

SAXIFAGE ET CASSE-PIERRE : QUELLES PROPRIÉTÉS DISTINCTIVES DES MOTS COMPOSÉS NV ET VN EN FRANÇAIS ?

Fiammetta Namer
Université Nancy 2 &
UMR « ATIF »
fiammetta.namer@univ-nancy2.fr

Florence Villoing
Université Paris 8 &
UMR 7023
Florence.Villoing@u-paris10.fr

1 Introduction

Données

La morphologie lexicale du français construit, par le procédé de la composition, des noms ou des adjectifs. Ceux-ci résultent de formations variées, faisant intervenir deux noms (1a), deux adjectifs (1b), un verbe et un nom (1d).

- (1) a [homme_N-grenouille_N]_N
[voiture_N-balai_N]_N
b [gris_A-bleu_A]_A
c [porte_V-plume_N]_N
[gratte_V-papier_N]_N
[perce_V-oreille_N]_N
(porte) [coupe_V-feu_N]_A

Parallèlement, le lexique du français comprend des unités lexicales tels que (2), dont on reconnaît qu'elles sont souvent composées de deux Eléments de Formation d'origine grecque ou latine¹, identifiables essentiellement comme deux noms (2a), un adjectif et un nom (2b), deux adjectifs (2c) ou encore un nom et un verbe (2d)².

- (2) a [anthropo_Nmorphe_N]_A
b [érythro_Acyte_N]_N
c [afro_A-cubain_A]_A
d [anthropo_Nphage_V]_A
[mélono_Nmane_V]_N
[caverni_Ncole_V]_{A/N}
[géné_Nocide_V]_{A/N}

Nous proposons, dans cette étude, d'examiner les composés construits selon les modèles illustrés en (1c) et (2d), pour tenter d'en proposer une analyse contrastive.

Notre étude repose sur l'hypothèse que le mode de formation des mots en (2) fait aujourd'hui partie du système morphologique du français, et ce malgré le fait que certains soient empruntés aux langues anciennes (grec et latin) et, le cas échéant, communs à plusieurs langues contemporaines (principalement les langues romanes, germaniques et slaves cf. Darmesteter (1877)). Les néologismes qui apparaissent quotidiennement dans les médias témoignent de cette productivité (par exemple *bibliophobe*, *tabacolâtre*, *Jocondovore*). En outre, un certain nombre d'arguments

¹ Nous empruntons le terme 'Élément de formation' notamment à Fradin (2000) et Iacobini (1999). Dans les exemples, les Eléments de Formation sont précédés de '°' ; voir à ce sujet le paragraphe 2.3.

² Un tableau, au paragraphe 6.1, donne une formulation de la définition de chaque composé néoclassique mentionné, en fonction du sens des éléments de formation qui le constituent.

viennent étayer l'idée que les composés néoclassiques ne relèvent pas des règles de construction appartenant aux langues classiques fournissant les Eléments de Formation :

- (i) l'association d'Eléments de Formation d'origine diverse et non systématiquement issus de la même langue, latine ou grecque (par exemple, combinaison d'Eléments de Formation Français+Latin : *insecticide* ; Grec+Français : *macromolécule* ; Français+Grec : *bureaucrate* ; Latin+Grec *planisphère*, ou d'unités dont l'une correspond à la troncation d'un mot du français : *cinéphile*, *anglophone*, *alcootest*) ;
- (ii) l'apparition d'une voyelle thématique impropre entre les deux composants : par exemple, la voyelle thématique « o », spécifique au grec, est également utilisée à la jonction de composants d'origine française (*franco-allemand*, *anglo-saxon*) ou d'origine latine, à la place du « i » attendu (*cérébro-spinal*, *génito-urinaire*) (Darmesteter (1894 : 256)) ;
- (iii) les entorses aux règles de la composition grecque ou latine que l'on observe dans la formation de multiples mots composés « néoclassiques ». Cf. par exemple,
 - a. le mot *microbe* dont Benveniste (1974 : 163-170) analyse la genèse pour montrer en quoi son sens construit (« petite vie » c'est à dire « organisme microscopique ») n'est pas accessible à partir d'un modèle de composition grec (duquel on construirait l'interprétation « vie courte », scientifiquement incohérente).
 - b. les mots *oxygène* et *hydrogène* analysés par Darmesteter (1894 : 253) pour montrer leur « mauvaise » formation sémantique puisqu'en grec, l'élément correspondant à *-gène* ne signifie pas « qui produit », comme ici («oxygène : qui produit de l'acide », « hydrogène : qui produit de l'eau »), mais « qui est produit ».

Plus précisément, on reconnaît derrière les mots en (2d) une composition Nom+Verbe similaire à la composition Verbe + Nom en français³, à la différence apparente près que les composants sont issus du lexique grec ou latin, et que l'ordre des composants est inversé (composant régi (Y)-composant recteur (X), tandis que la composition ordinaire adopte l'ordre XY). Notre recherche s'inscrit ainsi dans la lignée des analyses traditionnelles en morphologie qui reconnaissent dans ce cas une composition dite « néoclassique » par contraste avec une composition ordinaire, sur bases françaises (cette distinction traditionnelle a déjà cours au 19^{ème} siècle, cf. par exemple Hatzfeld et al. (1890)).

Problématique générale

En partant de l'hypothèse que les unités lexicales en (2d) relèvent d'une règle de composition NV, l'objet de notre recherche est de déterminer dans quelle mesure les mots composés VN et NV sont construits, en français, par deux règles de composition distinctes ou, au contraire, s'ils relèvent d'une seule et même règle. En d'autres termes, les compositions VN et NV correspondent-elles à

³ La composition Verbe+Nom existe dans les autres langues romanes ; voir à ce sujet Rainer & Varela (1992), Scalise (1992), Villalava (1992)

- un procédé unique de construction morphologique du français aux propriétés identiques hormis l'origine des composants et, par voie de conséquence, l'ordre de ses composants⁴ ;
- ou à deux procédés de formation différents du français qui ne construisent pas le même type de mot, et ne répondent pas aux mêmes contraintes sémantiques et catégorielles.

Méthodologie/plan

Pour répondre à ces questions, nous nous sommes attachées à mettre au jour les différences et les ressemblances entre les deux compositions VN et NV, en particulier en identifiant puis en comparant les contraintes catégorielles et sémantiques qui portent sur la formation de ces composés.

L'étude est menée à partir d'un corpus dont les principes d'élaboration sont présentés au paragraphe 2, qui, en outre, explicite les difficultés théoriques que pose l'identification des Eléments de Formation composant les supposés NV du type (2d). Le paragraphe 3, quant à lui, est voué à la comparaison des propriétés des mots composés VN et NV, comparaison qui aboutira finalement, au paragraphe 4, à montrer que les différences observées entre les composés VN et NV nous conduisent plus probablement à reconnaître deux règles de composition en français qu'une seule.

2 Les données

2.1 Constitution du corpus VN

Le corpus de mots composés VN comprend environ 2000 à 3000 unités. Il est principalement constitué de données lexicographiques issues de dictionnaires de langue ou de dictionnaires encyclopédiques. Les dictionnaires ont été choisis de sorte à être suffisamment représentatifs du lexique général et technique du français (*Trésor de la Langue Française*, *Grand Robert de la Langue Française*, *Dictionnaire de la langue française* (Littré), *Dictionnaire Général de la Langue française* (Darmesteter et Hartzfeld)).

D'autres sources sont venues compléter ces données lexicographiques : des corpus étiquetés permettant l'extraction de mots composés VN, des données issues du Web et enfin, la collecte fortuite de mots composés VN néologiques dans des journaux, des magazines et des catalogues.

2.2 Constitution du corpus NV

Le corpus de mots composés NV appartient au lexique du français sans être restreint à un domaine de spécialité, restriction qui constitue un risque fréquemment encouru à l'occasion d'études de procédés de construction utilisant des Eléments de Formation (puisque la construction du lexique scientifique exploite maximale ces Eléments de Formation).

Les sources sont essentiellement le *Trésor de la Langue Française informatisé* et le *Dictionnaire des structures du vocabulaire savant* de Henri Cottez.

Le 1^{er} état du corpus est constitué à partir de l'intuition sémantique que nous avons de l'étiquetage catégoriel 'V' du second composant. Par la suite, ce corpus initial a été

⁴ Comme l'observent de nombreux auteurs (voir par exemple Iacobini (2004), ou Corbin (2005)), l'ordre des composants en composition « néoclassique » est emprunté au grec ancien et au latin.

amendé en fonction d'une analyse plus raisonnée des composants des NV. C'est cette analyse que nous présentons en 2.3.

Au final, nous avons collecté environ 500 adjectifs et noms composés de type NV qui répondent aux critères catégoriels énoncés ci-dessous.

2.3 Identification des constituants dans le NV

La constitution du corpus des mots composés NV a posé d'importants problèmes théoriques liés à l'identification de ses composants. Plusieurs questions se sont posées :

1) comment reconnaître, dans les mots construits à partir de bases grecques ou latines, un mot composé. En d'autres termes, comment reconnaître que les Eléments de Formation qui constituent ces mots construits sont bien des lexèmes-bases et non des affixes⁵.

2) corollairement, comment reconnaître la catégorie de ces lexèmes-base : quels critères permettent d'identifier un V et un N dans ces unités dont la propriété est de ne jamais apparaître de façon autonome dans une phrase.

2.3.1 Cadre théorique

Ces questions prennent sens dans le cadre théorique de la morphologie lexématique (Anderson (1992), Aronoff & Fudeman (2004), Booij (2005), Fradin (2003)) au sein duquel les unités de base des opérations de construction de lexème (cf. Matthews (1991)) appartiennent à une des catégories N, V, Adj et se distinguent clairement des affixes qui ne sont que les exposants phonologiques de règles (cette approche n'associe donc pas aux affixes une catégorie lexicale majeure, à la différence d'autres propositions cf Williams (1981), Lieber (1981), Lüdeling et al. (2002)). La position que nous adoptons ici consiste à étendre les principes de la morphologie lexématique aux bases non autonomes, c'est-à-dire à envisager que les Eléments de Formation néoclassiques puissent être des lexèmes catégorisés N, V ou Adj au même titre que les bases autonomes. De ce point de vue, nous nous détachons des approches qui reconnaissent l'existence d'Eléments, sans les catégoriser (cf. Rey-Debove (2004), et Cottez (1988 : VII-XX), qui identifie des « formants du vocabulaire savant » équivalents à des « signes », c'est-à-dire des morphèmes, sans leur assigner le statut de lexème-base ou d'affixe :

Les signes dont nous avons dressé un inventaire méthodique sont donc des formants de mots de construction savante : nous nous servons de ce terme très général de formants (à la fois éléments et modèles de formation) parce qu'il est le plus neutre, exprimant le fait que 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. (Cottez (1988 : VIII))

En revanche, nous suivons la ligne adoptée notamment par Corbin (1985), Warren (1990), Fradin (2000 ; 2003) et Iacobini (2004) qui distinguent parmi les formants du vocabulaire savant, les lexèmes-bases d'un côté et les affixes de l'autre.

L'élaboration de notre corpus repose sur le fait d'avoir reconnu des lexèmes-bases dans les Eléments de Formation des mots du type *anthropophage*.

⁵ Nous employons le terme **lexème-base** plutôt que lexème pour le distinguer de l'acception du terme lexème conçu comme unité du lexique incluant déterminant, pronom, préposition (cf. Aronoff & Fudeman (2004)), autant de catégories qui ne servent pas de base à la morphologie constructionnelle.

2.3.2 Lexème-base ou affixe ?

Les composants des mots construits en (2) sont-ils des lexèmes-base (c'est-à-dire des Noms, des Verbes ou des Adjectifs) ou des affixes ?

La question est pertinente dans la mesure où les Eléments de Formation composant les mots du type *anthropophage* présentent des propriétés qui les rapprochent aussi bien des affixes que des lexèmes-bases. Des affixes, ils partagent la non-autonomie syntaxique ; en effet, ces Eléments de Formation occupent une position dans la structure des unités lexicales mais pas dans celle d'un énoncé. Des lexèmes-bases, ils ont en commun une propriété sémantico-référentielle et une propriété catégorielle.

Nous présentons ci-dessous un certain nombre de critères que nous avons exploités dans le but de reconnaître deux lexème-bases N et V dans les mots construits en (2d).

1) Critère sémantique

Le critère sémantique s'appuie sur la dichotomie fréquemment mise en avant dans la littérature entre « sens lexical » et « sens grammatical » (cf. Corbin (2001) pour une application à la morphologie du français, et Kleiber (1999) à la sémantique lexicale).

Un lexème-base est une unité lexicale abstraite, hors énoncé, qui présente, au minimum, trois propriétés : une forme phonologique, une catégorie syntaxique, un sens lexical. Les Eléments de Formation, à la différence des affixes qui ne réfèrent pas mais auxquels une instruction sémantique est associée, et à l'image des lexèmes-bases, renvoient de façon stable à une catégorie référentielle. Aussi peut-on assigner le statut de lexème-base à un Elément de Formation dont on distingue, au travers d'un ensemble de mots construits, un sens de type référentiel.

Néanmoins, l'utilisation de ce critère s'avère délicat,

- soit parce que le mode d'identification de la valeur référentielle d'un composant est fondé sur sa traduction, laquelle traduction manque de fiabilité. On observe, en effet, une grande diversité entre les dictionnaires (et, parfois même en leur sein), en particulier lorsqu'un Elément est susceptible d'être traduit par le verbe ou le nom déverbal correspondant ; par exemple, l'Elément de Formation *°graph(o)* est défini, dans le Robert Brio, par les deux termes 'écrire' et 'écriture' ;
- soit parce que l'identification de la valeur référentielle s'appuie sur le sens construit du composé et la relation sémantique entre les différents éléments qui le constitue, et que les informations sémantiques fournies par les dictionnaires restent insuffisantes (par manque de précision ou par inclinaison à décrire plutôt le référent que le sens). Par exemple, la définition lexicographique des mots *pleuronecte* (« Genre de Poissons au corps plat et dissymétrique comprenant les genres sole, turbot, plie, flétan... ») et *bathyergue* (« Mammifère rongeur d'Afrique australe, ayant des mœurs ressemblant à ceux de la taupe ») ne permet pas d'avoir accès au sens construit, parce qu'elle s'attache à décrire uniquement le référent.

En outre, ce critère du sens est difficile à manipuler lorsque les lexèmes-bases sont en cours de grammaticalisation et tendent à se comporter comme des affixes ; dans ce cas, leur valeur sémantique oscille entre les deux pôles référentiel et instructionnel (cf. Amiot&Dal (2005) à propos de *°-logue*).

2) Critère de la catégorie.

Ce critère s'appuie sur la propriété qu'a un lexème-base d'appartenir à l'une des catégories nom, verbe et adjectif (et, dans une certaine mesure, adverbe). Aussi, reconnaître l'appartenance d'un Élément de Formation à l'une de ces catégories suffit-il à garantir son identité de lexème-base. Cependant, les Éléments de Formation n'apparaissant jamais de façon autonome dans une phrase, les moyens d'identification traditionnels (distribution, comportement morphologique et syntaxique) s'avèrent totalement inopérants. Pour répondre à cette spécificité, le recours à d'autres moyens s'avère nécessaire. Nous en présentons quelques-uns ci-dessous.

- **Moyen 1** : Un Élément de Formation est catégorisable et interprétable en référence à sa traduction ou à son interprétation en français. Il obtiendra ainsi la même catégorie que le lexème-base par lequel il est traduit. Mais cette solution présente des limites liées au caractère incertain de la traduction (cf. ci-dessus).
- **Moyen 2** : Un Élément de Formation allomorphe d'un lexème-base est théoriquement catégorisable sans difficulté. Mais la notion d'allomorphie recouvre parfois des phénomènes de supplétion et les relations entre le lexème-base et son hypothétique allomorphe manquent parfois de transparence, de sorte que le problème reste parfois entier. Nous ne traiterons pas de cette problématique ici. (cf. Haspelmath (2002 :26-28) pour une description synthétique du problème, et Boyé (à paraître), pour une présentation de la notion de supplétion).
- **Moyen 3** : Un Élément de Formation utilisé comme base d'une affixation qui présente des contraintes de sélection est aussi catégorisable (Corbin (1985 : 62-64 ; 1987 : 182-sq)). Mais beaucoup d'affixes restent peu contraints catégoriellement et ne permettent pas de déterminer avec certitude la catégorie de la base.
- **Moyen 4** : Un Élément de Formation est aussi catégorisable en fonction du type de construction dans laquelle elle apparaît. En particulier, la relation sémantique que le premier Élément de Formation entretient avec le second permet de mettre au jour sa catégorie. Par exemple, si on identifie dans un composé YX une relation de type argument-verbe entre Y et X, on est en droit de supposer que X est un verbe et Y un nom.

L'exploitation combinée des critères présentés ci-dessus nous a permis de reconnaître dans les construits de type *anthropophage* présentés en (2d) des Éléments de Formation de catégorie verbe (élément de droite : X) et nom (éléments de gauche : Y). Les exemples (3) et (4) donnent un aperçu de l'application de ces moyens respectivement aux adjectifs construits *oculogyre* « qui commande la rotation du globe oculaire » et *anthropophage* « mangeur de chair humaine ».

(3) oculogyre_A
(3a) °*ocul(o)*

L'emploi de °*ocul* comme base dans l'adjectif *oculaire* permet de déterminer, *via* le Moyen 3, que cet EF est un nom. En effet, le suffixe *-aire* forme des adjectifs

exclusivement dénominaux : $pôle_N \rightarrow polaire_A$. Comme $oculaire_A$ est construit au moyen de ce procédé à partir de la base $^\circ ocul$, celle-ci est forcément nominale.

(3b) $^\circ gyre$

De la même manière, on va se servir des contraintes catégorielles qu'imposent les suffixes *-oire* et *-ion* pour identifier la catégorie de $^\circ gyre$, par le biais du Moyen 3. Le français dispose de l'adjectif *giratoire* (également orthographié *gyratoire*), dont le sens est : 'qui tourne autour d'un axe' et du nom *giration* (ou *gyration*) : 'action de tourner'. Faisons l'hypothèse que la base de ces lexèmes, construits par suffixation, respectivement, en *-oire* et en *-ion*, est la variante allomorphique $^\circ gir(at)$ de $^\circ gyre$. Chacun de ces suffixes sélectionne une base verbale de façon privilégiée ($meritoire_A$: $meriter_V$, $agression_N$: $agresser_V$), bien que cette sélection soit non exclusive ($republicatoire_A$: $republique_N$, $baladoire_A$: $balade_N$, $zonation_N$: $zone_N$). Cependant, quand la même base est employée avec les deux affixes, elle est exclusivement verbale ($revendicatoire_A$, $revendication_N$: $revendiquer_V$). C'est ce que l'on observe avec $^\circ girat$ ($giration_N$, $giratoire_A$), qui est donc une base verbale. De plus, les bases des noms en *-ion* ou des adjectifs en *-oire* se terminant par *-at*, sont identifiées comme la forme supin de verbes latins : $^\circ girat$, issu d'une forme verbale latine, est bien un allomorphe de $^\circ gyre$. $^\circ Gyre$ est bien un verbe.

Ainsi, le Moyen 3 a permis de reconnaître dans $oculogyre_A$, respectivement, un constituant nominal ($^\circ ocul$) et un constituant verbal ($^\circ gyre$) non autonomes. L'utilisation de Moyen 3 conduit ainsi à analyser $oculogyre$ comme un adjectif composé NV.

(4) Anthropophage_A

(4a) Moyen 1

Le moyen 1 permet de supposer que $^\circ anthrop$ est un nom et que $^\circ phage$ un verbe.

(4b) Ceci est confirmé par le Moyen 4.

$^\circ phage$ doit être analysé comme un verbe parce que sa relation sémantique à $^\circ anthrop$ est de type prédicat-argument : $^\circ phage$ dénote un procès qui s'applique sur son argument proto-patient, réalisé par $^\circ anthrop$.

Finalement, l'association des Moyens 1 et 4 conduit à analyser $anthropophage$ comme un mot composé NV.

3 Analyse

Une fois le corpus constitué, le cœur de l'analyse a consisté en la comparaison des mots composés « néoclassiques » NV et des composés VN sur base française. Les trois critères (c1) à (c3) ont servi cette comparaison :

(c1) les propriétés catégorielles et sémantiques des composés VN et NV.

(c2) le type de procès que dénote V

(c3) la relation sémantique entre V et N

Le croisement systématique de ces trois critères a eu pour objectif de répondre aux questions suivantes : les propriétés catégorielles et sémantiques des VN et des NV sont-elles identiques, partiellement juxtaposées ou en distribution complémentaire ? Corrélativement, existe-t-il des combinaisons non réalisées dans les critères (c1) à (c3) ?

3.1 Propriétés des mots composés VN ⁶

3.1.1 Critère 1 : catégorie et type sémantique des mots composés VN

• **Catégorie** : la composition VN construit principalement des noms et rarement des adjectifs ; ces derniers sont illustrés en (5) :

- (5) (papier) tue-mouche_A
(porte) coupe-feu_A

• **Type sémantique du composé.**

La plupart des mots composés VN nominaux dénotent des **artefacts**

- (6) tournevis_N
ouvre-boîte_N
porte-plume_N

Les autres réfèrent à des entités **animées** (humains ou animaux)

– soit par leur **fonction** ou leur occupation habituelle

- (7) garde-barrière_N
gratte-papier_N
aide-ménagère_N

– soit par une propriété ou un **comportement caractéristique**

- (8) trouble-fête_N
rabat-joie_{A/N}
pique-bœufs_N (oiseau)
perce-oreille_N (insecte)

Enfin, certains d'entre eux dénotent des **lieux**

- (9) coupe-gorge_{A/N}
garde-meuble_N,
pince-fesse_N

ou des **événements**

- (10) lèche-vitrine_N,
(jouer à) saute-mouton_N,
baise-main_N,
(faire du) remue-ménage_N

3.1.2 Critère 2 : type de procès du V

Le type de procès qu'exprime le prédicat verbal d'un mot composé VN est fortement contraint : selon la terminologie de Vendler (1967) et de ses successeurs (notamment Dowty (1979)), il est **dynamique**, c'est-à-dire compatible avec les contextes linguistiques « être en train de », « se mettre à », « s'arrêter de » (cf. Villoing, (2003) pour plus de détails) (cf. 11), et ne peut que très difficilement être statif (cf. 12) :

- (11) essuie-glace_N
Jean est en train d'essuyer la glace/ s'est mis à essuyer la glace/s'arrête d'essuyer la glace.

⁶ voir entre autres à ce sujet Villoing (2003) et Fradin (2005)

- (12) *connaît-allemand
*Jean est en train de connaître l'allemand/s'est mis à connaître l'allemand/s'arrête de connaître l'allemand.

3.1.3 Critère 3 : relation sémantique entre le V et le N

La relation sémantique entre le verbe et ses participants est également contrainte. Dans la terminologie de Dowty (1991) qui réfute les étiquettes des rôles thématiques et propose un continuum Proto-Agent/Proto-Patient, le N d'un composé VN, en particulier, est typiquement du côté du **Proto-Patient** ; il répond au critère du « changement d'état » et de « l'affectitude » (cf. aussi Foley & Van Valin (1984), Jackendoff (1990) pour la définition des critères).

Répondent également à cette analyse les quelques cas comme (13) qui semblent construits sur un verbe intransitif et pour lesquels le N semble se situer du côté du Proto-Agent :

- (13) trotte-bébé_N / 'le bébé trotte'
pense-bête_N / 'le bête pense'

Cependant, le sens du prédicat verbal est causatif et N est un argument du verbe qui se situe du côté du Proto-Patient :

- (14) trotte-bébé_N / 'fait trotter le bébé'
pense-bête_N / 'fait penser le bête'

On observe en outre, et de manière plus rare, d'autres relations sémantiques entre le V et le N. C'est ainsi que dans certains mots composés VN tels que (15) le N ne répond pas aux critères d'un Proto-patient mais plutôt à ceux d'un Proto-agent :

- (15) cuit-vapeur_N (type de casserole dont la cuisson se fait par la vapeur)
pousse-pied_N (bateau léger que l'on pousse avec le pied)
gobe-mouton_N (pilule empoisonnée qui fait mourir les animaux –que le mouton gobe)

Par ailleurs, on relève un certain nombre de mots composés VN tels que (16) où la relation entre le V et le N ne répond à aucun des critères d'un Proto-patient ou d'un Proto-agent, mais plutôt à ceux d'un modifieur temporel (16a) ou locatif (16b) :

- (16) a réveille-matin_N
grille-midi_N (plante)
b croque-télé_N

3.2 Propriétés des mots composés NV

3.2.1 Critère 1 : catégorie et type sémantique des mots composés NV

• Catégorie

La plupart du temps, la composition NV produit des adjectifs qui s'emploient indifféremment comme noms (Kerleroux (1991, 1996), Fradin (1997)).

- (17) ventriloque_{A/N}

• Type sémantique du composé.

Dans les cas où seul le type catégoriel nominal est réalisé, les mots composés NV réfèrent

- le plus fréquemment à un **objet concret**
- (18) odontoclaste_N
saxifrage_N
- à une **entité animée**
- (19) biographe_N
notonecte_N
- ou à un **événement**
- (20) lipolyse_N
infanticide_N

3.2.2 Critère 2 : type de procès du V

Le prédicat verbal décrit un procès

- **dynamique**
- (21) °ambule_V = ‘marcher’ (noctambule_{A/N})
- ou **statif**.

Parmi les **procès statifs**, on reconnaît des prédicats verbaux dénotant

- majoritairement des **relations spatiales** (Talmy (2000), Vandeloise (1986))
- (22) °fère_V = ‘contenir’ (carbonifère_A)
°phore_V = ‘porter’ (mélanophore_N)
- ou des **sentiments** (Levin (1993))
- (23) °phile_V = ‘apprécier’ (russophile_A)
°mane_V = ‘aimer’ (mélomane_A)
°lâtre_V = ‘adorer’ (wagnerolâtre_A)
°phobe_V = ‘craindre, détester’ (anglophobe_A)
- On reconnaît aussi des verbes de **perception** (Levin & Rappaport (1995))
- (24) °op_V = ‘voir’ (nyctalope_A)
- ou des verbes d’**existence** (Levin & Rappaport (1995))
- (25) °cole_V = ‘vivre’ (limicole_A)
- (26) °bie_V = ‘vivre’ (phyllobie_N)
- Et un verbe d’**émission de lumière** (Levin & Rappaport (1995))
- (27) °luque_V = ‘briller’ (noctiluque_A)

3.2.3 Critère 3 : relation sémantique entre le V et le N

Les relations sémantiques entre le V et le N dans les mots composés NV relèvent de deux dimensions différentes, selon le type de procès qu’exprime le verbe.

• Dans le cas des prédicats verbaux **dynamiques**, les critères de Dowty (1991) sont pertinents, au même titre que dans le cas des mots composés VN.

- N répond soit aux critères d’un **Proto-agent**
- (28) °psych(o)_N = ‘esprit’ dans psychogène_A
- soit à ceux d’un **Proto-patient**

- (29) °lip(o)_N = 'graisse' dans lipolyse_N
– soit ne répond à **aucun des critères** prototypiques d'un agent ou d'un patient
- (30) °ventr(i)_N = 'ventre' dans ventriloque_A,
- (31) °héli(o)_N = 'soleil' dans héliotrope_A
- (32) °noct_N = 'nuit' dans noctambule_A

• Dans le cas des prédicats verbaux **statifs**, les critères d'assignement des rôles de Dowty n'apparaissent pas pertinents. La relation entre le N et le V peut, en revanche, être exprimée au moyen des notions de **Figure** et **Ground** (Talmy (2000)⁷). N désigne alors

- principalement la **Figure**
- (33) °mélan(o)_N = 'mélanine' dans mélanophore_N
°carbon(i)_N = 'charbon' dans carbonifère_A
- ou, parfois, le **Ground**
- (34) °vas(i)_N = 'vaisseau sanguin' dans vasiducte_N

• Les prédicats verbaux exprimant un **sentiment** mettent en jeu une relation entre deux participants, l'**expérienteur** et le **stimulus**. Nous proposons d'apparenter ces verbes psychologiques à des relations spatiales et ainsi d'interpréter l'**expérienteur** comme le **Ground** (le site) et le **stimulus** comme la **Figure** (cible).

Ces prédicats verbaux ne se rencontrent, à notre connaissance, qu'avec des **N-Figure** (cible) :

- (35) *mél(o)* dans mélomane_A
angl(o) dans anglophobe_A

4 Résultats

Le croisement des critères 1, 2 et 3 est représenté dans le **Tableau 1** au paragraphe 6.1; il met en évidence les contrastes suivants entre VN et NV.

• Les mots composés NV couvrent un spectre beaucoup plus large que les mots composés VN du français, quel que soit le critère examiné :

- (i) Catégoriel : leur appartenance catégorielle correspond quasiment indifféremment à un adjectif ou un nom (tandis que les mots composés VN sont presque exclusivement des noms) ;
- (ii) Type de procès du prédicat verbal : de très nombreux prédicats verbaux de composés NV désignent des procès statifs de relation spatiale, à la différence des mots composés VN ;
- (iii) La relation sémantique NV : alors que le N se situe typiquement du côté du Proto-patient dans les mots composés VN, il remplit davantage de rôles sémantiques dans les composés NV dont le prédicat verbal exprime un procès dynamique ; de même, il peut jouer le rôle de Figure ou de ground lorsque le prédicat verbal est statif.

⁷ Talmy (2000) étudie les relations de type non agentifs entre des participants dans un **événement dit « spatial »** ou de **localisation**. Il reconnaît que ces événements mettent prioritairement en jeu les rôles de la Figure (ou Cible) et du Ground (ou Site), et accessoirement celles de Manner, Cause, Trajet (Path). La **Figure** (Cible) est reconnaissable comme étant l'objet en mouvement ou localisé, envisagé par rapport à un autre objet (objet de référence = **Ground** ou site).

En d'autres termes les composés NV instancient une très grande variété de combinaisons (c1-3) alors que les composés VN correspondent majoritairement aux combinaisons où V est dynamique et N est typiquement patient.

5 Conclusion

Même si, dans certains cas (lorsque le verbe est dynamique), les mots composés VN et NV semblent partager des contraintes catégorielles et sémantiques similaires (ce qui conduirait à n'envisager qu'une unique règle de composition sélectionnant des bases d'origine diverse), ils se distinguent nettement les uns des autres dans d'autres cas (lorsque le verbe est statif). En effet, seule la composition NV peut mettre en jeu un verbe statif et, en conséquence, construire une relation V-N du type spatiale. 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. Cf. : « un mélomane » = « un fou de musique ».

Les résultats de cette étude conduisent plutôt à penser que deux règles de composition sont en jeu en français. En conséquence, deux questions majeures émergent :

- Qu'est ce qui explique que la composition VN est si contrainte ?
- Quelle est l'origine des règles de formation de NV ? Peut-on les considérer comme des règles du français ou des règles empruntées aux langues dont les composants sont originaires (latin, grec) ?

Cette dernière question s'inscrit dans une problématique plus large qui cherche à déterminer si la composition néoclassique exploite ou non des règles héritées des langues dont les composants sont originaires. La doxa étant plutôt favorable à l'hypothèse de l'héritage (cf., en faveur de cette idée, Darmesteter (1894) et contre Benveniste (1974)), nous avons cherché à évaluer sa validité dans le cas spécifique des mots composés NV dont les composants sont d'origine grecque. L'hypothèse sous-jacente à cette recherche étant que, si la construction de noms ou d'adjectifs composés NV était reconnue en grec ancien, alors nous disposerions d'arguments en faveur de l'emprunt et, de fait, pourrions avancer une explication historique à la formation contemporaine de ce type de composés. Or, les divers travaux examinés portant sur la composition en grec ancien (Chantraine (1930), Smyth (1920), et les travaux récents de A. Ralli⁸) ne convergent pas : en effet, contrairement à Chantraine (1930) qui reconnaît en grec une composition NV, Smyth (1920) et A. Ralli considèrent, quant à eux, pour des raisons différentes, que cette formation est inexistante en grec ancien. Ces divergences d'analyse montrent la fragilité de l'hypothèse d'une formation des composés néoclassiques par emprunt des règles aux langues classiques.

Pour autant, le grec ou le latin n'a-t-il eu aucune influence sur la construction de ces composés ? Les travaux qui se sont intéressés aux formations néoclassiques en français, qu'ils soient favorables ou défavorables à l'hypothèse de l'héritage des règles latines ou grecques, tendent finalement tous vers une solution hybride. En témoignent par exemple, les analyses qu'A. Darmesteter (1894 : 252-256) donne des composés néoclassiques sur des bases d'origine grecque. Darmesteter, dans un paragraphe qui déplore l'entrée massive en français, au 19^{ème} siècle, de nouveaux composés

⁸ communication personnelle

néoclassiques (souvent irrespectueux des règles des langues d'origine⁹), montre comment ces composés résultent d'une réanalyse (parfois incorrecte) du grec ancien en français. Ainsi, la création d'un composé néoclassique (e.g. *adénographie*) passerait par l'invention de l'équivalent qu'il aurait en grec (e.g. $\alpha\delta\epsilon\nu\omicron\gamma\rho\alpha\phi\iota\alpha$ = *adénographia*), équivalent construit à partir de composants grecs attestés (y compris des verbes) ($\alpha\delta\eta\iota$ _N = *aden-*, $\gamma\rho\alpha\phi\omega\upsilon$ _V = *-grapho*), et en conformité, réelle ou supposée, avec les règles du grec ancien. L'impact du grec (et du latin qui a servi de relais) est donc certain, mais le rôle du système linguistique français est également prépondérant. On le voit, c'est par des locuteurs du français, ayant une certaine connaissance des langues anciennes, que se sont élaborées les règles de construction des composés néoclassiques.

Cette idée d'une réanalyse du grec en français est partagée par différents linguistes tout au long du 20^{ème} siècle. On l'a retrouvée, par exemple chez Benveniste (1974 :170) :

« Certains des néologismes scientifiques de forme gréco-latine créés en français, et tout particulièrement les composés [...] n'ont de grec ou de latin que la forme matérielle. Ce sont en réalité, des composés bâtis en français et seulement transposés –souvent d'une manière assez lâche- en lexèmes gréco-latins. ».

puis chez Bouffartigue et Delrieu (1996).

« Les fabricants de mots grecs ont soin de se rapprocher le plus possible de l'hypothétique mot grec qu'il aurait pu être. »

Pour aboutir notre recherche, il resterait encore à déterminer quelle a été l'influence du modèle de la composition VN, déjà très productif à la fin du 18^{ème} siècle, sur la formation ou la réanalyse des composés néoclassiques NV en français.

6 Annexes

6.1 Références

- Amiot, D. & Dal, G.**, 2005, « Integrating Combining Forms into a Lexeme-based Morphology », in *Fifth Mediterranean Morphology Meeting*, 15-18 sept, Fréjus
- Anderson, S.R.**, 1992, *A-morphous Morphology*, Cambridge: Cambridge University Press.
- Aronoff, M., & Fudeman, K.**, 2005, *What is morphology?*, Blackwell Publishing, Oxford.
- Benveniste, E.**, 1974, "Formes nouvelles de la composition nominale", *Problèmes de linguistique générale II*, Paris, Gallimard, p. 163-176.
- Booij, G.**, 2005, *The Grammar of Words: An Introduction to Linguistic Morphology*. Oxford, Oxford University Press.
- Bouffartigue, J. et Delrieu A.-M.**, 1996, *Les racines grecques*. Paris, Belin.
- Boyé, G.**, (à paraître), "Suppletion", *Encyclopedia of Language and Linguistics, 2nd Edition*, K. Brown, London/Amsterdam, Elsevier.
- Chantraine, P.**, 1933 (1979), *La formation des noms en grec ancien*. Paris, Champion (Klincksieck).

⁹ « Lavoisier était meilleur chimiste qu'helléniste. La nomenclature scientifique offre encore plus d'un de ces composés que désavoueraient les Grecs. » (Darmesteter 1894 : 253)

- Corbin, D.**, 1985, „Les bases non-autonomes en français ou comment intégrer l'exception dans le modèle lexical“, *Langue française* n°66, Paris, Larousse, pp. 54-76.
- Corbin, D.**, 1987, *Morphologie dérivationnelle et structuration du lexique*. Lille, Presses Universitaires de Lille.
- Corbin, D.**, 2001, "Préfixes et suffixes: du sens aux catégories", *Faits de Langue*, Paris, Ophrys, **15**, pp. 41-69.
- Corbin, D.**, 2005, "French (Indo European: Romance) " in Booij, Lehmann & Mugdan (eds) *Morphology An International Handbook on Inflection and Word Formation*, Berlin/New York, Walter de Gruyter, vol1, art. 121.
- Cottez, H.**, 1988, *Dictionnaire des Structures du vocabulaire savant. Eléments et modèles de Formation*, 4ème édition. Paris, Dictionnaires Le Robert.
- Darmesteter, A.** 1894. *Traité de la Formation des mots composés dans la langue française*, (1ère édition: 1875). Paris: Honoré Champion.
- Darmesteter, A.**, 1877, *De la création des mots nouveaux dans la langue française et des lois qui la régissent*. Paris, Wieveg.
- Dowty, D.**, 1979, *Word Meaning and Montague Grammar*. Synthese Language Library, n°7, Dordrecht: Reidel.
- Dowty, D.**, 1991, « Thematic Proto-roles and Argument Selection ». *Language*, 67(3), pp. 547-619.
- Fradin, B.** 1997, "Esquisse d'une sémantique de la préfixation en anti-." *Recherches Linguistique de Vincennes*, 26: pp. 87-112.
- Fradin, B.**, 2000, "Combining forms, blends and related phenomena", in Dolsescha Ursula, Thornton Anna (eds). *Extragrammatical and Marginal Morphology*. Munich, Lincom Europa. pp 11-59.
- Fradin, B.**, 2003, *Nouvelles approches en morphologie*, Paris, Puf.
- Fradin, B.** 2005., « On a semantically grounded difference between derivation and compounding". In *Morphology and its Demarcations*, Dressler W. U., D. Katovsky & F. Rainer (eds) Amsterdam / Philadelphia: John Benjamins.
- Foley, W. A., and Van Valin, R. D.**, 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Haspelmath, M.**, 2002, *Understanding Morphology*. London, Arnold.
- Hatzfeld, A., Darmesteter, A., et al**, 1890, *Traité de la formation de la langue*, Livre premier, Paris, Delagrave.
- Iacobini, C.**, 1999, "Distinguishing derivational prefixes from initial combining forms", in *Proceedings of First Mediterranean Morphology Meeting*, Mytilene (Grecia), 19-21 septembre 1997, G. Booij, A. Ralli, S. Scalise (eds.), pp. 132-140.
- Iacobini, C.**, 2004, "Composizione con elementi neoclassici", in M. Grossmann / F. Rainer (a c. di) *La formazione delle parole in italiano*, Tübingen, Niemeyer.
- Jackendoff, R.**, 1990, *Semantic Structures*. Cambridge Mass., MIT Press.
- Kerleroux, F.**, 1991, "Les formes nues sont-elles simples ?", *Linx* 25, p. 21-44.
- Kerleroux, F.** 1996, *La coupure invisible: études de syntaxe et de morphologie*. Villeneuve d'Ascq, Presses Universitaires du Septentrion.
- Kleiber, G.**, 1999, *Problèmes de sémantique - la polysémie en questions*. Villeneuve d'Ascq, Presses Universitaires du Septentrion.
- Levin, B.** 1993. *English Verbs Classes and Alternations*. Chicago, IL: University of Chicago Press.
- Levin, B. & Rappaport H.M.**, 1995. *Unaccusativity*. Cambridge, MA: MIT Press.

- Lieber, R.**, 1981. On the organization of the lexicon. Doctoral dissertation, MIT.
- Lüdeling, A., T. Schmidt, et al.**, 2002, "Neoclassical word formation in German." *Yearbook of Morphology 2001*, pp. 253-283.
- Matthews, P. H.** 1991, *Morphology, An introduction to the theory of word-structure*, Cambridge University Press, Cambridge.
- Rainer, F. & Varela, S.**, 1992, « Compounding in Spanish », *Rivista di linguistica* 4 (1), Pisa.
- Rey-Lebove, J.**, 2004, *Le Robert brio*. Paris, Dictionnaires le Robert.
- Scalise, S.**, 1992, « Compounding in Italian », *Rivista di linguistica* 4 (1), Pisa.
- Smyth, H. W.**, 1920, *A Greek Grammar for Colleges* (libre accès à l'URL: <http://www.ccel.org/s/smyth/grammar/html/toc.htm>). New York, American Book Company.
- Talmy, L.** 2000, *Toward a cognitive semantics (2 volumes)*. Cambridge, MA, Massachusetts Institute of Technology.
- Vandeloise, C.** 1986, *L'espace en français: sémantique des prépositions spatiales*, Les Editions du Seuil, Paris.
- Vendler, Z.**, 1967, "Verbs and Times", *Linguistics and Philosophy*, Z. Vendler, Ithaca, Cornell University Press, pp. 97-121.
- Villalava, A.**, 1992, « Compounding in Portuguese », *Rivista di linguistica* 4 (1), Pisa.
- Villoing, F.**, 2002, "Les mots composés [VN]N/A du français: réflexions épistémologiques et propositions d'analyse." Nanterre, Université de Paris 10.
- Villoing, F.**, 2003, « Les mots composés VN du français : arguments en faveur d'une construction morphologique », *Cahiers de Grammaire* 28, « Morphologie et Lexique », pp. 183-196.
- Warren, B.**, 1990. "The importance of combining forms" in Dressler W. U. et al. (eds.) 1990: 111-132.
- Williams, E.**, 1981, "On the Notions 'Lexically Related' and 'Head of a Word.'" *Linguistic Inquiry*, 12,(2), pp. 245-274.

6.2 Dictionnaires

- Gaffiot, F., 2000, *Le Grand Gaffiot - Dictionnaire Latin-Français*. Paris, Hachette.
- Bailly, A., 2000, *Dictionnaire Grec-Français, Le Grand Bailly*. Paris, Hachette Education.
- Dictionnaire historique de la langue française*, 1993, Dictionnaires le Robert, Paris.
- Cottez, Henri, 1988, Dictionnaire des structures du vocabulaire Savant, Eléments et modèles de formation, Les usuels de Robert, Paris, 4e édition.*
- Le Grand Robert Electronique, 1994, CD-ROM.
- Le Robert Brio, Analyse des mots et régularités du lexique*, 2004, Dictionnaires Le Robert-VUEF, Paris.
- Trésor de la Langue Française*, Dictionnaire de la langue du XIXe et du XXe siècle, 16 volumes, 1971-1994, CNRS, Gallimard, Paris. (TLFi : <http://atilf.atilf.fr/tlf.htm> et CDROM : CNRS édition, 2004.)

6.3 Tableau comparatif

	Nom				Adjectif
	VN/NV concret +animé	VN/NV concret -animé	VN/NV événement	VN/NV lieu	
V Dynamique + N Patient	trouble-fête garde-barrière perce-oreille pique-boeufs biographe anthropophage	tournevis ouvre-boîte porte-plume trotte-bébé pense-bête odontoclaste saxifrage	lèche-vitrine saute-mouton baise-main remue-ménage lipolyse neurolyse infanticide	coupe-gorge garde-meuble pince-fesse	tue-mouche rabat-joie coupe-feu anthropophage oculogyre
V Dynamique + N Agent		cuit-vapeur gobe-mouton pousse-pied			psychogène
V Dynamique + N autre participant	ventriloque notonecte xylographe plantigrade noctambule	réveille-matin grille-midi électrographe héliotrope			plantigrade ventriloque héliotrope noctambule
V Statif, Relation spatiale + N Ground (site)		vasiducte			
V Statif, Relation spatiale + N Figure (cible)		mélanophore			carbonifère
V Statif Sentiment + N Figure (cible)	mélomane wagnerolâtre russophile anglophobe				mélomane wagnerolâtre russophile anglophobe
V Statif (Autres)	nyctalope limicole				nyctalope limicole phyllobie noctiluque

6.4 Définitions

anglophobe _A	Qui °phobe _V = 'détester' les °anglo _N = 'anglais'
anthropomorphe _A	Qui a la °morphé _N = 'forme' d'un °anthropo _N = 'être humain'
bibliophobe _A	Qui °phobe _V = 'détester' les °biblio _N = 'livre'
anthropophage _A	Qui °phage _V = 'manger' les °anthropo _N = être(s) humain(s)'
biographe _N	Celui qui °graphe _V = 'écrire' la bio _N = 'vie'
carbonifère _A	Qui °fère _V = '(trans)porter' le °carboni _N = 'carbone'
cavernicole _A	Qui °cole _V = 'vivre' dans une caverne _N
électrographe _N	(Instrument qui) °graphe _V = 'écrire' au moyen de °électro _N = 'électricité'
érythrocyte _N	°cyte = 'cellule' °érythr = 'rouge' (« globule rouge »)

génocide _N	Action de °cide _V = 'tuer' une °gen _N = 'race'
héliotrope _A	Qui °trope _V = 'se tourner' vers le °hélio _N = 'soleil'
infanticide _N	Action de °cide _V = 'tuer' un °infanti _N = 'enfant'
Jocondovore _A	Qui °vore _V = 'se nourrir' de la Joconde _{NPR}
limicole	Qui °cole _V = 'vivre' dans la °limi _N = 'boue'
lipolyse _N	Action de °lyse _V = 'désintégrer' la °lipo _N = 'graisse'
mélanophore _N	(Cellule qui) °phore _V = 'porter' la °mélano _N = 'mélanine'
mélomane _A	Qui °mane _V = 'aimer' la °mélo _N = 'musique'
neurolyse _N	Action de °lyse _V = 'désintégrer' les °neuro _N = 'nerfs'
noctambule _A	Qui °ambul _V = 'marcher' pendant la °noct _N = 'nuit'
notonecte _N	(Insecte qui) °necte _V = 'nager' sur le °noto _N = 'dos'
noctiluque	Qui °luque _V = 'brille' la °noct _N = 'nuit'
nyctalope	Qui °lope _V = 'voir' la °nyct _N = 'nuit'
oculogyre _A	Qui fait °gyre _V = 'tourner' le °ocul _N = 'globe oculaire'
odontoclaste _N	(Cellule qui) °claste _V = 'casser' les °odonto _N = 'dent'
plantigrade _A	Qui °grade _V = 'marcher' sur ses °planti _N = 'plante(s) de pied'
psychogène _A	Qui est °gène _V = 'généré' par le °psycho _N = 'esprit'
russophile _A	Qui °phile _V = 'aime' les °russo _N = 'Russes'
saxifrage _N	(Plante qui) °frage _V = 'briser' la °saxi _N = 'pierre'
tabacolâtre _A	Qui °lâtre _V = 'adorer' le tabac _N
vasiducte _N	°vasi _N = 'canal' qui °ducte _V = 'conduit'
ventriloque _A	Qui °loque _V = 'parler' avec le ventre _N
wagnerolâtre _A	Qui °lâtre _V = 'adorer' Wagner _{NPR}
xylographe _N	(Personne qui) °graphe _V = 'écrire' sur du °xylo _N = 'bois'