and exponents as 'units of analysis' but these units are purely derivative. They are not 'meaningful' in isolation, and they are morphotactically dependent on surface word forms » (Blevins, 2006, p.563). Ainsi, dans

5. La parenté de ce terme revient à Gilles Boyé ; il a souvent été utilisé dans la littérature, Roché (2011b) en fait d'ailleurs un titre de partie.

la formation des mots complexes, les constituants utilisés ont certes un rôle à jouer, mais les relations existant entre les mots complexes également. Par exemple, un mot comme *royaume*, s'il n'est pas décomposable en une base lexématique et un affixe, appartient tout de même à la même famille morphologique de ROI ou ROYAL. De plus, comme nous l'avons vu un peu plus haut (7), des affixes différents peuvent aboutir, à partir d'une même base, à des lexèmes interprétables sémantiquement de la même manière. J'aborderai précisément la notion de lexème dans la partie 5.1 ; mais il faut d'ores et déjà voir cette entité comme étant multidimensionnelle : il s'agit d'une entité abstraite pourvue d'une forme phonologique, d'une catégorie grammaticale et d'un sens lexical. Une opération de construction de lexème prend donc pour input un lexème, mais celui-ci est intégré dans un réseau lexical : une famille morphologique par exemple. Une théorie basée sur le lexème permet de faire dériver ROYALISTE de ROI, alors que le radical a été emprunté à l'adjectif ROYAL (Roché, 2011a, p.75).

Cependant, une théorie qui prend le lexème comme unité de base de toute opération morphologique n'enterre pas pour autant la notion de morphème, ou du moins ne renie pas la présence d'unités plus petites que le lexème. En effet, si le statut d'affixe est renié, que faire, comme le souligne Roché (2011a), des constructions telles que celles impliquant le suffixe *-isme* ou *-age*, qui montrent une grande variété sémantique ? Si on abandonne la notion d'affixe pour n'en faire qu'un exposant de règle ou de construction, cela revient à nier le lien qui existe entre *atterrissage*, nom d'action, et *feuillage*, nom collectif, lien qui existe puisque, quelque part dans le lexique, *carrelage* signifie et l'action de carreler et l'ensemble des carreaux (Roché, 2011a, p. 137). Comme le dit Aronoff (2007, p.805), « antidecompositionnalists do not deny the value of morphemes, just as chemists do not deny the value of atoms and particles ». Cependant, il y a une différence entre reconnaître leur existence et les voir comme des unités dotées de sens. Les affixes, par exemple, ne sont pas des unités linguistiques dotées de sens dans le sens saussurien mais « de même qu'un déictique tire sa capacité référentielle d'une situation d'énonciation, un affixe doit sa dimension sémantique à la confrontation de deux séries lexicales dans lesquelles les lexèmes sont associés deux à deux » (Roché, 2011a, p.139). C'est ce que Montermini (2010a, p.48), entre autres, nomme la « dérégulation des morphèmes » : les affixes n'ont pas de substance sémantique, et les opérations morphologiques sont mieux caractérisées comme étant « output-based », c'est-à-dire qu'un lexème construit est le résultat optimal de la combinaison de plusieurs facteurs : les propriétés du lexème base, de la construction, de contraintes diverses que nous verrons dans les parties suivantes.

L'opération de construction de lexèmes a pris la forme, chez de nombreux auteurs, de règles de formation (Word Formation Rules (WFR) dans la littérature anglophone, Règles de Constructions de Mots/Lexèmes (RCM/RCL) dans la littérature francophone). La notion de règle en morphologie a été introduite par Halle (1973) qui représente le processus de formation de lexèmes par la figure 2.1.

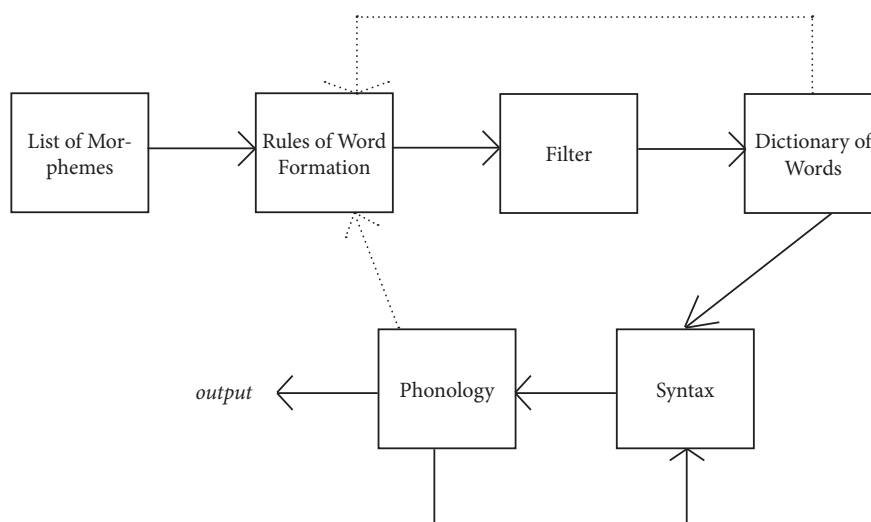


FIGURE 2.1 – Modèle selon Halle (1973)

Ces règles prennent comme input des morphèmes et construisent tous les mots potentiels qui peuvent apparaître dans une langue donnée. Le filtre, qui contient toutes les informations idiosyncratiques, permet lui de générer les mots actuels, ceux qui vont effectivement être utilisés par les locuteurs. Et c'est cet ensemble de mots actuels qui va former le *dictionary of the language*, l'appellation que Halle donne à ce que nous appelons lexique. Cependant, « word formation rule define, in part, the content of the dictionary » mais « they must have access to the dictionary » (Halle, 1973, p.10).

Dans une perspective basée sur le lexème et non sur le morphème, Aronoff propose cette définition des WFR :

such a rule specifies a set of words on which it can operate. This set, or any member of this set, we will term the *base* of that rule. Every WFR specifies a unique phonological operation which is performed on the base. Every WFR also specifies a syntactic label and subcategorization for the resulting word, as well as semantic reading for it, which is a function of the reading base. (Aronoff, 1976,

p.22)

Les règles de formation de lexèmes sont, selon les auteurs, globales ou scindées. Par exemple, Jackendoff (1975) propose de formuler séparément les règles morphologiques et les règles sémantiques puisque plusieurs règles morphologiques différentes peuvent exprimer une même relation sémantique, par exemple plusieurs affixes peuvent nominaliser un verbe :

- (9) enfourner > enfournage, enfournement
 ébarber > ébarbage, ébarbement (exemples de Fradin (2014))

De même, une même règle morphologique peut exprimer différentes relations sémantiques. C'est le cas de la règle morphologique $[NN]_N$ qui, pour Jackendoff (1975), peut avoir différentes interprétations qui sont autant de règles sémantiques. Au contraire, pour Aronoff (1976) ou Scalise (1986), la règle englobe tout ce qui relève de la phonologie, de la sémantique et de la syntaxe :

- (10) WFR formalisée par Scalise (1986, p.42) :
 $[W]_X \rightarrow [[W]_X + Af]_Y$ 'semantics of Y'
 $[F\alpha] \quad [F\alpha] \quad [F\beta]$

La base W a la catégorie syntaxique X et les traits α ; le mot complexe créé par la WFR, de catégorie Y et comprenant les traits β , a comme structure interne la base, une frontière « + » et un affixe. La part sémantique de la règle est donnée sous la forme d'une paraphrase. Cette formalisation est exemplifiée ainsi :

- (11) a. $[read]_V \rightarrow [[read]_V + able]_A$ 'capable of being read'
 $[+tr] \quad [+tr]$
 b. $[boy]_N \rightarrow [[boy]_N + hood]_N$ 'the quality of being a boy'
 $[-abstr] \quad [-abstr] \quad [+abstr]$

Cependant, une vision de la construction des lexèmes comme étant opérée par des règles morphologiques est en fait utopique. Les règles affichent en effet une rigidité certaine : il s'agit d'opérations binaires qui s'appliquent à un lexème, l'input, pour former un autre lexème, l'output. Donc, tout mot formé selon une règle morphologique à partir d'un input qui ne serait pas acceptable ou tout output qui serait en compétition avec un autre output pour une même règle seraient traités comme des exceptions. Et ces exceptions ne peuvent pas être traitées par la règle; Halle (1973), comme c'est illustré dans la Figure 2.1, postule un filtre qui comprendrait toutes les exceptions et idiosyncrasies. Donc les trois lexèmes *anti-cancer*, *anti-cancéreux*

et *anti-cancérogène* présentés en (7) dans la partie précédente comme parfaitement acceptables quoique répondant à la même opération sémantique opérée sur le lexème CANCER par différentes opérations morphologiques, seraient à considérer comme des exceptions ? La même observation peut être faite par exemple sur la base FIÈVRE qui peut donner les dérivés *anti-fièvre*, *anti-fiévreux* ou *antipyrétique*. Le problème principal qui se pose dans une conception dans laquelle les mots sont formés au moyen de règles est qu'elle ne rend pas compte de la réalité du lexique : « La notion même de règle, d'une part, est inadaptée à la réalité du lexique. (...) Dans une dérivation donnée, l'association des opérations sémantique et catégorielle avec tel ou tel affixe ne se fait pas de façon uniforme comme le voudrait une règle, qui peut souffrir des exceptions mais n'est pas modulable » (Roché, 2011b, p.18). De plus, les règles sont trop restrictives quand on les confronte à des données massives. Prenons par exemple l'*Unitary base hypothesis* prônée par Aronoff (1976) mais rejetée par exemple par Corbin (1999, 2001), Montermini (2001) ou Plag (2004). Ces auteurs montrent en effet qu'un même processus de formation peut prendre pour base des lexèmes appartenant à des catégories lexicales différentes, alors que pour Aronoff, « the base is always specified syntactically » (Aronoff, 1976, p.47). Or, au même titre que les exemples donnés par Montermini ou Plag, des constructions impliquant un seul et même élément néoclassique peut avoir des bases appartenant à des catégories lexicales différentes. Ainsi, pour la construction *Xlogie*, X peut être un nom commun (12-a), un nom propre (12-b), un adjectif (12-c) ou même un verbe(12-d) :

- (12) a. Une encyclopédie « magique » sur la « **dragonologie** », cette science hors du commun ! Grâce à cet ouvrage « de référence » en la matière, les dragons les plus rares des régions les plus reculées n'auront plus de secrets pour les lecteurs curieux ! (<http://www.editionsmilan.com/9397b219/Dragonologie.html>)
- b. Je sais, étant d'origine alsacienne je peux sans prétention avancer quelques compétences en **alsacologie** et confirmer tes propos. (<http://hublog.canalblog.com/archives/2007/08/04/5804422.html>)
- c. La **Chiantologie**, est-elle l'art d'emmerder son prochain... ? (<http://fr.answers.yahoo.com/question/index?qid=20080223031839AAYHLgi>)
- d. Aux dernières nouvelles, le professeur travaille sur la «**tombologie**», étant persuadé que la théorie de l'attraction universelle de Newton est erronée. Pour lui, en effet, les objets n'ont aucune raison de tomber. (<http://www.systella.fr/~bertrand/shadok>)

s/shadok.html)

Les règles sont par ailleurs orientées unidirectionnellement. Ainsi, comme, par exemple, Aronoff (1976) le dit dans la définition de la règle donnée plus haut, une règle s'applique sur une base qui est un mot existant. Cela présume de savoir toujours quelle mot est le premier, lequel est construit sur l'autre. Or cela n'est pas toujours aussi simple. En effet, si *esclavagisme* et *esclavagiste* ont pour base *esclavage*, quelle est la base de *pessimisme* ou *pessimiste* (l'exemple est de Roché (2011a)) ?

Au niveau sémantique les règles sont également inadéquates, car souvent la question du sens est traitée à part, comme nous l'avons vu chez Jackendoff (1975). Une règle peut difficilement rendre compte de la polysémie d'un élément ou d'une opération. Les auteurs postulent soit des règles différentes pour chacun des sens attestés, soit, comme Corbin (1989), des règles qui seraient sémantiquement sous-spécifiées, abstraites, la sémantique s'appliquant par la suite. Corbin analyse donc comme des distorsions les différences effectives entre le sens prédictible d'un mot complexe, celui compris dans la règle, et le sens attesté, celui qu'a effectivement le mot construit. Comme la question de la construction du sens des mots construits occupe une place centrale dans ce travail de thèse, je reviendrai dessus plus précisément dans la partie suivante 2.1.3.

Les schémas constructionnels de Booij (2010) se rapprochent des « redundancy rules » de Jackendoff (1975) : en effet, ils sont d'une part la généralisation des propriétés des lexèmes attestés et d'autre part les patrons servant à la formation de nouveaux mots. Ces deux fonctions sont rendues possibles par le fait que le lexique contient non seulement les lexèmes construits, même ceux qui sont forgés de manière tout à fait régulière et/ou qui sont lexicalisés, mais également les schémas de construction. Cela va à l'encontre de la « rule vs. list fallacy » (Langacker, 1987), c'est-à-dire l'idée selon laquelle un mot est soit construit de façon régulière par la règle, soit listé parmi les exceptions. Booij (2010) s'inscrit dans une théorie de type « full entry » selon laquelle schémas et mots complexes instanciant ces schémas sont inclus dans le lexique à un même niveau, ce qu'illustre la Figure 2.2. Les schémas, tout comme les lexèmes (voir à ce sujet la section 5.1.1), spécifient trois types d'information : une rubrique phonologique, une rubrique syntaxique et une rubrique sémantique. Cependant, lorsque je ferai référence à un schéma dans ce travail, afin de simplifier les notations, la rubrique phonologique ne sera pas donnée de façon explicite dans les représentations (comme on peut en trouver dans l'exemple (17) plus bas, et tout le long de cette thèse); les rubriques sémantiques et syntaxiques seront en effet mises en avant. C'est pourquoi ce qui relève de la rubrique phonologique, par exemple la voyelle de

liaison des constructions néoclassiques que je postule comme faisant partie de la construction même (voir mes arguments dans la section 3.1.2), ne sera pas explicitée dans les représentations schématiques, à moins qu'elle ne soit nécessaire pour la démonstration.

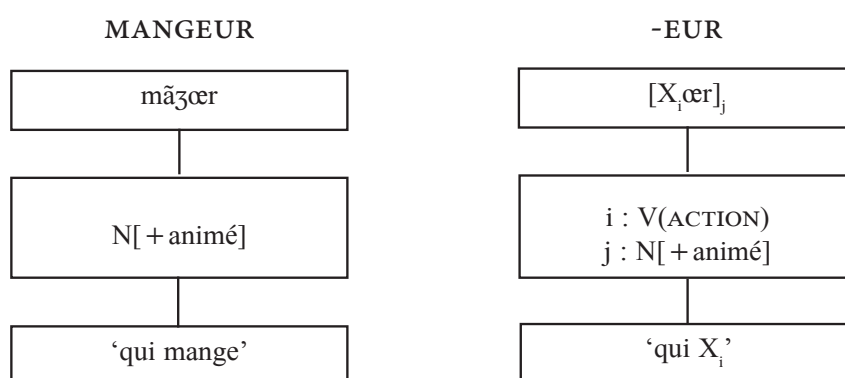


FIGURE 2.2 – Exemple de schéma et mot complexe mémorisés selon la théorie de Booij (2010)

Le lexique tel qu'il est conçu par Booij est un lexique hiérarchisé : un schéma peut donner lieu à des sous-schémas, liés entre eux sur la base d'un principe d'héritage par défaut : un sous-schéma a les mêmes propriétés que le schéma qui le domine, sauf s'il présente des propriétés plus spécifiques. Le schéma qui figure en 2.3 représente des constructions $[NN]_N$ dont les seconds éléments sont spécifiés, et les relie à des schémas plus généraux, dont elles héritent certaines propriétés, par exemple la position de l'élément sémantiquement et syntaxiquement recteur.

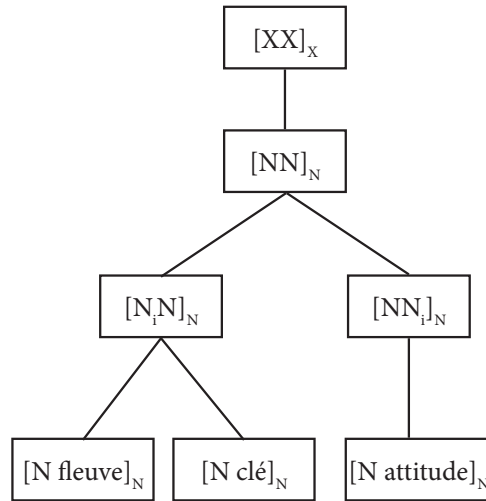


FIGURE 2.3 – Exemple de constructions hiérarchisées

Les constructions telles que $[X\text{fleuve}]_N$, $[X\text{ clé}]_N$ ou $[X\text{ attitude}]_N$, sur lesquelles je reviendrai dans la partie 5.2.2 mais dont je donne déjà quelques exemples ci-dessous, sont ce que Booij nomme des idiomes constructionnels, c'est-à-dire des constructions dans lesquelles certains éléments sont lexicalement spécifiés et d'autres constituent des variables :

- (13) a. $[X\text{ fleuve}]_N$: roman fleuve, discours fleuve
 b. $[X\text{ clé}]_N$: idée clé, notion clé
 c. $[X\text{ attitude}]_N$: beauf attitude, Bayrou attitude

La notion de réseau est centrale dans la théorie de Booij, puisqu'il entend conceptualiser les séries de relations existant entre les mots du lexique : « a lexicon has to be conceived of as a web of words (and phrasal lexical units), a module of the grammar that contains a network of relationships between individual words and morphological schemas. In addition, the schemas themselves are also related in three ways : in terms of instantiation (a schema can be instantiated by subschemas), unification (creating derived schemas), and paradigmatically (with systematic asymmetries between form and meaning). Furthermore, morphological schemas are to be seen as constructions, since they exhibit holistic properties (both formal and semantic ones) that do not derive from their constituents » (Booij, 2010, p.50).

Booij représente les schémas constructionnels avec des variables et des indices lexicaux :

$$(14) \quad \begin{array}{ccc} [[a]_{x_i} [b]_{y_j}]_{x_k} & \leftrightarrow & [SEM_i \text{ entretenant une relation R avec } SEM_j]_k \\ | & & | \\ [\alpha F] & & [\alpha F] \end{array}$$

Dans ce schéma, X est une variable qui représente une catégorie lexicale majeure (N, V, A et P), *a* et *b* représentent des séquences de sons arbitraires, *i*, *j* et *k* sont des indices lexicaux qui indiquent les propriétés phonologiques, syntactiques et sémantiques des mots, ainsi que la position que les mots ont à l'intérieur du lexique. Enfin, $[\alpha F]$ représente l'ensemble des traits pertinents pour la sous-catégorie (par exemple le genre ou la classe de flexion).

Un mot construit, selon Booij, forme un tout, et ne dérive pas toutes ses propriétés des constituants qui sont impliqués. Des propriétés sémantiques peuvent ainsi faire partie de la construction en elle-même, sans pouvoir être déduites de ses constituants. Booij (2010, p.36-38) cite par exemple des composés exocentriques dans les langues romanes, exemplifiés ici en (15), dont l'interprétation ne peut pas se faire uniquement à partir des éléments formants, mais qui est rendue possible à l'intérieur de la construction reportée en (16).

- (15) a. Fr. chauffe-eau
 b. It. lava-piatti 'lave-vaisselle'
 c. Esp. come-curas 'lit. mange-curés, anti-clérical'

$$(16) \quad [[V_k] [N_i]]_{N_j} \leftrightarrow [\text{agent / instrument}_j \text{ of action}_k \text{ on object}_i]_j$$

L'aspect hiérarchique du lexique permet de rendre compte de la polysémie systématique dans les mots construits, ce que Booij (2010, p.78) nomme polysémie régulière : un sens prototypique sert de point de départ pour dériver d'autres sens moins prototypiques, par exemple par métaphore ou métonymie. La figure 2.4 montre ainsi que si les constructions impliquant l'élément néoclassique *-phage* comportent prototypiquement la notion de manger, elles peuvent toutefois donner lieu à des sens moins prototypiques. Ces constructions sont exemplifiées en (17).

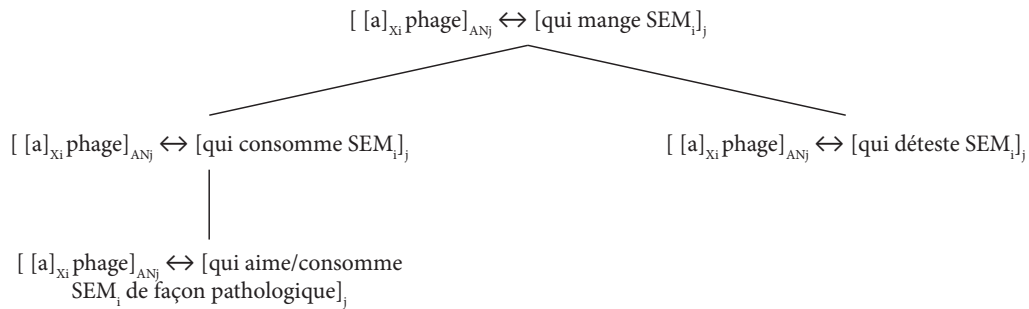


FIGURE 2.4 – La polysémie de la construction [Xphage]

- (17)
- a. $[[a]_{Xi} \text{phage}]_{ANj} \leftrightarrow [\text{qui mange SEM}_i]_j$: anthropophage, trollophage
 - b. $[[a]_{Xi} \text{phage}]_{ANj} \leftrightarrow [\text{qui déteste SEM}_i]_j$: turcophage, francophage
 - c. $[[a]_{Xi} \text{phage}]_{ANj} \leftrightarrow [\text{qui consomme SEM}_i]_j$: chronophage, pétrophage
 - d. $[[a]_{Xi} \text{phage}]_{ANj} \leftrightarrow [\text{qui aime, consomme SEM}_i \text{ de façon pathologique}]_j$: téléphage, cinéphage

Les schémas se situent donc à différents niveaux d'abstraction : ils possèdent plus ou moins de variables et de positions spécifiées lexicalement. Une de leurs propriétés est leur pouvoir d'imbrication : un schéma peut être inclus dans un autre schéma. Ainsi, des mots complexes potentiels mais non existants peuvent servir d'étapes intermédiaires dans la formation de mots nouveaux. Les opérations morphologiques ne sont donc pas obligées de s'appliquer sur des mots existants, déjà présents dans le lexique.

Un des avantages des schémas tels qu'ils sont décrits par Booij est qu'ils permettent de rendre compte de la composition et de la dérivation grâce à des outils similaires et ainsi d'éviter une distinction affixe / élément de composition, ce que nous verrons plus en détail dans le dernier chapitre de la thèse. Des éléments néoclassiques tels que ceux qui sont étudiés dans cette thèse pourront ainsi être décrits plus aisément, malgré leur apparente hétérogénéité. En effet, ils peuvent parfois constituer, comme nous le verrons tout au long de ce travail, des exposants de construction, c'est-à-dire qu'ils ne servent qu'à faire rentrer un lexème dans un réseau lexical. Par exemple, la séquence *-ier*, dont le comportement est décrit par Roché (2011b), sert à faire rentrer des lexèmes dans un réseau qui comporte, entre autres, des noms d'arbres. Il n'est pas nécessaire de postuler qu'elle a elle-même une signification du type 'arbre qui porte X'. Inversement, un même schème de

nomination peut être comporter des mots issus de procédés morphologiques différents. *Bibliophage*, *bibliophile*, *bibliologue* ou *bibliovore* désignent tous, à quelques nuances près, une personne qui aime passionnément les livres, en lit beaucoup, y porte un grand intérêt. Les séquences *-phage*, *-phile*, *-logue* et *-vore* font donc rentrer ces lexèmes dans un réseau lexical sans pour autant avoir de contribution sémantique propre. Je reviendrai donc sur ces questions de statut de l'élément néoclassique dans la dernière partie de cette thèse.

2.1.3 La construction du sens des mots construits

Dans le cadre que je viens de décrire, la sémantique est traitée à l'intérieur même de la construction, et les opérations morphologiques ne sont pas compositionnelles puisque le sens de l'output ne repose pas (uniquement) sur les propriétés sémantiques des constituants de la construction (cf. Aronoff (1976), Booij (2010), entre autres). Cependant, la construction du sens des mots construits n'a pas toujours été conçue comme telle.

Les modèles basés sur les règles ont longtemps conçu la formation du sens des mots construits comme un processus compositionnel. Ils s'appuient en général sur le principe de Frege (1980) qui avance que le sens d'une expression linguistique peut être déduit des parties qui le composent ainsi que de la manière dont ces parties sont assemblées. Cela suppose donc que le sens d'un lexème construit est prédictible puisqu'il est entièrement décrit par le résultat d'une opération sémantique faite sur le sens du lexème base par une règle de formation de mots donnée. Cependant, l'épreuve des données met à mal cette vision idéaliste, puisque, même pour des processus de formation de mots très productifs, il n'y a pas de rapport systématique entre une forme et un sens ; une même opération peut produire des mots complexes affichant une grande variété sémantique, certains sens étant connectés alors que pour d'autres il peut être difficile d'établir un rapport⁶. Roché (2011a) montre ainsi un exemple de construction polysémique, puisque trois modèles peuvent être dégagés pour la dérivation en *-isme*. Le premier construit une relation axiologique (18-a), le deuxième un nom d'action ou d'activité (15-c) et le troisième un nom de qualité (15-a) :

- (18) a. esclavage → esclavagisme
b. parachute → parachutisme

6. C'est le cas pour les mots de la langue courante, utilisés tous les jours par les locuteurs, contrairement à un système terminologique qui se doit d'être compositionnel : « chemical terminology is compositional, not for linguistic reasons, but because the terminology mirrors chemical theory, which is a logical world. But when the world that the words reflect is not entirely logical (and few are), then neither are the meanings of the words. » (Aronoff, 2007, p.814).

c. parallèle → parallélisme

Pour traiter cette polysémie, certains auteurs avancent que la construction a un sens très général, vague, abstrait, nommé *Gesamtbedeutung*, terme inventé par Roman Jakobson (1936) qui signifie en allemand ‘sens général’, sens duquel découleraient tous les sens attestés dans le lexique pour cette construction. C’est ce que Booij (2010) nomme approche monosémique. Selon cette approche, le sens d’une construction est prédictible, et tout sens attesté non conforme au sens prédictible le serait pour des raisons stylistiques, pragmatiques, etc. La théorie associative de Danielle Corbin, présentée par exemple dans Corbin (1989), qui implique que chaque règle de formation de mots génère en même temps une structure morphologique et un sens, prône cette distinction entre sens prédictible et sens attesté : « I will call the meaning of a constructed word as it may occur in its definition in one or more dictionaries of contemporary language the *attested meaning*. In opposition to the directly observable AM we have the *predictable meaning*, i.e. the meaning of a constructed word as it can be formulated on the basis of three factors that contribute to its semantic interpretation : the semantic operation characteristic of the rule by which the word has been constructed (WFR), the semantic characteristics of the morphological process used, and the meaning of its base. Thus, the predictable meaning is a hypothetical construct of the linguist, related to his conception of the general economy of the derivational grammar of the language, i.e. his theory of the lexicon. In my theory, the predictable meaning has two fundamental properties : it is a compositional function of the morphological structure of a constructed word, and it is not fixed by the attested meaning » (Corbin, 1989, p.32). À l’intérieur de ce qu’elle appelle sens prédictible, Corbin différencie le sens prédictible construit par la règle de formation de mots, le sens prédictible spécifié par le processus morphologique et le sens prédictible hérité de la base, chacun de ces sens constituant une strate de plus en plus fine, le premier, construit par la règle, implique tous les outputs générés par cette règle de formation, le dernier étant spécifique à chaque base particulière. Corbin explique donc les distorsions qui existent entre sens prédictible et sens attesté par exemple par des facteurs pragmatiques (une *fourchette* n’est pas une ‘petite fourche’ en raison du domaine extralinguistique dans lequel ce mot est employé) ou par la notion d’intégrateur paradigmatique : si *anti-alcoolique* existe, alors que le suffixe *-ique* dérive des adjectifs de bases nominales mais que *anti-alcool* n’existe pas, c’est que l’adjectif doit avoir une forme qui ressemble à un adjectif, et le suffixe est choisi pour cela en fonction de la forme suffixée du lexème non préfixé, *alcoolique*. Corbin propose d’autres explications pour des distorsions observées sur lesquelles je n’ai pas la place de revenir. Ce-

pendant, il est intéressant de retenir que chaque opération comporte un sens particulier, prédictible et unique, et que les sens attestés sont soit conformes à ce sens, soit dérivés de ce sens par des facteurs autres que morphologiques.

A cette approche monosémique Booij (2010) oppose l'idée d'une polysémie régulière : les constructions ont un sens prototypique duquel sont dérivés d'autres sens, par métaphore ou métonymie. Ces sens sont appelés par Booij des extensions (p.79). Ces sens étendus pourront être conventionnalisés puis être étendus à leur tour, par métaphore ou métonymie. Tous ces sens sont donc hiérarchisés, comme nous l'avons vu dans la figure 2.4. Cette approche permet ainsi d'aborder le sens des mots construits d'un point de vue non compositionnel. Elle permet aussi de rendre effectives les relations qui existent entre les lexèmes qui forment alors un réseau, comme nous l'avons vu dans la partie précédente. Ainsi, pour Rainer (2003, p.207) « the meaning of the derivative thus does not have to be assembled on the basis of the meanings of the constituents (base plus affix) and the meaning of a rule, but is derived from the lexical meaning of the leader word or leader group ». La dimension lexicale prend ici tout son sens ; le sens ne se construit pas uniquement en fonction de la base et de l'opération en jeu, il se construit également par rapport au lexique dans son ensemble. Dans une vision topologique (Hathout, 2009), le sens d'un mot est créé en fonction des distances relatives qu'il entretient avec les autres mots du lexique. Pour expliquer cette vision, la notion de mot leader apparaît comme particulièrement utile : pour chacune des opérations morphologiques, un ou plusieurs mots leader émergent du lexique, car ils sont plus saillants, plus fréquents, plus disponibles dans la compétence des locuteurs, et le mot construit se situera à une certaine distance de ce leader : très proche, il aura un sens similaire, plus éloigné, un sens un peu différent, très éloigné, un autre sens. Ce mot contenant donc un sens étendu, pourra à son tour servir de mot leader dans la formation de nouveaux mots. Chez Rainer également, cette distance qu'entretient un mot par rapport au mot leader est liée à des processus de métaphore ou de métonymie : « in this search-process, speakers may content themselves with an approximation rather than with a full match, especially if the model and the concept to be expressed are related metaphorically or metonymically » (Rainer, 2003, p.208).

La notion de mot leader peut être conçue diachroniquement ou synchroniquement. Diachroniquement, c'est le mot qui a émergé avec un sens particulier et qui a ouvert la voie à la création d'autres mots sur le même modèle. Roché (2011a, p.97) identifie ainsi *christianisme* comme mot leader pour la création de lexèmes *Xisme* selon le modèle de dérivation axiologique (*absolutisme, esclavagisme,...*). Cependant, en synchronie, le caractère saillant de ce

mot leader peut s'être estompé. Par ailleurs, un mot ressenti synchroniquement comme leader, c'est-à-dire un mot fréquent, servant de modèle pour la création de néologismes, peut ne pas être l'instigateur historique du modèle. Je reviendrai sur des exemples détaillés dans la partie 3.2.2.1. Cependant, nous pouvons d'ores et déjà observer que, concernant les formations *Xcratie* deux mots leader, tant diachroniques que synchroniques, se dégagent. Comme nous le verrons dans le Tableau 3.7, les lexèmes *démocratie* et *aristocratie* sont non seulement les premiers à être attestés en français avec l'élément final *-cratie*, mais également les plus fréquents aujourd'hui sur la Toile. Ils fonctionnent tous deux comme modèles en fonction desquels, par analogie, d'autres lexèmes seront créés. Ils exemplifient ce que Booij nomme polysémie régulière, puisqu'ils peuvent être liés entre eux, même si le lien est semblait-il distendu. A l'origine, en grec ancien, *ἀριστοκρατίας* signifiait 'le pouvoir des meilleurs'. En français, le trait sémantique 'pouvoir' s'est estompé et les aristocrates ne forment plus qu'un groupe social, un groupe de personnes. Cependant, ce sens est une extension du sens contenu dans *démocratie*, 'pouvoir du peuple'. On pourrait en effet tout à fait imaginer, même si une étude diachronique plus poussée devra le confirmer, un glissement de sens tel que celui présenté en (19), qui a conduit du sens d'un lexème comme *démocratie* à celui d'un lexème comme *aristocratie*.

- (19) pouvoir de X → groupe de personnes détenant un pouvoir → groupe de personnes

Cette polysémie régulière entre les deux constructions a donné lieu à deux pôles d'attraction ayant pour centre *démocratie* et *aristocratie* et en fonction desquels, dans l'espace du lexique, chaque mot construit va se placer à une distance plus ou moins grande. La Figure 3.31 du chapitre suivant propose une représentation d'un espace lexical de ce type.

Concernant les lexèmes *Xphobie* par contre, les mots qui semblent fonctionner comme modèles pour la création lexicale ne sont pas des mots leader tel qu'on peut l'entendre diachroniquement. En effet, si l'on regarde les lexèmes signifiant 'hostilité envers X' (une analyse sémantique détaillée de ces constructions sera donnée en 3.2.1), les premiers à avoir été construits sont *anglophobie* (1829), *gallophobie* (1845) ou *prétrophobie* (1865). Cependant, ils ne font pas partie des lexèmes les plus fréquents rencontrés aujourd'hui. *Xénophobie*, apparu en 1904, et *homophobie*, apparu en français en 1980 (ces dates ainsi que les occurrences relevées sur Google sont données dans le Tableau 3.8 du chapitre suivant), sont cependant bien plus fréquents aujourd'hui. La Figure

2.5, qui a été obtenue grâce à Google Books Ngrams Viewer⁷, montre en effet que ces deux lexèmes, bien que formés après les trois premiers cités, sont aujourd’hui plus fréquents, dans une large mesure.



FIGURE 2.5 – Représentation de la fréquence de plusieurs lexèmes *Xphobie* dans Google Books, de 1830 à 2008

Cette fréquence d’apparition importante les rend plus disponibles pour les locuteurs dans la formation de nouveaux mots. Bien qu’ils ne soient pas les instigateurs du schéma de construction [Xphobie] ↔ [hostilité envers X], ils sont aujourd’hui au centre du pôle d’attraction en fonction duquel vont se positionner les nouvelles créations.

Cette vision du lexique faite de pôles d’attraction remplis par des mots leaders en fonction desquels les nouvelles créations se placent pour construire un sens (mais également pour ordonner des contraintes d’ordre formel, comme le montre Burzio (2002) qui prône une vision de la morphologie « surface-to-surface »), s’inscrit dans une approche abstraite, selon les termes de Blevins (2006) : la formation des mots dépend d’abstractions qui sont faites sur des séries de mots existants. Les schémas sont des processus actifs, comme des recettes pour former des mots nouveaux, mais qui s’appuient sur des généralisations faites sur le lexique existant. Montermini (2010a, p.59) parle de « morphologie émergente » puisque les opérations morphologiques « émergent dans la compétence morphologique des locuteurs en tant que généralisations à partir du lexique, et que le lexique émerge lui-même comme une généralisation à partir des manifestations concrètes de la langue (énon-

7. <https://books.google.com/ngrams>

cés) auxquels un locuteur est exposé ». Le point de vue non compositionnaliste du lexique renforce donc la différence fondamentale qui existe entre la syntaxe et la morphologie : « the scope of syntax-based logical compositionality should not be extended below the lexeme because lexeme-internal structure (however it is described) is different from lexeme-external structure » (Aronoff, 2007, p.805). Contrairement à la syntaxe, les lexèmes créés par la morphologie sont stockés dans le lexique mental des locuteurs, au même titre que les schémas permettant de les créer.

2.2 L'utilisation de la Toile en morphologie

L'accès aux nouvelles formations qui sont le témoin de la vitalité d'un schéma de construction de lexèmes est un problème récurrent en morphologie et la Toile offre la source de données possible. Il faut cependant utiliser cet outil avec précaution afin d'en retirer tous les avantages en évitant les inconvénients qui lui sont inhérents. Ces avantages et inconvénients ont été plusieurs fois exposés dans la littérature, par exemple par Grefenstette (1999), Namer (2003), Lüdeling *et al.* (2006), Hathout *et al.* (2008), Hathout *et al.* (2009), Dal et Namer (2012) et bien d'autres. Cette section reprend donc l'essentiel de leurs arguments.

2.2.1 Une ressource inestimable

Avant la diffusion massive d'Internet et la naissance du World Wide Web, la morphologie était appliquée soit aux mots présents dans les dictionnaires, soit à ceux relevant de l'intuition du linguiste ou bien à des formations rencontrées par hasard. La Toile a rendu possible non seulement la récolte de nombreux mots mais aussi la vérification d'hypothèses à grande échelle. En effet, comparée à des corpus ou à des bases de données déjà existants, la Toile offre la possibilité d'accéder à un nombre de constructions linguistiques incommensurable. Grefenstette (1999) montre ainsi que pour des recherches identiques opérées en 1998 sur le British National Corpus (qui contient 100 millions de mots) et sur le Web grâce au moteur de recherche Altavista, la recherche sur la Toile recueillait 145 à 1450 fois plus de résultats. Ce grand réservoir de données que constitue la Toile va donc permettre de pouvoir capturer des phénomènes rares qui seraient soit absents des dictionnaires, soit si peu présents qu'ils sont considérés comme des phénomènes anecdotiques. Hathout *et al.* (2008) tracent la chronologie des découvertes de Marc Plénat concernant les formations en *-esque*, au fil de ses trouvailles sur le

Net. Plénat a pu recenser 3000 formes différentes sur la Toile quand les données lexicographiques ne lui en rapportaient même pas 100. L'hypothèse que des contraintes dissimilatives étaient en jeu dans la formation des dérivés en *-esque* quand les lexèmes bases se terminaient en / ϵ / suivi d'une consonne a pu être vérifiée, puis affinée aux lexèmes bases de plus de quatre syllabes. Plus la taille du corpus est importante, plus la chance de rencontrer des formes rares est grande, et ainsi de pouvoir confirmer ou infirmer des hypothèses ou même formuler empiriquement des nouvelles hypothèses. Comme le disent Hathout *et al.* (2009, p. 275) « l'accroissement de la quantité de données disponibles a un rôle analogue à celui de l'introduction du microscope dans les sciences de la nature ».

Par ailleurs, ce qui se faisait manuellement auparavant a pu être automatisé. L'accès à la version informatisée du TLF a par exemple permis d'extraire automatiquement de grandes listes de mots. De plus, la qualité des données a changé. Les corpus traditionnels étaient souvent cantonnés à des types de textes particuliers : des textes littéraires (par exemple Frantext), des textes journalistiques, etc. Sur la Toile, les productions sont plus spontanées, on écrit souvent comme on aurait parlé et « many interesting findings occur in forums, blogs, and other types of pages where a community of speakers freely express themselves in informal contexts » (Hathout *et al.*, 2008). Les types de textes et les contenus sont également variés, on accède tant à des textes officiels qu'à certains articles scientifiques, à des recettes de cuisine ou à des pages interdites aux moins de dix-huit ans, ce qui procure encore une fois la possibilité de recueillir des formes très variées.

Enfin, un autre avantage de la Toile est sa constante évolution : un terme absent à une période donnée pourra apparaître quelques temps plus tard et être ainsi le témoin de la vitalité d'un schéma constructionnel.

2.2.2 Problèmes qui se posent

Cependant, nombre des avantages qui viennent d'être cités sont également des inconvénients. L'inconstance de la Toile empêche par exemple de renouveler une expérience dans des conditions similaires à la première fois. Si des formes peuvent apparaître et d'autres disparaître au gré des pages indexées par un moteur de recherche, cela signifie qu'une même expérience exécutée à des semaines voire des jours différents ne donnera pas les mêmes résultats. Par ailleurs, Lüdeling *et al.* (2006) donnent de nombreux arguments qui vont à l'encontre de l'utilisation du Web en morphologie. Tout d'abord, ce n'est pas un corpus construit comme peuvent l'être les corpus traditionnels, il ne peut donc pas être représentatif. Faire des généralisations sur le Web,

par exemple sur la vitalité de tel processus morphologique dans tel type de discours est donc abusif. Sa taille est certes démesurée mais non quantifiable. Les moteurs de recherche ne nous donnent à voir qu'une partie de ce qui existe sur le Web. De plus, le plus souvent aucune information n'est donnée sur les auteurs des textes : leur âge, leur niveau d'étude, leur situation géographique par exemple ne sont pas connus. On peut donc accéder par exemple à des régionalismes sans en avoir conscience. Enfin, certains textes ne sont que des traductions, opérées soit par des locuteurs potentiellement bilingues, soit par des traducteurs automatiques. Quel poids attribuer à une formation qui n'est qu'une traduction et donc pas forcément pertinente en ce qui concerne la dynamique lexicale de la langue étudiée ?

D'autres problèmes se posent, concernant le caractère écrit des mots recueillis : comment savoir que le mot capturé ne résulte pas soit d'une faute d'orthographe, soit d'une faute de frappe ? De plus, un même mot prononcé de la même manière pourra être écrit de différentes manières. C'est le cas par exemple de *belgeophone*⁸ et *belgophone*⁹. Si le premier, selon la graphie, devrait se prononcer [bɛlʒɔfɔ̃n], ce qui est attendu oralement, la base se prononçant [bɛlʒ] et la séquence finale, dans tous les autres items la comprenant [ɔfɔ̃n], le deuxième est écrit comme s'il se prononçait [bɛlɡɔfɔ̃n], ce qui paraît difficilement croyable. Graphiquement cependant, *belgophone* s'aligne sur des mots comme *celtophone*, construit sur *celte* dont le <e> final s'efface avant la séquence <ophone>. *Belgeophone* et *belgophone* doivent-ils être considérés comme deux mots différents ou comme un seul et même mot écrit différemment par les locuteurs mais construit sur le même constituant initial (C1), selon le même schéma constructionnel, ayant le même C1 et la même forme phonologique ? Je pencherais pour la deuxième solution, mais à part demander à chaque auteur de chaque *belgeophone* et *belgophone* comment ils prononcent les mots qu'ils ont écrits, rien ne me dit que je fais le bon choix. D'autres problèmes graphiques peuvent également se poser, par exemple, pour les constructions néoclassiques, la présence d'un trait d'union pour isoler la voyelle de liaison. Cela permet, en (20) d'éviter la cédille sous le <c> qui, au même titre que le <e> dans *belgeophone*, invite le locuteur à prononcer [s], collant ainsi à la prononciation du C1. En (21), la base étant une lexie, les traits d'union entre les deux constituants du mot construit semblent être dans la continuation des traits d'union qui permettent à une base multiple de servir de base unique dans une dérivation (cf. par exemple *droits-de-l'homme* cité par Roché (2011a, p.93)). En (22), le trait d'union semble être présent pour inscrire paradigmatiquement le mot construit *clien-*

8. 98 occurrences sur Google le 26 juin 2015.

9. 1900 occurrences sur Google le 26 juin 2015.

tocratie sur le mot leader *démocratie*.

- (20) Et même si j'aime beaucoup les possibilités d'Eclipse, n'étant pas à l'origine développeur Java, je trouve dommage de devoir installer un soft aussi **ressourc-o-phage** pour ne finalement utiliser qu'un seul de ses plugins, soit tout au plus 10 à 15% de ses possibilités, le reste ne me parlant absolument pas du tout! (<http://forum.ubuntu-fr.org/viewtopic.php?id=310257>)
- (21) Adressez une réclamation au service après vente de votre fournisseur d'accès à l'éducation, car ils ont omis de vous signaler que l'adjectif « national » prend ici un « e » final, à cause de l'accord de genre (Ouh! le vilain mot!), « éducation » étant du genre féminin. C'est pareil pour « préférence nationale », mais vous êtes peut-être **théorie-du-genre-o-phobe**? (<http://uneanneeaulycee.blog.lemonde.fr/2015/03/27/si-lenseignement-fonctionnait-comme-la-telephonie-mobile/>)
- (22) **Client-o-cratie** mieux que la **Démo-cratie**? (<http://cocreation.blogs.com/alban/2009/11/clientocratie-mieux-que-la-d%C3%A9mocratie.html>)

Un des problèmes majeurs qui se posent avec l'utilisation du Web comme corpus est que justement, ce n'est pas un corpus à proprement parler : il n'a pas de frontières bien définies ce qui empêche d'opérer des calculs similaires à ceux que l'on peut effectuer dans un corpus traditionnel. En effet, si l'on fait une recherche sur la présence d'une occurrence sur le moteur de recherche Google, les chiffres donnés sur le nombre d'apparitions d'un mot ne peut pas être pris en compte en tant que tel. Si un même texte apparaît sur des sites différents, pour des raisons de citation, de plagiat, ou tout simplement de pages dupliquées (c'est-à-dire qu'un même contenu se trouve sur plusieurs pages différentes), ce texte sera compté autant de fois qu'il est utilisé, ainsi que les mots qui le composent. C'est le cas par exemple de *arménianophone* dont la recherche sur Google signale trois occurrences¹⁰ mais qui renvoient toutes à la même page. Parallèlement, Google, comme les autres moteurs de recherche, ayant des visées avant tout commerciales, a tout intérêt à afficher le plus grand nombre possible de résultats de recherches. Les chiffres donnés sont non seulement des approximations (les nombres toujours ronds au delà de 1000 occurrences en sont le reflet) mais révèlent également des aberrations, par exemple lorsque l'on emploie des opérateurs booléens, comme l'a montré

10. Recherche effectuée le 26 juin 2015.

Jean Véronis dans son blog¹¹ : la recherche « Chirac » donne moins de résultats que la recherche « Chirac OR Sarkozy », censée afficher toutes les pages contenant le mot Chirac ou le mot Sarkozy, soit au moins autant que celles contenant uniquement le mot Chirac. J'ai également été confrontée à ce genre de chiffres surprenants dans mes recherches. En cherchant (entre guillemets) le mot « homicide », Google affiche, le 8 mai 2015, 1 480 000 résultats. A la même date, à quelques secondes d'intervalles, j'ai cherché « commettre * homicide » (où « * » est un caractère joker, signifiant que n'importe quel mot peut apparaître à cette place), recherche normalement plus restrictive et pour laquelle on s'attend à recueillir moins de résultats, Google m'en a donné 10 600 000, soit plus de sept fois plus. Jean Véronis, dans d'autres articles¹² tente de trouver des raisons à ces comptes faussés, et fait l'hypothèse que Google gonflerait ses chiffres en n'ayant qu'une partie de sa base de données réellement indexée, et que, à partir de cet index, le moteur ferait une estimation du nombre de résultats que pourrait donner la partie non indexée. Les comptes de Google seraient donc extrapolés. Sans rentrer dans des détails très fins concernant la façon dont Google indexe ses pages, une étude quantitative se basant sur ce moteur de recherche semble exclue. Un corpus comme FrWac, qui est construit sur des pages Web, peut aider à certains calculs ; son utilisation exclut cependant un grand nombre de trouvailles qui ne sont accessibles que via Google.

2.2.3 Écueils à éviter et précautions à prendre lors de l'utilisation de la Toile

Si la Toile n'est pas un outil optimal pour accéder à des données en morphologie, c'est toutefois la solution la moins insatisfaisante. En dépit de tout ce qui vient d'être dit, aucune autre ressource n'est capable, à l'heure actuelle, de donner accès au morphologue à un nombre de données aussi conséquent, même si ses nombreux défauts font que cet outil doit être utilisé avec des précautions.

Une étude qualitative sera plus aisément menée qu'une étude quantitative. La Toile peut offrir des informations sur l'absence *vs.* la présence d'un mot. Si un mot est présent sur la Toile, cela signifie que sa présence est possible. S'il est absent, cela veut soit dire que sa formation est impossible, soit qu'il est possible mais absent sur Internet, soit possible mais pas encore créé, puisque ne répondant, pour l'instant, à aucun besoin onomasiologique.

11. <http://blog.veronis.fr/2005/01/web-google-perd-la-boole.html>

12. <http://blog.veronis.fr/2005/01/web-comptes-bidons-chez-google.html> ; <http://blog.veronis.fr/2005/02/web-le-mystre-des-pages-manquantes-de.html>

Cependant, la présence d'une séquence de lettres sur la Toile peut résulter, comme nous l'avons vu, d'une faute de frappe ou d'une faute d'orthographe. Pour s'en prémunir, le contexte offert par la page sur laquelle apparaît le mot peut nous aider : par exemple, un mot de la même série dérivationnelle ou de la même famille morphologique sera un bon indice pour nous dire que ce mot a été créé sciemment. Il faut aussi faire confiance à sa propre intuition de locuteur natif ; comme le disent Hathout *et al.* (2008), il faut que le mot sonne naturel et plausible. Chaque mot qui sera accepté dans la base de données que j'ai créée et que je présente dans la partie suivante 2.3 a donc été vérifié manuellement.

De plus, quand les chiffres sont utilisés, ce n'est pas pour étudier la productivité d'un processus de formation, puisque ce genre de calculs ne peut se faire qu'à l'intérieur d'un corpus fini. Cependant, les chiffres donnés par Google, même approximatifs, peuvent nous donner une information sur la fréquence d'un mot par rapport à un autre¹³. Le Tableau 2.1 montre ainsi que, d'un point de vue comparatif, *italophone* est beaucoup plus fréquent que *italanophone*.

	Singulier	Pluriel	Total
italophone	69 800	41 700	111 500
italianophone	1 760	6720	8 480
italionophone	5	9	14
italanophone	6	1	7

Tableau 2.1 – Différence de fréquence, basée sur le nombre d'occurrences relevées par Google le 26 juin 2015, des mots construits désignant des locuteurs de l'italien

L'utilisation de la Toile, faite en connaissance de ses limites et de ses défauts et en prenant certaines précautions, est donc bénéfique pour la recherche en morphologie. Comme le disent Hathout *et al.* (2009, p. 278), « il n'y a pas lieu, à notre sens, d'opposer la recherche systématique d'attestations de formes nouvelles et l'intuition. (...) les recherches sur la Toile permettent de conférer le statut de fait véritable à ce qui, autrement, serait resté en l'état de simple conjecture ». La Toile est un moyen, parmi d'autres que sont l'intuition, les expérimentations sur des locuteurs, l'utilisation de corpus traditionnels, de dictionnaires, etc., de pouvoir formuler des hypothèses sur la formation des mots, de les valider ou de les infirmer. C'est pourquoi la base de données que

13. Lindsay et Aronoff (2015) proposent ainsi un calcul de fréquence basé sur ces nombres d'occurrences.

j'ai constituée s'appuie en grande partie sur des informations recueillies sur la Toile.

2.3 Création de la base de données NeoClassy

Afin d'étudier de façon systématique l'intégration d'éléments néoclassiques en français, illustrant l'intégration d'éléments allogènes dans une langue, j'ai constitué une base de données regroupant des mots formés au moyen d'éléments néoclassiques finaux. Après avoir argumenté les choix que j'ai faits dans la sélection des éléments étudiés, je présenterai chacun d'eux puis j'exposerai quelle a été ma manière de procéder pour leur collecte. Enfin, je décrirai la base de données que j'ai constituée et nommée NeoClassy.

2.3.1 Choix des éléments néoclassiques à étudier

Ce travail ayant des limites dans sa longueur ainsi que dans le temps qui lui est consacré, je n'ai pas pu étudier tous les éléments néoclassiques utilisés en français. J'ai dû opérer plusieurs choix dans leur sélection.

Le premier choix a été de privilégier des éléments très souvent impliqués dans des formations de lexèmes. En effet, comme le disent Guilbert *et al.* (1986, p.LIX) dans la préface du *Grand Larousse de la Langue Française*, « la motivation de la relation entre les éléments composants tend à disparaître quand une base verbale se révèle particulièrement productive. L'accumulation des termes construits à partir d'un même élément, comme *graphie*, *logie*, lui substitue un mécanisme créateur par analogie. La relation paradigmatique l'emporte sur la relation syntagmatique ». Même si le calcul de la fréquence de chacun des éléments néoclassiques pour déterminer les plus récurrents n'a pas été exécuté en amont de l'analyse, chaque élément choisi fait partie des plus fréquents par exemple dans le TLFi. Cependant, tous les éléments néoclassiques très fréquents n'ont pas été choisis pour l'analyse.

Je me suis ainsi intéressée uniquement aux éléments finaux. L'exemple (23) montre que les ENs peuvent être attachés aux mêmes bases que des suffixes prototypiques. Or les suffixes font partie, selon Corbin (2001), des affixes les plus prototypiques : « les préfixes sont bien, dans une langue comme le français, des affixes moins prototypiques que les suffixes parce qu'il y a une certaine perméabilité, en synchronie, ou une analogie de comportement entre la classe des préfixes et d'autres classes lexicales. » (Corbin, 2001, p.56). Il sera donc intéressant d'étudier en quoi certains ENs peuvent rentrer dans des schémas similaires à ceux impliquant des affixes prototypiques.

- (23) a. américaniste / américaniser / américanologue
 b. cellulaire / cellulose / cellulophage
 c. vinaire / vinasse / vinicole
 d. bactérien / bactériodie / bactéricide

De plus, même si les éléments grecs sont les plus fréquents dans la formation de composés néoclassiques (une recherche sur le TLFi donne 400 éléments formants tirés du grec pour 94 tirés du latin), j'ai tout de même souhaité représenter les deux langues classiques. Parmi les onze éléments étudiés, huit sont d'origine grecque et trois d'origine latine.

La liste des ENs choisis est donnée en (24) :

- (24) -logue (gr.)
 -logie (gr.)
 -crate (gr.)
 -cratie (gr.)
 -phobe (gr.)
 -phobie (gr.)
 -phone (gr.)
 -phage (gr.)
 -vore (lat.)
 -cide (lat.)
 -cole (lat.)

Comme on peut le déduire de cette liste, un élément [+ie]¹⁴ n'est pas considéré comme étant construit sur un élément [-ie] mais bien étant un élément à part entière. Le suffixe *-ie* est le descendant du suffixe grec $-\iota\alpha$ qui construisait des noms abstraits sur des noms et adjectifs [+humains]. Il se retrouve en français dans des paires telles que celles données en (25) :

- (25) a. fou_A ↔ folie_N
 b. boucher_N ↔ boucherie_N

Cependant, plusieurs faits nous empêchent de considérer les lexèmes en *-logie* comme étant synchroniquement dérivés de ceux en *-logue* ou ceux en *-cratie* de ceux en *-crate*. Tout d'abord, comme nous le verrons dans la partie 3.2, chaque lexème [+ie] n'a pas de correspondant [-ie] et vice-versa. Par exemple, le lexème *macdocratie* (62-b) est construit selon un schéma [X cratie]_N dé-

14. Cette notation me permet de regrouper les lexèmes formés avec les éléments *-cratie*, *-phobie* et *-logie* qui forment des noms de qualité [-humain], sans présupposer que la séquence *-ie* est un suffixe.

signant une omniprésence de X ; ce schéma n'a pas de schéma $[X \text{ crate}]_N$ correspondant.

- (26) a. macdocratie 'omniprésence des restaurants Macdonald's'
b. *macdocrate

De plus, si l'on considérait par exemple les noms en *-logie* comme étant construits sur les noms en *-logue*, non seulement ceci impliquerait une modification phonologique de la consonne précédent le suffixe ($/g/ > /ʒ/$)¹⁵, mais cela irait aussi à l'encontre des autres schémas construisant des noms de spécialistes, par exemple au moyen des suffixes *-iste* ou *-ien*, pour lesquels le nom de spécialiste est construit sur celui de la spécialité :

- (27) violon > violoniste
musique > musicien
cardiologie < cardiologue
biologie > biologiste

Nous voyons bien dans cet exemple que lorsque le nom de spécialiste est construit grâce à la séquence *-logiste*, la relation entre le lexème en *-logie* et celui en *-logiste* se fait dans le même sens que celles en *-iste* ou *-ien*. Un élément comme *-logie* ne peut donc pas être considéré comme construit sur l'élément *-logue*. Il faut plutôt considérer ces deux éléments comme faisant partie d'un réseau complexe, qui comprendrait également l'élément *-logique*¹⁶. En effet, les lexèmes en *-logique* ne sont pas forcément construits sur des lexèmes en *-logie* ou en *-logue*, ils peuvent également, comme nous le verrons dans la partie 4.3.1, être construits directement sur la base : en (28), *météorologique* n'est pas construit sur *météorologie* ni sur *météorologue* mais directement sur la base TEMPS représentée ici par le thème supplétif *météo-*, et signifie donc 'relatif au temps' :

- (28) On est tellement habitués aux tempêtes métaphoriques (financières, médiatiques) qu'on est dépourvus face à une tempête **météorologique**. (http://www.liberation.fr/chroniques/2009/01/31/quand-la-tempete-tempete_306669)

Les lexèmes construits grâce aux éléments *-logie*, *-logue* et *-logique* sont certes

15. De même, une modification phonologique ($/t/ > /s/$) serait à prévoir pour le passage de *-crate* à *-cratie*.

16. Cet élément n'a pas fait l'objet d'une recherche systématique. Pour les besoins de l'analyse effectuée en 4.3.1, ont été recherchés les lexèmes en *-logique* construits grâce aux mêmes C1 que ceux en *-logie* et ceux en *-logue*, ainsi que ceux présents dans les dictionnaires. Cette liste de 1042 lexèmes n'a pas été annotée et ne fait donc pas partie de la base de données NeoClassy.

reliés et la nature de ces relations sera à définir, mais on ne peut pas dire que les uns servent de base pour dériver les autres. Chacune de ces séquences est donc analysée comme un élément à part entière.

Voyons maintenant les choix qui m'ont poussée à sélectionner les EN de la liste (24). Les premiers éléments de la liste, *-logue* et *-logie* ont été choisis pour leur très forte fréquence. En effet, si l'on regarde la fréquence des mots-types se terminant par la séquence *-logie* dans le corpus FrWac¹⁷, comparée à celle d'autres éléments néoclassiques également très courants, on remarque qu'elle est nettement supérieure (voir Tableau 2.2). Il est donc intéressant d'observer la dynamique de schémas impliquant des éléments néoclassiques particulièrement disponibles.

Mots se terminant par la séquence :	Fréquences (types) sur FrWac
<i>-logie</i>	5204
<i>-graphie</i>	1915
<i>-phobie</i>	829
<i>-cratie</i>	440

Tableau 2.2 – Comparaison des fréquences de différents éléments néoclassiques dans le corpus FrWac

L'élément *-cratie* présente l'intérêt particulier d'entrer dans des schémas qui ont deux mots leaders bien déterminés, que ce soit historiquement et en termes de fréquence : *démocratie* et *aristocratie*, comme nous le verrons plus en détail dans la partie 3.2. Le sens de ces mots leaders est enseigné dès l'école, en particulier celui de *démocratie*, et l'étude de ces schémas semble pertinente pour mettre au jour des relations particulières entre des schémas présents dans le lexique mental des locuteurs.

Pour *-logie* et *-cratie*, j'ai choisi de recenser également les mots en *-logue* et *-crate* correspondant qui me serviront à analyser les relations qui existent entre des schémas impliquant des éléments néoclassiques issus de la même racine classique.

Le même intérêt m'a poussée à étudier *-phobe* et *-phobie* avec la différence que *-phobie* peut également être employé comme lexème autonome : quelles relations s'opèrent entre les schémas impliquant d'une part un segment toujours attaché et d'autre part un segment pouvant être autonome ? La même analyse aurait pu être conduite sur la paire *-mane/-manie* mais cette paire semble moins présente (829 mots en *-phobie* sur FrWac contre 381 en *-manie*).

17. http://nl.ijs.si/noske/wacs.cgi/corp_info?corpname=frwac

Concernant l'élément *-phone*, l'intérêt était d'observer la relation qu'il existait entre un élément néoclassique et des bases qui ne sont pas que rarement classiques, en particulier lorsque *-phone* sert à la dénomination de noms de locuteurs de langues.

Les éléments *-phage* et *-vore* ont l'avantage de correspondre à des lexèmes ayant le même sens mais originaires de langues différentes. Nous pourrions donc avoir une comparaison pertinente entre des lexèmes construits sur des éléments grecs d'un côté et latins de l'autre.

L'élément *-cide* quant à lui a été choisi en fonction de la comparaison interlingue qu'il pouvait permettre : en effet, si les éléments latins *-cida* et *-cidium* ont donné lieu à un syncrétisme en français, où les deux formes se sont phonologiquement réduites à *-cide*, ce n'est pas le cas dans une langue comme l'italien, où deux formes sont encore utilisées, *-cida* et *-cidio*. Ces éléments feront en particulier l'objet d'une étude sémantique dans la partie 4.1.1, qui reprend, entre autres, les travaux présentés dans Lasserre et Montermini (2012).

Enfin, l'élément *-cole* a été choisi parce qu'à mon sens, il fait partie de ces éléments dont un locuteur français n'ayant jamais pratiqué le latin pourrait difficilement retrouver le sens originel. De plus, il est le continuateur du lexème latin COLEO qui était déjà polysémique, signifiant 'habiter', 'cultiver' ou 'honorer' et dont chacun de ces sens se retrouve dans des lexèmes en *-cole* du français. Il sera intéressant de voir les relations que ces schémas différents partagent.

L'ensemble des éléments listés en (24) permet ainsi d'observer un large éventail de phénomènes qui apparaissent dans des schémas impliquant des éléments néoclassiques récurrents ainsi que dans les relations qui existent entre ces schémas. Cependant, la volonté n'est pas ici d'être exhaustive dans l'analyse des faits relatifs aux éléments néoclassiques mais plutôt de proposer une étude de cas qui constituerait un préliminaire pour une analyse plus vaste.

2.3.2 Présentation des éléments néoclassiques choisis

Cette partie propose une présentation sommaire de chaque élément choisi : le sens qu'ils avaient à l'origine, la date d'introduction en français des premiers mots les comprenant, etc. Ce n'est qu'après avoir présenté la base de données que j'ai constituée, que je reviendrai plus longuement sur les questions de sens (partie 3.2), à partir d'observations que j'aurai pu alors effectuer sur les mots récoltés.

2.3.2.1 *-logie, -logue*

Les éléments *-logie* et *-logue* sont les descendants de deux lexèmes grecs : le nom λόγος, ‘discours’, d’une part, et le verbe λέγω, ‘choisir, recueillir’ d’autre part, représentés dans les composés grecs par la séquence finale *-λογία*. Cette séquence était déjà très présente en grec ancien : Perseus, ressource en ligne regroupant plusieurs dictionnaires¹⁸, recense 170 mots se terminant en *-λογία*. Cottez (1980) dénombre six types de constructions *Xlogie* et trois types de constructions *Xlogue* (voir Tableau 2.3).

Les premiers mots en *-logie* et *-logue* qui sont apparus en français, par emprunt au grec ancien, sont *étymologie* (1175), *généalogie* (fin du XII^{ème} siècle), *prologue* (1225), *astrologie* (1260), *théologie* (1270), *tropologie* (1295), *épi-logue* (1339-1348) et *astrologue* (14^{ème} siècle).

Le nom λόγος, ‘discours’ peut également être représenté à l’initiale de mots grecs (λογογραφία) dont certains ont été introduits en français (*logographie*). D’autres sont des créations modernes (*logopathie*). Lorsque *logo-* est à l’initiale, il n’aura que le sens de ‘mot’ ou ‘discours’ : la vitalité de *logo-* à l’initiale de mots construits est bien moindre comparé à celle de *-logue* ou *-logie* à la finale.

2.3.2.2 *-cratie, -crate*

Les éléments *-cratie* et *-crate* viennent du nom grec κράτος, ‘pouvoir’, représenté par *-κρατης* en composition (*ισοκρατής* ‘qui possède un pouvoir égal’). Les deux premiers mots en *-cratie* introduits en France sont *aristocratie* (1361) et *démocratie* (1370). Cottez (1980) distingue d’une part les lexèmes *Xcratie*, signifiant ‘gouvernement, pouvoir exercé par X’, pour lesquels les lexèmes *Xcrate* correspondant désignent des partisans de ce gouvernement ; les mots *aristocrate* et *démocrate* n’existaient pas en grec ancien et sont des créations modernes. Rey (1992) note ainsi qu’*aristocrate*, s’il a été introduit en français à la Renaissance, en 1550, est resté inusité jusqu’à la Révolution, ce que sembler montrer cette citation reportée par Brunot (1927b, p.113) :

« Jouant sur le mot *aristocratie*, ils disaient de M. d’Éprémesnil que c’était un “aristocrate”, d’un autre qu’il était “aristocrâne”, d’un troisième que c’était un “aristocroc”, d’un quatrième que c’était un “aristocruche” ». Monjoies, *Histoire de la Révolution de France*, 2^{ème} partie, p.154

18. Perseus (<http://www.perseus.tufts.edu/>) inclut, pour le grec ancien, le *Liddell-Scott-Jones Greek-English Lexicon* et sa version abrégée *An Intermediate Greek-English Lexicon* (Middle Liddell) et pour le latin, le dictionnaire *A Latin Dictionary* (Lewis and Short) et *An elementary Latin dictionary* (Charlton T. Lewis).

	<i>Xlogie</i>	<i>Xlogue</i>
1.	‘manière de parler, forme de discours caractérisée par X’ ex : néologie	
2.	‘mode d’expression, type de langage, utilisant les moyens désignés par X’ ex : dactylologie	
3.	‘étude (plus ou moins) scientifique de X’ ex : psychologie	‘savant, spécialiste de X’ ex : psychologue
4.	‘ensemble d’œuvres dénombré par X’ ex : tétralogie	
5.	‘recueil, collection de X’ ex : anthologie	
6.	noms abstraits se rapportant à → ex : homologie	‘qui est dans un certain rapport (exprimé par X)’ ex : homologue
7.		‘discours, partie de discours, produit dans les conditions exprimées par X’ ex : dialogue

Tableau 2.3 – Types de constructions *Xlogie* et *Xlogue* selon Cottez (1980)

Bauer (1983, p.275) fait l'hypothèse que le sens 'partisan d'un système politique' a dérivé en 'membre d'un groupe dirigeant ou d'une élite', sens recouvert par *bureaucrat*. Il note d'ailleurs (p.276) que, en anglais du moins, sous l'influence de *bureaucrat*, un autre type est apparu, qui affiche le sens de 'bureaucrate caractérisé par X' (par exemple *Eurocrat*).

Dans la partie 3.2, nous verrons que, en effet, deux types sémantiques principaux se dégagent pour les constructions *Xcratie* et *Xcrate*, 'pouvoir de X' et 'partisan de ce pouvoir' d'un côté, et 'groupe de X auto-suffisant' et 'membre de ce groupe' de l'autre.

2.3.2.3 -*phobe*, -*phobie*

Les éléments -*phobe* et -*phobie* existaient en grec ancien sous les formes -*φοβος* et -*φοβία*, toutes deux représentant dans les composés le nom *φόβος*, 'peur panique'. Les adjectifs en *X-φοβος* désignaient 'qui a peur de X' et le nom correspondant *X-φοβία* la 'peur ou horreur de X'. Cependant, ces derniers étaient peu nombreux, Perseus par exemple ne recense que deux *X-φοβία*, *ἀφοβία* 'qualité de ne pas avoir peur, courage' et *ὕδροφοβία* 'peur de l'eau', le second seulement étant un composé. Par ailleurs, Perseus dénombre 20 *X-φοβος*. Contrairement au grec, où **φοβία* n'existait pas comme élément autonome mais *φόβος* était, lui, un lexème autonome, le français a deux éléments attachés, -*phobe* et -*phobie* et un élément autonome, *phobie*. Cet élément a acquis son autonomie vers 1880, selon le *Dictionnaire Historique de la langue française* (DHLF) (Rey, 1992). Cottez (1980) distingue deux éléments -*phobe* (et -*phobie* correspondants) : le premier désigne une peur panique, le second une hostilité. Bouffartigue et Delrieu (1996a, p. 50) distinguent, en plus de ces deux sens, un troisième : une aversion naturelle d'un être ou d'une matière pour un autre être ou une autre matière. Le premier mot en -*phobie* introduit en français recouvre d'ailleurs ce sens, *hydrophobie* (1314).

2.3.2.4 -*phone*

L'élément -*phone* vient du nom grec *φωνή*, 'son, voix, parole, langage'. Dès le grec ancien, de nombreux mots complexes affichaient -*φωνος* comme deuxième constituant (Perseus en dénombre 92). Beaucoup sont des composés dans lesquels le premier constituant qualifie un type de voix (par exemple *λεπτόφωνος*, 'qui a une faible voix'). *φωνή* est aussi, plus rarement, représenté dans le constituant initial de mots complexes (*φωνομαξία*, 'dispute au sujet de mots'). En français, *phono-* peut également se rencontrer à l'initiale de mots

complexes, mais a toujours le sens de ‘son’ : *phonogramme*, *phonolithe*, *phonothécaire*, *phonologie*,... Concernant l’élément final *-phone*, Cottez (1980) recense trois types de mots qu’il peut construire, chacun affichant un des sens différents du nom grec originel. Tout d’abord, des mots dans lesquels *Xphone* désigne ce ‘qui a une voix, un langage, caractérisé par X’ (*aphone*, *homophone*). Dans ce type seraient compris les noms de locuteurs (*germanophone*, *francophone*), déjà présents en grec ancien (*βαρβαρόφωνος*, ‘qui parle une langue étrangère’). Le second type, toujours selon Cottez, regroupe les noms d’instruments de musique, dont le premier élément caractérise le type de son produit, ou le moyen de produire le son (*xylophone*, *saxophone*). Enfin, Cottez dégage un troisième type, où *Xphone* désigne un appareil de traitement de son, dans lesquels X indique par exemple l’intensité ou la transmission du son (*microphone*, *téléphone*). Selon Le Grand Robert, le premier mot à avoir été introduit en français semble être *téléphone*, dans une traduction de l’allemand donnée ici :

« Dans la troisième addition, Mr Huth recommande les porte-voix pour transmettre des nouvelles à des grandes distances au moyen de stations intermédiaires ; un tel téléphone pourrait être utile, surtout quand, à cause du brouillard, etc..., on ne peut pas se servir des télégraphes¹⁹. » E. F. Chladni, *Traité d’acoustique*, p. 289 (trad. de l’allemand, 1809).

Viendront ensuite *homophone* (1822), *aérophone* (1834) et *aphone* (1834). Cependant, la plupart seront créés plus tard. En effet, Rey (1992) note dans le DHLF que les lexèmes *Xphone* ont connu une très grande vitalité au début du XX^{ème} siècle, suite aux forts progrès en acoustique et en retransmission de sons. De plus, « la plupart des mots ainsi formés sont sortis d’usage, mais connus historiquement (*gramophone*...) ».

Certains lexèmes *Xphone* sont en relation avec des lexèmes *Xphonie*. Ainsi, si *Xphonie* est en relation avec un locuteur de langue *Xphone*, le terme désignera l’ensemble des locuteurs de la langue (par exemple *francophonie*).

2.3.2.5 *-phage*

L’élément *-phage* vient du verbe grec φαγειν, ‘manger’, déjà représenté en grec ancien dans des composés par la séquence *-φαγος* (voir partie 1.2.1). En grec ancien, cet élément verbal lié était déjà bien présent, Perseus dénombre 139 mots se terminant par *-φαγος*. Par contre, un seul mot commençant par

19. Le GR note ici : « Les télégraphes mentionnés ici sont les télégraphes aériens du type inventé par Chappe ».

l'élément φαγο- est recensé dans Perseus : φαγολοίδορος, 'avaleur d'insultes'. Cette proportion se retrouve en français, puisque dans cette langue aussi, très peu de mots sont formés avec l'élément *phago-* à l'initiale : le TLFi recense uniquement *phogocyte* 'cellule qui a le pouvoir de détruire, en les digérant, des particules étrangères' et le dérivé *phagocytose* qui désigne le mécanisme par lequel des cellules ont un tel comportement, le Grand Robert les mêmes, plus *phagocytaire*, l'adjectif relatif au phagocyte et à la phagocytose et le verbe *phagocyter* 'détruire par phagocytose', toujours de la même famille, ainsi que l'adjectif *phagotrophe* qui signifie 'qui se nourrit de proies introduites dans la bouche'. Cottez (1980) note également le nom *phagomanie*. Le premier mot construit avec l'élément *-phage* qui apparaît en français semble être *anthropophage* daté en 1375 par le TLFi et le Grand Robert.

Bouffartigue et Delrieu (1996a, p.115) notent que les mots *Xphage* sont des noms ou des adjectifs qui qualifient ou désignent des êtres en fonction de ce qu'ils mangent, ce qui est désigné par le premier élément, le sens de 'manger' pouvant être présent de façon métaphorique, par exemple dans *sarcophage*, emprunté au grec, qui signifie littéralement 'qui détruit les chairs' mais désigne un cercueil. Bouffartigue et Delrieu (1996a) dénombrent également ce qu'ils appellent deux intrus, *macrophage*, où *macro-* ne désigne pas ce qui est mangé mais celui qui mange, puisque le mot signifie 'une grosse cellule qui englobe et détruit des particules cellulaires', et *æsophage* qui signifie littéralement 'qui porte à manger'.

Cottez (1980), quant à lui, désigne deux éléments *-phage*, venant tous deux de l'élément lié grec *-φαγος* : le premier signifie 'qui mange' et le second 'qui détruit', qui est le *-phage* qu'on retrouve dans *sarcophage*.

Les mots que j'ai pu collecter dans ma base de données affichent en fait plusieurs sens, comme je l'ai montré dans la figure 2.4. Je reviendrai sur ces différents sens dans la partie 3.2.

2.3.2.6 *-vore*

A l'élément grec *-phage* fait pendant l'élément latin *-vore*, descendant du verbe *voro*, 'avaler, dévorer, engloutir'. En français, il n'est présent en composition que comme second élément. En latin, les deux seuls noms recensés par Perseus ayant *-vorus* comme second élément sont *carnivorus* et *omnivorus*. *Carnivore* est d'ailleurs le premier mot qui est apparu en français, en 1536. Cottez (1980), s'appuyant sur la polysémie du lexème originel, distingue deux éléments *-vore* : le premier construit des noms d'animaux par le type d'alimentation qui les caractérise, alimentation désignée par le premier élément du mot construit (*carnivore*, *granivore*, *frugivore*). Le second

s'appuie sur le sens 'avalier, détruire' et désigne un élément 'qui avale, qui détruit' ce qui est représenté par le premier constituant (*fumivore, ignivore*).

2.3.2.7 -cide

L'élément *-cide* est le continuateur du verbe latin *caedere*, 'frapper, abattre, tuer, fendre', déjà présent dans des composés latins, comme second élément verbal; *Xcida* signifiait une 'personne qui tue X' et *Xcidium* le 'meurtre de X'. Perseus recense 26 mots en *-cida* et 16 en *-cidium*. Le sens 'fendre' était également présent, par exemple *lignicida*, 'lit. qui fend le bois, bûcheron'. Si en français, quelques mots recouvrent ce sens de 'fendre' (le TLFi en donne quatre : *denticide, coricide, septicide* et *valvicide*), la plupart des mots construits avec l'élément *-cide* mettent en jeu le sens de 'tuer'. En français, les continuateurs de *-cida* et *-cidium* se sont réduits en une seule forme, *-cide*, désignant et l'auteur du meurtre, et le procès en lui-même. Nous verrons dans les parties 3.2 ou 4.1.1 que dans d'autres langues, l'italien par exemple, les deux formes sont toujours distinctes phonologiquement. Sur ce schéma, le lexème *homicide* semble être le premier à avoir été introduit en français (en 1150). Un autre sens a émergé, celui de 'tuer, anéantir, quelque chose de nuisible (une plante, un animal, etc.)'. Le premier lexème créé selon ce schéma semble être *germicide* (1784 selon le Grand Robert). Enfin, plus tardivement, un autre sens a émergé, celui de 'tuerie à grande échelle', quand le mot *génocide* a été créé à l'ONU, en 1954, d'abord en anglais puis le terme s'est propagé dans de nombreuses langues dont le français.

2.3.2.8 -cole

L'élément *-cole* vient du verbe latin *colo* qui avait plusieurs sens 'cultiver, habiter, honorer'. Ce verbe était représenté dans les composés latins par la séquence *-cola* qui portait les trois sens du verbe : *agricola* 'qui cultive les champs, agriculteur', *silvicola* 'habitant de la forêt' ou *Marticola* 'fidèle de Mars'. Ces trois sens se retrouvent dans les mots construits avec l'élément *-cole* en français, même si le sens de 'honorer' est moins représenté, malgré le fait que, selon le TLFi, le premier terme en *-cole* introduit en français soit *christicole* (XV^{ème} siècle). Cependant, si *-cola* formait des noms en latin, l'élément français *-cole* construit aujourd'hui des adjectifs de relation aux noms en *-culture* et en *-culteur*.

2.3.3 Recueil des données et création de la base de données NeoClassy

La récolte des lexèmes formés avec ces onze éléments néoclassiques finaux a permis la création de la base de données NeoClassy. Ce nom a été choisi car il intègre deux dimensions : d'une part, les lexèmes présents dans la base sont traditionnellement appelés *néoclassiques*, d'autre part, ils sont créés selon un processus utilisé dans un registre de langue recherché, soit dans le domaine scientifique, soit avec la volonté d'imiter la langue scientifique, et ainsi de donner un caractère *classy* à ces mots. Avant de présenter la base et son contenu, quelques précisions sur la notion de lexicalisation me semblent nécessaires.

2.3.3.1 Sur la question de la lexicalisation

Une grande proportion de lexèmes étudiés dans cette thèse sont des néologismes, voire des hapax ou des occasionnalismes. Ce sont des créations de l'instant, qui ont pu, par exemple par leur emploi dans un site Internet auquel j'ai pu accéder, être relevées et ainsi rentrer dans la base de données. Cependant, comme je l'ai déjà amorcé dans la partie 2.1 et que nous reverrons plus précisément dans la partie 3.2, le lexique engendre le lexique, c'est-à-dire que les lexèmes créés le sont en fonction des lexèmes déjà présents dans le lexique des locuteurs. Ceci étant dit, quels sont les lexèmes qui sont stockés dans le lexique mental des locuteurs, quels sont ceux qui ne le sont pas ? Selon Fradin (2003, p.233), « la lexicalisation est le phénomène par quoi une expression intègre le lexique », c'est la « dernière étape d'un processus qui se marque par le fait qu'une expression devient unité lexicale en acquérant au moins une propriété idiosyncrasique » (Fradin, 2003, p.220). Bauer (1983) propose que les étapes dans la lexicalisation d'un mot soient d'abord une création de circonstance, qui est ensuite institutionnalisée, la lexicalisation demeurant effectivement la dernière étape du processus.

Cependant, comment déterminer si un mot est lexicalisé ou s'il ne l'est pas ? Selon Montermini (2010a, p.88), la lexicalisation « est une dimension clairement externe au fonctionnement de la langue et en tout cas difficilement mesurable avec précision ». Fradin (2003, p.229) dit, en substance, la même chose : « le processus de lexicalisation n'est pas du ressort de la grammaire, ce qui permet justement le recyclage des expressions en dehors de la composante où elles sont engendrées ». En effet, différents facteurs poussent à la lexicalisation d'une unité : des facteurs sociologiques, pragmatiques, etc. Fradin (2003, p.221) fait « l'hypothèse qu'une expression a d'autant plus de chance d'être lexicalisée qu'elle répond à un besoin de dénomination et

s'avère adéquate sémiotiquement ». La morphologie n'a donc pas la capacité de traiter de la lexicalisation des unités puisqu'elle « propose des possibles mais n'a pas le contrôle de l'existant » (Fradin, 2003, p.228).

Donc comment déterminer si un mot est existant, c'est-à-dire, présent dans le lexique des locuteurs, et va ainsi pouvoir servir de modèle aux nouvelles créations ? Corbin (1997) propose que « la notion de mot existant implique une instance d'actualisation qui permette d'authentifier le mot. Cette instance peut être, au choix, les dictionnaires, les productions langagières primaires ou encore le stock lexical que croit connaître le linguiste en situation d'effectuer une classification de ce point de vue. Par rapport à cette instance, l'existence d'un mot est une notion discrète. Pour des raisons de commodité pratique (seuls les dictionnaires offrent un accès objectivable aux unités actualisées, même s'ils ne peuvent répertorier que celles qui figurent dans les documents à partir desquels ils sont constitués), j'entendrai ici par mot existant un mot répertorié dans l'un ou l'autre des grands dictionnaires actuels de référence, à savoir, pour l'essentiel, le *Grand Robert de la langue française* et le *Trésor de la langue française*. J'appellerai lexique actuel l'ensemble – nécessairement partiellement aléatoire des mots existants ainsi définis. »

Ici, Danielle Corbin montre bien que la définition des frontières de l'ensemble que constitue le lexique sont forcément aléatoires : elles sont en constant mouvement, varient selon les locuteurs ; même les dictionnaires n'ont pas le même nombre d'entrées, font des choix par exemple dans l'acceptation de nouveaux mots au sein d'une nouvelle édition. Cependant, comme il a bien fallu déterminer quels lexèmes sont lexicalisés, existants, présents dans le lexique, quels lexèmes servent de modèle aux autres, je suivrai Corbin et je m'appuierai sur leur présence ou leur absence dans deux dictionnaires : le *Grand Robert* (version en ligne) et le *TLFi*. Mais comme cet ensemble ne définit qu'aléatoirement les mots réellement lexicalisés, j'emploierai plutôt, pour décrire cet ensemble, le terme de dictionnarisé, emprunté à Michel Roché. Si un mot est présent dans l'un des deux dictionnaires de référence, je dirai donc qu'il est dictionnarisé, les autres seront non dictionnarisés.

2.3.3.2 NeoClassy : présentation

La base de données NeoClassy est constituée de mots extraits d'une part des dictionnaires de référence le *Grand Robert* et le *TLFi*, d'autre part de mots récoltés sur la Toile. Si recueillir les lexèmes nécessaires présents dans les dictionnaires en ligne a été aisé, récolter ceux présents sur la Toile a présenté davantage de difficultés : il n'est en effet pas possible de questionner des moteurs de recherche tels que Google selon la finale d'un mot. Des listes de mots se terminant par les éléments néoclassiques choisis ont cependant pu

été extraites grâce au vocabulaire de la ressource distribuée par le Linguistic Data Consortium (LDC) (<https://catalog.ldc.upenn.edu/LDC2006T13>) et la base FrWac (http://nl.ijs.si/noske/wacs.cgi/wordlist_form?corpname=frwac)²⁰. Chacun des mots des listes ainsi obtenues a été vérifié sur Google : non seulement sa présence a été confirmée ou infirmée, mais également sa faisabilité : comme je l’ai expliqué dans la partie 2.2, afin d’éviter de collecter des mots résultant de fautes de frappe par exemple, une observation des contextes de chaque mot de la liste a été effectuée. Par ailleurs, pour les lexèmes formés avec certains ENs finaux, des estimations ont été faites sur des bases potentielles. Ainsi, pour les lexèmes *Xphone* désignant des locuteurs de langues, une liste de langues, trouvée sur Wikipédia, a été testée : chaque langue, suivie de l’élément *-phone* a été recherchée sur Google, en essayant différentes configurations (présence ou absence de la voyelle de liaison, allomorphie thématique, etc.). La même démarche a été entreprise pour les lexèmes *Xcide* : comme nous le verrons dans les chapitres suivants, ces lexèmes peuvent désigner soit le meurtre d’une personne, en particulier un membre de la famille (*fratricide*), soit une tuerie à grande échelle (*génocide*), soit un instrument servant à la destruction d’animaux ou de plantes nuisibles (*insecticide*). Différents constituants pouvant précéder l’élément *-cide* ont donc été testés : des noms de membres de la famille (*belle-mèricide*), des noms de peuples opprimés (*tutsicide*) ou des noms de nuisibles (*miticide*).

Combiner des données dictionnaires et des données qui sont soit en passe d’être dictionnarisées, soit qui n’ont pas vocation à l’être mais ont été créés pour répondre à un besoin ponctuel (on parle alors d’occasionnalisme), est nécessaire. En effet, selon Roché (2011b, p.27), « ou bien on veut décrire la créativité actuelle, le lexique tel qu’il naît, et il faut ne retenir que les néologismes – sélectionner, dans les dictionnaires, les formations les plus récentes, les rechercher sur la Toile et dans les journaux. Ou bien on veut décrire le lexique construit existant, le lexique tel qu’il est, tel qu’il est perçu et utilisé par les locuteurs, et il faut prendre tous les mots construits qui sont restés motivés, quelle que soit leur origine et leur ancienneté ». Comme le but est ici de décrire le lexique tel qu’il est, et ainsi tel qu’il est utilisé en vue de la création de nouveaux mots, tous les lexèmes contenant les éléments néoclassiques finaux choisis présents dans les dictionnaires ou ayant pu être approchés sur la Toile ont été sélectionnés, pourvu qu’ils soient motivés, c’est-à-dire que le mot entre dans un schéma constructionnel et entretient des relations avec les autres mots du lexique construits selon le même schéma ou des schémas dérivés. C’est pourquoi même des termes empruntés au grec dans leur ensemble et non pas formés en français, pourvu qu’ils restent motivés,

20. Je remercie infiniment Nabil Hathout pour m’avoir aidée dans ce travail.

ont été acceptés dans la base de données, contrairement à ce qu'avancent Amiot et Dal (2008) qui excluent de leurs investigations des termes comme *philosophe* ou *quadrilatère*, sur la base de l'hypothèse qu'ils n'ont pas été créés selon les règles de formation effectives dans la langue moderne. Or la formation des lexèmes telle que je la conçois ne se fait pas au moyen de règles, mais grâce à des schémas dégagés à partir de ce qui est existant. Par exemple, la présence dans le lexique des locuteurs de termes comme *aristocratie* et *démocratie*, même s'ils ont été créés en grec puis empruntés à cette langue par le français, permet aux locuteurs de former des termes comme *mafiaocratie* ou *cathocratie* car ils engendrent des schémas dont le sens serait pour le premier 'pouvoir de X', pour le second 'groupe auto-suffisant constitué de X', ce sens d'*aristocratie* n'étant d'ailleurs pas celui du lexème grec ἀριστοκρατία, 'pouvoir des meilleurs'. Au contraire, un lexème comme *æsoophage* n'a pas été sélectionné, car il n'est, selon moi, plus motivé aujourd'hui, il n'entre plus dans des schémas constructionnels faisant intervenir l'élément néoclassique *-phage*.

La base de données NeoClassy contient 4499 lexèmes (cf. Tableau 2.4).

La base sera bientôt accessible en ligne. En attendant, la liste de tous les lexèmes qu'elle contient est donnée en annexe. Dans sa version en ligne, chaque mot construit est donné accompagné des informations suivantes :

- le mot construit ;
- sa catégorie lexicale ;
- s'il est dictionnarisé ;
- sa date de création (si elle est identifiable) ;
- la langue dans laquelle il a été créé (il peut s'agir d'une formation française ou d'un emprunt au grec par exemple) ;
- le nombre d'occurrences relevées sur Google ainsi que la date à laquelle la recherche a été effectuée ;
- le constituant initial ainsi que son origine ;
- le constituant final ;
- si le mot n'est pas dictionnarisé, un ou deux exemples dans lesquels le mot est utilisé, ainsi que l'adresse du site duquel l'exemple est extrait.

Les informations sémantiques ne sont donc pas indiquées à l'intérieur de la base de données, aucune glose des mots construits n'est fournie. En effet, si l'on part de l'hypothèse, que je développerai plus loin, que le sens des mots se construit en contexte et que donc les mots construits, en particulier les occasionnalismes, n'ont pas de sens *per se*, il serait difficile voire arbitraire de leur attribuer dans la base de données un sens en particulier. La base NeoClassy

	Dictionarisés	Non dictionarisés	Total
<i>X-logie</i>	535 52.97%	475 47.03%	1010
<i>X-logue</i>	227 28.16%	579 71.84%	806
<i>X-cratie</i>	30 11.54%	230 88.46%	260
<i>X-crate</i>	18 10.98%	146 89.02%	164
<i>X-phobie</i>	73 13.80%	456 86.20%	529
<i>X-phobe</i>	43 7.72%	514 92.28%	557
<i>X-phone</i>	68 20.92%	257 79.08%	325
<i>X-phage</i>	79 46.20%	92 53.80%	171
<i>X-vore</i>	41 23.30%	135 76.70%	176
<i>X-cide</i>	112 36.48%	195 63.52%	307
<i>X-cole</i>	152 78.35%	42 21.65%	194

Tableau 2.4 – Nombre de lexèmes présents dans la base, selon leur dictionnarité

devra être étendue ensuite à des constructions impliquant d'autres éléments néoclassiques. Cependant, telle qu'elle est actuellement, elle permet de réaliser des analyses précises sur le fonctionnement des constructions relevées, ce que nous allons voir dans les chapitres suivants.

Bilan. Les éléments néoclassiques ne forment pas une classe d'objets homogène. En effet, sur la seule base d'un critère étymologique, nous ne pouvons pas rendre compte de tous ces éléments de la même manière : certains ont par exemple un sens pleinement lexical alors que d'autres se rencontrent dans des distributions similaires à celles d'affixes. Le cadre théorique de la Morphologie Constructionnelle permet de rendre compte de ces différents types de comportements : cette théorie postule des schémas de constructions qui sont stockés dans le lexique et qui sont d'une part la généralisation de lexèmes existants, d'autre part une recette pour la formation de nouveaux lexèmes. Dans ces schémas, certains éléments peuvent être variables, et d'autres spécifiés.

Ces schémas peuvent afficher une polysémie régulière. Ainsi, le sens d'un schéma peut être étendu, par le biais de métaphores ou de métonymies : les schémas sont alors reliés et hiérarchisés. Parmi les lexèmes instanciant un schéma, un ou plusieurs peuvent servir de mot leader : ils fonctionnent comme un pôle d'attraction en fonction duquel les lexèmes créés selon le même schéma construisent leur sens, par analogie. Ces mots leader peuvent jouer le rôle d'initiateurs d'un modèle de formation, ou d'épicentre synchronique autour duquel gravitent les autres lexèmes instanciateurs du schéma. Afin d'observer la formation de lexèmes selon des modèles néoclassiques, j'ai choisi d'analyser des mots issus aussi bien de dictionnaires que de la Toile, moyen qui a certes de nombreux inconvénients mais qui a l'énorme avantage d'être le témoin de la vitalité d'un processus de formation. Je me suis attachée à recueillir les lexèmes formés grâce à onze éléments néoclassiques finaux intervenant dans de longues séries de mots. Cela a donné lieu à la création d'une base de données, NeoClassy, qui sera mise à la disposition de la communauté.

Grâce à cette base de données, je vais maintenant pouvoir tirer des observations tant formelles que sémantiques sur ces constructions néoclassiques. Le chapitre suivant sera en effet l'occasion d'analyser le comportement des lexèmes de NeoClassy afin de tirer des généralisations sur les contraintes pesant sur leur formation.

Chapitre 3

Description des données

Sommaire

3.1	Contraintes en jeu	120
3.1.1	Le constituant initial	121
3.1.2	La voyelle de liaison	138
3.1.3	Taille des mots construits	155
3.2	Organisation du lexique	162
3.2.1	Description sémantique	163
3.2.1.1	Du discours et des spécialistes	163
3.2.1.2	Du pouvoir et des coalitions	171
3.2.1.3	De la peur à la haine	175
3.2.1.4	De la voix qui sonne	179
3.2.1.5	Manger, dévorer, consommer	183
3.2.1.6	Manger, dévorer, consommer (bis)	188
3.2.1.7	Du meurtre à l'extermination	191
3.2.1.8	Habiter, cultiver, honorer	194
3.2.2	Le lexique s'organise	197
3.2.2.1	Organisation par mot leader	198
3.2.2.2	Organisation distribuée	202
3.2.2.3	Des liens qui se tissent	206

Les lexèmes qui ont été recueillis dans la base de données NeoClassy vont être étudiés dans un premier temps du point de vue de leur forme, en 3.1, puis du point de vue de leur sens, en 3.2.1. Du point de vue formel, des contraintes respectées dans la formation des constructions néoclassiques seront dégagées : certaines pèsent sur le constituant initial du mot construit, d'autres sur le segment à la jonction des deux constituants et, enfin, des contraintes pèsent

sur le mot construit dans son ensemble. Du point de vue du sens, des classes sémantiques seront mises au jour en fonction de chaque EN impliqué dans les constructions. Ces classes seront ensuite mises en perspective afin de faire ressortir un réseau lexical complexe. Nous verrons ainsi, en 3.2.2, que les constructions peuvent s'organiser autour de mots leader (notion que j'avais abordée dans le chapitre précédent, en 2.1.3) ou de façon plus distribuée.

3.1 Contraintes en jeu

Le choix du terme de contrainte qui a été fait ici rappelle la Théorie de l'Optimalité (désormais OT) (Prince et Smolensky, 2008). Cependant, même si je m'en inspire pour mon analyse, je ne fais pas le choix de m'inscrire pleinement dans cette théorie. Le principe premier d'OT est que toutes les langues répondent aux mêmes contraintes universelles, partiellement contradictoires, puisqu'elles sont le reflet de la grammaire universelle. Cependant, chaque langue est spécifique du point de vue de l'interaction de ces contraintes. Un générateur permet de créer pour un input donné, plusieurs outputs. Un évaluateur permet ensuite d'obtenir le candidat optimal pour l'input parmi tous les outputs potentiels, en fonction de celui qui viole le moins de contraintes, ou du moins des contraintes moins haut placées dans la hiérarchie. Il existe deux grands types de contraintes : les contraintes de fidélité (l'output doit présenter autant que possible les mêmes segments, structures prosodiques, etc. que l'input) et les contraintes anti-marques qui visent l'émergence de structures phonologiques optimales. Toutes ces contraintes sont, pour chaque langue, organisées selon un ordre spécifique, et elles peuvent être violées.

Ici, je m'inspire des contraintes telles qu'elles sont définies par exemple par Hathout (2009). Ces contraintes n'entrent pas dans le cadre d'OT, car elles ne sont pas hiérarchisées, sont gradables et leur universalité n'est pas discutée comme le dit Hathout (2009, p.29). Nous retiendrons donc que les contraintes dont il est question ici sont des conditions sur la formation des lexèmes qui peuvent être soit remplies, soit partiellement remplies, soit ignorées. Hathout (2009, p. 42 et suiv.) avance que les contraintes pesant sur la production lexicale sont de trois types. Les premières ont trait à la position lexicale, c'est-à-dire le lexème en tant qu'objet entretenant des relations avec les autres lexèmes du lexique. Par exemple, la contrainte de préservation de la proximité sémantique :

CONTRAİNTE DE PRÉSERVATION DE LA PROXIMITÉ SÉMANTIQUE :
La position lexicale p qui convient le mieux à un sens visé s est celle qui est morphologiquement plus proche des positions q₁, ...,

q_n associées aux mots m_1, \dots, m_n qui sont sémantiquement les plus proches de s .

Le second type est constitué de contraintes morphophonologiques, par exemple des contraintes de dissimilation ou des contraintes de taille, qui font partie des contraintes anti-marque.

Troisièmement, Hathout (2009) dégage des contraintes lexicales. Parmi celles-ci, des contraintes de recyclage ou de transparence.

CONTRAİNTE DE RECYCLAGE : *La forme optimale d'un mot est aussi proche que possible d'une forme mémorisée dans une position lexicale aussi saillante que possible.*

Les contraintes de transparence sont à ranger parmi les contraintes de fidélité. Elles stipulent qu'un dérivé doit afficher la plus grande transparence du point de vue des relations qu'il entretient avec les autres mots du lexique. Par exemple, un dérivé doit ressembler le plus possible aux autres dérivés formés selon le même schéma :

CONTRAİNTE DE TRANSPARENCE DE LA SÉRIE DÉRIVATIONNELLE : *La forme optimale d'un dérivé est formellement la plus proche possible du plus grand nombre possible de formes de mots appartenant à la série dérivationnelle du dérivé.*

De même, la base doit être facilement identifiable :

CONTRAİNTE DE TRANSPARENCE DE LA BASE : *La forme optimale permettant l'identification de la position lexicale d'un dérivé est formellement la plus proche possible de l'une des formes d'un mot qui est morphologiquement aussi proche que possible de la base du dérivé.*

Sans faire une étude exhaustive de toutes les contraintes qui entrent en jeu dans la formation des CNS, je vais principalement m'attacher à décrire celles qui semblent jouer un rôle plus important : les contraintes qui pèsent sur la sélection du premier constituant (C1), la contrainte responsable de l'apparition d'une voyelle de liaison et enfin des contraintes pesant sur la taille du mot construit.

3.1.1 Le constituant initial

J'appellerai C1 le constituant rentrant en jeu à l'initiale de la formation, soit le thème correspondant au lexème qui occupe la partie gauche¹. Le radical,

1. Les notions de lexème, thème et radical seront précisées dans la partie 5.1. La dénomination générale de C1 me permet ici de différer ces questions.

au sens de Roché (2010), c'est-à-dire la séquence phonologique qui représente le C1 dans le mot construit, sera notée C1bis. Enfin, je noterai X la séquence à gauche de l'EN final, voyelle de liaison comprise. Dans *mafiocrate*, le C1 est donc /mafja/, le C1bis /mafj/ et X /mafjo/. La question de l'appartenance de la voyelle de liaison à l'un ou l'autre des constituants, qui sera abordée dans la partie suivante, est ainsi mise de côté pour le moment. D'autre part, est ici considéré comme C1 d'origine classique tout constituant ne correspondant pas à une forme du français. Par exemple, dans *théâtrologie*, le C1 est /teatr/, c'est donc un C1 natif en français, et non pas un élément néoclassique d'origine grecque. Cependant, une séquence comme /dɛlfin/ est prise ici comme C1 classique, puisqu'elle ne se trouve pas en français de manière autonome. Cependant, il serait également juste de l'analyser comme un thème savant du lexème DAUPHIN. Pour ne pas devancer l'analyse de ce type de constituant et avoir à traiter des cas problématiques en amont, j'ai pris la décision de considérer ces séquences comme des C1 classiques.

Dans mon corpus, j'ai tout d'abord calculé la proportion de C1 classiques, c'est-à-dire grecs et latins, et de C1 non classiques, français pour la quasi totalité, parmi tous les mots de NeoClassy. La Figure 3.1 donne un résumé de la proportion de C1 de chaque origine, selon chaque EN étudié.

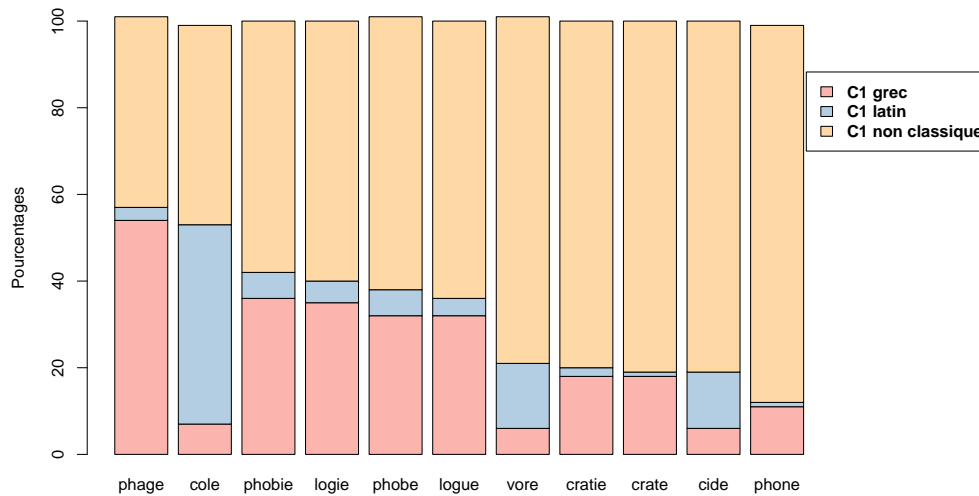


FIGURE 3.1 – Origine de C1, pour tous les lexèmes de NeoClassy, selon l'EN final

La plupart des ENs sélectionnent en majorité des C1 non classiques. C'est surtout le cas pour les ENs *-cide*, *-crate*, *-cratie* et *-phone* qui sélectionnent plus de 80% de C1 non classiques (respectivement 81%, 81%, 80% et 87%). Pour l'EN *-phone*, cela s'explique en grande partie par le fait que, comme

nous le verrons dans la partie 3.2.1 de ce chapitre, une grande partie des lexèmes *Xphone* servent à la dénomination des locuteurs d’une langue (*francophone*, *germanophone*, *anglophone*) ; le C1 est donc constitué d’un nom de langue. Or peu de constituants classiques désignant des langues sont disponibles, mis à part pour les langues européennes les plus parlées et/ou les plus anciennes (*hispanophone*, *germanophone*, etc.). La très forte proportion de C1 natifs dans la constructions des lexèmes *Xphone* est donc peu surprenante. Une telle explication n’est pas si aisée pour les lexèmes *Xcide*, *Xcrate* et *Xcratie*. Nous pourrions regarder du côté de la dictionnarité des mots construits, c’est-à-dire de leur présence dans un dictionnaire. En effet, nous pouvons faire l’hypothèse que les créations spontanées recueillies sur le Web sont le fait de locuteurs n’ayant pas de connaissances étymologiques particulières. Les mots non dictionnarisés sont donc plus à même de contenir des C1 natifs. C’est en effet ce que montre la Figure 3.2, puisque 77% des mots présents uniquement sur la Toile affichent un C1 natif, contre 40% pour les mots dictionnarisés.

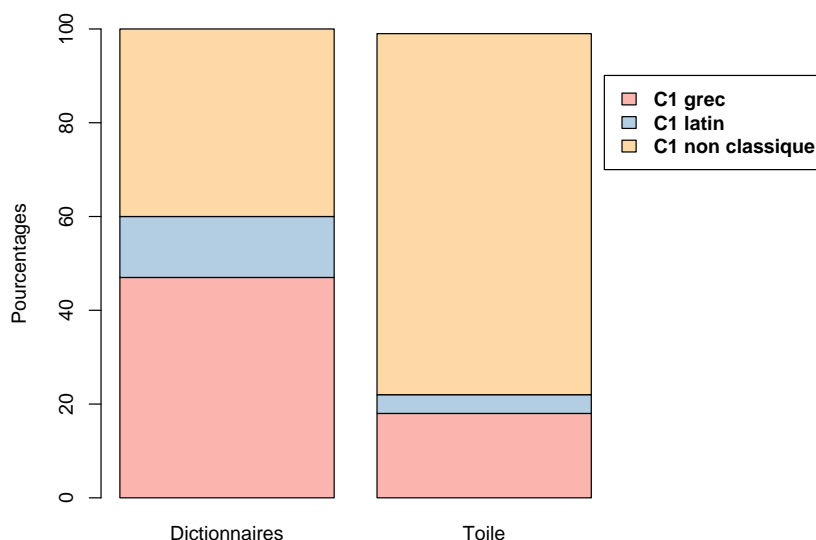


FIGURE 3.2 – Origine de C1, pour tous les lexèmes de NeoClassy, selon la dictionnarité des mots construits

Or si on regarde la Figure 3.3 qui illustre la table 2.4 du chapitre précédent, les mots formés avec les éléments *-crate* et *-cratie* sont en effet plutôt rares dans les dictionnaires, puisque respectivement 89% et 88% de ceux-ci sont uniquement présents sur la Toile.

De plus, la Figure 3.4 montre que, pour les lexèmes *Xcrate* et *Xcratie*, la

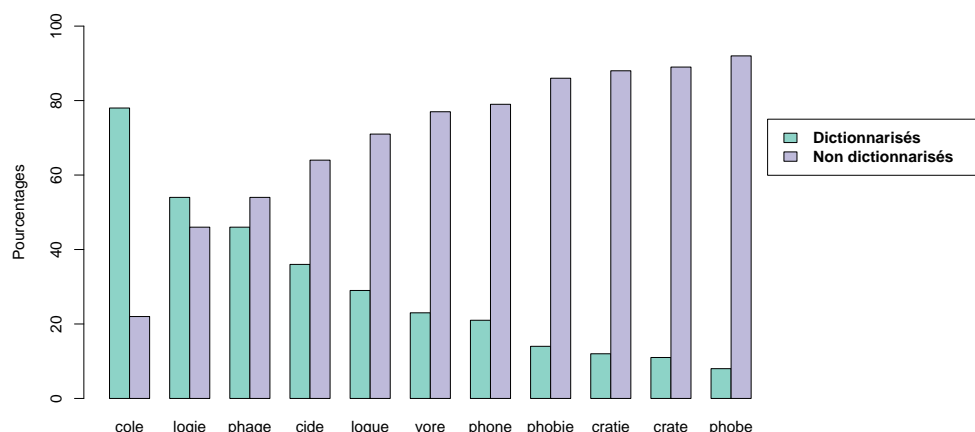


FIGURE 3.3 – Dictionnarité des lexèmes de NeoClassy, selon chaque EN

plupart de ceux qui ont un C1 non classique, ne sont pas dictionnarisés. L’hypothèse qui voudrait que les créations non dictionnarisées, étant le fait de locuteurs qui ont peu de connaissances étymologiques, contiendraient davantage de C1 non classiques, tient donc pour les mots formés avec ces éléments. La Figure 3.3 montre cependant que, concernant les lexèmes *Xcide*, la part de mots dictionnarisés est quasiment similaire à celle des *Xlogue* (36% de *Xcide* sont dictionnarisés et 28% des *Xlogue*) alors que, selon la Figure 3.1, la proportion de C1 non classiques est supérieure pour les premiers que pour les seconds (81% des *Xcide* ont un C1 non classique contre 64% des *Xlogue*). D’autre part, la Figure 3.1 montre un fort taux de bases classiques pour les lexèmes *Xcole*, *Xphage*, *Xlogue*, *Xlogie*, *Xphobe* et *Xphobie*. Concernant les premiers, l’hypothèse de la dictionnarité pourrait être défendue, puisque, comme la Figure 3.3 le montre, presque 80% des lexèmes de la base de données sont dictionnarisés. Cependant, la Figure 3.5 (A.), qui représente les liens qui existent entre dictionnarité du mot construit et origine de C1 montre qu’un tiers des *Xcole* non dictionnarisés sélectionnent un C1 d’origine classique ; la proportion est similaire pour les *Xphage* alors que seuls la moitié d’entre eux sont dictionnarisés. Cette forte proportion de C1 d’origine classique n’est donc pas uniquement liée à la dictionnarité des constructions néoclassiques. Plutôt, je fais l’hypothèse ici que le C1 d’une construction néoclassique telle que celles étudiées dans cette thèse répond à la contrainte suivante :

Contrainte 1 *Contrainte de sélection du premier constituant*

Le premier constituant d’une construction néoclassique doit être soit d’ori-

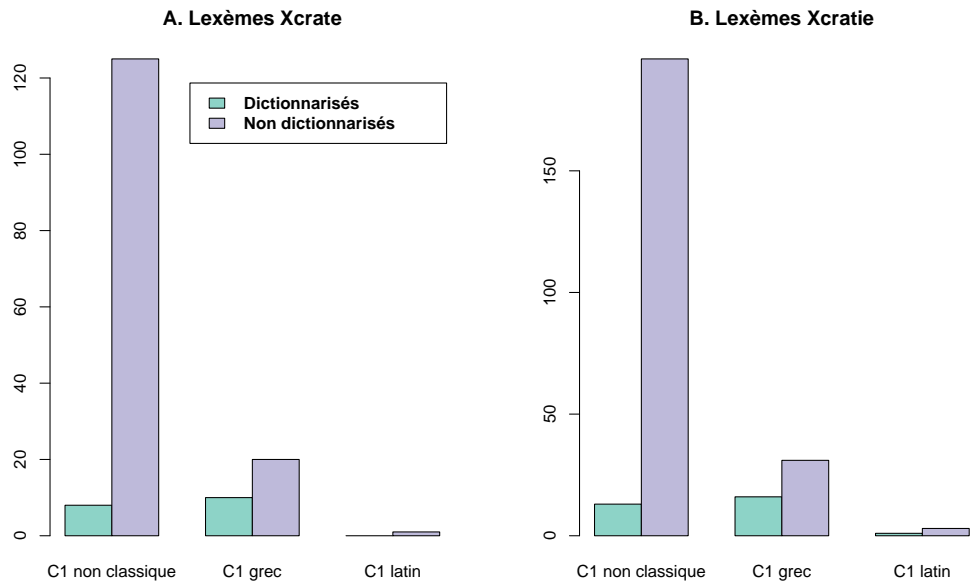


FIGURE 3.4 – Dictionnarité et origine de C1, pour les lexèmes *Xcrate* et *Xcratie*

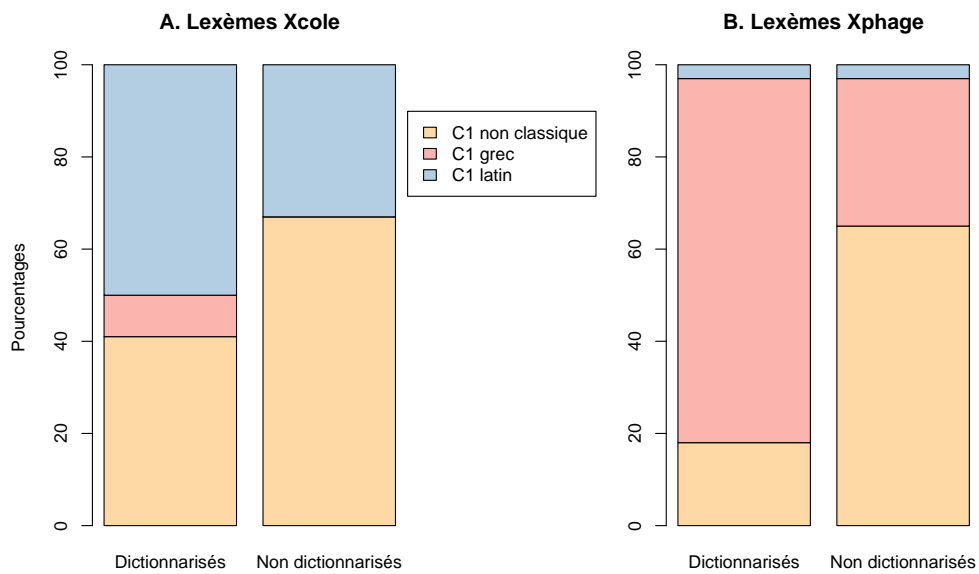


FIGURE 3.5 – Dictionnarité et origine de C1, pour les lexèmes *Xcole* et *Xphage*

gine classique, soit, quand aucune forme classique n'est répertoriée ou disponible, être néoclassique (c'est-à-dire qu'il doit ressembler à un constituant d'origine classique).

Les ENs finaux étudiés ici ont donc le même comportement que certains suffixes décrits par Roché (2010) ou Plénat (2008), par exemple *-ité* ou *-at*, qui préfèrent souvent sélectionner les thèmes supplétifs des lexèmes. Plénat parle de thème L, pour « latin » alors que Roché parle de thème S, pour « savant », qui englobe non seulement les thèmes réellement empruntés aux langues classiques, mais également les thèmes « pseudo-savants », c'est-à-dire construits par analogie avec des thèmes latins, mais qui n'existaient pas dans les langues classiques (*ivor-* dans *ivoréen* par exemple). On retrouve ici, par exemple en (1), des C1 semblables aux bases que l'on peut trouver avec ces suffixes.

- (1) a. génitoricide (< géniteur)
- b. pastoricide (< pasteur) (voir *pastorat*)
- c. odoricide (< odeur) (voir *odorat*)

Globalement, parmi les C1 d'origine classique de NeoClassy, 80% sont d'origine grecque, et 20% d'origine latine. Si ces chiffres généraux laissent transparaître une prédominance d'ENs grecs dans la base de données que d'ENs latins, la Figure 3.1 montre tout de même que les ENs finaux latins tendent à sélectionner davantage de C1 grecs que les ENs finaux grecs ne le font avec les C1 latins : parmi les C1 des mots à ENs finaux latins, 32% sont d'origine latine (2-a-i) et 27% d'origine grecque (2-a-ii), alors que parmi les C1 des ENs finaux grecs, 47% sont d'origine grecque (2-b-i) contre 6% d'origine latine (2-b-ii) :

- (2) a. (i) piscivore
- (ii) conchylicole
- b. (i) bibliophage
- (ii) somnologie

Par ailleurs, il est intéressant de remarquer la grande différence, du point de vue de l'origine du C1, entre les lexèmes formés avec les éléments *-phage* et *-vore*, synonymes mais issus de langues différentes : les premiers ont une proportion de 53% de C1 natifs contre 80% pour les seconds. Une hypothèse pourrait être que, de par la graphie, par exemple, l'élément *-phage* apparaît comme davantage néoclassique que l'élément *-vore*, dont tous les graphèmes font partie du français. Par ailleurs, l'élément *-vore* est présent dans des lexèmes du français qui ne sont pas considérés comme néoclassiques : DÉVO-

RER, VORACE, VORACITÉ. Même si cette hypothèse ne peut pas expliquer en totalité cette différence entre ces deux éléments, je pense qu'elle joue tout de même un certain rôle ; moins un élément a l'air néoclassique, graphiquement ou lexicalement, moins il va sélectionner de C1 néoclassique.

La contrainte 1 a plusieurs explications. D'un point de vue diachronique, les mots qui ont d'abord été créés dans les langues classiques puis importés en français (et dans d'autres langues), par exemple *démocratie* ou *anthropologie*, ont de fait une base classique. L'analogie qui s'est ensuite opérée a donc comme fondement des mots contenant deux constituants classiques. D'autre part, comme je l'avais abordé dans la partie 1.2.3, les premiers mots à avoir été créés dans les langues modernes selon des schémas néoclassiques l'ont été par des scientifiques qui avaient à cœur d'une part de nommer les choses d'une façon la plus juste et univoque, d'autre part de communiquer au moyen d'un jargon compris par eux seuls. C'est sur cette dernière base que nombre de mots sont créés, souvent à visée humoristique. Le premier constituant est souvent le lieu, dans le langage courant, de recherches de la part des locuteurs de la forme qui aura l'air le plus « scientifique », qui sera suffisamment obscure pour paraître savante. Ainsi, souvent la sélection de ce premier constituant répond également à une contrainte que j'ai nommé de « crypticité » : son interprétation, dans un registre courant, doit être possible mais difficile pour l'interlocuteur. Si dans un registre médical ou plus largement scientifique, la présence de constituants classiques permet, comme nous l'avons déjà vu, de nommer des concepts de façon très précise dans un vocabulaire partagé par la communauté, dans la langue courante l'emploi de ces constituants relève d'une imitation, souvent humoristique, de ce discours scientifique.

Contrainte 2 *Contrainte de crypticité*

La forme du premier constituant doit être suffisamment opaque pour être comprise par les seuls initiés ou faire l'objet d'un discours ludique ou humoristique.

Dans l'exemple (3), le lexème *pogonologue* (< *pogon-* 'poil, barbe'), qui veut donner un caractère scientifique à son référent, nécessite une paraphrase pour que l'interlocuteur puisse comprendre les propos.

- (3) Pourtant, après avoir discuté avec Allan Peterkin, un **pogonologue** (ou spécialiste de la barbe), auteur et expert de l'entretien du visage pour Dove Men+Care, on s'est rendu compte que les types à moustache sûrs d'eux – comme Charlie Hunnam dans *Sons of Anarchy* – nous émoustillent aussi. (<https://fr.news.yahoo.com/dis-barbe-portes-te-dirai-062546708.html>)

L'EN *-phobie*, quand il sert à la construction de noms de peurs, donne lieu à des lexèmes qui sont souvent de vraies énigmes. Nombre de pages Internet répertorient des listes de phobies, toutes plus obscures les unes que les autres (par exemple en (4)), certaines consistant en des devinettes.

- (4) ailurophobie : peur des chats
 algophobie : peur de la douleur
 apiphobie : peur des abeilles
 asthénophobie : peur de la faiblesse

<http://www.psychoweb.fr/news/insolite/575-listes-de-phobies-reelles-ou-pas.html>

Quelques fois, un C1 non natif est utilisé pour la formation de mots agissant comme des traits d'esprit, comme en (5), dont le caractère crypté demande à l'interlocuteur un certain savoir, au même titre que les traductions d'expressions populaires données en (6). Les constructions néoclassiques créées dans la langue courante servent ainsi à marquer une distance entre locuteurs, dans le langage courant, comme le faisaient les scientifiques avec la constitution de leur jargon.

- (5) a. C'est pas la seul émission qu'il faut déprogrammer d'NRJ12! Vi-
 rez les émissions **céphalophages** et rendez nous de vrai pro-
 grammes. Le voyeurisme et les programme abrutissants, à la
 trappe. (<http://communaute.lexpress.fr/membre/benjamin536>)
 b. Etant confrontée à deux traits complètement contradictoires de
 ma personnalité : le consumérisme et la **panurgeophobie**, j'es-
 saie de trouver des compromis entre les angoisses que je me paye
 (genre serai-je une bonne mère si je n'achète pas le meilleur des
 transats et sa farandole d'accessoires et de jeux), les névroses
 que la planète me flanque (genre si j'achète tout neuf, c'est quoi
 le bilan carbone?), ma haine pour la tyrannie de la société de
 consommation (Genre Et dans les régions où ils n'ont pas Au-
 tour de Bébé à côté, ils font comment pour élever dignement
 leurs gosses?) et les cheveux que mon financier s'arrache (style
 « Est-ce vraiment incontournable, chérie, d'acheter un mouche-
 bébé? »). (<http://carnet-ultime.blogspot.fr/2010/11/pu-ericulture-comparaison-de-prix.html>)
- (6) a. vos collègues de travail vous agressent dès le lundi matin après

- une grosse **encéphalite rectale** (Tête enfoncée dans un endroit précis de votre anatomie) (https://www.facebook.com/permalink.php?id=159700157445392&story_fbid=522060464542691)
- b. Wahou, je suis étonné de la capacité de certains à "deviner" le futur. Ton raisonnement est **capillotracté** mais je pense que tu as raison. (<http://www.universfreebox.com/article/21939/Evolution-pour-les-commandes-en-multiligne-sur-le-site-de-Free-Mobile>)
 - c. Ce texte parodie avec brio tous les excès érudisant et la **tétratrichotomie** des critiques universitaires pompeuses et formalisantes. (<http://www.oniris.be/nouvelle/flupke-exegese-d-un-poeme-flash-1655.html>)

Cependant, cette contrainte de crypticité est contrebalancée par une contrainte de fidélité :

Contrainte 3 *Contrainte de transparence du premier constituant*

La construction doit rester la plus fidèle possible au premier constituant ; celui-ci doit apparaître intact dans le mot construit, ne subissant ni ajout, ni effacement, ni transformation.

Certains lexèmes de NeoClassy forment des doublons avec d'une part, un C1 classique et d'autre part un C1 natif. Dans le premier cas, c'est la contrainte de crypticité qui est mise en avant, dans l'autre c'est la transparence du C1 qui est privilégiée et donc la contrainte 3, contrainte de fidélité, qui est respectée :

(7) bibliophage / livrophage (< livre)

Le calcul du respect de cette contrainte de transparence n'est pas forcément aisé. En effet, du point de vue des transformations, par exemple, le fait de considérer une séquence comme un allomorphe d'un lexème, un de ses thèmes, ou un lexème différent fait débat dans la littérature. Doit-on, par exemple, considérer *lus-*, dans *lusophone* comme un thème de PORTUGAIS, un de ses allomorphes ou un constituant latin ? Comme, au moment de la récolte des données, le choix d'un a priori théorique n'avait pas été fait, la mesure de la distance de Levenshtein a été choisie pour mesurer le taux de respect de la contrainte de transparence du C1. Il s'agit d'un algorithme élaboré par Vladimir Levenshtein (1966) qui mesure la similarité entre deux chaînes de caractères. Ici, chaque séquence a été transcrite en phonèmes. Ensuite, la distance entre les chaînes de phonèmes qui constituent C1, C1bis et X, a été calculée ; dans cette mesure, chaque troncation, ajout ou transformation

de phonème équivaut à 1 (pour une présentation détaillée de la distance de Levenshtein et son application à la morphologie, voir Pascoli (2015)). Pour rappel, C1 est le lexème initial de la construction, C1bis sa représentation dans le mot construit et X la séquence à gauche de l'EN final ; (8) illustre la différence entre ces séquences dans le mot *accidentologie*.

- (8) *accidentologie* /aksidãtɔlɔʒi/ :
 C1 : /aksidã/
 C1bis : /aksidãt/
 X : /aksidãtɔ/

Si la distance entre le C1 et le C1bis est la mesure la plus intéressante puisque, comme nous le verrons plus loin, la présence de la voyelle de liaison est une contrainte forte qui pèse sur l'output de la construction, le calcul de la distance entre le C1bis et X permet de faire apparaître l'importance de la voyelle de liaison, les cas où une troncation permet de faire apparaître la voyelle de liaison ainsi que les quelques cas dans lesquels une consonne épenthétique est ajoutée (voir l'exemple (10) plus bas) et le calcul de la distance entre le C1 et X, de son côté, permet de faire apparaître les lexèmes pour lesquels la construction n'a nécessité aucune voyelle de liaison (11-a). Pour garder le même exemple qu'en (8), le Tableau 3.1 représente le calcul de la distance de Levenshtein entre les C1, C1bis et X de ce même lexème *accidentologie*. La distance est de 1, puisque il y a ajout du phonème /t/ entre le C1 et le C1 bis. Elle est de 1 également entre le C1bis et X, puisque un /ɔ/ est ajouté (c'est la voyelle de liaison que nous étudierons dans la section suivante). Enfin, la distance est de 2 entre le C1 et X, puisqu'il y a cette fois ajout de deux phonèmes, /t/ et /ɔ/.

		C1	C1bis	X
		/aksidã/	/aksidãt/	/aksidãtɔ/
C1	/aksidã/	–	1	2
C1bis	/aksidãt/	1	–	1
X	/aksidãtɔ/	2	1	–

Tableau 3.1 – Exemple du calcul de la distance de Levenshtein, pour le lexème *accidentologie* (/aksidãtɔlɔʒi/)

Selon ces calculs, plus la distance de Levenshtein est élevée entre C1 et X, plus la contrainte 3 est violée. Une distance de 0 correspond à un respect total de cette contrainte. La figure 3.6 représente les distances calculées pour chaque

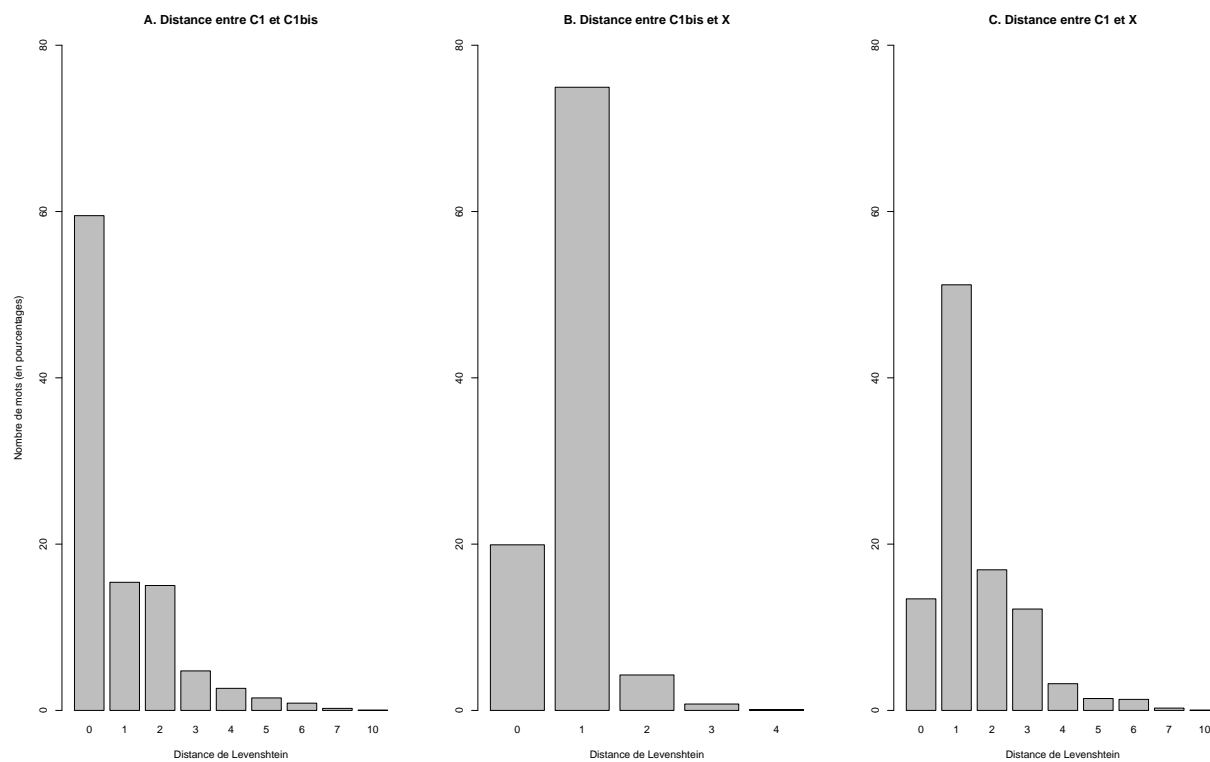


FIGURE 3.6 – Mesure de la distance de Levenshtein entre C1 et C1bis, C1bis et X et C1 et X, pour les C1 non classiques ($n = 2862$)

lexème de NeoClassy ayant un C1 non classique. En effet, j'ai considéré que la distance entre les C1 classiques et X étaient toujours de 0, puisqu'il paraît compliqué d'estimer si la voyelle est présente ou non dans le lexème original ainsi que dans sa représentation dans la langue moderne. De plus, il s'agit ici d'observer le comportement des locuteurs n'ayant pas nécessairement de connaissances étymologiques face à la création et la manipulation de constructions néoclassiques ; ainsi, leur comportement face à des C1 non classiques est plus intéressant que celui face à des C1 classiques. D'autre part, j'ai retiré de ces calculs de distance les mots ayant plus d'un constituant (voir Tableau 3.2).

Si l'avantage qu'a la mesure de la distance de Levenshtein de ne pas trancher entre le statut de thème ou allomorphe, cela constitue également un inconvénient. Ainsi, sur quel thème / forme opérer le calcul de la distance ? Pour le *couillonicide* que nous verrons en (9-b), faut-il calculer la distance sur le thème /kujõ/ ou /kujøn/ du lexème COUILLON ? Dans le premier cas, la distance est de 2, dans le second elle sera de 0. Comme chaque distance doit être calculée sur une seule séquence, j'ai fait le choix arbitraire de calculer les distances sur le thème 1 des noms et adjectifs, et sur le thème 3 des verbes (sur l'espace thématique des adjectifs, voir par exemple Bonami et Boyé (2005), sur celui des noms Roché (2010) et sur celui des verbes Bonami et Boyé (2003)). Quant au choix du lexème ayant la place du C1 lui-même, je me suis appuyée sur la sémantique des lexèmes, même si cela suppose des choix fondés sur mon intuition de locutrice native : par exemple, j'ai estimé que *corruptocratie* était formé sur le nom CORRUPTION et non sur le verbe CORROMPRE puisqu'il s'agit d'un 'pouvoir exercé au moyen de la corruption'. Nous voyons donc, dans la Figure 3.6 (A.), que la majorité des C1 (60%) sont représentés de façon intacte dans le mot construit, c'est à dire qu'ils n'ont pas subi de troncation ni de transformation de type allomorphie. La contrainte de fidélité est donc souvent respectée par les mots construits avec des C1 non classiques. Concernant les 40% restant, quels accidents² s'observent davantage ? La Figure 3.7 résume les types de modifications en fonction de la distance de Levenshtein calculée entre le C1 et le C1bis.

L'effacement d'un ou plusieurs phonèmes, contigus ou non, est le type d'accident le plus fortement représenté (52% de tous les accidents). Le plus souvent, cet effacement consiste en la troncation de la finale du C1 (9-a) ; quand seul le segment final est effacé, dans 75% des cas c'est une voyelle. Quand

2. J'emprunte cette appellation d'accident à Roché (2010) pour englober toutes les modifications subies par le C1, que ce soit des effacements, ajouts ou transformation. Ce terme est en effet pour moi plus englobant que celui de « modification » qui peut plus facilement être assimilé uniquement à « transformation ».

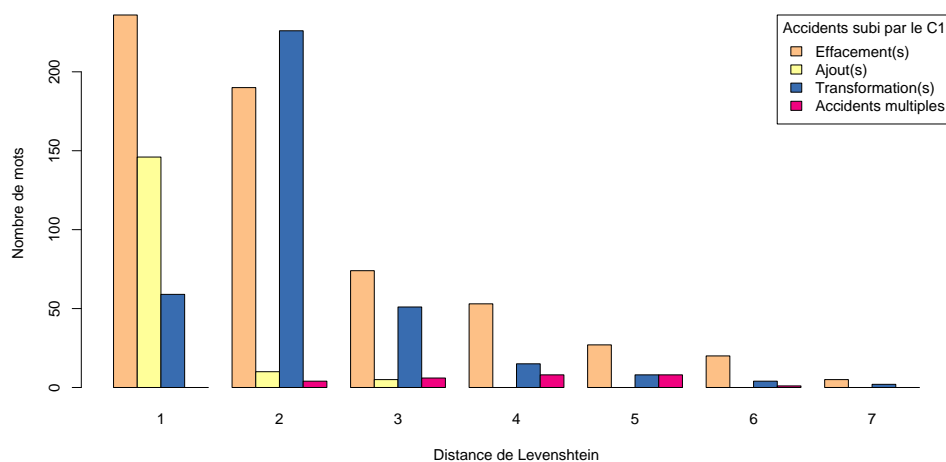


FIGURE 3.7 – Types d’accidents morphophonologiques subis par le C1, selon la distance de Levenshtein mesurée entre le C1 et le C1bis (n = 2862)

une consonne est effacée, c’est dans la majorité des cas pour laisser apparaître la voyelle de liaison déjà présente au sein du C1 (*musivore*) ; quand c’est une voyelle qui est effacée, c’est surtout pour laisser la place à une voyelle de liaison plus compatible avec l’output optimal d’une construction néoclassique, donc la voyelle /ɔ/ ou /i/ (*lingalophone*). Un deuxième type d’accident fréquent (31% des accidents) est la transformation. Sous ce terme se cachent généralement des allomorphies très fréquentes en français, comme celles représentées en (9-b). Enfin, lorsque l’accident consiste en l’ajout d’un ou plusieurs phonèmes, il s’agit là aussi souvent d’allomorphies fréquentes, la consonne ajoutée étant une consonne appartenant à un des thèmes du lexème (9-c-i). Très rarement, cette consonne est purement épenthétique (9-c-ii), n’étant présente dans aucune des formes du paradigme du lexème.

- (9) a. lingalophone (< lingala)
 musivore (< musique) cosmétophage (< cosmétiques)
 ordinophobe (< ordinateur)
- b. /ɛ̃/ → /in/ : requinophobie (< requin)
 /ɑ̃/ → /an/ : vaticanologie (< Vatican)
 /ɔ̃/ → /on/ : couillonicide (< couillon)
 /sjɔ̃/ → /t/ : corruptocratie (< corruption)
- c. (i) parenticide (< parent)
 cafardocide (< cafard)

- (ii) vouyoutocratie (< voyou)

Le graphique (B.) dans la Figure 3.6 affiche une majorité d'items ayant une distance de 1 ; ce graphique implique en effet en général l'adjonction de la voyelle de liaison. Cette voyelle peut être ajoutée à un C1 resté intact (10-a-i) ou déjà modifié (10-a-ii). Cependant, on remarque que pour 20% des mots, aucune voyelle de liaison n'est présente, soit que la finale du C1 est directement en contact avec l'EN final (10-b-i), soit qu'aucune voyelle n'a été rajoutée après troncation du C1 ((10-b-ii)), puisque la distance est de 0. Une petite minorité affiche des distances de 2 ou plus, ce qui indique l'insertion d'un segment épenthétique en plus de la voyelle de liaison (10-c).

- (10) a. (i) aspirateurophobie (< aspirateur)
trufficole (< truffe)
(ii) acnophobie (< acné)
soudanologie (< Soudan)
b. (i) tutsicide (< tutsi)
tortuphobe (< tortue)
(ii) écolophobie (< écolo(gie))
allerlogue (< allergie)
c. cannabinologie (< cannabis)
ombraticole (< ombre)

Le graphique (C.) de la Figure 3.6 fait ressortir 13% de lexèmes pour lesquels le C1 intact est directement suivi de l'EN final (11-a). De plus, la très forte proportion de lexèmes ayant une distance de 1 entre le C1 et X marque l'importance de la contrainte 3 (contrainte de transparence) puisque ces lexèmes sont principalement des C1 qui n'ont subi aucun accident qui sont liés à l'EN final par une voyelle de liaison (10-a-i). Ce peut toutefois être aussi des lexèmes dont la dernière voyelle est modifiée afin que la construction puisse répondre à la contrainte 4 que nous aborderons dans la section suivante (11-b), ou des lexèmes dont le dernier segment est tronqué pour faire apparaître justement cette voyelle (11-c).

- (11) a. escargologie (< escargot)
hébreuphone (< hébreu)
b. mafocratie (< mafia)
anglophobe (< anglais)
c. nématologie (< nématodes)
mousticide (< moustique)

La mesure de la distance de Levenshtein permet ainsi de faire ressortir des

cas de troncation, d'épenthèse ou d'allomorphie. Elle ne fait cependant pas ressortir les quelques 35 cas d'haplogogie entre le C1 et l'EN final recensés dans NéoClassy, par exemple :

- (12) ploucrate (< plouc)
 wolophone (< wolof)
 virucide (< virus)

Ces haplogogies permettent de respecter des contraintes dissimilatives qui défavorisent des formes comme *virusicide ou *wolofophone.

EN	Présence de plus d'un constituant à gauche de l'EN final
logie	133
logue	81
phobe	15
phobie	15
cole	6
cratie	5
crate	2
cide	0
phage	0
phone	0
vore	0

Tableau 3.2 – Nombre de lexèmes de NéoClassy ayant plus d'un constituant à gauche de l'EN final

Par ailleurs, ont été soustraits de cette étude de la distance de Levenshtein les lexèmes contenant plus d'un constituant à gauche de l'EN final. Ils sont au nombre de 257 ; le Tableau 3.2 en donne la répartition en fonction de chaque EN. Les ENs à afficher le plus de constituants initiaux multiples sont *-logue* et *-logie*. Cela s'inscrit dans une tradition de dénomination dans le domaine médical dont Namer (2007) étudie de nombreux exemples. Ces constituants initiaux multiples, en général au nombre de deux, peuvent s'ordonner syntaxiquement de différentes façons. Soit Z, Y et X les trois constituants du mot complexe (X étant l'EN final) : Z peut être l'objet de [YX] (13-a) ; Y peut être l'objet de [ZX] (13-b) ; [ZY] peuvent être coordonnés et soit former à eux deux l'objet de l'étude (13-c), soit l'objet d'étude est le rapport existant entre Z et Y (13-d) :

- (13) a. neurophysiologie ‘physiologie du système nerveux’
 b. histopathologie ‘histologie des tissus et organes malades’
 c. dermato-vénérologie ‘spécialité médicale qui regroupe l’étude et le traitement des maladies de peau et des maladies vénériennes’
 d. géobiologie ‘science des rapports de l’évolution géologique de la Terre avec celle de la vie et des organismes vivants’

Par ailleurs, concernant les mots de plus de deux constituants formés avec des ENs autres que *-logue* et *-logie*, les deux premiers constituants peuvent être également coordonnés (14-a). En (14-b), les deux premiers constituants sont redondants, ce qui apporte un effet d’accentuation.

- (14) a. hortifruticole (< hort- ‘jardin’ + fruit)
 médiabobocratie (< média + bobo (BOurgeois BOhème))
 b. caviarbobocratie (< caviar + bobo (id.))
 vitivinicole (< viti- ‘vigne’ + vin)

Enfin, les mots formés avec les constituants *-phobe* et *-phobie*, les plus à même de respecter la contrainte de crypticité, peuvent aussi contenir deux voire trois constituants à gauche de l’EN final.

- (15) a. hydroluxophobe (< hydr- ‘eau’ ; lux- ‘lumière’) ‘qui fuit l’eau et la lumière, fêtard’
 caligynephobie (< cali- ‘beau’ ; gyn- ‘femme’) ‘peur des femmes très séduisantes’
 b. autocheirothanatophobie (< auto- ‘soi-même’ ; cheir- ‘main’ ; thanat- ‘mort’) ‘peur de se tuer de sa propre main, peur du suicide’

Concernant la catégorie lexicale du C1, les lexèmes de NeoClassy affichent, dans une écrasante majorité, des noms en position initiale. La Figure 3.8 reporte la catégorie des C1 natifs des lexèmes de NeoClassy. Plus rarement, le C1 peut être un adjectif (16-a), un verbe intransitif (16-b), un verbe transitif (16-c), un syntagme nominal (16-d), une locution (16-e), une interjection (16-f) ou un pronom (16-g).

- (16) a. crudivore (< cru)
 b. bossocratie (< bosser)
 c. mixologie (< mixer)
 d. drageusedemonmecophobe (< dragueuse de mon mec)
 e. cestpasgravologue (< c’est pas grave)
 f. abracadabralogie (< abracadabra)

g. jecratie (< je)

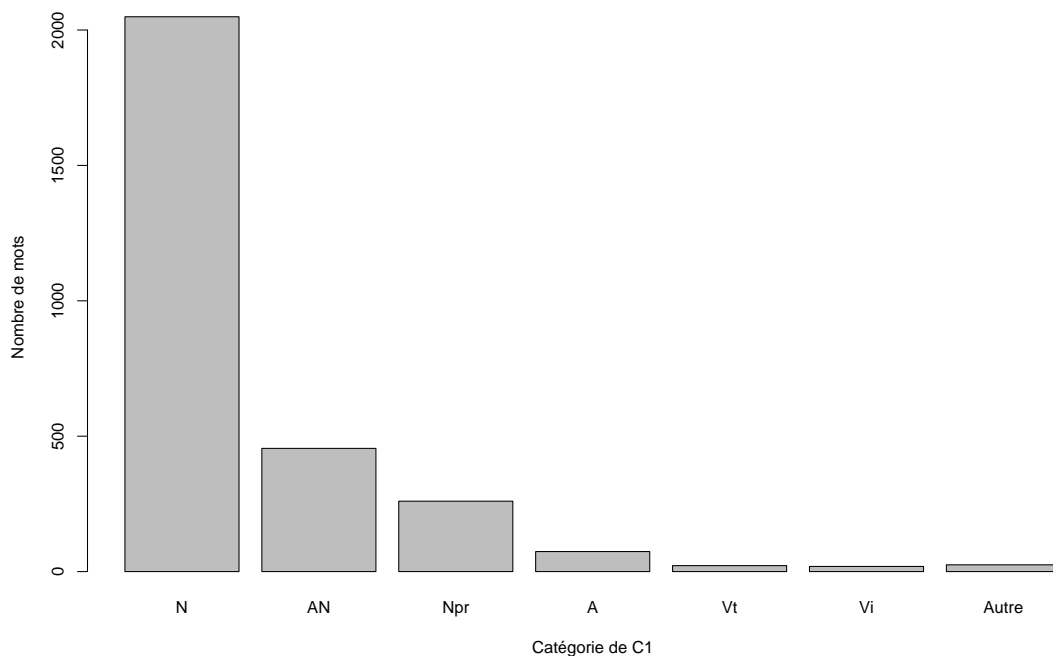


FIGURE 3.8 – Catégorie des C1 natifs des lexèmes de NeoClassy (n=2904)

Cependant, les catégories ici sont telles qu'elles sont observées dans le mot construit : elles relèvent donc du radical et non de la base (voir Roché (2010)). Ainsi, certains C1 ont été classés parmi les adjectifs alors que la base est nominale : par exemple, *marinicide* a été classé comme ayant un C1 adjectival alors que la base de la construction est très certainement le nom MER qui a emprunté son radical à la forme adjectivale correspondante. De plus, comme nous le verrons dans le dernier chapitre de cette thèse, dans la section 5.1.1.1, les C1 ici considérés comme éléments non natifs seront traités, après analyse, comme des thèmes S de lexèmes. Si, par manque de place, une étude de la catégorie du C1 après définition du statut de ces éléments ne sera pas conduite, ces premiers résultats poussent toutefois à considérer comme nominal le C1 prototypique d'une construction néoclassique telle que celles présentes dans NeoClassy.

Comme j'ai commencé à l'évoquer tout au long de cette partie sur le constituant initial, la voyelle de liaison a un rôle central dans la formation des

constructions néoclassiques.

3.1.2 La voyelle de liaison

Le segment qui est le plus communément appelé voyelle de liaison dans la littérature a suscité et suscite toujours de nombreux questionnements quant à son statut ou son appartenance à l'un ou l'autre des deux constituants de la construction néoclassique, comme je l'ai abordé dans la section 1.3.3. Je ne reviendrai donc pas ici sur les différents traitements qui sont faits de ce segment. Voyons plutôt ce qui se passe du côté de la base de données NeoClassy. Dans les analyses qui suivent, la voyelle <i> est considérée comme étant déjà présente dans le premier constituant si celui-ci est un lexème du français affichant cette voyelle (ex. *taxi*) ou un constituant d'origine latine (ex. *hort(i)-*). De même, avec la voyelle <o> qui sera considérée comme présente dans des lexèmes français, même si ce n'est pas graphiquement (ex. *bureau*), ou si la troncation du premier constituant se fait après la voyelle (ex. *choco-* < *chocolat*), et présente dans les constituants d'origine grecque (*phyt(o)-*). Cependant, dans ce travail, les voyelles <i> et <o> ne seront pas inscrites lors de la citation des constituants classiques, puisqu'elle ne sont pas forcément obligatoires ; *phyt(o)-* sera par exemple noté *phyt-*. De plus, comme il a été précisé précédemment, un lexème est considéré comme français dès lors qu'il fait partie du lexique de cette langue, même s'il a des origines classiques : *théâtre*, dans *théâtreologie*, est considéré comme un lexème du français et non pas un constituant *théâtr-* d'origine grecque. Le <o> de liaison dans *théâtreologie* par exemple, sera donc considéré comme un <o> rajouté, au même titre que le <o> présent dans, par exemple, *culturologie* (< *culture*).

La Figure 3.9 nous montre que ce que j'appelle voyelle de liaison n'est pas forcément une voyelle. En effet, moins de 1% des mots de NeoClassy (soit un effectif de 25³) présentent une consonne à la jonction des deux constituants du mot construit. Le Tableau 3.3 recense la distribution du segment intercalaire quand celui-ci est une consonne ou semi-consonne.

Les consonnes présentes à l'intersection des constituants des CNs de NeoClassy sont toutes des consonnes appartenant au premier constituant. Il n'y a donc, dans la base NeoClassy, aucune consonne épenthétique à la jonction des constituants. Ces consonnes permettent une conformité à la contrainte 3, contrainte de fidélité au C1, abordée dans la partie 3.1.1. On remarque toutefois que la majorité des consonnes exposées dans le Tableau 3.3 (/j/, /r/,

3. Un mot présentant une semi-consonne avant l'EN final, *magouillecratie* est compris dans ces 25 mots.

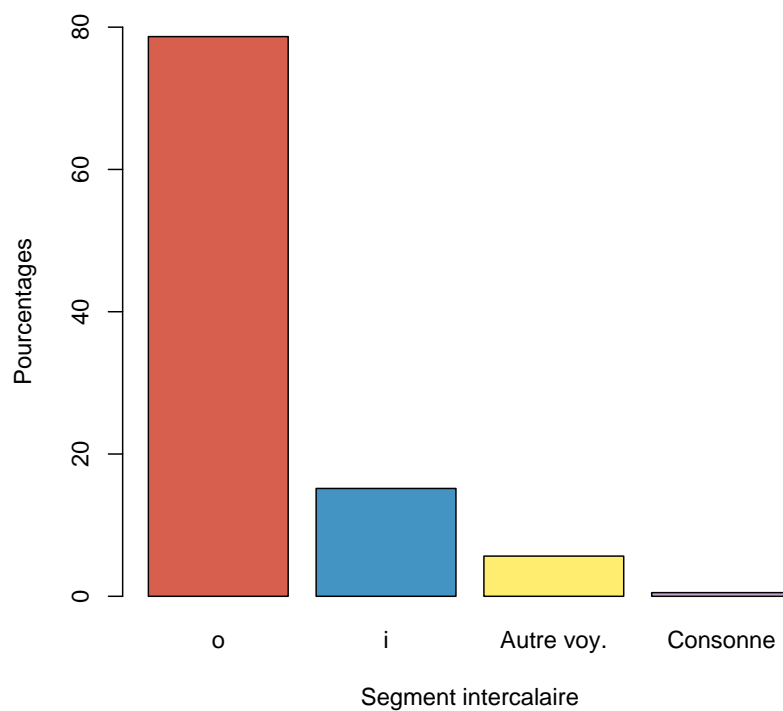


FIGURE 3.9 – Distribution du segment précédant l’EN, tous ENs de NeoClassy confondus (n=4459)

Segment intercalaire	Exemple	Nombre de mots de NeoClassy présentant cette consonne
/r/	papelarcratie	12
/s/	transphobie	4
/m/	aérodromphobe	3
/b/	webphage	2
/l/	multinationalecratie	2
/k/	ocphone	1
/j/	magouillecratie	1

Tableau 3.3 – Segment intercalaire quand celui-ci est une consonne ou semi-consonne

/m/, /l/) sont des sonantes, donc proches des voyelles. Par ailleurs, la plupart des mots présentant une consonne comme segment intercalaire ont un doublon affichant, lui, une voyelle ; ce doublon est souvent beaucoup plus fréquemment attesté, par exemple sur la Toile (Tableau 3.4). Les plus longs mots affichant une consonne comme segment intercalaire (*aérodromphobie*, *multinationalecratie*, *langueétrangèrophobie*) n'ont pas de doublons. On peut supposer que c'est afin de respecter une contrainte de taille : l'ajout d'une voyelle allongerait en effet la construction d'une syllabe.

Segment intercalaire : consonne	Occurrences recensées sur Google (le 03/09/15)	Segment intercalaire : voyelle	Occurrences recensées sur Google (le 03/09/15)
allerlogue	766	allergologue	3 718 000
cybercratie	303	cyberocratie	7
arabephone	227	arabophone	2 650 000
webphage	87	webophage	650
peoplecratie	69	peoplocratie	7
pulaarphone	8	pulaarophone	8 071
ocphone	5	ocophone	4

Tableau 3.4 – Exemples de doublons dont l'un présente une consonne comme segment intercalaire et l'autre une voyelle

Toutefois, la présence d'une consonne comme segment intercalaire restant anecdotique, je conserverai l'appellation de voyelle de liaison pour désigner le segment se trouvant à la jonction des deux constituants du mot construit dans la suite de ce travail.

Si nous regardons maintenant ce qui se passe du côté des voyelles dans la Figure 3.9, nous voyons que, dans la base NeoClassy, la voyelle <o> est clairement prédominante comme segment intercalaire. Comme la majorité des ENs finaux étudiés sont d'origine grecque, je formule donc l'hypothèse que la nature du segment intercalaire est davantage liée à l'origine de l'EN final qu'à l'origine du premier constituant.

En effet, si l'on compare la voyelle de liaison présente dans les mots construits avec les ENs *-logie* d'une part et *-phone* d'autre part, lorsque cet élément construit des noms de locuteurs d'une langue (par exemple *anglophone* ou *italophone*), le <o> est présent dans des proportions comparables pour les deux ensembles : autour de 95% pour les mots en *-logie* et autour de 82% pour

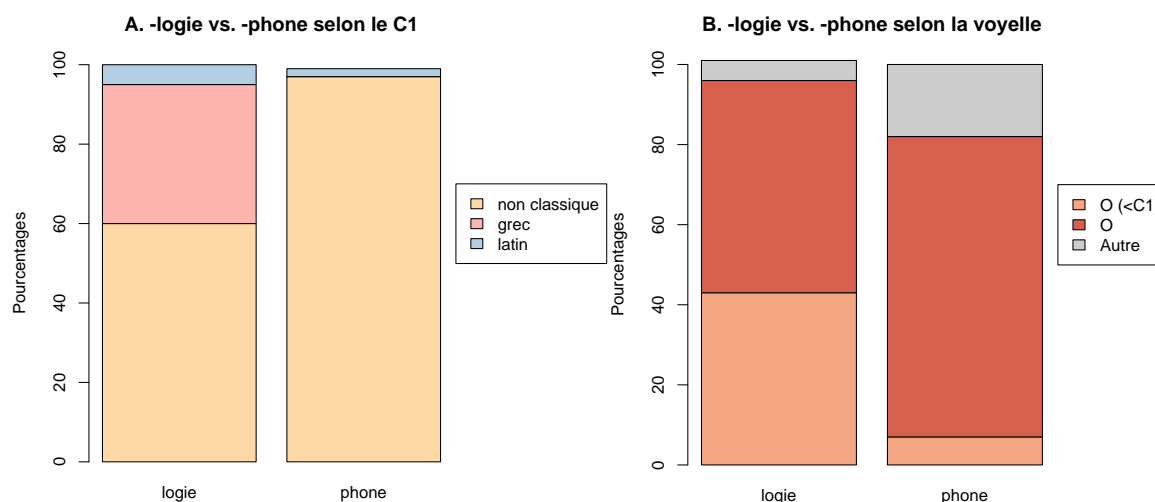


FIGURE 3.10 – Comparaison des EN *-logie* et *-phone* du point de vue de l’origine de C1 (A.) et du segment de liaison (B.)

ceux en *-phone* (Figure 3.10, B.). Or les lexèmes construits avec l’élément *-phone* désignant des locuteurs d’une langue prennent comme constituant de gauche le nom de la langue, qui sont, pour la quasi-totalité (c’est-à-dire 96,79%), des lexèmes du français et non des constituants d’origine classique. Cela explique pourquoi seulement 8,43% des mots contenant la voyelle <o> l’héritent du constituant de gauche (17) :

- (17) cantophone < cantonais
 estophone < estonien
 hellénophone < hellén-

Les mots construits avec l’élément *-logie* prennent, pour 40% d’entre eux, un C1 d’origine classique : 44,68% des mots contenant la voyelle <o> la tirent du premier constituant. Ainsi, en dépit de la différence des constituants de gauche quant à leur origine, les lexèmes construits avec les éléments *-logie* et *-phone* présentent la voyelle intercalaire <o> dans des proportions similaires, qu’elle soit ou non héritée du C1 : cette voyelle ne semble donc pas forcément corrélée à la nature du constituant de gauche.

La prédominance de la voyelle <o> malgré l’origine du C1 se vérifie sur l’ensemble des lexèmes de NeoClassy. La Figure 3.11 montre en effet que, si 90% des lexèmes formés avec un C1 d’origine grecque affichent la voyelle <o>, c’est également le cas pour près de 80% des lexèmes construits avec des C1 non classiques et 40% des lexèmes formés sur des C1 d’origine latine. La Figure 3.12 examine plus précisément ce qui se passe pour les lexèmes

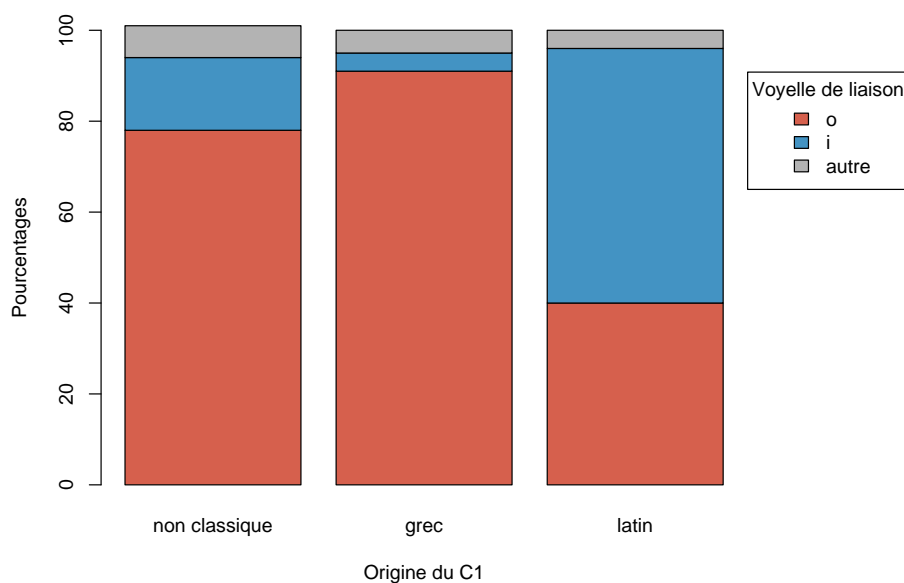


FIGURE 3.11 – Voyelle de liaison selon l’origine du C1

formés avec des C1 non classiques.

Ainsi, pour 65% des mots formés avec un C1 non classique, la voyelle <o>, non présente dans la base, a été rajoutée. La forme qui correspond au C1 dans la construction ressemble aux formations en *-o* abordées par Kilani-Schoch et Dressler (1992) et exemplifiées en (18-a). Ces formations se différencient des abréviations du type (18-b) pour lesquelles le <o> est déjà présent dans le lexème de départ. En (18-a), la séquence <o> est rajoutée après troncation du lexème originel.

- (18) a. intello, gaucho
b. gynéco, magnéto

Cependant, des différences significatives sont observables entre les données présentées par Kilani-Schoch et Dressler (1992) et les lexèmes de NeoClassy. D’un point de vue phonologique par exemple, les auteurs notent d’un côté qu’une semi-voyelle non syllabique n’est jamais maintenue avant un *-o* (*pharmacien* → *pharmaco*) et de l’autre l’impossibilité de finales en /so/ ou /kso/ (Kilani-Schoch et Dressler, 1992, p.73) : c’est ainsi que l’on rencontre *dico* et non pas **dicso*. Or plusieurs exemples contredisent ces principes lorsqu’il s’agit de C1 de constructions néoclassiques :

- (19) a. mafocratie, malariologie, ukrainiophone

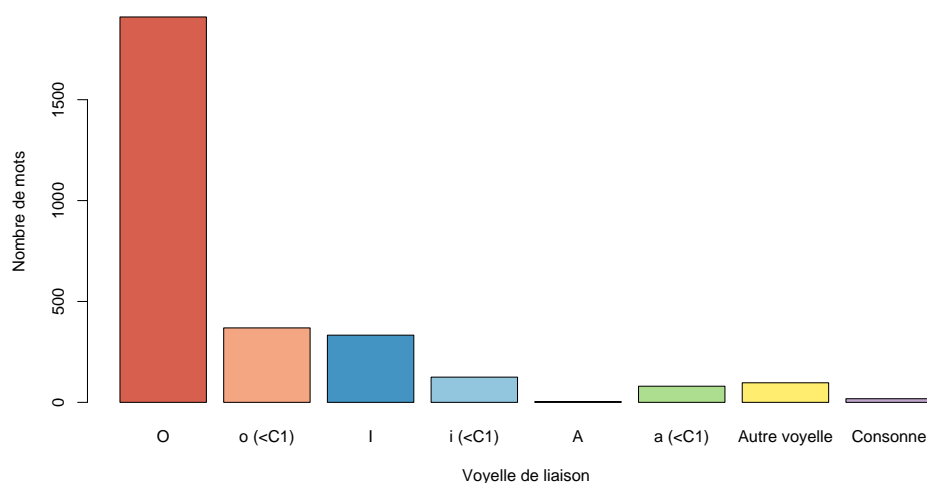


FIGURE 3.12 – Voyelle de liaison affichée pour les C1 non classiques

b. doxocratie

Par ailleurs, les formations en *-o* seraient formées sur les lexèmes construits, ou au moins complexes, dotés d’une structure morphologique interne ; la troncation serait opérée juste après ce que les auteurs appellent le « thème nu » ou « pseudo-racine » :

- (20) *gauch-iste* → *gaucho*
intell-ect-uel (cf. *dir-ect-eur* et *intell-igent*) → *intello* (Kilani-Schoch et Dressler, 1992, p.75)

Concernant les C1 des constructions de NeoClassy, ce principe n’est souvent pas respecté : les C1 ne sont pas forcément construits morphologiquement (21-a) et quand ils le sont, la troncation n’est pas obligatoirement faite après la « pseudo-racine » (21-b) :

- (21) a. blogologie (< blog), arabophone (< arabe), gamétocide (< gamète), islamophobie (< islam)
 b. minoritairocratie (< minorit-aire), architecturologie (< architecture), boulettologie (< boul-ette), princessologie (< princ-esse), racistologie (< rac-iste)

De même, concernant les abréviations du type (18-b), Kilani-Schoch et Dressler (1992, p.66) avancent que, si le lexème base contient un /o/ non initial, la troncation se fait toujours à sa suite. Si quelques lexèmes de NeoClassy

ont en effet un C1 tronqué après un /o/ non initial (22-a), certains ignorent ce /o/ (22-b), ou parfois plusieurs alternatives sont possibles pour un même C1 (22-c).

- (22) a. consocratie (< consommateur); tortiocratie (< tortionnaire); dieudophobe (< Dieudonné)
- b. géorgianophone (< géorgien); marocanophone (< marocain); spirochétocide (< spirochète)
- c. estophone / estonianophone (< estonien); bricologie / bricolophobe (< bricolage)

Ces quelques observations montrent que le <o> présent en grande majorité dans les constructions de NeoClassy peut difficilement être considéré comme faisant partie du C1, c'est-à-dire que le C1 a subi une troncation et l'ajout d'un suffixe /o/, comme les mots en (18-a), avant d'être utilisé comme constituant initial d'une construction néoclassique. Comme on le voyait en (21-a), le C1 n'est pas forcément tronqué avant l'adjonction du /o/. De plus, une autre contrainte entre en jeu dans la formation des constructions néoclassiques, que nous verrons dans la section suivante, celle de taille. Les CNs tendant à avoir des gabarits prosodiques plus longs que les formations en -o. Par exemple, même si Kilani-Schoch Dressler ont dans leur corpus 46% de lexèmes bases trisyllabiques, contre 24% dans NeoClassy, ces trisyllabiques donnent pour eux une moitié de dissyllabiques et une moitié de trisyllabiques; concernant mes données, les C1 trisyllabiques sont représentés dans les CNs par 23% de dissyllabiques, 49% de trisyllabiques et 25% de quadrisyllabiques. De même, les lexèmes quadrisyllabiques de Kilani-Schoch et Dressler donnent une grande majorité de formations en -o trisyllabiques; dans mes données, les C1 quadrisyllabiques donnent 37% de séquences trisyllabiques, 30% quadrisyllabiques et 17% pentasyllabiques.

Je reviendrai sur ces questions de taille dans la partie suivante, mais ces observations renforcent le fait que les C1 des CNs ne doivent pas être confondus avec des formations en -o comme celles étudiées par Kilani-Schoch et Dressler (1992). La voyelle <o> des CNs ne peut pas être uniquement liée au C1 de la construction. Regardons si l'EN final joue un plus grand rôle concernant le choix de la voyelle dans la construction.

Si nous observons la distribution du segment précédant l'EN final, en fonction de l'origine de ce dernier (Figure 3.13), nous voyons que du côté des EN d'origine grecque, la voyelle <o> comme élément de liaison est présente dans une écrasante majorité (90,93%). Cependant, alors qu'on s'attend à la même proportion du côté des EN d'origine latine avec la voyelle <i> la majorité, quoiqu'indiscutable, laisse tout de même davantage de place à la voyelle <o>

(11,96%) qui semble donc gagner du terrain pour devenir la voyelle marquant la composition néoclassique, aussi bien pour des éléments grecs que latins.

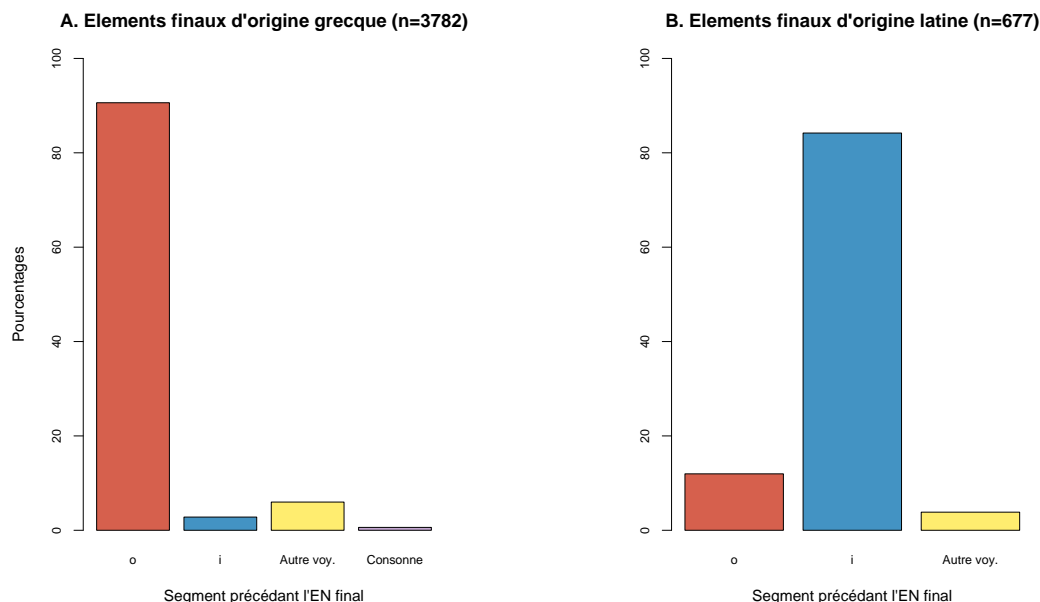


FIGURE 3.13 – Distribution du segment précédant l'EN, selon l'origine de l'EN final, aperçu

De plus, comme la Figure 3.14 le met en évidence, la voyelle <i> est non seulement peu présente devant un EN d'origine grecque, mais, à part quelques exceptions (23-b), quand elle précède un EN d'origine grecque, c'est qu'elle était déjà présente dans le premier constituant (23-a) :

- (23) a. papicratie < papi
taxiphobe < taxi
b. nectariphobe < nectar
pipolicratie < people

Or concernant les EN d'origine latine, ils font non seulement davantage de place à la voyelle <o> comme élément de liaison, mais cette voyelle est, pour 5,17% d'eux, présente dans le premier constituant (24-a) et pour 6,79% une pure épenthèse (24-b).

- (24) a. chocovore < chocolat
phytocide < phyt- 'plante'

- b. cheddarovore < cheddar
culturocide < culture

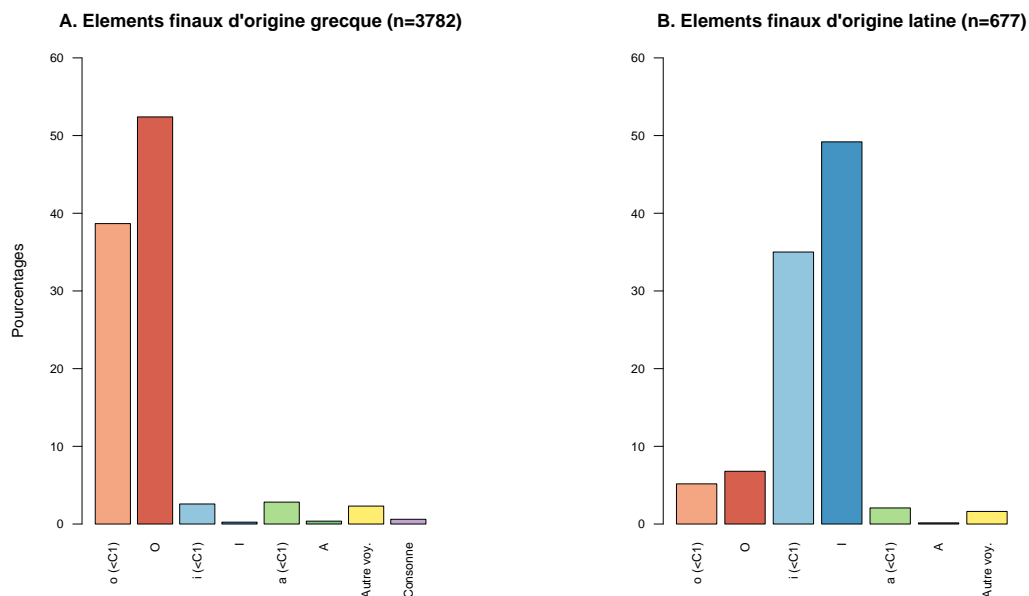


FIGURE 3.14 – Distribution du segment précédant l'EN, selon l'origine de l'EN final

La voyelle <o> est donc plus à même d'apparaître, même lorsqu'elle ne fait pas partie de C1 et que l'EN final n'est pas d'origine grecque. Par ailleurs, la voyelle <a> est la seule voyelle pouvant apparaître alors qu'elle ne fait pas partie du C1 (25) :

- (25) larvacide < larve
aromalogie < arôme
septalogie < sept

Cela s'explique certainement par le fait que, comme je l'avais abordé dans la partie 1.2.1, en grec ancien le *a* faisait partie des voyelles privilégiées pour marquer la composition. Certains mots affichant cette voyelle ont donc été transmis au français (26-a). D'autres mots, créés en français au XIX^{ème} siècle sur le modèle grec et devenus fréquents dans la langue courante, ont pris comme C1 un thème grec se terminant en *-a* (26-b). Ces mots servent donc de modèles aux nouvelles créations dont les exemples cités en (25) font partie.

- (26) a. décalogue, analogue
b. agoraphobie, mégaphone

Maintenant, regardons plus en détail se qui se passe pour chacun des EN de NeoClassy (Figure 3.15).

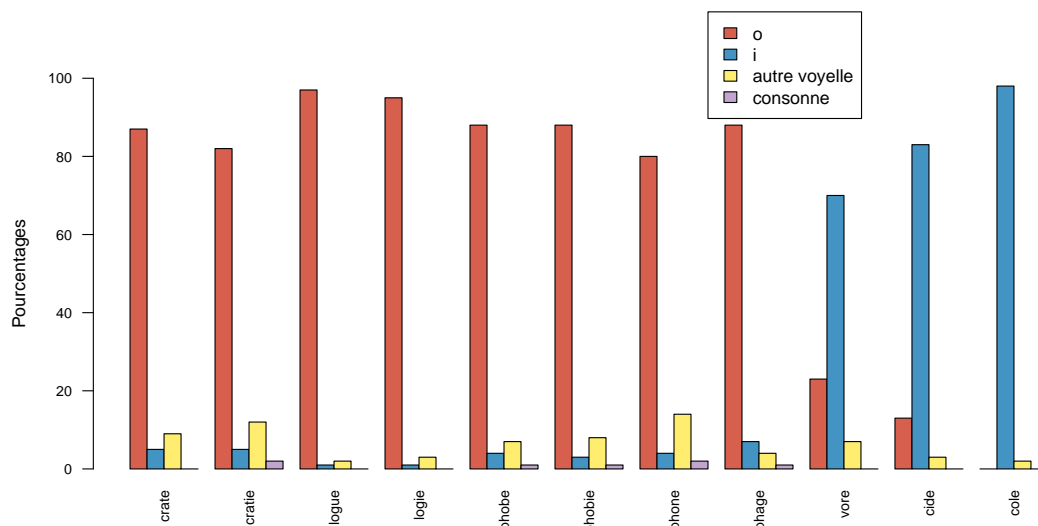


FIGURE 3.15 – Distribution du segment précédant l’EN, pour chacun des EN de NeoClassy

Concernant les mots construits avec les ENs finaux d’origine grecque, la voyelle <o> est présente dans une très grande proportion (de 82 à 97%). Les ENs *-logue* et *-logie* sont ceux qui l’affichent le plus (respectivement 97 et 95%). Cette prédominance apparaissait déjà dans les Figures 3.13 et 3.14. Les ENs finaux d’origine latine montrent quant à eux une moins grande homogénéité en ce qui concerne la voyelle de liaison. L’EN *-cole* est précédé dans 98% des cas par la voyelle <i>, jamais par la voyelle <o>. L’EN *-vore* quand à lui présente 70% de <i> et 23% de <o>. L’élément *-cide* a une distribution de la voyelle de liaison comparable à ce qui se passe pour les ENs d’origine grecque, avec 83% de <i> et 13% de <o>.

Regardons plus précisément ce qu’il se passe du côté des ENs d’origine latine. Concernant l’élément *-cide*, la voyelle <o> est présente surtout lorsque le C1 est un mot français. Cela correspond à ce que nous avons vu précédemment, à savoir que la voyelle <o> tend à être préférée. L’élément *-cole* a autant de bases françaises que latines (90 à chaque fois), mais ne présente quasiment que la voyelle <i>, à l’exception des deux lexèmes donnés en (114-b), qui présentent un <a> issu du C1.

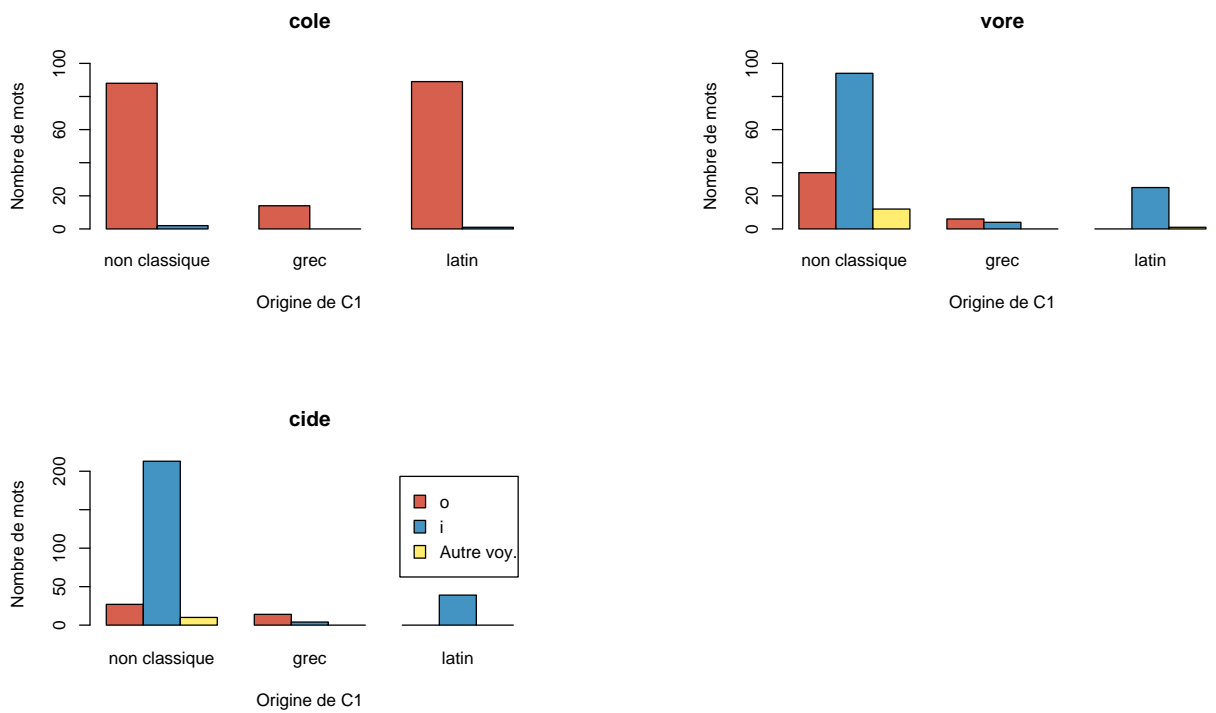


FIGURE 3.16 – Distribution du segment précédant l'EN latin, selon l'origine de C1

- (27) tabacole (< tabac_{FR})
 aquacole (< aqua_{LAT})

Je fais l'hypothèse que le sens de l'EN *-cole* est peu disponible pour des locuteurs ayant peu ou pas de connaissances étymologiques. Les lexèmes formés grâce à cet EN sont essentiellement forgés par des scientifiques ayant une connaissance du latin ou étant en contact avec des mots créés sur des modèles latins, qui eux-mêmes servent de modèle, ce qui induit une grande homogénéité. Ainsi, très peu de lexèmes en *-cole* ont pénétré la langue commune, ce qui n'est pas le cas par exemple des lexèmes en *-vore*. En effet, comme le montre le Tableau 3.5, la plupart des lexèmes en *-cole* sont dictionnarisés, et la tendance est inversée pour les lexèmes en *-vore*. Si l'on considère que la Toile est un témoin de la vivacité d'un processus de formation, nous pouvons dire que le processus de formation de mots en *-vore* est plus courant que celui en *-cole*.

	Dictionnarisés	Non dictionnarisés
-cole	152 78,35%	42 21,65%
-vore	41 23,30%	135 76,70%

Tableau 3.5 – Proportion de lexèmes en *-cole* et *-vore* dictionnarisés

Cela expliquerait donc la plus forte proportion de voyelle <o> parmi les lexèmes en *-vore*. De plus, *-vore* est le pendant latin de l'EN grec *-phage*. Il a pu donc être influencé par le schéma de formation de mots en *-phage* qui présente en majorité la voyelle <o>. Il y a d'ailleurs des doublets (voire des triplets) dans la base NeoClassy, qui présentent une même base couplée avec chacun des deux ENs synonymes. La Table 5.28 en annexe (p.460) recense ces doublets. Ceux qui sont reportés en (28) montrent que, si un lexème en *-vore* est le pendant d'un lexème en *-phage* bien plus fréquent, il pourra prendre un <o> comme voyelle de liaison (ce qui est renforcé par un C1 grec), alors que lorsque c'est le lexème en *-vore* qui est le plus fréquent, comprenant la voyelle de liaison <i>, le lexème en *-phage* pourra plus facilement conserver le <o>, peu importe l'origine du C1 :

- (28) a. anthropophage / anthropovore
 bibliophage / bibliovore

- b. carnivore / carnophage
fumivore / fumophage

Nous voyons donc que là encore, la voyelle <o> est souvent privilégiée, en dépit du C1 et de l'EN final des mots construits.

Enfin, pour conclure cette observation sur la nature de la voyelle de liaison, examinons sa corrélation avec une dernière variable : la présence des mots dans le dictionnaire. Si l'on considère que la voyelle <o> est la voyelle privilégiée dans la construction d'un néoclassique formé grâce aux ENs finaux étudiés dans cette thèse, et que les mots non dictionnarisés, étant des productions plus spontanées (cf. partie 2.2), alors la présence de la voyelle <o> devrait être plus forte pour les mots non dictionnarisés.

Vérifions d'abord si la fréquence qui s'observe avec les mots dictionnarisés en termes d'appariement voyelle <i> + EN d'origine latine s'applique sur les mots non dictionnarisés. Pour cela, regardons la fréquence obtenue avec les mots dictionnarisés. Sur 305 mots dictionnarisés formés avec un EN d'origine latine, 284 présentent la voyelle <i>, soit une fréquence de 93,11%. Faisons l'hypothèse que cette fréquence s'applique également aux mots non dictionnarisés et testons cette hypothèse à l'aide d'un loi de probabilité : la loi binomiale. Chaque mot formé avec un EN latin affiche soit un <i> (succès de probabilité 93,11%), soit une autre voyelle ou consonne (échec de probabilité 6,89%) : c'est une épreuve de Bernoulli de paramètre $p = 0,9311$. On répète cette épreuve de Bernoulli pour chacun des 372 mots non dictionnarisés. La répétition indépendante de cette épreuve de Bernoulli constitue alors un schéma de Bernoulli de paramètres $n = 372$ et $p = 0,9311$. Or dans un schéma de Bernoulli, la variable aléatoire qui compte le nombre de succès obtenus en répétant n fois de manière indépendante la même épreuve de Bernoulli de paramètre p suit une loi binomiale $\mathcal{B}(n; p)$. Donc, ici, si X est la variable aléatoire qui compte le nombre de mots non dictionnarisés affichant un <i> avec un EN latin, alors d'une part X peut prendre en théorie toutes les valeurs entières entre 0 succès et 372 succès et, d'autre part, cette variable aléatoire X suit la loi binomiale $\mathcal{B}(372; 0,9311)$. La Figure 3.17 fournit les prévisions de distribution des probabilités pour chaque valeur possible de X selon cette loi binomiale.

Les calculs démontrent qu'en pratique, en répétant 372 fois la même expérience aléatoire qui possède une probabilité de succès de 93,11%, la probabilité d'obtenir moins de 327 succès est quasiment nulle comme celle d'obtenir plus de 363 succès et que le pic de probabilité est atteint pour 347 succès (8,16% de chances).

Dans le but de tester l'hypothèse que nous avons faite, évaluons l'intervalle de

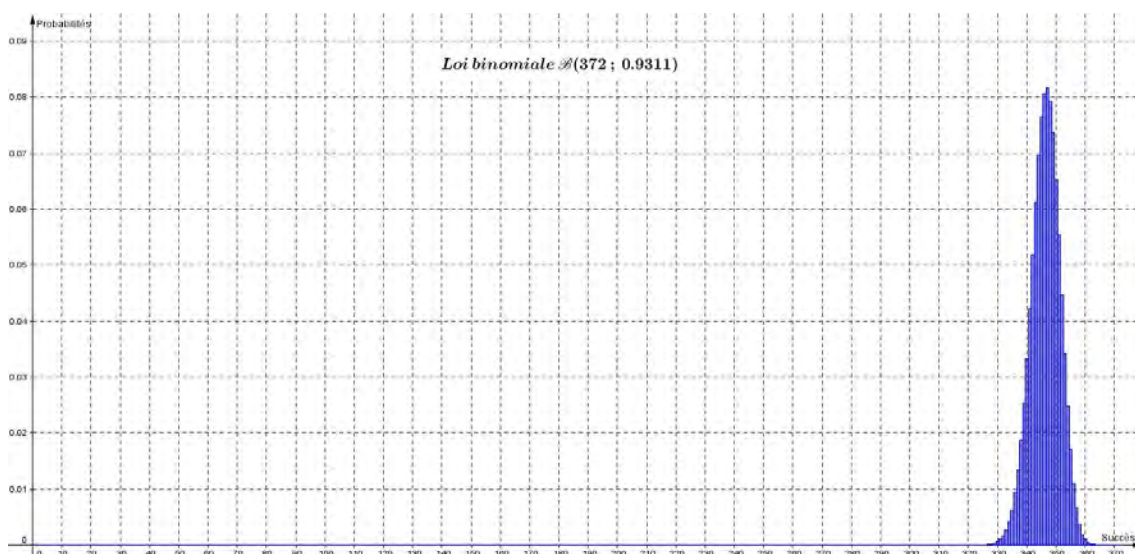


FIGURE 3.17 – Prévisions de distribution des probabilités pour chaque valeur possible de X selon la loi binomiale $\mathcal{B}(372; 0,9311)$

fluctuation asymptotique à 95% permettant d’encadrer avec une certitude de 95% la fréquence de succès que l’on doit obtenir si la probabilité de succès est, pour chaque mot non dictionnarisé et comme nous en avons fait l’hypothèse, la même que celle des mots dictionnarisés. Cet intervalle est le suivant :

$$I = \left[p - 1,96 \frac{\sqrt{p(1-p)}}{\sqrt{n}}; p + 1,96 \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$$

Ici, l’intervalle de fluctuation asymptotique est donc [90,54% ; 95,68%].

Or pour les EN d’origine latine, sur notre échantillon de 372 mots, nous obtenons 286 succès et donc une fréquence mesurée de 76,88%. Cette fréquence mesurée est très éloignée de l’intervalle qui possède pourtant 95% de chances de la contenir. Ce test nous autorise, avec une certitude mathématique de 95%, à rejeter notre hypothèse de 93,11% et donc de considérer qu’il y a une différence significative sur l’occurrence du <i> avant un EN d’origine latine entre les mots dictionnarisés et les mots non dictionnarisés. Le <i> est donc en fort recul parmi les mots non dictionnarisés.

Adoptons la même démarche maintenant concernant la présence de la voyelle <o> parmi tous les mots de NcoClassy. Sur 1376 mots dictionnarisés, 1033 affichent la voyelle <o>, soit une fréquence de 75,07%. Nous faisons donc l’hypothèse que cette fréquence s’applique aussi aux mots non dictionnarisés.

Nous mettons donc à nouveau en place une loi binomiale. Selon l'épreuve de Bernoulli, soit c'est la voyelle <o> qui est affichée (succès), soit c'est une autre voyelle ou une consonne (échec). Le schéma de Bernoulli a ici comme paramètres $n = 3083$ et $p = 0,7507$. X est ici la variable aléatoire qui compte le nombre de mots non dictionnarisés affichant une <o> peu importe l'origine de l'EN. La figure 3.18 fournit les prévisions de probabilités pour chaque valeur possible de X selon la loi binomiale $\mathcal{B}(3083; 0,7507)$.

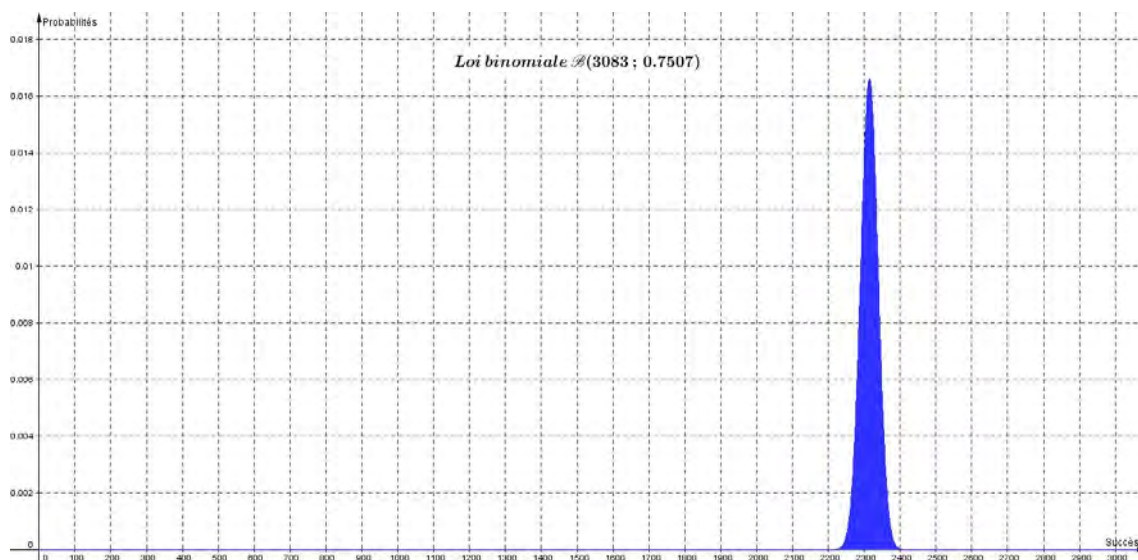


FIGURE 3.18 – Prévisions de distribution des probabilités pour chaque valeur possible de X selon la loi binomiale $\mathcal{B}(3083; 0,7507)$

L'intervalle de fluctuation asymptotique à 95% est ici [73,78% ; 76,36%].

Les calculs démontrent qu'en pratique, en répétant 3083 fois la même expérience aléatoire qui possède une probabilité de succès de 75,07%, la probabilité d'obtenir moins de 2232 succès est quasiment nulle, comme celle d'obtenir plus de 2396 succès, et que le pic de probabilité est atteint pour 2315 succès (1,66% de chances). L'épreuve de Bernoulli mise en place nous apporte 2475 succès, soit une fréquence de 80,28%. Cette fréquence ne se situant pas dans l'intervalle calculé, nous pouvons rejeter l'hypothèse avec une certitude mathématique de 95% : là encore, il y a une différence significative entre la présence du <o> parmi les mots dictionnarisés et les mots non dictionnarisés. La présence du <o> affiche une fréquence plus forte parmi les mots non dictionnarisés, ce qui est le signe que, parmi les créations spontanées, la voyelle <o> semble de plus en plus privilégiée.

C'est ce que montre Figure 3.19 qui expose la proportion de chaque élément

de liaison en fonction de la dictionnarité des mots de NeoClassy. Elle montre qu'il y a en effet davantage de mots contenant la voyelle <o> sur la Toile que dans les dictionnaires. De plus, la Figure illustre le recul significatif de la voyelle <i> dans les formations recueillies sur la Toile; ce recul se fait au profit d'autres voyelles ou même de consonnes, phénomènes très anecdotiques ou inexistant dans les dictionnaires. Ce pourrait être en faveur d'une contrainte de fidélité au premier constituant, comme nous l'avons vu dans la partie 3.1.1.

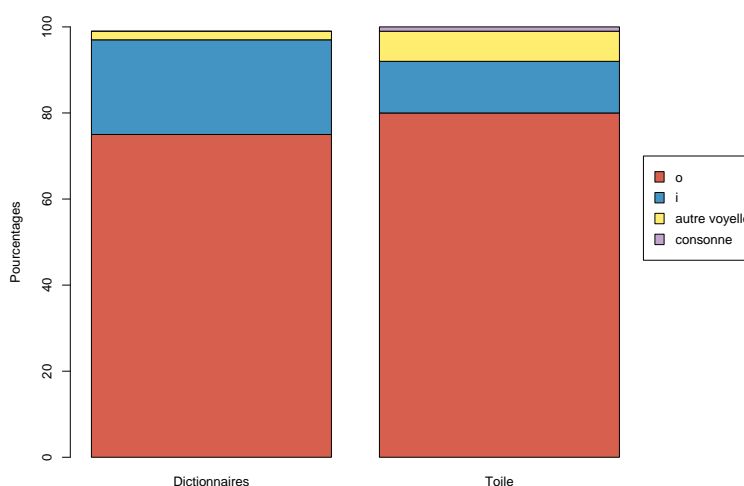


FIGURE 3.19 – Proportion, en pourcentages, de chaque élément de liaison selon que les lexèmes sont dictionnarisés ou non

Nous venons donc de voir que la voyelle de liaison n'est pas forcément liée à l'origine du C1, ni à l'origine de l'EN final. De plus, la voyelle <i> est en fort recul parmi les mots non dictionnarisés tandis que le <o> semble de plus en plus préféré. Comme je l'avais exposé dans la partie 1.3, Bauer (1998) émettait quatre propositions quant à la nature de la voyelle de liaison : soit elle appartient au premier constituant, soit au deuxième, soit aux deux ou encore à aucun. Finalement, aucune n'est satisfaisante : la présence du <o> peut ainsi être décidée par l'emploi d'un EN final d'origine grecque (29-a) ou la voyelle peut aussi être héritée du constituant de gauche (29-b). Parfois, elle est seulement une voyelle épenthétique n'étant dictée ni par le constituant initial, ni par le constituant final (29-c). Postuler son appartenance à l'un ou l'autre devient donc difficile.

- (29) a. bièrologue (< bière)
 b. chocovore (< chocolat)

c. trollocide (< troll)

Plutôt que d'essayer vainement de faire appartenir cette voyelle à l'un ou à l'autre des constituants, je fais l'hypothèse que la présence du <o> fait partie de la construction elle-même. Sa présence constitue une contrainte qui pèse sur l'output, contrainte à rapprocher de la contrainte de transparence de la série dérivationnelle déagée par Hathout (2009), dont je parlais en introduction de cette section. Le <o>, d'où qu'il provienne, fait rentrer une construction dans le paradigme des CNs.

Contrainte 4 *Contrainte de conformité de l'output*

L'output optimal d'un processus de formation de constructions néoclassiques doit afficher un <o> avant l'élément néoclassique final.

Si cette condition ne peut être remplie, le segment de liaison sera, dans l'ordre préférentiel : o > i > a > V > C.

La voyelle <o> fait partie de la construction même. Elle est d'ailleurs souvent mise en avant à l'écrit, par exemple entourée de tirets :

- (30) a. Le rôle de Mediaflex est de mettre en œuvre des solutions publicitaires pour ses annonceurs. C'est-à-dire, d'optimiser leur investissements dans ce domaine très « **budget-o-phage** » . (http://www.mediaflex.be/media_marketing_communication_mission.php)
- b. Je suis **Fischler-o-phobe**. (<http://caflim.blogspot.fr/2011/10/louvreuse-ma-dit-je-suis-fischler-o.html>)

Cette contrainte de conformité à l'output semble très forte, si bien que, puisque la voyelle <o> apparaît très souvent avant un élément néoclassique, elle tend à faire partie de celui-ci. Ainsi, il arrive fréquemment qu'un seul un tiret se trouve avant la voyelle, la rendant indissociable de l'EN final dans le processus de construction néoclassique :

- (31) a. Nous ne tolérerons aucune insulte (racisme, homophobie, **geek-ophobie** etc.) (<http://bulledegeekettes.blogspot.fr/p/regle-du-bon-usager.html>)
- b. **J'ai-rien-à-foutre-de-ma-vie-ophobie** (<http://www.europe1.fr/mediacenter/emissions/malik-bentalha/videos/j-ai-rien-a-foutre-de-ma-vie-ophobie-1035325>)

De plus, on trouve sur la Toile quelques exemples dans lesquels les éléments néoclassiques finaux sont employés en tant qu'items autonomes comprenant la voyelle de liaison :

- (32) a. « L' "**ophobie**" devient la norme, le mode d'expression privilégié de tous les occidentaux mentalement, voir physiquement castrés. Leur cerveau semble avoir été remplacé une tirelire vide, dans laquelle on place, au gré de l'actualité et des manipulations de nos maîtres, les désinfos médiatiques qui vont conditionner les robots humains qu'ils sont devenus. Comme je n'échappe pas à la règle, je crée le mot "ophobe-ophobe", qui consiste en une allergie viscérale à tous les "ophobes" de la terre. Bienvenue sur ce blog à tous les ophobes-ophobes de la Terre! »(<http://echelledejacob.blogspot.fr/2013/11/ophobie.html>)
- b. « Quel que soit l'issue d'un procès qui étale au grand jour le bavardage de tous les... **ologues** de service, la question fondamentale restera sans réponse »(http://www.lavie.fr/blog/josephine-bataille/le-proces-bonnemaison-sur-un-fil,3428?contexte=view_comment)

Maintenant que nous avons étudié la sélection et le traitement du premier constituant ainsi que le statut de la voyelle de liaison, regardons le mot construit dans son ensemble au travers du gabarit qui lui est imposé, à savoir des contraintes de taille.

3.1.3 Taille des mots construits

Les contraintes prosodiques que sont les contraintes de taille régissent la formation d'autres types de mots construits; par exemple Villoing (2009) avance que le mot composé VN dans son ensemble, doit avoir au minimum deux syllabes et au maximum quatre syllabes. Plénat (2009, p.63) évoque lui un idéal dissyllabique que doivent atteindre les dérivés ou les radicaux : « se dessinent dans le lexique deux grandes zones, l'une, marginale en termes sociologiques, où l'idéal dissyllabique tend à s'imposer par lui-même à la forme globale (formations argotiques et populaires), l'autre, central au regard du bon usage, où il doit s'allier à d'autres contraintes pour triompher, et où il ne triomphe que dans la forme qu'il donne au radical (conversation courante ou langage soutenu) ». Si on poursuit un peu la réflexion de Plénat, en français populaire l'idéal, même pour un mot construit, est dissyllabique; pour le français courant, central, il est trisyllabique pour les mots construits puisque le radical est dissyllabique et que la majorité des affixes sont monosyllabiques; peut-on en déduire que, pour un français plus soutenu, ou plus savant, cet idéal est encore plus long? La Figure 3.20 de gauche⁴ résume

4. Dans une boîte à moustaches, ou diagramme à boîte, le trait gras représente la médiane, la boîte centrale les deux quartiles du centre (la boîte contient donc 50% des observations les plus

la taille des mots de NéoClassy. On voit donc que la taille médiane est de 4 syllabes. Par ailleurs, la moyenne est aussi autour de 4 syllabes. Les constructions présentes dans la base de données sont donc de taille plus grande que celles étudiées par Villoing (2009). Elles sont également au-delà de l'idéal di- ou trisyllabique avancé par Plénat. Cependant, comme les éléments néoclassiques étudiés ici sont de tailles différentes, regardons également la taille de la séquence se situant à gauche de l'EN final, soit X dans *Xcide* par exemple (Figure 3.20 B.)⁵. La médiane est de trois syllabes, et les observations les plus centrales oscillent entre deux et trois syllabes. Là encore, X est de taille supérieure à la base idéale que doit avoir un dérivé c'est-à-dire deux syllabes.

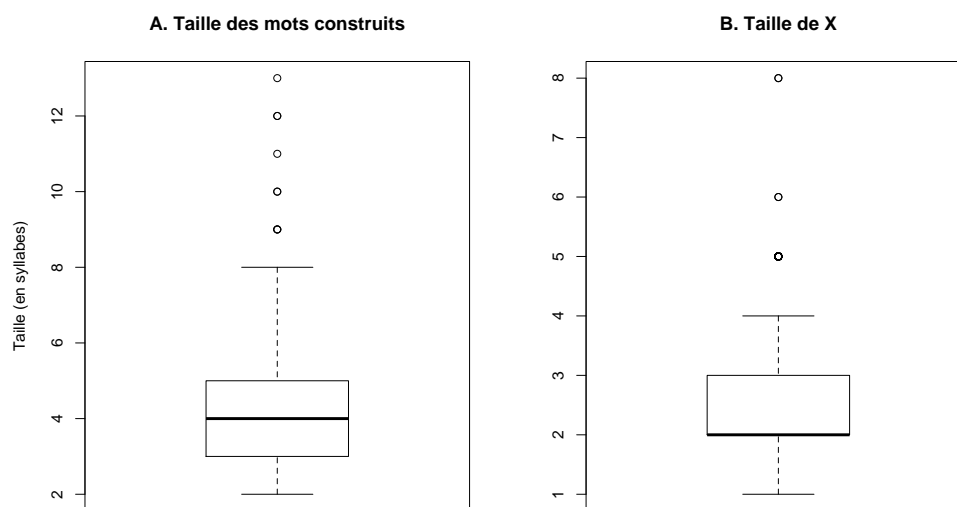


FIGURE 3.20 – Taille des mots construits de NéoClassy (A.) et de la séquence à gauche de l'EN, soit X (B.) (boxplot)

Regardons maintenant ce qu'il se passe pour les mots construits en fonction de chacun des ENs étudiés ici (Figure 3.21), du point de vue de la taille du mot construit dans son ensemble et de la taille de X.

Les mots les plus longs sont ceux qui sont construits avec les éléments *-phobe* et *-phobie*. Si on se rappelle que ces mots sont souvent créés dans des listes

centrales). Les moustaches représentent les données extrêmes.

5. Dans X est donc comprise la voyelle de liaison. Cela ne veut cependant pas dire que je considère la voyelle comme faisant partie du C1, comme je l'ai expliqué dans la section précédente. Cependant, écarter l'EN final des comptes permet de rapprocher *accidentologie* et *accidentologue*, respectivement 6 et 5 syllabes bien que construits de façon parallèle, et non pas *accidentlogie* et *accidentologue*, de 5 syllabes chacun mais qui n'affichent pas la même voyelle de liaison et dont le C1 a subi des traitements différents.

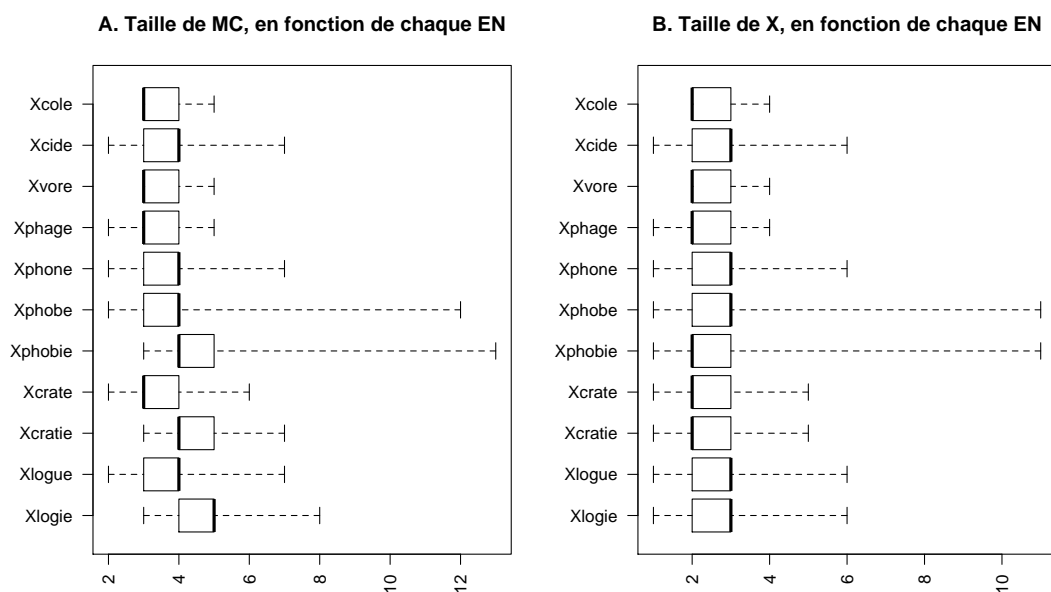


FIGURE 3.21 – Taille des mots construits de NeoClassy (A.) et de la séquence à gauche de l’EN, soit X (B.), en fonction de chaque EN final étudié (boxplot)

de noms de peur, souvent à visée ludique puisqu’il faut que les internautes déterminent de quelle peur il s’agit, certains de ces mots ont une contrainte forte de « crypticité » (voir section 3.1.1) :

- (33) hexakosioihexekontahexaphobie (< hexakosioi ‘six cents’ + hexakonta ‘soixante’ + hexa ‘six’) ‘peur du nombre 666’
 hippopotomonstrosesquipedaliophobie (< hippopotame + monstre + sequi ‘un demi en plus’ + ped- ‘pied’) ‘peur des mots très longs’
 paraskevidekatriaphobie (< paraskevi ‘vendredi’ + dekatria- ‘treize’) ‘peur du vendredi 13’

Cependant, quand on regarde les médianes de l’ensemble des mots, ce sont les mots construits avec l’élément *-logie* qui sont les plus longs. En revanche, les mots les plus courts sont ceux construits avec l’EN *-phage* (moyenne : 3,33 syllabes), l’EN *-crate* (moyenne : 3,45 syllabes), l’EN *-cole* (moyenne : 3,46 syllabes), l’EN *-vore* (moyenne : 2,48 syllabes). D’où proviennent ces disparités ?

Regardons la distribution des lexèmes de chaque taille en fonction de chaque EN (les valeurs extrêmes ont été retirées) avec la Figure 3.22.

Les résultats affichent une répartition entre des lexèmes trisyllabiques et des

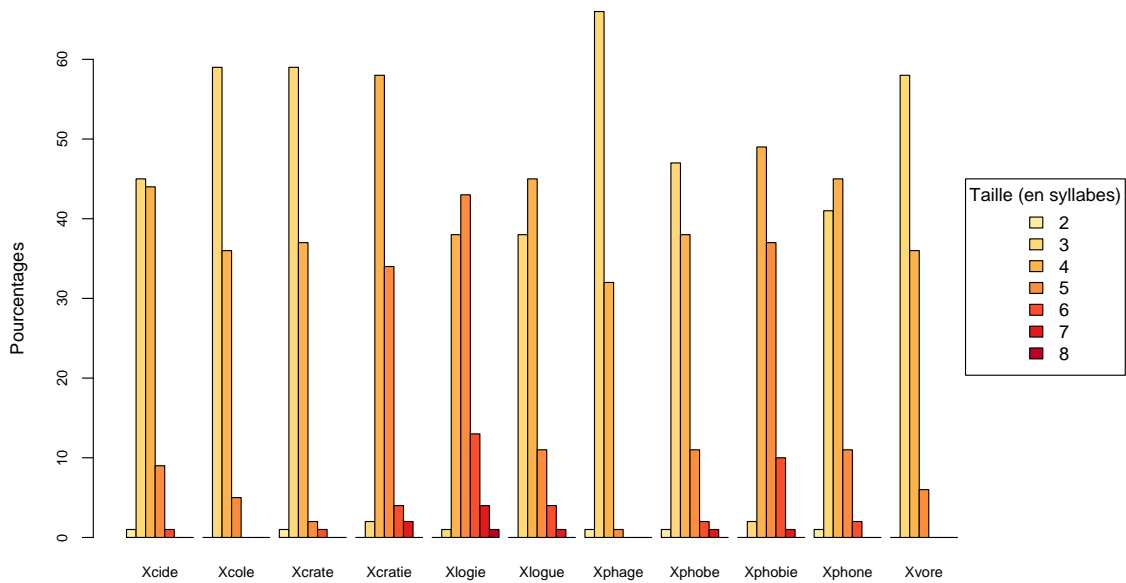


FIGURE 3.22 – Taille des lexèmes, en syllabes, en fonction de chaque EN

lexèmes quadrisyllabiques pour les ENs finaux monosyllabiques, et entre des lexèmes quadrisyllabiques et pentasyllabiques pour les ENs finaux bisyllabiques, soit un tiraillement entre $X = 2$ syllabes et $X = 3$ syllabes. Cette répartition évoque un tiraillement entre un idéal di ou trisyllabique, idéal que suivent de nombreuses constructions du français et une contrainte à laquelle semblent obéir les constructions néoclassiques :

Contrainte 5 *Contrainte de conformité de l'output à un gabarit*

L'output optimal d'un processus de formation de constructions néoclassiques doit être un lexème de trois ou quatre syllabes minimum.

Cette contrainte est à relier à la contrainte de crypticité que j'ai abordée précédemment.

L'explication de ce grand gabarit pourrait venir de la taille des C1 qui entrent en jeu dans nos constructions : mais ils font, pour la moitié d'entre eux deux syllabes, et pour un peu plus de 20% trois syllabes, comme le montre la Figure 3.23. Dans ce graphique, on voit également que les C1 bisyllabiques sont en majorité impliqués dans des constructions de 4 syllabes minimum, de même que les C1 trisyllabiques.

La taille du C1 n'est donc pas la seule variable déterminant la taille de la construction finale. Ainsi, si l'on compare plus précisément les lexèmes

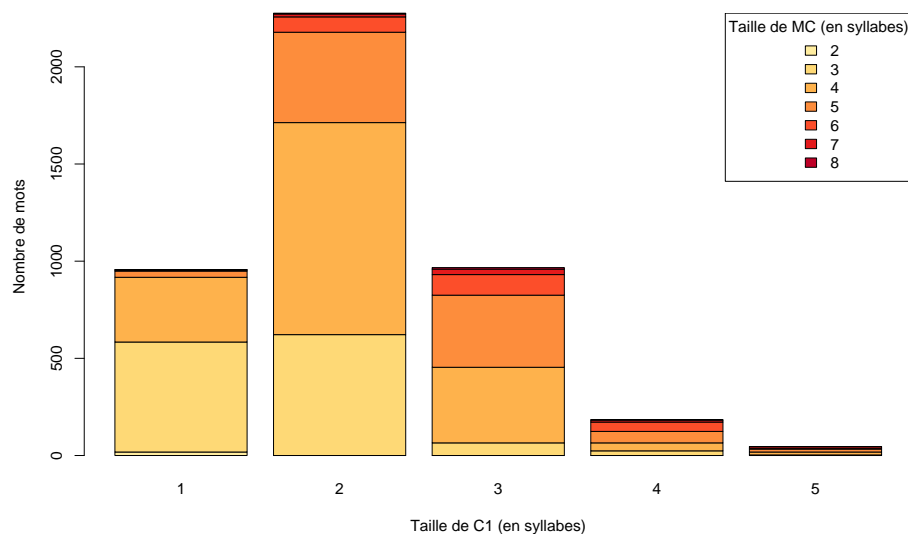


FIGURE 3.23 – Taille des C1, mis en perspective avec la taille des MC, pour les lexèmes ayant un seul constituant à gauche de l’EN final (n=4182)

Xlogue et *Xcrate* (Figure 3.24), la proportion de C1 de deux et trois syllabes est similaire, alors que les lexèmes *Xlogue* sont en moyenne plus longs que les lexèmes *Xcrate*.

Quelles différences y a-t-il selon d’autres variables entre les lexèmes construits avec ces deux éléments ? On avait vu dans le Tableau 2.4, reporté dans la Figure 3.3, que les lexèmes *Xcrate* étaient peu dictionnarisés. La dictionnarité joue-t-elle un rôle dans la taille des constructions néoclassiques ? Est-ce que les mots non dictionnarisés, étant souvent des créations spontanées, témoins de la vitalité d’un processus de formation, ont un comportement différent des mots dictionnarisés ? La Figure 3.25 nous montre des répartitions similaires de mots selon leur taille, parmi les mots dictionnarisés et non dictionnarisés, les non dictionnarisés étant à peine plus longs. Cependant, une différence est nette entre les lexèmes *Xlogue* et les lexèmes *Xcrate* : alors que, pour les premiers, les lexèmes dictionnarisés sont à peu près autant dissyllabiques que trisyllabiques, les lexèmes non dictionnarisés sont davantage trisyllabiques ; pour les lexèmes *Xcrate*, les mots dictionnarisés et non dictionnarisés sont tous davantage dissyllabiques, bien que les mots dictionnarisés aient tendance à être plus longs que les dictionnarisés. La différence de taille entre les *Xlogue* et les *Xcrate* ne semble donc pas due à leur dictionnarité.

D’autres variables jouent sur la taille de l’output. La Figure 3.26 (A.) nous

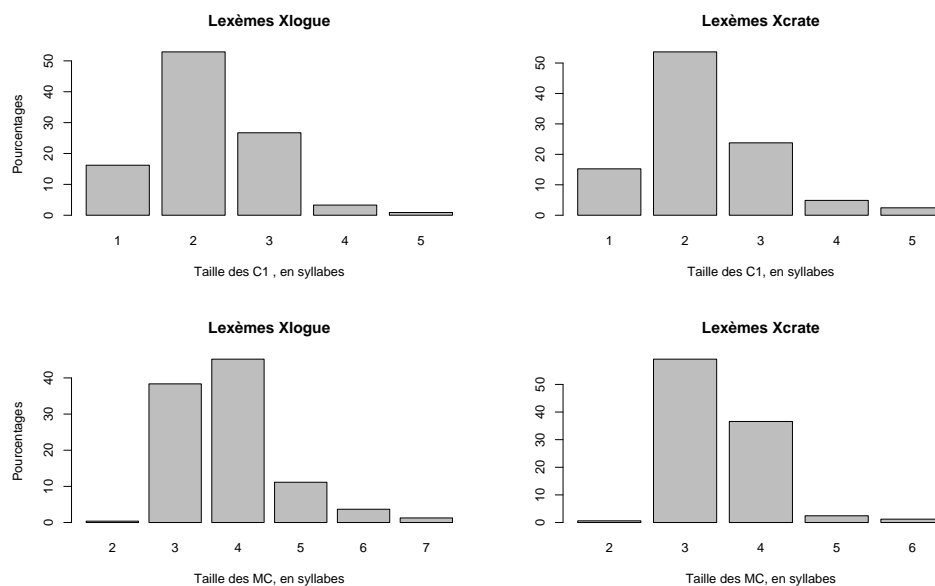


FIGURE 3.24 – Comparaison, pour les lexèmes *Xlogue* et *Xcrate*, de la taille des C1 et de la taille des mots construits

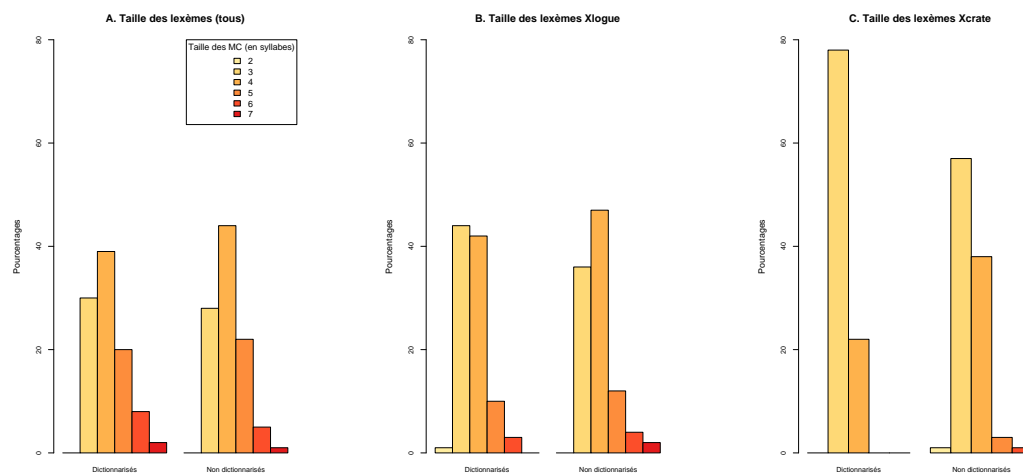


FIGURE 3.25 – Taille des mots construits selon leur dictionnarité, en pourcentages : pour l'ensemble des lexèmes de NeoClassy (A.), pour les lexèmes *Xlogue* (B.) et les lexèmes *Xcrate* (C.)

montre ainsi qu'un C1 natif aura tendance à donner un X plus long qu'un C1 d'origine classique : la médiane se trouve en effet à 3 syllabes pour les premiers contre 2 syllabes pour les seconds.

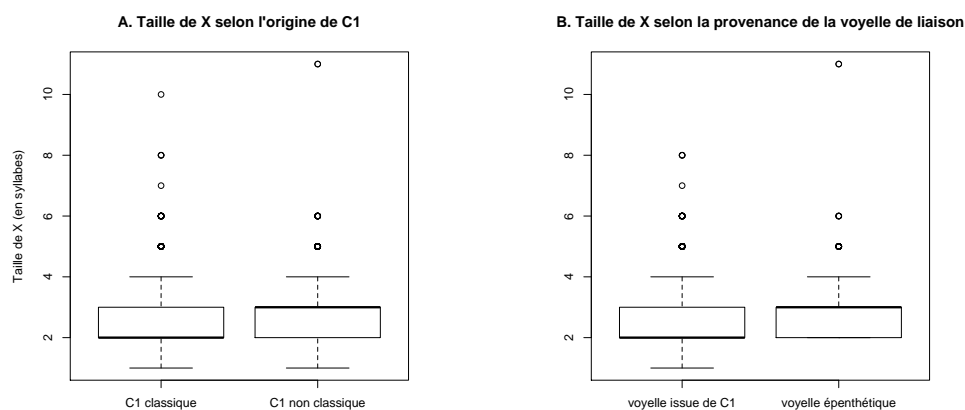


FIGURE 3.26 – Taille de X selon l'origine du C1 (A.) et la provenance de la voyelle de liaison (B.) (boxplot)

Cependant, la Figure B. montre que cette différence est due souvent à l'adjonction de la voyelle de liaison : celle-ci entraîne dans la plupart des cas une resyllabation qui allonge X d'une syllabe (34). Or dans la base de données, comme je l'ai expliqué dans la partie précédente, la voyelle de liaison qui suit un C1 classique est toujours considérée comme issue de celui-ci.

(34) chaussette ($\sigma\sigma$) + *-logue* (σ) → chaussettologue ($\sigma\sigma\sigma\sigma$)

Les accidents morphophonologiques étudiés dans la partie 3.1.1 tels que les insertions, les effacements ou les transformations ont donc un rôle dans la contrainte 5 de mise au gabarit de la construction néoclassique. En effet, la Figure 3.27, qui prend en compte uniquement les C1 natifs, montre que lorsque le C1 est augmenté d'un ou plusieurs segments (le plus souvent la voyelle de liaison), c'est pour construire des lexèmes majoritairement quadrisyllabiques. De plus, peu d'effacements conduisent à des lexèmes courts. Quant aux transformations, elles donnent là aussi lieu majoritairement à des lexèmes quadrisyllabiques.

Il y a ainsi la volonté de faire rentrer les C1 natifs dans des gabarits particuliers, imitant par là les lexèmes créés dans le domaine scientifique. Le trait [+savant] semble donc requérir des lexèmes plus longs. La contrainte de taille dégagée ici est donc davantage lexicale que phonologique.

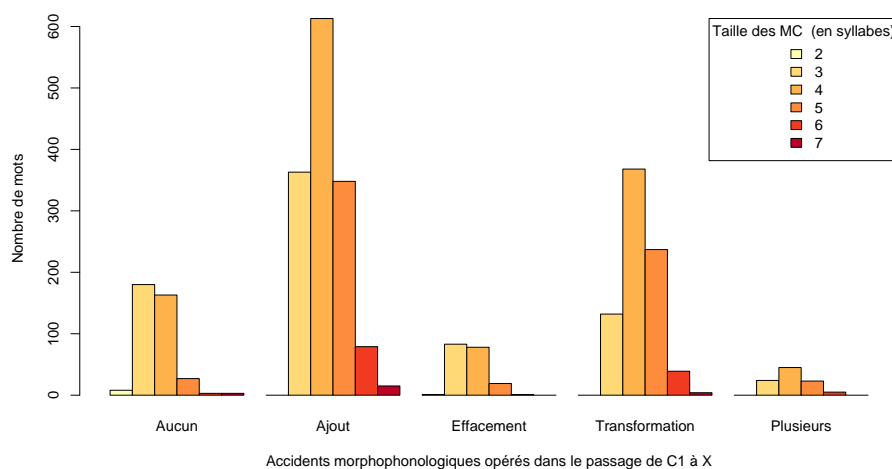


FIGURE 3.27 – Taille des lexèmes de NéoClassy ayant un C1 natif et unique, en fonction des accidents morphophonologiques subis par le C1 (n=2862)

Nous avons vu que, formellement, les constructions néoclassiques telles que celles regroupées dans la base NéoClassy répondent à différentes contraintes. Ainsi, une CN prototypique présente un constituant initial classique ou pseudo-classique, cela pour répondre à une contrainte de crypticité. De plus, la voyelle <o> doit apparaître dans la construction à la jonction des deux constituants. Enfin, une CN est plus longue que les autres mots du lexique, avec quatre syllabes en moyenne. Ces différentes contraintes apportent à la construction ce que j’ai déjà appelé un air savant, scientifique. En effet, des mots longs, constitués d’éléments difficilement interprétables, qui reprennent les codes de la composition classique (la voyelle de liaison par exemple), participent à l’intégration de lexèmes dans un paradigme lexical.

3.2 Organisation du lexique

Après avoir décrit les aspects formels des données de NéoClassy, je vais maintenant m’attacher à les décrire du point de vue de leur sens. Dans un premier temps, les constructions dans lesquelles chaque EN final étudié entre seront considérées séparément. Dans un second temps, j’observerai comment ces différentes constructions s’organisent en différents réseaux lexicaux

3.2.1 Description sémantique

Comme nous le verrons ultérieurement (section 4.1), le sens des lexèmes ne se construit pas en isolation, mais en fonction du contexte d'utilisation. Il sera donc difficile ici de donner le nombre de mots de NeoClassy transmettant un sens particulier. Chaque mot, selon les contextes dans lesquels il est utilisé, pourra avoir des sens différents parfois même contradictoires. Les nombres qui seront donnés dans cette partie sont donc une tendance dégagée pour un schéma en particulier comme étant potentiellement présent dans la base de données : ils ne peuvent pas être considérés comme représentatifs d'une productivité quelconque de ce schéma en français.

3.2.1.1 Du discours et des spécialistes

Concernant l'élément *-logie*, quatre schémas principaux sont à distinguer. Le premier (35) est l'originel. Ce schéma semble peu disponible pour de nouvelles créations. En effet, la quasi totalité des lexèmes en *-logie* de la base NeoClassy qui ont le sème 'discours' ont été créés avant le milieu du XX^{ème} siècle et sont dictionnarisés. Plusieurs sous-schémas se dégagent, selon la façon dont le C1 intervient dans l'acte de discours. Le C1 peut ainsi désigner un moyen par lequel le discours est produit (35-a), moyen qui peut être tant physique (*dactylogologie*) que stylistique (*paradoxologie*). Il peut également qualifier le discours (35-b). Le C1 peut enfin désigner la finalité qu'a le discours (35-c) : annoncer une mort (*nécrologie*), chanter des louanges (*doxologie*). Dans tous les cas, le mot construit aura un sens résultatif, décrivant le discours tel qu'il est produit selon les modalités envisagées.

- (35) $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{discours caractérisé par SEM}_i]_j$
- a. $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{discours exprimé au moyen de SEM}_i]_j$
 - *dactylogologie* (< dactyl- 'doigt') 'langage avec les doigts utilisé par les sourds-muets'
 - *paradoxologie* (< paradoxe) 'art d'user du paradoxe'
 - b. $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{discours qualifié par SEM}_i]_j$
 - *hypologie* (< hyp- 'en dessous') 'incapacité à prononcer des mots de plus de deux syllabes'
 - *périssologie* (< périss- 'superflu, redondant') 'redondance inutile employée pour créer un effet de style'
 - c. $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{discours exprimé dans le but SEM}_i]_j$
 - *nécrologie* (< nécr- 'mort')
 - *doxologie* (< dox- 'gloire') 'prière à la gloire de dieu'

Cependant, si ‘discours’ est le sens du nom grec λόγος, les noms en *-logie* dénotent aujourd’hui surtout des disciplines dont la finalité est l’étude de ce que désigne le C1 (36).

$$(36) \quad [X_i \text{ logie}]_{Nj} \leftrightarrow [\text{étude de SEM}_i]_j$$

Le C1 peut être de natures très diverses, comme l’illustre le Tableau 3.6 qui recense les sens les plus courants que X peut prendre dans le schéma (36).

maladie	cancérologie	< cancer
	ampoulologie	< ampoule
partie du corps	cardiologie	< cardio- ‘cœur’
	chevillologie	< cheville
objet concret	sédimentologie	< sédiment
	chaussettologie	< chaussette
objet abstrait	irénologie	< irén- ‘paix’
	problémologie	< problème
phénomène	courantologie	< courant (marin)
	neigeologie	< neige
individu	mozartologie	< Mozart
	derrickologie	< Derrick
activité, discipline	technologie	< techn- ‘technique’
	régimologie	< régime
animal	primatologie	< primate
	delphinologie	< delphin- ‘dauphin’
plante	algologie	< algue
	orchidologie	< orchidée
lieu, pays, peuple	américanologie	< américain
	balkanologie	< Balkans

Tableau 3.6 – Sens que peut recouvrir X dans le schéma (36)

De plus, des niches de spécialisation peuvent apparaître. Par exemple, parmi les objets, X peut dénoter plus particulièrement une boisson alcoolisée :

$$(37) \quad [X_i \text{ logie}]_{Nj} \leftrightarrow [\text{étude de la boisson SEM}_i]_j$$

- *œnologie* < œno- ‘vin’
- *vinologie* < vin

- *biérologie* < bière
- *cidrologie* < cidre

De même, parmi les individus, le schéma peut s'affiner et X dénoter des êtres surnaturels :

- (38) $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{étude des êtres surnaturels SEM}_i]_j$
- *démonologie* (< démon)
 - *elficologie* (< elfe)
 - *vampirologie* (< vampire)

Les mots donnés dans l'exemple (37) sont formés par analogie avec le mot leader *œnologie*, ceux en (38) avec le mot leader *démonologie*, comme nous le verrons dans la partie suivante 3.2.2.1.

Si le schéma (36) construit des mots se référant à l'étude de ce qui est dénoté par le C1, ils réfèrent également à la discipline qui résulte de la pratique de cette étude. Ce schéma, plus particulièrement lorsque X dénotait une activité, s'est étendu à d'autres pratiques ne nécessitant pas forcément une étude ou une recherche, et le schéma (39) a émergé. A l'intérieur de celui-ci j'ai pu dégager deux sous-schémas, selon qu'il s'agit de la pratique d'une activité ou d'un discours. Les exemples sont donnés en contexte en (40).

- (39) $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{pratique de SEM}_i]_j$
- a. $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{pratique de l'activité SEM}_i]_j$
 - *siestologie* (< sieste)
 - *piscinologie* (< piscine)
 - b. $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{pratique du discours SEM}_i]_j$
 - *excusologie* (< excuse)
 - *rirologie* (< rire)

- (40)
- a. 13h30 Sieste et séance de **siestologie** organisée par le Ministère de la Sieste et des Doigts de Pieds en Éventail. Remise des diplômes aux siesteurs. (<http://www.marseille2013.com/2012/fokonsvoi-sur-le-frioul-samedi-prochain/>)
 - b. La **piscinologie** m'avait été fortement déconseillée ces derniers mois par le corps médical, en des termes laissant peu de place à la fantaisie : "c'est la piscine ou la médecine". (<http://yaelleliane.blogspot.fr/2008/11/comme-une-sirne-sur-le-sable.html>)
 - c. Au fait, Mammouth, j'espère que tu as pris une bonne leçon en "**excusologie**" avec le message de Jéjé donnant plein d'excuses pour dire qu'il ne participerait pas. Je m'inquiète, le temps passe

et je n'ai toujours pas reçu ton excuse. (<http://www.christophegodin.com/forum/viewtopic.php?p=213218&sid=05b218f9d93a44f1a7c53de7234ac0fc>)

- d. Il s'agit d'une méthode de développement personnel par le rire, plus exactement par la "**Riologie**", qui ferait revenir l'optimisme dans la vie de ceux qui la pratiquent. (http://www.pointsdactu.org/article.php3?id_article=1425)

Enfin, un dernier schéma se dégage pour les lexèmes *Xlogie* (41), celui qui provient à l'origine du verbe grec λέγω, 'choisir, recueillir', comme je l'avais abordé dans la partie 2.3.2.1. Le sous-schéma (41-a) sous spécifie ce schéma, en précisant que c'est un ensemble d'œuvres dénombrées par le C1. Un exemple plus anecdotique rentre également dans ce schéma mais en spécifiant que l'ensemble est une communauté de spécialistes (41-b), plaçant ce sous-schéma à la jonction du schéma construisant des noms d'étude et celui construisant des noms d'ensemble.

- (41) $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{ensemble (organique) de SEM}_i]_j$
 ➤ *anthologie* (< anth- 'fleur') 'recueil de poèmes'
 ➤ *critérogie* (< critère)
 ➤ *bouletologie* (< boulet (Définition du GR : 'Personne qui ne sert à rien, ralentit l'activité des autres, constitue une charge pénible.'))
 (voir exemple(42))
- a. $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{ensemble de SEM}_i \text{ œuvres}]_j$
 ➤ *trilogie* (< tri- 'trois')
 ➤ *tétralogie* (tétra- 'quatre')
 ➤ *septalogie* (< sept)
- b. $[X_i \text{ logie}]_{N_j} \leftrightarrow [\text{communauté de spécialistes en SEM}_i]_j$
 ➤ *rockologie* (< rock) (voir exemple (43))
- (42) Mes chers amis, comme vous le savez, la **bouletologie** est en plein essor. Vous avez tous constaté à quel point les serveurs francophones sont à présent saturés de boulets, ces derniers semblant se reproduire à une vitesse hallucinante. . . (<http://www.jeuxvideo.com/forums/1-9725-153063-1-0-1-0-0.htm>)
- (43) Ainsi, on trouve des articles hilarants sur des quasi-inconnus que la **rockologie** a élevé au rang d'artistes cultes tel le fameux Alexander 'Skip' Spence, batteur de Jefferson Airplane et surtout guitariste de Moby Grape (il est à noter que Spence est un artiste fétiche de ce monument de rockologie que sont Les Inrockuptibles). (<http://www.inside-rock.fr/Le-Dictionnaire-Snob-Du-Rock>)

A chaque schéma *Xlogie* que je viens d'énumérer ne correspond pas forcément un schéma *Xlogue*. Ainsi, je n'ai, par exemple, trouvé qu'un seul exemple de lexème *Xlogue* qui contient le sens de 'pratique' sans celui d' 'étude', comme c'est le cas pour le schéma (39).

- (44) typtologue (< typt- 'frapper') : 'qui pratique la typtologie' (définition du TLFi : 'Communication des esprits au moyen de coups frappés par les tables tournantes ou sur un objet quelconque.')

De même, les exemples pour lesquels *Xlogue* signifie 'ensemble de X' sont plutôt anecdotiques. On peut penser au *décalogue*, 'l'ensemble des dix lois, dix commandements'. En (45-a), le lexème *heptalogue* a le sens de 'ensemble constitué par les sept jours de la semaine'. L'exemple (45-b) semble être le pendant de l'exemple (42) : les *chiantologues* constituent la *chiantologie*, quand les *bouletologues* pourraient être les membres de la *bouletologie*. L'exemple (45-c), où *emmerdologues* signifie bien l' 'ensemble des emmerdeurs' est intéressant du fait que le lexème fait suite à une énumération de spécialistes dont les noms sont tous formés avec l'EN *-logue* ; *emmerdologues* englobe ici la liste de ces spécialistes et est donc à interpréter comme l' 'ensemble des spécialistes dont le nom se termine en *-logue* qui sont des emmerdeurs'.

- (45) a. Attendu par certains, surprise pour d'autres, « Mardi » est le deuxième volet de notre "**heptalogue**". À l'hystérie du « Lundi » au travail succède donc la folie douce de la course au pouvoir d'achat. (<http://archipelrouge.blogspot.fr/2008/01/mardi-heptalogue-par-charly-dupuis.html>)
- b. Même si les **Chiantologues** peuvent être divisés du fait de leurs goûts, ils ont des points communs, qui les rassemblent et les unissent pour l'éternité : ils sont chiants, ils ont toujours raison, et ils aiment rien, parce qu'il faut bien le dire, c'était mieux avant, et c'est encore mieux quand c'est pas populaire. (<http://www.jeuxvideo.com/forums/1-36-13390618-1-0-1-0-1-eglise-de-chiantologie.htm>)
- c. On a donc vu, comme l'écrivait hier un journaliste, des associations, des politiciens, des artistes, un président étranger, des "philosophes" en chemise blanche, des psychologues, des sociologues, des politologues (bref, des "**emmerdologues**") s'emparer de "l'affaire". (<http://www.rmcsport.fr/blogs/bourdinaudco.php?post/2006/07/13/Reagissez-a-l-emission-du-jour>)

Comme on vient de le voir avec ce dernier exemple, l'EN *-logue* rentre majoritairement dans le schéma (46), qui est le correspondant du schéma (36).

$$(46) \quad [X_i \text{ logue}]_{N_j} \leftrightarrow [\text{spécialiste de SEM}_i]_j$$

La plupart des mots de NeoClassy sont ainsi des doublets avec d'un côté la discipline en *-logie* et de l'autre le spécialiste de cette discipline en *-logue* :

$$(47) \quad \begin{array}{l} \text{cardiologie / cardiologue} \\ \text{œnologie / œnologue} \\ \text{saucissologie / saucissologue} \end{array}$$

L'EN *-logue* devient donc un élément privilégié pour la formation de noms de spécialistes⁶. Ainsi, dans l'exemple (48), *dentistologue* affiche une double affixation, l'EN *-logue* semblant ici fonctionner comme un intégrateur paradigmatique, au sens de Corbin (1999) : il sert à faire rentrer le mot *dentiste* dans la série des spécialistes, même si le mot contient déjà un affixe *-iste*, déjà créateur de noms de spécialistes (voir section 4.3.2).

$$(48) \quad \text{Bon courage pour le } \mathbf{\text{dentistologue}} \text{ demain, je penserai à toi ma titi et aussi pour le dermato . Moi normalement j'y vais fin février pour me faire retirer un quist sur le tibia (<http://forum.doctissimo.fr/nutrition/surpoids-obesite/laetinou-binomettes-combatt-antes-sujet_169614_55.htm>)$$

Enfin, si des lexèmes *Xlogie* référaient à des discours (schéma (35)), des lexèmes *Xlogue* réfèrent à des formes ou à des parties de discours. Tous les termes formés selon ce schéma ont été soit créés en grec ancien et importés en français, soit créés en français il y a plusieurs siècles. Ce schéma ne semble plus disponible pour les locuteurs actuels.

$$(49) \quad \begin{array}{l} \text{a. } [X_i \text{ logue}]_{N_j} \leftrightarrow [\text{discours caractérisé par SEM}_i]_j \\ \quad \triangleright \textit{dialogue} \text{ (< dia- 'deux')} \\ \quad \triangleright \textit{monologue} \text{ (< mono 'un seul')} \\ \quad \triangleright \textit{tautologue} \text{ (< tauto- 'le même')} \\ \text{b. } [X_i \text{ logue}]_{N_j} \leftrightarrow [\text{partie SEM}_i \text{ d'un discours}]_j \\ \quad \triangleright \textit{épilogue} \text{ (< épi- 'sur, en plus')} \\ \quad \triangleright \textit{prologue} \text{ (< pro- 'avant')} \end{array}$$

6. Comme le notent Villoing et Namer (2012), *-logue* est un concurrent de *-logiste* dans la formation des noms de spécialistes. Selon elles, *-logue* permet la création de noms plus humoristiques quand *-logiste*, qui est en fait une suffixation en *-iste* d'un nom en *-logie*, reste préféré pour le domaine réellement scientifique. Je reviendrai sur la formation des noms de spécialistes dans la partie 4.3.2.

Par ailleurs, deux lexèmes créés avec l'EN *-logue* désignent les auteurs de discours. En (50-a), le lexème *Xlogue* désigne directement l'auteur de X, alors que en (50-b), le mot en *-logue* est lié à celui en *-logie*, puisque *nécrologue* est l'auteur d'une *nécrologie*.

- (50) a. hymnologue 'auteur d'hymnes'
b. nécrologue 'auteur de nécrologie'

Enfin, *-logue* intervient dans un dernier schéma, construisant des noms et des adjectifs, lui aussi peu disponible pour les locuteurs actuels, dans lequel il permet d'établir un certain rapport, souvent caractérisé par le C1 (51) :

- (51) homologue
analogue

La Figure 3.28 résume les différents schémas qui ont été dégagés pour les éléments *-logie* et *-logue*.

Les adjectifs en *-logique* de la base NeoClassy ont tous un correspondant en *-logie* : il ne semble pas y avoir de lexème en *-logique* quand il n'y a pas de lexème en *-logie*, quand bien même le lexème en *-logue* existerait. Cependant, le sens d'un adjectif *Xlogique* peut soit être relié au sens d'un nom *Xlogie*, soit au sens d'un nom *Xlogue*, soit être construit directement sur la base X.

- (52) a. La Fédération Française de Cardiologie apporte son soutien à la recherche **cardiologique** (<http://www.fedecardio.org/qui-sommes-nous/financer-la-recherche/bourses-fedecardio>)
b. Je sais uniquement que j'étais myope à moins 4.75 a l'oeil gauche avec de l'astigmatie (dont je ne connais pas le degré exact étant donné que je ne comprends pas le jargon **ophtalmologique** et que je ne sais pas lire une ordonnance d'ophtalmologue (<http://www.ophtalmologie.fr/operation-myopie/hopital-public-resultat-decevant-quelle-solution-t3840.html>)
c. On est tellement habitués aux tempêtes métaphoriques (financières, médiatiques) qu'on est dépourvus face à une tempête **météorologique**. (http://www.liberation.fr/chroniques/2009/01/31/-_306669)

L'utilisation de l'EN *-logique* pour la construction d'adjectifs relationnels est l'objet de Lasserre et Montermini (2015) et sera abordée dans la partie 4.3.1.

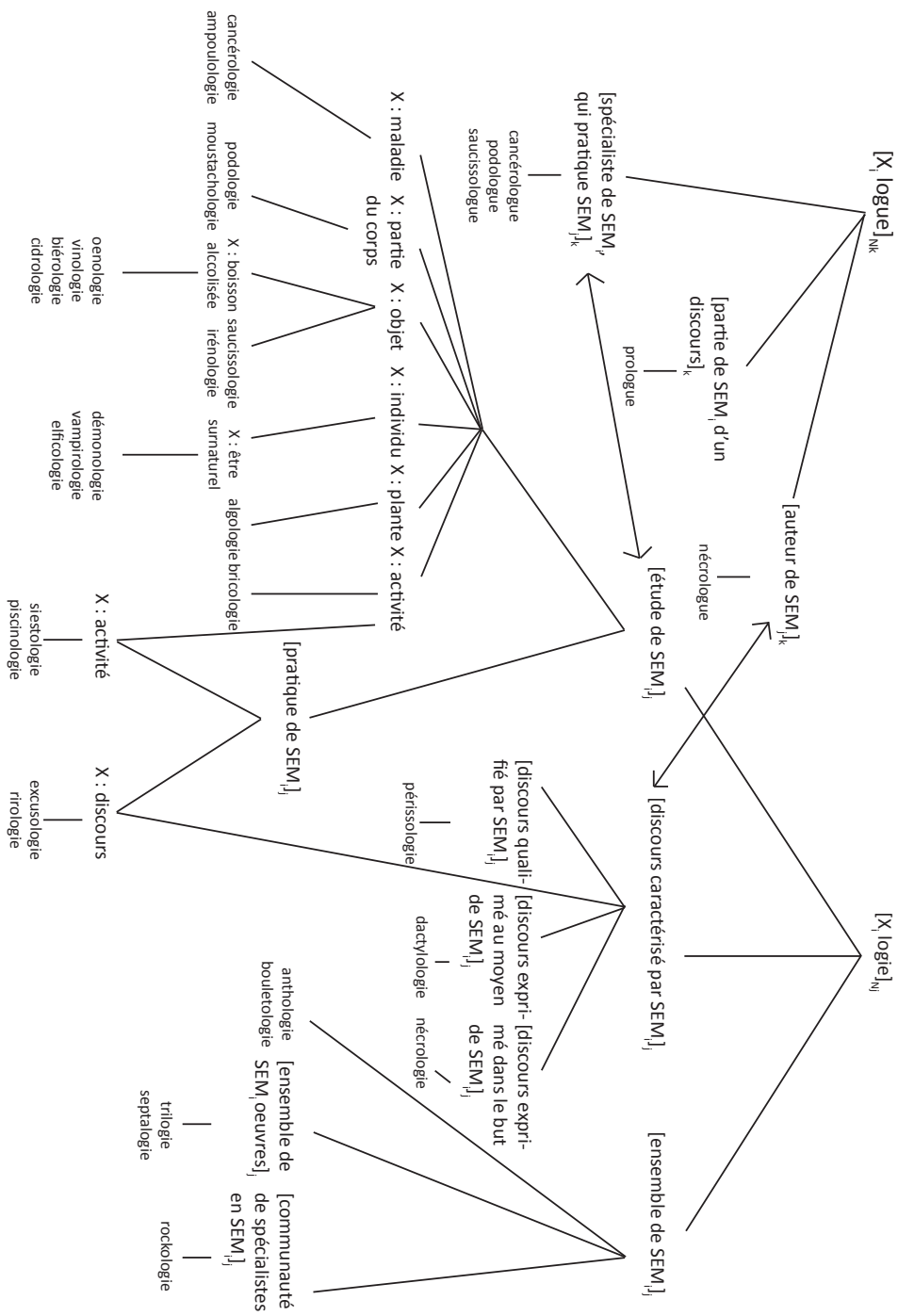


FIGURE 3.28 – Différents schémas impliquant l'élément néoclassique -logie

3.2.1.2 Du pouvoir et des coalitions

Comme je l'avais abordé dans la section 2.3.2, *κράτος* signifie en grec 'pouvoir'. Ce sens se retrouve dans un bon nombre de lexèmes en *-cratie* ou *-crate* du français moderne. Les lexèmes *X-cratie* désignent dans ce cas un mode de gouvernance ou un état dans lequel ce mode est mis en œuvre, et les lexèmes *Xcrate* nomment soit les personnes au pouvoir dans cet état, soit les partisans de cet état ou de ce mode de gouvernance. Quant *Xcratie* désigne un mode de gouvernance, le pouvoir est caractérisé par le C1. Celui-ci désigne soit l'agent (la personne, l'institution, la communauté,... qui exerce le pouvoir) (53-a-i), soit l'instrument (53-a-ii). Plus rarement, C1 peut désigner un but (53-a-iii) ou un lieu (53-a-iv).

- (53) a. (i) $[X_i \text{ cratie}]_{N_j} \leftrightarrow [\text{pouvoir exercé par SEM}_i]_j$
 > *démocratie* (< dém- 'peuple')
 > *particratie* (< parti) (exemple (54-a))
 (ii) $[X_i \text{ cratie}]_{N_j} \leftrightarrow [\text{pouvoir exercé au moyen de SEM}_i]_j$
 > *clicocratie* (< clic) 'pouvoir exercé grâce à Internet'
 (exemple (54-b))
 (iii) $[X_i \text{ cratie}]_{N_j} \leftrightarrow [\text{pouvoir dans le but SEM}_i]_j$
 > *pognocratie* (< pognon) (exemple (54-c))
 (iv) $[X_i \text{ cratie}]_{N_j} \leftrightarrow [\text{pouvoir dans le lieu SEM}_i]_j$
 > *thalassocratie* (< thalass- 'mer') (exemple (54-d))
 b. $[X_i \text{ crate}]_{AN_k} \leftrightarrow [\text{qui exerce / partisan de SEM}_j]_k$
- (54) a. La **particratie** sera considéré ici comme l'un des modèles de régimes politiques existant dans les démocraties parlementaires. (<http://www.larevuetoudi.org/fr/story/la-particratie-handicap-de-lid%C3%A9e-wallonne>)
 b. Pour cela, il a su s'entourer de jeunes professionnels de l'Internet qui apparaissent comme un nouveau centre de pouvoir dénomé "**Clicocratie**" par les médias. (<http://www.les-infostrategies.com/actu/0902652/la-clicocratie-au-service-de-barack-obama-qualifie-de-president-web-20>)
 c. Fricarchie, **pognocratie** et église d'oseillologie. L'argent est devenu, depuis longtemps d'ailleurs, le maître du monde, le dieu devant lequel tous doivent plier le genou et courber le dos. Aujourd'hui, l'intérêt financier est l'"ultima ratio", l'argument de toute politique individuelle et collective déterminée, le critère de décision majuscule des personnes et le squelette des relations humaines de quelque inspiration que ce soit. (<http://lemeilleur.sdesmondes.blogs.nouvelobs.com/archive/2013/01/13/fric>)

[archie-et-pognocratie.html](#))

- d. Agamemnon, bien que le centre de sa puissance ait été sur le continent, aurait aussi exercé une véritable **thalassocratie** grâce à la flotte dont il avait le contrôle et qui lui aurait permis de mener l'expédition contre Troie. (<http://www.universalis.fr/encyclopedie/thalassocratie/>)

Un autre schéma, construisant des composés exocentriques et désignant l'état dans lequel le pouvoir est exercé, se dégage également. Le C1 peut également désigner un individu particulier :

- (55) a. (i) $[X_i \text{ cratie}]_{N_j} \leftrightarrow [\text{état dans lequel le pouvoir est détenu par SEM}_i]_j$
 > *Sarkocratie* (< Sarkozy) : « Bienvenue en Sarkocratie » (<http://www.pantheresroses.org/Pourquoi-pas-encore-plus-de.html>)
 (ii) $[X_i \text{ cratie}]_{N_j} \leftrightarrow [\text{état dans lequel le pouvoir est exercé au moyen de SEM}_i]_j$
 > *corruptocratie* (< corruption)
 b. $[X_i \text{ crate}]_{AN_k} \leftrightarrow [\text{partisan de SEM}_j]_k$

Un même lexème peut avoir été construit selon ces deux schémas, par exemple le plus fréquent, *démocratie*, qui peut désigner soit un 'pouvoir exercé par le peuple' (56-a), soit l'état dans lequel le pouvoir est exercé par le peuple (56-b).

- (56) a. La **démocratie** est le pire des régimes – à l'exception de tous les autres déjà essayés dans le passé. (Winston Churchill) (http://www.lemonde.fr/idees/article/2015/04/29/la-democratie-un-regime-malade_4624979_3232.html)
 b. Et la Suisse est une **démocratie**, contrairement à la France qui est une corruptocratie. C'est d'ailleurs ce qui fait d'elle un pays du Tiers-Monde. (<http://www.jeuxvideo.com/forums/1-51-60282496-2-0-1-0-pour-moi-la-belgique-suisse-luxembourg.htm>)

De plus, des lexèmes construits sur des bases soit synonymes, soit d'origines diverses mais ayant le même sens, peuvent désigner le même référent. L'exemple (57) présente des lexèmes signifiant 'le pouvoir des femmes', le C1 étant soit un allomorphe de femme (*fémocratatie*), soit un élément néoclassique signifiant 'femme' (*gynécocratie*, *gynocratie*), soit une partie de l'ana-

tomie des femmes représentant celles-ci par métonymie ⁷ :

- (57) gynocratie, gynécocratie, féminocratie, clitocratie, vaginocratie, vulvocratie

Des schémas (53) a dérivé un sous-schéma dans lequel le trait ‘abusif’ s’est ajouté :

- (58) a. $[X_i \text{ cratie}]_{Nj} \leftrightarrow [\text{pouvoir exercé de manière abusive par SEM}_i]_j$
 ➤ *bureaucratie* (< bureau)
 b. $[X_i \text{ crate}]_{Nk} \leftrightarrow [\text{qui exerce SEM}_j]_k$
 ➤ *bureaucrate*

Du schéma (58-b) découle un autre schéma qui, lui, ne semble pas avoir de correspondant *Xcratie* :

- (59) $[X_i \text{ crate}]_{Nk} \leftrightarrow [\text{SEM}_i \text{ s'attribuant un pouvoir non mérité ou démesuré}]_k$
 ➤ *fonctionnaucrate* (< fonctionnaire) : « cette petite association de **fonctionnaucrate** me reproche d’avoir gagné deux sous et me font un redressement (décidément je dois les emmerder!) alors et merde, je renonce définitivement à toute réinsertion sociale » (<http://www.sotty.com/economie.html>)

Cependant, si la notion de pouvoir est au cœur même du lexème grec *κράτος*, elle peut être amenée à disparaître dans des lexèmes français construits avec l’EN *-cratie* : si *aristocratie* signifiait à l’origine ‘le pouvoir des meilleurs’, le sens de ‘pouvoir’ a aujourd’hui disparu, et le mot ne désigne plus qu’un ‘groupe fermé’. Un nouveau schéma a émergé (60-a) :

- (60) a. $[X_i \text{ cratie}]_{Nj} \leftrightarrow [\text{groupe social fermé constitué de SEM}_i \text{ s'auto-valorisant}]_j$
 ➤ *aristocratie* (< arist- ‘meilleur’)
 ➤ *cathocratie* (< catho(lique))
 ➤ *bobocratie* (< bobo (BOurgois BOhème))
 b. $[X_i \text{ crate}]_{ANk} \leftrightarrow [\text{membre de SEM}_j]_k$

Les lexèmes *Xcrate* correspondants désignent les membres de ce groupe .

Un autre schéma n’ayant pas non plus la notion de ‘pouvoir’ se dégage parmi les données de NeoClassy. Il concerne uniquement des lexèmes *Xcratie*, n’ayant pas de correspondance *Xcrate* (61), tout comme le schéma (59)

7. Ces lexèmes ne sont pas de parfaits synonymes puisque des nuances sont apportées, avec par exemple des traits plus ou moins péjoratifs.

n'avait pas de correspondant *Xcratie* :

- (61) $[X_i \text{ cratie}]_{N_j} \leftrightarrow [\text{omniprésence de SEM}_i]_j$
 ➤ *footocratie* (< foot(ball)) : « Ca fait 2 semaines qu'on se bouffe du foot à la télé, dans les bars, PARTOUT! NON A LA **FOOTOCRATIE!** » (<http://www.jeuxvideo.com/forums/1-8173-2716311-1-0-1-0-0.htm>)
 ➤ *festivocratie* (< fête) : « C'est donc, selon Muray, le confort intellectuel procuré par l'idée d'une « **festivocratie** », à savoir une société submergée par le Spectacle au point d'en adopter les codes et la didascalie, qui fait naître chez Homo festivus une résignation inconsciente face aux événements qui se déroulent sous ses yeux et qui contredisent sa pensée. » (<http://comptoir.org/2014/12/01/du-spectacle-au-festivisme-petite-histoire-de-la-diversion-contemporaine/comment-page-1/>)

Ainsi, une absence de correspondance, au sein de NeoClassy, entre un lexème *Xcratie* et un lexème *Xcrate* peut être fortuite, comme en (62-a) où la création *argencrate* est possible mais soit ne répond pas à un besoin de dénomination ou a été construite mais pas accessible sur Internet via les moteurs de recherche, ou être la conséquence d'une instruction sémantique uniquement attachée à l'un ou l'autre des éléments, comme c'est le cas en (62-b) où *macdocratie* est formé selon le schéma (61) pour lequel il n'y a pas de correspondant *Xcrate*.

- (62) a. argencratie (< argent) 'état dans lequel le pouvoir est exercé au moyen de l'argent'
 °argencrate 'partisan de cet état'
 b. macdocratie (<MacDo) 'omniprésence des restaurants McDonald's'
 *macdocrate

Pour reprendre les termes de Corbin (1997), l'exemple (62-a) reflète une lacune lexicale accidentelle, *argencrate* fait partie du lexique virtuel, alors que l'exemple (62-b) représente une lacune lexicale non accidentelle, puisque *macdocrate* n'est pas actualisable, du moins avec un sens correspondant à celui de *macdocratie* 'omniprésence des restaurants McDonald's'⁸.

Par ailleurs, un même lexème peut avoir été formé selon des schémas diffé-

8. Cependant, on peut tout à fait imaginer l'actualisation de *macdocrate* 'personne exerçant un pouvoir grâce à MacDonald's' ou 'personne exerçant un pouvoir au sein des restaurants MacDonald's'.

rents, impliquant ou non la notion de pouvoir : en (63-a), *mafia-cratie* désigne le ‘pouvoir exercé par la mafia’ (schéma (53-a-i)), en (63-b), le mot est construit selon le schéma (55-a-ii) et signifie un ‘état gouverné par la mafia’ et en (63-c) le mot désigne un ‘groupe composé de mafieux’ (schéma (60-a)).

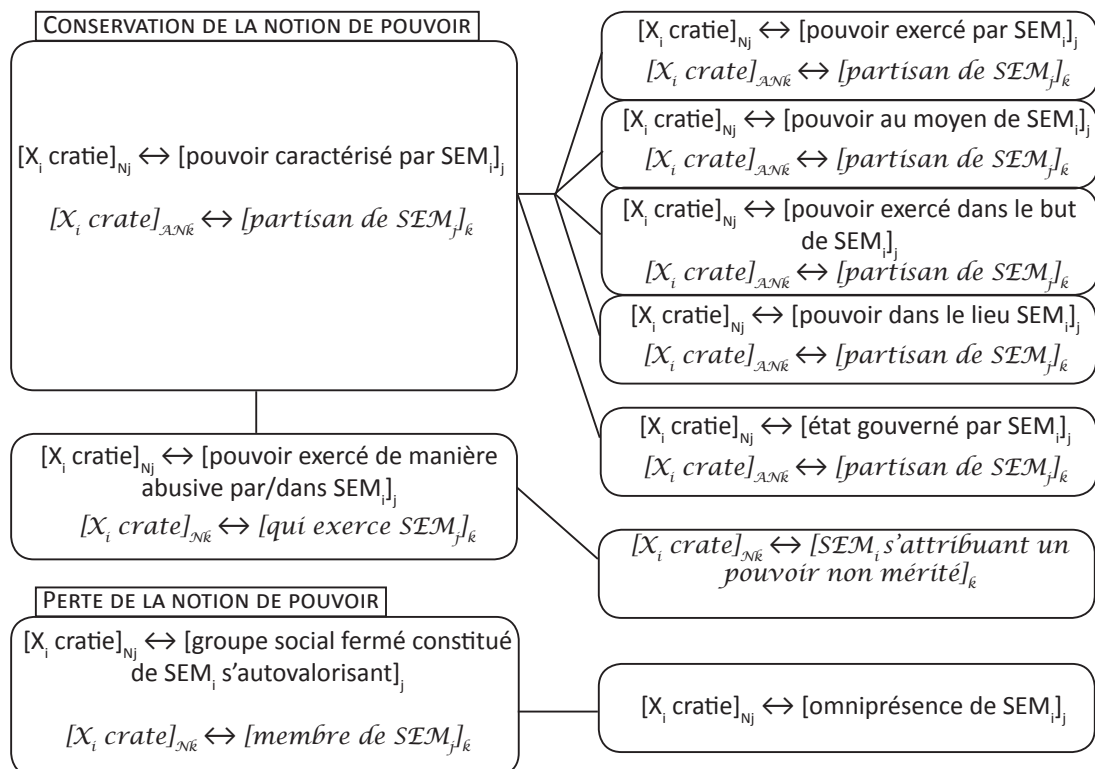
- (63) a. Quand on a des convictions, on les défend à visage découvert, pas caché derrière un masque : dans une vraie démocratie tout à fait d’accord. Hélas pas dans une **mafia-cratie** ou un état qui ne reconnaît pas son peuple après chaque élection, ce serait un geste suicidaire et stupide. (http://www.lepoint.fr/debats/anonymous-bandits-ou-justiciers-18-02-2012-1432682_34.php)
- b. La France est une vraie oligarchie et très près d’une **mafia-cratie** et c’est ça la différence, le peuple n’a absolument rien à dire. (<http://plus.lefigaro.fr/article/la-france-perd-son-aaa-20120113-661943/commentaires/8397506>)
- c. En rapides Tableaux, la caméra montre le ‘casse légal du siècle’ opéré par l’ex-président argentin, avec la complicité active de la ‘**mafia-cratie**’ : ministres, parlementaires, industriels, syndicalistes... (http://web.radicalparty.org/pressreview/print_250.php?func=detail&par=9458)

La Figure 3.29 synthétise les schémas qui ont été dégagés pour les lexèmes *Xcratie* et *Xcrate*.

3.2.1.3 De la peur à la haine

L’élément *-phobie* est le seul de la base NeoClassy à correspondre à un lexème qui est utilisé de façon autonome. Le mot *phobie* signifie ‘peur, crainte angoissante spécifiquement liée à certains objets, certains actes, certaines situations, certaines idées’ (définition du GR). Le schéma (64-a) construit des noms de qualité ayant le sens de ‘peur’; le schéma (64-b) qualifie ou réfère aux personnes atteintes par cette peur.

- (64) a. $[X_i \text{ phobie}]_{N_j} \leftrightarrow [\text{peur de SEM}_i]_j$
 ➤ *agoraphobie* (< agora ‘place publique, ici foule’)
 ➤ *cafardophobie* (< cafard)
 ➤ *nyctophobie* (< nyct- ‘nuit’)
- b. $[X_i \text{ phobe}]_{AN_j} \leftrightarrow [\text{qui a peur de SEM}_i]_j$
 ➤ *claustrophobe* (< claustr- ‘enfermé, verrouillé’)
 ➤ *thalassophobe* (< thalass- ‘mer’)
 ➤ *cancérophobe* (< cancer)

FIGURE 3.29 – Différents schémas impliquant l'élément néoclassique *-cratie*

La correspondance entre le nom autonome *phobie* et le schéma (64-a) a fait dire à Véronique Genest, comédienne tenant la tête d’affiche dans la série policière française *Julie Lescaut* et célèbre pour ses publicités pour du jambon :

« Je suis islamophobe, et phobie veut dire peur. Donc oui, peut-être oui, probablement, je suis islamophobe comme beaucoup de Français. J’ai peur de l’islam comme on a peur d’une chose que l’on ne connaît pas. » (Véronique Genest)⁹

Morgan Freeman, acteur américain autrement plus célèbre à l’international et surtout autrement plus talentueux, aurait pu lui répondre :

« I hate the word homophobia. It’s not a phobia, you’re not scared. You are just an asshole. » (Morgan Freeman sur Twitter)

En effet, les noms en *-phobie* peuvent aussi être des noms de sentiment dénotant une aversion (65-a) ou, quand C1 est [+humain], une réelle hostilité (66-a). Ce passage s’est fait par métonymie, puisque une peur panique de quelque chose sous-entend une répulsion. Cottez (1980) fait remonter le sens ‘hostilité’ en français à Joseph de Maistre, en 1821, avec le terme *théophobie* (< théo- ‘dieu’).

- (65) a. $[X_i \text{ phobie}]_{N_j} \leftrightarrow [\text{aversion envers SEM}_i]_j$
 ➤ *théâtrophobie* (< théâtre)
 ➤ *technophobie* (< technologie)
 ➤ *hiverophobie* (< hiver)
- b. $[X_i \text{ phobe}]_{AN_j} \leftrightarrow [\text{qui n’aime pas SEM}_i]_j$
 ➤ *accentophobe* (< accent)
 ➤ *cadeauphobe* (< cadeau)
 ➤ *rapophobe* (< rap)
- (66) a. $[X_i \text{ phobie}]_{N_j} \leftrightarrow [\text{hostilité envers SEM}_i]_j$
 ➤ *xénophobie* (< xén- ‘étranger’)
 ➤ *juvénophobie* (jeune)
 ➤ *bobophobie* (< bobo (acronyme de BOurgois BOhème))
- b. $[X_i \text{ phobe}]_{AN_j} \leftrightarrow [\text{qui est hostile envers SEM}_i]_j$
 ➤ *roxanophobe* (< Roxane)
 ➤ *musulmanophobe* (< musulmans)
 ➤ *autrophobe* (< autres)

9. http://www.lexpress.fr/actualites/1/styles/veronique-genest-auto-proclamee-islamophobe-cree-la-polemique-sur-le-net_1162627.html

Le schéma (66-a) peut prendre comme C1 des noms collectifs (noms de peuples, de pratiquants d'une religion,...) ou individuels (le schéma désignera alors une hostilité envers une personne pleinement désignée). On peut distinguer des sous-schémas qui se sont formés, par exemple quand le C1 désigne une orientation sexuelle (67-a) :

- (67) a. $[X_i \text{ phobie}]_{Nj} \leftrightarrow [\text{hostilité envers des personnes ayant SEM}_i \text{ comme orientation sexuelle}]_j$
 ➤ *homophobie* (< homosexuels)
 ➤ *lesbophobie* (< lesbiennes)
 ➤ *transphobie* (< transsexuels)
 ➤ *biphobie* (< bisexuels)
 ➤ *allophobie* (< allosexuels 'personne qui a une autre sexualité que l'hétérosexualité')
 ➤ *gayphobie* (< gay)
- b. $[X_i \text{ phobe}]_{ANj} \leftrightarrow [\text{qui est hostile envers des personnes ayant SEM}_i \text{ comme orientation sexuelle}]_j$

A côté de ces trois schémas principaux, fortement utilisés dans la langue courante, un schéma spécifique au domaine chimique a également émergé (68-a), contenant là aussi le sème 'répulsion' mais non plus comme un sentiment, plutôt comme une répulsion naturelle.

- (68) a. $[X_i \text{ phobie}]_{Nj} \leftrightarrow [\text{répulsion chimique envers SEM}_i]_j$
 ➤ *hydrophobie* (< hydr- 'eau')
 ➤ *calciphobie* (< calc- 'calcaire')
- b. $[X_i \text{ phobe}]_{ANj} \leftrightarrow [\text{qui est chimiquement non attiré par SEM}_i]_j$
 ➤ *chromophobe* (< chrom- 'couleur') 'Se dit de cellules qui possèdent une affinité réduite ou nulle pour les colorants, comme certaines cellules du lobe antérieur de l'hypophyse' (définition du TLFi)
 ➤ *lipophobe* (lip- 'graisse')

Parfois, ce n'est pas le référent direct du C1 qui est l'objet de la peur ou de l'hostilité. Par exemple, *bagnolophobie* (69) ne fait pas ici référence à une peur panique des voitures, mais à la peur de ne plus avoir sa voiture. Le sens devient donc antinomique par rapport au schéma (64-a).

- (69) Ou alors, il faut la faire évoluer dans le sens de la **bagnolophobie**, la peur d'être séparé de sa voiture. . . Avec un litre d'essence bientôt à 2 euros, on sera tous guéris dans très peu de temps. . . (<http://www.franceinfo.fr/decryptage/il-etait-une-mauvaise-foi/etes-vous>)

s-nomophobe-537795-2012-02-24)

De plus, comme c'est le cas pour les autres ENs étudiés, un même mot peut être formé selon deux schémas différents ; *bagnolophobie* peut ainsi désigner, par métonymie, une hostilité envers les automobilistes :

- (70) C'est vrai que c'est une magnifique voiture, que bon nombre d'entre nous (moi le premier) n'auront jamais les moyens d'acheter. Ça fait surtout plaisir de voir que malgré les radars, le CO2, et toute la "**bagnolophobie**" actuelle, il y a encore des constructeurs assez dingues pour nous faire rêver avec de telles machines. (<http://www.caradisiac.com/Mercedes-SL65-AMG-Black-Series-dans-la-nature-12-photos-26291.htm>)

Un autre exemple parmi d'autres est celui d'*acnophobe*(71) :

- (71) acnophobe (< acné) :
- a. 'peur d'avoir de l'acné'
moi jchui **acnophobe** l comme toute adolescente. (<http://smismi67.skyrock.com/998245092-Phobie.html>)
 - b. 'hostilité envers les personnes ayant de l'acné'
C'est du racisme anti-boutonneux, voilà ce que c'est ! De l'acnophobie. Vous êtes des **acnophobes**, tous autant que vous êtes, j'en suis persuadé ! (<http://me-myself-and-us-lea.skyrock.com/3.html>)

La peur n'étant jamais très éloignée de l'hostilité, ces exemples sont nombreux. La construction du sens des mots formés avec l'élément *-phobie* sera étudiée plus en détails dans la partie 4.1.1.

3.2.1.4 De la voix qui sonne

L'élément néoclassique *-phone* rentre dans différents schémas. Le sens de 'son' est toujours présent, mais selon des modalités différentes. Les constructions *Xphone* peuvent d'abord former des noms d'instruments. Ces noms désigneront soit des instruments de musique (72), soit des appareils (73). Pour les premiers, le C1 sera en général soit un nom dont le référent désigne ce qui cause le son émis de l'instrument (72-a), soit un nom propre désignant l'inventeur de l'instrument (72-b). Plus anecdotiquement, le C1 peut désigner le type de musique auquel est destiné l'instrument (72-c) ou ce à quoi ressemble le son émis par l'instrument (72-d).

- (72) a. $[X_i \text{ phone}]_{Nj} \leftrightarrow [\text{instrument dont la musique est provoquée par SEM}_i]_j$
 ➤ *xylophone* (< xyl- ‘bois’)
 ➤ *vibraphone* (< vibration) ‘Instrument à percussion dérivé du xylophone, en usage surtout dans les orchestres de jazz, dont les vibrations émises par les tubes de résonance placés sous chacune des lames métalliques sont amplifiées électriquement.’ (définition du TLFi)
 ➤ *bouteillophone* (< bouteille)
- b. $[X_i \text{ phone}]_{Nj} \leftrightarrow [\text{instrument de musique inventé par SEM}_i]_j$
 ➤ *sudrophone* (< Sudre)
 ➤ *sarussophone* (< Sarus)
- c. $[X_i \text{ phone}]_{Nj} \leftrightarrow [\text{instrument dont le but est de faire une musique SEM}_i]_j$
 ➤ *jazzophone* (< jazz) : Face à la «boîte à frissons» (accordéon) ample et généreuse de Galliano, Portal joue la dynamique des contrastes et l’alternance des timbres : ardeur de la clarinette, velouté de la clarinette basse (...), tendresse du **jazzophone**, instrument des années 30 méconnu (Ten Years Ago, jolie composition de Richard pour Barbara, qu’il a accompagnée durant une décennie). (http://www.lexpress.fr/informations/souffler-c-est-jouer_624471.html)
- d. $[X_i \text{ phone}]_{Nj} \leftrightarrow [\text{instrument dont le son ressemble à SEM}_i]_j$
 ➤ *brontosauophone* (< brontosauure) : Instrument de musique fictif inventé par Gaston Lagaffe « « Mammouthesque instrument de musique qui serait à la valse musette ce que le char Leclerc est au tourisme vert », le gaffophone est physiquement de très grande taille et d’une esthétique douteuse. En le voyant pour la première fois, Fantasio proposa de le baptiser « **brontosauophone** » » (<https://fr.wikipedia.org/wiki/Gaffophone>)

Xphone comme nom d’instrument peut aussi désigner un appareil, en général un appareil électro-acoustique. Le C1 dénotera en général ce qui est enregistré par l’appareil (73-a). Il pourra aussi, plus rarement, désigner la personne ayant inventé l’appareil (73-b), ou une caractéristique physique de l’appareil (73-d). Le nom *Xphone* peut également dénommer un appareil dont le C1 dénote le son destiné à être enregistré (73-c) ou la finalité de l’appareil (73-e).

- (73) a. $[X_i \text{ phone}]_{Nj} \leftrightarrow [\text{appareil destiné à enregistrer SEM}_i]_j$
 ➤ *gramophone* (< gram- ‘lettre’)

- *bordelophone* (< bordel)
- *graphophone* (< graph- ‘écrire’)
- *aviophone* (< avion)
- *parlophone* (< parler)
- b. $[X_i \text{ phone}]_{N_j} \leftrightarrow [\text{appareil inventé par SEM}_i]_j$
 - *dupinophone* (< Dupin)
- c. $[X_i \text{ phone}]_{N_j} \leftrightarrow [\text{appareil destiné à écouter le son provoqué par SEM}_i]_j$
 - *cassetophone* (< cassette)
 - *géophone* (< gé- ‘terre’) ‘Instrument d’écoute servant à détecter les bruits provenant du sol (travaux de sape, prospection, localisation de mineurs ensevelis); sismographe utilisé dans la prospection sismique’ (définition du TLFi)
- d. $[X_i \text{ phone}]_{N_j} \leftrightarrow [\text{appareil qui ressemble à SEM}_i]_j$
 - *beaubourophone* (< Beaubourg)
- e. $[X_i \text{ phone}]_{N_j} \leftrightarrow [\text{appareil dont le résultat est SEM}_i]_j$
 - *bruitophone* (< bruit)
 - *optophone* (< opt- ‘vue, vision’) ‘Appareil qui transforme les ondes lumineuses en ondes sonores et qui peut être utilisé par des aveugles, notamment pour leur permettre de lire’ (définition du TLFi)
 - *audiophone* (< audi- ‘entendre’) ‘Appareil acoustique destiné aux personnes atteintes de surdité, utilisant les os du crâne pour la transmission des vibrations sonores au labyrinthe’ (définition du TLFi)

Deux lexèmes de NeoClassy, dont le C1 désigne une activité (*cinéma*) ou un objet permettant une activité (*livre*) sont construits selon un schéma dans lequel l’EN *-phone* sert à qualifier le C1 de sonore.

- (74) $[X_{N_i} \text{ phone}]_{N_j} \leftrightarrow [\text{SEM}_i \text{ qualifié de sonore}]_j$
- *cinéphone* ‘cinéma parlant’
 - *livrophone* ‘livre sonore’

Dans le domaine de la linguistique, l’EN *-phone* est également présent (75). Le C1 désigne ici en général une modalité de réalisation du son :

- (75) *homophone* : ‘dont le son est retranscrit à l’identique’
allophone : ‘variante combinatoire d’un phonème’
polyphone : ‘qui représente plusieurs sons différents’
idéophone : ‘expression imitative’

Un schéma particulièrement disponible, sinon le plus présent, au moins dans la base NeoClassy, est celui pour lequel le C1 est un ethnique, c'est-à-dire un gentilé construit sur un nom de pays ainsi que l'adjectif correspondant, sur lequel est construit, par conversion, un nom de langue (*français, anglais, espagnol,...*) ; le lexème *Xphone* est un nom (77-a) ou un adjectif qui se rapporte au locuteur de la langue (77-b), ou à la langue elle-même (77-c).

- (76) a. $[X_i \text{ phone}]_{ANj} \leftrightarrow [\text{qui parle, où l'on parle SEM}_i / \text{locuteur de SEM}_i]_j$
 ➤ *francophone* (< français)
 ➤ *catalanophone* (< catalan)
 ➤ *hellenophone* (< hallen- 'grec')
- (77) a. Un **francophone** sur trois en Amérique du Nord n'est pas Québécois (<http://www.rcinet.ca/fr/2015/06/24/un-francophone-sur-trois-en-amerique-du-nord-est-pas-quebecois/>)
- b. Le stage non rémunéré en milieu **francophone** est conçu pour donner l'occasion aux étudiants de mettre en pratique leurs compétences de communication et créer des relations avec les membres de la communauté **francophone** de la Saskatchewan. (http://institutfrancais.uregina.ca/index.php/general/list/menu_id/442)
- c. l'approche d'enseignement d'une école d'immersion française est axée sur le français langue seconde, alors que les écoles **francophones** enseignent en français langue première (<http://developpement-langagier.fpfcb.bc.ca/fr/r%C3%B4le-des-parents-quand-les-deux-parents-sont-francophones>)

Si dans le schéma (76-a) X désigne en général une langue en particulier, il peut aussi désigner un type de langue (78-a), un type de personne (78-b), un lieu ou la communauté attachée à ce lieu (78-c), ou alors, dans l'exemple (78-d), les lexèmes sont construits par opposition à d'autres lexèmes *Xphone*, désignant des locuteurs d'autres langues que celles citées dans le contexte d'apparition du mot :

- (78) a. *dialectophone* (< dialecte)
patoisphone (< patois)
 b. *barbarophone* (< barbare)
 c. *banlieuphone* (< banlieue)
 d. *alterphone* (< alter- 'autre') 'qui parle une autre langue'
autophone (< autre) 'qui parle une autre langue' : « Preuve

en est les divers billets édités par les autres blogs consacrés au sujet, qu'ils soient francophones ou anglophones ou « **autrophone** ». » (<http://www.actulogo.fr/2012/04/nettoyage-de-printemps/>)

Enfin, l'EN *-phone* peut être l'aphérèse du mot *téléphone* (79), se rapprochant par là des phénomènes de mot-valisage ou de sécrétion que j'aborderai dans la partie 5.2.2.2 de cette thèse.

- (79) a. $[X_i \text{ phone}]_{Nj} \leftrightarrow [\text{téléphone qui est un SEM}_i]_j$
 ➤ *gadgetophone* (< gadget) : Les entreprises font de la réplique de données sur le Nuage et les employés autorisés y accèdent actuellement par BlackBerry, iPhone et autres «**gadgetophones**» (<http://www.directioninformatique.com/di/client/fr/DirectionInformatique/Nouvelles.asp?id=57946>)
- b. $[X_i \text{ phone}]_{Nj} \leftrightarrow [\text{téléphone qui a pour but SEM}_i]_j$
bisouphone (< bisou) : « Vous avez vu les nouvelles annonces publicitaires ? SOS fantasmes ? La ligne brûlante ? Minouphone ? **Bisouphone** ? Téléphone câlin ? Et si j'appelais Marilyn ? » SOLLERS Philippe, *Le Coeur absolu*, 1987, p. 115, III

3.2.1.5 Manger, dévorer, consommer

Le sens prototypique des constructions impliquant l'EN final *-phage* est celui indiqué dans le schéma (80). Le premier constituant est l'objet de l'élément verbal *-phage*.

- (80) $[X_i \text{ phage}]_{ANj} \leftrightarrow [\text{qui mange SEM}_i]_j$:
 ➤ *anthropophage* (< anthrop- 'homme')
 ➤ *planctophage* (< plancton)
 ➤ *entomophage* (< entom- 'insecte')

Ces mots construits désignent souvent un critère classificatoire d'espèces animales ou caractérisent une espèce :

- (81) phytophage (< phyt- 'plante') 'qui se nourrit de matières végétales'
 coprophage (< copr- 'excrément') 'qui se nourrit d'excréments'
 larviphage (< larve) 'qui se nourrit de larves'

Les constructions formées sur le schéma (80) peuvent aussi désigner des régimes ou habitudes alimentaires suivis par des hommes :

- (82) anthropophage (< anthrop- ‘homme’) ‘qui mange d’autres hommes’
 hippophage (< hipp- ‘cheval’) ‘qui mange de la viande de cheval’

Par extension, les constructions en *-phage* peuvent désigner des amateurs d’un aliment particulier :

- (83) choucroutophage (< choucroute) ‘qui aime la choucroute’
 patatophage (< patate) ‘qui aime les pommes de terre’

Cependant, le verbe *manger* ne s’emploie pas uniquement au sens propre de ‘se nourrir’. Déjà, en grec, le verbe φαγειν signifiait ‘manger, dévorer, dépenser’. Aujourd’hui, pour le français, le Grand Robert ne donne pas moins de neuf définitions :

Manger (GR) :

1. (a) Avaler pour se nourrir (un aliment solide ou consistant) après avoir mâché.
 (b) Porter à la bouche, mâchonner sans avaler.
2. Dévorer (un être vivant, une proie).
3. Absorber, prendre des aliments (en général, et notamment au cours d’un repas).
4. Prendre un repas, ses repas.
5. Ronger (insectes, rongeurs).
6. Faire disparaître en cachant, en recouvrant.
7. *Manger ses mots, la moitié des mots*, les prononcer indistinctement.
8. Absorber, consommer, user.
9. *Manger qqn, son argent*, le ruiner, lui prendre toute sa fortune.

Ainsi, selon le sens métaphorique que prend l’élément *-phage*, d’autres schémas ont vu le jour. Par exemple, si *manger* est pris dans le sens 5, cela donne lieu au schéma (84), désignant, là encore, une caractéristique d’un animal, ici un insecte :

- (84) $[X_i \text{ phage}]_{ANj} \leftrightarrow [\text{qualifie un insecte dont la particularité est de ronger SEM}_i]_j$:
- lithophage (< lith- ‘pierre’) : *mollusques lithophages*, ‘qui creusent des cavités dans les rochers pour s’y loger’ (GR)
 - xylophage (< xyl- ‘bois’) : *insecte xylophage*, ‘qui ronge, perce le bois’ (GR)

Si manger est pris au sens 8 du GR, les constructions en *-phage* répondront au schéma (85) formant des adjectifs qualifiant des noms en général [-humains]. Le C1 n'est pas ici un aliment dont on se nourrit mais souvent une substance, concrète ou abstraite et l'emploi de l'EN *-phage* semble indiquer qu'une quantité importante de cette substance est consommée ou absorbée :

- (85) $[X_i \text{ phage}]_{Aj} \leftrightarrow [\text{qui consomme beaucoup de SEM}_i]_j$:
- *pétrophage* (< pétrole)
 - *chronophage* (< chron- 'temps')
 - *lipophage* (< lip- 'graisse') 'qui brûle, absorbe, la graisse'

Dans ce schéma, un mot comme *lipophage* est un peu en périphérie puisque ce n'est pas vraiment la notion de 'consommer' mais plutôt celle de 'absorber' qui est véhiculée.

Par extension, quand les adjectifs répondant au schéma (85) qualifient des noms [+humains], le trait sémantique [pathologique] se rajoute, puisque la notion de quantité importante semble demeurer. Ce schéma construit également les noms désignant les personnes entretenant la passion de l'entité représentée par le C1, qui est souvent une activité :

- (86) $[X_i \text{ phage}]_{ANj} \leftrightarrow [\text{qui est passionné par SEM}_i, \text{ consommant, utilisant SEM}_i \text{ de façon pathologique}]_j$:
- *téléphage* (< télévision)
 - *bibliophage* (< bibli- 'livre')
 - *trenetophage* (< (Charles) Trénet)

Quand le C1 désigne un aliment, il peut être difficile de choisir entre le schéma (80) et le schéma (86) ; en effet, l'adjectif peut qualifier une personne qui non seulement mange cet aliment mais de plus le fait par passion, voire de façon pathologique (87) :

- (87) Nous sommes des adeptes des crêpes, enfin ma femme est Crêpophile donc du coup **Crêpophage** (http://www.tripadvisor.fr/ShowUserReviews-g187147-d718650-r219668950-La_Creperie_de_Josselin-Paris_Ile_de_France.html)

Plus rare, les constructions en *-phage* peuvent avoir un sens antinomique de tous les autres, puisqu'on a d'un côté des sens comme 'extrêmement passionné par X' et de l'autre 'qui est hostile à X' :

- (88) $[X_i \text{ phage}]_{ANj} \leftrightarrow [\text{qui est contre SEM}_i]_j$:
- *turcophage* (< turc)

➤ *basiléophage* (< basilé- ‘roi’), ‘républicain’

La Figure 3.30 illustre les différents schémas dégagés pour les mots construits avec l’EN *-phage*.

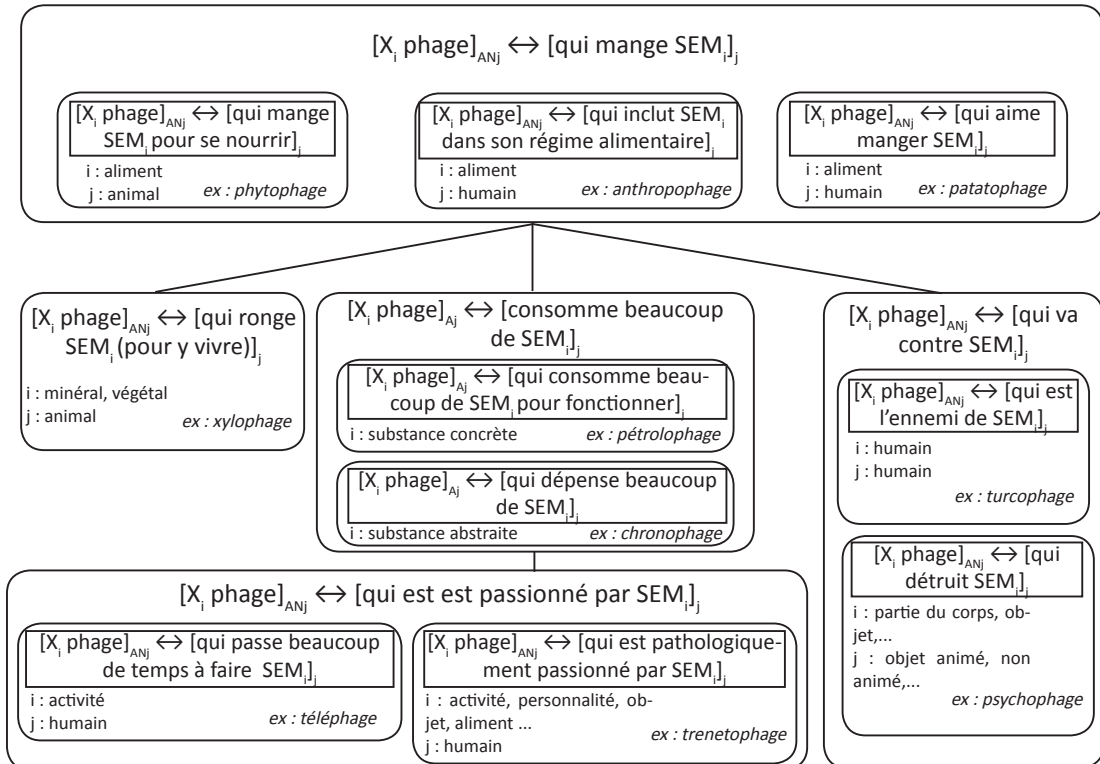


FIGURE 3.30 – Différents schémas impliquant l’élément néoclassique *-phage*

On pourra également noter quelques formations ne répondant pas aux schémas ci-dessus. Une des formations les plus anciennes de NeoClassy est le mot *sarcophage* (< sarc- ‘chair’) qui signifie littéralement ‘qui détruit les chairs’. Un autre exemple isolé est le mot *métrophage* (< métro) (89) dans lequel, à l’inverse des autres constructions, le premier constituant est le sujet de l’élément verbal *-phage* quand c’est l’objet dans tous les autres mots de la base.

- (89) « Parti cherché ma bagnole en métro à l’opposé de chez moi (logique, je suis un type logique je vous dis) l’après-midi du samedi je me suis retrouvé face à un flot de voyageurs, que dis- je, une marée humaine

de **métrophages**. » (<http://ludverglas.blogspot.fr/2007/11/croire-que-certains-collgues-lisent-ce.html>)

En (90), le lexème *onychophage* ne correspond pas aux schémas cités plus haut ; si l'adjectif qualifie un nom [+humain], il n'est pas construit comme *anthropophage* (80) ou *téléphage* (86), puisqu'il ne qualifie pas quelqu'un qui mange ses ongles ou qui est pathologiquement passionné par les ongles, même s'il semble contenir les sèmes 'manger' et 'pathologie'. Même s'il se rapproche du schéma (84), il ne désigne pas un animal, et l'action de ronger n'a pas pour but de creuser un habitat :

- (90) onychophage (< onych- 'ongle') 'qui se ronge compulsivement les ongles' (définition du GR)

Dans l'exemple (91), le mot *mélophage*, présent dans le TLFi et le GR, ne contient pas le sème 'manger', même métaphoriquement, puisqu'il désigne un parasite vivant sur les moutons.

- (91) mélophage (< μέλον 'mouton') : 'parasite vivant sur la toison des moutons'

De plus, comme c'est le cas pour tous les ENs du corpus, des mots peuvent appartenir à des schémas différents. Par exemple, *bibliophage* peut désigner soit un animal qui mange les pages d'un livre (schéma (80)), soit une personne qui lit énormément (schéma (86)). Cependant, il semble plus difficile de trouver *livrophage*, le synonyme formé avec un C1 natif, avec un autre sens que celui métaphorique, peut-être parce que le sens en (92-a-i) est davantage lexicalisé :

- (92) a. (i) L'humidité et la chaleur favorisent la formation de champignons, alimentation de nombreux insectes **bibliophages** et attirent des bactéries qui acidifient les papiers. (<http://www.archivosgenbriand.com/preservation-maladies.html>)
- (ii) J'adore surtout lire, je suis **bibliophage**. Je m'intéresse aussi à l'altermondialisme. (<http://voyageforum.com/membres/guachito/>)
- b. Il y a un peu plus d'un an, cette « **livrophage** » - c'est son expression - a ouvert la librairie jeunesse Mots et merveilles, rue des Épéers à Saint-Omer. Pour réaliser un rêve. (http://www.lavoixdunord.fr/Locales/Saint_Omer/actualite/Secteur_Saint_Omer/2011/07/12/article_sandrine-lenglet-livrophag)

e-amoureuse-d.shtml)

3.2.1.6 Manger, dévorer, consommer (bis)

Certains mots construits avec un C1 d'origine grecque et l'EN *-phage* ont leur exact synonyme avec un C1 d'origine latine et l'EN *-vore* :

- (93) phytophage / herbivore
 xylophage / lignivore
 chronophage / temporivore

De plus, des mots construits avec les ENs *-phage* et *-vore* sont fréquemment coordonnés :

- (94) Les cichlidés peuvent être herbivores, carnivores, limivores, macrophages, conchyliophages ou planctophages. (<http://aquariozone.free.fr/alimentation.php>)

Les schémas impliquant l'élément *-phage* ont ainsi pour beaucoup leur correspondant avec l'élément *-vore*. Le sens premier est donc celui de manger le référent du C1 :

- (95) $[X_i \text{ vore}]_{AN_j} \leftrightarrow [\text{qui mange SEM}_i]_j$:
 > *carnivore* (< carn- 'viande')
 > *truffivore* (< truffe)
 > *cheddarovore* (< cheddar)

Dans ce cas, les mots construits désignent soit un critère classificatoire quand ils caractérisent des animaux (96-a), soit un régime alimentaire quand le référent ou le nom qualifié est [+humain] (96-b) :

- (96) a. herbivore (< herbe), carnivore (< carn- 'viande'), piscivore (< pisc- 'poisson')
 b. panivore (< pain), lactivore (< lait), pâtivore (< pâtes)

Le C1 désigne donc l'aliment qui est consommé, à l'exception de l'exemple (97), où il ne désigne pas l'aliment en lui-même mais plutôt le lieu où l'on s'est procuré cet aliment :

- (97) locavore (< local) : qui consomme des aliments locaux

Comme c'était le cas pour les mots en *-phage* (84), *-vore* peut signifier 'ronger' :

- (98) $[X_i \text{ vore}]_{AN_j} \leftrightarrow$ [qualifie un insecte dont la particularité est de ronger SEM_i]_j :
- *lignivore* (< lign- ‘bois’)
 - *palmivore* (< palmier) : « PAPILLON **PALMIVORE** :Il est magnifique mais sa chenille est très dévastatrice pour les palmiers. » (http://www.homejardin.com/papillon_palmivore_nuisibles/detection_soins_traitement_des_nuisibles_au_jardin.html)

De même, des schémas *Xvore* peuvent porter le sens ‘manger’ de manière métaphorique quand C1 désigne soit une substance, soit une activité, comme c’était le cas pour les schémas *Xphage* (85) et (86) :

- (99) $[X_i \text{ phage}]_{Aj} \leftrightarrow$ [qui consomme beaucoup de SEM_i]_j :
- *énergivore* (< énergie)
 - *temporivore* (< tempor- ‘temps’)
 - *budgetivore* (< budget) ‘qui vit aux dépens de l’Etat’ (définition du GR)
- (100) $[X_i \text{ vore}]_{Aj} \leftrightarrow$ [qui est passionné par SEM_i , consommant, utilisant SEM_i de façon pathologique]_j :
- *télévore* (< télévision)
 - *infovore* (< informations)
 - *mangavore* (< manga)

Encore une fois, un même mot construit peut contenir les sens de ‘manger’ et ‘aimer de façon pathologique’ :

- (101) Je ne sais pas vous mais moi la tarte au Chocolat est une de mes tartes préférées étant une vraie **Chocolatovore**!! (<http://www.cookinglili.com/2012/04/tartelette-chocolat-sans-complexe-0-d-e-mat-grasses/>)

Enfin, à de très rares occasions, l’élément *-vore* peut être impliqué dans des schémas antinomiques des précédents, puisque c’est le sens ‘hostilité’ qui est ici mis en avant :

- (102) $[X_i \text{ vore}]_{AN_j} \leftrightarrow$ [qui est contre SEM_i]_j :
- *conductivore* (< conducteur) (voir exemple (103-a))
 - *stagivore* (< stagiaire) (voir exemple (103-b))
- (103) a. qui n’a pas aperçu, au bord d’une autoroute droite et vide, trois gendarmes mollement installés au soleil, Radar en batterie

au bord de la chaussée, attendant l'ignoble possesseur d'une "grosse" voiture s'autorisant - toujours sur l'autoroute vide - à rouler à l'allure que permettent -sans danger pour personne- son expérience et les caractéristiques de son véhicule? Hélas, hélas, l'imposture a la vie dure, et toutes les sources de revenus sont bienvenues pour un état "**conductivore**"! (<http://www.aquadesign.be/actu/article-493.php>)

- b. UNE PRESSE **STAGIVORE** Ce jour-là, donc, à la une de «Libé», Génération précaire est placardée, son combat médiatisé et sa cause, soutenue. La veille, le gouvernement a rendu public une Charte des stages étudiants en entreprise. Gérard Dupuy, éditorialiste du quotidien, regrette que cette charte ne soit qu'un «catalogue de bonnes volontés» et que les stages se situent toujours «en dehors du droit social». (http://www.bastamag.net/spip.php?page=imprimer&id_article=176)

Dans la même lignée que le schéma (102), l'exemple (104) implique une notion de destruction du référent du C1 :

- (104) chaussettivore (< chaussette)
 Quelques hypothèses pour expliquer ces mystérieuses disparitions : ma machine à laver est en fait une dangereuse **chaussettivore** / la poudre de lessive contient des agents chimiques destructeurs de chaussettes/ certaines chaussettes se réunissent et décident de fuir par le siphon pour aller vivre leur vie. (<http://leselucubrationsdenatacha.wordpress.com/2008/12/07/le-mystere-des-chaussettes/>)

Un exemple isolé (105) montre une construction dans laquelle l'élément *-vore* ne semble apporter aucune contribution sémantique :

- (105) ferrovivore (< ferroviaire)
 « Je vous souhaite aussi à tous les membres du forum une heureuse année **ferrovivore** 2008. » (<http://traincontroller.forum-actif.net/t176-et-deux-de-plus>)

Je reviendrai sur des constructions de ce type dans la partie 4.3.1. Enfin, un mot apparemment unique peut avoir été construit selon deux schémas, comme je l'avais montré déjà en (92) pour les constructions en *-phage* :

- (106) florivore (< fleur) :
 a. Je suis une **florivore**, lol, je fais sécher et réduire en poudre

les fleurs, comme la rose, la lavande, la violette...etc...etc Je voudrais savoir si les primevères, dont notre jardin regorge, sont comestibles? (http://www.forums.supertoinette.com/recettes_266785.manger_des_fleurs.html)

- b. Il existe 50 variétés dont les comportements sont extrêmement variés. Du plus petit primate du monde le Microcèbe nocturne et insectivore (30g), à l'Indri-indri diurne et **florivore** dont le poids dépasse les 7kg. (<http://www.makay-tours.com/decouvrir-circuit-sud-madagascar/voyage-pas-cher.html>)
- c. Une façade élégante et quelques potirons disposés sur les pavés du Vieux-Lille attirent mon attention. En grande **florivore**, je ne peux résister à une vitrine aussi alléchante. Me voilà alors au beau milieu de fleurs éclatantes aux senteurs délicates. (??)

Par ailleurs, une des différences notables entre les constructions *Xvore* et celles *Xphage* est qu'il existe pour les deuxièmes des constructions *Xphagie* qui leur sont reliées, mais qu'aucune construction [+ie] n'est répertoriée pour les lexèmes *Xvore*, du moins à ma connaissance.

3.2.1.7 Du meurtre à l'extermination

Trois schémas principaux impliquent l'EN *-cide* :

- (107)
- a. Type 1 :
 - (i) $[X_{N[+ANIMÉ]}]_i \text{ cide}]_{Nj} \leftrightarrow [\text{auteur du meurtre de SEM}_i]_j$
 - (ii) $[X_{N[+ANIMÉ]}]_i \text{ cide}]_{Nj} \leftrightarrow [\text{meurtre de SEM}_i]_j$
 - *homicide* (< hom- 'homme')
 - *belle-mèricide* (< belle-mère)
 - *hamstericide* (< hamster)
 - b. Type 2 : $[X_{N[+HUMAIN/+COLLECTIF]}]_i \text{ cide}]_{Nj} \leftrightarrow [\text{destruction méthodique de SEM}_i]_j$:
 - *génocide* (< gén- 'race')
 - *hutucide* (< hutu)
 - *pauvricide* (< pauvres)
 - c. Type 3 : $[X_{N[-HUMAIN]}]_i \text{ cide}]_{Nj} \leftrightarrow [\text{instrument servant à la destruction de SEM}_i]_j$:
 - *insecticide* (< insecte)
 - *fongicide* (< fong- 'champignon')
 - *raticide* (< rat)

Le premier regroupe en fait des lexèmes qui désignent à la fois un procès (le fait de tuer quelqu'un) et l'agent du procès (l'auteur du meurtre). Ces deux

schémas étaient distingués phonologiquement en latin, et le sont toujours dans d'autres langues romanes (108), mais le syncrétisme en français conduit à considérer soit qu'il s'agit d'un schéma polysémique, soit de deux schémas homonymiques.

		Procès	Agent
	Latin	homicidium	homicida
(108)	Français	homicide	homicide
	Espagnol	homicidio	homicida
	Italien	omicidio	omicida

Le constituant de gauche désigne souvent un membre de la famille (109-a), une personne désignée par sa fonction ou même son nom propre (109-b), un animal (109-c), mais peut occasionnellement être rempli par un non animé (109-d).

- (109)
- a. parricide, fratricide, grand-méricide, cousinicide, conjointicide, onclicide, tontonicide
 - b. sarkocide, tzaricide, présidenticide
 - c. hamstericide, brebicide
 - d. nounoursicide

Ce constituant de gauche désigne toujours une personne voire un objet pris dans son individualité.

Quand le constituant de gauche est un nom collectif, le mot en *-cide* sera construit selon le schéma (107-b). Le mot construit désignera alors un meurtre à grande échelle, souvent une destruction méthodique d'un groupe ethnique (110-a). Sont également construits selon ce schéma des mots désignant une tuerie à grande échelle d'animaux (110-b), voire d'un groupe d'humains désignés selon une caractéristique (âge, sexe, niveau de vie) (6-a).

- (110)
- a. hutucide, arabicide, juificide, rwandacide, sémicide
 - b. pigeonicide, chienicide, bovinicide
 - c. gérontocide, gynocide, pauvricide

Dans le schéma (107-c), le constituant de gauche désigne en général soit un animal (111-a) soit une plante (111-b), mais toujours des nuisibles à détruire. Le mot construit désigne l'instrument qui sert à détruire ces êtres indésirables. Sur le même schéma, l'instrument peut servir à faire une barrière sinon à détruire une substance (111-c), ou à contrer un micro-organisme (111-d) voire une maladie (111-e) :

- (111) a. raton laveuricide, lombricide, poissonicide
 b. herbicide, graminicide, sylvicide
 c. spermicide, spermatoicide, toxinicide
 d. bactéricide, parasiticide
 e. cancéricide

Par métaphore, d'autres schémas, qui prennent le sème 'tuer' au figuré, ont émergé :

- (112) a. $[X_{N[-ANIMÉ]}i \text{ cide}]_{Aj} \leftrightarrow [(qui \text{ provoque l'} \text{ anéantissement de } SEM_i)_j]$
 > *liberticide* (< liberté)
 > *amouricide* (< amour) : 'qui met des barrières à l'amour'
 > *tempslibricide* (< temps libre)
 > *écocide* (< écosystème)
- b. $[X_{N[-ANIMÉ]}i \text{ cide}]_{Aj} \leftrightarrow [qui \text{ est contre, ne respecte pas } SEM_i)_j]$
 > *intermittenticide* (< intermittent)
 > *légicide* (lég- 'loi) 'qui ne respecte pas la loi'
 > *argenticide* (< argent) : 'qui dépense trop d'argent (ex : des traders argentocides)

De plus, selon si le constituant de gauche est pris en tant qu'individu ou plutôt en tant que collectif, si le mot construit désigne un procès ou un instrument, des mots construits sur la même base pourront être construits sur des schémas différents. C'est le cas par exemple pour les mots construits à partir de noms d'animaux :

- (113) a. pigeonicide (< pigeon)
 (i) Type 1 : « J'ai commis un **pigeonicide** ! Je ne suis pas fière mais les fientes sur le balcon, shit ! It's no cool. » (<http://www.trappeursdepresent.fr/vos-present-s-du-17-fevrier/>)
 (ii) Type 2 : « Un **pigeonicide** est organisé par le conseil municipal sous le nom barbare de "dépigeonnisation" ça fait froid dans le dos, non ? » (<http://actualites.forum.orange.fr/messages/index/40198/rome-le-pape-francois-et-benoit-xvi-discrimination.html>)
- b. souricide (< souris)
 (i) Type 3 : « Pour reduire leur utilisation je suis obligé d'utiliser du **souricide** et raticide mais je crains pour la santé de mon chien, un chien de montagne des pyrenes. » (<http://forum.doctissimo.fr/sante/medeci>)

- ne-veterinaire/empoisonement-souricide-raticide-sujet_144864_1.htm)
- (ii) Type 1 : « Petit rectificatif Yoann, tu n’as pas causé un meurtre (c’est à dire donner la mort de façon intentionnelle et dans ton cas il n’y a pas de préméditation sinon cela aurait été un assassinat) mais un **souricide** involontaire ... J’espère que tu dormiras mieux » (<http://www.notre-planete.info/forums/discussion.php?id=72249>)

L’élément *-cide* est l’objet de l’article Lasserre et Montermini (2012) ; il fera également l’objet d’une analyse plus détaillée en 4.1.1.

3.2.1.8 Habiter, cultiver, honorer

L’élément *-cole* peut recouvrir trois sens principaux : celui d’‘honorer’ (114-a), celui de ‘cultiver’ (114-b) et celui d’‘habiter’ (114-c) :

- (114) a. $[X_i \text{ cole}]_{Aj} \leftrightarrow [\text{qui honore SEM}_i]_j$:
- *christicole* (< Christ)
 - *papicole* (< pape)
 - *serpenticole* (< serpent)
- b. $[X_i \text{ cole}]_{Aj} \leftrightarrow [\text{relatif à la culture de SEM}_i]_j$:
- *tabacole* (< tabac)
 - *fraisicole* (< fraise)
 - *cuniculicole* (< cunicul- ‘lapin’)
- c. $[X_i \text{ cole}]_{Aj} \leftrightarrow [\text{qui habite dans / sur SEM}_i]_j$:
- *arboricole* (< arbor- ‘arbre’)
 - *sabulicole* (< sabul- ‘sable’)
 - *torrenticole* (< torrent)

Le premier sens concerne uniquement 10 lexèmes de NeoClassy, tous dictionnarisés ; ce schéma ne semble plus disponible aujourd’hui pour la création de nouveaux mots. Le second sens recouvre en fait deux sens : celui de cultiver quand X désigne une plante et celui d’élever quand X désigne un animal. Les mots créés selon ce schéma entrent dans un réseau avec des mots formés selon les schémas *Xculture* et *Xculteur*. Des lexèmes *Xcole* et *Xculteur* peuvent parfois être synonymes (115) :

- (115) agent sylvicole, agent sylviculteur

Dans ce réseau, un lexème *Xcole* pourra entretenir diverses relations : il

pourra par exemple signifier ‘relatif à la culture de X’ (116-a), ‘relatif au lieu où est cultivé X’ (116-b) ou encore ‘relatif à X’ ((116-c) et (116-d)) :

- (116) rizicole :
- a. ‘relatif à la culture du riz ou riziculture’ :
 Cette conférence permettra aux riziculteurs et acteurs de la filière **rizicole** biologique française d’échanger avec leurs homologues européens (<http://www.sud-et-bio.com/evenements/premiere-conference-internationale-les-systemes-de-production-rizicole-biologique>)
 - b. ‘relatif à un lieu dans lequel le riz est cultivé’ :
 Dans ces plaines **rizicoles**, la forêt a souvent disparu (www.assistancescolaire.com/eleve/6e/geographie/reviser-une-notion/un-delta-rizicole-de-l-asie-du-sud-est-le-bengladesh-6gpr02)
 - c. ‘relatif au riz (en tant que plante cultivée)’ :
 De bonnes récoltes en vue. Les inondations causées par le passage des cyclones Funda et Chedza n’ont pas eu de gros impacts négatifs sur la production **rizicole** dans l’Alaotra-Mangoro, par conséquent le rendement n’a globalement pas baissé. (<http://www.lexpressmada.com/blog/information-live/alaotra-mangoro-une-bonne-recolte-rizicole-en-perspective-31886/>)
 - d. ‘relatif au riz (en tant qu’aliment)’ :
 Cette hausse des prix a remis en exergue la vulnérabilité de ces pays vis-à-vis du marché mondial, dans un contexte où la croissance de la consommation **rizicole** en milieu urbain en particulier a reposé pour une large part sur une disponibilité de surplus exportables à l’échelle mondiale. (<http://agritrop.cirad.fr/572271/>)

Dans le schéma (114-b), X ne désigne pas forcément une plante ou un animal. En effet, un mot comme *polycole* (117) est formé sur ce schéma : ici, X ne désigne pas une plante qui est cultivée mais une pluralité de plantes.

- (117) une exploitation polycole

Brass-, dans le lexème *brassicole*, désigne une action, brasser, et le lexème signifie en général ‘qui est relatif à la production de la bière’ dans son ensemble et non pas uniquement à la culture du houblon ou du malt. Là encore, plusieurs sens sont possibles, comme le montre le TLFi qui dénombre trois

sens pour l'entrée *brassicole* :

1. Qui se consacre à la fabrication de la bière : « En France, on cultive le houblon dans 6 ou 7 départements spécialisés, fertiles et brassicoles, situés dans la partie nord et est du pays » (*L'Industr. fr. de la brasserie*, 1955, p. 13).
2. Relatif à la fabrication et à l'industrie de la bière : « Un dictionnaire de termes brassicoles » (*L'Industr. fr. de la brasserie*, 1955, p. 29)
3. Apte à la fabrication de la bière. « Orges brassicoles » (*Lar. encyclop.*)

Le dernier schéma (114-c) signifie non seulement 'qui habite dans le milieu désigné par X' mais également 'qui croît, se développe dans le milieu désigné par X' (118) :

(118) calcicole (< calc- 'calcaire') 'qui croît dans les milieu calcaires'

- a. « Les plantes **calcicoles** se plaisent mieux dans un sol dont le taux d'acidité est supérieur à 7. Elles aiment un sol bien drainant et riche en éléments nutritifs. » (www.dcm-info.be/fr/hobby/conseil/520/entre-acidophile-et-calcicole/)
- b. « Les pelouses **calcicoles** mésophiles sont des formations herbacées basses et denses qui se développent sur des sols carbonatés ou basiques formés sur les calcaires secondaires du Jurassique (...) » (<http://www.poitou-charentes-nature.asso.fr/Pelouses-calcicoles-mesophiles.html>)

Si le schéma (114-c) construit des syntagmes du type *Y X-cole* où 'Y vit dans X', l'exemple (119) affiche la construction inverse 'X vit dans Y' :

(119) rivière cyprinicole (< cyprinidés 'famille de poissons') : rivière peuplée de cyprinidés

Par ailleurs, de nombreux mots *Xcole* sont construits parallèlement sur les schémas (114-b) et (114-c). Cependant, seul le lexème *Xcole* construit sur le schéma (114-b) est en relation avec un lexème *Xculture* :

(120) a. floricole 'qui vit sur les fleurs'
 ex. « Comme toutes les abeilles, ce bourdon est un insecte **floricole** appartenant à l'ordre des hyménoptères apoïdes, experts en pollinisation. » (<http://mediamateur.fr/role-dec-lin-des-insectes-pollinisateurs/>)

- b. floricole ‘relatif à la culture des fleurs’
ex. « Avant 1994, Tây Tuu était une zone de riziculture en banlieue de Hanoi. Les villageois se sont mis à la culture des fleurs quand les villages **floricoles** de Ngoc Hà, Quang Ba et Xuân Dinh ont progressivement perdu leurs terres en raison de l’urbanisation galopante. » (<http://vietnam.vnnet.vn/french/tay-tuu-village-floricole-high-tech-a-hanoi/106931.html>)
- c. floriculture ‘culture des fleurs’
ex. « En termes de ventes, l’hégémonie néerlandaise sur le marché mondial de la **floriculture** tend par ailleurs à se réduire. La part de marché des Pays-Bas à l’export sur les fleurs coupées est passée de 58% en 2003 à 52% en 2013 » (<http://indices.usinenouvelle.com/produits-agricoles/le-transport-maritime-nerf-de-la-guerre-du-commerce-floricole.5943>)

Enfin, comme nous le verrons dans la partie 4.3.1, l’EN *-cole* sert parfois à construire des adjectifs relationnels qui ne sont plus mis en relation avec un lexème *X culture* mais directement avec *X*. Si pour certains, on peut encore déceler le sens de ‘culture’ ou ‘habitat’ (121-a), pour d’autres l’élément *-cole* n’apporte plus aucune charge sémantique (121-b).

- (121) a. steppe savanicole ‘steppe constituée de savane’
b. gastronomie ostréicole ‘gastronomie relative aux huîtres’

3.2.2 Le lexique s’organise

Comme je l’ai abordé dans la partie 2.1.3, les schémas constructionnels peuvent afficher un sens prototypique duquel sont dérivés d’autres sens, par métaphore ou métonymie. Ces sens dérivés sont pris en charge par des sous-schémas. Tous ces schémas et sous-schémas s’organisent en réseau lexical. Les schémas affichent donc une polysémie régulière (Booij, 2010, p.78) : il y a une extension sémantique d’un sens prototypique grâce aux forces conductrices que sont la métaphore et la métonymie. Ainsi, tous les schémas impliquant un même élément néoclassique tissent entre eux des liens plus ou moins lâches : certains auront des sens très proches, d’autres des sens plus éloignés. La façon dont ces liens se tissent et donc la façon dont le réseau prend forme peut varier. D’un côté, le sens prototypique du schéma principal peut être pris en charge par un mot leader, ou un groupe de mots leader (Rainer, 2003), et chaque mot construit avec le même procédé morphologique, le même EN, construira son sens à une certaine distance du sens de ce mot leader (3.2.2.1).

De l'autre, aucun mot leader n'émerge, et, même si le sens de chaque mot se construit toujours en fonction du sens des mots existants, un éventail de sens possibles est offert pour la création de nouveaux mots (3.2.2.2).

3.2.2.1 Organisation par mot leader

Le premier type d'organisation du lexique, l'organisation par mot leader, a été mise en lumière par exemple par Rainer (2003) et Roché (2011a). Un ou plusieurs mots servent de modèle à la création de nouveaux mots, ce sont les prototypes d'un schéma de construction et, dans une vision topologique, chaque mot qui se crée se place à une certaine distance de ses mots leader : la distance est très faible lorsque l'apport sémantique du processus morphologique est identique, elle est plus grande lorsque le mot construit un sens plus différent. Le mot leader constitue dans cette vision un pôle d'attraction : plus la distance par rapport à ce pôle est grande, plus le sens diffère. Lorsque plusieurs mots sont à une grande distance d'un mot leader mais sont très proches les uns par rapport aux autres, un nouveau schéma émerge. Le prototype de ce nouveau schéma peut alors devenir leader, et servir à son tour de pôle d'attraction. Cette émergence de nouveaux schémas est nommée fragmentation sémantique par Rainer (2003, p.198) : « *semantic fragmentation* here means that a once semantically homogeneous word-formation process is split up in the course of time into a series of different processes ». Chaque nouveau mot trouvera dans cet espace sa place, à plus ou moins grande distance de chaque mot leader, de chaque pôle d'attraction.

Le mot leader est donc un mot particulièrement saillant, il est disponible dans la compétence lexicale des locuteurs. Son caractère saillant est marqué par sa fréquence élevée. Souvent, parmi les mots formés selon un processus morphologique, c'est un des premiers à avoir été créé ou importé. On peut aussi penser que le mot leader, puisque c'est un mot fréquent, connu des locuteurs, est un mot qui est dictionnarisé.

Prenons l'exemple des mots créés avec l'EN *-cratie*. Le Tableau 3.7 résume les dix mots de NeoClassy les plus fréquents, selon les chiffres donnés par Google (recherche effectuée le 10 septembre 2014, sur les occurrences au singulier et au pluriel de chaque lexème de la base), et les dix mots qui ont été les premiers importés ou créés avec cet élément. Tous ces chiffres sont bien sûr donnés pour indication : concernant les occurrences sur Google, comme je l'ai expliqué dans la partie 2.2, ces chiffres sont à prendre avec précaution. Ils ne sont utilisables qu'en comparaison les uns par rapport aux autres ; il ne semble pas abusif, pour donner un exemple extrême, de dire que *démocratie* (plus de 49 millions d'occurrences le 10/09/14) est plus fréquent que *blondocratie* (5 résultats le même jour). De même, les dates d'apparition

et de création sont celles qui sont données dans les dictionnaires utilisés, à savoir le TLFi et le GR, lorsqu'elle est indiquée.

		Occurrences Google	Date d'apparition ou de création	Dictionnarisé ?
Dix mots les plus fréquents de NeoClassy	démocratie	49 240 000	1370	TLFi, GR
	bureaucratie	2 945 000	1790	TLFi, GR
	aristocratie	1 569 800	1361	TLFi, GR
	autocratie	411 800	1794	TLFi, GR
	méritocratie	348 530	1971	GR
	théocratie	321 300	1679	TLFi, GR
	technocratie	205 600	1934	TLFi, GR
	médiocratie	164 080	1844	TLFi, GR
	voyoucratie	118 902	1865	TLFi, GR
	pornocratie	116 328	1865	TLFi, GR
Dix mots les plus anciens de NeoClassy	aristocratie	1 569 800	1361	TLFi, GR
	démocratie	49 240 000	1370	TLFi, GR
	gynécocratie	6 037	1565	TLFi, GR
	ochlocratie	8 343	1568	TLFi, GR
	ploutocratie	45 300	1582	TLFi, GR
	théocratie	321 300	1679	TLFi, GR
	thalassocratie	18 440	1730	TLFi, GR
	physiocratie	100 540	1758	TLFi, GR
	bureaucratie	2 945 000	1790	TLFi, GR
	canaillocratie	7 512	1793	TLFi, GR

Tableau 3.7 – Lexèmes *Xcratie* les plus fréquents et les plus anciens

Le Tableau dégage assez nettement plusieurs mots leader, du point de vue de la fréquence et de la date d'apparition. Les deux premiers mots introduits en français, *démocratie* et *aristocratie* font aussi partie des plus fréquents. Ce sont des mots que les locuteurs utilisent très souvent, que les enfants apprennent à l'école : ils ont donc un fort pouvoir d'attraction. Un troisième mot leader a émergé un peu plus tard, *bureaucratie*. Chacun de ces mots leader est un prototype des principaux schémas qui ont été cités dans la partie précédente. En termes de fréquence, leur caractère saillant ne fait pas de doute, puisqu'au moment de la recherche, Google renvoyait plus d'un million de pages les contenant alors que tous les autres affichaient moins de 500 000 occurrences. La différence importante entre ces différents résultats est donc susceptible d'être significative.

Cependant, les choses ne sont pas toujours aussi nettes. Rainer (2003) parle d'ailleurs de mot leader ou de groupe leader : ce rôle de prototype peut être rempli par un groupe de mots construits selon le même schéma sans que l'on puisse dire qu'un en particulier est plus saillant que qu'un autre. C'est ce qu'on observe pour les mots construits avec l'EN *-phobie*, dont les dix les plus fréquents et les plus anciens sont reportés dans le Tableau 3.8. Les mots créés ou importés les premiers ne font pas partie, pour la plupart, de ceux qui sont listés parmi les plus fréquents. Parmi les plus fréquents, certains ne sont pas dictionnarisés. Cependant, de par leur fréquence importante (plus d'un million de résultats sur Google), on peut considérer que les mots *homophobie*, *xénophobie* et *islamophobie*, même s'ils n'ont été créés que dans des temps relativement récents, ont acquis une importance significative et font partie d'un groupe de mots leader pour le schéma (66-a). De leur côté, *agoraphobie* et *claustrophobie*, qui font partie des mots les plus fréquents et les plus anciens (*claustrophobie* s'il ne fait pas partie des dix plus anciens, a été toutefois créé à une date proche de ceux-ci) formés avec l'élément *-phobie*, font partie du groupe de mots leader attaché au schéma (64-a). Enfin, *photophobie* ou *aquaphobie* doivent faire partie du groupe leader attaché au schéma (65-a). Ensuite, par un processus d'analogie, chaque mot créé selon un schéma se positionne par rapport à un des mots leader. Ce processus d'analogie n'est autre qu'une contrainte lexicale de série, selon, par exemple, Roché (2009) :

Contrainte 6 *Contrainte de série*

Le sens d'un mot se crée par analogie avec le sens d'un ou des mots leader formés selon un schéma donné.

Il arrive de fait que les mots leader soient explicitement cités comme référence pour la construction d'un mot. En (122), les néologismes *pipolicratie* (< people) et *éconocratie* (< économie) affichent leur lien avec respectivement les mots leader *aristocratie* et *démocratie*, le premier étant construit comme un hyponyme de *aristocratie*, le deuxième comme un antonyme de *démocratie*.

- (122) a. Une nouvelle forme d'aristocratie populaire : la **pipolicratie**. (http://www.noosphere.org/caza/lettresouvertes/lettresouvertes_141_150.html)
- b. Ah ! La sacrosainte économie, celle qui nous a privé de démocratie pour une **éconocratie** aveugle et qui affame! (<http://construire-lavenir.blogspot.com/p/la-partie-economie.html>)

Un mot leader peut également se détacher et devenir mot leader d'un sous-schéma. Par exemple, sur le modèle d'*homophobie*, d'autres hostilités envers

		Occurrences Google	Date d'apparition ou de création	Dictionnarisé ?
Dix mots les plus fréquents de NeoClassy	homophobie	5 516 500	1980	GR
	xénophobie	3 231 100	1904	TLFi, GR
	islamophobie	1 344 890	1994	GR
	dysmorphophobie	602 170		TLFi
	agoraphobie	241 830	1865	TLFi, GR
	christianophobie	197 397		Non
	transphobie	162 400		Non
	claustrophobie	154 640	1880	TLFi, GR
	photophobie	124 520	1812	TLFi, GR
	aquaphobie	78 695		Non
Dix mots les plus anciens de NeoClassy	hydrophobie	46 790	1314	TLFi, GR
	aérophobie	5 120	1751	TLFi, GR
	nécrophobie	4 070	1793	GR
	pantophobie	3 708	1808	TLFi, GR
	photophobie	124 520	1812	TLFi, GR
	anglophobie	17 300	1829	TLFi, GR
	gallophobie	3 790	1845	TLFi, GR
	agoraphobie	241 830	1865	TLFi, GR
	prétrophobie	1 070	1865	GR
	sitiophobie	6 730	1871	GR

Tableau 3.8 – Lexèmes *Xphobie* les plus fréquents et les plus anciens

des groupes d'orientation sexuelle se sont construites (123) :

- (123) Évoquant des actes inacceptables, ces témoignages expriment les souffrances vécues par ceux et celles qui sont confronté.e.s à la lesbophobie, la gayphobie, la biphobie et la transphobie. (<http://www.foleffet.com/SOS-HOMOPHOBIE>)

Souvent, les néologismes sont aussi des créations humoristiques. Au moment des débats sur le mariage homosexuel en France, pour contrer les discours homophobes de certaines personnes dont la politicienne Christine Boutin, elle-même mariée à son cousin, François Morel, humoriste, a dit dans une de ses chroniques :

- (124) Il dénonce les propos « **cousinophobes** », évoquant les pancartes rappelant que le mariage, c'est « un papa et une maman, un papa et un papa, une maman et une maman mais sûrement pas une maman et un tonton ou un papa et une tante » (<http://yagg.com/2013/06/14/lhommage-de-francois-morel-a-la-republicaine-christine-boutin-et-au-mariage-entre-cousins/>)

Le contexte nous montre ici que *cousinophobe* est construit en réaction à *homophobe*, et donc que le mot leader a joué ici son rôle de pôle attraction dans l'élaboration du sens du néologisme. En effet, comme je le verrai plus en détail dans la partie 2.1.3, la construction sémantique des mots ainsi que leur interprétation ne se fait pas en isolation mais toujours dans un contexte.

La Figure 3.31 donne un aperçu de ce à quoi peut ressembler une organisation lexicale composée de pôles d'attraction gouvernés par des mots leader.

Cependant, parfois aucun mot leader ou groupe leader ne se dégage, les schémas se multiplient, et le lexique peut s'organiser autrement.

3.2.2.2 Organisation distribuée

Comme nous l'avons vu dans la section précédente, certains éléments néo-classiques entrent dans des schémas qui peuvent être variés, affichant un éventail de sens possibles mais pas forcément liés. Nous n'avons plus de mots leader qui constituent des pôles d'attraction mais des schémas divers, plus ou moins proches sémantiquement et même non forcément liés. Même si des pôles peuvent se détacher, l'organisation lexicale prend la forme de radiants, au sens de Rainer (2003, p.206) : « a radial structure generally, but not necessarily, consists of a central member – the prototype – surrounded by a bunch of less central members. A less central or peripheral member is not

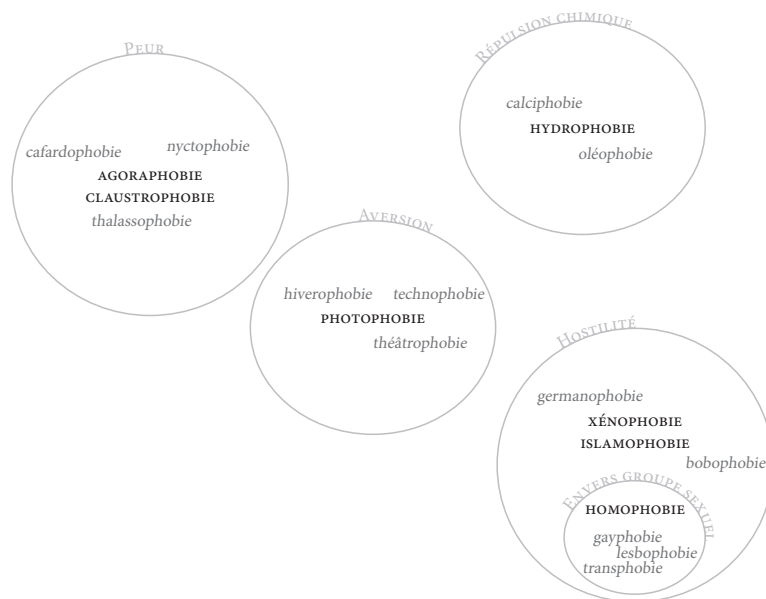


FIGURE 3.31 – Aperçu d'une organisation lexicale avec des mots leader

predictable from the central member, but is motivated by it, i.e., there is a some independently existing link that makes sense of their relationship ». Par exemple, si nous extrayons, comme nous venons de le faire, les dix mots les plus fréquents et les dix mots les plus anciens parmi les constructions avec *-phage* (Tableau 3.9) et les constructions avec *-logie* (Tableau 3.10), il devient vraiment ardu de définir un ou plusieurs mots leader.

Comme le montrait la Figure 3.30, l'élément *-phage* est impliqué dans de nombreux schémas et sous-schémas. Chacun de ces schémas peut être relié sémantiquement aux autres, grâce à la mise au jour de liens métaphoriques ou de glissements de sens. Le Tableau 3.9 ne distingue aucun mot leader, si ce n'est *anthropophage* et *sarcophage*. Ce dernier, comme je l'ai signalé dans la partie précédente, est une formation ancienne, qui existait en grec ancien et qui ne rentre dans aucun des schémas qui ont pu être dégagés. Le mot *anthropophage* quant à lui peut être vu comme le prototype dont parle Rainer dans la citation ci-dessus : il est non seulement fréquent, a été introduit en français il y a maintenant longtemps et sa construction sert de point de départ à toutes les autres : chaque mot construit avec *-phage* peut être attaché par un lien motivé à celui-ci, comme je l'avais montré dans la section précédente, puisque le sens de 'manger' peut en fait recouvrir nombre de variantes : 'consommer', 'dévorer', 'ronger', etc. Le sens de ces

		Occurrences Google	Date d'apparition ou de création	Dictionnarisé ?
Dix mots les plus fréquents de NeoClassy	sarcophage	1 427 000	1496	TLFi, GR
	chronophage	1 070 900	1961	TLFi, GR
	macrophage	604 000	1891	TLFi, GR
	xylophage	267 500	1808	TLFi, GR
	anthropophage	176 400	1375	TLFi, GR
	hippophage	169 600	1827	GR
	librophage	114 380		Non
	nécrophage	91 300	1802	TLFi, GR
	ampélophage	89 270	1874	TLFi, GR
	bactériophage	85 800	1918	TLFi, GR
Dix mots les plus anciens de NeoClassy	anthropophage	176 400	1375	TLFi, GR
	sarcophage	1 427 000	1496	TLFi, GR
	scatophage	26 690	1552	TLFi, GR
	ichtyophage	9 960	1552	TLFi, GR
	acridophage	1 670	1566	TLFi
	polyphage	33 700	1578	TLFi, GR
	lithophage	13 120	1694	TLFi, GR
	ophiophage	8 050	1721	TLFi, GR
	galactophage	3 120	1721	TLFi, GR
	rhizophage	3 030	1732	TLFi, GR

Tableau 3.9 – Lexèmes *Xphage* les plus fréquents et les plus anciens

lexèmes construits, comme le souligne Rainer, n'est pas prédictible à partir du sens de *anthropophage* mais chaque lexème est lié, puisque tous les schémas construisant ces lexèmes le sont.

L'éventail de sens des constructions en *-logie* est encore plus vaste (Lasserre et Montermini, 2014a). La même étude de NeoClassy concernant les mots les plus fréquents et les mots les plus anciens fait ressortir des mots qui ont été créés, ou importés, il y a très longtemps (dès le XII^{ème} siècle), et des mots très fréquents (beaucoup sont bien au-delà du million de résultats sur Google) (Tableau 3.10). Il est donc encore plus difficile de faire ressortir un mot leader ou même un prototype. Les pôles d'attraction qui se dégagent ne sont pas définis autour de mots leader, mais selon les différents points d'une structure radiante que nous pourrions imaginer comme représentée dans la Figure 3.32.

		Occurrences Google	Date d'apparition ou de création	Dictionnarisé ?
Dix mots les plus fréquents de NeoClassy	technologie	224 300 000	1656	TLFi, GR
	biologie	90 321 000	1802	TLFi, GR
	psychologie	48 360 000	1690	TLFi, GR
	généalogie	40 370 000	XII ^{ème}	TLFi, GR
	anthropologie	35 837 600	1507	TLFi, GR
	écologie	28 735 200	1874	TLFi, GR
	astrologie	19 579 000	1260	TLFi, GR
	sociologie	16 361 000	1830	TLFi, GR
	chronologie	15 720 000	1579	TLFi, GR
	archéologie	15 574 000	1599	TLFi, GR
Dix mots les plus anciens de NeoClassy	généalogie	40 370 000	XII ^{ème}	TLFi, GR
	étymologie	6 469 000	XIII ^{ème}	TLFi, GR
	astrologie	19 579 000	1260	TLFi, GR
	théologie	5 558 000	1270	TLFi, GR
	tropologie	14 340	1295	TLFi, GR
	mythologie	7 040 000	XIV ^{ème}	TLFi, GR
	analogie	13 120	1694	TLFi, GR
	anthropologie	35 837 600	1507	TLFi, GR
	amphibologie	56 420	1533	TLFi, GR
	météorologie	3 438 090	1547	TLFi, GR

Tableau 3.10 – Lexèmes *Xlogie* les plus fréquents et les plus anciens

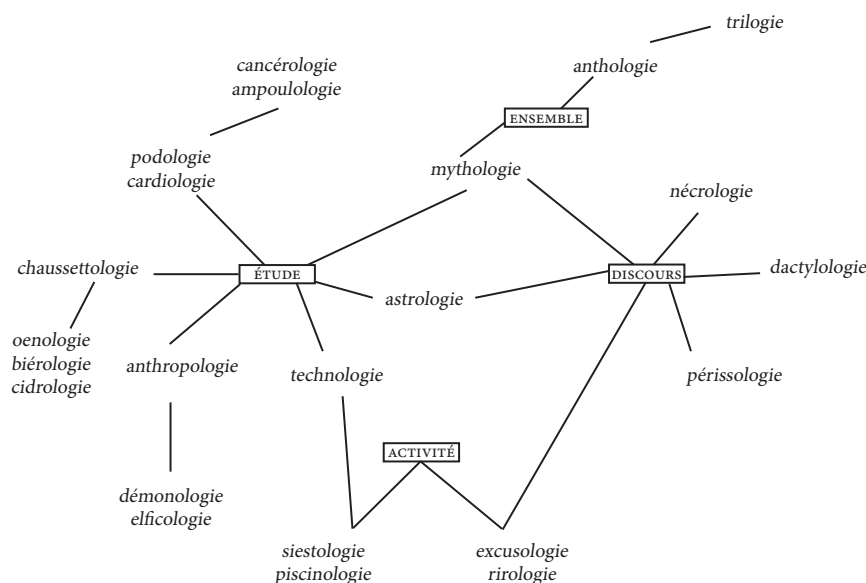


FIGURE 3.32 – Aperçu d’une organisation lexicale du type distribué

Si l’on compare les Figures 3.31 et 3.32, la différence dans l’organisation du lexique est évidente. D’un côté, des mots leader qui créent des pôles par rapport auxquels les mots nouvellement construits se positionnent plus ou moins proches en fonction du schéma selon lequel ils sont formés, de l’autre une structure radiante sur laquelle se positionnent également, toujours dans une vision topologique, les mots qui seront créés. Parmi les constructions néoclassiques étudiées dans cette thèse, celles formées avec l’EN *-cide* s’organisent en fonction de mots leader, celles avec *-phone* présentent un agencement plus distribué.

3.2.2.3 Des liens qui se tissent

Si les schémas impliquant un même EN entretiennent des relations, des liens se tissent également entre certains schémas mettant en jeu des ENs différents, parmi lesquels certains qui n’ont pas été choisis dans cette étude. Par exemple, certaines constructions sont synonymes et affichent même souvent des synonymes avec une base identique et un EN qui varie (125) :

$$(125) \quad [X_i \text{ vore}]_{A_j} \leftrightarrow [X_i \text{ cide}]_{A_j} \leftrightarrow [X_i \text{ fuge}]_{A_j} \leftrightarrow [\text{qui détruit SEM}_i]_j$$

➤ *tartrivore* = *tartrifuge*
 ➤ *vermicide* = *vermifuge*

➤ *ténicide* = *ténifuge*

De même, un réseau de schémas s'est formé entre avec les ENs *-phage*, *-phile*, *-mane* voire *-lâtre* et *-phobe* :

- (126) a. $[X_i \text{ phage}]_{ANj} \leftrightarrow [X_i \text{ phile}]_{ANj} \leftrightarrow [X_i \text{ mane}]_{ANj} \leftrightarrow [X_i \text{ lâtre}]_{ANj}$
 $\leftrightarrow [\text{amateur de SEM}_i]_j$
 b. $[X_i \text{ phobe}]_{ANj} \leftrightarrow [\text{qui n'aime pas SEM}_i]_j$

Là encore, les différents ENs finaux peuvent s'attacher à la même base pour former synonymes et antonymes.

- (127) a. (i) Mais ce sont les dates auxquelles vivaient ces deux **bibliophiles** qui donnent un intérêt particulier à cet ouvrage. Jean Jouffroy était de cette première génération de lecteurs qui pouvaient avoir accès aux volumes provenant des presses de l'imprimerie naissante. (<https://medievaux.revues.org/7054>)
 (ii) Je propose encore un petit jeu aux lecteurs du blog - sur le thème de la BD et du voyage. Il s'agit de se mesurer aux **bibliophages** de Babelio.com, qui sont assez furieux dans le genre, dévorant dix ou quinze livres par semaine en moyenne. (<http://fanzine.hautetfort.com/tag/bibliophage>)
 (iii) Le salon du livre 2015 de Paris était l'occasion pour les **bibliomanes** compulsifs comme pour les lecteurs occasionnels de rencontrer certains auteurs bien connus de vaporistes français. (<http://www.french-steampunk.fr/professeur-etienne.html>)
 (iv) Ces **bibliolâtres** essaient de ne vendre que des livres qu'ils ont lus, l'un ou l'autre, parfois l'un et l'autre. (<http://www.grainsdesel.com/page.php?art=3162&rub=#.Vfle1s79it8>)
 b. À chaque destination est associé un ouvrage y faisant référence, avec en prime le résumé de l'histoire, le niveau de concentration nécessaire pour en venir à bout, mais il est aussi précisé si une adaptation cinématographique a déjà eu lieu ou non. Les **bibliophobes** pourront ainsi se rattraper en se programmant une soirée ciné. (<http://www.grazia.fr/culture/news/l-infographie-du-jour-quel-livre-emporter-selon-sa-destination-de-vacances-774954>)

Souvent, on retrouve ces synonymes dans un même contexte (128).

- (128) C'est ce sentiment erroné qui fait que les **bibliomanes**, **bibliophiles** et autres **bibliolâtres** ne sont après tout que des imposteurs qui ne vibreront jamais pour la tournure poétique d'une phrase, pour la justesse d'une métaphore ou pour l'émotion suscitée par un récit. (http://lebibliomane.blogspot.fr/2011_01_01_archive.html)

Mais sont-ce vraiment des synonymes ? Souvent, il s'agit plutôt de gradients, comme le montre l'intitulé de ce test trouvé sur Internet (129), dans lequel *-phage* et *-phile* ne sont pas mis sur le même plan, puisque les *téléphages* sont plus consommateurs de télé que les *téléphiles*.

- (129) TEST : ÊTES-VOUS TÉLÉPHILE, TÉLÉPHOBE OU TÉLÉPHAGE ? (http://www.schools.ac.cy/eyliko/mesi/themata/gallika/didaktiko_yliko/la_telelevision_le_cinema/b_gym/4_competences/televisioncinema_3_c.o.c.e._b.docx)

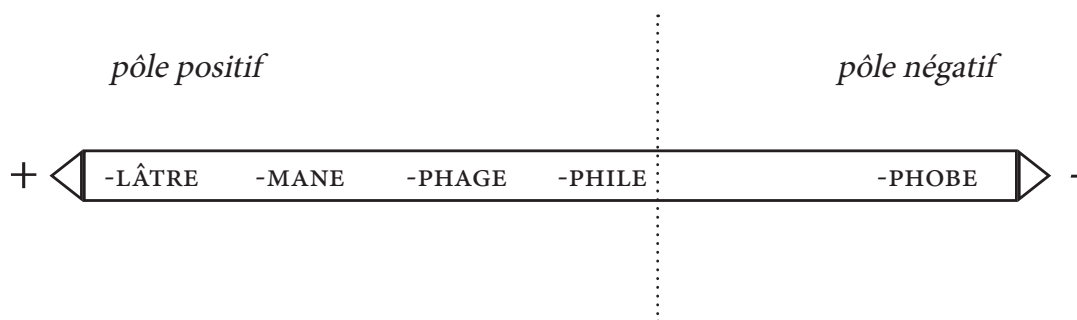
De même, dans l'exemple (130), *discophage* semble désigner un *discophile* à un degré supérieur.

- (130) Cette magnifique pianiste a déjà cinq disques à son actif (lire sa discographie), sans avoir commis le moindre faux-pas : de Rachmaninov à Granados, de Gershwin à Chopin, Lydia sait brillamment donner vie aux heureux compositeurs qu'elle honore de son talent, pour la plus grande joie des **discophiles** (voire des "**discophages**", comme c'est notre cas). (http://www.arre-se.com/arrese-articl-piano_fr.html)

Ce gradient pourrait être représenté comme dans la Figure 3.33.

Cependant, nous avons vu en (88) que l'élément *-phage* pouvait entrer en jeu dans des constructions antinomiques. On voit ainsi qu'en (131-a), le lexème *francophage* signifie 'hostile aux français', placé dans ce contexte en opposition par rapport à *anglophile*. En (131-b), il est même employé comme amplificateur par rapport à *francophobe*. Il serait donc, dans ce cas, à placer également à droite de *-phobe* dans la Figure 3.33.

- (131) a. Quant à votre commentaire, je demande à tous ceux qui vous liront de faire attention au fait que vous soyez très anglophile - donc forcément un peu **Francophage** (<http://www.agoravox.fr/tribune-libre/article/est-ce-bien-raisonnable-de-s>

FIGURE 3.33 – Gradabilité des ENs *-lâtre*, *-mane*, *-phage*, *-phile* et *-phobe*

- peak-137183)
- b. Dans sa critique du nationalisme teutomane et francophobe, voire **francophage**, héritier de 1813-1815 et prolongé jusqu'aux années 1830 et 1840, ... (Aux origines du couple franco-allemand, de Arnold Ruge, 2004, p.110)

Plus anecdotiquement, certaines constructions peuvent être mises en relation uniquement pour certaines occurrences. Dans l'exemple (132), *intermittenticide* signifie 'hostile aux intermittents', tout comme l'aurait été *intermittentophobe*.

- (132) Selon les intermittents, le ministre sait très bien que le protocole est **intermittenticide**, et il serait gêné d'avoir à l'expliquer en face de ses propres victimes. (http://archives.rezo.net/archives/agi_b.mbox/FSC4H4XHLFEBNTT7GBH2KQNG7FJRYBRA/)

Les constructions *Xphile* et *Xcrate* peuvent par ailleurs chacune désigner des partisans d'un type de régime, ou d'une personne, comme en (133)

- (133) a. Presse trop **sarkophile** : encore une photo présidentielle truquée dans le Figaro (<http://sarkozix.canalblog.com/archives/2009/05/05/13617690.html>)
- b. La justice **sarkocrate** et bourgeoise a une fois de plus édicté sa vérité. (<http://pasdebavure.eklablog.net/proces-du-8-mars-2012-la-justice-justifie-la-police-a44813584>)

Les éléments *-phone* et *-graphie* construisent tous deux des noms d'instruments :

- (134) a. téléphone, grammophone
 b. télégraphe, oscillographe, ellipsographe

A l'inverse, certaines constructions peuvent être antonymes quand elles impliquent une base en particulier (135) :

- (135) a. nidicole : se dit des oiseaux dont les jeunes ne quittent le nid que lorsqu'ils savent voler. (définition du GR)
 b. nidifuge : se dit des oiseaux dont les jeunes sont élevés peu de temps dans le nid. (définition du GR)

Ainsi, des liens se tissent non seulement entre des constructions impliquant un seul et même EN, mais également entre les constructions impliquant des ENs différents. Souvent, l'enjeu pour le locuteur est de former un mot qu'il veut savant, donc avec un élément néoclassique. Et le choix de cet élément importe moins que le fait que ce soit un élément néoclassique. La Figure 3.34 tente une illustration d'une partie réseau créé par les constructions qui ont été abordées dans cette partie. Nous verrons dans le chapitre suivant que des schémas impliquant des ENs peuvent également être mis en relation avec des schémas impliquant des suffixes, par exemple *Xlogue* forme des noms de spécialistes comme le suffixe *-iste* et *Xphobe* peut être synonyme de *anti X*.

Bilan. Ce chapitre a mis en lumière un jeu de contraintes pesant sur la formation des constructions néoclassiques de NeoClassy. Il y a tout d'abord des contraintes pesant sur la forme des lexèmes. Au niveau du constituant initial, deux contraintes opposées sont mises en balance : une contrainte de transparence qui facilite l'interprétation du mot construit et une contrainte de crypticité qui fait rentrer le lexème dans un jargon (pseudo-)scientifique. Du côté de la taille, un même type de conflit entre deux contraintes a été mis au jour : les lexèmes de NeoClassy respectent soit le gabarit des lexèmes construits du français, soit un gabarit plus long, signe, là encore, de l'appartenance à un registre savant. Ces contraintes opposées qui se jouent montrent ainsi un tiraillement entre des constructions qui sont à l'origine destinées à la langue savante mais qui pénètrent de plus en plus la langue courante. Cependant, une contrainte forte pèse sur la formation de ces constructions, peu importe le registre de langue pour lesquelles elles sont destinées, c'est la voyelle de liaison, qui n'appartient ni au premier constituant, ni au second, mais qui est indissociable du schéma de construction : la voyelle <o> est ainsi un marqueur fort de ces constructions, tendant à suppléer la voyelle <i> qui est encore majoritaire pour les ENs d'origine latine.

Par ailleurs, des schémas de construction ont permis des généralisations faites

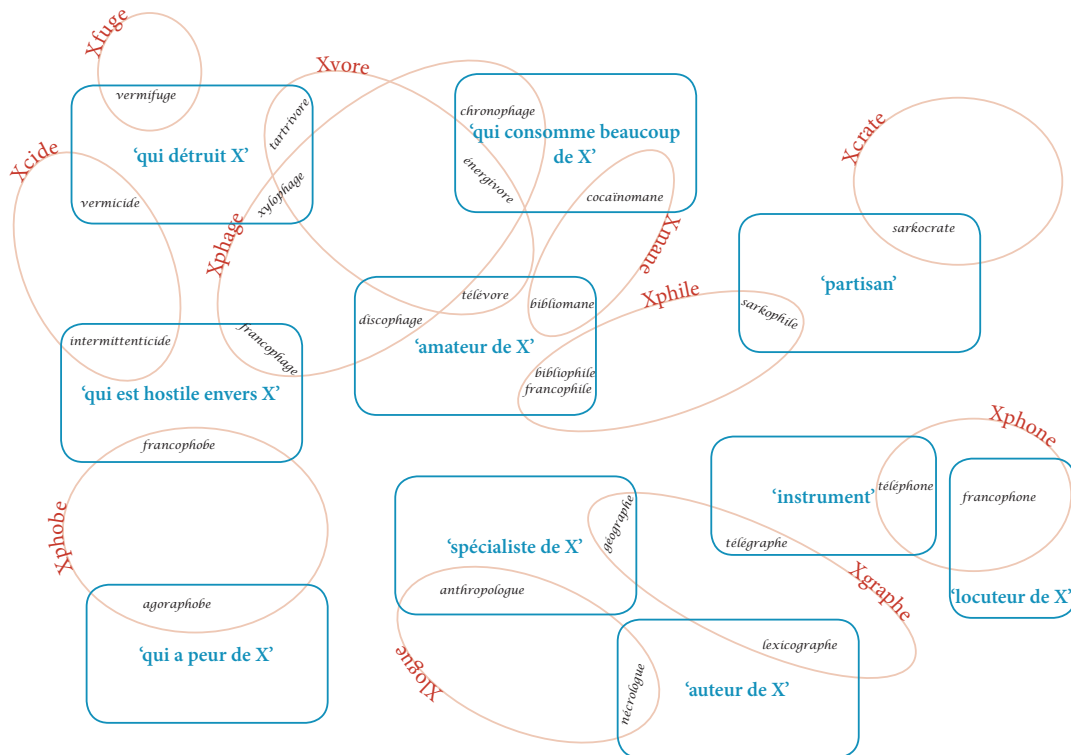


FIGURE 3.34 – Réseau formé par les constructions impliquant certains éléments néoclassiques

sur le sens des lexèmes de NeoClassy. Chaque EN final récurrent est en fait impliqué dans une série de schémas qui sont plus ou moins fortement connectés d'un point de vue sémantique, relevant de ce que Booij (2010) nomme polysémie régulière. Un sens prototypique, souvent celui originel, est étendu, par métaphore ou métonymie, à d'autres sens. Les schémas de construction, qui généralisent les lexèmes existants tout en permettant la création de nouveaux, sont instanciés par des mots particulièrement saillants dont le pouvoir d'attraction résulte en des contraintes de séries : tout mot est construit par analogie avec les mots présents dans le lexique. Ces mots saillants, les mots leader, permettent une structuration du lexique et chaque mot formé construit son sens à plus ou moins grande distance de ces mots leader. Pour certains schémas, la structuration du lexique est plus distribuée, l'éventail de sens est plus large mais la création lexicale trouve toujours sa place à l'intérieur de cet éventail.

Le chapitre suivant sera l'occasion d'analyser plus en détail la construction du sens de mots relevant de ces schémas de construction. Nous pourrons ainsi voir le rôle que jouent les éléments néoclassiques, éléments allogènes, dans la construction du sens d'un mot français.

Les éléments néoclassiques dans la construction du sens

Sommaire

4.1	Construction du sens et interprétation	214
4.1.1	Proposition d'analyse distributionnelle	217
4.1.1.1	Fondements et méthode	217
4.1.1.2	Ce que le contexte des lexèmes <i>Xcide</i> nous apprend	219
4.1.1.3	Les <i>-phobie</i> en contexte	228
4.2	Les relations inter-schématiques singulières . . .	237
4.2.1	Dérivés possibles et impossibles	237
4.2.2	Des liens privilégiés dans le réseau lexical	242
4.3	La charge sémantique des éléments néoclassiques	246
4.3.1	Les constructions néoclassiques comme adjectifs relationnels	246
4.3.2	Les éléments néoclassiques, des suffixes comme les autres?	257
4.3.2.1	La construction de noms de spécialistes <i>Xlogue</i>	257
4.3.2.2	Les anti-étrangers sont-ils xénophobes?	263

Le chapitre précédent a montré que les ENs finaux étudiés dans cette thèse sont impliqués dans des schémas divers, parfois antinomiques. Ces schémas sont cependant connectés et structurent l'espace du lexique. Chaque mot créé trouve sa place dans le lexique en se positionnant à plus ou moins grande distance des mots qui s'y trouvent déjà et qui jouent le rôle de pôle d'attraction sémantique : un locuteur qui crée un mot construit son sens par rapport

au sens des mots qu'il connaît déjà. Mais comment parvient-il à construire son sens, à utiliser ce mot selon le schéma voulu? Comment l'interlocuteur qui reçoit ce mot parvient-il à l'interpréter selon le sens que le locuteur lui a donné? En 4.1 nous verrons que le sens d'un lexème ne se construit pas en isolation mais que le contexte joue un rôle primordial dans son identification. Pour cela, j'appliquerai la méthode de l'analyse distributionnelle, souvent utilisée en sémantique mais très rarement en morphologie. En 4.2, je m'appuierai sur les relations qui existent à l'intérieur du lexique entre des schémas impliquant des ENs finaux différents afin de montrer que les liens unissant certains schémas sont des aides dans l'identification du sens des lexèmes. Enfin, en 4.3 je m'interrogerai sur la charge sémantique des éléments néoclassiques. En effet, dans certaines constructions, les ENs, qui sont traditionnellement rapprochés des lexèmes en raison d'un sens qui serait d'ordre lexical, ne contiennent plus une telle charge sémantique : ils ne sont en cela pas différents de suffixes prototypiques.

4.1 Construction du sens et interprétation

La partie 3.2.1 du chapitre précédent proposait une description des différents schémas dans lesquels les ENs finaux étudiés sont impliqués, et nous avons vu que certains schémas sélectionnent des C1 appartenant à des catégories sémantiques bien définies. On pourrait donc en déduire qu'observer le type sémantique de la base nous indiquerait le schéma de construction auquel correspond le mot complexe auquel nous faisons face. Cependant, ce n'est pas toujours aisé.

Par exemple, dans les schémas qui construisent des lexèmes *Xcide* j'avais dégagé trois types principaux (1) :

- (1) a. Type 1 :
- (i) $[X_{N[+ANIMÉ]}i \text{ cide}]_{Nj} \leftrightarrow [\text{auteur du meurtre de SEM}_i]_j$
 - (ii) $[X_{N[+ANIMÉ]}i \text{ cide}]_{Nj} \leftrightarrow [\text{meurtre de SEM}_i]_j$
- b. Type 2 : $[X_{N[+HUMAIN/+COLLECTIF]}i \text{ cide}]_{Nj} \leftrightarrow [\text{destruction méthodique de SEM}_i]_j$
- c. Type 3 : $[X_{N[-HUMAIN]}i \text{ cide}]_{Nj} \leftrightarrow [\text{instrument servant à la destruction de SEM}_i]_j$

Les types sémantiques du premier constituant de chaque schéma sont assez clairement définis. On peut donc imaginer que lorsqu'on est confronté à un C1 [-animé], le mot ne peut pas avoir été construit selon le schéma de type 1 ; s'il s'agit d'un C1 [-humain], il ne peut pas avoir été construit selon le schéma

de type 2. Or les exemples (2) nous montrent que certains mots peuvent avoir été construits selon des schémas avec des C1 inattendus.

- (2) a. Y'a une dizaine de jours je me suis rendue coupable de **pullicide** involontaire envers mon super grand gilet Maje en laine. (<http://www.dailyelle.fr/etre-fonelle/12-avril-2010-qui-a-deja-foutu-en-lair-une-fringue-super-1015>)
- b. Et tandis qu'Agnès songeait à la manière la plus efficace de pénétrer un **mouchicide** d'envergure, la situation prit un tour nouveau. (<http://www.univers-rr.com/RPartage/index.php?page=rp&id=4498&start=4>)

En (2-a), le lexème *pullicide* (< *pull*) est construit selon le schéma de type 1, et en (2-b), le lexème *mouchicide* (< *mouche*), dont le C1 est pris ici sous son aspect collectif, est construit selon le schéma de type 2. De plus, comme nous l'avons vu par exemple avec les exemples (113) du chapitre précédent reportés ici, une même séquence C1 + EN peut avoir été construite selon des schémas différents.

- (113) a. pigeonicide (< pigeon)
- (i) Type 1 : « J'ai commis un **pigeonicide** ! Je ne suis pas fière mais les fientes sur le balcon, shit ! It's no cool. » (<http://www.trappeursdepresent.fr/vos-present-s-du-17-fevrier/>)
- (ii) Type 2 : « Un **pigeonicide** est organisé par le conseil municipal sous le nom barbare de "dépigeonnisation" ça fait froid dans le dos, non ? » (<http://actualites.forum.orange.fr/messages/index/40198/rome-le-pape-francois-et-benoit-xvi-discrimination.html>)
- b. souricide (< souris)
- (i) Type 3 : « Pour reduire leur utilisation je suis obligé d'utiliser du **souricide** et raticide mais je crains pour la santé de mon chien, un chien de montagne des pyrenes. » (http://forum.doctissimo.fr/sante/medecine-veterinaire/empoisement-souricide-raticide-sujet_144864_1.htm)
- (ii) Type 1 : « Petit rectificatif Yoann, tu n'as pas causé un meurtre (c'est à dire donner la mort de façon intentionnelle et dans ton cas il n'y a pas de préméditation sinon cela aurait été un assassinat) mais un **souricide** involontaire ... J'espère que tu dormiras mieux

» (<http://www.notre-planete.info/forums/discussion.php?id=72249>)

De même, concernant les lexèmes *Xphage*, un C1 [+humain] peut servir à construire les sens ‘qui mange X’ (3-a), ‘qui est passionné par X’ (3-b) ou ‘qui est hostile à X’ (3-c).

- (3) a. anthropophage (< anthrop- ‘homme’)
 b. trénetophage (< Trénet (Charles))
 c. turcophage (< Turcs)

Concernant les mots formés avec l’EN *-phobie*, le type sémantique du C1 peut nous aiguiller sur l’interprétation du mot construit ; par exemple, quand il s’agit d’un animal, il est fort probable que le mot désigne une peur, et quand c’est un humain, il doit a priori s’agir d’une hostilité. Or la *clownophobie* (< *clown*) désigne le plus souvent la peur des clowns, et non pas une hostilité envers les clowns. En outre, quand le C1 ne réfère ni à un animal, ni à un humain, comment peut-on interpréter le mot construit ? Si l’on prend ainsi le mot *autophobie* (4), il est impossible, pris en isolation, de déterminer s’il s’agit de la peur des voitures ou d’une haine envers les voitures et plus largement les automobilistes. Cependant, grâce au contexte d’énonciation, l’interlocuteur, sans aucun doute, parviendra à comprendre en (4-a) que le locuteur parle d’une peur et en (4-b) qu’il parle d’une haine.

- (4) autophobie (< automobile)
 a. Mais je n’ai jamais été à l’aise en auto, je suis comme ça. (...) Je crois sincèrement souffrir d’**autophobie** depuis sa naissance. (<http://www.mamanpourlavie.com/forum/sujet/nouvelles-maman-s-de-mars-2007/page/10/voir/videos>)
 b. Si vous vous sentez concernés (...), par l’**autophobie** ambiante savamment entretenue par certains dirigeants et lobbys et par le ras le bol lié à la pression incessante qui règne sur nos vies de conducteur,... (<http://blogautomobile.fr/40-millions-d-automobilistes-dit-non-baisse-limitations-de-vitesse-206186#axzz3n2ExZMmi>)

Le sens des mots ne se construit pas en isolation. Le contexte sert à l’actualisation d’un lexème et participe pleinement à la construction de son sens. Il s’agit maintenant de trouver une méthode qui sera capable d’étudier le sens des mots sans laisser le contexte de côté.

4.1.1 Proposition d'analyse distributionnelle

L'analyse distributionnelle nous offre cette possibilité, puisqu'elle fait l'hypothèse qu'on peut accéder au sens d'un mot en observant les contextes dans lesquels ce mot est employé : « you shall know a word by the company it keeps » (Firth, 1957, p.11).

4.1.1.1 Fondements et méthode

Cette approche prend sa source dans le courant distributionnaliste dans les années 1930 aux Etats-Unis. Bloomfield (1933), s'inscrivant dans la lignée de la psychologie béhavioriste, avance une approche anti-mentaliste qui propose d'étudier le langage grâce à ce qui est directement observable. La psychologie béhavioriste étudie le comportement humain comme un ensemble d'excitations et de réponses, les premières étant des stimuli, les secondes des actions. Décrire le comportement humain, c'est donc décrire l'interaction de l'homme avec son milieu. Et ce comportement est entièrement explicable et prévisible à partir des situations dans lesquelles ce comportement est observé. Bloomfield applique la même méthode d'analyse au langage : toute unité linguistique peut se définir en fonction des contextes dans lesquels elle apparaît. Il se propose ainsi d'établir un modèle taxinomique : au moyen d'une procédure de découverte, chaque unité est classée en fonction de ses contextes d'apparition. La procédure de découverte consiste en une analyse en constituants immédiats : il s'agit d'une décomposition successive des unités appartenant au niveau le plus haut (la phrase) pour faire, niveau après niveau, apparaître des unités de plus en plus petites jusqu'aux morphèmes. Harris (1954, 1968, entre autres), continuateur de Bloomfield, s'attache à opérer une analyse distributionnelle, c'est-à-dire la reconnaissance puis l'identification de chacune des unités dégagées :

The distribution of an element will be understood as the sum of all its environments. An environment of an element A is an existing array of its co-occurents, i.e. the other elements, each in a particular position, with which A occurs to yield an utterance. A's co-occurents in a particular position are called its selection for that particular position. (Harris, 1954, p.146)

La distribution d'un élément est donc révélée par une méthode inductive : le linguiste part des faits pour induire des généralisations. Des unités ayant la même distribution seront regroupées ensuite dans des classes distributionnelles. Ces classes peuvent être grammaticales (la classe des noms, des adjectifs par exemple) ou sémantiques. Bloomfield refusait de s'intéresser à

la sémantique puisque dans une approche anti-mentaliste, le sens n'est pas quelque chose que l'on peut connaître. En effet, le sens dépend de tous les paramètres de situation d'énonciation du message, c'est-à-dire de différents stimuli et actions qui font que, à moins d'être omniscient, le linguiste ne peut atteindre le sens par l'observation et doit se restreindre à la seule description. Harris ne recourt pas non plus au sens dans l'analyse distributionnelle qui est « complete, without intrusion of other features such as history or meaning » (Harris, 1970). Le sens comme facteur extralinguistique n'aide pas à l'analyse. Cependant, le sens comme fait linguistique peut être un sujet de l'analyse, et donc de l'analyse distributionnelle : « the linguistic meanings which the structure carries can only be due to the relations in which the elements of the structure take part » (Harris, 1968, p.2). Comme le résume Sahlgren (2008, p.36), pour Harris « linguistic meaning is inherently differential, and not referential (since that would require an extra-linguistic component) » :

if we consider words or morphemes A and B to be more different in meaning than A and C, then we will often find that the distributions of A and B are more different than the distributions of A and C. In other words, difference of meaning correlates with difference of distribution. (Harris, 1970, p.786)

La distribution d'un mot est donc la représentation de son sens. Le sens est extrait, dans l'analyse distributionnelle, syntagmatiquement, en dégageant les mots co-occurents, et paradigmatiquement, en rapprochant le mot, et donc son sens, de mots partageant les mêmes voisins, et donc leur sens (Sahlgren, 2008). Une analyse distributionnelle se base donc sur l'usage. Les corpus sont ici des outils essentiels : « as repositories of linguistic usages, they represent the primary source of information to identify the word distributional properties » (Lenci, 2008, p.9).

Lenci (2008) distingue deux types d'hypothèses distributionnelles : dans la version faible, la méthode reste seulement quantitative. En effet, elle consiste à observer des contextes distributionnels jugés comme pertinents afin d'identifier certains aspects sémantiques qui pourraient être partagés par les mots qui ont des distributions similaires. La version faible, donc, « only assumes the existence of a correlation between semantic content and linguistic distributions, and exploits such correlation to get at a better understanding of the semantic behavior of lexical items ». Dans sa version forte, l'hypothèse distributionnelle est une hypothèse cognitive, puisqu'elle attribue un rôle causal aux distributions d'un mot : « the distributional behavior of a word in contexts is not only taken as a way to get at its semantic behavior, but indeed as a way to explain its semantic content at the cognitive level ».

Si l'analyse distributionnelle a été beaucoup utilisée en sémantique et en traitement automatique des langues, elle a pour l'instant suscité peu d'intérêt en morphologie. Lazaridou *et al.* (2013) font partie des rares à avoir abordé la morphologie dérivationnelle sous l'angle de l'analyse distributionnelle. Ils évaluent la capacité de différentes méthodes forgées en sémantique à rendre compte de phénomènes morphologiques. Ils rejettent également le point de vue compositionnel pour l'étude des mots construits (voir section 2.1.2).

L'analyse distributionnelle m'a permis, dans différents travaux (Lasserre et Montermini, 2012, 2015; Lasserre, 2015) d'aborder la construction du sens des constructions néoclassiques. Cette analyse n'est pas opérée de manière automatique, comme le sont celles pratiquées en sémantique ou en linguistique computationnelle. Ici, les analyses sont plus fines et, actuellement, au vu des moyens dont on dispose, difficiles à exécuter de manière automatique, d'où le caractère quelque peu artisanal de la méthode que je vais maintenant présenter.

4.1.1.2 Ce que le contexte des lexèmes *Xcide* nous apprend

Les sections 2.3.2.7 et 3.2.1.7 ont décrit diachroniquement et synchroniquement les lexèmes *Xcide* rencontrés en français. Comme je l'esquissais à la fin de la section 3.2.2.2, les schémas constructionnels impliquant l'élément *-cide*, déjà dégagés et reportés en (1), semblent s'organiser en fonction de mots leader. Regardons plus en détails, grâce au Tableau 4.1¹, quels sont les mots les plus fréquents et les plus anciens.

- (5) a. Type 1 :
- (i) $[X_{N[+ANIMÉ]}i \text{ cide}]_{Nj} \leftrightarrow [\text{auteur du meurtre de SEM}_i]_j$
 - (ii) $[X_{N[+ANIMÉ]}i \text{ cide}]_{Nj} \leftrightarrow [\text{meurtre de SEM}_i]_j$
- b. Type 2 : $[X_{N[+HUMAIN/+COLLECTIF]}i \text{ cide}]_{Nj} \leftrightarrow [\text{destruction méthodique de SEM}_i]_j$
- c. Type 3 : $[X_{N[-HUMAIN]}i \text{ cide}]_{Nj} \leftrightarrow [\text{instrument servant à la destruction de SEM}_i]_j$

Un mot leader manifeste se dégage pour le schéma de type 2, *génocide*. Ce mot a d'abord été créé en anglais par le juriste polonais Raphael Lemkin dans son ouvrage *Axis Rule in Occupied Europe*, publié en 1944. Le Grand Robert date son apparition en français l'année suivante. Il est donc clairement identifié comme étant le premier mot à avoir été créé avec le sens de 'destruction méthodique', et sa très forte fréquence sur la Toile est un des signes de son fort pouvoir d'attraction. Concernant le type 1, plusieurs mots

1. Recherche Google le 22/10/12

		Occurrences Google	Date d'apparition ou de création	Dictionnarisé ?
Dix mots les plus fréquents de NeoClassy	suicide	27 650 000	1734	TLFi, GR
	génocide	8 705 000	1945	TLFi, GR
	pesticide	5 550 000	1960	TLFi, GR
	homicide	3 952 000	1150	TLFi, GR
	insecticide	3 380 000	1838	TLFi, GR
	herbicide	1 562 000	1930	TLFi, GR
	fongicide	1 449 000	1867	TLFi, GR
	biocide	1 133 000	1969	GR
	liberticide	828 000	1791	TLFi, GR
infanticide	822 000	1564	TLFi, GR	
Dix mots les plus anciens de NeoClassy	homicide	3 952 000	1150	TLFi, GR
	parricide	285 000	1190	TLFi, GR
	fratricide	790 000	1450	TLFi, GR
	tyrannicide	26 570	1487	TLFi, GR
	uxoricide	4 671	1531	GR
	infanticide	822 000	1564	TLFi, GR
	matricide	56 400	1580	TLFi, GR
	déicide	54 280	1585	TLFi, GR
	régicide	212 500	1594	TLFi, GR
prêtricide	157	1651	GR	

Tableau 4.1 – Identification des mots leader pour le schéma *Xcide*, par fréquence et par date d'apparition

ressortent dans le tableau, tant au point de vue de la fréquence que de la date d'apparition : si *homicide* est sans conteste un mot leader bien défini, gravitent autour de lui un groupe de mots qui semblent également faire partie du pôle d'attraction. Nous aurions donc là un groupe de mots leader, dont certains sont les prototypes du sous-schéma dans lequel C1 est spécialisé en 'membre de la famille' : *parricide, fratricide, matricide*,... Enfin, le type 3 dégage pour les lexèmes *Xcide*, qui construit des instruments, est apparu plus tard, à partir du XIX^{ème} siècle. Cependant, certains lexèmes construits selon ce schéma sont très fréquents (*pesticide, insecticide, herbicide*,...) et forment ainsi un groupe de mots leader.

Selon l'hypothèse distributionnelle, on peut donc avancer que la proximité de chaque nouvelle formation avec les mots leader dégagés peut se calculer grâce à la quantité de contextes partagés. Il s'agit ici de la version faible de l'hypothèse distributionnelle (Lenci, 2008), c'est-à-dire que la quantité des contextes syntaxiques partagés constitue un indice de la proximité sémantique des constructions partageant ces contextes.

J'ai donc identifié les contextes les plus récurrents avec lesquels les mots leader identifiés sont employés ainsi que les contextes partagés par le plus grand nombre de ces mots leader. Cependant, je n'ai pas sélectionné les contextes qui étaient partagés par les mots leader de types différents, par exemple **TENTATIVE DE_Xcide** qui est utilisé autant avec les mots *génocide* que *homicide* ou *parricide*. Pour cela, je me suis appuyée sur plusieurs ressources. D'un côté j'ai regardé les contextes des mots leader sur la base de données Frantext qui recense en majorité des textes littéraires. Par ailleurs, j'ai utilisé les ressources *Les voisins de Le Monde*² (vdM) et *Les voisins de Wikipédia*³ (vdW). Ces ressources, développées au sein du laboratoire CLLE-ERSS de l'Université de Toulouse, sont basées respectivement sur 10 ans du journal *Le Monde* (1991-2000) et sur le Wikipédia français. Elles permettent de rechercher non seulement les contextes syntaxiques privilégiés d'un lexème, ou co-occurents, mais aussi les lexèmes qui partagent les mêmes contextes, les voisins. Concernant les co-occurents, deux mesures permettent d'extraire les contextes syntaxiques privilégiés : la fréquence et l'information mutuelle. La seconde, qui indique un rapport d'exclusivité entre un mot et ses contextes, est plus significative (cf. Morlane-Hondère (2013, p.50-51)) : la fréquence ne permet pas l'émergence des contextes les plus pertinents, puisqu'un même mot peut apparaître de façon très fréquente dans un contexte donné, de la même façon que beaucoup d'autres mots dans ce même contexte : ce contexte ne sera donc pas propre au mot étudié. Au contraire, plus l'information mu-

2. <http://redac.univ-tlse2.fr/voisinsdelemonde/>

3. <http://redac.univ-tlse2.fr/voisinsdewikipedia/>

tuelle est élevée, plus les co-occurents entretiennent un rapport d'exclusivité. La formule permettant de mesurer l'information mutuelle est la suivante ($freqTot$ est le nombre total d'occurrences dans le corpus et $freq(n)$ la fréquence de n dans le corpus) :

$$IM(x,y) = \log \left(freqTot \cdot \frac{freq(xy)}{freq(x) \cdot freq(y)} \right)$$

Grâce à ces différentes ressources, j'ai donc extrait certains contextes parmi les plus pertinents pour les mots leader des différents types de schémas construisant des lexèmes *Xcide*. Cette analyse se faisant manuellement, j'ai dû faire une sélection pour ne conserver que quelques contextes apparaissant comme les plus pertinents ; ils sont répertoriés dans la deuxième colonne du Tableau 4.2⁴. La deuxième étape a donc été de tester la co-occurrence de ces contextes et de chacun des mots *Xcide* de NeoClassy. Pour cela, les recherches listées dans la colonne de droite du Tableau 4.2 ont été soumises au moteur de recherche Google. Pour chacun des contextes, un échantillon de formes fléchies a été sélectionné pour la recherche.

Le Tableau 4.3 regroupe les résultats de l'analyse. A chaque fois qu'une requête donnait au moins un résultat, le lexème a été considéré comme ayant été construit selon le type auquel le contexte était attribué⁵. Malheureusement, plus de 40% des requêtes ont été infructueuses. En effet, d'une part, beaucoup des nouvelles formations recueillies dans NeoClassy n'ont qu'un nombre réduit d'occurrences sur Google, ce qui amoindrit les chances de les retrouver avec les memes contextes que d'autres, et d'autre part, n'ont été pris en considération qu'un nombre limite de contextes, l'analyse ayant été faite manuellement ; élargir la liste présentée au Tableau 4.2, voire conduire une analyse sur la totalité des contextes observables pour chacun des lexèmes de notre corpus, produirait certainement des résultats plus fiables.

Les types qui se dégagent le plus nettement sont les types 1 et 3, ceux qui désignent des agents / instruments ; respectivement 53 et 50 lexèmes de NeoClassy ont pu être rattachés à ces seuls schémas, sans partager aucun contexte réservé à un autre schéma. Que pouvons-nous dire du point de vue des types sémantiques des C1 au sein des types de schémas que nous venons de dégager ? La Figure 4.1 illustre la répartition des C1 selon leur type sémantique, dans les schémas qui ont pu être dégagés grâce à l'analyse distributionnelle.

4. Dans ce tableau, O est mis pour Objet, et DE regroupe la préposition *de* des l'ensemble des formes sous lesquelles elle peut apparaître (*de*, *d'*, *du*, etc.).

5. Le Tableau 4.3 regroupe donc les types et non les tokens.

Type	Contextes les plus fréquents	Recherches Google
1	<p>Xcide (IN)VOLONTAIRE</p> <p>COMMETTRE0_Xcide</p> <p>GESTE_Xcide</p>	<p>« Xcide volontaire »</p> <p>« Xcides volontaires »</p> <p>« Xcide involontaire »</p> <p>« Xcides involontaires »</p> <p>« commettre * Xcide »</p> <p>« commettre * Xcides »</p> <p>« commettent * Xcide »</p> <p>« commettent * Xcides »</p> <p>« commet * Xcide »</p> <p>« commet * Xcides »</p> <p>« commis * Xcide »</p> <p>« commis * Xcides »</p> <p>« Xcide commis »</p> <p>« Xcides commis »</p> <p>« geste Xcide »</p> <p>« gestes Xcides »</p>
2	<p>PERPÉTRER0_Xcide</p> <p>XCIDAIRE</p>	<p>« perpétrer * Xcide »</p> <p>« perpétrer * Xcides »</p> <p>« Xcidaire »</p> <p>« Xcidaires »</p>
3	<p>PRODUIT_Xcide</p> <p>UTILISATIONDE_Xcide</p> <p>UTILISER0_Xcide</p>	<p>« produit Xcide »</p> <p>« produits Xcides »</p> <p>« utilisation * Xcide »</p> <p>« utiliser * Xcide »</p>

Tableau 4.2 – Contextes sélectionnés pour l’analyse distributionnelle des CNs *Xcide*

Types dégagés à partir des contextes recherchés			Nombre	% sur le total
1	2	3		
X	–	–	53	17,32%
–	–	X	50	16,3%
X	X	X	19	6,21%
X	X	–	18	5,88%
X	–	X	15	4,90%
–	X	–	14	4,58%
–	X	X	10	3,27%
–	–	–	127	41,50%

Tableau 4.3 – Résultats de l’analyse distributionnelle effectuée pour les lexèmes *Xcide*

La quasi totalité des bases $N_{[+humain]}$ rentrent dans des schémas du type 1 (6-a) et, parmi celles-ci, presque toutes celles désignant des liens de parenté (6-b).

- (6) a. Pourquoi ai-je une telle envie de commettre un **gérontocide**? (http://forum.doctissimo.fr/sante/arreter-fumer/cabane-p-alabres-jours-sujet_177114_263.htm)
 b. combien on risque pour un **belle-mèricide** volontaire? (http://forum.doctissimo.fr/viepratique/informatique-photo-video-mp3/prison-combien-temps-sujet_3362_1.htm)

Dans le type 1, on trouve également des C1 $N_{[+animé]}$ désignant des animaux (7-a) et des C1 $N_{[+abstrait]}$ (7-b). Ces derniers mots construits utilisent le plus souvent des procédés métaphoriques ou métonymiques.

- (7) a. C’est vrai qu’il y va fort avec ces pauvres chatons, et ma belle-soeur à fort à faire pour empêcher un **chaticide** involontaire! (<http://www.vivelesrondes.com/BlueKat/22306>)
 b. L’ouverture d’un casier peu judicieux par Maître culpabilité et ses associés les regrets, où le pathétique acte manqué restera consigné à jamais en lettres de honte : « A tué sa plus belle

histoire armé d'un oubli » Bien sûr je plaiderai l'**amouricide** involontaire et non prémédité, j'invoquerai au juge Cupidon mes nombreuses années de service, mais comme il est vrai que l'on pourrait presque parler de récidive... (<https://ecritslibres.wordpress.com/2008/03/13/semaine-2-loubli-ludo/>)

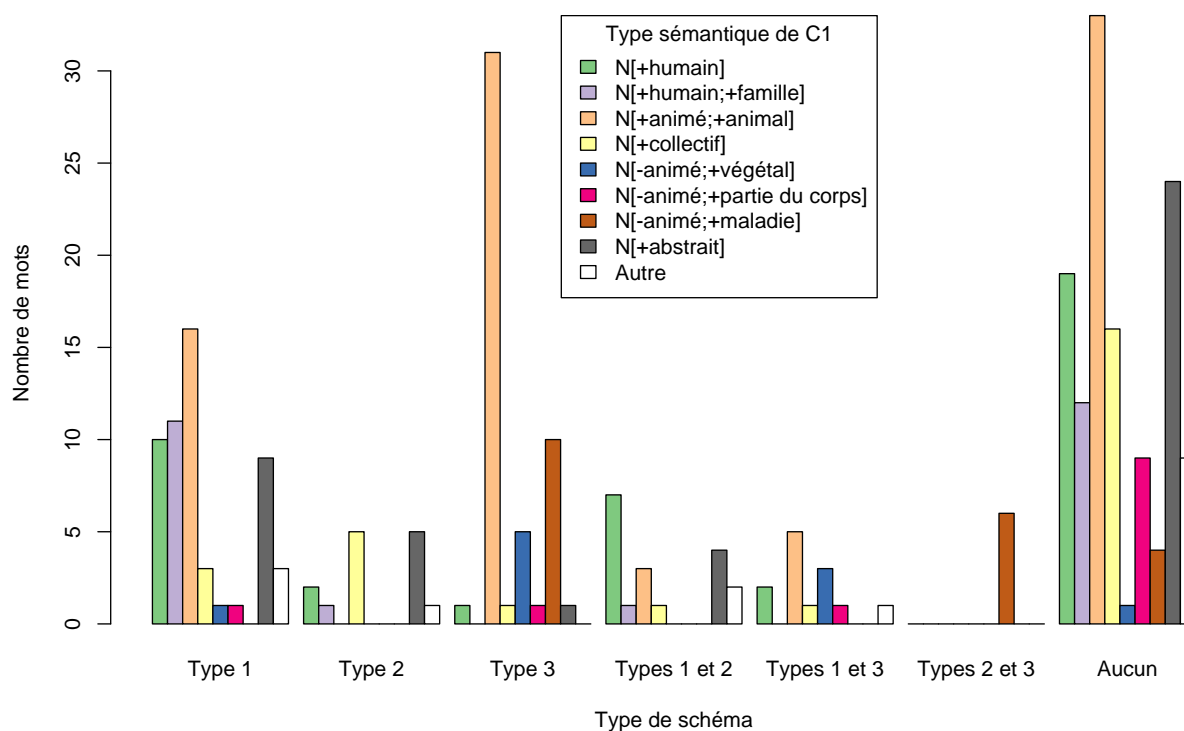


FIGURE 4.1 – Type sémantique du C1, selon le type de schéma impliquant l'EN *-cide*

Concernant le type 3, le schéma sélectionne en majorité des C1 désignant des animaux, des plantes ou des maladies (8), le plus souvent considérés comme nuisibles ou potentiellement nuisibles :

- (8) a. Les insecticides tuent les abeilles (et seraient présents dans le doux nectar ?) Et on courbe le dos devant le lobbying économique, on projette d'attendre 2005 pour y voir clair !! (le temps d'écouler les stocks du produit **abeillicide** Je veux bien être zen, mais

comme dit l'autre "faut pas pousser mémé dans les orties" tout de même!) (http://www.bertignac.com/forums/index.php?thisfor=BLEU&theme_id=&debut=17600)

- b. L'utilisation d'un **alguicide** est particulièrement recommandée pour les piscines en béton (en particulier les piscines carrelées) ou en cas de développement d'algues noires. (http://bleupiscines.nc/manuel/quesrep_tourne.htm)
- c. A défaut de la visite chez le dermatologue, demandez conseil à votre pharmacien. Il pourra vous proposer un produit **verrucide**, très efficace, à appliquer directement sur la verrue. (<http://www.e-sante.fr/verrue-plantaire-reagissez-tout-suite/2/actualite/1560>)

Contrairement aux deux types précédents, les analyses permettent plus difficilement de tracer les contours du type 2. Cela s'explique par le fait que le mot leader unique *génocide* présente moins de contextes qui lui sont réservés. Seuls trois lexèmes ont ici comme contexte exclusif, parmi ceux considérés, **PERPÉTRER0_X** : *linguocide*, *tutsicide* et *avunculicide* (< avuncul- 'oncle maternel'), le dernier ne désignant même pas, apparemment, une tuerie à grande échelle. Quant au contexte Xcidaire, s'il sert à désigner un agent, il est relativement répandu parmi les lexèmes de NeoClassy, certainement par l'influence de *suicidaire*. Cependant, il sert tout de même fréquemment, sinon exclusivement, à interpréter une occurrence comme ayant le sens de 'destruction méthodique' :

- (9) Et vous voulez nous deratiser ??? Bande de **raticidaires**! (<http://fr.answers.yahoo.com/question/index?qid=20090221164135AAbnBNM>)

Le type 2 est donc plus facilement identifiable par l'observation du type de C1, puisqu'il privilégie des noms collectifs. En fait, si le schéma désigne une 'tuerie à grande échelle', l'aspect collectif est contenu dans le C1, et non dans le schéma. De plus, l'analyse distributionnelle est plus difficile à mettre en place pour ce troisième type puisqu'il forme moins de lexèmes, par rapport aux deux autres.

Si l'on se place maintenant du côté des C1, certains d'entre eux se prêtent davantage à des interprétations diverses. C'est le cas pour les C1 N_[+animé] désignant des animaux. Sans surprise, l'analyse distributionnelle effectuée ici montre que 70% des C1 désignant un animal, parmi les 179 lexèmes qui ont pu être interprétés grâce à l'analyse distributionnelle, sont sélectionnés dans les contextes définis dans le type 3, et que 50% le sont de manière exclusive.

Cependant, il ressort que 45% de ces C1 sont utilisés dans des constructions attribuées au type 1, dont 25% de manière exclusive. Contrairement à ce qui se passe pour le type 2, l'animal désigné par le C1 est ici perçu individuellement, et l'acte de tuer comme unique. Les exemples (10) montrent cette différence de traitement : en (10-a), c'est l'identité de la souris en tant qu'animal de compagnie qui est saillante, alors qu'en (10-b) ce sont ses propriétés d'animal potentiellement nuisible qui le sont :

- (10) a. Quand à ma souris blanche je l'ai bien assassinée mais c'était un **souricide** involontaire, monsieur le Juge, je vous le jure. (<http://www.femiboard.com/threads/nouveau-jeu-2v%C3%A9rit%C3%A9s-et-un-mensonge.40924/page-33>)
- b. je n'aime pas trop utiliser de **souricide**, les bêtes souffrent terriblement... mieux vaut utiliser une tapette. (<http://www.graines-et-plantes.com/index.php?forum=jardin-jardinage&question=campagnols-sous-plancher>)

Ainsi, lorsque des mots sont formés avec des C1 appartenant à un même type sémantique, par exemple un $N_{[+\text{humain}]}$ désignant un peuple, le contexte syntaxique peut nous orienter sur l'interprétation même du C1, qui peut être collectif ou individuel. En (11-a), où le nom *Xcide* est le complément du verbe *commettre*, le C1 est interprétable comme référant à un individu, alors qu'en (11-b), où il est complément du verbe *perpétrer*, il est interprétable comme un nom collectif :

- (11) a. Il est vrai que les **arabicides** ne sont pas nouveaux. Ce qui est nouveau, à mon avis, c'est la fréquence nouvelle des **arabicides** commis par des policiers. (<http://blogs.mediapart.fr/blog/naja/010709/devant-cette-mort-la-serons-nous-aussi-habitués-resignés-soumis-qu-nous-le-dem>)
- b. Il accuse le FRODEBU de s'être allié au PALIPEHUTU pour perpétrer un « **tutsicide** ». Même si l'auteur reconnaît qu'au Burundi des Tutsi et des Hutu sont morts uniquement à cause de leur appartenance ethnique. (<http://repositories.lib.utexas.edu/bitstream/handle/2152/4119/2304.pdf?sequence=1>)

Si l'analyse distributionnelle nous donne ici quelques outils pour identifier le sens des CNs et les interpréter, elle ne permet pas de rendre compte de la totalité des données. En effet, il faudrait envisager des moyens permettant d'opérer une analyse distributionnelle de plus grande envergure, de manière automatique. Cependant, une telle analyse demanderait non seulement une

base de données pouvant englober les exemples que seule la Toile m'a fournis, ainsi qu'une annotation précise des contextes, ce que la Toile n'offre évidemment pas. Seuls des corpus ne contenant qu'une partie des lexèmes de NeoClassy sont ainsi annotés, ce qui restreindrait donc l'analyse.

Cependant, on peut déjà remarquer que le contexte a un rôle prépondérant dans l'étude du sens des mots construits; il permet par exemple de désambiguïser certains emplois, comme nous l'avons vu pour *souricide* en (10). J'aborderai dans la section 5.1.1.2 du chapitre suivant la question de la polysémie; mais si l'on considère que « une opération constructionnelle sélectionne une acception, quand la base est polysémique » (Roché, 2009), alors le contexte peut guider l'identification de la base de la construction, ainsi que celle du schéma utilisé. Le contexte *Xcide INVOLONTAIRE* sous-entend ainsi une individualisation de X; *souricide*, en (10-a), est donc très certainement construit sur l'acception 'animal domestique' de SOURIS, et non pas sur celle 'animal nuisible'.

Dans la section suivante, je présente le même type d'analyse, en prenant comme objet d'étude les constructions impliquant l'élément *-phobie* qui a la particularité de pouvoir être utilisé de façon autonome.

4.1.1.3 Les *-phobie* en contexte

J'ai donc effectué une analyse similaire à celle que je viens de présenter concernant les lexèmes *Xphobie* de NeoClassy. Pour rappel, trois types principaux avaient été dégagés précédemment, ils sont reportés en (12). Le type 2 se divise en deux sous-schémas, distingués par le type sémantique du C1.

- (12) a. Type 1 : $[X_i \text{ phobie}]_{N_j} \leftrightarrow [\text{peur de SEM}_i]_j$
 b. Type 2 :
 (i) $[X_i \text{ phobie}]_{N_j} \leftrightarrow [\text{aversion envers SEM}_i]_j$
 (ii) $[X_{[+\text{humain}]_i} \text{ phobie}]_{N_j} \leftrightarrow [\text{hostilité envers SEM}_i]_j$
 c. Type 3 : $[X_i \text{ phobie}]_{N_j} \leftrightarrow [\text{répulsion chimique envers SEM}_i]_j$

Cette analyse distributionnelle s'applique uniquement aux types 1 et 2, puisque le troisième est réservé à un domaine particulier, celui de la chimie, et que le type de discours dans lequel les lexèmes formés selon ce schéma est potentiellement suffisant pour les désambiguïser.

Le premier pas de cette analyse est identique à ce qui vient d'être opéré pour les lexèmes *Xcide*. La méthode est appelée descendante puisqu'elle part des contextes préférentiels avec lesquels les mots leader sont utilisés pour les rechercher ensuite avec les autres mots du corpus. Elle s'oppose à la méthode inductive qui lui fera suite, qui, elle, part des contextes avec lesquels sont

employés tous les mots du corpus pour ensuite rapprocher ces mêmes mots selon les contextes qu'ils partagent.

Méthode descendante. Nous avons vu dans la section 3.2.2.1 que les schémas impliquant l'EN *-phobie* s'organisent en pôles distincts, avec au centre de chacun d'eux un groupe de mots leader. Le Tableau 4.4 résume le Tableau 3.8 du chapitre précédent et liste les dix mots les plus fréquents (sur la Toile) ; par chance cinq appartiennent au type 1 et cinq autres au type 2. Je considérerai ici que ces dix mots font partie des groupes de mots leader servant de pôles d'attractions pour la création de nouveaux lexèmes.

Mot leader	Type 1	Type 2
homophobie		✓
xénophobie		✓
islamophobie		✓
dysmorphophobie	✓	
agoraphobie	✓	
christianophobie		✓
transphobie		✓
claustrophobie	✓	
photophobie	✓	
aquaphobie	✓	

Tableau 4.4 – mots leader correspondant aux deux schémas principaux construisant des *Xphobie*

Une fois ces mots leader identifiés, j'ai examiné leurs contextes d'apparition privilégiés. Pour cela, j'ai étendu la méthode employée pour les lexèmes *Xcide*. En plus de l'utilisation de Frantext, vdW et vdM, j'ai sélectionné un échantillon constitué des deux premières pages de résultats obtenus sur Google, pour des requêtes au singulier et au pluriel, pour extraire la distribution des mots leader. En effet, comme je l'avais noté dans le Tableau 3.8 du chapitre précédent, plusieurs mots parmi les plus fréquents ne sont pas dictionnarisés. Trouver leurs contextes d'apparition dans les ressources plus traditionnelles ne sera donc pas possible. *Christianophobie*, par exemple, n'est pas contenu dans Frantext, vdW ou vdM. On ne peut donc pas avoir accès par ces biais là à ses contextes d'apparition. L'investigation des deux premières pages de résultats Google pour les requêtes « christianophobie » et « christianophobies » a fait quant à elle ressortir le contexte **MANIFESTER CONTRE_Xphobie**, caractéristique des lexèmes *Xphobie* de type 2.

Là encore, étant donné qu'il s'agit d'une analyse effectuée de façon manuelle, il a fallu sélectionner certains contextes qui étaient partagés parmi les mots leader de chaque groupe et uniquement par ceux du même groupe : par exemple, le contexte `LUTTER CONTRE_Xphobie` est employé de manière très fréquente avec les lexèmes des deux types, il n'a donc pas été sélectionné pour l'analyse. Les contextes choisis sont listés dans le Tableau 4.5.

Type	Contextes les plus fréquents	Recherches Google
1	<code>SOUFFRIRDE_Xphobie</code> <code>SOIGNERO_Xphobie</code> <code>CRISEDE_Xphobie</code>	« souffrir de (d') Xphobie » « souffre de (d') Xphobie » « souffres de (d') Xphobie » « soigner * Xphobie » « soigne * Xphobie » « soignes * Xphobie » « crise de (d') Xphobie » « crises de (d') Xphobie »
2	<code>VAGUEDE_Xphobie</code> <code>Xphobie_AMBIANTE</code> <code>Xphobie_LATENTE</code>	« vague de (d') Xphobie » « vagues de (d') Xphobie » « Xphobie ambiante » « Xphobie latente »

Tableau 4.5 – Contextes sélectionnés pour l'analyse distributionnelle des CNs *Xphobie*

Une fois ces contextes sélectionnés, la démarche a été la même que précédemment pour les lexèmes *Xcide* : chaque lexème en *-phobie* de NeoClassy a été recherché avec chacun des contextes sélectionnés dans Google, selon la requête décrite dans la colonne de droite du Tableau 4.5. Chaque fois qu'il y avait au moins un résultat, le lexème *Xphobie* était noté comme pouvant être utilisé avec le contexte. Les résultats sont donnés dans le Tableau 4.6.

Encore une fois, la liste des contextes sélectionnés devrait être étendue et l'analyse faite de façon automatique pour pouvoir donner de meilleurs résultats : en effet, la proportion de lexèmes n'ayant pas pu être traités sémantiquement grâce à cette analyse distributionnelle est à peu près similaire que celle que nous avons pour les lexèmes *Xcide*, à savoir 45%. Cependant, une recherche automatique est difficile à exécuter sur Google, et on a déjà vu que la plupart des lexèmes du corpus ont des fréquences très faibles ; il serait difficile d'y avoir accès sur d'autres corpus, même les plus étendus. 223 lexèmes du corpus, soit 42,15% des *Xphobie* ont moins de 1000 occurrences

Types dégagés à partir des contextes recherchés		Nombre	% sur le total
1	2		
X	–	163	31%
–	X	43	8%
X	X	85	16%
–	–	238	45%

Tableau 4.6 – Résultats de l’analyse distributionnelle effectuée pour les lexèmes *Xphobie* (méthode descendante)

sur Google, leur absence dans les collocations avec les contextes sélectionnés peut donc être fortuite ; la probabilité de trouver des néologismes dans un contexte donné est en effet souvent faible. Par exemple, *kurdophobie* n’a que 840 occurrences sur Google⁶ ; même si on ne le trouve avec aucun des contextes privilégiés par le schéma de type 2, l’exemple ne laisse pas beaucoup de doutes quant à son interprétation.

- (13) La **kurdophobie** des Turcs conduit donc a de nombreuses aberrations de ce type. (...) la Turquie a encore beaucoup de chemin a faire dans le respect des droits du peuple kurde. (http://istanbul.blog.lemonde.fr/2006/03/22/2006_03_la_lettre_inter/)

Le type de C1 est également, comme c’était le cas pour les lexèmes *Xcide*, un indice pour l’interprétation des lexèmes *Xphobie* : dans *kurdophobie*, le C1 est un ethnique, la probabilité pour que le trait ‘hostilité’ soit présent dans le mot construit est donc forte.

Puisque la méthode descendante n’est pas pleinement satisfaisante, j’ai essayé d’appliquer la démarche inverse, c’est-à-dire de partir des contextes avec lesquels chaque lexème *Xphobie* est employé pour ensuite faire des recoupements entre ces contextes et regrouper les lexèmes grâce aux contextes qu’ils partagent.

Méthode inductive. Il ne s’agit donc plus ici de rechercher la co-occurrence d’un lexème avec un contexte privilégié dans la construction d’un type de schéma mais au contraire de partir des contextes dans lesquels chaque lexème

6. Recherche du 3 octobre 2015.

est effectivement utilisé pour ensuite en dégager des généralités. Dans cette méthode, le contexte est pleinement considéré comme faisant partie du réseau lexical dans lequel les C1, les mots construits (par exemple les mots leader) et les schémas de construction sont déjà reliés. En effet, dans un modèle de morphologie constructionnelle, les contextes particulièrement fréquents peuvent être intégrés dans le schéma de construction ou du moins être pris en compte, par exemple comme facteur déclenchant l'utilisation d'un schéma de construction. Le Tableau 4.7 illustre l'insertion des contextes dans le réseau lexical.

	X	Contextes de X
a.	claustrophobie	SOUFFRIRDE_X ; SOIGNERO_X
	homophobie	LUTTERCONTRE_X ; X_PRIMAIRE
	christianophobie	LUTTECONTRE_X ; MANIFESTERCONTRE_X ; OBSERVATOIREDE_X
	couplrophobie	SOUFFRIRDE_X ; X_AIGU ; X_LÉGER ; ATTEINDREDE_X
	microsoftphobie	X_PRIMAIRE ; X_AIGU
b.	lésion cérébrale ; affection neurologique	SOUFFRIRDE_X
	maladie mortelle ; mal incurable	ATTEINDREDE_X
	glaucome ; rhumatisme ; infarctus	X_AIGU
	terrorisme international ; trafic de stupéfiant	LUTTECONTRE_X
	suppression d'emploi ; insécurité ; racisme	MANIFESTERCONTRE_X
	anticléricisme ; antiaméricanisme	X_PRIMAIRE

Tableau 4.7 – Réseau de contextes

Ce tableau s'appuie sur des données relevées dans vdW et vdM ainsi que, pour les mots non contenus dans ces ressources, sur les contextes d'apparition relevés sur la Toile. En (a.), sont affichés des lexèmes *Xphobie* : *claustrophobie* et *homophobie* sont des mots leader des schémas de type 1 et 2, *christianophobie* a été noté comme un des lexèmes *Xphobie* les plus fréquents, mais n'est pas dictionnarisé, *couplrophobie* et *microsoftphobie* sont des néologismes recueillis sur la Toile. La colonne de droite affiche certains

de leurs contextes recensés dans vdW et vdM⁷, et sélectionnés en fonction de la mesure de l'information mutuelle, ou des contextes relevés sur la Toile. En (b.), sont exposés des lexèmes ou des syntagmes qui ne sont pas formés avec l'EN *-phobie* et qui sont employés avec les mêmes contextes que les lexèmes en (a.). Le tableau nous montre donc qu'un néologisme comme *couplophobie* est utilisé avec les contextes `SOUFFRIRDE_X`, `X_aigu` et `ATTEINDREDE_X`, exactement comme des noms de maladies. *Couplophobie* appartient donc à la même classe distributionnelle que *glaucome* ou *infarctus*, puisque ces lexèmes partagent les mêmes contextes. Le mot leader *claustrophobie* partage avec *couplophobie* le contexte `SOUFFRIRDE_X`; ils appartiennent donc également à la même classe de distribution. *Couplophobie* a certainement été construit par analogie avec le modèle de *claustrophobie*. *Christianophobie* est quant à lui utilisé avec les contextes `LUTTECONTRE_X` ou `MANIFESTERCONTRE_X` qui sont également des contextes privilégiés pour non pas des noms de maladies mais des noms à connotation péjorative; *christianophobie*, *insécurité*, *racisme*, appartiennent donc ainsi à une classe distributionnelle différente de la précédente. De plus, *christianophobie* et *homophobie* partagent certains contextes qui les font également appartenir à la même classe de distribution : ils ont des sens liés. *Microsoftophobie* appartient aux deux classes de distribution qui viennent d'être dégagées : il partage des contextes aussi bien avec *homophobie* qu'avec *couplophobie*, comme le montre l'exemple (14) :

- (14) a. Souffrant de **Microsoftophobie** aigue, j'essaie de ne pas utiliser les programmes mis "gracieusement" à disposition par Bill Gates. (<http://www.info-sumo.net/forum/archive/index.php/t-13.html?s=46041d46889aa266a3a574269f3ea878>)
- b. Je ne veux pas défendre Microsoft pour sa position monopolistique, mais c'est un peu lassant cette **Microsoftophobie** primaire et systématique. (<http://www.cousinsfringants.net/forum3//viewtopic.phpf=26&t=1251&st=0&sk=t&sd=a&hilit=contact&start=5100>)

La complexité du réseau lexical est due aussi bien aux contextes partagés par les différents lexèmes qu'à la distance existant entre le sens de chaque lexème.

Partant de cette observation, j'ai tenté de formaliser les liens existant entre les lexèmes *Xphobie* à travers les contextes dans lesquels chacun d'entre eux ap-

7. Les contextes affichant le verbe `LUTTERCONTRE_X` et le nom déverbal `LUTTECONTRE_X` apparaissent dans des entrées séparées dans ces ressources. C'est pourquoi ils sont ici différenciés. De même, `ATTEINDREDE_X`, même s'il est employé uniquement sous sa forme participiale, apparaît dans vdW et vdM sous la forme de citation du verbe. Je respecte ici cette notation.

paraît. Pour cela, pour chaque lexème *Xphobie* de NeoClassy, j’ai sélectionné un échantillon constitué des deux premières pages de résultats Google au singulier et au pluriel, et j’ai relevé les contextes avec lesquels les lexèmes *Xphobie* étaient utilisés à l’intérieur de cet échantillon. Chacun de ces contextes a ensuite été indexé selon ses traits sémantiques avec l’aide des vdW et vdM : par exemple, les vdW présentent le contexte **ACCUSATIONDE_X** comme étant préférentiellement utilisé avec $X = \textit{abus de pouvoir, corruption ou racisme}$, c’est-à-dire $X = \textit{‘attitude à combattre’}$. Le contexte **ACCUSATIONDE_X** sera donc considéré comme étant utilisé pour des lexèmes *Xphobie* du type 2 qui appartiennent à la même classe de distribution. Le contexte **ATTEINDREDE_X** apparaît seulement avec des noms de maladies tels que *daltonisme* ou *tuberculose* ; les lexèmes *Xphobie* qui sont également employés avec **ATTEINDREDE_X** appartiennent donc à la même classe distributionnelle que ces noms de maladie, ils sont donc construits selon le schéma de type 1. Si le traitement de ces deux contextes est assez évident, certains comme **SE PLAINDREDE_X** prêtent à davantage d’ambiguïtés : il est en effet utilisé avec des lexèmes comme *injustice*, c’est-à-dire une ‘attitude à combattre’ ou *douleur abdominale*, une maladie. Ce contexte apparaîtra donc potentiellement avec des lexèmes *Xphobie* de type 1 et 2. Enfin, il a été impossible pour certains contextes de les attribuer à un type en particulier, puisqu’ils n’étaient utilisés ni avec des attitudes à combattre, ni avec des maladies : c’est le cas par exemple de **VÉRITABLE_X**, qui est employé avec plusieurs lexèmes du corpus mais qui n’est propre à aucun type.

Les résultats de l’analyse distributionnelle dans sa version inductive sont donnés dans le Tableau 4.8.

Contextes	Nombre de lexèmes		
	Exclusivement employés avec ce type de contextes	Employés avec ce type de contextes ainsi que d’autres	Total
Type 1	86	62	148
Type 2	44	100	144
Types 1 et 2	15	50	65
autre contexte	47	101	148
aucun contexte	203		

Tableau 4.8 – Résultats de l’analyse distributionnelle effectuée pour les lexèmes *Xphobie* (méthode inductive)

On remarque tout d’abord que 203 lexèmes, soit 38% des *Xphobie*, n’ont pas

pu être interprétés selon cette méthode, car aucun contexte n'a pu être dégagé dans l'échantillon de pages Google sélectionné. Il s'agit en général soit de lexèmes très courants, lexicalisés, et les deux premières pages n'affichent que des entrées de dictionnaires et donc aucun contexte syntaxique, soit de lexèmes qui n'apparaissent que dans des listes de phobies (voir l'exemple (4) du chapitre précédent), ne laissant donc transparaître également aucun contexte syntaxique. Cette méthode a cependant l'avantage de pouvoir décrire sémantiquement davantage de lexèmes que la méthode descendante (62% contre 55%).

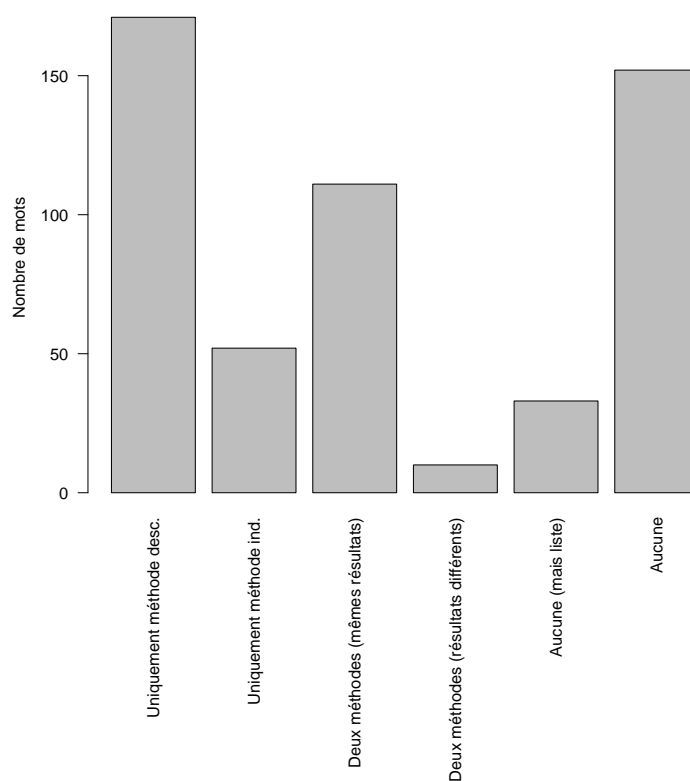


FIGURE 4.2 – Nombre de résultats en fonction de la méthode employée

Ces méthodes sont de plus complémentaires : 344 lexèmes sont captés par les deux méthodes, 171 uniquement par la méthode inductive et 52 uniquement par la méthode descendante. Il ne reste donc que 185 lexèmes, soit 35% du corpus, qui n'ont pu être décrits par aucune des deux méthodes. Si l'on exclut les 33 lexèmes qui ne se trouvent que dans des listes de phobies, et par

conséquent sont construits selon le schéma de type 1, 152 lexèmes ne peuvent pas être interprétés grâce à l'analyse distributionnelle, soit 28% du corpus. La Figure 4.2 résume ces résultats. On voit que pour 10 lexèmes seulement qui sont décrits par les deux méthodes, les résultats sont différents, c'est-à-dire que l'une attribue le sens de 'peur' et l'autre celui de 'hostilité'. Cela peut être dû au fait que, comme nous l'avons déjà vu à plusieurs reprises, un même lexème peut être construit selon des schémas différents ; il suffira que l'une ou l'autre des méthodes n'ait réussi à capter qu'une seule de ces constructions.

Nous avons vu avec les lexèmes *Xcide* que le type sémantique des bases était corrélé avec le type de schéma dans lequel entrait le mot construit. Pour aller plus loin, regardons s'il existe une corrélation entre la forme et le sens : y a-t-il des différences étymologiques et stylistiques entre les *Xphobie* 'peur de X' et *Xphobie* 'hostilité envers X' ? La Figure 4.3⁸ met en lumière des différences du point de vue du C1 selon le schéma utilisé dans la formation des lexèmes.

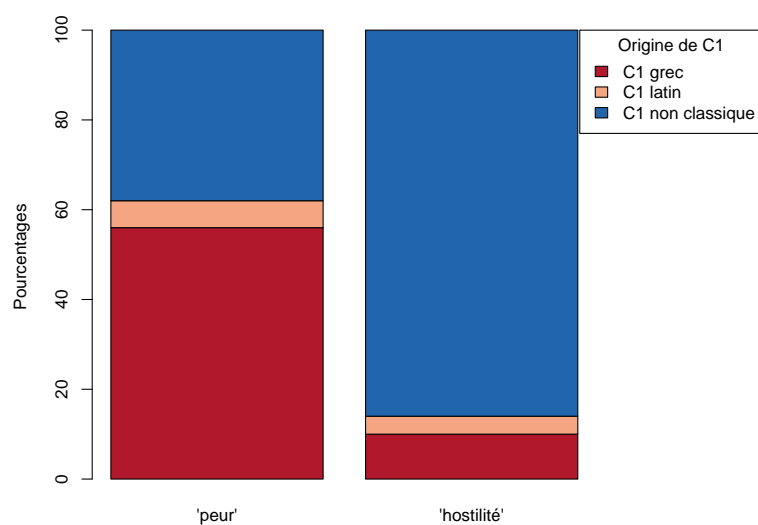


FIGURE 4.3 – Origine du C1 en fonction du sens de *Xphobie* (n = 377)

8. Cette figure prend en compte tous les lexèmes auxquels j'ai pu attribuer un sens par l'une ou l'autre des deux méthodes distributionnelles que j'ai employées, additionnés de ceux faisant partie de listes et notés comme des peurs.

Les lexèmes *Xphobie* portant le sens de ‘peur’ ont une plus forte proportion de C1 classiques (62%) alors que ceux portant le sème de ‘hostilité’ affichent une préférence pour des C1 non classiques (seuls 14% sont classiques). La contrainte de crypticité évoquée dans la partie 3.1.1 peut expliquer cette différence : lorsque *Xphobie* porte le sens de ‘peur’, le lexème dénote une maladie ; le mot doit donc avoir de préférence l’apparence d’un terme scientifique. Quant il s’agit d’une hostilité envers une personne ou un groupe de personnes, la cible a besoin d’être comprise facilement et par tous : la probabilité d’avoir un C1 non classique est donc plus forte.

4.2 Les relations inter-schématiques singulières

Si le contexte est un élément essentiel dans la construction et l’interprétation du sens des mots construits, les relations qui existent entre certains mots à l’intérieur du lexique en est un autre. Toutes les constructions impliquant un même EN peuvent ainsi ne pas être dérivées de manière identique. Roché (2011a) montre par exemple que, pour les dérivés en *-isme* construits dans le modèle qu’il désigne comme le modèle 3, construisant des noms de qualité, il n’y a normalement pas de dérivé en *-iste* (15-a), contrairement au modèle formant des dérivés axiologiques (15-b) et celui formant des dérivés processifs et agentifs (15-c).

- (15) a. $\text{parallèle}_A \rightarrow \text{parallélisme}_N$
 b. $\text{esclavage}_N \rightarrow \text{esclavagisme}_N \leftrightarrow \text{esclavagiste}_{N/A}$
 c. $\text{parachute}_N \rightarrow \text{parachutisme}_N \leftrightarrow \text{parachutiste}_{N/A}$

Une même observation peut être faite pour certains des ENs étudiés ici. De plus, les relations qui existent entre certains schémas impliquant des ENs différents peuvent aussi être un signe d’appartenance à un schéma en particulier, comme nous le verrons par la suite.

4.2.1 Dérivés possibles et impossibles

Pour rester avec les lexèmes *Xphobie*, il existe parfois des dérivés *Xphobique*. Tout comme le lexème PHOBIE existe de façon autonome avec le sens ‘peur panique’, le lexème PHOBIQUE existe : c’est un adjectif signifiant ‘relatif à la phobie, à une phobie’ ou un nom référant à une personne ‘atteinte de phobie’ (les définitions sont extraites du GR)⁹. Le GR recense sept lexèmes

9. Contrairement à l’analyse que je ferai de la séquence *-logique* plus bas (section 4.3.1), je ne pense pas que l’on puisse considérer *-phobique* comme un EN à part entière. En effet, il

Xphobique, le TLFi trois. Le Tableau 4.9 les liste tous, ainsi que les *Xphobie* correspondants et le schéma selon lequel ils sont construits.

	Xphobique	Cat.	Xphobie	Sens
GR	agoraphobique	A	agoraphobie	‘peur’
	hydrophobique	AN	hydrophobie	‘peur’ ‘répulsion chimique’
	nécrophobique	A	nécrophobie	‘peur’
	nosophobique	AN	nosophobie	‘peur’
	photophobique	A	photophobie	‘peur’
	sociophobique	A	sociophobie	‘peur’
	toxicophobique	A	toxicophobie	‘peur’
TLFi	hydrophobique	AN	hydrophobie	‘peur’
	photophobique	A	photophobie	‘peur’
	préphobique	A		

Tableau 4.9 – *Xphobique* répertoriés dans les dictionnaires de référence

Les *Xphobique* dictionnarisés sont tous dérivés de lexèmes *Xphobie* construits selon le schéma désignant des peurs et formant des noms de maladie. Beaucoup sont en fait des synonymes de lexèmes *Xphobe* : *agoraphobique*, par exemple, dans le Grand Robert, est indiqué dans l’entrée *agoraphobe* comme synonyme. De même, les deux mots se retrouvent dans des contextes similaires :

- (16) a. je suis **agoraphobique** et suis suivi par une psy ! et je n’ai pas de remède à part les calmants mais ma psy n’a pas voulu m’en prescrire :/ (http://forum.doctissimo.fr/sante/thyroide-problemes-endocrinologiques/aidez-comprendre-sujet_150934_1.htm)
- b. bonjour, je suis **agoraphobe** depuis 14 ans et suivi par un psychiatre . (<https://www.entrepaticients.net/fr/communautes/groupes/fibromyalgie/discussion/69659>)

Or, si les lexèmes *Xphobe* ne sont pas uniquement construits sur le schéma ‘qui a peur de X’, mais également sur celui ‘qui est hostile envers X’, ceux

semble uniquement dériver de noms *Xphobie*, auxquels il se rapporte nécessairement. Alors que les lexèmes *Xlogique* peuvent se rapporter directement à X, cela n’a pas l’air possible pour les lexèmes *Xphobique*. Cependant, il faut garder à l’esprit que les lexèmes *Xphobique* font partie du même réseau que les lexèmes *Xphobie* et *Xphobe*.

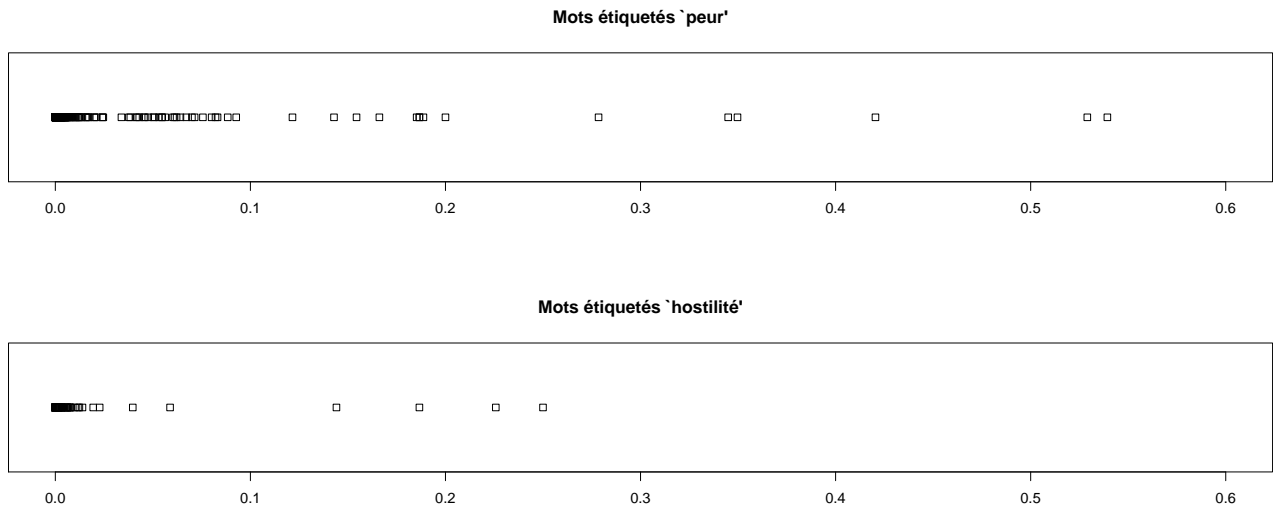


FIGURE 4.4 – Taux de *Xphobique* par rapport aux *Xphobie* étiquetés comme ‘peur’ ou ‘hostilité’

Construction Xphob-	Nombre d'entrées	Catégorie
Xphobia	126	N
Xphobic	47	A/N
Xphobe	46	A/N
Xphobist	11	A/N
Xphobous	7	A
Xphoby	7	N
Xphobiac	5	A/N
Xphobian	3	A/N
Xphobial	1	A

Tableau 4.10 – Constructions *Xphob-* présentes dans l'OED

Deux constructions, *Xphobe* et *Xphobic* semblent être en compétition pour former des noms et des adjectifs ; Lindsay et Aronoff (2015) parlent en effet de compétition lorsque deux processus de construction sont synonymes. Pour eux, cela ne peut se produire synchroniquement. Pour résoudre cette compétition, l'un des deux processus doit s'éteindre, c'est-à-dire être bloqué par l'autre, ou chacun doit se spécialiser, se différencier sémantiquement : ils s'évitent. Comment la compétition entre *Xphobe* et *Xphobic* est-elle envisagée en anglais ? Pour répondre à cette question, j'ai opéré une analyse en m'appuyant sur l'OED. Cette ressource montre que les lexèmes *Xphobia* de l'anglais sont construits selon les mêmes schémas que les *Xphobie* du français. Certains sont d'ailleurs des internationalismes (voir section 1.1.2).

- (19) a. $[X_i \text{ phobia}]_{N_j} \leftrightarrow [\text{peur de SEM}_i]_j$
 > *nyctophobia*
 b. $[X_i \text{ phobia}]_{N_j} \leftrightarrow [\text{aversion envers SEM}_i]_j$
 > *teatrophobia*
 c. $[X_i \text{ phobia}]_{N_j} \leftrightarrow [\text{hostilité envers SEM}_i]_j$
 > *Anglophobia*
 d. $[X_i \text{ phobia}]_{N_j} \leftrightarrow [\text{répulsion chimique envers SEM}_i]_j$
 > *lipophobia*

Les lexèmes *Xphobe* sont-ils davantage reliés à un schéma de *Xphobia* en particulier ? Et les lexèmes *Xphobic* ? Chaque lexème *Xphobia* présent dans l'OED a été étiqueté sémantiquement en fonction de la définition donnée dans l'OED. Ensuite, pour chacun, les correspondants *Xphobe* et *Xphobic* ont été recherchés sur Google et leur nombre de résultats relevés. Or Google n'est pas un corpus annoté syntaxiquement et il est donc difficile de comparer les *Xphobe*_A et les *Xphobic*_A ou les *Xphobe*_N et les *Xphobic*_N. Or pour beaucoup de paires, les noms et les adjectifs utilisant la même séquence ne sont pas présents en quantité similaire et les résultats s'en retrouvent affectés. Afin de neutraliser cette différence catégorielle et comparer uniquement les *Xphobe*_N et les *Xphobic*_N, je n'ai interrogé Google que sur les pluriels ; en effet, les adjectifs ne prennent aucune marque graphique au pluriel en anglais. Les séquences <Xphobes> et <Xphobics> ne peuvent pas conséquemment être que des noms. Le Tableau 4.11 rend compte des résultats de la recherche effectuée le 22 novembre 2013. Les séquences qui offraient le plus grand nombre de résultats ont été considérées comme gagnantes de la compétition entre les deux processus de construction. Le nombre de gagnants a ensuite été regroupé sur la base du schéma selon lequel le lexème *Xphobia* était construit.

Les résultats affichent donc sans conteste que *Xphobic*_N correspondent à des

	Sens principal (selon l'OED)					
	'peur'	'aversion'	'hostilité'	'répulsion chimique'	'peur' ou 'aversion'	'peur' ou 'hostilité'
Winner : <Xphobes>	56	13	28	10	10	0
Winner : <Xphobics>	19	3	0	0	1	1

Tableau 4.11 – Séquence, entre <Xphobes> et <Xphobics>, la plus présente sur Google, en fonction du sens de *Xphobia* correspondant

lexèmes *Xphobia* dénotants des peurs voire des aversions, jamais quand il s'agit de *Xphobia* dénotant une hostilité ou une répulsion chimique.

Ce que ces études sur le français et l'anglais nous montrent est que les schémas qui ont été distingués dans la section 3.2.1 ne diffèrent pas uniquement au niveau du sens attribué à chaque construction ; plus que ça, la position de chaque schéma se singularise selon les relations que ce schéma entretient avec d'autres dans le lexique. Ainsi, la possibilité de rencontrer un dérivé – ou l'impossibilité – constitue un indice pour interpréter le sens d'un mot construit. Par ailleurs, plus que dans la dérivation, les liens entretenus par différents schémas impliquant un EN particulier avec des schémas impliquant d'autres ENs peuvent également différer.

4.2.2 Des liens privilégiés dans le réseau lexical

Dans le chapitre précédent, la Figure 3.34 montrait un réseau de schémas qui, quoiqu'impliquant les mêmes ENs finaux, affichaient des sens différents voire antinomiques. Nous avons ainsi vu que, lorsque *Xphage* signifie 'amateur de X', il est synonyme, ou du moins sémantiquement proche de *Xphile* ou *Xmane*, mais lorsqu'il signifie 'hostile envers X', il est comparable à *Xphobe*. L'hypothèse ici est que les liens existant entre certains schémas permettent de déterminer sémantiquement selon lequel d'entre eux est construit un mot donné, et par conséquent quel est son sens. Par exemple, pour rester avec l'élément *-phobe*, une paire antinomique comme *francophobe/francophile* est plus attendue que la paire *agoraphobe/agoraphile*. Intuitivement, un lexème *Xphile* est en effet le contraire d'un lexème *Xphobe* dénotant une hostilité mais pas d'un lexème *Xphobe* dénotant une peur. Pour vérifier cette intuition, j'ai eu la même démarche que celle précédemment pour vérifier le lien entre

les *Xphobique* et les *Xphobie* : pour chaque lexème *Xphobe* de NeoClassy, j'ai recherché sur Google son contraire *Xphile*, et j'ai calculé son taux de présence :

$$P = \frac{(\text{occurrences} < Xphile > + \text{occurrences} < Xphiles >)}{(\text{occurrences} < Xphobe > + \text{occurrences} < Xphobes >)}$$

Les lexèmes *Xphile* étant bien plus présents sur la Toile que les lexèmes *Xphobique*, les valeurs de P obtenues sont par conséquent plus importantes¹². Ensuite, j'ai comparé le taux de présence de *Xphile* pour les *Xphobe* portant le sens de 'peur' et pour ceux portant le sens 'hostilité'. Cependant les résultats ne sont pas ceux escomptés (Figure 4.5).

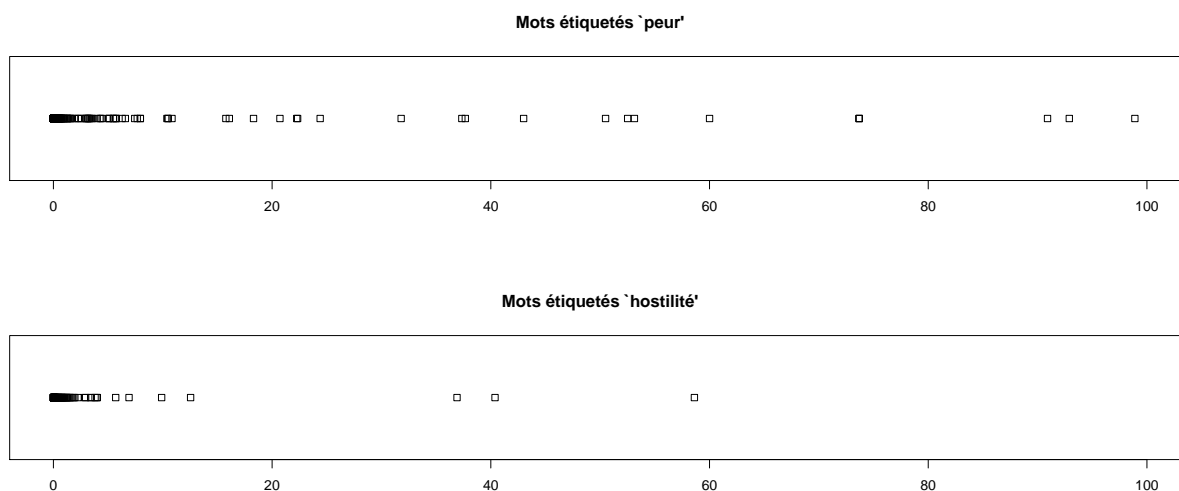


FIGURE 4.5 – Taux de *Xphile* par rapport aux *Xphobe* étiquetés comme 'peur' ou 'hostilité'

Comme il était attendu, les valeurs de P lorsqu'un *Xphile* partage le C1 d'un *Xphobe* portant le sens 'hostilité' ne sont pas égales à zéro, ce qui implique une présence du doublet. Cependant, la Figure 4.5 montre des valeurs de P supérieures lorsque *Xphobe* porte le sens de 'peur'. Même si la présence de *Xphile* pour un *Xphobe* portant le sens de 'hostilité' est plus fréquente que lorsqu'il s'agit de 'peur' (99% contre 94%), les valeurs plus fortes de P montrent qu'il existe un schéma impliquant l'EN *Xphile* qui est également lié au schéma *Xphobe* 'qui a peur de X'. Ce schéma est celui pour lequel

12. Les six valeurs extrêmes $P > 100$ ont été retirées des analyses.

X peut désigner aussi bien une entité susceptible de causer une peur qu'un objet fétiche que l'on peut être amené à collectionner :

- (20) hippophobe 'qui a peur des chevaux' *vs.* hippophile 'qui adore les chevaux'
 pyrophobe 'qui a peur du feu' *vs.* pyrophile 'qui adore le feu'
 insectophobe 'qui a peur des insectes' *vs.* insectophile 'qui aime les insectes, qui collectionne les insectes'

Ainsi, chaque schéma impliquant l'élément *-phobe* est relié à un schéma différent reliant l'élément *-phile* :

- (21) a. $[X_i \text{ phobe}]_{AN_j} \leftrightarrow$ [qui est hostile envers $SEM_i]_j$
 \updownarrow
 b. $[X_i \text{ phile}]_{AN_k} \leftrightarrow$ [qui entretient des relations amicales avec $SEM_i]_k$
 $[X_i \text{ phobe}]_{AN_j} \leftrightarrow$ [qui a peur de $SEM_i]_j$
 $[X_i \text{ phile}]_{AN_k} \leftrightarrow$ [qui est amateur, collectionneur de $SEM_i]_k$

Une étude uniquement basée sur les occurrences recensées sur Google ou dans un corpus ne peut donc pas être suffisante. Cependant, une analyse distributionnelle qui étudierait les contextes partagés par chacun des *Xphobe* et *Xphile* serait intéressante à réaliser.

Ces relations singulières qui s'instaurent entre schémas ou sous-schémas se retrouvent à plusieurs endroits du lexique. Par exemple, à certains lexèmes *Xcole* correspondent des lexèmes *Xculture* et *Xculteur* :

- (22) apicole / apiculture / apiculteur

Pour d'autres, aucun lexème *Xculture* ne correspond :

- (23) cavernicole / *caverniculture / *caverniculteur

Dans la partie 3.2.1 j'avais dégagé trois schémas principaux qui impliquaient l'élément *-cole*, que je reproduis en (24).

- (24) a. $[X_i \text{ cole}]_{A_j} \leftrightarrow$ [qui honore $SEM_i]_j$
 b. $[X_i \text{ cole}]_{A_j} \leftrightarrow$ [relatif à la culture de $SEM_i]_j$
 c. $[X_i \text{ cole}]_{A_j} \leftrightarrow$ [qui habite dans / sur $SEM_i]_j$

Or si l'on observe l'ensemble des lexèmes *Xcole*, *Xculture* et *Xculteur* présents dans les dictionnaires (TLFi et GR), on remarque que seul le schéma (24-b) est relié aux schémas impliquant les éléments *-culture* et *-culteur* :

- (25) a. $[X_i \text{ culture}]_{N_j} \leftrightarrow [\text{culture}/\text{élevage de SEM}_i]_j$
 b. $[X_i \text{ cole}]_{N_k} \leftrightarrow [\text{relatif à SEM}_j]_k$
 c. $[X_i \text{ cultivateur}]_{N_l} \leftrightarrow [\text{qui cultive SEM}_i, \text{ qui pratique SEM}_j]_l$

Ainsi, lorsqu'un lexème *Xcole* pouvait avoir été construit selon des schémas distincts, le lexème *Xculture*, s'il existe, ne correspond qu'au *Xcole* construit selon le schéma (24-b) :

- (26) a. arboricole 'qui vit dans les arbres'
 *arboriculture
 b. arboricole 'qui a rapport à la culture des arbres'
 arboriculture

Inversement, un lexème *Xculture* peut ne pas avoir de *Xcole* correspondant, ni même de *Xculteur*, dans le domaine de la médecine :

- (27) hémoculture : 'ensemencement d'un milieu de culture avec du sang pour la recherche des microbes susceptibles de s'y trouver'
 myéloculture : 'mise en culture d'un fragment de moelle osseuse pour faire pousser les germes pathogènes présumés responsables d'une infection'
 pyoculture : 'culture de bactéries faites avec un prélèvement de pus'¹³

Le GR fait également état d'un couple de lexèmes *Xculture-Xculteur* pour lequel aucun *Xcole* ne correspond :

- (28) capilliculture 'ensemble des soins donnés à la chevelure'
 capilliculteur 'coiffeur spécialiste des soins de la chevelure'
 *capillicole

Ainsi, chaque lexème construit son sens en fonction de ses constituants, de son contexte, mais aussi des liens entretenus par le schéma dans lequel il est construit avec d'autres éléments lexicaux, qu'il s'agisse de schémas dérivés ou de schémas impliquant d'autres éléments. La position occupée par un lexème construit à l'intérieur du lexique dépend donc (au moins) de ces trois facteurs. Chacun sera un indice, une clé, dans l'élaboration de son sens ainsi que son interprétation. Les schémas impliquant des éléments néoclassiques ne semblent pas si différents de ceux impliquant des affixes prototypiques. Même, comme nous allons maintenant le voir, les éléments néoclassiques et les affixes, du point de vue de leur charge sémantique ou de leur distribution, pourraient tout à fait être étudiés de façon similaire.

13. Les définitions sont extraites du GR.

4.3 La charge sémantique des éléments néo-classiques

Une des principales différences identifiées dans la littérature entre les affixes et les constituants de composés se trouve du côté du sens : les seconds auraient un sens lexical quand les premiers auraient un sens instructionnel (voir section 1.3). Or les éléments néoclassiques peuvent apporter au mot construit une charge sémantique identique à celle d'un affixe. En effet, certains ENs peuvent fonctionner comme des affixes construisant des adjectifs relationnels, et peuvent également rentrer dans les mêmes schèmes de nomination que des affixes, étant en distribution complémentaire avec ceux-ci.

4.3.1 Les constructions néoclassiques comme adjectifs relationnels

Si, comme nous venons de le voir, les lexèmes *Xcole* peuvent rentrer dans différents schémas, comme celui reporté en (29-a) dans lequel ils sont des adjectifs représentant le nom *Xculture*, il arrive que l'adjectif *Xcole* représente uniquement le C1, sans que l'EN n'apporte une quelconque information sémantique.

- (29) a. (i) $[X_i \text{ culture}]_{N_j} \leftrightarrow [\text{culture}/\text{élevage de SEM}_i]_j$
 (ii) $[X_i \text{ cole}]_{N_k} \leftrightarrow [\text{relatif à SEM}_j]_k$
 b. $[X_i \text{ cole}]_{N_k} \leftrightarrow [\text{relatif à SEM}_i]_k$

Ainsi, dans l'exemple (30-a), l'adjectif *brassicole* signifie 'relatif à la bière' et ne porte ni le sens de 'cultiver' ni celui d'habiter'. De même, dans l'exemple (30-b), aucun de ces deux sens n'est décelable dans l'adjectif *ostréicole* qui signifie ici 'relatif aux huîtres'.

- (30) a. LIBRAIRIE MOLLAT : sélection de livres **brassicoles** (<https://www.facebook.com/notes/135184543501566/>)
 b. Le Foyer Rural de Bouzigues organise le 21ème Salon de la Gastronomie **Ostréicole**. Pendant 2 jours, démonstration et dégustation de recettes cuisinées avec les Moules et des Huîtres de Bouzigues, stands de produits du terroir, animations variées, restauration sur place. (<http://www.ideesgo.com/agenda/salon-de-la-gastronomie-ostreicole>)

L'EN *-cole* sert donc non seulement à la mise en relation d'un nom recteur (Nr) avec un nom *Xculture* mais également d'un nom recteur avec un

nom base (Nb) X, qui n'a pas forcément le sens de 'cultiver'. Les lexèmes construits ont donc un emploi relationnel : « sémantiquement, ces adjectifs sont relationnels parce qu'ils instaurent une relation entre la sémantique de leur nom base et celle de leur nom recteur » (Fradin, 2008, p.71). Les adjectifs relationnels ont été amplement décrits dans la littérature (entre autres Barning (1976), Monceaux (1997), Roché (2006), Bisetto (2010), Fradin (2008), Strnadová (2014)). Parmi les propriétés traditionnellement données pour les différencier des adjectifs qualificatifs, il y a le fait qu'ils ne peuvent pas être antéposés au Nr, qu'ils ne peuvent pas être en position prédicative ou qu'ils ne sont pas gradables. Récemment, Roché (2006) ou Fradin (2008) ont montré la perméabilité de ces propriétés. Pour les besoins de cette étude qui n'est en aucun cas une étude sur les adjectifs relationnels mais plutôt sur l'emploi relationnel de certaines formations, nous retiendrons que les adjectifs relationnels résultent d'une opération formelle et d'une opération catégorielle sur un Nb, sans qu'aucune opération sémantique n'intervienne. Selon Monceaux (1997, p.40), l'adjectif détermine plutôt que qualifie le Nr. Il s'agit en fait, comme le dit Roché (2006, p.380), d'un « nom adjectivé lexicalement comme il peut l'être syntaxiquement par la préposition *de* ». Ces adjectifs sont en général construits grâce à des suffixes spécialisés pour la formation de relationnels (*-el/-al, -ien, -ique, -aire, -ais/-ois*) ou par des suffixes dont ce n'est pas le rôle principal (*-eux, -ier, -esque, etc.*) (Roché, 2006, p.381). En (30), les adjectifs *Xcole* mettent donc en relation les Nb *bière* et *huître* avec les Nr *livre* et *gastronomie*. Ce sont des « relationnels non marqués », au sens de Roché (2006, p.380) qui cite Maurel (1993, p. 24) : l'adjectif est susceptible de prendre toutes sortes de valeurs en fonction du contexte.

Cet emploi d'un EN sans charge sémantique n'est pas réservé à l'élément *-cole*. L'exemple (31) donne d'autres adjectifs formés avec des éléments néo-classiques finaux : accompagnés de certains noms recteurs, en (a.), les adjectifs contiennent le sens, ou un des sens possibles, de l'EN final : *Xfère* 'qui contient X', *Xphile* 'qui aime X', *Xphone* 'qui parle X', etc. En (b.), avec d'autres noms recteurs, ces mêmes adjectifs ne contiennent plus le sens traditionnellement attribué aux ENs, ils servent uniquement à représenter le nom en position initiale de l'adjectif : *pétrolifère* signifie alors 'relatif au pétrole', *cinéphile* 'relatif au cinéma', etc.

	terrain pétrolifère		dictature pétrolifère
	‘terrain contenant du pétrole’		‘dictature sur le pétrole’
	personne cinéphile		connaissances cinéphiles
	‘personne qui aime le cinéma’		‘connaissances sur le cinéma’
(31)	a. communauté anglophone	b.	consonance anglophone
	‘communauté parlant anglais’		‘qui ressemble à l’anglais’
	personne locavore		consommation locavore
	‘personne qui mange local’		‘consommation locale’
	institut œnologique		contrôle œnologique
	‘institut où l’on étudie le vin’		‘contrôle de la consommation de vin’

Pour Rainer (2013, p.191), l’utilisation des adjectifs *Xfère* comme adjectifs relationnels « semble dû[e] au fait que des locuteurs sans connaissance du latin avaient procédé à une réinterprétation sur la base d’emplois ambigus comme *gisement petrolifère*, interprétable non seulement comme ‘gisement qui contient du pétrole’, mais aussi plus généralement comme ‘gisement de pétrole, qui a quelque chose à voir avec le pétrole’. Une telle réinterprétation a été rendue possible parce que le sens possessif de *-ifère* n’est plus étayé, dans les langues romanes, par un verbe correspondant, *ferre* ayant été remplacé par *porter* et d’autres verbes synonymes ». Si cette explication vaut pour *-fère* ou peut-être pour *-cole*, le sens de *-phile* ou *-logique* semble bien plus disponible pour les locuteurs sans connaissance des langues classiques. Pourquoi cette utilisation des constructions néoclassiques comme adjectifs relationnels alors que d’autres processus morphologiques sont disponibles pour les locuteurs ?

Pour répondre à cette question je vais m’intéresser ici aux lexèmes *Xlogique*¹⁴. Et puisque, comme le montrent les exemples (31), ce sont les Nr qui guident l’interprétation de l’adjectif comme qualifiant ou relationnel, je vais étudier la distribution des adjectifs *Xlogique*. Ces adjectifs ne font pas partie, en tout cas pour l’instant, du fonds de lexèmes présents dans NeoClassy. En effet, les 1042 lexèmes *Xlogique* étudiés ici ont été répertoriés à partir des lexèmes *Xlogue* et *Xlogique* de NeoClassy, ainsi qu’à partir des données des dictionnaires (GR et TLFi). Comme ces lexèmes n’ont pas été annotés comme l’ont été les lexèmes de NeoClassy, je n’ai pas pu encore les faire figurer dans la base.

Dans la section 3.2.1, j’avais dégagé trois schémas principaux construisant des noms *Xlogie* ; je les reproduis ici :

14. Cette étude s’inspire de ce que Fabio Montermini et moi avons proposé lors du 9th Mediterranean Morphology Meeting (Lasserre et Montermini, 2015). Je remercie Marta Cagliani pour l’aide qu’elle nous a apportée.

- (32) a. $[X_i \text{ logie}]_{Nj} \leftrightarrow [\text{discours caractérisé par SEM}_i]_j$
 b. $[X_i \text{ logie}]_{Nj} \leftrightarrow [\text{étude de SEM}_i]_j$
 c. $[X_i \text{ logie}]_{Nj} \leftrightarrow [\text{ensemble de SEM}_i]_j$

Pour le GR, « le suffixe *-logique* sert à former des adjectifs (« relatif à une science ; à une façon de parler ») » ; pour le TLFi, l'élément *-logique* est « formateur d'adj. correspondant à des subst. en *-logie* ». Les deux dictionnaires qui servent ici de référence envisagent donc l'élément *-logique* comme créant uniquement des adjectifs en relation avec des des noms *Xlogie*. Or, au même titre qu'un adjectif *Xcole* pouvait être relié soit au nom *Xculture* soit directement à X, un adjectif *Xlogique* peut ne pas être relié uniquement à un nom *Xlogie* :

- (33) a. recherches météorologiques : 'recherches liées à la météorologie en tant qu'étude du temps'
 b. prévisions météorologiques : 'prévisions liées au temps comme objet d'étude'
 c. tempête météorologique : 'tempête liée au temps' :
 On est tellement habitués aux tempêtes métaphoriques (financières, médiatiques) qu'on est dépourvus face à une tempête **météorologique**. (http://www.liberation.fr/chroniques/2009/01/31/-_306669)

Si en (33-a) et en (33-b), l'adjectif *météorologique* porte le sens de 'étude', il en est dénué en (33-c), exemple dans lequel il sert à relier le Nr *tempête* avec le nom *temps* ; ce dernier, quand il a le sens de 'état atmosphérique' et non de 'durée', n'a pas d'adjectif relationnel construit sur l'un ou l'autre de ses thèmes avec un suffixe comme ceux qui ont été cités plus haut disponible. L'adjectif *météorologique*, par un procédé de recyclage, devient l'adjectif de relation manquant ici.

Par ailleurs, un adjectif *Xlogique* peut aussi être lié à un nom *Xlogie*, comme le montre l'exemple (34) où le *jargon ophtalmologique* est 'le jargon des ophtalmologues'. Mais cette relation est plus marginale et ambiguë puisque ça peut aussi référer au 'jargon utilisé en ophtalmologie'.

- (34) Je sais uniquement que j'étais myope à moins 4.75 a l'oeil gauche avec de l'astigmatie (dont je ne connais pas le degré exact étant donné que je ne comprends pas le jargon **ophtalmologique** et que je ne sais pas lire une ordonnance d'ophtalmologue (<http://www.ophtalmologie.fr/operation-myopie/hopital-public-resultat-decevant-quelle-solution-t3840.html>))

Afin d'évaluer la capacité des adjectifs *Xlogique* à être employés comme des adjectifs mettant en relation deux noms, le Nb (qui est donc le C1 de la formation *Xlogique*) et le Nr de l'adjectif, je me suis concentrée sur les lexèmes non dictionnarisés ; en effet, étant des créations plus spontanées, ils sont le témoin de la vitalité d'un processus de formation ou d'un emploi particulier d'une construction. Pour cela, j'ai recherché la séquence *Xlogique* correspondant à chacun des *Xlogie* et *Xlogue* non dictionnarisés de NéoClassy. En appliquant la méthode inductive présentée en 4.1.1.3, j'ai recherché pour chacun les Nr avec lesquels ils étaient employés dans un échantillon constitué des deux premières pages de résultats Google, au singulier et au pluriel. Au total, 329 adjectifs en *Xlogique* non dictionnarisés ont présenté au moins un Nr dans l'échantillon. L'étude s'appuiera donc sur ces 329 adjectifs. Jusqu'à 35 Nr différents ont pu être collectés pour chacun d'eux ; au total, 2 279 paires NA ont été dégagées, avec 927 Nr différents. En (35) sont listés les Nr les plus fréquents parmi ces 927 accompagnés de leur taux de présence :

	étude	87
	recherche	62
	analyse	42
	approche	38
	science	30
	connaissance	25
	donnée	22
	littérature	20
	aventure	17
(35)	théorie	16
	travail	16
	méthode	15
	problème	15
	aspect	14
	examen	14
	expérience	13
	perspective	13
	savoir	13
	tradition	13
	réflexion	12

Chacun des 927 Nr a ensuite été sémantiquement codé, à l'aide des vdW et vdM, selon sa compatibilité avec les trois sens que peut recouvrir *Xlogie* donnés en (32), à savoir 'discours', 'étude' et 'ensemble'. Certains Nr ne sont compatibles avec aucun de ces sens, d'autres sont génériques, par exemple

plan ou *niveau* dans les expressions *sur le plan Xlogique* ou *au niveau Xlogique*. Le Tableau 4.12 affiche le nombre de Nr selon les sens qui ont été dégagés.

Catégorie sémantique du Nr	Nombre de Nr ayant ce sens	Exemples
‘étude’ : ➤ objet / concept ➤ activité ➤ propriété ➤ résultat ➤ lieu, institution ➤ humain	656	<i>donnée, théorie</i> <i>étude, recherche</i> <i>connaissance, savoir</i> <i>découverte, ouvrage</i> <i>centre, institut</i> <i>expert, confrère</i>
autre sens	183	<i>aventure, moment</i>
➤ dont [+concret]	129	<i>arbre, véhicule</i>
sens générique	63	<i>plan, niveau</i>
‘discours’	16	<i>discussion, débat</i>
‘ensemble’	9	<i>inventaire, patrimoine</i>

Tableau 4.12 – Catégories sémantiques des Nr

Les adjectifs *Xlogique* sont donc employés à 70% avec un Nr recouvrant le sens de ‘étude’ ; plus surprenant, pour 20% d’entre eux, le sens du Nr n’est compatible avec aucun des sens de *Xlogie* canoniques. L’hypothèse ici est que ces Nr n’ayant ni le sens de ‘étude’, ni celui de ‘discours’, ni celui de ‘ensemble’ sont révélateurs d’emplois des *Xlogique* comme adjectifs relationnels où le Nb n’est plus *Xlogie* mais X.

- (36) a. . . . chacun saura rendre grâce à ces infatigables mineurs du savoir **discologique**. (<http://www.pop-hits.net/article-les-promesses-de-l-aube-episode-2012-starring\protect\kernel+.2222em\relax-jerry-mengo-96203143.html>)
- b. Le tout autour et avec les excellentissimes choix **discologiques** de Sebastien Tison, dont la playlist figure dans le meme blog. (<http://patwhite.com/node/4314>)

Ainsi, en (36-a), le Nr est *savoir*, il recouvre le sens de ‘étude’, et le Nb de *discologique* est *discologie* : la paraphrase de l’ensemble NA pourrait être ‘le savoir en discologie’. Inversement, en (36-b), le Nr, *choix* ne comporte pas le sens de ‘étude’ ni aucun des deux autres sens attendus ; le Nb de la

construction est *disque*, et le NA peut être paraphrasé en ‘choix en matière de disques’.

Les sens que peuvent avoir les Nr que nous appellerons non canoniques peuvent être multiples :

- (37)
- a. activité / période de temps
 - aventure bistrologique
 - soirée vinologique
 - b. propriété
 - vitesse escargologique
 - rhétorique injurologique
 - c. objet concret
 - bougie aromacologique
 - kit bobologique
 - d. humain
 - coach tarologique

Dans tous ces exemples, l’élément *-logique* n’apporte aucune contribution sémantique au mot construit ; il est à l’image des suffixes construisant des adjectifs relationnels, un exposant de mise en relation d’un Nb avec un Nr. Comment cette perte sémantique a-t-elle été possible ?

Comme le montrait la liste en (35), le Nr le plus fréquemment rencontré est *étude*. Or déjà, l’emploi d’un adjectif *Xlogique* avec cet Nr offre une information sémantique redondante. En effet, dans *étude architecturologique* par exemple, le sens de ‘étude’ est porté et par le Nr, et par le Nb *architecturologie*. Dans le contexte (38), les NA *étude architecturologique* peuvent être paraphrasés en ‘étude de l’architecture’ et non pas ‘étude en architecturologie’. La présence du Nr qui a pourtant le sens canonique ‘étude’ n’empêche pas l’emploi de l’adjectif *Xlogique* comme relationnel représentant X et non *Xlogie*.

- (38) Une étude **architecturologique** de l’Institut du Monde Arabe montre que Jean Nouvel et ses collaborateurs ont tiré de la même « échelle de modèle » plusieurs mesures distinctes (<https://hal.archives-ouvertes.fr/halshs-00006245/document>)

L’analyse est la même concernant l’exemple (39) : il y a une redondance entre le sens du Nr *inventaire* et la construction *Xlogique* ‘ensemble de X’ ; ici, il s’agit de l’‘inventaire des pommes’. L’EN *-logique* n’apporte aucune information sémantique supplémentaire.

- (39) En effectuant l’inventaire **pomologique** de la petite commune de

Trémereuc (22), à quelques kilomètres au nord de Dinan, nous avons retrouvé une quatrième variété : Sonne Sonne. (<http://www.mordusdelapomme.fr/spip.php?article447>)

Inversement, un *Xlogique* utilisé avec un Nr ayant un sens canonique peut être employé de manière relationnelle ; c'est le cas de (40), où *déclinologique* ne contient pas le sens de 'discours' mais signifie uniquement ici 'à propos du déclin'.

- (40) La souffrance sociale, la souffrance au travail est le support d'un discours **déclinologique** qui estime que « la vraie société c'était avant » (<http://www.metiseurope.eu/alain-ehrenberg-pour-une-politique-de-l-autonomie-en-70-art-29022.html>)

De plus, certains lexèmes *Xlogique* sont fortement attachés à un Nr en particulier. C'est le cas des NA présentés dans le Tableau 4.13 qui ont des indices d'information mutuelle très forts dans les deux ressources vdM et vdW.

Co-occurents	IM (vdM)	IM (vdW)
fouilleMOD_archéologique	10.051	8.772
arbreMOD_généalogique	10.425	8.008
signeMOD_astrologique	9.324	9.134
prévisionMOD_météorologique	8.045	9.525

Tableau 4.13 – NA présentant des mesures d'information mutuelle fortes

On assiste à une analogie qui prend pour modèle ces NA dans leur ensemble, comme en témoignent le Tableau 4.14.

NA co-occurents	Nbre recensés dans le corpus	Exemples
PRÉVISIONMOD_Xlogique	20	prévision tarologique
SIGNEMOD_Xlogique	19	signe arbrologique
FOUILLEMOD_Xlogique	11	fouille blogologique
ARBREMOD_Xlogique	6	arbre linguistologique

Tableau 4.14 – NA construits par analogie avec des NA co-occurents fréquents

L'utilisation de l'élément *-logique* dans la construction de certains adjectifs de relation trouverait ici une de ses explications. Les processus d'analogie

au niveau de co-occurents fréquents ne peuvent pourtant pas forcément expliquer tous les emplois qui ont été recensés.

Une autre explication que l'on peut trouver à l'emploi de l'EN *-logique* pour la formation d'adjectifs relationnels est qu'il permet un allongement des dérivés. En effet, le français préfère des bases dissyllabiques et des dérivés trisyllabiques (Plénat, 2009). Or les exemples (41) sont tous formés sur des bases unisyllabiques (*bière*, *rhum*, *blog*, *rock*) ; les suffixes traditionnels permettant la construction d'adjectifs relationnels faisant en général une syllabe, les dérivés seraient trop courts d'une syllabe par rapport à l'idéal prosodique. L'élément *-log-* pourrait presque être considéré ici comme un interfixe, au sens de Roché (2003), c'est-à-dire un élément sémantiquement neutre, qui apparaît après des bases monosyllabiques et dont la raison d'être est phonologique et non sémantique. En (41-d), l'élément *-log-* permet de plus de satisfaire une contrainte dissimilative en évitant la présence de deux vélaires dans deux syllabes adjacentes, par exemple dans ^o*rockique*.

- (41) a. Bière et Malt : passion « **biérolologique** » (imageck.com/168866072-biere-et-malt-passion-bierologique.html)
- b. Seule information véritablement inédite : cette promesse de Sarkozy à Patrick Ricard de plaider devant Obama la cause du Havana Club, rhum franco-cubain produit par une joint venture Pernod-Ricard-Fidel Castro, dans la sombre bataille **rhumologique** qui l'oppose à la maison Bacardi. (<http://rue89.nouvelobs.com/2015/06/24/espionnage-lart-lenquete-terrain-259924?imprimer=1>)
- c. La deuxième raison est ma récente arrivée chez Heaven, une très belle agence de communication/marketing on-line. Il était donc de bon ton de faire mes débuts "**blogologiques**" suite à cette arrivée. (<http://www.paperblog.fr/users/tichose/>)
- d. Ce collectif à six têtes, entièrement masculin demeure assez peu charismatique en comparaison des Mamas the Papas, et souffrira rapidement d'un manque de crédibilité, voire d'un véritable snobisme **rockologique** au point de se voir assimilé aux pires niaiseries et compromissions de l'époque. (<https://rockernaute.wordpress.com/tag/ted-bluechel-jr/>)

Cependant, cette explication ne vaut que pour les bases monosyllabiques, qui ne sont pas les seules à être construites avec l'élément *-logique* pour construire des adjectifs de relation. En fait, plus que des contraintes phonologiques ou sémantiques, des contraintes lexicales guident cet emploi de l'élément *-logique* ; comme Roché (2009, 2011b, parmi d'autres) le montre,

le dérivé le plus adéquat est celui qui s'insère dans un réseau lexical dense, il doit ressembler aux autres dérivés de la série. Or si l'on regarde les segments se situant à gauche du suffixe *-ique* dans les lexèmes *Xique* du corpus lexic3¹⁵ (Figure 4.6), la séquence /lɔʒ/ est la plus fréquente parmi celles qui précèdent la séquence /ik/ : 7,2% des adjectifs *Xique* du corpus se terminent par la séquence /lɔʒik/ (ensuite, les seconds plus fréquents sont les adjectifs en /matik/ qui représentent 3,8% du corpus). D'un point de vue purement phonologique, la séquence /lɔʒ/ est activée dans la compétence des locuteurs comme étant la plus à même de précéder la séquence /ik/. La séquence /lɔʒik/ est donc à considérer comme une niche morphologique pour la formation d'adjectifs relationnels (pour approfondir la notion de niche morphologique et rencontrer d'autres exemples, voir Lindsay et Aronoff (2015) et Strnadová (2014)).

Il est évident que le choix de cette séquence pour marquer une relation entre deux noms a aussi des explications d'ordre pragmatique, comme je l'ai déjà noté précédemment : l'utilisation d'un élément néoclassique est souvent corrélée à un discours humoristique, pseudo-scientifique. La séquence *-logique* permet d'allonger les lexèmes et de satisfaire la contrainte de crypticité évoquée dans le chapitre précédent.

Ici, l'EN *-logique* perd donc sa charge sémantique pour ne servir plus qu'à mettre deux noms en relation, comme on aurait pu le faire en syntaxe au moyen d'un syntagme prépositionnel. C'est pourquoi je le considère comme étant un EN à part entière, et non pas l'EN *-logie* suffixé au moyen de *-ique*, selon l'analyse que j'ai pu faire précédemment de *-phobique*. Faut-il en déduire qu'un nouveau schéma a émergé, qu'on pourrait formaliser comme en (42) ?

$$(42) \quad [X_i \text{ logique}]_{Aj} \leftrightarrow [\text{qui a un rapport avec SEM}_i]_j$$

Même s'il est possible qu'un tel schéma ait émergé, on assiste surtout à la réponse à la contrainte de recyclage que j'avais citée dans l'introduction de la partie 3.1 ; cette contrainte préfère, plutôt que créer de nouveaux mots, recycler les formes déjà existantes dans un souci d'économie. C'est ce que j'avais évoqué en début de cette analyse pour le lexème *météorologique* : aucun adjectif relationnel lié au lexème TEMPS n'existe ; quand le locuteur cherche à construire un tel adjectif, il va se retourner vers une forme existante et lui conférer une dimension relationnelle.

L'élément *-logique*, perdant sa contribution sémantique, fonctionne de la même façon qu'un suffixe prototypique. Nous allons maintenant voir que cer-

15. www.lexique.org

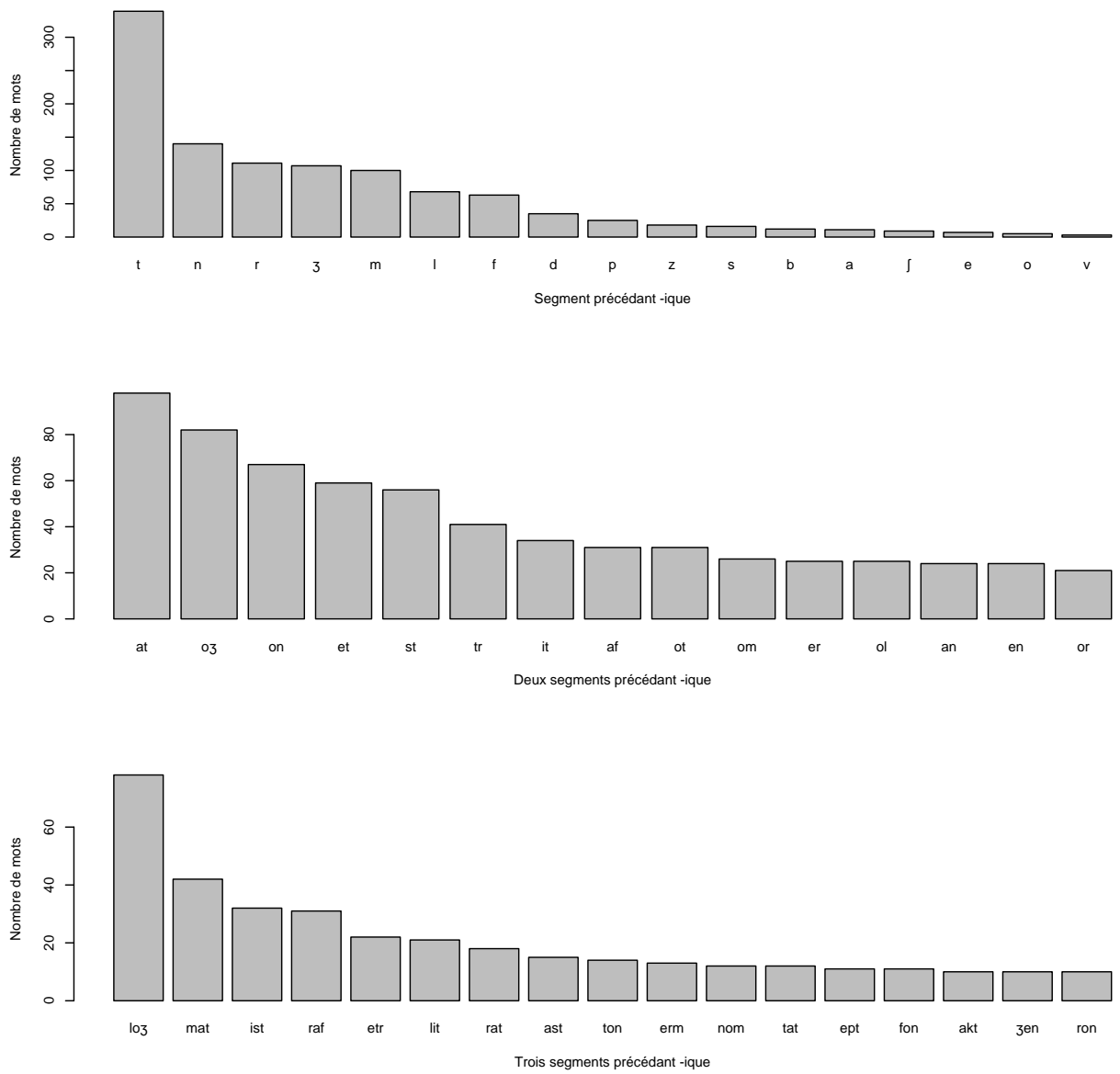


FIGURE 4.6 – Uni-, bi- et trigrammes précédant la séquence /ik/ (source : www.lexique.org)

tains ENs ont le même comportement que certains suffixes, même lorsqu'il ne s'agit pas de construire des adjectifs relationnels.

4.3.2 Les éléments néoclassiques, des suffixes comme les autres ?

4.3.2.1 La construction de noms de spécialistes *Xlogue*

L'EN *-logue* a déjà été traité à plusieurs reprises dans la littérature. Par exemple, Amiot et Dal (2007) se posent la question de son statut. Pour elles, le fait que *-logue* ait le même comportement que *-iste* dans la création des noms de spécialistes donne des arguments pour le considérer comme un suffixe. En fait, le schème de nomination¹⁶ des spécialistes est associé à de multiples procédés morphologiques en français. Ces procédés concurrents sont recensés dans le Tableau 4.15.

Domaine	Spécialiste du domaine
chimie	chimiste
mathématiques	mathématicien
psychologie	psychologue
biologie	biologiste
géographie	géographe
échographie	échographiste
pédiatrie	pédiatre

Tableau 4.15 – Procédés morphologiques associés au schème de nomination des noms de spécialistes

Les procédés concurrents énumérés dans le tableau impliquent soit des suffixes (*-iste*, *-ien*), soit des éléments néoclassiques (*-logue*, *-graphe*, *-iatre*), soit des combinaisons de suffixes et d'éléments néoclassiques (*-logiste*, *-graphiste*). L'élément *-iatre*, qui vient du grec où il signifiait 'médecin' est peu fréquent (voir Tableau 4.16 plus bas), et uniquement employé dans des lexèmes fortement institutionnalisés ; un seul néologisme est recensé dans le corpus FrWac (44), clairement construit par analogie avec *pédiatre*, comme l'atteste l'emploi du /d/ épenthétique qui semble en fait emprunté à ce dernier ; *bébédiatre* pourrait d'ailleurs être analysé comme un mot-valise construit sur *bébé* et *pédiatre*, qui signifierait 'pédiatre pour bébé' (sur les mots-valises, voir la section 5.2.2.2).

16. Sur la notion de schème de nomination, voir Roché (2011b, p.21).

- (43) pédiatre, gériatre, phoniatre, psychiatre
- (44) Voilà , pour le moment . Ce soir je fais ma première sortie dehors pour aller voir ma « **bébédiatre** » ou un truc comme ça , je n' ai pas bien compris ce que m' ont expliqué papa et maman . (<http://www.theotim.fr/>)

Villoing et Namer (2012) analysent les lexèmes *Xlogiste* comme étant formés sur des lexèmes *Xlogie* grâce à une suffixation en *-iste*. La même analyse pourrait être faite concernant les lexèmes *Xgraphiste* construits sur les noms *Xgraphie*. Cependant, il est intéressant de remarquer que, contrairement au doublet *Xlogue/Xlogiste* qui sont, du moins dans les données dictionnaires, synonymes ou quasi-synonymes (selon Villoing et Namer (2012)), les doublets *Xgraphie/Xgraphiste*, quand ils sont tous les deux existants, ne désignent pas les mêmes référents, puisque dans ce cas-là *Xgraphie* peut désigner un instrument ou un résultat alors que *Xgraphiste* désigne le spécialiste de *Xgraphie* :

- (45) épigraphe 'inscription placée sur un édifice' *vs.* épigraphiste 'spécialiste de l'épigraphe'
 échographe 'appareil qui permet, au moyen d'ultra-sons, l'enregistrement des échos des milieux réfringents du corps humain' *vs.* échographiste 'praticien qui se spécialise dans la technique de l'échographie'

De plus, les lexèmes *Xgraphie* semblent nécessiter pour leur construction un lexème *Xgraphie* correspondant alors que les lexèmes *Xlogue*, comme nous l'avons déjà vu et comme Villoing et Namer (2012), peuvent exister sans *Xlogie* correspondant.

Les procédés de constructions de noms de spécialistes concurrents peuvent être coordonnés au sein d'un même contexte :

- (46) a. Les scientifiques présents ont rappelé que les avancées des connaissances passent avant tout par des approches pluridisciplinaires associant des chercheurs de diverses disciplines : biologistes, chimistes, épidémiologistes, expologues, généticiens, médecins, toxicologues, etc. (<http://www.anses.fr/index.htm>)
- b. Enfin après les avoir épaté, j'ai pris mon gavage et attendu la visite des spécialistes (respirologue, nutritionniste, pédiatre, neurologue... qui eux aussi sont restés sans mots devant mes prouesses.). (<http://audreyannedeslandes.wordpress.com/category/uncategorized/page/20/>)

De même, des lexèmes *Xlogue* et *Xiste* peuvent être formés sur les mêmes bases et être donc, a priori, de parfaits synonymes :

- (47) a. Il organise la formation des **archivistes** et **archivologues** depuis 1977. (http://www.lepotentiel.com/afficher_article.php?id_article=124667&id_edition=100161)
 b. Un **urgentologue** ou **urgentiste** est un médecin spécialisé dans la médecine d'urgence (ou urgentologie). La médecine de catastrophe peut aussi faire partie de cette spécialité. (<http://www.ch-bailleul.fr/?Urgentologie.html>)

Des procédés morphologiques concurrents sont donc en compétition pour la dénomination d'un même type d'entité, les noms de spécialistes. Qu'est-ce qui dicte, dans l'exemple (48), l'utilisation du suffixe *-iste* avec la base *latin* et celle de l'EN *-logue* pour la base *grec* alors que le contexte est le même, et les deux bases disponibles pour des constructions néoclassiques pour ce même locuteur, comme en témoigne la présence de *grécophone* et *latinophone* dans la même phrase ?

- (48) Il est anormal d'avoir des Départements qui forment des **Grécologues** et des **Latinistes** et de briller par l'absence des publications, voire de simples thèses sur les sources grecques ou grécophones et latines ou latinophones de la Pensée Africaine. (<http://www.grioo.com/forum/viewtopic.php?t=11810>)

La compétition entre les suffixes *-iste* et *-ien* a été traitée par Lignon (2008) qui a mis au jour des contraintes morphophonologiques expliquant l'existence de deux formes compétitrices. Qu'en est-il pour *Xlogue* et *Xiste* ? Si je ne dispose pas de données suffisantes, surtout concernant les mots *Xiste*, pour faire une analyse comparée détaillée de ces deux processus morphologiques, plusieurs remarques peuvent toutefois être formulées.

Tout d'abord, des contraintes phonologiques sont certainement en jeu, par exemple en (48), l'EN *-logue* permet de former un dérivé trisyllabique sur une base monosyllabique, *grec*, là où le suffixe *-iste* aurait formé un lexème dissyllabique, *?gréquiste* / *?gréciste*¹⁷. La base *latin*, dissyllabique, ne pose pas ce genre de problème de taille. L'utilisation de deux procédés morphologiques distincts permet ici de coordonner deux mots de même taille.

Ensuite, les suffixes *-iste* et *-ien* ne sont pas uniquement utilisés pour la création de noms de spécialistes. Roché (2011a) et Lignon (2000), par exemple,

17. *Helléniste* est bien-sûr possible mais dans ce cas-là, la contrainte de transparence du C1 serait violée.

décrivent les différents schémas impliquant respectivement les suffixes *-iste* et *-ien*. Ainsi, quand la base est anthroponymique (49), les suffixations en *-iste* et *-ien* peuvent dénoter des noms de spécialistes ou des noms de partisans, là où l'emploi de *-logue* lève toute ambiguïté et ne peut désigner qu'un spécialiste.

(49) chiraquiste, chiraquien, chiracologue

Également, des contraintes pragmatiques sont sans doute en jeu. Villoing et Namer (2012) ont par exemple montré que les lexèmes *Xlogue* non lexicalisés peuvent créer des noms de spécialistes dans des domaines non scientifiques voire non conventionnels. Ce procédé morphologique est en effet souvent usité dans des créations humoristiques ; cela montre ainsi sa prégnance dans l'esprit des locuteurs pour la construction de noms de spécialistes auxquels sont conférés un caractère scientifique.

Le Tableau 4.16 recense la fréquence de chaque finale recensée dans le Tableau 4.15 dans les corpus FrWac et Lexique3 ainsi que dans le dictionnaire GR. Le fort nombre de résultats concernant les suffixes *-iste* et *-ien* s'explique par le fait que les résultats n'ont pas été nettoyés, et qu'ils englobent également des lexèmes qui présentent des finales homophones mais non suffixoïdes (*schiste*, *rien*) et des lexèmes construits certes avec ce suffixe mais ne dénotant pas des noms de spécialistes (*communiste*, *galérien*).

Procédé	FrWac	Lexique3	GR
<i>-iste</i>	12288	678	1439
<i>-ien</i>	16040	494	1003
<i>-logue</i>	1681	82	198
<i>-graphe</i>	947	51	173
<i>-logiste</i>	381	34	115
<i>-iatre</i>	125	4	17
<i>-graphiste</i>	108	4	11

Tableau 4.16 – Fréquence des différentes finales construisant des noms de spécialistes dans différents corpus et dictionnaire

Ce tableau montre toutefois, en dehors des suffixes canoniques *-iste* et *-ien*, que l'élément *-logue* est privilégié pour la construction de noms de spécialistes. Un effet de série fait émerger des contraintes paradigmatiques : un nom de spécialiste idéal devrait se terminer en *-logue*. C'est ainsi que l'on trouve sur la Toile des constructions dans lesquelles *-logue* ne sert qu'à inté-

grer un mot déjà construit dans la série dérivationnelle. J'ai ainsi rencontré des lexèmes présentant et l'EN *-graphie*, et l'EN *-logue* (50) : si, comme je l'indiquais en (45), le lexème *échographologue* est plausible puisque *échographe* signifie l'instrument servant à l'échographie, le lexème *géographologue* est plus surprenant car, si on considère qu'il est construit sur *géographe*, il présente une double exponence : littéralement, le lexème signifierait 'spécialiste des spécialistes en géographie'.

- (50) a. Les prospections géophysiques ont pour but de trouver une trace de présence humaine sur un site. Cela peut se faire avec l'aide d'un **géographologue** qui va repérer les différences de reliefs, de couleurs des arbres... (<http://lemercier-prehistoire.blogspot.fr/2006/06/rions-un-peu-1-grande-encyclopedie-de.html>)
- b. D'ailleurs, le sénologue que j'avais consulté (je te dis pas le succès que j'ai eu dans la salle d'attente) m'a, après la compression mentionnée dans ton billet (mais non, pas des roubignoles!), renvoyé chez un **échographologue** qui m'a rassuré : malgré une gynécomastie évidente, je ne produirais jamais de lait (il n'était néanmoins pas exclu que je doive porter des bonnets B). (<http://berthoise.canalblog.com/archives/2011/11/19/22730997.html>)

C'est un phénomène similaire que j'ai relevé en (51) : un lexème déjà lexicalisé avec la finale *-iatre* est additionné de l'EN *-logue*.

- (51) quand vous et votre Adjointe expliquez que personne ne veut de cette réforme, c'est faire insulte aux associations de parents d'élèves qui se sont toutes prononcées pour ! Et la PEEP est aussi pour ! Voilà, comme ça, c'est dit ! D'accord ? La PEEP et la FCPE ! Au moins, c'est dit ! Et c'est surtout ignorer que tous les experts "**pédiatologues**" et chronobiologistes sont pour ! (http://www.marseille.fr/sitevdm/document?id=21578&id_attribute=48)

De même, des noms de métiers / spécialistes / passionnés déjà existants (*linguiste, dentiste, graphiste, damiste*) ont été augmentés de l'EN *-logue* dans les exemples (52).

- (52) a. Quand un **linguistologue** parle de compétence, ce n'est pas la même chose que performance. (https://fr.wikipedia.org/wiki/Wikipédia:Oracle/semaine_6_2006)
- b. Bon courage pour le **dentistologue** demain, je penserai à toi

ma titi et aussi pour le dermato . Moi normalement j'y vais fin février pour me faire retirer un quist sur le tibia (http://forum.doctissimo.fr/nutrition/surpoids-obesite/laetinou-binomettes-combattantes-sujet_169614_55.htm)

- c. Vous cherchez une boîte de type Freelance pour vous accompagner à concevoir et dessiner des maquettes de flyers ou d'autres illustrations digitales (print ou web) à Orleans ? Hervé ridet **graphistologue** (Freelance à Orleans), met en avant ses prestations pour des dessins papiers de toute sorte et des prestations de graphisme (<http://www.flyer-flyers.fr/professionnel/herve-ridet-graphistologue/>)
- d. Et le combat cessa , faute de combattants ! comme disait un champion du XIe siècle . Dieu merci , un **damistologue** n'est jamais pris au dépourvu : - Si les Blancs ont gagné , affirma -t-il , c' est parce qu' ils ont... (http://damierlyonnais.free.fr/joueurs_post_fantaisie_pion.htm)

Tous ces exemples montrent bien la force d'attraction du modèle *Xlogue* pour la construction de noms de spécialistes dès lors qu'un locuteur souhaite ajouter une dimension plus recherchée au terme, que ce soit de manière ironique ou non. En (52-b), le lexème *dentistologue* semble être créé sur le modèle des noms de médecins en *-logue* ; un autre terme de spécialiste médical, quoiqu'apocopé, est également présent dans le même contexte, *dermato*. En (52-c), l'exemple a été recueilli sur la page personnelle d'une personne proposant ses services : *graphistologue*, encore une fois par l'utilisation d'un élément néoclassique, par sa longueur (voir les contraintes dégagées en 3.1) confère une fonction plus importante au référent. L'élément *-logue*, marqueur des noms de spécialité dans les domaines (pseudo-)scientifiques, peut être également utilisé pour railler ces (pseudo-)spécialistes, comme c'est le cas en (53).

- (53) L'on se souviendra, avec ou sans peine, de ce brillant **juristologue**, qui est aussi **criminologistologue**, **journalistologue**, mandarin national autoproclamé et bien sûr, professeur des universités, voulant partager sa science avec son auditoire, déclara sur les ondes et sur un ton nettement péremptoire : « la constitution a été respectée en partie ». (<https://groups.yahoo.com/neo/groups/grandsdebats/conversations/messages/119098>)

Les différents processus morphologiques qui entrent en jeu dans la création de noms de spécialistes semblent donc être en distribution complémentaire,

tant du point de vue phonologique que pragmatique. Des schémas distincts sont donc disponibles pour les locuteurs, par exemple :

- (54) a. $[X_i \text{ iste}]_{N_j} \leftrightarrow [\text{spécialiste de SEM}_i]_j$
 b. $[X_i \text{ ien}]_{N_j} \leftrightarrow [\text{spécialiste de SEM}_i]_j$
 c. $[X_i \text{ logue}]_{N_j} \leftrightarrow [\text{spécialiste de SEM}_i]_j$

Ainsi, l'élément *-logue* n'a pas un comportement différent des suffixes *-iste* et *-ien*. Le dernier chapitre de cette thèse s'attachera à traiter du statut de cet élément. En effet, les derniers exemples (50), (51) et (52) montrent que l'élément néoclassique n'apporte au mot construit aucune contribution sémantique, répondant uniquement à des contraintes pragmatiques et paradigmatiques. J'en conclus que *-logue* est à considérer comme un exposant de construction, au même titre que n'importe quel suffixe plus prototypique.

Afin d'élargir ces observations à un autre EN, observons maintenant le comportement des lexèmes *Xphobe* comparés aux *anti X*.

4.3.2.2 Les anti-étrangers sont-ils xénophobes ?

Des exemples comme (55), (56) ou (57) montrent le parallèle qu'il existe entre certaines constructions *Xphobe* et *anti X*.

- (55) a. Face à la montée d'un sentiment **anti-étranger** à Singapour, le gouvernement doit limiter l'immigration (<http://www.24h-en-asie.com/2014/02/immigration-singapour-sous-pression.html>)
 b. Le problème de l'exacerbation du sentiment **xénophobe** des populations comme on le voit en A.F est le résultat du maintien d'une partie du monde dans la pauvreté et l'insécurité (http://www.lemonde.fr/afrique/video/2015/04/17/le-climat-xenophobe-affecte-toutes-les-couches-de-l-afrique-du-sud_4618316_3212.html)
- (56) a. La France devra renoncer à ses lois **anti-étranger** et anti-Islam (<http://lesactualitesdudroit.20minutes-blogs.fr/archives/2012/09/24/la-france-devra-renoncer-a-ses-lois-anti-etranger-et-anti-is.html>)
 b. Dès le 17 juillet, sept jours après avoir reçu les pleins pouvoirs, Pétain fait adopter une première loi **xénophobe** excluant tout individu non français d'origine de la Fonction publique et des professions libérales. (<http://www.humanite.fr/node/76553>)

- (57) a. Le journaliste **anti-Bush** risque alors jusqu'à quinze ans de prison pour "agression contre un chef d'Etat étranger en visite officielle". (<http://www.republique-des-lettres.fr/10874-mountazer-zaidi.php>)
- b. Le Président a été interrogé par un journaliste de NBC sur l'affaire. "C'est la première fois que je vois ce magazine", a-t-il déclaré naïvement, confirmant malgré lui l'analyse de Newsweek. "Ne demandez jamais à un type dans une bulle s'il est dans une bulle. Il ne peut répondre, parce qu'il est dans une bulle", persiflait ce matin Maureen Dowd, la chroniqueuse **bushophobe** du New York Times. (http://usa.blogs.liberation.fr/2005/12/cette_semaine_1.html)

Dans ces exemples, *anti étranger* et *xénophobe* d'un côté et *anti-Bush* et *bushophobe* de l'autre fonctionnent comme des synonymes : ils sont interchangeables dans des contextes similaires. Ces dérivés en *anti-* sont interprétés par Hathout (2011) comme des adversatifs exocentriques : « (i) leurs bases sont interprétées soit comme des nuisances, soit comme des causes de nuisances ; (ii) ces dérivés nomment la propriété qui consiste à s'opposer à ces nuisances » (Hathout, 2011, p.296) ; ils ont une valeur axiologique dépréciative. Pour moi, les mots construits au moyen de l'élément *-phobe* ayant comme mot leader *xénophobe* ou *homophobe* ont la même interprétation. Les schémas *Xphobe* et *anti X* sont donc à mettre sur le même plan, lorsque la base des dérivés est un anthroponyme :

- (58) $[X_{[+\text{humain}]_i} \text{phobe}]_{\text{AN}_j} \leftrightarrow [\text{anti } X_{[+\text{humain}]_i}]_{\text{AN}_j} \leftrightarrow [\text{qui est hostile envers} / \text{qui s'oppose à SEM}_i]_j$

Au sein d'un même contexte, les deux formes concurrentes peuvent d'ailleurs être employées, peut-être dans une volonté d'éviter une répétition (59), ou bien des mots issus des deux schémas peuvent être coordonnés (60) :

- (59) La Suisse n'est certainement pas un pays **anti-étrangers** avec l'un des pourcentages les plus élevés au monde d'étrangers par rapport à sa population et à la taille de son territoire. Savoir dire « non » ne signifie pas être **xénophobe** mais peut-être juste réaliste ? (<http://www.tdg.ch/suisse/Le-vote-du-9-fevrier-est-antiétrangers/story/12383247?comments=1>)
- (60) Beaucoup sur le forum sont trop jeune pour avoir connu les années 80 mais moi je n'oublierai jamais l'agitation **xénophobe, anti-jeunes** et **anti-islam** de l'hypocrite Charles Piqué (<http://www.bladi.in>)

fo/threads/dinosaure-moureaux-vire-molenbeek.333937/page-3)

Ainsi, si *anti-* est un affixe, comme l'affirment Montermini (2002) ou Hathout (2011) par exemple, je ne vois pas quelles raisons pourraient empêcher *-phobe* de recevoir la même analyse. Je reviendrai sur la notion d'affixe dans le chapitre suivant ; cependant, nous pouvons d'ores et déjà dire que l'EN *-phobe* ici n'apporte pas de contribution sémantique d'ordre lexical, il ne correspond pas à un lexème dont il serait le représentant ; tout comme *anti-*, il s'agit ici de l'exposant d'une construction qui forme des lexèmes ayant un sens adversatif.

Le Tableau 4.17 affiche certains exemples de formations *Xphobe* et *anti X* ainsi que le nombre de résultats rapportés par Google¹⁸ ; même si ces quelques exemples ne peuvent pas généralisés, on voit tout de même que, excepté pour les mots leader *Xphobe* qui avaient été identifiés dans la section 3.2.2.1, les formations *anti X* sont plus fréquentes. Ainsi, à moins qu'un lexème soit particulièrement saillant pour les locuteurs, la préfixation en *anti-* est en général préférée pour une interprétation adversative.

Le Tableau montre donc que les adjectifs *Xphobe* rentrent en concurrence avec des séquences du type *anti-X-suff*. Comme le montrait Corbin (2001, p.51), un lexème tel que *anti-grippe*_A répond à la contrainte de conformité sémantique qui dit que la forme optimale d'un lexème affixé est la concaténation de tous les « constituants concourant à la construction de son sens » alors que *anti-grippal*_A respecte la contrainte de conformité catégorielle qui stipule que tout adjectif doit avoir une finale suffixiforme. Ici, *anti-islam* respecte la contrainte de conformité sémantique, *anti-islamique* celle de conformité catégorielle et *islamophobe* répond non seulement à cette dernière contrainte mais apporte également un trait [+ savant] au mot construit. Comme nous l'avons déjà vu, au schéma *Xphobe* correspond le schéma *Xphobie* construisant des noms de qualité. De par la variété des finales des mots *anti X*, aucun schéma ne semble privilégié pour former des noms de qualité reliés à ces adjectifs, mis à part peut-être [anti-Xisme]_N, quand un nom *Xisme* est disponible.

a.	[X _i phobe] _{AN_j}	↔	[qui est hostile envers SEM _i] _j
		↓	
(61)	[X _i phobie] _{N_k}	↔	[hostilité envers SEM _i] _k
	b. [anti X _i] _{AN_j}	↔	[qui est hostile envers SEM _i] _j
		↓	
	?	↔	[hostilité envers SEM _i] _k

18. Recherche effectuée le 19 octobre 2015.

anti-étranger(s)	9 850
xénophobe(s)	968 000
anti-arabe(s)	77 100
arabophobe(s)	8 500
anti-musulman(s)	202 600
musulmanophobe(s)	1 998
anti-Islam(s)	288 834
anti-islamique(s)	63 000
islamophobe(s)	1 017 000
anti-homosexuel(s)	18 870
anti-homo(s)	11 070
homophobe(s)	1 051 000
anti-américain(s)	102 200
américanophobe(s)	5 100
anti-Bush(s)	21 625
anti-bushiste(s)	533
anti-bushien(s)	84
bushophobe(s)	191

Tableau 4.17 – Comparaison du nombre de résultats sur Google pour certains lexèmes *anti X* et *Xphobe*

Ainsi, certains adjectifs *anti X*, en particulier ceux formés sémantiquement sur une base nominale, n'ont pas dans la série dérivationnelle de la base de formes pouvant être utilisées pour dénoter des noms de qualité. C'est le cas de *anti étranger*.

- (62) xénophobe ↔ xénophobie
 antiaméricain ↔ antiaméricanisme
 anti étranger ↔ ?

Donc, pour référer à ce nom de qualité, le locuteur peut soit utiliser une périphrase (*racisme anti étranger* en (63-a)), soit utiliser un nom construit grâce à l'EN *-phobie* :

- (63) a. L'antisémitisme est certes un racisme particulier, mais les racismes anti-étranger (xénophobie), anti-arabe, anti-noir, anti-islam (islamophobie), anti-gitan, anti-rom... le sont tout autant. (<http://www.reperes-antiracistes.org/2015/10/motion-sur-les-racismes-du-mrap-lunellois-petite-camargue.html>)
 b. Les années Hollande qui s'annoncent ne seront pas, du moins au vu des engagements publics, celles d'une stigmatisation et d'une xénophobie assumée au plus haut sommet de l'Etat. Et ce n'est pas rien de ne plus avoir des ministres qui se sentent obligés d'avoir un «quart d'heure» anti-étrangers ou anti-musulman chaque semaine. (<http://www.slateafrique.com/86847/avec-hollande-etranger-sarko-gueant>)

Ainsi, aux deux procédés concurrents *Xphobe* et *anti X* peut correspondre un seul procédé de formation de noms de qualité :

$$(64) \quad [X_{[+\text{humain}]_i} \text{ phobe}]_{\text{AN}_j} \leftrightarrow [\text{anti } X_{[+\text{humain}]_i}]_{\text{AN}_j} \leftrightarrow \begin{array}{l} \text{[qui est hostile envers /} \\ \text{qui s'oppose} \\ \text{à SEM}_i]_j \end{array}$$

$$\quad \quad \quad \updownarrow$$

$$\quad \quad \quad [X_i \text{ phobie}]_{\text{NK}} \leftrightarrow [\text{hostilité envers SEM}_i]_k$$

Par ailleurs, Hathout (2011) évoque d'autres dérivés en *anti-* avec une interprétation adversative dont la base n'est pas un anthroponyme, par exemple *anti-limace*. La base est toujours un nuisible, « l'agent d'un événement qui est considéré comme une nuisance » (Hathout, 2011, p.296) et le dérivé dénote une entité qui a la propriété d'éloigner ce nuisible. Ce type de mot construit rappelle les CNs *Xcide* du type *insecticide*. En effet, les deux types

de constructions peuvent être employées dans des contextes similaires :

- (65) a. Si malgré cela de nombreuses limaces sont présentes, il est possible d'utiliser un **antilimace** bio durant les 15 jours qui suivent la levée. (<http://www.bioactualites.ch/fr/cultures/grandes-cultures-bio/colza/colza-nouveaux.html>)
- b. Pourquoi avez-vous tant de réticence à utiliser un **limacide** biologique, en l'occurrence à base de ferramol d'autant plus que les autres méthodes de lutte ont échoué? (<http://forum.jardiner-malin.fr/viewtopic.php?id=10866>)

Cependant, une différence sémantique distingue les deux constructions. Quand un *anti limace* repousse les limaces par des moyens qui peuvent être divers, un *limacide* résout le problème de la nuisance des limaces par leur destruction. Contrairement à ce qui se passait pour *xénophobe* et *anti étranger*, *limacide* et *anti limaces* ne sont pas synonymes :

- (66) a. $X_i \text{ cide}]_{AN_j} \leftrightarrow [(instrument) \text{ qui sert à la destruction de SEM}_i]_j$
- b. $[anti X_i]_{AN_j} \leftrightarrow [(instrument) \text{ qui sert à repousser SEM}_i]_j$

Ces deux schémas ne sont pas pour autant déconnectés dans le lexique : en effet, *limacide* peut être considéré comme un hyponyme de *anti limace*, puisque la destruction des nuisibles fait partie des possibles pour repousser les limaces.

La différence entre les comportements des lexèmes *Xcide* et *Xphobe* par rapport à celui des lexèmes *anti X* se remarque aussi dans la possibilité de dérivation en *anti-*. Hathout (2011, p.304) fait remarquer que les lexèmes *anti X* peuvent avoir des finales variées : le radical sélectionné fait en effet partie de la famille dérivationnelle de la base. Or il est possible de trouver des lexèmes *anti Xcide* qui ont pour base X (67) : un *anti-insecticide* est un produit qui repousse les insectes. En fait, il y a comme une redondance dans la construction des lexèmes, de la même façon que ce que nous avons observé précédemment avec des lexèmes comme *dentistologue*.

- (67) a. Peut-on mettre de l'engrais, un **anti-insecticide** et traiter les maladies des plantes en même temps ou bien doit-on respecter un temps entre chaque application. (<http://jardin.gerbeaud.com/insecticide-maladies.html>)
- b. Si vous possédez un stock de chaussures à désinfecter, nous pouvons proposer d'effectuer un traitement **anti-bactéricide** efficace pour l'entretien des chaussures (<http://www.apcsprl.com/Apcsprl.php?state=Desinfection>)

Cependant, un lexème *anti Xphobe* n'aura pas comme base *X*, mais *Xphobe* ; ainsi, un *anti xénophobe* est contre les *xénophobes* et non pas contre les étrangers. Les constructions *anti X* et *Xphobe* ne peuvent donc pas être cumulées.

Nous voyons donc que la Figure 3.34 du chapitre précédent peut être élargie et contenir également des affixes prototypiques. Dans la Figure 4.7, les schémas impliquant des affixes et ceux impliquant des éléments néoclassiques sont en effet représentés comme participant à des mêmes schèmes de nomination.

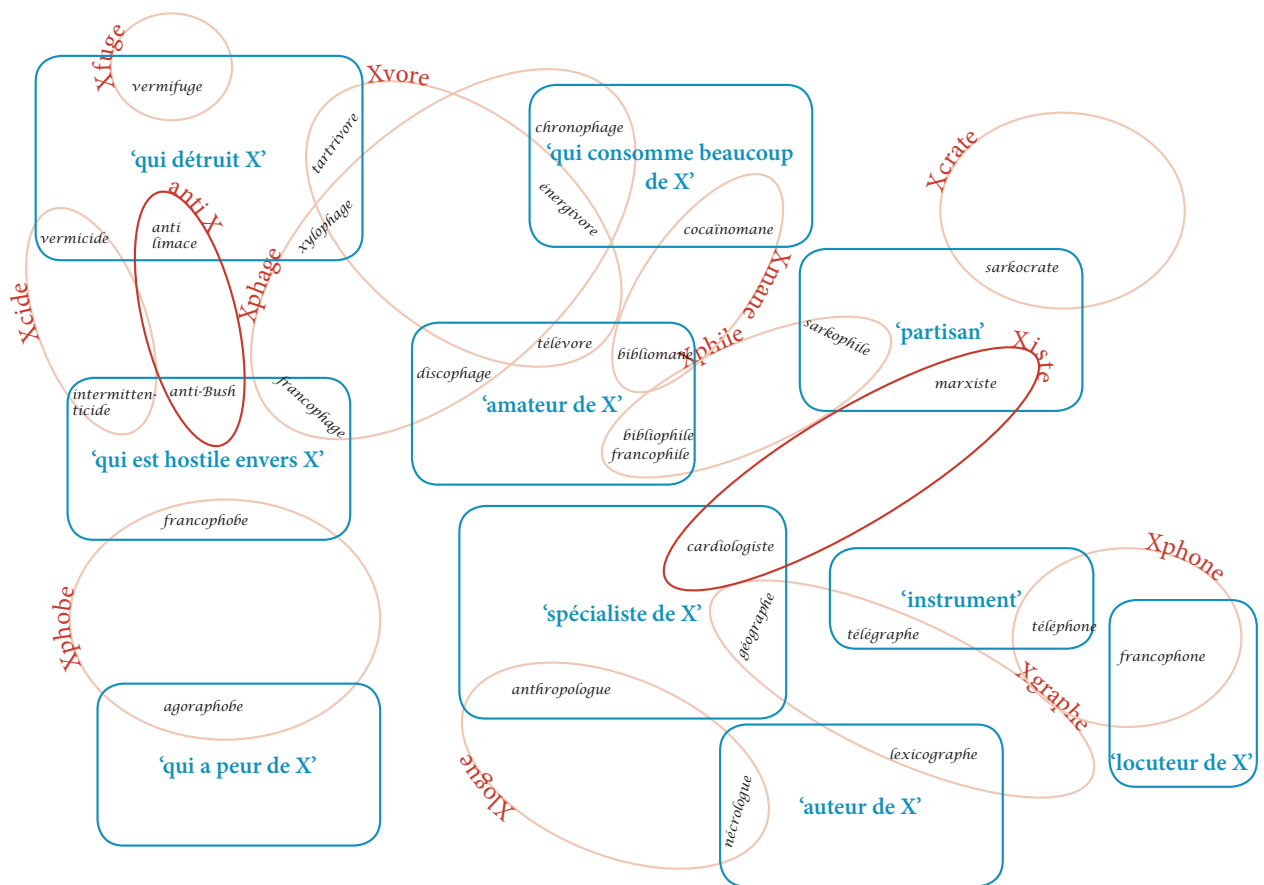


FIGURE 4.7 – Réseau formé par les constructions impliquant certains éléments néoclassiques ainsi que certains affixes

Bilan. Cette partie a offert certaines perspectives d'analyse des constructions impliquant des éléments néoclassiques, en particulier du point de vue

sémantique. Nous avons tout d'abord vu que les constructions néoclassiques ne sont pas différentes d'autres types de constructions : une analyse compositionnelle de leur sens n'est pas possible, puisque la construction de celui-ci dépend de différents facteurs : le type de C1, le contexte dans lequel le mot est employé, les relations entretenus avec les autres mots du lexique. Le contexte d'emploi d'une construction néoclassique peut ainsi participer à l'identification de son sens. Pour le démontrer, j'ai utilisé la technique de l'analyse distributionnelle que l'on retrouve usuellement en sémantique ou en linguistique computationnelle. Ce type d'analyse, rarement utilisé en morphologie, offre des perspectives nouvelles pour l'analyse du sens des mots construits.

Par ailleurs, des réseaux lexicaux complexes ont été mis au jour : des schémas impliquant un même élément néoclassique peuvent entretenir des relations avec des schémas différents dans le lexique, certaines dérivations sont par exemple possibles pour certains schémas mais pas pour d'autres.

Enfin, une comparaison du comportement des éléments néoclassiques dans certaines constructions avec le comportement de suffixes prototypiques a mis en évidence une similarité des schémas de formation ; les éléments néoclassiques étudiés dans cette thèse, donnant lieu à de longues listes de lexèmes et étant impliqués dans des schémas disponibles pour les locuteurs, semblent acquérir le même comportement que certains affixes. Faut-il en déduire que ce sont eux-mêmes des affixes, ou du moins qu'ils sont en voie d'affixation ?

Afin de pouvoir répondre à cette question, les concepts de lexème et d'affixe doivent maintenant être définis, et avec eux les concepts de composition et dérivation. C'est ce que le prochain chapitre va s'attacher à faire. Ensuite, le cadre posé, nous serons plus à même d'étudier le statut des éléments néoclassiques et donc des constructions dans lesquelles ils sont amenés à apparaître.

Un modèle unifié

Sommaire

5.1	Sur la question du lexème	272
5.1.1	Le lexème, une unité multidimensionnelle	272
5.1.1.1	Forme(s) du lexème	277
5.1.1.2	Signification(s) du lexème	286
5.1.1.3	Catégorie lexicale	294
5.1.2	Lexème et affixe : porosité des classes	297
5.2	Les schémas constructionnels	302
5.2.1	La composition, un type de dérivation	302
5.2.2	Perméabilité des frontières	317
5.2.2.1	De l’idiome constructionnel à la grammaticalisation	317
5.2.2.2	Les lexèmes tronqués	330
5.2.2.3	Aux frontières des éléments néoclassiques	341
5.3	Quelle place pour les éléments néoclassiques ? .	343
5.3.1	Critères de classification	344
5.3.1.1	Les éléments finaux	344
5.3.1.2	Les constructions à élément final spécifié	352
5.3.2	Une typologie des éléments finaux	358

La section 1.3 faisait état d’une très grande variété de traitements des constructions néoclassiques et des éléments impliqués dans ces constructions. Selon certains auteurs, les ENs sont des lexèmes, ou des thèmes de lexèmes ; pour d’autres, ils affichent des caractéristiques affixales qu’il ne faut pas négliger. De l’identification du statut de ces éléments dépend l’identification de

la construction en elle-même : opter pour une analyse des éléments néo-classiques comme lexèmes amène à considérer ces constructions comme des composés. Inversement, si les ENs sont des affixes, le mot sera construit par dérivation. Grâce aux analyses effectuées dans les chapitres précédents, il s'agit maintenant de répondre à ces questions en intégrant les éléments néo-classiques dans un modèle incluant d'autres types de relations. Pour cela, les notions de lexème et d'affixe seront définies, dans la partie 5.1. Nous verrons que ces catégories ne présentent pas des frontières totalement imperméables et que, par exemple, des lexèmes peuvent entrer dans des schémas de construction avec un rôle comparable à celui d'un affixe. La partie 5.2 proposera ensuite une définition de la composition et de la dérivation ; à l'intérieur du cadre de la Morphologie Constructionnelle (Booij, 2010), ces types de formations lexicales ne sont pas nécessairement opposables. En effet, la composition sera vue comme un type de dérivation. Et c'est à l'intérieur de ce cadre dans lequel les frontières entre lexème et affixe et entre composition et dérivation sont floues et poreuses que je proposerai, en 5.3 une typologie des éléments finaux et des constructions renfermant ces éléments, en m'appuyant sur les analyses faites dans les chapitres précédents pour lister des critères de classification.

5.1 Sur la question du lexème

Comme je l'envisageais dans la section 2.1.2, la construction d'un mot n'est pas une simple concaténation de morceaux ou morphèmes, mais plutôt une opération réalisée sur un mot et qui a pour résultat un autre mot. Cette opération peut impliquer différentes manipulations morphophonologiques dont l'adjonction d'un affixe à un thème. En quoi ces différentes unités, lexème d'un côté, affixe de l'autre, sont-elles différentes ? Desquelles doit-on rapprocher les éléments néoclassiques ?

5.1.1 Le lexème, une unité multidimensionnelle

Depuis le début de ce travail, j'utilise indifféremment les termes de *mot* et de *lexème*. Or le terme de *mot* recouvre différentes réalités alors que le terme de lexème est plus restreint : un lexème est un mot, mais un mot n'est pas forcément un lexème.

Matthews (1974), et après lui de nombreux linguistes (Aronoff, 1976, 1994; Blevins, 2014, pour ne citer qu'eux), distinguent trois façons de faire référence au mot. Tout d'abord le *wordform*, le mot phonologique, est une séquence de sons dépourvue de sens : « the three lexemes that we spell PARE, PAIR, and

PEAR share the one sound form [per]. The notion of a homophone is rooted in the sound form of a word » (Aronoff, 1994, p.9). Ensuite, les *grammatical words* (« mots grammaticalisés » selon la traduction de Kerleroux (2003)) sont les mots tels qu'ils sont employés en syntaxe. Ce sont les représentations syntaxiques des lexèmes, troisième type de mots, qui sont, eux, des unités abstraites. Fradin (2003) les nomme *mots formes*, Amiot et Dal (2008) *mots syntaxiques*. Le lexème n'est donc ni un lemme, ni un mot tel qu'on peut le rencontrer dans une phrase.

Le lexème est un signe, au sens saussurien, c'est-à-dire la combinaison d'une forme et d'un sens qui sont reliés de façon arbitraire (pour Fradin (2003) un substrat et un abstrat). En tant que signe linguistique minimal, c'est la plus petite unité pourvue de signification. Cependant, c'est « une unité hors emploi par le fait même qu'elle est abstraite des variations flexionnelles et des altérations qui sont le lot des unités figurant dans les phrases » (Fradin, 2003, p.102). Ainsi, un lexème n'a pas une forme mais une pluralité de formes : la face signifiante du signe est donc en fait un ensemble de formes appelées thèmes (Roché, 2010; Bonami et Boyé, 2003, ...). La face du signifié est quant à elle pleinement spécifiée, comme le dit Fradin (2003), et n'est donc pas remplie par un sens très général pouvant rendre compte de toutes les acceptions que peut recouvrir un lexème. Enfin, en plus d'une forme et d'un sens, il est communément admis qu'un lexème appartient à une catégorie lexicale majeure, nom, verbe ou adjectif (parmi les adjectifs sont souvent compris les adverbes) et qu'il a des propriétés morpho-syntaxiques (par exemple verbe transitif ou nom massif).

Un lexème regroupe donc ces trois types d'informations que Booij (2010) nomme PHON (représentation phonologique), SYN (représentation syntaxique) et SEM (représentation sémantique) ; pour lui, les deux premières relèvent de la forme (Booij, 2010, p.5). Pour Fradin (2003, p.102 et suiv.), un lexème regroupe cinq rubriques. Concernant le substrat du lexème, sa face formelle, la rubrique graphémique (G) contient la forme de citation du lexème (notée par un dièse en exposant à la fin du mot), la rubrique phonologique (F) regroupe les différents thèmes du lexème, les formes attachées au lexème sur lesquelles les opérations morphologiques interviennent (voir la section suivant 5.1.1.1). Du côté de l'abstrat, Fradin dégage trois rubriques : la rubrique SYN de Booij est ainsi divisée en deux ici, la rubrique syntactique (SX) qui apporte les informations sur l'insertion du lexème dans une phrase, et la rubrique morphologique (M) recense des informations idiosyncrasiques (déclinaisons, voyelles thématiques,...) et concerne donc le syntactique interne. Enfin, une rubrique sémantique (S) donne la représentation sémantique du lexème. (1) donne la représentation du lexème DORMIR, selon Fradin (2003, p.105).

- (1) (G) dormir #
 (F) (dɔrm)
 (SX) cat :v, str-arg<N0>
 (M)
 (S) (dormir' · ev · x)

Pour ce travail, je m'en tiendrai aux trois rubriques principales proposées par Booij (et que l'on retrouve dans une majorité de travaux) : la rubrique phonologique qui regroupe donc les thèmes du lexème, la rubrique sémantique et la rubrique syntaxique, pour laquelle je m'intéresserai surtout à la catégorie lexicale des lexèmes. Je reviendrai sur chacune de ces trois rubriques de lexème dans les prochaines sous-sections.

Parmi les nombreux aspects du lexème décrits dans la littérature, deux en particulier retiennent mon attention. Pour le premier, regardons cette définition du lexème d'Aronoff :

a lexeme is a (potential or actual) member of a major lexical category, having both form and meaning but being neither, and existing outside of any particular syntactic context. In actual use, though, any instance of a lexeme appears in a sentence, a grammatical and pragmatic context. It is here that the notion of a grammatical word comes into play : a *grammatical word* is a lexeme in a particular syntactic context, where it will be provided with morphosyntactic features (like case and number) and with the morphophonological realization of these morphosyntactic features as bound forms. Grammatical words are the members of the paradigm of a particular lexeme. (Aronoff, 1994, p.11)

Pour lui, un lexème n'a pas besoin d'être listé dans le lexique pour être considéré comme tel. Cela rajoute d'ailleurs à son caractère abstrait : « the lexeme is also abstract in that it is not necessary to assume for any given lexeme that it be lexical (in the sense of this term that is more or less equivalent in meaning to *arbitrary*) or listed in a lexicon » (p.10). A l'opposé, Gaeta et Ricca (2009) estiment qu'un occasionnalisme n'a pas pour vocation à être listé dans le lexique et recevoir le statut de lexème. Par exemple, les formations [+morphologique] [-lexical] du type *gheddafiano*¹, qui, bien que construites selon des schémas réguliers, ne répondent pas à des besoins onomasiologiques suffisants, sont qualifiées par eux de « non-lexeme-forming » (Gaeta et Ricca, 2009, p.41). Si l'on envisage un lexème comme faisant partie du lexique, et que le lexique comporte tous les lexèmes existants, attestés,

1. Cet exemple de Gaeta et Ricca (2009) est également commenté par Montermini (2010a, p.90).

ainsi que les schémas permettant de construire d'autres lexèmes, il n'est pas utile d'envisager tous les mots qui n'ont pas été construits, et qui ne le seront peut-être jamais, comme des lexèmes. Comme nous venons de le voir, un lexème a un sens pleinement attesté. Parmi les mots non dictionnarisés que j'ai étudiés dans les parties précédentes, on a vu que certains pouvaient avoir été formés selon différents schémas impliquant un même EN, comme l'ont montré les analyses distributionnelles effectuées en 4.1. L'exemple (10) du chapitre précédent reporté ci-dessous affiche une même formation, *souricide*, construite selon des schémas différents quoiqu'incluant le même EN final.

- (10) a. Quand a ma souris blanche je l'ai bien assassinee mais c'était un **souricide** involontaire, monsieur le Juge, je vous le jure. (<http://www.femiboard.com/threads/nouveau-jeu-2v%C3%A9rit%C3%A9s-et-un-mensonge.40924/page-33>)
- b. je n'aime pas trop utiliser de **souricide**, les betes souffrent terriblement... mieux vaut utiliser une tapette. (<http://www.graines-et-plantes.com/index.php?forum=jardin-jardinage&question=campagnols-sous-plancher>)

Ici, il faudrait soit décider que *souricide* ne répond pas à des besoins onomasiologiques suffisants pour être stocké dans le lexique, soit envisager deux lexèmes homonymiques ayant chacun un sens pleinement spécifié (je tenterai de répondre à cette question dans la section 5.1.1.2) :

		Lexème1	Lexème2
(2)	PHON	/surisid/	/surisid/
	SYN	N	N
	SEM	événement	instrument

Montermini (2010a) est moins catégorique que Gaeta et Ricca (2009), et considère comme étant un lexème toute unité [+morphologique], qu'elle soit [+lexical] ou [-lexical]. Pour Rainer (2012), il faut distinguer en fait trois modes d'existence pour les mots : les mots existants, actuels, c'est-à-dire ceux qui sont stockés dans le lexique, les mots potentiels, qui peuvent être amenés à exister car sont construits selon un schéma de formation, et les mots virtuels qui, eux, ne seront jamais stockés dans le lexique, quoique formés de façon régulière, car ils sont bloqués par l'existence d'un autre mot répondant aux mêmes besoins onomasiologiques. C'est d'ailleurs ce qui est sous-entendu dans la dénomination RCL, Règle de Construction de Lexème, employée par de nombreux morphologues, et qui réfère à des règles permettant de construire une multitude de mots, qui ne sont pas tous destinés à

être stockés dans le lexique, mais à qui, par le nom même du procédé, est conféré le nom de lexème. A part pour les mots incontestablement lexicalisés, le sens d'un mot se construit lorsqu'il est employé, comme l'a montré l'analyse distributionnelle des CNs dont l'exemple (10) est extrait. Il paraît donc difficile de postuler un statut de lexème pour chaque formation, même si elle est forgée par des moyens morphologiques réguliers et/ou productifs. Il faudrait dans ce cas imaginer presque autant de lexèmes que de mots construits de façon occasionnelle sans qu'aucun ne soit [+lexical], du moins jusqu'à ce qu'ils s'institutionnalisent avec un sens plein et une forme fixe, comme c'est proposé en (2). Je reviendrai sur ces questions de sens en 5.1.1.2.

Le second aspect problématique relevé dans la littérature est pointé par Amiot et Montermini (2009) :

« tout lexème ne dispose pas nécessairement d'un mot-forme correspondant en syntaxe. Dans la plupart des langues (du moins pour ce qui regarde les langues indo-européennes), il existe des lexèmes complexes qui ne peuvent pas être fléchis mais qui servent en revanche de bases à des lexèmes complexes, dérivés ou composés. Il s'agit le plus souvent d'éléments savants (d'origine latine ou grecque) que la tradition grammaticale appelle aussi "mots non autonomes" ou encore "éléments de composition" » (Amiot et Montermini, 2009, p.127)

Est-ce qu'un lexème sans mot grammaticalisé est un lexème ? Comme le dit Fradin (2003, p.93), le lexème, en tant qu'unité abstraite, hors emploi, ne doit pas être confondu avec sa forme de citation : « il n'y a pas de sens à dire qu'un lexème est un élément libre (alors que, par exemple, les affixes seraient des éléments liés) ». Comme nous le verrons dans la section 5.1.1.1 juste après, la rubrique PHON du lexème regroupe la liste des thèmes du lexème ; si certains correspondent effectivement à la forme de citation du lexème, par exemple /bɔ̃/ pour l'adjectif BON représenté en (3), c'est presque fortuit.

(3)	BON :	PHON	/bɔ̃/ ; /bɔn/
		SYN	A
		SEM	'...'

Ainsi, pour un verbe en français, la forme de citation, qui est son infinitif, ne correspond pas à un thème mais à une forme fléchie, elle ne fait donc pas partie de la représentation du verbe :

(4)	BOIRE :	PHON	/bwa/ ; /bwav/ ; /byv/ ; ...
		SYN	V
		SEM	'...'

Il est donc tout à fait envisageable qu'un lexème, en tant qu'entité abstraite, n'ait pas de forme de citation correspondant à un mot syntaxique autonome mais ait dans la rubrique PHON de sa représentation un ou plusieurs thème(s), qui sont sélectionnés pour la formation d'autres lexèmes. En fait, pour se soustraire à l'énumération des propriétés intrinsèques au lexème, une autre façon de le définir est de s'en tenir à ses propriétés combinatoires, en particulier à celle qui le situe comme base des constructions morphologiques : « nous appelons lexème l'objet qui sert de base aux règles morphologiques de construction (que nous appelons Règles de Construction de Lexème) » (Fradin et Kerleroux, 2009, p.83). Pour Fradin (2003, p.102), il sert également de « catégorie de sortie aux procédés de morphologie constructionnelle ».

Il convient toutefois de se demander si les unités dont parlent Amiot et Montermini (2009) dans la citation ci-dessus sont des thèmes supplétifs de lexèmes ou au contraire des lexèmes à part entière sans autonomie syntaxique. C'est ce à quoi je vais m'attacher à présent.

Les sous-sections suivantes ont pour objectif d'envisager chacune des rubriques du lexème en tant que constituant de mot construit mais également de mot construit dans sa globalité.

5.1.1.1 Forme(s) du lexème

Considérons les C1 des unités suivantes :

- (5) a. mycétologie < mycét- 'champignon'
 b. delphinologie < delphin- 'dauphin'
 c. piscivore < pisc- 'poisson'
 ichtyophage < ichty- 'poisson'

Ce sont des éléments empruntés au grec (5-a), au latin (5-b), et même des éléments synonymes empruntés au grec et au latin (5-c). Comment doit-on les considérer ? Deux traitements en particulier peuvent en être fait : (i) ce sont des lexèmes dépourvus de mots syntaxiques, il peuvent être représentés comme en (6)² :

		MYCÉT-	DELPHIN-	ICHT(Y)-	PISC-
(6)	PHON	/miset/	/dɛlfin/	/ikt(j)/	/pis/
	SYN	?	?	?	?
	SEM	'organisme ...'	'animal ...'	'animal ...'	'animal ...'

2. Il est difficile d'attribuer une catégorie lexicale à des éléments non autonomes. La rubrique SYN sera pour l'instant laissée vide dans ces conditions. Je reviendrai sur l'aspect catégoriel en 5.1.1.3.

(ii) ce sont des thèmes des lexèmes CHAMPIGNON, DAUPHIN et POISSON dont la représentation est donnée en (7) :

		CHAMPIGNON	DAUPHIN	POISSON
		/ʃãpiɲõ/	/dofẽ/	/pwasõ/
		/ʃãpiɲõn/	/dofin/	/pwasõn/
(7)	PHON	/miset/	/dɛlfin/	/ikt(j)/
		/pis/
				...
	SYN	N	N	N
	SEM	‘organisme ...’	‘animal ...’	‘animal ...’

Amiot et Dal (2007) traitent ainsi *lud-* comme un radical du lexème JEU (ce qu’ici nous appellerions plutôt un thème), en mettant en avant la distribution complémentaire des radicaux /lyd/ et /ʒø/, le premier étant réservé à des positions initiales et le second à des positions finales dans des mots construits. L’exemple (8) montre que, à la différence de JEU et EAU, les thèmes A et B des lexèmes CHAMPIGNON, DAUPHIN et POISSON peuvent être utilisés dans des positions initiales de mots construits.

- (8) champignonnière, champignonniste, ...
 dauphinelle
 poissonnier, poissonnaille, ...

Le critère de distribution ne semble pas effectif ici. Afin d’évaluer lequel des deux traitements est le plus adéquat, il faut d’abord revenir sur la définition de thème et de radical. Pour cela, je m’appuie sur les dires de Roché (2010). La différence fondamentale entre ces deux notions est que « le *thème* regarde du côté du lexème, le *radical* de la forme fléchie » (Roché, 2010, p.99). En effet, les thèmes sont des « formes phonologiques associées à un même lexème » (Bonami et Boyé, 2003, p.104) quand « le radical est la manifestation phonologique du lexème-base » (Bonami *et al.*, 2009, p.104). En fait, le thème correspond à la séquence qui est stockée dans la rubrique phonologique du lexème ; si souvent la formation du radical revient à sélectionner l’un des thèmes du lexème, des modifications peuvent toutefois être réalisées sur ce thème ; Roché (2010) donne ainsi l’exemple d’ infixations qui ne se jouent pas au niveau du thème mais plutôt du radical. En (9), les thèmes de PUIITS sont /pɥi/ (*avant-puits*) et /pɥiz/ (*puisard*) et non pas /pɥizat/.

- (9) PUIITS → avant-puits, puisard, puisatier

On ne peut pas rendre compte de cette infixation en se plaçant uniquement à

l'intérieur de l'espace thématique du lexème (Bonami et Boyé, 2003), puisque « les différents thèmes d'un même lexème ne sont jamais totalement inter-prédictibles, mais les relations phonologiques qui les associent obéissent à certaines régularités. Les variations proprement thématiques, celles qui sont intrinsèques au lexème, sont circonscrites à des limites qui ici sont outrepassées » (Roché, 2010, p.100). De plus, « phonologiquement, les variations entre thème A et thème B portent uniquement sur la dernière rime et sont bien connues puisqu'elles correspondent à la variation en genre depuis longtemps étudiée » (p.102) et « les types d'alternances possibles sont en nombre restreint » (p.103). Pour reprendre l'exemple *delphinologie*, l'alternance en finale de thème /*ẽ*/ ~ /in/ est bien connue (*copain/copine* ; *voisin/voisine* ; *marin/marine*,...). Cependant, l'alternance /o/ ~ /el/ est non seulement bien plus rare, mais apparaît également ici en milieu de séquence et non pas sur la finale. Peut-on alors considérer /dɛlfɪn/ comme un thème de DAUPHIN ? On parle de supplétisme³ lorsque l'alternance thématique n'est ni fréquente, ni systématique, souvent réservée à un lexème en particulier.

« the relationship between any linguistic units A and B which meet the following condition : the semiotic distinction between A and B is regular, while the formal distinction between them is not regular » (Mel'čuk, 1976, p.50)

Différents supplétismes peuvent être distingués : Roché (2010) parle de supplétisme intégré et de supplétisme lexical, Dressler (1985) de suppléance faible ou forte. Roché limite le supplétisme intégré à la dernière syllabe, et fait ainsi référence à des alternances qui font partie du système mais ne sont plus productives, et qui affectent des séries plus ou moins importantes de lexèmes (10-a). Pour Dressler, la suppléance faible n'est pas limitée à la dernière rime du thème, c'est une « alternance de segments uniques, qui ne dépend d'aucune règle » (Dressler, 1985, p.42) (11-a). Pour les deux auteurs, le supplétisme lexical (10-b) ou la suppléance forte (11-b) concernent l'intégralité du lexème.

- (10) a. RIGOUREUX → rigourosité
IVOIRE → ivoréen
b. RIGOUREUX → rigorosité
IVOIRE → éburnéen
- (11) a. REIMS → Rémois
b. LE PUY → Anicien

3. Le terme *suppléance* est également rencontré dans la littérature, par exemple chez Dressler (1985), ou encore le terme *suppletion* dans la littérature anglophone principalement.

Le supplétisme lexical concerne uniquement les emprunts : « chaque forme est spécifique à un lexème donné, donc imprévisible » (Roché, 2010, p.110). Faut-il intégrer ces éléments supplétifs dans l'espace thématique des lexèmes ou alors en rendre compte lors de la formation du radical? Roché a donné différents points de vue. En premier lieu,

« *Equestre* est à *cheval*, *aqueux* est à *eau* ce que *terrestre* et *terreux* sont à *terre*. Lexicalement, fonctionnellement, *aqueux* fait partie du paradigme dérivationnel de *eau*. *Cavalier*, *hippique*, *équestre* font partie de celui de *cheval* au même titre que *chevalin*. Mais on peut difficilement soutenir, si l'on veut garder la notion de thème sa spécificité, que /ak/ est un thème du lexème EAU, que /kaval/, /ek/, /ip/ sont des thèmes du lexème CHEVAL. (...) l'adjectif méronymique correspondant à *eau* est construit sur le radical /ak/ emprunté à une forme supplétive non autonome ». (Roché, 2009)

Cependant, dans Roché (2010, p.111), les formes relevant du supplétisme lexical ont leur place dans l'espace thématique du lexème, tant du point de vue lexical et dérivationnel, mais « une place à part » ; par exemple, contrairement aux autres cases de l'espace thématique, le thème S (selon la dénomination de Roché) peut contenir plusieurs formes. Le Tableau 5.1 est une représentation de lexème, telle qu'elle est envisagée par Roché : du côté du thème S, les formes représentées dans les cases supérieures relèvent du supplétisme intégré, celles des cases inférieures du supplétisme lexical.

Lexème	Thème A		Thème B		Thème S	
IVOIRE	ivwar	<i>ivoire</i>	ivwar	<i>ivoirier</i>	ivɔr	<i>ivoréen</i>
					ebyrn	<i>éburnéen</i>

Tableau 5.1 – Espace thématique de POISSON

Le Tableau 5.2 est une proposition de représentation de l'espace thématique du lexème POISSON, calquée sur les représentations de lexèmes présentées dans Roché (2010) telles celle du Tableau 5.1.

Ce lexème ne présente pas de forme supplétive intégrée, ce qui explique la case vide au-dessus des formes supplétives lexicales. L'intégration des formes supplétives dans l'espace thématique du lexème est appuyée par l'équivalence sémantique de chacune des formations impliquant chacun des thèmes énumérés dans le Tableau 5.2 : un *poissonnier* peut vendre du *poisson-chat* dans sa *poissonnerie*, un *poisson* élevé dans des fermes *piscicoles* et destiné aux *ichthyophages*. Le Tableau 5.3 recense la plupart des formations construites

Lexème	Thème A		Thème B		Thème S	
POISSON	pwasõ	<i>poisson-chat</i>	pwasøn	<i>poissonnier</i>	pis ikt(j)	<i>piscicole</i> <i>ichtyophage</i>

Tableau 5.2 – Espace thématique de POISSON

sur le lexème POISSON présentes dans le GR. Chacune a pour base le lexème POISSON, c'est-à-dire que chacune résulte d'une opération effectuée sur les rubriques PHON, SYN et SEM du lexème, comme le montre l'exemple (12).

		POISSON	POISSONNIER	POISSON-CHAT
(12)		pwasõ		
	PHON	pwasøn	pwasønje	pwasõ ja
		pis		
		ikt		
	SYN	N	N	N
	SEM	[animal aquatique] _i	[vendeur de SEM _i] _j	[sorte de SEM _i] _k
		PISCICOLE	ICHTYOPHAGE	
	PHON	pisikøl	iktjɔfag	
	SYN	A	AN	
	SEM	[relatif à la culture de SEM _i] _l	[qui se nourrit de SEM _i] _m	

Je suis donc en faveur de la deuxième solution énoncée en début de cette section : plutôt que de traiter les éléments *ichty-* ou *pisc-* comme des lexèmes auxquels aucun mot syntaxique ne correspond, je penche pour considérer les séquences /pis/ et /ikt/ comme des thèmes du lexème POISSON, cette solution permettant de mettre en évidence les liens sémantiques évidents existant entre toutes les formations citées. La représentation des lexèmes CHAMPIGNON, DAUPHIN et POISSON adoptée est donc celle en (7).

Il convient maintenant de se demander ce qui prédispose à la sélection d'un thème S au lieu d'un thème A ou B. Roché et Plénat (2014, p.1870) évoquent la contrainte suivante :

CONTRAİNTE DE REGISTRE : *le registre (populaire ou savant) du thème sélectionné doit être identique à celui du mode de formation.*

Thème	Exemples	Définition du GR
A : /pwasõ/	poisson-chat poisson-scie	‘animal aquatique à corps fusiforme, à membres non apparents’ ‘poisson au corps semblable à celui du requin, portant de chaque côté une rangée de dents qui lui donne l’aspect d’une scie.’
B : /pwasõn/	poissonnaille poissonnerie poissonneux poissonnier poissonnière	‘menu poisson’ ‘commerce du poisson et des produits animaux de la mer et des rivières’ ‘qui contient de nombreux poissons’ ‘personne qui fait le commerce de détail des poissons, coquillages, crustacés, fruits de mer’ ‘ustensile de cuisine, de forme allongée, servant à faire cuire le poisson au four ou au court-bouillon’
S : /pis/	piscivore pisciforme pisciculture	‘qui se nourrit ordinairement de poissons’ ‘qui a la forme d’un poisson’ ‘ensemble des procédés et techniques de production et d’élevage du poisson’
S : /ikt/	ichtyophage ichtyologie ichtyomorphe ichtyosaure ichtyose ichtyotoxine	‘qui se nourrit principalement ou exclusivement de poisson’ ‘partie de la zoologie qui traite des poissons’ ‘en forme de poisson’ ‘grand reptile fossile de l’époque secondaire’ ‘maladie congénitale de la peau (caractérisée par la sécheresse des téguments épaissis, rugueux et couverts de grosses écailles)’ ‘toxine contenue dans le corps de certains poissons’

Tableau 5.3 – Lexèmes ayant pour base POISSON

Selon leur expression, l'utilisation de thèmes supplétifs pour la formation de néoclassiques permet de « faire savant ». Puisque les constructions néoclassiques sont des formations savantes, de par le type de discours dans lesquelles elles étaient à l'origine réservées, l'utilisation de schémas impliquant des éléments néoclassiques induit la sélection de thèmes S.

Si la composition native sélectionne le thème A, la dérivation au moyen des suffixes *-aille*, *-erie*, *-eux* ou *-ier* sélectionne en général le thème B des lexèmes-bases :

- (13) a. blanchaille (< (POISSON) BLANC); chiennaille (< CHIEN); cochonnaille (COCHON); ...
 b. amidonnerie (< AMIDON); croissanterie (< CROISSANT); déchetterie (< DÉCHET); ...
 c. limoneux (< LIMON); goudronneux (< GOUDRON); poussiéreux (< POUSSIÈRE); ...
 d. argentier (< ARGENT); chansonnier (< CHANSON); magasinier (< MAGASIN); ...

Lorsque le thème S est sélectionné dans ces affixations, c'est pour suppléer un thème populaire qui n'apparaît dans aucun lexème construit :

- (14) lumineux (< LUMIÈRE)
 aqueux (< EAU)

Quant aux formations en *-forme*, *-morphe*, *-saure*, *-ose* ou *-toxine*, elles sélectionnent en priorité le thème S, quand celui-ci est disponible :

- (15) a. anthropoforme (< HOMME); arboriforme (< ARBRE); scuti-forme (< ÉCU); ...
 b. pithécomorphe (< SINGE); cynomorphe (< CHIEN); zoomorphe (< ANIMAL); ...
 c. brachiosaure (< BRAS); brontosauure (< TONNERRE); ptérosaure (< AILE); ...
 d. dermatose (< PEAU); cyanose (< BLEU); arthrose (< ARTICULATION); ...
 e. entérotoxine (< INTESTIN); phytotoxine (< PLANTE); mytilotoxine (< MOULE); ...

Certaines formations sélectionnent un thème populaire quand il n'y a pas de thème S qui soit phonologiquement différent du thème populaire (en (16-c), si on suit l'argumentation de Roché (2010), la réduction agit sur un des thèmes pour construire un nouveau radical) :

- (16) a. acnéiforme (< ACNÉ) ; gélatiniforme (< GÉLATINE)
 b. salmonellose (< SALMONELLE) ; bitumose (< BITUME)
 c. quinotoxine (< QUININE) ; amattoxine (< AMANITE)

On observe toutefois une gradabilité : on peut rencontrer une construction *Xforme* avec un thème populaire alors que le GR ne recense aucune construction *Xmorphe* qui aurait sélectionné un thème A ou B.

Nous avons vu dans la section 3.1.1 que l'utilisation de ce que j'avais alors nommé « C1 grec », « C1 latin » ou plus généralement « C1 classique », n'était pas majoritaire dans la construction des lexèmes impliquant les ENs finaux sélectionnés (revoir Figure 3.1). En effet, on trouve de nombreux lexèmes dans NeoClassy qui sont construits sur des thèmes A ou B et non sur des thèmes S :

- (17) catalanophone < CATALAN
 astrologie < ASTRE
 collèguicide < COLLÈGUE
 crétinophobie < CRÉTIN

Même si on pourrait objecter que le choix du thème dans ces exemples est dû à une non disponibilité des thèmes S, les doublons affichant d'un côté une construction avec un thème populaire et de l'autre une construction avec un thème savant montre que la disponibilité du thème S n'est pas une condition suffisante pour son utilisation :

- (18) arachnicide *vs.* araignicide < ARAIGNÉE
 ornithophobie *vs.* oiseau-phobie < OISEAU

Selon Roché et Plénat (2014, p.1872), « le thème populaire tend à remplacer le thème savant quand un suffixe passe de la langue des clercs à celle de tous les jours ». Ainsi, lorsqu'un mode de formation savant tend à être démocratisé, la contrainte de registre s'affaiblit et les locuteurs auront davantage tendance à former un lexème sur un thème populaire, plus disponible. Cependant, lorsqu'elle est disponible, les locuteurs vont avoir tendance à privilégier une forme supplétive intégrée, qui rend le lexème base transparent en lui conférant toutefois un « aspect savant ». Le Tableau 5.4 montre ainsi que, bien que deux thèmes supplétifs soient disponibles pour le lexème ODEUR, /ɔdɔʁ/ (*odorat, odorité, odorisation, odorigène,...*) et /ɔlfakt/ (*olfactif, olfactène, olfactomètre...*), celui qui relève d'un supplétisme intégré est préféré à celui relevant d'un supplétisme lexical.

Le choix de ce thème est symptomatique d'un tiraillement chez le locuteur qui n'a pas de connaissance étymologique, et donc pour qui les thèmes sup-

Thème	Lexème construit	Nbre d'occurrences du lexème construit (recherche effectuée le 11/11/15)
A / B	odeuricide	1
S (intégré)	ordoricide	274
S (lexical)	olfacticide	13

Tableau 5.4 – Formations *Xcide* construites sur le lexème ODEUR

plétifs lexicaux sont peu disponibles, entre la volonté de créer un lexème selon un mode de formation savant et la seule possibilité d'utiliser des thèmes populaires ou relevant de supplétismes intégrés.

Si les C1 des lexèmes regroupés dans NeoClassy sont des lexèmes du français, représentés par des thèmes populaires ou supplétifs, qu'en est-il des ENs finaux dont il est question dans cette thèse? Il est très difficile de les considérer comme des thèmes supplétifs de lexèmes. Ainsi, si le sens de 'animal aquatique' est constant dans toutes les formations énumérées dans le Tableau 5.3, nous avons vu à maintes reprises, que le sens des éléments néo-classiques finaux des lexèmes de NeoClassy n'est pas constant (chapitres 3 et 4). Formellement, on pourrait imaginer des représentations de lexèmes telles qu'elles sont affichées dans les Tableaux 5.5 et 5.6.

Lexème	Thème 1		Thème 3		Thème S	
MANGER	mãʒ	<i>mange-disque</i>	mãʒ	<i>mangeaille</i>	faʒ vɔʀ	<i>ichtyophage</i> <i>piscivore</i>
TUER	ty	<i>tue-mouche</i>	ty	<i>tueur</i>	sid	<i>insecticide</i>

Tableau 5.5 – Espaces thématiques hypothétiques de MANGER et TUER

Cependant, si l'on regarde du côté du sens, on se rend compte que les constructions impliquant les éléments néoclassiques finaux présentent un éventail sémantique plus large que celles impliquant les thèmes populaires des lexèmes correspondants. (19) est un exemple de différences qui peuvent exister entre ces types de constructions :

Lexème	Thème A		Thème B		Thème S	
VOIX	vwa	<i>porte-voix</i>	vwa		vɔk fɔn	<i>vocal</i> <i>phoniatre</i> <i>francophone</i>
DISCOURS	diskur		diskur	<i>discourir</i>	lɔg	<i>logopathie</i> <i>psychologue</i>
POUVOIR	puvwar	<i>contre-pouvoir</i>	puvwar		krat	<i>démocrate</i>

Tableau 5.6 – Espaces thématiques hypothétiques de VOIX, DISCOURS et POUVOIR

	PHON	[ty X _i] _j	[X _m sid] _n
	SYN	(A)N	(A)N
(19)	SEM	[qui sert à tuer SEM _i] _j	[auteur du meurtre de SEM _m] _n [meurtre de SEM _m] _n [destruction méthodique de SEM _m] _n [instrument servant à la destruction de SEM _m] _n

Les schémas constructionnels *Xcide* affichent en effet une polysémie régulière, au sens de Booij (2010, p.78) : « a prototypical meaning forms the starting point for deriving other meanings through the semantic extension mechanisms of metaphor and metonymy » (voir section 2.1.3). Si un lexème a un sens pleinement spécifié (Fradin, 2003), comment traiter cette polysémie ?

5.1.1.2 Signification(s) du lexème

Nous partons donc du principe qu'un lexème a un sens spécifique et non pas un sens général abstrait : « lorsque l'entité est polysémique (...) tout dérivé se forme sur une des significations associées à cette entité et non pas sur une valeur sémantique abstraite sous-spécifiée qui serait le plus petit commun dénominateur des significations distinguées » (Fradin et Kerleroux, 2009, p.84) ; « une opération constructionnelle sélectionne une acception, quand la base est polysémique » (Roché, 2009, p.4) Le lexème, en tant qu'input et output des constructions, a donc un sens unique et pleinement spécifié. Comment donc aborder la polysémie dans ce cadre là ? En effet, si un lexème ne peut contenir qu'une seule et même signification dans sa rubrique SEM, comment traiter les cas que nous avons déjà vus précédemment, reportés ci-dessous, pour lesquels des lexèmes sont formés sur la même base, avec un même EN final, mais par des schémas constructionnels différents :

- (20) a. J'ai commis un **pigeonicide** ! Je ne suis pas fière mais les fientes sur le balcon, shit ! It's no cool. (<http://www.trappeursdeprezent.fr/vos-presents-du-17-fevrier/>)
- b. il est vrai que dans les codes des chasses des siècles précédents, et même dans le code qui est encore en vigueur dans la Pologne, l'Allemagne, etc. l'on coupe la tête et la main aux hommes convaincus d'être cervicides, lièvricides, **pigeonicides**, etc. (*L'Histoire si curieuse des mines de Brandes*, par F. Peloux, M.-C. Bailly-Maitre et H. Viallet p.218)
- c. Un **pigeonicide** est organisé par le conseil municipal sous le nom barbare de "dépigeonnisation" ça fait froid dans le dos, non ? (<http://actualites.forum.orange.fr/messages/index/40198/rome-le-pape-francois-et-benoit-xvi-discriminat ion.html>)
- d. C'est une structure pigeonicide à double circonvolution broyeuse... C'est pas énorme ça comme concept ??? (https://www.youtube.com/all_comments?v=fc_E2EeifY0)
- (21) a. Pour réduire leur utilisation je suis obligé d'utiliser du **souricide** et raticide mais je crains pour la santé de mon chien, un chien de montagne des pyrenes. (http://forum.doctissimo.fr/sante/medecine-veterinaire/empoisonement-souricide-raticide-sujet_144864_1.htm)
- b. Petit rectificatif Yoann, tu n'as pas causé un meurtre (c'est à dire donner la mort de façon intentionnelle et dans ton cas il n'y a pas de préméditation sinon cela aurait été un assassinat) mais un **souricide** involontaire ... J'espère que tu dormiras mieux (<http://www.notre-planete.info/forums/discussion.php?id=72249>)
- (22) a. (i) L'humidité et la chaleur favorisent la formation de champignons, alimentation de nombreux insectes **bibliophages** et attirent des bactéries qui acidifient les papiers. (<http://www.archivosgenbriand.com/preservation-maladies.html>)
- (ii) J'adore surtout lire, je suis **bibliophage**. Je m'intéresse aussi à l'altermondialisme. (<http://voyageforum.com/membres/guachito/>)
- b. Il y a un peu plus d'un an, cette « **livrophage** » - c'est son expression - a ouvert la librairie jeunesse Mots et merveilles, rue des Épéers à Saint-Omer. Pour réaliser un rêve. (http://www.lavoixdunord.fr/Locales/Saint_Omer/actualite/Secteur_Sai

nt_Omer/2011/07/12/article_sandrine-lenglet-livrophage-amoureuse-d.shtml)

Les Tableaux 5.7 et 5.8 représentent les schémas qui ont été dégagés dans les chapitres précédents, pour les constructions *Xcide* et *Xphage*. Comme il a été dit dans la section 2.1.2, les schémas sont stockés dans le lexique, tout comme les lexèmes. Leur représentation est donc identique.

	$[X_i\text{CIDE}]_j$	$[X_k\text{CIDE}]_l$
PHON	X(i)sid	X(i)sid
SYN	i : N[+animé],[+individuel] j : N[+procès]	k : N[+animé],[+individuel] l : N[+animé],[+individuel]
SEM	‘meurtre de SEM _i ’	‘auteur du meurtre de SEM _k ’
	$[X_m\text{CIDE}]_n$	$[X_o\text{CIDE}]_p$
PHON	X(i)sid	X(i)sid
SYN	m : N[+collectif] n : N[+procès],[+résultat du procès]	o : N[+générique],[+nuisible] p : AN[+instrument]
SEM	‘destruction méthodique de SEM _m ’	‘qui sert à la destruction de SEM _o ’

Tableau 5.7 – Quatre schémas principaux correspondant aux $[XCIDE]$ stockés dans le lexique

En (20), PIGEONICIDE (20-a) est construit selon le schéma $X\text{CIDE}_j$, PIGEONICIDE (20-c) selon le schéma $X\text{CIDE}_n$ et PIGEONICIDE (20-d) selon le schéma $X\text{CIDE}_p$, tous trois sur le lexème base PIGEON. Les schémas requièrent cependant des types de N-base différents, un N [+individuel] pour l’un, un N[+collectif] pour le deuxième et un N référant à une classe de nuisibles pour le dernier.

Doit-on postuler trois lexèmes PIGEON dans ce cas là ? La proposition (23) établit trois lexèmes, dont le premier représente l’individu, le deuxième une collection de ces individus et le dernier le type représenté comme nuisible.

	$[X_i\text{PHAGE}]_j$	$[X_k\text{PHAGE}]_l$
PHON	X(o)faʒ	X(o)faʒ
SYN	i : N[+concret] j : AN[+animé]	k : N[+abstrait],[+activité] l : AN[+humain]
SEM	‘qui mange SEM _i ’	‘qui est passionné par SEM _k , consommant, utilisant SEM _k souvent de façon pathologique’

Tableau 5.8 – Deux schémas principaux correspondant aux $[X\text{PHAGE}]$ stockés dans le lexique

	PIGEON1	PIGEON2
	piʒõ	piʒõ
PHON	piʒɔn	piʒɔn
	kɔlɔb	kɔlɔb
SYN	N[+comptable]	N[+collectif]
SEM	‘oiseau au bec grêle, aux ailes courtes, aux pattes à quatre doigts, ... (GR)’	‘oiseau qui prospère beaucoup dans les villes’
(23)	<hr/> <hr/>	
	PIGEON3	
	piʒõ	
PHON	piʒɔn	
	kɔlɔb	
SYN	N[+nuisible]	
SEM	‘oiseau nuisible très présent dans les villes’	

Distinguer ces trois acceptions est difficile, puisque toutes trois ont un seul et même référent possible. Cependant, c’est la manière d’envisager ce référent qui change. Selon l’hypothèse de spécificité maximale (Fradin et Kerleroux, 2009), une opération morphologique sélectionne l’information sémantique la plus spécifique dont elle a besoin pour former un output. Les schémas de construction vont donc sélectionner ici une part de l’information sémantique du lexème base⁴, qui a un seul et unique sens spécifique représenté en (24) :

4. Cette information sémantique pourrait être décomposés en ce que Booij et Lieber (2004)

		PIGEON
		piʒõ
	PHON	piʒõn
		kõlõb
(24)	SYN	N
	SEM	‘oiseau au bec grêle, aux ailes courtes, aux pattes à quatre doigts, ... souvent considéré comme nuisible car il prospère beaucoup dans les villes’

Est-on dans le même cas de figure concernant SOURICIDE ? Les deux exemples (21) renvoient-ils au même référent ? Dans l’entrée *souris*, le GR distingue différents types de souris dont la souris au pelage gris, qui fait des ravages dans les habitations, et la souris blanche, qui est utilisée comme animal de compagnie. Ces deux types de souris sont donc des hyponymes du nom générique *souris*. En (21-a), SOURICIDE est synonyme de SOURISGRISICIDE et en (21-b), SOURICIDE aurait pu être SOURISBLANCHICIDE. Contrairement à ce qu’on avait pour *pigeon*, les référents sont ici distincts, et donc les lexèmes bases ne doivent pas être confondus :

		SOURIS1	SOURIS2
		suri	suri
(25)	PHON	suris	suris
	SYN	N	N
	SEM	‘animal nuisible’	‘animal de compagnie’

Ainsi, dans l’exemple (21-a), SOURICIDE est construit sur le lexème SOURIS1 selon le schéma XCIDE_p. Ce schéma nécessite comme X un N référant à un nuisible, ce que remplit tout à fait SOURIS1. En (21-b), SOURICIDE est construit selon le schéma XCIDE_l requérant comme base un N[+animé], [+comptable], rôle que satisfait SOURIS2. Les deux SOURICIDE sont donc non seulement construits selon des schémas différents quoiqu’affichant un EN final similaire, mais également sur deux lexèmes différents quoiqu’étant des co-hyponymes.

Concernant l’exemple (22), même si on a affaire à la même entité, d’un côté le schéma sélectionne le contenant, de l’autre le contenu.

appellent *squelette* (*skeleton*), c’est-à-dire la part de la description sémantique permettant de faire des généralisations inter-catégories, et le *corps* (*body*), qui comprend les informations sémantiques de type encyclopédiques qui ne sont que partiellement formalisables. Je ne ferai pas de distinction aussi détaillée ici.

		LIVRE1	LIVRE2
	PHON	livr bibli	livr bibli
(26)	SYN	N[+concret]	N[+abstrait]
	SEM	‘objet constitué de feuilles de papier reliées’	‘ensemble de mots, de phrases, ayant une signification, racontant une histoire’

Fradin et Kerleroux (2009, p.85) mettaient au jour l'impossibilité pour LIVRE1 de servir de base pour la formation d'un adjectif dénominal :

- (27) Une (culture + éducation) livresque (< LIVRE2)
 *Les industries livresques (< LIVRE1)
 *Une reliure livresque (< LIVRE1)

Doit-on cependant postuler des lexèmes distincts ? La situation est différente de celle, toujours exposée par Fradin et Kerleroux (2009), des lexèmes FILLE :

		FILLE1	FILLE2
	PHON	fij	fij
(28)	SYN	N	N
	SEM	‘descendant immédiat de sexe féminin’	‘enfant de sexe féminin’

Si FILLE2 implique forcément *fille1*, le contraire n'est pas vrai. La dérivation met encore en évidence la spécificité de chacun de ces deux lexèmes :

- (29) fillette < FILLE2

Or concernant LIVRE, le livre (contenant) et le livre (contenu) sont comme les deux faces d'un signe linguistique ou le recto et le verso d'une feuille de papier : l'un implique forcément l'autre. Cependant, si « la morphologie constructionnelle tranche et délimite » (Fradin et Kerleroux, 2009, p.84), il est difficile d'imaginer qu'un schéma requérant un Nbase [+concret] et un schéma requérant un Nbase [+abstrait] aient recours à un seul et unique lexème. Il faut donc accepter de distinguer deux lexèmes LIVRE comme deux réalités du même référent.

On voit donc qu'un lexème ne peut pas avoir un sens sous-spécifié ; le lexème qui sert de base à un schéma de construction doit avoir accès à un sens spécifique et non pas sous-jacent, puisque le schéma en lui-même contient des informations sur le type de base à sélectionner. Ici, trois configurations

différentes ont été abordées : un seul lexème PIGEON, non polysémique, à partir duquel les différents schémas sélectionnent une partie de l'information sémantique mais deux lexèmes LIVRE puisque les opérations constructionnelles nécessitent des Nbases opposés, [+abstrait] d'un côté et [+concret] de l'autre. SOURICIDE a révélé qu'on avait affaire à deux référents distincts, représentés par deux lexèmes co-hyponymes quoique homophones.

Dans tous les cas, la question demeure : doit-on envisager des lexèmes PIGEONICIDE, SOURICIDE et BIBLIOPHAGE polysémiques ou des lexèmes distincts ? La question de la polysémie dépasse le cadre de cette thèse. Nous retiendrons simplement ici que la polysémie se joue au niveau de la construction. Il s'agit de polysémie régulière dont parle Booij (2010) : si on peut faire dériver le sens des constructions de celui d'une construction prototypique, il est toutefois difficile de relier les sens des lexèmes construits grâce à ces schémas. La Figure 5.1 représente cette polysémie régulière : chacun des schémas XCIDE pourrait par exemple être relié au schéma prototypique $[X_i \text{ cide}]_j$ ⁵.

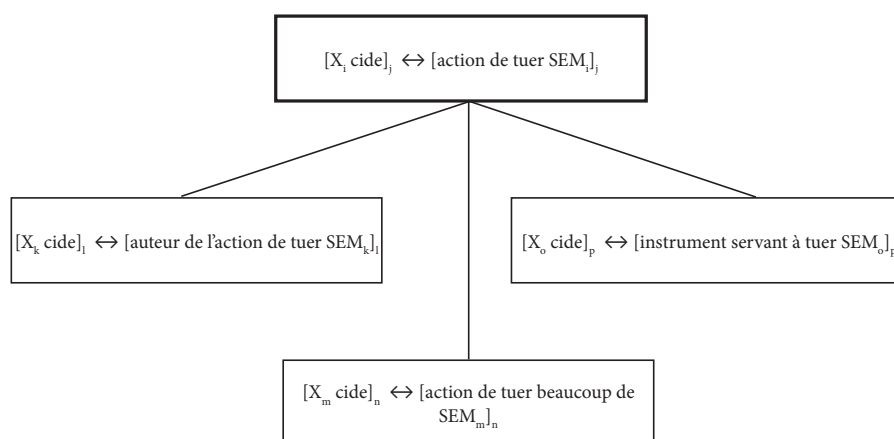


FIGURE 5.1 – Polysémie régulière des constructions XCIDE

Cependant, les différents *pigeonicide*, construits par des schémas différents empruntant à un même lexème base (PIGEON) des informations spécifiques, ne peuvent pas être reliés de la même façon. Il faudrait en fait reprendre la distinction que fait Fradin (2014, p.112-113) pour le verbe ÉLEVER entre objet morphologique et objet lexical : la première proposition désigne le verbe tel qu'il est réalisé dans le discours, au travers de mots-formes ; il est dans

5. J'ai choisi de représenter ici le sens 'action de tuer' comme étant le sens prototypique. Cependant, ce sens prototypique dont découlent les autres sens devra être sélectionné par des analyses sémantiques plus poussées.

ce cas-là unique. Le verbe comme objet lexical est quant à lui le lexème tel qu'il est corrélé à une construction : deux lexèmes sont alors à distinguer, ÉLEVER₁, 'prendre soin du développement d'un enfant, d'un animal' qui est nominalisé une suffixation en *-age* (*élevage*) et ÉLEVER₂, 'placer plus haut', dont la nominalisation est *élévation*. Ici, si PIGEON est bien une seule et même unité morphologique, elle peut être divisée en différents objets lexicaux qui sont autant de bases pour les différents schémas [Xcide]. Pour ces lexèmes ainsi construits, on ne peut plus parler de polysémie mais bien de lexèmes distincts, qui certes partagent la même base et qui sont construits selon des schémas affichant des liens sémantiques forts.

- (30) PIGEONICIDE1 'meurtre d'un pigeon' (< PIGEON ; [X_iCIDE]_j)
 PIGEONICIDE2 'auteur du meurtre d'un pigeon' (< PIGEON ; [X_kCIDE]_l)
 PIGEONICIDE3 'processus d'extermination de tous les pigeons d'une ville' (< PIGEON ; [X_mCIDE]_n)
 PIGEONICIDE4 'instrument servant à la destruction des pigeons comme espèce nuisible' (< PIGEON ; [X_oCIDE]_p)

Cette question de la polysémie est problématique et je suis consciente que la position que j'adopte ici peut poser question. Il faut garder toutefois à l'esprit que *pigeonicide*, peu importe le mot-occurrence en question, n'a pas vocation à être lexicalisé (il s'agit, selon les termes de Corbin (1997), d'une unité lexicale à faible probabilité d'actualisation). En début de cette partie 5.1.1, je m'interrogeais sur la possibilité pour une occurrence d'avoir un statut de lexème sans qu'elle ne soit lexicalisée, stockée dans le lexique. Le lexème, unité abstraite, hors emploi, est pourvu d'une seule et unique signification contenue dans la rubrique SEM de sa représentation : « par "signification", j'entends une interprétation stabilisée en langue, c'est-à-dire reconnue intersubjectivement et qui n'est pas soumise à chaque fois à renégociation » (Fradin, 2003, p.236). Ainsi, deux solutions sont envisageables quant au traitement de l'entité *pigeonicide* : (i) il s'agit de différents lexèmes homonymes, (ii) aucune des occurrences de *pigeonicide* ne peut acquérir le statut de lexème puisque le sens de *pigeonicide* ne peut pas être déterminé de façon unanime par une communauté de locuteurs. Cette dernière proposition, qui sous-entend que le sens de *pigeonicide* est recréé chaque fois qu'il est employé, fait écho aux résultats de l'analyse distributionnelle effectuée dans le chapitre 4 : l'entité non lexicalisée nécessite une distribution analogue à celle de lexèmes lexicalisés pour pouvoir créer un sens comparable à ces derniers. Si la question du sens du constituant de gauche des constructions en jeu dans ce travail ainsi que la question du sens des lexèmes construits ont été traitées, je laisse la question du sens des ENs finaux en suspens pour le moment, avant

de l'aborder dans la partie 5.3.

5.1.1.3 Catégorie lexicale

Différents moyens pour déterminer la catégorie lexicale d'un EN ont été envisagés dans la littérature. Villoing et Namer (2006) en répertorient quatre : (i) si on traduit un EN, ou si on l'interprète, la catégorie lexicale de la traduction donne la catégorie lexicale de l'EN ; (ii) si on considère qu'un EN est un allomorphe d'un lexème du français, la question de la catégorie ne présente plus aucune difficulté ; (iii) si l'EN est utilisé comme base lors d'une opération de dérivation, les contraintes de sélection propres à l'opération sont un indice sur la catégorie de l'EN ; (iv) le type de construction faisant apparaître un EN est un autre indice sur sa catégorie.

Le premier moyen, elles le reconnaissent d'ailleurs, est un peu bancal : en effet, comment déterminer si *-graph-* est la traduction du verbe *écrire* ou du nom *écriture* (voire du nom *scripteur*) ? Le TLFi donne comme définition « Élém. initial tiré du gr. *γράφειν* 'écrire' et formant qq. mots sav. relatifs à l'écriture » et le GR « Premier élément de mots savants, tiré du grec *graphein* 'écrire' ». Ce dernier dictionnaire donne également les définitions suivantes :

- (31) graphothérapie 'rééducation de l'écriture à objectif thérapeutique'
graphomanie 'besoin pathologique d'écrire'

Ainsi, la traduction n'aide pas ici à la catégorisation de l'élément *graph-*. Les trois autres moyens considérés par Villoing et Namer (2006) sont en fait complémentaires. En effet, si on considère un EN comme une des représentations phonologiques d'un lexème, la question de la catégorie ne se pose pas : en effet, le thème supplétif n'a pas de catégorie syntaxique à proprement parler, mais est la représentation d'un lexème qui, lui, est pourvu d'une catégorie lexicale. Cependant, comme disent Fradin et Kerleroux (2009, p.100), « la règle de construction morphologique peut et doit être exploitée comme un test révélateur de l'identité lexémique » ; ainsi, les schémas dans lesquels les ENs sont impliqués peuvent être révélateurs du lexème dont ils sont la représentation. Par exemple, le suffixe *-ique* crée des adjectifs dénominaux ; dans *hydrique* la séquence *hydr-* est un thème du lexème EAU qui est un nom. De même, Lüdeling *et al.* (2002, p.270) rapportent qu'en allemand, la tête des composés est à droite, ce qui signifie que la catégorie de l'élément de droite est identique à la catégorie du composé ; ainsi, comme *-morph* est la tête du composé *anthropomorph_A*, il faut analyser *-morph* comme un adjectif. Pour reprendre les exemples (31), après examination des données du GR⁶, le

6. Il s'agit ici d'une première observation qui demanderait évidemment d'être approfondie, ce

schéma *Xthérapie* selon lequel *graphothérapie* semble avoir été construit sélectionne des Nbases (32), donc *graph-* dans *graphothérapie* peut être analysé comme un nom. Cependant, les schémas qui ont pu être dégagés lors d'une première observation des lexèmes *Xmanie* ne permettent pas de limiter la sélection d'une catégorie unique de lexèmes, puisque X peut être soit un nom, soit un verbe (33). Il n'est donc pas possible ici d'identifier la catégorie de *graph-* dans *graphomanie* au moyen des schémas de construction.

- (32) $[X_{Ni}thérapie]_j \leftrightarrow [thérapie \text{ au moyen de } SEM_i]_j$
 ➤ musicothérapie (< MUSIQUE_N)
 ➤ théâtrothérapie (< THÉÂTRE_N)
- (33) a. $[X_imanie]_j \leftrightarrow [excès, condamnable \text{ ou } ridicule \text{ de } SEM_i]_j$
 ➤ bronzomanie (< BRONZER_V)
 ➤ musicomanie (< MUSIQUE_N)
 b. $[X_kmanie]_l \leftrightarrow [état \text{ pathologique limité à } SEM_k]_l$
 ➤ cocaïnomanie (< COCAÏNE_N)
 ➤ cleptomane (< VOLER_V ? VOLEUR_N ?)

Néanmoins, nous conviendrons que seuls les lexèmes peuvent appartenir à une catégorie lexicale majeure. Or, même si la question a été quelque peu laissée de côté depuis le début de ce chapitre, question sur laquelle je reviendrai longuement dans la dernière section (5.3), tous les ENs ne peuvent pas être analysés comme des lexèmes et ainsi se voir attribuer une catégorie lexicale. Ainsi, un traitement unifié des éléments néoclassiques est impossible puisque certains ENs finaux conservent une interprétation, quand ce n'est pas le cas pour d'autres.

Par exemple, le sens de 'tuer' ou de 'manger' est bel et bien présent respectivement dans les constructions *Xcide* et *Xphage/Xvore* dans lesquelles les rôles thématiques sont identifiables :

- (34) ➤ $[X_iCIDE]_j \leftrightarrow [meurtre \text{ de } SEM_i]_j$:
 i : patient
 j : procès
 ➤ $[X_kCIDE]_l \leftrightarrow [auteur \text{ du meurtre de } SEM_k]_l$:
 k : patient
 l : agent
 ➤ $[X_mCIDE]_n \leftrightarrow [destruction \text{ méthodique de } SEM_m]_n$:
 m : patient
 n : procès

qui pourrait faire l'objet de travaux ultérieurs.

- $[X_oCIDE]_p \leftrightarrow$ [qui sert à la destruction de SEM_o]_p :
 o : patient
 p : instrument
- (35) ➤ $[X_iPHAGE]_j \leftrightarrow$ [qui mange SEM_i]_j :
 i : patient
 j : agent
- $[X_kPHAGE]_l \leftrightarrow$ [qui est passionné par SEM_k]_l
 k : stimulus
 l : expérian

Cependant, dès lors qu'on sort des sens prototypiques, l'interprétation verbale s'estompe : un verbe psychologique a en effet un sens moins prototypiquement verbal qu'un verbe d'action (35). La construction $[X_kPHAGE]_l$ ne présente pas de différences majeures avec des constructions pouvant afficher un suffixe canonique :

- (36) a. $[X_{Ni}ier]_{Nj} \leftrightarrow$ [personne qui réalise SEM_i]_j
 ➤ bijoutier (< bijou)
 ➤ crêpier (< crêpe)
- b. $[X_{Nk}phage]_{Nl} \leftrightarrow$ [personne qui aime SEM_i]_j
 ➤ téléphage (< télé(vision))
 ➤ cinéphage (< ciné(ma))

De même, il paraît abusif d'attribuer une interprétation verbale prototypique aux constructions en (37), alors même que l'EN a pour origine un nom et non un verbe. Cependant, la construction $[X_iPHOBE]_j$ entretient bel et bien des liens avec la construction $[X_{Ni}phage]_{Nj}$, comme nous l'avons vu dans le chapitre 3.

- (37) ➤ $[X_iPHOBE]_j \leftrightarrow$ [qui a peur de SEM_i]_j :
 i : stimulus
 j : expérian
- $[X_kPHOBE]_l \leftrightarrow$ [qui est hostile à SEM_k]_l
 k : stimulus
 l : expérian

Concernant l'élément *-logue*, Amiot et Dal (2007, p.333) affirment : « today, nouns such as PSYCHOLOGUE, DÉCLINOLOGUE or FUTUROLOGUE are not perceived as having a verb in their structure, nor do they have an agentive interpretation ». En effet, on peut faire remonter l'élément *-logue* au nom

grec λόγος ou au verbe λέγω ‘parler’. Et la glose donnée par un même dictionnaire pour deux lexèmes construits avec ce même élément met tantôt en avant une interprétation verbale, tantôt une interprétation nominale :

- (38) graphologue ‘personne qui étudie, qui pratique la graphologie’ (GR)
 addictologue ‘spécialiste de l’addictologie’ (GR)

Un *addicologue* peut en effet être un ‘spécialiste des addictions’, une ‘personne qui étudie les addictions’, une ‘personne qui fait courir des recherches sur les addictions’ ou encore une ‘personne qui parle au sujet des addictions’. De même un *anesthésiste* peut être un ‘spécialiste de l’anesthésie’ ou une ‘personne qui pratique des anesthésies’. Ce parallélisme entre les constructions *Xlogue* et *Xiste* a été examiné dans le chapitre 4.

Pour Amiot et Dal (2007), lorsque l’EN perd son sens verbal, il y a dès lors possibilité de le considérer comme un suffixe. Cependant, est-ce une condition suffisante ? Est-ce que *-cide*, qui conserve son sens verbal dans les constructions dans lesquelles il est employé, doit avoir un traitement différent de celui de *-logue* ? Afin de répondre à cela, nous avons besoin de définir la classe des affixes.

5.1.2 Lexème et affixe : porosité des classes

Après avoir défini le lexème comme une unité multidimensionnelle servant de base aux opérations de construction de lexèmes, je vais maintenant m’attacher à définir l’affixe qui est l’exposant de ces opérations. Si cette présentation propose une observation des affixes en général, elle sera tout de même fortement orientée vers les suffixes : leur définition me permettra par la suite de les mettre en perspective avec les ENs finaux qui m’occupent dans ce travail. Aronoff et Fudeman (2011, p.258) définissent l’affixe comme « a bound morpheme that attaches to a root or stem to form a new lexeme (derived form) or an inflected form or stem of an existing lexeme ». Je ne m’intéresse ici qu’aux affixes dérivationnels.

Un affixe est en général défini selon, d’un côté, des propriétés intrinsèques, par exemple sa forme ou le sens des constructions dans lesquelles il intervient, de l’autre des propriétés extrinsèques, des critères combinatoires par exemple. C’est ce qui est en quelque sorte sous-entendu par Scalise *et al.* (2005, p.137) : « derivational suffixes have in their representation two different sets of linguistic information : (a) Information (consisting of a skeleton and a body) that ‘belongs’ to the suffix itself and percolates to the upper node in the course of a derivation (this kind of information is not relevant for the mechanism of selection) and (b) Information that specifies the properties that a word

must have in order to qualify as a base for a grammatical derivation (this information is relevant for selection) ». Pour eux, les propriétés intrinsèques sont surtout liées au sens des affixes ; ils reprennent ainsi la représentation sémantique proposée par Lieber (2003) sur laquelle je ne reviendrai pas ici. Cependant, certains auteurs se sont attachés à en donner des caractéristiques formelles. Par exemple, pour Montermini (2002, p.50), un affixe dérivationnel prototypique « a une forme phonologique monosyllabique ou en tout cas pas plus que dissyllabique ». Pour Roché et Plénat (2014, p.1874), en français, « la plupart des suffixes, et les désinences verbales, commencent par une voyelle et constituent la rime de la dernière syllabe du dérivé ». Le fait de commencer par une voyelle permet une fonte de l’affixe dans le lexème construit, grâce à une resyllabification, les thèmes des lexèmes sur lesquels s’attachent les suffixes se terminant en général par une consonne (Roché, 2010). L’affixe dérivationnel est donc contraint dans sa forme. Qu’en est-il de son sens ?

Le sens est le biais que la plupart des analyses ayant pour but de différencier affixes et lexèmes empruntent. Pour revenir à la célèbre distinction avancée par Danielle Corbin (dans Corbin (2001) par exemple), les lexèmes ont un sens lexical alors que les affixes ont un sens instructionnel, c’est-à-dire qu’ils n’ont pas la capacité de référer à un objet du monde au sens large : ils « sont porteurs d’une instruction sémantique leur permettant, en combinaison avec tout ou partie des propriétés sémantiques de leur base, de donner à voir d’une certaine façon le référent désigné par le mot construit ». Kastovsky (2009) propose quant à lui une distinction entre les lexèmes qui seraient « lexically specific » et les affixes « lexically non-specific », c’est-à-dire porteurs d’un sens général, abstrait. Cependant, Aronoff (1984, p.48) ou Lehrer (2000, p.149) font remarquer que certains affixes ont un sens qui est le même que celui d’un membre d’une catégorie lexicale majeure. Aronoff donne ainsi les exemples que je reporte en (39) sur l’anglais, et les exemples reportés en (40) qui concernent le tagalog et pour lesquels le préfixe *mag-* accompagné d’une reduplication de l’initiale CV de la base a le sens de ‘vendeur du produit désigné par la base’.

	<u>préfixe</u>	<u>adverbe</u>	<u>exemples</u>
	re-	again	redo, reinvestigate
	mis-	incorrectly	misdiagnose, misconstrue
(39)	over-	excessively	overreach, overcorrect
	under-	insufficiently	under react, undercorrect
	un-	in reverse	undo, untie
	out-	in a manner that exceeds	outrun, outdo

	<u>base</u>	<u>sens</u>	<u>forme dérivée</u>	<u>sens</u>
(40)	baboy	‘cochon’	magbababoy	‘vendeur de cochon’
	bulaklak	‘fleur’	magbubulaklak	‘vendeur de fleurs’
	...			

Si ces exemples ne concernent que des préfixes, Kastovsky (2009) propose également une analyse similaire concernant le suffixe *-age* en anglais qui peut être analysé comme ayant le sens de ‘prix, frais’ :

	corkage	(< CORK ‘bouchon’)	‘prix du bouchon’
(41)	housage	(< HOUSE ‘maison’)	‘frais payés pour les biens et les produits de la maison’

En français, on trouve des exemples comparables :

	<u>suffixe</u>	<u>nom</u>	<u>exemples</u>
	-aie	plantation	bananeraie, châtaigneraie
(42)	-ier	arbre	pommier, cerisier
	-iste	partisan	sarkozyste, marxiste
	-iste	spécialiste	pianiste, cardiologiste

Un affixe peut donc être vu comme ayant un contenu sémantique similaire à celui d’un lexème appartenant à une catégorie lexicale majeure (nom, adjectif, adverbe). De même, des constructions impliquant un affixe peuvent être tout à fait comparables avec une opération impliquant un lexème dans un rôle de modifieur :

(43)	maisonnette <i>vs.</i> bébé-maison
------	------------------------------------

Le constituant *bébé-* est analysé par Grandi (2002, p.249-254) comme constituant de composé et par Van Goethem et Amiot (2009) comme un préfixoïde, même s’il garde toutefois un rapport sémantique transparent avec le lexème BÉBÉ. Nous verrons dans la partie suivante non seulement que les constituants de composés et les affixes peuvent entrer en concurrence au sein d’un même schème de nomination mais également que des lexèmes peuvent être en voie d’affixation. Toutefois, nous pouvons dès lors remarquer que, même du point de vue sémantique, les différences entre lexème et affixe ne sont pas nettes et bien marquées.

Quant aux propriétés extrinsèques souvent attribuées aux affixes, elles sont de plusieurs ordres. Nous avons tout d’abord des propriétés combinatoires. Pour reprendre les propriétés de l’affixe dérivationnel prototypique énoncées par Montermini (2002, p.50), un affixe a une position fixe, il est soit préfixe,

soit suffixe, soit circonfixe, mais en aucun cas deux ou trois de ces propositions. De plus, il « ne se lie pas à des unités plus petites ou plus grandes que des lexèmes abstraits (non fléchis) ». Ainsi, la combinaison de deux seuls affixes est impossible. Un affixe permet la mise en place d'une opération morphologique sur une base lexématique : il « s'amalgame avec sa base, en provoque la resyllabification et est intégré dans le même mot prosodique ». Par ailleurs, des contraintes peuvent être activées concernant la sélection d'un type de base. En général, on considère que les affixes peuvent sélectionner un type de base particulier, du point de vue sémantique par exemple, ou sélectionner un thème de lexème (par exemple un thème savant ou un thème populaire). Cependant, je considère ici que ce n'est pas l'affixe qui a ce pouvoir là, mais plutôt l'affixe fait partie d'une construction qui nécessite un type de base particulier (voir par exemple Aronoff (1976)). Ainsi, un affixe ne peut pas être attaché à chaque lexème présent dans le lexique.

Deux autres propriétés sont souvent mises en avant. La première est que les affixes ne sont pas des unités syntaxiquement autonomes. Corbin (2001) parle ainsi d'« unité infralexicale ». Cette propriété sert souvent, encore une fois, à différencier les affixes des lexèmes. Cependant, pour Fradin (2003, p.93), cela n'a pas de sens d'affirmer qu'un lexème est autonome alors qu'un affixe est un élément lié. En effet, le lexème est par définition une entité abstraite : il n'a pas les marques de flexion permettant, en tant que lexème, d'être utilisé en syntaxe. Il serait plus adéquat de parler de l'autonomie des formes fléchies du lexèmes, des mots grammaticaux ou mots occurrences. Cependant, comme nous le verrons ultérieurement, certains lexèmes, quand ils sont employés dans des types de constructions particuliers, peuvent se voir attribuer par exemple une interprétation sémantique spécifique. Il s'agit par exemple des constructions *X fleuve* dans lesquelles de l'entité *fleuve* est extrait uniquement le sème 'long' .

(44) discours fleuve, roman fleuve

Nous verrons dans les parties suivantes comment qualifier ces constructions ; cependant, il est difficile d'affirmer ici que *-fleuve* est une entité autonome puisqu'elle n'est pas dotée du même sens que le lexème FLEUVE employé en (45) :

(45) La Garonne est un très beau fleuve.

Enfin, la dernière propriété met encore l'affixe en balance avec le lexème : le premier appartiendrait à une liste fermée quand le second appartient à une liste ouverte. En effet, si l'on exclut les processus d'affixation qui ont

lieu sur des périodes très longues, la classe des affixes les plus prototypiques est difficilement augmentable. Nous verrons cependant dans la suite de ce chapitre que des procédés tels que l’affixation sécrétive permettent un enrichissement de la classe des affixes sur des périodes de temps courtes, même si les affixes ainsi créés ne peuvent pas être qualifiés de prototypiques :

(46) Watergate → chevalgate, jupegate, tweetgate, lasagne-gate

Lexèmes et affixes sont donc traditionnellement décrits en opposition, selon plusieurs critères. La difficulté de classification des éléments néoclassiques tient à ce qu’ils partagent, a priori, des propriétés autant avec les premiers qu’avec les seconds. Le Tableau 5.9 résume les différences de chacun de ces éléments selon les certains des critères qui viennent d’être énoncés.

	Est autonome	A un sens lexical	Appartient a une liste ouverte
Lexème	✓	✓	✓
Affixe	-	-	-
Élément néoclassique	-	✓	-
Grammème	✓	-	-

Tableau 5.9 – Comparaison des lexèmes, affixes, éléments néoclassiques et grammèmes du point de vue de différentes propriétés

J’ai choisi dans ce tableau de rajouter la catégorie des grammèmes en ce qu’ils sont également opposés aux lexèmes, bien qu’étant autonomes. Ils sont un peu le pendant autonome des affixes flexionnels (pour une définition et un résumé des propriétés, voir par exemple Fradin (2003, p.89)).

Les données de ce Tableau 5.9 relèvent quelque peu d’une vision utopique de la morphologie. Aronoff (1984) parle ainsi d’un système idéal qui comprendrait trois types d’items : les membres des catégories lexicales majeures, les items variant syntaxiquement (items flexionnels) et les morphèmes dérivationnels : « in this best system, each affix has no meaning other than to specify its base and derived category » (Aronoff, 1984, p.270-46).

Je pense cependant qu’il faut cesser de penser qu’un affixe a une réalité autre que phonologique : il n’a de sens ni instructionnel (Corbin, 2001), ni général ou abstrait (Kastovsky, 2009). Il est seulement une marque formelle servant à relier des lexèmes entre eux. En cela, il est l’exposant de schémas constructionnels. Le sens est à retrouver à l’intérieur de ces schémas mêmes, et non pas à l’intérieur de l’affixe. Ainsi, l’affixe *-ier* n’est pas porteur de

sens, il est seulement l'opérateur qui permet l'expression du schéma (47) ; il permet ainsi de mettre les lexèmes POMME et POMMIER en relation.

$$(47) \quad [[X]_{N_i} \text{ier}]_{N_j} \leftrightarrow [\text{arbre dont le fruit est SEM}_i]_j$$

En cela, pour en revenir aux propos de Booij (2010), l'affixe n'étant pas un item lexical, il ne bénéficie pas dans les schémas d'un indice qui correspond à la position occupée dans le lexique à tout élément stocké à l'intérieur de celui-ci. Un affixe n'a donc de réalité qu'à l'intérieur d'un schéma. Il est uniquement un moyen morphologique mais n'a pas de propriétés sémantiques. Il a une position fixe puisqu'il existe dans des schémas dans lesquels les positions sont fixées de façon générale. Cependant, cette définition d'affixe peut englober un nombre plus important d'éléments qui ne sont pas considérés traditionnellement comme tels. Ainsi certains lexèmes peuvent avoir une position fixe dans des schémas constructionnels. Et les notions de dérivation, processus employant prototypiquement des affixes, et de composition, processus employant prototypiquement des lexèmes, doivent être repensés.

5.2 Les schémas constructionnels

Les opérations de construction de lexèmes sont traditionnellement divisées en deux grands types : la dérivation, dans laquelle l'opération de construction est canoniquement satisfaite par un affixe et la composition, c'est-à-dire la combinaison de deux lexèmes. Cependant, les notions de lexème et affixe, dans certains cas, semblent se chevaucher : nous venons par exemple de voir qu'un affixe peut avoir un sens lexical, ou du moins avoir une charge sémantique similaire à celle d'un lexème, et un lexème peut voir sa charge sémantique s'alléger. Comment dans ces cas-là rendre compte des schémas utilisant ces items ? Le cadre de la Morphologie Constructionnelle (voir Booij (2010), ainsi que le chapitre 2 de cette thèse) permet d'englober ces deux types d'opérations de construction de lexèmes sous une même étiquette, celle de constructions pouvant avoir des variables ou des positions fixes.

5.2.1 La composition, un type de dérivation

Si nous concevons la dérivation et la composition comme deux pôles opposés dans ce qui pourrait constituer l'ensemble des opérations morphologiques de construction de lexèmes, essayons de voir qu'est-ce que la dérivation canonique d'un côté, et la composition canonique de l'autre ; canonique plutôt que prototypique car, selon Corbett (2010, p.142), « a prototype, at least

in the term's prototypical use, has an exemplar, is in speakers' heads (it is claimed to be psychologically real), and can vary across cultures. Canonical instances need have no exemplar, they are not claimed to be part of speakers' competence (they are theoretical constructs of linguists), and they are ideally invariant ». Nous allons donc identifier les critères de la dérivation et de la composition canoniques.

Pour la première, je m'appuie toujours sur l'article de Corbett (2010). Selon lui, les dérivés canoniques répondent à deux grands principes. Le premier est qu'ils ont synchroniquement des indicateurs clairs de leur statut : ainsi, un dérivé comprend une base (un lexème) et un marqueur (un affixe), chacun pouvant être substitué par un élément de même nature, afin de former un autre dérivé :

- (48) danseur (< DANSER)
 a. chanteur, bricoleur
 b. danserie, dansable

De plus, le sens du dérivé canonique peut être obtenu de manière régulière à partir du sens de la base ainsi que du sens additionnel de la dérivation. Ainsi, « the canonical ideal would be full compositionality in meaning » (Corbett, 2010, p.144). Pour reprendre l'exemple (48), *danseur* est obtenu de manière régulière à partir du sens de DANSER 'bouger en rythme' et du schéma (49) :

- (49) $[x]_{Vi} \rightarrow [[x]_{Vi} \text{ eur}]_{Nj} \leftrightarrow [\text{qui fait SEM}_i \text{ de façon habituelle ou professionnelle}]_j$

Enfin la structure même du dérivé canonique est transparente, c'est-à-dire que l'ensemble est parfaitement segmentable : dans *danseur*, la base DANSER fournit le thème /dās/ parfaitement transparent et l'affixe *-eur* est également immédiatement identifiable : il n'y a rien de plus, rien de moins.

Le deuxième principe concernant les dérivés canoniques, pour Corbett, est qu'ils sont complètement distincts de leur base : base et mot dérivé ont des indices lexicaux propres, n'occupant pas les mêmes positions dans le lexique, et les mots dérivés comprennent une information sémantique supplémentaire (a « semantic predicate ») par rapport à leur base.

De plus, parmi les affixes, les suffixes sont traditionnellement considérés comme étant dotés d'une plus grande canonicité, de par leur plus grande présence, et à travers différentes langues, et à l'intérieur d'une même langue utilisant différents types d'affixes. Cette distribution a été expliquée par le fait que le début d'un mot est plus saillant en ce qui concerne sa reconnaissance et son interprétation (Montermini, 2008, 2010a, pour plus de détails). Trouver une définition de la composition, même canonique, est bien moins

aisé. Comme le font remarquer Guevara et Scalise (2009, p.106), « one's views and beliefs regarding the fundamental notions of morphology – and of linguistics in general – are critical in shaping a working definition for compounding ». Un composé, impliquant uniquement des objets lexicaux et non pas des marqueurs, dépend crucialement de la définition du lexème que l'on adopte, et du cas que l'on fait des éléments plus petits que ceux-ci comme les thèmes, les racines, les radicaux. De plus, aucun marqueur ne contraint la relation sémantique qui unit les deux items lexicaux. Deux lexèmes construits selon des schémas catégoriels identiques pourront recevoir ainsi des interprétations variées.

- (50) a. $[VN]_N$:
 ouvre-boîte_N (< OUVRIR_V ; BOÎTE_N) 'instrument permettant d'ouvrir les boîtes de conserve'
vs.
 pisse-mamie_N (< PISSER_V ; MAMIE_N) 'tisane, lit. boisson qui fait pisser les mamies'
- b. $[NN]_N$:
 poisson-chat_N (< POISSON_N ; CHAT_N) 'poisson qui a des moustaches semblables à celles du chat'
vs.
 pause-café_N (< PAUSE_N ; CAFÉ_N) 'pause destinée à boire un café'

Guevara et Scalise (2009) décrivent donc le composé canonique ainsi : « in a compound there are typically two constituents that are held together by a relation which is not explicitly (phonetically) realized » (Guevara et Scalise, 2009, p.107) ; ils lui donnent la représentation que j'ai donnée dans le chapitre 1 et que je reproduis ici :

$$(51) \quad [X \mathfrak{R} Y]_Z$$

X, Y et Z représentent les catégories lexicales majeures et \mathfrak{R} la relation sémantique implicite entre les constituants. Pour Guevara et Scalise (2009), un composé canonique est donc une unité formée de plusieurs mots qui doit répondre à au moins quatre critères. Tout d'abord, il constitue un atome syntaxique, ses éléments ne sont pas sujets à, par exemple, des relations d'anaphore avec un élément extérieur au composé. Ensuite, il affiche une intégrité lexicale, entendue au sens de Anderson (1992, p.84) : « the syntax neither manipulates nor has access to the internal structure of words ». Par exemple, les éléments ne peuvent pas être réordonnés. Le troisième critère est que les constituants du composé canonique sont de nature lexicale, ils sont donc membres d'une des catégories lexicales majeures. C'est ce qui est

représenté par X et Y en (51). Enfin, le dernier critère est que le composé dans son ensemble fait lui aussi partie d'une des catégories lexicales majeures. Il s'agit de Z dans la représentation (51).

Les opérations de construction canoniques correspondent en substance, en morphologie constructionnelle, à des schémas plus généraux auxquels correspondent des schémas plus spécifiques, qui en héritent par défaut des propriétés. Les schémas spécifiques peuvent se diviser ensuite en sous-schémas pour parvenir, par exemple, à des idiomes constructionnels et en bout de branche à des unités lexicales spécifiques.

Concernant la composition, les sous-schémas précisent le type de relation qui unit les deux constituants. Pour reprendre la classification de Bisetto et Scalise (2009) qui s'appuient sur la relation grammaticale unissant les deux éléments d'un composé, trois types principaux peuvent être dégagés : les composés subordonnés, les composés attributifs et les composés coordonnés. Dans chacun de ces types, les composés peuvent être endocentriques ou exocentriques. La Figure 5.2 affiche cette proposition de typologie, accompagnée d'exemples du français.

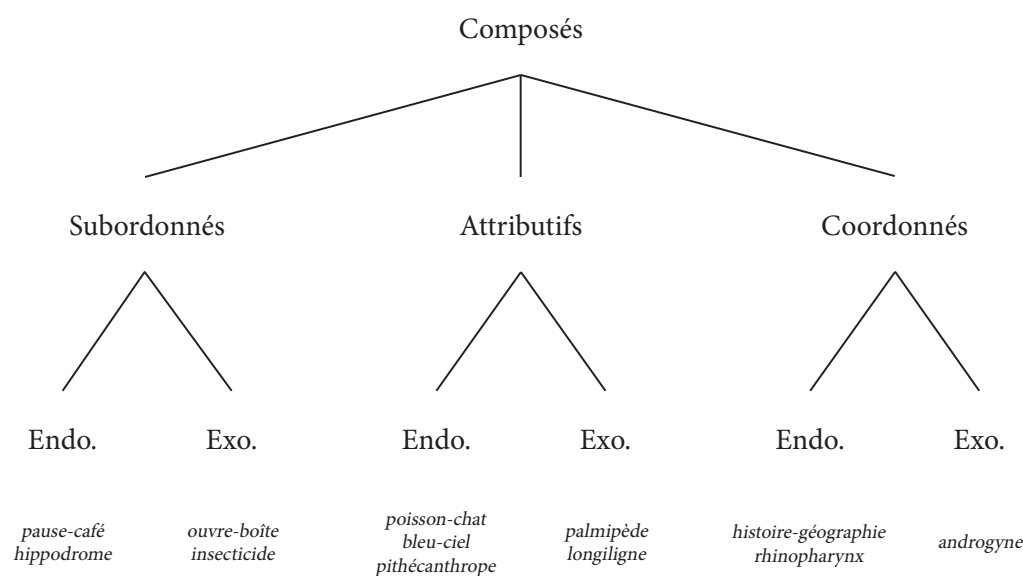


FIGURE 5.2 – Types de composés proposés par Bisetto et Scalise (2009) illustrés par les exemples de Villoing (2012)

Pour les composés subordonnés, la relation \mathfrak{R} est une relation de complémentation, pour les composés attributifs il s'agit d'une relation de modification

et pour les composés coordonnés d'une relation de coordination.

Dans les composés attributifs français, l'élément de droite qualifie l'élément de gauche en lui attribuant une propriété. Si c'est un adjectif, il gardera donc son pouvoir qualifiant, si c'est un nom, la valeur attributive de l'élément de droite est conférée par une valeur métaphorique : c'est une propriété saillante du N qui est extraite pour qualifier le N de gauche. Ainsi, pour Bisetto et Scalise (2009), c'est cette métaphoricité qui permet de différencier *mushroom cloud* 'champignon atomique, lit. nuage qui a la forme d'un champignon' (composé attributif) et *mushroom soup* 'soupe aux champignons' (composé subordonné). C'est cette métaphoricité qui semble également différencier en français *photolyse éclair* et *visite-éclair* : le premier est un composé subordonné et l'élément de gauche conserve tous ses traits sémantiques ('lumineux', 'rapide', 'intense' par exemple) alors que le second serait un composé attributif puisque *éclair* ne conserve ici plus que le sens de 'rapide'. Le fait que l'élément *-éclair* dans cette figuration soit utilisé dans bon nombre de composés pose la question de son statut, question sur laquelle je reviendrai dans la section suivante.

- (52) a. Une technique d'une grande portée et bien adaptée à la détection d'espèces de courte durée de vie, formées consécutivement à l'absorption d'un photon, est celle qui a été mise au point et développée par R. G. W. Norrish et G. Porter à partir de 1950, et connue sous le nom de **photolyse éclair** ou spectroscopie par éclair. Le principe en est simple : il s'agit de créer dans un temps très court, par un éclair bref mais intense, une concentration d'espèces transitoires suffisante pour qu'elles soient observées par spectroscopie et que l'on puisse suivre leur évolution dans le temps. Les éclairs étaient produits par la décharge de condensateurs dans des tubes remplis de gaz sous de faibles pressions fournissant des impulsions lumineuses dont la durée de vie pouvait difficilement être inférieure à quelques microsecondes. (<http://www.universalis.fr/encyclopedie/photochimie/3-techniques-et-applications/>)
- b. Quelques heures avant le passage en 2015, le président de la République, François Hollande était en **visite éclair** au poste de sécurité des Champs-Élysées. (<http://lci.tf1.fr/videos/2014/saint-sylvestre-hollande-en-visite-eclair-au-poste-de-securite-85402708.html>)

Concernant la question de la tête du composé, et donc du caractère endocentrique ou exocentrique de chaque instance, Guevara et Scalise (2009)

proposent deux types de tête : il y a tout d'abord une tête formelle, c'est-à-dire qu'elle partage avec le composé ses traits formels. De ce point de vue, le composé a les mêmes propriétés distributionnelles que la tête. Ainsi, *bleu-ciel* a les mêmes propriétés distributionnelles que *bleu*, il peut donc qualifier un nom ou être en position prédicative :

- (53) a. Aujourd'hui je mets mon pantalon bleu.
Aujourd'hui je mets mon pantalon bleu-ciel.
b. Ce pantalon est bleu.
Ce pantalon est bleu-ciel.

La tête peut également être sémantique, c'est-à-dire qu'elle offre au composé ses informations lexicales et conceptuelles; le composé est ainsi un hyponyme de la tête. *Poisson-chat* est ainsi une sorte de *poisson*, *bleu-ciel* un type de *bleu*. La question de l'exocentricité est traitée ainsi : « An *endocentric compound* has at least one formal head and at least one semantic head. If an endocentric compound has only one formal head and only one semantic head, then the two must coincide. If a compound has one or more formal heads and no semantic head, it will be considered *exocentric*. If a compound has one or more semantic heads and no formal head, it will also be considered *exocentric* » (Guevara et Scalise, 2009, p.113-114). Cependant, pour reprendre l'exemple des composés exocentriques $[VN]_N$ du français, la tête n'est ni formelle, ni sémantique. Dans la relation de complémentation, le N est un argument du V, souvent un argument interne. Syntaxiquement, le V gouverne donc le N. Cependant, *ouvre-boîte* ne partage pas les propriétés distributionnelles du verbe OUVRIR et n'en est pas non plus un hyponyme. Si, en tant que nom, il partage les propriétés distributionnelles de BOÎTE, il est difficilement concevable, au vu des données syntaxiques, de faire de BOÎTE la tête du composé. De plus, BOÎTE et OUVRE-BOÎTE ne partagent pas de propriétés sémantiques spécifiques.

Par ailleurs, pour Booij, les schémas de construction permettent d'attribuer au mot construit des propriétés qui ne sont pas directement déductibles de ses constituants : « there is room for holistic properties » (Booij, 2010, p.37). Par exemple, il propose le schéma suivant pour rendre compte des composés du type *ouvre-boîte* permettant ainsi de faire des généralisations sur les composés $[VN]_N$ des langues romanes, traditionnellement traités comme des exocentriques :

- (54) $[[V_k] [N_i]]_{N_j} \leftrightarrow [\text{agent / instrument}_j \text{ of action}_k \text{ on object}_i]_j$

La question de la tête n'est donc plus ici pertinente. Concernant les composés

dits endocentriques, la question de la tête est traitée de la façon suivante dans les schémas constructionnels :

$$(55) \quad \begin{array}{ccc} [[a]_{xi} [b]_{yj}]_{xk} & \leftrightarrow & [SEM_i \text{ entretenant une relation R avec } SEM_j]_k \\ | & & | \\ [\alpha F] & & [\alpha F] \end{array}$$

Dans ce schéma, X et Y représentent des catégories lexicales majeures, et $[\alpha F]$ représente une série de traits pertinents. Le schéma montre que $[a]$ est la tête du composé puisque cet élément fournit à ce dernier non seulement sa catégorie lexicale (ce qui peut relever de la tête formelle de Guevara et Scalise (2009)) et un ensemble de traits qui peuvent être sémantiques ou syntaxiques.

Le schéma (55) représente un composé canonique du français. Comme dans les autres langues romanes, l'élément fournissant ses propriétés à l'ensemble du composé est l'élément de gauche.

Cependant ce schéma reste canonique et des formations peuvent s'en éloigner. Par exemple, dans les formations en (56), l'élément fournissant ses propriétés au composé est l'élément de droite.

- (56)
- a. auto-école 'école pour apprendre à conduire'
 - autoberge 'berge destinée aux automobilistes'
 - auto-stop 'activité qui consiste à demander l'arrêt des voitures'
 - b. ciné-club 'club d'amateurs de cinéma'
 - ciné-parc 'parc dans lequel est projeté des films'
 - c. télé-achat 'offre de vente de produits par le biais de la télévision'
 - téléthéâtre 'théâtre retransmis à la télévision'
 - téléréalité 'programme télévisuel filmant la vie quotidienne de candidats'
 - d. eurosceptique 'personne sceptique à propos de l'Union européenne'
 - euromissile 'missile basé en Europe'
 - eurodéputé 'député européen'

Ces cas de constructions sont limités à certains éléments initiaux, qui résultent en général de la troncation de lexèmes plus longs :

- (57)
- a. auto < AUTOMOBILE
 - b. ciné < CINÉMA

- c. télé < TÉLÉVISION
- d. euro < EUROPE

Certaines de ces troncations peuvent toutefois parfaitement être utilisées de façon autonome. Pour Rainer et Varela (1992), ces constituants initiaux sont à analyser comme des préfixes, compte-tenu par exemple de leur nombre restreint et de leur gabarit phonologique (ce sont des éléments dissyllabiques se terminant par des voyelles). Cette position est toutefois démentie d'une part par le fait que ces éléments ont un sens identique à celui du lexème dont ils sont issus, d'autre part par le fait que la relation sémantique reliant les constituants n'est pas fixe, ce qui est caractéristique des composés : « compounds express a relation between two objects or notions, but say nothing of the way in which the relation is to be understood. That must be inferred from the context or otherwise » (Jespersen, 1943, p.137).

Pour Booij (2010, p.67), ils peuvent être analysés comme des sous-schémas de composés, puisque les schémas peuvent comprendre des éléments spécifiés :

- (58)
- a. [[auto]_{Ni} [x]_{Nj}]_{Nk} ↔ [SEM_j entretenant une relation R avec SEM_i]_k
 - b. [[ciné]_{Ni} [x]_{Nj}]_{Nk} ↔ [SEM_j entretenant une relation R avec SEM_i]_k
 - c. [[télé]_{Ni} [x]_{Nj}]_{Nk} ↔ [SEM_j entretenant une relation R avec SEM_i]_k
 - d. [[euro]_{Ni} [x]_{Nj}]_{Nk} ↔ [SEM_j entretenant une relation R avec SEM_i]_k

« This schema states that right-headed compounds are possible with a number of specific, listed constituents » (Booij, 2010, p.67).

D'autres composés affichent également un élément de droite fournissant les propriétés sémantico-syntaxiques à l'ensemble du mot construit, sans que l'élément de gauche ne soit nécessairement tronqué. Dans ces cas-là, l'élément récurrent est l'élément de droite, celui-là même dont le composé hérite ses propriétés.

- (59)
- a. théâtrothérapie, thalassothérapie, ...
 - b. caféinodépendant, cocaïnodépendant, ...
 - c. twittosphère, Dieudosphère, ...
 - d. technopôle, canceropôle, ...
 - e. Hollywood compatible, Sarko-compatible, ...
 - f. beauf attitude, Bayrou attitude, ...

Ces exemples peuvent également être généralisés à travers des schémas de

construction, sur lesquels je reviendrai plus longuement dans la section 5.2.2.1. Certains de ces exemples ((59-a), (59-b), (59-c) ou (59-d)) sont généralement traités dans la littérature comme des composés néoclassiques. Cependant, peut-on rendre compte de toutes les constructions néoclassiques grâce à des sous-schémas d'un schéma principal généralisant les composés ?

En effet, les constructions traditionnellement appelées composés néoclassiques sont différenciées en français des composés natifs (ou composés ordinaires) par le fait qu'elles affichent un élément de droite gouvernant l'élément de gauche. Par exemple Villoing (2012, p.51) affirme : « the head of native endocentric compounds is to the left (XY), as in other Romance languages, while the head of neoclassical compounds is on the right (YX) ». Mais peut-on vraiment considérer cet élément de droite comme étant la tête des composés ? Les constructions impliquant des éléments néoclassiques font montre de davantage d'exocentricité que les composés natifs. Villoing (2012) relève quatre types de composés néoclassiques exocentriques (60-a) contre un type de composés natifs (60-b) :

- (60) a. $NN_{A/N}$: androgyne
 NN_A : anthropomorphe
 AN_A : macrocéphale
 $NV_{A/N}$: biographe
 b. $VN_{(A/)N}$: ouvre-boîte

De plus, si on considère qu'un composé endocentrique doit être un hyponyme d'un de ses constituants, certains composés néoclassiques classés par Villoing (2012) parmi les endocentriques seraient plutôt à considérer, selon moi, comme des exocentriques :

- (61) pithécantrope (< *pithéc-* 'singe' ; *-antrope* 'homme') : mammifère primate qui a des caractéristiques du singe et des caractéristiques de l'homme
 mégalithe (< *méga-* 'grosse' ; *-lithe* 'pierre') : monument constitué d'une grosse pierre
 hippodrome (< *hippo-* 'cheval' ; *-drome* 'course') : 'lieu où se déroulent des courses de chevaux'

Les exemples restants, parmi ceux classés dans les composés exocentriques, dans le tableau qui apparaît à la fin de l'article de Villoing (2012, p.56), sont ceux reproduits ci-dessous :

- (62) a. rhinopharynx
 franco-allemand
 médico-social
 anarcho-capitaliste
- b. micro-organisme
 macromolécule
- c. hydrothérapie
 eurodéputé
 photosensible
 éco-compatible
 séropositif

Il est intéressant de remarquer que chacun de ces lexèmes a au moins un de ses constituants qui est syntaxiquement autonome. Si la question de la tête se pose moins en (62-a), puisqu'il s'agit de composés coordonnés⁷, les constructions en (62-b) et (62-c) présentent quant à elle un élément de droite syntaxiquement autonome en français ; c'est également l'élément hyperonyme du composé dans son ensemble, que nous pourrions donc considérer comme la tête de la construction.

Au vu de ces éléments, il semble, même si cela doit être confirmé par des études plus approfondies, qu'il est difficile pour un élément syntaxiquement non autonome de constituer la tête d'une construction, du moins selon la définition de Guevara et Scalise (2009). Ce serait une des raisons pour lesquelles, selon Villoing et Namer (2006), « la morphologie 'courante', avec la composition VN, est incapable de construire un sens équivalent à celui qu'offre la composition NV et la langue doit, pour cela, faire appel à la syntaxe ». Cependant, il est possible de rendre compte de généralisations de ces constructions à travers des schémas constructionnels. Par exemple, le schéma (63) généralise des constructions comme *hippodrome*, *aérodrome* ou *vélodrome* :

- (63) [[x]_{Ni} [drome]_{Nj}]_{Nk} ↔ [lieu où une course caractérisée par SEM_i a lieu]_k

Ce schéma, dont l'un des éléments est spécifié, est comparable aux schémas reproduits en (58), même si c'est ici l'élément de droite qui est spécifié. Il est également possible de généraliser les constructions dénotant des propriétés

7. Le statut de composé coordonné de *rhinopharynx* me pose ici problème, puisque que rhinopharynx est une partie du pharynx, autrement appelée *cavum*, pharynx qui comprend trois « étages » : le rhinopharynx est l'étage supérieur. Le rhinopharynx est donc la 'partie du pharynx constituée des fosses nasales'. Cependant, je n'approfondirai pas ici cette question.

ou des noms [+animé] telles que *noctiluque*, *herbivore* ou *anthropophage* :

- (64) $[[a]_i [b]_k]_{AN_j} \leftrightarrow [SEM_j \text{ caractérisé par accomplissant l'action } SEM_k \text{ sur } SEM_i]_j$

Cette construction est issue des constructions du grec ancien formant des N [+humain] ou des adjectifs, dont le second élément est un élément nominal déverbal, que nous avons vues dans le chapitre 1, exemple (28), que je reproduis ici :

- (65) a. φαγειν ‘manger’ > -φάγος ex. ἄρτοφάγος (< ἄρτος ‘pain’) ‘celui qui mange du pain’
 b. γράφω ‘écrire’ > -γράφος ex. λογογράφος (< λόγος ‘discours’) ‘celui qui écrit les discours’
 c. φερῶ ‘apporter’ > -φόρος ex. ἀνθοφόρος (< ἄνθος ‘fleur’) ‘celui qui porte des fleurs’
 d. δέμω ‘construire’ > -δόμος ex. οἰκοδόμος (< οἶκος ‘maison’) ‘constructeur, architecte’
 e. τρέχω ‘courir’ > -δρόμος ex. ὄρειδρόμος (< ὄρος ‘colline’) ‘celui qui court sur les collines’
 f. ποιέω ‘faire’ > -ποιός ex. ἀγαλατομοιός (< ἄγαλμα ‘statue’) ‘sculpteur’

J’avais déjà analysé ces formations comme des constructions exocentriques. L’exocentricité se retrouve également dans des noms ne dénotant pas des entités [+animé] tels que *anthropologie* ou *cartographie*. En effet, postuler *anthropologie* comme étant endocentrique, pouvant être paraphrasé ‘étude de l’homme’ ou ‘discours sur l’homme’, selon le lexème dont on considère que *-logie* est l’allomorphe suppose d’analyser ce nom comme un nom processif; or *anthropologie* signifie la ‘discipline qui consiste à étudier l’homme, à discourir sur l’homme’. De même, *cartographie* n’est pas une ‘écriture de cartes’ mais plutôt une ‘théorie et technique de l’établissement des cartes géographiques’ (TLFi).

Par ailleurs, la récurrence d’un schéma de construction provoque des glissements de sens. Prenons l’exemple du schéma *Xthèque*. L’élément *-thèque* est issu du grec θήκη, ‘coffre, boîte’ et *bibliothèque* signifiait à l’origine un ‘meuble servant à ranger des livres’. Par extension, déjà en grec ancien, le nom a référé à l’endroit regroupant des livres. Hors le sens métonymique de ‘collection de livres, recueil’, premier sens employé en français selon le DHLF, le sens de ‘armoire contenant des livres’ précède celui de ‘bâtiment renfermant des livres’ en français. C’est cependant *bibliothèque* avec ce dernier sens

qui a servi de mot leader pour la formation des mots suivants :

(66) ludothèque, cinémathèque, cassetothèque

Le schéma (67) représente la généralisation de ces constructions :

(67) [[a]_{X_i} [thèque]]_{N_j} ↔ [lieu où SEM_i est entreposé / conservé / susceptible d'être emprunté]_j

Dans ce schéma, il est difficile de soutenir que *-thèque* est un allomorphe de BOÎTE ou COFFRE. Si βιβλιοθήκη est en grec un composé [NN]_N de type subordonné dont la fonction de tête sémantique et formelle est remplie par l'élément de droite, il n'en est pas de même en français. Pouvons-nous alors parler d'un cas de grammaticalisation, c'est-à-dire du passage d'un élément lexical dans la classe des éléments affixaux ? Pour répondre à cette question, il faut envisager la composition et la dérivation d'un point de vue comparatif.

A partir des définitions de la dérivation et de la composition canoniques données en début de section, mises à part les questions de nature des éléments, quelles sont les différences que l'on peut dégager entre ces deux types de constructions ?

La principale différence se situe au niveau de la construction du sens, ou plutôt de sa représentation. Le sens d'un dérivé peut être généralisé dans une construction alors que le sens d'un composé dépendra davantage d'un savoir encyclopédique et contextuel. Ainsi, un mot construit selon le schéma (49), si la base sélectionnée par la construction est adéquate (c'est-à-dire par exemple, un verbe ayant un argument agentif), aura la même relation sémantique entre ses éléments qu'un autre mot construit selon le même schéma. Au contraire, un mot construit selon un des schémas donnés en (58) n'affichera pas forcément une relation sémantique entre ses constituants identique à celle d'un autre mot construit selon le même schéma. Même si les composés peuvent afficher des relations plus prédictibles que d'autres, par exemple les [VN]_N construisent majoritairement des noms d'instruments et des noms d'agent, la partie sémantique de la construction ne pourra pas forcément représenter l'ensemble des lexèmes pouvant être construits sur le même patron morphologique.

Cependant, doit-on absolument opposer composition et dérivation ? Fradin (2000, p.52), à propos de certains éléments finaux, les classe comme composés car : « compounding, first because formatives like *-tique*, *télé-*, *-tel*, *-ware*, etc. are semantically not affixes, (...) ». Donc, ce qui n'est pas de la dérivation doit-il être considéré forcément comme de la composition, et

vice-versa ? D'un point de vue onomasiologique, pourtant, composition et dérivation ne sont pas à opposer. En effet, des opérations morphologiques différentes peuvent participer à la construction de mêmes « schèmes de nomination » (Roché, 2011b, p.21-22). Les schèmes de nomination assurent, via la référence, un lien entre les lexèmes en tant qu'objets linguistiques et les objets du monde⁸. Plusieurs configurations sont envisageables : un même schème peut être réalisé par des opérations morphologiques multiples et imprévisibles, suite à des évolutions diachroniques particulières par exemple (Tableau 5.10, a.) ou plusieurs procédés spécifiques peuvent être en concurrence pour un seul schème (Tableau 5.10, b.). Autre possibilité, une opération unique peut se dégager pour remplir un seul schème (Tableau 5.10, c.). Un

	Schème de nomination	Opération morphologique	Exemples
a.	[sorte de SEM _i] _j	[[x] _{Ni} SUFF] _{Nj}	coquillage caveau portière veston
b.	[instrument servant à SEM _i] _j	[[x] _{Vi} oir] _{Nj} [[x] _{Vi} eur] _{Nj} [[x] _{Vi} [x] _{Nk}] _{Nj}	semoir compteur lave-linge
c.	[arbre sur lequel pousse SEM _i] _j	[[x] _{Ni} ier] _{Nj}	pommier cerisier poirier

Tableau 5.10 – Différentes configurations de l'intersection entre schèmes de nomination et opérations morphologiques selon Roché (2011b)

même schème de nomination peut donc être réalisé par un affixé, un composé ou une construction néoclassique (Tableau 5.11).

Štekauer (2005) propose un modèle qui se rapproche par certains aspects des schèmes de nomination de Roché. Ce modèle cognitif-onomasiologique s'articule sur différents niveaux : la réalité extra-linguistique, la communauté des locuteurs et le signe linguistique ; la procédure de *Morpheme-to-Seme-Assignment Principle* affecte un morphème à un sème. En cela, elle permet

8. Même si ces notions ne sont pas vraiment superposables, mais parce qu'il a fallu les représenter, les schèmes de nomination seront représentés ici comme la partie des schémas de construction s'occupant de la construction du sens de l'output à partir de celui de l'input.

Schème de nomination	Opération morphologique	Exemples
[produit servant à anéantir SEM _i] _j	[anti [x] _{Ni}] _{Nj} [[x] _{Vk} [x] _{Ni}] _{Nj} [[x] _{Ni} cide] _{Nj}	anti mouche tue-mouche mouchicide

Tableau 5.11 – Réalisations du schème [produit servant à anéantir SEM_i]_j

de générer un nombre de possibilités morphologiques pour répondre à un unique besoin onomasiologique ; des composés et des affixés peuvent être des « outcomes of one and the same onomasiological type, a common, universal word-formation process, which is reflected in the identical, conceptually based, onomasiological structure » (Štekauer, 2005, p.155).

Une opération morphologique de construction de lexème est une opération qui a pour base un lexème dans sa tridimensionnalité (une forme, un sens, une catégorie syntaxique) et qui aboutit aussi à un lexème tridimensionnel. Dans le cas de l’affixation, l’opération morphologique consiste en l’adjonction d’un affixe à une base lexicale et dans le cas de la composition, en la coalescence de deux items lexicaux. La morphologie constructionnelle s’occupe de, selon Aronoff et Fudeman (2011, p.262 (glossaire)), « the creation of a new lexeme from one or more other lexemes through the application of some morphological process, such as affixation or compounding. Also called lexeme formation and word formation. Derivation contrasts with inflection ». Ainsi, à une même opération sémantique peuvent correspondre différentes opérations formelles : l’adjonction d’un suffixe ou la coalescence de deux lexèmes (ou plus)⁹. Dans la Morphologie Constructionnelle telle qu’elle est proposée par Booij (2010), une des différences entre lexème et affixe est que les premiers sont stockés dans le lexique (et reçoivent donc un indice lexical) quand les seconds ne le sont pas. Mais cette différence ne doit pas pour autant nous pousser à tracer une frontière imperméable entre les deux : « there is simply no way to exclude from derivational semantics types of meanings which are more typical of lexical categories like noun, verb, and adjective » (Aronoff, 1984, p.49). Les constituants de composés sont donc des objets à étudier sur le même plan que les affixes : « derivational affixes are pieces of morphological structure, just like the constituents of compounds » (Booij, 2005, p.130).

9. D’autres opérations, comme la conversion, sont également possibles, mais hors de propos ici.

Pour Roché (2009), « ce que l'on appelle « composition » est une forme de dérivation, à laquelle on peut appliquer les mêmes méthodes d'analyse (...) le corollaire est que la notion de base est également pertinente ». Il propose ainsi d'abandonner la notion de tête dans l'analyse de la composition.

Dans le Tableau 5.11, j'avais mis en parallèle *anti mouche*, *mouchicide* et *tue-mouche* comme répondant au même schème de nomination. Puisque la base de *anti mouche* est sans conteste MOUCHE, il n'y a aucune raison sémantique qui empêcherait de considérer ce lexème comme étant également la base de *mouchicide* ou *tue-mouche*. Le fait que, dans *tue-mouche*, le constituant *mouche* soit un argument interne du verbe TUER ne semble pas empêcher cette analyse. En effet, le schéma (68) se montre disponible :

(68) [[tue]_k [x]_{N_i}]_{N_j} ↔ [produit servant à anéantir SEM_i]_j

- a. Il attire les limaces, évitant d'utiliser un **tue limace** chimique. (http://www.alterre-bourgogne.org/arkotheque/client/alterre_bourgogne/_depot_arko/articles/136/guidepeda-alimresp-fiche-15-composition-de-jardins_doc.pdf)
- b. je ne pars jamais en vacances sans ma bombe de **tue-araignées** (<http://ceciel.unblog.fr/2008/01/14/une-araignee-au-plafond/>)

De même, dans le schéma (69), la base est l'élément variable, soit [x] : *mouchicide* est construit sur le lexème base MOUCHE, tout comme *insecticide* est construit sur INSECTE.

(69) [[x]_{N_i} cide]_{N_j} ↔ [produit servant à anéantir SEM_i]_j

Pour les schémas affichant des éléments récurrents dans des positions fixes, schémas que Booij (2010) nomme idiomes constructionnels, la base est l'élément qui est dans la position variable.

La récurrence de certains éléments lexicaux dans des positions fixes de certains schémas provoque ainsi la formation de paradigmes plus ou moins importants. Ces éléments lexicaux peuvent de plus voir leur sens se spécialiser quand ils interviennent dans ces schémas particuliers. Ces schémas sont nommés idiomes constructionnels par Booij :

The schema at the intermediate level is a constructional idiom : a construction in which one of the two positions is occupied by a specific lexical item. In this respect this schema is completely identical to that for deverbal nouns (Booij, 2005, p.125).

Booij (2010, p.56) donne ainsi l'exemple de l'usage productif de lexèmes comme intensificateurs en néerlandais. Des sous-schémas se dégagent, dans lesquels l'exposant de construction, c'est-à-dire l'élément marquant formellement l'opération sémantique, est un lexème, ici le nom *bere* 'ours', l'adjectif *dol* 'fou' ou le verbe *kots* 'vomir' :

- (70)
- a. $[[\text{bere}]_N [X]_{A_i}]_{A_j} \leftrightarrow [\text{très SEM}_i]_j$
 ➤ *bere-sterk* 'très fort, lit. ours-fort'
 - b. $[[\text{dol}]_A [X]_{A_i}]_{A_j} \leftrightarrow [\text{très SEM}_i]_j$
 ➤ *dol-blij* 'très heureux, lit. fou-heureux'
 - c. $[[\text{kots}]_V [X]_{A_i}]_{A_j} \leftrightarrow [\text{très SEM}_i]_j$
 ➤ *kots-misselijk* 'très malade, lit. vomir-malade'

La présence de ces idiomes constructionnels pose la question de la grammaticalisation d'items qui passent d'un statut lexical à un statut affixal. Cette question est particulièrement pertinente dans cette thèse puisque, si les ENs sont à l'origine des lexèmes autonomes et certains peuvent être aujourd'hui considérés comme des thèmes S de lexèmes français (revoir partie 5.1.1.1), d'autres semblent uniquement avoir la place d'exposants de constructions.

5.2.2 Perméabilité des frontières

5.2.2.1 De l'idiome constructionnel à la grammaticalisation

Nous venons donc de voir qu'un idiome constructionnel est un sous-schéma dans lequel un élément est spécifié à l'intérieur même de la construction. Ces sous-schémas sont instanciés par des séries plus ou moins longues de lexèmes. Le constituant récurrent dans ces listes peut recouvrir un sens spécifique à l'intérieur de la construction.

« we can say that an element can occur both as part of a complex word and as a free morpheme. In such cases, only a careful analysis of its linguistic properties can reveal whether the element in question is really the same in both cases. If (and only if) there are significant differences between the two usages we can safely assume that we are dealing with two different items. If there are no significant differences, the element should be treated as a free morpheme and the pertinent complex word as a compound » (Plag, 2003)

Je propose de regarder ici certaines de ces listes de lexèmes affichant un élément récurrent, afin d'observer les différences sémantiques éventuelles entre le lexème autonome et le lexème tel qu'il apparaît dans les constructions.

Ces observations sont la continuation de celles que nous avons amorcées dans Lasserre et Montermini (2014b). La Figure 5.3 propose l'intégration de certains sous-schémas affichant des N récurrents dans une structure hiérarchique du lexique. Pour rappel, ce genre de représentation implique l'héritage par défaut, c'est-à-dire que chaque niveau hérite par défaut des propriétés du niveau supérieur, à moins qu'une propriété plus spécifique soit affichée.

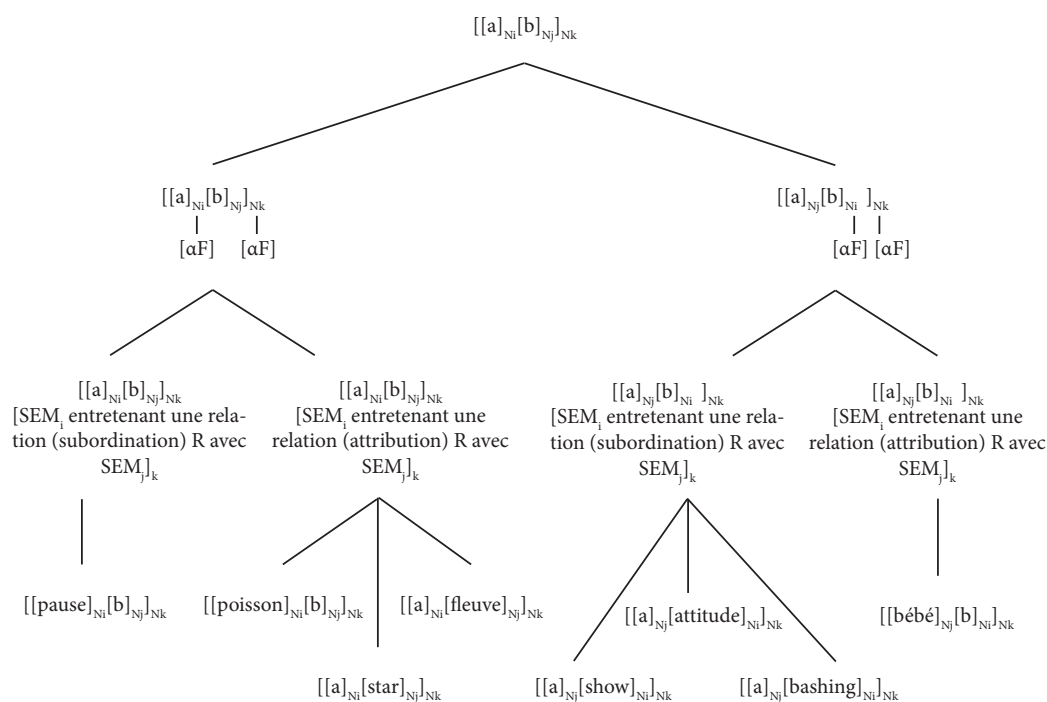


FIGURE 5.3 – Sous-schémas $[NN]_N$

Regardons en détails chacun de ces sous-schémas. Sur les branches de gauche de la Figure 5.3 sont présentées des constructions $[NN]_N$ dont le constituant léguant ses propriétés au composé (la tête dans les traitements traditionnels) est l'élément de gauche.

Parmi les constructions dont la relation entre les constituants est une relation de subordination, l'élément de gauche, c'est-à-dire l'hyperonyme du composé créé, peut être récurrent (71)¹⁰ :

- (71) a. pause café, pause cigarette

10. Koga (2012) traite de ce type de composés. Je remercie l'auteur de m'avoir permis d'accéder à ses données.

- b. espace détente, espace jeux
- c. coin cuisine, coin terrasse, coin parents

Dans ces exemples, même si l'élément initial est récurrent, les composés fonctionnent comme des composés subordonnés prototypiques : l'élément de droite complète l'élément de gauche. Cependant, la relation unissant les deux éléments semble prédéfinie dans la construction même ; l'interprétation du composé ne demande pas de recours au contexte ou à des connaissances extra-linguistiques :

$$(72) \quad [[\text{espace}]_{N_i} [\text{b}]_{N_j}]_{N_k} \leftrightarrow [\text{SEM}_i \text{ destiné à SEM}_j]_k$$

Parmi les constructions de type attributif, l'élément récurrent peut être l'élément de gauche ou l'élément de droite. Dans la première configuration, c'est, comme précédemment, l'élément hyperonyme du composé qui est récurrent ; il ne subit aucun changement sémantique. L'élément de droite, quant à lui, spécifie la base par une des propriétés saillantes de son référent. C'est un schéma de construction privilégié pour répondre au schème de nomination de certains animaux :

$$(73) \quad \text{poisson-scie, poisson-chat, poisson-coffre}$$

Quand c'est l'élément de droite qui est récurrent, c'est-à-dire l'élément qui qualifie la base du composé, cet élément peut apporter l'intégralité de ses traits sémantiques (74), ou bien certains de ses traits sémantiques peuvent être sélectionnés par métaphore (75).

$$(74) \quad [[a]_{N_i} [\text{star}]_{N_j}]_{N_k} \leftrightarrow [\text{SEM}_i \text{ caractérisé par le fait qu'il est SEM}_j]_k$$

- pâtissier star
- acteur star
- enfant star

$$(75) \quad \text{a. } [[a]_{N_i} [\text{fleuve}]_{N_j}]_{N_k} \leftrightarrow [\text{long SEM}_i]_k$$

- roman fleuve
- discours fleuve

$$\text{b. } [[a]_{N_i} [\text{clé}]_{N_j}]_{N_k} \leftrightarrow [\text{SEM}_i \text{ primordial}]_k$$

- idée clé
- mot clé

$$\text{c. } [[a]_{N_i} [\text{éclair}]_{N_j}]_{N_k} \leftrightarrow [\text{SEM}_i \text{ express}]_k$$

- visite éclair
- campagne éclair

Les constructions données en (75) ont été étudiées par exemple par Noailly

(1990) ou Amiot et Van Goethem (2010), mais pas de façon univoque. Pour Noailly, ces N récurrents doivent être analysés comme des adjectifs et ainsi la construction comme relevant de la syntaxe : elle s'appuie pour cela sur l'extension métaphorique des N et propose des tests syntaxiques tels que :

(76) Une visite tout-à-fait éclair (Noailly, 1990, p.43)

En effet, comme le notent Amiot et Van Goethem (2010), un élément comme *-clé* peut être modifié par un adverbe ou utilisé comme attribut. De plus, *-éclair* peut être coordonné avec un adjectif :

(77) COP21 : visite éclair et symbolique de Laurent Fabius à Bonn (http://www.lemonde.fr/cop21/article/2015/10/20/cop21-laurent-fabius-en-visite-eclair-et-symbolique-a-bonn_47932709_27027432.html)

Les éléments *-fleuve*, *-éclair* et *-clé* sont utilisés de façon non référentielle. Mais, comme le font remarquer Amiot et Van Goethem (2010, p.851), cet emploi de noms de façon non référentielle est très fréquent, non seulement dans les composés attributifs, comme nous l'avons vu pour *poisson-chat* par exemple, mais également de façon autonome :

(78) Pierre est très ours en ce moment. / Marie devient de plus en plus femme. (Amiot et Van Goethem, 2010, p.851)

Pour elles, ce type d'emploi ne demande pas un changement de catégorie du N modifieur. Cependant, la différence relevée par Amiot et Van Goethem entre *poisson-chat* et *idée clé* est que, dans le premier cas, le N2 modifieur fournit une propriété permanente et définitoire au N1 : par exemple, avoir des moustaches. C'est ce qui permet d'inscrire la formation *poisson-chat* parmi les espèces de poisson, et ce de manière permanente. Dans le second cas, le N2 offre une propriété non essentielle : une idée est cruciale dans un contexte donné, une visite très rapide pour tel événement, etc.

Regardons maintenant ce qu'il se passe dans la partie droite de la Figure 5.3. Ici, ce sont les éléments de droite du composé qui constituent les hyperonymes de celui-ci. Cet ordre est souvent appelé ordre néoclassique (voir par exemple Amiot et Dal (2008)) ; en effet, comme nous l'avons vu dans ce travail, les constructions néoclassiques ont prototypiquement l'élément gouvernant à droite ; cette propriété est héritée des langues classiques, en particulier du grec, comme nous l'avons vu dans le chapitre 1. Amiot et Dal (2008) soulignent que cet ordre est souvent marqué par la présence de la voyelle /ɔ/ ou /i/. Dans ce type de constructions, le N1 est souvent un

thème S de lexème (soit un EN), mais ce n'est pas une condition *sine qua non*. Ces composés sont en général de type subordonné.

- (79) a. théâtrothérapie, phytothérapie
 b. cocaïnodépendant, caféinodépendant
 c. cancéropôle, technopôle
 d. blogosphère, twittosphère, dieudosphère (< Dieudonné)
 e. Hollande show, Aubry show, Chirac show, Hollandoshow, Chiracoshow

Contrairement aux constructions du type *Xfleuve*, ici l'élément récurrent est celui qui fournit au composé ses propriétés sémantico-syntaxiques (c'est donc une construction parallèle à celle *poisson-X*). La relation entre les constituants peut être multiple, comme l'ont montré Amiot et Dal (2008) concernant les formations de types *Xthérapie* :

- (80) a. [[a]_{Nj} [thérapie]_{Ni}]_{Nk} ↔ [SEM_i au moyen de SEM_j]_k
 > hydrothérapie, théâtrothérapie
 b. [[a]_{Nj} [thérapie]_{Ni}]_{Nk} ↔ [SEM_i effectuée sur SEM_j]_k
 > dermothérapie, vertébrothérapie

Dans d'autres cas, la relation sémantique unissant les deux constituants est unique :

- (81) [[a]_{Nj} [dépendant]_{(A)Ni}]_{Nk} ↔ [personne dépendante de SEM_j]_k

De plus, si certains de ces éléments récurrents conservent pleinement leur sens (*thérapie, dépendant*), pour d'autres un sens en particulier est sélectionné dans ces constructions ; pour *-pôle* c'est le sens de « point central qui exerce une attraction, un rayonnement » et pour *-sphère* celui de « domaine limite où se manifestent et s'exercent l'activité, les attributions ou l'influence de quelqu'un » (définitions du TLFi). Cependant, contrairement aux formations du type *Xfleuve* qui sélectionnaient un trait saillant du N2, ici c'est un des sens que peut recouvrir le lexème qui est sélectionné.

Les N2 récurrents rentrant dans ce type de constructions sont en général non natifs, et surtout d'origine grecque ou latine. Cela explique sans doute la présence de la voyelle héritée de la composition gréco-latine et se trouvant aujourd'hui dans les constructions mêmes (revoir à ce sujet la partie 3.1.2). Cependant, on voit l'émergence de N2 d'origine anglophone, tels que *show*. L'ordre des constituants s'explique également par l'ordre des constituants dans la langue d'origine. Même si la voyelle de liaison n'est pas ici diachroniquement explicable, on remarque qu'elle semble parfois nécessaire

pour certains locuteurs pour marquer l'ordre inverse de cette composition, comparé aux composés natifs. Le lexème *SHOW* est également entré dans la langue française de façon autonome, en dehors de son apparition dans ces constructions. Ce n'est pas le cas d'autres constructions, formées elles aussi à partir d'un lexème anglophone, telles que :

$$(82) \quad [[a]_{N_j} [bashing]_{N_i}]_{N_k} \leftrightarrow [ressentiment \text{ anti } SEM_j]_k$$

Le verbe anglais *bash* signifie 'critiquer, descendre'. Le mot *leader* ayant servi de modèle à la création du schéma (82) semble être ici *French bashing*, créé en anglais et importé dans sa globalité :

- (83) a. « La France, un pays fini » : petite histoire des grands classiques du **French bashing** (<http://www.atlantico.fr/decryptage/france-pays-fini-petite-histoire-grands-classiques-french-bashing-1793043.html#WLawur0QCgiwKfkE>.99)
- b. Le « **French bashing** » qui a connu un pic au début de la Guerre d'Irak en 2003 à laquelle était opposé Paris, est une habitude largement répandue des républicains aux Etats-Unis. Surtout lorsqu'il sont en difficulté dans les sondages. (http://www.lexpress.fr/actualite/monde/amerique-nord/etats-unis-le-french-bashing-diversion-favorite-des-republicains-en-campagne_1731335.html)

Sur ce modèle, ont été créés de nombreux lexèmes, souvent avec des noms propres comme *N1* mais pas uniquement :

- (84) a. « Il y a un **Blagnac bashing** chez les militants PRG » (<http://www.lejournaltoulousain.fr/politique/interview-xxl-bernard-keller-il-y-a-un-blagnac-bashing-chez-les-militants-prg-43395>)
- b. La porte-parole du gouvernement, Valérie Pécresse, a en tout cas volé au secours de sa collègue, affirmant sur Europe1 qu'elle ne participerait pas au "**Morano bashing**", jeu qui fait fureur sur Internet et les réseaux sociaux. (http://www.huffingtonpost.fr/2012/02/08/morano-tacle-joly_n_1261914.html)
- c. Plutôt que de verser dans un « **service public bashing** » simpliste comme d'autres versent dans le « French bashing », innovons ensemble pour l'emploi! (<http://www.lesechos.fr/idees-debats/editos-analyses/0212702249767-pole-emploi-pour-que-cesse-le-service-public-bashing-1174286.php?sZjwT4QeuIToaLwT>.99)

- d. Pour remobiliser les troupes de gauche, les spin doctors du Président ont eu l'idée d'invoquer un **Hollande bashing**. (<http://www.lopinion.fr/17-juin-2013/hollande-bashing-est-fantasme-1147>)
- e. Après le référendum chez Smart et les violences à Air France, des voix se lèvent à droite pour contester la légitimité des syndicats, un "**syndicat bashing**" dangereux dénoncé par les centrales. (<http://www.la-croix.com/Actualite/France/Apres-les-violences-a-Air-France-les-syndicats-face-a-un-proces-en-illegitimite-2015-10-12-1367589>)

Comme pour les constructions *Xshow* la voyelle de liaison /ɔ/ peut apparaître :

- (85) a. Si tu étais une personnalité politique, tout le monde éclaterait de rire ou d'admiration devant ce geste d'auto-persuasion ou de langue de bois mêlé à un mouvement de **franco-bashing** sympathique. (<http://actu-politique.info/la-bc-achetera-des-dettes-etatiques-t7934-30.html>)
- b. Il y a eu un « sarkobashing », relayé depuis l'avenue de Ségur (siège de campagne de François Hollande), comme désormais il y a un « **hollandobashing** », encouragé du côté de l'UMP. (<http://revuecivique.eu/articles-et-entretiens/responsabilite-des-medias/francis-balle-le-reseau-des-reseaux/#.V1NE8cohWHk>)

Cet exemple montre donc que, si ces constructions affichent bel et bien un élément lexical en position finale, cet élément peut ne pas faire partie des lexèmes autonomes du français. Son interprétation est toutefois nominale.

Le dernier type de schémas de la Figure 5.3 sont les composés de type attributif, toujours avec l'élément recteur à droite. Ces exemples sont assez rares et isolés. Ils n'impliquent pas, contrairement aux précédents, d'éléments allogènes. Dans les premiers, c'est l'élément de droite, l'élément recteur donc, qui est récurrent :

- (86) a. (i) Vous ne savez pas qui voter ? Vous hésitez ? Ou encore vous êtes de gauche et vous avez envie de voter Bayrou ? Bref, vous êtes dans la merde. Une seule solution : lire Libération de ce lundi. Vous trouverez en page 36 dans la rubrique Rebonds, un texte de Laurent Joffrin, directeur de Libération sur la « **Bayrou-attitude** ». (<http://carz>)

- on.20minutes-blogs.fr/tag/vote)
- (ii) l'oreillette bluetooth est devenu le must-have de la **beauf-attitude** (http://www.acuite.fr/F3/forum_posts.asp?TID=27030&get=last)
 - (iii) La **Poulpe Attitude**, utiliser son intuition pour prendre les bonnes décisions (<http://journaldesgrandesecoles.com/la-poulpe-attitude-utiliser-son-intuition-pour-prendre-les-bonnes-decisions/>)
 - (iv) j'assume ma **boulet attitude!** ça vous est déjà arrivé de vous couper en vous rasant? (<http://gab-featcat.skyrock.com/68961402-j-assume-ma-boulet-attitude.html>)
- b. (i) Lambert Wilson "je ne suis pas **Hollywood compatible**" (<https://www.youtube.com/watch?v=W10ApCSru5k>)
 - (ii) François Baroin, un fidèle de Chirac devenu **sarko-compatible** (http://www.liberation.fr/france/2011/06/29/francois-baroin-un-fidele-de-chirac-devenu-sarko-compatible_746074)

Ce type de formation n'affiche pas non plus de préférence pour des N1 non natifs, ni ne met en avant de voyelle de liaison entre ses deux constituants. Une étude plus approfondie de ces exemples permettrait peut-être de faire ressortir les mots-leaders qui ont servi de modèles à ces formations et ont pu permettre la généralisation des schémas :

- (87) a. [[a]_{Nj} [attitude]_{Ni}]_{Nk} ↔ [attitude typique de SEM_j]_k
 b. [[a]_{Nj} [compatible]_{Ai}]_{Ak} ↔ [dont la qualité SEM_k est d'être compatible avec SEM_j]_k

Concernant le schéma (87-a), l'origine pourrait remonter au début des années 2000, alors que la chanteuse Lorie chantait la *Positive attitude*, expression reprise par le Premier ministre d'alors, Jean-Pierre Raffarin. Le modificateur du nom *attitude* était alors, comme attendu, un adjectif. Le passage à une construction [NN]_N disponible s'est-il opéré à la suite de ce buzz médiatique? Ce serait à définir mais la réponse à cette question dépasse les limites de cette thèse.

Toutefois, nous voyons que l'ordre dit néoclassique n'est pas limité aux éléments allogènes et ne nécessite pas forcément de marqueurs tels qu'une voyelle de liaison signifiant le sens de l'interprétation; cet exemple contredit donc la généralisation donnée par Amiot et Dal (2008) : « en français, l'ordre YX s'observe dans les composés dès lors que l'un des constituants qu'il met

en jeu ne coïncide pas avec une forme de lexème du français ». Pour comprendre les exemples (88) qu'elles citent dans cette généralisation, elles sont contraintes à « poser un radical en /i/ ou /ɔ/, dès lors qu'un lexème apparaît dans un composé à ordre néoclassique, et uniquement dans ce type de composés, très particuliers car marqués (journalisme politique, valeur ironique / péjorative) ».

(88) sarko-lepénisme ; Mitterando-nostalgie

La voyelle de liaison n'est donc pas nécessaire pour marquer un ordre inversé par rapport aux composés natifs, même si elle ne semble pouvoir apparaître que lorsque cet ordre inverse des constituants est adopté. Comme nous l'avons vu dans la section 3.1.2, la présence, dans l'idéal, du <o> dans les constructions néoclassiques est une contrainte qui pèse sur l'output de ces formations. Cette voyelle marque ainsi l'insertion de la construction dans un ensemble de constructions qui ont été créées dans un registre de langue qui se veut savant, même si c'est de manière humoristique ; les exemples (86) montrent donc que cette voyelle n'est pas liée à l'ordre des constituants.

Cette voyelle n'est pas non plus présente dans le dernier schéma de la Figure 5.3 :

(89) [[bébé]_{Nj} [b]_{Ni}]_{Nk} ↔ [petit SEM_i]_k

Cet exemple est proposé par Van Goethem et Amiot (2009) ou Grandi (2002). Pour Van Goethem et Amiot (2009), il n'appartient pas au lexique standard, puisque il est surtout utilisé par des enfants ou par des adultes avec une valeur affective :

- (90) a. bébé chien *vs.* chiot ; bébé chat *vs.* chaton
 b. bébé phoque *vs.* ?
 c. bébé plante *vs.* ?
 d. bébé maison *vs.* maisonnette

Van Goethem et Amiot (2009) analysent *bébé-* en (90-d) non plus comme un constituant de composé mais comme un préfixoïde, suite à une désémantisation, une décatégorisation et la constitution d'une série (je reviendrai sur ces critères de grammaticalisation juste après). Cependant, elles analysent toujours des éléments tels que *-fleuve* comme des éléments lexicaux.

Un autre exemple d'analyse d'un élément initial récurrent comme préfixe ou préfixoïde est donné par Villoing (2002) : *aide-*, qui peut avoir l'interprétation de 'assistant', par exemple dans *aide-infirmier*.

Qu'est-ce qui détermine donc le passage d'un élément lexical récurrent dans

un schéma constructionnel à un affixe ? Quels critères sont à retenir ?

Tous ces cas peuvent être vus comme des cas de grammaticalisation, qui peut être définie comme :

« An originally free word that has entered into a compound can serve as the basis for an entire pattern of like compounds. Once such a pattern takes hold and becomes productive, the original constituent may begin to deviate from its free equivalent in form or meaning and develop into an affix-like element » (Olsen, 2000, p.901-902)

ou comme :

« the change whereby in certain linguistic contexts speakers use parts of a construction with a grammatical function. Over time the resulting grammatical item may become grammatical by acquiring more grammatical functions and expanding its host-classes » (Brinton et Traugott, 2005, p.99).

Dans chacune de ces définitions, il s'agit du passage d'un statut lexical à un statut grammatical, étalé dans le temps, où un niveau intermédiaire serait que l'item lexical est utilisé comme un item grammatical ; dans ce niveau intermédiaire, l'élément en cours de grammaticalisation est souvent appelé *affixoïde*¹¹. La grammaticalisation pourrait être schématisée ainsi :

(91) lexème → affixoïde → affixe

Van Goethem et Amiot (2009) ne parlent pas de grammaticalisation mais d'affixisation, puisque la grammaticalisation consiste ici en l'émergence d'un affixe. J'utiliserai les deux termes indifféremment.

Comme nous venons de le voir dans la section précédente, dans le cadre de la morphologie des constructions (Booij, 2010), un lexème aussi bien qu'un affixe peut être l'exposant d'une construction. La principale différence qui existe entre les deux est que le lexème est stocké dans le lexique, il est donc accompagné d'indice lexical dans sa représentation à l'intérieur des schémas alors que l'affixe n'est pas stocké dans le lexique, il existe uniquement à l'intérieur de schémas constructionnels. Pour Booij (2005, p.117), lors de la grammaticalisation d'un élément, « semantic change precedes formal changes ». Un affixoïde a donc une forme proche de celle d'un lexème, sinon identique,

11. Dalton-Puffer et Plag (2000) refusent de postuler pour un stade intermédiaire de ce type. Pour eux, si un élément est qualifié de « semi-suffix », la nature du mot construit avec cet élément ne pourra pas être définie.

mais un sens de type affixal, c'est-à-dire qu'il fonctionne comme exposant de la construction : « morphemes which look like parts of compounds, and do occur as lexemes, but have a specific and more restricted meaning when used as part of a compound » (Booij, 2005, p.114). L'affixoïde est donc le stade intermédiaire, l'élément récurrent et spécifié lexicalement intervenant dans des sous-schémas, autrement dit l'élément spécifié d'un idiome constructionnel :

an affixoid is a lexeme that occurs in a subschema for compounds in which the other position is still a variable, that is, without a lexical specification. Such schemas are intermediate between concrete individual compounds and fully abstract schemes for compound structures. The specific and recurrent meaning of a lexeme in the compound structure is specified at this intermediate level (Booij, 2005, p.130).

Tout comme les affixes, les affixoïdes sont intégrés à un schéma ; cependant, ils sont encore reliés à un lexème, et ont donc toujours une catégorie lexicale, contrairement aux affixes. Dans la littérature, le terme de semi-affixe est aussi rencontré pour rendre compte de ce niveau intermédiaire, par exemple chez Marchand (1969, p.326) : « such elements as stand midway between full words and suffixes. Some of them are used only as second words of compounds, though their word character is still recognizable ».

Pour reprendre le schéma (91), selon Van Goethem et Amiot (2009) les trois stades de l'affixation sont les suivants : tout d'abord le lexème est employé dans un composé ; il devient ensuite récurrent, et constitue ainsi l'élément spécifié d'un idiome constructionnel. En tant qu'affixoïde il a toujours une catégorie lexicale, mais son sens est différent de celui du lexème autonome. Le dernier stade est donc celui de l'affixe, qui n'a pas de sens conceptuel, pas de catégorie lexicale, et qui permet la formation de séries de lexèmes plus ou moins longues. Pour Van Goethem et Amiot (2009), il entre également souvent en compétition avec d'autres affixes.

Pour résumer les critères d'affixation de Van Goethem et Amiot, pour qu'un lexème devienne un affixe, l'élément doit donc tout d'abord être désémantisé, par exemple par extension métonymique, métaphore ou réanalyse. Un sous-schéma doit être formé dans lequel il prend la place de l'élément spécifié, l'autre élément restant variable. Ce sous-schéma permet la constitution de plus ou moins longues séries de lexèmes. L'élément finit par être décatégorisé et entre en compétition avec d'autres affixes dérivationnels. Cette affixation est permise par une utilisation répétée d'un lexème comme constituant de composé. Warren (1990, p.123-124) par exemple, fait de cette utilisation répétée la cause d'une érosion phonétique et sémantique. Bauer (2005) dégage ainsi des exemples qui, en synchronie, peuvent illustrer différents stades. Le

premier stade est celui où des lexèmes sont employés de plus en plus souvent comme constituants de composés, et même parfois plus souvent dans des composés que de façon indépendante, et le sens quand il est lié ou indépendant peut être différent. On retrouve encore la notion de récurrence d'une formation possédant un élément spécifié :

(92) *idée-* dans *idée cuisine* ; *-clef* dans *mot clef*

Le deuxième stade est celui où un affixe peut être homophone ou orthographiquement identique à un lexème indépendant et qu'une relation sémantique peut être perçue entre les deux, même s'ils ne sont pas synonymes. Bauer (2005) reporte ici le lien qui existe entre le suffixe *-able* et le lexème ABLE, qui n'ont aucun lien diachronique mais qui ont subi une réinterprétation de la part des locuteurs :

(93) en. ABLE 'capable' vs. *Xable* 'pouvant être X' (*transportable, printable*)

Le dernier stade est celui où la relation sémantique entre le lexème et l'affixe n'est plus décelable, et l'élément est facilement classifiable comme affixe. Bauer (2005) donne l'exemple de l'élément *-wise* ; Dalton-Puffer et Plag (2000) voient quant à eux toujours un lien sémantique entre cet élément lié et l'élément autonome correspondant, mais cela ne les empêche pas de les classer parmi les suffixes, au même titre que des éléments comme *-ful* ou *-less*.

(94) en. *-wise*, dans *ressource-wise*

Selon Warren (1990, p.123), « affixes were originally independent words ». Même s'il faudrait vérifier avec exactitude la véracité de ces propos pour chacun des affixes existant, prenons tout de même les exemples des suffixes *-ifier* et *-ment* en français. Tous deux sont originellement des lexèmes latins. Le premier est issu du verbe latin FACIO qui, déjà en latin, était utilisé comme deuxième élément de constructions, sous la forme (à l'infinitif) *-ficio* :

(95) a. $[[x]_{Ni} \text{ ficio}_{Vj}]_k \leftrightarrow [\text{transformer en SEM}_i]_k$
 ➤ *deificare* < *deus* 'dieu'
 b. $[[x]_{Ai} \text{ ficio}_{Vj}]_k \leftrightarrow [\text{rendre SEM}_i]_k$
 ➤ *fortificare* < *fortis* 'fort'

Les deux schémas se retrouvent en français et sont disponibles pour la création de nouveaux lexèmes, mais le deuxième élément est alors considéré comme un suffixe.

- (96) a. $[[x]_{Ni} \text{ ifier}]_k \leftrightarrow [\text{transformer en SEM}_i]_k$
 > exemplifier < exemple (dat. 1810)
- b. $[[x]_{Ai} \text{ ifier}]_k \leftrightarrow [\text{rendre SEM}_i]_k$
 > densifier < dense (dat. 1896)

Selon les critères de Van Goethem et Amiot (2009), le verbe latin s'est décatégorisé, désémantisé (même si le sens n'est pas si éloigné d'un sens lexical, mais, pour reprendre la citation d'Aronoff (1984, p.49), « there is simply no way to exclude from derivational semantics types of meanings which are more typical of lexical categories like noun, verb, and adjective »), a donné lieu à de longues séries de mots, et s'est paradigmaté, devenant concurrent d'autres suffixes comme *-iser* (Lignon, 2013).

Le suffixe *-ment*¹² servant aujourd'hui à la construction d'adverbes sur des bases adjectivales était en latin un nom, MENS, signifiant 'esprit, disposition d'esprit'. La forme *mente* en est l'ablatif. Selon le TLFi, en latin, des constructions du type $[A_{abl.} + mente]_{abl.}$ formaient des compléments circonstanciels dans lequel le nom *mente* a peu à peu pris le sens de 'manière d'être', jusqu'à être considéré comme un suffixe. Le TLFi fait remonter la première utilisation de cet élément comme un suffixe à la seconde partie du VIII^{ème} siècle, dans le glossaire de Reichenau, qui traduit *singulariter* en langue vulgaire *solamente*. Ensuite, le suffixe connut l'essor décrit par Brunot :

Dès le latin de la décadence, un certain nombre d'adverbes se formaient par l'adjonction d'un nom latin *mente*, qui signifiait 'esprit' (...) l'analogie avait rapidement étendu ce type à des adjectifs auxquels le mot esprit ne s'appliquait plus. *Ment* équivalait là à 'façon', et il était devenu rapidement une sorte de suffixe capable d'enrichir sans cesse la langue de toutes sortes d'adverbes (Brunot, 1922, p.596).

Là encore, l'élément *mente* a subi une décatégorisation, une désémantisation (bien plus forte que ce que l'on avait vu avec *-ifier*) ; il est à présent un suffixe canonique, qui change par exemple la catégorie de la base des formations.

Ces exemples diachroniques illustrent donc que le passage d'un item lexical à un item grammatical est possible et même attesté à travers le temps. Certaines formations que j'ai abordées dans cette section se trouvent peut-être à différents stades d'affixation aujourd'hui et pourraient être, qui sait, un jour considérées comme de réelles affixations. La distance que j'ai prise ici avec les éléments néoclassiques m'a permis d'envisager nombre de formations

12. Pour une revue détaillée de ce suffixe, voir le premier numéro de la revue *Lexique* (1982) qui lui est consacré.

dont le statut des constituants est questionnable. Toutes ces formations impliquaient cependant des éléments qui peuvent par ailleurs être autonomes syntaxiquement, ce qui n'est pas le cas, à quelques exceptions, des éléments néoclassiques. Comment des unités non autonomes peuvent participer à la formation de lexèmes ? Avant de voir, dans la dernière partie de ce chapitre, comment on peut rendre compte du statut des éléments néoclassiques, je vais d'abord observer l'utilisation de formes lexicales non autonomes dans la formation de lexèmes.

5.2.2.2 Les lexèmes tronqués

Si je me suis intéressée dans la section précédente aux lexèmes qui entraient tout entiers dans des sous-schémas, d'autres lexèmes ne le sont qu'après une troncation de leur représentation phonologique.

Abréviations. Cette troncation peut être seulement une abréviation, c'est-à-dire que si l'intégrité phonologique du lexème est atteinte, le sens est conservé intact. Ce sont les éléments que Corbin (2001, p.44) nomme *fracto-constituants*, « les représentants tronqués et lexicalisés comme tels d'unités lexicales françaises », ou que Tournier (1985) nomme *fractomorphèmes*. Ce sont par exemple les éléments *auto-*, *ciné-*, *euro-* ou *télé-* qui entrent dans les schémas décrits en (58). Certains peuvent être employés de manière autonome alors que pour d'autres c'est impossible, du moins avec le même sens. *Euro* utilisé de manière autonome ne désigne pas une réalité géographique mais une monnaie ou un tournoi sportif. Un élément comme *pétro-* (< PÉTROLE) n'est quant à lui jamais rencontré de façon autonome. Contrairement aux cas de sécrétion que j'aborderai juste après, les bases des phénomènes d'abréviation ne subissent aucune modification sémantique. Ces abréviations interviennent dans la formation de lexèmes dont l'élément hyperonyme du tout est en général l'élément de droite, autrement dit des constructions affichant un ordre néoclassique :

- (97) auto-école, euro-sceptique, pétrodollar, ciné-club, télé-achat, écobilan (< ÉCOLOGIE), sociolinguistique (< SOCIÉTÉ), photocomposition (< PHOTOGRAPHIE)

Parmi ces éléments abrégés, Iacobini (2004, p.73) distingue ceux qui correspondent à un EN dans le lexème d'origine (*éco-* et *photo-* par exemple étaient des ENs dans *écologie* et *photographie*), et ceux qui sont des raccourcissements ne correspondant pas, dans le lexème à un élément séparable. On peut également remarquer que, pour *socio-* par exemple, l'élément abrégé a

non seulement subi une troncation, mais également l'ajout d'un <o>. Même si cette question devra être étoffée à l'aide d'un corpus plus fourni, ces éléments abrégés semblent correspondre à un gabarit phonologique : deux syllabes, un segment final /ɔ/ voire /e/. Par ailleurs, ce sont des éléments utilisés en position initiale dans les mots construits. Leur gabarit phonologique les fait se rapprocher des ENs, comme le fait remarquer Iacobini (2004). On les trouve d'ailleurs utilisés comme C1 dans des lexèmes de NéoClassy :

(98) autophobe, cinéphage, télévore, eurocrate, sociologie, photovore

Certains sont d'ailleurs homophones d'ENs :

(99) télé- < TÉLÉVISION (ex. téléphage) vs. télé- 'distance' (ex. télécommunication)
photo- (< PHOTOGRAPHIE (ex. photocomposition) vs. photo- 'lumière' (ex. photophobie)

Cependant, concernant le lexème PÉTROLE, quand il est C1 d'une CN, il n'affiche pas le radical *pétro-* :

(100) pétrolivore, pétrolophage

Une explication pourrait être la confusion qu'il peut y avoir entre *pétro-1* (< PÉTROLE) et *pétro-2* (< PIERRE), le deuxième étant celui que l'on retrouve dans *pétrologie*. Une autre explication pourrait être le respect de la contrainte de taille des CNs qui demande en général des bases trisyllabiques.

Fradin (2000) distingue les formes abrégées (abbreviated forms) et les formes amputées (amputated forms) : ces dernières fonctionnent comme des lexèmes complets (*bibli* pour *bibliothèque*) et donc utilisables de manière autonome, et ont des usages pragmatiques particuliers ainsi que des connotations sémantiques. Certaines formes abrégées et formes amputées peuvent correspondre (*télé*, *photo*, ...) mais ce n'est pas forcément le cas : la forme abrégée *socio-* représente SOCIÉTÉ dans les mots construits, la forme amputée *socio* est un raccourcissement du lexème SOCIOLOGIE, bien que celui-ci soit formé en employant la forme abrégée *socio-*.

Pour moi, suivant l'analyse d'Amiot et Dal (2007), ces éléments sont des radicaux de lexèmes, au sens de Roché (2010) (revoir la section 5.1.1.1 de ce chapitre) : ils correspondent donc à la représentation phonologique de lexèmes. Aucun changement d'ordre sémantique ou syntaxique n'a été opéré. Ces radicaux sont formés pour répondre à des gabarits phonologiques. Ainsi, il n'y a aucune différence significative entre le lexème *xénophobe* et le lexème *europhobe*. Dans le premier cas, comme on l'a vu dans la section 5.1.1.1, *xén-*

est le thème S du lexème ÉTRANGER ; la construction *Xphobe* sélectionne, quand il existe et qu'il est disponible, le thème S du lexème X. La construction du radical de *xénophobe* se fait donc par la sélection du thème S. La présence de la voyelle <o> se joue au niveau de la construction dont la rubrique phonologique requiert de préférence un /ɔ/. Concernant *europphobe*, la construction du radical sélectionne l'unique thème du lexème EUROPE mais opère une troncation afin que le lexème réponde aux contraintes phonologiques pesant sur la base des constructions *Xphobe*, faisant ainsi apparaître la voyelle <o> avant l'EN final et une base dissyllabique.

Mots-Valises. Les mots-valises, ces formations que Aronoff (1976, p.20) nomme *oddities* en raison de leur opacité, ont fait l'objet de mon mémoire de Master 2 (Lasserre, 2011). Ces formations, décrites, pour le français, entre autres par Grésillon (1984); Fradin (1997); Fradin *et al.* (2009); Leturgie (2012), mettent en jeu deux lexèmes qui, par un enchevêtrement de leurs formes, permettent un enchevêtrement de leurs sens ; ils se rapprochent ainsi des composés en ce que leur sens n'est pas prédictible dans la construction :

- (101) alicament (< ALIMENT ; MÉDICAMENT) 'aliment servant de médicament'

Les fragments de lexèmes se retrouvant dans le mot-valise sont souvent appelés dans la littérature anglophone *splinters*, par exemple chez Bauer (2005). Fradin *et al.* (2009) parlent quant à eux d'*éléments d'amalgame*. Si Grésillon les appelle des monstres et Aronoff des « oddities », c'est parce qu'aucune généralité ne peut être faite concernant ces mots. Les schémas de construction de mots de Booij illustrent les généralités qui peuvent être opérées dans la langue : les mots-valises ne peuvent donc rentrer dans aucun d'eux. En effet, tant du point de vue de la forme que du sens, les mots-valises sont dépendants des lexèmes à partir desquels ils sont formés. D'un point de vue phonologique, le chevauchement des lexèmes-bases, quand il a lieu, dépend des séquences que partagent ces lexèmes. De plus, pour faciliter cet enchevêtrement, chacun des lexèmes peut être tronqué, ou aucun, ou les deux. Dans Lasserre (2011), j'avais dégagé sept types de mots-valises, selon ces deux critères. Ils sont reportés dans le Tableau 5.12¹³.

Certaines contraintes phonologiques sont respectées, comme la contrainte de fidélité positionnelle à l'intérieur de la syllabe qui stipule que les segments conservent dans le mot-valise la même position syllabique que celle qu'ils

13. mécontemporain < MÉCONTENT, CONTEMPORAIN ; célibattante < CÉLIBATAIRE, BATTANTE ; caméscope < CAMÉRA, MAGNÉSCOPE ; ahurifié < AHURI, HORRIFIÉ ; mocallerines < MOCASSINS, BALLERINES ; apéricube < APÉRITIF, CUBE ; fruitillante < FRUIT, PÉTILLANTE

	Séquence commune	Troncation Lexème 1	Troncation Lexème 2	Exemple
(i)	✓	-	-	mécontemporain
(ii)	✓	✓	-	célibattante
(iii)	-	✓	✓	caméscope
(iv)	✓	-	✓	ahurifié
(v)	✓	✓	✓	mocallerines
(vi)	-	✓	-	apéricube
(vii)	-	-	✓	fruitillante

Tableau 5.12 – Types de mots-valises

avaient dans le lexème base. Cependant, ces contraintes dépendent encore une fois des lexèmes entrant en jeu et ne peuvent pas être généralisées.

D'un point de vue sémantique, Grésillon (1984) analyse dans chaque mot-valise une relation de co-prédication, c'est à dire que le mot-valise est à la fois le lexème-base1, à la fois le lexème-base2; un mot-valise est « le fait d'énoncer deux choses en même temps, d'effectuer deux prédications absolument simultanées, l'une dans l'autre, l'autre dans l'une ». Le sens des lexèmes-bases s'imbrique donc autant que leur forme. Mais cette imbrication se fait de différentes manières, là encore non prévisibles; il n'y a par exemple aucune correspondance entre un type de mot-valise parmi ceux exposés dans le Tableau 5.12 et la façon dont le sens se construit. Les interprétations des mots-valises ne peuvent donc être faites qu'en contexte. Et toutes les interprétations sont envisageables : « toutes les combinaisons des sens des lexèmes-bases sont permises, pourvu qu'elles soient possibles » (Fradin *et al.*, 2009, p.39). Dans Lasserre (2011), j'ai par exemple relevé les interprétations suivantes :

- Interprétations coordonnées :
 - interprétation contrastive : *délicieuse* 'substance délicieuse mais vicieuse'
 - interprétation nulle : *enfant* 'ni enfant ni adulte'
 - interprétation équivalente : *prostituée* 'prostituée'
 - interprétation médiane : *gris* 'entre le gris et le beige'
 - interprétation balancée : *métro* 'le métro ou le trolleybus'
- interprétations non coordonnées :

- interprétation équative¹⁴ : *alicament* ‘un aliment qui est un médicament’
- interprétation attributive : *célibattante* ‘célibataire battante’
- interprétation locative : *angoulag* ‘goulag en Angola’
- interprétation causale : *mécontemporain* ‘mécontent au sujet des choses contemporaines’

De plus, si dans la plupart des mots-valises les lexèmes bases importent leur sens dans son intégrité, certains, à la manière des constructions du type *Xfleuve*, ne contribuent au sens du mot-valise que par une sélection de certains de leurs traits sémantiques. Tout comme l’élément modifieur d’un composé de type attributif, ce lexème-base est utilisé de façon non référentielle dans le mot-valise :

- (102) *ordinosaure* (< ORDINATEUR, DINOSAURE) ‘vieil ordinateur’
 infobésité (< INFORMATION, OBÉSITÉ) ‘abondance d’informations’

Ce type de mot-valise semble être marginal. On peut toutefois remarquer que, comme pour les lexèmes du type *Xfleuve*, c’est ici l’élément de droite qui se voit amputé d’une partie de son sens.

Pour résumer, le fait qu’il n’y ait pas de rapport constant entre forme et sens dans les mots-valises empêche de les considérer comme des composés canoniques, même s’ils sont des mots complexes mettant en jeu deux lexèmes. Le fait que ces formations soient des jeux de mots est pour Fradin *et al.* (2009, p.34) une preuve de leur extragrammaticalité. Cependant, de cette extragrammaticalité même peut naître une liste de lexèmes dont la construction peut être généralisée au sein d’un schéma. En effet, lorsqu’un splinter devient récurrent, son sens peut être spécifié à l’intérieur d’un schéma constructionnel. Prenons les exemples (103) :

- (103) *risotto* : riz cuit par une absorption progressive de liquide bouillant
 coquisotto : coquillettes cuites comme un risotto
 lentisotto : lentilles cuites comme un risotto
 quinotto : quinoa cuit comme un risotto

Le lexème RISOTTO n’est pas un lexème construit, du moins il n’est pas perçu comme tel en français. Même si chacun des mots-valises en (103) peut être considéré comme un mot-valise pour lequel, comme c’était le cas en (102), l’élément de droite ne fournit que quelques uns de ses traits sémantiques, l’existence de cette série s’oppose à l’appellation de mot-valise : en effet, « la

14. Selon Fradin *et al.* (2009).

fabrication des mots-valises ne peut être un procédé productif, puisque par définition chaque mot-valise est créé au coup par coup, sans suivre un patron quelconque » (Fradin *et al.*, 2009, p.34). Or dès lors qu'une liste se forme dans laquelle un élément contribue sémantiquement de la même manière, un schéma peut être envisagé :

$$(104) \quad [[x]_{Ni} (s)otto]_{Nj} \leftrightarrow [SEM_i \text{ cuit par une absorption progressive de liquide bouillant}]_j$$

Pour Bauer (2005, p.104), un splinter est d'ailleurs « a fragment of a word used repetitively in the formation of new words », qui apparaît par le biais du mot-valisage.

L'emploi de splinters de manière récurrente et avec un sens restreint rapproche ces formations des affixes sécrétifs.

Sécrétion. Le phénomène de sécrétion est différent de celui de l'abréviation en ce qu'une sélection dans les traits sémantiques du lexème tronqué est opérée. Pour Fradin (2000, p.35)¹⁵, la sécrétion repose sur deux critères : (i) le processus fondamental consiste en une abstraction (ii) une sélection est ensuite effectuée dans les prédicats du lexème modèle, dont certains seront abandonnés.

Prenons l'exemple du schéma *Xgate*, décrit par exemple par Warren (1990) ou Fradin (2000) pour l'anglais, mais qui est également disponible en français, notamment dans le domaine journalistique. Ce schéma a un mot leader clairement défini, *Watergate*. Ce mot, par métonymie, fait référence au scandale politique qui a eu son origine dans l'immeuble du Watergate. Les exemples (105) ont été créés par analogie avec ce mot :

- (105) a. « **Chevalgate** » : Le scandale s'étend aux boîtes de conserve (<http://www.20minutes.fr/economie/1123753-20130322\discrétionary{-}{-}{-}20130322-chevalgate-scandale-setend-boites-conserve>)
- b. Pas de répit dans le « **lasagnegate** ». Après l'Irlande, le Royaume-Uni, la France et la Pologne, la Suisse est touchée par le scandale de la viande chevaline utilisée dans des lasagnes mais frau-

15. Fradin (2000) distingue deux types de sécrétion : la sécrétion standard dont le schéma *Xgate* est un exemple, et la sécrétion complexe, dans laquelle non pas une fonction est dégagée mais une famille de fonctions : X dans *Xstroïka* peut ainsi désigner l'instigateur du renouveau, le lieu où le renouveau a lieu ou le domaine dans lequel le renouveau a lieu. Je ne m'attarderai pas ici sur ces différences, voulant uniquement décrire le processus d'affixation dont la sécrétion est un moyen.

- duleusement étiquetée comme du bœuf. (<http://www.letemps.ch/economie/2013/02/13/suisse-touchee-lasagnagate>)
- c. La Guerre des étoiles a bien eu lieu. Ce mardi, un événement notable dans l'histoire de Twitter a bien failli « casser l'Internet » : le « **FAVgate** ». Soit le jour où la petite étoile céda sa place au petit cœur, et où le « fav » dut s'incliner face au « like ». (<http://www.20minutes.fr/web/1723475-20151104-favgate-nuls>)
 - d. Sept candidatures pour une présidence, le destin de Platini toujours en suspens : le **FIFAGATE** n'a pas fini d'agiter la planète foot. (<http://www.francefootball.fr/news/Le-fifagate-a-tout-chamboule/603143>)
 - e. Landes : une agression à la pelle et un "**slipgate**" viral. Des riverains à Audon (Landes) ont vivement pris à partie des membres de la Ligue de Protection des oiseaux (LPO) venus dénoncer le braconnage des pinsons. L'image d'un riverain en slip et armé d'une pelle a été détournée pour le plus grand plaisir des internautes. (<http://www.leprogres.fr/france-monde/2015/11/09/des-militants-de-la-lpo-agresses-a-coups-de-pelle>)
 - f. Le **Shy'mgate** : Les mecs du RAID ils ont crié "TOUS À TERRE", Shy'm¹⁶ elle a pas cherché à comprendre. (<http://twog.fr/le-comptwoir-du-19-novembre-2015/>)

Comment l'analogie s'est-elle mise en place ? Contrairement aux cas d'abréviations que j'ai abordés précédemment, la partie tronquée de *Watergate* présente dans *chevalgate* ne représente pas le lexème dans son intégrité sémantique : seul le trait sémantique 'scandale' est retenu ; cela répond au critère (ii) de Fradin (2000). Cependant, on peut remarquer qu'il y a également eu un glissement sémantique d'ordre métonymique pour le lexème autonome : *Watergate* est en effet passé du sens de 'hôtel situé à Washington' au sens de 'scandale provoqué par une écoute téléphonique', dans des exemples tels que :

- (106) C'est un peu le **Watergate** de Sainte-Maxime. Le maire cette commune du Var, Vincent Morisse, a déposé plainte contre X après avoir trouvé un micro-espion dans son bureau. (<http://www.rtl.fr/actu/politique/qui-en-veut-au-maire-de-sainte-maxime-dans-le-var-2298512>)

16. Shy'm est une chanteuse française à la mode, qui a eu la malheureuse expérience de chuter en se jetant dans le public lors d'un concert.

Cependant, les exemples comprenant le lexème autonome sont bien plus limités sémantiquement que ceux présentant la forme apocopée. Le schéma permettant la généralisation des exemples (105) est le suivant :

$$(107) \quad [[x]_{Ni} \text{ gate}]_{Nj} \leftrightarrow [\text{scandale impliquant SEM}_i]_j$$

Le critère (i) de Fradin (2000), celui de l'abstraction qui « transforms an expression into a function », est appuyé dans notre cadre théorique par le fait que l'élément *-gate* est ici l'exposant de la construction, la manifestation formelle de l'opération sémantique. De plus, l'élément *a* ici perdu ses propriétés lexématiques, il est par exemple dépourvu de catégorie lexicale et par conséquent d'indice lexical. On remarque que dans le schéma (107) le sens de 'scandale' est à prendre dans un sens très large, puisque si *chevalgate* évoque bien une « grave affaire qui émeut l'opinion publique, à la fois par son caractère immoral et par la personnalité des gens qui y sont compromis » (GR), *Shy'mgate* n'a pas ce degré de gravité. Ce dernier exemple montre peut-être que la désémantisation du schéma n'est pas encore achevée et que le sens de 'politique' peut s'affaiblir.

On voit donc qu'ici la notion de mot leader est bien illustrée : la formation de ce schéma étant contemporaine, puisque le scandale du Watergate a eu lieu en 1972, le mot leader est facilement identifiable et même certainement bien présent dans l'esprit des locuteurs, du moins certains d'entre eux, qui créent un mot *Xgate*. De plus, comme le fait remarquer Warren (1990), les constructions impliquant un affixe sécrétif sont des co-hyponymes du mot leader, et non pas des hyponymes de leur base, ce qui les différencie par exemple des composés tels que nous les avons abordés précédemment. Par ailleurs, comme le dit Fradin (2000, p.37), puisque *-gate* a un sens stable, une représentation phonologique fixe et que sémantiquement, c'est une fonction (nous dirons ici exposant de construction), cet élément est affixal ; il ne présente pas de différence majeure avec des suffixes plus prototypiques. Cependant, l'intégrité de la base à laquelle s'attache *-gate* est telle qu'on est en présence de séquences phonologiques peu fréquentes en français, par exemple /ɲg/ ou /pg/ ; il est probable qu'une affixation prototypique éviterait ce genre de séquence. Comme on l'a vu précédemment, un suffixe français affiche en général une voyelle à l'initiale permettant une resyllabation avec le thème à finale consonnantique.

Regardons maintenant un cas de sécrétion un peu différent. Tout comme FLEUVE dans les schémas *Xfleuve* ou ÉCLAIR dans *Xéclair*, le lexème MATHON 'discipline de l'athlétisme consistant en une course à pied longue de 42,195 kilomètres' a donné lieu à un sous-schéma de composés [NN]_N dans lequel certains traits sémantiques du lexème ont été sélectionnés. Le

sous-schéma (108) est l'opposé du sous-schéma que nous avons vu en (75-c), comme l'utilisation du même N à l'initiale, visite, en (75-c) et (109-c) le suggère.

(108) [[a]_{Ni} [marathon]_{Nj}]_{Nk} ↔ [SEM_i long/dense/chargé]_k

- (109) a. Et c'est le Croate, tête de série numéro 9, qui a eu le dernier mot en dominant l'Américain au terme d'un **match marathon** de cinq sets, après plus de 4h30' de jeu (<https://fr.sports.yahoo.com/news/cilic-remporte-un-match-marathon-contre-isner-141432204--ten.html>)
- b. Le ministre de l'Agriculture, de l'Agroalimentaire et de la Forêt, était en visite à Dijon ce vendredi. Une **journée marathon** où Stéphane Le Foll a rencontré pour un petit-déjeuner et une réunion en préfecture, les organisations syndicales agricoles et consulaires. (<http://www.bienpublic.com/politique/2015/10/30/dijon-la-folle-journee-dijonnaise-de-stephane-le-foll>)
- c. François Hollande sera resté en tout huit heures au Salon de l'Agriculture. Une **visite marathon** placée sous le signe de la sécurité en raison du plan Vigipirate renforcé. (http://www.francetvinfo.fr/societe/salon-de-l-agriculture/francois-hollande-en-visite-marathon-au-salon-de-l-agriculture_831059.html)

MARATHON peut également être utilisé avec ce sens restreint de façon autonome :

- (110) Les négociations ont tourné au marathon ; le marathon budgétaire (exemples tirés du GR)

C'est certainement la possibilité d'utiliser ce lexème avec un sens métaphorique qui a permis la création du mot-valise *téléthon*, contraction de TÉLÉVISION et MARATHON et qui désigne une émission de télévision se déroulant sur plusieurs jours et qui a pour but de récolter des fonds pour des associations caritatives. Sur le modèle *téléthon*, de la même façon qu'un affixe a été secrété grâce à une sélection de traits sémantiques et une abstraction, un sous-schéma a été généré, dans lequel le premier constituant désigne une activité :

- (111) [[x]_{Ni} thon]_{Nj} ↔ [événement de vaste envergure ayant pour but de rassembler des fonds pour des associations caritatives grâce à / au

moyen de SEM_i]j

Ici, la longueur de l'opération n'est plus un critère, mais plutôt, dans les exemples suivants, ce qui est mis en avant est (i) une opération de vaste envergure, (ii) une collecte de fonds à but caritatif. De plus, le trait sémantique 'télévision' présent dans TÉLÉTHON a été abandonné : nous sommes bien dans un cas de sécrétion.

- (112)
- a. La 3e édition du ' **Blogathon** ' se déroulera demain samedi 26 juillet. Plusieurs centaines de blogueurs écriront pour soutenir des associations caritatives. (<http://www.01net.com/actualites/blogathon-24-heures-de-blog-pour-les-bonnes-causes-2127058.html>)
 - b. L'**internethon** est un programme lancé par l'Association Francophone des utilisateurs d'Internet (AFNeT1) dont le but est la réutilisation des ordinateurs normalement destinés au recyclage. (<https://fr.wikipedia.org/wiki/Internethon>)
 - c. Le **Masturbathon** vise à financer un Centre pour le sexe et la culture, une association à but non lucratif oeuvrant en faveur d'une meilleure éducation sexuelle (<http://funnynews.free.fr/actus.php?n=720>)
 - d. Au seuil de l'hiver, la marque de smoothies (boisson aux fruits super frais) innocent lance pour la deuxième fois une vaste opération, un « **tricothon** », pour aider les Petits Frères des Pauvres. (http://www.aiguille-en-fete.com/Tricothon-pour-les-Petits-Freres-des-Pauvres-_a342.html)
 - e. Le moment est venu de se préparer pour notre **Basketthon** - saison 2015/2016 et surtout de vous permettre de faire exploser vos scores personnels, alors tous et toutes à vos ballons ! Cet événement, important dans la vie de notre club, connaît, grâce à votre participation active, un succès grandissant. Il nous permet également de soutenir l'Association « Enfants des Collines ». (<http://www.del-basket.ch/del2013/node/13>)
 - f. Le **Cyclothon**, c'est un grand rassemblement fraternel ayant pour objectif principal de récolter des fonds pour permettre aux personnes que nous accompagnons de partir en vacances. (https://www.petitsfreresdespauvres.fr/nos-implantations/527000-01-fraternite-de-lille/evt_20eme-cyclothon,-inscrivez-vous-!.html)

Le schéma (111) s'est également étendu : dans les exemples (113), la base dé-

signe le destinataire de la collecte de fonds, et la notion de collecte caritative a disparu, ou du moins demeure de façon ironique.

- (113) a. Ce pauvre Cyril, complètement déconnecté de la réalité, nous ferait presque pleurer. A quand un **Hanounathon** pour l'aider à terminer ses fins de mois? (<http://www.fn42.fr/article-cyril-hanouna-je-gagne-25-000-mois-par-rapport-a-d-autres-c-est-tres-peu-121417974.html>)
- b. Allez, sèche tes larmes, Serge Dassault, c'est pas demain la veille qu'on verra un **Dassaulthon** sur France 2 à Noël. Et c'est à Hervé Morin que tu le dois. (<http://enattendanth5n1.20minutes-blogs.fr/tag/armement>)

Ces exemples pourraient être schématisés ainsi :

- (114) [[x]_{Ni} thon]_{Nj} ↔ [événement de vaste envergure ayant pour but de rassembler des fonds destinés à SEM_i]_j

En outre, les exemples (115) n'affiche plus le sens de 'récolte de fonds', mais uniquement celui de 'événement de vaste envergure'.

- (115) a. **Bovéthon** : déjà 4 300 signatures ; La pétition pour la candidature de José Bové à la présidentielle avait recueilli hier 4 300 signatures, dont celle de Jean Ferrat, chanteur proche du PCF. (http://www.liberation.fr/france/2007/01/10/bovethon-deja-4-300-signatures_81512)
- b. Au moment du Grenelle de l'Environnement, qui doit nous livrer fin octobre son plan d'action en faveur de l'écologie, du développement et de l'aménagement durables, Agir Pour l'Environnement lance le **Grenelléthon**. Il s'agit de 40 mesures pour agir concrètement, chacune assortie d'une fiche détaillée sur l'état des lieux et l'action à mener, et de cartes-réponses à envoyer au Président de la République, au ministre de l'Écologie, du Développement et de l'Aménagement Durables ainsi qu'à la secrétaire d'État à l'Écologie. (<http://www.ouverture-facile.net/environnement/article/participez-au-grenellethon>)

Ici, le rôle de la base dans le procès ne semble pas déterminé. Le schéma ayant permis la formation de *bovéthon* et *Grenelléthon* pourrait être :

- (116) [[x]_{Ni} thon]_{Nj} ↔ [événement de vaste envergure ayant un lien avec

SEM_i]j

Cependant, on pourrait toutefois considérer que, dans le mot *Grenelléthon*, la présence du /e/, qui n'apparaît pas dans la base, provient de *téléthon* et que nous sommes donc en présence d'un mot-valise; dans *Bovéthon*, c'est plus difficile à déceler car le /e/ peut être attribué au constituant initial, *Bové*.

Dans les trois schémas (111), (114) et (116), le mot leader à partir duquel la sécrétion a débuté est bien *téléthon* et non *marathon*. La désémantisation qui va de pair avec l'abstraction est de plus en plus avancée entre le schéma (111) et le schéma (116).

Les affixes sécrétifs sont à différencier des éléments tels que *-ciel* qui représente le lexème LOGICIEL dans les formations (117) :

(117) ludiciel, didacticiel

Dans ce cas-là, aucune abstraction n'a été faite, ni aucune sélection sémantique. En fait, ces formations seraient à rapprocher des mots-valises, mais la formation d'une série qu'il est possible de généraliser semble l'interdire. Fradin (2000) parle de composition cachée.

5.2.2.3 Aux frontières des éléments néoclassiques

Pour finir cet inventaire d'éléments oscillant entre lexèmes et affixes, revenons vers les éléments qui nous occupent particulièrement dans ce travail : les éléments néoclassiques finaux rentrant dans des schémas disponibles pour les locuteurs du français et générateurs de longues listes de lexèmes. Mis à part leurs origines allogènes, en quoi sont-ils différents, synchroniquement, des éléments j'ai abordés dans cette partie 5.2.2 ?

Regardons par exemple l'élément *-cratie*. Dans le chapitre 3, nous avons vu que cet élément rentrait dans le schéma que nous reportons ici, et que ce schéma a été formé sur la base du mot leader *démocratie* :

(118) [X_i cratie]_{Nj} ↔ [pouvoir exercé par SEM_i]j

Cependant, dans l'exemple (119), *dictocratie* doit-il être analysé comme instanciant le schéma (118) ou comme étant un mot-valise formé de la contraction de DÉMOCRATIE et de DICTATURE ?

(119) La **dictocratie** (ou démocrature) désigne un régime qui, sous l'apparence d'une démocratie, fonctionne en réalité comme une dicta-

ture. (http://mobile.avocats.fr/space/laurent.epailly/content/_398F98FC-8BCF-40F0-A09B-F6C543F5B8D3/comment/1)

Ici, la réponse se trouve dans l'exemple même puisque l'auteur propose comme formation alternative à *dictocratie* la formation opposée, *démocrature*. Cependant, la question peut se poser dans d'autres exemples :

- (120) le marketing politique en ligne en est encore à ses tout premiers balbutiements, ne donnant ainsi qu'une faible épaisseur à la « **clico-cratie** » — cette démocratie numérique qui a fait les beaux jours de la campagne de Barack Obama — et dont les fruits au Québec ne seraient finalement pas encore mûrs. (<http://web1.ledevoir.com/societe/actualites-en-societe/267749/1-info-a-sens-unique-est-un-veritable-cul-de-sac>)

Dans l'exemple (120), *clico-cratie* instancie-t-il le schéma (118) ou est-il un mot-valise? C'est plus difficile à déterminer. Dans la première solution, le sens de 'démocratie' ne doit pas être détecté dans le mot *clico-cratie*, uniquement celui de 'pouvoir', contenu dans le schéma. Or, ce schéma a justement été généré suite aux analogies opérés sur le lexème DÉMOCRATIE. Pour Marchand (1969, p.212), *-ocracy*, en anglais, a été obtenu par sécrétion : « a certain structural element of one linguistic system is isolated and introduced into another linguistic system. The speaker with a knowledge of Greek isolates *-ocracy* 'rule' in a series of Greek-coined words and introduces it as a derivative element into a structural system of English ». En effet, on peut considérer qu'un affixe sécrétif, *-(o)cratie*, a été obtenu sur le mot leader DÉMOCRATIE ; on a en effet tous les ingrédients de la sécrétion : une partie tronquée de cet élément, *-ocratie*, une sélection sémantique 'pouvoir exercé' avec l'abandon de certaines traits ('peuple') et une abstraction, l'emploi de l'élément comme exposant de construction. Cette solution est envisageable puisque les locuteurs aujourd'hui, dans la créativité lexicale, ne manipulent pas du grec ancien mais bien du français. Cette analyse de *-(o)cratie* peut poser problème puisque, dans *démocratie*, le sens de l'élément originaire du grec, 'pouvoir' a non seulement le même sens que celui qu'il avait dans cette langue mais a également certainement été appris aux enfants des écoles, quand on aborde la notion de démocratie. Cependant, elle a l'air d'être plus soutenable pour un élément comme *-thèque*. Cet élément, que j'ai déjà abordé plus haut dans ce chapitre (67), s'il est également issu du grec ancien, n'est plus transparent pour les locuteurs actuels, du moins ceux n'ayant pas de connaissance étymologique. Il apparaît surtout dans le lexème BIBLIOTHÈQUE dont sont issus les exemples que j'ai donnés en (66). L'élément *-thèque*, même s'il dé-

signait à l'origine un élément signifiant, est devenu, par une abstraction, un exposant de construction ; une sélection sémantique s'est opérée également, ce qui m'amène à considérer cet élément comme un affixe sécrétif, au même titre que *-gate*.

Ces questionnements m'amènent à penser qu'il y a une place pour les éléments néoclassiques au sein du modèle qui comprend chacun des éléments que je viens de citer. Il reste maintenant à savoir sur quels critères s'appuyer pour rapprocher ou différencier ces éléments finaux afin d'élaborer une typologie qui rendrait compte de chacun.

5.3 Quelle place pour les éléments néoclassiques ?

Après cette revue de formations présentant des caractéristiques éloignées des affixés ou composés canoniques, nous voyons que le cadre théorique de la morphologie des constructions (Booij, 2010), permet d'accueillir ces formations : il permet de considérer par exemple une unité lexicale comme un exposant de construction, sans pour autant rompre le lien qu'il peut exister entre les différentes manifestations de cette unité quand elle est attachée et quand elle est autonome. Les éléments néoclassiques tels que je les ai abordés tout au long de ce travail oscillent entre unité lexicale et exposant de construction. L'EN *-logue* par exemple entre en concurrence avec un suffixe comme *-iste* et présente des caractéristiques proches de celles des affixes. Et si dans *hydrologue*, *-logue* est un élément de type affixal, *hydr-* constitue la base de la construction, et est par conséquent un élément lexical : il s'agit en effet du thème S du lexème EAU. Une analyse unifiée de tous les éléments néoclassiques est donc illusoire. Si une étude diachronique peut montrer une parenté entre chacun (quoique certains, dès les langues classiques, avaient déjà le comportement d'exposants de construction), une étude synchronique ne peut pas en rendre compte uniformément. Même à l'intérieur du groupe d'EN finaux fortement disponibles pour la formation de lexèmes que j'étudie dans cette thèse, une seule et même analyse ne peut pas rendre compte de tous les comportements.

Cette dernière section a pour finalité l'élaboration d'une typologie des éléments finaux spécifiés dans des schémas ou des sous-schémas. Elle reprend en l'affinant la typologie proposée dans Lasserre et Montermini (2014b). Avant de présenter cette typologie, je vais d'abord énumérer l'ensemble des critères sur lesquels je me suis appuyée.

5.3.1 Critères de classification

La liste des critères présentée ici se divise en deux parties. Tout d'abord, je listerai les critères permettant de caractériser les éléments spécifiés dans la position de droite des constructions que j'ai abordées dans cette partie ainsi que plus généralement dans cette thèse. Ensuite, les constructions en elles-mêmes seront observées en fonction d'autres critères. Chacune de ces listes de critères sera récapitulée dans un tableau en fin de section.

5.3.1.1 Les éléments finaux

Hétérolexicalité. Le critère de l'hétérolexicalité est proposé par Warren (1990). Il est difficilement discriminant, étant donné que de nombreux lexèmes du français sont bien sûr empruntés au grec, au latin, ou à d'autres langues. Par exemple, faut-il considérer *thérapie* comme étant hétérolexical ? Ce critère sera ici considéré comme effectif pour tout élément ne pouvant pas être employé de manière autonome en français et n'étant pas formé à partir d'un élément autonome du français. L'hétérolexicalité distingue ici les éléments *-gate* et *-ciel* ou les éléments *-attitude* et *-bashing*. La question de l'hétérolexicalité d'un élément comme *-show* pose problème, puisque nous avons vu qu'il peut être utilisé de manière autonome en français ; je pense toutefois que nous sommes dans un cas de xénisme, et que ce lexème n'est pas vraiment intégré au lexique français : il ne prend pas la marque du pluriel, et sa prononciation n'est pas fixée entre, par exemple, [ʃo] et [ʃou].

Autonomie. Le critère de l'autonomie est un de ceux souvent présentés pour différencier lexèmes et affixes. Parmi les ENs étudiés ici, seuls les éléments *phobie* et *logique* peuvent être utilisés de manière autonome. Cependant, comme nous l'avons déjà vu, *-phobie* offre un éventail de sens bien plus large quand il est attaché que quand il est autonome. De son côté, *logique*, quand il est un nom autonome, désigne une science dont l'objet d'étude sont les normes de la vérité et quand il est un adjectif, se rapporte à cette science et est passé dans le langage courant avec le sens 'cohérent, conforme au bon sens'. Il a donc plus à voir avec le sens originel de 'rapport' que l'on retrouve dans *analogue* et donc dans *analogique*. Nous avons déjà vu que ce sens n'est plus disponible en français moderne et sa présence est uniquement due à des emprunts faits au grec. Il est donc plus difficile de relier LOGIQUE à *biologique* que PHOBIE à *xénophobie*, même si dans ce dernier le sens de *-phobie* n'est pas le même que celui de PHOBIE. Dans le cas de *-phobie* nous sommes en présence d'une polysémie régulière, alors que dans le cas de *-logique* plutôt du côté d'une démotivation du lien.

Intégrité Lexicale. Ce critère est proposé par Warren (1990). Il discrimine les éléments tronqués des éléments non tronqués. Dans le cas des éléments en position finale, cette troncature semble équivaloir systématiquement à une perte d'autonomie. Ce n'est pas le cas des éléments initiaux, comme nous l'avons vu avec *télé* ou *ciné* qui peuvent, en tant que lexèmes tronqués, aussi bien intervenir dans des formations qu'être utilisés comme unité lexicale au sein d'une phrase.

Distribution. Un des arguments souvent donnés contre l'attribution du statut d'affixe aux éléments néoclassiques est que leur position peut varier selon les formations. D'une part, d'un point de vue du sens, le constituant apparaissant en position initiale a généralement un sens plus restreint. Prenons l'exemple de *-log-*¹⁷ : comme le soulignent Amiot et Dal (2007, p.332), quand il est en position initiale, il est restreint au sens de 'discours' (121-a). Ce sens de 'discours' est également présent dans des lexèmes formés en français où *-log-* est en position initiale (121-b) mais, lorsqu'il est en position finale, il l'est uniquement dans des lexèmes empruntés au grec ancien (121-c).

- (121) a. logomachie, logopathie
 b. logosphère, logophasie
 c. dialogue, misologue

Quand *-log-* est en position finale, comme nous l'avons vu, l'éventail de sens est bien plus large (voir la description qui en a été faite en 3.2.1).

La même analyse peut être faite pour l'élément *-phon-*. Lorsqu'il est employé de façon autonome, sa sémantique est restreinte aux sens de 'voix' ou 'son', dans le domaine linguistique (dans lequel il peut également être la forme raccourcie de PHONOLOGIE) (122-a), de la physique acoustique (122-b), de la médecine (122-c) ou de la biologie (122-d) :

- (122) a. phonostylistique
 b. phonocapteur
 c. phonospasme, phonocardiographe
 d. phonocinèse

Si cet élément rentre dans des formations de noms d'instruments autant en première position (*phonomètre*) qu'en seconde (*vibraphone*), quand il est en seconde position l'éventail de types d'instruments est plus large (par exemple

17. Cette notation permet de référer à la fois à l'élément initial *logo-* et aux éléments finaux *-logue* ou *-logie* (dans leurs différents sens) qui partagent une même origine.

des instruments de musique); *-phone* est par ailleurs particulièrement disponible pour la formation de noms de locuteurs, ce qui est impossible pour *phono-* (revoir la partie 3.2.1 pour une étude sémantique détaillée des formations *Xphone*).

Pour un élément comme *-morph-*, la position inverse s'observe : c'est lorsque l'élément est en position initiale qu'il présente un éventail de sens plus large. En effet, *-morph-* rentre uniquement dans la formation de lexèmes *Xmorph-* signifiant 'qui a la forme de X' (123-a), alors que *morpho-*, s'il présente ce sens (123-b), peut également représenter le lexème *morphologie* dans le domaine linguistique (123-c) ou dans celui des sciences de la vie (123-d).

- (123) a. anthropomorphe
 b. morphogénèse
 c. morphosyntaxe
 d. morphopsychologie

De même, ces éléments sont largement plus utilisés avec des lexèmes autonomes quand ils sont en position finale qu'en position initiale. Le Tableau 5.13 donne, pour indication, le nombre d'éléments dans chaque position, ainsi que leur association avec des items autonomes, selon les données du TLFi.

élément	1 ^{ère} position	avec C2 autonome	2 ^{de} position	avec C1 autonome
-log-	28	6	170	50
-phon-	24	6	42	16
-morph-	20	12	40	3

Tableau 5.13 – Proportion d'éléments néoclassiques en première et seconde position

Ces différences de comportements, tant sémantiques que distributionnels, font penser que nous sommes en présence d'éléments distincts quoique diachroniquement reliés. Concernant l'ensemble des éléments finaux récurrents abordés ici, seul *show* semble avoir un comportement identique quand il est en position initiale ou finale, du moins sémantiquement, puisqu'il apporte toujours le sens de 'spectacle' (quand il est en position initiale, il s'agit uniquement d'emprunts faits à l'anglais) :

- (124) a. show-business, showroom, show-tv
 b. Hollande show, Sarko-show

Cependant, phonologiquement, des différences apparaissent puisque, comme nous l'avons vu, une voyelle de liaison <o> peut apparaître entre les deux

constituants lorsque *-show* est en position finale.

Catégorie lexicale. Le critère de la catégorie lexicale est difficile à déterminer dès lors que les éléments en question n'ont pas d'autonomie syntaxique. Si certains confèrent une catégorie lexicale aux éléments néoclassiques en fonction d'une traduction possible ou de leurs propriétés combinatoires (voir plus haut en 5.1.1.3), une généralisation à tous les éléments néoclassiques est abusive. Pour s'en tenir à ceux qui nous préoccupent, certains affichent des propriétés clairement prédicatives alors que pour d'autres celles-ci sont exclues. Par exemple, lorsque l'EN était dans les langues classiques un élément verbal, il peut avoir conservé certaines propriétés de cette catégorie. Regardons-les chacun plus en détails.

L'élément *-cide* n'intervient que dans des constructions dans lesquelles le C1 est l'objet du meurtre/de la tuerie/de la destruction. L'élément conserve toujours une interprétation verbale ; si le C1 a toujours le rôle de patient, le MC peut désigner l'agent, le procès ou l'instrument.

L'élément *-cole* ne rentre que dans la formation d'adjectifs. Les lexèmes du type *christicole*, où *-cole* a le sens de 'honorer', ne sont pas générateurs de nouvelles formations en français. Si l'EN a une interprétation verbale, elle ne semble pas se répandre aux autres formations. Dans les adjectifs du type *agricole*, même si *-cole* était à l'origine un verbe, cette interprétation verbale semble avoir cédé sa place à des relations lexicales pour la lecture du mot construit : en effet, *agricole* se rapporte à *agriculture* ou *agriculteur* qui n'affichent pas de relation prédicative entre les constituants. Pour moi, ce type de relation n'est pas non plus présent dans *agricole* où *-cole* n'est plus perçu comme étant un verbe. Concernant les lexèmes dans lesquels *-cole* a le sens de 'habiter' (*arboricole*), le C1 désigne la position du prédicat. L'élément *-cole* peut être, dans ce cas, encore perçu comme un verbe.

Les éléments *-phage* et *-vore* entrent dans des formations dans lequel le C1 désigne le patient (*anthropophage*, *carnivore*), tout comme l'élément *-cide*. Cependant, la relation prédicative s'étioule dès lors qu'une extension métaphorique s'installe (*téléphage*, *BDvore*).

Pour Amiot et Dal (2007, p.333), les constructions *Xlogue* ne sont pas perçues par les locuteurs du français comme ayant un verbe dans leur structure. Pour moi, il en va de même pour les ENs *-crate*, *-phone* et *-phobe*. Il est dès lors difficile de leur attribuer une catégorie lexicale. S'il s'agissait de noms, leur sens serait constant, ce qui n'est pas le cas.

Indice lexical. L'indice lexical représente l'endroit du lexique où est stocké le lexème. Un indice lexical est attaché aux trois rubriques du lexème, PHON,

SYN et SEM. Par exemple, si on suppose que BAKE, ‘cuire’, a l’indice lexical 82, on peut faire référence à ses propriétés comme PHON₈₂, SEM₈₂ et SYN₈₂ (Booij, 2010, p.7). Chaque lexème a un indice propre. Un affixe ne peut donc pas avoir d’indice lexical. Ce critère est rempli lorsqu’un élément est relié à un lexème (il en est le thème et/ou le radical).

Spécification sémantique. Ce critère est pertinent pour les éléments ayant un indice lexical, c’est-à-dire reliés à une position lexicale dans le lexique, position qui affiche une forme phonologique, un sens et une catégorie lexicale. Le critère de spécification sémantique s’intéresse aux éléments dont le sens lorsqu’ils sont attachés est plus spécifique que lorsqu’ils sont autonomes. Plus précisément, leur sens peut varier dans certaines constructions, mais rester intact dans d’autres. Le lexème CLEF, ‘objet servant à ouvrir des portes’, a un sens spécifié dans les constructions [N clef]_N (*notion clef*) mais pas dans les constructions [V clef]_N (*porte clef*).

Abstraction. L’abstraction est un critère avancé par Fradin (2000, p.36) au sujet de la sécrétion : « abstraction amounts to replacing an element by a variable of the same type and to binding it by the lambda operator. It transforms an expression into a function ». C’est difficile ici de considérer l’abstraction en tant que processus, puisque certains éléments, les ENs étudiés dans cette thèse, ont été introduits en français déjà par le biais de constructions forgées dans les langues classiques. C’est un processus différent de celui de la sécrétion pour lequel le phénomène de l’abstraction peut être décrit dans le passage d’un lexème autonome à un affixe. Ici, je considérerai comme répondant au critère de l’abstraction les éléments qui sont considérés dans le cadre théorique de la morphologie constructionnelle comme exposants de construction. Nous avons déjà vu que ce rôle d’exposant peut être rempli par des items lexicaux ou affixaux, même si l’affixe est l’exposant de construction prototypique. Un exposant de construction est spécifié à l’intérieur du schéma constructionnel, et représente formellement une opération sémantique. L’autre élément est donc la base de la construction. Le sens d’un lexème construit par un schéma impliquant un exposant de construction n’est pas apporté par l’exposant lui-même, mais il est construit selon le schéma employé, la base sélectionnée, le contexte d’emploi. Cependant, analyser un élément comme exposant de construction n’empêche pas de lui conférer une existence. C’est ce pour quoi militent Aronoff (2007) ou Roché (2011a) : « Si les “exposants” associés a des règles qui produisent respectivement, par exemple, un nom processif et un nom de qualité étaient seulement des chaînes segmentales identiques, on serait devant un cas d’homophonie

fortuite. On ne pourrait pas rendre compte des interférences et des éléments de continuité que nous avons observés. Ils attestent que l’affixe a une réalité, une existence en lui-même » (Roché, 2011a, p.138). Ainsi, la formation du sens du mot construit est pris en charge par, entre autres, le schéma de construction ; l’output n’est pas la concaténation des sens des éléments à l’input. Cependant, ces éléments dont le sens est devenu abstrait ne sont pas eux mêmes abstraits : ils sont le témoin de relations existant dans le lexique, par exemple dans des cas de polysémie régulière, ou marquent l’appartenance à une série dérivationnelle.

Par exemple, l’élément *fleuve* a vu son sens se spécifier et devenir abstrait dans la construction *Xfleuve*, il peut être considéré comme un exposant de construction ; cependant, l’exemple (125) montre bien que nier son existence empêche de comprendre la métaphore présente dans la phrase.

- (125) C’est sans doute la promesse politique la plus célèbre à propos de la Seine. Pas un **discours fleuve**, juste une petite phrase : « Dans cinq ans, on pourra à nouveau se baigner dans la Seine », avait déclaré Jacques Chirac le lundi 28 novembre 1988, alors qu’il brigait un troisième mandat à la Mairie de Paris. (http://www.la-croix.com/Ethique/Sciences-Ethique/Sciences/Se-baigner-dans-la-Seine-_NG_-2010-07-05-553851)

Envisager les ENs comme exposants de construction, c’est-à-dire des éléments spécifiés dans des schémas, schémas qui ont été réalisés sur la base de généralisations faites sur le lexique, ne revient pas à nier leur existence et à en faire uniquement des objets abstraits. Par exemple, l’utilisation de *-logue* plutôt que *-iste* confère au mot construit une dimension plus savante ; mais *Xiste* et *Xlogue* répondent au même schéma de nomination.

Charge sémantique nulle. Je propose ce critère pour distinguer les éléments pouvant perdre toute charge sémantique. Nous avons vu que certains éléments étaient susceptibles de ne représenter qu’une intégration dans un paradigme ou qu’une adjectivisation.

Ainsi, j’avais déjà observé certains emplois de *-logue* comme intégrateur du mot construit dans un paradigme :

- (126) dentistologue, graphistologue

De plus, certains affixes sont uniquement une marque de changement catégoriel sans opération sémantique :

- (127) rayonnement solaire
 personne abstentionniste
 élection présidentielle

Nous avons vu dans le chapitre 4 que certains éléments néoclassiques pouvaient avoir la même fonction. Je reporte ici les exemples que nous avons vu :

- (128) gastronomie ostréicole
 consonance anglophone
 contrôle œnologique
 consommation locavore

Certains éléments sont particulièrement disponibles pour ce genre d'emploi (nous avons vu l'exemple de *-logique*) alors que d'autres semblent l'être beaucoup moins (je n'ai par exemple trouvé aucun emploi de *-cide* comme marqueur de relation). Cependant, comme il est plus facile de démontrer la possibilité d'un emploi que son impossibilité, selon les exemples rencontrés, les éléments pour lesquels cet emploi sans charge sémantique n'est noté dans le Tableau 5.14 que lorsque des exemples ont été effectivement trouvés, et la case reste vide quand ce n'est pas le cas.

Bilan. Le Tableau 5.14 répertorie chacun des éléments finaux décrits dans ce chapitre selon les neuf critères envisagés ici ; il résume ainsi ce qui a été dit dans le début de cette section. Ce tableau constitue une proposition initiale : des études plus approfondies sur chaque critère permettrait d'affiner sa réalisation. De plus, il est parfois difficile de déterminer si un élément remplit un critère ou non, ce qui explique les points d'interrogation. Enfin, certains critères ne sont pas pertinents pour certains éléments (notés n.p. dans le tableau).

	Hétérolexicalité	Autonomie	Intrégrité lexicale	Distribution	Catégorie lexicale	Indice lexical	Spécification sémantique	Abstraction	Charge sémantique nulle
<i>-fleuve</i>	-	✓	✓	-	✓	✓	✓	✓	
<i>-star</i>	-	✓	✓	-	✓	✓	-	-	
<i>-thérapie</i>	-	✓	✓	-	✓	✓	-	✓	
<i>-show</i>	✓	✓	✓	✓?	✓	✓	-	✓	
<i>-attitude</i>	-	✓	✓	-	✓	✓	-	✓	
<i>-bashing</i>	✓	✓ (?)	✓	-	✓	✓	-	✓	
<i>-gate</i>	✓	-	-	-	-	-	n.p.	✓	
<i>-ciel</i>	-	-	-	-	✓	✓	-	-	
<i>-phobie</i>	-	✓	n.p.	-	?	✓	✓	✓	
<i>-phobe</i>	✓	-	n.p.	-	-	-	n.p.	✓	
<i>-cide</i>	✓	-	n.p.	-	✓	-	n.p.	✓	
<i>-cole</i>	✓	-	n.p.	-	-	-	n.p.	✓	✓
<i>-vore</i>	✓	-	n.p.	-	✓/-	-	n.p.	✓	✓
<i>-phage</i>	✓	-	n.p.	-	✓/-	-	n.p.	✓	
<i>-phone</i>	✓	-	n.p.	-	-	-	n.p.	✓	✓
<i>-cratie</i>	✓	-	n.p.	-	-	-	n.p.	✓	
<i>-crate</i>	✓	-	n.p.	-	-	-	n.p.	✓	
<i>-logie</i>	✓	-	n.p.	-	-	-	n.p.	✓	
<i>-logue</i>	✓	-	n.p.	-	-	-	n.p.	✓	✓
<i>-logique</i>	✓	✓ (?)	n.p.	-	-	-	n.p.	✓	✓
<i>-ifier</i>	n.p.	-	n.p.	n.p.	-	-	n.p.	✓	
<i>-iste</i>	n.p.	-	n.p.	n.p.	-	-	n.p.	✓	✓

Tableau 5.14 – Caractérisation d’éléments finaux récurrents selon 9 critères

5.3.1.2 Les constructions à élément final spécifié

A présent, comparons les constructions impliquant ces éléments finaux récurrents selon dix critères. Les observations ci-dessous seront résumées dans le Tableau 5.16.

Héritage des propriétés syntaxico-sémantiques. Ce critère n'est pertinent que pour les constructions impliquant des éléments finaux autonomes. Il permet de faire la distinction entre les composés traditionnellement dits natifs et ceux dits néoclassiques.

Position de la base de la construction. Nous avons vu précédemment que la notion de tête n'est pas la plus adéquate. Un lexème est construit sur la base d'un autre lexème en instanciant un schéma. Même si l'ordre interne d'un mot construit peut être guidé par les catégories lexicales des constituants et les relations qui les unissent, l'élément qui gouverne l'autre n'est pas forcément la base du lexème construit. Par exemple, dans le schéma (129), X est la base de la construction. Les lexèmes *mouchicide*, *limacide* ou *lombricide*instancient ce schéma en ayant respectivement pour base MOUCHE, LIMACE et LOMBRIC.

(129) [[X]_{N_i} cide]_{N_j} ↔ [instrument permettant la destruction de SEM_i]_j

Même si X a le rôle de patient dans l'interprétation verbale de *-cide*, nier qu'il est également la base du mot construit empêche de considérer le schéma (129) comme générateur d'une liste de lexèmes, liste pourtant présente dans NeoClassy.

L'élément base de la construction peut être différent de l'élément qui lègue ses propriétés sémantico-syntaxiques ; c'est ce que l'on voit par exemple avec les constructions *Xphobie*.

Contraintes sur le constituant de gauche. Ce critère distingue les constructions opérant une contrainte dans la sélection de la base, que ce soit une contrainte sur son type sémantique ou sur le thème sélectionné. Ainsi, les constructions suffixales imposent une sélection d'un thème de la base, voire un réajustement du thème au niveau du radical (voir par exemple Roché (2011a) sur les constructions *Xiste*).

Nous avons vu, au travers d'exemples, que la construction *Xgate* prenait systématiquement le thème A du lexème base, peu importe sa finale (revoir l'exemple (105) de ce chapitre).

Par ailleurs, les constructions néoclassiques privilégient des thèmes S, comme nous l'avons vu précédemment.

Contraintes de taille. Le critère de la taille des constructions demanderait une étude plus poussée. Il s'agit ici de donner une sensibilité que peut avoir le constituant de gauche à se conformer à un gabarit particulier. Nous avons toutefois vu que les constructions impliquant les ENs étudiés ici sont soumises à des contraintes de taille. Plénat (2009) a également dégagé des contraintes de taille auxquels sont soumis les dérivés. En particulier, Lignon (2013) a démontré la préférence des constructions *Xifier* pour des bases monosyllabiques.

Contraintes sur la structure segmentale. Ce critère distingue les constructions ayant par exemple une préférence pour un segment déterminé à la jonction des deux éléments, /ɔ/ ou /i/ ou requérant un radical à finale consonnantique. Nous avons vu que, même pour les ENs d'origine latine affichant préférentiellement la voyelle /i/ dans la construction, la voyelle /ɔ/ gagnait peu à peu du terrain. Cette voyelle est également parfois présente dans les constructions *Xshow* ou *Xbashing*, même si aucun de ces éléments n'est d'origine classique.

Participation à un réseau lexical comprenant des éléments non autonomes. Ce critère distingue, parmi les constructions impliquant des éléments finaux qui peuvent être autonomes, celles qui entretiennent des relations avec des constructions impliquant des éléments finaux non autonomes. Ce critère permet ainsi de différencier les constructions *Xphobie* des constructions *Xthérapie*, mais de les rapprocher des constructions *Xlogie* :

- (130) théâtrelogie / théâtrelogue
 théâtrophobie / théâtrophobe
 théâtrothérapie / théâtrothérapeute

L'élément *-phobe* est en effet un élément non autonome, au même titre que *-logue* ou *-logie* mais forme des lexèmes qui correspondent à d'autres lexèmes qui sont eux formés avec l'élément autonome *-phobie*. Nous avons cependant vu précédemment, par exemple en 3.2.1, que le constituant *-phobie* proposait un éventail de sens plus important que le lexème autonome PHOBIE.

Mot leader identifié. Dans le chapitre 3, nous avons vu que certains schémas étaient initiés par des phénomènes d'analogie avec un mot leader.

Cette notion est abordée surtout d'un point de vue diachronique par Roché (2011a). Recourir à l'analogie permet d'identifier les mots qui sont à l'origine d'un effet de série. Roché identifie les mots leader desquels des séries sont nées (par exemple, les noms en *-alisme*). Cependant, si ces lexèmes ont bel et bien été leader à un moment de l'histoire, il me paraît difficile, synchroniquement, de déterminer lequel est encore leader. Concernant les constructions forgées à partir des ENs étudiés dans cette thèse, certaines sont clairement formées sur le modèle de mots particulièrement saillants :

- (131) démocratie, aristocratie
 homicide, génocide, insecticide
 arachnophobe, claustrophobe, xénophobe, homophobe

Preuve en est que, souvent, des néologismes sont syntagmatiquement reliés au mot leader sur le modèle duquel ils ont été formés :

- (132) Un système politique avec zéro peuple et un Président zéro, qui ne représente rien sauf lui même et un ramassis de corrompus, ce n'est pas de la Démocratie mais de la **Zérocra**tie! (http://blaisap.ty pepad.fr/mon_weblog/2011/07/le-bnin-et-la-zrocratie.html)

Pour certains ENs, des mots leader sont saillants uniquement pour certains schémas, c'est pourquoi le critère sera noté comme respecté mais entre parenthèses, dans le Tableau 5.16 :

- (133) carnivore, herbivore, omnivore
 agricole

Les lexèmes formés par sécrétion ont un mot leader clairement identifié, ceux relevant de la composition cachée également.

Paradigmatisation. La taille de la série morphologique est un autre critère : certains schémas impliquant des suffixes prototypiques, par exemple les schémas *Xiste*, forment de très longues séries de lexèmes. C'est également le cas pour les ENs étudiés ici (c'était d'ailleurs un des critères de sélection de ces ENs en particulier, voir chapitre 2). Plus une série dérivationnelle est longue, plus elle a un pouvoir d'attraction fort. Nous avons eu un exemple avec les lexèmes en *-logue* formés sur des noms qui désignaient déjà des spécialistes, pour lesquels l'insertion dans la série paradigmaticque passait par l'affichage du même élément final. Très anecdotiquement, sur la version PDF

téléchargeable en ligne de Villoing et Namer (2006), le titre annonce *saxifage* au lieu de *saxifrage* ; les auteurs, ou éditeurs, semblent avoir été attirés par l'élément *-phage* bien plus répandu que l'élément *-frage* : à titre comparatif, le GR ne mentionne qu'un seul autre lexème construit avec l'élément *-frage*, *ossifrage*, contre 53 avec l'élément *-phage*.

Parce que la paradigmatization d'un schéma est difficilement quantifiable, j'ai relevé les occurrences pour chacun des éléments finaux considérés ici dans le corpus FrWac. Les fréquences, données dans le Tableau 5.15, sont ici données à titre indicatif¹⁸ ; en effet, les listes de mots obtenues n'ont pas subi de tri ultérieur et contiennent donc beaucoup de bruit : pour ne donner qu'un exemple, la recherche des termes en *-gate* donne également comme résultat *régate*. Concernant les mots en *-cole* j'ai retiré tous ceux en *-école* qui doubleraient la taille de la liste. Les mots en *-ciel* et en *-star* ne sont pas affichés ici, ne contenant à première vue que très peu de résultats pertinents.

Élément	Nombre sur FrWac	Élément	Nombre sur FrWac
Xbashing	20	Xcole	554
Xfleuve	71	Xcide	717
Xvore	234	Xifier	771
Xattitude	244	Xthérapie	798
Xphage	253	Xphobie	829
Xcrate	268	Xphone	1065
Xphobe	282	Xlogue	1681
Xshow	322	Xlogique	2224
Xgate	403	Xlogie	5204
Xcratie	440	Xiste	12288

Tableau 5.15 – Fréquence de type des formations à élément final spécifié

Si ces chiffres demandent à être affinés et chaque occurrence vérifiée manuellement, nous voyons que, comparativement, certaines séquences sont bien plus fréquentes que d'autres et que la force d'attraction de certains ENs est comparable à celle de suffixes.

Nombre de schémas. Il n'est pas rare, et même plutôt fréquent, que des suffixes rentrent dans la formation de différents schémas constructionnels,

18. Concernant les formations impliquant des lexèmes par ailleurs autonomes, comme *Xfleuve*, le corpus FrWac ne fournit que ceux dont les constituants sont reliés par un trait d'union. Cela peut expliquer le faible nombre qui a été reporté dans le tableau.

sémantiquement reliés ou non. Des sous-schémas impliquant un même élément peuvent en fait afficher une polysémie régulière (Booij, 2010, p.77) ; cela se vérifie en général avec des suffixes prototypiques (Booij propose une étude de l’affixe néerlandais *-er*, Roché (2011a) celle des suffixes *-isme* et *-iste* par exemple) : « un même affixe peut être associé à plusieurs modèles dérivationnels » (Roché, 2011a, p.138).

L’implication dans plusieurs schémas semble donc un critère pouvant témoigner qu’un élément est en voie d’affixation. Pour ne pas risquer la non exhaustivité, dans le Tableau 5.16 je m’en tiendrai à discriminer les éléments entrant dans un schéma et ceux dans plusieurs, sans en donner le nombre. Concernant les éléments néoclassiques étudiés ici, la description des schémas identifiés à partir des données de NeoClassy est donnée dans la partie 3.2.1.

Concurrence avec d’autres procédés. Un des critères donnés par Van Goethem et Amiot (2009) pour décrire l’affixation d’un élément lexical est qu’il rentre en compétition avec d’autres affixes dérivationnels. En effet, des schémas impliquant des affixes divers peuvent remplir des schèmes de nomination identiques. Par exemple, Lignon (2013) explore la compétition entre les suffixes *-iser* et *-ifier*. Concernant les éléments néoclassiques, on peut identifier les compétitions suivantes :

- (134) *Xiste* / *Xcrate* / *Xphile* ‘partisan de X’
Xiste / *Xien* / *Xlogue* ‘spécialiste de X’
anti-X / *Xcide* / *tue-X* ‘instrument destructeur de X’
anti-X / *Xphobe* ‘qui est hostile à X’
Xphage / *Xphile* / *Xvore* / *Xmane* ‘amateur de X’

Même si toutes ces compétitions ne se jouent pas avec des affixes prototypiques, leur existence même est le témoin d’une possibilité de, quasiment à des nuances près parfois, interchanger les éléments finaux. Cela est le témoin d’une abstraction sémantique telle que je l’ai décrite dans la section précédente.

Bilan. Le Tableau 5.16 caractérise donc chacune des constructions envisagées en fonction des critères qui viennent d’être proposés. Je vais maintenant, à partir de ce tableau, présenter une typologie de ces constructions.

Héritage propriétés sém-syn
 Position base
 Contraintes sur CI
 Contraintes de taille
 Contraintes structurales
 Réseau lexical
 mot leader
 Paradigmatisation
 Nombre de schémas
 Concurrence

<i>Xfleuve</i>	g	g	-	-	-	-	-	-	-	+	1	-
<i>Xstar</i>	g	g	-	-	-	-	-	-	-	-	1	-
<i>Xthérapie</i>	d	g	✓	✓	/ə/	/ə/	-	-	-	+++	>1	-
<i>Xshow</i>	d	?	-	-	(/ə/)	(/ə/)	-	-	-	++	1	-
<i>Xattitude</i>	d	d	-	-	/ə/	/ə/	-	-	-	++	1	-
<i>Xbashing</i>	d	d	-	-	(/ə/)	(/ə/)	-	✓	✓	+	1	-
<i>Xgate</i>	n.p.	g	-	-	/ə/	/ə/	n.p.	✓	✓	++	1	-
<i>Xciel</i>	n.p.	d	-	-	/i/	/i/	n.p.	✓	✓	-	1	-
<i>Xphobie</i>	d	g	✓	✓	/ə/	/ə/	✓	✓	✓	+++	>1	-
<i>Xphobe</i>	n.p.	g	✓	✓	/ə/	/ə/	n.p.	✓	✓	++	>1	✓
<i>Xcide</i>	n.p.	g	✓	✓	/i/ > /ə/	/i/ > /ə/	n.p.	✓	✓	+++	>1	✓
<i>Xcole</i>	n.p.	g	✓	✓	/i/ > /ə/	/i/ > /ə/	n.p.	(✓)	(✓)	+++	>1	✓
<i>Xvore</i>	n.p.	g	✓	✓	/i/ > /ə/	/i/ > /ə/	n.p.	(✓)	(✓)	++	>1	✓
<i>Xphage</i>	n.p.	g	✓	✓	/ə/	/ə/	n.p.	-	-	++	>1	✓
<i>Xphone</i>	n.p.	g	✓	✓	/ə/	/ə/	n.p.	-	-	++++	>1	✓
<i>Xcratie</i>	n.p.	g	✓	✓	/ə/	/ə/	n.p.	✓	✓	++	>1	-
<i>Xcrate</i>	n.p.	g	✓	✓	/ə/	/ə/	n.p.	✓	✓	++	>1	✓
<i>Xlogie</i>	n.p.	g	✓	✓	/ə/	/ə/	n.p.	-	-	++++	>1	✓
<i>Xlogue</i>	n.p.	g	✓	✓	/ə/	/ə/	n.p.	-	-	++++	>1	✓
<i>Xlogique</i>	n.p.	g	✓	✓	/ə/	/ə/	n.p.	-	-	++++	>1	✓
<i>Xifier</i>	n.p.	g	✓	✓	-	-	n.p.	-	-	+++	>1	✓
<i>Xiste</i>	n.p.	g	✓	✓	-	-	n.p.	-	-	++++	>1	✓

Tableau 5.16 – Caractérisation de constructions à élément final récurrent selon 10 critères

5.3.2 Une typologie des éléments finaux

Pour résumer les informations données dans les Tableaux 5.14 et 5.16, les ENs finaux récurrents, tout comme les suffixes, sont non autonomes, font preuve d'une abstraction sémantique, entrent en concurrence avec d'autres éléments finaux dans la réalisation de schèmes de nomination, sont impliqués dans des schémas obéissant à des contraintes phonologiques, témoignent d'une force d'attraction manifeste en rentrant dans de longues séries de lexèmes et sont la manifestation formelle d'opérations sémantiques affichant une polysémie régulière.

Ainsi, nous pourrions schématiser, comme nous l'avons fait dans Lasserre et Montermini (2014b), les différentes formations comme en (135).

- (135) composés canoniques > X fleuve > X thérapie > X EN > suffixés canoniques

Du point de vue du sens, une abstraction s'opère au niveau de *Xfleuve*. Du point de vue de la forme, des contraintes sur la sélection du C1 ainsi que des contraintes phonologiques sur le mot construit commencent à peser au niveau de *Xthérapie*, par exemple la présence d'une voyelle de liaison dans la construction. Du point de vue syntaxique, on assiste à une perte de la catégorie lexicale au niveau des *X EN*. Nous l'avons vu avec par exemple les constructions *Xphage* dans lesquelles le constituant n'a plus que la fonction de verbe psychologique, moins prototypique qu'un verbe d'action :

- (136) cinéphage, téléphage,...

Enfin, du point de vue de la structuration du lexique, on assiste à une paradigmatisme ainsi qu'à une concurrence à l'intérieur de schèmes de nomination au niveau des *X EN*. En fait, le passage du stade du composé canonique à celui de suffixé canonique se fait par le biais non seulement des différentes rubriques que l'on trouve dans le lexème (PHON, SEM et SYN), mais également par le biais de l'intégration du schéma de construction dans le lexique. Chacun des éléments finaux abordés dans cette dernière partie peuvent être à des stades différents selon chacune des modalités. La Figure 5.4 propose une typologie, en reprenant les différents critères déjà évoqués, de chacune des constructions étudiées, en essayant de les placer chacune sur des gradients reliant composés canoniques et suffixés canoniques.

La classe des *X EN* n'est donc pas homogène : ainsi, les constructions *Xlogue* ou *Xlogique* ont davantage de propriétés suffixales que les constructions *Xcide* dans lesquelles une relation prédicative relie les constituants ou que les constructions *Xcrate*, qui sont toujours en relation avec un mot leader

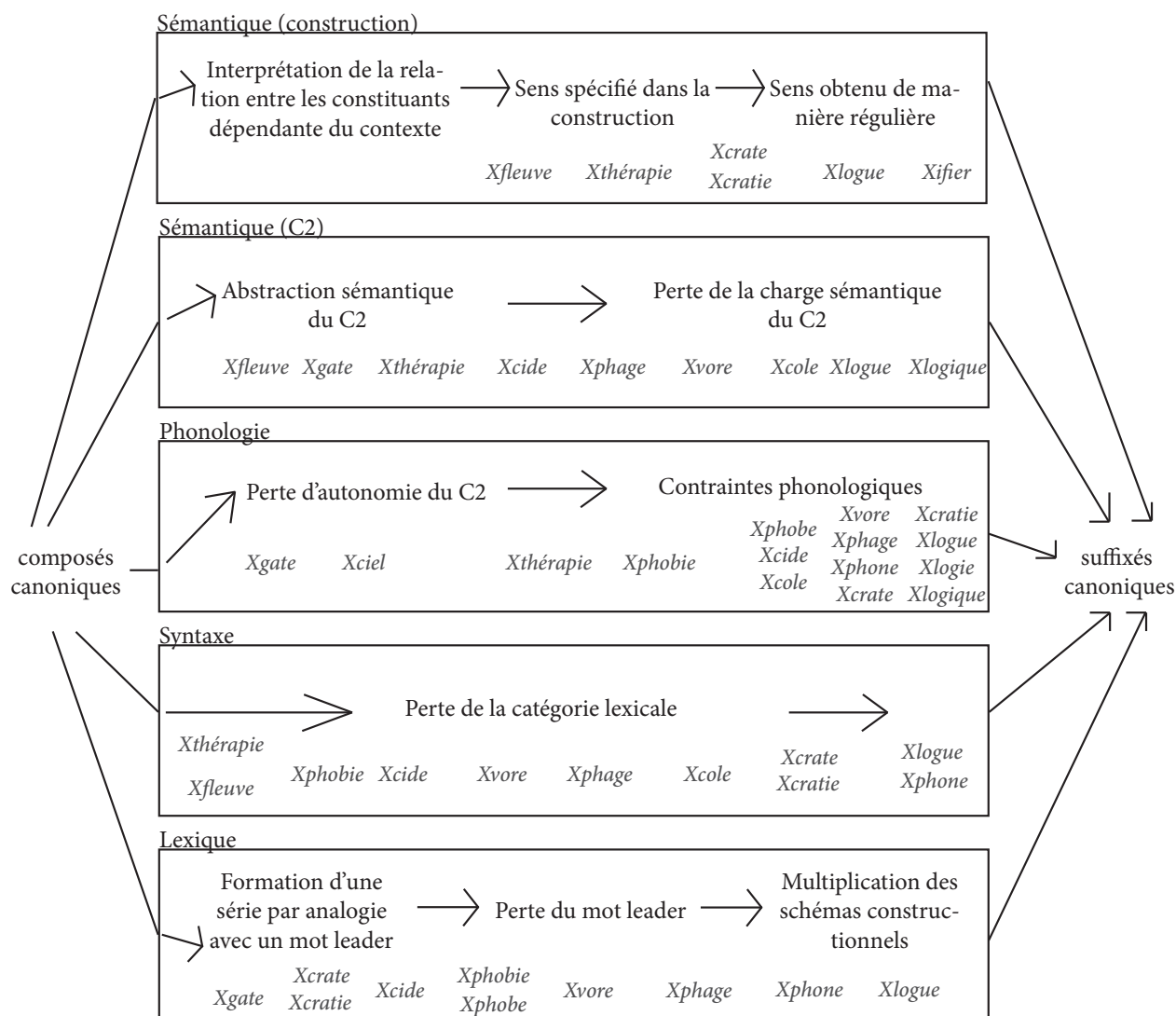


FIGURE 5.4 – Typologie des constructions présentant un élément final récurrent

saillant.

Toutefois, à des degrés différents, chacun des ENs étudiés dans cette thèse se rapproche davantage de la catégorie des suffixes que de celle des constituants composés. Roché et Plénat (2014) parlent d'ailleurs de *thème présuffixal saillant* pour qualifier le constituant de gauche des constructions néoclassiques, sous-entendant par là que le constituant final est un suffixe. Néanmoins, l'appellation d'élément néoclassique ne rend pas compte d'une classe homogène ; elle reste purement diachronique et lexicographique. A l'intérieur de cette classe, il y a aussi bien des thèmes de lexèmes que des éléments de type suffixal. Ainsi, si *hydrologie* peut être décrit comme étant formé grâce à deux ENs, ce lexème instancie en fait le schéma $[X_{N_i} \text{ logie}]_{N_j}$ dans lequel *-logie* est l'exposant, sur la base EAU ; ce schéma sélectionne préférentiellement des thèmes supplétifs, et c'est pourquoi le thème S *hydr-* est ici sélectionné. Postuler une classe hétérogène de ENs permet de rendre compte des constructions néoclassiques telles que *hydrologie* dans le même cadre qu'une analyse de composés ou suffixés canoniques.

Bilan. Ce dernier chapitre a repris les notions de lexème et d'affixe et avec elles celles de composition et de dérivation pour démontrer que, dans une morphologie non idéalisée, les frontières entre chaque élément et entre chaque type de construction ne peuvent pas être nettes et clairement définies. Un suffixe peut afficher un sens qui serait considéré comme lexical, un lexème peut avoir une place similaire à celle d'un affixe dans un schéma de construction. Les éléments néoclassiques, dans ce cadre, trouvent une place qui ne peut pas être définie comme vérifiée pour chacun des éléments d'une classe qui n'aurait de raison d'être que diachronique. En effet, certains éléments néoclassiques ne sont autres que des thèmes supplétifs de lexèmes, servant de base à des opérations de constructions, une affixation par exemple. D'autres, au contraire, et en particulier ceux qui nous intéressent dans cette thèse, ont une existence à l'intérieur de schémas constructionnels et servent à la structuration du lexique, tout comme des suffixes canoniques. Les notions d'élément néoclassique ou de combining form doivent être réservées à des fins lexicographiques mais ne rendent pas compte de la disparité des données.

Les ENs finaux récurrents ont ainsi été inscrits dans une typologie comprenant d'autres éléments finaux récurrents : ces éléments peuvent être des lexèmes qui ont conservé leur intégrité, des lexèmes tronqués ou des suffixes. Tous ces éléments, ainsi que les constructions qui les renferment, pourront être rapprochés selon certains critères : par exemple, phonologiquement *Xthérapie* et *Xlogie* partagent la contrainte de la présence de la voyelle <o> dans la construction, même si le premier comprend un élément final autonome,

mais pas le second. Cet éventail de critères qui a été proposé permet ainsi de ne pas avoir une vision tranchée des classes d'éléments, chacun pouvant être rapproché d'un autre selon certains critères mais éloigné selon d'autres critères. Cependant, les éléments néoclassiques finaux récurrents se rapprochent des suffixes, partageant avec ces derniers éléments de nombreux critères, tant phonologiques, que sémantiques ou lexicaux.

Conclusion et perspectives

L'objet de ce travail de thèse a été de traiter les constructions françaises impliquant des éléments néoclassiques finaux, spécifiés à l'intérieur de schémas de construction et disponibles pour la création lexicale. Cette thèse a montré que ces éléments souvent considérés comme faisant partie d'une catégorie homogène peuvent en fait afficher des caractéristiques très différentes. En effet, si certains d'entre eux sont des éléments lexicaux, et plus précisément des thèmes supplétifs de lexèmes, d'autres n'ont plus que le rôle de marqueur morphophonologique d'une opération sémantique et ainsi se rapprochent des affixes prototypiques. Le mot *palynologue*, 'spécialiste des pollens', n'est pas construit selon un procédé différent de ceux par lesquels sont construits *mathématicien*, 'expert en mathématiques' ou *linguiste*, 'qui étudie les langues'. Si l'élément *-logue* est à classer parmi les suffixes, tout comme *-ien* et *-iste*, *palyn-* ne peut donc pas être considéré comme autre chose qu'un élément lexical, puisque la base d'une opération morphologique ne peut être autre chose qu'un lexème. *Palynologie* instancie le schéma $[X_N\text{logue}]_N \leftrightarrow [\text{spécialiste de X}]$, prenant pour base le lexème POLLEN, ici représenté radicalement par le thème supplétif /palin/. Cette séquence n'est en effet plus reliée à son étymon grec, le verbe *παλύνειν* 'répandre de la farine' : puisque le schéma *Xlogue* prend pour base un nom, la séquence *palyn-* ne peut plus être rattachée au verbe originel mais doit être considérée comme étant un élément de l'espace thématique du lexème nominal POLLEN.

Synchroniquement, en dépit de l'origine allogène des éléments néoclassiques, il n'y a donc pas lieu d'analyser ces éléments comme appartenant à une autre langue. L'existence d'un schéma comme $[X_N\text{logue}]_N \leftrightarrow [\text{spécialiste de X}]$ a été rendue possible par une généralisation qui a été faite à partir des lexèmes existants. La récurrence de lexèmes *Xlogue* référant à des spécialistes de l'objet ou du concept désigné par X a en effet donné lieu à une abstraction de la base X et à l'intégration du schéma dans le lexique, aux côtés des lexèmes et

des autres schémas. La théorie de la Morphologie Constructionnelle (Booij, 2010) apparaît comme particulièrement adaptée pour traiter les différentes formations qui vont sous le nom de composés néoclassiques. Dans cette théorie, la distinction entre lexème et affixe n'est pas posée de manière dichotomique puisque chacun de ces éléments peut apparaître de manière spécifiée dans un schéma. Il n'y a donc pas forcément lieu de poser les éléments néoclassiques comme soit faisant partie de la catégorie des lexèmes, soit de celles des affixes : tout comme les éléments natifs, les éléments allogènes peuvent être spécifiés à l'intérieur d'un schéma.

Une étude extensive des constructions néoclassiques n'avait pas encore été conduite, du moins en français et sur ce type de constructions impliquant des éléments néoclassiques spécifiés. Cette thèse a permis la création d'une base de données, NeoClassy, qui comporte près de 4500 entrées, et qui est mise à la disposition de la communauté scientifique. Elle a également été l'occasion d'amorcer l'application à l'analyse morphologique d'une méthode traditionnellement utilisée en sémantique et en traitement automatique des langues, celle de l'analyse distributionnelle : cette méthode permet de qualifier sémantiquement un mot construit grâce aux contextes d'emploi des mots. En conclusion, les résultats des analyses opérées sur la base de données NeoClassy ont rendu possible l'élaboration d'une typologie des constructions affichant un élément final spécifié, intégrant ainsi non seulement les constructions impliquant des éléments allogènes mais également des schémas dont l'élément spécifié est un élément indigène. Cette typologie est construite sur un grand nombre de critères phonologiques, sémantiques et lexicaux, et permet d'envisager non pas les schémas de construction comme appartenant à des catégories parfaitement différenciées, par exemple des composés d'un côté, des dérivés de l'autre, mais plutôt de rapprocher chacune de ces constructions selon certains critères et de les distinguer selon d'autres, dessinant ainsi un gradient plus que des classes distinctes.

* * *

Ce travail pourra être prolongé dans trois directions principales.

Tout d'abord, je souhaiterais réaliser sur mes données une analyse précise de productivité. En effet, j'ai tout au long de cette thèse parlé de disponibilité des processus de formation étudiés. Cette disponibilité est une notion qualitative (comme le dit Dal (2003)), et ne repose sur aucun calcul précis. Ces calculs de productivité, pour pouvoir être faits, doivent, si l'on se réfère à la méthode de Baayen (Baayen et Lieber, 1991; Baayen, 1992, 1993), être exécutés sur un corpus fini, délimité : la Toile ne peut donc pas servir de base à ces calculs. Ainsi, les calculs de productivité au sens strict proposés

par Baayen divisent, pour l'étude d'un processus morphologique, le nombre d'hapax legomena construits selon ce processus présents dans un corpus par le nombre d'occurrences des types construits par ce même procédé, à l'intérieur du même corpus. Comme le dit Dal (2003), ce calcul de productivité fait aujourd'hui office de standard et d'autres calculs devraient être envisagés pour faire courir ce genre d'analyse sur des données récoltées sur la Toile.

Concernant l'analyse distributionnelle présentée dans le chapitre 4, j'envisage d'effectuer une analyse automatique qui pourrait permettre de comparer le comportement des opérations impliquant des éléments néoclassiques et celles impliquant des affixes prototypiques. Même si cette analyse devra être effectuée sur la base d'un corpus permettant l'autonomisation des requêtes et donc comprendre moins de mots construits, des similitudes sémantiques entre certains schémas pourront être mises en avant grâce à des contextes partagés par différents schémas. On peut imaginer par exemple que le contexte idéologieMOD_X pourra rapprocher des lexèmes *Xphobe* et *antiX*.

Par ailleurs, il serait très intéressant de questionner le statut des éléments néoclassiques d'un point de vue psycholinguistique. Les travaux d'Hélène Giraudo (repris dans, par exemple Giraudo (2015)) ont montré, au travers des expérimentations permettant de contrôler des effets morphologiques et des effets orthographiques, que les thèmes de lexème et les affixes n'ont pas le même effet d'amorçage dans la reconnaissance des mots. Une analyse effectuée sur des lexèmes comme ceux de NeoClassy permettrait de comparer le stockage dans le lexique des éléments néoclassiques par rapport aux lexèmes prototypiques et aux affixes prototypiques. Une analyse de ce type pourra mettre en lumière la différence qui existe entre un élément comme *palyn-* et un autre comme *-logie*.

Je souhaiterais également augmenter les données de NeoClassy et étendre les analyses à d'autres éléments néoclassiques récurrents : les lexèmes en *-lâtre*, *-graphie*, *-phile* ou *-mane* me paraissent par exemple également très intéressants. Le premier est un élément néoclassique qui ne semble plus aujourd'hui disponible pour la création lexicale et qui est, à mon sens, particulièrement proche des affixes. Les lexèmes *Xgraphie* partagent avec les lexèmes en *-logue* le même schème de nomination, celui des noms de spécialistes ; toutefois, l'élément *-graphie* semble conserver une charge sémantique plus importante que l'élément *-logue*. L'élément *-phile* me paraît également très intéressant : il s'oppose tout d'abord à *-phobe* qui est étudié ici et les doublets *Xphile/Xphobe* pourraient être appréhendés de manière plus approfondie. De plus, les formations *Xphile* semblent déployer un éventail de sens très large, se rapprochant par là de l'organisation lexicale des éléments *Xlogue* que j'ai abordés dans cette thèse : 'amateur de X' (*apérophile*), 'ama-

teur (sexuellement) de X' (*asiatophile*), 'collectionneur de X' (*astérixophile*), 'consommateur de X' (*alcoolophile*), dans le vocabulaire de la chimie 'qui est chimiquement attiré par X' (*hydrophile*), 'qui vit dans le milieu X' (*alcalinophile*), 'qui nécessite X pour vivre' (*aérophile*), etc. Enfin, l'élément *-mane*, tout comme l'élément *-phobe* avec l'élément *(-)phobie*, appartient au même réseau lexical que l'élément *(-)manie* qui, lui, peut être employé de façon autonome.

A côté de ces autres éléments néoclassiques finaux, je souhaiterais également étendre mon analyse à des éléments néoclassiques initiaux récurrents : *néo-*, *pseudo-*, *homo-*, *hétéro-* et les éléments servant à l'expression de la dimension du type *méga-*, *macro-*, *micro-*, etc., afin d'évaluer les caractéristiques qu'ils partagent avec les préfixes et de constituer, de la même façon que je l'ai fait dans cette thèse pour les éléments finaux, une typologie des schémas dans lesquels l'élément initial est spécifié ; cette typologie pourra par exemple intégrer des constructions du type *pause X* ou *bébé X* que j'ai évoquées dans le chapitre 5.

Cette thèse a donc permis, à travers l'étude d'éléments spécifiques, une analyse plus large des constructions morphologiques en français. Elle est le premier pas vers une étude plus vaste sur les procédés productifs de construction de lexèmes. Elle amorce ainsi une volonté de traiter les constructions du français de manière unifiée, à l'intérieur d'un modèle théorique qui postule que les locuteurs opèrent des généralisations sur le lexique et partent de ces généralisations mêmes pour créer de nouveaux lexèmes, en s'affranchissant de l'origine ou même du sens premier d'un élément.

Annexes

La base de données NeoClassy

Je liste ici tous les lexèmes classés par élément final puis par ordre alphabétique. La première colonne indique le mot construit, la seconde la base. La troisième colonne indique, dans les cas où la base est représentée par un thème supplétif, le lexème grec ou latin originel. Enfin, figure dans la quatrième colonne la ressource dans laquelle j'ai trouvé le lexème construit : Le Grand Robert (GR), le Trésor de la Langue Française informatisé (TLFi) ou le Web. Il est évident que les mots recensés dans les ressources dictionnaires sont également présentes, en règle générale, sur la Toile. Cette base de données sera bientôt mise en ligne.

Xcide

abeillicide	abeille		Web
acaricide	acarien		GR
adoprépubèreicide	ado prépubère		Web
adulticide	(insecte) adulte		TLFi
africacide	africain / Afrique		Web
âgicide	âge		Web
algicide	algue		GR
alguicide	algue		Web
altruicide	autrui		GR
amibicide	amibe		TLFi
amiticide	amitié		Web
amoricide	amour		Web
amour propicide	amour propre		Web
amouricide	amour		Web
anguicide	serpent	lat. <i>anguis</i>	TLFi
anophélide	anophèle		TLFi
aphagocide	ne mange pas	gr. <i>α-φαγειν</i>	TLFi
arabicide	arabe		Web
arabocide	arabe		Web
arachnicide	araignée	gr. <i>ἀράχνη</i>	Web
araignicide	araignée		Web
arboricide	arbre		TLFi
argenticide	argent		Web
arménicide	arménien / Arménie		Web
ascaricide	ascaride		TLFi
autricide	auteur		Web
autruicide	autrui		GR
avicide	oiseau	lat. <i>avis</i>	Web
avunculicide	oncle	lat. <i>avunculus</i>	Web
bacillicide	bacille		TLFi
bactéricide	bactérie		TLFi, GR
bactériocide	bactérie		Web
beau-matricide	belle-mère		Web

beau-parricide	beau-père		Web
beau-pèricide	beau-père		Web
belle-matricide	belle-mère		Web
belle-mèricide	belle-mère		Web
biocide	vie	gr. <i>βίος</i>	GR
blatticide	blatte		Web
blog(g)icide	blog		Web
blogocide	blog		Web
boeuficide	bœuf		Web
bovicide	bovin		Web
brebicide	brebis		Web
cafardicide	cafard		Web
cafaricide	cafard		Web
campagnolicide	campagnol		Web
canaricide	canard		Web
cancéricide	cancer		TLFi
canicide	chien	lat. <i>canis</i>	Web
cellulicide	cellule		TLFi
cercaricide	cercaire		TLFi
cérébralicide	cerveau		TLFi
chaticide	chat		Web
chatricide	chat		TLFi
chevreuilicide	chevreuil		Web
chienicide	chien		Web
christicide	christ		Web
ciconicide	cigogne	lat. <i>ciconia</i>	TLFi
cimicide	punaise	lat. <i>cimex</i>	TLFi
civicide	civilisation		Web
civilicide	civilisation		Web
civilisicide	civilisation		Web
clochardicide	clochard		Web
clodocide	clodo (clochard)		Web
colléguicide	collègue		Web
colombocide	colombe		Web
communicide	commune		Web
compagnicide	compagnie (aérienne)		Web
congolicide	congolais / Congo		Web
conjointicide	conjoint		Web
conjugicide	conjoint		TLFi
coquicide	coque		TLFi
corbicide	corbeau		Web
coricide	cor		GR
corneilicide	corneille		Web
corvicide	corbeau	lat. <i>corvus</i>	GR
couillonicide	couillon		Web
cousinicide	cousin		Web
crétinicide	crétin		Web
culturicide	culture		Web
culturocide	culture		Web
cytocide	cellule	gr. <i>κύτος</i>	TLFi, GR
déicide	dieu	lat. <i>deus</i>	TLFi, GR

despoticide	despote		Web
dramaticide	dramatique (art)		TLFi
écocide	écosystème		GR
embryonicide	embryon		Web
enfanticide	enfant		TLFi
épithélicide	épithélium		TLFi
épizoïcide	épizoaire		TLFi
escargoticide	escargot		Web
espèceide	espèce		Web
étacide	état		Web
ethnicide	ethnie		Web
ethnocide	ethnie		GR
étournicide	étourneau		Web
eurocide	Europe		Web
familicide	famille (membres de la)		Web
félicide	nourrisson		Web
félinicide	félin		Web
fémicide	femme		Web
filaricide	filaire		TLFi
flicide	fil		GR
flsicide	fls		Web
flavicide	jaune	lat. <i>flavus</i>	TLFi
foeticide	foetus		Web
foncticide	fonctionnaire		Web
fongicide	champignon	lat. <i>fungus</i>	TLFi, GR
fourmicide	fourmi		Web
fratricide	frère	lat. <i>frater</i>	TLFi, GR
gaméticide	gamète		TLFi
gamétocide	gamète		TLFi
gendricide	genre		Web
génitoricide	géniteur		Web
génocide	race	gr. <i>γένος</i>	TLFi, GR
genicide	genre		Web
germicide	germe		TLFi, GR
géronticide	vieux	gr. <i>γέρον</i>	Web
gérontocide	vieux	gr. <i>γέρον</i>	Web
globulicide	globule		TLFi
glottocide	langue	gr. <i>γλώττα</i>	Web
graminicide	graminées		Web
grand-mèricide	grand-mère		Web
grèvicide	grève		Web
groupicide	groupe		Web
guépicide	guépe		Web
gynocide	femme	gr. <i>γυνή</i>	Web
hamstericide	hamster		Web
hélicide	escargot	lat. <i>helix</i>	TLFi
helminthicide	helminthe		TLFi
herbicide	herbe		TLFi, GR
hérissocide	hérisson		Web
hirudicide	sangsue	lat. <i>hirudo</i>	TLFi
historicide	histoire		Web

hobbiticide	hobbit		Web
hollandicide	Hollande (Français)		Web
homicide	homme	lat. <i>homo</i>	TLFi, GR
hutucide	hutu		Web
identicide	identité		Web
idéocide	idée		Web
imagocide	imago		TLFi
infanticide	enfant	lat. <i>infans</i>	TLFi, GR
informaticide	informatique		Web
insecticide	insecte		TLFi, GR
intelligicide	intelligence		Web
intermittenticide	intermittent		Web
irakicide	irakien / Irak		Web
iulicide	iule		TLFi
ivoiricide	ivoirien / Ivoire (Côte d')		Web
judicide	juif		Web
juificide	juif		Web
jumeaucide	jumeau		Web
lampricide	lamproie		Web
lapinicide	lapin		Web
larvacide	larve		Web
larvicide	larve		TLFi, GR
légicide	loi	lat. <i>lex</i>	TLFi
libéricide	enfant	lat. <i>liber</i>	TLFi
liberticide	liberté		TLFi, GR
libricide	livre	lat. <i>liber</i>	TLFi
limacide	limace		Web
limacide	limace		Web
linguocide	langue		Web
listericide	listériose		Web
lombricide	lombric		Web
lombricide	lombric		Web
lutinicide	lutin		Web
magnicide	grand homme	lat. <i>magnus</i>	TLFi, GR
malicide	mal		TLFi
maricide	mari		Web
marinicide	marin		Web
mariticide	mari		Web
marricide	mère		Web
matricide	mère	lat. <i>mater</i>	TLFi, GR
médusicide	méduse		Web
mémoricide	mémoire		Web
menticide	esprit	lat. <i>mens</i>	GR
microbicide	microbe		TLFi, GR
miticide	mite		Web
molluscicide	mollusque		TLFi
molluscide	mollusque		GR
motaricide	motard		Web
mouchicide	mouche		Web
mousticide	moustique		Web
muloticide	mulot		Web

myopicide	myopie		TLFi
nationicide	nation		Web
nécrocide	mort	gr. <i>νεκρός</i>	Web
négrocide	nègres		Web
nématocide	nématodes		TLFi
nématocide	nématodes		TLFi
néocide	nouveau	gr. <i>νέος</i>	TLFi
néonaticide	nouveau-né		Web
népoticide	neveu		TLFi
nervicide	nerf		TLFi
niécicide	nièce		Web
noraicide	Nora		Web
nounoursicide	nounours		Web
odoricide	odeur		Web
oeuficide	œuf		Web
oncicide	oncle		Web
oniricide	rêve	gr. <i>ὄνειρος</i>	TLFi
oursicide	ours		Web
ovicide	œuf	lat. <i>ovum</i>	TLFi
ovinicide	ovin		Web
pacificide	paix		Web
papicide	pape		Web
papyicide	papy		Web
parasiticide	parasite		TLFi, GR
parenticide	parent		Web
parricide	parent		TLFi, GR
passionicide	passion		Web
pastoricide	pasteur		TLFi
pauvricide	pauvres		Web
pédiculicide	poux	lat. <i>pedis</i>	TLFi
pédocide	enfant	gr. <i>παις</i>	Web
pesticide	parasite nuisible	ang. <i>pest</i>	TLFi, GR
peuplicide	peuple		Web
phytocide	plante	gr. <i>φυτάς</i>	Web
pigeonicide	pigeon		Web
pigeonicide	pigeon		Web
piscicide	poisson	lat. <i>piscis</i>	Web
plasmodicide	plasmodium		TLFi
poissonicide	poisson		Web
pomme-de-terricide	pomme-de-terre		Web
populicide	peuple	lat. <i>populus</i>	TLFi
poulicide	poule		Web
présidenticide	président		Web
prêtricide	prêtre		GR
princicide	prince		Web
proficide	prof		Web
protozoicide	protozoaire		TLFi
psychiatricide	psychiatrie		Web
publicide	publicité		Web
pullicide	pull		Web
putocide	pute		Web

putricide	putridité		Web
raticide	rat		TLFi, GR
raton laveuricide	raton laveur		Web
raticide	rat		Web
régicide	roi	lat. <i>rex</i>	TLFi, GR
renardicide	renard		Web
renaricide	renard		Web
républicide	république		Web
rodenticide	rongeur	lat. <i>rodere</i>	GR
rodonticide	rongeur	lat. <i>rodere</i>	GR
romicide	rom		Web
romicide	Rome		Web
rongeuricide	rongeur		Web
rongicide	rongeur		GR
rwandacide	rwandais / Rwanda		Web
sanglicide	sanglier		Web
sangliericide	sanglier		Web
sarkocide	Sarkozy (Nicolas)		Web
scabicide	gale (parasite de la)	lat. <i>scarcoptes scabiei</i>	TLFi
schistosomicide	schistosome		TLFi
schizonticide	schizonte		TLFi
schizontocide	schizonte		TLFi
scorpionicide	scorpion		Web
sémicide	juif		Web
sexocide	sexe		Web
somnicide	sommeil	lat. <i>somnus</i>	TLFi
sorciéricide	sorcière		Web
sororicide	sœur	lat. <i>soror</i>	TLFi
souricide	souris		Web
spermatocide	spermatozoïde		GR
spermatocide	spermatozoïde		TLFi
spermicide	spermatozoïde		TLFi, GR
spirillicide	spirille		TLFi
spirochéticide	spirochète		TLFi
spirochétocide	spirochète		TLFi
sporicide	spore		TLFi
sporontocide	spore		TLFi
sporticide	sport		Web
suicide	soi	lat. <i>sui</i>	TLFi, GR
sujeticide	sujet		Web
sylvicide	forêt	lat. <i>sylva</i>	Web
taupicide	taupe		Web
tauricide	taureau		TLFi
tempslibricide	temps libre		Web
ténicide	ténia		TLFi, GR
terricide	Terre		Web
téticide	tête		Web
tiquicide	tique		Web
tontonicide	tonton		Web
topocide	lieu	gr. <i>τόπος</i>	Web
touticide	tout		Web

toxicicide	toxine		TLFi
tréponémicide	tréponème		TLFi
trollocide	troll		Web
trypanocide	trypanosome		TLFi
tumoricide	tumeur		Web
tutsicide	tutsi		Web
tyrannicide	tyran		TLFi, GR
tzaricide	tzar		Web
urbicide	ville	lat. <i>urbs</i>	Web
uxoricide	épouse	lat. <i>uxor</i>	GR
vermicide	ver	lat. <i>vermis</i>	GR
verrucide	verrue		Web
verticide	verts		Web
vibriocide	vibrion		Web
viesocialicide	vie sociale		Web
virocide	virus		GR
virucide	virus		GR
virulicide	virus		TLFi
zoocide	animal	gr. ζῷος	Web

Xcole

abiéticole	sapin	lat. <i>abies</i>	Web
acéricole	érable	lat. <i>acer</i>	GR
acicole	aciculaire		TLFi
aéricole	air	lat. <i>aer</i>	TLFi, GR
agaricole	agaric		TLFi
agricole	champs	lat. <i>ager</i>	TLFi, GR
agrumicole	agrumes		Web
alpicole	Alpes		TLFi, GR
amnicole	rivière	lat. <i>amnis</i>	TLFi, GR
andicole	Andes		TLFi
apicole	abeille	lat. <i>apis</i>	TLFi, GR
aquacole	eau	lat. <i>aqua</i>	GR
aquicole	eau	lat. <i>aqua</i>	TLFi, GR
arboricole	arbre		TLFi, GR
arénicole	sable	lat. <i>arena</i>	TLFi, GR
argilicole	argile		TLFi
arvicole	champs	lat. <i>arvum</i>	TLFi, GR
ascidicole	ascidie		TLFi
astacicole	écrevisse	lat. <i>astacus</i>	TLFi
avicole	oiseau	lat. <i>avis</i>	TLFi, GR
basicole	basique		Web
bétulicole	bouleau	lat. <i>betula</i>	Web
blatticole	blatte		TLFi
brassicole	brasserie		TLFi, GR
buzeinicole	Buzeins		Web
calcaricole	calc		Web
calcicole	calcaire		TLFi, GR
campanicole	champs	lat. <i>campus</i>	TLFi

canalicole	canal		TLFi
canaricole	canari		Web
capillicole	cheveu		Web
carbonicole	carbonisé		Web
caricole	laïche	lat. <i>carex</i>	TLFi
carpinicole	charme	lat. <i>carpinus</i>	TLFi
castanéicole	chataîgne	lat. <i>castanea</i>	Web
caulicole	tige	lat. <i>caulis</i>	TLFi, GR
cavernicole	caverne		TLFi, GR
cavicole	cavité		TLFi
cédricole	cèdre		Web
célicole	ciel		TLFi
cellicole	cellier		TLFi
cericole	cerise		Web
chresticole	christ		TLFi
cidricole	cidre		TLFi, GR
citricole	citronnier		TLFi
colombicole	colombe		Web
conchicole	coquille	lat. <i>concha</i>	TLFi
conchylicole	coquillage	gr. <i>κογχύλιον</i>	TLFi
confervicole	conferve		TLFi
convolvulicole	liseron	lat. <i>convolvulus</i>	TLFi
cordicole	cœur	lat. <i>cor</i>	TLFi
cornicole	cornouiller		TLFi
corticicole	écorce	lat. <i>cortex</i>	TLFi, GR
corticole	écorce	lat. <i>cortex</i>	Web
cunicole	lapin	lat. <i>cuniculus</i>	GR
cuniculicole	lapin	lat. <i>cuniculus</i>	GR
cupricole	cuivre	lat. <i>cuprum</i>	Web
cuticole	peau	lat. <i>cutis</i>	TLFi, GR
cyprinicole	cyprinidés		Web
cystidicole	vessie	gr. <i>κύστις</i>	TLFi
déicole	dieu	lat. <i>deus</i>	TLFi, GR
démonicole	démon		GR
déserticole	désert		TLFi
détriticole	détritus		TLFi
discicole	disque	lat. <i>discus</i>	TLFi
dulcaquicole	eau douce	lat. <i>dulcis, aqua</i>	TLFi, GR
dulcicole	douce (eau)	lat. <i>dulcis</i>	Web
dumicole	buisson	lat. <i>dumus</i>	TLFi
éricicole	bruyère	lat. <i>erice</i>	TLFi
fagicole	hêtre	lat. <i>fagus</i>	TLFi
fimicole	fumier	lat. <i>fumum</i>	TLFi
floricole	fleur	lat. <i>flos</i>	TLFi, GR
fluvicole	fleuve	lat. <i>fluvius</i>	TLFi
follicole	feuille	lat. <i>folium</i>	TLFi
fongicole	champignon	lat. <i>fungus</i>	GR
fraisicole	fraise		Web
fraxinicole	frêne	lat. <i>frazinus</i>	TLFi
frondicole	fronde		GR
fruiticole	fruit		Web

fruticole	arbrisseau	lat. <i>frutex</i>	TLFi
fucicole	varek	gr. <i>φύκος</i>	TLFi
gallicole	galle		TLFi, GR
gastricole	estomac	gr. <i>γαστήρ</i>	TLFi
graminicole	herbe	lat. <i>gramen</i>	TLFi
granitique	granit		Web
gypsicole	gypse		Web
hélicicole	escargot	lat. <i>helix</i>	Web
herbicole	herbe		GR
hominicole	homme	lat. <i>homo</i>	TLFi
horticole	jardin	lat. <i>hortus</i>	TLFi, GR
hortifruticole	jardin + fruit	lat. <i>hortus</i>	Web
hortofruticole	jardin + fruit	lat. <i>hortus</i>	Web
humicole	humus		TLFi, GR
ignicole	feu	lat. <i>ignis</i>	TLFi, GR
ilicicole	houx	lat. <i>ilex</i>	TLFi
joncicole	jonc		TLFi
kiwicole	kiwi		Web
lapidicole	pierre	lat. <i>lapis</i>	TLFi
larvicole	larve		TLFi
lavandicole	lavande		Web
lemnicole	Lemnos		TLFi
lichenicole	lichen		Web
lignicole	bois	lat. <i>lignum</i>	TLFi, GR
limicole	limon		TLFi, GR
linicole	lin		Web
lombricole	lombric		Web
lunicole	Lune		Web
maïsicole	maïs		GR
méléagricole	pintade	lat. <i>maleagris</i>	Web
merdicole	merde		TLFi
messicole	moisson	lat. <i>messis</i>	Web
monticole	montagne		TLFi, GR
muscicole	mousse	lat. <i>muscus</i>	TLFi
mytilicole	moule	gr. <i>μύτιλος</i>	GR
némoricole	forêt	lat. <i>nemus</i>	TLFi, GR
nidicole	nid		TLFi, GR
nivicole	neige	lat. <i>nives</i>	TLFi
nodicole	nœud	lat. <i>nodus</i>	TLFi
nucicole	noix	lat. <i>nux</i>	Web
oenicole	vin	gr. <i>οἶνος</i>	TLFi
oléicole	olivier	lat. <i>olea</i>	TLFi
ombraticole	ombre		TLFi
orbicole	orbe		TLFi, GR
orduricole	ordure		TLFi
osiéricole	osier		TLFi
ostréicole	huître	lat. <i>ostrea</i>	TLFi, GR
palmicole	palmier		TLFi
paludicole	marais	lat. <i>palus</i>	TLFi
papicole	pape		TLFi
perlicole	perle		Web

pétricole	pierre	lat. <i>etra</i>	TLFi
phréaticole	phréatique		TLFi, GR
pinguicole	gras	lat. <i>pinguis</i>	TLFi
pinicole	pin		TLFi
piscicole	poisson	lat. <i>piscis</i>	TLFi, GR
plébicole	plèbe		TLFi
polycole	plusieurs	gr. <i>πολύς</i>	TLFi
pom(m)icole	pomme		Web
populicole	peuplier	lat. <i>populus</i>	TLFi
praticole	pré	lat. <i>pratium</i>	TLFi
prunicole	prune		Web
quercicole	chêne	lat. <i>quercus</i>	TLFi
radicicole	racine	lat. <i>radix</i>	TLFi, GR
ranicole	grenouille	lat. <i>rana</i>	Web
régnicole	royaume	lat. <i>regnum</i>	TLFi, GR
républicole	république		TLFi
ripicole	rive	lat. <i>ripa</i>	GR
rizicole	riz		TLFi, GR
rubicole	ronce	lat. <i>rubus</i>	TLFi
rupicole	rocher	lat. <i>rupes</i>	TLFi, GR
ruricole	campagne	lat. <i>rus</i>	TLFi
ruscicole	fragon	lat. <i>ruscum</i>	TLFi
rusticole	campagne	lat. <i>rusticus</i>	TLFi
sabulicole	sable	lat. <i>sabulum</i>	TLFi, GR
saccharicole	sucre	gr. <i>σάκχαρ</i>	TLFi
salicicole	saule	lat. <i>salix</i>	TLFi
salicole	sel	lat. <i>sal</i>	TLFi, GR
salmonicole	salmonidés		TLFi
sanguicole	sang		GR
sanguinicole	sang		TLFi
saprolignicole	pourri + bois	lat. <i>saprus, lignum</i>	TLFi
savanicole	savane		Web
saxicole	pierre	lat. <i>saxum</i>	GR
sépicole	haie	lat. <i>sepes</i>	TLFi
séricicole	soie	lat. <i>sericus</i>	TLFi, GR
séricole	soie	lat. <i>sericus</i>	TLFi, GR
serpenticole	serpent		TLFi
silicicole	silice		GR
solénicole	tuyau	gr. <i>σωλήν</i>	TLFi
spongicole	éponge	lat. <i>spongia</i>	TLFi
stagnicole	stagnante (eau)		TLFi
subcorticole	écorce	lat. <i>cortex</i>	TLFi
sylvicole	forêt	lat. <i>sylva</i>	TLFi, GR
tabacole	tabac		Web
taxicole	if	lat. <i>taxus</i>	TLFi
ténébricole	ténèbres		TLFi
terramaricole	terramare		TLFi
terricole	terre		TLFi, GR
thalassicole	thalass	gr. <i>θάλασσα</i>	TLFi, GR
théicole	thé		Web
torrenticole	torrent		TLFi, GR

trufficole	truffe		Web
trui(t)icole	truite		TLFi, GR
tubicole	tube		TLFi, GR
turbicole	tourbe		GR
urbicole	ville	lat. <i>urbs</i>	TLFi
vaginicole	gaine	lat. <i>vagina</i>	TLFi
vasicole	vase		GR
ventricole	ventre		TLFi
vinicole	vin		TLFi, GR
viticole	vigne	lat. <i>vitis</i>	TLFi, GR
vitivinicole	vigne + vin	lat. <i>vitis, vinum</i>	Web
vitulicole	veau	lat. <i>vitellus</i>	TLFi
zostéricole	zostère		TLFi

Xcrate

administocrate	administrateur		Web
administrocrate	administrateur		Web
architectocrate	architecte		Web
aristobobocrate	aristocrate + bobo	(Bourgeois Bohème)	Web
aristocrate		gr. <i>ἀριστοκρατία</i>	TLFi, GR
athéocrate	athée		Web
autocrate	soi	gr. <i>αὐτός</i>	TLFi, GR
avocrate	avocat		Web
bancocrate	banque		TLFi
bandocrate	bande		Web
bobocrate	bobo	(Bourgeois Bohème)	Web
bourreaucrate	bourreau		Web
branlocrate	branleur		Web
bureaucrate	bureau		TLFi, GR
cathocrate	catho(lique)		Web
caviar-bobocrate	caviar + bobo	(Bourgeois Bohème)	Web
censurocrate	censure		Web
chameaucrate	chameau		Web
cleptocrate	voler	gr. <i>κλέπτω</i>	Web
cléocrate	tirage au sort	gr. <i>κλήπος</i>	Web
climatocrate	climat		Web
clitocrate	clitoris		Web
connocrate	con		Web
consocrate	consommateur		Web
copinocrate	copain		Web
corpocrate	corporation		Web
corruptocrate	corruption		Web
crétinocrate	crétin		Web
démocrate	peuple	gr. <i>δῆμος</i>	TLFi, GR
dictocrate	dictateur		Web
dictocrate	dictateur		Web
dixiecrate	Dixie		Web
doxocrate	doxa		Web
dromocrate	course	gr. <i>δρόμος</i>	Web

écolocrate	écologie		Web
éconocrate	économie		Web
egocrate	ego		Web
esclavocrate	esclavagiste		Web
escrocrate	escroc		Web
ethnocrate	peuple	gr. <i>ἔθνος</i>	Web
eurocrate	Europe		Web
expertocrate	expert		Web
farocrate	Faro (bière)		Web
feminocrate	femme		Web
femmocrate	femme		Web
festivocrate	fête		Web
financrate	financier		Web
fonctionnaucrate	fonctionnaire		Web
footocrate	foot		Web
fricocrate	fric		Web
gaspillocrate	gaspiller		Web
gemmocrate	gemme		Web
geniocrate	génie		Web
gérontocrate	vieux	gr. <i>γέρον</i>	TLFi, GR
gourmetocrate	gourmet		Web
gouvernocrate	gouverner		Web
gueulocrate	gueuler		Web
gynécocrate	femme	gr. <i>γυνή</i>	Web
gynocrate	femme	gr. <i>γυνή</i>	Web
historiocrate	historien		Web
humanocrate	humain		Web
iconocrate	icône		Web
idéocrate	idée		Web
idiocrate	idiot		Web
intellocrate	intellectuel		Web
internaucrate	internaute		Web
islamocrate	Islam		Web
jargonocrate	jargon		Web
judéocrate	juif		Web
jurocrate	juriste		Web
krytocrate	juge	gr. <i>κριτής</i>	Web
lobbycrate	lobby		Web
logocrate	discours	gr. <i>λόγος</i>	Web
lotocrate	loto		Web
mafiacrate	mafia		Web
mafocrate	mafia		Web
mageocrate	magicien		Web
magiocrate	mage		Web
magocrate	mage		Web
magouillocrate	magouille		Web
martyrocrate	martyr		Web
mediacrate	média		Web
médiocrate	moyenne (classe)	lat. <i>medius</i>	TLFi
megalocrate	mégalo		Web
mercantocrate	mercantile		Web

merdocrate	merde		Web
méritocrate	mériter		GR
modocrate	modérateur		Web
mollahcrate	mollah		Web
mondialocrate	mondialisation		Web
monnaicrate	monnaie		Web
monocrate	seul	gr. <i>μόνος</i>	Web
narcocrate	narcotique		Web
netocrate	Internet		Web
nucleaiocrate	nucléaire		Web
nucléocrate	nucléaire		Web
ochlocrate	foule	gr. <i>ὄχλος</i>	Web
oligocrate	petit nombre	gr. <i>ὀλίγος</i>	Web
onucrate	ONU		Web
ordinocrate	ordinateur		Web
papicrate	papi		Web
parentocrate	parent		Web
particrate	parti (politique)		Web
partitocrate	parti (politique)		Web
partocrate	parti (politique)		Web
pathocrate	sociopathe/psychopathe		Web
pecnocrate	péquenot		Web
pédagocrate	pédagogue		Web
pédantocrate	pédant		Web
pédocrate	enfant	gr. <i>παῖς</i>	Web
personocrate	personne		Web
phallocrate	phallus		TLFi, GR
pharmacrate	médicament	gr. <i>φάρμακον</i>	Web
phobocrate	peur	gr. <i>φόβος</i>	Web
photocrate	photo		Web
physiocrate	nature	gr. <i>φύσις</i>	TLFi, GR
pipocrate	pipeau	(« c'est du pipeau »)	Web
plantocrate	planteur		Web
ploucocrate	plouc		Web
ploucrate	plouc		Web
ploutocrate	richesse	gr. <i>πλοῦτος</i>	TLFi, GR
pognocrate	pognon		Web
politocrate	politique		Web
pollutocrate	pollution		Web
polycrate	plusieurs	gr. <i>πολύς</i>	Web
pornocrate	porno		TLFi, GR
psychocrate	psychologue		Web
publicrate	publicité		Web
réacrate	réac		Web
rebellocrate	rebelle		Web
ripoucrate	ripou		Web
robinocrate	robin		Web
sarkocrate	Sarkozy		Web
sarkozycrate	Sarkozy		Web
scientocrate	science		Web
sidacrate	sida		Web

sociocrate	société		TLFi
sonocrate	sondage		Web
sophocrate	sagesse	gr. σοφία	Web
statocrate	état	lat. <i>status</i>	Web
stochocrate	stochastique	(hasard)	Web
stratocrate	armée	gr. στρατός	Web
syllabocrate	syllabique (système)		Web
syndicalocrate	syndicat		Web
syndicrate	syndicat		Web
technocrate	technique		TLFi, GR
téléocrate	télévision		TLFi
thalassocrate	mer	gr. θάλασσα	Web
thanatocrate	mort	gr. θάνατος	Web
théâtrocrate	théâtre		Web
théocrate	Dieu	gr. θεός	TLFi, GR
tortiocrate	tortionnaire		Web
trouillocrate	trouille		Web
urbanocrate	urbanisation		Web
vaginocrate	vagin		Web
victimocrate	victime		Web
vidéocrate	vidéo		Web
vinocrate	vin		Web
voyoucrate	voyou		TLFi
voyoutocrate	voyou		Web
vulvocrate	vulve		Web
webocrate	web		Web
zoocrate	animal	gr. ζώος	Web

*X*cratie

actionnariacratie	actionnariat		Web
adhocratie	ad hoc		Web
aidocratie	aide		Web
allah-cratie	allah		Web
alpha-cratie	alpha		Web
androcratie	homme	gr. άνήρ	Web
anglocratie	anglais		Web
anthropocratie	homme	gr. άνθρωπος	Web
antifricocratie	antifric		Web
argencratie	argent		Web
argentocratie	argent		Web
aristocratie		gr. άριστοκρατία	TLFi, GR
athéocratie	athée		Web
autocratie	soi	gr. αυτός	TLFi, GR
avocratie	avocat		Web
ayraultcratie	Ayrault		Web
bancocratie	banque		TLFi
bandocratie	bande		Web
bayroucratie	Bayrou		Web
bibliocratie	livre	gr. βιβλιόν	Web

bibliocratie	Bible		Web
biocratie	vie	gr. βίος	Web
bling-blingocratie	bling-bling		Web
blingocratie	bling-bling		Web
blogocratie	blog		Web
blondocratie	blond		Web
bobocratie	bobo	(Bourgeois Bohème)	Web
bossocratie	bossier		Web
bouddhocratie	Bouddha		Web
bourreaucratie	bourreau		Web
boursocratie	bourse		Web
bureaucratie	bureau		TLFi, GR
bushocratie	Bush		Web
canaillocratie	canaille		TLFi, GR
castocratie	caste		Web
cathocratie	catho(lique)		Web
caviar-bobocratie	caviar + bobo	(Bourgeois Bohème)	Web
célinocratie	Céline		Web
censurocratie	censure		Web
chameaucratie	chameau		Web
chiracocratie	Chirac		Web
christocratie	Christ		Web
cleptocratie	voler	gr. κλέπτω	Web
cléocratie	tirage au sort	gr. κλήπος	Web
clicocratie	clic		Web
climatocratie	climat		Web
clitocratie	clitoris		Web
commissarocratie	commissariat		Web
connocratie	con		Web
conservocratie	conserver		Web
consocratie	consommateur		Web
copinocratie	copain		Web
corpocratie	corporation		Web
corporacratie	corporation		Web
corporatocratie	corporation		Web
corruptocratie	corruption		Web
cosmocratie	cosmos		Web
cretinocratie	crétin		Web
cryptopathocratie	caché + maladie	gr. κρύπτος, παθός	Web
culocratie	cul		Web
cybercratie	cyber		Web
cyberocratie	cyber		Web
délocratie	humiliation	arab. dhollo	Web
democratie	peuple	gr. δῆμος	TLFi, GR
demoncratie	démon		Web
dictacratie	dictateur		Web
dictocratie	dictateur		Web
doxocratie	doxa		Web
dromocratie	course	gr. δρόμος	Web
écolocratie	écologie		Web
econocratie	économie		Web

égocratie	ego		Web
ergocratie	travail	gr. <i>ἔργον</i>	TLFi
érotocratie	érotique		Web
esclavocratie	esclavagiste		Web
escrocratie	escroc		Web
etatocratie	état		Web
ethnocratie	peuple	gr. <i>ἔθνος</i>	Web
ethocratie	mœurs	gr. <i>ἦθος</i>	Web
eugénocratie	bonne famille	gr. <i>ἔυ, γενεά</i>	Web
eurocratie	Europe		Web
expertocratie	expert		Web
feminocratie	femme		Web
femmocratie	femme		Web
festivocratie	fête		Web
filoucratie	filou		Web
financratie	financier		Web
flicocratie	flic		Web
fonctionnaucratie	fonctionnaire		Web
footocratie	foot		Web
francocratie	francs		Web
fricocratie	fric		Web
gemellocratie	jumeau		Web
gémocratie	jumeau		Web
gemmocratie	gemme		Web
geniocratie	génie		Web
gérontocratie	vieux	gr. <i>γέρον</i>	TLFi, GR
gourmetocratie	gourmet		Web
gouvernocratie	gouverner		Web
grammatocratie	grammaire		Web
gueulocratie	gueuler		Web
gynéocratie	femme	gr. <i>γυνή</i>	TLFi, GR
gynocratie	femme	gr. <i>γυνή</i>	Web
gypocratie	vautour	gr. <i>γύπη</i>	Web
humanocratie	humain		Web
hypocritocratie	hypocrite		Web
iconocratie	icône		Web
idéocratie	idée		TLFi
idiocratie	idiot		Web
infocratie	information		Web
intellocratie	intellectuel		Web
Internaucratie	internaute		Web
internetocratie	Internet		Web
islamocratie	Islam		Web
jargonocratie	jargon		Web
je(-)cratie	je		Web
judéocratie	juif		Web
jurocratie	juriste		Web
krytocratie	juge	gr. <i>κριτής</i>	Web
lobbycratie	lobby		Web
logocratie	discours	gr. <i>λόγος</i>	Web
lotocratie	loto		Web

ludocratie	jeu	lat. <i>ludus</i>	Web
macdocratie	macdo(nald's)		Web
mafiacratie	mafia		Web
mafioocratie	mafia		Web
mageocratie	magicien		Web
magioocratie	mage		Web
magocratie	mage		Web
magouillecratie	magouille		Web
magouilloocratie	magouille		Web
marchéocratie	marché		Web
martyrocratie	martyr		Web
médiabobocratie	média + bobo	(Bourgeois Bohème)	Web
mediacratie	média		Web
médiocratie	moyenne (classe)	lat. <i>medius</i>	TLFi, GR
médiocroocratie	médiocre		TLFi
megalocratie	mégalo		Web
mélanocratie	mélanine		Web
mercantocratie	mercantile		Web
merdocratie	merde		Web
méritocratie	mériter		GR
miaoucratie	miaou		Web
minoritairocratie	minoritaire		Web
mitterrandocratie	Mitterrand		Web
modocratie	modérateur		Web
mollahcratie	mollah		Web
mondialocratie	mondialisation		Web
monnaicratie	monnaie		Web
monnaiecratie	monnaie		Web
monocratie	seul	gr. <i>μόνος</i>	TLFi, GR
multicratie	plusieurs	lat. <i>multus</i>	Web
multinationalecratie	multinationale		Web
narcocratie	narcotique		Web
nazi(-)cratie	nazi		Web
netocratie	Internet		Web
neuneucratie	neuneu		Web
nomocratie	loi	gr. <i>νόμος</i>	TLFi
nonnoncratie	non		Web
nucléocratie	nucléaire		Web
ochlocratie	foule	gr. <i>ὄχλος</i>	TLFi, GR
œnocratie	vin	gr. <i>οἴνος</i>	Web
oligocratie	petit nombre	gr. <i>ὀλίγος</i>	Web
onucratie	ONU		Web
ordinocratie	ordinateur		Web
papelarcratie	papelar		Web
papicratie	papi		Web
pardocratie	noir colombien		Web
parentocratie	parent		Web
particratie	parti (politique)		Web
particulocratie	particularisme		Web
partitocratie	parti (politique)		Web
partocratie	parti (politique)		Web

pathocratie	sociopathe/psychopathe		Web
pecnocratie	péquenot		Web
pédagogocratie	pédagogue		Web
pédantocratie	pédant		TLFi
pédocratie	enfant	gr. παῖς	Web
peoplecratie	people		Web
peoplocratie	people		Web
personocratie	personne		Web
phallocratie	phallus		TLFi, GR
pharmacratie	médicament	gr. φάρμακον	Web
phobocratie	peur	gr. φόβος	Web
photocratie	photo		Web
physiocratie	nature	gr. φύσις	TLFi, GR
pillocratie	piller		Web
pipocratie	pipeau	(« c'est du pipeau »)	Web
pipolicratie	people		Web
pipolocratie	people		Web
plantocratie	planteur		GR
ploucocratie	plouc		Web
ploucratie	plouc		Web
ploutocratie	richesse	gr. πλοῦτος	TLFi, GR
pluricratie	plusieurs	lat. <i>plures</i>	Web
pognocratie	pognon		Web
policratie	cité	gr. πόλις	Web
politocratie	politique		Web
pollutocratie	pollution		Web
polycratie	plusieurs	gr. πολὺς	Web
pomocratie	pompom girl		Web
pornocratie	porno		TLFi, GR
psychocratie	psychologue		Web
psychocratie	psychologue		Web
publicratie	publicité		Web
réacratie	réaction		Web
rebellocratie	rebelle		Web
reggaecratie	reggae		Web
ringardocratie	ringard		Web
ripoucratie	ripou		Web
robinocratie	robin		Web
royalocratie	Royal		Web
rue(-)cratie	rue		Web
sacerdocratie	sacerdoce		Web
sarkocratie	Sarkozy		Web
sarkozycratie	Sarkozy		Web
savantocratie	savant		Web
scientocratie	science		Web
ségogocratie	Ségolène		Web
ségolocratie	Ségolène		Web
sentimentalocratie	sentiment		Web
sidocratie	sida		Web
sinocratie	Chine/chinois		Web
sociocratie	société		TLFi, GR

sondageocratie	sondage		Web
sondocratie	sondage		Web
sophocratie	sagesse	gr. σοφία	Web
statocratie	état	lat. <i>status</i>	TLFi
stochocratie	stochastique	(hasard)	Web
stratocratie	armée	gr. στρατός	TLFi
stupidocratie	stupide		Web
syndicalocratie	syndicat		Web
syndicratie	syndicat		Web
synocratie	ensemble	gr. σύν	Web
technocratie	technique		TLFi, GR
télécratie	télévision		TLFi
thalassocratie	mer	gr. θάλασσα	TLFi, GR
thanatocratie	mort	gr. θάνατος	Web
théâtrocratie	théâtre		Web
théocratie	Dieu	gr. θεός	TLFi, GR
tigrocratie	tigre		Web
tortiocratie	tortionnaire		Web
trigonocratie	trigone		Web
trouillocratie	trouille		Web
turcocratie	Turquie/turc		Web
urbanocratie	urbanisation		Web
vaginocratie	vagin		Web
victimocratie	victime		Web
vidéocratie	vidéo		Web
vieillocratie	vieux		Web
vinocratie	vin		Web
viriocratie	homme	lat. <i>vir</i>	Web
voyoucratie	voyou		TLFi, GR
voyoutocratie	voyou		Web
vulvocratie	vulve		Web
webocratie	web		Web
xénocratie	étranger	gr. ξένος	Web
zérocration	zéro		Web
zoocratie	animal	gr. ζώος	Web

Xlogie

abracadabralogie	abracadabra		Web
absinthologie	absinthe		Web
acarologie	acarien		Web
accentologie	accent		Web
accidentologie	accident		Web
accidentologie	accident		GR
acoulogie	entendre	gr. ἀκούω	Web
acribologie	exact	gr. ἀκριβής	Web
acridologie	acridés		Web
acrologie	extrémité	gr. ἄκρος	Web
acropathologie	extrémité + maladie	gr. ἄκρος, πάθος	TLFi
actinologie	rayon	gr. ἀκτίς	TLFi, GR

actologie	acte		Web
addictologie	addiction		GR
adénologie	glande	gr. <i>ἀδὴν</i>	TLFi
aérobiologie	air + vie	gr. <i>ἀήρ, βίος</i>	GR
aérologie	air	gr. <i>ἀήρ</i>	TLFi, GR
africologie	Afrique		Web
afrologie	Afrique		Web
agnotologie	agnosie		Web
agression	agression		Web
agrobiologie	champ + vie	gr. <i>ἀγρός, βίος</i>	GR
agroclimatologie	champ + climat	gr. <i>ἀγρός</i>	Web
agroécologie	champ + habitat	gr. <i>ἀγρός, οἶκος</i>	Web
agrogéologie	champ + terre	gr. <i>ἀγρός, γη</i>	TLFi, GR
agrologie	champ	gr. <i>ἀγρός</i>	TLFi, GR
agrométéorologie	champ + temps	gr. <i>ἀγρός</i>	Web
agropédologie	champ + sol	gr. <i>ἀγρός, πεδόν</i>	TLFi, GR
agrostologie	agrostis		Web
albanologie	Albanie		Web
albuminocytologie	albumine + cellule	gr. <i>κύτος</i>	Web
alcoologie	alcool		GR
alcoologie	alcool		Web
algologie	algue		TLFi, GR
algologie	douleur	gr. <i>ἄλγος</i>	GR
allergologie	allergie		TLFi, GR
alsacologie	Alsace		Web
alzheimerologie	alzheimer		Web
américanologie	américain		Web
amourologie	amour		Web
ampélogie	vigne	gr. <i>ἄμπελος</i>	GR
amphibiologie	amphibie		TLFi
amphibologie	ambigu	gr. <i>ἀμφίβολος</i>	TLFi, GR
amphorologie	amphore		Web
ampoulologie	ampoule (cloque)		Web
analogie		gr. <i>ἀναλογία</i>	TLFi, GR
anatomohistologie	anatomie + tissu	gr. <i>ἴστος</i>	Web
anatomopathologie	anatomie + maladie	gr. <i>ἴστος, πάθος</i>	GR
anatomophysiologie	anatomie + nature	gr. <i>φυσίς</i>	TLFi, GR
anatomoradiologie	anatomie + radio		Web
andrologie	homme	gr. <i>ἀνήπ</i>	GR
anémiologie	vent	gr. <i>ἄνεμος</i>	Web
anesthésiologie	anesthésie		TLFi, GR
angéiologie	vaisseau	gr. <i>ἀγγεῖον</i>	TLFi, GR
angélogie	ange	lat. <i>angelus</i>	TLFi, GR
angéologie	ange	lat. <i>angelus</i>	TLFi, GR
angiologie	vaisseau	gr. <i>ἀγγεῖον</i>	TLFi, GR
anorganologie	a + organe		TLFi
anthologie		gr. <i>ἀνθολογία</i>	TLFi, GR
anthracologie	charbon de bois	gr. <i>ἄνθραξ</i>	Web
anthropobiologie	homme + vie	gr. <i>ἄνθρωπος, βίος</i>	Web
anthropologie	homme	gr. <i>ἄνθρωπος</i>	TLFi, GR
anthroposociologie	homme + société	gr. <i>ἄνθρωπος</i>	Web

anthropozoologie	homme + animal	gr. <i>ἄνθρωπος, ζῴος</i>	TLFi
antilogie		gr. <i>ἀντιλογία</i>	TLFi, GR
aphasiologie	aphasie		Web
aphrodisiologie	aphrodisiaque		Web
apologie	abeille	lat. <i>apis</i>	TLFi
aptéologie	aptère (sans ailes)		Web
aquariologie	aquarium		TLFi
arachnologie	araignée	gr. <i>ἀράχνη</i>	TLFi, GR
aranéologie	araignée	lat. <i>aranea</i>	GR
arbrologie	arbre		Web
archéologie	ancien	gr. <i>ἀρχαῖος</i>	TLFi, GR
archéozoologie	ancien + animal	gr. <i>ἀρχαῖος, ζῴος</i>	Web
archétypologie	archétype		Web
architecturologie	architecture		Web
archivologie	archive		Web
argotologie	argot		Web
arithmologie	nombre	gr. <i>ἀριθμός</i>	TLFi, GR
aromac(h)ologie	arôme		Web
aromalogie	arôme		Web
aromat(h)ologie	arôme		Web
artériologie	artère		TLFi, GR
arthrologie	articulation	gr. <i>ἄρθρον</i>	GR
arthurologie	Arthur		Web
artiologie	art		Web
ascidiologie	ascidie		Web
assyriologie	Assyrie		TLFi, GR
assyrologie	Assyrie		Web
astérosismologie	étoile + secousse	gr. <i>ἀστήρ, σεισμός</i>	GR
asthmologie	asthme		Web
astragalologie	astragale		TLFi
astrobiologie	astre		GR
astrologie	astre		TLFi, GR
athéologie	athéisme		Web
atlantologie	Atlantide		Web
atlasologie	atlas		Web
audiologie	ouïe	lat. <i>audio</i>	GR
audiophonologie	ouïe + voix	lat. <i>audio</i> , gr. <i>φωνή</i>	GR
auriculairologie	auriculaire		Web
auriculologie	oreille	lat. <i>auricula</i>	Web
autécologie	soi + habitat	gr. <i>ἄντος, οἶκος</i>	Web
autoécologie	soi + habitat	gr. <i>ἄντος, οἶκος</i>	Web
autologie	automobile		Web
auxinologie	auxine		GR
auxologie	croissance	gr. <i>αὕξις</i>	Web
avrilologie	avril (1er)		Web
axiologie	valeur	gr. <i>ἄξις</i>	TLFi, GR
bactériologie	bactérie		TLFi, GR
bagadologie	Bagad		Web
baguettologie	baguette		Web
balkanologie	Balkans		Web
balnéologie	bains	lat. <i>balneum</i>	Web

bambouologie	bambou		Web
banliologie	banlieu		Web
bantologie	Bantul		Web
baraminologie	baramine		Web
baratinologie	baratin		Web
barbologie	barbe		Web
barologie	pesanteur	gr. βαρός	TLFi
barologie	bar (lieu)		Web
barthologie	Barthes (Roland)		Web
bascologie	Basque		Web
bateaulogie	bateau		Web
batmanologie	Batman		Web
batrachologie	grenouille	gr. βάτραχος	Web
battologie	Battos		TLFi, GR
bédéologie	BD		Web
bélemnitologie	bélemnite		TLFi, GR
belgologie	Belgique/belge		Web
bendrologie	bendré		Web
benthologie	benthos		Web
berberologie	berbère		Web
bibliologie	livre	gr. βιβλίον	TLFi, GR
bibliopsychologie	livre + esprit	gr. βιβλίον, ψυχή	GR
biéologie	bière		Web
bimbologie	bimbo		Web
biocénologie	biocénose		TLFi
bioclimatologie	vie + climat	gr. βίος	TLFi, GR
biologie	vie	gr. βίος	TLFi, GR
biométéorologie	vie + temps	gr. βίος	GR
biopsychologie	vie + esprit	gr. βίος, ψυχή	GR
biotechnologie	vie + technologie	gr. βίος	GR
biotypologie	vie + type	gr. βίος	TLFi, GR
bistrologie	bistrot		Web
blablatologie	blabla		Web
blaguologie	blague		TLFi
blogologie	blog		Web
blondologie	blonde		Web
bluesologie	blues		Web
bobologie	bobo		Web
botanologie	herbe	gr. βοτάνη	Web
bouddhologie	bouddhisme		Web
bouletologie	boulet		Web
boulettologie	boulette		Web
brachylogie		gr. βραχυλογία	TLFi, GR
bradylogie	lent	gr. βραδύς	Web
brassensologie	Brassens		Web
brassicologie	brassage (bière)		Web
bricabracologie	bric à brac		TLFi
bricolage	bricolage		Web
bromatologie	nourriture	gr. βρώμα	GR
brouillonologie	brouillon		Web
bruitologie	bruit		Web

bryologie	mousse	gr. βρύον	TLFi, GR
bulbologie	bulbe		Web
buzzologie	buzz		Web
byzantinologie	byzantin		GR
cacalogie	caca		Web
cacologie	kakos	gr. κακός	TLFi
caféologie	café		Web
cailloutologie	caillou		Web
calcéologie	chaussure	lat. calceo	Web
calvitologie	calvitie		Web
caméologie	camélidé		Web
campanologie	cloche	lat. campana	TLFi
cancérologie	cancer		TLFi, GR
cannabinologie	cannabis		Web
capillologie	cheveu	lat. capillus	Web
capoeirologie	capoiera		Web
caractériologie	caractère		TLFi, GR
caractérologie	caractère		TLFi, GR
carbuologie	carburant		Web
carcinologie 1	crabe	gr. καρκίνος	TLFi, GR
carcinologie 2	cancer	gr. καρκίνος	TLFi, GR
cardiologie	cœur	gr. καρδιά	TLFi, GR
carphologie	brindille	gr. κάρφη	Web
carpologie	fruit	gr. καρπός	Web
cartologie	carte		Web
caryologie	caryo	gr. κάρυον	TLFi
castellologie	château		Web
caucasologie	Caucase		Web
celtologie	celte		Web
centilogie	cent		TLFi
céramologie	céramique		GR
cérébrologie	cerveau		TLFi
cervologie / cerveaulogie	cerveau		Web
cestpasgravologie	c'est pas grave		Web
cétologie	cétacé		TLFi
champignologie	champignon		Web
chaotologie	chaos		Web
chevillologie	cheville		Web
chiantologie	chiant		Web
chiracologie	Chirac		Web
chirologie	main	gr. χείρ	TLFi, GR
chocologie	chocolat		Web
chondrologie	cartilage	gr. χόνδρος	GR
choréologie	danse	gr. χορεία	Web
chorologie	contrée	gr. χώρα	GR
christologie	Christ		TLFi, GR
chromatologie	couleur	gr. χρώμα	Web
chronobiologie	temps + vie	gr. χρόνος, βίος	GR
chronologie	temps	gr. χρόνος	TLFi, GR
chronopathologie	temps + maladie	gr. χρόνος, πάθος	GR
cidrologie	cidre		Web

cinésiologie	mouvement	gr. κίνησις	Web
classologie	classement		TLFi
clavologie	clou	lat. <i>clavus</i>	Web
climatologie	climat		TLFi, GR
climatopathologie	climat		TLFi
clitologie	clitoris		Web
clodologie	clodo		Web
cocot(t)ologie	cocotte		Web
codicologie	code	lat. <i>codex</i>	TLFi, GR
coenologie	commun	gr. κοινός	TLFi
coenontologie	commun	gr. κοινός	TLFi
coeurologie	cœur		Web
coloproctologie	colon + anus	gr. πρωκτός	Web
colorologie	couleur		Web
comicologie	comics		Web
comitologie	comité		GR
communicologie	communication		Web
conceptologie	concept		Web
conchyliologie	coquillage	gr. κογχύλιον	TLFi, GR
conchyologie	coquille	gr. κόγχος	TLFi
conologie	con		Web
contactologie	contact (lentilles)		GR
coprologie	excrément	gr. κόπρος	TLFi
coptologie	copte		Web
cosmétologie	cosmétique		TLFi, GR
cosmobiologie	cosmos + vie	gr. βίος	GR
cosmologie	cosmos		TLFi, GR
cosmopathologie	cosmos + maladie	gr. πάθος	GR
couacologie	couac		Web
courantologie	courant (marin)		GR
couturologie	couture		Web
crabologie	crabe		Web
cradologie	crade (ce qui est)		Web
crampologie	crampe		Web
craniologie	crâne		TLFi, GR
cranologie	crâne		TLFi
crapaulogie	crapaud		Web
crénologie	source (eau)	gr. κρήνη	GR
crimilogie	crime		Web
criminalogie	criminel		Web
criminologie	criminel		TLFi, GR
critériologie	critère		TLFi, GR
critérologie	critère		Web
criticologie	critique		Web
crustacéologie	crustacé		TLFi
cryologie	froid	gr. κρύος	GR
cryptologie	cryptogramme		TLFi, GR
cryptologie	caché (ce qui est)	gr. κρυπτός	TLFi, GR
cuissologie	cuisse		Web
culologie	cul		Web
culturologie	culture		Web

cyclogie	cycle		Web
cynologie	chien	gr. <i>κύων</i>	TLFi
cytobiologie	cellule + vie	gr. <i>κύτος, κύτος</i>	GR
cytologie	cellule	gr. <i>κύτος</i>	TLFi, GR
cytopathologie	cellule + malade	gr. <i>κύτος, πάθος</i>	GR
cytophysiologie	cellule + nature	gr. <i>κύτος, φύσις</i>	GR
dactylogie	doigt	gr. <i>δάκτυλος</i>	TLFi, GR
dahologie	dahu		Web
décalogie	dix	gr. <i>δέκα</i>	TLFi
décimologie	notation	lat. <i>decimus</i> ('dixième')	Web
décimologue	notation	lat. <i>decimus</i> ('dixième')	Web
déclinologie	déclin		Web
défectologie	défaut	lat. <i>defectio</i>	Web
delphinologie	dauphin	gr. <i>δελφίς</i>	Web
démonologie	démon		TLFi, GR
dendrochronologie	arbre + temps	gr. <i>δένδρον, χρόνος</i>	TLFi, GR
dendrologie	arbre	gr. <i>δένδρον</i>	GR
denimologie	denim		Web
dentistologie	dentiste		Web
denturologie	denture		Web
déontologie	devoir	gr. <i>δεόντως</i>	TLFi, GR
dermato-vénérologie	peau + vénérien	gr. <i>δέρμα</i>	GR
dermatologie	peau	gr. <i>δέρμα</i>	TLFi, GR
derrickologie	Derrick		Web
desmologie	ligament	gr. <i>δεσμός</i>	GR
diabétologie	diabète		TLFi, GR
diablogie	diable		Web
diabologie	diable		Web
dialectologie	dialecte		TLFi, GR
dianologie	Diana		Web
diarrhologie	diarrhée		Web
didactologie	didactique		Web
dieudonnologie	Dieudonné		Web
dingologie	dingue (ce qui est)		Web
dinosaurologie	dinosaure		Web
diptérologie	diptère		TLFi
discologie	disque		Web
divinologie	divin (ce qui est)		Web
docimologie	épreuve	gr. <i>δοκιμή</i>	TLFi
documentologie	document		TLFi
dodologie	dodo		Web
domologie	maison	lat. <i>domus</i>	Web
donjonologie	donjon		Web
dosologie	dose		Web
doxologie		gr. <i>δοξολογία</i>	TLFi, GR
doxologie	doxa		GR
dracologie	dragon	gr. <i>δράκων</i>	Web
draconologie	dragon	gr. <i>δράκων</i>	Web
dragonologie	dragon		Web
drapeaulogie	drapeau		Web
dromologie	course	gr. <i>δρόμος</i>	Web

dylanologie	Dylan (Bob)		Web
dynamologie	force	gr. <i>δύναμις</i>	TLFi
eaunologie	eau		Web
ecclésiologie	église	gr. <i>ἐκκλησία</i>	TLFi
éco-éthologie	habitat + mœurs	gr. <i>οἶκος, ἦθος</i>	GR
écoclimatologie	habitat + climat	gr. <i>οἶκος</i>	GR
écologie	habitat	gr. <i>οἶκος</i>	TLFi, GR
écophysiologie	habitat + nature	gr. <i>οἶκος, φύσις</i>	GR
écotoxicologie	habitat + toxique (ce qui est)	gr. <i>οἶκος</i>	GR
écureuillogie	écureuil		Web
égologie	soi	gr. <i>ἐγώ</i>	TLFi
égyptologie	Egypte		TLFi, GR
électrobiologie	électricité + vie	gr. <i>βίος</i>	GR
électrologie	électricité		TLFi, GR
électrophysiologie	électricité	gr. <i>φύσις</i>	TLFi, GR
électroradiologie	électricité + rayon	lat. <i>radius</i>	TLFi, GR
elficologie	elfe		Web
elvisologie	Elvis		Web
embryologie	embryon		TLFi, GR
emmerdologie	emmerde		Web
émotiologie	émotion		Web
endermologie	endermique (méthode)		Web
endocrinologie	endocrine		TLFi, GR
entomologie	insecte	gr. <i>ἔντομος</i>	TLFi, GR
entreprenologie	entreprenariat		Web
enzymologie	enzyme		TLFi, GR
éphéologie	éphèbe		GR
épidémiologie	épidémie		TLFi, GR
épileptologie	épilepsie		Web
épistémologie	science	gr. <i>ἐπιστήμη</i>	TLFi, GR
épizootologie	épidémie + animal	gr. <i>ζῴος</i>	TLFi
équitologie	équitation		Web
ergologie	travail	gr. <i>ἔργον</i>	TLFi, GR
érotologie	érotique (ce qui est)		TLFi, GR
(h)erpétologie	serpent	gr. <i>έρπετόν</i>	TLFi, GR
escapologie	escapade		Web
escargologie	escargot		Web
eschatologie	fin du monde	gr. <i>ἔσχατος</i> ('extrême')	TLFi, GR
escrologie	escroc		Web
esquimaulogie	esquimau		Web
esthésiologie	sensation	gr. <i>αἴθησις</i>	GR
esthétique	esthétique		Web
ethnobiologie	peuple + vie	gr. <i>ἔθνος, βίος</i>	TLFi, GR
ethnologie	peuple	gr. <i>ἔθνος</i>	TLFi, GR
ethnométhodologie	peuple + méthode	gr. <i>ἔθνος</i>	GR
ethnomusicologie	peuple + musique	gr. <i>ἔθνος</i>	TLFi, GR
ethnopsychologie	peuple + esprit	gr. <i>ἔθνος, ψυχή</i>	GR
ethnozoologie	peuple + animal	gr. <i>ἔθνος, ζῴος</i>	TLFi, GR
éthologie	mœurs	gr. <i>ἦθος</i>	TLFi, GR
étiologie	cause	gr. <i>αἰτία</i>	TLFi, GR
étruscologie	Etrusques		GR

étymologie		gr. <i>ἐτυμολογία</i>	TLFi, GR
eurovisionologie	Eurovision		Web
euthanasologie	euthanasie		Web
eutonologie	bonne tension	gr. <i>ἔυ, τόνος</i>	GR
excusologie	excuse		Web
expologie	exposition		Web
facturologie	facture		Web
fantôologie	fantôme		Web
fessologie	fesse		Web
festologie	fête		Web
fillologie	filles		Web
filmologie	film		TLFi, GR
fixologie	fixation		Web
flamencologie	flamenco		Web
foetologie	foetus		Web
forumologie	forum		Web
francologie	France / français		Web
freudologie	Freud		Web
frontologie	fronts		GR
fruitologie	fruit		Web
futurologie	futur		TLFi, GR
gadgétologie	gadget		Web
gastro-entérologie	estomac + intestin	gr. <i>γαστήρ, ἔντερον</i>	TLFi, GR
gastrologie	estomac	gr. <i>γαστήρ</i>	TLFi, GR
gemmaologie	gemme		TLFi, GR
généalogie		gr. <i>γενεαλογία</i>	TLFi, GR
généthliologie	naissance	gr. <i>γενέθλη</i>	GR
géobiologie	terre + vie	gr. <i>γῆ, βίος</i>	GR
géochronologie	terre + temps	gr. <i>γῆ, χρόνος</i>	TLFi, GR
géologie	terre	gr. <i>γῆ</i>	TLFi, GR
géomorphologie	terre + forme	gr. <i>γῆ, μορφή</i>	TLFi, GR
germinologie	germe	lat. <i>germen</i>	Web
gérontologie	vieux	gr. <i>γέρων</i>	TLFi, GR
gestuologie	gestuelle		Web
girafologie	girafe		Web
girouettologie	girouette		Web
gîtologie	gîte		GR
glaciologie	glacier		TLFi, GR
glossologie	langue	gr. <i>γλῶττα</i>	TLFi, GR
glottochronologie	langue + temps	gr. <i>γλῶττα, χρόνος</i>	GR
glyptologie	Pierre gravée	gr. <i>γλυπτός</i> ('gravé')	GR
gnathologie	mâchoire	gr. <i>γνάθος</i>	TLFi, GR
gnoséologie	connaissance	gr. <i>γνώσις</i>	TLFi, GR
grammatologie	écriture	gr. <i>γράμμα</i>	GR
graphologie	écriture	gr. <i>γράφω</i> ('écrire')	TLFi, GR
grecologie	grec		Web
grottologie	grotte		Web
guzzologie	guzzi		Web
gynécologie	femme	gr. <i>γυνή</i>	TLFi, GR
habitologie	habitat		Web
hagiologie	saint	gr. <i>ἅγιος</i>	TLFi, GR

hamburgerologie	hamburger		Web
hamsterologie	hamster		Web
handicapologie	handicap		Web
haplogie	simple	gr. ἀπλός	TLFi, GR
hébélogie	jeunesse	gr. ἥβη	GR
héliométéorologie	soleil + temps	gr. ἥλιος	GR
héliosismologie	hélio + secousse	gr. ἥλιος, σεισμός	GR
helminthologie	helminthe		TLFi, GR
hématocytologie	sang + cellule	gr. αἷμα, κύτος	GR
hématologie	sang	gr. αἷμα	TLFi, GR
hépatologie	foie	gr. ἥπαρ	TLFi, GR
herbologie	herbe		Web
hérésiologie	hérésie		TLFi, GR
hippologie	cheval	gr. ἵππος	TLFi, GR
histoenzymologie	tissu + enzyme	gr. ἴστος	TLFi
histologie	tissu	gr. ἴστος	TLFi, GR
histopathologie	tissu + maladie	gr. ἴστος, πάθος	TLFi, GR
histophysologie	tissu + nature	gr. ἴστος, φύσις	TLFi
hodologie	route	gr. ὁδός	TLFi
holmesologie	Holmes (Sherlock)		Web
homologie	semblable	gr. ὁμός	TLFi, GR
hontologie	honte		Web
hormonologie	hormone		TLFi, GR
humanologie	humain		Web
humourologie	humour		Web
hydrobiologie	eau + vie	gr. ὕδρων, πάθος	TLFi, GR
hydrogéologie	eau + terre	gr. ὕδρων, γῆ	TLFi, GR
hydrologie	eau	gr. ὕδρων	TLFi, GR
hydrométéorologie	eau + temps	gr. ὕδρων	GR
hygrologie	humidité	gr. ὑγρός	GR
hymnologie	hymne		TLFi, GR
hypnologie	sommeil	gr. ὕπνος	Web
hypologie	en-dessous	gr. ὑπό	GR
hystérologie		gr. ὑστερολογία	TLFi, GR
ichnologie	empreinte	gr. ἴχνος	GR
ichtyolithologie	poisson + pierre	gr. ἰχθύς, λίθος	GR
ichtyologie	poisson	gr. ἰχθύς	TLFi, GR
iconologie	icône		TLFi, GR
idéologie	idée		TLFi, GR
idiologie	individu	gr. ἴδιος	GR
imagologie	imago		Web
imamologie	imam		Web
immunohématologie	immunité		TLFi
immunologie	immunité		TLFi, GR
impactologie	impact		Web
implantologie	implants		GR
indologie	Inde		GR
infectiologie	infection		GR
informatologie	information		Web
initialologie	initiale		Web
injurologie	injure		Web

insectologie	insecte		TLFi, GR
insultologie	insulte		Web
internetologie	internet		Web
interpsychologie	entre + esprit	lat. <i>inter</i> , gr. <i>ψυχή</i>	TLFi, GR
intonologie	intonation		Web
intrapshologie	intérieur + esprit	lat. <i>intra</i> , gr. <i>ψυχή</i>	TLFi
inuitologie	inuit		Web
irakologie	Irak		Web
iranologie	Iran		Web
irénologie	paix	gr. <i>εἰρήνη</i>	Web
iridiologie	iris	gr. <i>ἴρις</i>	Web
iridologie	iris	gr. <i>ἴρις</i>	TLFi, GR
islamologie	Islam		GR
jackyologie	Jacky		Web
japonologie	Japon		Web
jésulogie	Jésus		Web
jésuologie	Jésus		Web
jurologie	juron		GR
kafkalogie	Kafka		Web
karstologie	karst		Web
khmérologie	khmer		Web
kinésiologie	mouvement	gr. <i>κίνησις</i>	Web
kremlinologie	Kremlin		TLFi, GR
kurdologie	Kurde		Web
labyrinthologie	labyrinthe		Web
lacrymologie	larme	lat. <i>lacrima</i>	Web
lalopathologie	parole + maladie	gr. <i>λαλέω</i> ('parler'), <i>πάθος</i>	GR
lapointologie	Lapointe (Bobby)		Web
laryngologie	larynx		TLFi, GR
lépidoptérologie	lépidoptère		GR
léprologie	lèpre		GR
lexicologie	lexique		TLFi, GR
librettologie	libretto		Web
limnobiologie	étang + vie	gr. <i>λίμνη</i> , <i>βίος</i>	GR
limnologie	étang	gr. <i>λίμνη</i>	TLFi, GR
lipidologie	lipide		Web
listologie	liste		Web
lithologie	Pierre	gr. <i>λίθος</i>	TLFi, GR
ludologie	jeu	lat. <i>ludus</i>	Web
luminologie	lumière	lat. <i>lumen</i>	Web
lunologie	Lune		Web
lutinologie	lutin		Web
lymphologie	lymphe		Web
macroclimatologie	macroclimat		TLFi
macrosociologie	grand + société	gr. <i>μακρός</i>	TLFi, GR
magicologie	magie		Web
magmatologie	magma		Web
magnétologie	magnétisme		Web
malacologie	mollusque	gr. <i>μαλάκος</i> ('mou')	TLFi, GR
malariaologie	malaria		Web
malariologie	malaria		GR

malherbologie	mauvaise herbe		GR
mammalogie	mammifère	lat. <i>mammalia</i> ('animaux à mamelle')	TLFi, GR
mammologie	mammifère	lat. <i>mammalia</i> ('animaux à mamelle')	TLFi
manchologie	manchot		Web
mandalalogie	mandala		Web
mangologie	mangue		Web
manuscriptologie	manuscript		GR
mariologie	Marie		TLFi, GR
martyriologie	martyr		Web
martyrologie	martyr		TLFi
marxologie	Marx		TLFi, GR
massologie	massage		Web
mastologie	sein	gr. <i>μαστός</i>	GR
masturbologie	masturbation		Web
matériologie	matériel		GR
maternologie	maternité		GR
mathématologie	mathématiques		Web
médialogie	média		TLFi
médiologie	médiation		Web
mégéthologie	grandeur	gr. <i>μέγεθος</i>	TLFi
merdologie	merde		Web
mésologie	milieu	gr. <i>μέσος</i>	GR
métabiologie	après + vie	gr. <i>μετά, βίος</i>	TLFi
métallogie	métal		Web
métapsychologie	après + esprit	gr. <i>μετά, ψυχή</i>	TLFi, GR
météorologie	temps		TLFi, GR
météoropathologie	temps + maladie	gr. <i>πάθος</i>	TLFi, GR
méthodologie	méthode		TLFi, GR
métrologie	mesure	gr. <i>μέτρον</i>	TLFi, GR
microbiologie	petit + vie	gr. <i>μικρός, βίος</i>	TLFi, GR
microclimatologie	petit + climat	gr. <i>μικρός</i>	TLFi, GR
micrologie	petit (ce qui est)	gr. <i>μικρός</i>	TLFi
microsociologie	petit + société	gr. <i>μικρός</i>	TLFi, GR
mimologie	mime		TLFi, GR
minéralogie	minéral		TLFi, GR
misologie	haine	gr. <i>μίσος</i>	TLFi
missiologie	mission		TLFi
missionologie	mission		Web
mixologie	mixer (action de)		Web
mobilologie	mobile		Web
modemologie	Modem		Web
modernologie	modernité		Web
molinologie	moulin		Web
molysmologie	pollution	gr. <i>μόλυσμα</i> ('souillure')	GR
monacologie	moine		Web
monadologie	monade		TLFi, GR
mondiologie	monde		Web
morphologie	forme	gr. <i>μορφή</i>	TLFi, GR
morphonologie	forme + son	gr. <i>μορφή, φωνή</i>	TLFi, GR

morphophonologie	forme + son	gr. <i>μορφή, φωνή</i>	TLFi, GR
morphopsychologie	forme + esprit	gr. <i>μορφή, ψυχή</i>	TLFi, GR
motivologie	motivation		Web
moulinologie	moulin		Web
moustachologie	moustache		Web
mozartologie	Mozart		Web
murphologie	Murphy (loi de)		Web
muscologie	mousse	lat. <i>muscus</i>	GR
muséologie	musée		TLFi, GR
musicologie	musique		TLFi, GR
mycétologie	champignon	gr. <i>μύκης</i>	TLFi
mycologie	champignon	gr. <i>μύκης</i>	TLFi, GR
myologie	muscle	gr. <i>μῦς</i>	TLFi, GR
myriologie	très nombreux	gr. <i>μυρίος</i>	TLFi
myrmécologie	fourmi	gr. <i>μύρμηξ</i>	TLFi, GR
mystériologie	mystère		TLFi
mythologie	mythe		TLFi, GR
nanardologie	nanar		Web
nanarologie	nanar		Web
narcissologie	narcisse		Web
narcologie	narcotiques		Web
narratologie	narration		TLFi, GR
naturologie	nature		Web
navetologie	navet		Web
nécrologie	mort	gr. <i>νεκρός</i>	TLFi, GR
neigeologie	neige		Web
nématologie	nématodes		TLFi
néologie	nouveau (ce qui est)	gr. <i>νέος</i>	TLFi, GR
néonatalogie	nouveau-né		TLFi
néonatalogie	nouveau-né		TLFi, GR
néphrologie	rein	gr. <i>νεφρός</i>	TLFi, GR
neuneulogie	neuneu		Web
neuroendocrinologie	nerf + endocrines (glandes)	gr. <i>νεῦρον</i>	GR
neurohistologie	nerf + tissu	gr. <i>νεῦρον, ἵστός</i>	GR
neurologie	nerf	gr. <i>νεῦρον</i>	TLFi, GR
neuropathologie	nerf + maladie	gr. <i>νεῦρον, πάθος</i>	GR
neuropharmacologie	nerf + médicament	gr. <i>νεῦρον, φάρμακον</i>	TLFi, GR
neurophysiologie	nerf + nature	gr. <i>νεῦρον, φύσις</i>	TLFi, GR
neuropsychologie	nerf + esprit	gr. <i>νεῦρον, ψυχή</i>	TLFi, GR
névrologie	nerf	gr. <i>νεῦρον</i>	TLFi, GR
nigologie	nigaud		Web
nintendologie	nitendo		Web
nipiologie	enfant	gr. <i>νήπιος</i>	GR
nivologie	neige	lat. <i>nix</i>	TLFi, GR
nombrologie	nombre		Web
nomologie	loi	gr. <i>νόμος</i>	TLFi
noologie	pensée	gr. <i>νόος</i>	TLFi
nosologie	maladie	gr. <i>νόσος</i>	TLFi, GR
nubiologie	Nubie		Web
numérologie	nombre	lat. <i>numerus</i>	GR
nutellalogie	Nutella		Web

nutellatologie	Nutella		Web
obésiologie	obésité		Web
obésologie	obésité		Web
occlusologie	occlusion		Web
océanologie	océan		TLFi, GR
odologie	chant	gr. <i>ὠδή</i>	GR
odonatologie	odonates		Web
odontologie	dent	gr. <i>ὀδούς</i>	TLFi, GR
odontostomatologie	dent + bouche	gr. <i>ὀδούς, στόμα</i>	TLFi, GR
œnologie	vin	gr. <i>οἴνος</i>	TLFi, GR
oignonologie	oignon		Web
oiseaulogie	oiseau		Web
oncologie	masse	gr. <i>ῥγκος</i>	TLFi, GR
onirologie	rêve	gr. <i>ὄνειρος</i>	TLFi, GR
onomasiologie	appellation	gr. <i>ὀνομασία</i>	TLFi, GR
ontologie	être	gr. <i>ὄτος</i>	TLFi, GR
onychologie	ongle	gr. <i>ὄνυξ</i>	Web
oologie	œuf	gr. <i>ὠόν</i>	TLFi, GR
ophiologie	serpent	gr. <i>ὄφης</i>	TLFi, GR
ophtalmologie	œil	gr. <i>ὀφθαλμός</i>	TLFi, GR
orchidologie	orchidée		Web
organologie	organe		TLFi, GR
organologie	instrument de musique	lat. <i>organum</i>	TLFi, GR
orientologie	orient		Web
ornithologie	oiseau	gr. <i>ὄρνις</i>	TLFi, GR
orologie	montagne	gr. <i>ὄρος</i>	TLFi
orthologie	droit (ce qui est)	gr. <i>ὀρθός</i>	TLFi, GR
oryctologie	fossile	gr. <i>ὀρυκτός</i>	TLFi, GR
osmologie	odeur	gr. <i>ὄσμή</i>	TLFi, GR
ostéologie	os	gr. <i>ὀστέον</i>	TLFi, GR
ostracologie	coquille	gr. <i>ὄστρακον</i>	GR
otologie	oreille	gr. <i>οὖς</i>	TLFi, GR
oursologie	ours		Web
ovniologie	OVNI		GR
ovniologie	OVNI		Web
ovologie	œuf	lat. <i>ovum</i>	TLFi, GR
paléanthropologie	ancien + homme	gr. <i>παλαιός, ἄνθρωπος</i>	TLFi
paléoanthropologie	ancien + homme	gr. <i>παλαιός, ἄνθρωπος</i>	Web
paléoarchéologie	ancien + préhistoire	gr. <i>παλαιός, ἀρκαῖος</i>	GR
paléobiologie	ancien + vie	gr. <i>παλαιός, βίος</i>	TLFi, GR
paléoclimatologie	ancien + climat	gr. <i>παλαιός</i>	TLFi, GR
paléocytologie	ancien + cellule	gr. <i>παλαιός, κύτος</i>	GR
paléoécologie	ancien + habitat	gr. <i>παλαιός, οἶκος</i>	TLFi, GR
paléoethnologie	ancien + peuple	gr. <i>παλαιός, ἔθνος</i>	TLFi
paléohistologie	ancien + tissu	gr. <i>παλαιός, ἴστος</i>	GR
paléoneurologie	ancien + nerf	gr. <i>παλαιός, νεῦρον</i>	TLFi
paléontologie	ancien + être	gr. <i>παλαιός, ὄτος</i>	TLFi, GR
paléopalynologie	ancien + pollen	gr. <i>παλαιός, παλύνω</i> ('répandre')	GR
paléopathologie	ancien + maladie	gr. <i>παλαιός, πάθος</i>	GR
paléopsychologie	ancien + esprit	gr. <i>παλαιός, ψυχή</i>	GR
paléospéléologie	ancien + grotte	gr. <i>παλαιός, σπήλαιον</i>	GR

paléozoologie	ancien + animal	gr. <i>παλαιός, ζός</i>	TLFi, GR
palestinologie	Palestine		Web
paethnologie	ancien + peuple	gr. <i>παλαιός, ἔθνος</i>	TLFi
palmologie	palme		Web
paludologie	paludisme		GR
palynologie	pollen	gr. <i>παλύνω</i> ('répandre')	TLFi, GR
papyrologie	papyrus		TLFi, GR
paradigmatologie	paradigme		Web
paradoxologie	paradoxe		TLFi
paralogie	à côté	gr. <i>παρά</i>	GR
parapsychologie	à côté + esprit	gr. <i>παρά, ψυχή</i>	TLFi, GR
parasitologie	parasite		TLFi, GR
parémiologie	proverbe	gr. <i>παροιμία</i>	TLFi, GR
parentologie	parent		Web
parkinsonologie	parkinson		Web
parodontologie	parodonte		TLFi, GR
parthénologie	vierge	gr. <i>παρθένος</i>	GR
parturologie	parturiente		Web
patatologie	patate		Web
pathologie	maladie	gr. <i>πάθος</i>	TLFi, GR
patrimoniologie	patrimoine		Web
patrologie	père	lat. <i>pater</i>	TLFi, GR
pédologie	enfant	gr. <i>παῖς</i>	TLFi, GR
peinologie	peine (de prison)		Web
pékinologie	Pékin		Web
pelviologie	pelvis		Web
pendologie	pendule		Web
pénologie	peine		TLFi, GR
perecologie	Perec		Web
périlogie	autour (ce qui est)	gr. <i>περί</i>	GR
périnatalologie	autour + natalité	gr. <i>περί</i>	TLFi, GR
perinéologie	périnée		Web
périssologie	superflu	gr. <i>περισσός</i>	TLFi, GR
personnologie	personne		GR
pétassologie	pétasse		Web
pétrologie	Pierre	gr. <i>πέτρος</i>	TLFi, GR
peurologie	peur		Web
phallologie	phallus		Web
pharmacologie	médicament	gr. <i>φάρμακον</i>	TLFi, GR
pharmacopsychologie	médicament + esprit	gr. <i>φάρμακον, ψυχή</i>	TLFi, GR
phénologie	phénomène		TLFi, GR
phénoménologie	phénomène		TLFi, GR
philologie	qui aime	gr. <i>φιλέω</i>	TLFi, GR
phlébologie	veine	gr. <i>φλέψ</i>	TLFi, GR
phono-audiologie	son + ouïe	gr. <i>φωνή</i> , lat. <i>audio</i>	GR
phonologie	son	gr. <i>φωνή</i>	TLFi, GR
photobiologie	lumière + vie	gr. <i>φῶς, βίος</i>	GR
photologie	lumière	gr. <i>φῶς</i>	GR
phraséologie	phrase		TLFi, GR
phréatobiologie	phréatique (nappe) + vie	gr. <i>βίος</i>	GR
phrénologie	âme	gr. <i>φρήν</i>	TLFi, GR

phtisiologie	phtisie		TLFi, GR
phycologie	algue	gr. φύκος	TLFi, GR
physiologie	nature	gr. φύσις	TLFi, GR
physiopathologie	nature + maladie	gr. φύσις	TLFi, GR
phytobiologie	plante + vie	gr. φυτόν, φύσις	TLFi, GR
phytoécologie	plante + habitat	gr. φυτόν, οἶκος	GR
phytologie	plante	gr. φυτόν	TLFi
phytopathologie	plante + maladie	gr. φυτόν, πάθος	TLFi, GR
phytosociologie	plante + société	gr. φυτόν	TLFi, GR
piafologie	piaf		Web
picturologie	peinture		Web
pifologie 1	pif		Web
pifologie 2	Pif (Gadget)		Web
pinacologie	tableau	gr. πίναξ	GR
pingouinologie	pingouin		Web
pingouinologie	pingouin		Web
pipeaulogie (pipologie)	pipeau		Web
piscinologie	piscine		Web
pixelologie	pixel		Web
plaisiologie	plaisir		Web
planctonologie	plancton		TLFi, GR
planétologie	planète		TLFi, GR
pluviologie	pluie		Web
pneumatologie	souffle	gr. πνεῦμα	TLFi, GR
pneumologie	souffle	gr. πνεῦμα	TLFi, GR
pneumonologie	souffle	gr. πνεῦμα	TLFi
pneumophtisiologie	souffle + phtisie	gr. πνεῦμα	TLFi, GR
podologie	pied	gr. πούς	TLFi, GR
pogonologie	barbe	gr. πώγων	Web
polémologie	guerre	gr. πόλεμος	TLFi, GR
politicologie	politique		TLFi, GR
politologie	politique		TLFi, GR
pollutologie	pollution		Web
pommologie	pomme		Web
pomologie	fruit	lat. <i>potum</i>	TLFi, GR
pornoologie	porno		Web
positologie	position		Web
posologie	quantité	gr. ποσός ('combien grand')	TLFi, GR
posturologie	posture		Web
potamobiologie	fleuve + vie	gr. ποταμός, βίος	GR
potamologie	fleuve	gr. ποταμός	TLFi, GR
potionologie	potion		Web
praxéologie	praxis		TLFi, GR
prénomologie	prénom		Web
préventologie	prévention		GR
primatologie	primates		GR
princessologie	princesse		Web
problématologie	problématique		TLFi
problémologie	problème		Web
proctologie	anus	gr. πρωκτός	TLFi, GR
prophétologie	prophète		GR

prosopologie	personne	gr. <i>πρόσωπον</i>	TLFi, GR
protistologie	protiste		TLFi, GR
protozoologie	protozoaire		TLFi, GR
psychobiologie	psycho	gr. <i>ψυχή, βίος</i>	TLFi, GR
psychologie	esprit	gr. <i>ψυχή</i>	TLFi, GR
psychomorphologie	esprit + forme	gr. <i>ψυχή, μορφή</i>	TLFi, GR
psychoneurologie	esprit + nerf	gr. <i>ψυχή, νεῦρον</i>	Web
psychopathologie	esprit + maladie	gr. <i>ψυχή, πάθος</i>	TLFi, GR
psychopharmacologie	esprit + médicament	gr. <i>ψυχή, φάρμακον</i>	TLFi, GR
psychophysiology	esprit + nature	gr. <i>ψυχή, φύσις</i>	TLFi, GR
psychopolémologie	esprit + guerre	gr. <i>ψυχή, πόλεμος</i>	GR
psychosociologie	esprit + société	gr. <i>ψυχή</i>	TLFi, GR
ptéridologie	fougère	gr. <i>πτέρις</i>	TLFi
publicitologie	publicité		Web
pulmonologie	poumon		Web
pulsologie	pouls		Web
putologie	pute		Web
pyramidologie	pyramide		Web
pyrétiologie	fièvre	gr. <i>πυρετός</i>	TLFi
pyrologie	feu	gr. <i>πῦρ</i>	Web
quadrilogie	quatre	lat. <i>quattuor</i>	TLFi
qualitologie	qualité		Web
questionologie	question		Web
racaillogie	racaille		Web
racialogie	race		Web
raciologie	race		TLFi, GR
racistologie	raciste		Web
radiobiologie	rayon + vie	lat. <i>radius</i> , gr. <i>βίος</i>	TLFi, GR
radiochronologie	rayon + temps	lat. <i>radius</i> , gr. <i>χρόνος</i>	GR
radioécologie	rayon + habitat	lat. <i>radius</i> , gr. <i>οἶκος</i>	TLFi, GR
radiogéologie	rayon + terre	lat. <i>radius</i> , gr. <i>γῆ</i>	TLFi
radiologie	rayon	lat. <i>radius</i>	TLFi, GR
radiopathologie	rayon + maladie	lat. <i>radius</i> , gr. <i>πάθος</i>	GR
radiumbiologie	radium		GR
ratologie	rater		Web
ratologie	rat		Web
reboutologie	rebouter		Web
réflexologie	réflexe		TLFi, GR
régimologie	régime		Web
relaxologie	relaxation		Web
requinologie	requin		Web
respirologie	respiration		Web
rhabdologie	baguette magique	gr. <i>ράβδος</i>	Web
rhéologie	fluide	gr. <i>ρέον</i>	TLFi, GR
rhinologie	nez	gr. <i>ρίς</i>	TLFi, GR
rhodologie	rose	gr. <i>ρόδον</i>	GR
rhumatologie	rhumatisme		TLFi, GR
rhumologie	rhum		Web
rhumologie	rhum		Web
rienologie	rien		TLFi
rigologie	rigoler		Web

rirologie	rire		Web
risquologie	risque		Web
robotologie	robot		Web
rockologie	rock		Web
romologie	Rom		Web
rudologie	décombres	lat. <i>rudus</i>	GR
runologie	rune		TLFi, GR
russologie	Russe		Web
rwandologie	Rwanda/rwandais		Web
rythmologie	rythme (cardiaque)		Web
sacramentologie	sacramentel		Web
saisonnologie	saison		Web
sarcologie	chair	gr. <i>σάρξ</i>	GR
sarkologie	Sarkozy		Web
saucissologie	saucisse		Web
savourologie	saveur		Web
scannologie	scanner		Web
scatologie	excrément	gr. <i>σκόρος</i>	TLFi, GR
scénologie	scène		TLFi, GR
scientologie	sciences		TLFi, GR
sédimentologie	sédiment		TLFi, GR
séismologie	séisme		TLFi
sélénologie	Lune	gr. <i>Σελήνη</i>	TLFi, GR
sémasiologie	signification	gr. <i>σημαία</i>	TLFi, GR
séméiologue	signe	gr. <i>σημαία</i>	Web
sémiologue	signe	gr. <i>σημαία</i>	TLFi, GR
sémitologie	sémité		Web
sénégalologie	Sénégal		Web
sénologie	sein		TLFi, GR
sensologie	sens		Web
septalogie	sept		Web
sérologie	sérum		TLFi, GR
sexologie	sexe		TLFi, GR
sidologie	sida		GR
siestologie	sieste		Web
sindologie	sindon		GR
sindonologie	sindon		Web
sinologie	chinois		TLFi, GR
sismologie	secousse	gr. <i>σεισμός</i>	TLFi, GR
sitiologie	nourriture	gr. <i>σιτίον</i>	Web
sitologie	site		GR
slipologie	slip		Web
sociobiologie	société + vie	gr. <i>βίος</i>	TLFi, GR
sociologie	société		TLFi, GR
sociopathologie	société + maladie	gr. <i>πάθος</i>	GR
somatologie	corps	gr. <i>σῶμα</i>	TLFi, GR
somnologie	sommeil	lat. <i>somnus</i>	Web
sondologie	sondage		Web
sophrologie	sain d'esprit	gr. <i>σώφρων</i>	TLFi, GR
sotériologie		gr. <i>σωτήριον</i>	TLFi, GR
soudanologie	Soudan		Web

souritologie	souris		Web
soviétologie	soviet		Web
spatiologie	espace		TLFi, GR
spectrologie	spectre		Web
spéléologie	grotte	gr. σπήλαιον	TLFi, GR
spermatologie	sperme		Web
spermiologie	sperme		Web
spiralogie	spirale		Web
splanchnologie	viscères	gr. σπλάγχνον	TLFi, GR
standardologie	standard		Web
statisticologie	statistiques		Web
stomatologie	bouche	gr. στόμα	TLFi, GR
strabologie	strabisme		Web
structurologie	structure		GR
stupidologie	stupide		Web
suaudologie	Suaudeau		Web
subsidologie	subsidier		Web
suggestologie	suggestif		Web
suicidologie	suicide		TLFi, GR
symbologie	symbole		TLFi, GR
sympathologie	sympathique		Web
symptologie	symptôme		Web
symptomatologie	symptôme		TLFi, GR
symptomologie	symptôme		Web
synanthérogologie	synanthéré		TLFi
syndesmologie	ligament	gr. σύνδεσμος	GR
syndicalogie	syndicat		Web
syphilologie	syphilis		GR
tabacologie	tabac		GR
tachylogie	rapide	gr. ταχύς	GR
tarologie	tarot		Web
tautologie	même chose	gr. ταυτό	TLFi, GR
taxilogie	classification	gr. τάξις	GR
taxologie	classification	gr. τάξις	TLFi, GR
technobiologie	technique		TLFi
technologie	technique		TLFi, GR
technopsychologie	technique		TLFi
tectologie	charpente	gr. τεκτονία	GR
téléologie	fin	gr. τέλος	TLFi, GR
tellurologie	tellurique		Web
tendançologie	tendance		Web
téphrochronologie	cendre + temps	gr. τέφρα, χρόνος	GR
tératologie	monstre	gr. τέρας	TLFi, GR
terminologie	terme		TLFi, GR
termitologie	termite		Web
terrariologie	terrarium		Web
testologie	test		TLFi, GR
tétralogie	quatre	gr. τέτταρες	TLFi, GR
textologie	texte		TLFi, GR
texturologie	texture		GR
thanatologie	mort	gr. θάνατος	TLFi, GR

théologie	Dieu	gr. <i>θεός</i>	TLFi, GR
thermologie	chaud (ce qui est)	gr. <i>θερμός</i>	TLFi, GR
thoracologie	thorax		Web
thyrologie	thyroïde		Web
tibétologie	Tibet		Web
tibiologie	tibia		Web
timbrologie	timbre		Web
tintinologie	Tintin		Web
titrologie	titre		Web
tocologie	accouchement	gr. <i>τόκος</i>	GR
tolkienologie	Tolkien		Web
tombologie	tomber		Web
topologie	lieu	gr. <i>τόπος</i>	TLFi, GR
totologie	Toto		Web
toxicologie	toxique (ce qui est)		TLFi, GR
toxinologie	toxine		Web
tracéologie	trace		Web
traductologie	traduction		Web
traficologie	trafic		Web
transitologie	transition		Web
transplantologie	transplanter		Web
traumatologie	blessure	gr. <i>τραῦμα</i>	TLFi, GR
travaillogie	travail		Web
tribologie	frottement	gr. <i>τρίβος</i>	GR
trichologie	poil	gr. <i>θρίξ</i>	GR
trilogie	trois	gr. <i>τρεις</i>	TLFi, GR
trollogie	troll		Web
trophologie	nutrition	gr. <i>τροφή</i>	GR
tropicologie	tropiques		Web
tropologie	trope		TLFi, GR
trouducologie	trou du cul		Web
trouducologie	trou du cul		Web
tubologie	tube		Web
turcologie	turc		Web
typologie	type		TLFi, GR
typtologie	esprit frappeur	gr. <i>τύπτειν</i> ('frapper')	TLFi, GR
tziganologie	Tzigane		Web
ufologie	ufo	ang. unidentified flying object	TLFi, GR
uniformologie	uniforme		Web
uranologie	ciel	gr. <i>οὐρανός</i>	TLFi, GR
urbanologie	urbanisation		GR
urgentologie	urgences		Web
urologie	urine		TLFi, GR
usologie	usages		Web
vaccinologie	vaccin		Web
vaguologie	vague		Web
vampirologie	vampire		Web
vaticanologie	Vatican		Web
vectorologie	vecteur		Web
veinologie	veine		Web
vénérologie	vénérien		TLFi, GR

vénérologie	vénérien		TLFi, GR
venimologie	venin		Web
verbologie	verbe		Web
versaillogie	Versailles		Web
vertébrologie	vertèbre		Web
vexillologie	vexille		GR
vibratologie	vibration		Web
victimologie	victime		TLFi, GR
vietnamologie	Vietnam		Web
vinologie	vin		Web
virologie	virus		TLFi, GR
vitaminologie	vitamine		TLFi, GR
volcanologie	volcan		TLFi, GR
vulcanologie	volcan		TLFi, GR
xylogie	bois	gr. ξύλον	TLFi
zodiacologie	zodiaque		Web
zombilogie	zombie		Web
zombiologie	zombie		Web
zoologie	animal	gr. ζώος	TLFi, GR
zoopsychologie	animal + esprit	gr. ζώος, ψυχή	TLFi, GR
zymologie	levure	gr. ζύμη	TLFi, GR

Xlogue

absinthologue	absinthe		Web
acarologue	acarien		Web
accentologue	accent		Web
accidentologue	accident		Web
acribologie	exact	gr. ἀκριβής	Web
acridologie	acridés		Web
acrologie	extrémité	gr. ἄκρος	Web
addictologue	addiction		GR
aérologue	air	gr. ἀήρ	TLFi
africologue	Afrique		Web
afrologue	Afrique		Web
agroclimatologue	champ + climat	gr. ἀγρός	Web
agroécologue	champ + habitat	gr. ἀγρός, οἶκος	Web
agrométéorologue	champ + temps	gr. ἀγρός	Web
albanologue	Albanie		Web
alcoologue	alcool		GR
alcooologue	alcool		Web
algologue	algue		TLFi, GR
algologue	douleur	gr. ἄλγος	GR
allergologue	allergie		GR
allerlogue	allergie		Web
alsacologue	Alsace		Web
alzheimerologue	alzheimer		Web
américanologue	américain		Web
amourologue	amour		Web
ampélogue	vigne	gr. ἄμπελος	Web

amphorologue	amphore		Web
ampoulogue	ampoule (objet)		Web
analogue		gr. <i>ἀναλογία</i>	TLFi, GR
andrologue	homme	gr. <i>ἀνήπ</i>	GR
anémologue	vent	gr. <i>ἄνεμος</i>	Web
anthologue		gr. <i>ἀνθολογία</i>	TLFi
anthracologue	charbon de bois	gr. <i>ἄνθραξ</i>	Web
anthropobiologue	homme + vie	gr. <i>ἄνθρωπος, βίος</i>	Web
anthropologue	homme	gr. <i>ἄνθρωπος</i>	TLFi, GR
anthroposociologue	homme + société	gr. <i>ἄνθρωπος</i>	Web
anthropozoologue	homme + animal	gr. <i>ἄνθρωπος, ζῴος</i>	Web
antilogue		gr. <i>ἀντίλογος</i>	TLFi
aphasiologue	aphasie		Web
apidologue	abeille	lat. <i>apis</i>	TLFi
apologue		gr. <i>ἀπολόγος</i>	GR
aptérologue	aptère (sans ailes)		Web
aquariologue	aquarium		Web
aqueducologue	aqueduc		Web
arachnologue	araignée	gr. <i>ἀράχνη</i>	TLFi, GR
aranéologue	araignée	lat. <i>aranea</i>	GR
arbrologue	arbre		Web
archéologue	ancien	gr. <i>ἀρχαῖος</i>	TLFi, GR
archéozoologue	ancien + animal	gr. <i>ἀρχαῖος, ζῴος</i>	Web
archétypologue	archétype		Web
architecturologue	architecture		Web
archivologue	archive		Web
argotologue	argot		Web
arithmologue	nombre	gr. <i>ἀριθμός</i>	Web
aromac(h)ologue	arôme		Web
aromat(h)ologue	arôme		Web
artériologue	artère		Web
arthrologue	articulation	gr. <i>ἄρθρον</i>	Web
artiologue	art		Web
ascidiologue	ascidie		Web
assyrientologue	Assyrie		Web
assyriologue	Assyrie		TLFi, GR
assyrologue	Assyrie		Web
astérosismologue	étoile + secousse	gr. <i>ἀστήρ, σεισμός</i>	Web
asthmologue	asthme		Web
astrologue	astre		TLFi, GR
athéologue	athéisme		Web
atlasologue	atlas		Web
audiologue	ouïe	lat. <i>audio</i>	Web
auriculairologue	auriculaire		Web
auriculologue	oreille	lat. <i>auricula</i>	Web
autologue	automobile		Web
auxologue	croissance	gr. <i>αὕξις</i>	Web
axiologue	valeur	gr. <i>ἄξιος</i>	Web
bactériologue	bactérie		TLFi, GR
balkanologue	Balkans		Web
balnéologue	bains	lat. <i>balneum</i>	Web

bamboulogue	bambou		Web
bantulogue	Bantul		Web
baratinologue	baratin		Web
barbologue	barbe		Web
barologue	bar (lieu)		Web
bascologue	Basque		Web
bateaulogue	bateau		Web
batmanologue	Batman		Web
batrachologue	grenouille	gr. <i>βάτραχος</i>	Web
battologue	Battos		Web
bédéologue	BD		Web
bélemnitologue	bélemnite		Web
belgologue	Belgique/belge		Web
bendrologue	bendré		Web
benthologue	benthos		Web
berberologue	berbère		Web
bibliologue	livre	gr. <i>βιβλίον</i>	Web
bibliopsychologue	livre + esprit	gr. <i>βιβλίον, ψυχή</i>	Web
biérologue	bière		Web
bioclimatologue	vie + climat	gr. <i>βίος</i>	Web
biologue	vie	gr. <i>βίος</i>	TLFi, GR
biométéorologue	vie + temps	gr. <i>βίος</i>	Web
biopsychologue	vie + esprit	gr. <i>βίος, ψυχή</i>	Web
biotechnologue	vie + technologie	gr. <i>βίος</i>	Web
biotypologue	vie + type	gr. <i>βίος</i>	Web
bistrologue	bistrot		Web
blaguologue	blague		Web
blogologue	blog		Web
blondologue	blonde		Web
bluesologue	blues		Web
bobologue	bobo		Web
botanologue	herbe	gr. <i>βοτάνη</i>	Web
bouddhologue	bouddhisme		Web
boudinologue	boudin		Web
bouletologue	boulet		Web
boulettologue	boulette		Web
brassensologue	Brassens		Web
brassicologue	brassage (bière)		Web
bricologue	bricolage		Web
bromatologue	nourriture	gr. <i>βρώμα</i>	Web
brouillonologue	brouillon		Web
bruitologue	bruit		Web
bryologue	mousse	gr. <i>βρύον</i>	GR
bulbologue	bulbe		Web
buzzologue	buzz		Web
byzantinologue	byzantin		GR
cacalogue	caca		Web
cacologue	kakos	gr. <i>κακός</i>	Web
caféologue	café		Web
cailloutologue	caillou		Web
calcéologue	chaussure	lat. <i>calceo</i>	Web

calvitologue	calvitie		Web
camélogue	camélidé		Web
campanologue	cloche	lat. <i>campana</i>	Web
canadologue	Canada		Web
canadologue	Canada		Web
cancérologue	cancer		TLFi, GR
canelenologue	canelé		Web
cannabinologue	cannabis		Web
capillologue	cheveu	lat. <i>capillus</i>	Web
caractériologue	caractère		Web
caractérologue	caractère		TLFi, GR
carbuologue	carburant		Web
carcinologue 1	crabe	gr. <i>καρκίνος</i>	TLFi, GR
carcinologue 2	cancer	gr. <i>καρκίνος</i>	TLFi, GR
cardiologue	cœur	gr. <i>καρδιά</i>	TLFi, GR
carpologue	fruit	gr. <i>καρπός</i>	Web
cartologue	carte		Web
castellologue	château		Web
caucasologue	Caucase		Web
celtologue	celte		Web
céramologue	céramique		GR
cervologue / cerveaologue	cerveau		Web
cestpasgravologue	c'est pas grave		Web
cétologue	cétacé		TLFi
champignologue	champignon		Web
chaotologue	chaos		Web
chaussetologue	chaussette		Web
chevillologue	cheville		Web
chiantologue	chiant		Web
chiracologue	Chirac		Web
chirologue	main	gr. <i>χείρ</i>	Web
chocologue	chocolat		Web
choréologue	danse	gr. <i>χορεία</i>	Web
christologue	Christ		Web
chromatologue	couleur	gr. <i>χρῶμα</i>	Web
chronobiologue	temps + vie	gr. <i>χρόνος, βίος</i>	Web
chronologue	temps	gr. <i>χρόνος</i>	TLFi
cidrologue	cidre		Web
cinésiologue	mouvement	gr. <i>κίνησις</i>	Web
climatologue	climat		TLFi, GR
clodologue	clodo		Web
cocot(t)ologue	cocotte		Web
codicologue	code	lat. <i>codex</i>	TLFi, GR
coenologue	Coen (frères)		Web
coeurologue	cœur		Web
coloproctologue	colon + anus	gr. <i>πρωκτός</i>	Web
colorologue	couleur		Web
comicologue	comics		Web
communicologue	communication		Web
comptologue	comptes		Web
conceptologue	concept		Web

conchyliologue	coquillage	gr. κογχύλιον	TLFi, GR
conologue	con		Web
contactologue	contact (lentilles)		Web
coprologue	excrément	gr. κόπρος	Web
coptologue	copte		Web
cosmétologue	cosmétique		TLFi, GR
cosmobiologue	cosmos	gr. βίος	Web
cosmologue	cosmos		TLFi
couacologue	couac		Web
courantologue	courant (marin)		Web
crabologue	crabe		Web
cradologue	crade (ce qui est)		Web
crampologue	crampe		Web
craniologue	crâne		Web
cranologue	crâne		Web
crapaologue	crapaud		Web
crimilogue	crime		Web
criminologue	criminel		TLFi, GR
criticologue	critique		Web
cryologue	froid	gr. κρύος	Web
cryptologue	caché (ce qui est)	gr. κρυπτός	Web
cuirologue	cuir		Web
culologue	cul		Web
culturologue	culture		Web
cyclogue	cycle		Web
cynologue	chien	gr. κύων	Web
cytologue	cellule	gr. κύτος	TLFi
cytopathologue	cellule + malade	gr. κύτος, πάθος	Web
dahulogue	dahu		Web
dakarologue	Dakar		Web
dangérosologue	dangérosité		Web
décalogue	dix	gr. δέκα	TLFi, GR
déclinologue	déclin		Web
défectologue	défaut	lat. defectio	Web
delphinologue	dauphin	gr. δελφίς	Web
démonologue	démon		TLFi, GR
dendrochronologue	arbre + temps	gr. δένδρον, χρόνος	Web
dendrologue	arbre	gr. δένδρον	GR
denimologue	denim		Web
dentistologue	dentiste		Web
denturologue	denture		Web
déontologue	devoir	gr. δέοντως	GR
dermato-vénérologue	peau + vénérien	gr. δέρμα	Web
dermatologue	peau	gr. δέρμα	TLFi, GR
derrickologue	Derrick		Web
diabétologue	diabète		TLFi, GR
dialectologue	dialecte		TLFi, GR
dianologue	Diana		Web
diarrhologue	diarrhée		Web
didactologue	didactique		Web
dingologue	dingue (ce qui est)		Web

dinosaurologue	dinosaure		Web
diptérologue	diptère		Web
divinologue	divin (ce qui est)		Web
docimologue	épreuve	gr. <i>δοκιμή</i>	Web
dodologue	dodo		Web
domologue	maison	lat. <i>domus</i>	Web
dracologue	dragon	gr. <i>δράκων</i>	Web
draconologue	dragon	gr. <i>δράκων</i>	Web
dragonologue	dragon		Web
dromologue	course	gr. <i>δρόμος</i>	Web
dylanologue	Dylan (Bob)		Web
eaunologue	eau		Web
ecclésiologue	église	gr. <i>ἐκκλησία</i>	TLFi, GR
écologue	habitat	gr. <i>οἶκος</i>	TLFi, GR
écophysiologue	habitat + nature	gr. <i>οἶκος, φύσις</i>	Web
écotoxicologue	habitat + toxique (ce qui est)	gr. <i>οἶκος</i>	Web
écureuillologue	écureuil		Web
égologue	soi	gr. <i>ἐγώ</i>	Web
égyptologue	Egypte		TLFi, GR
électrologue	électricité		Web
électrophysiologue	électricité	gr. <i>φύσις</i>	Web
électroradiologue	électricité + rayon	lat. <i>radius</i>	Web
elficologue	elfe		Web
elvisologue	Elvis		Web
embryologue	embryon		Web
emmerdologue	emmerde		Web
émotologue	émotion		Web
endermologue	endermique (méthode)		Web
endocrinologue	endocrine		TLFi, GR
entomologue	insecte	gr. <i>ἔντομος</i>	TLFi, GR
enzymologue	enzyme		Web
épidémiologue	épidémie		Web
épileptologue	épilepsie		Web
épilogue		gr. <i>ἐπίλογος</i>	TLFi, GR
épistémologue	science	gr. <i>ἐπιστήμη</i>	TLFi, GR
équitologue	équitation		Web
ergologue	travail	gr. <i>ἔργον</i>	Web
(h)erpétologue	serpent	gr. <i>έρπετόν</i>	Web
érotologue	érotique (ce qui est)		TLFi, GR
escapologue	escapade		Web
escargologue	escargot		Web
eschatologue	fin du monde	gr. <i>ἔσχατος</i> ('extrême')	Web
escrologue	escroc		Web
esquimaologue	esquimau		Web
esthétologue	esthétique		Web
ethnobiologue	peuple + vie	gr. <i>ἔθνος, βίος</i>	Web
ethnologue	peuple	gr. <i>ἔθνος</i>	TLFi, GR
ethnométhodologue	peuple + méthode	gr. <i>ἔθνος</i>	GR
ethnomusicologue	peuple + musique	gr. <i>ἔθνος</i>	GR
ethnopsychologue	peuple + esprit	gr. <i>ἔθνος, ψυχή</i>	Web
ethnozoologue	peuple + animal	gr. <i>ἔθνος, ζῴος</i>	Web

éthologue	mœurs	gr. ἦθος	TLFi
étiologue	cause	gr. αἰτία	Web
étruscologue	Etrusques		GR
étymologue		gr. ἔτυμολογία	Web
eurovisionologue	Eurovision		Web
euthanasiologue	euthanasie		Web
eutonologue	bonne tension	gr. ἔυ, τόνος	Web
excusologue	excuse		Web
expologue	exposition		Web
facturologue	facture		Web
fantômologue	fantôme		Web
fessologue	fesse		Web
fictionologue	fiction		Web
filmologue	film		GR
fixologue	fixation		Web
flamencologue	flamenco		Web
foetologue	foetus		Web
formatologue	formation		Web
forumologue	forum		Web
francologue	France / français		Web
freudologue	Freud		Web
fruitologue	fruit		Web
futurologue	futur		TLFi, GR
gadgétologue	gadget		GR
galactologue	lait	gr. γάλα	Web
gastro-entérologue	estomac + intestin	gr. γαστήρ, ἔντερον	TLFi, GR
gemmologue	gemme		TLFi, GR
généticologue	génétique		Web
géobiologue	terre + vie	gr. γῆ, βίος	Web
géochronologue	terre + temps	gr. γῆ, χρόνος	Web
géologue	terre	gr. γῆ	TLFi, GR
géomorphologue	terre + forme	gr. γῆ, μορφή	TLFi, GR
gérontologue	vieux	gr. γέρων	TLFi, GR
gestuologue	gestuelle		Web
girafologue	girafe		Web
girouettologue	girouette		Web
gîtologue	gîte		GR
glaciologue	glacier		TLFi, GR
glottochronologue	langue + temps	gr. γλωττα, χρόνος	Web
gnathologue	mâchoire	gr. γνάθος	Web
grammatologue	écriture	gr. γράμμα	Web
graphistologue	graphiste		Web
graphologue	écriture	gr. γράφω ('écrire')	TLFi, GR
grecologue	grec		Web
grottologue	grotte		Web
gynécologue	femme	gr. γυνή	TLFi, GR
habituologue	habitat		Web
hagiologue	saint	gr. ἅγιος	TLFi, GR
hamburgerologue	hamburger		Web
hamsterologue	hamster		Web
handicapologue	handicap		Web

hébélogue	jeunesse	gr. ἡβη	Web
héliométéorologue	soleil + temps	gr. ἥλιος	Web
héliosismologue	hélio + secousse	gr. ἥλιος, σεισμός	Web
helminthologue	helminthe		Web
hématologue	sang	gr. αἷμα	GR
hépatologue	foie	gr. ἥπαρ	TLFi, GR
heptalogue	sept	gr. ἑπτά	Web
herbologue	herbe		Web
hérésiologue	hérésie		TLFi, GR
hétérologue	autre	gr. ἕτερος	TLFi, GR
hippologue	cheval	gr. ἵππος	GR
histologue	tissu	gr. ἵστός	GR
histopathologue	tissu + maladie	gr. ἵστός, πάθος	Web
historiologue	histoire		TLFi
holmesologue	Holmes (Sherlock)		Web
homologue	semblable	gr. ὁμός	TLFi, GR
hontologue	honte		Web
hormonologue	hormone		Web
humanologue	humain		Web
humourologue	humour		Web
hydrobiologue	eau + vie	gr. ὕδρων, πάθος	Web
hydrogéologue	eau + terre	gr. ὕδρων, γῆ	GR
hydrologue	eau	gr. ὕδρων	TLFi, GR
hydrométéorologue	eau + temps	gr. ὕδρων	Web
hymnologue	hymne		TLFi
hypnologue	sommeil	gr. ὕπνος	Web
ichnologue	empreinte	gr. ἵχνος	Web
ichtyologue	poisson	gr. ἰχθύς	TLFi, GR
iconologue	icône		TLFi, GR
idéologue	idée		TLFi, GR
immunohématologue	immunité		Web
immunologue	immunité		Web
impactologue	impact		Web
implantologue	implants		Web
indologue	Inde		GR
infectiologue	infection		Web
informatologue	information		Web
injectologue	injection		Web
injurologue	injure		Web
insectologue	insecte		Web
insultologue	insulte		Web
internetologue	internet		Web
intonologue	intonation		Web
inuitologue	inuit		Web
irakologue	Irak		Web
iranologue	Iran		Web
irénologue	paix	gr. εἰρήνη	Web
iridiologue	iris	gr. ἴρις	Web
iridologue	iris	gr. ἴρις	GR
islamologue	Islam		GR
jackyologue	Jacky		Web

japanologue	Japon		Web
japonologue	Japon		Web
jazzologue	jazz		Web
jurologue	juron		Web
kafkologue	Kafka		Web
karstologue	karst		Web
khmérologue	khmer		Web
kinésiologue	mouvement	gr. <i>κίνησις</i>	Web
kremlinologue	Kremlin		TLFi, GR
kurdologue	Kurde		Web
lacrymologue	larme	lat. <i>lacrima</i>	Web
lapointologue	Lapointe (Bobby)		Web
laryngologue	larynx		TLFi, GR
lépidoptérologue	lépidoptère		Web
léprologue	lèpre		GR
lexicologue	lexique		TLFi, GR
limnologue	étang	gr. <i>λίμνη</i>	Web
linguistologue	linguiste		Web
lipidologue	lipide		Web
listologue	liste		Web
lithologue	Pierre	gr. <i>λίθος</i>	TLFi, GR
ludologue	jeu	lat. <i>ludus</i>	Web
luminologue	lumière	lat. <i>lumen</i>	Web
lunettologue	lunettes		Web
lunologue	Lune		TLFi
lutinologue	lutin		Web
lymphologue	lymphe		Web
macrosociologue	grand + société	gr. <i>μακρός</i>	Web
magicologue	magie		Web
magmatologue	magma		Web
magnétologue	magnétisme		Web
malacologue	mollusque	gr. <i>μαλάκος</i> ('mou')	Web
malariaologue	malaria		GR
malherbologue	mauvaise herbe		Web
mammalogue	mammifère	lat. <i>mammalia</i> ('animaux à mamelle')	Web
mammologue	mammifère	lat. <i>mammalia</i> ('animaux à mamelle')	Web
manchologue	manchet		Web
manuscriptologue	manuscrit		GR
mariologue	Marie		Web
martyrologue	martyr		TLFi
marxologue	Marx		TLFi, GR
massologue	massage		Web
mastologue	sein	gr. <i>μαστός</i>	GR
masturbologue	masturbation		Web
matériologue	matériel		Web
maternologue	maternité		Web
mathématologue	mathématiques		Web
médialogue	média		TLFi
médiologue	médiation		Web

mélogue	chant	gr. μέλος	TLFi
merdologue	merde		Web
météorologue	temps		TLFi, GR
méthodologue	méthode		Web
métrologue	mesure	gr. μέτρον	TLFi, GR
mexicologue	Mexique		Web
microbiologue	petit + vie	gr. μικρός, βίος	Web
microclimatologue	petit + climat	gr. μικρός	Web
microsociologue	petit + société	gr. μικρός	Web
mielologue	miel		Web
minéralogue	minéral		TLFi
misologue	haine		TLFi
missiologue	mission		Web
mixologue	mixer (action de)		Web
modemologue	Modem		Web
modernologue	modernité		Web
modilogue	mode		TLFi
molinologue	moulin		Web
molysmologue	pollution	gr. μόλυσμα ('souillure')	Web
monacologue	Monaco		Web
monadologue	monade		TLFi
mondiologue	monde		Web
monologue	seul	gr. μόνος	TLFi, GR
morphologue	forme	gr. μορφή	TLFi, GR
morphopsychologue	forme + esprit	gr. μορφή, ψυχή	Web
motivologue	motivation		Web
moulinologue	moulin		Web
moustachologue	moustache		Web
mozartologue	Mozart		Web
muscologue	mousse	lat. muscus	Web
muséologue	musée		TLFi, GR
musicologue	musique		TLFi, GR
mycétologue	champignon	gr. μύκης	TLFi
mycologue	champignon	gr. μύκης	TLFi, GR
myologue	muscle	gr. μῦς	Web
myrmécologue	fourmi	gr. μύρμηξ	GR
mythologue	mythe		TLFi, GR
nanardologue	nanar		Web
nanarologue	nanar		Web
narcologue	narcotiques		Web
narratologue	narration		Web
naturologue	nature		Web
navetologue	navet		Web
nécrologue	mort	gr. νεκρός	TLFi, GR
neigeologue	neige		Web
nématologue	nématodes		Web
néologue	nouveau (ce qui est)	gr. νέος	TLFi, GR
néonatalogue	nouveau-né		Web
néonatologue	nouveau-né		Web
néphrologue	rein	gr. νεφρός	TLFi, GR
neuneulogue	neuneu		Web

neuroendocrinologue	nerf + endocrines (glandes)	gr. νεῦρον	Web
neurologue	nerf	gr. νεῦρον	TLFi, GR
neuropathologue	nerf + maladie	gr. νεῦρον, πάθος	Web
neuropharmacologue	nerf + médicament	gr. νεῦρον, φάρμακον	Web
neurophysiologue	nerf + nature	gr. νεῦρον, φύσις	Web
neuropsychologue	nerf + esprit	gr. νεῦρον, ψυχή	TLFi
névrologue	nerf	gr. νεῦρον	TLFi
nintendologue	nitendo		Web
nivologue	neige	lat. <i>nix</i>	GR
nombrologue	nombre		Web
nomologue	loi	gr. νόμος	Web
noologue	pensée	gr. νόος	Web
nosologue	maladie	gr. νόσος	Web
nubiologue	Nubie		Web
numérologue	nombre	lat. <i>numerus</i>	Web
nutellalogue	Nutella		Web
nutellatologue	Nutella		Web
obésologue	obésité		GR
obésologue	obésité		GR
occlusologue	occlusion		Web
océanologue	océan		TLFi, GR
odologue	chant	gr. ᾠδή	Web
odonatologue	odonates		Web
odontologue	dent	gr. ὀδούς	Web
odontostomatologue	dent + bouche	gr. ὀδούς, στόμα	Web
œnologie	vin	gr. οἶνος	TLFi, GR
oiseaulogue	oiseau		Web
oncologue	masse	gr. ὄγκος	GR
onirologue	rêve	gr. ὄνειρος	TLFi, GR
onomasiologue	appellation	gr. ὀνομασία	Web
ontologue	être	gr. ὄντος	TLFi
oologie	œuf	gr. ὄον	Web
ophiologue	serpent	gr. ὄφης	Web
ophtalmologue	œil	gr. ὀφθαλμός	TLFi, GR
orchidologue	orchidée		Web
organologue	instrument de musique	lat. <i>organum</i>	Web
orientologue	orient		Web
ornithologue	oiseau	gr. ὄρνις	TLFi, GR
oryctologue	fossile	gr. ὄρυκτός	Web
osmologue	odeur	gr. ὀσμὴ	Web
ostéologue	os	gr. ὀστέον	Web
otologue	oreille	gr. οὖς	TLFi
oulipologue	Oulipo		Web
oursologue	ours		Web
ovnilogue	OVNI		Web
ovniologue	OVNI		Web
paléanthropologue	ancien + homme	gr. παλαιός, ἄνθρωπος	Web
paléoanthropologue	ancien + homme	gr. παλαιός, ἄνθρωπος	Web
paléoarchéologue	ancien + préhistoire	gr. παλαιός, ἀρκαῖος	Web
paléoclimatologue	ancien + climat	gr. παλαιός	Web
paléoécologue	ancien + habitat	gr. παλαιός, οἶκος	Web

paléoethnologue	ancien + peuple	gr. παλαιός, ἔθνος	Web
paléohistologue	ancien + tissu	gr. παλαιός, ἰστός	Web
paléologue	ancien (ce qui est)	gr. παλαιός	GR
paléoneurologue	ancien + nerf	gr. παλαιός, νεῦρον	Web
paléontologue	ancien + être	gr. παλαιός, ὄτος	TLFi, GR
paléopalynologue	ancien + pollen	gr. παλαιός, παλύνω ('répandre')	GR
paléopathologue	ancien + maladie	gr. παλαιός, πάθος	Web
paléopsychologue	ancien + esprit	gr. παλαιός, ψυχή	Web
paléozoologue	ancien + animal	gr. παλαιός, ζῴος	Web
palestinologue	Palestine		Web
paethnologue	ancien + peuple	gr. παλαιός, ἔθνος	Web
palmologue	palme		Web
paludologue	paludisme		GR
palynologue	pollen	gr. παλύνω ('répandre')	GR
papyrologue	papyrus		TLFi, GR
parapsychologue	à côté + esprit	gr. παρά, ψυχή	Web
parasitologue	parasite		TLFi, GR
parémiologue	proverbe	gr. παροιμία	GR
parentologue	parent		Web
parfumologue	parfum		Web
parkinsonologue	parkinson		Web
parodontologue	parodonte		GR
parturologue	parturiente		Web
patatologue	patate		Web
pathologue	maladie	gr. πάθος	TLFi
pédagogue	pédagogie		Web
pédologue	enfant	gr. παῖς	TLFi, GR
pékinologue	Pékin		GR
pendologue	pendule		Web
pénologue	peine		GR
perecologue	Perec		Web
périnatalogue	autour + natalité	gr. περί	Web
pétrologue	Pierre	gr. πέτρος	Web
peurologue	peur		Web
pharmacologue	médicament	gr. φάρμακον	TLFi, GR
pharmacopsychologue	médicament + esprit	gr. φάρμακον, ψυχή	Web
phénologue	phénomène		Web
phénoménologue	phénomène		TLFi, GR
philologue	qui aime	gr. φιλέω	TLFi, GR
phlébologue	veine	gr. φλέψ	GR
phonologue	son	gr. φωνή	TLFi, GR
phraséologue	phrase		TLFi
phrénologue	âme	gr. φρήν	TLFi, GR
phtisiologue	phtisie		TLFi, GR
phycologue	algue	gr. φῦκος	GR
physiologue	nature	gr. φύσις	TLFi, GR
phytoécologue	plante + habitat	gr. φυτόν, οἶκος	Web
phytologue	plante	gr. φυτόν	Web
phytopathologue	plante + maladie	gr. φυτόν, πάθος	Web
phytosociologue	plante + société	gr. φυτόν	Web
piafologue	piaf		Web

pietologue	pied		Web
pifologue 1	pif		Web
pifologue 2	Pif (Gadget)		Web
pingouinlogue	pingouin		Web
pingouinologue	pingouin		Web
pipeaulogue (pipologue)	pipeau		Web
piscinologue	piscine		Web
pixelologue	pixel		Web
plaisiologue	plaisir		Web
planctonologue	plancton		Web
planétologue	planète		TLFi, GR
pluviologue	pluie		Web
pneumatologue	souffle	gr. πνεῦμα	Web
pneumologue	souffle	gr. πνεῦμα	TLFi, GR
pneumonologue	souffle	gr. πνεῦμα	Web
pneumophtisiologue	souffle + phtisie	gr. πνεῦμα	GR
podologue	pied	gr. πούς	GR
pogonologue	barbe	gr. πάγων	Web
poirologue	poire		Web
polémologue	guerre	gr. πόλεμος	GR
politicologue	politique		TLFi, GR
politologue	politique		TLFi, GR
pommologue	pomme		Web
pomologue	fruit	lat. <i>pomum</i>	TLFi, GR
pornologue	porno		Web
positologue	position		Web
posturologue	posture		Web
potamologue	fleuve	gr. ποταμός	Web
potionologue	potion		Web
prénomminologue	prénom		Web
préventologue	prévention		GR
primatologue	primates		GR
princessologue	princesse		Web
problémologue	problème		Web
proctologue	anus	gr. πρωκτός	TLFi, GR
prologue		gr. πρόλογος	TLFi, GR
protistologue	protiste		TLFi, GR
psychologue	esprit	gr. ψυχή	TLFi, GR
psychomorphologue	esprit + forme	gr. ψυχή, μορφή	Web
psychoneurologue	esprit + nerf	gr. ψυχή, νεῦρον	TLFi
psychopathologue	esprit + maladie	gr. ψυχή, πάθος	Web
psychopharmacologue	esprit + médicament	gr. ψυχή, φάρμακον	GR
psychophysiologue	esprit + nature	gr. ψυχή, φύσις	TLFi
psychosociologue	esprit + société	gr. ψυχή	TLFi, GR
ptéridologue	fougère	gr. πτέρις	Web
publicitologue	publicité		Web
pulmonologue	poumon		Web
pulsologue	pouls		Web
putologue	pute		Web
pyramidologue	pyramide		Web
pyrologue	feu	gr. πῦρ	Web

qualitologue	qualité		Web
questionologue	question		Web
racaillogue	racaille		Web
raciologue	race		Web
racistologue	raciste		Web
radiochronologue	rayon + temps	lat. <i>radius</i> , gr. <i>χρόνος</i>	Web
radiologue	rayon	lat. <i>radius</i>	TLFi, GR
ratologue	rat		Web
reboutologue	rebouter		Web
réflexologue	réflexe		Web
régimologue	régime		Web
relaxologue	relaxation		Web
requinologue	requin		Web
respirologue	respiration		Web
rhéologue	fluide	gr. <i>ῥέον</i>	GR
rhinologue	nez	gr. <i>ῥίς</i>	Web
rhodologue	rose	gr. <i>ῥόδον</i>	Web
rhumatologue	rhumatisme		TLFi, GR
rhumologue	rhum		Web
rhumologue	rhume		Web
rienologue	rien		TLFi
rigologue	rigoler		Web
rirologue	rire		Web
risquologue	risque		Web
robotologue	robot		Web
rockologue	rock		Web
romologue	Rom		Web
rudologue	décombres	lat. <i>rudus</i>	Web
runologue	rune		TLFi
russologue	Russe		Web
rwandologue	Rwanda/rwandais		Web
rythmologue	rythme (cardiaque)		Web
saisonologue	saison		Web
sarcologue	chair	gr. <i>σάρξ</i>	Web
sarkologue	Sarkozy		Web
saucissologue	saucisse		Web
scannologue	scanner		Web
scatologue	excrément	gr. <i>σκώρ</i>	TLFi, GR
scénologue	scène		Web
scientologue	sciences		TLFi
sédimentologue	sédiment		TLFi, GR
séismologue	séisme		TLFi
sélénologue	Lune	gr. <i>Σελήνη</i>	GR
séméiologie	signe	gr. <i>σημαία</i>	TLFi
sémiologie	signe	gr. <i>σημαία</i>	TLFi, GR
sémitologue	sémite		TLFi
sénégalologue	Sénégal		Web
sénologue	sein		GR
sensologue	sens		Web
sérologue	sérum		Web
sexologue	sexe		TLFi, GR

sidénologue	sida		TLFi
sidologue	sida		TLFi, GR
siestologue	sieste		Web
sindologue	sindon		Web
sindonologue	sindon		Web
sinologue	chinois		TLFi, GR
sismologue	secousse	gr. <i>σεισμός</i>	TLFi, GR
sitologue	site		GR
slipologue	slip		Web
sociologue	société		TLFi, GR
sociopathologue	société + maladie	gr. <i>πάθος</i>	Web
somatologue	corps	gr. <i>σῶμα</i>	Web
somnologue	sommeil	lat. <i>somnus</i>	Web
sondologue	sondage		Web
sophrologue	sain d'esprit	gr. <i>σώφρων</i>	TLFi, GR
sotériologue		gr. <i>σωτήριοι</i>	Web
soudanologue	Soudan		Web
souritologue	souris		Web
soviétologue	soviet		TLFi, GR
spatiologue	espace		GR
spectrologue	spectre		Web
spéléologue	grotte	gr. <i>σπήλαιον</i>	TLFi, GR
spermatologue	sperme		Web
statisticologue	statistiques		Web
stomatologue	bouche	gr. <i>στόμα</i>	TLFi, GR
strabologue	strabisme		Web
stupidologue	stupide		Web
suaudologue	Suaudeau		Web
suggestologue	suggestif		Web
suicidologue	suicide		Web
symbologue	symbole		Web
symptomatologue	symptôme		Web
syndicalogue	syndicat		Web
syphilologue	syphilis		Web
tabacologue	tabac		GR
tarologue	tarot		Web
tautologue	même chose	gr. <i>ταυτό</i>	GR
taxologue	classification	gr. <i>τάξις</i>	GR
technologue	technique		TLFi, GR
téléologue	fin	gr. <i>τέλος</i>	Web
tendançologue	tendance		Web
tératologue	monstre	gr. <i>τέρας</i>	TLFi, GR
terminologue	terme		TLFi, GR
termitologue	termite		Web
terrariologue	terrarium		Web
testologue	test		Web
textologue	texte		Web
thanatologue	mort	gr. <i>θάνατος</i>	GR
théologue	Dieu	gr. <i>θεός</i>	TLFi
thermologue	chaud (ce qui est)	gr. <i>θερμός</i>	Web
thoracologue	thorax		Web

thyrologue	thyroïde		TLFi
tibétologue	Tibet		Web
timbrologue	timbre		Web
tintinologue	Tintin		Web
titrologue	titre		Web
tolkienologue	Tolkien		Web
topologue	lieu	gr. τόπος	GR
totologue	Toto		Web
toxique	toxique (ce qui est)		TLFi, GR
toxinologue	toxine		Web
tracéologue	trace		Web
traductologue	traduction		Web
traficologue	trafic		Web
transitologue	transition		Web
transplantologue	transplanter		Web
traumatologue	blessure	gr. τραῦμα	TLFi, GR
travailleologue	travail		Web
tribologue	frottement	gr. τρίβος	GR
trichologue	poil	gr. θρίξ	GR
trophologue	nutrition	gr. τροφή	Web
tropologue	trope		Web
trouducologue	trou du cul		Web
trouducologue	trou du cul		Web
tuilologue	tuile		Web
turcologue	turc		Web
twingologue	Twingo		Web
typologue	type		Web
typtologue	esprit frappeur	gr. τύπτειν ('frapper')	TLFi
tziganologue	Tzigane		Web
ufologue	ufo	ang. unidentified flying object	TLFi, GR
uniformologue	uniforme		Web
urbanologue	urbanisation		GR
urgentologue	urgences		Web
urologue	urine		TLFi, GR
usologue	usages		Web
vaccinologue	vaccin		Web
vaguologue	vague		Web
vampirologue	vampire		TLFi
vaticanologue	Vatican		TLFi
vectorologue	vecteur		Web
veinologue	veine		Web
vénérologue	vénérien		TLFi
vénérologue	vénérien		TLFi, GR
venimologue	venin		Web
verbologue	verbe		Web
vertébrologue	vertèbre		Web
vexillologue	vexille		GR
vibratologue	vibration		Web
victimologue	victime		Web
vietnamologue	Vietnam		Web
vinologue	vin		Web

virologue	virus		TLFi, GR
vitaminologue	vitamine		Web
volcanologue	volcan		TLFi, GR
vulcanologue	volcan		TLFi, GR
xylogue	bois	gr. ξύλον	Web
zenologue	zen		Web
zodiacologue	zodiaque		Web
zombilogue	zombie		Web
zombiologue	zombie		Web
zoologue	animal	gr. ζώος	TLFi, GR
zoopsychologue	animal + esprit	gr. ζώος, ψυχή	GR
zymologue	levure	gr. ζύμη	Web

Xphage

acanthophage	épine	gr. ἄκανθος	TLFi
acridophage	sauterelle	gr. ἀκρίς	TLFi
actinophage	actinomycète		Web
aériphage	air	gr. ἀήρ	TLFi
aérophage	air	gr. ἀήρ	TLFi
algophage	algue		TLFi
amourophage	amour		Web
ampélophage	vigne	gr. ἄμπελος	TLFi, GR
androphage	homme	gr. ἀνήπ	Web
anthophage	fleur	gr. ἄνθος	GR
anthropophage	homme	gr. ἄνθρωπος	TLFi, GR
arachnophage	araignée	gr. ἀράχνη	Web
arrhénoophage	mâle	gr. ἀρρενικός	Web
arsénicophage	arsenic		TLFi
artophage	pain	gr. ἄρτος	Web
artophage 2	art		Web
autophage	soi	gr. αὐτός	TLFi, GR
bactériophage	bactérie		TLFi, GR
basiléophage	roi	gr. βασιλεύς	TLFi
bédéphage	BD		Web
benthophage	benthos		Web
bibliophage	livre	gr. βιβλίον	GR
biophage	vie	gr. βίος	TLFi, GR
bouquinophage	bouquin		Web
bovinophage	bovin		Web
bovinphage	bovin		Web
brasilophage	Brésil		Web
brocophage	brocoli		Web
cacophage	caca		Web
carnophage	chair	lat. caro	Web
carpophage	carp	gr. καρπός	TLFi
caryophage	noyau (cellule)	gr. κάρυον	Web
cellulophage	cellule		Web
céphalophage	tête	gr. κεφαλή	Web
cerveauphage	cerveau		Web

chélonophage	tortue	gr. <i>χελωνός</i>	GR
choucrouthophage	choucroute		Web
chromophage	couleur	gr. <i>χρῶμα</i>	Web
chronophage	temps	gr. <i>χρόνος</i>	TLFi, GR
cinéphage	cinéma		GR
climatophage	climat		Web
coccophage	graine	gr. <i>κόκκος</i>	TLFi
conchyliophage	coquillage	gr. <i>κογχύλιον</i>	Web
coprophage	excrément	gr. <i>κόπρος</i>	TLFi, GR
cordophage	corde		Web
cosmétophage	cosmétique		Web
créophage	viande	gr. <i>κρέας</i>	GR
crépophage	crêpe		Web
culturophage	culture		Web
cyanophage	cyanobactéries		Web
cynophage	chien	gr. <i>κύων</i>	Web
dendrophage	arbre	gr. <i>χρόνος</i>	TLFi, GR
détritiphage	détritus		Web
discophage	disque		Web
dvdphage	DVD		Web
électrophage	électricité		Web
endophage	en dedans	gr. <i>ἔνδον</i>	TLFi, GR
entomophage	insecte	gr. <i>ἔντομος</i>	TLFi, GR
érythrophage	érythrocyte		TLFi
europophage	Europe		Web
euryphage	large	gr. <i>εὐρύς</i>	GR
francophage	France/français		Web
fumophage	fumée		Web
galactophage	lait	gr. <i>γάλα</i>	TLFi, GR
géophage	terre	gr. <i>γῆ</i>	TLFi, GR
glottophage	langue	gr. <i>γλῶττα</i>	Web
glyciphage / glycyphage	sucre	gr. <i>γλυκύς</i>	TLFi
gynophage	femme	gr. <i>γυνή</i>	Web
harpactophage	capturée (proie)	gr. <i>ἄρπακτός</i>	Web
hématophage	sang	gr. <i>αἷμα</i>	TLFi, GR
hémophage	sang	gr. <i>αἷμα</i>	TLFi
herpétophage	serpent	gr. <i>ἕρπετόν</i>	Web
heurophage	heure		Web
hippophage	cheval	gr. <i>ἵππος</i>	GR
histophage	tissu	gr. <i>ἰστός</i>	Web
homophage	semblable	gr. <i>ὁμός</i>	GR
humanophage	humain		Web
hyperphage	beaucoup	gr. <i>ὑπέρ</i>	Web
ichthyophage	poisson	gr. <i>ἰχθύς</i>	TLFi, GR
infophage	information		Web
judéophage	juif		TLFi
larviphage	larve		TLFi
librophage	livre	lat. <i>liber</i>	Web
lipophage	graisse	gr. <i>λίπος</i>	Web
lithophage	Pierre	gr. <i>λίθος</i>	TLFi, GR
livrophage	livre		Web

logophage	discours	gr. λόγος	Web
macrophage	grand	gr. μακρός	TLFi, GR
mallophage	toison	gr. μαλλός	TLFi
médiophage	média		Web
mélophage	mouton	gr. μῆλον	TLFi, GR
merdophage	merde		Web
métrophage	méto		Web
microphage	petit (ce qui est)	gr. μικρός	TLFi, GR
monophage	seul	gr. μόνος	TLFi, GR
mucophage	mucus		Web
musicophage	musique		Web
musophage	banane	lat. <i>musa</i>	TLFi
mycétophage	champignon	gr. μύκης	TLFi, GR
mycophage	champignon	gr. μύκης	TLFi, GR
myrmécophage	fourmi	gr. μύρμηξ	TLFi, GR
mythophage	mythe		Web
nécrophage	mort	gr. νεκρός	TLFi, GR
nectariphage	nectar		Web
nematophage	nématodes		Web
neuronophage	neurone		TLFi
neurophage	neurone		TLFi
neutrophage	neutron		Web
nomophage	loi	gr. νόμος	TLFi
oligophage	petit nombre	gr. ὀλίγος	TLFi
omniphage	tout	lat. <i>omnis</i>	Web
omophage	viande crue	gr. ὠμός ('cru')	TLFi, GR
onirophage	rêve	gr. ὄνειρος	Web
onthophage	bouse	gr. ὄνθος	Web
onychophage	ongle	gr. ὄνυξ	GR
oophage	œuf	gr. ὠόν	Web
ophiophage	serpent	gr. ὄφης	TLFi, GR
opiophage	opium		TLFi, GR
ornithophage	oiseau	gr. ὄρνις	Web
palynophage	pollen	gr. παλύνω ('répandre')	Web
papyrophage	papyrus		Web
patatophage	patate		Web
pédophage	enfant	gr. παῖς	Web
pélophage	boue	gr. πηλός	GR
pétrolophage	pétrole		Web
photophage	photo		Web
phycophage	algue	gr. φῦκος	TLFi
phyllophage	feuille	gr. φύλλον	TLFi
phytophage	plante	gr. φυτόν	TLFi, GR
pithécophage	singe	gr. πίθηκος	TLFi
planctophage	plancton		GR
polyphage	plusieurs	gr. πολύς	TLFi, GR
pornophage	porno		Web
psychophage	esprit	gr. ψυχή	Web
publiphage	publicité		Web
pyrophage	feu	gr. πῦρ	GR
radiophage	radio		Web

rhizophage	racine	gr. <i>ρίζα</i>	TLFi
rhizophage	racine	gr. <i>ρίζα</i>	TLFi, GR
saprophage	matière en décomposition	gr. <i>σαπρός</i> ('pourri')	TLFi, GR
sarcophage	chair	gr. <i>σάρξ</i>	TLFi, GR
sarkophage	Sarkozy		Web
scatophage	excrément	gr. <i>σκώρ</i>	TLFi, GR
séri(e)phage	série		Web
sériophage	série		Web
sidérophage	fer	gr. <i>σιδηρος</i>	Web
sitophage	blé	gr. <i>σίτος</i>	TLFi
sténophage	étroit	gr. <i>στενός</i>	TLFi, GR
tachyphage	rapide	gr. <i>ταχύς</i>	Web
technophage	technologie (nouvelle)		Web
téléphage	télévision		Web
termitophage	termite		TLFi, GR
teutophage	calamar	gr. <i>εuthis</i>	Web
théatrophage	théâtre		Web
théophage	Dieu	gr. <i>θεός</i>	Web
thermophage	chaud (ce qui est)	gr. <i>θερμός</i>	Web
toxicophage	toxique		GR
trenetophage	Trenet		Web
trollophage	troll		Web
turcophage	Turquie/turc		TLFi
végétophage	végétal		Web
virophage	virus		GR
vitrophage	vitre		Web
webophage	Web		Web
webphage	Web		Web
xénophage	étranger	gr. <i>ξένος</i>	TLFi
xylophage	bois	gr. <i>ξύλον</i>	TLFi, GR
zéophage	épeautre	gr. <i>ζέα</i>	GR
zoophage	animal	gr. <i>ζοός</i>	TLFi, GR

Xphobe

ablutophobe	se laver	lat. <i>abluo</i>	Web
acarophobe	acare (gale)		Web
accentophobe	accent		Web
acerophobe	Acer		Web
achluophobe	obscurité		Web
achmophobe	aiguille		Web
acnophobe	acné		Web
acousticophobe	ascoutique		Web
acoustophobe	ascoutique		Web
acrophobe	haut	gr. <i>ἄκρος</i>	Web
aérodromphobe	aérodrome		Web
aérophobe	air	gr. <i>ἀήρ</i>	TLFi, GR
africanophobe	Afrique/africain		TLFi
agoraphobe	ville	gr. <i>ἀγείρειν</i> ('rassembler')	TLFi, GR
aichmophobe	pointe	gr. <i>αἰχμή</i>	Web

ailourophobe	chat	gr. <i>αἴλουρος</i>	Web
albophobe	blanc	lat. <i>album</i>	Web
algériophobe	Algérie/algérien		Web
algophobe	douleur	gr. <i>ἄλγος</i>	Web
allophobe	allosexuel		Web
alonsophobe	Alonso		Web
altérophobe	autre	lat. <i>alter</i>	Web
altomophobe	haut	lat. <i>altus</i>	Web
amatophobe	poussière	gr. <i>ἄμιμος</i> ('sable') (?)	Web
amaxophobe	véhicule	gr. <i>ἀμάξιον</i> ('wagon')	Web
amazighophobe	amazigh		Web
américanophobe	américain		GR
analphabètephobe	analphabète		Web
androphobe	homme	gr. <i>ἀνήπ</i>	GR
anémophobe	vent	gr. <i>ἄνεμος</i>	Web
anginophobe	angine		Web
anglicismophobe	angliscisme		Web
anglophobe	anglais		TLFi, GR
anglophonophobe	anglophone		Web
anglosaxophobe	anglosaxon		Web
anorexiquophobe	anorexique		Web
anorgasmophobe	sans + orgasme		Web
anthropophobe	homme	gr. <i>ἄνθρωπος</i>	TLFi
anuptaphobe	sans + mariage	lat. <i>nuptiae</i>	Web
apopathodiaphulatophobe	constipation	gr. <i>ἀπο</i> ('loin'), <i>πάθος</i> ('souffrance'), <i>διά</i> ('à travers'), <i>φυλάσσω</i> ('garder'),	Web
aquaphobe	eau	lat. <i>aqua</i>	Web
aquarophobe	ce qui concerne l'eau	lat. <i>aquarius</i>	Web
arabophobe	arabe		TLFi
arachnéophobe	araignée	gr. <i>ἀράχνη</i>	Web
arachnéphobe	araignée	gr. <i>ἀράχνη</i>	Web
arachnophobe	araignée	gr. <i>ἀράχνη</i>	GR
aragnophobe	araignée		Web
arénaphobe	araignée	lat. <i>aranea</i>	Web
arménophobe	Arménie/arménien		Web
artisanautophobe	artisanate		Web
artistophobe	artiste		Web
aspirateurophobe	aspirateur		Web
aspiratophobe	aspirateur		Web
associationophobe	association		Web
asthénophobe	faiblesse	gr. <i>ἀσθενής</i>	Web
astraphobe	éclair	gr. <i>ἀστραπή</i>	Web
astrophobe	astre		Web
atélophobe	sans + fin	gr. <i>τέλος</i>	Web
athéophobe	athée		Web
audiophobe	ouïe	lat. <i>audio</i>	Web
aulophobe	flûte	gr. <i>αὐλός</i>	Web
austrophobe	Autriche / autrichien		Web
autocheirothanatophobe	soi + main + mort	gr. <i>αὐτός, χεῖρ, θάνατος</i>	Web
automysophobe	soi + sale	gr. <i>αὐτός, μύσος</i>	Web
autophobe	soi	gr. <i>αὐτός</i>	Web

autophobe	auto(mobile)		Web
autrophobe	autres		Web
aviophobe	avion		Web
bacillophobe	bacille		Web
bactériophobe	bactérie		Web
bactérophobe	bactérie		Web
bagnolophobe	bagnole		Web
balophobe	BAL (boîte aux lettres)		Web
barbarophobe	barbare		Web
barbarophobe	barbarie		Web
barbophobe	barbe		Web
barophobe	bar (pression)		Web
bascophobe	basque		Web
basophobe	marche	gr. <i>βάσις</i>	Web
bathophobe	profondeur	gr. <i>βάθος</i>	Web
batrachophobe	grenouille	gr. <i>βάτραχος</i>	Web
belgeophobe	Belgique / belge		Web
belgophobe	Belgique / belge		Web
bellicophobe	guerre	lat. <i>bellicum</i>	Web
bélonéphobe	aiguille	gr. <i>βελόνη</i>	Web
bibliophobe	livre	gr. <i>βιβλίον</i>	TLFi
bidouilleurophobe	bidouilleur		Web
biennophobe	vase	?	Web
bigornophobe	bigornau		Web
biphobe	bi(sexuel)		Web
bitophobe	bite		Web
blairophobe	Blair		Web
blemmophobe	regard (des autres)	gr. <i>βλέμμα</i>	Web
blogophobe	blog		Web
blondophobe	blonde		Web
bobophobe	bobo		Web
bretonophobe	breton		Web
bricolophobe	bricolage		Web
bricophobe	bricolage		Web
brittophobe	breton	lat. <i>britto</i>	Web
brontophobe	tonnerre	gr. <i>βροντή</i>	Web
bronzophobe	bronzé		Web
bushophobe	Bush		Web
cacaphobe	caca		Web
cadeauphobe	cadeau		Web
cafardophobe	cafard		Web
caissonophobe	caisson		Web
calciphobe	calcaire		GR
calendriophobe	calendrier		Web
caligynephobe	belle femme	gr. <i>κάλλος, γυνή</i>	Web
câlinophobe	câlin		Web
cancérophobe	cancer		Web
cardiophobe	cœur	gr. <i>καρδιά</i>	Web
carnophobe	chair	lat. <i>caro</i>	Web
cartophobe	carte		Web
catalanophobe	catalan		Web

cataphobe	catalan		Web
catapédaphobe	grimper haut	gr. <i>κατά</i> ('vers le bas'), <i>πεδόν</i> ('sol')	Web
cathisophobe	s'asseoir	gr. <i>καθίζω</i>	Web
catholicophobe	catholique		Web
cathophobe	catho(lique)		Web
catoptrophobe	miroir	gr. <i>κάτοπτρον</i>	Web
cdphobe	cd		Web
chaetophobe	longs cheveux	gr. <i>χαίτη</i>	Web
chatophobe	chat		Web
chaussurophobe	chaussure		Web
cheimatophobe	hiver	gr. <i>χειμα</i>	Web
cheimophobe	mauvais temps	gr. <i>χειμή</i>	Web
chérophobe	être heureux	gr. <i>καίρω</i>	Web
chionophobe	neige	gr. <i>ξιών</i>	Web
christianophobe	chrétien		Web
christophobe	Christ		Web
chromatophobe	couleur	gr. <i>χρῶμα</i>	Web
chromophobe	couleur	gr. <i>χρῶμα</i>	TLFi
chronophobe	temps	gr. <i>χρόνος</i>	Web
claustrophobe	être enfermé	lat. <i>claustra</i> ('verrou')	TLFi, GR
clésiophobe	être enfermé	gr. <i>κλείς</i> ('clé')	Web
cléricophobe	clergé		Web
cliquophobe	clic		Web
clopophobe	clope		Web
clownophobe	clown		Web
communistophobe	communiste		Web
connardophobe	connard		Web
connophobe	con		Web
copain-de-sa-fillophobe	copain de sa fille		Web
coquillophobe	coquille		Web
corsophobe	corse		Web
coulrophobe	clown		Web
couplophobe	couple		Web
cousinophobe	cousin (famille)		Web
cousinophobe	cousin (insecte)		Web
crétinophobe	crétin		Web
cryophobe	froid	gr. <i>κρύος</i>	Web
culinarophobe	cuisine		Web
cumulophobe	cumulus		Web
cumulophobe	cumul (mandats)		Web
cuniculophobe	lapin	lat. <i>cuniculus</i>	Web
cyclophobe	cycliste		Web
cynophobe	chien	gr. <i>κύων</i>	Web
démonophobe	démon		Web
dendrophobe	arbre	gr. <i>δένδρον</i>	Web
dentiphobe	dentiste		Web
dermatophobe	peau	gr. <i>δέρμα</i>	Web
dextrophobe	droite	lat. <i>dextra</i>	Web
diabétophobe	diabète		Web
dieudonné	Dieudonné		Web
dipsophobe	soif	gr. <i>δίψα</i>	Web

diversitophobe	diversité		Web
djeunophobe	djeun's		Web
doraphobe	Dora (l'exploratrice)		Web
dragueusedemonmecophobe	drageuse de mon mec		Web
dysmorphobe	trouble + forme	gr. <i>δυσ, μορφή</i>	Web
dysmorphophobe	trouble + forme	gr. <i>δυσ, μορφή</i>	Web
dyspondéromorphophobe	trouble + poids + forme	gr. <i>δυσ</i> , lat. <i>pondus</i> , gr. <i>μορφή</i>	Web
écologhobe	écologie		Web
écophobe	écologie		Web
écophobe	économie		Web
ectophobe	dehors	gr. <i>ἐκτός</i>	Web
educophobe	éducation (nationale)		Web
efféminophobe	efféminé		Web
egophobe	soi	gr. <i>ἐγώ</i>	Web
electricitophobe	électricité		Web
électrophobe	électricité		Web
électrophobe	électrique (voiture)		Web
électrophobe	électro (musique)		Web
électrophobe	électricité		Web
éleuthérophobe	liberté	gr. <i>ἐλεύθερος</i>	Web
émétophobe	vomir	gr. <i>ἐμετιάω</i>	Web
encrophobe	encre		Web
entomophobe	insecte	gr. <i>ἔντομος</i>	Web
éosophobe	aurore	gr. <i>ἠώς</i>	Web
éphébofobe	éphèbe		Web
érémitophobe	ermite	lat. <i>eremita</i>	Web
éreuthophobe	rougir	gr. <i>ἐρεύτω</i>	GR
ergasiophobe	travail	gr. <i>ἐργασία</i>	Web
ergophobe	travail	gr. <i>ἔργον</i>	Web
érotophobe	érotisme		Web
érythrophobe	rouge	gr. <i>ἐρυθρός</i>	GR
espérantophobe	espéranto		Web
ethnophobe	peuple	gr. <i>ἔθνος</i>	Web
ethylophobe	éthylisme		Web
europophobe	Europe		Web
europophobe	Europe		Web
fabiusophobe	Fabius		Web
fascistophobe	fasciste		Web
fatwaphobe	fatwa		Web
féliphobe	félin		Web
fellophobe	fellation		Web
fibulanophobe	bouton	lat. <i>fibula</i>	Web
fillophobe	filles		Web
footballophobe	football		Web
footophobe	foot		Web
francophobe	France / français		TLFi, GR
fritophobe	frite		Web
fromageophobe	fromage		Web
gadgetophobe	gadget		Web
galactophobe	lait	gr. <i>γάλα</i>	Web
galéphobe	belette	gr. <i>γαλέη</i>	Web

gallophobe	gaulois	lat. <i>gallus</i>	TLFi, GR
gamétophobe	personne mariée	gr. <i>γαμετή</i>	Web
gay(-)phobe	gay		Web
génophobe	gène		Web
géphyrophobe	pont	gr. <i>γέφυρα</i>	Web
gériaphobe	vieux	gr. <i>γεραιός</i>	Web
germanophobe	Allemagne / allemand		TLFi, GR
gérontophobe	vieux	gr. <i>γέρον</i>	Web
gouinophobe	gouine		Web
graphophobe	écriture	gr. <i>γράφω</i> ('écrire')	Web
grévophobe	grève		Web
grophobe	gros		Web
grosophobe	gros		Web
grossophobe	gros		Web
groszizophobe	gros zizi		Web
gymnophobe	nu	gr. <i>γύμνος</i>	Web
gynécophobe	femme	gr. <i>γυνή</i>	Web
gynécophobe	gynécologue		Web
gynéphobe	femme	gr. <i>γυνή</i>	Web
gynophobe	femme	gr. <i>γυνή</i>	Web
halitophobe	halitose	(mauvaise haleine)	Web
handiphobe	handicapé		Web
haptophobe	attacher	gr. <i>ἅπτω</i>	Web
harpaxophobe	enleveur	gr. <i>ἄρπαξ</i>	Web
hédonophobe	hédonisme		Web
héliophobe	soleil	gr. <i>ἥλιος</i>	TLFi, GR
hellenophobe	Grèce / grec	gr. <i>Ἑλληνικός</i>	Web
hématophobe	sang	gr. <i>αἷμα</i>	Web
hémophobe	sang	gr. <i>αἷμα</i>	Web
herpétophobe	serpent	gr. <i>ἕρπεντόν</i>	Web
hétérophobe	hétéro(sexuel)		Web
hexakosioihexekontahexaphobe	666	gr. <i>ἑξακόσιοι</i> (600), <i>ἑξήκοντα</i> (60), <i>ξξ</i> (6)	Web
hippiophobe	hippie		Web
hippophobe	cheval	gr. <i>ἵππος</i>	Web
hippopotomonstrosesquipedaliophobe	mot très long		Web
hispanophobe	espagnol		Web
hivrophobe	hiver		Web
hodophobe	route	gr. <i>ὁδός</i>	Web
homichlophobe	brouillard	gr. <i>ὁμίκλη</i>	Web
homophobe	homosexuel		GR
homophobophobe	homophobe		Web
homosexophobe	homosexuel		Web
homosexuellophobe	homosexuel		Web
hoplophobe	arme	gr. <i>ὄπλον</i>	Web
humanophobe	humain		Web
hongarophobe	hongrois		Web
hydroluxophobe	eau + lumière	gr. <i>ὑδρῶν</i> , lat. <i>lux</i>	Web
hydrophobe	eau	gr. <i>ὑδρῶν</i>	TLFi, GR
hygrophobe	humidité	gr. <i>ὕγρῶς</i>	GR
hylophobe	bois	gr. <i>ῥυλα</i>	Web

hypégiaphobe	reponsabilité	gr. ὑπέγγυος	Web
hypnophobe	sommeil	gr. ὕπνος	Web
iatrophobe	médecin	gr. ἱατρός	Web
iconophobe	icône		TLFi, GR
impairophobe	impair (nombre)		Web
incestophobe	inceste		Web
indophobe	Inde/indien		Web
infidéliphobe	infidèle		Web
insectophobe	insecte		Web
iranophobe	Iran/iranien		Web
islamistophobe	islamiste		Web
islamophobe	Islam		GR
israélophobe	Israël/israélien		Web
italophobe	Italie/italien		Web
ithyphallophobe	sexe en érection	gr. ἰθύς	Web
jaguarophobe	jaguar		Web
japonophobe	Japon/japonais		Web
joeystarrophobe	Joey Starr		Web
johnnyphobe	Johnny (Halliday)		Web
judéophobe	juif		TLFi
juvénophobe	jeune		Web
kabylophobe	Kabyle		Web
kaïnophobe	nouveau	gr. καινός	Web
kantophobe	Kant		TLFi
katagélophobe	dérision	gr. κατάγελως	Web
kathisophobe	s'asseoir	gr. καθίζω	Web
kénophobe	vide	gr. κενός	Web
kévinophobe	kévin	(internaute nuisible et immature)	Web
kleptophobe	voler	gr. κλέπτω	Web
koniphobe	poussière	gr. κονία	Web
kurdophobe	kurde		Web
laboriphobe	travail	lat. labor	Web
lachanophobe	légume	gr. λάχανον	Web
lactophobe	lait	lat. lac	Web
lalphobe	parler	gr. λαλέω ('parler')	Web
lapinophobe	lapin		Web
latinophobe	latin		Web
latinophobe	latino		Web
lépidophobe	lépidoptère		Web
lesbophobe	lesbienne		Web
leuc(k)ophobe	blanc	gr. λευκός	Web
lévophobe	gauche	lat. laeva	Web
libanophobe	Liban/libanais		Web
libertiphobe	liberté		Web
libertophobe	liberté		Web
linonophobe	fil de lin	gr. λινόν	Web
lipophobe	graisse	gr. λίπος	Web
listophobe	liste		Web
lituanophobe	Lituanie/lituanien		Web
logophobe	discours	gr. λόγος	Web
logophobe	logo		Web

lupinophobe	loup	lat. <i>lupus</i>	Web
lycanophobe	loup	gr. λύκος	Web
lycanthrophobe	loup + homme	gr. λύκος, άνθρωπος	Web
lyophobe	dissolution	gr. λύη	GR
machinphobe	machin		Web
mammouthphobe	mammouth		Web
maophobe	Mao		Web
marocophobe	Maroc/marocain		Web
mastigophobe	fouetter	gr. αστιγίω	Web
mathildophobe	Mathilde		Web
mathophobe	math		Web
mayonnaisophobe	mayonnaise		Web
médiaphobe	média		Web
mélrophobe	chant	gr. μέλος	TLFi
ménageophobe	ménage		Web
météorophobe	météore		Web
microbophobe	microbe		Web
microsoftphobe	microsoft		Web
migrophobe	migrant		Web
mnémophobe	mémoire	gr. μνήμη	Web
mochophobe	moche		Web
modophobe	mode		Web
modophobe	modo	(modérateur des forums)	Web
mondialophobe	mondialisation		Web
monophobe	seul	gr. μόνος	Web
motophobe	moto		Web
multiphobe	plusieurs	lat. <i>multus</i>	Web
musicophobe	musique		TLFi
musulmanophobe	musulman		Web
mycophobe	champignon	gr. μύκης	Web
myrmécophobe	fourmi	gr. μύρμηξ	Web
mysophobe	saleté	gr. μύσος	Web
mythophobe	mythe		Web
naturophobe	nature		Web
naziophobe	nazi		Web
naziophobe	nazi		Web
nécrophobe	mort	gr. νεκρός	Web
néerlandophobe	néerlandais		Web
négrophobe	négro		Web
néophobe	nouveau	gr. νέος	TLFi, GR
néphophobe	nuage	gr. νέφος	Web
neurophobe	neurologie		Web
nipponphobe	nippon		Web
nippophobe	nippon		Web
noctiphobe	nuit	lat. <i>nox</i>	Web
nomophobe	loi	gr. νόμος	Web
nomophobe	no mobile		Web
nosophobe	maladie	gr. νόσος	GR
nutriphobe	nutriment		Web
nyctalophobe	nyctal	gr. νύξ	Web
nyctaphobe	nyctal	gr. νύξ	Web

nyctophobe	nyctal	gr. νύξ	Web
occidentalophobe	occidental		Web
occidentophobe	occident		Web
ochlophobe	foule	gr. ὄχλος	Web
odontophobe	dent	gr. ὀδούς	Web
œnophobe	vin	gr. οἶνος	Web
oikophobe	maison	gr. οἶκος	Web
oiseauphobe	oiseau		Web
oléophobe	gras	lat. olea	Web
olfactophobe	odeur	lat. olfactus	Web
onéirophobe	rêve	gr. ὄνειρος	Web
onomatophobe	nom	gr. ὄνομα	Web
ophiophobe	serpent	gr. ὄφις	Web
ordinophobe	ordinateur		Web
ornithophobe	oiseau	gr. ὄρνις	Web
orthographobe	orthographe		Web
paladinophobe	paladin		Web
palestinophobe	Palestine(nien)		Web
panophobe	tout	gr. πᾶν	Web
pantophobe	tout	gr. πᾶν	Web
panurgeophobe	panurge (mouton de)		Web
papierphobe	papier		Web
papillophobe	papillon		Web
papophobe	pape		Web
papyphobe	papy		Web
papyrophobe	papyrus		Web
paraskevidékatriaphobe	vendredi 13	gr. παρασκευή, δεκατρεῖς	Web
patatophobe	patate		Web
pauvrophobe	pauvre		Web
pécéphobe	PC		Web
pédiculophobe	pou	lat. pediculus	Web
pédophobe	enfant	gr. παις	Web
peignirophobe	peignoir		Web
péladophobe	pelade		Web
péniaphobe	pauvreté	gr. πενία	Web
perrinophobe	Perrin		Web
petassophobe	pétasse		Web
phagophobe	manger	gr. φαγεῖν	Web
phanérophobe	visible	gr. φανερός	Web
phantasmophobe	fantôme		Web
phasmophobe	fantôme	gr. φάσμα	Web
philatélophobe	philatélie		Web
phobophobe	phobie		Web
phonophobe	son	gr. φωνή	Web
photophobe	lumière	gr. φῶς	TLFi
phronémophobe	réfléchir	gr. φρονέω	Web
pilophobe	poil	lat. pilus	Web
pipophobe	pipe		Web
piqueusedemecophobe	piqueuse de mec		Web
pisciphobe	poisson	lat. piscis	Web
plangonophobe	poupée		Web

plombophobe	plombier		Web
ploutophobe	richesse	gr. πλοῦτος	Web
plumophobe	plume		Web
pogonophobe	barbe	gr. πώγων	Web
poilophobe	poil		Web
poissonrougeophobe	poisson rouge		Web
polyphobe	plusieurs	gr. πολύς	Web
ponctuationphobe	ponctuation		Web
ponophobe	travail	gr. πόνος	Web
pornophobe	porno		Web
portugophobe	Portugal/portugais		Web
potophobe	boire	gr. ποτής	Web
prêtraphobe	prêtre		GR
privatophobe	privatisation		Web
prostiphobe	prostituée		Web
psophobe	PS		Web
publiphobe	publicité		TLFi, GR
putophobe	pute		Web
pyrophobe	feu	gr. πῦρ	Web
québécoisphobe	Québec		Web
qwertyphobe	qwerty		Web
racistophobe	raciste		Web
radarophobe	radar		Web
rapophobe	rap		Web
réformophobe	réforme		Web
religiophobe	religion		Web
reptilophobe	reptile		Web
requinophobe	requin		Web
richophobe	riche		Web
riscophobe/risquophobe	risque		Web
ritalphobe	rital		Web
robophobe	robot		Web
rollerphobe	roller		Web
romanophobe	romano		Web
romanophobe	romain		Web
roséphobe	rosé		Web
roxanophobe	roxane		Web
rupophobe	être sale	gr. ρυπάω	Web
russophobe	Russie / russe		TLFi, GR
sabinophobe	Sabine		Web
saccharophobe	sucre	gr. σάκχαρον	Web
sarkophobe	Sarko(zy)		Web
scatophobe	excrément	gr. σκώρ	Web
schtroumpfophobe	Shtroumpf		Web
schtroumpfophobe	Shtroumpf	(n'importe quoi)	Web
scienc(ç)ophobe	science		Web
sciophobe	ombre	gr. σκία	Web
sciophobe	scie		Web
scopophobe	regarder	gr. σκοπέω	Web
scotophobe	obscurité	gr. σκότιος	Web
scoutophobe	scout		Web

sectophobe	secte		Web
segaphobe	sega		Web
serbophobe	Serbie/serbe		Web
serpentphobe	serpent		Web
seulophobe	seul		Web
sexophobe	sexe		GR
shortmoulantophobe	short moulant		Web
sidérodromophobe	fer + course	gr. <i>σίδηρος, δρόμος</i>	Web
sinistrophobe	gauche	lat. <i>sinister</i>	Web
sinophobe	Chine / chinois		GR
sitiophobe	nourriture	gr. <i>σιτίον</i>	Web
sociophobe	société		Web
sodophobe	sodomie		Web
sonyphobe	Sony		Web
spermatophobe	sperme		Web
spielbergophobe	Spielberg		Web
squalophobe	squale		Web
stalinophobe	Staline		Web
suiophobe	soi	lat. <i>sui</i>	Web
suitophobe	suite		Web
superhydrophobe	super + eau	gr. <i>ὑδρῶν</i>	Web
syphilophobe	syphilis		TLFi, GR
tabacophobe	tabac		Web
tachophobe	rapide	gr. <i>ταχύς</i>	Web
tachophobe	tache		Web
taphophobe	enterrement	gr. <i>ταφή</i>	Web
taxiphobe	taxi		Web
taxophobe	taxe		Web
technocratophobe	technocrate		Web
technophobe	technologie (nouvelles)		TLFi, GR
téléphobe	télé(vision)		Web
téléphonophobe	téléphone		Web
tératophobe	montre	gr. <i>τέρας</i>	Web
tétraphobe	quatre	gr. <i>τετρα</i>	Web
teutonophobe	teuton		Web
thaasophobe	être assis	gr. <i>θαάσσω</i>	Web
thalassophobe	mer	gr. <i>θάλασσα</i>	Web
thanathophobe	mort	gr. <i>θάνατος</i>	Web
théophobe	Dieu	gr. <i>θεός</i>	TLFi
thermophobe	chaleur	gr. <i>θερμός</i>	Web
thixophobe	toucher	gr. <i>θίξις</i>	Web
tintinophobe	Tintin		Web
tocophobe	accoucher	gr. <i>τοκάω</i>	Web
tofuphobe	tofu		Web
topophobe	lieu	gr. <i>τόπος</i>	Web
tortuphobe	tortue		Web
tourophobe	tour		Web
toutphobe	tout		Web
toxicophobe	toxique		Web
transistorophobe	transistor		Web
transphobe	trans(sexuel)		Web

traumatophobe	blessure	gr. τραῦμα	Web
trichophobe	poil	gr. θρίξ	Web
triskaidékaphobe	treize	gr. τρισκαίδεκα	Web
trypanophobe	percer	gr. τρυπάω	Web
tsiganophobe	tsigane		Web
tuberculophobe	tuberculose		Web
tuberculophobe	tubercule		Web
turcophobe	Turquie / Turc		TLFi, GR
tuyauphobe	tuyau		Web
urbanophobe	ville	lat. <i>urbs</i>	Web
urbaphobe	ville		Web
vaissellophobe	vaisselle		Web
végétophobe	végétarien		Web
vélibophobe	vélib		Web
vélophobe	vélo		GR
vermiphobe	ver	lat. <i>vermis</i>	Web
vidophobe	vide		Web
vinophobe	vin		Web
virusophobe	virus		Web
webophobe	Web		Web
windophobe	Windows		Web
windowsophobe	Windows		Web
xénophobe	étranger	gr. ξένος	TLFi, GR
zélophobe	jalousie	gr. ζήλος	Web
zoophilophobe	zoophile		Web
zoophobe	animal	gr. ζώος	Web

*X*phobie

ablutophobie	se laver	lat. <i>abluo</i>	Web
acarophobie	acare (gale)		TLFi
accentophobie	accent		Web
acerophobie	Acer		Web
acerophobie	acerbe		Web
achluophobie	obscurité		Web
achmophobie	aiguille		Web
acnophobie	acné		Web
acouphobie	entendre	gr. ἀκούω	Web
acousticophobie	ascoutique		Web
acoustophobie	ascoutique		Web
acrophobie	haut	gr. ἄκρος	TLFi, GR
aérodromphobie	aérodrome		Web
aérophobie	air	gr. ἀήρ	TLFi, GR
africanophobie	Afrique/africain		Web
agoraphobie	ville	gr. ἀγείρειν ('rassembler')	TLFi, GR
aichmophobie	pointe	gr. αἰχμή	Web
ailourophobie	chat	gr. αἴλουρος	Web
albophobie	blanc	lat. <i>album</i>	Web
algériophobie	Algérie/algérien		Web
algophobie	douleur	gr. ἄλγος	GR

allophobie	allosexuel		Web
allophobie	all	gr. ἄλλος	Web
alonsophobie	Alonso		Web
altérophobie	autre	lat. <i>alter</i>	Web
altomorphobie	haut	lat. <i>altus</i>	Web
amatophobie	poussière	gr. ἄμιμος ('sable') (?)	Web
amaxophobie	véhicule	gr. ἀμάξιον ('wagon')	Web
amazighophobie	amazighe		Web
américanophobie	américain		GR
analphabétophobie	analphabète		Web
androphobie	homme	gr. ἀνήπ	TLFi, GR
anémosphobie	vent	gr. ἄνεμος	GR
anginophobie	angine		Web
angliscismophobie	angliscisme		Web
anglophobie	anglais		TLFi, GR
anglophonophobie	anglophone		Web
anglosaxophobie	anglosaxon		Web
anorgasmophobie	sans + orgasme		Web
anthelmophobie	anthelm... (anti + helminthe)		Web
anthropophobie	homme	gr. ἄνθρωπος	TLFi
anuptaphobie	sans + mariage	lat. <i>nuptiae</i>	Web
apopathodiaphulatophobie	constipation	gr. ἀπο('loin'), πάθος('souffrance'), διά('à travers'), φυλάσσω('garder'),	Web
aquaphobie	eau	lat. <i>aqua</i>	Web
aquarophobie	ce qui concerne l'eau	lat. <i>aquarius</i>	Web
arabophobie	arabe		Web
arachnéophobie	araignée	gr. ἀράχνη	Web
arachnéphobie	araignée	gr. ἀράχνη	Web
arachnophobie	araignée	gr. ἀράχνη	GR
aragnophobie	araignée		Web
arénaphobie	araignée	lat. <i>aranea</i>	Web
arménophobie	Arménie/arménien		Web
artistophobie	artiste		Web
aspirateurophobie	aspirateur		Web
aspiratophobie	aspirateur		Web
asthénophobie	faiblesse	gr. ἀσθενής	Web
astraphobie	éclair	gr. ἀστραπή	Web
astrapophobie	éclair	gr. ἀστραπή	Web
astrophobie	astre		TLFi
atélophobie	sans + fin	gr. τέλος	Web
athéophobie	athée		Web
audiophobie	ouïe	lat. <i>audio</i>	Web
aulophobie	flûte	gr. αὐλός	Web
austrophobie	Autriche / autrichien		Web
autocheirothanatophobie	soi + main + mort	gr. αὐτός, χεῖρ, θάνατος	Web
automysophobie	soi + sale	gr. αὐτός, μύσος	Web
autophobie	soi	gr. αὐτός	Web
autophobie	auto(mobile)		Web
autrophobie	autres		Web
aviophobie	avion		Web
bacillophobie	bacille		TLFi

bactériophobie	bactérie		Web
bactérophobie	bactérie		Web
bagnolophobie	bagnole		Web
balophobie	BAL (boîte aux lettres)		Web
barbarophobie	barbare		Web
barbophobie	barbe		Web
barophobie	bar (lieu)		Web
barophobie	bar (pression)		Web
bascofobie	basque		Web
basophobie	marche	gr. <i>βάσις</i>	TLFi, GR
bathophobie	profondeur	gr. <i>βάθος</i>	GR
batrachophobie	grenouille	gr. <i>βάτραχος</i>	Web
belgeophobie	Belgique / belge		Web
belgophobie	Belgique / belge		Web
bélonéphobie	aiguille	gr. <i>βελόνη</i>	GR
biennophobie	vase	?	Web
biphobie	bi(sexuel)		Web
bitophobie	bite		Web
blairophobie	Blair		Web
blemmophobie	regard (des autres)	gr. <i>βλέμμα</i>	Web
blogophobie	blog		Web
blondophobie	blonde		Web
bobophobie	bobo		Web
brettonophobie	breton		Web
bricolophobie	bricolage		Web
brittophobie	breton	lat. <i>britto</i>	Web
brontophobie	tonnerre	gr. <i>βροντή</i>	Web
bronzophobie	bronzé		Web
bushophobie	Bush		Web
cacaphobie	caca		Web
cadeauphobie	cadeau		Web
cafardophobie	cafard		Web
calciphobie	calcaire		Web
caligynephobie	belle femme	gr. <i>κάλλος, γυνή</i>	Web
câlinophobie	câlin		Web
cancérophobie	cancer		GR
cardiophobie	cœur	gr. <i>καρδία</i>	TLFi
carnophobie	chair	lat. <i>caro</i>	Web
cartophobie	carte		TLFi
catalanophobie	catalan		Web
catapédaphobie	grimper haut	gr. <i>κατά</i> ('vers le bas'), <i>πεδόν</i> ('sol')	Web
cathisophobie	s'asseoir	gr. <i>καθίζω</i>	Web
catholicophobie	catholique		Web
cathophobie	catho(lique)		Web
catoptrophobie	miroir	gr. <i>κάτοπτρον</i>	Web
chaetophobie	longs cheveux	gr. <i>χάιτη</i>	Web
chatophobie	chat		Web
chaussurophobie	chaussure		Web
cheimatophobie	hiver	gr. <i>χειμα</i>	Web
heimophobie	mauvais temps	gr. <i>χειμιά</i>	Web
chérophobie	être heureux	gr. <i>καίρω</i>	Web

chionophobie	neige	gr. <i>ξιών</i>	Web
christianophobie	chrétien		Web
christophobie	Christ		Web
chromatophobie	couleur	gr. <i>χρῶμα</i>	Web
chromophobie	couleur	gr. <i>χρῶμα</i>	Web
chronophobie	temps	gr. <i>χρόνος</i>	Web
claustrophobie	être enfermé	lat. <i>claustra</i> ('verrou')	TLFi, GR
clésiophobie	être enfermé	gr. <i>κλείς</i> ('clé')	Web
cléricophobie	clergé		Web
cliquophobie	clic		Web
clopophobie	clope		Web
clownophobie	clown		Web
communistophobie	communiste		Web
connardophobie	connard		Web
connophobie	con		Web
contagiophobie	contagion		Web
corsophobie	corse		Web
coulrophobie	clown		Web
couplophobie	couple		Web
cousinophobie	cousin		Web
crétinophobie	crétin		Web
cryophobie	froid	gr. <i>κρύος</i>	Web
cumulophobie	cumulus		Web
cuniculophobie	lapin	lat. <i>cuniculus</i>	Web
cyclophobie	cycliste		Web
cynophobie	chien	gr. <i>κύων</i>	GR
démonophobie	démon		Web
dendrophobie	arbre	gr. <i>δένδρον</i>	Web
dentiphobie	dentiste		Web
dermatophobie	peau	gr. <i>δέρμα</i>	Web
dextrophobie	droite	lat. <i>dextra</i>	Web
diabétophobie	diabète		Web
dieudophobie	Dieudonné		Web
dipsophobie	soif	gr. <i>δίψα</i>	GR
diversitophobie	diversité		Web
doraphobie	fouurrure	gr. <i>δορά</i> ('peau écorchée')	Web
dysmorphobie	trouble + forme	gr. <i>δυσ, μορφή</i>	Web
dysmorphophobie	trouble + forme	gr. <i>δυσ, μορφή</i>	TLFi
dyspondéromorphophobie	trouble + poids + forme	gr. <i>δυσ</i> , lat. <i>pondus</i> , gr. <i>μορφή</i>	Web
écologophobie	écologie		Web
écophobie	écologie		Web
écophobie	habitat	gr. <i>οἶκος</i>	GR
ectophobie	dehors	gr. <i>ἐκτός</i>	Web
efféminophobie	efféminé		Web
egophobie	soi	gr. <i>ἐγώ</i>	Web
eibohphobie	eiboh	(palindrome de phobie)	Web
électrophobie	électricité		Web
électrophobie	électrique (voiture)		Web
électrophobie	électro (musique)		Web
éleuthérophobie	liberté	gr. <i>ἐλεύθερος</i>	Web
émétophobie	vomir	gr. <i>ἐμετιάω</i>	Web

encrophobie	encre		TLFi
énétophobie	épingle	gr. ἐνετή	Web
entomophobie	insecte	gr. ἔντομος	Web
éosophobie	aurore	gr. ἠώς	Web
éphétophobie	éphèbe		Web
éremíttophobie	ermite	lat. <i>eremita</i>	Web
éreuthophobie	rougir	gr. ἐρεύτω	TLFi, GR
ergasiophobie	travail	gr. ἐργασία	TLFi
ergophobie	travail	gr. ἔργον	TLFi
érottophobie	érotisme		Web
érythrotophobie	rouge	gr. ἐρυθρός	GR
espéranttophobie	espéranto		Web
ethnophobie	peuple	gr. ἔθνος	Web
ethylophobie	éthylisme		Web
europophobie	Europe		Web
europophobie	Europe		Web
fabiusophobie	Fabius		Web
fascisttophobie	fasciste		Web
fellophobie	fellation		Web
feràrepasserttophobie	fer à repasser		Web
fibulanophobie	bouton	lat. <i>fibula</i>	Web
fillophobie	filles		Web
footballophobie	football		Web
foottophobie	foot		Web
francophobie	France / français		TLFi, GR
frittophobie	frite		Web
fromageophobie	fromage		Web
gadgettophobie	gadget		Web
galactophobie	lait	gr. γάλα	Web
galétophobie	belette	gr. γαλέη	Web
gallophobie	gaulois	lat. <i>gallus</i>	TLFi, GR
gaméttophobie	personne mariée	gr. γαμετή	Web
gay(-)tophobie	gay		Web
génophobie	gène		Web
géphyrotophobie	pont	gr. γέφυρα	Web
gériatophobie	vieux	gr. γεραιός	Web
germanophobie	Allemagne / allemand		TLFi, GR
géronttophobie	vieux	gr. γέρων	TLFi
gonadtophobie	gonade		Web
gouinophobie	gouine		Web
graphtophobie	écriture	gr. γράφω ('écrire')	TLFi
grévtophobie	grève		Web
grophobie	gros		Web
grosophobie	gros		Web
grossophobie	gros		Web
gymnophobie	nu	gr. γύμνος	Web
gynécophobie	femme	gr. γυνή	TLFi, GR
gynétophobie	femme	gr. γυνή	TLFi
gynophobie	femme	gr. γυνή	TLFi, GR
halitophobie	halitose	(mauvaise haleine)	Web
handiphobie	handicapé		Web

haptophobie	attacher	gr. ἅπτω	Web
harpaxophobie	enleveur	gr. ἄραξ	Web
hédonophobie	hédonisme		Web
héliophobie	soleil	gr. ἥλιος	TLFi
hellenophobie	Grèce / grec	gr. Ἑλληνικός	Web
hématophobie	sang	gr. αἷμα	GR
hémophobie	sang	gr. αἷμα	GR
herpétophobie	serpent	gr. ἕρετόν	Web
hétérophobie	hétéro(sexuel)		Web
hexakosioihexekontahephobie	666	gr. ἑξακόσιοι (600), ἑξήκοντα (60), ξξ (6)	Web
hippophobie	cheval	gr. ἵππος	Web
hippopotomonstrosesquipedaliophobie	mot très long		Web
hispanophobie	espagnol		Web
hiverophobie	hiver		Web
hodophobie	route	gr. ὁδός	Web
homichlophobie	brouillard	gr. ὁμίκλη	Web
homophobie	homosexuel		GR
homophobophobie	homophobe		Web
homosexophobie	homosexuel		Web
homosexuellophobie	homosexuel		Web
hoplophobie	arme	gr. ὄπλον	Web
horméphobie	assaut	gr. ὄρμη	Web
humanophobie	humain		Web
hongrois	hongrois		Web
hydroluxophobie	eau + lumière	gr. ὕδρων, lat. lux	Web
hydrophobie	eau	gr. ὕδρων	TLFi, GR
hygrophobie	humidité	gr. ὑγρός	Web
hylophobie	bois	gr. ὕλα	Web
hypégiaphobie	responsabilité	gr. ὑπέγγυος	Web
hypnophobie	sommeil	gr. ὕπνος	TLFi
iatrophobie	médecin	gr. ἰατρός	Web
iconophobie	icône		TLFi
impairphobie	impair (nombre)		Web
incestophobie	inceste		Web
indophobie	Inde/indien		Web
infidéliphobie	infidèle		Web
insectophobie	insecte		Web
iranophobie	Iran/iranien		Web
islamistophobie	islamiste		Web
islamophobie	Islam		GR
israélophobie	Israël/israélien		Web
italophobie	Italie/italien		Web
ithyphallophobie	sexe en érection	gr. ἰθύς	Web
jaguarophobie	jaguar		Web
japonophobie	Japon/japonais		Web
johnnyphobie	Johnny (Halliday)		Web
judéophobie	juif		Web
juvénophobie	jeune		Web
kabylophobie	Kabyle		Web
kaïnophobie	nouveau	gr. καινός	GR

kantophobie	Kant		Web
katagélaphobie	dérision	gr. κατάγελωσ	Web
kathisophobie	s'asseoir	gr. καθίζω	Web
kénophobie	vide	gr. κενός	Web
kévinophobie	kévin	(internaute nuisible et immature)	Web
kleptophobie	voler	gr. κλέπτω	Web
koniphobie	poussière	gr. κονία	Web
kopophobie	fatigue	gr. κοπώω	Web
kurdophobie	kurde		Web
lachanophobie	légume	gr. λάχανον	Web
lactophobie	lait	lat. lac	Web
lalophobie	parler	gr. λαλέω ('parler')	Web
langueetrangerephobie	langue étrangère		Web
lapinophobie	lapin		Web
latinophobie	latin		Web
latinophobie	latino		Web
lépidophobie	lépidoptère		Web
lesbophobie	lesbienne		Web
leuc(k)ophobie	blanc	gr. λευκός	Web
lévophobie	gauche	lat. laeva	Web
libanophobie	Liban/libanais		Web
libertiphobie	liberté		Web
libertophobie	liberté		Web
linonophobie	fil de lin	gr. λινόν	Web
lipophobie	graisse	gr. λίπος	Web
listophobie	liste		Web
logophobie	discours	gr. λόγος	Web
logophobie	logo		Web
lupinophobie	loup	lat. lupus	Web
lycanophobie	loup	gr. λύκος	Web
lycanthrophobie	loup + homme	gr. λύκος, άνθρωπος	Web
lygophobie	crépuscule	gr. λύγη	Web
lyophobie	dissolution	gr. λύη	GR
lyssophobie	rage	gr. λύσσα	Web
machinphobie	machin		Web
maophobie	Mao		Web
marocophobie	Maroc/marocain		Web
mastigophobie	fouetter	gr. αστιγώω	Web
mathophobie	math		Web
méchanophobie	machine	gr. μηχανήμα	Web
médiaphobie	média		Web
mélaphobie	chant	gr. μέλος	Web
ménageophobie	ménage		Web
météorophobie	météore		Web
microbophobie	microbe		Web
microsoftophobie	microsoft		Web
migrophobie	migrant		Web
mnémophobie	mémoire	gr. μνήμα	Web
mochophobie	moche		Web
modophobie	modo	(modérateur des forums)	Web
mondialophobie	mondialisation		Web

monophobie	seul	gr. <i>μόνος</i>	Web
motophobie	moto		Web
multiphobie	plusieurs	lat. <i>multus</i>	Web
musicophobie	musique		Web
musulmanophobie	musulman		Web
mycophobie	champignon	gr. <i>μύκης</i>	Web
myrmécophobie	fourmi	gr. <i>μύρμηξ</i>	Web
mysophobie	saleté	gr. <i>μύσος</i>	Web
mythophobie	mythe		Web
naturophobie	nature		Web
naziophobie	nazi		Web
naziophobie	nazi		Web
nécrophobie	mort	gr. <i>νεκρός</i>	GR
néerlandophobie	néerlandais		Web
négrophobie	négro		Web
néophobie	nouveau	gr. <i>νέος</i>	TLFi, GR
néphrophobie	nuage	gr. <i>νέφος</i>	Web
neurophobie	neurologie		Web
nipponphobie	nippon		Web
nippophobie	nippon		Web
noctiphobie	nuit	lat. <i>nox</i>	Web
nomophobie	loi	gr. <i>νόμος</i>	Web
nomophobie	no mobile		Web
nosophobie	maladie	gr. <i>νόσος</i>	GR
nyctalophobie	nyctal	gr. <i>νύξ</i>	Web
nyctaphobie	nyctal	gr. <i>νύξ</i>	Web
nyctophobie	nyctal	gr. <i>νύξ</i>	Web
occidentalophobie	occidental		Web
occidentophobie	occident		Web
ochlophobie	foule	gr. <i>ὄχλος</i>	Web
odontophobie	dent	gr. <i>ὀδούς</i>	Web
œnophobie	vin	gr. <i>οἴνος</i>	Web
oïkophobie	maison	gr. <i>οἶκος</i>	Web
oiseauphobie	oiseau		Web
oléophobie	gras	lat. <i>olea</i>	Web
olfactophobie	odeur	lat. <i>olfactus</i>	Web
ommatophobie	œil	gr. <i>ὄμμα</i>	Web
onéirophobie	rêve	gr. <i>ὄνειρος</i>	Web
onomatophobie	nom	gr. <i>ὄνομα</i>	Web
ophiophobie	serpent	gr. <i>ὄφις</i>	Web
ordinophobie	ordinateur		Web
ornithophobie	oiseau	gr. <i>ὄρνις</i>	Web
orthographobie	orthographe		Web
paladinophobie	paladin		Web
palestinophobie	Palestine(nien)		Web
panophobie	tout	gr. <i>πᾶν</i>	TLFi
pantophobie	tout	gr. <i>πᾶν</i>	TLFi, GR
panurgeophobie	panurge (mouton de)		Web
papierphobie	papier		Web
papillophobie	papillon		Web
papophobie	pape		Web

papyphobie	papy		Web
papyrophobie	papyrus		Web
paraskevidékatriaphobie	vendredi 13	gr. παρασκευή, δεκατρεῖς	Web
parthénophobie	jeune fille	gr. παρθένος	GR
patatophobie	patate		Web
pathophobie	maladie	gr. πάθος	GR
pauvrophobie	pauvre		Web
peccatophobie	faute	lat. peccatus	Web
pécéphobie	PC		Web
pédiculophobie	pou	lat. pediculus	Web
pédophobie	enfant	gr. παῖς	Web
péladophobie	pelade		Web
péniaphobie	pauvreté	gr. πενία	Web
perrinophobie	Perrin		Web
phagophobie	manger	gr. φαγεῖν	Web
phanérophobie	visible	gr. φανερός	Web
phantasmophobie	fantôme		Web
phasmophobie	fantôme	gr. φάσμα	Web
philatélophobie	philatélie		Web
phobophobie	phobie		GR
phonophobie	son	gr. φωνή	Web
photophobie	lumière	gr. φῶς	TLFi, GR
phronémophobie	réfléchir	gr. φρονέω	Web
phtisiophobie	phtisie		Web
pilophobie	poil	lat. pilus	Web
pipophobie	pipe		Web
pisciphobie	poisson	lat. piscis	Web
plangonophobie	poupée		Web
plombophobie	plombier		Web
plombophobie	plomb		Web
ploutophobie	richesse	gr. πλοῦτος	Web
plumophobie	plume		TLFi
pogonophobie	barbe	gr. πώγων	Web
poilophobie	poil		Web
poïnéphobie	châtiment	gr. ποινή	Web
poissonrougeophobie	poisson rouge		Web
polyphobie	plusieurs	gr. πολύς	Web
ponophobie	travail	gr. πόνος	Web
pornophobie	porno		Web
portugophobie	Portugal/portugais		Web
potophobie	boire	gr. ποτής	Web
prêtrophobie	prêtre		GR
privatophobie	privatisation		Web
prostiphobie	prostituée		Web
psophobie	PS		Web
publiphobie	publicité		Web
putophobie	pute		Web
pyrophobie	feu	gr. πῦρ	Web
québécoisphobie	Québec		Web
quovadisophobie	quo vadis		TLFi
qwertyphobie	qwerty		Web

racistophobie	raciste		Web
radarophobie	radar		Web
rapophobie	rap		Web
réformophobie	réforme		Web
religiophobie	religion		Web
requinophobie	requin		Web
richophobie	riche		Web
riscophobie/risquophobie	risque		Web
ritalophobie	rital		Web
robophobie	robot		Web
romanophobie	romano		Web
rupophobie	être sale	gr. <i>ρύπάω</i>	Web
russophobie	Russie / russe		TLFi
saccharophobie	sucre	gr. <i>σάκχαρον</i>	Web
sarkophobie	Sarko(zy)		Web
scatophobie	excrément	gr. <i>σκώρ</i>	Web
schtroumpfophobie	Shtroumpf		Web
schtroumpfophobie	Shtroumpf	(n'importe quoi)	Web
scienc(ç)ophobie	science		Web
sciophobie	ombre	gr. <i>σκία</i>	Web
scopophobie	regarder	gr. <i>σκοπέω</i>	Web
scotophobie	obscurité	gr. <i>σκότιος</i>	Web
scoutophobie	scout		Web
sectophobie	secte		Web
segaphobie	sega		Web
sélaphobie	éclair	gr. <i>σέλας</i>	Web
serbophobie	Serbie/serbe		Web
serpentphobie	serpent		Web
seulophobie	seul		Web
sexophobie	sexe		Web
sidérodromophobie	fer + course	gr. <i>σιδηπος, δρόμος</i>	Web
sidérophobie	astre	lat. <i>sidus</i>	Web
sinistrophobie	gauche	lat. <i>sinister</i>	Web
sinophobie	Chine / chinois		Web
sitiophobie	nourriture	gr. <i>σιτίον</i>	GR
sociophobie	société		GR
sodophobie	sodomie		Web
sonyphobie	Sony		Web
spermatophobie	sperme		Web
spielbergophobie	Spielberg		Web
squalophobie	squale		Web
stalinophobie	Staline		Web
stygiophobie	enfer	gr. <i>Στύξ</i>	Web
suiophobie	soi	lat. <i>sui</i>	Web
superhydrophobie	super + eau	gr. <i>ὑδρων</i>	Web
syphilidophobie	syphilis		TLFi
syphilophobie	syphilis		TLFi, GR
tabacophobie	tabac		Web
tachophobie	rapide	gr. <i>ταχύς</i>	Web
tachophobie	tache		Web
taphophobie	enterrement	gr. <i>ταφή</i>	GR

taxophobia	taxe		Web
technocratophobia	technocrate		Web
technophobia	technologie (nouvelles)		Web
téléphobie	télé(vision)		Web
téléphonophobie	téléphone		Web
tératophobie	montre	gr. τέρας	Web
tétraphobie	quatre	gr. τετρα	Web
teutonophobie	teuton		Web
thaasophobia	être assis	gr. θαάσσω	Web
thalassophobia	mer	gr. θάλασσα	Web
thanatophobia	mort	gr. θάνατος	GR
théophobie	Dieu	gr. θεός	TLFi
thermophobia	chaleur	gr. θερμός	GR
thixophobia	toucher	gr. θίξις	Web
tintinophobia	Tintin		Web
tocophobia	accoucher	gr. τοκάω	Web
tofophobia	tofu		Web
topophobia	lieu	gr. τόπος	TLFi, GR
tortuphobie	tortue		Web
tourophobie	tour		Web
toutphobie	tout		Web
toxicophobia	toxique		TLFi, GR
transphobie	trans(sexuel)		Web
traumatophobia	blesure	gr. τραῦμα	Web
trémophobia	trembler	gr. τρέμω	Web
trichophobia	poil	gr. θριξ	Web
triskaidékaphobie	treize	gr. τρισκαίδεκα	Web
trypanophobia	percer	gr. τρυπάω	Web
tsiganophobia	tsigane		Web
tuberculophobia	tuberculose		Web
turcophobia	Turquie / Turc		TLFi
tuyauphobie	tuyau		Web
urbanophobia	ville	lat. <i>urbs</i>	Web
urbaphobie	ville		Web
vaissellophobie	vaisselle		Web
végétophobia	végétarien		Web
vélibophobia	vélib		Web
vélophobie	vélo		Web
vermiphobie	ver	lat. <i>vermis</i>	Web
vidophobia	vide		Web
vinophobia	vin		Web
virusophobia	virus		Web
webophobia	Web		Web
windophobia	Windows		Web
xénophobia	étranger	gr. ξένος	TLFi, GR
zélophobie	jalousie	gr. ζήλος	Web
zoophilophobia	zoophile		Web
zoophobia	animal	gr. ζοός	GR

Xphone

aéréphone	air	gr. <i>ἀήρ</i>	TLFi
aérophone	air	gr. <i>ἀήρ</i>	TLFi, GR
afarophone	afar		Web
afrikaanophone	afrikaans		Web
albanophone	albanais		Web
algérianophone	algérien		Web
allophone	all	gr. <i>ἄλλος</i>	TLFi, GR
alsacianophone	alsacien		Web
alsacienophone	alsacien		Web
alsacienphone	alsacien		Web
alternophone	alternative		Web
alterphone	autre	lat. <i>alter</i>	Web
amazighophone	amazight		Web
amharicophone	amharique		Web
androphone	Androïd		Web
angevinophone	angevin		Web
anglophone	anglais		TLFi, GR
anglosaxophone	anglosaxon		Web
arabophone	arabe		Web
arabophone	arabe		TLFi, GR
araméophone	araméen		GR
archeophone	ancien	gr. <i>ἀρχαῖος</i>	Web
arménianophone	arménien		Web
audiophone	ouïe	lat. <i>audio</i>	TLFi, GR
audiphone	ouïe	lat. <i>audio</i>	TLFi
autophone	autre		Web
aviophone	avion		Web
aymarophone	aymara		Web
aymarophone	aymara		Web
bablophone	bablos		Web
baladophone	baladeur		Web
bambaraphone	bambara		Web
bambouphone	bambou		Web
banlieuphone	banlieue		Web
barbarophone	barbare		Web
bascophone	basque		Web
béarnophone	béarnais		Web
beaubourophone	beaubourg		Web
belg(e)ophone	belge		Web
bengalophone	bengali		GR
berbérophone	berbère		TLFi, GR
biélorussophone	biélorusse		Web
bigophone	Bigot		TLFi, GR
birmanophone	birman		Web
bisouphone	bisou		Web
blogophone	blog		Web
bordelophone	bordel		Web
bourbonnophone	bourbonnais		Web
bouteillophone	bouteille		TLFi
breizophone	breton	bret. <i>breizhek</i>	Web
bretonophone	breton		Web

brit(t)ophone	breton	lat. <i>britto</i>	Web
brontosauophone	brontosauure		Web
bruitophone	bruit		GR
bulgarophone	bulgare		Web
cancerphone	cancer		Web
cantophone	cantonais		Web
cassetophone	cassette		Web
castillanophone	castillan		Web
catalanophone	catalan		Web
cathodophone	cathode		TLFi
celtophone	celte		Web
champhenophone	champenois		Web
charentophone	charentais		Web
chronophone	temps	gr. <i>χρόνος</i>	Web
chtiophone	chti		Web
cinéphone	cinéma		Web
comorophone	comorien		Web
comtophone	franc-comtois		Web
coptophone	copte		Web
cordophone	corde		TLFi, GR
coréanophone	coréen		Web
corsophone	corse		Web
créolophone	créole		TLFi, GR
croatophone	croate		Web
cryptophone	caché	gr. <i>κρύπτος</i>	Web
danophone	danois		Web
darijophone	darija		Web
dauphinophone	dauphinois		Web
déconophone	déconner		GR
dialectophone	dialecte		Web
dictaphone	dicter		TLFi, GR
dioulaphone	dioula		Web
dodécaphone	douze	gr. <i>δώδεκα</i>	Web
dunkerquophone	dunkerquois		Web
dupinophone	Dupin		Web
dynamophone	dynamo		Web
électrophone	électricité		TLFi, GR
espagnophone	espagnol		Web
estonianophone	estonien		Web
estophone	estonien		Web
éthérophone	éther		Web
europophone	Europe		Web
farsiphone	farsi		Web
fennophone	finnois		Web
finnophone	finnois		Web
flamandophone	flamand		Web
flamandphone	flamand		Web
flamanphone	flamand		Web
flamophone	flamand		Web
franco-provençalophone	franco-provençal		Web
francophone	français		TLFi, GR

franglophone	franglais		Web
frisonophone	frison		Web
frisophone	frison		Web
gadgetophone	gadget		Web
gallophone	gallois		Web
gasconophone	gascon		Web
géophone	terre	gr. γῆ	TLFi, GR
géorgianophone	géorgien		Web
germanophone	allemand		TLFi, GR
gramophone	écriture	gr. γράμμα	Web
grammophone	écriture	gr. γράμμα	Web
gramophone	écriture	gr. γράμμα	TLFi, GR
graphophone	écrire	gr. γράφω	TLFi, GR
grecophone	grec		Web
groenlandophone	groenlandais		Web
guaraniphone	guarani		Web
guaranophone	guarani		Web
gueulophone	gueuler		Web
gulmancémaphone	gulmancéma		Web
haoussaphone	haoussa		Web
hassanophone	hassani		Web
hausaphone	hausa		Web
hébreophone	hébreu		GR
hébreuphone	hébreu		Web
hellenophone	grec		Web
hindiphone	hindi		Web
hispanophone	espagnol		GR
homophone	semblable	gr. ὁμός	TLFi, GR
hongrophone	hongrois		Web
hongrophone	hongrois		Web
hydrophone	eau	gr. ὕδρων	GR
hygiaphone	sain	gr. ὑγιής	GR
ibophone	ibo		Web
icophone	icône		GR
idéophone	idée		TLFi, GR
idiomophone	idiome		Web
idiophone	propre	gr. ἴδιος	GR
indonésianophone	indonésien		Web
indophone	indien		Web
inductophone	induction		Web
interphone	intérieur		TLFi, GR
iranophone	iranien		Web
irlandophone	irlandais		Web
islandophone	islandais		Web
italanophone	italien		Web
italianophone	italien		Web
italionophone	italien		Web
italophone	italien		Web
japanophone	japonais		Web
japonaiphone	japonais		Web
japonophone	japonais		Web

javanophone	javanais		Web
jazzophone	jazz		Web
kannadaphone	kannada		Web
kannadophone	kannada		Web
kazakhophone	kazakh		Web
khmerophone	khmer		Web
kirghizophone	kirghiz		Web
krasslandophone	krasslandais		Web
kurdophone	kurde		Web
ladinophone	ladin		Web
lamellophone	lamelle		Web
languedocianophone	languedocien		Web
languedocienophone	languedocien		Web
laryngophone	larynx		TLFi, GR
latinophone	latin		Web
lettonophone	letton		Web
ligurophone	ligure		Web
limniphone	étang	gr. λίμνη	Web
limousinophone	limousin		Web
lingalaphone	lingala		Web
lingalophone	lingala		Web
lithophone	pierre	gr. λίθος	Web
livrophone	livre		Web
lorrainophone	lorrain		Web
luophone	luo		Web
lusitanophone	portugais		Web
lusitophone	portugais		Web
lusophone	portugais		Web
luxembourgeoisophone	luxembourgeois		Web
luxembourgeophone	luxembourgeois		Web
luxembourgophone	luxembourgeois		Web
lyonnaisophone	lyonnais		Web
macédonophone	macédonien		Web
magnétophone	magnétisme		TLFi, GR
magyarophone	hongrois		Web
maithilophone	maithili		Web
malayophone	malais		Web
malgachophone	malgache		Web
malinképhone	malinké		Web
maltophone	maltais		Web
mandarinophone	mandarin		Web
mandinguophone	mandingue		Web
maohiphone	maohi		Web
marathiphone	marathi		Web
marimbaphone	marimba		Web
marocanophone	marocain		Web
martinophone	Martin		Web
mégaphone	grand	gr. μέγας	TLFi, GR
mélophone	chant	gr. μέλος	TLFi
membranophone	membrane		GR
membraphone	membrane		Web

métallophone	métal		TLFi, GR
microphone	petit	gr. μικρός	TLFi, GR
minouphone	minou		Web
moldavophone	moldave		Web
mongolophone	mongol		Web
monophone	seul	gr. μόνος	Web
monotéléphone	seul + distance	gr. μόνος, τῆλε	TLFi
mooréphone	mooré		Web
nahuatlophone	nahuatl		Web
néerlandophone	néerlandais		GR
népalophone	nepali		Web
neurophone	neurone		Web
nipponophone	japonais		Web
nippophone	japonais		Web
nissartophone	nissart		Web
normandophone	normand		Web
normannophone	normand		Web
occitanophone	occitan		TLFi
occophone	oc		Web
ocophone	oc		Web
ocphone	oc		Web
oïlphone	oïl		Web
opiniophone	opinion		Web
optophone	vue	gr. ὀπτός ('visible')	TLFi
ourdophone	ourdou		Web
ourdouphone	ourdou		Web
ouzbekophone	ouzbek		Web
paléophone	ancien	gr. παλαιός	Web
parisienophone	parisien		Web
parlophone	parler		TLFi, GR
patoiphone	patois		Web
patoisophone	patois		Web
patoisphone	patois		Web
percuphone	percuter		Web
persanophone	persan		Web
peulophone	peul		Web
phonophone	son	gr. φωνή	TLFi
photophone	photo		Web
photophone	lumière	gr. φῶς	TLFi
picardophone	picard		Web
picarophone	picard		Web
pictophone	poitevin		Web
plattophone	platt		Web
pluviophone	pluie		Web
poitevinophone	poitevin		Web
polonophone	polonais		Web
polyphone	plusieurs	gr. πολύς	TLFi, GR
portugophone	portugais		Web
provençalophone	provençal		Web
provençophone	provençal		Web
publiphone	public		GR

pulaarphone	pulaar		Web
pulaarphone	pulaar		Web
quechuaphone	quechua		Web
quechuophone	quechua		Web
radiophone	rayon	lat. <i>radius</i>	TLFi
radiotéléphone	rayon + distance	lat. <i>radius</i> , τῆλε	TLFi, GR
romandophone	romand		Web
romaniphone	romani		Web
romanophone	roman		Web
romophone	rom		Web
roumanophone	roumain		Web
russophone	russe		TLFi, GR
rwandophone	rwandais		Web
saintongeophone	saintongeais		Web
sangophone	sango		Web
sardophone	sarde		Web
sarrusophone	Sarrus		TLFi, GR
sarrussophone	Sarrus		TLFi, GR
saxophone	Sax		TLFi, GR
serbophone	serbe		Web
sérérophone	sérère		Web
sindhophone	sindhi		Web
sinophone	chinois		Web
sismophone	secousse	gr. σεισμός	Web
slavophone	slave		Web
slovacophone	slovaque		Web
slovanophone	slovène		Web
slovénophone	slovène		Web
somaliphone	somali		Web
sommerophone	Sommer		Web
soninképhone	soninké		Web
soudanophone	soudanais		Web
sphygmophone	pulsation	gr. σφυγμός	TLFi
sudrophone	Sudre		Web
suédophone	suédois		Web
sundaphone	sunda		Web
swahiliphone	swahili		Web
tadjikophone	tadjik		Web
tamoulophone	tamoul		Web
taxiphone	taximètre		TLFi, GR
tchécophone	tchèque		Web
télégraphophone	télégraphe		Web
téléphone	distance	gr. τῆλε	TLFi, GR
télougouphone	télougou		Web
tensiophone	tension		TLFi
théâtrophone	théâtre		TLFi, GR
thermophone	chaleur	gr. θερμός	TLFi, GR
tibetophone	tibétain		Web
tourangeaophone	tourangeau		Web
turcophone	turc		GR
turkménophone	turkmène		Web

ukrainiophone	ukrainien		Web
ukrainophone	ukrainien		Web
valaisanophone	valaisan		Web
verlanophone	verlan		Web
verlanphone	verlan		Web
vibraphone	vibrer		TLFi, GR
vidéophone	vidéo		TLFi, GR
vietnamophone	vietnamien		Web
visiophone	vision		TLFi, GR
wallonophone	wallon		Web
wallonphone	wallon		Web
wallophone	wallon		Web
wolophone	wolof		Web
xylophone	bois	gr. ξύλον	TLFi, GR
yoroubaphone	yourouba		Web
zarmaphone	zarma		Web
zoulouphone	zoulou		Web

Xvore

aérivore	air	lat. <i>aer</i>	TLFi
agoravore	agora		Web
agrumovore	agrume		Web
algivore	algue		Web
alguivore	algue		Web
animévore	animé (dessin)		Web
anthropovore	homme	gr. ἄνθρωπος	Web
apivore	abeille	lat. <i>apis</i>	TLFi
aquavore	eau	lat. <i>aqua</i>	Web
arénivore	sable	lat. <i>arena</i>	TLFi
baccivore	baie	lat. <i>bacca</i>	GR
bactérovore	bactérie		Web
bananivore	banane		Web
bandepassantivore	bande passante		Web
bédévore / bdvore	BD		Web
bibliovore	livre	gr. βιβλίον	Web
biblivore	livre	gr. βιβλίον	Web
blogovore	blog		Web
bouquinovore	bouquin		Web
budgetivore	budget		TLFi, GR
carnivore	chair	lat. <i>caro</i>	TLFi, GR
carotivore	carotte		TLFi
céréalivore	céréales		Web
chaussettivore	chaussette		Web
cheddarovore	cheddar		Web
chocolatovore	chocolat		Web
chocolatovore	chocolat		Web
chocolivore	chocolat		Web
chocovore	chocolat		Web
cinémavore	cinéma		Web

cinévore	cinéma		Web
clownivore	clown		Web
comburivore	comburant		Web
concertivore	concert		Web
conductivore	conducteur		Web
croquettivore	croquette		Web
crudivore	cru		Web
déchetivore	déchet		Web
dépositivore	déposer		Web
détritivore	détritus		TLFi, GR
détrivore	détritus		Web
divixovore / divxovore	divix/divx		Web
dvdivore	DVD		Web
encrivore	encre		TLFi
énergétivore	énergie		Web
énergivore	énergie		Web
énergovore	énergie		Web
eurovore	euro		Web
ferrovivore	ferroviaire		Web
feuillivore	feuille		Web
ficellovore	ficelle		Web
filmivore	film		Web
filmovore	film		Web
filovore	fil		Web
florivore	fleur		Web
follivore	feuille	lat. <i>folium</i>	Web
footivore	foot		Web
forumovore	forum		Web
fricovore	fric		Web
fructivore	fruit	lat. <i>fructus</i>	TLFi
frugivore	fruit	lat. <i>frux</i>	TLFi, GR
fruitivore	fruit		Web
fumivore	fumée		TLFi, GR
glandivore	gland		TLFi, GR
graminivore	graminée		Web
granivore	grain		TLFi, GR
grignotivore	grignoter		Web
gryllivore	gryllidés		Web
heliovore	soleil	gr. ἥλιος	Web
herbivore	herbe		TLFi, GR
hydrovore	eau	gr. ὕδρων	Web
ignivore	feu	lat. <i>ignis</i>	TLFi, GR
imagivore	image		Web
imagovore	image		Web
informativore	information		Web
infovore	information		Web
insectivore	insecte		TLFi, GR
instinctivore	instinct		Web
knackivore	Knacki		Web
lactivore	lait	lat. <i>lac</i>	Web
lapinivore	lapin		Web

larvivre	larve		TLFi, GR
lectivore	lecture		Web
lectovore	lecture		Web
léguminivore	légume		TLFi
légumivore	légume		TLFi
lettrivore	lettre		Web
librivore	livre	lat. <i>liber</i>	Web
lignivore	bois	lat. <i>lignum</i>	TLFi
limivore	limon		TLFi, GR
limnivore	étang	gr. <i>λίμνη</i>	TLFi, GR
listivore	liste		Web
livrivore	livre		Web
livrovore	livre		Web
locavore	local		Web
ludivore	jeu	lat. <i>ludus</i>	Web
macdonaldovore	macdonald's		Web
maindœuvrivore	main d'œuvre		Web
mangavore	manga		Web
mathovore	maths		Web
mazdavore	mazda		Web
médiavore	média		Web
mémovore	mémoire		Web
merdivore	merde		Web
microbivore	microbe		Web
moussivore	mousse		TLFi
musicovore	musique		Web
musivore	musique		Web
mycovore	champignon	gr. <i>μύκης</i>	Web
naheulbeukivore	Naheulbeuk		Web
nanarovore	nanar		Web
narutovore	Naruto		Web
nécovore	mort	gr. <i>νεκρός</i>	Web
nectarivore	nectar		Web
nucivore	noix	lat. <i>nux</i>	TLFi
numerivore	numérique		Web
numerivore	numéro		Web
omnivore	tout	lat. <i>omnis</i>	TLFi, GR
oryzivore	riz	gr. <i>ὄρυζα</i>	TLFi
ossivore	os		Web
oursivore	ours		Web
oursivore	oursin		Web
ovnivore	OVNI		Web
pagivore	page		Web
palmivore	palmier		Web
panivore	pain		Web
papivore	papier		TLFi
pâtivore	pâte		Web
pétrolivore	pétrole		Web
pétrolovore	pétrole		Web
photovore	photo		Web
piscivore	poisson	lat. <i>piscis</i>	TLFi, GR

pizzativore	pizza		Web
pizzavore	pizza		Web
planctonivore	plancton		Web
plantivore	plante		Web
pommivore	pomme		Web
productivore	production		Web
psammivore	sable	gr. ψάμμος	TLFi
pubivore	pub(licité)		Web
publivore	publicité		Web
pupivore	pupe		TLFi, GR
radicivore	racine	lat. <i>radix</i>	TLFi, GR
ramenovore	ramen		Web
ramovore	RAM (mémoire)		Web
reptilivore	reptile		TLFi
ressourcivore	ressource		Web
rienivore	rien		Web
rockovore	rock		Web
romanrussivore	roman russe		Web
sanguinivore	sang		GR
sanguivore	sang		Web
séri(e)vore	série (tv)		Web
serpentivore	serpent		Web
sexovore	sexe		Web
soldivore	soldes		Web
sonovore	son		Web
sourivore	souris		Web
spermivore	sperme		Web
spermovore	sperme		Web
stagivore	stage		Web
sucettivore	sucette		Web
sucrivore	sucre		Web
suspensivore	suspension		TLFi
tartrivore	tartre		TLFi
technivore	technologies (nouvelles)		Web
télévore	télévision		TLFi
temporivore	temps	lat. <i>tempus</i>	Web
truffivore	truffe		TLFi
tubérovore	tubercule		TLFi, GR
végétivore	végétaux		TLFi
vermivore	ver	lat. <i>vermis</i>	GR
videovore	vidéo		Web
vinivore	vin		Web
webivore	web		Web
yaourtovore	yaourt		Web

Doublets et triplets *-phage / -vore*

C1	Origine de la base	Doublet	Occurrences	Voyelle affichée	Voyelle attendue ?
aer-	latin	aérivore	3079	<i>	✓
		aériphage	1	<o>	X
alg-	latin	algivore	17520	<i>	✓
		algophage	170	<o>	✓
anthrop-	grec	anthropophage	176400	<o>	✓
		anthropovore	29	<i>	X
bactérie	français	bactériophage	85800	<o>	✓
		bactérovore	2617	<i>	✓
BD	français	bédévore	27330	<i>	X
		bédéphage	24180	<o>	X
bibli-	grec	bibliophage	24940	<o>	✓
		bibliovore	9910	<i>	X
bouquin	français	bouquinovore	133 510	<i>	X
		bouquinophage	4 967	<o>	✓
carn-	latin	carnivore	2 950 000	<i>	✓
		carnophage	63 203	<o>	✓
ciné(ma)	français	cinéphage	85 100	<o>	X
		cinévore	35 400	<i>	X
détritus	français	détritivore	240 800	<i>	✓
		détritiphage	7 490	<o>	X
dvd	français	dvdvore	92 900	<i>	X
		dvdphage	1 352	<o>	X
euro	français	eurovore	45 341	<i>	X
		europophage	1 306	<o>	✓
fumée	français	fumivore	13 420	<i>	✓
		fumophage	5	<o>	✓
larve	français	larvivore	4 028	<i>	✓
		larvipophage	121	<o>	X
libr-	latin	librophage	114 380	<o>	✓
		librivore	19 077	<i>	✓
livre	français	livrovore	8 420	<i>	X
		livrophage	6 050	<o>	✓

		livrivore	1 632	<i>	√
média	français	médiavore	7 810	<i>	X
		médiaphage	3 188	<o>	X
merde	français	merdophage	2 602	<o>	√
		merdivore	313	<i>	√
myc-	grec	mycophage	6 980	<o>	√
		mycovore	26	<i>	X
necr-	grec	nécrophage	91 300	<o>	√
		nécrovore	<i>	10 200	X
nectar	français	nectarivore	22 100	<i>	√
		nectariphage	292	<o>	X
omn-	latin	omnivore	445 000	<i>	√
		omniphage	334	<o>	X
pétrole	français	pétrolivore	1019	<i>	√
		pétrolophage	487	<o>	√
		pétrolovore	2	<i>	X
publicité	français	publivore	114 800	<i>	√
		publiophage	1 159	<o>	X
technologies (nouvelles)	français	technophage	4 720	<o>	√
		technivore	4 202	<i>	√
télé(vision)	français	téléphage	65400	<o>	X
		télévore	34 020	<i>	X
végétal	français	végétivore	2 783	<i>	√
		végétophage	126	<o>	√
web	français	webivore	1 722	<i>	√
		webophage	650	<o>	√
		webphage	87	<o>	X

Tableau 5.28: Comparaison des doublets (ou triplets) en *-vore* et *-phage* du point de vue du segment de liaison

Bibliographie

- ADAMS, V. (1973). *An Introduction to Modern English Word-Formation*, volume 7. Longman, London.
- ALLARD, J. (1971). *Grammaire grecque : à l'usage des classes de la 4^e aux classes préparatoires*. Classiques Hachette. Hachette, Paris.
- AMIOT, D. (2004). Préfixes ou prépositions? Le cas de *sur* (-), *sans* (-), *contre* (-) et les autres. *Lexique*, 16:67–83.
- AMIOT, D. (2011). Paradigmes, radicaux supplétifs et constituants néoclassiques en morphologie constructionnelle. *Recherches ACLIF : Actes du Séminaire de Didactique Universitaire*, (8):21–36.
- AMIOT, D. et DAL, G. (2007). Integrating Neoclassical Combining Forms into a Lexeme-Based Morphology. In BOOIJ, G., DUCCESCHI, L., FRADIN, B., GUEVARA, E., SCALISE, S. et RALLI, A., éditeurs : *On-Line Proceedings of Fifth Mediterranean Morphology Meeting*, pages 323–336, Bologna : Università di Bologna.
- AMIOT, D. et DAL, G. (2008). La composition néoclassique en français et ordre des constituants. In AMIOT, D., éditeur : *La composition dans les langues dans une perspective typologique*, pages 89–113. Artois Presses Université, Arras.
- AMIOT, D. et MONTERMINI, F. (2009). Affixes et mots grammaticaux. In FRADIN, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Aperçus de morphologie du français*, pages 127–141. Presses Universitaires de Vincennes, Saint-Denis.
- AMIOT, D. et VAN GOETHEM, K. (2010). Le statut de *-clé* et de *sleutel-* dans *mot-clé* / *sleutelwoord* : une analyse unifiée? In NEVEU, F., MUNI TOKE,

- V., KLINGER, T., DURAND, J., MONDADA, L. et PRÉVOST, S., éditeurs : *Actes du 2e Congrès Mondial de Linguistique Française*, Paris. Institut de Linguistique Française.
- ANDERSON, S. R. (1992). *A-morphous Morphology*. Cambridge University Press, Cambridge.
- ANDRÉ, J. (1986). Sur la constitution des langues techniques en latin. *Sciences et techniques à Rome, Études de lettres*, (1):5–18.
- ARCODIA, G. F. et BASCIANO, B. (2012). Neoclassical compounding beyond Europe : the case of East Asia. *Verbum*, XXXIV(2):233–256.
- ARONOFF, M. (1976). *Word Formation in Generative Grammar*. The MIT Press, Cambridge, Massachusetts and London, England.
- ARONOFF, M. (1984). Word formation and lexical semantics. *Quaderni di semantica*, 5(1):45–49.
- ARONOFF, M. (1994). *Morphology by Itself : Stems and Inflectional Classes*. MIT press, Cambridge, MA.
- ARONOFF, M. (2007). In the beginning was the word. *Language*, (83):803–830.
- ARONOFF, M. et FUEDEMAN, K. (2011). *What is Morphology ?* John Wiley & Sons, New York.
- ARONOFF, M. et SRIDHAR, S. N. (1988). Prefixation in Kannada. In HAMMOND, M. et NOONAN, M., éditeurs : *Theoretical Morphology*, pages 179–191. Academic Press, San Diego and London.
- BAAYEN, H. (1992). Quantitative aspects of morphological productivity. In BOOIJ, G. E. et VAN MARLE, J., éditeurs : *Yearbook of Morphology 1991*, pages 109–149. Kluwer Academic Publishers, Dordrecht.
- BAAYEN, H. (1993). On frequency, transparency and productivity. In BOOIJ, G. E. et VAN MARLE, J., éditeurs : *Yearbook of Morphology 1992*, pages 181–208. Kluwer Academic Publishers, Dordrecht.
- BAAYEN, H. et LIEBER, R. (1991). Productivity and English derivation : a corpus-based study. *Linguistics*, 29(5):801–844.
- BARTNING, I. (1976). *Remarques sur la syntaxe et la sémantique des pseudo-adjectifs dénominaux en français*. Almqvist & Wiksell Int., Stockholm.

- BAUER, L. (1979). Against word-based morphology. *Linguistic Inquiry*, 10(3):508–509.
- BAUER, L. (1983). *English Word-Formation*. Cambridge University Press, Cambridge.
- BAUER, L. (1998). Is there a class of neoclassical compounds, and if so is it productive? *Linguistics*, 36(3):403–422.
- BAUER, L. (2005). The borderline between derivation and compounding. In DRESSLER, W. U., RAINER, F., KASTOVSKY, D. et PFEIFFER, O., éditeurs : *Morphology and its Demarcations*, pages 97–108. John Benjamins, Amsterdam / Philadelphia.
- BENVENISTE, E. (1966). *Problèmes de linguistique générale*, volume 2. Gallimard, Paris.
- BISETTO, A. (2010). Relational adjectives crosslinguistically. *Lingue e linguaggio*, 9(1):65–85.
- BISETTO, A. et SCALISE, S. (2009). The classification of compounds. In LIEBER, R. et ŠTEKAUER, P., éditeurs : *The Oxford Handbook of Compounding*, pages 34–54. Oxford University Press, Oxford.
- BIVILLE, F. (1985). Les éléments grecs du vocabulaire français. *L'information grammaticale*, 24(1):3–8.
- BIVILLE, F. (2002). La Dimension grecque de la dérivation latine. Interférences et emprunts. In KIRCHER, C., éditeur : *Grammaire fondamentale du latin, tome IX : Création lexicale, la formation des noms par dérivation suffixale.*, pages 353–390. Peeters, Louvain.
- BLEVINS, J. P. (2006). Word-based morphology. *Journal of Linguistics*, 42:531–573.
- BLEVINS, J. P. (2014). The morphology of words. In GOLDRICK, M., MIOZZO, M. et FERREIRA, V., éditeurs : *The Oxford Handbook of Language Processing*, pages 152–164. Oxford University Press, Oxford.
- BLOOMFIELD, L. (1933). *Languages*. Allen and Unwin, London. [Traduction française (1970) : *Le langage*. Paris : Payot].
- BONAMI, O. et BOYÉ, G. (2003). Supplétion et classes flexionnelles dans la conjugaison du français. *Langages*, 152:102–126.

- BONAMI, O. et BOYÉ, G. (2005). Construire le paradigme d'un adjectif. *Recherches Linguistiques de Vincennes*, 34:77–98.
- BONAMI, O., BOYÉ, G. et KERLEROUX, F. c. (2009). L'allomorphie radicale et la relation flexion-construction. In FRADIN, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Aperçus de morphologie du français*, pages 103–125. Presses Universitaires de Vincennes, Saint-Denis.
- BOOIJ, G. (2002). Prosodic restrictions on affixation in Dutch. In BOOIJ, G. et VAN MARLE, J., éditeurs : *Yearbook of Morphology 2001*, pages 183–201. Kluwer Academic Publishers, Dordrecht.
- BOOIJ, G. (2005). Compounding and derivation. Evidence for Construction Morphology. In DRESSLER, W. U., RAINER, F., KASTOVSKY, D. et PFEIFFER, O., éditeurs : *Morphology and its Demarcations*, pages 109–132. John Benjamins, Amsterdam / Philadelphia.
- BOOIJ, G. (2010). *Construction Morphology*. Oxford University Press, Oxford.
- BOOIJ, G. et LIEBER, R. (2004). On the paradigmatic nature of affixal semantics in English and Dutch. *Linguistics*, 42(2):327–358.
- BOOIJ, G. E. (1992). Compounding in Dutch. *Rivista di Linguistica*, 4(1):37–59.
- BOTTÉRO, J. (1987). *Mésopotamie : l'écriture, la raison et les dieux*. Gallimard, Paris.
- BOUFFARTIGUE, J. et DELRIEU, A.-M. (1996a). *Les racines grecques*, volume 1. Belin, Paris.
- BOUFFARTIGUE, J. et DELRIEU, A.-M. (1996b). *Les racines latines*, volume 2. Belin, Paris.
- BOWKER, L. et PEARSON, J. (2002). *Working with Specialized Language : a Practical Guide to Using Corpora*. Routledge, London.
- BRINTON, L. J. et TRAUGOTT, E. C. (2005). *Lexicalization and Language Change*. Cambridge University Press, Cambridge.
- BRUCALÉ, L. (2012). Latin compounds. *Probus, International Journal of Latin and Romance Linguistics*, 24(1):93–117.

- BRUNOT, F. (1922). *La pensée et la langue : méthode, principes et plan d'une théorie nouvelle du langage appliquée au français*. Masson et cie, Paris.
- BRUNOT, F. (1927a). *Histoire de la langue française des origines à 1900. Tome IX. La Révolution et l'Empire*. Armand Colin, Paris.
- BRUNOT, F. (1927b). *Histoire de la langue française des origines à 1900. Tome X. La langue classique dans la tourmente*. Armand Colin, Paris.
- BURZIO, L. (2002). Surface-to-surface morphology : when your representations turn into constraints. In BOUCHER, P., éditeur : *Many morphologies*, pages 142–177. Cascadilla Press, Somerville, MA.
- CABRÉ, M. T. (1998). *La terminologie : théorie, méthode et applications*. Presses de l'Université d'Ottawa/Armand Colin, Ottawa/Paris.
- CHANTRAINE, P. (1933). *La formation des noms en grec ancien*, volume 38. Champion, Paris.
- CORBETT, G. G. (2010). Canonical derivational morphology. *Word Structure*, 3(2):141–155.
- CORBIN, D. (1989). Form, structure and meaning of constructed words in an associative and stratified lexical component. In BOOIJ, G. et VAN MARLE, J., éditeurs : *Yearbook of Morphology 1989*, pages 31–54. Kluwer Academic Publishers, Dordrecht.
- CORBIN, D. (1992). Hypothèses sur les frontières de la composition nominale. *Cahiers de grammaire*, 17:26–55.
- CORBIN, D. (1997). Entre les mots possibles et les mots existants : les unités lexicales à faible probabilité d'actualisation. In CORBIN, D., FRADIN, B., HABERT, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Mots possibles, mots existants. Actes du 1er forum de morphologie*, volume 1, pages 79–90, Villeneuve d'Ascq. Université de Lille.
- CORBIN, D. (1999). Pour une théorie sémantique de la catégorisation affixale. *Faits de langues*, 14(7):65–77.
- CORBIN, D. (2001). Préfixes et suffixes : du sens aux catégories. *Journal of French Language Studies*, 11(01):41–69.

- CORBIN, D. (2004). Français (Indo-européen : Roman). In BOOIJ, G., LEHMANN, C., MUGDAN, J. et SKOPETEAS, S., éditeurs : *Morphologie/Morphology. Ein internationales Handbuch zur Flexion und Wortbildung/An International Handbook on Inflection and Word Formation*, volume 2, pages 1285–1299. Mouton de Gruyter, Berlin.
- COTTEZ, H. (1980). *Dictionnaire des structures du vocabulaire savant : éléments et modèles de formation*. Le Robert, Paris.
- DAL, G. (2003). Productivité morphologique : définitions et notions connexes. *Langue française*, 140(1):3–23.
- DAL, G. et NAMER, F. (2012). Faut-il brûler les dictionnaires ? Ou comment les ressources numériques ont révolutionné les recherches en morphologie. In NEVEU, F., MUNI TOKE, V., BLUMENTHAL, P., KLINGER, T., LIGAS, P., PRÉVOST, S. et TESTON-BONNARD, S., éditeurs : *Actes du 3e Congrès Mondial de Linguistique Française*, pages 1261–1276, Paris. Institut de Linguistique Française.
- DALTON-PUFFER, C. et PLAG, I. (2000). Categorywise, some compound-type morphemes seem to be rather suffix-like : On the status of *-ful*, *-type*, and *-wise* in present day English. *Folia Linguistica*, 34(3-4):225–244.
- DARMESTER, A. (1874). *Traité de la formation des mots composés dans la langue française : Comparée aux autres langues romanes et au latin*, volume 19. A. Franck, Paris.
- DI SPALDRO, J., AUGER, P. et LADOUCEUR, J. (2010). Le calque technoscientifique : un procédé néologique avantageux pour la terminologie française ? *Neologica*, 4:163–184.
- DRESSLER, W. U. (1985). Sur le statut de la suppléance dans la Morphologie Naturelle. *Langages*, 78:41–56.
- FIRTH, J. R. (1957). *Papers in Linguistics*. Oxford University Press, Oxford.
- FRADIN, B. (1997). Les mots-valises : une forme productive d'existants impossibles. In CORBIN, D., FRADIN, B., HABERT, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Mots possibles, mots existants. Actes du 1er forum de morphologie*, volume 1, pages 101–110, Villeneuve d'Ascq. Université de Lille.
- FRADIN, B. (2000). Combining forms, blends and related phenomena. In DOLESCHAL, U. et THORNTON, A. M., éditeurs : *Extragrammatical and Marginal Morphology*, volume 12, pages 11–57. Lincom, Munich.

- FRADIN, B. (2003). *Nouvelles approches en morphologie*. Presses universitaires de France, Paris.
- FRADIN, B. (2008). Les adjectifs relationnels et la morphologie. In FRADIN, B., éditeur : *La raison morphologique. Hommage à la mémoire de Danielle Corbin.*, pages 69–92. John Benjamins, Amsterdam/Philadelphia.
- FRADIN, B. (2014). La variante et le double. In VILLOING, F., DAVID, S. et LEROY, S., éditeurs : *Foisonnements morphologiques. Études en hommage à Françoise Kerleroux*, pages 109–147. Presses Universitaires de Paris Ouest, Nanterre.
- FRADIN, B. et KERLEROUX, F. c. (2009). L'identité lexémique. In FRADIN, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Aperçus de morphologie du français*, pages 83–102. Presses universitaires de Vincennes, Saint-Denis.
- FRADIN, B., MONTERMINI, F. et PLÉNAT, M. (2009). Morphologie grammaticale et extra-grammaticale. In FRADIN, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Aperçus de morphologie du français*, pages 21–45. Presses Universitaires de Vincennes, Saint-Denis.
- FREGE, G. (1980). Letter to Jourdain. In GABRIEL, G., HERMES, H., KAMBARTEL, F., THIEL, C. et VERAART, A., éditeurs : *Philosophical and mathematical correspondance*, pages 78–80. Chicago University Press, Chicago.
- FRUYT, M. (1987). L'emprunt suffixal du latin au grec. *Bulletin de la Société de Linguistique de Paris*, 82(1):227–255.
- FRUYT, M. (1996). Lexique et conscience linguistique en latin : sens fonctionnel et sens parallèle. In FRUYT, M. et MOUSSY, C., éditeurs : *Structures Lexicales du latin*, numéro 3 de *Lingua Latina*. P.U.P.S, Paris.
- FRUYT, M. (2002). Constraints and productivity in Latin nominal compounding. *Transactions of the Philological Society*, 100(3):259–287.
- FRUYT, M. (2011). Word-Formation in Classical Latin. In CLACKSON, J., éditeur : *A Companion to the Latin Language*, pages 157–175. Wiley-Blackwell, Malden, MA.
- GAETA, L. et RICCA, D. (2009). *Composita solvantur* : Compounds as lexical units or morphological objects. *Italian Journal of Linguistics*, 21(1):35–70.

- GAMBIER, Y. (1991). Travail et vocabulaire spécialisés : prolégomènes à une socio-terminologie. *Meta : Journal des traducteurs/ Meta : Translators' Journal*, 36(1):8–15.
- GIRAUDO, H. (2015). *Où est la morphologie ? Regards croisés de la linguistique et de la psycholinguistique*. Mémoire d'habilitation à diriger des recherches, Université de Toulouse.
- GRANDI, N. (2002). *Morfologia in contatto. Le costruzioni valutative nelle lingue del Mediterraneo*. Franco Angeli, Milano.
- GRANDI, N. et POMPEI, A. (2010). Per una tipologia dei composti del greco. In PUTZU, I., PAULIS, G., NIEDDU, G. et CUZZOLIN, P., éditeurs : *La morfologia del greco tra tipologia e diacronia*, pages 209–232, Milano. Franco Angeli.
- GREFENSTETTE, G. (1999). The World Wide Web as a resource for example-based machine translation tasks. *Translating and the Computer*, 21:4–7.
- GRÉSILLON, A. (1984). *La règle et le monstre : le mot-valise : interrogations sur la langue, à partir d'un corpus de Heinrich Heine*. M. Niemeyer, Tübingen.
- GUEVARA, E. et SCALISE, S. (2009). Searching for universals in compounding. In *Universals of Language Today*, pages 101–128. Springer, Amsterdam.
- GUILBERT, L. (1975). *La créativité lexicale*. Larousse, Paris.
- GUILBERT, L., LAGANE, R. et NIOBEY, G. (1986). *Grand Larousse de la langue française : en sept volumes*. Larousse, Paris.
- HALLE, M. (1973). Prolegomena to a theory of word formation. *Linguistic inquiry*, 4:3–16.
- HARRIS, Z. S. (1954). Distributional structure. *Word*, 10(23):146–162.
- HARRIS, Z. S. (1968). *Mathematical Structures of Language*. John Wiley & Sons, New York.
- HARRIS, Z. S. (1970). *Distributional Structure*. Springer, Amsterdam.
- HASPELMATH, M. (2001). The European linguistic area : standard average European. In HASPELMATH, M., KÖNIG, E., OESTERREICHER,

- W. et RAIBLE, W., éditeurs : *Language Typology and Language Universals. (Handbücher zur Sprach- und Kommunikationswissenschaft)*, pages 1492–1510. Mouton de Gruyter, Berlin.
- HATHOUT, N. (2009). *Contributions à la description de la structure morphologique du lexique et à l'approche extensive en morphologie*. Mémoire d'habilitation à diriger des recherches, Université de Toulouse.
- HATHOUT, N. (2011). Une approche topologique de la construction des mots : propositions théoriques et application à la préfixation en *anti*. In ROCHÉ, M., BOYÉ, G., HATHOUT, N., LIGNON, S. et PLÉNAT, M., éditeurs : *Des unités morphologiques au lexique*, pages 251–318. Hermès Science-Lavoisier, Paris.
- HATHOUT, N., MONTERMINI, F. et TANGUY, L. (2008). Extensive data for morphology : using the World Wide Web. *Journal of French Language Studies*, 18(01):67–85.
- HATHOUT, N., NAMER, F., PLÉNAT, M. et TANGUY, L. (2009). La collecte et l'utilisation des données en morphologie. In FRADIN, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Aperçus de morphologie du français*, pages 267–287. Presses Universitaires de Vincennes, Saint-Denis.
- HATZFELD, A., DARMESTETER, A., THOMAS, A. et OTHERS (1890). *Traité de la formation de la langue. Livre premier*. Delagrave, Paris.
- HOCKETT, C. F. (1958). A course in modern linguistics. *Language Learning*, 8(3-4):73–75.
- HOFFMANN, L. (1979). Toward a Theory of LSP. Elements of a methodology of LSP analysis. *Fachsprache.*, 1(2):12–17.
- HUMBLEY, J. (1974). Vers une typologie de l'emprunt linguistique. *Cahiers de lexicologie*, 25(2):46–70.
- IACOBINI, C. (1999). I prefissi dell'italiano. In BENINCÁ, P., MIONI, A. et VANELLI, L., éditeurs : *Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI Congresso della Società di Linguistica Italiana*, pages 369–399.
- IACOBINI, C. (2004). Composizione con elementi neoclassici. In GROSSMAN, M. et RAINER, F., éditeurs : *La formazione delle parole in italiano*, pages 69–95. Niemeyer, Tübingen.

- JACKENDOFF, R. (1975). Morphological and semantic regularities in the lexicon. *Language*, 51:639–671.
- JAKOBSON, R. (1936). Beitrag zur Allgemeinen Kasuslehre : Gesamtbedeutungen der russischen Kasus. In HAMP, E., HOUSEHOLDER, F. et AUSTRERLITZ, R., éditeurs : *Readings in Linguistics II*, pages 51–95. University of Chicago Press, Chicago.
- JESPERSEN, O. (1943). *A Modern English Grammar on Historical Principles, part 6 : Morphology*. George Allen and Unwin, London.
- KASTOVSKY, D. (2009). *Astronaut, astrology, astrophysics* : about combining forms, classical compounds and affixoids. In MCCONCHIE, R., HONKAPOHJA, A. et TYRKKÖ, J., éditeurs : *Selected proceedings of the 2008 symposium on new approaches in English historical lexis (HEL-LEX 2)*., pages 1–13, Somerville, MA. Cascadilla Proceedings Project.
- KERLEROUX, F. c. (2003). Morpho-logie : la forme et l'intelligible. *Langages*, 37(152):12–32.
- KILANI-SCHOCH, M. et DRESSLER, W. U. (1992). *Pro-l-o, intell-o, gauch-o* et les autres. Propriétés formelles de deux opérations du français parlé. *Romanistisches Jahrbuch*, 43:65–86.
- KOCOUREK, R. (1982). *La langue française de la technique et de la science*. John Benjamins, Amsterdam / Philadelphia.
- KOGA, K. (2012). *Enquête sur la productivité des noms composés du type pause-café en français*. Mémoire de M2, Université de Bordeaux Montaigne.
- LANGACKER, R. W. (1987). *Foundations of Cognitive Grammar : Theoretical Prerequisites*, volume 1. Stanford University Press, Stanford.
- LASSERRE, M. (2011). *La formation des mots-valises en français*. Mémoire de M2, Université de Bordeaux Montaigne.
- LASSERRE, M. (2015). What are you afraid of? The construction of meaning in *X-(o)phobie* lexemes. *Acta Linguistica Hungarica*, 62(4):477–495.
- LASSERRE, M. et MONTERMINI, F. (2012). Une approche distributionnelle pour l'analyse des composés néoclassiques. *Verbum*, XXXIV(2):195–212.
- LASSERRE, M. et MONTERMINI, F. (2014a). How is the meaning of complex lexemes constructed? A study of neoclassical compounds. *Italian Journal of Linguistics*, 26(2):157–181.

- LASSERRE, M. et MONTERMINI, F. (2014b). Pour une typologie des lexèmes construits : entre composition, composition néoclassique et affixation. In NEVEU, F., BLUMENTHAL, P., HRIBA, L., GERSTENBERG, A., MEIN-SCHAEFER, J. et PRÉVOST, S., éditeurs : *Actes du 4e Congrès Mondial de Linguistique Française (CMLF 2014)*, pages 1797–1812, Paris. Institut de Linguistique Française.
- LASSERRE, M. et MONTERMINI, F. (2015). Neoclassical compounds as relational adjectives. In AUDRING, J., KOUTSOUKOS, N., MASINI, F. et RAFFAELLI, I., éditeurs : *Morphology and Semantics, MMM9 On-line Proceedings*, pages 64–76.
- LAZARIDOU, A., MARELLI, M., ZAMPARELLI, R. et BARONI, M. (2013). Compositional-ly Derived Representations of Morphologically Complex Words in Distributional Semantics. In *Proceedings of ACL*, Sofia.
- LE FEUVRE, C. (2012). Hippocrate, la mégalospanchnie et la splanchnomégalie. *Verbum*, XXXIV(2):179–193.
- LEHRER, A. (1998). *Scapes, holics, and thons* : The semantics of English combining forms. *American Speech*, 73(1):3–28.
- LEHRER, A. (2000). Are Affixes Signs : The semantic relationships of English derivational affixes. In DRESSLER, W. U., PFEIFFER, O. E., PÖCHTRAGER, M. A. et RENNISON, J. R., éditeurs : *Morphological Analysis in Comparison*, pages 143–154. John Benjamins, Amsterdam/Philadelphia.
- LENCI, A. (2008). Distributional semantics in linguistic and cognitive research. *Italian Journal of Linguistics*, 20(1):1–31.
- LETURGIE, A. (2012). *L'amalgamation lexicale en français : approches lexicologique et morphologique : Vers une grammaire de l'amalgamation lexicale en français*. Thèse de doctorat, Université de Cergy Pontoise.
- LEVENSHTEIN, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics doklady*, 10(8):707–710.
- LIEBER, R. (2003). Compound interpretation : Lexical semantics, not syntax. In BOOIJ, G., DE CESARIS, J., RALLI, A. et SCALISE, S., éditeurs : *Topics in Morphology : Selected Papers from the Third Mediterranean Morphology Meeting.*, pages 241–253, Barcelona. IULA, Universitat Pompeu Fabra.
- LIGNON, S. (2000). *La suffixation en -ien : aspects sémantiques et phonologiques*. Thèse de doctorat, Université Toulouse II - Le Mirail.

- LIGNON, S. (2008). Les noms de spécialistes en *-iste* et en *-ien* : le chimiste perturbé ou comment le physicien se réajuste. In VAXELAIRE, B., MARSAC, F., SOCK, R. et KLEIBER, G., éditeurs : *Perturbations et Réajustements*, pages 287–295. Presses Universitaires de l'Université de Marc Bloch, Strasbourg.
- LIGNON, S. (2013). *-iser* and *-ifier* suffixations in French : Verify data to verize hypotheses ? In MONTERMINI, F., BOYÉ, G. et TSENG, J., éditeurs : *Décembrettes 7*, pages 109–132. Lincom Europa.
- LINDSAY, M. et ARONOFF, M. (2015). Partial organization in languages : la langue est un système où la plupart se tient. In MONTERMINI, F., BOYÉ, G. et TSENG, J., éditeurs : *Selected Proceedings of the 7th Décembrettes.*, München. Lincom Europa.
- LÜDELING, A. (2006). Neoclassical compounding. In BROWN, K., ANDERSON, A., BAUER, L., BERNS, M., MILLER, J. et GRAEME, H., éditeurs : *Encyclopedia of Language and Linguistics, 2nd Edition.*, pages 580–582. Elsevier, Amsterdam.
- LÜDELING, A., EVERT, S. et BARONI, M. (2006). Using Web data for linguistic purposes. *Language and Computers*, 59(1):7–24.
- LÜDELING, A., SCHMID, T. et KIOKPASOGLU, S. (2002). Neoclassical word formation in German. In BOOIJ, G. et VAN MARLE, J., éditeurs : *Yearbook of Morphology 2001*, pages 253–283. Kluwer Academic Publishers, Dordrecht.
- MARCHAND, H. (1969). *The Categories and Types of Present-day English Word-Formation : A Synchronic-Diachronic Approach*. Numéro 13. C.H. Beck'sche Verlagsbuchhandlung, München.
- MATTHEWS, P. (1974). *Morphology*. Cambridge University Press, Cambridge.
- MAUREL, J.-P. (1993). Des adjectifs de relation en latin. *L'information grammaticale*, 58(1):23–26.
- MELLONI, C. et BISETTO, A. (2010). Parasynthetic compounds : data and theory. In SCALISE, S. et VOGEL, I., éditeurs : *Cross-Disciplinary Issues in Compounding*, pages 199–217. John Benjamins, Amsterdam / Philadelphia.
- MEL'ČUK, I. A. (1976). On suppletion. *Linguistics*, 14(170):45–90.

- MEYER, I. et MACKINTOSH, K. (2000). L'étirement du sens terminologique : aperçu du phénomène de la déterminologisation. In BÉJOINT, H. et THOIRON, P., éditeurs : *Le sens en terminologie*, pages 198–217. Presses Universitaires de Lyon, Lyon.
- MONCEAUX, A. (1997). Adjectifs de relation, complémentation et sous-classification. *Langages*, 31(126):39–59.
- MONTERMINI, F. (2001). The Unitary Base Hypothesis and the semantics of word formation rules. In BOUILLON, P. et KANZAKI, K., éditeurs : *First International Workshop on Generative Approaches to the Lexicon.*, pages 26–28, Université de Genève.
- MONTERMINI, F. (2002). *Le système préfixal en italien contemporain*. Thèse de doctorat, Université de Paris X-Nanterre, Università degli Studi di Bologna.
- MONTERMINI, F. (2008). *Il lato sinistro della morfologia. La prefissazione in italiano e nelle lingue del mondo*. FrancoAngeli, Milano.
- MONTERMINI, F. (2010a). *La morphophonologie de l'italien : une approche lexicaliste*. Mémoire d'habilitation à diriger des recherches, Université de Toulouse.
- MONTERMINI, F. (2010b). Units in compounding. In SCALISE, S. et VOGEL, I., éditeurs : *Cross Disciplinary Issues in Compounding.*, pages 77–92. John Benjamins, Amsterdam/Philadelphia.
- MORLANE-HONDÈRE, F. c. (2013). *Une approche linguistique de l'évaluation des ressources extraites par analyse distributionnelle automatique*. Thèse de doctorat, Université Toulouse le Mirail-Toulouse II.
- MOUNIN, G. (1979). La linguistique comme science auxiliaire dans les disciplines juridiques. *Meta : Journal des traducteurs/Meta :Translators' Journal*, 24(1):9–17.
- NAMER, F. (2003). WaliM : Valider les unités morphologiques complexes par le Web. In FRADIN, B., DAL, G., HATHOUT, N., KERLEROUX, F. c., PLÉNAT, M. et ROCHÉ, M., éditeurs : *Les unités morphologiques. Actes du 3e forum de morphologie*, volume 3, pages 142–150, Villeneuve d'Ascq. Presses Universitaires du Septentrion.
- NAMER, F. (2007). Composition néoclassique : est-on dans l'«hétéromorphosémie»? In HATHOUT, N. et MONTERMINI, F.,

- éditeurs : *Morphologie à Toulouse - Actes du colloque international de Morphologie*, pages 185–206, München. Lincom Europa.
- NICOLAS, C. (1994). Le procédé du calque sémantique. *Cahiers de lexicologie*, 65(2):75–101.
- NICOLAS, C. (1996). *Utraque lingua : le calque sémantique : domaine gréco-latin*. Peeters, Louvain.
- NOAILLY, M. (1990). *Le substantif épithète*. Presses universitaires de France, Paris.
- OLSEN, S. (2000). Composition. In BOOIJ, G., LEHMANN, C. et MUGDAN, J., éditeurs : *Morphology*, volume 1, pages 897–916. Mouton de Gruyter, Berlin.
- ONIGA, R. (1992). Compounding in Latin. *Rivista di linguistica*, 4(1):97–116.
- PASCOLI, M. (2015). *A computational morphological analysis of Italian verbal system*. Thèse de doctorat, Università degli studi di Verona & Université Toulouse II - Jean Jaurès.
- PLAG, I. (1996). Selectional restrictions in English suffixation revisited : a reply to Fabb (1988). *Linguistics*, 34(4):769–798.
- PLAG, I. (2003). *Word-Formation in English*. Cambridge University Press, Cambridge.
- PLAG, I. (2004). Syntactic category information and the semantics of derivational morphological rules. *Folia Linguistica*, 38(3-4):193–226.
- PLÉNAT, M. (2008). Le Thème L de l'adjectif et du nom. In DURAND, J., HABERT, B. et LAKS, B., éditeurs : *Actes du Congrès Mondial de Linguistique Française*, volume 2008, pages 1613–1626, Paris. Institut de Linguistique Française.
- PLÉNAT, M. (2009). Les contraintes de taille. In FRADIN, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Aperçus de morphologie du français*, pages 47–63. Presses Universitaires de Vincennes, Saint-Denis.
- POPPER, K. (1973). *La logique de la découverte scientifique*. Payot, Lausanne.
- PRINCE, A. et SMOLENSKY, P. (2008). *Optimality Theory : Constraint Interaction in Generative Grammar*. John Wiley & Sons, New York.

- PRĆIĆ, T. (2005). Prefixes vs initial combining forms in English : A lexicographic perspective. *International Journal of Lexicography*, 18(3):313–334.
- PRĆIĆ, T. (2008). Suffixes vs final combining forms in English : A lexicographic perspective. *International Journal of Lexicography*, 21(1):1–22.
- RAINER, F. (2003). Semantic fragmentation in word-formation : The case of Spanish *-azo*. In SINGH, R. et STAROSTA, S., éditeurs : *Explorations in seamless morphology*, pages 197–211. SAGE, New Delhi/Thousand Oaks/London.
- RAINER, F. (2012). Morphological metaphysics : virtual, potential, and actual words. *Word Structure*, 5(2):165–182.
- RAINER, F. (2013). L'interprétation des composés latins du type *aurifer* en français. In DAL, G. et AMIOT, D., éditeurs : *Repères en morphologie*, pages 182–214. Lille. [http://stl.recherche.univ-lille3.fr/textesenligne/Reperes-Morphologie/Rainer-Repères_morphologie_p182-214.pdf]. Réédition en ligne du laboratoire STL de Rainer F. (2002), L'intégration des composés latins du type *aurifer* en français, *Silexicales* 3, pp. 151-168.
- RAINER, F. et VARELA, S. (1992). Compounding in Spanish. *Rivista di linguistica*, 4(1):117–142.
- RALLI, A. (2013). *Compounding in Modern Greek*. Springer, Amsterdam.
- REY, A. (1992). *Dictionnaire historique de la langue française*. Dictionnaires le Robert, Paris.
- REY-DEBOVE, J. (1973). La sémiotique de l'emprunt lexical. *Travaux de Linguistique et de Littérature*, 11(1):109–23.
- RIEGEL, M., PELLAT, J.-C. et RIOUL, R. (1994). *Grammaire méthodique du français*. Presses universitaires de France, Paris.
- ROCHÉ, M. (1997). *Briard, bougeoir et camionneur*. Dérivés aberrants, dérivés possibles. In CORBIN, D., FRADIN, B., HABERT, B., KERLEROUX, F. o. et PLÉNAT, M., éditeurs : *Mots possibles, mots existants. Actes du 1er forum de morphologie*, volume 1, pages 241–250, Villeneuve d'Ascq. Université de Lille.
- ROCHÉ, M. (2003). L'interfixe est-il une unité morphologique? In FRADIN, B., DAL, G., HATHOUT, N., KERLEROUX, F. c., PLÉNAT, M. et ROCHÉ,

- M., éditeurs : *Les unités morphologiques. Actes du 3e forum de morphologie*, volume 3, pages 169–178, Villeneuve d'Ascq. Presses Universitaires du Septentrion.
- ROCHÉ, M. (2006). Comment les adjectifs sont sémantiquement construits. *Cahiers de grammaire*, (30):373–387.
- ROCHÉ, M. (2009). Pour une morphologie *lexicale*. *Mémoires de la Société de Linguistique de Paris. Nouvelle série*, (17):65–87.
- ROCHÉ, M. (2010). Base, thème, radical. *Recherches linguistiques de Vincennes*, (1):95–134.
- ROCHÉ, M. (2011a). Quel traitement unifié pour les suffixations en *-isme* et en *-iste*? In ROCHÉ, M., BOYÉ, G., HATHOUT, N., LIGNON, S. et PLÉNAT, M., éditeurs : *Des unités morphologiques au lexique.*, pages 69–143. Hermès-Lavoisier, Paris.
- ROCHÉ, M. (2011b). Quelle morphologie? In ROCHÉ, M., BOYÉ, G., HATHOUT, N., LIGNON, S. et PLÉNAT, M., éditeurs : *Des unités morphologiques au lexique.*, pages 15–38. Hermès-Lavoisier, Paris.
- ROCHÉ, M. et PLÉNAT, M. (2014). Le jeu des contraintes dans la sélection du thème présuffixal. In NEVEU, F., BLUMENTHAL, P., HRIBA, L., GERSTENBERG, A., MEINSCHAEFER, J. et PRÉVOST, S., éditeurs : *Actes du 4e Congrès Mondial de Linguistique Française*, pages 1863–1878, Paris. Institut de Linguistique Française.
- RONDEAU, G. (1981). *Introduction à la terminologie*. Centre éducatif et culturel, Montréal.
- SAGER, J. C., DUNGWORTH, D. et McDONALD, P. F. (1980). *English Special Languages : Principles and Practice in Science and Technology*. Brandstetter, Wiesbaden.
- SAHLGREN, M. (2008). The distributional hypothesis. *Italian Journal of Linguistics*, 20(1):33–54.
- SAULNIER, S. (2012). Quels problèmes posent les lexèmes du type *quadripède, tétraèdre* à la composition néoclassique? *Verbum*, XXXIV(2):159–177.
- SCALISE, S. (1986). *Generative Morphology*. Foris, Dordrecht.

- SCALISE, S. (1990). *Morfologia e lessico : una prospettiva generativista*. Il Mulino, Bologna.
- SCALISE, S., BISETTO, A. et GUEVARA, E. (2005). Selection in compounding and derivation. In DRESSLER, W. U., RAINER, F., KASTOVSKY, D. et PFEIFFER, O., éditeurs : *Morphology and its Demarcations*, pages 133–150. John Benjamins, Amsterdam / Philadelphia.
- SGROI, S. C. (2003). Per una ridefinizione di 'confisso' : composti confissati, derivati confissati, parasintetici confissati vs etimi ibridi e incongrui. *Quaderni di semantica : rivista internazionale di semantica teorica e applicata*, 24(1):81–154.
- SMYTH, H. W. (1920). *A Greek Grammar for Colleges*. American Book Company, New York.
- ŠTEKAUER, P. (2005). Compounding and affixation. Any difference? In DRESSLER, W. U., RAINER, F., KASTOVSKY, D. et PFEIFFER, O., éditeurs : *Morphology and its Demarcations*, pages 151–159. John Benjamins, Amsterdam / Philadelphia.
- STRNADOVÁ, J. (2014). *Les réseaux adjectivaux. Sur la grammaire des adjectifs dénominaux en français*. Thèse de doctorat, Université Paris Diderot (Paris 7) et Univerzita Karlova V Praze.
- THOMPSON, L. C. (1965). *A Vietnamese Grammar*. University of Washington Press, Washington.
- TOURNIER, J. (1985). *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Champion/Slatkine, Paris/Genève.
- VAN GOETHEM, K. et AMIOT, D. (2009). Affixation processes in Dutch and French. Chypre. Université de Nicosie.
- VILLOING, F. (2002). *Les mots composés [VN]_{A/N} du français : réflexions épistémologiques et propositions d'analyses*. Thèse de doctorat, Université de Paris X-Nanterre.
- VILLOING, F. (2009). Les mots composés VN. In FRADIN, B., KERLEROUX, F. c. et PLÉNAT, M., éditeurs : *Aperçus de morphologie du français*, pages 175–197. Presses Universitaires de Vincennes, Saint-Denis.
- VILLOING, F. (2012). French compounds. *Probus, International Journal of Latin and Romance Linguistics*, 24(1):29–60.

- VILLOING, F. et NAMER, F. (2006). *Saxifrage et casse-pierre* : quelles propriétés distinctives des mots composés VN et NV en français ? In KELLING, C., MEINSCHAEFER, J. et MUTZ, K., éditeurs : *XXIX Romanistentages in Saarbrücken*, pages 177–197. Universität Konstanz.
- VILLOING, F. et NAMER, F. (2012). Composition néoclassique en *-logue* et en *-logiste* : les noms en *-logue* sont-ils encore des noms de spécialistes ? *Verbum*, XXXIV(2):213–231.
- WARREN, B. (1990). The importance of combining forms. In DRESSLER, W. U., LUSCHÜTZKY, H. C., PFEIFFER, O. E. et RENNISON, J. R., éditeurs : *Contemporary Morphology*, pages 111–132. Mouton de Gruyter, Berlin/New York.
- WHITEHEAD, B. N. (2011). The alleged Greek influence on Latin compounding. In ONIGA, R., IOVINO, R. et GIUSTI, G., éditeurs : *Formal Linguistics and the Teaching of Latin : Theoretical and Applied Perspectives in Comparative Grammar*, pages 215–224. Cambridge Scholars Publishing, newcastle upon tyne édition.
- WHORF, B. et CARROLL, J. (1956). *Language, Thought, and Reality : Selected Writings of Benjamin Lee Whorf*. Language/Anthropology. Technology Press of Massachusetts Institute of Technology.
- WILMANN, W. (1899). *Deutsche Grammatik. Gotisch, Alt-, Mittel- und Neuhochdeutsche. Zweite Abteilung : Wortbildung*. Mouton de Gruyter, Berlin.

Résumé

Les éléments empruntés au grec ancien ou au latin entrant dans les formations traditionnellement appelées composés néoclassiques sont souvent, dans la littérature, regroupés dans une même catégorie d'éléments. Cette thèse se concentre sur les formations impliquant onze éléments finaux (en français), *-logie*, *-logue*, *-cratie*, *-crate*, *-phobie*, *-phobe*, *-phone*, *-phage*, *-vore*, *-cide* et *-cole*. Ces formations donnent lieu en français à de longues listes de lexèmes et semblent disponibles pour la création lexicale en dehors du vocabulaire savant. L'étude de ces éléments est effectuée dans le cadre de la Morphologie Constructionnelle (Booij 2010). Une base de données a été constituée pour analyser ces éléments : baptisée NeoClassy, elle regroupe les lexèmes formés au moyen de ces éléments finaux et récoltés dans des ressources dictionnaires et sur la Toile. Les analyses effectuées sur cette base de données ont mis au jour des contraintes morphophonologiques et des contraintes sémantiques pesant sur ces formations. Une analyse distributionnelle a par ailleurs montré un comportement similaire entre certaines constructions néoclassiques de NeoClassy et les dérivés. Les schémas impliquant des éléments néoclassiques finaux et les schémas dans lesquels des éléments natifs finaux sont spécifiés sont intégrés dans un même modèle, sur la base de critères phonologiques, sémantiques et lexicaux. Les éléments néoclassiques ne font ainsi pas partie d'une classe homogène ; certains sont considérés comme des thèmes supplétifs de lexèmes quand d'autres ont un comportement similaire à celui des affixes.

Abstract

Such elements as those borrowed from Ancient Greek or Latin, which enter into so-called neoclassical compounds, are often, in the literature, grouped together in a unique class of elements. This dissertation focuses on formations that involve eleven French final elements : *-logie*, *-logue*, *-cratie*, *-crate*, *-phobie*, *-phobe*, *-phone*, *-phage*, *-vore*, *-cide* and *-cole*. In French, those constructions generate large sets of lexemes and appear to be available for lexical creation aside from learned vocabulary. The study of these elements is conducted within the framework of Constructed Morphology (Booij 2010). A database, called NeoClassy, was built in order to analyse these elements. It gathers the lexemes, formed with these final elements, which were collected in dictionaries and on the Web. The analyses that were conducted on this database have brought morphophonological, semantic and lexical constraints to light. A distributional analysis has also shown that some neoclassical constructions from NeoClassy and derivatives have a similar behaviour. The schemas that involve final neoclassical elements and those in which native final elements are specified are integrated in a same model, on the basis of phonological, semantic and lexical criteria. Neoclassical elements do not constitute a homogeneous class : some are considered as suppletive stems of lexemes when others have a similar behaviour to that of affixes.
